Security Analysis of Quantum Key Distribution: Methods and Applications

by

Jie Lin

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

Waterloo, Ontario, Canada, 2021

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Anthony Leverrier Research Scientist INRIA Paris
Supervisor:	Norbert Lütkenhaus Professor Department of Physics and Astronomy University of Waterloo
Internal Member:	Thomas Jennewein Associate Professor Department of Physics and Astronomy University of Waterloo
Internal-External Member:	Michele Mosca Professor Department of Combinatorics and Optimization University of Waterloo
Committee Member:	Jon Yard Associate Professor Department of Combinatorics and Optimization University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of contributions

This thesis is based on the following published articles, of which I am a primary author under the supervision of Norbert Lütkenhaus.

[A]. **Jie Lin** and Norbert Lütkenhaus, Phys. Rev. A 98, 042332 (2018).

[B]. **Jie Lin**, Twesh Upadhyaya and Norbert Lütkenhaus, Phys. Rev. X 9, 041064 (2019).

[C]. **Jie Lin** and Norbert Lütkenhaus, Phys. Rev. Applied 14, 064030 (2020).

This thesis is also based on the following unpublished article under preparation, of which I am a co-first author.

[D]. Thomas van Himbeeck*, **Jie Lin***, Ian George*, Kun Fang and Norbert Lütkenhaus, in preparation.

I contribute equally with Thomas van Himbeeck and Ian George. T. V. H. came up with main ideas of the algorithms as well as contributing to the implementation of algorithms. I implemented the algorithms with help of T. V. H. and worked out details in each protocol example. I. G. worked out details of the security proof claim using entropy accumulation. K. F. organized early part of the project and contributed to discussions. N. L. supervised this project.

This thesis also adapts some small portion of the following publications, of which I am a coauthor. Each of these works was led by the corresponding first author who did major parts of the work. I contributed to the discussions of ideas, refinement of details and improvement of manuscript presentations. In addition, I made partial contributions to computer codes.

[E]. Ian George, **Jie Lin** and Norbert Lütkenhaus, Phys. Rev. Research 3, 013274 (2021)

[F]. Twesh Upadhyaya, Thomas van Himbeeck, **Jie Lin** and Norbert Lütkenhaus, PRX Quantum 2, 020325 (2021)

See Section 1.2 for detailed explanations about how materials from these papers are used in this thesis.

Abstract

Quantum key distribution (QKD) can be proved to be secure by laws of quantum mechanics. In this thesis, we review security proof methods in Renner's framework and discuss numerical methods to calculate asymptotic and finite key rates. These methods are highly versatile and applicable to general device-dependent QKD protocols. We also discuss analytical tools that extend the applicability of these numerical methods. We then present the asymptotic security proof against collective attacks for a variant of the twin-field QKD protocol, which can overcome the repeaterless secret-key capacity bound. Our variant reduces the sifting cost and uses non-phase-randomized coherent states as both signals and test states. We confirm the loss scaling of this protocol. Another important family of protocols that we investigate here are discrete-modulated continuous-variable QKD protocols. They are interesting due to their experimental simplicity and their great potential for massive deployment in the quantum-secured networks. Our security proof method can provide tight asymptotic key rates. We demonstrate that the postselection of data in combination with reverse reconciliation can improve the key rates. We analyze both untrusted and trusted detector noise scenarios. Our results in the trusted detector noise scenario show that we can thus cut out most of the effect of detector noise and obtain asymptotic key rates similar to those had we access to ideal detectors. Finally, we present several simple examples to illustrate our newly developed method for the numerical finite-key analysis against the most general attacks via the entropy accumulation theorem.

Acknowledgements

First of all, I would like to thank my supervisor Professor Norbert Lütkenhaus who gave me the opportunity to pursue my Master of Science and PhD study in the Institute for Quantum Computing. I want to thank him for his support, helpful discussions and advice in the past six years.

I would also like to thank past and present members of the OQCT research group. I would like to give special thanks to all my coauthors: Dr. Patrick Coles, Dr. Fang Kun, Ian George, Dr. Hao Hu, Haesol Im, Dr. Thomas van Himbeeck, Professor Norbert Lütkenhaus, Adam Winick, Professor Henry Wolkowicz, Dr. Yanbao Zhang. I have learned a lot from all of them. I would also like to thank many other friends from IQC. In particular, thank Sara Zafar Jafarzadeh and Vlad Gheorghiu for both scientific discussions and many helpful advice in the daily life. Thank Anqi Huang, Hao Qin and Poompong Chaiwongkhot for many interesting discussions. I would like to thank my PhD advisory committee members: Professor Thomas Jennewein, Professor Jon Yard, Professor John Watrous and my supervisor for giving helpful feedback on my research progress. I also would like to thank my PhD exam committee members for agreeing in serving in the committee and spending time reading this thesis. In particular, I would like to thank my external examiner, Dr. Anthony Leverrier, for pointing out many typographical errors in the previous version of the thesis and for providing helpful suggestions that improve the final version of this thesis. The remaining errors and omissions are entirely the author's responsibility.

I would like to thank my family members and close relatives for their encouragement and love, and particularly thank Charlie for many helps.

Dedication

This is dedicated to my family.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

RSA Rivest-Shamir-Adleman

QKD quantum key distribution

DV discrete variable

CV continuous variable

BB84 Bennett-Brassard 1984

POVM positive operator-valued measure

PVM projection-valued measure

EPR Einstein, Podolsky and Rosen

SDP semidefinite program

EAT entropy accumulation theorem

CP completely positive

TP trace-preserving

TNI trace-nonincreasing

TF-QKD twin-field QKD

DI device-independent

MDI measurement-device-independent

PM-MDI QKD phase-matching measurement-device-independent QKD

DMCVQKD discrete-modulated continuous-variable QKD

QPSK quadrature phase-shift keying

QND quantum non-demolition

CSS Calderbank-Shor-Steane

QECC quantum error-correcting code

PLOB Pirandola, Laurenza, Ottaviani and Banchi

GLLP Gottesman, Lo, Lütkenhaus and Preskill

i.i.d. independent and identically distributed

AEP asymptotic equipartition property

SNU shot noise unit

MUB mutually unbiased basis

List of Symbols

$\mathcal{H}$	Hilbert space
$\langle \cdot, \cdot \rangle$ or $\langle \cdot \vert \cdot \rangle$	Inner product
$\mathbb{1}$	Identity map
$\mathrm{L}(\mathcal{H})$	Space of linear maps from $\mathcal{H}$ to itself
$\mathrm{Herm}(\mathcal{H})$	Set of Hermitian operators on $\mathcal{H}$
$\mathrm{Pos}(\mathcal{H})$	Set of positive operators on $\mathcal{H}$
$\mathrm{D}(\mathcal{H})$	Set of density matrices on $\mathcal{H}$
$\mathrm{D}_{\leq}(\mathcal{H})$	Set of subnormalized states on $\mathcal{H}$
$P \geq 0$	Positive operator (positive semidefinite matrix) P
$\Vert \cdot \Vert_p$	Schatten p-norm
$\mathrm{id}_{\mathcal{H}}$	Identity channel on $\mathrm{L}(\mathcal{H})$
$J(\Phi)$	Choi representation of the quantum channel Φ
$\mathrm{Sym}(\mathcal{H}^{\otimes n})$	Symmetric subspace of $\mathcal{H}^{\otimes n}$
$\mathrm{F}(\cdot, \cdot)$	Fidelity function
$P(\cdot, \cdot)$	Purified distance
$\mathbb{P}(\mathcal{X})$	Space of probability vectors over the alphabet $\mathcal{X}$
$\mathrm{dom}(f)$	Domain of a function f
$\mathrm{relint}(C)$	Relative interior of a set C
$\rho_{\mathrm{th}}(\bar{n})$	A thermal state with its mean photon number $\bar{n}$
A_1^n	A collection of registers $A_1, \cdots, A_n$
boldface letters like $\boldsymbol{q}$	A vector in some vector space
$\mathbf{1}$	The vector that contains all 1's
Greek letters like Σ, Λ	Typical names for alphabets. Exceptions may exist.
Greek letters like Φ, Ψ	Typical names for quantum channels. Exceptions may exist.

Chapter 1

Introduction

In the current digital world, cryptography provides an indispensable safeguard to protect our data for both integrity and confidentiality. Classical cryptography is typically built on unproven mathematical assumptions and certain restrictions of an adversary's computational powers. In particular, classical cryptographic protocols are designed by assuming that some underlying mathematical problems are difficult to solve; i.e., no efficient algorithm currently exists. A well-known example is Rivest-Shamir-Adleman (RSA), a public-key cryptographic encryption protocol based on the difficulty of prime factorization. However, there is no guarantee that no efficient algorithm can be discovered in the future to render such a cryptographic protocol insecure. In fact, if one looks at the history of cryptography, the chase between code makers and code breakers has been lasting for many centuries. A lot of cryptographic protocols have been broken by discovery of efficient algorithms to invalidate the underlying assumptions. Since then, those protocols have been replaced by new cryptographic protocols that rely on new unproven assumptions. The repetition of this process continues.

In recent years, quantum information has become a rapidly developing field to build emerging technologies based on the laws of quantum mechanics. The idea of quantum computers was firstly envisioned by a famous physicist, Richard Feynman. It is arguably to say that quantum computing started to attract more attentions and became a more serious research area after the discovery of Shor's algorithm [1] for prime factorization and discrete logarithm by Peter Shor in 1994. While quantum computers are expected to bring us advantages in solving difficult problems, they also bring challenges for cryptographic systems. With the invent of Shor's algorithm, current cryptographic protocols based on prime factorization and discrete logarithm will become insecure soon after a practical quantum computer is invented to realize Shor's algorithm. Recognizing the threats posed by quantum computers calls the need of quantum-safe cryptography. Currently, there are two major solutions: post-quantum cryptography, which consists of classical cryptographic algorithms that are thought to be resistant to quantum attacks; quantum cryptography that directly uses quantum mechanics to achieve cryptographic goals. Both approaches have attractive advantages and unfortunate limitations. Roughly speaking, post-quantum cryptography is a software solution in the sense that switching from the current existing cryptographic protocols to post-quantum cryptographic protocols can be done mainly on the software level. However, the underlying security assumptions for post-quantum cryptography are still unproven and there remains the possibility that one needs to constantly replace these algorithms by a newer generation of algorithms. On the other hand, quantum cryptography can be viewed as a hardware solution since quantum cryptographic protocols usually require new hardware infrastructures. Unlike classical cryptography, many quantum cryptographic protocols can be proven to be secure in the information-theoretic sense, which is an attractive feature. Moreover, quantum cryptography researchers have been investigating protocols that can reuse the existing classical communication infrastructures with the hope that these quantum solutions can be more cost-effective, thereby becoming more competitive to post-quantum cryptography. Nevertheless, it is highly expected that both post-quantum and quantum cryptography will be useful and will complement each other in different use cases. In this thesis, we focus on quantum cryptography and specifically quantum key distribution (QKD).

1.1 Quantum key distribution

An early idea of quantum cryptography was proposed by Stephen Wiesner in 1970s. In his work, he proposed a quantum money scheme based on conjugate basis encoding. However, his work remained unpublished until 1983 [2]. The idea of conjugate basis encoding was later explored by Bennett and Brassard, which led to the invention of Bennett-Brassard 1984 (BB84) protocol [3] published in 1984, the first QKD protocol. QKD is a key establishment protocol that allows two distant honest parties, traditionally known as Alice and Bob, to establish a shared secret key even if an eavesdropper, Eve, can tamper the communication channel. Because their work was published in the proceeding of a computer science conference, it remained largely unnoticed by the physics community until 1991 when Artur Ekert independently invented an equivalent version of the BB84 protocol in his paper [4] published in *Physical Review Letters*. This equivalence was shown by Bennett, Brassard and Mermin [5]. More attentions from the physics community have been drawn to quantum cryptography. Despite some common confusion, it should be noted that the notion of *quantum cryptography* is not equivalent to QKD since QKD is not the only task in quantum cryptography. There are many other quantum cryptographic tasks such as quantum fingerprinting [6–10], blind quantum computing [11–15] and quantum digital signature [16, 17]. Among all quantum cryptographic protocols, QKD is arguably the most developed one. Many QKD protocols have been designed and implemented with or without full security proofs. There are many good reviews on the topic of QKD [18–23].

According to their detection technology, QKD protocols can be categorized into two families:

discrete variable (DV) and continuous variable (CV). In particular, DVQKD protocols like the BB84 protocol [3] are realized by encoding the information into qubit-like degrees of freedom of photons, such as polarization and time bin, and by measuring with single-photon detectors. On the other hand, CVQKD (e.g., see Refs. [24–26]) uses detection technology that is widely used in modern optical (classical) communication methods, which turns those classical methods and the CVQKD apparatus into nearly identical devices. This technological similarity gives CVQKD a competitive edge for large-scale deployment in quantum-secured networks. DVQKD enjoys great success in experimental implementations and corresponding security analyses, and can currently reach longer distances than CVQKD. Among the direct point-to-point QKD protocols, DVQKD can currently reach over 421 km [27] while CVQKD recently achieves 202.81 km [28]. Many experiments of CVQKD on both Gaussian modulation schemes such as Refs. [28–34] and discrete modulation schemes like Refs. [35–39] have been demonstrated.

For conventional QKD protocols, one usually assumes that sources and detectors are trusted and characterized in the security proofs. Under assumptions about behaviors of those devices, QKD protocols can be proven to be information-theoretically secure. However, the physical implementations of QKD usually deviate from the theoretical models. Thus, the deviations open up side-channel attacks for Eve. The field of quantum hacking is an active research field to explore gaps between the theory and the experiment, thereby breaking the implementation security of QKD. One possible approach to solve this issue is (i) to revise models for real devices so that security proofs use more reasonable assumptions and (ii) to implement necessary countermeasures to safeguard the devices at the same time. However, this approach can be costly in practice since one may not be able to completely close the gap between the theory and the implementation. Countermeasures may also open up more side-channel attacks. Fortunately, QKD provides the forward secrecy; that is, keys generated from QKD remain secure if they were secure at the time they were generated. Future discovery of side-channel attacks cannot break the security of existing keys generated by QKD. Another approach is to design device-independent (DI) protocols. In DIQKD [40–42], one makes minimal assumptions about devices: devices do not maliciously leak the private information. In particular, one needs not to trust devices. Instead, one typically performs non-local games to test violation of Bell's inequalities and to verify that those devices operate in the quantum regime before any secret key is generated. Although recent theoretical works [43, 44] make the implementation more feasible, a realistic demonstration of DIQKD is still missing at the time of writing.

Notably, measurement devices are more vulnerable to side-channel attacks than sources as there are more attacks targeted at measurement devices (see e.g. [45–53]; also see [54–56]). Following this observation, the proposal of measurement-device-independent (MDI) protocols [57, 58] aims to provide a more practical solution compared to DI protocols. While security proofs still assume trusted and characterized sources, MDIQKD protocols have been implemented to reach a couple hundred of kilometers in fiber [59, 60] and recently in free space [61].

The original MDI protocols are based on two-photon interference events. Recently, twin-field

QKD (TF-QKD) [62, 63] is proposed based on single-photon interference events. A main motivation of this proposal is related to one major limitation of QKD protocols. As QKD protocols are usually implemented with quantum optical systems, photons are used as the information carriers. Unfortunately, photons can be lost during the transmissions and thus the secret key rate is limited. By now, it is known that there are key rate limits on the point-to-point direct link of QKD protocols [64, 65], which are called repeaterless bounds. In particular, the bound by Pirandola, Laurenza, Ottaviani and Banchi (PLOB) [65] is the tightest bound for the lossy channel and is dubbed PLOB bound by the names of these authors. For the single-photon transmittance η, the PLOB bound is $-\log_2(1-\eta) = O(\eta)$, where we use the big O notation from computer science. This means that for all direct point-to-point QKD protocols, the secret key rate scales at most linearly with the channel transmittance. Protocols like the single-photon BB84 and Gaussian-modulated CVQKD have the key rate scaling of $O(\eta)$. The BB84 protocol with weak coherent pulses scales as $O(\eta^2)$ if no decoy state is used and scales as $O(\eta)$ when the decoy state method [66–68] is applied. Even though the MDI protocols involve an intermediate station and thus are not subject to the PLOB bound, the original MDI protocols based on two-photon interference events scale as $O(\eta)$. In this case, although the probability of a photon arriving from each party (assuming a symmetric setup where the intermediate station is at the middle of Alice and Bob) is $O(\sqrt{\eta})$, a successful measurement outcome at the middle station requires two photons to arrive and thus the key rate scales as $O(\eta)$. The key observation in TF-QKD is that based on a single-photon interference, Alice and Bob can also generate secret keys. Because only one photon is needed for a successful detection, it is expected that TF-QKD can scale as $O(\sqrt{\eta})$. This bound is confirmed in several works [63, 69–73] that appeared around the same time. While TF-QKD is successful to beat the repeaterless bounds and current experiments can reach around 500 km, it is like a one-node repeater and thus limited by one-node repeater bound that scales like $O(\sqrt{\eta})$. To reach much longer distances, quantum repeaters are needed. The field of quantum repeaters is an active research field that aims at extending the quantum communication to arbitrarily long distances without trusting the intermediate nodes. Currently, it is still very challenging to realize a practical quantum repeater that can overcome the repeaterless bound (if we do not consider the TF-QKD as a special quantum repeater). If we sacrifice some security promise and trust the intermediate nodes, then we can extend QKD to arbitrarily long distances by the trusted-node relays (see e.g. [74]). Satellites can also be used for this purposes.

In summary, there are many progresses in the field of QKD. Novel protocols have been constantly proposed and analyzed. Many QKD experiments have been demonstrated to reach increasingly longer distances and/or to achieve higher key bits per second. The maturity of the field can be witnessed by development of commercial prototypes in several startup companies, successful launch of a QKD satellite in China [75] and many ongoing efforts to launch QKD satellites from all over the world (see [76, Table 2] for a list), and developments of chip-based QKD systems [77–80].

1.2 Contribution and structure of this thesis

This thesis covers in details these published papers [71, 81, 82] of which I am the first author and an ongoing, unpublished work [83] of which I am a co-first author. It also summarizes some important ideas from these papers [84, 85] of which I am a coauthor. It briefly mentions some results from those papers [86, 87] of which I am a coauthor. Another paper [88] that I coauthored during my Ph.D. study is not covered in this thesis.

The structure of the thesis is outlined as follows.

In Chapter 2, we review some background knowledge on quantum information theory, quantum optics, convex optimization as well as quantum key distribution. This chapter contains reviews only. Readers who are familiar with these topics can skip this chapter at the first reading and can refer to this chapter when certain materials are referenced in other chapters.

In Chapter 3, we then review the security definition of QKD as well as various proof techniques. Most parts of this chapter are literature review. It also contains some refinements of statements that are presented in our work [84]. The discussion here focuses on Renner's framework [89] for security proofs based on the leftover hashing lemma since all security proofs in this thesis use this framework. We briefly mention an alternative framework based on the phase error correction in Section 3.8.1.

In Chapter 4, we discuss numerical methods for the security proof. It includes the formulation of nonlinear semidefinite programs for both asymptotic key rate and finite key analysis. Some part of Section 4.1 is a literature review. Section 4.2 contains a review of the qubit-based squashing model [90–92], a summary of the flag-state squasher from our work [86], a summary of dimension reduction method from our paper [85] as well as a summary of the facial reduction idea from our paper [87]. The finite-key method based on a canonical approach in Section 4.3 is summarized from our work [84]. The algorithm discussion of finite key analysis via the entropy accumulation theorem in Section 4.4 is from our unpublished work [83].

We then move to the security analysis of individual protocols. These protocols that we provide asymptotic security analysis include phase-matching measurement-device-independent QKD (PM-MDI QKD) that is a variant of TF-QKD in Chapter 5, discrete-modulated continuous-variable QKD (DMCVQKD) with heterodyne detection in Chapter 6 and DMCVQKD with homodyne detection in Chapter 7. We also include finite key analysis of some simple entanglement-based device-dependent protocols using the entropy accumulation theorem (EAT) in Chapter 8.

In Chapter 5, the asymptotic security analysis of PM-MDI QKD is mostly from our paper [71]. Our variant of TF-QKD removes the need of phase randomization and uses properties of coherent states directly. Our proof was among early works to show that such a protocol can overcome the secret key capacity bound for point-to-point QKD. Since our initial protocol proposal involves some idealization, we discuss some ideas to prove the security of a practical version of the protocol and we review a related work by Primaatmaja *et. al.* [93] that proves the security of a practical version. Since TF-QKD has been a hot topic in the QKD community since its proposal in 2018, there are many progresses made after we published our work. Therefore, we also provide an additional, short literature review about recent progresses in Section 5.6.

In Chapter 6, we study the asymptotic security analysis of DMCVQKD where Bob applies the heterodyne detection. The security analysis in the ideal (or untrusted) detector scenario is from [81] and that in the trusted detector scenario is from [82]. Both analyses use the photon-number cutoff assumption as explained in Section 6.5. While this assumption is numerically plausible, it lacks rigorous justifications that one would like to have for a waterproof security proof. Our recent dimension reduction method [85] allows us to remove the photon-number cutoff assumption. The discussion about removing the photon-number cutoff assumption in Section 6.10 is based on [85].

In Chapter 7, we study the asymptotic security of DMCVQKD where Bob uses the homodyne detection. The presentation is based on [81]. For this variant, we study only the ideal detector scenario under the photon-number cutoff assumption. It is possible to extend the analysis of this protocol to trusted noise scenario and also possible to remove the photon-number cutoff assumption using methods similar to the heterodyne scheme. The exact modification is left for future work.

In Chapter 8, we discuss more about finite key analysis based on EAT for entanglement-based device-dependent QKD protocols. This discussion is based on [83].

We provide technical details in appendices.

In Appendix A, we discuss the formulation of classical postprocessing steps in the numerical framework. This presentation is based on [81].

In Appendix B, we present technical details related the security proof of PM-MDI QKD studied in Chapter 5. This presentation is from [71].

In Appendix C, we present additional technical details related to DMCVQKD. Specifically, in Appendix C.1, we present a derivation of POVM elements for the noisy heterodyne detector model used in Chapter 6. This presentation is from [82]. In Appendix C.2, we discuss how to represent operators under the photon-number cutoff assumption. This presentation is from [81, 82].

In Appendix D, we present technical details and proofs for the numerical finite-key analysis method via EAT. While the main idea is due to Thomas van Himbeeck, a co-first author of our work [83], the specific presentation in Appendix D.1 is mostly due to me. The proof in Appendix D.2 is mostly due to Ian George, a co-first author of our work [83].

Chapter 2

Background

To make this thesis more self-contained, we review some necessary background on quantum information theory, quantum optics, convex optimization and quantum key distribution. They are helpful to understand this thesis. This chapter also serves a purpose of defining common notations used throughout this thesis. Readers who are familiar with these topics can skip this chapter.

2.1 Mathematical preliminaries

We assume readers are familiar with basic linear algebra as well as basic concepts about norm, metric and inner product (see, e.g., [94, Chapter 1, Section 1.1] for a quick review). In this section, we start with some important mathematical concepts related to the quantum information theory.

We use physicists' convention for the inner product, that is, $\langle \cdot, \cdot \rangle$ is linear in the second variable and conjugate linear in the first variable. We often use Dirac notations, for example, $|\cdot\rangle$ for vectors, and $\langle \cdot|\cdot \rangle$ for inner products between vectors. We now review the definition of Hilbert space.

Definition 2.1.1 (Hilbert space)**.** A (complex) Hilbert space, denoted as $\mathcal{H}$, is a vector space over $\mathbb{C}$ equipped with an inner product $\langle \cdot|\cdot \rangle$ such that it is a Banach space (i.e. complete normed space) with the norm induced from the inner product.

In this thesis, we always assume a given Hilbert space is over $\mathbb{C}$. Furthermore, we only deal with separable Hilbert spaces, which can emit a countable orthonormal basis. The dual space $\mathcal{H}^*$ of $\mathcal{H}$ is the space of bounded linear functionals, that is, linear maps from $\mathcal{H}$ to $\mathbb{C}$. By Riesz representation theorem, it is natural to define a unique element in $\mathcal{H}^*$, denoted as $\langle \psi|$, a bra vector, for each ket vector $|\psi\rangle \in \mathcal{H}$ such that the linear functional $\langle\psi|$ takes the value $\langle\psi|\phi\rangle$ when evaluated at $|\phi\rangle \in \mathcal{H}$ and also $\|\langle\psi|\|_{\mathcal{H}^*} = \||\psi\rangle\|_{\mathcal{H}}$ where the norm on $\mathcal{H}^*$ is the operator norm.

As we often deal with multipartite systems, the concept of tensor products is important.

Definition 2.1.2 (Tensor product of Hilbert spaces)**.** Let $\mathcal{H}$ and $\mathcal{K}$ be two Hilbert spaces with inner products $\langle\cdot|\cdot\rangle_{\mathcal{H}}$ and $\langle\cdot|\cdot\rangle_{\mathcal{K}}$, respectively. Let $\mathcal{H}\otimes\mathcal{K}$ denote the tensor product of these two vector spaces. For every $|u\rangle, |v\rangle \in \mathcal{H}\otimes\mathcal{K}$ such that $|u\rangle = \sum_{i=1}^{n} h_i \otimes f_i$ and $|v\rangle = \sum_{i=1}^{k} x_j \otimes y_j$ where $h_i, x_j \in \mathcal{H}$ and $f_i, y_j \in \mathcal{K}$, we set

$$\langle u|v\rangle = \sum_{i,j=1}^{n,k} \langle h_i|x_j\rangle_{\mathcal{H}} \langle f_i|y_j\rangle_{\mathcal{K}}\,. \tag{2.1}$$

This defines an inner product on the vector space $\mathcal{H}\otimes\mathcal{K}$. We call the completion of the vector space $\mathcal{H}\otimes\mathcal{K}$ under this inner product as the tensor product of these two Hilbert spaces. With an abuse of notation, we denote this completion also as $\mathcal{H}\otimes\mathcal{K}$.

One can apply this definition inductively to define n-fold tensor product spaces. In particular, we write the n-fold tensor product of the same Hilbert space $\mathcal{H}$ as $\mathcal{H}^{\otimes n}$, that is, $\mathcal{H}^{\otimes n} = \mathcal{H}\otimes\cdots\otimes\mathcal{H}$ with n copies of $\mathcal{H}$.

It is interesting to study linear maps on the Hilbert space $\mathcal{H}$ to itself, which are called linear operators. In particular, let $\mathbb{1}_{\mathcal{H}}$ denote the identity map on $\mathcal{H}$. Let $\mathrm{L}(\mathcal{H})$ denote the vector space of linear operators that acts on $\mathcal{H}$. Let $\mathrm{B}(\mathcal{H}) \subseteq \mathrm{L}(\mathcal{H})$ denote the vector space of bounded linear operators. There are several important sets of operators that are relevant for our discussions.

Definition 2.1.3 (Hermitian operator)**.** An operator $H \in \mathrm{B}(\mathcal{H})$ is Hermitian or self-adjoint if $H = H^\dagger$, where $H^\dagger$ is the adjoint operator of H defined by the adjoint equation $\left\langle H^\dagger\phi|\psi\right\rangle = \langle\phi|H\psi\rangle\,\forall\,|\psi\rangle$ and $\forall\,\langle\phi|$. Let $\mathrm{Herm}(\mathcal{H})$ denote the set of Hermitian operators on $\mathcal{H}$.

Definition 2.1.4 (Positive operator)**.** An operator $P \in \mathrm{B}(\mathcal{H})$ is positive, denoted as $P \geq 0$, if $\forall h \in \mathcal{H}, \langle h|Ph\rangle \geq 0$. Let $\mathrm{Pos}(\mathcal{H})$ denote the set of positive operators on $\mathcal{H}$.

Remark 2.1.5. We note that an alternative characterization of Hermitian operators is that H is Hermitian if and only if $\langle\psi|H|\psi\rangle \in \mathbb{R}, \forall\psi \in \mathcal{H}$. We also note two alternative characterizations of positive operators: (1) $P \geq 0$ if and only if $P \in \mathrm{Herm}(\mathcal{H})$ and the spectrum of P lies in $[0,\infty)$; (2) $P \geq 0$ if and only if it can be written as $P = X^\dagger X$ for $X \in \mathrm{B}(\mathcal{H})$.

Definition 2.1.6 (Projection operator)**.** An operator P is a projection operator if $P \geq 0$ and $P^2 = P = P^\dagger$.

Definition 2.1.7 (Density operator)**.** A density operator ρ is a positive operator with $\mathrm{Tr}(\rho) = 1$. We denote the space of density operators on $\mathcal{H}$ by $\mathrm{D}(\mathcal{H})$.

If $\mathcal{H}$ is finite-dimensional, we can also call density operators as density matrices since they can be represented by finite-dimensional positive semidefinite matrices with unit trace. As we see later, it is sometimes interesting to talk about the set of positive operators with trace at most 1, which is the set of subnormalized states. We denote the set of subnormalized states by $\mathrm{D}_{\leq}(\mathcal{H})$.

We are often interested in functions of operators. We briefly remark how one can extend an ordinary function on real or complex numbers to a function on operators.

Remark 2.1.8. For a function of the form $f : \mathbb{C} \to \mathbb{C}$, one often would like to extend f to the set of Hermitian operators $\mathrm{Herm}(\mathcal{H})$[1]. By the spectral theorem [94, Theorem 1.3], such an extension can be done. If the function f is defined on a subset of $\mathbb{C}$, then the extension also needs to restrict to operators whose spectrum is restricted accordingly. In particular, the square root and logarithm of a positive operator can be defined. We omit the details here and direct interested readers to [94, Chapter 1, Section 1.1.3] for more explanations.

We review some important norms for an operator in $\mathrm{L}(\mathcal{H})$. In particular, we are interested in a family of norms called Schatten p-norms, which include the trace norm, Hilbert-Schmidt norm and operator norm as special cases.

Definition 2.1.9 (Schatten p-norm)**.** Let $\mathcal{H}$ be a Hilbert space. Let $p \in [1, \infty)$. For any $L \in \mathrm{L}(\mathcal{H})$, the Schatten p-norm of L is

$$\|L\|_p = \left[\mathrm{Tr}(|L|^p) \right]^{\frac{1}{p}}, \tag{2.2}$$

where $|L| = \sqrt{L^\dagger L}$. One can take the limit $p \to \infty$ to define $\|\cdot\|_\infty$.

Remark 2.1.10. For $p = 1$, the Schatten 1-norm is also called the trace norm. For $p = 2$, the Schatten 2-norm coincides with the Frobenius norm and also the Hilbert-Schmidt norm, which is the induced norm from the Hilbert-Schmidt inner product on $\mathrm{L}(\mathcal{H})$. For $p = \infty$, the Schatten ∞-norm coincides with the operator norm. For each $p \in [1, \infty]$, the dual norm of Schatten p-norm is Schatten q-norm where $1/p + 1/q = 1$. In general, one can also define Schatten p-norm for linear transformations between two different Hilbert spaces.

The Schatten p-norm and its dual norm satisfy the Hölder's inequality for Schatten norms, which is a quite useful relation.

Lemma 2.1.11 (Hölder's inequality for Schatten norms)**.** Let $\mathcal{H}$ be two Hilbert spaces. For each $p, q \in [1, \infty]$ such that $1/p + 1/q = 1$, the following inequality holds for any $A, B \in \mathrm{L}(\mathcal{H})$:

$$|\langle A, B \rangle| \leq \|A\|_p \|B\|_q. \tag{2.3}$$

[1] More generally, the set of normal operators, i.e., $\{L \in \mathrm{L}(\mathcal{H}) : L^\dagger L = L L^\dagger\}$.

It is useful in many practical scenarios to have some ways to measure closeness between density operators or subnormalized states, or more generally, positive operators. We first define trace distance and the fidelity function for this purpose.

Definition 2.1.12 (Trace distance)**.** Let $P, Q \in \mathrm{Pos}(\mathcal{H})$. The trace distance between P and Q is defined as

$$\delta(P, Q) = \frac{1}{2}\|P - Q\|_1. \tag{2.4}$$

It is worth noting that some authors define the trace distance without the factor 1/2.

Definition 2.1.13 (Fidelity)**.** Let $P, Q \in \mathrm{Pos}(\mathcal{H})$. The fidelity $\mathrm{F}(P, Q)$ between P and Q is defined as

$$\mathrm{F}(P, Q) = \left\|\sqrt{P}\sqrt{Q}\right\|_1. \tag{2.5}$$

It is worth noting that some authors (e.g. [95]) define the fidelity function as $\mathrm{F}(P, Q)^2$ in our definition here. Also, it is helpful to know some relationship between the trace distance and the fidelity function.

Theorem 2.1.14 (Fuchs-van de Graaf inequalities, [94, Theorem 3.33])**.** Let $\rho, \sigma \in \mathrm{D}(\mathcal{H})$. It holds that

$$1 - \frac{1}{2}\|\rho - \sigma\|_1 \leq \mathrm{F}(\rho, \sigma) \leq \sqrt{1 - \frac{1}{4}\|\rho - \sigma\|_1^2}. \tag{2.6}$$

Equivalently,

$$2 - 2\,\mathrm{F}(\rho, \sigma) \leq \|\rho - \sigma\|_1 \leq 2\sqrt{1 - \mathrm{F}(\rho, \sigma)^2}. \tag{2.7}$$

Since the fidelity function is not a metric, it is sometimes handy to have a metric that is closely related to fidelity. The purified distance is a metric that is related to the fidelity function and is defined for subnormalized states.

Definition 2.1.15 (Purified distance)**.** Let $\rho, \sigma \in \mathrm{D}_{\leq}(\mathcal{H})$. The purified distance between ρ and σ is defined as

$$P(\rho, \sigma) = \sqrt{1 - \left(\left\|\sqrt{\rho}\sqrt{\sigma}\right\|_1 + \sqrt{(1 - \mathrm{Tr}(\rho))(1 - \mathrm{Tr}(\sigma))}\right)^2}. \tag{2.8}$$

The trace distance can be upper bounded by the purified distance, which follows from the Fuchs-van de Graaf inequalities. We state it as a lemma here.

Lemma 2.1.16. Let $\rho, \sigma \in \mathrm{D}_{\leq}(\mathcal{H})$. The following inequality holds:

$$\frac{1}{2}\|\rho - \sigma\|_1 \leq P(\rho, \sigma). \tag{2.9}$$

In our discussion, we implicitly assume the Hilbert spaces are finite-dimensional. Many definitions and statements can be generalized to infinite-dimensional Hilbert spaces, but some care is needed. We briefly remark some complications in the infinite-dimensional Hilbert spaces. Readers interested in infinite-dimensional Hilbert spaces may find it useful to learn more about functional analysis and in particular operator algebra.

Remark 2.1.17. When $\mathcal{H}$ is infinite-dimensional, one often deals with compact operators instead of general linear operators. Compact operators are limits of finite rank operators (with respect to the operator norm). A finite rank operator, as the name suggests, is a linear operator whose range is finite-dimensional. Let $\mathbb{K}(\mathcal{H})$ denote the set of compact operators on $\mathcal{H}$. It is the case that $\mathbb{K}(\mathcal{H}) \subseteq \mathrm{B}(\mathcal{H})$. If $\mathcal{H}$ is finite-dimensional, $\mathrm{L}(\mathcal{H}) = \mathrm{B}(\mathcal{H}) = \mathbb{K}(\mathcal{H})$. The definition of density operator implicitly restricts to the trace-class operators, a subset of compact operators, since only trace-class operators have a well-defined trace. Trace-class operators are operators whose Schatten 1-norm is finite. Similarly, the definition of quantum channel that we define in the next section (in Definition 2.2.23) is restricted to a mapping from trace-class operators to trace-class operators instead of general linear operators when the relevant Hilbert spaces are infinite-dimensional. To avoid unnecessary complications, we assume the dimension of $\mathcal{H}$ is finite unless stated otherwise. The generalization to the infinite-dimensional spaces can be straightforward in many cases by simply replacing $\mathrm{L}(\mathcal{H})$ by either $\mathrm{B}(\mathcal{H})$, $\mathbb{K}(\mathcal{H})$ or some subspace of $\mathbb{K}(\mathcal{H})$. However, it can also be complicated and some results are not generalized to infinite-dimensional spaces. The reader is cautioned that some of results stated in this thesis may rely on the assumption that one works with the finite-dimensional spaces.

2.2 Basic quantum information theory

We review the definitions of quantum states, measurements, quantum channels and relevant entropic quantities in this section. This section is mainly based on the textbook by Watrous [94] and the textbook by Nielsen and Chuang [96]. More details can be found there.

When we discuss physical systems, we typically use the term *register*. Intuitively, registers are mathematical abstractions of physical objects, or parts of physical objects which can store information. We use the term *alphabet* to mean a finite and nonempty set, whose elements may be considered as symbols. We use capital Greek letters such as Σ, Λ to denote alphabets. One can also think of an alphabet as an index set.

Definition 2.2.1 (Register, [94, Definition 2.1])**.** A *register* X is either one of the following two objects:

1. An alphabet Σ.
2. An n-tuple $X = (Y_1, \cdots, Y_n)$, where n is a positive integer and $Y_1, \ldots, Y_n$ are registers.

For a register X, we often write the identity operator as $\mathbb{1}_X$ in place of $\mathbb{1}_{\mathcal{H}_X}$ for simplicity.

2.2.1 Quantum states and measurements

In quantum mechanics, the state space of a physical system is a complex Hilbert space $\mathcal{H}$. Pure states are unit vectors in the Hilbert space.

Definition 2.2.2 (Pure state)**.** Suppose $\mathcal{H}$ is the Hilbert space of the physical system of interest. We call $|\psi\rangle$ a pure state if $|\psi\rangle \in \mathcal{H}$ with $\||\psi\rangle\| = 1$.

In general, an arbitrary state of a system is not necessarily pure. It can be given as a probabilistic mixture of pure states. This leads to the concept of mixed state.

Definition 2.2.3 (Mixed state)**.** Let $\{|\psi_i\rangle : i = 1, \ldots, n\}$ be a collection of pure states. Let $p_i \geq 0$ such that $\sum_{i=1}^n p_i = 1$. We call $\{|\psi_i\rangle, p_i : 1 \leq i \leq n\}$ a mixed state.

We typically represent states by density operators. A pure state ψ can be represented by a density operator $\rho = |\psi\rangle \langle\psi|$. Given an ensemble $\{|\psi_i\rangle, p_i : 1 \leq i \leq n\}$, this mixed state can be represented by a density operator $\rho = \sum_{i=1}^n p_i |\psi_i\rangle\langle\psi_i|$.

Definition 2.2.4 (Ensemble of states)**.** Let $\{\rho_i \in \mathrm{D}(\mathcal{H}) : i = 1, \ldots, n\}$ be a collection of states. Let $p_i \geq 0$ such that $\sum_{i=1}^n p_i = 1$. We call $\{\rho_i, p_i : 1 \leq i \leq n\}$ an ensemble of states.

When we deal with bipartite or multipartite systems, it is often of interest to discuss the marginal state on some parts of a system. A reduced density operator can be obtained by taking the partial trace.

Definition 2.2.5 (Partial trace)**.** Let $\mathcal{H}_{A_1}, \cdots, \mathcal{H}_{A_n}$ be Hilbert spaces. The partial trace mapping that removes the k-th system is the unique linear map that satisfies the equation

$$\mathrm{Tr}_{A_k}(X_1 \otimes \cdots \otimes X_n) = \mathrm{Tr}(X_k) X_1 \otimes \cdots X_{k-1} \otimes X_{k+1} \cdots \otimes X_n \tag{2.10}$$

for all $X_i \in \mathcal{H}_{A_i}$.

Definition 2.2.6 (Reduced density operator)**.** Let $\rho_{AB} \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. The reduced density operator on system A is $\rho_A = \mathrm{Tr}_B(\rho_{AB})$ and the reduced density operator on system B is $\rho_B = \mathrm{Tr}_A(\rho_{AB})$.

Theorem 2.2.7 (Schmidt decomposition)**.** Let $\rho_{AB} = |\psi\rangle\langle\psi| \in \mathrm{D}(\mathcal{H}_{AB})$ be a pure state. Then one can write

$$|\psi\rangle_{AB} = \sum_{i=1}^{r} \sqrt{\lambda_i}\, |i\rangle_A |i\rangle_B \tag{2.11}$$

such that $\rho_A = \mathrm{Tr}_B(\rho_{AB}) = \sum_i \lambda_i |i\rangle\langle i|_A$ and $\rho_B = \mathrm{Tr}_A(\rho_{AB}) = \sum_i \lambda_i |i\rangle\langle i|_B$, where $\{|i\rangle_A\}$ and $\{|i\rangle_B\}$ are orthonormal sets of vectors in $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively.

In many practical scenarios, it is more convenient to deal with pure states instead of mixed states. One can always find a larger space such that there exists a pure state whose reduced density operator is the original mixed state.

Theorem 2.2.8 (Purification)**.** Let $\rho_A \in \mathrm{D}(\mathcal{H}_A)$. Then there exists a reference space $\mathcal{H}_R$ with $\dim(\mathcal{H}_R) \geq \dim(\mathcal{H}_A)$ and a pure state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_R$ such that $\rho_A = \mathrm{Tr}_R(|\psi\rangle\langle\psi|)$.

Entanglement is an important concept in quantum information. We now define separable and entangled states.

Definition 2.2.9 (Separable states)**.** A state $\rho \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is separable if it can be written as a convex combination of product states

$$\rho_{AB} = \sum_{x\in\Sigma} p(x)\rho_A^x \otimes \rho_B^x, \tag{2.12}$$

where Σ is some alphabet, $p(x)$ is an entry of a probability vector $\boldsymbol{p}$, and $\rho_A^x \in \mathrm{D}(\mathcal{H}_A), \rho_B^x \in \mathrm{D}(\mathcal{H}_B)$ for all $x \in \Sigma$.

Definition 2.2.10 (Entangled states)**.** A state $\rho \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is entangled if it is not separable.

When we discuss about de Finetti theorems in Section 3.5, it is useful to know symmetric subspaces and exchangeable states. We briefly define them here and direct interested readers to [89, Chapter 4] and [94, Chapter 7] for details. Let $\mathcal{S}_n$ be the set of permutations on $\{1, \ldots, n\}$. For any $\pi \in \mathcal{S}_n$, let W_π be the unitary operation on $\mathcal{H}^{\otimes n}$ which permutes the n subsystems, i.e.,

$$W_\pi(|\theta_1\rangle \otimes \cdots \otimes |\theta_n\rangle) := \left|\theta_{\pi^{-1}(1)}\right\rangle \otimes \cdots \otimes \left|\theta_{\pi^{-1}(n)}\right\rangle \tag{2.13}$$

for every $|\theta_1\rangle, \ldots, |\theta_n\rangle \in \mathcal{H}$.

Definition 2.2.11 (Symmetric subspace)**.** Let $\mathcal{H}$ be a Hilbert space and let $n \in \mathbb{N}$ be a positive integer. The symmetric subspace of $\mathcal{H}^{\otimes n}$, denoted by $\mathrm{Sym}(\mathcal{H}^{\otimes n})$, is the subspace of $\mathcal{H}^{\otimes n}$ that is spanned by all vectors which are invariant under any permutations W_π of the subsystems, i.e.,

$$\mathrm{Sym}(\mathcal{H}^{\otimes n}) := \{|\Psi\rangle \in \mathcal{H}^{\otimes n} : W_\pi |\Psi\rangle = |\Psi\rangle \ \ \forall \pi \in \mathcal{S}_n\}. \tag{2.14}$$

Definition 2.2.12 (Exchangeable states)**.** A density operator $\rho \in \mathrm{D}(\mathcal{H}^{\otimes n})$ is exchangeable if and only if

$$\rho = W_\pi \rho W_\pi^\dagger \tag{2.15}$$

for every permutation $\pi \in \mathcal{S}_n$.

Definition 2.2.13 ($n+k$-exchangeable states)**.** A density operator $\rho_n \in \mathrm{D}(\mathcal{H}^{\otimes n})$ is $n+k$-*exchangeable* for some $k \in \mathbb{N}$ if there exists $\rho_{n+k} \in \mathrm{D}(\mathcal{H}^{\otimes n+k})$ such that ρ_{n+k} is exchangeable and $\mathrm{Tr}_{\mathcal{H}^{\otimes k}}(\rho_{n+k}) = \rho_n$. Moreover, $\rho_n \in \mathrm{D}(\mathcal{H}^{\otimes n})$ is called *infinitely exchangeable* if it is $n+k$-exchangeable for all $k \in \mathbb{N}$.

In a physical scenario, one often performs some measurements on some quantum state of interest to obtain some classical measurement outcomes. We review the mathematical description of measurements.

Definition 2.2.14 (POVM)**.** Let $\mathcal{H}$ be a Hilbert space. Let Σ be an alphabet. A collection of positive operators $\{E_i \in \mathrm{Pos}(\mathcal{H}) : i \in \Sigma\}$ is called a positive operator-valued measure (POVM) if

$$\sum_{k\in\Sigma} E_k = \mathbb{1}_{\mathcal{H}} \,. \tag{2.16}$$

A special type of measurements is projective measurements or projection-valued measure.

Definition 2.2.15 (PVM)**.** Let $\mathcal{H}$ be a Hilbert space. Let Σ be an alphabet. A collection of positive operators $\{P_i \in \mathrm{Pos}(\mathcal{H}) : 1 \leq i \leq n\}$ is called a projection-valued measure (PVM) if $\sum_{k\in\Sigma} P_k = \mathbb{1}_{\mathcal{H}}$ and $P_i^2 = P_i = P_i^\dagger$ for each $1 \leq i \leq n$.

While a general POVM is not necessarily a projective measurement, one can always construct a PVM from POVM on a larger space, a consequence of the Naimark's dilation theorem.

Theorem 2.2.16 (Naimark's dilation theorem)**.** Let $\mathcal{H}_A$ be a Hilbert space. Let $\{E_i \in \mathrm{Pos}(\mathcal{H}_A) : 1 \leq i \leq n\}$ be a POVM. There exists a Hilbert space $\mathcal{H}_R$, an isometry $V : \mathcal{H}_A \to \mathcal{H}_A \otimes \mathcal{H}_R$ and a PVM $\{P_i : 1 \leq i \leq n\}$ such that $E_i = V^\dagger P_i V$ for each i.

2.2.2 Quantum channels

Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be some Hilbert spaces. We are interested in physical processes that transform one state (density operator) in $\mathrm{D}(\mathcal{H}_A)$ to some density operator in $\mathrm{D}(\mathcal{H}_B)$. Such a process can be described mathematically as linear maps. There are some additional requirements if we demand the output of such a linear map to be a quantum state whenever the input is a valid quantum state. In particular, they are quantum channels which are linear maps that map density operators in $\mathrm{D}(\mathcal{H}_A)$ to density operators in $\mathrm{D}(\mathcal{H}_B)$. We now define several interesting linear maps that act on linear operators on some Hilbert spaces, including quantum channels.

Definition 2.2.17 (Hermitian-preserving map)**.** A linear map $\Phi : \mathrm{L}(\mathcal{H}_A) \to \mathrm{L}(\mathcal{H}_B)$ is Hermitian-preserving if for every Hermitian operator $H \in \mathrm{Herm}(\mathcal{H}_A)$, it is the case that $\Phi(H) \in \mathrm{Herm}(\mathcal{H}_B)$.

Definition 2.2.18 (Positive map)**.** A linear map $\Phi : \mathrm{L}(\mathcal{H}_A) \to \mathrm{L}(\mathcal{H}_B)$ is positive if for every positive operator $P \in \mathrm{Pos}(\mathcal{H}_A)$, it is the case that $\Phi(P) \in \mathrm{Pos}(\mathcal{H}_B)$.

Definition 2.2.19 (CP map)**.** A linear map $\Phi : \mathrm{L}(\mathcal{H}_A) \to \mathrm{L}(\mathcal{H}_B)$ is completely positive (CP) if for every Hilbert space $\mathcal{H}_R$, the map $\Phi \otimes \mathrm{id}_R$ is a positive map.

Definition 2.2.20 (TP map)**.** A linear map $\Phi : \mathrm{L}(\mathcal{H}_A) \to \mathrm{L}(\mathcal{H}_B)$ is trace-preserving (TP) if for every $X \in \mathrm{L}(\mathcal{H}_A)$, it is the case that $\mathrm{Tr}(\Phi(X)) = \mathrm{Tr}(X)$.

Definition 2.2.21 (TNI map)**.** A linear map $\Phi : \mathrm{L}(\mathcal{H}_A) \to \mathrm{L}(\mathcal{H}_B)$ is trace-nonincreasing (TNI) if for every $P \in \mathrm{Pos}(\mathcal{H}_A)$, it is the case that $\mathrm{Tr}(\Phi(P)) \leq \mathrm{Tr}(P)$.

Note that in the definition of TNI map, the condition is restricted to only positive operators. This restriction is necessary since if one still required that the condition holds for all $\mathrm{L}(\mathcal{H}_A)$, this definition would be identical to the definition of TP maps. One can see this by taking $P \in \mathrm{Pos}(\mathcal{H}_A)$ and $-P$, which leads to conditions $\mathrm{Tr}(\Phi(P)) \leq \mathrm{Tr}(P)$ and $-\mathrm{Tr}(\Phi(P)) \leq -\mathrm{Tr}(P) \implies \mathrm{Tr}(\Phi(P)) = \mathrm{Tr}(P)\ \forall P \in \mathrm{Pos}(\mathcal{H}_A)$. Since any operator $X \in \mathrm{L}(\mathcal{H}_A)$ can be written as a linear combination of positive operators, by linearity of Φ, it would be the case that $\mathrm{Tr}(\Phi(X)) = \mathrm{Tr}(X)$, the precise definition of TP maps.

Definition 2.2.22 (Unital map)**.** A linear map $\Phi : \mathrm{L}(\mathcal{H}_A) \to \mathrm{L}(\mathcal{H}_B)$ is unital if $\Phi(\mathbb{1}_{\mathcal{H}_A}) = \mathbb{1}_{\mathcal{H}_B}$.

Definition 2.2.23 (Quantum channel)**.** A quantum channel Φ between two registers A and B is a linear map from $\mathrm{L}(\mathcal{H}_A)$ to $\mathrm{L}(\mathcal{H}_B)$ such that Φ is CPTP.

In particular, a CPTP map maps density operators in $\mathrm{D}(\mathcal{H}_A)$ to density operators in $\mathrm{D}(\mathcal{H}_B)$. A special channel is the identity channel which leaves the state intact. We denote the identity channel acting on $\mathrm{L}(\mathcal{H})$ by $\mathrm{id}_{\mathcal{H}}$. We now look at different characterizations of a quantum channel.

Proposition 2.2.24 (Kraus representation)**.** Let $\mathcal{H}_1, \mathcal{H}_2$ be Hilbert spaces. Let $\Phi : \mathrm{L}(\mathcal{H}_1) \to \mathrm{L}(\mathcal{H}_2)$ be a linear map. Let Σ be an alphabet. Let $\{A_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{H}_1, \mathcal{H}_2)$ be a collection of operators indexed by Σ. If Φ can be written as

$$\Phi(X) = \sum_{a \in \Sigma} A_a X A_a^\dagger \tag{2.17}$$

for all $X \in \mathrm{L}(\mathcal{H}_1)$, then Φ is CP. Furthermore, if $\sum_{a \in \Sigma} A_a^\dagger A_a = \mathbb{1}_{\mathcal{H}_1}$, then Φ is TP.

Conversely, if Φ is CP, then there exists an alphabet Σ and $\{A_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{H}_1, \mathcal{H}_2)$ be a collection of operators such that Φ can be written in the form of Eq. (2.17). If Φ is TP, then $\sum_{a \in \Sigma} A_a^\dagger A_a = \mathbb{1}_{\mathcal{H}_1}$. We note that a map Φ can have many different Kraus representations.

For any map Φ, another interesting and related map is its adjoint map.

Definition 2.2.25 (Adjoint map)**.** Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two Hilbert spaces. Let $\Psi : \mathrm{L}(\mathcal{H}_1) \to \mathrm{L}(\mathcal{H}_2)$ be a linear map. The adjoint map of Ψ, denoted by $\Psi^\dagger$, is the unique linear map that satisfies the adjoint equation:

$$\left\langle \Psi^\dagger(B) \middle| A \right\rangle_{\mathcal{H}_1} = \langle B | \Psi(A) \rangle_{\mathcal{H}_2} \tag{2.18}$$

for all $A \in \mathrm{L}(\mathcal{H}_1)$ and $B \in \mathrm{L}(\mathcal{H}_2)$.

Given a Kraus representation of a quantum channel Φ with $\Phi(X) = \sum_{a\in\Sigma} A_a X A_a^\dagger$, one can write a Kraus representation of its adjoint map $\Phi^\dagger$ as $\Phi^\dagger(Y) = \sum_{a\in\Sigma} A_a^\dagger Y A_a$. As Φ is TP, $\Phi^\dagger$ is unital.

Another useful representation of quantum channel is the Choi representation.

Definition 2.2.26 (Choi matrix)**.** Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two Hilbert spaces with $\dim(\mathcal{H}_1) = d$. Let $\Phi : \mathrm{L}(\mathcal{H}_1) \to \mathrm{L}(\mathcal{H}_2)$ be a linear map. The Choi matrix of Φ is

$$J(\Phi) = \sum_{i,j=1}^{d} \Phi(E_{i,j}) \otimes E_{i,j}, \tag{2.19}$$

where $E_{i,j}$ is a d-by-d matrix with (i, j)-entry being 1 and all other entries being 0.

Proposition 2.2.27. Let $\Phi : \mathrm{L}(\mathcal{H}_1) \to \mathrm{L}(\mathcal{H}_2)$ be a linear map with $\mathcal{H}_1$ and $\mathcal{H}_2$ being two Hilbert spaces. The action of the map Φ can be recovered from the Choi matrix $J(\Phi)$ by

$$\Phi(X) = \mathrm{Tr}_{\mathcal{H}_1}(J(\Phi)(\mathbb{1}_{\mathcal{H}_2} \otimes X^T)). \tag{2.20}$$

Moreover, Φ is CP if and only if $J(\Phi) \geq 0$. Φ is TP if and only if $\mathrm{Tr}_{\mathcal{H}_2}(J(\Phi)) = \mathbb{1}_{\mathcal{H}_1}$.

We describe a particular type of channel called depolarizing channel.

Definition 2.2.28 (Depolarizing channel)**.** Let $p \in [0, 1]$. Let $\mathcal{H}$ be a Hilbert space with $\dim(\mathcal{H}) = d$. The depolarizing channel $\mathcal{D}_p : \mathrm{L}(\mathcal{H}) \to \mathrm{L}(\mathcal{H})$ is defined as

$$\mathcal{D}_p(\rho) = (1 - p)\rho + p\,\mathrm{Tr}(\rho)\frac{\mathbb{1}_{\mathcal{H}}}{d} \tag{2.21}$$

for every $\rho \in \mathrm{L}(\mathcal{H})$.

An interesting special case is for a qubit system, i.e., $\mathcal{H} = \mathbb{C}^2$. We can characterize this channel for a qubit by Pauli matrices as Kraus operators:

$$\mathcal{D}_p(\rho) = (1 - \frac{3p}{4})\rho + \frac{p}{4}(X\rho X + Y\rho Y + Z\rho Z), \tag{2.22}$$

where X, Y and Z are Pauli matrices:

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},\ Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix},\ Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{2.23}$$

In particular, we typically use the depolarizing channel to simulate noise on qubit systems (see Chapter 8).

Finally, we end this section with the definition of the diamond norm for a linear map.

Definition 2.2.29 (Diamond norm)**.** Let $\mathcal{E} : \mathrm{L}(\mathcal{H}_A) \to \mathrm{L}(\mathcal{H}_B)$ be a linear map. The *diamond norm* of $\mathcal{E}$ is defined as

$$\|\mathcal{E}\|_\diamond = \max_{\sigma \in \mathrm{D}_{\leq}(\mathcal{H}_A \otimes \mathcal{H}_R)} \|(\mathcal{E} \otimes \mathrm{id}_R)(\sigma)\|_1. \tag{2.24}$$

2.2.3 Entropy

Entropy is a powerful tool in information theory since it has operational significance in many information-processing tasks and is often an indispensable tool for quantitative analyses. In this thesis, we deal with security proofs of QKD protocols and we use many entropic quantities to calculate the secret key rate. We review those important entropic quantities. In addition to [94, 96], some of the material presented here is also based on [95]. While our focus is on quantum entropies, it is often instructive to mention the classical entropies where quantum entropies are generalizations of their classical counterparts.

Definition 2.2.30 (Shannon entropy)**.** Let X be a discrete random variable whose value is taken from a finite set $\mathcal{X}$ and which is distributed according to the probability distribution P_X. The Shannon entropy $H(X)$ of X is defined as

$$H(X) = -\sum_{x\in\mathcal{X}} P_X(x)\log_2(P_X(x)). \tag{2.25}$$

The von Neumann entropy is a generalization of the Shannon entropy.

Definition 2.2.31 (von Neumann entropy)**.** Let $\mathcal{H}_X$ be the Hilbert space that describes the physical system of X. Let $\rho \in \mathrm{D}(\mathcal{H}_X)$ be the state of the system. Then von Neumann entropy $H(X)_\rho := H(\rho)$ of the system X in the state ρ is defined as

$$H(X)_\rho = -\operatorname{Tr}(\rho\log_2(\rho)). \tag{2.26}$$

The following lemma, known as Fannes' inequality, shows that the von Neumann entropy is continuous and its proof can be found in [96].

Lemma 2.2.32 (Fannes' inequality)**.** Let $\mathcal{H}$ be a Hilbert space with $\dim(\mathcal{H}) = n$. Let $\rho, \sigma \in \mathrm{Pos}(\mathcal{H})$ such that $\|\rho - \sigma\|_1 \leq \kappa \leq 1/e$, where e is the base of the natural logarithm. Then,

$$|H(\rho) - H(\sigma)| \leq \kappa\log_2(n/\kappa). \tag{2.27}$$

Quantum relative entropy is a useful aid for studying the von Neumann entropy.

Definition 2.2.33 (Quantum relative entropy)**.** Let $P, Q \in \mathrm{Pos}(\mathcal{H})$. The quantum relative entropy of P with respect to Q is defined as

$$D(P||Q) = \begin{cases} \operatorname{Tr}(P\log_2(P)) - \operatorname{Tr}(P\log_2(Q)) & \text{if } \operatorname{im}(P) \subseteq \operatorname{im}(Q) \\ \infty & \text{otherwise.} \end{cases} \tag{2.28}$$

For classical systems, one can represent the probability distribution P_X by a diagonal matrix with entries specified by P_X. Because P_X is a valid probability distribution, this diagonal matrix is a valid density matrix. Therefore, one can easily see the quantum entropies are generalizations of classical entropies. For this reason, we state only quantum versions of different entropic quantities and assume the classical version can be deduced as a special case directly from the quantum versions.

Definition 2.2.34 (Conditional entropy)**.** Let $\rho_{AB} \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. The conditional von Neumann entropy of A given B is defined as

$$H(A|B)_{\rho_{AB}} = H(A,B)_{\rho_{AB}} - H(B)_{\rho_B}. \tag{2.29}$$

We note that the conditional entropy can be written in terms of quantum relative entropy as

$$H(A|B)_{\rho_{AB}} = -D(\rho_{AB}||\, \mathbb{1}_A \otimes \rho_B). \tag{2.30}$$

Definition 2.2.35 (Mutual information)**.** Let $\rho_{AB} \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. The (quantum) mutual information between A and B is defined as

$$I(A\!:\!B)_{\rho_{AB}} = H(A)_{\rho_A} + H(B)_{\rho_B} - H(A,B)_{\rho_{AB}}. \tag{2.31}$$

Definition 2.2.36 (Holevo quantity)**.** Let X be a classical register with an alphabet Σ. Let Y be a register and $\{\rho_a \in \mathrm{D}(\mathcal{H}_Y), p_a : a \in \Sigma\}$ be an ensemble of states. The Holevo information (also called the Holevo quantity), denoted by $\chi(X\!:\!Y)$, is defined as the quantum mutual information $I(X\!:\!Y)$ between registers X and Y with respect to the state σ where

$$\sigma = \sum_{a\in\Sigma} p_a \, |a\rangle\langle a|_X \otimes \rho_a. \tag{2.32}$$

The accessible information is related to the scenario where classical information is encoded in a quantum system with an ensemble of quantum states. It is the amount of classical information that can be extracted from such a quantum system when an optimal measurement is applied.

Definition 2.2.37 (Accessible information)**.** Let $E := \{\rho_x, p_x : x \in \Sigma\}$ be an ensemble of states. Let X be a random variable that takes value $x \in \Sigma$. Let Y_M be the random variable that denotes the measurement outcomes from a POVM M applied to this ensemble. The accessible information is defined as

$$I_{\mathrm{acc}}(E) = \max_M I(X\!:\!Y_M), \tag{2.33}$$

where the maximization is over all possible POVMs.

Definition 2.2.38 (Conditional mutual information)**.** Let $\rho_{ABC} \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$. The conditional mutual information between A and B conditioned on C is defined as

$$I(A\!:\!B|C)_{\rho_{ABC}} = H(A|C)_{\rho_{AC}} + H(B|C)_{\rho_{BC}} - H(AB|C)_{\rho_{ABC}}. \tag{2.34}$$

With the help of the conditional mutual information, one can define the Markov chain condition, a useful concept when we discuss the entropy accumulation theorem in Section 3.7 and in Chapter 8.

Definition 2.2.39 (Markov chain condition)**.** Let $\rho_{ABC} \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C))$. The tripartite quantum state ρ_{ABC} is said to fulfill the Markov chain condition $A \leftrightarrow C \leftrightarrow B$ if $I(A\!:\!B|C)_{\rho_{ABC}} = 0$.

Similar to the relative entropy, the α-Rényi divergence D_α is a useful tool to study the α-Rényi entropy. There are many different quantum generalizations of the classical α-Rényi divergence. We state two definitions that are commonly used.

Definition 2.2.40 (Sandwiched α-Rényi divergence)**.** Let $\alpha \in [\frac{1}{2}, 1) \cup (1, \infty)$, and let $P, Q \in \mathrm{Pos}(\mathcal{H})$ with $P \neq 0$. The sandwiched quantum Rényi divergence of P with Q is defined as

$$\tilde{D}_\alpha(P||Q) = \begin{cases} \frac{1}{\alpha-1} \log_2 \left(\frac{\left\| Q^{\frac{1-\alpha}{2\alpha}} P Q^{\frac{1-\alpha}{2\alpha}} \right\|_\alpha^\alpha}{\mathrm{Tr}(P)} \right) & \text{if } (\alpha < 1 \wedge P \not\perp Q) \vee \mathrm{im}(P) \subseteq \mathrm{im}(Q) \\ \infty & \text{otherwise.} \end{cases} \tag{2.35}$$

For $\alpha = \infty$ and a density operator ρ, one can define $D_\infty(\rho||Q) = \log_2 \inf\{\lambda : \rho \leq \lambda Q\}$.

Definition 2.2.41 (Petz α-Rényi divergence)**.** Let $\alpha \in [0, 1) \cup (1, 2]$, and let $P, Q \in \mathrm{Pos}(\mathcal{H})$ with $P \neq 0$. The Petz Rényi divergence of P with Q is defined as

$$\bar{D}_\alpha(P||Q) = \begin{cases} \frac{1}{\alpha-1} \log_2 \left(\frac{\mathrm{Tr}\left(P^\alpha Q^{1-\alpha}\right)}{\mathrm{Tr}(P)} \right) & \text{if } (\alpha < 1 \wedge P \not\perp Q) \vee \mathrm{im}(P) \subseteq \mathrm{im}(Q) \\ \infty & \text{otherwise.} \end{cases} \tag{2.36}$$

We note that for $\alpha \to 1$ and for $P \in \mathrm{D}(\mathcal{H})$, both $\tilde{D}_\alpha(P||Q)$ and $\bar{D}_\alpha(P||Q)$ converge to the quantum relative entropy $D(P||Q)$.

One can define various conditional Rényi entropies from each α-Rényi divergence. There are two useful variants of quantum conditional entropy from the sandwiched Rényi divergence. These entropies will be useful if readers are interested in details of EAT in [97–99].

Definition 2.2.42. For $\alpha \in [\frac{1}{2}, \infty]$ and for any density operator ρ_{AB}, the sandwiched α-Rényi entropy of A conditioned on B is defined as

$$H_\alpha(A|B)_\rho = -\tilde{D}_\alpha(\rho_{AB}||\, \mathbb{1}_A \otimes \rho_B). \tag{2.37}$$

Definition 2.2.43. For $\alpha \in [\frac{1}{2}, 1) \cup (1, \infty]$ and for any density operator ρ_{AB}, we define

$$H_\alpha^{\uparrow}(A|B)_\rho = - \inf_{\sigma_B \in \mathrm{D}_{\leq}(\mathcal{H}_B)} \tilde{D}_\alpha(\rho_{AB}||\, \mathbb{1}_A \otimes \sigma_B). \tag{2.38}$$

There are two special conditional entropies of interest: conditional min-entropy which corresponds to $H_{\infty}^{\uparrow}$ and max-entropy which corresponds to $H_{\frac{1}{2}}^{\uparrow}$.

Definition 2.2.44 (Conditional min-entropy)**.** Let $\rho_{AB} \in \mathrm{D}_{\leq}(\mathcal{H}_A \otimes \mathcal{H}_B)$. The min-entropy of A conditioned on B of the state ρ_{AB} is

$$H_{\min}(A|B)_{\rho} = \sup_{\sigma_B \in \mathrm{D}_{\leq}(\mathcal{H}_B)} \sup\{\lambda \in \mathbb{R} : \rho_{AB} \leq \exp(-\lambda)\, \mathbb{1}_A \otimes \sigma_B\} \tag{2.39}$$

Definition 2.2.45 (Conditional max-entropy)**.** Let $\rho_{AB} \in \mathrm{D}_{\leq}(\mathcal{H}_A \otimes \mathcal{H}_B)$. The max-entropy of A conditioned on B of the state ρ_{AB} is

$$H_{\max}(A|B)_{\rho} = \sup_{\sigma_B \in \mathrm{D}_{\leq}(\mathcal{H}_B)} 2 \log_2 \mathrm{F}(\rho_{AB}, \mathbb{1}_A \otimes \sigma_B) \tag{2.40}$$

We use the convention that the max-entropy corresponds to $\alpha = \frac{1}{2}$ since this max-entropy satisfies the duality relation with the min-entropy [95, Section 5.3]. As we heavily rely on results from Renner's Ph.D. thesis [89], it should be noted that a different definition of max-entropy is used there. It corresponds to $\alpha = 0$, and is the quantum version of the Hartley entropy. We denote this as H_0.

Definition 2.2.46 (Conditional Hartley entropy)**.** Let $\rho_{AB} \in \mathrm{D}_{\leq}(\mathcal{H}_A \otimes \mathcal{H}_B)$. The quantum Hartley entropy of A conditioned on B of the state ρ_{AB} is

$$H_0(A|B)_{\rho} = \sup_{\sigma \in \mathrm{D}(\mathcal{H}_B)} \log_2 \mathrm{Tr}[(\mathbb{1}_A \otimes \sigma_B)\Pi_{\rho_{AB}}], \tag{2.41}$$

where $\Pi_{\rho_{AB}}$ is the projection onto the support of ρ_{AB}.

The min-entropy and the max-entropy can be discontinuous, which can be undesirable in some applications. This leads to the idea of smoothed entropies.

We first define a ball centered around a state $\rho \in \mathrm{D}(\mathcal{H})$ with a radius r as

$$B_r(\rho) := \{\sigma \in \mathrm{D}_{\leq}(\mathcal{H}) : P(\rho, \sigma) \leq r\}. \tag{2.42}$$

In this definition, we use the purified distance as a way to measure how close the state σ is to the state ρ. Different authors may use different distance measures in defining such a ball. In particular, the trace distance is used in [89]. According to Lemma 2.1.16, the ball defined in terms of the trace distance is contained in the ball defined in terms of the purified distance. This means the derivation of any lower bound (upper bound) on the smooth min-entropy (max-entropy) from [89] remains valid for the smooth min-entropy (max-entropy) defined with the purified distance.

Definition 2.2.47 (Smooth min-entropy)**.** Let $\varepsilon \in [0, 1]$. For any density operator $\rho_{AB} \in \mathrm{D}(\mathcal{H}_{AB})$, the ε-smooth min-entropy of A conditioned on B is defined as

$$H^{\varepsilon}_{\min}(A|B)_\rho = \sup_{\sigma \in B_\varepsilon(\rho)} H_{\min}(A|B)_\sigma, \tag{2.43}$$

where $B_\varepsilon(\rho)$ is defined in Eq. (2.42) in terms of the purified distance.

Definition 2.2.48 (Smooth max-entropy)**.** Let $\varepsilon \in [0, 1]$. For any density operator $\rho_{AB} \in \mathrm{D}(\mathcal{H}_{AB})$, the ε-smooth max-entropy of A conditioned on B is defined as

$$H^{\varepsilon}_{\max}(A|B)_\rho = \inf_{\sigma \in B_\varepsilon(\rho)} H_{\max}(A|B)_\sigma, \tag{2.44}$$

where $B_\varepsilon(\rho)$ is defined in Eq. (2.42) in terms of the purified distance.

2.3 Introduction to quantum optics

QKD protocols are usually realized by quantum optical systems. Therefore, to prove the security of a practical QKD protocol, it is necessary to understand some basics about quantum optics. We review some essential concepts from quantum optics. This section is mainly based on [100, 101].

2.3.1 Optical modes

In classical electrodynamics, an optical mode is referred to some orthonormal basis solution to the *Maxwell's Equations* for the vector potential in the vacuum. A general solution can be expressed as a linear combination of those modes. One can show that the energy of a classical electromagnetic field mode is of the form of a harmonic oscillator with canonical position and moment variables (also called field quadratures), q and p, respectively, for the classical system. In quantum mechanics, the field amplitudes of orthonormal modes are promoted to mode operators through canonical quantization. Each mode is converted into a quantum harmonic oscillator. In terms of the canonical variables, these classical variables q and p for each mode are replaced by their corresponding field quadrature operators $\hat{q}$ and $\hat{p}$ where they satisfy the canonical commutation relation

$$[\hat{q}, \hat{p}] = i\hbar \mathbb{1}\,. \tag{2.45}$$

We follow custom to drop the identity operator $\mathbb{1}$ and choose the natural units (i.e. $\hbar = 1$) throughout this thesis. The reader is cautioned that different authors may adopt different conventions by setting $\hbar = \frac{1}{2}, 1$ or 2.

Let $\hat{a}$ and $\hat{a}^\dagger$ be annihilation and creation operators of a mode, respectively. They satisfy the commutation relation

$$[\hat{a}, \hat{a}^\dagger] = 1. \tag{2.46}$$

One can write the photon number operator $\hat{n}$ as $\hat{n} = \hat{a}^\dagger \hat{a}$. We can write field quadrature operators $\hat{q}$ and $\hat{p}$ in terms of annihilation and creation operators $\hat{a}$ and $\hat{a}^\dagger$ as

$$\hat{q} = \frac{1}{\sqrt{2}}(\hat{a} + \hat{a}^\dagger),\ \hat{p} = \frac{i}{\sqrt{2}}(\hat{a}^\dagger - \hat{a}). \tag{2.47}$$

We can also define *quadrature states* which are eigenstates of the quadrature operators, $\hat{q}\left|q\right\rangle = q\left|q\right\rangle$ and $\hat{p}\left|p\right\rangle = p\left|p\right\rangle$. They are normalized under the conditions $\langle q|q'\rangle = \delta(q-q')$, $\langle p|p'\rangle = \delta(p-p')$ where $\delta(\cdot)$ is the Dirac delta function. In particular, the set $\{\left|q\right\rangle : q \in \mathbb{R}\}$ and the set $\{\left|p\right\rangle : p \in \mathbb{R}\}$ each form a complete basis, i.e., $\int |q\rangle\langle q|\, dq = \mathbb{1}$ and $\int |p\rangle\langle p|\, dp = \mathbb{1}$.

2.3.2 Important classes of quantum states of light

We review several important single-mode states and their properties. We also review an important two-mode state: two-mode squeezed vacuum state, which is the continuous-variable version of Einstein, Podolsky and Rosen (EPR) states, maximally entangled states.

Fock state

Fock states, or photon-number states, are eigenstates of the number operator $\hat{n}$. We denote them by $\left|n\right\rangle$ where n is the corresponding eigenvalue of $\hat{n}$:

$$\hat{n}\left|n\right\rangle = n\left|n\right\rangle. \tag{2.48}$$

Moreover, the set of Fock states $\{\left|n\right\rangle : n \in \mathbb{N}\}$ forms a complete basis for the Hilbert space of a single-mode system.

Coherent state

A coherent state $\left|\alpha\right\rangle$ is an eigenstate of the annihilation operator $\hat{a}$ with a complex eigenvalue α. It can be written in the Fock state basis as

$$\left|\alpha\right\rangle = e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} \left|n\right\rangle. \tag{2.49}$$

Coherent states can be generated by the displacement operator, $\hat{D}(\alpha)$, acting on the vacuum.

Definition 2.3.1 (Displacement operator)**.** The displacement operator with a displacement $\alpha \in \mathbb{C}$ is defined as

$$\hat{D}(\alpha) = \exp\Big(\alpha \hat{a}^\dagger - \alpha^* \hat{a}\Big). \tag{2.50}$$

We also define s-ordered displacement operator by the relation

$$\hat{D}(\alpha, s) = \hat{D}(\alpha) e^{s|\alpha|^2/2}, \tag{2.51}$$

where $s \in \mathbb{C}$.

Then $|\alpha\rangle = \hat{D}(\alpha)\,|0\rangle$. It is interesting to note the relationship

$$\hat{D}(\alpha)\hat{D}(\beta) = e^{i\,\mathrm{Im}(\alpha\beta^*)}\hat{D}(\alpha+\beta). \tag{2.52}$$

It is also worth noting that coherent states form an overcomplete basis for the Hilbert space of a single-mode system, i.e., $\int_{\mathbb{C}} |\alpha\rangle\langle\alpha|\, d^2\alpha = \pi\,\mathbb{1}$. The overlap between two possibly different coherent states $|\alpha\rangle$ and $|\beta\rangle$ is

$$\langle\beta|\alpha\rangle = e^{i\,\mathrm{Im}(-\beta\alpha^*)} e^{-|\alpha-\beta|^2/2} = e^{-(|\alpha|^2+|\beta|^2)/2+\alpha\beta^*}. \tag{2.53}$$

For a coherent state $|\alpha\rangle$, its expectation values for first- and second-moment of field quadrature operators $\hat{q}$ and $\hat{p}$ as follows:

$$\langle\hat{q}\rangle := \langle\alpha|\,\hat{q}\,|\alpha\rangle = \frac{1}{\sqrt{2}}(\alpha+\alpha^*) \tag{2.54}$$

$$\langle\hat{q^2}\rangle := \langle\alpha|\,\hat{q}^2\,|\alpha\rangle = \frac{1}{2}(\alpha^2+\alpha^{*2}+2|\alpha|^2+1) \tag{2.55}$$

$$\langle\hat{p}\rangle := \langle\alpha|\,\hat{p}\,|\alpha\rangle = \frac{i}{\sqrt{2}}(\alpha^*-\alpha) \tag{2.56}$$

$$\langle\hat{p^2}\rangle := \langle\alpha|\,\hat{p}^2\,|\alpha\rangle = -\frac{1}{2}(\alpha^2+\alpha^{*2}-2|\alpha|^2-1). \tag{2.57}$$

The mean photon number of a coherent state $|\alpha\rangle$ is

$$\langle\hat{n}\rangle := \langle\alpha|\,\hat{n}\,|\alpha\rangle = |\alpha|^2. \tag{2.58}$$

The photon number distribution of a coherent state follows the Poisson distribution. The probability of finding n photons is

$$p_n(|\alpha|^2) = \frac{(|\alpha|^2)^n e^{-|\alpha|^2}}{n!}. \tag{2.59}$$

Another interesting property about coherent states is that they are minimum uncertainty states, that is, $\Delta q \Delta p = 1/2$, where $\Delta q = \sqrt{\langle\hat{q}^2\rangle - \langle\hat{q}\rangle^2}$ and Δp is defined similarly.

Thermal state

Thermal states are the states of the light that do not evolve in time in thermal equilibrium. They are the statistical mixture of Fock states with maximal disorder; i.e. they maximize the von Neumann entropy for a given energy E. They can be written as

$$\rho_{\text{th}}(\bar{n}) = \sum_{n=0}^{\infty} \frac{\bar{n}^n}{(1+\bar{n})^{n+1}} |n\rangle\langle n| \,, \tag{2.60}$$

where $\bar{n}$ is the mean photon number of the thermal state. In particular, the photon number distribution of the thermal state (which corresponds to the coefficients in Eq. (2.60)) is a thermal distribution, that is,

$$p_n(\bar{n}) = \frac{\bar{n}^n}{(\bar{n}+1)^{n+1}}. \tag{2.61}$$

The uncertainty in photon number is $\Delta n = \sqrt{\bar{n}^2 + \bar{n}}$.

One can define a thermal operator $\hat{T}(\theta)$ which creates thermal states out of the vacuum [102].

Definition 2.3.2 (Thermal operator)**.** Let $\hat{a}$ be the annihilation operator for the mode of interest. Let $\hat{b}$ be the annihilation operator for a fictitious ancillary mode. Let $\theta \in \mathbb{R}$ be a parameter. The thermal operator is

$$\hat{T}(\theta) = \exp\Big[\theta(\hat{a}^\dagger \hat{b}^\dagger - \hat{a}\hat{b})\Big]. \tag{2.62}$$

The thermal operator is of the form of a two-mode squeezing operator (see Definition 2.3.4). A thermal state can be created by tracing out the ancillary mode of $\hat{T}(\theta)\, |0\rangle_a \, |0\rangle_b$. The mean photon number $\bar{n}$ of a thermal state is related to the parameter θ in the thermal operator by $\bar{n} = \sinh^2(\theta)$.

Squeezed state

Another important class of minimum uncertainty states are squeezed states. Unlike coherent states where the uncertainty is equally spread between q and p quadratures, they distribute the uncertainty unequally between quadratures. One can use the squeezing operator to generate squeezed states.

Definition 2.3.3 (Squeezing operator)**.** Let $\xi \in \mathbb{C}$. The squeezing operator with a squeezing parameter ξ is defined as

$$\hat{S}(\xi) = \exp\left[\frac{1}{2}\xi^* \hat{a}^2 - \frac{1}{2}\xi(\hat{a}^\dagger)^2\right]. \tag{2.63}$$

An important squeezed state is the squeezed vacuum state, $\hat{S}(\xi)\left|0\right\rangle$. It has the following expectation values:

$$\begin{aligned}
\langle\hat{q}\rangle &:= \left\langle 0\right| \hat{S}(\xi)^\dagger \hat{q} \hat{S}(\xi) \left|0\right\rangle = 0 && (2.64)\\
\langle\hat{q^2}\rangle &:= \left\langle 0\right| \hat{S}(\xi)^\dagger \hat{q}^2 \hat{S}(\xi) \left|0\right\rangle = \frac{1}{2}[\cosh^2(r) + \sinh^2(r) - 2\cosh(r)\sinh(r)\cos(\theta)] && (2.65)\\
\langle\hat{p}\rangle &:= \left\langle 0\right| \hat{S}(\xi)^\dagger \hat{p} \hat{S}(\xi) \left|0\right\rangle = 0 && (2.66)\\
\langle\hat{p^2}\rangle &:= \left\langle 0\right| \hat{S}(\xi)^\dagger \hat{p}^2 \hat{S}(\xi) \left|0\right\rangle = \frac{1}{2}[\cosh^2(r) + \sinh^2(r) + 2\cosh(r)\sinh(r)\cos(\theta)]. && (2.67)
\end{aligned}$$

where we write $\xi = re^{i\theta}$. The uncertainty Δq is minimized when $\theta = 0$. For this value of θ, we have

$$\Delta q = \frac{1}{\sqrt{2}}e^{-r},\ \Delta p = \frac{1}{\sqrt{2}}e^{r}. \tag{2.68}$$

We can also express the squeezed vacuum state in the Fock state basis as

$$\left|\xi\right\rangle := \hat{S}(\xi)\left|0\right\rangle = \frac{1}{\sqrt{\cosh(r)}}\sum_{n=0}^{\infty}(-1)^n\frac{\sqrt{(2n)!}}{2^n n!}e^{in\theta}\tanh^n(r)\left|2n\right\rangle. \tag{2.69}$$

The probability of finding $2n$ photons is

$$p_{2n}(r) = \frac{(2n)!}{2^{2n}(n!)^2}\frac{\tanh^{2n}(r)}{\cosh(r)}, \tag{2.70}$$

and the probability of finding an odd number of photons is 0.

Two-mode squeezed vacuum state

The continuous-variable version of the maximally entangled states (EPR states) are two-mode squeezed vacuum states. They can be generated using the two-mode squeezing operator $\hat{S}_2(\xi)$.

Definition 2.3.4 (Two-mode squeezing operator)**.** Let $\hat{a}, \hat{b}$ be annihilation operators for two modes. The two-mode squeezing operator with a squeezing parameter $\xi \in \mathbb{C}$ is

$$\hat{S}_2(\xi) = \exp\Big(\xi^*\hat{a}\hat{b} - \xi\hat{a}^\dagger\hat{b}^\dagger\Big). \tag{2.71}$$

One can expand the two-mode squeezed vacuum state, $\hat{S}_2(\xi)\left|0,0\right\rangle$, in the Fock state basis as

$$\hat{S}_2(\xi)\left|0,0\right\rangle = \frac{1}{\cosh(r)}\sum_{n=0}^{\infty}(-1)^n e^{in\theta}\tanh^n(r)\left|n,n\right\rangle. \tag{2.72}$$

We can see that it has perfect photon number correlations in two modes.

2.3.3 Quasi-probability distributions

We review three types of quasi-probability distributions that are useful for our discussions. They are Wigner function, Husimi Q-function and Glauber-Sudarshan P-function. We use a general framework to define them (see [103, 104] for more discussions). They can be defined in terms of characteristic functions with the s-ordered displacement operator. In particular, the Wigner function corresponds to $s = 0$, the Q function corresponds to $s = -1$ and the P function corresponds to $s = 1$. We first define the s-parameterized quasiprobability distributions $W(\alpha; s)$ and then recover these three quasiprobability distributions as special cases. This discussion follows from [101, Section 4.2.3] with an alternative representation in terms of a complex number α instead of field quadratures q and p. It is often convenient to define the characteristic function $\widetilde{W}(\eta; s)$ first, and then the s-parameterized quasiprobability distributions $W(\alpha; s)$ can be obtained by the inverse Fourier transformation of the characteristic function. The characteristic function $\widetilde{W}(\eta; s)$ is defined as

$$\widetilde{W}(\eta; s) = \mathrm{Tr}\Big(\rho\hat{D}(\eta; s)\Big), \tag{2.73}$$

and the s-parameterized quasiprobability distributions $W(\alpha; s)$ are

$$W(\alpha; s) = \frac{1}{\pi^2}\int \widetilde{W}(\eta; s)\exp(\eta^*\alpha - \eta\alpha^*)d^2\eta, \tag{2.74}$$

where $d^2\eta = d\,\mathrm{Re}(\eta)d\,\mathrm{Im}(\eta)$. Note that the quasiprobability distribution is normalized, i.e., $\int W(\alpha; s)d^2\alpha = 1$.

Wigner function

By convention, the Wigner function is denoted by $W(\alpha)$ for $\alpha \in \mathbb{C}$. It corresponds to $s = 0$.

To calculate $\mathrm{Tr}(FG)$ for two operators F and G in terms of their Wigner functions W_F and W_G, the overlap formula is

$$\mathrm{Tr}(FG) = \pi\int d^2\alpha\; W_F(\alpha)W_G(\alpha). \tag{2.75}$$

We can easily generalize the formula to multi-mode cases. The input-output Wigner functions under a beam splitter transformation whose transmittance is η are related by

$$W_{\mathrm{out}}(\alpha, \beta) = W_{\mathrm{in}}(\sqrt{\eta}\alpha + \sqrt{1-\eta}\beta, \sqrt{1-\eta}\alpha - \sqrt{\eta}\beta). \tag{2.76}$$

We list the Wigner functions for some important quantum states below. The Wigner function of a vacuum state $|0\rangle$ is

$$W_{|0\rangle}(\gamma) = \frac{2}{\pi}e^{-2|\gamma|^2}. \tag{2.77}$$

The Wigner function of a thermal state $\rho_{\mathrm{th}}(\bar{n})$ with the mean photon number $\bar{n}$ is

$$W_{\rho_{\mathrm{th}}(\bar{n})}(\gamma) = \frac{2}{\pi}\frac{1}{1+2\bar{n}}e^{-\frac{2|\gamma|^2}{1+2\bar{n}}}. \tag{2.78}$$

The Wigner function of a displaced thermal state (DTS) $\rho_{\mathrm{DTS}}(\alpha, \bar{n}) := \hat{D}(\alpha)\rho_{\mathrm{th}}(\bar{n})\hat{D}^{\dagger}(\alpha)$ with the amount of displacement α is

$$W_{\rho_{\mathrm{DTS}}(\alpha,\bar{n})}(\gamma) = \frac{2}{\pi}\frac{1}{1+2\bar{n}}e^{-\frac{2|\gamma-\alpha|^2}{1+2\bar{n}}}. \tag{2.79}$$

We notice that if we set $\alpha = 0$, it reduces to Eq. (2.78).

It is also useful to note the Wigner functions of a squeezed thermal state (STS) and of a displaced squeezed thermal state (DSTS). Let $\hat{S}(\xi)$ denote the squeezing operator with a squeezing parameter ξ. For our discussion, we restrict $\xi \in \mathbb{R}$. For a squeezed thermal state $\rho_{\mathrm{STS}}(\xi, \bar{n}) := \hat{S}(\xi)\rho_{\mathrm{th}}(\bar{n})\hat{S}^{\dagger}(\xi)$, its Wigner function reads (see, e.g., [105, Eq. (4.13)])

$$W_{\rho_{\mathrm{STS}}(\xi,\bar{n})}(\gamma) = \frac{2}{\pi}\frac{1}{1+2\bar{n}}\exp\left\{-2[\frac{e^{2\xi}\,\mathrm{Re}(\gamma)^2 + e^{-2\xi}\,\mathrm{Im}(\gamma)^2}{1+2\bar{n}}]\right\}. \tag{2.80}$$

The Wigner function of a displaced squeezed thermal state, $\rho_{\mathrm{DSTS}}(\alpha, \xi, \bar{n}) := \hat{D}(\alpha)\hat{S}(\xi)$ $\rho_{\mathrm{th}}(\bar{n})\hat{S}^{\dagger}(\xi)\hat{D}^{\dagger}(\alpha)$, can be similarly written as

$$W_{\rho_{\mathrm{DSTS}}(\alpha,\xi,\bar{n})}(\gamma) = \frac{2}{\pi}\frac{1}{1+2\bar{n}}\exp\left\{-2[\frac{e^{2\xi}\,\mathrm{Re}(\gamma-\alpha)^2 + e^{-2\xi}\,\mathrm{Im}(\gamma-\alpha)^2}{1+2\bar{n}}]\right\}. \tag{2.81}$$

Q function

The Husimi Q function of a state can be obtained by convolution of the Wigner function of that state with the Wigner function of a coherent state. This quasi-probability distribution corresponds to $s = -1$. For a density operator ρ, its Q-function can be written as

$$Q(\alpha) = \frac{1}{\pi}\langle\alpha|\,\rho\,|\alpha\rangle\,. \tag{2.82}$$

It is normalized, i.e., $\int Q(\alpha)d^2\alpha = 1$. This quasiprobability distribution is always nonnegative for density operators. For an operator F, the expectation value of F can be calculated by

$$\mathrm{Tr}(\rho F) = \int f_F^{(A)}(\alpha)Q(\alpha)d^2\alpha, \tag{2.83}$$

where $f_F^{(A)}(\alpha)$ is obtained from the antinormally ordered operator[2] $\hat{f}_F^{(A)}(\hat{a}, \hat{a}^\dagger)$ of the operator F after replacing $\hat{a}$ by α and $\hat{a}^\dagger$ by α^*.

P function

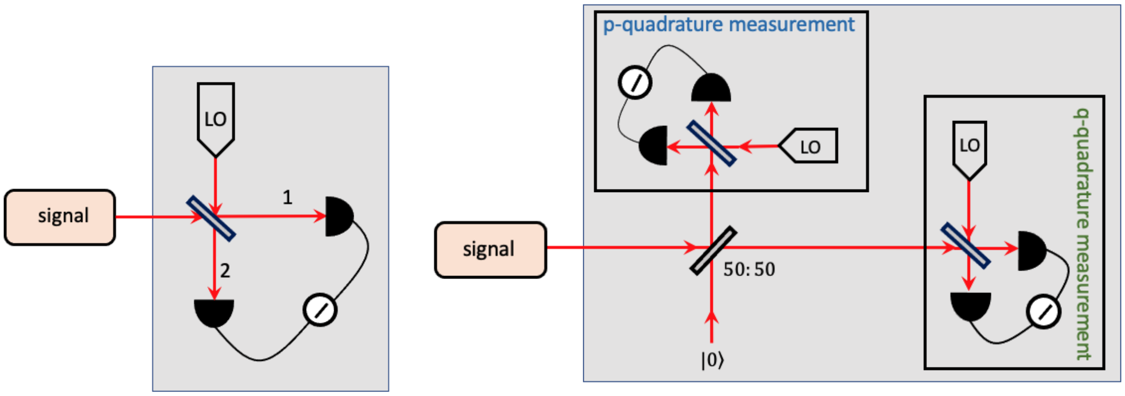
This quasi-probability distribution corresponds to $s = 1$. When the P function of a state is convolved with the Wigner function of a coherent state, it gives the Wigner function of that state. An arbitrary state can be expressed in terms of its P function as

$$\rho = \int P(\alpha) \, |\alpha\rangle\langle\alpha| \, d^2\alpha. \tag{2.84}$$

This means an arbitrary quantum state is diagonal in the coherent state representation due to the overcompleteness of the set of coherent states.

2.3.4 Coherent detection

We give a quick introduction to homodyne and heterodyne detection as they are important to understand the detection techniques used in CVQKD.



(a) Homodyne detection

(b) Heterodyne detection

Figure 2.1: Setups for (a) homodyne and (b) heterodyne measurements. Devices enclosed in each gray box represent the detector. LO stands for local oscillator.

[2] An operator is antinormally ordered if all annihilation operators are to the left of all creation operators in the product when it is written in terms of sums and products of creation and annihilation operators.

Homodyne measurements

The setup for a balanced homodyne detection is shown in Figure 2.1a. The quantum state of light to be measured is labeled as signal. The signal is mixed with a strong coherent state, called the local oscillator, on a 50/50 beam splitter. A pair of photodetectors are used to measure the intensity of the light in each output mode. The measurement outcome of interest is the difference in their detected intensities. In particular, the difference in the measured intensities in these two output modes 1 and 2 is proportional to the q_θ quadrature of the signal, where $q_\theta = q\cos(\theta) + p\sin(\theta)$ and θ is the phase of the local oscillator. By controlling the phase of the local oscillator, one may measure different quadratures of the signal.

Heterodyne measurements

A heterodyne detection is referred to a particular combination of balanced homodyne measurements described previously. Its setup is depicted in Figure 2.1b. In the CVQKD community, the meaning of heterodyne detection is conjugate homodyne detection; that is, a heterodyne detector consists of a 50/50 beam splitter and two homodyne detectors. After splitting the signal by a 50/50 beam splitter, one measures one half with q quadrature in one homodyne measurement and the other half with p quadrature in the other homodyne measurement. This setup allows us to measure both quadratures of the signal simultaneously at the cost of introducing an additional vacuum noise to the signal. The reader is cautioned that this terminology use of heterodyne detection is different from the classical communication community. Throughout this thesis, the heterodyne detection is always referred to this conjugate homodyne detection.

2.4 Introduction to convex optimization

Convex optimization has become an indispensable tool for theorists in quantum information theory to study many interesting problems. This is because many problems can be formulated as convex optimization or more specifically, semidefinite program problems. If a problem is a convex optimization problem, it means in principle there exists an efficient algorithm to solve this problem numerically.

We provide a very brief review of some key ideas in convex optimization that are needed to understand this thesis. We direct readers to [106, 107] for details.

2.4.1 Basic definitions

We review some basic definitions relevant for convex optimization.

Definition 2.4.1 (Affine hull)**.** For a set $C \subseteq \mathbb{R}^n$, the *affine hull* of C, denoted by $\mathrm{aff}(C)$, is defined as

$$\mathrm{aff}(C) = \Big\{ \sum_{i=1}^{k} \theta_i x_i : x_1, \ldots, x_k \in C, \sum_{i=1}^{k} \theta_i = 1 \text{ for some } k \in \mathbb{N} \Big\}. \tag{2.85}$$

We note that the affine hull is the smallest affine set that contains C.

Definition 2.4.2 (Relative interior)**.** The *relative interior* of a set $C \subseteq \mathbb{R}^n$, denoted by $\mathrm{relint}(C)$ is defined as

$$\mathrm{relint}(C) = \{x \in C : B_r(x) \cap \mathrm{aff}(C) \subseteq C \text{ for some } r > 0\}, \tag{2.86}$$

where $B_r(x)$ is a ball centered at x with a radius r.

Definition 2.4.3 (Convex set)**.** A set C is *convex* if for any $x_1, x_2 \in C$ and any t with $0 \leq t \leq 1$, it is the case that $tx_1 + (1-t)x_2 \in C$.

Definition 2.4.4 (Cone)**.** A set K is called a *cone* if for every $x \in K$ and $\alpha \geq 0$, it is the case that $\alpha x \in K$.

Definition 2.4.5 (Face)**.** A convex cone F is a *face* of a convex set, denoted $F \trianglelefteq K$, if

$$x, y \in K, x + y \in F \implies x, y \in F.$$

Definition 2.4.6 (Minimal face)**.** Let K be a closed convex cone and let $X \in K$. Then $\mathrm{face}(X) \trianglelefteq K$ is the *minimal face*, the intersection of all faces of K that contain X.

Definition 2.4.7 (Affine function)**.** A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is *affine* if it is a sum of a linear function and a constant, i.e., if it is of the form $f(x) = Ax + b$ for some $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$.

Definition 2.4.8 (Convex function)**.** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function with its domain $\mathrm{dom}(f)$. We say f is *convex* if $\mathrm{dom}(f)$ is a convex set and if for all $x, y \in \mathrm{dom}(f)$ and $0 \leq t \leq 1$, it is the case that

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y). \tag{2.87}$$

It is often convenient to let the function f be defined over the entire $\mathbb{R}^n$ and let it take the value $+\infty$ outside its domain. It is thus interesting to talk about the extended-value extension.

Definition 2.4.9 (Extended-value extension)**.** If f is convex, we define its *extended-value extension* $\tilde{f} : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ by

$$\tilde{f}(x) = \begin{cases} f(x) & x \in \mathrm{dom}(f) \\ +\infty & x \notin \mathrm{dom}(f). \end{cases} \tag{2.88}$$

It is sometimes useful to define the domain of an extension as

$$\mathrm{dom}(\tilde{f}) = \{x \in \mathbb{R}^n : \tilde{f}(x) < \infty\}. \tag{2.89}$$

Remark 2.4.10. Many results in standard convex optimization textbooks (e.g. [106, 107]) are stated for the Euclidean spaces over $\mathbb{R}$ (i.e., $\mathbb{R}^n$ for some n). In quantum information, we are often interested in Hermitian matrices as optimization variables. Hermitian matrices are complex matrices, but they form a vector space over $\mathbb{R}$. We note that one can always recast an optimization problem with Hermitian matrices to a problem over $\mathbb{R}^m$ for some suitable choice of $m \in \mathbb{N}$. For example, for any Hermitian matrix $H \in \mathbb{C}^{n\times n}$, we can write it as a matrix in $\mathbb{R}^{2n\times 2n}$ in the following way which preserves positive semidefinite ordering. Let $\mathrm{Re}(H)$ denote the real part of H and $\mathrm{Im}(H)$ denote its imaginary part. Then $H \geq 0$ if and only if

$$\begin{bmatrix} \mathrm{Re}(H) & -\mathrm{Im}(H) \\ \mathrm{Im}(H) & \mathrm{Re}(H) \end{bmatrix} \geq 0. \tag{2.90}$$

Moreover, for any complex vector $w = u + iv \in \mathbb{C}^n$, we have $w^\dagger H w \geq 0$ if and only if

$$\begin{bmatrix} u^T & v^T \end{bmatrix} \begin{bmatrix} \mathrm{Re}(H) & -\mathrm{Im}(H) \\ \mathrm{Im}(H) & \mathrm{Re}(H) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} \geq 0 \tag{2.91}$$

These relations imply that any optimization problem with complex vectors and Hermitian matrices (or positive semidefinite matrices) can be stated as a problem with only real vectors and symmetric matrices. Then those standard results stated for real symmetric matrices and real vectors can be directly applied.

2.4.2 Fenchel duality

Duality is a useful concept in convex optimization. We need to review one form of duality called *Fenchel duality* that is relevant for the algorithm in Section 4.4.2. See [107] for details about Fenchel duality.

Definition 2.4.11 (Fenchel conjugate)**.** The *Fenchel conjugate* of a function $h : \mathbb{R}^n \to [-\infty, +\infty]$ is the function $h^* : \mathbb{R}^n \to [-\infty, +\infty]$ defined by

$$h^*(v) = \sup_{x\in\mathbb{R}^n} \{\langle v, x\rangle - h(x)\}. \tag{2.92}$$

We now state Fenchel's duality theorem [107, Theorem 3.3.5] with the condition for strong duality replaced by Eq. (2.95) according to [107, page 74, Exercise 20(e)].

Theorem 2.4.12 (Fenchel duality, [107, Theorem 3.3.5])**.** For given functions $f : \mathbb{R}^n \to (-\infty, +\infty]$ and $g : \mathbb{R}^m \to (-\infty, +\infty]$ and a (bounded) linear map $A : \mathbb{R}^n \to \mathbb{R}^m$, let $p, d \in [-\infty, +\infty]$ be primal and dual values in the following Fenchel problems:

$$p = \inf_{x \in \mathbb{R}^n} \{f(x) + g(Ax)\}, \tag{2.93}$$

$$d = \sup_{v \in \mathbb{R}^m} \{-f^*(A^\dagger v) - g^*(-v)\}. \tag{2.94}$$

These values satisfy the weak duality $p \geq d$.

If f and g are convex and satisfy the condition

$$\operatorname{relint}(\operatorname{dom}(g)) \cap A \operatorname{relint}(\operatorname{dom}(f)) \neq \emptyset, \tag{2.95}$$

where $A \operatorname{relint}(\operatorname{dom}(f)) = \{Ax : x \in \operatorname{relint}(\operatorname{dom}(f))\}$, then the strong duality holds; i.e., $p = d$, and the supremum in the dual problem [Eq. (2.94)] is attained if $d < +\infty$.

We note that although the formulation of this theorem is for unconstrained optimization, one can easily handle constrained optimization by the extended-value extension [see Eq. (2.88)]. Similarly, one can define an indicator function for a set C by

$$\delta_C(x) = \begin{cases} 0 & \text{if } x \in C \\ \infty & \text{if } x \notin C, \end{cases} \tag{2.96}$$

and use this indicator function to convert a constraint set C into a part of the objective function.

2.4.3 Semidefinite program

Semidefinite program (SDP) is an important type of convex optimization problems. It is a special form of conic programs where the cone is the positive semidefinite cone. It plays a significant role in the theoretical study of quantum information processing tasks such as entanglement verification, coherence distillation and nonlocal games. Many important quantities like the fidelity function, trace norm and min-entropy can be formulated in terms of SDP. In Chapter 4, we discuss the formulation of key rate computation of quantum key distribution in terms of SDP and the nonlinear version of SDP. Hence, we review some basic facts about SDP. As we discuss previously, we can work with Hermitian matrices even though many standard textbooks on SDP discuss only real symmetric matrices instead of Hermitian matrices. Here we define SDP directly with Hermitian matrices, following the definition in [94][3].

[3] We state an equivalent definition instead of the exact form in [94].

Definition 2.4.13 (Semidefinite program)**.** Let $\Psi : \mathrm{L}(\mathcal{H}_A) \to \mathrm{L}(\mathcal{H}_B)$ be a Hermitian-preserving map, $A \in \mathrm{Herm}(\mathcal{H}_A)$ and $B \in \mathrm{Herm}(\mathcal{H}_B)$.

A semidefinite program is a triple (Ψ, A, B) with the following optimization problems:

$$\begin{aligned} &\text{minimize } \langle A, X\rangle \\ &\text{subject to } \Psi(X) = B \\ &\qquad X \in \mathrm{Pos}(\mathcal{H}_A) \end{aligned} \tag{2.97}$$

$$\begin{aligned} &\text{maximize } \langle B, Y\rangle \\ &\text{subject to } \Psi^\dagger(Y) \leq A \\ &\qquad Y \in \mathrm{Herm}(\mathcal{H}_B) \end{aligned} \tag{2.98}$$

where $\Psi^\dagger$ is the adjoint map of Ψ. We call Eq. (2.97) the *primal problem* and Eq. (2.98) the *dual problem*. The set $\mathcal{A} = \{X \in \mathrm{Pos}(\mathcal{H}_A) : \Psi(X) = B\}$ is called the *primal feasible set*, and the set $\mathcal{B} = \{Y \in \mathrm{Herm}(\mathcal{H}_B) : \Psi^\dagger(Y) = A\}$ is called the *dual feasible set*. We denote the primal optimal value as α and the dual optimal value as β.

Remark 2.4.14. Since SDP is a special form of convex optimization problems, by weak duality, it is the case that $\alpha \geq \beta$.

Theorem 2.4.15 (Slater's condition)**.** For a semidefinite program (Ψ, A, B), if $\mathcal{A} \neq \emptyset$ and there exists a Hermitian operator Y that is strictly dual feasible, i.e., $\Psi^\dagger(Y) < A$, then the strong duality holds for this problem, i.e., $\alpha = \beta$, and the optimal value is obtained in the primal problem.

Sometimes, it is convenient to express the SDP problem in another common form, which is equivalent to the standard form presented in Definition 2.4.13. A useful alternative form is given by

$$\begin{aligned} &\underset{X}{\text{minimize}} \langle A, X\rangle \\ &\text{subject to } \langle B_i, X\rangle = b_i,\ i = 1, \dots, n \\ &\qquad X \in \mathrm{Pos}(\mathcal{H}_A) \end{aligned} \tag{2.99}$$

$$\begin{aligned} &\underset{y_1,\dots,y_m}{\text{maximize}} \sum_{i=1}^m b_i y_i \\ &\text{subject to } \sum_{i=1}^m y_i B_i \leq A \\ &\qquad y_i \in \mathbb{R},\ i = 1, \dots, m\ . \end{aligned} \tag{2.100}$$

2.5 Introduction to quantum key distribution

2.5.1 Generic QKD protocols

To facilitate our discussions, we present generic (device-dependent) QKD protocols. Different protocols fit in with the generic description after some modifications of details. Roughly speaking, there are two major structures of protocols: prepare-and-measure and entanglement-based.

We present a prepare-and-measure scheme in detail and note how an entanglement-based scheme differs from it in the first two steps.

Protocol 2.1 Generic prepare-and-measure QKD protocol

(1) *State preparation and transmission:* Alice prepares a state from an ensemble $\{|\varphi_x\rangle, p_x : x \in \Sigma_A\}$ and sends it to Bob. She may label states $|\varphi_x\rangle$ as $|\varphi_{(a,\alpha)}\rangle$.

(2) *Measurement:* Bob receives the state in his register B, and applies his POVM $\{M_y^B : y \in \Sigma_B\} =: \{M_{(b,\beta)}^B\}$ to measure this state.

We call each repetition of the first two steps as one round. After N rounds, they continue to the following classical post-processing part of the protocol.

(3) *Parameter estimation:* Alice and Bob randomly select a subset of rounds to perform parameter estimation. Let m denote the number of rounds selected. For those rounds, they disclose their measurement outcomes (x, y) to construct the frequency distribution $F(x, y)$. If $F(x, y)$ is in the set of acceptable frequency distributions $\mathcal{Q}$ which they agree before the execution of the protocol, they continue. Otherwise, they abort the protocol.

(4) *Data partitioning and announcement:* They partition their respective data from each signal preparation and measurement, indexed by the round i, into private information α, β to store in registers $\overline{A}_i, \overline{B}_i$, and public information a, b in registers $\widetilde{A}_i, \widetilde{B}_i$. Alice and Bob announce the public information.

(5) *General sifting:* Alice and Bob throw away some rounds of the remaining $N - m$ signals according to the public announcement. Let n denote the number of remaining signals. The remaining private information forms their raw data strings $\tilde{\boldsymbol{x}} \in \{0, \dots, k_A - 1\}^n$ and $\tilde{\boldsymbol{y}} \in \{0, \dots, k_B - 1\}^n$, respectively, where k_A and k_B are the number of possible outcomes for Alice's and Bob's private information, respectively.

(6) *Key map:* Alice applies a key map, which is a function of her private data $\tilde{\boldsymbol{x}}$ as well as the public information of both parties to obtain a key string $\boldsymbol{z} \in \{0, 1, \dots, d - 1\}^n$, where d is the size of the alphabet for the key.

(7) *Error correction and error detection:* Alice and Bob use the public communication to try to match $\tilde{\boldsymbol{y}}$ with $\boldsymbol{z}$. In the end, Bob obtains $\boldsymbol{z'} \in \{0, 1, \ldots, d-1\}^n$. Then, they use the public communication to choose a (two-universal) hash function at random and apply it to each of $\boldsymbol{z}$ and $\boldsymbol{z'}$. They compare their hash values via the public communication. If their hash values disagree, they either redo the error correction or abort the protocol.

(8) *Privacy amplification:* By using a two-universal hash function on their key strings $\boldsymbol{z}$ and $\boldsymbol{z'}$, they obtain their final keys K_A and K_B, respectively. Let ℓ denote the length of the final key K_A.

Remark 2.5.1. In the parameter estimation step, to choose rounds for testing, there are two possible approaches. They may choose these rounds in a probabilistic way: For each round, they select it for testing with a probability $\gamma > 0$. Then the number of rounds selected is a random variable with its expectation value $N\gamma$. The other possibility is that they select randomly a fixed number of rounds, m. For example, they may apply a random permutation to all rounds and then choose the first m rounds for testing. However, if the number of signals N is large, implementing this random permutation can be computationally challenging for the classical post-processing units. Moreover, they may perform some additional data processing on $F(x, y)$ to obtain coarse-grained statistics in some cases if it is advantageous to do so. In some specific QKD protocols, one may postpone the parameter estimation step until after the error correction step so that the parameter estimation step can also use the information announced in the error correction step.

Remark 2.5.2. In the key map step, we present the direct reconciliation scheme by letting Alice apply the key map. Then in the error correction step, Bob needs to correct his string to match with Alice's. This direct reconciliation scheme is commonly used in discrete-variable protocols. On the other hand, if Bob applies a key map to his string and then Alice corrects her string to match with Bob's, this reconciliation procedure is called the reverse reconciliation. The reverse reconciliation scheme is more commonly used in continuous-variable protocols to overcome 3 dB loss limit for transmission.

To prove the security of a QKD protocol, it is often easier to work with an entanglement-based scheme. We describe a generic entanglement-based protocol, which only differs from a prepare-and-measure protocol in the first two steps. As we will see in Section 2.5.2, a prepare-and-measure protocol can be converted as an entanglement-based protocol by the source-replacement scheme [5, 108–110].

Protocol 2.2 Generic entanglement-based QKD protocol

---

(1) *State preparation and transmission:* Alice (or Eve) prepares a bipartite quantum state ρ_{AB}. Alice keeps the system A and sends the system B to Bob.

(2) *Measurement:* Alice applies her POVM $\{M^A_x : x \in \Sigma_A\} =: \{M^A_{(a,\alpha)}\}$ to her register A and obtains the measurement outcome x.

Similarly, Bob applies his POVM $\{M^B_y : y \in \Sigma_B\} =: \{M^B_{(b,\beta)}\}$ to his register B.

The rest of the protocol is identical to Protocol 2.1.

The protocols presented so far are conventional QKD schemes. Another family of protocols to be discussed in thesis are the MDI protocols [57, 58]. MDI protocols can be described as either entanglement-based or prepare-and-measure schemes. We highlight first three steps in the prepare-and-measure version.

Protocol 2.3 Generic measurement-device-independent QKD protocol

(1) *State preparation and transmission:* Alice prepares a state from an ensemble $\{|\varphi_{A,x}\rangle, p_A(x) : x \in \Sigma_A\}$ and sends it to an untrusted third-party Charlie. Similarly, Bob prepares a state from an ensemble $\{|\varphi_{B,y}\rangle, p_B(y) : y \in \Sigma_B\}$ and sends it to Charlie.

(2) *Measurement:* After receiving the states sent by Alice and Bob, Charlie applies some measurements and announce the measurement results to both Alice and Bob.

After N repetition of the first two steps, Alice and Bob continue to the classical post-processing part of the protocol.

(3) *Parameter estimation:* This step is almost identical to that in Protocol 2.1. The difference is that they may partition their data into several sets according to Charlie's announcement outcomes. Then they may process each set of data independently. If one set of data fails the test, they throw away. If all sets fail, they abort the protocol. Otherwise, they continue.

The rest of the protocol is identical to Protocol 2.1.

Similarly to conventional QKD protocols, one can obtain equivalent entanglement-based schemes by the source-replacement scheme discussed in Section 2.5.2.

2.5.2 Source-replacement scheme

The source-replacement scheme [5, 108–110] can be used to obtain an equivalent entanglement-based scheme for the given prepare-and-measure protocol.

Given Alice's state ensemble $\{|\varphi_x\rangle, p_x\}$ for her preparation in the prepare-and-measure scheme, Alice effectively prepares a bipartite state $|\Psi\rangle_{AA'}$ in the source-replacement scheme, which is defined as

$$|\Psi\rangle_{AA'} = \sum_x \sqrt{p_x}\, |x\rangle_A |\varphi_x\rangle_{A'}\,, \tag{2.101}$$

where $\{|x\rangle\}$ is an orthonormal basis for register A. Then Alice sends register A' to Bob via an insecure quantum channel and keeps register A for her measurement described by the POVM $M^A = \{M_x^A = |x\rangle\langle x|\}$. The quantum channel that maps register A' to Bob's register B is described by a CPTP map, $\mathcal{E}_{A'\to B}$, and is assumed to be under Eve's control. Thus, Alice and Bob's joint state ρ_{AB} before their measurements is

$$\rho_{AB} = (\mathrm{id}_A \otimes \mathcal{E}_{A'\to B})(|\Psi\rangle\langle\Psi|_{AA'}), \tag{2.102}$$

where id_A is the identity channel on Alice's system A.

When Alice performs a local measurement using her POVM $\{M_x^A\}$ on register A and obtains an outcome x, she effectively sends the state $|\varphi_x\rangle$ to Bob. Bob's received state ρ_B^x conditioned on Alice's choice of x is

$$\rho_B^x = \frac{1}{p_x} \mathrm{Tr}_A[\rho_{AB}(|x\rangle\langle x|_A \otimes \mathbb{1}_B)]. \tag{2.103}$$

Bob applies his POVM $M^B = \{M_y^B\}$ to register B to obtain his measurement outcomes.

To establish an equivalence between the original prepare-and-measure scheme and the entanglement-based scheme obtained by the source-replacement scheme for the security proof, it is important to note that Alice's reduced density operator ρ_A is unchanged by Eve since her system A never goes through Eve's domain. The reduced density operator ρ_A is given as

$$\rho_A = \sum_{x,x'} \sqrt{p_x p_{x'}}\, \langle\varphi_{x'}|\varphi_x\rangle\, |x\rangle\langle x'|_A\,. \tag{2.104}$$

Chapter 3

Security proof methods

The aim of this chapter is to provide an overview of several existing security proof methods for general QKD protocols. In this chapter we start with the security definition of QKD and then present several methods to prove the security based on Renner's framework [89]. In particular, under this framework, the security proofs are modular in the sense that one can prove the security against a class of restricted attacks, known as collective attacks (which is explained in Section 3.2), and then lift the security proof to general attacks by techniques like quantum de Finetti theorem [89, 111], postselection technique [112] or the entropy accumulation theorem [97, 98]. We briefly review these techniques. At the end of this chapter, we also provide remarks about an alternative framework - phase error correction method and briefly mention the security proof method based on entropic uncertainty relations.

3.1 Security definition of QKD

In the cryptographic setting of QKD, Eve may hold a quantum register E that is in some state ρ_E after the quantum communication of a QKD session. She also obtains a copy of the register that contains all the classical announcements since she can listen to all the classical communication. Her goal is to guess the final key string K produced from this QKD protocol or the secret information in the applications that use the key K. However, she is not forced to measure her system immediately. She can perform any measurements that are physically allowed on her system at any time of her choosing. Secret keys generated from a QKD protocol are supposed to be consumed by other cryptographic applications such as one-time pad encryption and digital signature schemes. This means that she may delay her measurements until the key is being used in those cryptographic applications. Hence, it is desirable that secure cryptographic applications which need perfectly secure keys remain secure if keys are supplied by QKD. This requirement is related to the notion of universal composability. It is also worth noting that it is unlikely to generate perfectly secure keys in practical protocols. As we can see from the well-known BB84 protocol based on qubits [3], there is a nonzero probability (which decays exponentially in the number of signals) that Eve guesses Bob's basis correctly in each round and then learns every single bit of the key without introducing any disturbance. Since this probability is negligible when Alice and Bob exchange a sufficiently large number of signals, it is reasonable to consider keys from a protocol like this to be secure. On the other hand, users should be aware of the risks associated with using such a key. Thus, it is also important to quantify the risk via a security parameter that has operational meanings. All these requirements are satisfied by the security definition proposed in [113, 114], which achieves the goal of composable security.

Under the universally composable security framework, one often compares the real protocol with an idealized protocol and quantifies the deviation from the ideal protocol. Both real and ideal protocols can abort. In this case, Alice and Bob both output a symbol $\perp$ to indicate the failure of the protocol. If the protocol does not abort, the ideal protocol for QKD is supposed to generate an identical key for Alice and Bob from the key space such that the key is uniformly distributed among all possible choices of keys and that Eve knows nothing about the key. The trace distance is used as the measure for deviation, which is a generalization of the variational distance used for classical systems. This security definition has been widely used in the literature. We state this definition here; see also Renner's Ph.D. thesis [89] as well as [115] .

Definition 3.1.1 (ε-security of QKD)**.** A QKD protocol consists of a state distribution process that establishes a tripartite state ρ_{ABE} among Alice, Bob and Eve, and a key distillation protocol which is described by a completely positive and trace non-increasing map $\mathcal{E}_{ABE\rightarrow K_AK_BE'}$. A QKD protocol is ε-secure if for any input state ρ_{ABE} to the key distillation protocol, the trace distance between the (subnormalized) output state $\rho_{K_AK_BE'} = \mathcal{E}_{ABE\rightarrow K_AK_BE'}(\rho_{ABE})$ and the ideal state $\omega_{K_AK_B} \otimes \rho_{E'}$ is at most ε:

$$\frac{1}{2}\left\|\rho_{K_AK_BE'} - \omega_{K_AK_B} \otimes \rho_{E'}\right\|_1 \leq \varepsilon, \tag{3.1}$$

where $\omega_{K_AK_B} = \frac{1}{|\mathcal{K}|}\sum_{k\in\mathcal{K}} |k\rangle\langle k|_{K_A} \otimes |k\rangle\langle k|_{K_B}$ is a perfectly classically correlated state, $\{|k\rangle\}$ is an orthonormal basis for each of K_A and K_B, $\rho_{E'} = \mathrm{Tr}_{K_AK_B}(\rho_{K_AK_BE'})$, and $\mathrm{Tr}(\rho_{K_AK_BE'}) = \mathrm{Tr}(\rho_{E'})$ is the probability of not aborting the protocol. Here, each $|k\rangle$ represents a possible key string in the key space $\mathcal{K}$.

In this definition, we use subnormalized states and the trace of each subnormalized state is the probability of not aborting. We notice that whenever the protocol aborts, Alice and Bob obtain a trivial key in both real and ideal protocol so that the trace distance between the real and ideal state is zero in this case. One can write any such subnormalized output state $\rho_{K_A K_B E'}$ of the real protocol conditioned on not aborting in the following generic form:

$$\rho_{K_A K_B E'} = \sum_{k,k' \in \mathcal{K}} p(k,k') \, |k\rangle\langle k|_{K_A} \otimes |k'\rangle\langle k'|_{K_B} \otimes \rho_{E'}^{(k,k')}, \tag{3.2}$$

where $\rho_{E'}^{(k,k')}$ is Eve's (normalized) state conditioned on the choice (k,k'). The state $\omega_{K_A K_B} \otimes \rho_{E'}$ in this security definition is the state given by an ideal protocol. It has the property of correctness $(k = k')$, and satisfies the uniform randomness requirement $(p(k,k') = \frac{1}{|\mathcal{K}|}\delta_{k,k'})$. Moreover, Eve's conditional state is independent of any particular value of k, which is the secrecy requirement.

Alternatively, if we define in terms of normalized output states $\tilde{\rho}_{K_A K_B E'} = \frac{1}{\mathrm{Tr}\left(\rho_{K_A K_B E'}\right)}\rho_{K_A K_B E'}$ and $\tilde{\rho}_{E'} = \frac{1}{\mathrm{Tr}(\rho_{E'})}\rho_{E'}$, then Eq. (3.1) becomes

$$(1 - p_{\text{abort}})\frac{1}{2}\|\tilde{\rho}_{K_A K_B E'} - \omega_{K_A K_B} \otimes \tilde{\rho}_{E'}\|_1 \leq \varepsilon, \tag{3.3}$$

where p_{abort} denotes the probability of not aborting.

We note that the trace distance of two states has an operational interpretation: the advantage in distinguishing these two states, which are given with equal *a priori* probability, compared to a random guessing. This means the ε-security definition of QKD can be interpreted as follows. Suppose two black-box devices, one that implements a real QKD protocol and the other one that implements an ideal protocol, are given to an adversary whose goal is to verify which device is for the real QKD protocol. The adversary can send any input to each of the devices and can examine the output states. Then the probability that this adversary guesses correctly which device implements the real protocol by looking at output states is at most $\frac{1}{2}(1+\varepsilon)$. Alternatively, based on Eq. (3.3), one can say that the probability of the event that Alice and Bob do not abort the protocol and that the key K is not perfect (i.e., random, correct and secret) is upper bounded by ε.

The security parameter ε can usually be subdivided into two parts by the use of the triangle inequality: correctness ε_{cor} and secrecy ε_{sec}.

Definition 3.1.2 (ε_{cor}-correctness)**.** Let $\varepsilon_{\text{cor}} \geq 0$. A QKD protocol is ε_{cor}-correct if for key strings K_A and K_B that are chosen according to the distribution defined by $\rho_{K_A K_B}$, it is the case that $\Pr(K_A \neq K_B \wedge \neg\text{abort}) \leq \varepsilon_{\text{cor}}$.

The correctness of a QKD protocol is usually guaranteed by the error correction subprotocol.

Definition 3.1.3 (ε_{sec}-secrecy)**.** For any $\varepsilon_{\text{sec}} \geq 0$, a key K is ε_{sec}-secret with respect to an adversary who holds a register E that captures her knowledge about the key K if the joint state ρ_{KE} satisfies

$$\frac{1}{2}\|\rho_{KE} - \tau_K \otimes \rho_E\|_1 \leq \varepsilon_{\text{sec}}, \tag{3.4}$$

where $\tau_K = \frac{1}{|\mathcal{K}|}\sum_{k\in\mathcal{K}} |k\rangle\langle k|_K$ is the completely mixed state on K.

As we discuss in more details in Section 3.3.1, the secrecy of a QKD protocol often comes from the two-universal family of hash functions used for the privacy amplification subprotocol.

If a QKD protocol is ε_{cor}-correct and outputs an ε_{sec}-secret key, then the QKD protocol is ε-secure where $\varepsilon = \varepsilon_{\text{cor}} + \varepsilon_{\text{sec}}$. This can be easily seen from the triangle inequality of trace norm:

$$\begin{aligned}\frac{1}{2}\|\rho_{K_AK_BE'} - \omega_{K_AK_B} \otimes \rho_{E'}\|_1 &\leq \frac{1}{2}\|\rho_{K_AK_BE'} - \tilde{\rho}_{K_AK_BE'}\|_1 + \frac{1}{2}\|\tilde{\rho}_{K_AK_BE'} - \omega_{K_AK_B} \otimes \rho_{E'}\|_1\\ &\leq \frac{1}{2}\|\rho_{K_AK_B} - \tilde{\rho}_{K_AK_B}\|_1 + \frac{1}{2}\|\tilde{\rho}_{K_AE'} - \omega_{K_A} \otimes \rho_{E'}\|_1\\ &= \Pr(K_A \neq K_B \wedge \neg\text{abort}) + \frac{1}{2}\|\tilde{\rho}_{K_AE'} - \omega_{K_A} \otimes \rho_{E'}\|_1\\ &= \varepsilon_{\text{cor}} + \varepsilon_{\text{sec}} = \varepsilon,\end{aligned} \tag{3.5}$$

where $\tilde{\rho}_{K_AK_BE'} = \sum_{k\in\mathcal{K}} p(k)\,|k\rangle\langle k|_{K_A} \otimes |k\rangle\langle k|_{K_B} \otimes \rho_{E'}^{(k)}$ and we use the fact that the trace distance is non-decreasing under the partial trace.

In some scenarios (e.g., see Section 3.6), it is also useful to consider a similar, related definition.

Definition 3.1.4 (ε-security of QKD in terms of diamond norm)**.** Let N be a given positive integer. Let $\mathcal{E}^{\text{real}} : \mathrm{L}((\mathcal{H}_A \otimes \mathcal{H}_B)^{\otimes N}) \to \mathrm{L}(\mathcal{H}_{K_A} \otimes \mathcal{H}_{K_B})$ be the completely positive, trace non-increasing map that represents the real QKD protocol. Let $\mathcal{E}^{\text{ideal}}$ be the corresponding CPTNI map that represents the ideal QKD protocol. The real QKD protocol is ε-secure if

$$\frac{1}{2}\left\|\mathcal{E}^{\text{real}} - \mathcal{E}^{\text{ideal}}\right\|_\diamond \leq \varepsilon. \tag{3.6}$$

We note that the ε-security under this alternative definition implies the ε-security under the Definition 3.1.1. It is because the diamond norm involves a maximization over all possible valid input states and also $(\mathcal{E}^{\text{ideal}} \otimes \mathrm{id}_E)(\sigma) = \omega_{K_AK_B} \otimes \rho_{E'}$ for any input state $\sigma \in \mathrm{D}_{\leq}(\mathcal{H}_{AB}^{\otimes N} \otimes \mathcal{H}_E)$.

We end this section by mentioning some history. A security definition of QKD based on Eve's accessible information about the key K was used in early works prior to [113, 114]. It was later shown that small accessible information does not imply universally composable security [116]. While early security proofs of QKD were based on this obsolete security definition, the key rate formula in the asymptotic limit was redeemed under the composable security definition.

3.2 Assumptions and eavesdropping strategies

On one hand, the security of QKD protocols is not based on computational assumptions used in classical cryptography. Eve may have unlimited computational powers including both classical and quantum computers. As long as she is bound by the laws of quantum mechanics, she is even assumed to have access to any advanced future technologies that are physically allowed. On the other hand, there are still assumptions in security proofs. For example, one may assume there exist some protected spaces (like Alice's laboratory) where Eve cannot access. One may also make assumptions on how Alice's and Bob's devices function. Before we proceed with a security proof, it is important to state all the assumptions. Some assumptions become requirements for experimental implementations of a QKD protocol. If an implementation fails to satisfy the assumptions in a security proof, then the given security proof is no longer applicable. If one ignores the differences between assumptions in a security proof and an experimental implementation and still applies that security proof to calculate secret key rates, it opens up side channels. These side channels can be explored by Eve. Side-channel attacks can lead to the breach of security of the QKD system. Quantum hackers for QKD search for gaps between theory and experiment, design eavesdropping strategies to explore side channels and propose countermeasures to prevent side-channel attacks.

Different assumptions on Alice's and Bob's devices usually also lead to variations in the protocol design. According to different levels of assumptions on Alice's and Bob's devices, there are three major types of QKD protocols: device-dependent, semi-device-independent and device-independent.

1. Device-dependent: Both Alice's and Bob's devices are fully characterized and trusted. Its security proof depends on the characterization of devices.

2. Semi-device-independent: One party's device is fully characterized and trusted while the other party's device can be uncharacterized and untrusted. Security proofs may use some entropic uncertainty relation to bound unknown measurement outcomes by the known one.

3. Device-independent: It is not required to characterize and trust Alice's and Bob's devices. They need to perform self-testing of the devices to certify sufficient quantum correlations for secret key generation, usually by violation of some Bell's inequalities. The security

proofs may still require some behaviors of the devices. For example, some proofs require the devices have no memory between different sessions.

We note that MDI QKD protocols are in the category of device-dependent protocols. In MDI protocols, Alice and Bob each have a source that is trusted and characterized. The measurement device is located in the third-party Charlie's domain and is untrusted. Eve can impersonate Charlie and perform the measurements.

In this thesis, we focus only on device-dependent QKD protocols. For the device-dependent protocols, according to structure of the protocol, we can put these protocols into two schemes: entanglement-based and prepare-and-measure schemes. The device assumptions of these two schemes as as follows:

1. Entanglement-based: Both Alice and Bob perform measurements on the state received. There is a source that is supposed to distribute bipartite entangled states. This source may be located in Alice's laboratory or in Eve's domain. If the source is assumed to be located in Alice's laboratory, then one may trust the source. Otherwise, the source is untrusted. Both Alice's and Bob's measurement devices are trusted and characterized.

2. Prepare-and-measure: Alice has a source that prepares states for Bob. Bob performs measurements. Both Alice's source and Bob's measurement devices are characterized and trusted.

Assumptions on the devices require experimental validations. On the protocol level, a security proof is typically done with a given mathematical description of devices. As one can typically draw the boundary between trusted parties and adversaries, these assumptions do not unreasonably restrict Eve's abilities. On the other hand, there are situations where we would like to make assumptions on possible eavesdropping strategies. These assumptions may or may not be reasonable. However, it turns out that in many practical cases, some additional assumption may help greatly simplify the security proof and can then be removed by some technical tools that essentially lift the previous proof to the general cases. Therefore, it is worth discussing these eavesdropping strategies. There are two major classes of eavesdropping strategies:

1. Collective attacks[4]: Eve interacts with each individual signal that goes through her domain with the same strategy. For each round, she may attach one ancillary register to each transmitted signal, perform a joint unitary operation on both the signal and her ancillary system. She stores all her ancillary registers in a quantum memory and then performs a

[4] The definition of collective attacks does not necessarily lead to the i.i.d. assumption since not all protocols use an i.i.d. source and perform i.i.d. measurements. For this reason, a more precise term for the main use of the terminology *collective attacks* in literature is i.i.d. collective attacks.

collective measurement on all of them at a time of her choosing. She may postpone her measurements until after listening to all the classical communication between Alice and Bob since she can use any additional information learned in the classical communication to help her decide how to perform her measurements on her ancillary systems. Under this type of eavesdropping strategies, when the source in a QKD protocol emits independent and identically distributed (i.i.d.) states and Alice's and Bob's interactions with the signals are also i.i.d., one can assume the joint state ρ^N_{ABE} among Alice, Bob and Eve after N rounds has an i.i.d. structure, i.e., $\rho^N_{ABE} = \sigma^{\otimes N}_{ABE}$ for some single-copy state σ_{ABE}.

2. Coherent attacks: They are the most general type of attacks. Instead of interacting with each signal individually, Eve may have one giant ancillary system that is attached to all the signals and then perform a joint unitary operation among her ancillary register and all the signals. She can then measure her ancillary register at any time of her choosing, possibly after listening to all the classical communication or even after the current QKD session is finished. There is no assumption about the structure of the joint state ρ^N_{ABE} after N rounds.

While one needs to prove the security against the general attacks - coherent attacks, this task is usually daunting to handle directly. One then considers a simplified problem by assuming collective attacks. Historically, individual attacks, a weaker type of eavesdropping strategy than both collective and coherent attacks, were also investigated when it was still challenging to prove security against collective attacks. While it helped to gain intuitions about potential performances of the protocol, the restriction to individual attacks is often unjustifiable and would lead to over-optimistic key rates, that is, insecure keys can be generated. On the other hand, it has been shown that collective attacks are asymptotically optimal for many protocols. Thus, when studying the asymptotic behavior of a protocol, it is often sufficient to assume collective attacks as long as the protocol satisfies some requirement (e.g. permutation invariant among different rounds). Of course, a realistic scenario is when the number of signals is finite. Fortunately, security proof techniques often allow one to lift a proof from collective attacks to coherent attacks in the finite-size regime by some changes of parameters or some small cost in the key rate, keeping the essential parts of the proof the same. We discuss these techniques in the later parts of this chapter.

3.3 Overview of Renner's security proof framework

Since the error correction (and error detection) subprotocol essentially guarantees the ε_{cor}-correctness of any QKD protocol, to prove the ε-security of the QKD protocol, the main task is to prove the ε_{sec}-secrecy of the output key from the protocol. The ε_{sec}-secrecy definition requires us to bound the trace distance between the output classical-quantum state ρ_{KE} of the real protocol and the output state of the ideal protocol. We focus on Renner's framework to achieve this task, which is fundamentally based on the leftover hashing lemma. We review several key ideas from Renner's Ph.D. thesis [89]. In particular, we look at each protocol step in the reverse order, and explain how the task of bounding the trace distance from the QKD security definition can be reduced to the calculation of some entropic quantities. Following the notation used in the generic protocol description (Section 2.5.1), in this section, we use $\boldsymbol{Z}$ to denote the raw key string (of some length n) after the key map step and let $\mathcal{Z}$ to denote the set of all possible values of the raw key string. We use $\mathcal{K}$ to denote the space of all possible final keys after an execution of the protocol. We adapt all statements to these notations.

3.3.1 Privacy amplification: leftover hashing lemma

The leftover hashing lemma establishes a relation between the amount of extractable secret keys from a string and the smooth min-entropy of the string given Eve's (quantum) side information. The essential part of the leftover hashing lemma is to use two-universal family of hash functions.

Definition 3.3.1 (Two-universal family of hash functions)**.** Let $\mathcal{Z}$ and $\mathcal{K}$ be two alphabets. Let $\mathcal{F}$ be a family of functions from $\mathcal{Z}$ to $\mathcal{K}$ and let $P_{\mathcal{F}}$ be a probability distribution[5] on $\mathcal{F}$. The pair $(\mathcal{F}, P_{\mathcal{F}})$ is called *two-universal* if for any distinct $z, z' \in \mathcal{Z}$, the probability that $f(z) = f(z')$, when f is chosen randomly from $\mathcal{F}$ according to the distribution $P_{\mathcal{F}}$, is at most $\frac{1}{|\mathcal{K}|}$, i.e.,

$$\Pr_{f\in\mathcal{F}}[f(z) = f(z')] \leq \frac{1}{|\mathcal{K}|}. \tag{3.7}$$

As seen in a generic QKD protocol in Section 2.5.1, the last step of the protocol is the privacy amplification step. If two-universal hash functions are used to achieve the task of privacy amplification, then the leftover hashing lemma establishes a bound on the trace distance between ideal key and the key from the real protocol. The statement of leftover hashing lemma is from [89, Corollary 5.6.1].

Theorem 3.3.2 (Leftover hashing lemma, [89, Corollary 5.6.1])**.** Let $\rho_{\boldsymbol{Z}E'} \in \mathrm{D}(\mathcal{H}_{\boldsymbol{Z}} \otimes \mathcal{H}_{E'})$ be classical with respect to an orthonormal basis $\{|z\rangle\}_{z\in\mathcal{Z}}$. Let $\mathcal{F}$ be a two-universal family of hash functions from the alphabet $\mathcal{Z}$ to $\{0,1\}^{\ell}$ for some non-negative integer ℓ and let $\varepsilon' \geq 0$. Then,

$$\frac{1}{2}\left\|\rho_{F(\boldsymbol{Z})E'F} - \frac{1}{|\{0,1\}^{\ell}|}\mathbb{1}_{F(\boldsymbol{Z})} \otimes \rho_{E'F}\right\|_1 \leq \varepsilon' + 2^{-\frac{1}{2}(H^{\varepsilon'}_{\min}(Z|E')-\ell-2)}, \tag{3.8}$$

where F is the register that stores the choice of hash function f chosen from $\mathcal{F}$ and $F(\boldsymbol{Z})$ is the register that stores the result of $f(\boldsymbol{Z})$.

[5] The probability distribution $P_{\mathcal{F}}$ is typically a uniform distribution.

We note that in Renner's Ph.D. thesis [89] from 2005, the ε'-ball $B_{\varepsilon'}(\rho)$ in the definition of smooth min-entropy (see Definition 2.2.47) is defined in terms of trace distance. More recent works usually define it in terms of the purified distance. Because the purified distance provides an upper bound on the trace distance (see Lemma 2.1.16), the statement of the leftover hashing lemma is unchanged when we use the purified distance instead of the trace distance.

To use this theorem, we set the target secrecy parameter ε_{sec} to be $\varepsilon' + 2^{-\frac{1}{2}(H_{\min}^{\varepsilon'}(\boldsymbol{Z}|E')-\ell-2)}$. Conventionally, we introduce a security parameter called ε_{PA} and set ε_{PA} as

$$\varepsilon_{\text{PA}} = 2^{-\frac{1}{2}(H_{\min}^{\varepsilon'}(\boldsymbol{Z}|E')-\ell-2)}. \tag{3.9}$$

Solving ℓ in terms of ε_{PA} and the smooth min-entropy leads to the the following corollary, which is stated in [117, Eq. (11)].

Corollary 3.3.3 ([117, Eq. (11)]). Let $\rho_{\boldsymbol{Z}E'} \in \mathrm{D}(\mathcal{H}_{\boldsymbol{Z}} \otimes \mathcal{H}_{E'})$ and $\mathcal{F}$ be the same as in Theorem 3.3.2. Let $\varepsilon' > 0$. Let $\varepsilon_{\text{PA}} := 2^{-\frac{1}{2}(H_{\min}^{\varepsilon'}(\boldsymbol{Z}|E')-\ell-2)}$. Then, the key obtained after applying a hash function f chosen from $\mathcal{F}$ is $\varepsilon_{\text{sec}} := \varepsilon' + \varepsilon_{\text{PA}}$-secret if the length of the key, ℓ, satisfies the following condition:

$$\ell \leq H_{\min}^{\varepsilon'}(\boldsymbol{Z}|E') - 2\log_2(\frac{2}{\varepsilon_{\text{PA}}}). \tag{3.10}$$

Note that the second term in [117, Eq. (11)] is $-2\log_2(1/\varepsilon_{\text{PA}})$, which can overestimate the key length ℓ by two extra bits due to a typographical error. This typographical error is corrected here.

From Eq. (3.10), our task to prove the security is reduced to obtaining a (tight) lower bound on the smooth min-entropy $H_{\min}^{\varepsilon'}(\boldsymbol{Z}|E')$ for some $\varepsilon' > 0$.

3.3.2 Error correction and error detection

Alice and Bob have some partially correlated strings before they reach the error correction step. The task of error correction is to correct errors in one party's string to match with the other party's key string. Without loss of generality, we take the direct reconciliation scheme as an example. In this case, Alice reveals some information about her key string through public communication and Bob needs to correct his string to match with Alice's. Some error correcting codes only require communication in one direction such as low-density parity check (LDPC) codes [118] and polar codes [119]. Other codes can require multiple-round two-way communications. We define the total number of bits needed in communication for the error correction purpose as leak_{EC}.

To achieve higher efficiency, error correcting codes are usually designed to correct errors with a high probability instead of correcting errors perfectly. Therefore, the error correction step can fail to produce two identical key strings. Even if Alice's and Bob's key strings are slightly different, their final keys after the privacy amplification step will then be dramatically different. Since it is much easier to detect errors, they can use an error detection step after the error correction subprotocol to minimize the failure probability of the overall error correction and error detection step. To do so, they choose a two-universal hash function at random by communicating through the authenticated classical channel, and then compute the hash values of their strings. If their hash values disagree, they either redo the error correction subprotocol or abort the QKD protocol. We note that this error detection step can also fail due to collision of hash functions. However, this failure probability of error detection is much smaller than the failure probability of error correction alone. In the literature, one typically introduces a security parameter ε_{EC}, which is actually the failure probability of the error detection if error detection is performed. To guarantee the QKD protocol is ε_{cor}-correct, we demand ε_{EC} to be ε_{cor}. From Definition 3.3.1, one needs to choose the length of hash to be at least $\lceil \log_2(1/\varepsilon_{\text{EC}}) \rceil \leq \log_2(1/\varepsilon_{\text{EC}}) + 1 = \log_2(2/\varepsilon_{\text{EC}})$.

In the expression of the smooth min-entropy $H^{\varepsilon'}_{\text{min}}(\boldsymbol{Z}|E')$, we note that $\boldsymbol{Z}$ is the raw key string of length n before the privacy amplification step, and the system E' contains not only Eve's initial system E after the quantum communication phase, but also public announcement registers including the register that stores transcripts of the error correction communication. Let C' denote the register that stores communication transcripts of the error correction step. Let C'' denote the register that contains all other public announcement other than error correction communication transcripts. That is, $E' = EC'C''$. The information leakage during the error correction and error detection subprotocol has been estimated above. Then, we can bound $H^{\varepsilon'}_{\text{min}}(\boldsymbol{Z}|EC'C'')$ from below by $H^{\varepsilon'}_{\text{min}}(\boldsymbol{Z}|EC'')$ and the information leakage term, which is stated in [117, Lemma 2]. That is,

$$H^{\varepsilon'}_{\text{min}}(\boldsymbol{Z}|EC'C'') \geq H^{\varepsilon'}_{\text{min}}(\boldsymbol{Z}|EC'') - \text{leak}_{\text{EC}} - \log_2(\frac{2}{\varepsilon_{\text{EC}}}), \tag{3.11}$$

where leak_{EC} is the communication cost of the error correction subprotocol and $\log_2(2/\varepsilon_{\text{EC}})$ is the communication cost of the error detection step. Combining Eq. (3.11) with Corollary 3.3.3, one can see that the key is ε_{sec}-secret if its length ℓ satisfies

$$\ell \leq H^{\varepsilon'}_{\text{min}}(\boldsymbol{Z}|EC'') - \text{leak}_{\text{EC}} - \log_2(\frac{2}{\varepsilon_{\text{EC}}}) - 2\log_2(\frac{2}{\varepsilon_{\text{PA}}}). \tag{3.12}$$

The task of calculating secure key length ℓ is then reduced to finding lower bound of $H^{\varepsilon'}_{\text{min}}(\boldsymbol{Z}|EC'')$. For simplicity of writing, let $[E]$ denote EC''. Thus, we can write $H^{\varepsilon'}_{\text{min}}(\boldsymbol{Z}|[E])$ for $H^{\varepsilon'}_{\text{min}}(\boldsymbol{Z}|EC'')$.

3.3.3 Parameter estimation

In the parameter estimation step of a QKD protocol, Alice and Bob perform some statistical analysis on their data to estimate Eve's knowledge about the key. This estimation helps them to make a decision about continuation of the protocol. It also helps to find a lower bound on the smooth min-entropy $H_{\min}^{\varepsilon'}(\boldsymbol{Z}|[E])$ where the input state is unknown but subject to constraints from the observed statistics.

Let $M = \{M_\lambda\}_{\lambda\in\Lambda}$ be Alice and Bob's joint POVM used for the parameter estimation subprotocol. To perform parameter estimation, they choose a random subset of signals. For the chosen subset, they announce all the information to bound Eve's knowledge about the key. Let m be the number of signals used for parameter estimation. When they announce all the information, they are able to construct a sequence of outcomes $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_m) \in \Lambda^{\times m}$, and then they are able to compute the frequency distribution over Λ:

$$\mathsf{freq}(\boldsymbol{\lambda}) = \sum_{\lambda\in\Lambda} \frac{|\{i : \lambda_i = \lambda\}|}{m} |\lambda\rangle\langle\lambda| \,. \tag{3.13}$$

We note that we write a frequency distribution as a diagonal matrix whose diagonal entries are frequencies of outcomes. We may also write it as a vector by replacing $|\lambda\rangle\langle\lambda|$ with $|\lambda\rangle$ in the above expression. For a frequency distribution $F \in \mathbb{P}(\Lambda)$, we abuse the notation to let F be either the vector version of the diagonal matrix version, which should be clear from the context.

Before the execution of the protocol, they have agreed on the set of acceptable frequency distributions $\mathcal{Q}$. If $\mathsf{freq}(\boldsymbol{\lambda}) \in \mathcal{Q}$, they continue the protocol. Otherwise, they abort.

ε_{PE}-securely filtered state

Statistics from parameter estimation give us constraints on possible eavesdropping attacks. Ideally, we need only to evaluate $H_{\min}^{\varepsilon'}(\boldsymbol{Z}|[E])$ for states that can produce the observed statistics, and consider the worst-case scenario among those states. However, when the number m is finite, all measured data are subject to statistical fluctuations. Due to statistical fluctuations, some extremely adverse states can also produce the observed statistics with an extremely small probability. It means that regardless of Eve's attacks, it is always possible that the observed statistics are produced by those states which lead to extremely low values of $H_{\min}^{\varepsilon'}(\boldsymbol{Z}|[E])$. If we cannot exclude such states, the worst-case scenario may always lead to a key with short or even zero length. For instance, let us consider the following scenario. Let $\mathcal{H}_A \otimes \mathcal{H}_B$ be the joint Hilbert space of Alice's and Bob's single-round systems. Suppose they accept only the statistic that would be produced by N copies of the maximally entangled state on $\mathcal{H}_A \otimes \mathcal{H}_B$, i.e., perfectly correlated data. However, if ρ_{AB} is N copies of the completely mixed state on $\mathcal{H}_A \otimes \mathcal{H}_B$ where Alice and Eve (as well as Bob and Eve) share N copies of the maximally entangled state in the joint state ρ_{ABE}, there is still a non-zero (but exponentially small in m) probability that Alice and Bob would obtain perfectly correlated data by sampling m from these N signals. It is clear that one cannot generate secret keys from this joint state ρ_{ABE}. Thus, in this scenario, we would trivially conclude that we could not generate any keys of nonzero length if we could not exclude the completely mixed state from consideration due to its nonzero probability of producing the observed statistics.

It motivates the definition of $\varepsilon_{\mathrm{PE}}$-securely filtered states, which are states that we want to ignore in the security analysis due to their low probabilities of producing the desired statistics.

Definition 3.3.4 ($\varepsilon_{\mathrm{PE}}$-securely filtered states, [89, Definition 6.2.1])**.** Let $\varepsilon_{\mathrm{PE}} \geq 0$. Let $\{M_\lambda \in \mathrm{Pos}(\mathcal{H}_A \otimes \mathcal{H}_B) : \lambda \in \Lambda\}$ be Alice and Bob's joint POVM, which is any measurement that can be performed by two remote parties with an authenticated classical channel. Let $\rho_{AB}^N \in \mathrm{D}((\mathcal{H}_A \otimes \mathcal{H}_B)^{\otimes N})$. We say the state ρ_{AB}^N is $\varepsilon_{\mathrm{PE}}$-securely filtered if the parameter estimation subprotocol, which applies the joint POVM $\{M_\lambda : \lambda \in \Lambda\}$ m times on m randomly selected signals from the state ρ_{AB}^N, aborts with a probability at least $1 - \varepsilon_{\mathrm{PE}}$.

In other words, the probability that the parameter estimation subprotocol accepts the statistics produced from measuring m signals of an $\varepsilon_{\mathrm{PE}}$-securely filtered state ρ_{AB}^N, by applying the POVM $\{M_\lambda\}_{\lambda\in\Lambda}$ to each of these m signals, is at most $\varepsilon_{\mathrm{PE}}$. One would like to exclude such states from security analysis at the cost of a failure probability $\varepsilon_{\mathrm{PE}}$, that is, the probability that the real state (after Eve's attacks) they sampled from to obtain the observed frequency distribution $\mathsf{freq}(\boldsymbol{\lambda})$ is actually excluded in the security analysis. Ideally, one would like to consider the complement of the following set

$$\mathbf{S}_{\leq\varepsilon_{\mathrm{PE}}} = \{\rho^N \in \mathrm{D}((\mathcal{H}_A \otimes \mathcal{H}_B)^{\otimes N}) : \Pr\big(\text{parameter estimation accepts} \mid \rho^N\big) \leq \varepsilon_{\mathrm{PE}}\} \tag{3.14}$$

in the security analysis. However, this set is in general difficult to characterize. If we are allowed to restrict the set of states to i.i.d. states, we are able to consider $\varepsilon_{\mathrm{PE}}$-securely filtered i.i.d. states and then consider single-copy states that are much easier to analyze. This concept is relevant for a proof against i.i.d. collective attacks. Thus, it is quite useful as in many scenarios we can prove the security against i.i.d. collective attacks first before applying appropriate techniques to lift the proof to general attacks.

Definition 3.3.5 ($\varepsilon_{\mathrm{PE}}$-securely filtered i.i.d. states)**.** Let $\varepsilon_{\mathrm{PE}} \geq 0$. Let $\rho_{AB} \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be a state from a single round. We say the state ρ_{AB} is an $\varepsilon_{\mathrm{PE}}$-securely filtered i.i.d. state if the parameter estimation subprotocol aborts with a probability at least $1 - \varepsilon_{\mathrm{PE}}$ on the state $\rho_{AB}^{\otimes m}$.

In [89], Renner defines the set $\Gamma_{M,\mathcal{Q}}^{\leq\mu}$ of density operators for which the measurement M leads to a frequency distribution which has (variational) distance at most μ to the set $\mathcal{Q}$ as

$$\Gamma_{M,\mathcal{Q}}^{\leq\mu} = \{\sigma : \min_{Q\in\mathcal{Q}} \|\Phi_{\mathcal{M}}(\sigma) - Q\|_1 \leq \mu\}, \tag{3.15}$$

where we define the measurement CPTP map $\Phi_{\mathcal{M}}$ as

$$\Phi_{\mathcal{M}}(\sigma) = \sum_{\lambda\in\Lambda} \mathrm{Tr}(\sigma M_\lambda) \left|\lambda\middle\rangle\!\middle\langle\lambda\right| , \tag{3.16}$$

and we write Q as a diagonal matrix. Sometimes, it is helpful to define another version of $\Phi_{\mathcal{M}}$ that outputs vectors instead of diagonal matrices:

$$\Phi_{\boldsymbol{\mathcal{M}}}(\sigma) = \sum_{\lambda\in\Lambda} \operatorname{Tr}(\sigma M_\lambda) \left|\lambda\right\rangle . \tag{3.17}$$

This set $\Gamma^{\leq\mu}_{M,\mathcal{Q}}$ contains all the states that one needs to consider in the security analysis assuming i.i.d. sampling when we allow some small probability of failure ε_{PE}. Any state that is not in this set is ε_{PE}-securely filtered. In order to relate the variational distance μ to the failure probability ε_{PE}, some concentration inequalities are needed. This connection is established by the law of large numbers in [89]. Although we assume the i.i.d. scenario here, as we will see, a security proof under the i.i.d. assumption can usually be lifted to a security proof against general attacks by techniques like the quantum de Finetti representation theorem [89, 111], the postselection technique [112, 120] or the entropy accumulation theorem [97, 98] at some cost of the key length. We also note that when the number of measurement outcome is two, one may find tighter bounds such as Chernoff bound [121], Hoeffding's inequality [122] or Serfling's inequality [123] depending on the assumptions used.

We note that a similar concept of ε_{PE}-securely filtered states exists when one applies the EAT to the QKD security proofs. See Remark 8.2.2 for the exact claim in the context of EAT.

Coarse-graining

Another important concept related to how data are used in the parameter estimation step is *coarse-graining*. As we will see later, for a fixed security parameter ε_{PE}, the variational distance bound μ typically depends on the size of alphabet for the measurement outcomes. It is thus sometimes helpful to consider a smaller alphabet in order to make the bound μ smaller. We discuss the concept of coarse-graining here. Coarse-grained data and fine-grained data are relative to each other. We call the data *fine-grained* if they have a larger alphabet. We call the data *coarse-grained* relative to the fine-grained data if they have a smaller alphabet and there is a (deterministic or probabilistic) function that maps from the larger alphabet to the smaller one. In general, this function can be described by a conditional probability distribution $p_{\Lambda|\Sigma}$ where we call the larger alphabet Σ and the smaller one Λ. For a given frequency distribution F over the alphabet Λ, one can obtain the coarse-grained statistics F^C by

$$F^C(i) = \sum_{j\in\Sigma} p_{\Lambda|\Sigma}(i|j)F(j). \tag{3.18}$$

One may define a classical-to-classical channel $\mathcal{N}$ to represent this linear transformation from $\mathbb{P}(\Sigma)$ to $\mathbb{P}(\Lambda)$. Then, one can write F^C as $\mathcal{N}(F)$ with

$$\mathcal{N}(F) = \sum_{x\in\Sigma, y\in\Lambda} p_{\Lambda|\Sigma}(y|x) \left\langle x\right| F \left|x\right\rangle \left|y\middle\rangle\middle\langle y\right| . \tag{3.19}$$

Here, we abuse the notation again to let F mean either the vector version or the diagonal matrix version of the frequency distribution.

If the original POVM is $\{\Gamma_j : j \in \Sigma\}$, the effective POVM $\{\widetilde{\Gamma}_i^C : i \in \Lambda\}$ corresponding to the coarse-grained statistics is then given by

$$\widetilde{\Gamma}_i^C = \sum_{j\in\Sigma} p_{\Lambda|\Sigma}(i|j)\Gamma_j. \tag{3.20}$$

In a QKD protocol, Alice has a POVM $\{M_x^A : x \in \Sigma_A\}$ with some alphabet Σ_A and Bob has a POVM $\{M_y^B : y \in \Sigma_B\}$ with some alphabet Σ_B. We typically call the statistics that they obtain directly by applying their joint POVM $\{M_x^A \otimes M_y^B : (x, y) \in \Sigma_A \times \Sigma_B\}$ as fine-grained since it captures the most detailed information that one can obtain via measurements allowed by the protocol. Coarse-graining then typically refers to processing data from their joint alphabet $\Sigma_A \times \Sigma_B$ to another joint alphabet Λ with $|\Lambda| \leq |\Sigma_A \times \Sigma_B|$.

Acceptance set of frequency distributions

For the rest of this thesis, we assume that the acceptance set $\mathcal{Q}$ is defined in the following way:

$$\mathcal{Q} = \{F \in \mathbb{P}(\Sigma) : \left\|\overline{\mathcal{N}}(F) - \overline{\mathcal{N}}(\overline{F})\right\|_1 \leq t\} \tag{3.21}$$

for some alphabet Σ, where $\overline{\mathcal{N}}$ is a classical-to-classical channel that models a coarse-graining and $\overline{F}$ is a preferred frequency distribution in $\mathbb{P}(\Sigma)$, (e.g. one may choose $\overline{F}$ to be an expected probability distribution in an honest implementation) and $t \geq 0$ is some acceptance threshold. We note that this description of the acceptance set is a simple and practical description of the acceptance set since it includes a typical scenario in the BB84-type protocols; that is, one typically defines an acceptance threshold for the phase error in those protocols. A protocol that accepts a unique frequency distribution $\overline{F}$ can set $t = 0$.

3.3.4 Collective attacks, coherent attacks, and proof lifting

We now see the task of key length calculation is reduced to the evaluation of the smooth min-entropy $H_{\min}^{\varepsilon'}(\boldsymbol{Z}|[E])$ from states which are not ε_{PE}-securely filtered by the parameter estimation subprotocol. However, this evaluation is still challenging since each state contains N signals as a QKD session typically has N rounds with N quite large (with typical values ranging from 10^5 to 10^{14}). Moreover, it involves Eve's system which is outside the control of Alice and Bob. On the other hand, a security proof against i.i.d. collective attacks is much easier to perform since under the i.i.d. assumption, the joint state ρ_{ABE}^N has a tensor product structure, that is, $\rho_{ABE}^N = \sigma_{ABE}^{\otimes N}$.

One can then use the asymptotic equipartition property (AEP) (see Theorem 3.4.2) or a similar tool (Theorem 3.4.3) to bound $H_{\min}^{\varepsilon'}(\boldsymbol{Z}|[E])$ by the von Neumann entropy of each individual copy σ_{ABE}. While it is reasonable to assume both the state preparation and the measurement process behave in the i.i.d. fashion, it seems too restrictive to assume that Eve's attacks are also i.i.d. at the first glance. Fortunately, for many protocols that are invariant under permutation of rounds, it is shown [89, 124, 125] that the i.i.d. assumption is not too restrictive since the optimal attacks in the asymptotic regime are i.i.d. collective attacks. Moreover, one can use some version of the quantum de Finetti representation theorem [89, 111] to link the security proof against i.i.d. collective attacks to a proof against coherent attacks in the finite-key regime. There are also other techniques like the postselection technique [112, 120] and the entropy accumulation theorem [97, 98] that establish links between the i.i.d. scenario and the most general case provided that the protocol meets certain conditions. We discuss collective attacks as well as these techniques to lift the security proof in the next several sections. It is worth noting that the lifting techniques can still be improved and are an active research subject (e.g., see a recent paper [99]).

3.4 Security against i.i.d. collective attacks

We now consider a much simplified analysis by assuming i.i.d. collective attacks. As we will see later, many essential ideas remain when we lift such a proof to the most general case.

3.4.1 Parameter estimation under i.i.d. assumption

Under i.i.d. assumption, we would like to consider the set of i.i.d. states that are not $\varepsilon_{\mathrm{PE}}$-securely filtered, which corresponds to the set $\Gamma^{\leq\mu}_{M,\mathcal{Q}}$. We now discuss the method to calculate μ from a given security parameter $\varepsilon_{\mathrm{PE}}$ in [89], which is also used in [84, 117, 126, 127]. In [89], the law of large numbers is used to relate the security parameter $\varepsilon_{\mathrm{PE}}$ to the variational distance bound μ as

$$\mu = \sqrt{2}\sqrt{\frac{\ln(1/\varepsilon_{\mathrm{PE}}) + |\Sigma| \ln(m+1)}{m}}. \tag{3.22}$$

In other words, if a state σ produces a frequency distribution $\Phi_{\mathcal{M}}(\sigma)$ by applying the measurement M, and $\|\Phi_{\mathcal{M}}(\sigma) - \mathsf{freq}(\boldsymbol{\lambda})\|_1 \leq \mu$ for some $\mathsf{freq}(\boldsymbol{\lambda}) \in \mathcal{Q}$, then the probability that the parameter estimation subprotocol does not abort, when sampling from $\sigma^{\otimes m}$, is at least $\varepsilon_{\mathrm{PE}}$.

Let $A_{\mathcal{T}}$ denote the event that the parameter estimation protocol does not abort. Let $\Pr(A_{\mathcal{T}}|\sigma^{\otimes m})$ denote the probability of the event $A_{\mathcal{T}}$ when the parameter estimation protocol samples from $\sigma^{\otimes m}$. We state our result [84, Theorem 1] on the security analysis with multiple coarse-grainings.

Theorem 3.4.1 (Security with multiple coarse-grainings, [84, Theorem 1])**.** Let $\varepsilon_{\text{PE}} > 0$. Let Ξ be a finite alphabet that indexes multiple coarse-grainings. For each $k \in \Xi$, let

$$\begin{aligned}\mathbf{S}_{\mu_k} = \{\rho \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B) : \exists F_k \in \mathbb{P}(\Sigma) \text{ such that } &\left\|\Phi_{\mathcal{M}_k}(\rho) - \mathcal{N}_k(F_k)\right\|_1 \leq \mu_k \\ \& \left\|\overline{N}_k(F_k) - \overline{N}_k(\overline{F})\right\|_1 \leq t_k\},\end{aligned} \tag{3.23}$$

where $\overline{F} \in \mathbb{P}(\Sigma)$ is a preferred frequency distribution to define the set of acceptable statistics $\mathcal{Q}$, each $\overline{N}_k$ is a classical-to-classical channel that realizes necessary coarse-grainings on $\overline{F}$ and F_k in the definition of $\mathcal{Q}$ in Eq. (3.21), and t_k is the acceptance threshold for the k-th coarse-graining. The measurement channel $\Phi_{\mathcal{M}_k}$ is defined in Eq. (3.16) with the same set of measurement outcomes as the coarse-graining $\mathcal{N}_k$. Each variational distance bound μ_k is calculated according to Eq. (3.22) for the given security parameter ε_{PE} so that $\Pr(A_{\mathcal{T}}|\sigma^{\otimes m}) \leq \varepsilon_{\text{PE}}\ \forall \sigma \notin \mathbf{S}_{\mu_k}$.

Then if $\sigma \notin \cap_k \mathbf{S}_{\mu_k}$, it is the case that $\Pr(A_{\mathcal{T}}|\sigma^{\otimes m}) \leq \varepsilon_{\text{PE}}$.

Proof. It follows simply by the construction of $\mathbf{S}_{\mu_k}$ which has the property that $\Pr(A_{\mathcal{T}}|\sigma^{\otimes m}) \leq \varepsilon_{\text{PE}}$ for every $\sigma \notin \mathbf{S}_{\mu_k}$. Thus, if $\sigma \notin \cap_k \mathbf{S}_{\mu_k}$, there exists some $j \in \Xi$ such that $\sigma \notin \mathbf{S}_{\mu_j}$. So, $\Pr(A_{\mathcal{T}}|\sigma^{\otimes m}) \leq \varepsilon_{\text{PE}}$. □

This result tells us that we can consider multiple coarse-grainings without modifying the security parameter ε_{PE}, an improvement from a similar statement in [127] where the security parameter is increased to allow multiple coarse-grainings. We also note that in this theorem we allow the set of acceptable frequency distributions to set different threshold values for different coarse-grainings, which is a generalization of the definition in Eq. (3.21). By this result, the set of i.i.d. states that we need to consider in the security proof is

$$\mathbf{S}_{\varepsilon_{\text{PE}}} = \{\rho \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B) : \rho \in \bigcap_k \mathbf{S}_{\mu_k}\}. \tag{3.24}$$

When there is ρ_A constraint, we add the promise that the reduced density operator on Alice's system is unaffected by Eve. This leads to a modified definition of $\mathbf{S}_{\varepsilon_{\text{PE}}}$ which we call $\mathbf{S}_{\text{PE}}$:

$$\mathbf{S}_{\text{PE}} = \{\rho \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B) : \rho \in \bigcap_k \mathbf{S}_{\mu_k}, \mathrm{Tr}_B(\rho) = \rho_A\}. \tag{3.25}$$

For the entanglement-based protocol where there is no ρ_A constraint, we have $\mathbf{S}_{\text{PE}} = \mathbf{S}_{\varepsilon_{\text{PE}}}$.

3.4.2 Key length under i.i.d. collective attacks

Under the i.i.d. assumption, the quantum states that we want to evaluate the smooth min-entropy are of a tensor product form. To evaluate the smooth min-entropy of an i.i.d. quantum state, one can relate it to the von Neumann entropy of the state on a single system. This relation can be achieved by a quantum version of AEP. We state the direct part of AEP [128, 129] that allows us to obtain a lower bound on the smooth min-entropy by the von Neumann entropy. A converse bound can be found in [128].

Theorem 3.4.2 (Quantum AEP (direct part) [128, Theorem 9])**.** Let $\varepsilon > 0$. Let $n \in \mathbb{N}$ such that $n \geq \frac{8}{5}\log_2(\frac{2}{\varepsilon^2})$. Let $\rho = \sigma_{AB}^{\otimes n}$ be an i.i.d. quantum state on $\mathcal{H}_{AB}^{\otimes n}$. Then,

$$\frac{1}{n}H_{\min}^{\varepsilon}(A^n|B^n)_\rho \geq H(A|B)_\sigma - \frac{\delta(\varepsilon,\nu)}{\sqrt{n}}, \tag{3.26}$$

where

$$\delta(\varepsilon,\nu) := 4\log_2(\nu)\sqrt{\log_2(\frac{2}{\varepsilon^2})}, \tag{3.27}$$

$$\nu := \sqrt{2^{-H_{\min}(A|B)_\sigma}} + \sqrt{2^{H_{\max}(A|B)_\sigma}} + 1. \tag{3.28}$$

Alternatively, we can use a result from Renner's Ph.D. thesis [89] that relates the smooth min-entropy and the von Neumann entropy. While this result is potentially looser than the quantum AEP, it gives an easy way to evaluate the lower-order terms since we do not need to calculate the smooth min- and max-entropy of the state on a single system. It is also sufficient for practical cases. We correct a known typographical error of [89, Corollary 3.3.7] and make a small improvement by taking an intermediate result from its proof. This result is stated in Theorem 3.4.3.

Theorem 3.4.3 (Variation of [89, Corollary 3.3.7])**.** Let $\varepsilon > 0$. Let $\rho = \sigma_{AB}^{\otimes n}$ be an i.i.d. quantum state on $\mathcal{H}_{AB}^{\otimes n}$. Then,

$$\frac{1}{n}H_{\min}^{\varepsilon}(A^n|B^n)_\rho \geq H(A|B)_\sigma - H(B)_\sigma - \delta \tag{3.29}$$

where $\delta = 2\log_2(\text{rank}(\sigma_A) + 3)\sqrt{\frac{\log_2(2/\varepsilon)}{2}}$.

We then state our result in [84] which is based on Theorem 3.4.3 and a generalization of the result from [89] by allowing multiple coarse-grainings. We label the smoothing parameter of the smooth min-entropy as $\bar{\varepsilon}$ in this following theorem.

Theorem 3.4.4. Let n be the number of rounds for key generation and m be the number of rounds used in the parameter estimation step. The QKD protocol is $\varepsilon = \varepsilon_{\text{PE}} + \bar{\varepsilon} + \varepsilon_{\text{EC}} + \varepsilon_{\text{PA}}$-secure against i.i.d. collective attacks when the protocol does not abort and the length ℓ of the output key satisfies

$$\ell \leq n\Big[\min_{\rho \in \mathbf{S}_{\text{PE}}} H(Z|[E]) - 2\log_2(d+3)\sqrt{\frac{\log_2(2/\bar{\varepsilon})}{n}}\Big] - \text{leak}_{\text{EC}} - \log_2(\frac{2}{\varepsilon_{\text{EC}}}) - 2\log_2(\frac{2}{\varepsilon_{\text{PA}}}), \tag{3.30}$$

where the feasible set $\mathbf{S}_{\text{PE}}$ is defined in Eq. (3.25) with μ defined in Eq. (3.22), leak_{EC} is the number of bits leaked during the error correction step, and d is the size of key alphabet after the key map step.

For DVQKD, the parameter leak_{EC} is characterized as $nf_{\text{EC}}H(Z|Y)$ where $f_{\text{EC}} \geq 1$ is the inefficiency of the error correction code used in the protocol. For CVQKD protocols, a different convention for the error correction efficiency is used, that is, $\text{leak}_{\text{EC}} = n[H(Z) - \beta_{\text{EC}} I(Z:Y)]$ for $\beta_{\text{EC}} \leq 1$. For the reverse reconciliation scheme, we simply replace Y by X in the above expressions.

Then the key rate R_N^{coll} against i.i.d. collective attacks, when N signals is transmitted, is ℓ/N, where ℓ is from Eq. (3.30). That is,

$$\begin{aligned} R_N^{\text{coll}} =& \frac{n}{N}\Big[\min_{\rho \in \mathbf{S}_{\text{PE}}} H(Z|[E]) - \delta_{\text{EC}}\Big] \\ &- \frac{1}{N}\Big[\Big(2\log_2(d+3)\sqrt{\log_2(2/\bar{\varepsilon})}\Big)\sqrt{n} - \log_2(\frac{2}{\varepsilon_{\text{EC}}}) - 2\log_2(\frac{2}{\varepsilon_{\text{PA}}})\Big], \end{aligned} \tag{3.31}$$

where we define

$$\delta_{\text{EC}} = \frac{1}{n}\text{leak}_{\text{EC}}. \tag{3.32}$$

When the number of signals N goes to infinity, we see the last several terms in Eq. (3.31) vanish. Thus, the asymptotic key rate R_∞ against collective attacks is

$$R_\infty = p_{\text{pass}}\Big[\min_{\rho \in \mathbf{S}} H(Z|[E]) - \delta_{\text{EC}}\Big], \tag{3.33}$$

where $p_{\text{pass}} = \lim_{N\to\infty} n/N$ is the probability that a given round is used for key generation (n is also a function of N), and the feasible set $\mathbf{S}$ is obtained from $\mathbf{S}_{\text{PE}}$ [in Eq. (3.25)] by taking each μ to be zero.

3.5 Security against coherent attacks via quantum de Finetti theorem

3.5.1 Finite quantum de Finetti theorem

We present the finite Quantum de Finetti theorem and its application to QKD security proofs [89].

Let $\mathcal{H}$ be a Hilbert space and let $0 \leq m \leq n$. For a pure state $|\theta\rangle \in \mathcal{H}$, we define the set of pure states that are partially i.i.d. pure states with respect to $|\theta\rangle$ as

$$\mathcal{V}(\mathcal{H}^{\otimes n}, |\theta\rangle^{n}) := \{W_\pi(|\theta\rangle^{\otimes m} \otimes |\Psi\rangle) : \pi \in \mathcal{S}_n, |\Psi\rangle \in \mathcal{H}^{\otimes n-m}\} \ . \tag{3.34}$$

(See Eq. (2.13) for the definition of W_π.) The subspace of symmetric pure states that are partially i.i.d. states with respect to a pure state $|\theta\rangle$, denoted by $\mathrm{Sym}(\mathcal{H}^{\otimes n}, |\theta\rangle^{\otimes n-r})$, is defined as

$$\mathrm{Sym}(\mathcal{H}^{\otimes n}, |\theta\rangle^{\otimes n-r}) := \mathrm{Sym}(\mathcal{H}^{\otimes n}) \cap \mathrm{span}(\mathcal{V}(\mathcal{H}^{\otimes n}, |\theta\rangle^{n})). \tag{3.35}$$

(See Eq. (2.14) for the definition of $\mathrm{Sym}(\mathcal{H}^{\otimes n})$.)

Theorem 3.5.1 (Finite quantum de Finetti theorem, [89, Theorem 4.3.2])**.** Let $\rho_{n+k} \in \mathrm{Sym}(\mathcal{H}^{\otimes n+k})$ be a pure density operator. Let $0 \leq r \leq n$. Then there exists a measure ν on the unit sphere $\mathcal{S}_1(\mathcal{H})$ and a pure state $\sigma^{|\theta\rangle} \in \mathrm{Sym}(\mathcal{H}^{\otimes n}, |\theta\rangle^{\otimes n-r})$ for each $|\theta\rangle \in \mathcal{S}_1(\mathcal{H})$ such that

$$\left\| \mathrm{Tr}_{\mathcal{H}^{\otimes k}}(\rho_{n+k}) - \int_{\mathcal{S}_1(\mathcal{H})} \sigma^{|\theta\rangle} \nu(|\theta\rangle) \right\|_1 \leq 2e^{-\frac{k(r+1)}{2(n+k)} + \frac{1}{2}\dim(\mathcal{H})\ln(k)}. \tag{3.36}$$

When it applies to QKD security proofs, the following theorem is an adaption of [89, Theorem 6.5.1] and is explicitly stated in [84].

Theorem 3.5.2 (Adaption of [89, Theorem 6.5.1])**.** Given a general entanglement-based protocol as in Protocol 2.2, let N be the total number of transmitted signals in a QKD session. Let m be the number of signals used for parameter estimation, and let n be the number of signals used for key generation. Let $k \in \mathbb{N}$ such that $bn + m + k = N$ where b accounts for block-wise processing.

Let $\varepsilon_{\mathrm{PE}}, \bar{\varepsilon}, \varepsilon_{\mathrm{EC}}, \varepsilon_{\mathrm{PA}}, \varepsilon_{\mathrm{QdF}} > 0$. Then the QKD is $\varepsilon = \varepsilon_{\mathrm{PE}} + \bar{\varepsilon} + \varepsilon_{\mathrm{EC}} + \varepsilon_{\mathrm{PA}} + \varepsilon_{\mathrm{QdF}}$-secure if it is $\varepsilon_{\mathrm{EC}}$-correct and if

$$\begin{aligned}\ell \leq & n[\min_{\rho \in \mathbf{S}_{\mathrm{PE}}} H(Z|[E]) - \delta(\bar{\varepsilon})] - 2(m+k)\log_2(\dim(\mathcal{H}_A \otimes \mathcal{H}_B)) \\ & - \mathrm{leak}_{\mathrm{EC}} - \log_2(\frac{2}{\varepsilon_{\mathrm{EC}}}) - 2\log_2(\frac{2}{\varepsilon_{\mathrm{PA}}}),\end{aligned} \tag{3.37}$$

where the set $\mathbf{S}_{\mathrm{PE}}$ is defined in Eq. (3.25) with a modified definition of the variational distance bound μ (for each μ_k) as

$$\mu = 2\sqrt{h(\frac{r}{m}) + \frac{\log_2(1/\varepsilon_{\mathrm{PE}}) + |\Sigma|\log_2(m/2+1)}{m}}, \tag{3.38}$$

$$\delta(\bar{\varepsilon}) = (\frac{5}{2}\log_2(d) + 4)\sqrt{h(r/n) + \frac{2}{n}\log_2(4/\bar{\varepsilon})}, \tag{3.39}$$

$$r = (\frac{bn+m}{k} + 1)\left[2\ln\left(\frac{2}{\varepsilon_{\mathrm{QdF}}}\right) + \dim(\mathcal{H}_A \otimes \mathcal{H}_B)^2 \ln(k)\right] - 1 \leq N. \tag{3.40}$$

For a prepare-and-measure protocol, when one applies the source-replacement scheme to obtain an equivalent entanglement-based scheme, there is one additional promise that Alice's reduced density operator ρ_A is unaffected by Eve. To guarantee that we need to consider only states with the desired reduced state ρ_A in the parameter estimation step, one needs to introduce another security parameter $\varepsilon_{\mathrm{SR}} > 0$ according to [89, Remark 4.3.3], which we state as a proposition here.

Proposition 3.5.3 ([89, Remark 4.3.3]). Let $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ be a bipartite Hilbert space. Let $\rho_{A^{n+k}B^{n+k}} \in \mathrm{Sym}(\mathcal{H}^{\otimes n})$ be a pure density operator such that $\rho_{A^{n+k}} = \sigma_A^{\otimes n+k}$ for some $\sigma_A \in \mathrm{D}(\mathcal{H}_A)$. Let ν be the measure in Theorem 3.5.1. Then for any $\delta \geq 0$, the set

$$\overline{\Gamma^{\delta}} := \{|\theta\rangle \in \mathcal{S}_1(\mathcal{H}) : \|\mathrm{Tr}_B(|\theta\rangle\langle\theta|) - \sigma_A\|_1 \geq \delta\} \tag{3.41}$$

has at most weight $\nu(\overline{\Gamma^{\delta}}) \leq e^{-\frac{1}{4}k\delta^2 + \dim(\mathcal{H})\ln(k)}$.

To use this proposition, we set $\varepsilon_{\mathrm{SR}} = e^{-\frac{1}{4}k\delta^2 + \dim(\mathcal{H})\ln(k)}$ and solve for δ to find that

$$\delta = \frac{2}{\sqrt{k}}\sqrt{\dim(\mathcal{H}_A \otimes \mathcal{H}_B)\ln(k) + \ln\left(\frac{1}{\varepsilon_{\mathrm{SR}}}\right)}. \tag{3.42}$$

Then, the set $\mathbf{S}_{\mathrm{PE}}$ in the finite-key rate problem is replaced by the set $\mathbf{S}_{\mathrm{PE}}^{\mathrm{SR}}$, which is defined as

$$\mathbf{S}_{\mathrm{PE}}^{\mathrm{SR}} = \{\sigma \in \mathbf{S}_{\varepsilon_{\mathrm{PE}}} : \|\mathrm{Tr}_B(\sigma) - \rho_A\|_1 \leq \delta\}. \tag{3.43}$$

In summary, for a prepare-and-measure protocol (as described in Protocol 2.1), one introduces an additional security parameter $\varepsilon_{\mathrm{SR}} > 0$ and modify the total security parameter ε as $\varepsilon = \varepsilon_{\mathrm{PE}} + \bar{\varepsilon} + \varepsilon_{\mathrm{EC}} + \varepsilon_{\mathrm{PA}} + \varepsilon_{\mathrm{QdF}} + \varepsilon_{\mathrm{SR}}$ in Theorem 3.5.2 as well as changing the set $\mathbf{S}_{\mathrm{PE}}$ to $\mathbf{S}_{\mathrm{PE}}^{\mathrm{SR}}$. Similarly, in the prepare-and-measure version of a measurement-device-independent protocol, systems AB together play the role of A in the discussion above and the classical register C plays the role of B here.

3.5.2 de Finetti theorem for infinite-dimensional systems

The finite quantum de Finetti theorem [89, Theorem 4.3.2] (Theorem 3.5.1) does not apply to infinite-dimensional systems since it requires the dimension, $\dim(\mathcal{H})$, of the joint Hilbert space $\mathcal{H}$ to be sufficiently smaller than the number N of signals (otherwise, the bound in Eq. (3.36) is trivial). Thus, it cannot be applied for CVQKD protocols. A version of the quantum de Finetti theorem is developed for states that live in infinite-dimensional Hilbert spaces in [111] by appending an additional test in the protocol. In particular, the test measures two canonical observables with an equal probability and then infers the measurement outcome of another (energy-like) measurement based on measurement outcomes of these two canonical observables. CVQKD protocols usually already perform such measurements. Thus, it is well suited for CVQKD. Not surprisingly, this test for CVQKD is dubbed as *energy test*. If the protocol passes this test, then most of its signals live in a finite-dimensional subspace except some small failure probability. One can then apply a de Finetti-type argument to show that the state is well approximately by almost i.i.d. states.

On the technical level, the energy test is based on [111, Lemma 1 and Lemma 2]. We consider two canonical observables X and Y such that they satisfy the commutation relation $[X, Y] = i$.

Lemma 3.5.4 ([111, Lemma 1]). Let $U_1 := \frac{1}{2}P^{X^2 \geq n_0/2} + \frac{1}{2}P^{Y^2 \geq n_0/2}$ and $V_1 := P^{X^2+Y^2 \geq 2n_0+1}$. Let

$$\gamma_{U_1 \to V_1}(\delta) := \sup\{\mathrm{Tr}(V_1\sigma) : \sigma \in \mathrm{D}(\mathcal{H}), \mathrm{Tr}(U_1\sigma) \leq \delta\}. \tag{3.44}$$

For any $n_0 \in \mathbb{N}$ and $\delta \geq 0$,

$$\gamma_{U_1 \to V_1}(\delta) \leq 4\delta + \frac{4}{\sqrt{c_0 \pi n_0}} e^{-n_0 c_0}, \tag{3.45}$$

with $c_0 = (1 - 1/\sqrt{2})^2$.

This lemma in the context of energy test tells us that for all states such that the probability that both observables lead to large values is small, then the probability that the maximum energy among those states is large can be found to be small. Based on this lemma, the following lemma tells us that we can do the security analysis in the finite-dimensional space under an experimentally verifiable condition.

Before we state the lemma, we first define a subspace $\mathcal{S}^M_{\bar{\mathcal{H}}^{\otimes M-k}}$ of $\mathcal{H}^{\otimes M}$ which consists of vectors that are contained in a low-dimensional subspace $\bar{\mathcal{H}} \subset \mathcal{H}$ except k subsystems. Formally,

$$\mathcal{S}^M_{\bar{\mathcal{H}}^{\otimes M-k}} := \operatorname{span} \bigcup_{\pi} W_\pi (\bar{\mathcal{H}}^{\otimes M-k} \otimes \mathcal{H}^{\otimes k}) W_\pi^{-1}, \tag{3.46}$$

where the union is taken over all permutations π. With this subspace defined, we state the result for the failure probability from [111, Lemma 2].

Lemma 3.5.5 ([111, Lemma 2])**.** Let $(z_1, \dots, z_k)$ be the measurement outcomes when one applies X and Y, chosen randomly with an equal probability $\frac{1}{2}$, to k subsystems of a permutation invariant state ρ^N of N systems. Let Ω be the event that a projection on the subspace $\mathcal{S}^M_{\bar{\mathcal{H}}^{\otimes M-k}}$ applied to the remaining $M = N - k$ systems fails. Then

$$\Pr\left[\Omega \wedge \max_{i=1}^{k} z_i^2 < \frac{n_0}{2}\right] \leq 8k^{3/2} e^{-\frac{k^3}{25N^2}} \tag{3.47}$$

for any $n_0 \geq 12 \ln(5N/k)$ and $N/k \gg 1$.

In other words, if from the test, one finds both quadratures of the state are consistently smaller than a certain threshold, then except some small probability, the rest of the state lives in some subspace $\mathcal{S}^M_{\bar{\mathcal{H}}^{\otimes M-k}}$. We note that the subspace $\bar{\mathcal{H}}$ in the definition of $\mathcal{S}^M_{\bar{\mathcal{H}}^{\otimes M-k}}$ can be chosen to be the Fock space with a photon-number cutoff in the context of CVQKD. It then allows us to perform the security analysis in the finite-dimensional subspace. With these lemmas, one can either use a de Finetti-type argument to reduce to the collective attacks as given in [111, Theorem 4] or combine with the postselection technique [112] discussed in Section 3.6.

Theorem 3.5.6 ([111, Theorem 4])**.** Let $\bar{\mathcal{H}}$ be a subspace of a (possibly infinite-dimensional) Hilbert space $\mathcal{H}$ with $\dim(\bar{\mathcal{H}}) = d$. Let $M, k \in \mathbb{N}$ with $M > 4k$ and let $\rho^M \in \mathrm{D}(\mathrm{Sym}^M(\mathcal{H}) \cap \bar{\mathcal{S}}^M_{\mathcal{H}^{\otimes M-2k}})$. Then there exists a probability distribution $p_{\mathcal{V}}$ on a finite set $\mathcal{V}$ of unit vectors $\nu \in \bar{\mathcal{H}}$ and a family of density operators $\{\hat{\rho}^{M-4k}_{\nu}\} \subset \mathrm{D}(\mathcal{S}^{M-4k}_{\nu^{\otimes M-8k}})$ such that, for $\rho^{M-4k} = \mathrm{Tr}_{\mathcal{H}^{\otimes 4k}}(\rho^M)$,

$$\mathrm{F}(\rho^{M-4k}, \sum_{\nu \in \mathcal{V}} p_{\mathcal{V}}(\nu) \hat{\rho}^{M-4k}_{\nu}) > 1 - k^d e^{-\frac{4k(k+1)}{M}}. \tag{3.48}$$

We note that the fidelity and trace distance are closely related as one can apply Fuchs-van de Graaf inequalities (see Theorem 2.1.14). In order to apply this theorem to QKD, the protocol needs to be invariant under permutations of N signal pairs held by Alice and Bob after the state distribution step and the privacy amplification step needs to use the two-universal hash functions. These two requirements are usually satisfied. In addition, the protocol needs to apply measurements to verify the dimension of the relevant subspace $\bar{\mathcal{H}}$ of the signal space $\mathcal{H}$ is small compared to N. For CVQKD protocols, one can verify this if all measurement outcomes from two canonical observables are small in the sense of Lemma 3.5.5. Then the de Finetti representation theorem (Theorem 3.5.6) allows us to conclude that for some $n \approx N$, the reduced state ρ^n is well approximated by a mixture of almost i.i.d. states. We can then lift a proof against i.i.d. collective attacks to the general attacks.

3.6 Security against coherent attacks via postselection technique

We review the postselection technique by Christandl, König and Renner [112]. The name of postselection comes from the fact (proven in [112, Lemma 2]) that any desired permutation invariant state can be *postselected* from a fixed state (see also [120, Lemma 3]).

Definition 3.6.1 (Permutation-invariant map, [120, Definition 2])**.** Let $\mathcal{H}, \mathcal{H}'$ be Hilbert spaces and let $N \in \mathbb{N}$ be a positive integer. A linear map $\mathcal{E} : \mathrm{L}(\mathcal{H}^{\otimes N}) \to \mathrm{L}(\mathcal{H}')$ is said to be permutation invariant if for any permutation $\pi \in \mathcal{S}_N$, there exists a CPTP map K_π such that $K_\pi \circ \mathcal{E} \circ \Phi_\pi = \mathcal{E}$, where Φ_π is the CPTP map that permutes the subsystems according to the permutation π.

The postselection theorem [112] is stated more generally for a permutation-invariant map Δ. In the application to QKD, the map Δ is the difference between the CPTP map that implements the real protocol and the CPTP map of the ideal protocol.

Theorem 3.6.2 (Postselection theorem [112])**.** Let $\Delta : \mathrm{L}(\mathcal{H}^{\otimes N}) \to \mathrm{L}(\mathcal{H}')$ be a linear map. If Δ is permutation invariant, then

$$\|\Delta\|_\diamond \leq g_{N,d} \|(\Delta \otimes \mathrm{id}_\mathcal{R})(\tau_{\mathcal{H}^N\mathcal{R}})\|_1, \tag{3.49}$$

where

$$g_{N,d} = \binom{N+d^2-1}{N} \leq (N+1)^{d^2-1} \tag{3.50}$$

for $d = \dim(\mathcal{H})$ and $\tau_{\mathcal{H}^N\mathcal{R}}$ is a purification of the state $\tau_{\mathcal{H}^n}$ defined as

$$\tau_{\mathcal{H}^N} = \int \sigma_\mathcal{H}^{\otimes N} \mu(\sigma_\mathcal{H}), \tag{3.51}$$

where $\mu(\cdot)$ is the measure on the space of density operators on a single subsystem $\mathcal{H}$ induced by the Hilbert-Schmidt metric [130] and the dimension of the purifying system $\mathcal{R}$ is sufficiently large (i.e. $\mathcal{R} \cong (\mathcal{H}^N)^{\otimes 3}$).

For an entanglement-based QKD protocol that satisfies the permutation invariant property, this theorem allows us to prove the ε-security against coherent attacks by first proving ε'-security against collective attacks for some ε' that is related to the security parameter ε, the number of signals N and the dimension d_{AB} of Alice and Bob's joint state on a single subsystem. We note that many protocols satisfy the permutation invariant requirement either by construction or by adding an additional symmetrization step.

Corollary 3.6.3. Let $\mathcal{E}^{\mathrm{real}} : \mathrm{D}(\mathcal{H}_{AB}^{\otimes N}) \to \mathrm{D}(\mathcal{H}_{K_A} \otimes \mathcal{H}_{K_B})$ be the quantum channel that models the physical implementation of a QKD protocol. Let $\mathcal{E}^{\mathrm{ideal}}$ be the quantum channel that models the ideal QKD protocol. Define $\Delta := \mathcal{E}^{\mathrm{real}} - \mathcal{E}^{\mathrm{real}}$. Suppose Δ is permutation invariant. If $\mathcal{E}^{\mathrm{real}}$ is ε'-secure against the collective attacks, then

$$\|\Delta\|_\diamond \leq \varepsilon'(N+1)^{d_{AB}^2-1} =: \varepsilon, \tag{3.52}$$

i.e., $\mathcal{E}^{\mathrm{real}}$ is ε-secure against the general attacks.

To the best of our knowledge, the postselection technique has not been rigorously applied to prepare-and-measure protocols using the source-replacement scheme. As explained in the context of finite quantum de Finetti theorem, the reason is that the source-replacement scheme only proves protocol security on states with a fixed reduced density matrix. However, the postselection technique proof requires security on arbitrary i.i.d. states. A simple solution is to ignore the promise on the reduced density operator. Alternatively, one proves the security of a virtual protocol so that its security implies the security of the original protocol. In the virtual protocol, one then introduces some extra testing on the reduced density operator. However, this solution will come at some extra cost in the key rate for the small block-size regime. An alternative possible solution is to prove a statement similar to Eq. (3.41) [89, Remark 4.3.3] for the de Finetti state in Eq. (3.51) used in the proof of the postselection theorem. If one can show that for the measure μ used in Theorem 3.6.2, the set $\overline{\Gamma^{\delta}}$ has at most weight $\varepsilon_{\mathrm{SR}}$ for some security parameter $\varepsilon_{\mathrm{SR}}$, then one can first show ε'-security against collective attacks with the feasible set in Eq. (3.43). This remains an open question for future work.

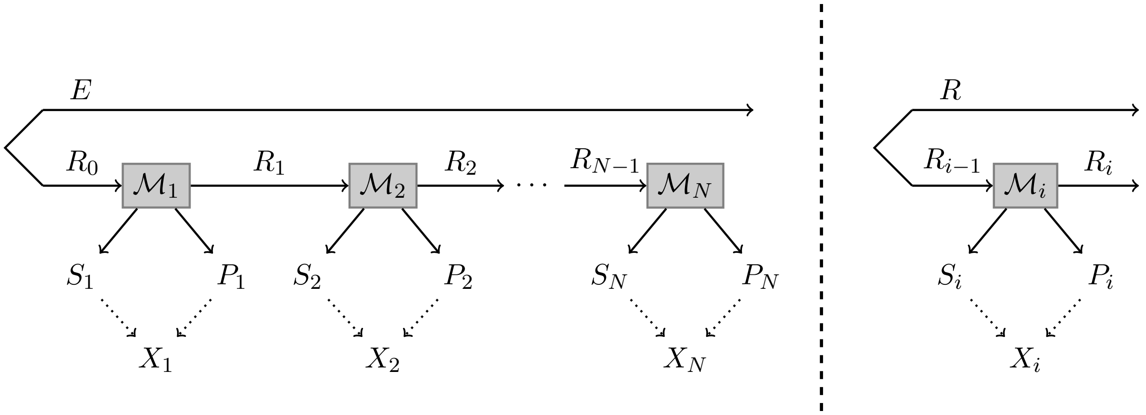
3.7 Security against coherent attacks via entropy accumulation theorem

Entropy accumulation theorem [97, 98] is a useful tool to remove the i.i.d. assumption. We summarize the entropy accumulation theorem (EAT) and adapt it in the context of device-dependent QKD. The general idea of entropy accumulation theorem is to bound the smooth min-entropy of an N-round process by the entropy of individual rounds with a much weaker condition than the i.i.d. assumption. The specific condition is Markov chain condition between the secret information and the public announcements. Effectively, the leading-order term in the key rate is the one obtained under the i.i.d. assumption.

In this section, we review the statements of the entropy accumulation theorem so that we can apply EAT to finite key analysis of device-dependent QKD. In particular, we are interested in two versions of entropy accumulation theorem. They differ in the second-order correction terms. Proofs of these two statements rely on novel chain rules of sandwiched α-Rényi entropies. They are beyond the scope of this thesis. We direct readers who are interested in technical details to [97, 98] for proofs.

We review definitions needed to understand the statement of entropy accumulation theorem. We use a shorthand notation A_1^N to denote a sequence of registers $A_1, \cdots, A_N$. The EAT process is depicted in Figure 3.1. In this diagram, $\{S_i\}$ are finite-dimensional quantum systems of dimension d_S, $\{X_i\}$ are finite-dimensional classical registers, $\{P_i\}$ and $\{R_i\}$ are arbitrary quantum registers.

Definition 3.7.1 (EAT channel)**.** EAT channels are CPTP maps $\mathcal{M}_i : R_{i-1} \to S_i P_i X_i R_i$ for all $i \in \{1, \ldots, N\}$. Each EAT channel can be decomposed as $\mathcal{M}_i = \mathcal{T}_i \circ \mathcal{M}'_i$ where $\mathcal{M}'_i : R_{i-1} \to$



(a) EAT process

(b) Single round

Figure 3.1: Diagrammatic depiction of the EAT process: (a) the overall process; (b) the process captured by the min-tradeoff function (Definition 3.7.2). They are related by the Entropy Accumulation Theorem (Theorems 3.7.3 and 3.7.4). Note that $\mathcal{M}_N$ may be viewed as outputting a trivial register, $R_N \cong \mathbb{C}$, which we have suppressed.

$S_iP_iR_i$ is any CPTP map and $\mathcal{T}_i : S_iP_i \to X_i$ is of the following form:

$$\mathcal{T}_i(W_{S_iP_i}) = \sum_{y\in\mathcal{Y},z\in\mathcal{Z}} (\Pi_{S_i,y} \otimes \Pi_{P_i,z}) W_{S_iP_i} (\Pi_{S_i,y} \otimes \Pi_{P_i,z}) \otimes |t(y,z)\rangle\langle t(y,z)|_{X_i}\,, \tag{3.53}$$

where $\{\Pi_{S_i,y} : y \in \mathcal{Y}\}$ and $\{\Pi_{P_i,z} : z \in \mathcal{Z}\}$ are families of mutually orthogonal projectors on S_i and P_i, the function $t : \mathcal{Y} \times \mathcal{Z} \to \mathcal{X}$ is a deterministic function, and $\mathcal{X}_i$ are finite-dimensional classical systems.

In order to apply the EAT, we demand the output state of the EAT maps

$$\rho_{S_1^N P_1^N X_1^N E} = (\mathcal{M}_N \circ \cdots \circ \mathcal{M}_1 \otimes \mathrm{id}_E)(\rho_{R_0E}), \tag{3.54}$$

where $\rho_{R_0E} \in \mathrm{D}(R_0 \otimes E)$ is any input state, needs to satisfy the Markov conditions (see Definition 2.2.39):

$$S_1^{i-1} \leftrightarrow P_1^{i-1}E \leftrightarrow P_i \qquad \forall i \in \{1,\ldots,N\}. \tag{3.55}$$

In the above definition, our choice of referring to the spaces by S_i and P_i is to stand for *secret* and *public* registers respectively as they are the natural interpretations of these registers in terms of cryptography. One can then see the X register is a *testing* register which is a function of the secret and public registers. The intuition of the above definition is that the testing structure characterizes how one constructs a test to determine the behavior of an on-going process per round, which is useful for determining how entropy accumulates after each round. The use of mutually orthogonal projectors generalizes the notion of X_i being a deterministic function of classical registers S_i and P_i to the quantum regime, where one, for example, may want to check subspaces without destroying coherence within the subspace. The Markovian behavior restriction [Eq. (3.55)] guarantees that the process (defined by the sequential application of $\mathcal{M}_i$) does not *a priori* destroy entropy being accumulated in the S_1^N registers. It does this by guaranteeing that, for all $i \in [N]$, the public information in round i, P_i, is not correlated to the generated secret information of previous rounds, S_1^{i-1}, even though it may be correlated to the public information of previous rounds, P_1^{i-1}, and the ancillary system E.

We notice that the Markov conditions can be trivially satisfied if all public announcements P_i's are seeded by randomness, which is the case for device-independent processes. However, one advantage of device-dependent QKD over device-independent QKD is its ability to have more complicated public announcement structures. Therefore, we would hope that the EAT can be applied to QKD protocols whose public announcements are not simply seeded with randomness.

Another important definition is the min- and max-tradeoff functions. Let $\mathbb{P}(\mathcal{X})$ denote the probability distribution over the alphabet $\mathcal{X}$.

Definition 3.7.2 (min- and max-tradeoff functions)**.** A function $f : \mathbb{P}(\mathcal{X}) \to \mathbb{R}$ is called a *min-* or *max-tradeoff function* for $\mathcal{M}_i$ if it satisfies

$$f(\boldsymbol{q}) \leq \inf_{\nu \in \Sigma_i(\boldsymbol{q})} H(S_i|P_iR)_\nu \text{ or } f(\boldsymbol{q}) \geq \sup_{\nu \in \Sigma_i(\boldsymbol{q})} H(S_i|P_iR)_\nu, \tag{3.56}$$

respectively, where

$$\Sigma_i(\boldsymbol{q}) := \{\nu_{S_iP_iX_iR_iR} = (\mathcal{M}_i \otimes \mathrm{id}_R)(\omega_{R_{i-1}R}) : \omega \in \mathrm{D}(R_{i-1} \otimes R) \text{ and } \nu_{X_i} = \boldsymbol{q}\}, \tag{3.57}$$

where R is isomorphic to R_{i-1}.

The second-order terms in the entropy accumulation theorem can depend on properties of the min-tradeoff function f. In particular, we define

$$\mathsf{Max}(f) := \max_{\boldsymbol{q} \in \mathbb{P}(\mathcal{X})} f(\boldsymbol{q}), \tag{3.58}$$

$$\mathsf{Min}(f) := \min_{\boldsymbol{q} \in \mathbb{P}(\mathcal{X})} f(\boldsymbol{q}), \tag{3.59}$$

$$\mathsf{Min}_\Sigma(f) := \min_{\boldsymbol{q} : \Sigma_i(\boldsymbol{q}) \neq \emptyset} f(\boldsymbol{q}), \tag{3.60}$$

$$\mathsf{Var}(f) := \max_{\boldsymbol{q} : \Sigma_i(\boldsymbol{q}) \neq \emptyset} \Big[\sum_{x \in \mathcal{X}} q(x) f(\delta_x)^2 - \Big(\sum_{x \in \mathcal{X}} q(x) f(\delta_x) \Big)^2 \Big], \tag{3.61}$$

where δ_x is the probability distribution over $\mathcal{X}$ defined by $\delta_x(x) = 1$ and $\delta_x(x') = 0$ for all $x' \neq x$.

We state two versions of the entropy accumulation theorem (EAT).

Theorem 3.7.3 (EAT version 1, [97, Theorem 4.4])**.** Let $\mathcal{M}_1, \ldots, \mathcal{M}_N$ and $\rho_{S_1^N P_1^N X_1^N E}$ be such that Eq. (3.54) and the Markov conditions [Eq. (3.55)] hold. Let $h \in \mathbb{R}$, let f be an affine min-tradeoff function for $\mathcal{M}_1, \ldots, \mathcal{M}_N$, and let $\varepsilon \in (0, 1)$. Then for any event $\Omega \subseteq \mathcal{X}^N$ that implies $f(\mathsf{freq}(X_1^N)) \geq h$,

$$H_{\text{min}}^{\varepsilon}(S_1^N | P_1^N E)_{\rho|\Omega} > Nh - c_1\sqrt{N} \tag{3.62}$$

holds for $c_1 = 2(\log_2(1 + 2d_S) + \lceil \|\nabla f\|_\infty \rceil)\sqrt{1 - 2\log_2(\varepsilon\rho(\Omega))}$, where d_S is the maximum dimension of the systems S_i.

The second-order correction term can be improved. The improvement leads to the second version of EAT. We also note that this version of EAT allows us to consider infrequent testing by the construction of crossover min-tradeoff functions as explained below.

Theorem 3.7.4 (EAT version 2, [98, Theorem V.2])**.** Let $\mathcal{M}_1, \ldots, \mathcal{M}_N$ and $\rho_{S_1^N P_1^N X_1^N E}$ be such that Eq. (3.54) and the Markov conditions [Eq. (3.55)] hold. Let $h \in \mathbb{R}$, let f be an affine min-tradeoff function for $\mathcal{M}_1, \ldots, \mathcal{M}_N$, and let $\varepsilon \in (0, 1)$. Let $\alpha \in (1, 2)$. Then for any event $\Omega \subseteq \mathcal{X}^N$ that implies $f(\mathsf{freq}(X_1^N)) \geq h$,

$$H_{\text{min}}^{\varepsilon}(S_1^N | P_1^N E)_{\rho|\Omega} > Nh - N\frac{(\alpha - 1)\ln 2}{2}V^2 - \frac{1}{\alpha - 1}\log_2 \frac{2}{\epsilon^2 \rho[\Omega]^2} - N(\alpha - 1)^2 K_\alpha, \tag{3.63}$$

with

$$V = \sqrt{\mathsf{Var}(f) + 2} + \log_2(2d_S^2 + 1) \tag{3.64}$$

$$K_\alpha = \frac{1}{6(2 - \alpha)^3 \ln 2} 2^{(\alpha - 1)[2\log_2 d_S + \mathsf{Max}(f) - \mathsf{Min}_\Sigma(f)]} \ln^3\left(2^{2\log_2 d_S + \mathsf{Max}(f) - \mathsf{Min}_\Sigma(f)} + e^2\right). \tag{3.65}$$

When the systems S_i are classical, $2\log_2 d_S$ in Eq. (3.65) can be replaced by $\log_2 d_S$. When α is chosen to be

$$\alpha = 1 + \frac{\sqrt{2\log_2 \frac{2}{\rho[\Omega]^2 \epsilon^2}}}{V\sqrt{N \ln 2}}, \tag{3.66}$$

the statement is simplified to be

$$H_{\text{min}}^{\varepsilon}(S_1^N | P_1^N E)_{\rho|\Omega} > Nh - c\sqrt{N} - c', \tag{3.67}$$

where

$$c = \sqrt{2\ln 2}\Big[\log_2(2d_S^2+1) + \sqrt{2+\mathsf{Var}(f)}\Big]\sqrt{1-2\log_2(\varepsilon\rho[\Omega])}\ , \tag{3.68}$$

$$c' = \frac{35\big[1-2\log_2(\varepsilon\rho[\Omega])\big]}{\big[\log_2(2d_S^2+1)+\sqrt{2+\mathsf{Var}(f)}\big]^2} 2^{2\log_2 d_S + \mathsf{Max}(f) - \mathsf{Min}_\Sigma(f)} \ln^3\Big(2^{2\log_2 d_S + \mathsf{Max}(f) - \mathsf{Min}_\Sigma(f)} + e^2\Big). \tag{3.69}$$

In this statement, the parameter α is related to α-Rényi divergence and a chain rule about conditional α-Rényi entropy is used. In the end, α can be optimized. The specific choice of α in Eq. (3.66) leads to a nicely looking statement in Eq. (3.67). However, it is in general not optimal for applications. In applications of this version of entropy accumulation theorem, one often needs to optimize the choice of α in Eq. (3.63).

To handle infrequent testing, according to [98, Definition V. 4], we can express an EAT channel $\mathcal{M}_{i,R_{i-1}\to X_iS_iP_iR_i}$ with testing probability $\gamma \in [0,1]$ as

$$\mathcal{M}_{i,R_{i-1}\to X_iS_iP_iR_i}(\cdot) = \gamma\mathcal{M}^{\text{test}}_{i,R_{i-1}\to X_iS_iP_iR_i}(\cdot) + (1-\gamma)\mathcal{M}^{\text{data}}_{i,R_{i-1}\to X_iS_iP_iR_i}(\cdot)\otimes |\bot\rangle\langle\bot|_{X_i}\ , \tag{3.70}$$

such that $\mathcal{M}_i^{\text{test}}$ never outputs the symbol $\bot$ on X_i. Let $\mathcal{X}'$ be the alphabet that is obtained from removing $\bot$ from the alphabet $\mathcal{X}$. A crossover min-tradeoff function is a type of min-tradeoff function that only uses the statistics from the testing parts $\mathcal{M}_i^{\text{test}}$ and is defined below.

Definition 3.7.5 (Crossover min-tradeoff function)**.** Let $\mathcal{M}_i$ be a channel with a testing probability γ as defined above. The crossover min-tradeoff function is an affine function $h : \mathbb{P}(\mathcal{X}') \to \mathbb{R}$ satisfying

$$h(\boldsymbol{q'}) \leq \min_{1\leq i\leq n}\ \min_{\nu\in\Sigma_i'(\boldsymbol{q'})} H(S_i|P_iR)_\nu, \quad \forall \boldsymbol{q'} \in \mathbb{P}(\mathcal{X}'), \tag{3.71}$$

where the set of quantum states is

$$\begin{aligned}\Sigma_i'(\boldsymbol{q'}) := \{\nu_{S_iP_iX_iR} = (\mathcal{M}_i\otimes \mathrm{id}_{R_i})(\omega_{Q_iR}) : \omega \in \mathrm{D}(Q_i\otimes R)\\ \text{and } [(\mathcal{M}_i^{\text{test}}\otimes \mathrm{id}_{R_i})(\omega_{Q_iR})]_{X_i} = \boldsymbol{q'}\}.\end{aligned} \tag{3.72}$$

The crossover min-tradeoff function h automatically defines a min-tradeoff function $f : \mathbb{P}(\mathcal{X}) \to \mathbb{R}$ by:

$$f(\delta_x) = \mathsf{Max}(h) + \frac{1}{\gamma}[h(\delta_x) - \mathsf{Max}(h)] \qquad \forall x\in\mathcal{X}' \tag{3.73}$$

$$f(\delta_\bot) = \mathsf{Max}(h). \tag{3.74}$$

Moreover, we have the relations

$$\mathsf{Max}(f) = \mathsf{Max}(h) \tag{3.75}$$

$$\mathsf{Min}(f) = (1 - \frac{1}{\gamma})\mathsf{Max}(h) + \frac{1}{\gamma}\mathsf{Min}(h) \tag{3.76}$$

$$\mathsf{Min}_{\Sigma}(f) \geq \mathsf{Min}(h) \tag{3.77}$$

$$\mathsf{Var}(f) \leq \frac{1}{\gamma}[\mathsf{Max}(h) - \mathsf{Min}(h)]^2. \tag{3.78}$$

It can also be seen that for $\boldsymbol{q'} \in \mathbb{P}(\mathcal{X}')$,

$$h(\boldsymbol{q'}) = f(\boldsymbol{q}) \tag{3.79}$$

for $\boldsymbol{q} = (\boldsymbol{q'}^T,\ 1-\gamma)^T$ by this construction. Moreover, every frequency distribution in the set of acceptable frequency distribution $\mathcal{Q}$ over $\mathcal{X}$ always has the value of $1-\gamma$ in the position corresponding to the $\perp$ symbol. This means for any $\boldsymbol{q} \in \mathcal{Q}$, one can always find $\boldsymbol{q'} \in \mathbb{P}(\mathcal{X}')$ such that Eq. (3.79) holds.

A recent work by Dupuis [99] establishes a version of the leftover hashing lemma (cf. Theorem 3.3.2) with the sandwiched α-Rényi entropy instead of the smooth min-entropy. As the proofs of two early versions of EAT use sandwiched α-Rényi entropy as an intermediate step before converting to the smooth min-entropy, a version of EAT based on the sandwiched α-Rényi entropy can potentially give a tighter key rate bound. This improved version of EAT is presented in [99] and is also discussed in detail in our forthcoming paper [83]. We direct readers there for precise theorem statements.

We state the application of EAT to the device-dependent QKD in Theorem 8.2.3 as well as in Theorem 8.2.4.

3.8 Remarks

3.8.1 Phase error correction framework

The phase error correction-based approach was initially developed by Mayers [131] and Shor, Preskill [132]. This framework was further developed by Koashi [133] and Hayashi [134].

The argument by Shor and Preskill involves a reduction to an entanglement distillation protocol based on Calderbank-Shor-Steane (CSS) quantum error correcting codes (QECC). This initial proof requires one to choose carefully methods for error correction and privacy amplification such that the CSS code structure is preserved. By encrypting the error correction step, it is shown [135] that one can decouple the error correction from the privacy amplification when a certain constraint from the CSS-QECC is satisfied. Reference [133] removes this constraint. It is based on a complementarity argument. The core of Koashi's qubit-based security proof [133] is to regard the key after the error correction as the outcome of the Z-basis measurements on N virtual qubits $\mathcal{K}^{\otimes N}$. It involves a reduction from the two-party private and random key distribution to a single-party private and random number generation by constructing a single-party virtual protocol called the squashing protocol. Formally, if there exists a quantum operation Λ called squashing operation that converts any state $\rho \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ to a state $\Lambda(\rho) \in \mathrm{D}(\mathcal{H}_R \otimes \mathcal{K}^{\otimes N})$, where $\mathcal{H}_R$ is some virtual ancillary system, and the standard Z-basis measurement on $\mathcal{K}^{\otimes N}$ gives the same result as the key after the error correction step in the actual protocol, then one can construct a single-party squashing protocol. If one can show that the squashing protocol is ε_{sec}-secret and the error correction protocol is ε_{cor}, then the actual protocol is $\varepsilon_{\text{sec}} + \varepsilon_{\text{cor}}$-secure. The security of the squashing protocol can be proved by phase error correction. The complementarity argument is basically the uncertainty principle of two complementary operators: If the X-basis measurement on $\mathcal{K}^{\otimes N}$ is completely certain, then the Z-basis measurement on $\mathcal{K}^{\otimes N}$ is completely random.

This framework can be easily adapted to BB84-type protocols and can give analytical solutions. The essential asymptotic key rate formula is a variation of the qubit-based formula:

$$R_\infty = 1 - h(e_{\text{ph}}) - h(e_{\text{bit}}), \tag{3.80}$$

where e_{ph} is the phase error rate and e_{bit} is the bit error rate. The core of a security proof is to have a correct definition of phase error and then bound the phase error rate. For practical DVQKD protocols, since the source is usually not a perfect single-photon source, one instead uses some version of Gottesman, Lo, Lütkenhaus and Preskill (GLLP) formula [136] in combination of the decoy state method [66, 67], which is used to estimate the contribution from the single-photon part. It is given [66] as

$$R_\infty = q\{Q_1[1 - h(e_1)] - Q_\mu f_{\text{EC}} h(E_\mu)\}, \tag{3.81}$$

where Q_μ and E_μ are the gain and the quantum bit error rate of the signal state, respectively, Q_1 and e_1 is the gain and quantum bit error rate of of the single-photon state, respectively, and q is a sifting factor. We note that a finite-key analysis based on this approach can also be done as Ref. [133] gives a tool to do so.

However, it is usually difficult to apply this framework to other protocols that are quite different from BB84, like CVQKD protocols and high-dimensional QKD. Another disadvantage is that in many security proofs based on this framework, the phase error rate is loosely bounded so that the key rate can also be loose. It is also arguable that the finite-key analysis based on this framework is less understandable since the proof is not modular and sometimes details are not clearly presented as some authors assume formulas like the GLLP formula can be trivially adapted to new protocols and they study only the parameter estimation step in the finite key analysis without explaining other finite-size effects.

Finally, we note that there has been an effort to unify two frameworks [137]. Reference [137] claims that both frameworks give essentially the same key rate (except an inessential constant). It is interesting to explore this direction and to understand more about the mathematical structures behind these two frameworks.

3.8.2 Entropic uncertainty relation

In addition to aforementioned proof techniques, entropic uncertainty relations (see [138] for review) can be used to prove security of QKD protocols in both frameworks. In the presence of quantum side information, there are two useful entropic uncertainty relations that can be used in the asymptotic or finite-key scenarios: the entropic uncertainty relation for von Neumann entropy [139–141] and the one for smooth entropies [142].

The setting is as follows. Let A, B, C be three quantum systems. These systems A, B and C can be in any quantum state (potentially an entangled state). Suppose there are two POVMs $X = \{M_x\}$ and $Z = \{N_z\}$ that act on system A. With an abuse of notations, we use X and Z to label measurement outcomes of the corresponding POVMs. The uncertainty relation for the von Neumann entropy [139–141] reads

$$H(X|B) + H(Z|C) \geq q, \tag{3.82}$$

where

$$q := \log_2(\frac{1}{c}) \text{ where } c := \max_{x,z} \left\| \sqrt{M_x}\sqrt{N_z} \right\|_\infty^2. \tag{3.83}$$

Similarly, the entropic uncertainty relation for smooth entropies [142] reads

$$H_{\text{min}}^{\varepsilon}(X|B) + H_{\text{max}}^{\varepsilon}(Z|C) \geq q. \tag{3.84}$$

In the context of QKD security proofs, entropic uncertainty relations are typically used for BB84-type protocols as there are naturally two POVMs that give nontrivial values of q. If Alice performs the key map, then A is typically referred as Alice's qubits used for key generation, B is Bob's system and C is Eve's system E. The key is obtained by measuring Alice's qubit using the POVM Z. Then in the asymptotic scenario, one bounds Eve's ignorance about the key (measurement outcomes associated with POVM Z) $H(Z|E)$ by $H(Z|E) \geq q - H(X|B)$. For the finite key analysis, we swap min- and max- entropies in Eq. (3.84) and bound the smooth min-entropy $H_{\text{min}}^{\varepsilon}(Z|E)$ by $q - H_{\text{max}}^{\varepsilon}(X|B)$.

Entropic uncertainty relations have also been used for CVQKD. Based on an extension of the entropic uncertainty relation of the smooth entropies to the infinite-dimensional scenario, Ref. [143] provides a security analysis of an entanglement-based CV protocol using two-mode squeezed vacuum states with homodyne detection. However, the key rate is pessimistic and does not reach the key rate obtained under Gaussian collective attacks when the number of signals increases to infinity.

In summary, entropic uncertainty relations can be used in QKD security proofs. Unfortunately, it either is limited to certain types of protocols or gives loose key rates.

Chapter 4

Numerical methods for key rate calculation

To aid our discussions about security proofs of specific protocols in the following chapters, we discuss numerical methods for the secret key rate calculation. We start with a review of the formulation of asymptotic key rate [144]. It provides the basics for the security proof presented in Chapters 6 and 7 and captures the essential components of the two numerical methods for finite key analysis. We summarize several analytical tools [85, 86, 90, 92] that can ease the numerical calculation by reducing the dimension of the optimization problem. We then present key ideas from our work [84] for numerical finite-key analysis which follows Renner's framework [89]. Finally, we present another finite-key analysis method using the entropy accumulation theorem based on our ongoing work [83], of which we discuss applications in Chapter 8.

4.1 Asymptotic key rate optimization

In the asymptotic limit, one can typically restrict to the i.i.d. collective attacks since it is asymptotically optimal for virtually any protocol (as long as the protocol is invariant under permutation of rounds). The asymptotic key rate against i.i.d. collective attacks is given by the Devetak-Winter formula [145]. In particular, one can rewrite it in terms of quantum conditional entropy. Recall in the generic QKD protocol described in Protocol 2.1 (similarly in Protocol 2.2 and Protocol 2.3), after the key map step, the party who applied the key map holds $\boldsymbol{z}$ while the other party holds $\tilde{\boldsymbol{y}}$ in a direct reconciliation scheme or $\tilde{\boldsymbol{x}}$ in a reverse reconciliation scheme. Due to the i.i.d. structure, each position in the key string $\boldsymbol{z}$ is independently and identically distributed according to some probability distribution. Let Z be a random variable which is distributed according to this probability distribution. Similarly, we can define random variables X and Y associated with $\tilde{\boldsymbol{x}}$ and $\tilde{\boldsymbol{y}}$, respectively.

For simplicity of writing, we use the direct reconciliation scenario as an example. For the reverse reconciliation protocols, one can simply exchange the role of Alice and Bob starting from the key map step, and thus replace Y by X. Thus, the asymptotic key rate is [145]

$$\begin{aligned} R_\infty &= I(Z\!:\!Y) - \max_{\rho\in\mathbf{S}} \chi(Z\!:\![E]) \\ &= \min_{\rho\in\mathbf{S}} H(Z|[E]) - H(Z|Y), \end{aligned} \tag{4.1}$$

where the set $\mathbf{S}$ contains all density operators that are compatible with observed statistics (see Section 4.1.1 for detailed explanations). We use again the notation $[E]$ to denote all Eve's registers including her initial register E attached in the quantum phase of the protocol and all classical communication registers. In this formula, it is assumed that the error correction protocol can reach the efficiency given by the Shannon limit. Since the practical error correction protocols cannot achieve such an efficiency, we replace the term $H(Z|Y)$ by the actual amount of information leakage in the protocol. As defined in Eq. (3.32), let δ_{EC} denote the number of bits of the communication cost in the error correction protocol per key generation round, i.e. $\text{leak}_{\text{EC}}/n$ for a total number of n signals used for the key generation. We rewrite this formula in the following generic way, which also includes the probability that a given signal is used to generate keys as

$$R_\infty = p_{\text{pass}}[\min_{\rho\in\mathbf{S}} H(Z|[E]) - \delta_{\text{EC}}]; \tag{4.2}$$

i.e., we express the key rate as the number of secret bits per signal sent. This expression is exactly the same as Eq. (3.33) which we derived from the finite key rate R_N^{coll} against collective attacks by taking the limit $N \to \infty$.

In the following subsections, we follow the works [144, 146] to express this key rate problem as a convex optimization problem. We also consider a variation of the problem formulation in [83].

4.1.1 Constraint set

The optimization variable is the joint density matrix ρ_{AB} for entanglement-based and prepare-and-measure schemes. For simplicity of writing, we drop the subscript when it causes no confusion. Besides the obvious requirement for ρ to be a valid density matrix, we can divide the constraints into two sets according to origins of the constraints: observational and reduced density operator constraint sets.

From the parameter estimation step, we have constraints from measurements. In the asymptotic limit, each observed frequency distribution is indeed a probability distribution. The set of states ρ satisfying the observational constraints is given by

$$S_O = \{\rho \geq 0 : \mathrm{Tr}\big(\rho M_x^A \otimes M_y^B\big) = p(x,y), x \in \Sigma_A, y \in \Sigma_B\}, \tag{4.3}$$

where Σ_A and Σ_B are alphabets for Alice's and Bob's measurement outcomes, respectively. We can also calculate expectation values of Hermitian observables from the probability distribution. In general, we can consider Hermitian observables $\{\Gamma_i \in \mathrm{Herm}(\mathcal{H}_A \otimes \mathcal{H}_B)\}$ with the corresponding expectation values $\{\gamma_i \in \mathbb{R}\}$.

For prepare-and-measure protocols, we have an additional promise that the reduced density operator ρ_A from the source-replacement scheme is unchanged by Eve.

$$\begin{aligned} S_R &= \{\rho \geq 0 : \mathrm{Tr}_B(\rho) = \rho_A\} \\ &= \{\rho \geq 0 : \mathrm{Tr}(\rho \ \Theta_j \otimes \mathbb{1}_B) = \theta_j, \forall j = 1, \ldots, \dim(A)^2\}, \end{aligned} \tag{4.4}$$

where $\theta_j = \mathrm{Tr}(\Theta_j \rho_A)$ and $\{\Theta_j : j = 1, \ldots, \dim(A)^2\}$ is an orthonormal basis for the real vector space of Hermitian matrices on system A. Thus, $\mathbf{S} = S_O \cap S_R$.

We note that in the measurement-device-independent scheme, one also includes a classical register C that stores Charlie's classical announcements and thus one optimizes the joint density matrix $\rho_{ABC} \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathbb{C}^{|\Lambda|})$ where Λ is the alphabet for Charlie's announcements. In this case, since we apply the source-replacement scheme to both Alice's and Bob's sources, the reduced density operator ρ_{AB} is unchanged by Eve. Therefore, registers A and B together play the role of A in Eq. (4.4).

As our constraints are linear constraints and a positive semidefinite constraint, the feasible set is a spectrahedron, i.e., the intersection of an affine manifold and the positive semidefinite cone.

4.1.2 Modeling of postprocessing

The parameter estimation step gives us possible states ρ_{AB} immediately after the state distribution but before any measurements. In the worst-case scenario, Eve holds a purification of ρ_{AB}. Let ρ_{ABE} denote a purification. In Eq. (4.2), each state that we want to evaluate the conditional entropy is after Alice's and Bob's measurements as well as the classical postprocessing steps including announcement, sifting and key map steps, and it is before the error correction step. Thus, we would like to describe the linear CPTP map that transforms a state ρ_{ABE} to a final state after aforementioned steps, which is a classical-classical-quantum state ρ_{ZPE}, where the register P holds all the public information.

In the measurement step, Alice holds a POVM $\{M_x^A : x \in \Sigma_A\} = \{M_{(a,\alpha)}^A : a \in S^A, \alpha \in O^A\}$, and Bob holds a POVM $\{M_y^B : y \in \Sigma_B\} = \{M_{(b,\beta)}^B : b \in S^B, \beta \in O^B\}$ for some index sets S^A, O^A, S^B and O^B.

After applying their POVMs to the state ρ_{ABE} and announcing public information a, b, if we keep the post-measurement state, then the joint state $\rho^{(1)}_{\widetilde{A}\overline{A}\widetilde{B}\overline{B}ABE}$ can be obtained by applying a CPTP map Φ_{M} as follows:

$$\begin{aligned}\rho^{(1)}_{\widetilde{A}\overline{A}\widetilde{B}\overline{B}ABE} := \Phi_{\mathrm{M}}(\rho_{ABE}) = &\sum_{a,b}\sum_{\alpha,\beta} |a\rangle\langle a|_{\widetilde{A}} \otimes |\alpha\rangle\langle\alpha|_{\overline{A}} \otimes |b\rangle\langle b|_{\widetilde{B}} \otimes |\beta\rangle\langle\beta|_{\overline{B}} \\ &\otimes\Big(\sqrt{M^A_{(a,\alpha)}} \otimes \sqrt{M^B_{(b,\beta)}} \otimes \mathbb{1}_E\Big)\rho_{ABE}\Big(\sqrt{M^A_{(a,\alpha)}} \otimes \sqrt{M^B_{(b,\beta)}} \otimes \mathbb{1}_E\Big).\end{aligned} \tag{4.5}$$

The general sifting step basically performs a partition of the set $S^A \times S^B$ as $S^A \times S^B = \mathbf{K} \cup \mathbf{D}$, where $\mathbf{K}$ is the set of announcements to be kept and $\mathbf{D}$ is the set of announcements to be discarded. We define a projector $\Pi_{\mathbf{K}}$ which projects onto the subspace of $\mathcal{H}_{\widetilde{A}} \otimes \mathcal{H}_{\widetilde{B}}$ corresponding to the set $\mathbf{K}$ and similarly the projector $\Pi_{\mathbf{D}} := \mathbb{1} - \Pi_{\mathbf{K}}$ projects onto the subspace related to the set $\mathbf{D}$. Explicitly,

$$\Pi_{\mathbf{K}} = \sum_{(a,b)\in\mathbf{K}} |a\rangle\langle a|_{\widetilde{A}} \otimes |b\rangle\langle b|_{\widetilde{B}}\,. \tag{4.6}$$

Thus, we can describe the sifting step that acts on $\rho^{(1)}_{\widetilde{A}\overline{A}\widetilde{B}\overline{B}ABE}$ as a linear map Φ_{S} and obtain the state $\rho^{(2)}_{\widetilde{A}\overline{A}\widetilde{B}\overline{B}ABES}$ after the map

$$\rho^{(2)}_{\widetilde{A}\overline{A}\widetilde{B}\overline{B}ABES} := \Phi_{\mathrm{S}}\big(\rho^{(1)}_{\widetilde{A}\overline{A}\widetilde{B}\overline{B}E}\big) = \Pi_{\mathbf{K}}\rho^{(1)}_{\widetilde{A}\overline{A}\widetilde{B}\overline{B}ABE}\Pi_{\mathbf{K}} \otimes |\mathbf{K}\rangle\langle\mathbf{K}|_S + \Pi_{\mathbf{D}}\rho^{(1)}_{\widetilde{A}\overline{A}\widetilde{B}\overline{B}ABE}\Pi_{\mathbf{D}} \otimes |\mathbf{D}\rangle\langle\mathbf{D}|_S\,. \tag{4.7}$$

This map is also CPTP.

The key map step applies the key map function f_{KM}, which is a function of one party's private information and all public information. For the direct reconciliation scheme, $f_{\mathrm{KM}} : O^A \times S^A \times S^B \to \{0, 1, \ldots, d-1\}$, where d is the size of key alphabet[6]. The action of the key map can be represented by a CPTP map Φ_{KM} with a single Kraus operator which is also an isometry:

$$V_{\mathrm{KM}} = \sum_{\alpha\in O^A}\sum_{a\in S^A, b\in S^B} |f_{\mathrm{KM}}(\alpha, a, b)\rangle_Z \otimes |a\rangle\langle a|_{\widetilde{A}} \otimes |\alpha\rangle\langle\alpha|_{\overline{A}} \otimes |b\rangle\langle b|_{\widetilde{B}}\,. \tag{4.8}$$

Therefore, the final state after the key map is

$$\rho^{(3)}_{Z\widetilde{A}\overline{A}\widetilde{B}\overline{B}ABES} = \Phi_{\mathrm{KM}}\big(\rho^{(2)}_{\widetilde{A}\overline{A}\widetilde{B}\overline{B}ABES}\big) = V_{\mathrm{KM}}\rho^{(2)}_{\widetilde{A}\overline{A}\widetilde{B}\overline{B}ABES}V^{\dagger}_{\mathrm{KM}}. \tag{4.9}$$

[6] In many protocols like BB84, there is already a natural way of partitioning measurement outcomes into private and public information (bit value and basis choice). Moreover, the private information has been mapped to the key alphabet directly in the data partition. In this case, the key map $f_{\mathrm{KM}}(\alpha, a, b) = \alpha$ just outputs the value of the private information.

The public announcement register P mentioned previously consists of registers $\widetilde{A}\widetilde{B}S$. Since Eve has access to all public announcements, $[E]$ denotes $\widetilde{A}\widetilde{B}ES$. The final state on which we evaluate the entropy is given by the overall postprocessing map Φ_{PP} as

$$\rho^{(3)}_{Z[E]} = \Phi_{\text{PP}}(\rho_{ABE}) := (\text{Tr}_{AB\overline{AB}} \circ \Phi_{\text{KM}} \circ \Phi_{\text{S}} \circ \Phi_{\text{M}})(\rho_{ABE}). \tag{4.10}$$

We note that in this formulation, Φ_{PP} is a CPTP map. To handle sifting, we allow the key map f_{KM} to output a special symbol $\perp$ whenever the round is discarded due to the sifting step. Intuitively, those rounds output a fixed outcome regardless the private information and thus no secret keys are expected to generated from those rounds. Formally, one can use properties of the objective function in the optimization to show that all rounds with $\perp$ symbol do not contribute to the entropy. In Ref. [144], the postprocessing map is described as a CPTNI map by removing rounds discarded in the sifting step. In Appendix A, one can also formulate this postprocessing map Φ_{PP} in an isometric fashion to preserve the purity of the output state $\rho^{(3)}$. This formulation is useful to rewrite the objective function in terms of quantum relative entropy. We note that these formulations give the same reduced density operator $\rho^{(3)}_{Z[E]}$ for the evaluation of $H(Z|[E])$. They are equivalent in this sense. When we drop the discarded rounds, we call this description of postprocessing map as $\mathcal{G}$ in the same way as presented in [144].

4.1.3 Objective function

We present two ways to write the objective function. The first one is based on the quantum relative entropy and is presented in [144, 146] following the work [147]. The second approach is directly based on the quantum conditional entropy and is presented in [83]. In both formulations, we use the fact that in the worst-case scenario, Eve holds a purification of ρ_{AB}. In this section, we write $[E]$ as two registers P, E where the register P contains all public information and the register E is Eve's system in the initial purification of ρ_{AB}.

Quantum relative entropy

The purpose of writing $H(Z|PE)$ to quantum relative entropy is two-fold: (a) The objective function can be directly seen as a convex function since quantum relative entropy is jointly convex in both arguments as mentioned in Section 2.2.3; (b) The state used in the quantum relative entropy only involves Alice's and Bob's systems without the need to find Eve's state. To rewrite $H(Z|PE)$ in terms of quantum relative entropy, we adapt [147, Theorem 1] to the QKD setting.

Theorem 4.1.1 ([147, Theorem 1]). Let ρ_{RABPE} be a pure state and let $Z = \{Z_j^R : j \in \{0,\dots,d-1\}\}$ be a projective measurement that only acts on register R. Define an isometry V_Z that models the Z measurement on system R and stores the measurement outcomes in a register M_Z as

$$V_Z := \sum_j |j\rangle_{M_Z} \otimes Z_j^R. \tag{4.11}$$

Define

$$\tilde{\rho}_{M_Z RABPE} := V_Z \rho_{RABPE} V_Z^\dagger \tag{4.12}$$

and $\tilde{\rho}_{M_Z PE} = \mathrm{Tr}_{RAB}(\tilde{\rho}_{M_Z RABPE})$. Then

$$H(Z|PE)_{\tilde{\rho}} = H(\tilde{\rho}_{M_Z PE}) - H(\rho_{PE}) = D\Big(\rho_{RABP} || \mathcal{Z}(\rho_{RABP})\Big), \tag{4.13}$$

where $\rho_{PE} = \tilde{\rho}_{PE}$ and $\mathcal{Z}$ is a quantum channel whose Kraus operators are given by $\{Z_j^R \otimes \mathbb{1}_{AB}\}$.

In Appendix A, we present an equivalent formulation of the postprocessing map $\mathcal{G}$ in terms of an isometric representation. This formulation allows us to apply this theorem in the QKD setting where there exists a postprocessing map $\mathcal{G}$ such that $\rho_{RABP} = \mathcal{G}(\rho_{AB})$.

Thus, we can write the objective function in terms of the quantum relative entropy. It then becomes obvious that our optimization problem is a convex optimization problem, and specifically, a nonlinear SDP problem. We write the objective function based on the quantum relative entropy as

$$\mathcal{W}(\rho) := D\Big(\mathcal{G}(\rho) || (\mathcal{Z} \circ \mathcal{G})(\rho)\Big). \tag{4.14}$$

The gradient of the our objective function at a point $\rho > 0$ is

$$\nabla \mathcal{W}(\rho) = \mathcal{G}^\dagger\Big(\log_2[\mathcal{G}(\rho)]\Big) + \mathcal{G}^\dagger\Big(\log_2[(\mathcal{Z} \circ \mathcal{G})(\rho)]\Big), \tag{4.15}$$

Note that our matrix representation of $\nabla \mathcal{W}(\rho)$ does not involve an additional transpose which was initially presented in [144] since we absorb the occurring transposition in the definition of the matrix representation of gradient; i.e., we represent the gradient in the standard basis $|k\rangle$ used for the density matrix ρ as

$$\nabla \mathcal{W}(\rho) = \sum_{j,k} d_{jk} |k\rangle\langle j|, \text{ with } d_{jk} := \left.\frac{\partial \mathcal{W}(\sigma)}{\partial \langle j| \sigma |k\rangle}\right|_{\sigma=\rho}. \tag{4.16}$$

We note that when $\mathcal{G}(\rho)$ is singular (which unfortunately happens almost all the time given the construction of the map $\mathcal{G}$), the gradient is not well-defined. One may overcome the issue by introducing a small perturbation as it is done in [144]:

$$\mathcal{G}_\epsilon(\rho) = \mathcal{D}_\epsilon \circ \mathcal{G}(\rho), \tag{4.17}$$

where $\mathcal{D}_\epsilon$ is a depolarizing channel with the depolarizing probability ϵ (see Definition 2.2.28). This perturbed definition $\mathcal{G}_\epsilon$ guarantees the positive definiteness of the output state so that the gradient is well defined. We write $\mathcal{W}_\epsilon(\rho)$ for $D\Big(\mathcal{G}_\epsilon(\rho)||(\mathcal{Z} \circ \mathcal{G}_\epsilon)(\rho)\Big)$. We note that an alternative solution exists by a suitable regularization of $\mathcal{G}$ map via facial reduction (see [87]).

Quantum conditional entropy

An alternative way to write the objective function is to directly show the quantum conditional entropy $H(Z|PE)$ in the key rate expression is a convex function and to find a way to represent Eve's conditional state in terms of Alice's and Bob's joint state. We write the objective function based on the quantum conditional entropy as

$$W(\rho) := H(Z|PE)_\rho, \tag{4.18}$$

for $\rho \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Proposition 4.1.2 gives us a more explicit expression of $W(\rho)$ in terms of $\rho \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and Alice's and Bob's joint POVM. The convexity of $W(\rho)$ is given in Proposition 4.1.3.

Proposition 4.1.2. Let $\{M_{zp} : z, p\}$ be Alice's and Bob's joint POVM which is regrouped according to be the public information p and the key value z. Let $M_p = \sum_z M_{zp}$. Then for $\rho \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$,

$$W(\rho) = \sum_{z,p} H\left(\mathcal{K}_{zp}(\rho)\right) - \sum_p H\left(\mathcal{K}_p(\rho)\right) \tag{4.19}$$

where $\mathcal{K}_{zp}(\rho) := K_{zp}\rho K_{zp}^\dagger$ with $K_{zp} = \sqrt{M_{zp}}$, and $\mathcal{K}_p(\rho) := K_p \rho K_p^\dagger$ with $K_p = \sqrt{M_p}$.

Proof. Recall from the definition of conditional entropy that for a classical-quantum state $\rho_{ZPE} = \sum_{z,p} |z\rangle\langle z| \otimes |p\rangle\langle p| \otimes \rho_E^{(z,p)}$, we have $H(Z|PE) = \sum_{z,p} H(\rho_E^{(z,p)}) - \sum_p H(\rho_E^{(p)})$, where $\rho_E^{(p)} = \sum_z \rho_E^{(z,p)}$. We need to find the marginal states $\rho_E^{(z,p)} = \mathrm{Tr}_{AB}[(M_{zp} \otimes \mathbb{1}_E)\rho_{ABE}]$. Without loss of generality, we can write the purifying state as $\rho_{ABE} = |\psi\rangle\langle\psi|$, where $|\psi\rangle = \sqrt{\rho} \otimes \mathbb{1}_E |\Phi^+\rangle$ and $|\Phi^+\rangle = \sum_i |i\rangle_{AB} |i\rangle_E$, from which we find

$$\begin{aligned} \rho_E^{(z,p)} &= \mathrm{Tr}_{AB}[(M_{zp} \otimes \mathbb{1}_E)\rho_{ABE}] \\ &= \mathrm{Tr}_{AB}[(\sqrt{\rho} M_{zp} \sqrt{\rho} \otimes \mathbb{1}_E) \left|\Phi^+\middle\rangle\middle\langle\Phi^+\right|] \\ &= \sum_{z,p} \left(\sqrt{\rho} M_{zp} \sqrt{\rho}\right)^T \ . \end{aligned} \tag{4.20}$$

By the fact that for any operator A, $(A^\dagger A)$ and $(A^\dagger A)^T$ have the same spectrum, it is the case that $H(\rho_E^{(z,p)}) = H\Big([(\sqrt{\rho}K_{zp})(K_{zp}\sqrt{\rho})]^T\Big) = H(K_{zp}\rho K_{zp}) = H(K_{zp}\rho K_{zp}^\dagger)$ and a similar expression holds for $H(\rho_E^{(p)})$. □

Proposition 4.1.3. The function $W(\rho)$ is convex.

Proof. Let $\bar{\rho}_{AB} = \sum_\lambda p_\lambda \rho_{AB}^\lambda$ be a convex decomposition of $\bar{\rho}_{AB}$. We want to show that

$$\sum_\lambda p_\lambda W(\rho_{AB}^\lambda) \geq W(\bar{\rho}_{AB}) \ . \tag{4.21}$$

Let ρ_{ABE}^λ be a purification of ρ_{AB}^λ. We define the state $\rho_{ABE\Lambda} = \sum_\lambda p_\lambda \rho_{ABE}^\lambda \otimes |\lambda\rangle\langle\lambda|_\Lambda$, which satisfies $\mathrm{Tr}_{E\Lambda}[\rho_{ABE\Lambda}] = \bar{\rho}_{AB}$. Then

$$\begin{aligned}\sum_\lambda p_\lambda W(\rho_{AB}^\lambda) &= \sum_\lambda p(\lambda) H(Z|PE)_{\Phi_{\mathrm{PP}}\otimes \mathrm{id}_E(\rho_{ABE}^\lambda)} \\ &= H(Z|PE\Lambda)_{\Phi_{\mathrm{PP}}\otimes \mathrm{id}_{E\Lambda}(\rho_{ABE\Lambda})} \\ &\geq H(Z|PE)_{\Phi_{\mathrm{PP}}\otimes \mathrm{id}_E(\bar{\rho}_{ABE})} \\ &= W(\bar{\rho}_{AB})\end{aligned} \tag{4.22}$$

where $\bar{\rho}_{ABE}$ is the purification of $\bar{\rho}_{AB}$ and where the inequality comes from sub-additivity because the entropy is minimized when Eve holds a purification of $\bar{\rho}_{AB}$. □

The gradient of W for $\rho > 0$ is given as

$$\nabla W(\rho) = -\sum_{z,p} \mathcal{K}_{zp}^\dagger\Big(\log_2 \mathcal{K}_{zp}(\rho)\Big) + \sum_p \mathcal{K}_p^\dagger\Big(\log_2 \mathcal{K}_p(\rho)\Big). \tag{4.23}$$

We can use the same perturbation technique in Eq. (4.17) that is used in the quantum relative entropy formulation to define the gradient for every $\rho \geq 0$. We similarly denote the perturbed version of this function as $W_\epsilon(\rho)$.

4.1.4 Numerical algorithm

The key rate optimization is a convex optimization problem since our objective function is convex and our feasible set is also convex. It can be written as

$$\begin{aligned}\underset{\rho}{\text{minimize}} \ & f(\rho) \\ \text{subject to } \ & \mathrm{Tr}(\rho\Gamma_i) = \gamma_i, \quad \forall i \in \Sigma \\ & \mathrm{Tr}_B(\rho) = \rho_A \\ & \mathrm{Tr}(\rho) = 1 \\ & \rho \geq 0\end{aligned} \tag{4.24}$$

where $f(\rho)$ is either the quantum relative entropy function $\mathcal{W}(\rho)$ in Eq. (4.14) or quantum conditional entropy function $W(\rho)$ in Eq. (4.18). As both $\text{Tr}(\rho) = 1$ and $\text{Tr}_B(\rho) = \rho_A$ can be written in the form of $\text{Tr}(\rho\Gamma_i) = \gamma_i$ by using $\mathbb{1}_{AB}$ and $\{\Theta_j \otimes \mathbb{1}_B\}$ in Eq. (4.4), we can consider the following form of the optimization problem:

$$\begin{aligned} &\underset{\rho}{\text{minimize}} \ f(\rho) \\ &\text{subject to} \ \text{Tr}(\rho\Gamma_i) = \gamma_i, \quad \forall i \in \Sigma' \\ &\qquad \rho \geq 0 \ . \end{aligned} \tag{4.25}$$

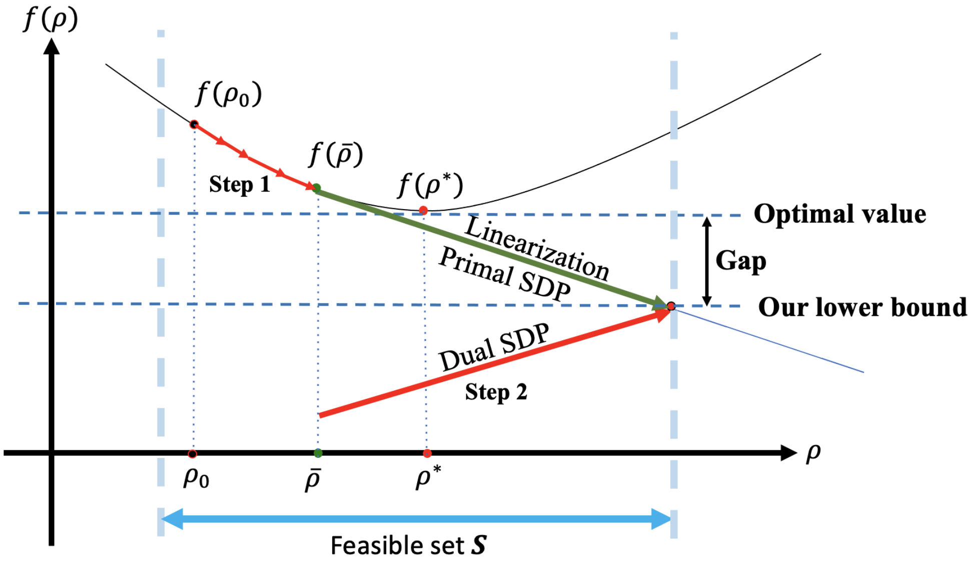


Figure 4.1: Illustration of the two-step method (Algorithm 4.1) in a one-dimensional abstraction. Red lines indicate the optimization steps performed in the algorithm. The gap between our lower bound from the second step and the true optimal value can be made small by finding $\overline{\rho}$ close to the true optimal point ρ^*.

In [144], a two-step procedure is proposed to provide a reliable lower bound on the secret key rate. In the first step, a nearly optimal solution $\overline{\rho}$ for ρ is found by the Frank-Wolfe method [148]. In the second step, one looks at the linearization of the objective function at the point $\overline{\rho}$. The algorithm is presented in Algorithm 4.1 and the main idea is illustrated in Figure 4.1. The motivation to adopt a two-step procedure can be explained as follows. The problem in Eq. (4.25) is a minimization problem. Due to finite precision in computers, an algorithm that solves this minimization problem can stop at a near-optimal point $\overline{\rho}$ and thus $f(\overline{\rho})$ is only an upper bound of the true key rate. We may overestimate the key rate if we simply use $f(\overline{\rho})$. This issue might be more severe given that our objective function is nonlinear and standard solvers might not be optimized to solve such problems. On the other hand, we are only interested in reliable lower bounds since they give us the desired physical meaning: number of secret key bits (per transmitted signal). To find a lower bound, in the second step of the procedure, we solve a linearized problem at the point $\overline{\rho}$ returned from the first step. Since our objective function is continuous and convex, such a linearization always gives us a lower bound on the key rate. The linearized problem at a point ρ is

$$\begin{aligned} \text{minimize:} \quad & \langle \nabla f(\rho), \sigma \rangle \\ \text{subject to:} \quad & \operatorname{Tr}(\sigma \Gamma_i) = \gamma_i \ \ \forall i \in \Sigma \\ & \sigma \geq 0. \end{aligned} \tag{4.26}$$

Its dual problem is

$$\begin{aligned} \text{maximize:} \quad & \boldsymbol{\gamma} \cdot \boldsymbol{y} \\ \text{subject to:} \quad & \sum_{i \in \Sigma} \boldsymbol{y}(i) \Gamma_i \leq \nabla f(\rho) \\ & \boldsymbol{y} \in \mathbb{R}^{|\Sigma|}, \end{aligned} \tag{4.27}$$

where $\boldsymbol{\gamma}(i) = \gamma_i$. We define the dual feasible set for this linearization at the point ρ as $\mathbf{S}^*(\rho)$. That is,

$$\mathbf{S}^*(\rho) = \{\boldsymbol{y} \in \mathbb{R}^{|\Sigma|} : \sum_{i \in \Sigma} \boldsymbol{y}(i) \Gamma_i \leq \nabla f(\rho)\}. \tag{4.28}$$

For any optimal or suboptimal attack σ, one can always obtain a reliable lower bound $\beta(\sigma)$ by [144, Theorem 1]

$$\beta(\sigma) := f(\sigma) - \operatorname{Tr}(\sigma \nabla f(\sigma)) + \max_{\boldsymbol{y} \in \mathbf{S}^*(\sigma)} \boldsymbol{\gamma} \cdot \boldsymbol{y}. \tag{4.29}$$

We note that $\beta(\sigma)$ is a lower bound since it solves the dual problem of the SDP from the linearization at a point σ. This equation provides a lower bound as long as the gradient $\nabla f(\sigma)$ is defined. We note that since the gradient may not be well-defined, a small perturbation is introduced as described in Eq. (4.17). Since our objective function is continuous, by a continuity bound, we introduce a small correction term. More explicitly, we use the following lemma which shows the closeness between the state before the perturbation and that after the perturbation.

Lemma 4.1.4 ([144, Lemma 7])**.** Let $\rho \in \mathrm{D}(\mathcal{H})$ and let $\mathcal{G}$ be a CPTNI map. Let $\mathcal{G}_\epsilon(\rho) = \mathcal{D}_\epsilon \circ \mathcal{G}(\rho)$ and let d' be the dimension of $\mathcal{G}_\epsilon(\rho)$. Let $\mathcal{Z}$ be a CPTP map. Then

$$\|\mathcal{G}(\rho) - \mathcal{G}_\epsilon(\rho)\|_1 \leq \epsilon(d' - 1) \ , \tag{4.30}$$

$$\|\mathcal{Z}(\mathcal{G}(\rho)) - \mathcal{Z}(\mathcal{G}_\epsilon(\rho))\|_1 \leq \epsilon(d' - 1) \ . \tag{4.31}$$

For the objective function $\mathcal{W}(\cdot)$ based on the quantum relative entropy, the continuity bound is given in the following lemma from [144, Lemma 8].

Lemma 4.1.5 ([144, Lemma 8])**.** Let $\mathcal{W}$ be defined in Eq. (4.14) and $\mathcal{W}_\epsilon$ be the perturbed version of $\mathcal{W}$ under the perturbation defined in Eq. (4.17). Let $\epsilon \in (0, \frac{1}{e(d'-1)}]$ where d' is the dimension of $\mathcal{G}_\epsilon(\rho)$. Let ρ be any density matrix. Then

$$|\mathcal{W}(\rho) - \mathcal{W}_\epsilon(\rho)| \leq \zeta_\epsilon \tag{4.32}$$

where $\zeta_\epsilon = \epsilon(d'-1)\log_2 \frac{d'}{\epsilon(d'-1)}$.

This lemma basically follows from Lemmas 2.2.32 and 4.1.4 and its proof can be found in [144, Appendix C. 2].

For the objective function $W(\cdot)$ based on the quantum conditional entropy, a similar continuity bound can be found and is given in the following lemma [83].

Lemma 4.1.6. Let W be defined in Eq. (4.18) and W_ϵ be the perturbed version. Let $\mathcal{H}$ be a Hilbert space with $\dim(\mathcal{H}) = d$. Let $\epsilon \in (0, 1/(e(d-1))]$. Then for any $\rho \in \mathrm{D}(\mathcal{H})$,

$$|W(\rho) - W_\epsilon(\rho)| \leq \eta_\epsilon \tag{4.33}$$

where $\eta_\epsilon = (|Z|+1)|P|\epsilon(d-1)\log_2 \frac{d}{\epsilon(d-1)}$, $|Z|$ is the size of the alphabet for the key string and $|P|$ is the size of the alphabet for all public announcements.

Proof. By Lemma 4.1.4 [144, Lemma 7], we know that

$$\|\mathcal{K}_{zp}(\rho) - \mathcal{K}^\epsilon_{zp}(\rho)\|_1 \leq \epsilon(d-1) \quad \text{and} \quad \|\mathcal{K}_p(\rho) - \mathcal{K}^\epsilon_p(\rho)\|_1 \leq \epsilon(d-1), \tag{4.34}$$

where $\mathcal{K}^\epsilon_{zp} = \mathcal{D}_\epsilon \circ \mathcal{K}_{zp}$ and similarly $\mathcal{K}^\epsilon_p = \mathcal{D}_\epsilon \circ \mathcal{K}_p$. By the continuity of von Neumann entropy (Lemma 2.2.32), it holds for any $\epsilon \in (0, 1/(e(d-1))]$ and density operator ρ,

$$\left|H\left(\mathcal{K}_{zp}(\rho)\right) - H\left(\mathcal{K}^\epsilon_{zp}(\rho)\right)\right| \leq \zeta_\epsilon \quad \text{and} \quad \left|H\left(\mathcal{K}_p(\rho)\right) - H\left(\mathcal{K}^\epsilon_p(\rho)\right)\right| \leq \zeta_\epsilon \tag{4.35}$$

with $\zeta_\epsilon = \epsilon(d-1)\log_2 \frac{d}{\epsilon(d-1)}$. Thus, Eq. (4.33) holds with η_ϵ by the triangle inequality following from the definition of W and W_ϵ. □

Numerical issues related to constraint violations and numerical precision are discussed in detail in [144]. We direct readers to [144] for more discussions about numerical precision and perturbation. We also note that by using facial reduction as is done in [87], one can guarantee the strict feasibility and thus removes the need of perturbation. Then one can also use an interior-point method for solving this optimization problem. See Section 4.2.4 for a high-level summary of facial reduction applied to the key rate calculation problem and [87] for technical details. With perturbation ϵ and constraint violation ϵ' due to both floating point representation and solver precision, a reliable lower bound that replaces Eq. (4.29) is then given as [144, Theorem 3]

$$\beta_{\epsilon\epsilon'}(\sigma) := f_\epsilon(\sigma) - \mathrm{Tr}(\sigma \nabla f_\epsilon(\sigma)) + \max_{(\boldsymbol{y},\boldsymbol{z}) \in \tilde{\mathbf{S}}^*_\epsilon(\sigma)} \Big(\boldsymbol{\gamma} \cdot \boldsymbol{y} - \epsilon' \sum_i \boldsymbol{z}(i) \Big) \tag{4.36}$$

where the set $\tilde{\mathbf{S}}^*_\epsilon(\sigma)$ is defined as

$$\tilde{\mathbf{S}}^*_\epsilon(\sigma) = \Big\{ (\boldsymbol{y}, \boldsymbol{z}) \in (\mathbb{R}^{|\Sigma|}, \mathbb{R}^{|\Sigma|}) - \boldsymbol{z} \leq \boldsymbol{y} \leq \boldsymbol{z}, \sum_i \boldsymbol{y}(i)\Gamma_i \leq \nabla f_\epsilon(\sigma) \Big\}. \tag{4.37}$$

Algorithm 4.1 Asymptotic key rate algorithm based on a two-step procedure [144]

```
0. Let $\epsilon_{\mathrm{th}} > 0$, $\rho_0 \in \mathbf{S}$, maxIter $\in \mathbb{N}$, and $i = 0$.
   *First Step*
1. Compute
       $\Delta\rho := \operatorname*{arg\,min}_{\delta\rho + \rho_i \in \mathbf{S}} \mathrm{Tr}[(\delta\rho)\nabla f(\rho_i)].$
2. If $\mathrm{Tr}[(\Delta\rho)\nabla f(\rho_i)] < \epsilon_{\mathrm{th}}$ or $i >$ maxIter, then proceed to *Second Step*.
3. Find $\lambda \in [0, 1]$ that minimizes $f(\rho_i + \lambda\Delta\rho)$.
4. Set $\rho_{i+1} = \rho_i + \lambda\Delta\rho$, $i \to i + 1$.
5. Go back to Item 1.
   *Second Step*
6. Let $\rho$ be the result of *First Step*. Let $\zeta \geq 0$ be the maximum constraint violation of $\rho$ from the original set $\mathbf{S}$ constraints.
7. Calculate $\nabla f(\rho)$ to use for constructing $\mathbf{S}^*$
8. Expand $\mathbf{S}^*$ such that states which violated the original constraints by $\zeta$ are included.
9. Calculate $\beta(\rho)$ using Eq. (4.29) with the result of the SDP in Eq. (4.27).
```

We also note that this algorithm is very flexible to include linear inequality constraints. With inequality constraints of the form $\mathrm{Tr}(\rho\Omega_j) \leq \omega_j$, we solve the following linearized problem in each iteration of the first step of the algorithm. The linearized problem at a point ρ (to be replaced by a different ρ_i in each iteration) is

$$\begin{aligned} \text{minimize:} \quad & \langle \nabla f(\rho), \sigma \rangle \\ \text{subject to:} \quad & \mathrm{Tr}(\sigma\Gamma_i) = \gamma_i \;\; \forall i \in \Sigma \\ & \mathrm{Tr}(\sigma\Omega_j) \leq \omega_j \;\; \forall j \in \Lambda \\ & \sigma \geq 0. \end{aligned} \tag{4.38}$$

The corresponding dual problem to be solved in the second step of the algorithm at a point ρ (to be replaced by a suboptimal solution $\bar{\rho}$ from the first step) is

$$\begin{aligned} \text{maximize:} \quad & \boldsymbol{\gamma} \cdot \boldsymbol{y_1} + \boldsymbol{\omega} \cdot \boldsymbol{y_2} \\ \text{subject to:} \quad & \sum_{i\in\Sigma} \boldsymbol{y_1}(i)\Gamma_i + \sum_{j\in\Lambda} \boldsymbol{y_2}(j)\Omega_j \leq \nabla f(\rho) \\ & \boldsymbol{y_1} \in \mathbb{R}^{|\Sigma|} \\ & \boldsymbol{y_2} \in \mathbb{R}^{|\Lambda|}, \boldsymbol{y_2} \geq 0 \ . \end{aligned} \tag{4.39}$$

4.1.5 On numerical performance and limitations

To solve linear SDP problems in both first and second steps, we employ the CVX package [149, 150] with either SDPT3 [151, 152] or MOSEK [153] solver in Matlab.

In principle, this numerical framework is general for many device-dependent QKD protocols. In practice, it can handle only low-dimensional protocols. The reason is as follows. In our key rate calculation problem, the optimization variable is the joint density matrix ρ and the size of ρ depends on the dimension of Alice and Bob's joint Hilbert space. When the dimension becomes large, the problem becomes more numerically challenging and requires more computational resources. The running time can also depend on the number of constraints although the dimension is the major factor. If too many constraints cause an issue, one may perform coarse-grainings at the potential cost of losing key rates. Typically, for a problem with dimension less than 100 and with fine-grained statistics, one Frank-Wolfe iteration may take several seconds to a minute on a personal computer. The running time largely depends on the number of Frank-Wolfe iterations. For some protocols in less numerically challenging parameter regimes (e.g. low loss scenario), the first step calculation can finish within a few iterations. For numerically challenging problems, it is possible to use several hundred of iterations. We typically set the maximum iteration number to be around 300 for many protocols as it is usually sufficient to produce relatively tight key rates.

However, there exist scenarios where the solver uses the maximal iterations and still produces loose key rates. As the gap between our lower bound and true optimal value depends on the performance of the first step of our algorithm, it is sometimes necessary to improve the result of the first step calculation in order to obtain a better lower bound. Possible ways to improve the first step result include:

(i) replacing the Frank-Wolfe algorithm by other optimization algorithms (e.g. Gauss-Newton interior point method with facial reduction [87]);

(ii) using a different SDP solver;

(iii) choosing a different initial point (ρ) for the first step;

(iv) increasing the number of iterations for the first step of Algorithm 4.1.

The main reason behind these alternatives is that different solvers and different algorithms can have different rates of convergence and thus can potentially give better results within the time limit.

To extend the applicability of this numerical method, we discuss some analytical tools that can help us to reduce the dimension of some protocols in Section 4.2. Developing a customized algorithm for this particular optimization problem can also help reduce the running time and improve the numerical performance as it is done in [87]. More research along this direction is desired.

In addition to the key rate calculation for a fixed protocol setting, one may be also interested in optimization of free parameters in the protocol description such as the probability distribution of signal states or intensities of coherent states (if the protocol uses coherent states). It is interesting to note that machine learning techniques have been adopted to the QKD setting to help reduce the computational resources for the optimization of protocol parameters when the key rate calculation for a single setting is computationally expensive [154, 155].

4.2 Analytical aids to numerical security proof

As mentioned previously, it becomes numerically challenging to deal with protocols with high-dimensional Hilbert spaces. Before we discuss finite-key rate optimization problem, we present some analytical tools developed so far that ease the numerical challenges by finding smaller-dimensional effective Hilbert spaces. These methods can then make the key rate optimization problem more numerically feasible. We note that in principle they may also be applicable for the finite-key rate calculation problems.

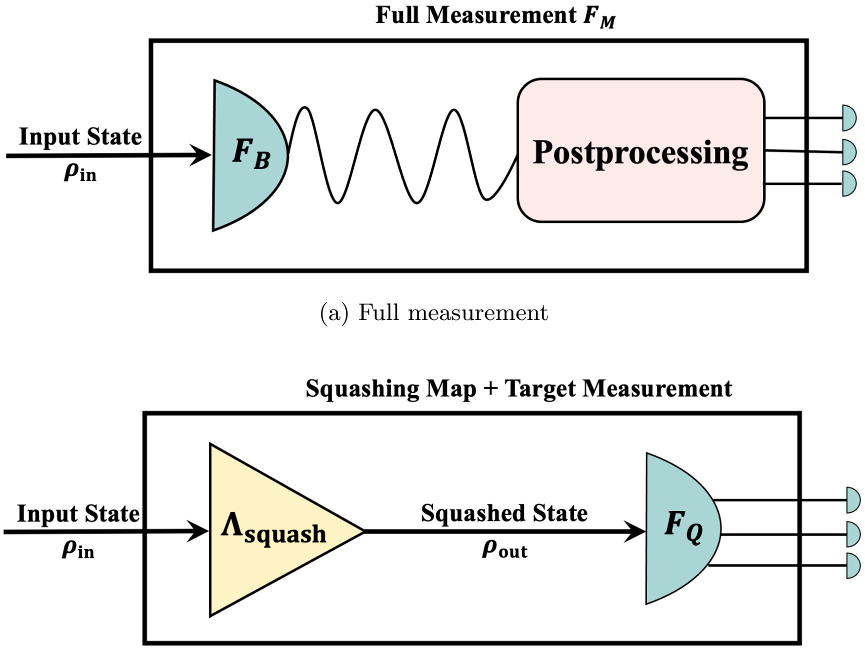
In this section, we start with the idea of squashing model for detectors [86, 90–92, 156] which relates a measurement on optical modes to a target measurement on a low-dimensional space. As will be explained below, these two measurements are equivalent in the sense that they produce the same statistics for any input states. The initial squashing model [90–92, 156] typically maps to a qubit space or a direct sum of a qubit space and the vacuum space (although it is not restricted to be so). However, this method is not universal for any DVQKD since it is proven by Beaudry *et al.* [90] that the six-state protocol does not admit an exact qubit-based squashing model. Another variant of the squashing model to overcome this drawback is the flag-state squasher [86] that is general for any DVQKD. For CV protocols, it is currently an open question whether this idea of squashing model can be extended to CV systems. One method to deal with infinite-dimensional systems is the dimension reduction method developed in [85], which connects the target infinite-dimensional problem with a solvable finite-dimensional problem via some small correction term that is typically obtained by a continuity bound of the relevant entropic quantity. We briefly review some key ideas of these methods and more details can be found in the corresponding references. Finally, for a general semidefinite program (SDP) problem, one can apply some preprocessing procedure such as facial reduction to reduce the problem size. We briefly summarize the idea of facial reduction in the context of the QKD key rate calculation at the end of this section.

4.2.1 Qubit-based squashing model

While early security proofs of QKD protocols like BB84 were based on qubit systems, optical implementations of these protocols typically use weak coherent pulses, states of which have a natural representation in the infinite-dimensional Fock space. The idea of squashing model was motivated to establish a connection between optical implementations and qubit-based security proofs, thereby enjoying the ease of implementation as well as simplicity in the security proof. The idea of squashing model was first mentioned by GLLP [136] to prove the security of BB84 protocol with imperfect devices. This idea was formalized in [90, 91] and subsequently developed in [92, 156].

The main idea of the qubit-based squashing model can be explained with Figure 4.2. In a QKD protocol, the measurement device labeled as B in this figure is described by a POVM F_B that acts on optical modes. This measurement is called the *basic* measurement in previous literature [90, 92]. Since optical modes have a natural infinite-dimensional representation, one can represent elements of F_B by positive operators in the infinite-dimensional Hilbert space. The measurement outcomes from the basic measurement can undergo a classical post-processing procedure. The overall effective measurement process is called the *full* measurement and is associated with the POVM F_M. On the other hand, security proofs based on qubit systems assume that the detectors are described by a POVM $F_Q = \{F_Q^{(i)} : i \in \Sigma\}$ for some alphabet Σ, which we call a *target* measurement for the squashing model. The target measurement is usually on a low-dimensional

Hilbert space such as a qubit space or a three-dimensional space consisting of a qubit and the vacuum in many applications (although not necessarily a qubit). A requirement for the squashing model is that for any input state ρ_{in} to the measurement devices, Figure 4.2a and Figure 4.2b are equivalent in the sense that the output statistics are identical. This requirement also implicitly demands that the full measurement and the target measurement to have the same number of outcomes, which can be guaranteed by choosing a suitable classical postprocessing procedure.



(a) Full measurement

(b) Target measurement

Figure 4.2: Illustration of the qubit-based squashing model [90]. (a) A schematic description of the full measurement with its POVM F_M. The full measurement consists of a POVM F_B that describes the measurement device followed by a postprocessing procedure. (b) A schematic description of the target measurement with its POVM F_Q. A squashing map is first applied to an input state before the target measurement. This squashing map will be given to Eve in the security analysis.

Before one can apply the qubit-based squashing model, one needs to show the existence of a squashing map Λ_{squash}. Formally, a squashing map is a quantum channel such that Λ_{squash} satisfies the following requirement:

$$\operatorname{Tr}\Big(\Lambda_{\text{squash}}(\rho)F_Q^{(i)}\Big) = \operatorname{Tr}\Big(\rho F_M^{(i)}\Big) \quad \forall \rho \in \mathrm{D}(\mathcal{H}),\ \forall i \in \Sigma. \tag{4.40}$$

This requirement can be translated to the requirement on the adjoint map $\Lambda^{\dagger}_{\text{squash}}$, which is a CP unital map such that

$$F_M^{(i)} = \Lambda^{\dagger}_{\text{squash}}(F_Q^{(i)}) \ \ \forall i \in \Sigma \ . \tag{4.41}$$

In order for the squashing model to be useful in the QKD security proofs, it is also natural to demand that the squashing map is not entanglement-breaking. The existence of this type of squashing map also depends on the postprocessing procedure. Once the existence of a squashing map is verified, we can give the squashing map to Eve since it only enhances Eve's power. Then our security analysis can assume that Eve gives the squashed state to Bob. Our analysis proceeds with the target measurement.

The verification of existence of a squashing map can be formulated as a convex optimization problem with linear constraints and semidefinite constraints with the help of Choi representation of the channel. More details can be found in [90, 92].

4.2.2 Flag-state squasher

The flag-state squasher is based on the observation that in many DVQKD protocols, Bob's measurement POVM elements are block-diagonal with respect to the subspaces associated with total photon numbers across all modes. Because of the block-diagonal structures of Bob's POVM, we can assume without loss of generality that Eve performs a quantum non-demolition (QND) measurement of the total photon number after her interaction with the signals and before their arrivals at Bob's side. Thus, the state of ρ_{AB} can be assumed to have the same block-diagonal structure. That is, the state takes the form

$$\rho_{AB} = \bigoplus_{n=0}^{\infty} p_n \rho_{AB}^{(n)}. \tag{4.42}$$

The weight of each subspace carrying a total number of n photons is given by the corresponding probability p_n, and the corresponding normalized conditional state is denoted by $\rho_{AB}^{(n)}$. By choosing a cutoff parameter $k \in \mathbb{N}$ in the security proof, we can write the state and Bob's measurement POVM as

$$\begin{aligned} \rho_{AB} &= p_{n\leq k}\rho_{AB}^{(n\leq k)} \bigoplus (1-p_{n\leq k})\rho_{AB}^{(n>k)}, \\ M_y^B &= M_{y,n\leq k}^B \bigoplus M_{y,n>k}^B, \end{aligned} \tag{4.43}$$

where $p_{n\leq k}$ is the probability that no more than k photons arrive at Bob. In general, one can find a flag-state squasher if the measurement POVM has a finite number of outcomes and all POVM elements are block-diagonal in some basis. The following theorem [86, Theorem 1] finds a target measurement such that a squashing map always exists.

Theorem 4.2.1 (Flag-state squasher, [86, Theorem 1])**.** Suppose that we have a POVM with elements M_y, where $y \in \{1, \dots, J\}$, such that each element can be written in a block-diagonal form $M^B_{y,n\leq k} \bigoplus M^B_{y,n>k}$, with an associated Hilbert space structure given by $\mathcal{H}_{n\leq k} \bigoplus \mathcal{H}_{n>k}$. Then there exists a CPTP map Λ_{squash} from $\mathcal{H}_{n\leq k} \bigoplus \mathcal{H}_{n>k}$ to $\mathcal{H}_{n\leq k} \bigoplus \mathcal{H}_J$ where $\dim(\mathcal{H}_J) = J$, such that $\mathrm{Tr}(\rho M_y) = \mathrm{Tr}\Big(\Lambda_{\text{squash}}(\rho)\tilde{M}_y\Big)$ $\forall \rho \in \mathcal{H}_{n\leq k} \bigoplus \mathcal{H}_{n>k}$ with

$$\tilde{M}_y = M_{y,n\leq k} \oplus |y\rangle\langle y|\,, \tag{4.44}$$

where the states $\{|y\rangle\}$ form an orthonormal basis of $\mathcal{H}_J$.

The idea of the theorem is illustrated in Figure 4.3. To use this theorem in a security proof, one also needs to bound the weight in the subspace $\mathcal{H}_{n>k}$ since the squashing map maps any input state in this subspace to a classical state that gives Eve complete information about the measurement outcomes. This means no keys can be distilled from this subspace. If the state mostly lives in the orthogonal subspace $\mathcal{H}_{n\leq k}$, then it remains possible to generate positive keys. Otherwise, we expect the secure key rate to be zero as the squashing model would be an entanglement-breaking channel in that case.

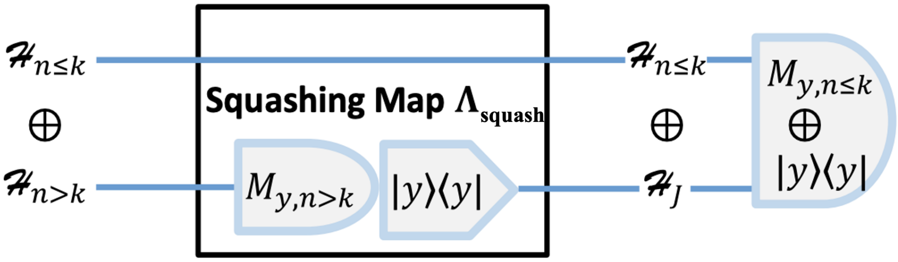
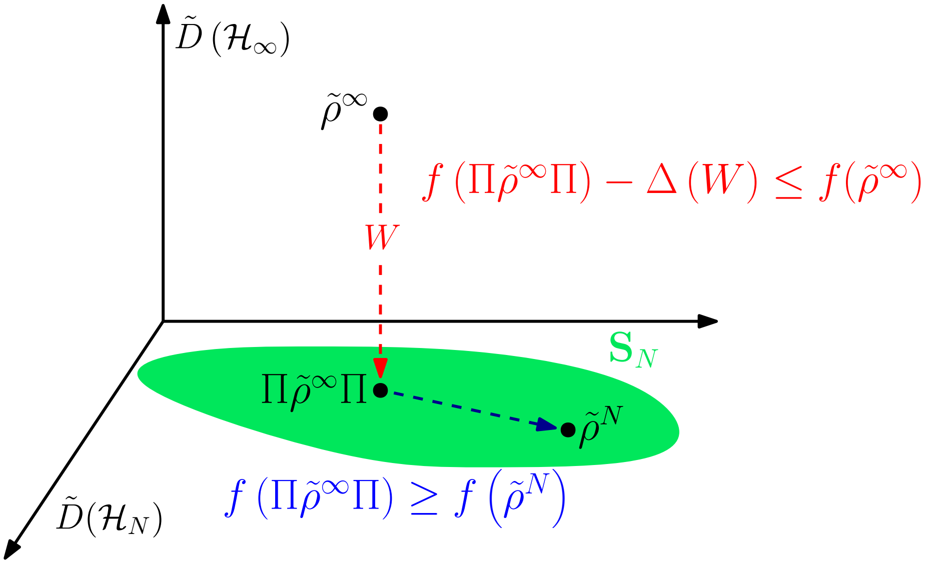


Figure 4.3: Illustration of the flag-state squasher [86]. The squashing map Λ_{squash} leaves the states in the subspace $\mathcal{H}_{n\leq k}$ untouched while measuring states in the subspace $\mathcal{H}_{n>k}$ and outputting classical states that flag the measurement outcomes.

4.2.3 Dimension reduction method

As mentioned previously, squashing models are not known for CVQKD protocols. The dimension reduction method looks at the infinite-dimensional key rate problem at a different perspective and connects it with a finite-dimensional problem. Thus, it can be used for some CVQKD protocols.

While the framework in [85] is more general than it is needed for QKD, for the purpose of QKD security proofs, we restrict our discussion to the application of the framework to QKD. The general idea is illustrated in Figure 4.4.



$\tilde{D}\left(\mathcal{H}_\infty\right)$
$\tilde{\rho}^\infty$
$f\left(\Pi\tilde{\rho}^\infty\Pi\right) - \Delta\left(W\right) \leq f(\tilde{\rho}^\infty)$
W
$\mathbf{S}_N$
$\Pi\tilde{\rho}^\infty\Pi$
$\tilde{\rho}^N$
$\tilde{D}(\mathcal{H}_N)$
$f\left(\Pi\tilde{\rho}^\infty\Pi\right) \geq f\left(\tilde{\rho}^N\right)$

Figure 4.4: Illustration of the dimension reduction method [85]. Pictorial representation of Theorem 4.2.4. $\tilde{\rho}^\infty$ and $\tilde{\rho}^N$ represent the optimal solutions to the infinite- and finite-dimensional optimization problem, respectively. The set $\mathbf{S}_N$ is chosen to contain the projection $\Pi\tilde{\rho}^\infty\Pi$, which is used to relate $f(\tilde{\rho}^\infty)$ and $f(\tilde{\rho}^N)$.

Let $\mathcal{H}_\infty$ be a separable Hilbert space, which may be infinite-dimensional. Let $\mathbf{S}_\infty$ be the feasible set of a QKD key rate problem, which is a convex subset of $\mathrm{D}(\mathcal{H}_\infty)$. Since our problem is linked to a physical scenario and we are typically interested in parameter regimes where the feasible set is nonempty, we assume without loss of generality that the set $\mathbf{S}_\infty$ is nonempty. Otherwise, we simply set the key rate to be zero if the problem is infeasible. The presumably infinite-dimensional optimization problem of interest is

$$\min_{\rho\in\mathbf{S}_\infty} f(\rho), \tag{4.45}$$

where f is the our convex objective function for QKD key rate calculation. Because f is continuous and $\mathbf{S}_\infty$ is typically chosen to be compact, there exists a feasible operator $\tilde{\rho}^\infty$ that achieves the minimum.

Let $\mathcal{H}_N$ be a finite-dimensional subspace of $\mathcal{H}_\infty$ that one chooses to use in a security proof. Let Π be the projector onto this subspace and define $\bar{\Pi} = \mathbb{1} - \Pi$. One then chooses $\mathbf{S}_N$ wisely such that it is a nonempty convex subset of $\mathrm{D}_{\leq}(\mathcal{H}_N)$ with the requirement that

$$\Pi \mathbf{S}_\infty \Pi \subseteq \mathbf{S}_N, \tag{4.46}$$

where $\Pi \mathbf{S}_\infty \Pi := \{\Pi \sigma \Pi : \sigma \in \mathbf{S}_\infty\}$. We note that it is always possible to choose such a set $\mathbf{S}_N$ since $\Pi \mathbf{S}_\infty \Pi \subseteq \mathrm{D}_{\leq}(\mathcal{H}_N)$. We then define the finite-dimensional optimization problem as

$$\min_{\tilde{\rho} \in \mathbf{S}_N} f(\tilde{\rho}). \tag{4.47}$$

Moreover, it is assumed that there exists a feasible operator $\tilde{\rho}^N$ that reaches this optimum. This assumption holds whenever we choose $\mathbf{S}_N$ to be compact.

One key ingredient of this theorem is an entropy continuity bound as stated in the following lemma, which is a generalized version of [157, Lemma 2] to subnormalized states. Its proof follows essentially the proof of [157, Lemma 2] by considering subnormalized states. See [85, Appendix A] for a complete proof.

Lemma 4.2.2 ([85, Lemma 1])**.** Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be two Hilbert spaces, where $\dim(\mathcal{H}_A) = |A| < \infty$ while $\mathcal{H}_B$ can be infinite-dimensional. Let $\tilde{\rho}_{AB}, \tilde{\sigma}_{AB} \in \mathrm{D}_{\leq}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be two subnormalized, classical-quantum states with $\mathrm{Tr}(\tilde{\rho}_{AB}) \geq \mathrm{Tr}(\tilde{\sigma}_{AB})$. If $\frac{1}{2}\|\tilde{\rho}_{AB} - \tilde{\sigma}_{AB}\|_1 \leq \epsilon$, then

$$H(A|B)_{\tilde{\sigma}_{AB}} - H(A|B)_{\tilde{\rho}_{AB}} \leq \epsilon \log_2 |A| + (1+\epsilon) h(\frac{\epsilon}{1+\epsilon}), \tag{4.48}$$

where $h(x)$ is the binary entropy function.

Another useful result is to bound the trace distance between a state ρ and its projected state under Π. This is given in [89, Lemma A.2.8]. This lemma can be easily generalized to infinite-dimensional Hilbert spaces by the same arguments.

Lemma 4.2.3 ([89, Lemma A.2.8])**.** Let $\rho, \tilde{\rho} \in \mathrm{D}_{\leq}(\mathcal{H})$ such that $\tilde{\rho} = \Pi \rho \Pi$ for some projector Π on $\mathcal{H}$. Then,

$$\frac{1}{2}\|\rho - \tilde{\rho}\|_1 \leq \sqrt{\mathrm{Tr}(\rho)(\mathrm{Tr}(\rho) - \mathrm{Tr}(\tilde{\rho}))}. \tag{4.49}$$

Combining these two results, we give a slightly improved version of [85, Theorem 2] combined with [85, Theorem 1].

Theorem 4.2.4. Let $\tilde{\rho}^N$ be an optimal solution to the finite-dimensional problem in Eq. (4.47) and $\tilde{\rho}^\infty$ be an optimal solution to the possibly infinite-dimensional problem in Eq. (4.45). Let W be chosen such that

$$W \geq \max_{\rho \in \mathbf{S}_\infty} \mathrm{Tr}\big(\rho \bar{\Pi}\big), \tag{4.50}$$

then for the QKD problem with its objective function f, it is the case that

$$f(\tilde{\rho}^N) - \Delta(W) \leq f(\tilde{\rho}^\infty), \tag{4.51}$$

where

$$\Delta(W) = \sqrt{W}\log_2|Z| + (1+\sqrt{W})h(\frac{\sqrt{W}}{1+\sqrt{W}}), \tag{4.52}$$

and $|Z|$ is the size of alphabet for the key.

Proof. Since $\Pi\mathbf{S}_\infty\Pi \subseteq \mathbf{S}_N$, it follows that $\Pi\tilde{\rho}^\infty\Pi \in \mathbf{S}_N$. Thus,

$$f(\tilde{\rho}^N) \leq f(\Pi\tilde{\rho}^\infty\Pi) \tag{4.53}$$

as $\tilde{\rho}^N$ achieves the minimum of the finite-dimensional optimization. Because $\mathrm{Tr}(\tilde{\rho}^\infty) = 1$ and $W \geq \max_{\rho\in\mathbf{S}_\infty}\mathrm{Tr}\big(\rho\bar{\Pi}\big)$, by Lemma 4.2.3, it is the case that

$$\begin{aligned}\frac{1}{2}\|\tilde{\rho}^\infty - \Pi\tilde{\rho}^\infty\Pi\|_1 &\leq \sqrt{\mathrm{Tr}(\tilde{\rho}^\infty)(\mathrm{Tr}(\tilde{\rho}^\infty) - \mathrm{Tr}(\Pi\tilde{\rho}^\infty\Pi))} \\ &\leq \sqrt{\mathrm{Tr}(\tilde{\rho}^\infty(\mathbb{1} - \Pi))} \leq \sqrt{W}.\end{aligned} \tag{4.54}$$

Because the trace distance is non-increasing under a CPTP map, after applying the postprocessing map Φ_{PP} which is a CPTP map, it is the case that

$$\frac{1}{2}\|\Phi_{\mathrm{PP}}(\tilde{\rho}^\infty) - \Phi_{\mathrm{PP}}(\Pi\tilde{\rho}^\infty\Pi)\|_1 \leq \frac{1}{2}\|\tilde{\rho}^\infty - \Pi\tilde{\rho}^\infty\Pi\|_1 \leq \sqrt{W}. \tag{4.55}$$

Let $\tau_{Z[E]} := \Phi_{\mathrm{PP}}(\tilde{\rho}^\infty)$ and $\tau^{\Pi}_{Z[E]} := \Phi_{\mathrm{PP}}(\Pi\tilde{\rho}^\infty\Pi)$. Note that $\mathrm{Tr}\big(\tau_{Z[E]}\big) \geq \mathrm{Tr}\Big(\tau^{\Pi}_{Z[E]}\Big)$. Then by Lemma 4.2.2 with $\epsilon = \sqrt{W}$, it is the case that

$$f(\Pi\tilde{\rho}^\infty\Pi) - f(\tilde{\rho}^\infty) = H(Z|[E])_{\tau_{Z[E]}} - H(Z|[E])_{\tau^{\Pi}_{Z[E]}} \leq \Delta(W) \tag{4.56}$$

with $\Delta(W)$ given in Eq. (4.52). Combining Eq. (4.55) with Eq. (4.53), we have Eq. (4.51). □

The remaining task is to choose an appropriate set $\mathbf{S}_N$ before we solve Eq. (4.47). Since the requirement is that for any $\sigma \in \mathbf{S}_\infty$, the state $\tilde{\sigma}^\Pi := \Pi\sigma\Pi$ must be in the set $\mathbf{S}_N$, we loosen each constraint in $\mathbf{S}_\infty$ under the projection of Π to guarantee that $\Pi\mathbf{S}_\infty\Pi \subseteq \mathbf{S}_N$. We note that $\mathbf{S}_\infty$ for QKD problems can be written in terms of linear equality and inequality constraints in addition to semidefinite ordering constraints as explained in Section 4.1.1.

For a linear equality constraint of the form $\mathrm{Tr}(\sigma\Gamma_i)=\gamma_i$, the loosened constraints are $\gamma_i^{\min} \leq \mathrm{Tr}\left(\tilde{\sigma}^{\Pi}\Gamma_i\right) \leq \gamma_i^{\max}$, where

$$\begin{aligned}\gamma_i^{\max} &= \max_{\rho\in\mathbf{S}_\infty} \mathrm{Tr}(\Gamma_i\Pi\rho\Pi),\\ \gamma_i^{\min} &= \min_{\rho\in\mathbf{S}_\infty} \mathrm{Tr}(\Gamma_i\Pi\rho\Pi).\end{aligned} \tag{4.57}$$

Determining of those bounds is tractable in practical scenarios. For example, Ref. [158] derives tight bounds for noncommuting, nonpositive and unbounded observables Γ_i's. In the following discussion, we restrict to an interesting special case where $\Gamma_i \geq 0$ with $[\Pi,\Gamma_i]=0$ for every constraint. This restriction is not too limiting since by a clever choice of the projection Π, this condition holds for many protocols where the linear constraints are from POVM. For example, one may satisfy this requirement for DVQKD protocols by choosing Π to be the projection onto the subspace where the total number of photons is less than a cutoff value. Such a projection commutes with all measurements from threshold detectors. An example of a CVQKD protocol can be found in Section 6.10, which is one of the examples in [85].

Theorem 4.2.5 ([85, Theorem 4])**.** Let $\Gamma_i \geq 0$ such that $[\Pi,\Gamma_i]=0$. If $\mathrm{Tr}\left(\rho\bar{\Pi}\right) \leq W$ and $\mathrm{Tr}(\rho\Gamma_i)=\gamma_i$, then

$$\gamma_i - W\|\Gamma_i\|_\infty \leq \mathrm{Tr}(\Pi\rho\Pi\Gamma_i) \leq \gamma_i. \tag{4.58}$$

Proof. From the commutation relation, it follows that

$$\mathrm{Tr}(\Pi\rho\Pi\Gamma_i) = \mathrm{Tr}\left(\sqrt{\Gamma_i}\Pi\rho\Pi\sqrt{\Gamma_i}\right) = \mathrm{Tr}\left(\Pi\sqrt{\Gamma_i}\rho\sqrt{\Gamma_i}\Pi\right). \tag{4.59}$$

Because the trace of a positive operator can only decrease under projection, the upper bound is

$$\mathrm{Tr}\left(\Pi\sqrt{\Gamma_i}\rho\sqrt{\Gamma_i}\Pi\right) \leq \mathrm{Tr}\left(\sqrt{\Gamma_i}\rho\sqrt{\Gamma_i}\right) = \gamma_i. \tag{4.60}$$

By using Hölder's inequality (Lemma 2.1.11), the lower bound can be derived:

$$\begin{aligned}\mathrm{Tr}\left(\Pi\sqrt{\Gamma_i}\rho\Sigma\right) &= \mathrm{Tr}\left(\sqrt{\Gamma_i}\rho\sqrt{\Gamma_i}\right) - \mathrm{Tr}\left(\bar{\Pi}\sqrt{\Gamma_i}\rho\sqrt{\Gamma_i}\bar{\Pi}\right)\\ &= \gamma_i - \mathrm{Tr}\left(\bar{\Pi}\rho\bar{\Pi}\Gamma_i\right)\\ &\geq \gamma_i - \left\|\bar{\Pi}\rho\bar{\Pi}\right\|_1 \|\Gamma_i\|_\infty\\ &\geq \gamma_i - W\|\Gamma_i\|_\infty\ .\end{aligned} \tag{4.61}$$

□

We note that this lower bound is trivial for unbounded observables. By applying this theorem or directly by the definition of W in Eq. (4.50), it is easy to see the trace constraint $\mathrm{Tr}(\sigma)=1$ in $\mathbf{S}_\infty$ becomes the constraint $1-W \leq \mathrm{Tr}\left(\tilde{\sigma}^{\Pi}\right) \leq 1$. From Lemma 4.2.3, since $\frac{1}{2}\left\|\sigma-\tilde{\sigma}^{\Pi}\right\|_1 \leq \sqrt{W}$

and the trace distance is monotonically nonincreasing under the partial trace, it is the case that $\left\|\rho_A - \mathrm{Tr}_B(\tilde{\sigma}^{\Pi})\right\|_1 \leq 2\sqrt{W}$ for the reduced density operator constraint in Eq. (4.4). In summary, the finite-dimensional optimization problem in Eq. (4.47) can be written as

$$\begin{aligned}
\underset{\tilde{\rho}}{\text{minimize}} \quad & f(\tilde{\rho}) \\
\text{subject to} \quad & \gamma_i - W\|\Gamma_i\|_\infty \leq \mathrm{Tr}(\tilde{\rho}\Pi\Gamma_i\Pi) \leq \gamma_i \\
& \|\mathrm{Tr}_B(\tilde{\rho}) - \rho_A\|_1 \leq 2\sqrt{W} \\
& 1 - W \leq \mathrm{Tr}(\tilde{\rho}) \leq 1 \\
& \tilde{\rho} \geq 0 \ .
\end{aligned} \tag{4.62}$$

By the construction, we satisfy the requirement in Eq. (4.46). Thus, we can solve this optimization and apply Theorem 4.2.4 to obtain key rates of an infinite-dimensional protocol. We remark that the determination of W in Eq. (4.50) may require some additional work. Reference [85] provides two concrete examples where W can be determined analytically. We discuss one of the examples in Section 6.10. We also note that when the POVM elements for the key map have some additional structure like block-diagonal structures, one can tighten up the correction terms and even make the correction term $\Delta(W)$ to be zero. We direct readers to [85] for more discussions.

4.2.4 Facial reduction

We provide only a brief high-level description about facial reduction, which is a general technique to preprocess an optimization problem. In our context, facial reduction is a process of identifying the minimal face (see Definition 2.4.6) of the positive semidefinite cone $\mathrm{Pos}(\mathcal{H})$ containing the affine subspace (defined by linear constraints) $\{\rho : \mathrm{Tr}(\Gamma_i\rho) = \gamma_i \ \forall i\}$. The general idea is to regularize the problem in Eq. (4.25) by removing redundant constraints and redundant unknown variables that always take a fixed value (in particular zero) in the entire feasible set. For example, if the reduced density operator ρ_A is singular (i.e., it has the zero eigenvalue), then we know that any feasible ρ must have the zero eigenvalue. Effectively, we can solve a smaller problem by restricting to the eigenspaces with nonzero eigenvalues. Similarly, while our construction of the postprocessing map $\mathcal{G}$ is physically motivated, this canonical representation typically unavoidably introduces zero eigenvalue to the resulting state $\mathcal{G}(\rho)$ even if we start with a positive definite ρ. It then creates a numerical challenge for defining the gradient of the objective function since the gradient is ill-defined for singular matrices. While the two-step procedure based on the Frank-Wolfe method chooses to introduce perturbation to avoid this issue, it may lead to numerical instability. By finding a mathematically equivalent way of defining $\mathcal{G}$, we can avoid introducing zeros. Facial reduction provides us such a mathematical tool to transform a physically motivated construction of the postprocessing map $\mathcal{G}$ to an equivalent optimization-friendly representation of the map. Through a sequence of facial reduction steps, the regularized problem for QKD key rate problem turns out to also be in the form of Eq. (4.25) but with fewer constraints and a smaller space where ρ lives. Since the running time of an optimization algorithm typically scales with the size of problem, it is always beneficial to apply such a preprocessing procedure to reduce the problem size. We direct readers to [87] for technical details about the application of facial reduction to our QKD key rate calculation problem.

4.3 Finite key rate optimization via a canonical approach

In this section, we discuss a finite key rate calculation method based on Ref. [84]. This method uses the framework developed in [89, 117, 126] and extends the asymptotic numerical method [144] to the finite-key scenario. In particular, our discussion focuses on the collective attacks since main ideas of the method remain the same when it is applied to general attacks via either the quantum de Finetti theorem (Section 3.5) or the postselection technique (Section 3.6).

4.3.1 Collective attacks

In the finite-key rate calculation problem, there are two types of constraints: constraints that are not subject to statistical fluctuations and ones that are. We refer to the set of constraints that are not subject to statistical fluctuations as certainty constraints and denoted by $\{\Gamma_i : i \in \Lambda\}$.We refer to the set of constraints that are subject to statistical fluctuations as uncertainty constraints and denoted by $\{\widetilde{\Gamma}_j : j \in \Sigma\}$. Certainty constraints are usually from the source-replacement scheme (see Section 2.5.2) since there is an additional promise that the reduced density operator is unchanged. Under the assumption of collective attacks, one can directly require all states in the feasible set to satisfy the reduced density operator constraint [Eq. (4.4)] exactly. We define a CPTP map for the certainty constraints in a similar way as in Eq. (3.16):

$$\Phi_0(\rho) = \sum_{i\in\Lambda} \mathrm{Tr}(\rho\Gamma_i) |i\rangle\langle i| . \tag{4.63}$$

We use Φ_M to refer to a similar CPTP map when the uncertainty constraints are used. The feasible set in the case of collective attacks is given in Eq. (3.25) which is defined in terms of the set in Eq. (3.23). Since there are constraints in the form of trace norm, it is useful to note that the trace norm of a Hermitian matrix A admits the following SDP:

$$\begin{aligned} \text{minimize} \quad & \mathrm{Tr}(Q) + \mathrm{Tr}(R) \\ \text{subject to} \quad & Q \geq A, \\ & R \geq -A, \\ & Q, R \geq 0. \end{aligned} \tag{4.64}$$

This SDP satisfies the strong duality [94]. This property is useful to show the tightness of the method in the sense that if an algorithm can solve the finite-key rate convex optimization problem precisely, the lower bound from the algorithm is precisely the primal optimal value. Our algorithm for the finite-key analysis is similar to the algorithm for the asymptotic key rate in Algorithm 4.1 with some modifications.

Unique acceptance

To illustrate main ideas, we start with a simple scenario where the set of acceptable frequency distribution contains only a single point, that is, $\mathcal{Q} = \{\boldsymbol{f}\}$ for some $\boldsymbol{f} \in \mathbb{P}(\mathcal{X})$. We use F to denote the diagonal matrix version of $\boldsymbol{f}$. In this case, the feasible set is

$$\mathbf{S}_{\mathrm{PE}}^{\mathrm{UA}} = \{\rho \in \mathrm{D}(\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}}) : \|\Phi_{\mathcal{M}}(\rho) - F\|_1 \leq \mu, \Phi_0(\rho) = \boldsymbol{\gamma}\}. \tag{4.65}$$

We use Eq. (4.64) to rewrite the trace norm constraint. At each iteration of the first step, we solve a linear SDP which follows from a linearization of the objective function at a point ρ with the feasible set $\mathbf{S}_{\mathrm{PE}}^{\mathrm{UA}}$:

$$\begin{aligned} \text{minimize} \quad & \langle \nabla f(\rho), \sigma \rangle \\ \text{subject to:} \quad & \Phi_0(\sigma) = \boldsymbol{\gamma}, \\ & \mathrm{Tr}(G) + \mathrm{Tr}(H) \leq \mu, \\ & G \geq \Phi_{\mathcal{M}}(\sigma) - F, \\ & H \geq -[\Phi_{\mathcal{M}}(\sigma) - F], \\ & \sigma,\ G,\ H \geq 0. \end{aligned} \tag{4.66}$$

The dual of this problem (see Definition 2.4.13) after some straightforward simplification is

$$\begin{aligned} \text{maximize} \quad & \boldsymbol{\gamma} \cdot \boldsymbol{y} + \boldsymbol{f} \cdot \boldsymbol{z} - a\mu \\ \text{subject to:} \quad & \sum_i \boldsymbol{y}(i)\Gamma_i + \sum_j \boldsymbol{z}(j)\widetilde{\Gamma}_j \leq \nabla f(\rho), \\ & -a\mathbf{1} \leq \boldsymbol{z} \leq a\mathbf{1} \\ & a \geq 0,\ \boldsymbol{y} \in \mathbb{R}^{|\Lambda|},\ \boldsymbol{z} \in \mathbb{R}^{|\Sigma|}. \end{aligned} \tag{4.67}$$

This dual problem which replaces Eq. (4.27) is solved in the second step of the algorithm for finite-key analysis.

Acceptance set $\mathcal{Q}$ with a single coarse-graining

Next, we consider the case with a more general acceptance set $\mathcal{Q}$, that is, $\mathcal{Q} = \{F \in \mathbb{P}(\Sigma) : \left\|\overline{\mathcal{N}}(F) - \overline{\mathcal{N}}(\overline{F})\right\|_1 \leq t\}$ according to Eq. (3.21), where $\overline{F} \in \mathbb{P}(\Sigma)$ is a preferred frequency distribution and $\overline{\mathcal{N}}$ is a classical-to-classical channel for coarse-graining. The measurement channel

$\Phi_{\mathcal{M}}$ may use a coarse-grained POVM with a smaller alphabet Σ' instead of the fine-grained alphabet Σ [see Eq. (3.20)]. It means that there is an additional classical-to-classical channel $\mathcal{N}$ that models the coarse-graining from the alphabet Σ to the new alphabet Σ'. The feasible set $\mathbf{S}_{\mathrm{PE}}$ in this scenario is

$$\mathbf{S}_{\mathrm{PE}} = \{\rho \in \mathrm{D}(\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}}) : \|\Phi_{\mathcal{M}}(\rho) - \mathcal{N}(F)\|_1 \leq \mu, \Phi_0(\rho) = \boldsymbol{\gamma}, \left\|\overline{\mathcal{N}}(F) - \overline{\mathcal{N}}(\overline{F})\right\|_1 \leq t, F \in \mathbb{P}(\Sigma)\}. \tag{4.68}$$

We note that F is a free variable subject to some constraints. Since there are two trace norm constraints, we apply Eq. (4.64) to each of the constraints. Then at each iteration of the first step, we solve the following linear SDP problem:

$$\begin{aligned}
\text{minimize} \quad & \langle \nabla f(\rho), \sigma \rangle \\
\text{subject to:} \quad & \Phi_0(\sigma) = \boldsymbol{\gamma}, \\
& \mathrm{Tr}(G) + \mathrm{Tr}(H) \leq \mu, \\
& G \geq \Phi_{\mathcal{M}}(\sigma) - F, \\
& H \geq -[\Phi_{\mathcal{M}}(\sigma) - F], \\
& \mathrm{Tr}\big(\overline{G}\big) + \mathrm{Tr}\big(\overline{H}\big) \leq t, \\
& \overline{G} \geq \overline{F} - F, \\
& \overline{H} \geq -(\overline{F} - F), \\
& \mathrm{Tr}(F) = 1, \\
& G, H, \overline{G}, \overline{H}, F \geq 0 \\
& \sigma \geq 0.
\end{aligned} \tag{4.69}$$

The dual of this problem is

$$\begin{aligned}
\underset{\boldsymbol{y}, a, \bar{a}, b, \boldsymbol{z}, \bar{\boldsymbol{z}}}{\text{maximize}} \quad & \boldsymbol{\gamma} \cdot \boldsymbol{y} + \bar{\boldsymbol{f}} \cdot \bar{\boldsymbol{z}} - a\mu - \bar{a}t - b \\
\text{subject to:} \quad & \sum_i \boldsymbol{y}(i)\Gamma_i + \sum_j \boldsymbol{z}(j)\widetilde{\Gamma}_j \leq \nabla f(\rho), \\
& \overline{N}^{\dagger}(\bar{\boldsymbol{z}}) - \overline{N}^{\dagger}(\boldsymbol{z}) \leq b\mathbf{1} \\
& -a\mathbf{1} \leq \boldsymbol{z} \leq a\mathbf{1} \\
& -a\mathbf{1} \leq \bar{\boldsymbol{z}} \leq a\mathbf{1} \\
& a, \bar{a} \geq 0 \\
& \boldsymbol{y} \in \mathbb{R}^{|\Lambda|}, \boldsymbol{z}, \bar{\boldsymbol{z}} \in \mathbb{R}^{|\Sigma'|},
\end{aligned} \tag{4.70}$$

where we abuse the notation $\overline{N}$ to denote the version of the map that outputs vectors instead of diagonal matrices, and $\bar{\boldsymbol{f}}$ is the vector version of $\overline{F}$.

Acceptance set $\mathcal{Q}$ with multiple coarse-grainings

We now consider the general case where we allow multiple coarse-grainings. The feasible set is defined in Eq. (3.25):

$$\begin{aligned}\mathbf{S}_{\mathrm{PE}} = \{\rho \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B) :& \left\|\Phi_{\mathcal{M}_k}(\rho) - \mathcal{N}_k(F_k)\right\|_1 \leq \mu_k, \Phi_0(\rho) = \boldsymbol{\gamma}, \\ & \left\|\overline{N}_k(F_k) - \overline{N}_k(\overline{F})\right\|_1 \leq t_k, F_k \in \mathbb{P}(\Sigma)\ \forall k\}.\end{aligned}$$

For each trace norm constraint labeled by k, we introduce two slack variables G_k and H_k (or $\overline{G}_k$ and $\overline{H}_k$). In each iteration of the first step of the algorithm, we solve

$$\begin{aligned}\text{minimize} \quad & \langle \nabla f(\rho), \sigma \rangle \\ \text{subject to:} \quad & \Phi_0(\sigma) = \boldsymbol{\gamma}, \\ & \mathrm{Tr}(G_k) + \mathrm{Tr}(H_k) \leq \mu_k, \\ & G_k \geq \Phi_M(\sigma) - F_k, \\ & H_k \geq -[\Phi_M(\sigma) - F_k], \\ & \mathrm{Tr}\big(\overline{G}_k\big) + \mathrm{Tr}\big(\overline{H}_k\big) \leq t_k, \\ & \overline{G}_k \geq \overline{F}_k - F_k, \\ & \overline{H}_k \geq -[\overline{F}_k - F_k], \\ & \mathrm{Tr}(F_k) = 1, \\ & G_k, H_k, \overline{G}_k, \overline{H}_k, F_k \geq 0 \ \ \forall k, \\ & \sigma \geq 0.\end{aligned} \tag{4.71}$$

The dual of this problem is

$$\begin{aligned}\underset{\boldsymbol{y},\boldsymbol{a},\bar{\boldsymbol{a}},\boldsymbol{b},\boldsymbol{z_k},\bar{\boldsymbol{z}}_{\boldsymbol{k}}}{\text{maximize}} \quad & \boldsymbol{\gamma} \cdot \boldsymbol{y} + \sum_k \bar{\boldsymbol{f}}_{\boldsymbol{k}} \cdot \bar{\boldsymbol{z}}_{\boldsymbol{k}} - \boldsymbol{a} \cdot \boldsymbol{\mu} - \bar{\boldsymbol{a}} \cdot \boldsymbol{t} - \|\boldsymbol{b}\|_1 \\ \text{subject to:} \quad & \sum_i \boldsymbol{y}(i)\Gamma_i + \sum_k \sum_j \boldsymbol{z_k}(j)\widetilde{\Gamma}_j^k \leq \nabla f(\rho), \\ & \overline{N}_k^\dagger(\bar{\boldsymbol{z}}_{\boldsymbol{k}}) - \overline{N}_k^\dagger(\boldsymbol{z_k}) \leq b_k \mathbf{1} \ \ \forall k \\ & -a_k \mathbf{1} \leq \boldsymbol{z_k} \leq a_k \mathbf{1} \ \ \forall k \\ & -\bar{a}_k \mathbf{1} \leq \bar{\boldsymbol{z}}_{\boldsymbol{k}} \leq \bar{a}_k \mathbf{1} \ \ \forall k \\ & \boldsymbol{a}, \bar{\boldsymbol{a}} \geq 0 \\ & \boldsymbol{y} \in \mathbb{R}^{|\Lambda|}, \boldsymbol{z_k}, \bar{\boldsymbol{z}}_{\boldsymbol{k}} \in \mathbb{R}^{\left|\Sigma_k'\right|},\end{aligned} \tag{4.72}$$

where each vector $\boldsymbol{x}$ consists of entries x_k for $\boldsymbol{x} = \boldsymbol{a}, \bar{\boldsymbol{a}}, \boldsymbol{\mu}, \boldsymbol{t}, \boldsymbol{b}$ and we abuse the notation $\overline{N}_k$ to denote the version of the map that outputs vectors instead of diagonal matrices.

4.3.2 Coherent attacks

We can easily adapt this numerical method to handle coherent attacks using either the finite quantum de Finetti theorem [89] (see Section 3.5) or the postselection technique [112, 120] (see Section 3.6).

In the case of using the finite quantum de Finetti theorem, for entanglement-based protocols, we only need to calculate each μ_k according to Eq. (3.38) instead of Eq. (3.22). The numerical algorithm for the optimization part of the finite-key rate calculation stays the same. After the algorithm, we then use the key length formula in Eq. (3.37) instead of Eq. (3.30). For prepare-and-measure protocols, we need to remove the certainty constraints about the reduced density matrix ρ_A and treat the reduced density matrix constraint in terms of trace norm constraint as suggested in Eq. (3.43).

If we use the postselection technique for entanglement-based protocols, we first calculate the security parameters ε's for collective attacks from the desired security parameters for coherent attacks. Then, we can directly use our finite-key solver to calculate the key rate. For the prepare-and-measure protocols, see the discussion in Section 3.6 for the complication that is different from entanglement-based protocols.

4.4 Finite key rate optimization via the entropy accumulation theorem

As mentioned in Section 3.7, we can apply the EAT to prove the security against general attacks. To do so, we need to guarantee the Markov chain condition [see Eq. (3.55)] as well as providing a valid min-tradeoff function (see Definition 3.7.2). In this section, we present numerical methods to find min-tradeoff functions.

We require any min-tradeoff function to be an affine function in order to apply EAT. Any affine linear function that takes an input $\boldsymbol{q} \in \mathbb{P}(\mathcal{X})$ is of the form

$$f(\boldsymbol{q}) = \sum_{x\in\mathcal{X}} \boldsymbol{f}(x)\boldsymbol{q}(x) + a \tag{4.73}$$

where $a \in \mathbb{R}$, $\boldsymbol{f} \in \mathbb{R}^{|\mathcal{X}|}$. Moreover, we can without loss of generality consider only linear functions since $\sum_{x\in\mathcal{X}} \boldsymbol{q}(x) = 1$ and one can absorb the constant a into coefficients $\boldsymbol{f}(x)$. Thus, a min-tradeoff function f is fully specified by a vector $\boldsymbol{f}$. It leads to

$$f(\boldsymbol{q}) = \boldsymbol{f} \cdot \boldsymbol{q}. \tag{4.74}$$

Moreover, $\mathsf{Min}(f) = \min_{x\in\mathcal{X}}(\boldsymbol{f}(x)) =: \min(\boldsymbol{f})$ and $\mathsf{Max}(f) = \max_{x\in\mathcal{X}}(\boldsymbol{f}(x)) =: \max(\boldsymbol{f})$.

4.4.1 Algorithm for the min-tradeoff function: first attempt

Our first algorithm is similar to the algorithm used in the asymptotic key rate problem. Recall that the optimization which defines the min-tradeoff function can be equivalently written as

$$g(\boldsymbol{q}) = \min_{\rho\in\mathbf{S}(\boldsymbol{q})} W(\rho), \tag{4.75}$$

with the objective function $W(\rho) = H(Z|PE) = \sum_{z,p} H(K_{zp}\rho K_{zp}^{\dagger}) - \sum_{p} H(K_p\rho K_p^{\dagger})$ and the feasible set $\mathbf{S}(\boldsymbol{q}) := \{\rho \geq 0 : \mathrm{Tr}(\Gamma_i\rho) = \boldsymbol{q}(i) \ \ \forall i\}$. Then any real-valued affine function f such that $f(\boldsymbol{q}) \leq g(\boldsymbol{q}), \forall \boldsymbol{q} \in \mathbb{P}(\mathcal{X})$ is a valid min-tradeoff function. Note that the optimization in Eq. (4.75) admits an exactly same formula for the asymptotic key rate computation whose evaluation can be divided into two steps [144]. First, one finds a near optimal solution for Eq. (4.75) via a certain convex optimization algorithm (e.g. Frank-Wolfe algorithm). Then one builds an SDP by considering the linearization of $W(\cdot)$ at the feasible point we get. Finally the dual SDP will lead to a reliable lower bound for Eq. (4.75). A *key observation* here is that the dual SDP in the second step gives not only the lower bound for the asymptotic key rate but also a valid min-tradeoff function.

As the algorithm involves the gradient ∇W, we need to work with a slightly perturbed version of the objective function to make the gradient always well-defined. For simplicity of writing, we denote the corresponding perturbed maps $\mathcal{K}_{zp}^{\epsilon} := \mathcal{D}_{\epsilon} \circ \mathcal{K}_{zp}$ and $\mathcal{K}_{p}^{\epsilon}(\cdot) := \mathcal{D}_{\epsilon} \circ \mathcal{K}_{p}$ where $\mathcal{D}_{\epsilon}$ is the depolarizing channel (see Definition 2.2.28). Let $\epsilon \in (0, 1)$ and consider the perturbed objective function

$$W_{\epsilon}(\rho) := \sum_{z,p} H\left(\mathcal{K}_{zp}^{\epsilon}(\rho)\right) - \sum_{p} H\left(\mathcal{K}_{p}^{\epsilon}(\rho)\right) \tag{4.76}$$

and the corresponding optimal value $g_{\epsilon}(\boldsymbol{q}) = \min_{\rho\in\mathbf{S}(\boldsymbol{q})} W_{\epsilon}(\rho)$. Then the gradient $\nabla W_{\epsilon}(\rho)$ always exists for all $\rho \geq 0$:

$$\nabla W_{\epsilon}(\rho) = -\sum_{z,p} (\mathcal{K}_{zp}^{\epsilon})^{\dagger}\left(\log \mathcal{K}_{zp}^{\epsilon}(\rho)\right) + \sum_{p} (\mathcal{K}_{p}^{\epsilon})^{\dagger}\left(\log \mathcal{K}_{p}^{\epsilon}(\rho)\right). \tag{4.77}$$

By Lemma 4.1.6, $W(\rho) - \eta_{\epsilon} \leq W_{\epsilon}(\rho) \leq W(\rho) + \eta_{\epsilon}$ with $\eta_{\epsilon} = (|Z| + 1)|P|\epsilon(d-1)\log_2 \frac{d}{\epsilon(d-1)}$. Taking minimization over $\rho \in \mathbf{S}(\boldsymbol{q})$, we have $g(\boldsymbol{q}) - \eta_{\epsilon} \leq g_{\epsilon}(\boldsymbol{q}) \leq g(\boldsymbol{q}) + \eta_{\epsilon}$. Equivalently,

$$|g(\boldsymbol{q}) - g_{\epsilon}(\boldsymbol{q})| \leq \eta_{\epsilon}. \tag{4.78}$$

This indicates that the optimal value of Eq. (4.75) is continuous around $\epsilon = 0$ when we consider the perturbed objective function.

The explicit algorithm for finding a min-tradeoff function is given below in Algorithm 4.2.

Algorithm 4.2 Algorithm for constructing the min-tradeoff function

Inputs:

$\boldsymbol{q_0}$	A given probability distribution in $\mathbb{P}(\mathcal{X})$
W_ϵ	Objective function with a perturbation error $\epsilon \in (0, 1/(e(d-1))]$
Γ_i	Observables

Output:

$\boldsymbol{y}^\star$ A vector in $\mathbb{R}^{|\mathcal{X}|}$ which defines a min-tradeoff function by $f_\epsilon(\boldsymbol{q}) := \langle \boldsymbol{q}, \boldsymbol{y}^\star \rangle - \eta_\epsilon$.

Algorithm:

1. Consider the convex optimization $g_\epsilon(\boldsymbol{q_0}) = \min_{\rho \in \mathbf{S}(\boldsymbol{q_0})} W_\epsilon(\rho)$ with the true optimal solution $\rho^\star_{\epsilon,\mathrm{opt}}$. Solve the optimization (e.g. by Frank-Wolfe algorithm) and obtain a near optimal solution $\rho^\star_\epsilon$.
2. Let $W_{\epsilon,\mathrm{lin}}(\rho) := W_\epsilon(\rho^\star_\epsilon) + \mathrm{Tr}[\nabla W_\epsilon(\rho^\star_\epsilon) \cdot (\rho - \rho^\star_\epsilon)]$ be the linearization of $W_\epsilon(\cdot)$ at the point $\rho^\star_\epsilon$. This can be equivalently written as

$$W_{\epsilon,\mathrm{lin}}(\rho) = \mathrm{Tr}[M_\epsilon \rho] \quad \text{with} \quad M_\epsilon := \left(W_\epsilon(\rho^\star_\epsilon) - \mathrm{Tr}\left[\nabla W_\epsilon(\rho^\star_\epsilon) \cdot \rho^\star_\epsilon\right]\right) \mathbb{1} + \nabla W_\epsilon(\rho^\star_\epsilon). \tag{4.79}$$

   Since $W_\epsilon(\rho)$ is a convex function in ρ, we know that $W_{\epsilon,\mathrm{lin}}(\rho) \leq W_\epsilon(\rho)$, $\forall \rho$, and $W_{\epsilon,\mathrm{lin}}(\rho^\star_\epsilon) = W_\epsilon(\rho^\star_\epsilon)$.
3. Consider the SDP $\min_{\rho \in \mathbf{S}(\boldsymbol{q}^\star)} W_{\epsilon,\mathrm{lin}}(\rho)$ whose dual SDP is given by

$$\max \langle \boldsymbol{q}^\star, \boldsymbol{y} \rangle \quad \text{such that} \quad \sum_{i \in \mathcal{X}} \boldsymbol{y}(i) \Gamma_i \leq M_\epsilon, \ \boldsymbol{y} \in \mathbb{R}^{|\mathcal{X}|}. \tag{4.80}$$

   Solve the dual program and obtain the optimal solution $\boldsymbol{y}^\star$.
4. Construct the min-tradeoff function by $f_\epsilon(\boldsymbol{q}) := \langle \boldsymbol{q}, \boldsymbol{y}^\star \rangle - \eta_\epsilon$.

Remark 4.4.1 (Slater's condition for Eq. (4.80))**.** Since $\mathbf{S}(\boldsymbol{q_0})$ corresponds to a set of valid density matrices, without loss of generality we may take $\Gamma_1 = \mathbb{1}$ and $\gamma_1 = 1$. Let λ_{min} be the smallest eigenvalue of M_ϵ. Then it follows that $(\lambda_{\mathrm{min}} - 1)\Gamma_1 < M_\epsilon$. So $\boldsymbol{y} = (\lambda_{\mathrm{min}} - 1, 0, \cdots, 0)^T$ is a strictly feasible solution for Eq. (4.80). With Slater's condition satisfied, the strong duality holds for Eq. (4.80) as long as $\mathbf{S}(\boldsymbol{q_0})$ is nonempty.

[6] Here we use the fact that ρ is a normalized density matrix. Sifting is handled by assigning a $\perp$ symbol to the discarded rounds.

Remark 4.4.2. Note that our construction of min-tradeoff function is not unique and it depends on the input $\boldsymbol{q_0} \in \mathbb{P}(\mathcal{X})$. Thus, a smart choice of $\boldsymbol{q_0}$ is required to get a relatively tight key rate in the end. For this we can numerically test a bunch of $\boldsymbol{q_0}$ and pick the best one, or we can try to optimize over $\boldsymbol{q_0}$ to get the best key rate possible if the dimension of $\boldsymbol{q_0}$ is not too large (similar to [43, Eq. (32)]).

4.4.2 Algorithm for the min-tradeoff function: second attempt

We notice that the objective function in Algorithm 4.2 is essentially the leading-order term in the key rate. It does not take into account the second-order correction terms in the construction of the min-tradeoff function. We optimize the choice of min-tradeoff functions by heuristically picking different starting points. As such, while it can reproduce asymptotic key rate, it may behave poorly in the small block-size regime. Therefore, it motivates us to include second-order correction terms related to f in the objective function when we search for optimal min-tradeoff functions. In particular, we focus on the version of EAT in Theorem 3.7.4. For a min-tradeoff function f, we collect all terms that depend on f after dividing the number of rounds N from both sides of Eq. (3.63). This leads to a function of min-tradeoff functions:

$$\mathcal{L}(f,\alpha) := \min_{\boldsymbol{q}\in\mathcal{Q}} f(\boldsymbol{q}) - \frac{(\alpha-1)\ln 2}{2}\sqrt{\mathsf{Var}(f)+2} - (\alpha-1)^2 K_\alpha, \tag{4.81}$$

where V is defined in Eq. (3.64) and K_α is defined in Eq. (3.66) with $\alpha \in (1,2)$.

One difficulty in optimizing this function $\mathcal{L}(f,\alpha)$ is that the parameter α needs to be optimized in addition to optimizing the choice of f and all frequency distributions $\boldsymbol{q} \in \mathcal{Q}$. For simplicity, we consider the process to optimize them independently. We first optimize the choice of min-tradeoff function f with the specific choice of α that leads to Eq. (3.67) and a fixed frequency distribution $\boldsymbol{q} \in \mathcal{Q}$ (typically the preferred frequency distribution that defines the set $\mathcal{Q}$). For the min-tradeoff function found in the previous optimization, we then optimize α and optimize $\boldsymbol{q} \in \mathcal{Q}$. We note that this way of optimization is not necessarily optimal and we leave it for future work to optimize all of the parameter α, the frequency distribution $\boldsymbol{q} \in \mathcal{Q}$ and the choice of min-tradeoff function f simultaneously. From Eq. (3.67), after collecting all terms that involve the min-tradeoff function f for a fixed $\boldsymbol{q}_0 \in \mathcal{Q}$, we get

$$\mathcal{L}(f) := f(\boldsymbol{q_0}) - \frac{1}{\sqrt{N}}\sqrt{1-2\log_2(\varepsilon\rho[\Omega])}\sqrt{2+\mathsf{Var}(f)} - \frac{1}{N}c', \tag{4.82}$$

where c' is defined in Eq. (3.69). For a general min-tradeoff function f, $\mathsf{Var}(f)$ can be upper bounded by a function of $\mathsf{Max}(f)$ and $\mathsf{Min}(f)$ as

$$\mathsf{Var}(f) \leq \frac{1}{4}[\mathsf{Max}(f) - \mathsf{Min}(f)]^2. \tag{4.83}$$

We note that the dependence of c' on the min-tradeoff function f is more complicated. However, its contribution to the key rate is much smaller than the first two terms of Eq. (4.82) due to the $1/N$ dependence. Therefore, for simplicity of our method, we focus on an objective function that maximize the first two terms in Eq. (4.82). We verify that in practice, the term c'/N is much smaller than other terms; that is, c' does not have a dependence on N that would be greater than $\sqrt{N}$.

As we see previously in Eq. (4.74), a min-tradeoff function f is fully specified by a vector $\boldsymbol{f}$ as $f(\boldsymbol{q}) = \boldsymbol{f} \cdot \boldsymbol{q}$. Moreover, $\mathsf{Min}(f) = \min(\boldsymbol{f})$ and $\mathsf{Max}(f) = \max(\boldsymbol{f})$. Moreover, in order for f to be a valid min-tradeoff function, it also needs to satisfy the constraint that $\sum_x \boldsymbol{f}(x) \operatorname{Tr}(\rho M_x) \leq W(\rho)\ \forall \rho \in \mathrm{D}(\mathcal{H})$.

In principle, we would like to consider the following optimization:

$$\begin{aligned} \underset{\boldsymbol{f}}{\text{maximize}}\ & \boldsymbol{f} \cdot \boldsymbol{q_0} - c_0\sqrt{1 + c_1^2[\max(\boldsymbol{f}) - \min(\boldsymbol{f})]^2} \\ \text{subject to}\ & \sum_x \boldsymbol{f}(x) \operatorname{Tr}(\rho M_x) \leq W(\rho)\ \forall \rho \in \mathrm{D}(\mathcal{H}), \end{aligned} \tag{4.84}$$

where c_0, c_1 are two constants to be set for generality. In particular, the set of values[7], $c_0 = 2\sqrt{\ln 2}\sqrt{1 - 2\log_2(\varepsilon\rho[\Omega])}/\sqrt{N}, c_1 = \frac{1}{2\sqrt{2}}$, corresponds to the optimization of Eq. (4.82) after dropping the term c'/N.

The reason that we introduce two constants c_0 and c_1 is to make our algorithm general enough to allow the construction of crossover min-tradeoff function as well as the normal min-tradeoff function in the statements of EAT. If our algorithm is used to find a crossover min-tradeoff function h (see Definition 3.7.5), which can be used to reconstruct a min-tradeoff function f by Eqs. (3.73) and (3.74), then $\mathsf{Var}(f)$ is upper bounded by $\frac{1}{\gamma}[\mathsf{Max}(h) - \mathsf{Min}(h)]^2$ according to Eq. (3.78). Moreover, $h(\boldsymbol{q'}) = f(\boldsymbol{q})$ for every $\boldsymbol{q} \in \mathcal{Q}$ where $\boldsymbol{q'}$ is renormalized after removing the position corresponding to the $\perp$ symbol. Thus, it is the case that the problem for finding a crossover min-tradeoff function still has the form of Eq. (4.84). For crossover min-tradeoff functions, these two constants take the following values: $c_0 = 2\sqrt{\ln 2}\sqrt{1 - 2\log_2(\varepsilon\rho[\Omega])}/\sqrt{N}$, $c_1 = 1/\sqrt{2\gamma}$.

The optimization problem in Eq. (4.84) has an infinite number of constraints that we cannot really handle. However, we can use the Fenchel duality (see Section 2.4.2) to show it is the dual problem of some primal problem that we can actually solve. Let $\{M_j\}$ denote the relevant bipartite POVM of a protocol. Appendix D.1 presents a detailed derivation of the primal-dual

[7] When EAT is used for QKD security proofs, the parameter $\rho[\Omega]$ is replaced by a security parameter ε_{EA} for the EAT subprotocol as explained in Theorem 8.2.1 and Remark 8.2.2.

problem relation including strong duality. The primal problem corresponding to Eq. (4.84) is

$$\begin{aligned} \underset{\rho,\xi}{\text{minimize}} \quad & W(\rho) - \sqrt{c_0^2 - \Big(\sum_j \xi_j\Big)^2/(4c_1^2)} \\ \text{subject to} \quad & \mathrm{Tr}(\rho) = 1 \\ & -\xi_j \le \mathrm{Tr}(\rho M_j) - \boldsymbol{q}_0(j) \le \xi_j \\ & \sum_j \xi_j \le 2c_0c_1 \\ & \rho \ge 0\ . \end{aligned} \tag{4.85}$$

The derivative of $\sqrt{c_0^2 - \Big(\sum_j \xi_j\Big)^2/(4c_1^2)}$ with respect to ξ_k is

$$\frac{\partial}{\partial \xi_k}\sqrt{c_0^2 - \Big(\sum_j \xi_j\Big)^2/(4c_1^2)} = -\frac{1}{4c_0^2c_1^2}\frac{1}{\sqrt{1-(\sum_j \xi_j)^2/(4c_0^2c_1^2)}}\sum_j \xi_j\ . \tag{4.86}$$

Algorithm 4.3 Improved algorithm for constructing min-tradeoff function

Inputs:

$\boldsymbol{q}_0$	A given probability distribution in $\mathbb{P}(\mathcal{X})$
c_0, c_1	Two constants related to EAT correction term
W	Quantum conditional entropy function
$\{M_j\}$	Bipartite POVM

Output:

$\boldsymbol{y}^\star$ A vector in $\mathbb{R}^{|\mathcal{X}|}$ which defines a min-tradeoff function by $f(\boldsymbol{q}) := \langle \boldsymbol{q}, \boldsymbol{y}^\star \rangle$.

Algorithm:

1. Use either the Frank-Wolfe method or an interior-point method to solve Eq. (4.85) and obtain a near optimal solution $\rho^\star$.
2. Solve the dual SDP problem of the linearization in Eq. (4.88) with $\rho^\star$ to obtain $\boldsymbol{y}^\star = \boldsymbol{f}$.

The dual problem of the linearization at a point $\rho^\star \in \mathrm{D}(\mathcal{H})$ is

$$\begin{aligned} \underset{\boldsymbol{f}}{\text{maximize}} \quad & \boldsymbol{q_0} \cdot \boldsymbol{f} - c_0 \sqrt{1 + c_1^2 [\max(\boldsymbol{f}) - \min(\boldsymbol{f})]^2} \\ \text{subject to} \quad & \sum_j \boldsymbol{f}(j) M_j \leq \nabla W(\rho^\star)\ . \end{aligned} \tag{4.87}$$

We rewrite this problem as an SDP by introducing slack variables u, v, t:

$$\begin{aligned} \underset{\boldsymbol{f}, u, v, t}{\text{maximize}} \quad & \boldsymbol{q_0} \cdot \boldsymbol{f} - t \\ \text{subject to} \quad & \sum_j \boldsymbol{f}(j) M_j \leq \nabla W(\rho^\star) \\ & v\mathbf{1} \leq \boldsymbol{f} \leq u\mathbf{1} \\ & \begin{pmatrix} t - c_0 & c_0 c_1 (u - v) \\ c_0 c_1 (u - v) & t + c_0 \end{pmatrix} \geq 0\ . \end{aligned} \tag{4.88}$$

Chapter 5

Asymptotic security analysis of PM-MDI QKD

A bottleneck for QKD applications, be it as individual link or as part of a network, is the scaling of the generated secret key rate with the loss in the channel represented by the single-photon transmissivity η. The best known QKD protocols have a scaling of their key rate in the limit of infinite channel uses (infinite key limit) as $R_\infty = O(\eta)$, and by now we have bounds on repeaterless optical channels that show that this is the optimal rate scaling that can be achieved [64, 65]. The tight bound on the performance of QKD in terms of secret key rate per employed optical mode is given by $R_\infty \leq \log_2 \frac{1}{1-\eta}$ [65]. In principle, inserting intermediate stations performing some operations can improve the performance, and quantum repeaters [159] aim at this. The field of quantum repeater research is very active and made conceptual and practical advances over the recent years, but as of today, no quantum repeater has been demonstrated yet that would outperform the direct use of optical channels, and thus breaking the repeaterless bounds.

While proposals have been made for simplest possible devices that allow a demonstration of quantum repeater action by beating repeaterless bounds using a simple single node layout [160], the corresponding quantum advantage has not been experimentally demonstrated yet. In a pleasant surprise to the field, the TF-QKD protocol [62] was recently shown to beat the repeaterless bound when using suitable test states with the phase-encoding MDIQKD protocols [161, 162]. This important observation justifiably creates quite an interest in the community. In the original paper [62] it has been argued that the asymptotic secret key rate scales as $\sqrt{\eta}$, where we keep η as the single-photon transmissivity of the total distance, rather than that of a segment. It is interesting to see that an MDI protocol can achieve that performance without the use of any quantum memory or similar advanced components. Remarkably, the only difference to previous MDI QKD protocols that show an η scaling of R_∞ is the change from single-photon signals (or mixture of photon number states) with two-photon interference events at the beamsplitter, to coherent states as signal states and single-photon interference events at the beamsplitter.

Before our work, the security analyses [63, 69] of the TF-QKD protocols were done in a framework based on the quantum error correction inspired approach by Shor-Preskill [132], which is improved by Koashi [133]. The goal of our work is two-fold: we propose a variant of the TF-QKD protocol that clearly distinguishes between test states, meant to probe potential eavesdropping activities, and signal states, which are meant to establish secret keys. For this modified protocol we then execute a security analysis which is expected to be tight as it uses the framework by Renner [89]. This framework is known to be flexible in terms of error correction and privacy amplification methods, and is general to be adaptable to any generic QKD protocol.

We analyze the security of the protocol in a setting with infinitely many different test states, similar to the initial discussion of decoy states in weak coherent pulse BB84 protocols [66]. In this setting, we can derive an analytical key rate formula for the scenario where Alice and Bob observe correlations coming from a loss-only scenario. We derive the general framework that includes also the noisy case, for which we then resort to numerical evaluations to demonstrate the stability of the proposed protocol. We also briefly review a related follow-up work [93] that proves the security of a protocol in a setting with a few test states.

This chapter is mostly based on the work [71].

5.1 Protocol description

This protocol uses the setup of MDI protocol and thus the protocol description is similar to Protocol 2.3.

Protocol 5.1 PM-MDI QKD

1. *Key-generation vs. test mode selection.*— Alice (Bob) randomly chooses a random bit m_A (m_B) according to *a priori* probability distribution $\{p_A, 1-p_A\}$ ($\{p_B, 1-p_B\}$). If $m_A = 0$, Alice then labels this round as in the key-generation mode. If $m_A = 1$, Alice labels this round as in the test mode. Similarly for Bob.

2. *State preparation.*— If the test mode is chosen, Alice (Bob) then randomly chooses a phase θ_A (θ_B) $\in [0, 2\pi)$ and randomly chooses an intensity μ_A (μ_B). Then she (he) prepares a coherent state $\left|\sqrt{\mu_A}e^{i\theta_A}\right\rangle$ ($\left|\sqrt{\mu_B}e^{i\theta_B}\right\rangle$) and sends it to the untrusted third party Charlie through the quantum channel.

   If the key-generation mode is chosen, Alice (Bob) randomly generates a bit value k_A (k_B) $\in \{0,1\}$ with a uniform probability distribution. Alice (Bob) chooses the pre-agreed intensity μ and sends a coherent state $\left|\sqrt{\mu}e^{i\pi k_A}\right\rangle$ ($\left|\sqrt{\mu}e^{i\pi k_B}\right\rangle$) to Charlie.

3. *Measurements.*— For each round, Charlie performs a joint measurement on the signals received from Alice and Bob, and then makes an announcement about the measurement outcome. If Charlie is honest, he is supposed to perform the measurement as shown in Figure 5.1a and announces one of the following outcomes {"Only detector D_+ clicks", "Only detector D_- clicks", "No detectors click", "Both detectors click"}, which, for the later convenience of notation, we abbreviate as $\{+, -, ?, d\}$, respectively. We denote Charlie's announcement as γ throughout this chapter.

   After steps 1-3 are repeated for many times, and after Charlie has made all the announcements, Alice and Bob then proceed with the following steps.

4. *Sifting.*— Alice and Bob use an authenticated classical channel to communicate and sort all rounds into two disjoint sets, where one set is used for the key generation and the other is for the parameter estimation. To do so, they disclose the choices of m_A and m_B for each round and also use the announcement γ. If $m_A = m_B = 0$, that is, they both selected the key-generation mode for a given round, and Charlie announced $\gamma \in \{+, -\}$, they save their data corresponding to this round for the key generation. All remaining rounds are used for parameter estimation.

5. *Parameter estimation.*— To perform parameter estimation, Alice and Bob disclose the choices of $\mu_A, \mu_B, \theta_A, \theta_B$ (also k_A, k_B if they have chosen one for that round) for the rounds in the set labeled for parameter estimation and also use the announcement result γ for each of these rounds to estimate how Eve has interacted with the signals during their exchange in the protocol. If, from their analysis, they find out that Eve has learned too much about the signals and no secret keys can be generated, then they abort the protocol. Otherwise, they continue.

6. *Key map.*— Alice forms a raw key using her bit value k_A from each of the rounds saved for the key generation. (In principle, Bob does not need to do anything in this step since he can correctly determine Alice's key by the error correction. In practice, depending on the choice of error correction code, it might be convenient for Bob to flip his bit value when the announcement is $\gamma =$ " $-$ ".)

7. *Error correction and privacy amplification.*— Alice and Bob then apply the procedures of error correction and privacy amplification as in a typical QKD protocol to generate a secret key.

We remark that since this protocol uses an MDI setup, it is inherently immune to all side channels in the measurement devices once its security is proven. However, Alice's and Bob's sources have to be trusted and protected. In our security analysis, we assume that Alice's and Bob's devices are fully characterized and Eve has no access. This assumption needs to be justified in the experimental implementations of the protocol. In particular, we want to remark that the choices of m_A and m_B (also k_A and k_B) should not be leaked to Eve by side channels before Charlie's announcement is made. In the implementation of the protocol, Alice and Bob need to make sure that Eve cannot distinguish the key-generation mode from test mode by any classical side information leaked from their devices before Charlie's announcement. Just like other MDI QKD protocols, this protocol can be vulnerable to side-channel attacks on the sources.

We also comment on the the choices of parameters p_A, p_B. While values of p_A, p_B need to be optimized in the finite-key regime, in the infinite key limit, we can choose p_A and p_B arbitrarily close to 1 so that the sifting factor is asymptotically 1, like the efficient BB84 protocol [163].

Finally, we remark on the choices of μ_A and μ_B and their corresponding probability distributions. Since states in the test mode essentially are used to perform a tomography on Eve's attacks on the subspace of signal states in the key-generation mode, for the purpose of this chapter, we initially use coherent states whose complex amplitudes cover the entire complex space. In the infinite key limit, the probability distribution (with no zeros) does not matter. We then remark on how a finite number of choices of test states can approximately accomplish the same task and the choices of μ_A and μ_B is closely related to the value of μ.

5.2 Security proof detail

5.2.1 Applying the source-replacement scheme

In the source-replacement scheme (see Section 2.5.2), Alice's and Bob's sources effectively prepare the following state

$$\begin{aligned} |\Psi\rangle_{ABA'B'} &= (\sum_x \sqrt{p_x}\,|x\rangle_A\,|\varphi_x\rangle_{A'}) \otimes (\sum_y \sqrt{q_y}\,|y\rangle_B\,|\varphi_y\rangle_{B'}) \\ &= \sum_{x,y} \sqrt{p_x q_y}\,|x,y\rangle_{AB}\,|\varphi_x,\varphi_y\rangle_{A'B'}\,, \end{aligned} \tag{5.1}$$

where register A records the choices of states prepared in register A' and similarly register B records the choices of states in register B'. We introduce an orthonormal basis $\{|x\rangle_A\}$ for register system A corresponding to states $\{|\varphi_x\rangle\}$, and an orthonormal basis $\{|y\rangle_B\}$ for register system B corresponding to states $\{|\varphi_y\rangle\}$. It is crucial that Eve has no access to registers A and B. Alice keeps register A and sends system A' to Charlie, and similarly, Bob keeps B and sends B'. To learn their choices of states sent to Charlie for each round, Alice performs a local measurement described by a POVM $M_A = \{|x\rangle\langle x|\}$ on her register A and likewise, Bob applies his POVM $M_B = \{|y\rangle\langle y|\}$ to his register B.

Importantly, we only apply the source-replacement scheme for the signal states in the key-generation mode since the test states in the test mode are only used to put constraints on how Eve acts in the subspace spanned by signal states. We denote the set of states in the key-generation mode as $\mathcal{S}$, that is,

$$\mathcal{S} =\{\left|+\sqrt{\mu}, +\sqrt{\mu}\right\rangle, \left|-\sqrt{\mu}, -\sqrt{\mu}\right\rangle, \left|+\sqrt{\mu}, -\sqrt{\mu}\right\rangle, \left|-\sqrt{\mu}, +\sqrt{\mu}\right\rangle\}, \tag{5.2}$$

where each state is a two-mode coherent state coming from both Alice and Bob, and we dropped the subscript $A'B'$ for the ease of writing. Since finitely many coherent states are linearly independent, we want to point out that $\mathcal{S}$ is indeed a basis of $\text{span}(\mathcal{S})$.

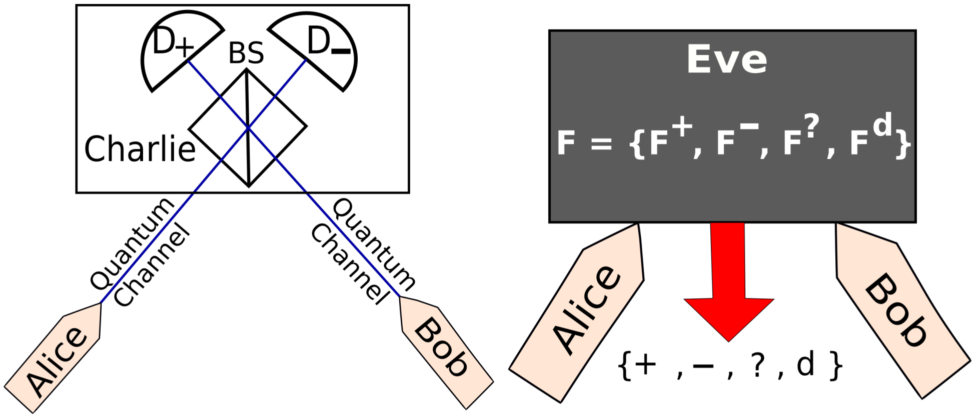
5.2.2 Description of Eve's attack

As an MDI QKD protocol, Eve has a full control of both the quantum channels connecting Alice, Bob and the intermediate node Charlie, and the measurement devices at the intermediate node. Since measurement devices are neither characterized nor trusted, Eve is assumed to play the role of Charlie to perform the measurement. Therefore, in the PM-MDI QKD protocol, we can view the protocol in an alternative and equivalent picture, as shown in Figure 5.1b. In order to make an announcement strategy, Eve performs some measurement, which can be described by a POVM F, directly on the states from Alice and Bob in the registers A' and B'. Moreover, without loss of generality, we can assume that F only has four elements since only $\{+, -, ?, d\}$ outcomes are meaningful for Alice and Bob, and all other outcomes are simply discarded in the protocol. (Even though Alice and Bob may only keep $\{+, -\}$ outcomes to distill keys, we are allowed to include $\{?, d\}$ outcomes for parameter estimation.) We write this POVM F as $F = \{F^+, F^-, F^?, F^d\}$, or abbreviate it as $\{F^\gamma\}$ for $\gamma \in \{+, -, ?, d\}$. The probability of announcing the outcome γ is $\text{Tr}(F^\gamma \sigma_{A'B'})$ for an input state $\sigma_{A'B'}$.

From Alice and Bob's point of view, they can only know the probability of each announcement, not the post-measurement states in Eve's hand. They can infer what POVM F that Eve applied from their observed correlations. However, Eve can perform a nondestructive measurement and keep her post-measurement states for further analysis. That is, Eve applies a CPTP map $\mathcal{E}_{A'B' \to EC}$ on the input quantum states in registers A' and B'. Her announcement about the measurement outcome is stored in the classical register C and she keeps the post-measurement state in register E. Here, we introduce an orthonormal basis $\{|\gamma\rangle\}$ for register C, each of which corresponds to every possible announcement outcome. In general, we can write $\mathcal{E}_{A'B' \to EC}$ as follows:

$$\mathcal{E}_{A'B' \to EC}(X) = \sum_{\gamma} \mathcal{E}_\gamma(X) \otimes |\gamma\rangle\langle\gamma|_C\,, \tag{5.3}$$

where each $\mathcal{E}_\gamma$ is a completely positive trace non-increasing map and X is an arbitrary linear operator on the systems $A'B'$.



(a)

(b)

Figure 5.1: (a). Schematic setup of the PM-MDI QKD protocol. Alice and Bob send coherent states to the untrusted third party Charlie in the middle, who performs measurements and broadcasts outcomes. BS: 50-50 beamsplitter. D_+, D_-: single-photon detectors. (b). Equivalent view of the protocol. Eve is assumed to perform the measurements in the middle. Effectively, Eve performs a 4-element POVM, denoted as $\{F^+, F^-, F^?, F^d\}$, corresponding to four possible announcements $\{D_+$ clicks, D_- clicks, no detectors click, both detectors click$\}$, which are abbreviated as $\{+, -, ?, d\}$.

In the Kraus representation, each $\mathcal{E}_\gamma$ can be written as

$$\mathcal{E}_\gamma(X) = \sum_{j \in \mathcal{I}(\gamma)} K_j^\gamma X (K_j^\gamma)^\dagger, \tag{5.4}$$

with $\sum_j (K_j^\gamma)^\dagger K_j^\gamma = F^\gamma$ and the summation going over some index set $\mathcal{I}(\gamma)$ that depends on γ. Without loss of generality we can use maps $\mathcal{E}_\gamma(X)$ with a single Kraus operator $K^\gamma = \sqrt{F^\gamma}$. The reason for this is that the general case of Eqs. (5.3) and (5.4) can be represented as a concatenation of two maps, the first one using the case of $K^\gamma = \sqrt{F^\gamma}$, followed by a second channel operation that is conditioned on the classical register C and uses Kraus operators $\tilde{K}_j^\gamma = K_j^\gamma (F^\gamma)^{-1/2}$. To see this we need only to verify two things: (a) the concatenation of both operations gives the general form, and (b) the Kraus operators $\tilde{K}_j^\gamma$ for each value of γ define a valid CPTP map. The proof of (a) is trivial, and for (b) we need only to verify that $\sum_{j \in \mathcal{I}_\gamma} \left(\tilde{K}_j^\gamma\right)^\dagger \tilde{K}_j^\gamma = \mathbb{1}_\gamma$, where $\mathbb{1}_\gamma$ is the identity on the support of F^γ and $(F^\gamma)^{-1/2}$ is the pseudoinverse of $\sqrt{F^\gamma}$ with respect to $\mathbb{1}_\gamma$, which defines the support of the signals given the outcome γ. We insert the definition to find

$$\begin{aligned}
&\sum_{j\in\mathcal{I}(\gamma)} \left(\tilde{K}_j^\gamma\right)^\dagger \tilde{K}_j^\gamma \\
&= \sum_{j\in\mathcal{I}(\gamma)} (F^\gamma)^{-1/2} \left(K_j^\gamma\right)^\dagger K_j^\gamma (F^\gamma)^{-1/2} \\
&=(F^\gamma)^{-1/2} \left(\sum_{j\in\mathcal{I}(\gamma)} \left(K_j^\gamma\right)^\dagger K_j^\gamma\right) (F^\gamma)^{-1/2} \\
&=(F^\gamma)^{-1/2} F^\gamma (F^\gamma)^{-1/2} \\
&=\mathbb{1}_\gamma
\end{aligned} \tag{5.5}$$

Clearly, since the general case can thus be considered a two-step procedure, where the first step gives rise to the announcement γ and the second step acts only on Eve's conditional states, it can only strengthen Eve's position by not forcing her to do this second step. Without loss of generality, we can thus assume that Eve's optimal strategy performs only the first step.

Since we assume the sources are protected, Eve cannot have the access to registers A and B and cannot modify the states in those registers. Therefore, when Eve directly acts on the state $|\Psi\rangle_{ABA'B'}$ shown in Eq. (5.1) from the source-replacement scheme, the joint state ρ_{ABEC} shared by Alice, Bob and Eve along with the classical register C for announcements is as follows:

$$\begin{aligned}
\rho_{ABEC} &= (\mathrm{id}_{AB} \otimes \mathcal{E}_{A'B'\to EC})(|\Psi\rangle\langle\Psi|_{ABA'B'}) \\
&= \sum_{x,y,x',y'} \sqrt{p_x p_{x'} q_y q_{y'}} \, |x,y\rangle\langle x',y'|_{AB} \otimes \sum_\gamma (\sqrt{F^\gamma} \, |\varphi_x,\varphi_y\rangle\langle\varphi_{x'},\varphi_{y'}| \, \sqrt{F^\gamma})_E \otimes |\gamma\rangle\langle\gamma|_C
\end{aligned} \tag{5.6}$$

5.2.3 Key rate evaluation with Devetak-Winter formula

To distill keys from ρ_{ABEC}, Alice and Bob perform measurements using POVMs M_A on the register A and M_B on the register B, respectively. Upon measurements, Alice stores her measurement outcomes in a classical register X and Bob stores his in a classical register Y. Alice then applies a key map that maps her measurement result in register X to a raw key bit in register K. We want to point out that the key map step is necessary, but the key map can be trivial, as it is in this PM-MDI QKD protocol. The key map here is an identity map from the register X to the register Z. Let $\mathcal{G}$ denote the effective CPTP map that transforms ρ_{ABEC} to ρ_{ZYEC}. In the end, we generate keys from the state ρ_{ZYEC}, which has the form

$$\rho_{ZYEC} =\mathcal{G}(\rho_{ABEC}) = \sum_{z,y,\gamma} p(\gamma)p(z,y|\gamma) \, |z\rangle\langle z|_Z \otimes |y\rangle\langle y|_Y \otimes \rho_E^{(z,y,\gamma)} \otimes |\gamma\rangle\langle\gamma|_C \, , \tag{5.7}$$

where $\rho_E^{(z,y,\gamma)}$ is Eve's conditional state conditioned on Alice holding z in the register Z, Bob having y in the register Y and the central node announcing γ. Here, $p(\gamma)$ is a marginal probability of the joint probability distribution $p(z,y,\gamma)$ and $p(z,y|\gamma) = \frac{p(z,y,\gamma)}{p(\gamma)}$ is a conditional probability.

Under collective attacks, we can evaluate the secret key generation rate using Devetak-Winter formula [145], which is expressed in terms of a single-copy state ρ_{ZYEC} shared by Alice, Bob and Eve.

As is typical in the MDI protocols, we can choose to generate keys from each announcement outcome γ independently as the announcement is available to all parties. We rewrite ρ_{ABEC} by defining conditional states of Alice, Bob and Eve conditioned on the announcement outcome γ as

$$\rho_{ZYE}^{\gamma} = \sum_{z,y} p(z,y|\gamma) \left|z\middle\rangle\!\middle\langle z\right|_Z \otimes \left|y\middle\rangle\!\middle\langle y\right|_Y \otimes \rho_E^{(z,y,\gamma)}, \tag{5.8}$$

and $\rho_{ZYEC} = \sum_\gamma p(\gamma) \rho_{ZYE}^{\gamma} \otimes \left|\gamma\middle\rangle\!\middle\langle\gamma\right|_C$.

We adapt the Devetak-Winter formula to a general case where the error correction is not necessarily performed at the Shannon limit. In that case the number of secret bits that we can distill from the state ρ_{ZYE}^{γ} is $r(\rho_{ZYE}^{\gamma})$, which is defined as

$$r(\rho_{ZYE}^{\gamma}) = \max(1 - \delta_{\mathrm{EC}}^{\gamma} - \chi(Z\colon E)_{\rho_{ZYE}^{\gamma}}, 0), \tag{5.9}$$

where $\delta_{\mathrm{EC}}^{\gamma}$ is the amount of information leakage per signal during the error correction step for the rounds corresponding to the announcement outcome γ, and

$$\chi(Z\colon E)_{\rho_{ZYE}^{\gamma}} = H(\rho_E^{\gamma}) - \sum_z p(z|\gamma) H(\rho_E^{(z,\gamma)}) \tag{5.10}$$

is the Holevo information, where $H(\rho) = -\operatorname{Tr}(\rho \log_2 \rho)$ is the von Neumann entropy (Definition 2.2.31). The states ρ_E^{γ} and $\rho_E^{(z,\gamma)}$ are defined as:

$$\begin{aligned} \rho_E^{(z,\gamma)} &:= \sum_y \frac{p(z,y|\gamma)}{p(k|\gamma)} \rho_E^{(z,y,\gamma)} = \sum_y p(y|z,\gamma) \rho_E^{(z,y,\gamma)} \\ \rho_E^{\gamma} &:= \sum_z p(z|\gamma) \rho_E^{(z,\gamma)} \end{aligned} \tag{5.11}$$

In the Shannon limit, we have $H(Z) - \delta_{\mathrm{EC}}^{\gamma} = I(Z:Y)_{\rho_{ZYE}^{\gamma}}$ and thus we recover the original Devetak-Winter formula in Eq. (5.9). Another important observation is that $\delta_{\mathrm{EC}}^{\gamma}$ is directly determined from the experimentally observed correlations.

The total number of secret bits that we can distill from the state ρ_{KYEC}, denoted by $\tilde{r}(\rho_{ZYEC})$, is defined as

$$\tilde{r}(\rho_{ZYEC}) = \sum_\gamma p(\gamma) r(\rho_{ZYE}^{\gamma}). \tag{5.12}$$

From Eq. (5.6), we can calculate Eve's conditional states $\rho_E^{(z,y,\gamma)}$ as

$$\rho_E^{(z,y,\gamma)} = \left|\Theta_{z,y}^{\gamma}\middle\rangle\!\middle\langle\Theta_{z,y}^{\gamma}\right|, \tag{5.13}$$

where we define

$$\left|\Theta_{z,y}^{\gamma}\right\rangle = \frac{\sqrt{F^{\gamma}}\,|\varphi_z,\varphi_y\rangle}{\sqrt{\langle\varphi_z,\varphi_y|\,F^{\gamma}\,|\varphi_z,\varphi_y\rangle}}. \tag{5.14}$$

Then, by substituting Eq. (5.13) into Eq. (5.11), we can calculate the conditional states ρ_E^{γ} and $\rho_E^{z,\gamma}$, and evaluate $\chi(Z\!:E)_{\rho_{ZYE}^{\gamma}}$ in Eq. (5.10) to obtain $r(\rho_{ZYE}^{\gamma})$ in Eq. (5.9).

From the relation between ρ_{ABEC} and $\{F^{\gamma}\}$ shown in Eq. (5.6), we notice that a full knowledge of $\{F^{\gamma}\}$ gives us a full knowledge of ρ_{ABEC} and thus we can determine the key rate using Eq. (5.12). However, if we cannot uniquely determine F^{γ}, then we cannot uniquely determine ρ_{ABEC}. In that case, we have a set of compatible density operators ρ_{ABEC}, that is,

$$\mathbf{S} = \{\rho_{ABEC} : \mathrm{Tr}_{EC}(\rho_{ABEC}) = \rho_{AB}, \mathrm{Tr}(\rho_{ABEC}\,|x\rangle\!\langle x|_A \otimes |y\rangle\!\langle y|_B \otimes \mathbb{1}_E \otimes |\gamma\rangle\!\langle\gamma|_C) = p(x,y,\gamma)\}. \tag{5.15}$$

Thus, we need to consider the worst-case scenario by taking the minimum of $\tilde{r}$ over the set $\mathbf{S}$, or equivalently, over the set $\mathbf{S}' = \{\rho_{ZYEC} : \rho_{ZYEC} = \mathcal{G}(\rho_{ABEC}),\ \text{where } \rho_{ABEC} \in \mathbf{S}\}$.

In this situation, the asymptotic key rate R_∞ should be expressed as

$$R_\infty = \min_{\rho_{ZYEC}\in\mathbf{S}'} \tilde{r}(\rho_{ZYEC}) = \min_{\rho_{ABEC}\in\mathbf{S}} \tilde{r}(\mathcal{G}(\rho_{ABEC})). \tag{5.16}$$

The essential part of the optimization is to optimize the Holevo information $\chi(Z\!:E)$ by finding the all possible Eve's conditional states, which are needed to evaluate Eq. (5.10).

We remark that most of the discussion so far is general to a generic MDI QKD protocol. In the next subsection, we adapt this procedure to our specific PM-MDI QKD protocol.

5.2.4 Determination of Eve's POVM for PM-MDI QKD

As discussed in the previous subsections, knowing Eve's POVM elements allows us to calculate the key rate, since the minimization in Eq. (5.16) is now over a set containing only one element. We will now explain how our choice of test states (coherent states with a continuum of complex amplitudes) allows in principle to determine Eve's POVM elements.

For simplicity, let us concentrate on the case of testing a measurement device acting on a single mode (rather than the two-mode case of our protocol). Knowing some POVM element $\tilde{F}$ is equivalent to being able to predict the probability $p(\tilde{F})$ of the associated outcome for any input state ρ as $p(\tilde{F}) = \mathrm{Tr}\Big[\rho\,\tilde{F}\Big]$. We can now use the phase-space formalism of quantum mechanics (see for example [103, 104]) where we use the P-function representation of $\rho = \int d^2\alpha \; P(\alpha) \left|\alpha\right\rangle\!\left\langle\alpha\right|$ so that we have

$$p(\tilde{F}) = \int d^2\alpha \; P(\alpha) \; \langle\alpha| \, \tilde{F} \, |\alpha\rangle \; . \tag{5.17}$$

As we see from this equation, knowledge of the function $p(\tilde{F}|\alpha) := \langle\alpha| \, \tilde{F} \, |\alpha\rangle$ allows the prediction of $p(\tilde{F})$ for all input states for which the P-function of the density matrix ρ exists. So testing the measurement device with all possible coherent states $|\alpha\rangle$ and observing the corresponding probabilities $p(\tilde{F}|\alpha)$ is equivalent to knowing $\tilde{F}$.

Actually, using results from [103, 104] one can reconstruct the operator $\tilde{F}$ explicitly also in cases where the P-function of ρ may not exist. Let us go through the arguments directly for the POVM elements F^γ for the outcome γ in the two-mode case. We adapt [104, Eqs. (3.4)-(3.6)] to our scenario.

By substituting [104, Eqs. (3.4) and (3.6)] into [104, Eq. (3.5)], we obtain a power series for each F^γ as:

$$F^\gamma = \sum_{n_1,n_2,m_1,m_2=0}^{\infty} \frac{(\partial_{\alpha_1}^{m_1}\partial_{\alpha_2}^{m_2}\partial_{\bar{\alpha}_1}^{n_1}\partial_{\bar{\alpha}_2}^{n_2} \langle\alpha_1,\alpha_2| \, F^\gamma \, |\alpha_1,\alpha_2\rangle)|_{\alpha_1=0,\alpha_2=0}}{m_1!m_2!n_1!n_2!} ({a_1}^\dagger)^{n_1} ({a_2}^\dagger)^{n_2} {a_1}^{m_1} {a_2}^{m_2}, \tag{5.18}$$

where α_1, α_2 and their complex conjugated counterparts $\bar{\alpha}_1, \bar{\alpha}_2$ are treated as independent variables, and $a_1, {a_1}^\dagger$, $a_2, {a_2}^\dagger$ are the annihilation and creation operators of the two modes. Since F^γ is a POVM element and thus has bounded eigenvalues, such series exist and converge [103]. Using the two-mode test states $|\alpha_1, \alpha_2\rangle$ and the associated observed probabilities $p(\gamma|\alpha_1, \alpha_2) = \langle\alpha_1,\alpha_2| \, F^\gamma \, |\alpha_1,\alpha_2\rangle$ thus uniquely determines F^γ.

Note that a full description of F^γ as shown above is more than what we actually need since we are only interested in how F^γ acts on the subspace span($\mathcal{S}$), which is only a four-dimensional space.

For this, we need to be able to calculate off-diagonal elements of the form $\langle\alpha_1,\alpha_2| \, F^\gamma \, |\beta_1,\beta_2\rangle$. It is an interesting question whether we can estimate these elements well enough with just a few number of coherent states. (The diagonal elements are directly accessible.) We present now the handle to attack this question.

We first notice that characterizing F^γ on span($\mathcal{S}$) is equivalent to the question whether the operator $\left|\beta_1,\beta_2\right\rangle\!\left\langle\alpha_1,\alpha_2\right|$ can be approximated to arbitrary precision in the Hilbert-Schmidt norm by the discrete diagonal coherent state representation [164, 165]:

$$\left|\beta_1,\beta_2\right\rangle\!\left\langle\alpha_1,\alpha_2\right| = \sum_{i=1}^{\infty} \lambda_i \left|\omega_1^{(i)},\omega_2^{(i)}\right\rangle\!\left\langle\omega_1^{(i)},\omega_2^{(i)}\right| \tag{5.19}$$

where we use sets of tensor products of coherent states $\left|\omega_1^{(i)},\omega_2^{(i)}\right\rangle$ and complex numbers $\lambda_i \in \mathbb{C}$.

Then, we can write $\langle\alpha_1,\alpha_2| F^\gamma |\beta_1,\beta_2\rangle$ as a sum of observed values $\left\langle \omega_1^{(i)},\omega_2^{(i)}\right| F^\gamma \left|\omega_1^{(i)},\omega_2^{(i)}\right\rangle$ as

$$\begin{aligned}\langle\alpha_1,\alpha_2| F^\gamma |\beta_1,\beta_2\rangle &= \mathrm{Tr}(F^\gamma |\beta_1,\beta_2\rangle\langle\alpha_1,\alpha_2|) \\ &= \sum_{i=1}^{\infty} \lambda_i \left\langle \omega_1^{(i)},\omega_2^{(i)}\right| F^\gamma \left|\omega_1^{(i)},\omega_2^{(i)}\right\rangle .\end{aligned} \tag{5.20}$$

By appropriate choices of $\{\left|\omega_1^{(i)},\omega_2^{(i)}\right\rangle\}_{i=1}^N$, we will be able to get a good approximation by terminating the summation at N. From the approximation, we will then determine a set of POVMs compatible with experimental correlations, which is a neighborhood of the POVM that Eve actually performed. When we calculate the key rate in this case, we need to perform the minimization in Eq. (5.16). In that case, we may apply numerical methods [144, 146] to perform the desired optimization. If such an approximation makes this set of compatible POVMs small enough, then the key rate with several choices of test states would be close to the key rate with infinite choices of test states. We leave the detailed analysis of finite choices of test states scenario to the future work.

In Appendix B, we will discuss how to represent F^γ in the four-dimensional subspace span($\mathcal{S}$) after knowing $\langle\alpha_1,\alpha_2| F^\gamma |\beta_1,\beta_2\rangle$ for $|\alpha_1,\alpha_2\rangle, |\beta_1,\beta_2\rangle \in \mathcal{S}$.

5.3 Simulation

We perform simulations to study the loss scaling of this PM-MDI QKD protocol and also the stability of the protocol.

5.3.1 Loss-only scenario

To show that the key rate of this protocol has a scaling of $\sqrt{\eta}$ with the single photon transmissivity η between Alice and Bob, we first study the loss-only scenario. We simulate the quantum channel as a lossy channel and we consider the normal situation where Charlie (Eve) performs the measurements so that the observed statistics during the parameter estimation step is compatible with Charlie performing the measurement shown in Figure 5.1a. That is, we calculate the POVM F corresponding to the real setup. Our protocol can verify via test states in the test mode that this is the actual POVM performed by Eve in the loss-only scenario. For the purpose of our presentation, we consider a symmetric setup, that is, Charlie is at an equal distance from Alice and Bob, and the loss in each path is the same. For a total transmissivity η between Alice and Bob, each path has a transmissivity $\sqrt{\eta}$.

In this situation, when Alice sends a coherent state $|\alpha_A\rangle$ and Bob sends a coherent state $|\alpha_B\rangle$ in the same optical mode, the state becomes $\left|\sqrt{\sqrt{\eta}}\alpha_A, \sqrt{\sqrt{\eta}}\alpha_B\right\rangle$ after the lossy channel. When

α_A, α_B	$+\sqrt{\mu}, +\sqrt{\mu}$	$-\sqrt{\mu}, -\sqrt{\mu}$	$+\sqrt{\mu}, -\sqrt{\mu}$	$-\sqrt{\mu}, +\sqrt{\mu}$
$P(+\vert\alpha_A, \alpha_B)$	$1 - e^{-2\sqrt{\eta}\mu}$	$1 - e^{-2\sqrt{\eta}\mu}$	0	0
$P(-\vert\alpha_A, \alpha_B)$	0	0	$1 - e^{-2\sqrt{\eta}\mu}$	$1 - e^{-2\sqrt{\eta}\mu}$
$P(?\vert\alpha_A, \alpha_B)$	$e^{-2\sqrt{\eta}\mu}$	$e^{-2\sqrt{\eta}\mu}$	$e^{-2\sqrt{\eta}\mu}$	$e^{-2\sqrt{\eta}\mu}$
$P(d\vert\alpha_A, \alpha_B)$	0	0	0	0

Table 5.1: Conditional probability distribution of announcement outcomes given the states from $\mathcal{S}$ in the loss-only scenario. η is the single photon transmissivity between Alice and Bob and μ is the intensity of coherent states in the key-generation mode.

Charlie performs the measurement on this state, the probability for each announcement outcome γ can be calculated as follows:

$$\begin{aligned}
\langle\alpha_A, \alpha_B| F^{+} |\alpha_A, \alpha_B\rangle &= (1 - e^{-\frac{\sqrt{\eta}|\alpha_A+\alpha_B|^2}{2}})e^{-\frac{\sqrt{\eta}|\alpha_A-\alpha_B|^2}{2}} \\
\langle\alpha_A, \alpha_B| F^{-} |\alpha_A, \alpha_B\rangle &= e^{-\frac{\sqrt{\eta}|\alpha_A+\alpha_B|^2}{2}}(1 - e^{-\frac{\sqrt{\eta}|\alpha_A-\alpha_B|^2}{2}}) \\
\langle\alpha_A, \alpha_B| F^{?} |\alpha_A, \alpha_B\rangle &= e^{-\frac{\sqrt{\eta}|\alpha_A+\alpha_B|^2}{2}}e^{-\frac{\sqrt{\eta}|\alpha_A-\alpha_B|^2}{2}} \\
\langle\alpha_A, \alpha_B| F^{d} |\alpha_A, \alpha_B\rangle &= (1 - e^{-\frac{\sqrt{\eta}|\alpha_A+\alpha_B|^2}{2}})(1 - e^{-\frac{\sqrt{\eta}|\alpha_A-\alpha_B|^2}{2}})
\end{aligned} \tag{5.21}$$

Specifically, the conditional probability of each announcement outcome for each state in the set $\mathcal{S}$ is summarized in Table 5.1. From this table, we can directly evaluate the classical mutual information $I(K:Y)$ as

$$\begin{aligned}
I(K:Y)_{\rho^{+}_{KYE}} &= I(K:Y)_{\rho^{-}_{KYE}} &= 1, \\
I(K:Y)_{\rho^{?}_{KYE}} &= I(K:Y)_{\rho^{d}_{KYE}} &= 0.
\end{aligned} \tag{5.22}$$

Clearly, we cannot distill keys from $\gamma =$ "?" and $\gamma =$ "d" announcements. So, we will only evaluate $\chi(K:E)$ for $\gamma =$ " $+$ " and $\gamma =$ " $-$ ". We first find conditional states $\rho_E^{k,y,+}$ and $\rho_E^{k,y,-}$ defined in Eq. (5.13).

As we can notice from Table 5.1, in the loss-only scenario, whenever Alice and Bob prepare coherent states with a π phase difference, Charlie will never announce $\gamma =$ " $+$ " and whenever they prepare coherent states with the same phase, Charlie will never announce $\gamma =$ "$-$". Because $p(0,1,+) = p(1,0,+) = 0$ and $p(0,0,-) = p(1,1,-) = 0$, each of the states $\rho_E^{k,+}$ and $\rho_E^{k,-}$ $\forall k \in \{0,1\}$ is a pure state so that $H(\rho_E^{k,+}) = H(\rho_E^{k,-}) = 0$.

Therefore, we only need to evaluate $H(\rho_E^+)$ and $H(\rho_E^-)$. In this loss-only case,

$$\begin{aligned}\rho_E^+ &= \frac{1}{2}(\left|\Theta_{0,0}^+\right\rangle\!\left\langle\Theta_{0,0}^+\right| + \left|\Theta_{1,1}^+\right\rangle\!\left\langle\Theta_{1,1}^+\right|)\\ \rho_E^- &= \frac{1}{2}(\left|\Theta_{0,1}^-\right\rangle\!\left\langle\Theta_{0,1}^-\right| + \left|\Theta_{1,0}^-\right\rangle\!\left\langle\Theta_{1,0}^-\right|)\end{aligned} \tag{5.23}$$

The eigenvalues of ρ_E^+ are $\frac{1}{2}(1 \pm \left|\left\langle\Theta_{0,0}^+\middle|\Theta_{1,1}^+\right\rangle\right|)$ and thus $H(\rho_E^+) = h(\frac{1-|\langle\Theta_{0,0}^+|\Theta_{1,1}^+\rangle|}{2})$. Similarly, the eigenvalues of ρ_E^- are $\frac{1}{2}(1 \pm \left|\left\langle\Theta_{0,1}^-\middle|\Theta_{1,0}^-\right\rangle\right|)$ and thus $H(\rho_E^-) = h(\frac{1-|\langle\Theta_{0,1}^-|\Theta_{1,0}^-\rangle|}{2})$. Using the definition of $\left|\Theta_{k,y}^\gamma\right\rangle$ in Eq. (5.14), we obtain

$$\begin{aligned}\left\langle\Theta_{0,0}^+\middle|\Theta_{1,1}^+\right\rangle &= \frac{\langle+\sqrt{\mu},+\sqrt{\mu}|\,F^+\,|-\sqrt{\mu},-\sqrt{\mu}\rangle}{\sqrt{\langle+\sqrt{\mu},+\sqrt{\mu}|\,F^+\,|+\sqrt{\mu},+\sqrt{\mu}\rangle\langle-\sqrt{\mu},-\sqrt{\mu}|\,F^+\,|-\sqrt{\mu},-\sqrt{\mu}\rangle}}\\ \left\langle\Theta_{0,1}^-\middle|\Theta_{1,0}^-\right\rangle &= \frac{\langle-\sqrt{\mu},+\sqrt{\mu}|\,F^-\,|+\sqrt{\mu},-\sqrt{\mu}\rangle}{\sqrt{\langle-\sqrt{\mu},+\sqrt{\mu}|\,F^-\,|-\sqrt{\mu},+\sqrt{\mu}\rangle\langle+\sqrt{\mu},-\sqrt{\mu}|\,F^-\,|+\sqrt{\mu},-\sqrt{\mu}\rangle}}\end{aligned} \tag{5.24}$$

Thus, we have $H(\rho_E^+) = H(\rho_E^-) = h(\frac{1-e^{-4\mu(1-\sqrt{\eta})}e^{-2\mu\sqrt{\eta}}}{2})$. We provide explicit expressions of F^γ for this loss-only scenario in Appendix B.2.1, using which the reader can check the result directly.

Finally, we obtain the expression of secret key generation rate as a function of η and the intensity μ in this loss-only scenario as

$$R_\infty = (1 - e^{-2\mu\sqrt{\eta}})[1 - h(\frac{1 - e^{-4\mu(1-\sqrt{\eta})}e^{-2\mu\sqrt{\eta}}}{2})]. \tag{5.25}$$

When we optimize μ, we find optimal $\mu = O(1)$. After substituting the optimal μ, we can easily verify R_∞ scales as $\sqrt{\eta}$.

In Figure 5.2, the blue dashed line is the asymptotic key rate of this loss-only scenario as a function of the transmission distance L, where we take $\eta = 10^{-\frac{0.2L}{10}}$ and μ is optimized for each distance L. The red solid line is the fundamental repeaterless bound $-\log_2(1-10^{-\frac{0.2L}{10}})$ [65]. This calculation gives an intuitive understanding on how the PM-MDI QKD can beat the repeaterless key rate bound. We see that this PM-MDI QKD protocol beats the repeaterless bound at around 150 km. Our key rate expression in Eq. (5.25) is tight for the loss-only scenario. Therefore, we expect this is the loss limit for PM-MDI QKD.

5.3.2 Realistic Imperfections

It is of practical interests to study how stable this protocol is in noisy scenarios. In particular, we simulate the scenario with realistic imperfections in experimental devices, including the dark counts of detectors, mode mismatch and phase mismatch, detector inefficiency, and error correction inefficiency. In this section, we briefly introduce sources of imperfections and corresponding simulation parameters, then explain the correlations that Alice and Bob would observe in our simulation model, and finally present the results of our key rate calculation. In Appendix B.2.2, we provide more detailed explanations for the physical model of each imperfection.

For the purpose of presentation, we assume that both detectors have the same detector efficiency η_d and the same dark count probability p_d. We remark that the simulation method described in Appendix B.2.2 is also applicable to more general situations.

In the ideal implementation of this protocol, Alice and Bob should prepare coherent states in the same optical mode, that is, with the same spectral, temporal profiles and the same polarization, in order to have single-photon interference at the beam splitter. In reality, since their states may come from different lasers and pass through different optical components before reaching the central node, the modes of their states are not necessarily perfectly matched. Thus, we consider the relative mode mismatch between their states with a simulation parameter V. In our simulation, if without any mode mismatch, the state arriving at the central node from Alice and Bob is supposed to be $|\alpha_A, \alpha_B\rangle$, then with the mode mismatch, the state becomes $\left|\alpha_A, \sqrt{V}\alpha_B\right\rangle$ in the original mode and $\left|0, \sqrt{1-V}\alpha_B\right\rangle$ in a second mode. Both modes enter Charlie's devices independently.

Another source of imperfection considered in our simulation model is the phase mismatch. In the key-generation mode, Alice and Bob are supposed to prepare states in the set $\mathcal{S}$, which are coherent states with the same global phase and with the encoding information in the relative phases. In reality, the global phase is not guaranteed to be the same when states reach the detectors. Therefore, we consider the situation where there is a relative phase mismatch between Alice's signal state and Bob's signal state. If without any phase mismatch, the state is supposed to be $|\alpha_A, \alpha_B\rangle$, then due to the phase mismatch, the state is changed to $\left|\alpha_A, \alpha_B e^{i\delta}\right\rangle$ with a simulation parameter δ.

Table 5.2 lists the choice of parameters in our simulation. In particular, We choose the same values for the efficiency of a detector η_d and the dark count probability of a detector p_d as those used in the Ref. [63] for comparison purpose. We also select pessimistic values for the mode mismatch and phase mismatch to demonstrate the feasibility of beating the repeaterless bound with currently available devices.

We give the expressions for the probability of each announcement outcome γ given each choice of the input state in terms of the simulation parameters V, δ, η_d, and p_d. We define the total

Detector efficiency η_d	14.5%
Detector dark count probability p_d	8×10^{-8}
Mode mismatch $(1 - V)$	5%
Phase mismatch δ	$\frac{\pi}{60}$
Error correction efficiency f_{EC}	1.15

Table 5.2: Values for simulation parameters. They are experimentally feasible and might be pessimistic values. See main text for more explanations.

transmissivity as $\eta = \eta_t \eta_d^2$, where η_t is the channel transmission probability between Alice and Bob.

$$\begin{aligned}
\langle \alpha_A, \alpha_B | \, F^{+} \, | \alpha_A, \alpha_B \rangle &= (1 - p_d)(1 - \xi_1 \xi_2) \xi_2 \xi_3 + (1 - p_d) p_d \xi_1 \xi_2^2 \xi_3 \\
\langle \alpha_A, \alpha_B | \, F^{-} \, | \alpha_A, \alpha_B \rangle &= (1 - p_d) \xi_1 \xi_2 (1 - \xi_2 \xi_3) + (1 - p_d) p_d \xi_1 \xi_2^2 \xi_3 \\
\langle \alpha_A, \alpha_B | \, F^{?} \, | \alpha_A, \alpha_B \rangle &= (1 - p_d)^2 \xi_1 \xi_2^2 \xi_3 \\
\langle \alpha_A, \alpha_B | \, F^{d} \, | \alpha_A, \alpha_B \rangle &= p_d (1 - \xi_1 \xi_2) \xi_2 \xi_3 + p_d \xi_1 \xi_2 (1 - \xi_2 \xi_3) \\
&\quad + p_d^2 \xi_1 \xi_2^2 \xi_3 + (1 - \xi_1 \xi_2)(1 - \xi_2 \xi_3),
\end{aligned} \tag{5.26}$$

where for the simplicity of writing, we have made the following definitions

$$\begin{aligned}
\xi_1 &= e^{-\frac{1}{2}\sqrt{\eta}\left|\alpha_A + \sqrt{V}\alpha_B e^{i\delta}\right|^2} , \\
\xi_2 &= e^{-\frac{1}{2}\sqrt{\eta}(1-V)\left|\alpha_B\right|^2} , \\
\xi_3 &= e^{-\frac{1}{2}\sqrt{\eta}\left|\alpha_A - \sqrt{V}\alpha_B e^{i\delta}\right|^2} .
\end{aligned} \tag{5.27}$$

From Eq. (5.26), it is straightforward to derive the conditional probability of each announcement outcome given the state in $\mathcal{S}$. Similarly to the loss-only scenario, we also discover that the mutual information $I(K\!:\!Y)$ is zero for $\gamma =$ "?" and $\gamma =$ "d" since the probability of making those outcomes is independent from the signal states sent by Alice and Bob. Thus, we only generate keys from $\gamma =$ " $+$ " and $\gamma =$ " $-$ " outcomes.

We define error rates e_+ and e_- given the announcement outcome $\gamma =$ " $+$ " and $\gamma =$ " $-$ ", respectively.

$$\begin{aligned}
e_+ &:= p(0,1|+) + p(1,0|+) = \frac{\zeta_2 - (1 - p_d)\zeta_1 \zeta_2}{\zeta_1 + \zeta_2 - 2(1 - p_d)\zeta_1 \zeta_2} \\
e_- &:= p(0,0|-) + p(1,1|-) \ = e_+,
\end{aligned} \tag{5.28}$$

where we define $\zeta_1 = e^{-\sqrt{\eta}\mu(1-\sqrt{V}\cos(\delta))}$ and $\zeta_2 = e^{-\sqrt{\eta}\mu(1+\sqrt{V}\cos(\delta))}$.

The rest of the task is to find each of $\rho_E^{k,y,+}$ and $\rho_E^{k,y,-}$. In Appendix B.2.3, we give explicit expressions for F in this scenario. Using Eqs. (B.3) and (B.21), we can again find the four-dimensional representation of each of $\rho_E^{k,y,+}$ and $\rho_E^{k,y,-}$. We numerically evaluate the Holevo information $\chi(K{:}E)$ for $\gamma =$ " $+$ " and $\gamma =$ " $-$ " (even though it is still possible but non-trivial to evaluate $\chi(K{:}E)$ analytically).

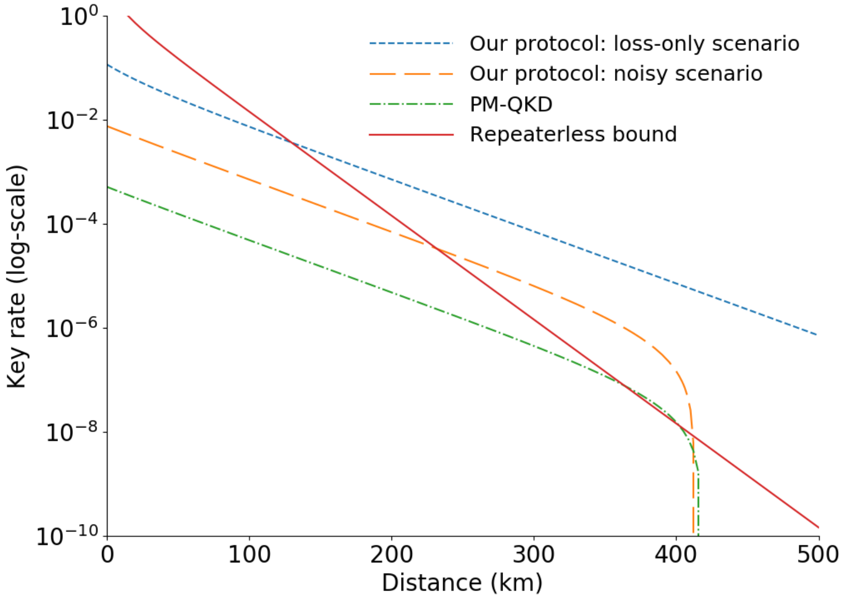


Figure 5.2: A log-linear plot for the key rate as a function of the transmission distance. The red solid line is the fundamental repeaterless secret key capacity bound $-\log_2(1-10^{-\frac{0.2L}{10}})$ for a transmission distance L (in kilometers) [65]. The blue dashed line is the loss-only key rate for PM-MDI QKD after optimizing μ in Eq. (5.25). The orange solid line is the key rate for PM-MDI QKD with experimentally feasible parameters listed in Table 5.2. The green dash-dotted line is the key rate for phase-matching QKD [63] with similar parameters.

In Figure 5.2, the orange solid line shows the result of our simulation. With those experimentally feasible parameters, we see this PM-MDI QKD protocol can still beat the repeaterless bound and this crossover happens at around 250 km.

5.4 About finite test states

As briefly mentioned in Section 5.2.4, the purpose of using test states is to determine each F^γ in the four-dimensional subspace span$(\mathcal{S})$ by determining $\langle \alpha_1, \alpha_2 | F^\gamma | \beta_1, \beta_2 \rangle$ for each $|\alpha_1, \alpha_2\rangle, |\beta_1, \beta_2\rangle \in \mathcal{S}$. While it remains an open question how to choose optimal test states, we discuss some observations and ideas. One observation is that due to the symmetric setup between Alice and Bob, it seems enough to consider only a single-mode scenario. Alice and Bob would then use the same choices of test states. From the definition of $\mathcal{S}$ in Eq. (5.2), it is then enough to approximate $|+\sqrt{\mu}\rangle\langle-\sqrt{\mu}|$ and $|-\sqrt{\mu}\rangle\langle+\sqrt{\mu}|$. These two are related by Hermitian conjugation. Therefore, our main task is to approximate $|+\sqrt{\mu}\rangle\langle-\sqrt{\mu}|$ by a finite collection of coherent states $\{|\omega^{(i)}\rangle\langle\omega^{(i)}|\}_{i=1}^{N}$. One simple idea is to consider a greedy algorithm by minimizing the difference between the actual operator $|+\sqrt{\mu}\rangle\langle-\sqrt{\mu}|$ and the diagonal approximation with N coherent states at each step, where N may start with 3 (or 1) and may increase during the execution of the algorithm. (We can either use the existing data from $|+\sqrt{\mu}\rangle\langle+\sqrt{\mu}|$ and $|-\sqrt{\mu}\rangle\langle-\sqrt{\mu}|$ or start without them.) The difference may be quantified by Hilbert-Schmidt norm, which is a natural choice that makes calculation simple. Once the set of test states is determined, we can then apply the numerical method [144] presented in Section 4.1 to calculate key rates. A main obstacle is the dimension of the problem is quite large if we do it in a canonical way. Since this protocol is an MDI protocol, we need to optimize ρ_{ABC} where C is a register that stores classical announcements made by Charlie. If N test states are used, then $\dim(\mathcal{H}_A) = \dim(\mathcal{H}_B) = 2 + N$. Even if Charlie makes only two announcements: successful or unsuccessful events, the size of ρ_{ABC} is $2(2+N)^2$. The constraint that ρ_{AB} is fixed gives $(2+N)^4$ linear constraints. Thus, this optimization problem becomes numerically challenging even for small N (say 4 test states). On top of each key rate calculation, we need to optimize protocol parameters and choices of test states to find regimes where the protocol can overcome the repeaterless bound. While it might be still possible to do the calculation, it requires a lot of computational resources to do so. We leave it as a future work to explore this direction.

Interestingly, we notice that shortly after our work [71], Primaatmaja et. al. [93] investigates this protocol with finite test states using a different numerical security proof method. This method is based on characterizing the Gram matrix G for Eve's quantum side information in the MDI setup. In particular, this Gram matrix G is positive semidefinite and its entries are related to the probability of successful measurement outcomes at the middle station as well as bit and phase error rates. It formulates an SDP problem to find the maximal phase error given observed statistics. Then it uses the phase error correction framework to find the asymptotic key rate against collective attacks. In particular, from their simulations, they show that two test states are sufficient to overcome the PLOB bound in the noiseless scenario; four test states are enough in a noisy scenario. In summary, they find that the repeaterless bound can be overcome with only a few test states. Their approach has an advantage in running time since the SDP problem is a standard linear SDP where our problem is nonlinear SDP. In principle, we can improve upon their key rates using our method as our method can typically give tight key rates while the phase error correction approach usually introduces looseness in finding the phase error rate. Therefore, an interesting direction is to combine our approach with theirs. We can use their approach to search for optimal parameters and with the fixed parameters, we then use our method to calculate key rates.

5.5 Other protocol variants

Many variants of the PM-MDI QKD protocol have been proposed and investigated. Different names have been assigned to different variations, such as, phase-encoding scheme for MDI [161], MDI-B92 [162], TF-QKD[62] and PM-MDI QKD [63]. However, they all have the essential components needed to achieve the rate scaling of $R_\infty = O(\sqrt{\eta})$, namely, they all use coherent states as signal states and rely on single-photon interference events at the beam splitter of an untrusted intermediate node, even though not all variations can indeed achieve this scaling.

We first describe the common features of all those protocols and then discuss how each variation differs in the following two aspects: choices of signal states used for establishing secret keys and choices of test states used to probe Eve's attacks.

In an ideal PM-MDI QKD protocol, Alice and Bob will only establish keys from the rounds where each of them has selected a state from the set $\{\left|+\alpha\right\rangle, \left|-\alpha\right\rangle\}$, where α can be an arbitrary complex number. In other words, Alice and Bob will only establish keys from the rounds that satisfies the phase-matching condition, that is, they have chosen the same global phase and same intensity for their states. We call two coherent states $\{\left|+\alpha\right\rangle, \left|-\alpha\right\rangle\}$ with only a π phase difference as a phase-matching pair. In addition, Alice and Bob may decide to send some states as test states to probe eavesdropping activities for randomly selected rounds and those rounds will be used in the parameter estimation step only. Alice and Bob will send their states to an untrusted party Charlie at the intermediate node for measurements. An honest Charlie will use the single-photon inference events at the beam splitter for his announcement.

Since this type of protocol is measurement-device independent and generates keys when Alice and Bob use the same phase-matching pair of coherent states, phase-matching measurement-device independent QKD is in our view a more descriptive name that captures important features of this type of protocol. In the literature, many authors prefer to use either TF-QKD following the proposal [62] or phase-matching QKD [63] as the name.

Now, we compare some variations of PM-MDI QKD. Different variations may use different number of phase-matching pairs as signal states and may use different types of states as test states, such as a mixture of photon number states (phase-randomized coherent states), partially phase-randomized coherent states, and coherent states without phase randomization. Some variations may use the same number of phase-matching pairs as signal states, but differ in how to handle them. We present those variations just for the comparison purpose. We note that the following list is not an exhaustive list of all variants and we mention only these works that appeared around the similar time as our work. In the next section, we discuss some recent progresses that appear after our work.

1. The variant proposed in [161] is called phase encoding scheme I for MDI. This protocol essentially uses two phase-matching pairs of coherent states. In the original description of the protocol, these two pairs are labeled as two bases, similar to a BB84-type protocol, due to the proof technique adopted. In an abstract description, we can view this protocol as essentially using one phase-matching pair of coherent states as signal states and an additional pair as test states. Due to the proof technique and a limited number of test states, the $\sqrt{\eta}$ scaling of the asymptotic key rate R_∞ was not found.

2. The variant studied in [162] is called MDI-B92 protocol. Ref. [162] analyzes different types of measurements for the intermediate node. Under the investigation of unambiguous state discrimination attacks, it basically proposes a variation of PM-MDI protocol with exactly one phase-matching pair of coherent states as the signal states and no test states. Because there are no test states, this protocol is not expected to have the $\sqrt{\eta}$ scaling for the asymptotic key rate R_∞.

3. The variant proposed in [62] has the name of TF-QKD protocol. This protocol uses infinitely many phase-matching pairs of coherent states (phase-randomized coherent states) as signal states. In addition, for the purpose of security analysis, each round is assigned to one of two bases to mimic a BB84-type protocol. Instead of achieving the perfect phase-matching conditions, this protocol allows some small errors in identifying whether Alice and Bob have chosen the same phase-matching pair. To distill keys, Alice and Bob disclose some partial information about the global phases. If their global phases only differ by a small amount, they assume they have used the same phase-matching pair. In this protocol, states used as test states are effectively the same as states used for signal states at a cost of introducing a sifting loss. These test states are partially phase-randomized coherent states as Eve knows some partial information about the global phase. Ref. [62] argued that this type of protocol can have the $\sqrt{\eta}$ scaling for the asymptotic key rate R_∞.

4. The variant investigated in [63] uses the name PM-QKD protocol. Similar to TF-QKD [62], it also uses infinitely many phase-matching pairs of coherent states as signal states and adopts a similar procedure as [62] in identifying whether Alice and Bob have chosen the same phase-matching pair for each round. It also uses partially phase-randomized coherent states as test states. The difference from TF-QKD is that there is no assignment of basis choice for each round. The security analysis does not use the standard decoy state methods. It confirms with the $\sqrt{\eta}$ scaling for the asymptotic key rate R_∞.

5. The variant studied in [69] is called TF-QKD* protocol. This protocol, similar to the original TF-QKD protocol, uses infinitely many phase-matching pairs as signal states and later post-selects on rounds where the global phases are different by less than a small amount. Effectively, by allowing some errors, Alice and Bob assume that they have chosen the same phase-matching pair when the difference in their global phases is small. This protocol also has an assignment of basis choice for each round in order to apply a BB84-type security argument. Different from the original TF-QKD protocol, this protocol uses a mixture of photon number states as test states. The security analysis applies the standard decoy state methods and also confirms the $\sqrt{\eta}$ scaling for the asymptotic key rate R_∞.

6. The variant proposed in [72] was initially called PM-QKD protocol and then renamed as TF-QKD. It uses exactly one phase-matching pair as signal states and uses a mixture of photon number states as test states. The $\sqrt{\eta}$ scaling is verified.

7. The variant studied in [73, Protocol 3] is referred as a TF-QKD type protocol. This variation essentially is the same as in [72]. It uses exactly one phase-matching pair as signal states and uses a mixture of photon number states as test states. These two works [72, 73] differ by the security proof methods. The $\sqrt{\eta}$ scaling is confirmed.

8. The variant proposed in [70] is called sending-or-not-sending TF-QKD. This variant is quite different from other variants since it does not use phase-matching pairs of the form $\{\left|+\alpha\right\rangle, \left|-\alpha\right\rangle\}$. Instead, it uses $\{\left|0\right\rangle, \left|\alpha\right\rangle\}$. To guarantee security, Alice (Bob) needs to send the vacuum state with a high probability (much close to 1). Instead of distinguishing the phase differences between Alice's and Bob's signals, the immediate node is supposed to detect presence or absence of photons. The security analysis of the work [70] is based on the phase error rate approach. It also shows the $\sqrt{\eta}$ scaling for the asymptotic key rate R_∞.

In the end, we remark that the advantages of different types of test states. Using a mixture of photon number states as test states allows the standard decoy state analysis, which has been investigated and well understood. In addition, using a small number of decoy states as test states has been investigated in many other protocols and might be readily adapted to some variations of PM-MDI QKD protocol. On the other hand, using coherent states as test states has the potential to give tighter key rates, as we will demonstrate in this paper when using infinitely many coherent states. Also, it does not require the phase randomization in the experimental implementations.

5.6 Recent progress

PM-MDI QKD (or TF-QKD or PM-QKD) has been a hot topic in the QKD community since its initial proposal [62, 63] to overcome the repeaterless bound. Soon after early theory works mentioned in the previous section, there have been quite a few experimental demonstrations [166–169] that overcome the repeaterless bound. Since early theory papers focus on the asymptotic key rates, there have been many papers on the finite key analysis of the protocol and they aim at providing tighter and tighter finite key rates [170, 171]. The initial works focus on the symmetric scenario where Alice and Bob are at an equal distance from the intermediate node for simplicity of calculation. It was expected that these security proofs could be generalized to the asymmetric scenario where Alice-Charlie and Bob-Charlie channels have different losses. Explicit calculation was done in two independent works about asymmetric TF-QKD [172, 173]. Initial experimental (proof-of-principle) demonstrations use a single laser for both Alice and Bob since it is quite challenging to lock phases of two independent lasers and phase-locking is needed in those protocol variants. Experimental demonstrations based on two independent lasers are done in Refs. [174, 175]. In Ref. [174], they implement the sending-or-not-sending variant of TF-QKD and apply a remote frequency-locking technique to lock two independent lasers. They are able to reach over 500 km and beat the repeaterless bound in the absolute limit, which is the PLOB bound that uses the channel transmittance as η instead of the overall transmittance that includes the efficiency of imperfect detectors. In Ref. [175], they implement the phase-matching protocol [63, 176] and employ the laser injection technique and the phase post-compensation method to match the modes of two independent lasers. They are able to overcome the repeaterless bound via 302 km and 402 km commercial-fiber channels.

In many variants (except our variant [71, 93]), while the signal states are non-phase-randomized, they require the test states to be fully phase-randomized in order to apply the decoy-state analysis [66, 68]. In practice, continuous phase randomization might not be possible. Failure to satisfy this assumption can open up security loopholes. There have been several works to study the discrete phase randomization for TF-QKD [177–179], which has previously been studied for BB84 [180]. We note that Ref. [93] prior to these works has already studied a practical version of the protocol without the need of continuous phase randomization. This variant is in some sense a discrete-phase-randomized protocol.

There are many interesting directions inspired by TF-QKD. For example, Ref. [181] studies a conference key agreement protocol that is based on the single-photon interference, a key idea for TF-QKD. TF-QKD can also be used in the quantum digital signature schemes based on the work [17].

5.7 Future directions

For our protocol variant, it is interesting to explore the scenario with finite test states using the numerical method [144] presented in Section 4.1. The current main challenge that prevents us from doing such an analysis is that the dimension of the problem is relatively large and for a fixed choice of parameters (e.g. choices of test states, signal state intensity, and probability of testing), the optimization problem takes a significantly amount of computer time (even with the improved algorithm in [87]). The parameter regimes for this situation are not well investigated. The exploration of parameter regimes seems computationally challenging if we just use the canonical treatment of the problem setup and current best algorithm. A meaningful exploration along this direction requires a dimension reduction for the representation on the source sides. We may use the approach from [93] to search for optimal parameters and then apply ours to get tighter key rates. It is also desirable to perform the finite-key analysis for a practical version of the protocol. Moreover, it is worth studying the source imperfections in this protocol and take into account potential side channels in a revised analysis.

Chapter 6

Asymptotic security analysis of DMCVQKD with heterodyne detection

CVQKD protocols with a discrete modulation are interesting due to their experimental simplicity and their great potential for massive deployment in the quantum-secured networks, but their security analysis is less advanced than that of Gaussian modulation schemes. We apply our numerical method to analyze the security of discrete-modulation protocols against collective attacks in the asymptotic limit, paving the way for a full security proof with finite-size effects. While our method is general for discrete-modulation schemes, we focus on two variants of the CVQKD protocol with quaternary modulation. Interestingly, thanks to the tightness of our proof method, we show that this protocol is capable of achieving much higher key rates over significantly longer distances with experimentally feasible parameters compared with previous security proofs of binary and ternary modulation schemes and also yielding key rates comparable to Gaussian modulation schemes. Furthermore, as our security analysis method is versatile, it allows us to evaluate variations of the discrete-modulated protocols, including direct and reverse reconciliation, and postselection strategies. In particular, we demonstrate that postselection of data in combination with reverse reconciliation can improve the key rates.

A main security proof technique for CVQKD is the optimality of Gaussian attacks [182, 183] for protocols with Gaussian modulation. In fact, there are many security proofs for CVQKD with a Gaussian modulation and some protocols have been proved to be secure against general attacks (see [20] for a review). However, this type of protocol puts a lot of demands on the modulation devices and classical error-correction protocols. In addition, the effect of a finite constellation needs to be taken into account carefully [184, 185]. In probing quantumness of devices using coherent states, we notice that even a small number of coherent states have the same quantumness verification power as a Gaussian modulation of states [186]. We thus expect that a DMCVQKD protocol will approach the performance of Gaussian-modulated CVQKD with just a few different modulation amplitudes. However, the corresponding security proof is more involved due to missing analytical tools. The binary [187] and ternary modulation schemes [188] have been proved secure against collective attacks. Unfortunately, the key rates obtained are not tight, and the proof technique is not expected to be generalizable to discrete-modulation schemes with more states. For the quaternary modulation scheme, also known as the quadrature phase-shift keying (QPSK) scheme, its security was previously analyzed under the assumption of linear bosonic channels[8] [189] or Gaussian attacks [38, 190], which restricts Eve's ability.

We apply a versatile numerical method to study the security of discrete-modulated schemes with a focus on the quaternary modulation scheme. Specifically, we analyze two variants of the QPSK modulation scheme: one with heterodyne detection in this chapter and the other one with homodyne detection in Chapter 7. Our method enables us to obtain tight key rates against collective attacks in the asymptotic limit. During the preparation of our work [81], we noticed an independent work [191] that analyzes the asymptotic security of the quaternary modulation scheme with heterodyne detection. In this security analysis [191], Ghorai *et al.* use a reduction to the Gaussian optimality proof method and apply an SDP technique with a photon-number cutoff assumption. We emphasize here that our proof technique is quite different from their work, and, in particular, we do not invoke the arguments of Gaussian optimality. For this reason, we also directly compare our results with their results in this work. Remarkably, compared with the similar heterodyne scheme considered in Ref. [191], we obtain higher key rates. Furthermore, our approach can be extended to variants of the protocol using homodyne measurements. Since our method does not rely on the arguments of Gaussian optimality, it also allows us to investigate the effects of the postselection of data [192, 193], which is not considered in Ref. [191] due to their proof technique. Postselection is commonly done for the classical telecommunication protocols and DVQKD protocols to discard noisy data and to improve the performance of the protocols. However, postselection strategies are currently not compatible with the Gaussian optimality proof technique, since the relevant states are non-Gaussian and Gaussian attacks are not expected to be optimal in the presence of postselection. Previously, postselection for discrete-modulation schemes was considered under a restricted class of attacks [38, 194]. We show that postselection can improve the key rates under collective attacks. Finally, we remark that our security proof method works for both direct reconciliation and reverse reconciliation protocols. However, we focus on reverse reconciliation in this work, since reverse reconciliation is known to have better performance than the direct reconciliation in terms of transmission distances.

For our security analysis, we rely on the numerical key rate optimization methods developed in Refs. [144, 146], and we use the version of Ref. [144] (Algorithm 4.1) to prove the security

[8] While a security analysis with additional assumptions usually can serve as an upper bound for the true key rate, unfortunately, the analysis [189] with the additional linear channel assumption is not expected to be tight and, thus, cannot be used as an upper bound of the key rate.

against collective attacks in the asymptotic limit. As we further develop the framework to handle the classical postprocessing for the numerical method presented in Ref. [144], this development makes it easier for us to study the postselection strategies and also simplifies some aspects of the numerical calculation. In order to perform such an optimization numerically, we impose the photon-number cutoff assumption, which is the same assumption considered in Ref. [191]. Although, ultimately, one would like to prove the security without this assumption, this assumption is numerically reasonable because we numerically verify that our key rate results do not depend on the choice of cutoff when the cutoff photon number is much larger than the mean photon number of each received state. It is also interesting to point out that even though we demonstrate our proof method only on the QPSK scheme here, our approach can be easily generalized to other discrete-modulation schemes beyond four coherent states.

After our initial work [81], we consider the trusted detector noise scenario where detector imperfections (detector inefficiency and electronic noise) are not accessible to Eve in [82]. We remark that Gaussian modulation schemes have been analyzed in the trusted detector noise scenario [29, 195–197] and it is known that the effects of electronic noise and detector inefficiency on the key rates are not very significant in the trusted detector noise scenario under realistic experimental conditions compared to the ideal detector scenario. As we show here, this observation also holds for discrete modulation schemes. However, we emphasize that our analysis is not a trivial application of the method used for Gaussian modulation protocols and instead we adopt a different approach. The reason is that the previous method used in the Gaussian modulation protocols relies on the fact [182, 183] that the optimal Eve's attacks for Gaussian modulation schemes correspond to Gaussian channels which make it easy to decouple the trusted detector noise from the channel noise when one looks at the covariance matrix. However, we cannot assume Gaussian channels here since Gaussian attacks are not expected to be optimal for discrete modulation schemes. In our analysis, based on a (commonly used) quantum optical model of the imperfect detector, we find its corresponding mathematical description in terms of POVM and then use this POVM to construct observables corresponding to quantities that are measured experimentally. These observables are then used in our security proof. We also point out the crucial difference between our analysis and Ref. [198] for discrete modulation schemes: Our asymptotic analysis is valid against arbitrary collective attacks while Ref. [198] uses the Gaussian channel assumption and thus its security analysis [198] is restricted to Gaussian collective attacks.

By developing a dimension reduction method, our recent work by Upadhyaya *et. al.* [85] removes the photon-number cutoff assumption. We summarize the main idea of the general framework in Section 4.2.3 and we briefly present its application of this framework to DMCVQKD with the heterodyne detection in Section 6.10. More details can be found in [85].

This chapter is based on [81, 82, 85].

6.1 Protocol description

In the following description, let $[N]$ denote the set of positive integers from 1 to N. A coherent state with an amplitude α or γ is denoted by $|\alpha\rangle$ or $|\gamma\rangle$. While our proof method is general for discrete modulation schemes, our focus is on the QPSK scheme.

Protocol 6.1 DMCVQKD with quadrature phase-shift keying scheme and heterodyne detection

(1). *State preparation.*—Alice randomly selects $x \in \{0, 1, 2, 3\}$ with a uniform probability distribution ($p_x = \frac{1}{4}$). She prepares one of those four signal states $\{|\alpha e^{i\theta_x}\rangle : \theta_x = \frac{x\pi}{2}\}$and sends it to Bob.

(2). *Measurement.*—Upon receiving Alice's state, Bob performs a heterodyne measurement on the state, which can be described by a POVM $\{E_\gamma = \frac{1}{\pi} |\gamma\rangle\langle\gamma| : \gamma \in \mathbb{C}\}$. After applying this POVM, he obtains the measurement outcome $y_k \in \mathbb{C}$.

(3). *Announcement and sifting.*—After N rounds of first two steps, Alice and Bob determine a small subset $\mathcal{I}_{\text{test}} \subset [N]$. Rounds indexed by the set $\mathcal{I}_{\text{test}}$ are used for parameter estimation. They use the remaining rounds indexed by $\mathcal{I}_{\text{key}} = [N]/\mathcal{I}_{\text{test}}$ to generate keys. Let m denote the size of the index set $\mathcal{I}_{\text{key}}$ and let f be a bijective function from $[m]$ to $\mathcal{I}_{\text{key}}$. After sifting, Alice obtains her string $\mathbf{X} = (x_1, \ldots, x_m)$ by the following rule:

$$\forall j \in [m],\ x_j = \begin{cases} 0 & \text{if } |\psi_{f(j)}\rangle = |\alpha\rangle\,, \\ 1 & \text{if } |\psi_{f(j)}\rangle = |i\alpha\rangle\,, \\ 2 & \text{if } |\psi_{f(j)}\rangle = |-\alpha\rangle\,, \\ 3 & \text{if } |\psi_{f(j)}\rangle = |-i\alpha\rangle\,. \end{cases} \tag{6.1}$$

(4). *Parameter estimation.*—Alice and Bob perform the parameter estimation by disclosing all the information in the rounds indexed by the test set $\mathcal{I}_{\text{test}}$. To perform such an analysis, they process the data by computing quantities like the first and second moments of q and p quadratures conditioned on each of four states that Alice sends. These quantities allow them to constrain their joint state ρ_{AB}. They then calculate the secret key rate according to the optimization problem in Eq. (6.11). If their analysis shows that no secret keys can be generated, then they abort the protocol. Otherwise, they proceed.

(5). *Reverse reconciliation key map.*—Bob performs a key map to obtain his raw key string. This key map discretizes his measurement outcome y_k to an element in the set $\{0, 1, 2, 3, \perp\}$ for each $k \in \mathcal{I}_{\text{key}}$. As $y_k \in \mathbb{C}$, we write $y_k = |y_k|e^{i\theta_k}$, where $\theta_k \in [-\frac{\pi}{4}, \frac{7\pi}{4})$.

Bob sets each z_j of his key string $\mathbf{Z} = (z_1, \dots, z_m)$ according to the rule

$$z_j = \begin{cases} 0 & \text{if } \theta_{f(j)} \in [-\frac{\pi}{4} + \Delta_p, \frac{\pi}{4} - \Delta_p) \text{ and } |y_{f(j)}| \geq \Delta_a, \\ 1 & \text{if } \theta_{f(j)} \in [\frac{\pi}{4} + \Delta_p, \frac{3\pi}{4} - \Delta_p) \text{ and } |y_{f(j)}| \geq \Delta_a, \\ 2 & \text{if } \theta_{f(j)} \in [\frac{3\pi}{4} + \Delta_p, \frac{5\pi}{4} - \Delta_p) \text{ and } |y_{f(j)}| \geq \Delta_a, \\ 3 & \text{if } \theta_{f(j)} \in [\frac{5\pi}{4} + \Delta_p, \frac{7\pi}{4} - \Delta_p) \text{ and } |y_{f(j)}| \geq \Delta_a, \\ \perp & \text{if } \theta_{f(j)} \text{ and } |y_{f(j)}| \text{ are none of the above.} \end{cases} \tag{6.2}$$

$\Delta_a \geq 0$ and $\Delta_p \geq 0$ are two parameters related to postselection of data. A protocol without postselection can set $\Delta_a = \Delta_p = 0$. This key map is depicted in Figure 6.1. Positions with the symbol $\perp$ are deleted from their strings. Again, we use $\mathbf{X}, \mathbf{Z}$ to mean the strings after removing the positions related to $\perp$. $\mathbf{Z}$ is called the raw key string.

(6). *Error correction and privacy amplification.*—They perform error correction and privacy amplification to generate a secret key.

Alice and Bob may decide to recast their strings to binary strings before or during the error-correction step depending on their choice of error-correcting code. For the consistency of our presentation, we use the alphabet $\{0, 1, 2, 3\}$ in the following discussion.

6.2 Detector modeling: noisy heterodyne detection

In this section, we present one physical model for a noisy heterodyne detector and give the corresponding POVM description. We start with a slightly more general model and then we make a simplification for the ease of calculation at the end of this section. This simplified model then reduces to a model commonly used in the literature.

6.2.1 Trusted detector noise model

As a heterodyne detector consists of two homodyne detectors and a beam splitter, we consider imperfections in each homodyne detector. A homodyne detector may have non-unity detector efficiency and also have some amount of electronic noise which is the additional noise introduced to the measured data by its electronic components. In an experiment, one is able to measure the amount of electronic noise and the value of detector efficiency by a calibration routine. To model a realistic homodyne detector with non-unity detector efficiency and some amount of electronic noise, we use a quantum optical model which is used in Refs. [29, 195–198], although the source of this electronic noise is in the actual electronics part of the detector. An alternative view of the

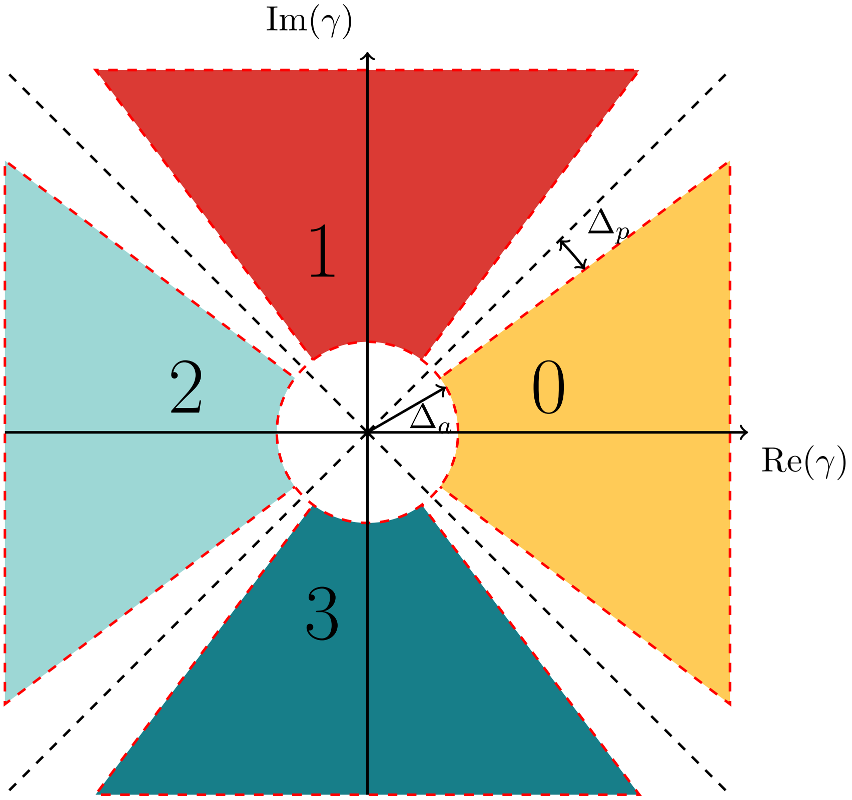


Figure 6.1: Key map for this protocol. When Bob has a measurement outcome $\gamma \in \mathbb{C}$, if γ is in one of the four shaded areas, then Bob maps the measurement outcome to the corresponding value of that area for his key string. If γ is not in the shaded areas, Bob obtains the symbol $\perp$. Δ_a and Δ_p are two parameters related to postselection.

electronic noise is that we can think about the detector as being a perfect detector followed by some classical postprocessing of the data, which adds noise. One should note that in a trusted device scenario, the characterization of the actual noise should be experimentally verified. Our physical model is chosen for convenience of calculating the POVM of the actual measurement. We depict this physical model of a noisy heterodyne detector in Figure 6.2. In this diagram, we consider a more general case where two homodyne detectors have different imperfections. We label the efficiency of the homodyne detector used for q quadrature measurement as η_1 and its electronic noise as ν_1 [expressed in shot noise unit (SNU)]. Similarly, the efficiency of the homodyne detector used for p quadrature measurement is labeled as η_2 and its electronic noise is labeled as ν_2.

Since our treatment for each homodyne detector in this heterodyne setup is the same, we take one homodyne detector (shown in each dashed box in Figure 6.2) as an example and treat the other one similarly by using its corresponding efficiency and electronic noise. An imperfect homodyne detector with its efficiency $\eta_j < 1$ and electronic noise $\nu_j \geq 0$ (for $j = 1$ or 2) can be modeled by a beam splitter placing before a perfect homodyne detector with the following specification. (1) The ratio of transmission to reflection of this beam splitter is $\eta_j : 1 - \eta_j$. (2) One input port of this beam splitter is the signal pulse and the other input port is a thermal state used to model electronic noise, which is equivalent to sending one mode of a two-mode squeezed vacuum state (EPR state) to the beam splitter. Each quadrature's variance of this ancillary thermal state is related to the value of electronic noise ν_j. More specifically, it is $[1 + \nu_j/(1 - \eta_j)]N_0$ [29], where $N_0 = 1/2$ denotes the shot-noise variance. In Figure 6.2, we choose to parametrize the thermal state in terms of its mean photon number as $\bar{n}_j = \frac{\nu_j}{2(1-\eta_j)}$ instead of the variance of each quadrature, which is convenient for writing of expressions in later sections[9]. We note that this way of modeling electronic noise is valid when $\eta_j \neq 1$. Furthermore, we will assume $\eta_j \neq 0$. That is, we consider the case $\eta_j \in (0, 1)$, which is the case of a realistic detector of our interest.

In the next subsection, we derive the POVM corresponding to this detector model. We then choose to consider a simplified scenario where these two homodyne detectors are identical for the purpose of illustration and the ease of numerical calculation. That is, we will later assume they both have the same detector efficiency $\eta_1 = \eta_2 =: \eta_d$ and the same electronic noise $\nu_1 = \nu_2 =: \nu_{\text{el}}$.

6.2.2 POVM description

We use the Wigner function formulation to find the POVM $\{G_y : y \in \mathbb{C}\}$ corresponding to this noisy heterodyne detector model. When two homodyne detectors give two real numbers q_s and p_s for q and p quadrature measurements, we label the outcome as $y = q_s + ip_s$. By considering $\text{Tr}(\rho G_y)$ for an arbitrary input density operator ρ to the noisy heterodyne detector, we are able

[9] The electronic noise ν_j is the thermal noise added by the detection electronics. In the quantum mechanical model of the detector shown in each dashed box of Figure 6.2 , the electronic noise is modeled by an ancillary thermal state added to the second input port of the beam splitter that models the detector efficiency. Since the value of electronic noise is unaffected by the detector efficiency, to simulate the desired amount of noise before this beam splitter, one then needs to scale it by the reflectance of the beam splitter which is $1 - \eta_j$. As the variance of a thermal state with a mean photon number $\bar{n}$ is $(1 + 2\bar{n})N_0$, one can easily see that the mean photon number of this ancillary thermal state is $\bar{n}_j = \frac{\nu_j}{2(1-\eta_j)}$.

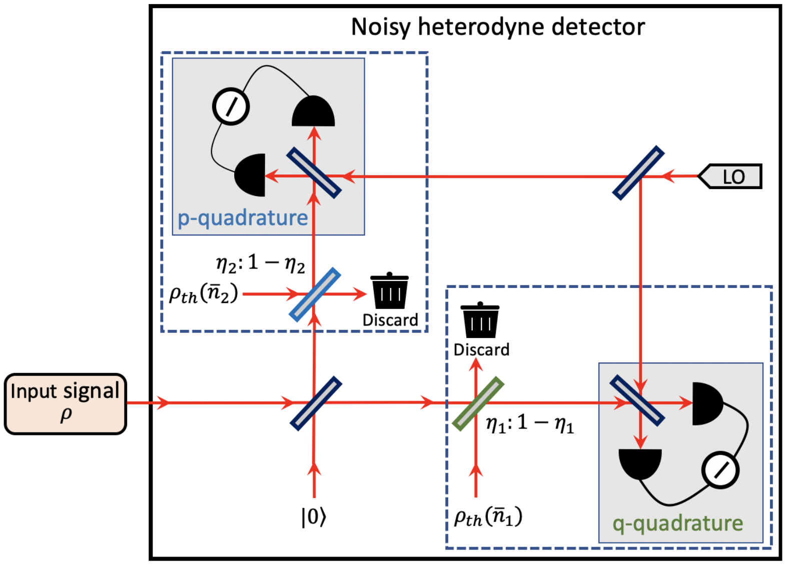


Figure 6.2: A physical model for a noisy heterodyne detector. The homodyne detector for the q quadrature measurement has detector efficiency η_1 and electronic noise ν_1. The homodyne detector for the p quadrature measurement has detector efficiency η_2 and electronic noise ν_2. The notation $\rho_{\text{th}}(\bar{n})$ stands for a thermal state with a mean photon number $\bar{n}$. In particular, $\bar{n}_1 = \frac{\nu_1}{2(1-\eta_1)}$ and $\bar{n}_2 = \frac{\nu_2}{2(1-\eta_2)}$ (see main text for more explanations). Beam splitters are 50:50 unless specified otherwise. Each homodyne detector inside a gray box is ideal. Each dashed box encloses the physical model for a noisy homodyne detector. LO stands for local oscillator.

to find the Wigner function W_{G_y} of the POVM element G_y as

$$\begin{aligned} W_{G_y}(\gamma) =& \frac{1}{\sqrt{\eta_1\eta_2}\pi}\frac{2}{\pi}\frac{1}{\sqrt{1+\frac{2(1-\eta_1+\nu_1)}{\eta_1}}}\frac{1}{\sqrt{1+\frac{2(1-\eta_2+\nu_2)}{\eta_2}}} \\ &\times \exp\left(\frac{-2[\text{Re}(\gamma)-\frac{1}{\sqrt{\eta_1}}\text{Re}(y)]^2}{1+\frac{2(1-\eta_1+\nu_1)}{\eta_1}}\right)\exp\left(\frac{-2[\text{Im}(\gamma)-\frac{1}{\sqrt{\eta_2}}\text{Im}(y)]^2}{1+\frac{2(1-\eta_2+\nu_2)}{\eta_2}}\right). \end{aligned} \tag{6.3}$$

By comparing this Wigner function with that of a displaced squeezed thermal state [see Eq. (2.81)], we can identify that the POVM element G_y is a projection onto a displaced squeezed thermal state up to a prefactor $\frac{1}{\sqrt{\eta_1\eta_2}\pi}$. We give a full derivation of this Wigner function and the explicit parameters for displacement, squeezing and thermal state mean photon number in terms of detector parameters η_1, η_2, ν_1 and ν_2 in Appendix C.2.3.

We restrict our discussion to a simpler scenario where we assume both homodyne detectors have the same imperfection for the ease of numerical calculation and for the illustration purpose. We discuss how to perform the calculation in the general case in Appendix C.2. In this simple case, we set $\eta_1 = \eta_2 = \eta_d$ and $\nu_1 = \nu_2 = \nu_{\text{el}}$ in Eq. (6.3). This equation is simplified to be

$$W_{G_y}(\gamma) = \frac{1}{\eta_d \pi} \frac{2}{\pi} \frac{1}{1 + \frac{2(1-\eta_d+\nu_{\text{el}})}{\eta_d}} \exp\left(\frac{-2\left|\gamma - \frac{y}{\sqrt{\eta_d}}\right|^2}{1 + \frac{2(1-\eta_d+\nu_{\text{el}})}{\eta_d}} \right). \tag{6.4}$$

One can observe that it is the Wigner function of a displaced thermal state [see Eq. (2.79)] apart from the prefactor $1/(\eta_d \pi)$. Therefore, the POVM element G_y in this case is a scaled projection onto a displaced thermal state. More precisely,

$$G_y = \frac{1}{\eta_d \pi} \hat{D}(\frac{y}{\sqrt{\eta_d}}) \rho_{\text{th}}(\frac{1-\eta_d+\nu_{\text{el}}}{\eta_d}) \hat{D}^\dagger(\frac{y}{\sqrt{\eta_d}}), \tag{6.5}$$

where $\hat{D}(\frac{y}{\sqrt{\eta_d}})$ is the displacement operator with the amount of displacement $y/\sqrt{\eta_d}$ and $\rho_{\text{th}}(\frac{1-\eta_d+\nu_{\text{el}}}{\eta_d})$ is a thermal state with the mean photon number $(1-\eta_d+\nu_{\text{el}})/\eta_d$, which can be expressed in the photon-number basis as

$$\rho_{\text{th}}(\bar{n}) = \sum_{n=0}^{\infty} \frac{\bar{n}^n}{(1+\bar{n})^{n+1}} |n\rangle\langle n| . \tag{6.6}$$

Later in Section 6.3.2, we need to express operators defined in terms of POVM elements G_y's in the photon-number basis for the numerical key rate calculation. Analytical expressions of matrix elements $\langle m| G_y |n\rangle$ are known in the literature [199] and shown in Appendix C.2.

Let us end this section with a few remarks about the simplification considered here. Firstly, as we later define operators involving integrals of POVM elements G_y's and need to find their matrix representations in the photon-number basis for the numerical key rate calculation, we are able to find efficiently computable analytical expressions for these operators under this simplification. Without this simplification, one may need to perform some numerical integration. We emphasize that the principles presented in this work also hold for the general case and we choose to present results based on this simplified case for the ease of calculation. Secondly, with this simplification, our detector model is then optically equivalent to the detector model used in other works [195, 197]. Thirdly, if two homodyne detectors in the heterodyne detection scheme do not have the same imperfection, one can instead use the POVM in the general case by following the procedure outlined in Appendix C.2.3 despite being more numerically challenging.

We remark on the generality of our method for treating trusted detector noises. If a different physical model of a detector is adopted (which needs to be verified experimentally), we expect that a similar method as described here can be used to find a correct POVM description for the given physical model and then this POVM can be used in the security analysis.

6.3 Key rate optimization

The first step of the security proof is to apply the source-replacement scheme (see Section 2.5.2) to recast the protocol into its equivalent entanglement-based protocol. In the entanglement-based version, Alice applies a POVM $\{M_x^A\}$.

We discuss the key rate calculation problems in the ideal (or untrusted) detector and trusted detector noise scenarios in this section.

6.3.1 Ideal detector scenario

We can obtain expectation values of the first and second moments of the quadrature operators $\langle\hat{q}\rangle$, $\langle\hat{q}^2\rangle$, $\langle\hat{p}\rangle$, and $\langle\hat{p}^2\rangle$. We can calculate the mean photon number of each conditional state ρ_B^x from the $\hat{n} = \frac{1}{2}(\hat{q}^2 + \hat{p}^2 - 1) = \hat{a}^\dagger\hat{a}$. In addition to $\hat{n}$, we define an operator $\hat{d} = \hat{q}^2 - \hat{p}^2 = \hat{a}^2 + (\hat{a}^\dagger)^2$ to utilize the second moment observations $\langle\hat{q}^2\rangle$ and $\langle\hat{p}^2\rangle$ to constrain ρ_{AB}.

For an input state ρ, heterodyne measurements give us the Husimi Q function $Q(\gamma) = \frac{1}{\pi}\langle\gamma|\,\rho\,|\gamma\rangle = \mathrm{Tr}(\rho E_\gamma)$, where $\{E_\gamma = \frac{1}{\pi}|\gamma\rangle\langle\gamma| : \gamma \in \mathbb{C}\}$ is the POVM description of heterodyne detection. From the Q function, we can obtain values of $\langle\hat{q}\rangle$, $\langle\hat{p}\rangle$, $\langle\hat{n}\rangle$, and $\langle\hat{d}\rangle$, whose operators are functions of $\hat{a}$ and $\hat{a}^\dagger$, by the following equation [200]:

$$\mathrm{Tr}\Big[\rho\hat{f}(\hat{a},\hat{a}^\dagger)\Big] = \langle\hat{f}^{(A)}(\hat{a},\hat{a}^\dagger)\rangle := \int d^2\gamma Q(\gamma) f^{(A)}(\gamma), \tag{6.7}$$

where $\hat{f}^{(A)}(\hat{a},\hat{a}^\dagger)$ is the antinormally ordered operator of an operator $\hat{f}$ written in terms of $\hat{a}$ and $\hat{a}^\dagger$, $f^{(A)}(\gamma)$ is the corresponding expression by replacing $\hat{a}$ by γ and $\hat{a}^\dagger$ by γ^*, and $d^2\gamma = d\,\mathrm{Re}(\gamma) d\,\mathrm{Im}(\gamma)$.

To write out the Kraus operator for the postprocessing map $\mathcal{G}$ including postselection, we define *region operators* that tell us in which region in Figure 6.1 Bob's measurement outcome lies. We express them using the polar coordinate for the integration as

$$\begin{aligned}
R_0 &= \frac{1}{\pi}\int_{\Delta_a}^{\infty}\int_{-\frac{\pi}{4}+\Delta_p}^{\frac{\pi}{4}-\Delta_p} \gamma\left|\gamma e^{i\theta}\middle\rangle\middle\langle\gamma e^{i\theta}\right| \, d\theta\; d\gamma, \\
R_1 &= \frac{1}{\pi}\int_{\Delta_a}^{\infty}\int_{\frac{\pi}{4}+\Delta_p}^{\frac{3\pi}{4}-\Delta_p} \gamma\left|\gamma e^{i\theta}\middle\rangle\middle\langle\gamma e^{i\theta}\right| \, d\theta\; d\gamma, \\
R_2 &= \frac{1}{\pi}\int_{\Delta_a}^{\infty}\int_{\frac{3\pi}{4}+\Delta_p}^{\frac{5\pi}{4}-\Delta_p} \gamma\left|\gamma e^{i\theta}\middle\rangle\middle\langle\gamma e^{i\theta}\right| \, d\theta\; d\gamma, \\
R_3 &= \frac{1}{\pi}\int_{\Delta_a}^{\infty}\int_{\frac{5\pi}{4}+\Delta_p}^{\frac{7\pi}{4}-\Delta_p} \gamma\left|\gamma e^{i\theta}\middle\rangle\middle\langle\gamma e^{i\theta}\right| \, d\theta\; d\gamma.
\end{aligned} \tag{6.8}$$

The area of integration for each operator corresponds to the relevant region shown in Figure 6.1. Again, Δ_a and Δ_p are parameters related to postselection.

In this case, the postprocessing map $\mathcal{G}(\sigma) = K\sigma K^\dagger$ is given by a single Kraus operator

$$K = \sum_{z=0}^{3} |z\rangle_R \otimes \mathbb{1}_A \otimes (\sqrt{R_z})_B. \tag{6.9}$$

The pinching quantum channel $\mathcal{Z}$ is given by the projections $Z_j = |j\rangle\langle j|_R \otimes \mathbb{1}_{AB}$ for $j = 0, 1, 2, 3$, that is, for a valid input state σ:

$$\mathcal{Z}(\sigma) = \sum_{j=0}^{3}(|j\rangle\langle j|_R \otimes \mathbb{1}_{AB})\sigma(|j\rangle\langle j|_R \otimes \mathbb{1}_{AB}). \tag{6.10}$$

The exact form of the key rate optimization problem for this protocol is

$$\begin{aligned}
\text{minimize} \quad & D\big(\mathcal{G}(\rho_{AB})||\mathcal{Z}[\mathcal{G}(\rho_{AB})]\big) \\
\text{subject to} \quad & \\
& \mathrm{Tr}[\rho_{AB}(|x\rangle\langle x|_A \otimes \hat{q})] = p_x\langle\hat{q}\rangle_x, \\
& \mathrm{Tr}[\rho_{AB}(|x\rangle\langle x|_A \otimes \hat{p})] = p_x\langle\hat{p}\rangle_x, \\
& \mathrm{Tr}[\rho_{AB}(|x\rangle\langle x|_A \otimes \hat{n})] = p_x\langle\hat{n}\rangle_x, \\
& \mathrm{Tr}\Big[\rho_{AB}(|x\rangle\langle x|_A \otimes \hat{d})\Big] = p_x\langle\hat{d}\rangle_x, \\
& \mathrm{Tr}[\rho_{AB}] = 1, \\
& \mathrm{Tr}_B[\rho_{AB}] = \sum_{i,j=0}^{3} \sqrt{p_i p_j}\, \langle\varphi_j|\varphi_i\rangle\, |i\rangle\langle j|_A\,, \\
& \rho_{AB} \geq 0,
\end{aligned} \tag{6.11}$$

where $x \in \{0, 1, 2, 3\}$ and $\langle\hat{q}\rangle_x, \langle\hat{p}\rangle_x, \langle\hat{n}\rangle_x$, and $\langle\hat{d}\rangle_x$ denote the corresponding expectation values of operators $\hat{q}, \hat{p}, \hat{n}$, and $\hat{d}$ for the conditional state ρ_B^x, respectively.

We remark that we make an additional simplification for the Kraus operator K. Unlike the general discussion in Section 4.1.2 or in Ref. [144], we do not introduce the registers $\widetilde{A}, \widetilde{B}, \overline{A}$, and $\overline{B}$ in the postprocessing map $\mathcal{G}$ for this protocol. The aim of such a simplification is to reduce the total dimension of the quantum states in the key rate optimization without affecting the calculated key rates. We provide a detailed analysis in Appendix A to explain why such a simplification can be made. Here, we discuss the ideas behind this simplification.

i) The quantum register $\overline{A}$ is Alice's private register that stores her measurement outcome after she performs her POVM $\{M_x^A\}$ on register A in a coherent fashion. Since Eve has no access to register $\overline{A}$, Alice can choose to first perform a coarse-grained measurement that introduces only the announcement register $\widetilde{A}$ and then perform a refined measurement conditioned on the announcements, which is described by a local isometry. Moreover, in the reverse reconciliation scheme, since the key map isometry V does not depend on Alice's measurement outcome, the isometry for the refined measurement commutes with both the key map isometry V and the pinching map $\mathcal{Z}$. As our objective function is invariant under this type of local isometries, we can choose not to apply this isometry and, thus, we do not introduce register $\overline{A}$.

ii) In the announcement step, Alice and Bob each announce whether a given round is kept for the key generation. Then the sifting process keeps only one announcement outcome, that is, when they both decide to keep the round. So, both classical registers $\widetilde{A}$ and $\widetilde{B}$ after applying the sifting projection Π are effectively one dimensional. We then use another property of the quantum relative entropy regarding quantum-classical states to show that the calculated key rates remain the same if we omit registers $\widetilde{A}$ and $\widetilde{B}$.

iii) The key map in this protocol uses only the coarse-grained information about Bob's measurement outcomes, that is, in which interval Bob's measurement outcome lies. As with the previous discussion about register $\overline{A}$, we can view Bob's measurement in two steps. At the first step, Bob performs a coarse-grained measurement in a coherent fashion to store the desired coarse-grained outcomes in register $\overline{B}$. At the second step, Bob performs a refined measurement conditioned on the coarse-grained information to update register $\overline{B}$, which is described by a local isometry (denoted by W). Since the key map uses only the coarse-grained information, the key map isometry V effectively needs to first undo the isometry W. So, we can choose not to perform the isometry W and let the key map isometry V use the coarse-grained information directly. The calculated key rates remain the same after we ignore the isometry W. In this case, the key map isometry V simply copies register $\overline{B}$ to register R in the standard basis. Thus, we combine these two registers and retain the name of R. The calculated key rates are unaffected, because copying register $\overline{B}$ to register R in the standard basis is done by a local isometry, which we can omit.

Finally, we explain how we derive the Kraus operator shown in Eq. (6.9). First, since we consider the reverse reconciliation schemes, we can omit the measurement outcome register $\overline{A}$. Second, since the key map of each protocol only uses the coarse-grained measurement outcomes, instead of using Bob's fine-grained POVM corresponding to heterodyne measurements, we use the coarse-grained POVM $(\{R_0, R_1, R_2, R_3, \mathbb{1} - \sum_{j=0}^{3} R_j\})$. Since the set $\mathbf{K}$ contains only one element, we are left with only one term in the summation of Eq. (A.8) after removing registers $\widetilde{A}$ and $\widetilde{B}$. Finally, the key map in this case is the identity map. Thus, we combine registers R and $\overline{B}$ and retain the name R for this combined register.

6.3.2 Revised optimization problem in the trusted detector noise scenario

We reconsider the key rate optimization problem in the untrusted detector noise scenario by rewriting region operators in Eq. (6.8) and observables in Eq. (6.11) in terms of the POVM of an ideal heterodyne detector $\{E_y\}$. In the case of ideal heterodyne detection, the POVM description of Bob's measurement $\{M_y^B\}$ is $M_y^B = E_y = \frac{1}{\pi} |y\rangle\langle y|$, the projection onto a coherent state $|y\rangle$. By writing $y = re^{i\theta}$ in the polar coordinate and integrating over the corresponding region $\mathcal{A}_j$, we obtain Eq. (6.8). If we rewrite Eq. (6.8) in terms of M_y^B, we see region operators R_j's are defined by

$$R_j = \int_{y\in\mathcal{A}_j} M_y^B d^2y, \tag{6.12}$$

where the region of integration $\mathcal{A}_j$ in the complex plane corresponds to the region with value j shown in Figure 6.1 and $d^2y = d\,\mathrm{Re}(y)d\,\mathrm{Im}(y)$.

In general, one may be interested in a quantity like $\int f(y,y^*)P(y)d^2y$ where $f(y,y^*)$ is a real-valued function on y and y^* such that the integral converges. Such a quantity can be described as the expectation value of an observable that is defined in the following way

$$\hat{O} = \int f(y,y^*)M_y^B d^2y \tag{6.13}$$

since

$$\begin{aligned}\mathrm{Tr}\Big[\rho\,\hat{O}\Big] &= \int d^2y\; f(y,y^*)\,\mathrm{Tr}\big(\rho M_y^B\big)\\ &= \int d^2y\; f(y,y^*)P(y).\end{aligned} \tag{6.14}$$

In other words, operators constructed in this way correspond to expectation values $\int f(y,y^*)P(y)d^2y$ obtained in an experiment.

In the previous subsection (also in Ref. [81]), we chose observables $\{\hat{O}\} = \{\hat{q},\hat{p},\hat{n},\hat{d}\}$ by using $M_y^B = E_y$ in Eq. (6.13) for the untrusted detector noise scenario. In the trusted detector noise scenario, we change to a new set of observables $\{\hat{q},\hat{p},\hat{n}+\hat{d}/2+\mathbb{1},\hat{n}-\hat{d}/2+\mathbb{1}\}$, which gives the same key rates as the old one since the last two observables in this new set are linear combinations of observables $\hat{n}$ and $\hat{d}$ as well as the identity operator. This new set of observables corresponds to the set of $\{f(y,y^*)\} = \{\sqrt{2}\,\mathrm{Re}(y), \sqrt{2}\,\mathrm{Im}(y), 2\,\mathrm{Re}(y)^2, 2\,\mathrm{Im}(y)^2\}$[10]. The sole purpose of this change compared with the previous subsection is to make the data postprocessing in an agreement with the typical classical postprocessing in an experiment. That is, in an experiment, when a heterodyne detection gives two real numbers q_s and p_s which we set $\mathrm{Re}(y) = q_s$ and $\mathrm{Im}(y) = p_s$,

[10] Due to our definition of quadrature operators, we include the factor $\sqrt{2}$ so that we can simply enter values reported in an experiment using shot noise units as expectation values of corresponding observables.

one usually computes variances of $\mathrm{Re}(y)$ and $\mathrm{Im}(y)$ by computing the expectation values of $\mathrm{Re}(y)^2$ and $\mathrm{Im}(y)^2$ in addition to expectation values of $\mathrm{Re}(y)$ and $\mathrm{Im}(y)$.

In the trusted detector noise scenario, we need to substitute M_y^B in Eqs. (6.12) and (6.13) by G_y. To distinguish operators defined in this way from first- and second-moment of quadrature operators $\hat{q}$ and $\hat{p}$, we shall call first-moment observables as $\hat{F}_Q$ and $\hat{F}_P$ and second-moment observables as $\hat{S}_Q$ and $\hat{S}_P$. More explicitly, they are defined as

$$\begin{aligned}\hat{F}_Q &= \int \frac{y+y^*}{\sqrt{2}} G_y d^2y,\\ \hat{F}_P &= \int \frac{i(y^*-y)}{\sqrt{2}} G_y d^2y,\\ \hat{S}_Q &= \int (\frac{y+y^*}{\sqrt{2}})^2 G_y d^2y,\\ \hat{S}_P &= \int [\frac{i(y^*-y)}{\sqrt{2}}]^2 G_y d^2y.\end{aligned} \tag{6.15}$$

Then the revised key rate optimization problem becomes

$$\begin{aligned} \text{minimize} \quad & D\big(\mathcal{G}(\rho_{AB})||\mathcal{Z}[\mathcal{G}(\rho_{AB})]\big)\\ \text{subject to} \quad & \\ & \mathrm{Tr}\Big[\rho_{AB}(|x\rangle\langle x|_A \otimes \hat{F}_Q)\Big] = p_x\langle \hat{F}_Q\rangle_x,\\ & \mathrm{Tr}\Big[\rho_{AB}(|x\rangle\langle x|_A \otimes \hat{F}_P)\Big] = p_x\langle \hat{F}_P\rangle_x,\\ & \mathrm{Tr}\Big[\rho_{AB}(|x\rangle\langle x|_A \otimes \hat{S}_Q)\Big] = p_x\langle \hat{S}_Q\rangle_x,\\ & \mathrm{Tr}\Big[\rho_{AB}(|x\rangle\langle x|_A \otimes \hat{S}_P)\Big] = p_x\langle \hat{S}_P\rangle_x,\\ & \mathrm{Tr}[\rho_{AB}] = 1,\\ & \mathrm{Tr}_B[\rho_{AB}] = \sum_{i,j=0}^{3} \sqrt{p_i p_j}\, \langle \alpha_j|\alpha_i\rangle\, |i\rangle\langle j|_A\,,\\ & \rho_{AB} \geq 0,\end{aligned} \tag{6.16}$$

where the index x runs over the set $\{0,1,2,3\}$ and the Kraus operator for the postprocessing map $\mathcal{G}$ has the same form as in Eq. (6.9) but now with the region operators defined in terms of G_y's in Eq. (6.12).

In Appendix C.2, we discuss how to represent these operators in the photon-number basis. Combining with the photon-number cutoff assumption (i.e. $\rho_{AB} = (\mathbb{1}_A \otimes \Pi_N)\rho_{AB}(\mathbb{1}_A \otimes \Pi_N)$, where N is the cutoff photon number and Π_N is the projection onto the subspace spanned by the photon-number states from 0 to N photons), we can directly solve this key rate optimization problem in Eq. (6.16) numerically.

6.4 Simulation method

To understand how the protocols behave in a realistic scenario, we simulate the quantum channel as a realistic physical channel in the absence of Eve. Such a channel in the context of optical fiber communication can be described by a phase-invariant Gaussian channel with transmittance η and excess noise ξ which is defined as

$$\xi = \frac{(\Delta q_{\text{obs}})^2}{(\Delta q_{\text{vac}})^2} - 1, \tag{6.17}$$

where $(\Delta q_{\text{vac}})^2$ is the variance of q quadrature for the vacuum state and $(\Delta q_{\text{obs}})^2$ is the variance of q quadrature observed for the signal state. Here we consider the case where both q and p quadratures have the same variance. With our definition of quadrature operators, $(\Delta q_{\text{vac}})^2 = \frac{1}{2}$. In the literature, the value of excess noise is usually reported in a couple of different ways, depending on who makes the observation of $(\Delta q_{\text{obs}})^2$. To avoid possible confusion when discussing the value of excess noise, we clarify these definitions. We use ξ to mean the excess noise in the case where Alice measures $(\Delta q_{\text{obs}})^2$ at the output of her lab and use δ in the case where Bob measures $(\Delta q_{\text{obs}})^2$ for the received signal state.

A natural way to simulate this phase-invariant Gaussian channel is that, when Alice prepares a coherent state $|\alpha\rangle$ and sends to Bob via this channel, the output state from the channel becomes a displaced thermal state centered at$\sqrt{\eta}\alpha$ with the variance $\frac{1}{2}(1+\delta)$ for each quadrature. An alternative but equivalent way is that, when Alice wants to prepare a coherent state $|\alpha\rangle$, the state after preparation becomes a displaced thermal state centered at α with the variance $\frac{1}{2}(1+\xi)$ for each quadrature at the output of her lab. Then, the state is transmitted via a pure-loss channel, and the final output state that reaches Bob's lab is a displaced thermal state centered at $\sqrt{\eta}\alpha$ with the variance $\frac{1}{2}(1+\eta\xi)$ for each quadrature. Therefore, we see that, for this physical channel, $\delta = \eta\xi$. In this work, we use the definition of ξ when we discuss the value of excess noise. Readers should be able to translate between these two definitions by the relation $\delta = \eta\xi$.

Given a displaced thermal state centered at $\sqrt{\eta}\alpha$ with the variance $\frac{1}{2}(1+\eta\xi)$ for each quadrature, we can then calculate our simulated values for $\langle\hat{q}\rangle$, $\langle\hat{p}\rangle$, $\langle\hat{n}\rangle$, and $\langle\hat{d}\rangle$ (by either using quasiprobability distribution like the Wigner function or Q function of the final state or expanding the final state in the photon-number basis). These values can then be supplied to the optimization problem in Eq. (6.11), which, in turn, can be solved numerically.

6.5 Photon-number cutoff assumption

The key rate optimization problem in Eq. (4.2) involves optimizing over all possible bipartite states ρ_{AB} in the feasible set $\mathbf{S}$. The number of free variables depends on the size of ρ_{AB}. In order to numerically perform the optimization by computer optimization packages, we can deal only with finite-dimensional ρ_{AB}. In our optimization problem, as we can see from the source-replacement scheme, the dimension of Alice's system A is determined by the number of different signal states that she prepares. For the QPSK scheme, the dimension of register A is four. However, since each state that Bob receives is an optical mode and, in principle, can be manipulated by Eve, Bob' state lives in an infinite-dimensional Hilbert space $\mathcal{H}_B$. A basis for this Hilbert space is the photon-number states $\{|n\rangle : n \in \mathbb{N}\}$. We immediately see that Bob's POVM elements are infinite-dimensional operators and ρ_{AB} is also infinite dimensional. For DVQKD, one method to reduce the dimension of the system is to apply a squashing model [90, 92, 156] for the protocol to obtain a lower-dimensional representation of his POVM. This reduction is possible for many DVQKD protocols, since one can explicitly formulate the squashing model. However, it is not clear how one can formulate a squashing model for CV systems. Instead, we have to impose an additional assumption in this work in order to perform the numerical optimization. This additional assumption is what we call the photon-number cutoff assumption. We impose the assumption that Bob's system lives in the Hilbert space $\mathcal{H}_B = \text{span}\{|0\rangle, |1\rangle, \dots, |N_c\rangle\}$ for some cutoff photon number N_c. Namely, if we define $\Pi_{N_c} = \sum_{n=0}^{N_c} |n\rangle\langle n|$ with a suitable choice of photon-number cutoff parameter N_c on Bob's system, we assume $\rho = \Pi_{N_c}\rho\Pi_{N_c}$ for the state ρ under consideration. This assumption allows us to truncate the infinite-dimensional Hilbert space. If N_c is chosen to be large enough, this assumption is a reasonable working assumption based on the following observations.

(i) Bob can obtain the mean photon number $n_x := \text{Tr}(\rho_B^x \hat{n})$ of each conditional state ρ_B^x via homodyne or heterodyne measurements, where $\hat{n}$ denotes the number operator.

(ii) Since n_x is known, we can pick $N_c \in \mathbb{N}$ such that N_c is much larger than n_x for each $x \in \{0, 1, 2, 3\}$. For such a choice of N_c, the probability of finding the state to have a photon number $n \leq N_c$ is close to 1. This probability suggests that the contribution from $n > N_c$ photon subspace becomes negligible. Similarly, the off-diagonal blocks $(\mathbb{1} - \Pi_{N_c})\rho\Pi_{N_c}$ and $\Pi_{N_c}\rho(\mathbb{1} - \Pi_{N_c})$ also have vanishing contributions.

(iii) We can increase N_c to have a numerical verification that the key rate is unchanged after we choose a large enough N_c.

This photon-number cutoff assumption renders our numerical optimization of the key rate problem feasible. Although this assumption sounds numerically reasonable as our numerical verification suggests, we emphasize that one has to deliver an exact analysis to remove this assumption for a watertight security proof. In this sense, our proof is restricted. Nevertheless, we expect this working assumption to have minimal effects on the key rates of these protocols when the cutoff is chosen suitably, and this expectation is confirmed by the results in our recent work [85]. The choice of the cutoff value depends on the observed statistics as mentioned in (i) above. For our simulations in the ideal detector scenario presented in this thesis, we find that choosing $N_c \geq 10$ can produce numerically stable results. For our simulations in the trusted (and untrusted) detector scenario, we choose to use a slightly larger cutoff due to a larger mean photon number of the state received and we use $N_c \geq 14$.

We now provide some insights for removing this assumption and also for extending our current analysis to include finite-size effects and general attacks. To remove the photon-number cutoff assumption, one needs to combine our numerical optimization approach with some appropriate analytical tools. One possible approach is to develop a CV version of the squashing model. If such a squashing model exists, the key rate optimization problem then becomes a finite-dimensional problem even without the photon-number cutoff assumption. Since the effective dimension is finite, it might also be possible to apply existing tools (which are valid for finite-dimensional systems) such as the quantum de Finetti representation theorem [201] or the postselection technique [112] to obtain the composable security [89] of the protocol against general attacks in the finite-size regime.

Another possible method for removing the photon-number cutoff assumption is to adopt a similar idea used in Ref. [185], that is, using the entropy continuity bounds [202] to provide a tight error analysis of the key rate due to the truncation of Bob's Hilbert space. After one can tightly bound the trace distance between the optimal state in the truncated subspace and the optimal state in the original infinite-dimensional space, the continuity bound for the relevant entropic quantities allows us to obtain a small correction term due to the photon-number cutoff. Then, one obtains a full security proof against collective attacks in the asymptotic limit. In Section 6.10, we discuss the dimension reduction method introduced in Ref. [85]. Reference [85] has successfully applied this method to DMCVQKD to remove the photon-number cutoff assumption. To reach a full composable security proof along this path, one may first manage to include the finite-size effects with collective attacks and then apply appropriate tools similar to the quantum de Finetti representation theorem for CVQKD [111] to take the general attacks into consideration.

6.6 Evaluation of loss-only key rates

We discuss how to evaluate the Devetak-Winter formula in the loss-only scenario in the absence of postselection. When Alice sends $|\alpha_x\rangle_{A'}$ to Bob, in the absence of noise, Bob can verify that he receives a pure coherent state via homodyne or heterodyne detection. In the case of homodyne detection, Bob can verify that the received state is a minimum uncertainty state with the same variance for both quadratures, which implies it is a pure coherent state. In the case of heterodyne detection, Bob performs a tomography to verify that the received state is a pure coherent state. In particular, if Bob verifies his state to be an attenuated coherent state $\left|\sqrt{\eta}\alpha_x\right\rangle$, it is shown [194] that Eve's optimal attack is the generalized beam-splitting attack for this pure-loss channel. Thus, the state shared by Bob and Eve becomes $\left|\sqrt{\eta}\alpha_x\right\rangle_B \left|\sqrt{1-\eta}\alpha_x\right\rangle_E$ after this channel. Because of the product state structure of Bob and Eve's joint state, Bob's measurement outcome does not influence Eve's state. Therefore, conditioned on the value x of Alice's string $\mathbf{X}$ and the value z in Bob's raw key $\mathbf{Z}$, Eve's conditional state $|\epsilon_{x,z}\rangle$ is

$$|\epsilon_{x,z}\rangle = \left|\sqrt{1-\eta}\alpha_x\right\rangle := |\epsilon_x\rangle\,, \tag{6.18}$$

which is independent from z and, thus, we call it $|\epsilon_x\rangle$ for simplicity.

For this protocol, since $\alpha_x \in \{\alpha, i\alpha, -\alpha, -i\alpha\}$, Eve's conditional states $|\epsilon_x\rangle$ are $\left|\sqrt{1-\eta}\alpha\right\rangle$, $\left|i\sqrt{1-\eta}\alpha\right\rangle$, $\left|-\sqrt{1-\eta}\alpha\right\rangle$, and $\left|-i\sqrt{1-\eta}\alpha\right\rangle$, which span only a four-dimensional subspace. Therefore, we can find an orthonormal basis $\{|f_0\rangle, |f_1\rangle, |f_2\rangle, |f_3\rangle\}$ for this subspace similar to the basis $\{|e_0\rangle, |e_1\rangle\}$ and find a four-dimensional matrix presentation for each of Eve's conditional states (see Ref. [203] for an explicit expression). With a four-dimensional matrix representation of Eve's conditional states $|\epsilon_x\rangle$ in the basis $\{|f_0\rangle, |f_1\rangle, |f_2\rangle, |f_3\rangle\}$, it is also straightforward to directly evaluate the Devetak-Winter formula.

6.7 Simulation results in the ideal detector scenario

6.7.1 Loss-only scenario: Comparison to analytical results

We first present the results for the loss-only scenario, that is, $\xi = 0$. For this scenario, we can also obtain an analytical result to have a direct comparison with our numerical result. A direct evaluation of the Devetak-Winter formula is possible in this scenario since we can determine Eve's relevant conditional states (up to irrelevant unitaries). As shown in Ref. [194], in the loss-only case, we need only to consider the generalized beam-splitting attack. When Alice sends $|\alpha_x\rangle_{A'}$ to Bob, the state becomes $\left|\sqrt{\eta}\alpha_x\right\rangle_B \left|\sqrt{1-\eta}\alpha_x\right\rangle_E$ after the pure-loss channel. Eve's conditional states conditioned on Alice's string value x and Bob's raw key string value z effectively live in a four-dimensional subspace for this protocol. This result makes the direct analytical evaluation possible. The procedure of this analytical evaluation is discussed in Section 6.6. For the numerical key rate optimization, the loss-only scenario follows as a special case of the noisy scenario (using $\xi = 0$), which we discuss in later subsections.

A pure-loss channel is characterized by its transmittance $\eta = 10^{-(\alpha_{\text{att}}L/10)}$ for each distance L in kilometers with the attenuation coefficient α_{att}, which is 0.2 dB/km for the relevant communication fiber. One may take the quantum efficiency of realistic homodyne and heterodyne detectors into account. A simple but pessimistic way to deal with the detector efficiency is that the loss due to the imperfect detector is also attributed to Eve. In such a worse-case scenario, we can define the total transmittance as $\eta = \eta_d 10^{-0.02L}$, where η_d is the quantum efficiency of the detectors. If one defines an effective distance L_0 for the detector inefficiency, that is, $\eta_d = 10^{-0.02L_0}$, then L_0 is less than 13 km for practical homodyne and heterodyne detectors with the quantum efficiency $\geq 55\%$ [30]. For the ease of presentation and convenience of comparison with other works using different values of detector efficiency, we set $\eta_d = 1$ in this section unless noted otherwise. One may obtain the key rate value corresponding to a realistic value of efficiency by subtracting the effective distance L_0 from all relevant figures.

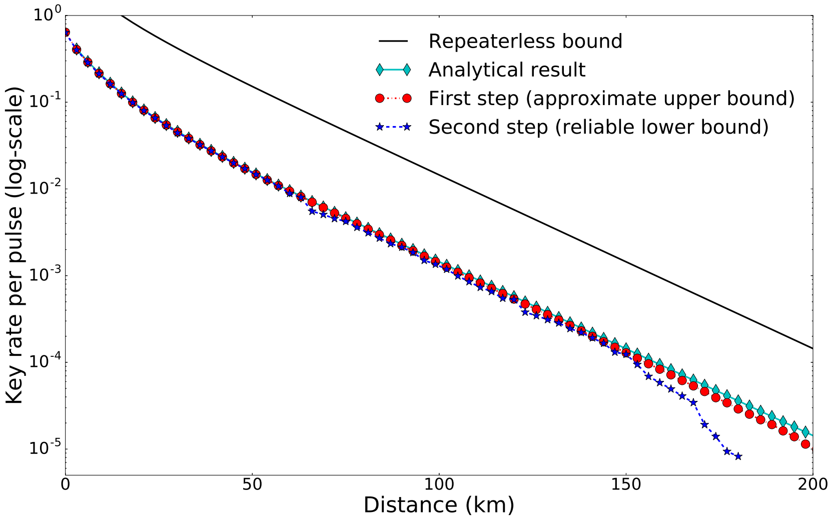


Figure 6.3: Secure key rates versus the transmission distance for the pure-loss channel to demonstrate the numerical behavior of our two-step key rate calculation procedure and to compare with the analytical results (direct evaluation of the Devetak-Winter formula). The transmittance is $\eta = 10^{-0.02L}$ for each distance L in kilometers, and the reconciliation efficiency is $\beta = 0.95$. The curve with circle markers is the approximate upper bound from the first step, and the curve with star markers is the reliable lower bound obtained from the second step. The curve with square markers is the analytical results presented in Section 6.6. The solid line with no markers is the repeaterless secret key capacity bound [65]. The coherent state amplitude α is optimized via a coarse-grained search in the interval [0.6, 0.95].

We plot the key rate of this protocol versus transmission distance in the loss-only scenario in Figure 6.3. We also plot both the numerical key rate calculation results and the key rate that can be obtained by a direct evaluation of the Devetak-Winter formula. Interestingly, we see that our numerical results are close to the analytical results for both protocols up to a distance around 120 km. Above 120 km, we notice that there is a visible gap between our approximate upper bound and the reliable lower bound, which indicates there is room for improvement on the numerical algorithm. We also notice that our first-step result is slightly lower than the analytical result as the feasible set **S** might be enlarged due to constraint violation and coarse-graining of data. We note that the new method developed in Ref. [87] allows us to reproduce the analytical result exactly with the numerical method. This result is shown in Figure 6.4.

We also include the repeaterless secret key capacity bound for the pure-loss channel [65] in Figure 6.3, that is, $R_\infty \leq -\log_2(1-\eta)$. With the reconciliation efficiency $\beta_{\text{EC}} = 0.95$, and the key rate for the this protocol is approximately 1/10 of the bound. Since Gaussian modulation schemes with the perfect reconciliation efficiency can reach 1/2 of the PLOB bound [65], we see that the performance of the QPSK modulation scheme is not far away from that of the Gaussian modulation schemes in the loss-only scenario.

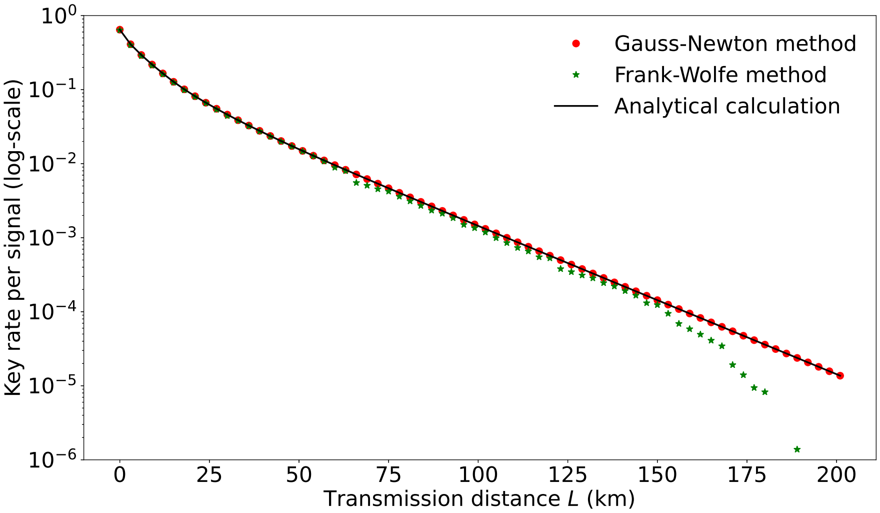


Figure 6.4: Comparison of key rate for DMCVQKD among the Gauss-Newton method [87], the Frank-Wolfe method (Algorithm 4.1) and analytical key rate for the loss-only case ($\xi = 0$).

6.7.2 Noisy scenario

Simulated statistics and error-correction cost

From the heterodyne measurements, for each conditional state ρ_B^x with $\alpha_x \in \{\alpha, i\alpha, -\alpha, -i\alpha\}$, we obtain a Q function Q_x as

$$Q_x(\gamma) = \frac{1}{\pi(1+\eta\xi/2)} \exp\left(-\frac{\left|\gamma - \sqrt{\eta}\alpha_x\right|^2}{1+\eta\xi/2}\right). \tag{6.19}$$

From each Q function, we can then calculate

$$\begin{aligned}
\langle\hat{q}\rangle_x &= \frac{1}{\sqrt{2}}\int(\gamma+\gamma^*)Q_x(\gamma)d^2\gamma = \sqrt{2\eta}\,\mathrm{Re}(\alpha_x),\\
\langle\hat{p}\rangle_x &= \frac{i}{\sqrt{2}}\int(\gamma^*-\gamma)Q_x(\gamma)d^2\gamma = \sqrt{2\eta}\,\mathrm{Im}(\alpha_x),\\
\langle\hat{n}\rangle_x &= \int(|\gamma|^2-1)Q_x(\gamma)d^2\gamma = \eta|\alpha_x|^2 + \frac{\eta\xi}{2},\\
\langle\hat{d}\rangle_x &= \int[\gamma^2+(\gamma^*)^2]Q_x(\gamma)d^2\gamma = \eta[\alpha_x^2+(\alpha_x^*)^2].
\end{aligned} \tag{6.20}$$

Note that those values are exactly the same as from the homodyne measurements, since we have the same state after the simulated quantum channel. We obtain those values here indirectly via the Q function.

We also present the procedure to calculate δ_{EC} for this protocol. For the error correction performed at the Shannon limit, we have $\delta_{\mathrm{EC}} = H(\mathbf{Z}|\mathbf{X}) = h(e)$, where $h(x)$ is the binary entropy function. To take into account the inefficiency of error correction, we first write $\delta_{\mathrm{EC}} = H(\mathbf{Z}|\mathbf{X}) = H(\mathbf{Z}) - I(\mathbf{X}\!:\!\mathbf{Z})$ in terms of $I(\mathbf{X}\!:\!\mathbf{Z})$ and then scale $I(\mathbf{X}\!:\!\mathbf{Z})$ to be $\beta_{\mathrm{EC}}I(\mathbf{X}\!:\!\mathbf{Z})$ where β_{EC} is the reconciliation efficiency whose value is usually reported in the CVQKD literature. Therefore,

$$\begin{aligned}
\delta_{\mathrm{EC}} &= H(\mathbf{Z}) - \beta_{\mathrm{EC}}I(\mathbf{X}\!:\!\mathbf{Z})\\
&= (1-\beta_{\mathrm{EC}})H(\mathbf{Z}) + \beta_{\mathrm{EC}}H(\mathbf{Z}|\mathbf{X})
\end{aligned} \tag{6.21}$$

We can numerically evaluate $H(\mathbf{Z}|\mathbf{X})$ via the probability distribution:

$$\begin{aligned}
P(z=j|x=k) &= \mathrm{Tr}\Big(R_j\rho_B^k\Big)\\
&= \int_{\Delta_a}^{\infty}\int_{\frac{2j-1}{4}\pi+\Delta_p}^{\frac{2j+1}{4}\pi-\Delta_p} \frac{\exp\left(-\frac{\left|\gamma e^{i\theta}-\sqrt{\eta}\alpha_k\right|^2}{1+\eta\xi/2}\right)}{\pi(1+\eta\xi/2)}\gamma\, d\theta\, d\gamma,
\end{aligned} \tag{6.22}$$

where $j, k \in \{0, 1, 2, 3\}$, R_j's are the region operators defined in Eq. (6.8), and the conditional state ρ_B^k is defined in Eq. (2.103). In the case of postselection, we renormalize this probability distribution by the probability of being postselected. Then, δ_{EC} can be calculated by Eq. (6.21) as we take into account that we have an alphabet of four symbols on both sides in the error-correction step. Throughout this thesis, we choose to use the error-correction efficiency $\beta_{\text{EC}} = 95\%$ for DMCVQKD unless stated otherwise[11].

Key rates

We start by investigating the optimal choice of coherent state amplitude α for this protocol. In Figure 6.5, we plot the key rates versus α for selected distances when the excess noise ξ is 0.01. Comparing to Figure 7.2, we see that the optimal choice of α for this protocol is around 0.7 for those selected distances, corresponding to a mean photon number around 0.49. We observe that the optimal value of α does not change significantly for a wide range of distances[12]. From the observation here, we limit our search for optimal choice of α in a restricted interval.

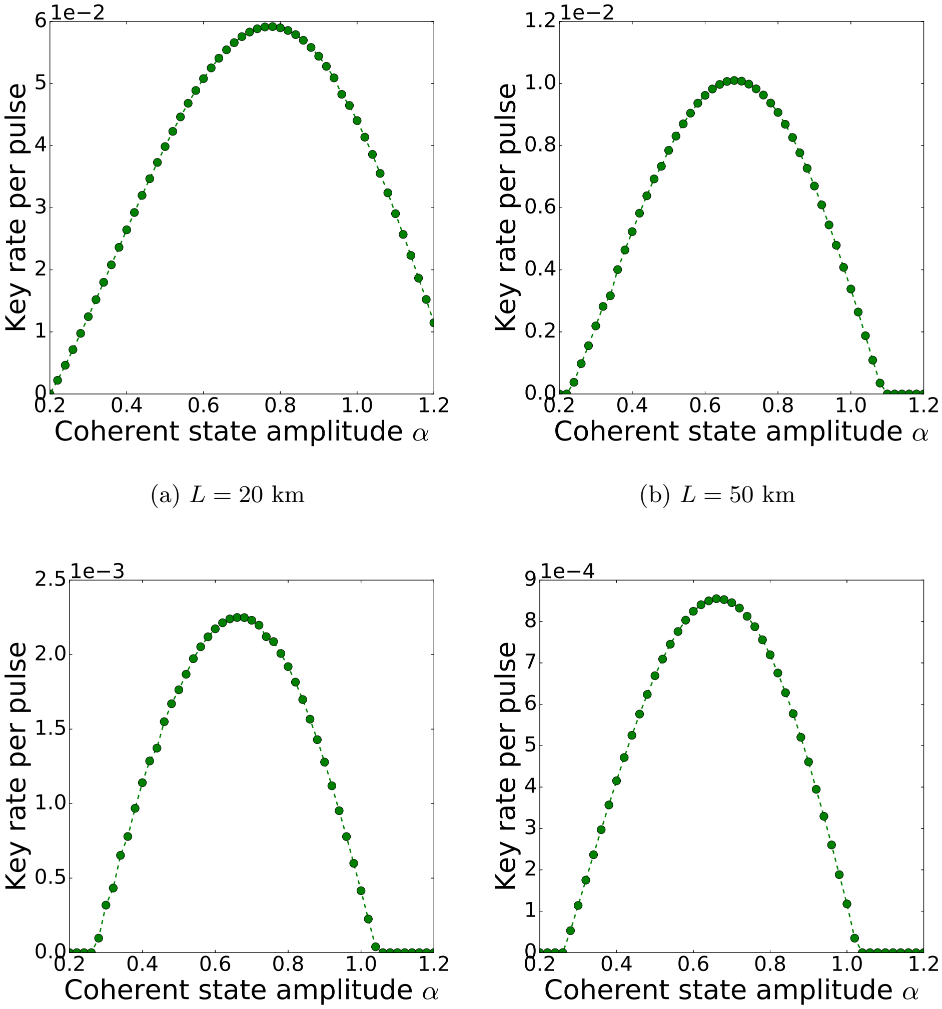
In Figure 6.6, we plot the secure key rate versus the transmission distance for different values of excess noise ξ. We optimize the coherent state amplitude α via a coarse-grained search over the interval $[0.6, 0.92]$. We observe that the heterodyne protocol can reach around 200 km even with an excess noise $\xi = 0.02$.

We then compare our results with an independent security analysis in Ref. [191] for a similar protocol. In addition to different proof methods, we differ from that protocol by how the error correction is done, which affects the calculation of the error-correction cost term δ_{EC}. In particular, our error-correction cost is higher because we discretize Bob's measurement results and consider only binary or quaternary error-correcting codes. In Ref. [191], the mutual information $I(\mathbf{X}\!:\!\mathbf{Z})$ is obtained by the channel capacity of the binary additive white Gaussian noise channel, which is approximated by the capacity of an additive white Gaussian noise channel

$$I(\mathbf{X}\!:\!\mathbf{Z}) \approx \log_2\Big(1 + \frac{2\eta\alpha^2}{2 + \eta\xi}\Big). \tag{6.23}$$

[11] This choice of value has been claimed by some experimentalists and theorists for the symmetric binary channel that is relevant for our situation. However, the author of this thesis is not aware of a specific construction of the error correction code that achieves this efficiency in the DMCVQKD parameter regime. The work [204] appears to give a similar value in the high error rate regime. In this sense, all key rates reported in this chapter and the next chapter are based on the assumption that this efficiency or a similar efficiency can be achieved in practice.

[12] The optimal amplitude α changes more significantly for small distances (less than 25 km). See [85, Figure 6].



(a) $L = 20$ km

(b) $L = 50$ km

(c) $L = 80$ km

(d) $L = 100$ km

Figure 6.5: Secure key rate for the QPSK protocol with heterodyne detection versus the coherent state amplitude α for selected choices of distances (a) $L = 20$ km, (b) $L = 50$ km, (c) $L = 80$ km, and (d) $L = 100$ km with the excess noise $\xi = 0.01$ and reconciliation efficiency $\beta_{\text{EC}} = 0.95$. Detectors are assumed to be ideal.

This result leads to a smaller value of δ_{EC} by the conversion formula in the first line of Eq. (6.21). In Figures 6.7 and 6.8, we plot the key rate results from both our work and Ref. [191] with a fixed choice of the coherent state amplitude $\alpha = 0.35$ for all distances plotted and with two different values of excess noise. As we can see, our security proof approach provides remarkably higher key rates compared with the approach with a reduction to the Gaussian optimality. Our security analysis shows that this protocol has a good tolerance on the excess noise and can extend to significantly longer distances. We emphasize that this choice of $\alpha = 0.35$ is not optimal for both works. While this value is closer to the optimal value found in Ref. [191], the optimal value of the coherent state amplitude found in our work is around 0.7 (for $\xi = 0.01$), as mentioned before. Thus, we also include two curves from Figure 6.6, where the coherent state amplitude α is optimized via a coarse-grained search in the interval $[0.6, 0.92]$ for comparisons.

As we can see from Figures 6.7 and 6.8, the key rate can be significantly improved after we optimize α. We summarize two factors that can boost the key rates. First, our security proof technique gives a tighter estimation of Eve's information compared with the reduction to the Gaussian optimality approach. Second, the key rate can be improved by using a slightly larger value of α than what is investigated in Ref. [191]. This regime of α is not explored in Ref. [191], because the reduction to the Gaussian optimality approach for discrete-modulation schemes gives tight key rates only in the limit of $\alpha \to 0$ and can give quite loose key rates for large values of α. In the same figure, we also compare our results for this protocol with a Gaussian-modulated CVQKD protocol using heterodyne detection [26], where the modulation variance is optimized for each distance. We observe that this quaternary modulation scheme has key rates comparable to the Gaussian modulation scheme.

6.8 Simulation results in the trusted detector noise scenario

6.8.1 Simulated statistics

From our simulation, the simulated state σ_B^x conditioned on the choice of x is a displaced thermal state whose Wigner function is

$$W_{\sigma_B^x}(\gamma) = \frac{1}{\pi}\frac{1}{\frac{1}{2}(1+\eta_t\xi)}\exp\left[-\frac{\left|\gamma-\sqrt{\eta_t}\alpha_x\right|^2}{\frac{1}{2}(1+\eta_t\xi)}\right]. \tag{6.24}$$

When Bob applies his heterodyne measurement described by the POVM $\{G_y\}$, the probability density function $P(y|x)$ for the measurement outcome y conditioned on Alice's choice x is

$$P(y|x) = \frac{1}{\pi(1+\frac{1}{2}\eta_d\eta_t\xi+\nu_{\text{el}})}\exp\left[-\frac{\left|y-\sqrt{\eta_d\eta_t}\alpha_x\right|^2}{1+\frac{1}{2}\eta_d\eta_t\xi+\nu_{\text{el}}}\right]. \tag{6.25}$$

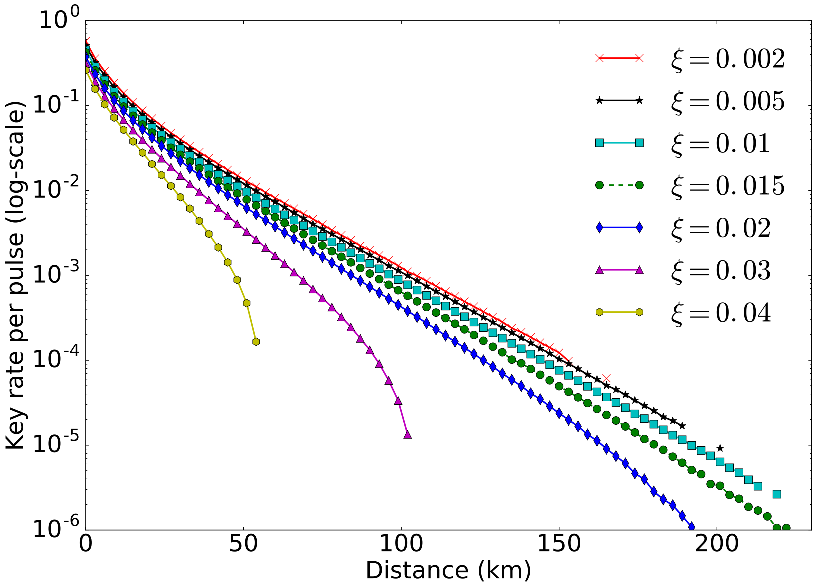


Figure 6.6: Secure key rate versus the transmission distance for the protocol with heterodyne detection for different values of the excess noise ξ, from top to bottom, $\xi = 0.002, 0.005, 0.01, 0.015, 0.02, 0.03, 0.04$. The coherent state amplitude is optimized via a coarse-grained search over the interval $[0.6, 0.92]$, the transmittance is $\eta = 10^{-0.02L}$ for each distance L in kilometers and the reconciliation efficiency is $\beta_{\text{EC}} = 0.95$.

The observables defined in Eq. (6.15) have the following expectation values from the simulation:

$$\begin{aligned}
\langle \hat{F}_Q \rangle_x &= \sqrt{2\eta_d \eta_t}\,\mathrm{Re}(\alpha_x),\\
\langle \hat{F}_P \rangle_x &= \sqrt{2\eta_d \eta_t}\,\mathrm{Im}(\alpha_x),\\
\langle \hat{S}_Q \rangle_x &= 2\eta_d \eta_t\,\mathrm{Re}(\alpha_x)^2 + 1 + \frac{1}{2}\eta_d \eta_t \xi + \nu_{\text{el}},\\
\langle \hat{S}_P \rangle_x &= 2\eta_d \eta_t\,\mathrm{Im}(\alpha_x)^2 + 1 + \frac{1}{2}\eta_d \eta_t \xi + \nu_{\text{el}}.
\end{aligned} \tag{6.26}$$

6.8.2 Comparison between trusted and untrusted detector noise scenarios

For this comparison, we supply the same set of simulated data from Eq. (6.26) to the optimization problem for the untrusted detector noise scenario in Eq. (6.11) and the one for the trusted detector

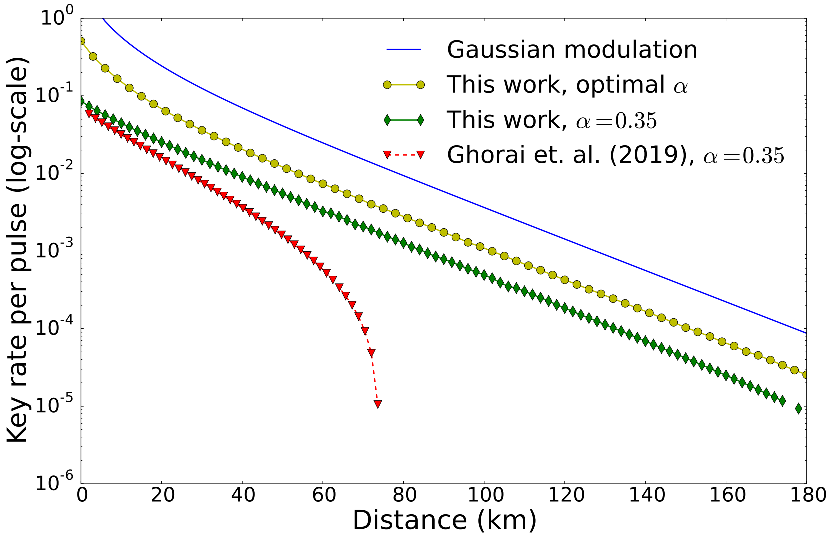


Figure 6.7: A comparison of key rates between our work and Ref. [191] for the heterodyne scheme with two different values of excess noise $\xi = 0.005$. The curve with triangle markers is from Ref. [191] with a fixed (not optimal) coherent state amplitude $\alpha = 0.35$, and the curve with diamond markers is from this work with the same value of α. The curve with circle markers is also from this work with an optimal value of α for each distance. The curve with no markers is the key rates of a Gaussian-modulated CVQKD protocol [26] with an optimal modulation variance for each distance. All curves use the reconciliation efficiency $\beta_{\text{EC}} = 0.95$.

noise scenario in Eq. (6.16). For simulation, we choose parameters $\eta_d = 0.719$, $\nu_{\text{el}} = 0.01$ from [32] for illustration. The result is shown in Figure 6.9.

As we can see from this figure, the key rate of the untrusted detector noise scenario drops quickly at a short distance less than 20 km even though the electronic noise is only 0.01 SNU, which is a low value compared to detectors used in many other CV experiments. On the other hand, the key rate in the trusted detector noise scenario extends to much longer distances, which exhibits a similar behavior as the results shown in Ref. [81] when the detector is treated as ideal. One explanation for this behavior is that in Ref. [81], we have observed that the key rate for the QPSK scheme drops quickly when the channel excess noise ξ is large. Since the value of ξ is reported at the input of the quantum channel while the value of ν_{el} is measured at Bob's side, to treat ν_{el} as a part of channel excess noise in the untrusted detector noise scenario, one needs to define the effective value of ξ to include the value of ν_{el}. For the effective value ξ_{eff}, the electronic noise ν_{el} needs to be scaled by a factor of $1/\eta_t$ (in addition to $1/\eta_d$), which is large for slightly long distances as η_t becomes small. As a result, the redefined value ξ_{eff} of ξ is quite large as shown

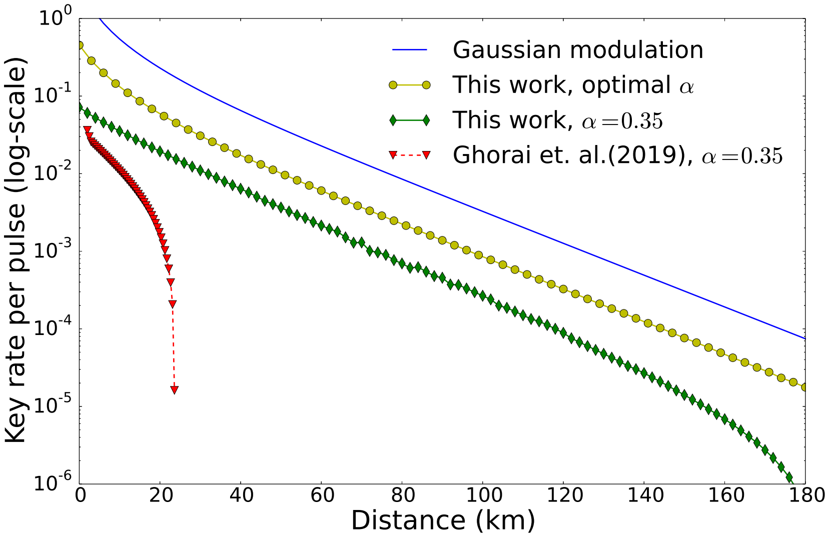


Figure 6.8: A comparison of key rates between our work and Ref. [191] for the heterodyne scheme, similar to Figure 6.7 but with an excess noise $\xi = 0.01$.

in Figure 6.9 and this behavior of key rate is then expected. By the observation made from this figure, it is not surprised that for a larger value of electronic noise, the key rate in the untrusted detector noise scenario would drop to zero at an even shorter distance.

6.8.3 Detector imperfection in the trusted detector noise scenario

To guide the experimental implementation of the QPSK scheme, we may be interested in the robustness of the protocol in the presence of detector inefficiency and electronic noise in the trusted detector noise scenario. For this purpose, we investigate the effects of different levels of detector efficiency and electronic noise on the key rate. For curves in Figures 6.10 and 6.11, our simulation uses the same channel model but different detector imperfections, that is, in Eq. (6.26), the same values of channel parameters η_t and ξ but different values of detector efficiency η_d and electronic noise ν_{el} (as specified in the captions) for different curves.

In Figure 6.10, we choose values of η_d and ν_{el} for a homodyne detector from two experiments [30, 32] and compare these results with the ideal detector. For the comparison, we optimize α via a coarse-grained search for each distance. We see that with a noisy heterodyne detector, the key rate drops moderately from the key rate of using an ideal detector. The amount of decrease is like a constant prefactor in the key rate. As the detector is noisier, the key rate becomes lower as expected.

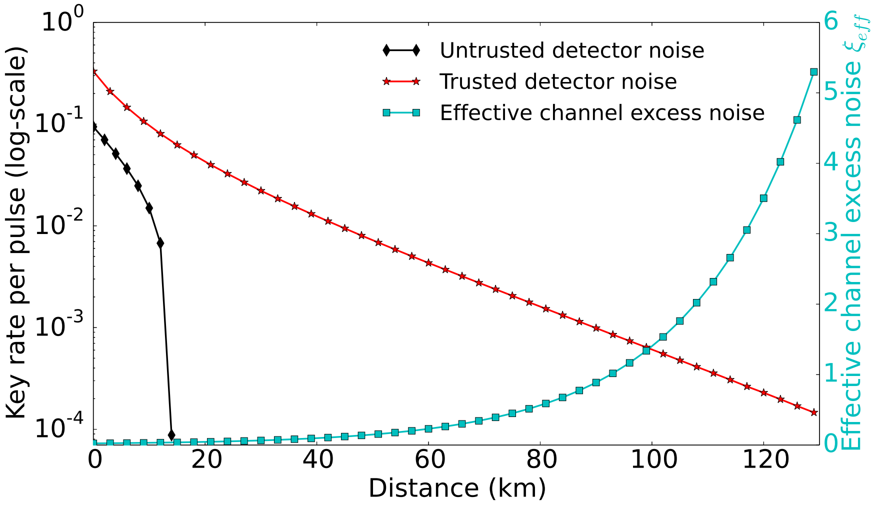


Figure 6.9: Secure key rate versus the transmission distance for untrusted detector noise (black diamonds) and trusted detector noise (red stars) scenarios. The excess noise is $\xi = 0.01$ at the input of the quantum channel. Parameters for detector are $\eta_d = 0.719$, $\nu_{\text{el}} = 0.01$[32]. The error correction efficiency is $\beta_{\text{EC}} = 0.95$. The coherent state amplitude is optimized via a coarse-grained search over the interval $[0.5, 0.9]$ with a step size of 0.05 and the channel transmittance is $\eta_t = 10^{-0.02L}$ for each distance L in kilometers. The effective channel excess noise in the untrusted detector scenario is shown with the y-axis on the right. At 20 km, the effective channel excess noise ξ_{eff} is roughly 0.045.

To show that different values of electronic noise have little impacts on the secure key rates in the trusted noise scenario, we compare key rates with two choices of the electronic noise value in Figure 6.11a while we fix the value of detector efficiency η_d to be 0.7. As the key rate difference is relatively small between the curve with $\nu_{\text{el}} = 0.05$ and that with $\nu_{\text{el}} = 0.08$, we also plot the difference of key rate (that is, the key rate with $\nu_{\text{el}} = 0.05$ minus the key rate with $\nu_{\text{el}} = 0.08$) in the same figure. (Note that the non-smoothness in the curve of difference is due to the coarse-grained search for the coherent state amplitude in the presence of the numerical performance issue discussed in Ref. [81].) We observe that when the electronic noise is trusted, its impact on the secure key rates is insignificant. This result eases the requirements of a detector in a CVQKD experiment with the QPSK scheme. Similarly, we investigate the effects of detector efficiency in Figure 6.11b. In particular, we fix the value of electronic noise ν_{el} to be 0.05 SNU and plot four choices of detector efficiency between 0.5 and 0.8. We see the key rate curves are close to each other.

In Figure 6.12, we investigate the tradeoff between trusting the detector efficiency and lumping

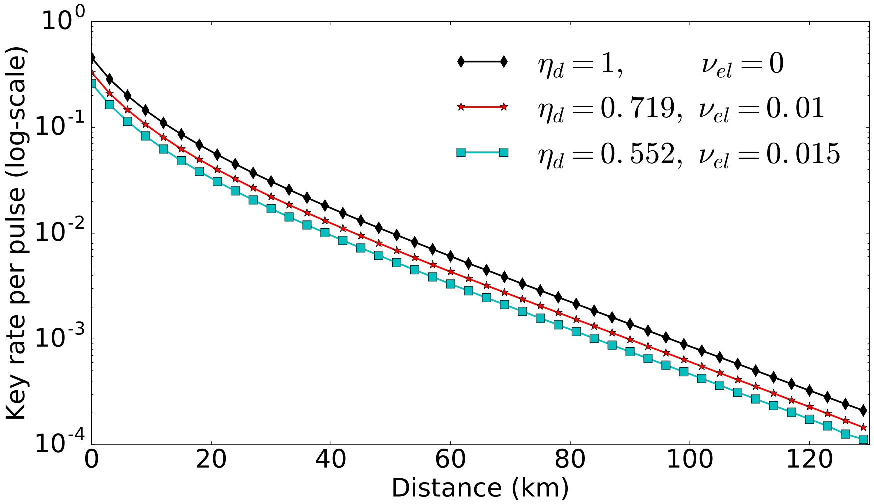
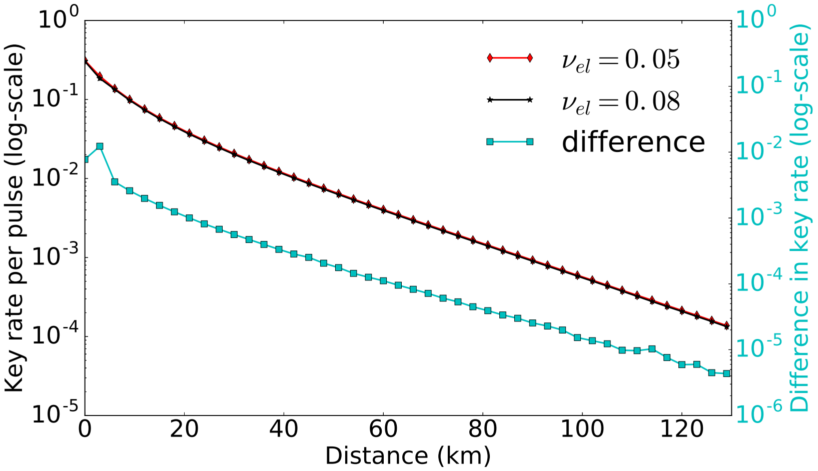


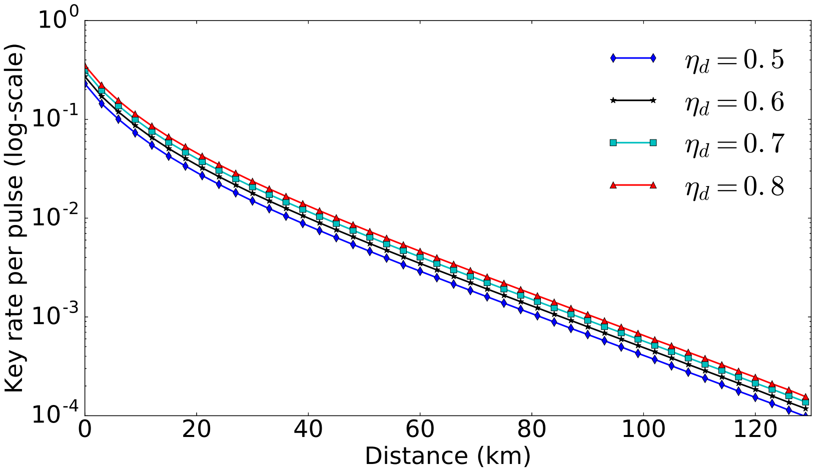
Figure 6.10: Secure key rate versus the transmission distance for different detector imperfections reported in experiments in a comparison to the ideal detector. Other parameters are the excess noise $\xi = 0.01$, error-correction efficiency $\beta_{\text{EC}} = 0.95$ and the transmittance $\eta_t = 10^{-0.02L}$ for each distance L in kilometers. For each distance, the coherent state amplitude α is optimized via a coarse-grained search in the interval $[0.5, 0.9]$ with a step size of 0.05. Black curve with diamond markers is for the ideal heterodyne detector; red curve with star markers is for the detector used in Ref. [32]; cyan curve with square markers is for the detector used in Ref. [30].

it together with the channel transmittance, similar to a scenario studied in Ref. [86] for discrete-variable systems. For the fixed amount of total transmittance $\eta := \eta_t \eta_d$, it is interesting to see how trusting different values of detector efficiency affects the key rate. We observe that when the value of the product of channel transmittance η_t and detector efficiency η_d is fixed, if the detector efficiency η_d is lower, meaning that if more contribution to the total transmittance η is trusted, then the key rate is higher. This observation is similar to the observation made for discrete-variable systems in Ref. [86].

To summarize, in a discrete modulation experiment, if one is able to obtain accurate values of η_d and ν_{el} by a suitable calibration procedure and able to maintain a low level of the effective channel excess noise ξ to a value like 0.01, then the QPSK scheme is able to extend to a distance beyond 100 km in the asymptotic regime. We remark that the optimal amplitude for the QPSK scheme in the trusted detector noise scenario is around 0.75 corresponding to a mean photon number of around 0.56, similar to the optimal amplitude in the ideal or untrusted detector noise scenario reported in our previous work [81]. This mean photon number is much lower than that for Gaussian modulation schemes.



(a)



(b)

Figure 6.11: Secure key rate versus transmission distance for different detector imperfections with the excess noise $\xi = 0.01$. For both plots, the coherent state amplitude is optimized via a coarse-grained search over the interval $[0.5, 0.9]$ with a step size 0.05 and $\beta_{\text{EC}} = 0.95$. (a) Comparison of key rates between two values of the electronic noise when the detector efficiency is set to be $\eta_d = 0.7$ for both curves. The difference of two curves is also plotted with the secondary y-axis on the right. (b) Comparison of key rates for different values of detector efficiency when the electronic noise is $\nu_{\text{el}} = 0.05$.

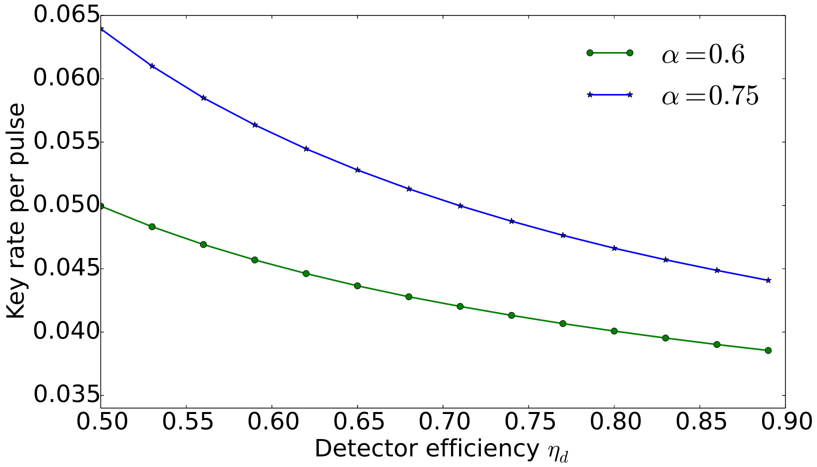


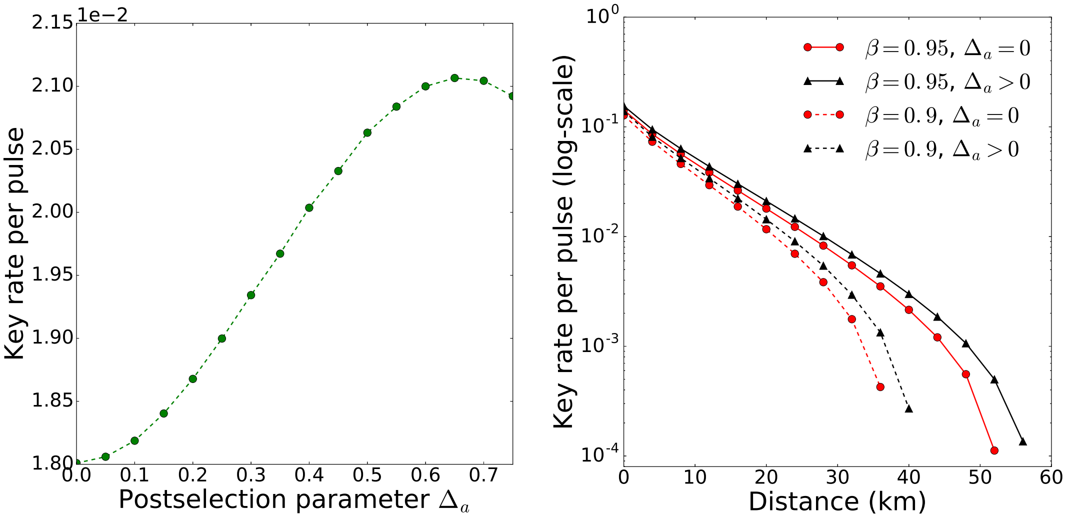
Figure 6.12: Secure key rate versus the detector efficiency η_d for a fixed value of total transmittance $\eta := \eta_t\eta_d = 0.3155$. This figure studies the tradeoff between the key rate and the amount of trusted loss. Other parameters are the excess noise $\xi = 0.01$, the electronic noise $\nu_{\text{el}} = 0.01$ and the error-correction efficiency $\beta_{\text{EC}} = 0.95$. We include two curves for different choices of coherent state amplitude α.

6.9 Postselection of data

Postselection is simple to implement in an experiment. It not only improves the key rate but also reduces the required volume of data postprocessing. Thus, it is advantageous to include a postselection step in the protocol. Postselection of data can improve the key rate of the QPSK scheme in both the untrusted and trusted detector noise scenarios.

6.9.1 Ideal detector scenario

We present the results on the effects of postselection. Our coarse-grained search for values of Δ_p suggests that the optimal value is $\Delta_p = 0$; that is, we do not postselect the data based on the phase. For the postselection parameter Δ_a related to the amplitude of the measured complex value from heterodyne detection, we then perform a coarse-grained search for its optimal value. In Figure 6.13, we consider the scenario with an excess noise $\xi = 0.04$ and with a fixed coherent state amplitude $\alpha = 0.6$ as an example. In Figure 6.13a, we plot the key rate versus this parameter Δ_a at the distance $L = 20$ km with the reconciliation efficiency $\beta_{\text{EC}} = 0.95$. From this plot, we observe that the optimal value of Δ_a is around 0.6 at this distance. We also obtain similar plots for various choices of the distance and find that the optimal value roughly falls in the interval $[0.4, 0.7]$. In Figure 6.13b, we compare key rates with or without postselection for two different values of reconciliation efficiency at different transmission distances, and, for this plot, we optimize the values of Δ_a via a coarse-grained search in the interval $[0.4, 0.7]$. We again notice that postselection with reverse reconciliation can improve the key rates. We remark that the improvement due to postselection in the reverse reconciliation scheme is more visible for less efficient error-correcting codes, larger excess noise, and longer transmission distances. This result agrees with the observation made in Ref. [194] under a restricted class of attacks.



(a)

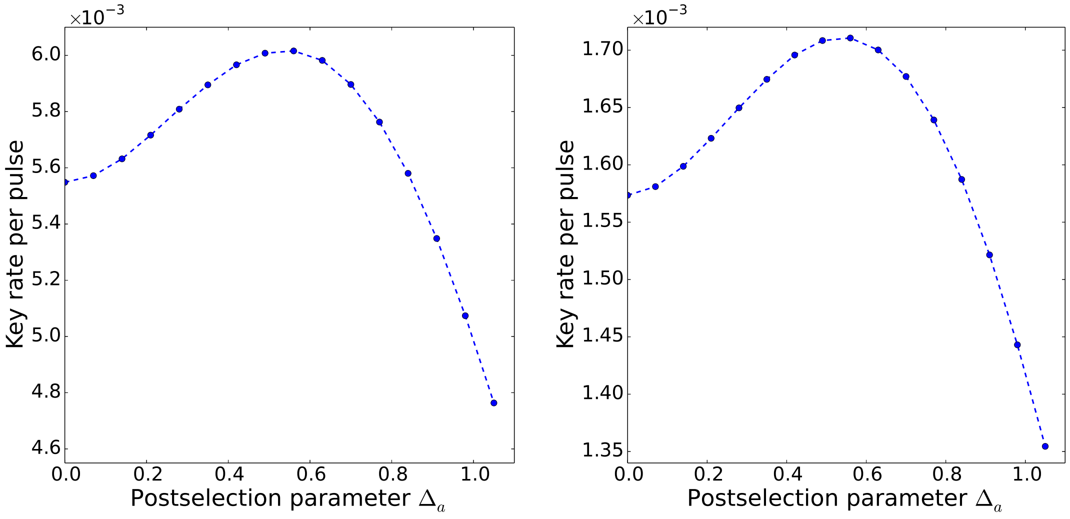
(b)

Figure 6.13: Secure key rate for the QPSK protocol using heterodyne detection with postselection in the ideal detector scenario. The excess noise is $\xi = 0.04$, the coherent state amplitude is $\alpha = 0.6$, and one of the postselection parameters is $\Delta_p = 0$. (a) Secure key rate versus the postselection parameter Δ_a at the distance $L = 20$ km. The reconciliation efficiency is $\beta_{\text{EC}} = 0.95$. (b) Secure key rate versus the transmission distance with or without postselection for two different values of reconciliation efficiency β_{EC}. Solid lines have $\beta_{\text{EC}} = 0.95$ and dashed lines have $\beta_{\text{EC}} = 0.9$. Lines with (red) circle markers have $\Delta_a = 0$, and lines with (black) triangle markers have Δ_a optimized via a coarse-grained search in the interval $[0.4, 0.7]$.

6.9.2 Trusted detector noise scenario

In this subsection, we investigate the effects of postselection in the trusted detector noise scenario. As demonstrated in our previous analysis [81]. As expected, we show here that this advantage also exists in the trusted detector noise scenario.

In Figure 6.14, we search for the optimal postselection parameter for different transmission distances and take the distances $L = 50$ km and $L = 75$ km as examples. For this figure, we also optimize the choice of coherent state amplitude via a coarse-grained search. The x axis in each plot is the postselection parameter Δ_a. We observe the optimal value of the postselection parameter Δ_a is around 0.6 for both $L = 50$ km and $L = 75$ km. We also observe that the optimal choice of the postselection parameter Δ_a does not change significantly for different distances.



(a)

(b)

Figure 6.14: (a) Secure key rate versus postselection parameter Δ_a for $L = 50$ km. (b) Secure key rate versus postselection parameter Δ_a for $L = 75$ km. For both plots, we study the trusted detector noise scenario. The channel excess noise is $\xi = 0.01$ and the error-correction efficiency $\beta_{\text{EC}} = 0.95$. The coherent state amplitude is optimized via a coarse-grained search in the interval [0.6, 0.8] with a step size of 0.05. Parameters for detectors are $\eta_d = 0.552$ and $\nu_{\text{el}} = 0.015$ from Ref. [30].

In Figure 6.15, we show the key rate as a function of transmission distance for two scenarios: with or without postselection. Since the optimal postselection parameter does not change significantly for different distances, we optimize the postselection parameter Δ_a via a coarse-grained search in a restricted interval. For this figure, we fix the coherent state amplitude to be 0.75 and the channel excess noise ξ to be 0.01. We see postselection can indeed improve the key rate. The percentage of improvement compared to the key rate without postselection is roughly between 5% to 8% and the probability of being postselected is around 70% to 80%. Thus, postselection can reduce the amount of data for postprocessing by around 20% to 30% while improving the key rate.

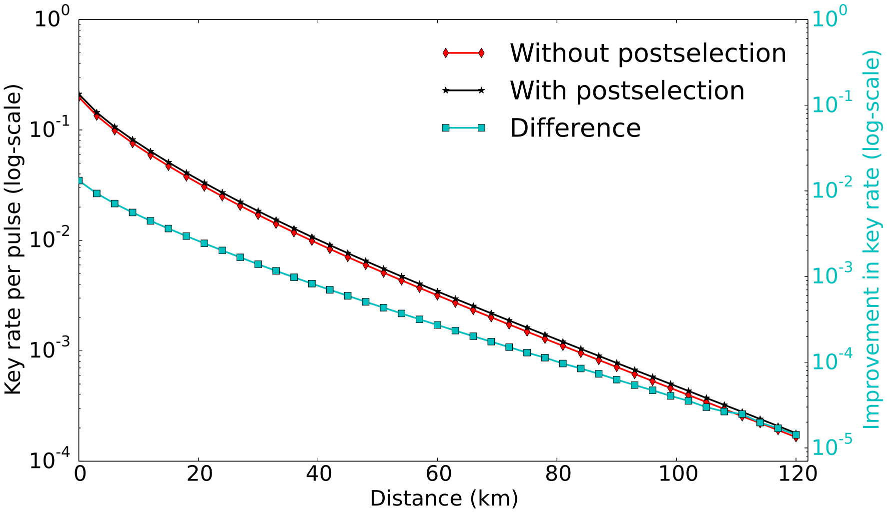


Figure 6.15: Comparison of key rates with or without postselection in the trusted detector noise scenario. Detector parameters are from Ref. [30] where $\eta_d = 0.552$ and $\nu_{\mathrm{el}} = 0.015$. The difference of two curves is also plotted with the secondary y axis on the right. Other parameters are the channel excess noise $\xi = 0.01$, the coherent state amplitude $\alpha = 0.75$, and the error-correction efficiency $\beta_{\mathrm{EC}} = 0.95$. The postselection parameter is optimized via a coarse-grained search in the interval [0.45,0.7] with a step size 0.05.

We end this section with a remark on the postselection pattern. The postselection pattern (see Figure 6.1) studied in this work is a simple, intuitive, and convenient choice when we evaluate the region operators. However, it is not necessarily the optimal way to postselect data [192, 194]. It is an interesting future work to investigate other patterns of postselection.

6.10 Application of dimension reduction method to remove the photon-number cutoff assumption

Finally, we discuss the application of the dimension reduction method (Section 4.2.3) in [85] to the heterodyne protocol to remove the photon-number cutoff assumption. To obtain a tight key rate and to ease the numerical computation, it is important to choose the subspace $\mathcal{H}_N$ carefully. In our work [85], we choose the displaced Fock space. Specifically, for each possible choice of states $|\alpha_x\rangle$ sent by Alice, we choose a displacement $\beta_x \in \mathbb{C}$ and consider the displaced Fock space $\{|n_{\beta_x}\rangle := \hat{D}(\beta_x)|n\rangle : 0 \leq n \leq N\}$ for a cutoff photon number N. The basis for Alice and Bob's joint subspace $\mathcal{H}_N$ is given by $\{|x\rangle \otimes |n_{\beta_x}\rangle : x \in \{0,1,2,3\}, 0 \leq n \leq N\}$. A main observation that leads to this choice of subspace is that if the quantum channel is a pure-loss channel, then each state that Bob receives is a coherent state with an amplitude $\sqrt{\eta}\alpha$ for the total transmittance η. This state is the displaced vacuum state with the displacement $\beta = \sqrt{\eta}\alpha$. Thus, for any $N \geq 0$, the chosen subspace $\mathcal{H}_N$ captures all the information about the state. This means the finite-dimensional optimization problem with this choice of $\mathcal{H}_N$ solves exactly the original infinite-dimensional problem in this special case. When the quantum channel is a lossy and noisy Gaussian channel, the expected state in an honest implementation is a displaced thermal state. Then there is some weight leaked into displaced Fock states with $n > 0$. We can use observed statistics from heterodyne detection to bound the weight outside the subspace $\mathcal{H}_N$ as we will discuss shortly.

As we work in the displaced Fock space, we use operators that are under conjugation of displacement operators. For any operator X, we use the short-hand notation $X_\gamma := \hat{D}(\gamma)X\hat{D}^\dagger(\gamma)$ where $\gamma \in \mathbb{C}$ is the amount of displacement. We choose the following set of constraints

$$\{\Gamma_i\} = \{|x\rangle\langle x| \otimes \hat{n}_{\beta_x}, |x\rangle\langle x| \otimes \hat{n}^2_{\beta_x} : x \in \{0,1,2,3\}\}. \tag{6.27}$$

The reason for this choice is that we use $\hat{n}_{\beta_x}, \hat{n}^2_{\beta_x}$ to bound the weight W outside the subspace as sufficiently small as we desire. It might be possible to choose other constraints that achieve a similar goal. We leave it as future work.

The relevant infinite-dimensional optimization is thus

$$\begin{aligned} \underset{\rho}{\text{minimize}} \quad & f(\rho) \\ \text{subject to} \quad & \text{Tr}(\rho) = 1 \\ & \text{Tr}_B(\rho) = \sum_{x,x'} \sqrt{p_x p'_x}\, \langle \alpha_{x'}|\alpha_x\rangle\, |x\rangle\langle x'| \\ & \text{Tr}[\rho(|x\rangle\langle x| \otimes \hat{n}_{\beta_x})] = p_x \langle \hat{n}_{\beta_x}\rangle_x \\ & \text{Tr}\big[\rho(|x\rangle\langle x| \otimes \hat{n}^2_{\beta_x})\big] = p_x \langle \hat{n}^2_{\beta_x}\rangle_x \\ & \rho \geq 0\ , \end{aligned} \tag{6.28}$$

where $\langle \hat{n}_{\beta_x} \rangle_x$ and $\langle \hat{n}^2_{\beta_x} \rangle_x$ denote the expectation values of the operator $\hat{n}_{\beta_x}$ and the operator $\hat{n}^2_{\beta_x}$ when Alice sends the state labeled by x, respectively. To apply the dimension reduction method, we need to bound the weight outside the chosen subspace $\mathcal{H}_N$. We define the projection on $\mathcal{H}_N$ as

$$\Pi^N = \sum_x |x\rangle\langle x|_A \otimes \Pi^N_{B_{\beta_x}}, \tag{6.29}$$

where

$$\Pi^N_{B_{\beta_x}} = \sum_{n=0}^{N} |n_{\beta_x}\rangle\langle n_{\beta_x}| \,. \tag{6.30}$$

One can bound the weight W outside the subspace by the dual problem of Eq. (4.50). Specifically, it is given in [85, Theorem 5].

Theorem 6.10.1 ([85, Theorem 5])**.** For the DMCVQKD protocol with Π^N defined in Eq. (6.29), the weight W outside the subspace $\mathcal{H}_N$ is bounded by

$$W = \sum_x p_x \frac{\langle \hat{n}^2_{\beta_x} \rangle_x - \langle \hat{n}_{\beta_x} \rangle_x}{N(N+1)}. \tag{6.31}$$

The proof of this theorem consists of two parts. The first part is to show that $W = \sum_x p_x W_x$ where $W_x := \mathrm{Tr}\Big[\rho^x_B(\mathbb{1}_B - \Pi^N_{B_{\beta_x}})\Big]$. The second part is to show $W_x = \frac{\langle \hat{n}^2_{\beta_x} \rangle_x - \langle \hat{n}_{\beta_x} \rangle_x}{N(N+1)}$ by looking at the dual problem of the optimization problem in Eq. (4.50) for ρ^x_B. A specific dual feasible solution gives the desired upper bound of W_x. We direct readers to [85] for the proof.

The finite-dimensional optimization problem that we can solve numerically is

$$\begin{aligned}
\underset{\tilde{\rho}}{\text{minimize}}\ & f(\tilde{\rho}) \\
\text{subject to}\ & 1 - W \le \mathrm{Tr}(\tilde{\rho}) \le 1 \\
& \frac{1}{2}\|\mathrm{Tr}_B(\tilde{\rho}) - \rho_A\|_1 \le \sqrt{W} \\
& \mathrm{Tr}[\tilde{\rho}(|x\rangle\langle x| \otimes \hat{n}_{\beta_x})] \le p_x \langle \hat{n}_{\beta_x} \rangle_x \\
& \mathrm{Tr}\big[\tilde{\rho}(|x\rangle\langle x| \otimes \hat{n}^2_{\beta_x})\big] \le p_x \langle \hat{n}^2_{\beta_x} \rangle_x \\
& \tilde{\rho} \in \mathrm{Pos}(\mathcal{H}_B)\ ,
\end{aligned} \tag{6.32}$$

where W is given in Eq. (6.31).

We use the same simulation method as in the ideal detector scenario. The result for this scenario is shown in Figure 6.16. In particular, we can write the asymptotic key rate R_∞ as $R_\infty = C_{\text{num}} - \Delta(W)$, where $\Delta(W)$ is given in Eq. (4.52) with $|Z| = 4$ for this protocol. In the

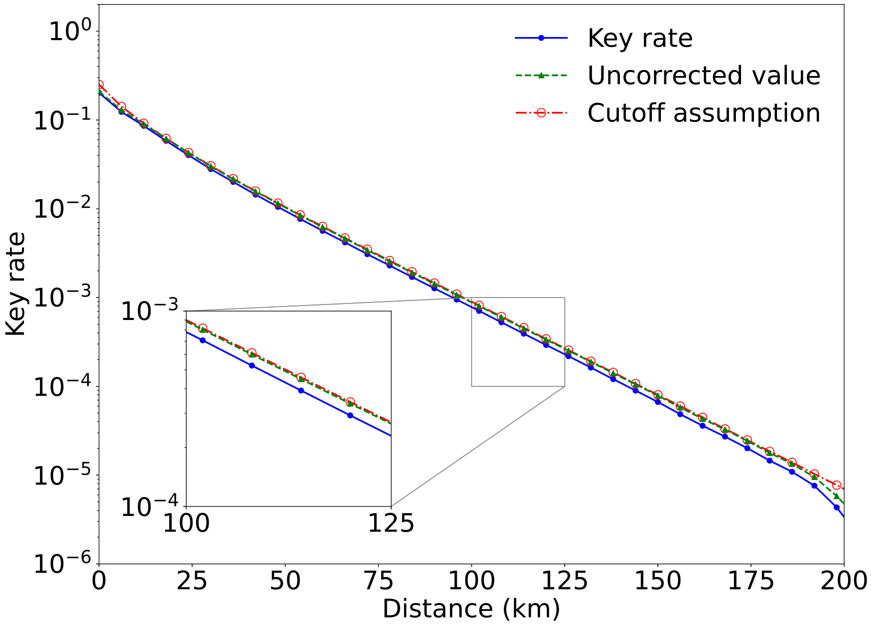


Figure 6.16: Key rate versus distance for comparing results obtained by the dimension reduction method and the one from Ref. [81] based on the photon-number cutoff assumption. Uncorrected value, i.e. C_{num}, is also plotted for comparison. Results are in the ideal detector scenario and with the channel excess noise $\xi = 0.01$. Postselection parameters and signal-state intensities from Ref. [81] are used: $\alpha = 0.6, \Delta_p = 0$, and Δ_a is optimized with a coarse-grained search over the interval $[0.5, 0.65]$. The subspace dimension parameter is $N = 20$.

same plot, we also show the uncorrected value C_{num}, and we can see the correction term $\Delta(W)$ does not affect the key rate much. It is interesting to note that the key rate obtained by the dimension reduction method is close to the key rate obtained under the photon-number cutoff assumption in previous sections [81]. In other words, all qualitative observations made in this chapter with the photon-number cutoff assumption remain valid after removing this assumption.

In the case of trusted detector noise scenario, we need to change Bob's POVM from the ideal detector POVM $\{\frac{1}{\pi}|\gamma\rangle\langle\gamma| : \gamma \in \mathbb{C}\}$ to the trusted noisy detector POVM $\{G_y : y \in \mathbb{C}\}$. This change affects both the objective function and the observables used for constraints. In particular, for the set of constraints used in Eq. (6.27), one can show [85, Appendix D] that the ideal and noisy observables are related by linear combinations. Following the notation used in Ref. [85], we denote the noisy counterpart of an operator A by $[A]'$. It can be shown that

$$[\hat{n}_{\sqrt{\eta_d}\beta_x}]' = \eta_d \hat{n}_{\beta_x} + \nu_{\text{el}}\,\mathbb{1}, \tag{6.33}$$

$$[\hat{n}^2_{\sqrt{\eta_d}\beta_x}]' = \eta_d^2 \hat{n}^2_{\beta_X} + \eta_d(4\nu_{\text{el}} + 1 - \eta_d)\hat{n}_{\beta_x}(2\nu_{\text{el}}^2 + \nu_{\text{el}})\,\mathbb{1}\,. \tag{6.34}$$

This means if we replace region operators used in the objective function of Eq. (6.32) and the expectation values $\langle \hat{n}_{\beta_x}\rangle_x$ and $\langle \hat{n}^2_{\beta_x}\rangle_x$ by the effective values $\langle \hat{n}_{\beta_x}\rangle_x^{\text{eff}}$ and $\langle \hat{n}^2_{\beta_x}\rangle_x^{\text{eff}}$, which can be obtained by the reverse relationship between the ideal and noisy observables as follows:

$$\langle \hat{n}_{\beta_x}\rangle_x^{\text{eff}} = \frac{\langle [\hat{n}_{\sqrt{\eta_d}\beta_x}]'\rangle_x - \nu_{\text{el}}}{\eta_d} \tag{6.35}$$

$$\langle \hat{n}^2_{\beta_x}\rangle_x^{\text{eff}} = \frac{1}{\eta_d^2}\Big\{\langle [\hat{n}^2_{\sqrt{\eta_d}\beta_x}]'\rangle_x - 2\nu_{\text{el}}^2 - \nu_{\text{el}} - (4\nu_{\text{el}} + 1 - \eta_d)(\langle [\hat{n}_{\sqrt{\eta_d}\beta_x}]'\rangle_x - \nu_{\text{el}})\Big\}, \tag{6.36}$$

then the finite-dimensional optimization problem in the trusted detector noise scenario is also in the form of Eq. (6.32). As expected, the key rate obtained in the trusted detector noise scenario is also close to the key rate obtained under the photon-number cutoff assumption in previous sections [81]. More results can be found in Ref. [85].

6.11 Future directions

We note that a recent work [205] applies our numerical approach to investigate the postselection patterns for DMCVQKD. It shows that the postselection based on the phase can improve the key rate for long distances. This observation is different from ours. While it is possible that we missed this result due to the looseness introduced by some numerical instability or due to our exploration of different parameter regimes, we have not been able to reproduce their results by the time of writing. It remains an ongoing effort to validate their results. Interestingly, they investigate an alternative postselection pattern: cross-shape in the phase space. They show that the cross-shape can perform similarly as the postselection with both amplitude and phase. Moreover, it can give better key rates in some scenarios. Investigation of optimal postselection patterns is an interesting direction to explore. Moreover, a recent work by Denys, Brown and Leverrier [206] gives an analytical solution for general discrete-modulation schemes. Their method is based on finding an analytical solution to the SDP in Ref. [191], which is based on the optimality of Gaussian attacks [182, 183]. Since numerical methods become more computationally challenging as the constellation size increases due to the fact that the dimension of the optimization problem depends on the number of signals sent by Alice, it is desirable to obtain analytical solutions for large constellation sizes. Despite that the key rate obtained by the analytical approach [206] is not tight for smaller constellation sizes, it approaches the key rate of Gaussian modulation as the size increases (around 64 states). It is also interesting to investigate along this direction and to try to find analytical solutions of our optimization problem. However, solving our nonlinear SDP problem analytically seems much more challenging. To improve the applicability of our numerical method to higher constellation sizes, one possible way is to improve the optimization algorithm as it is done in our work [87] by invoking a preprocessing technique to transform the optimization problem to a lower-dimensional one and applying an efficient algorithm. When the algorithm is generalized to include inequality constraints and semidefinite ordering constraints in a future work, we expect that our numerical method can handle much large constellation size than we currently can. However, it may remain challenging for this approach to analyze 64 states, which is done analytically in [206]. Another possible direction to explore is whether one can invent another dimension reduction method to reduce Alice's dimension.

We also note that the error correction protocol used in our security analysis does not use soft information such as the amplitude of Bob's measurement outcomes. As such, we do not use the channel capacity of binary additive white Gaussian noise channel that is used in Refs. [191, 206] for the mutual information between Alice and Bob. Instead, our estimation of the mutual information is related to the binary symmetric channel. While many error correction codes have been studied to achieve higher and higher reconciliation efficiency for the binary additive white Gaussian noise channel, the study of codes for the binary symmetric channel is usually outside the parameter interests in CVQKD. Unfortunately, the study of the error correction protocols is beyond the scope of this thesis. It remains an interesting future work to investigate how to use soft information in our current security proof method so that we can obtain higher mutual information and enjoy better reconciliation efficiency. It is also interesting to see concrete code constructions for the binary symmetric channel in the parameter regimes relevant for DMCVQKD.

Finally, our current analyses are all restricted to the asymptotic scenario. We notice that there is a recent work on the finite-key analysis of binary modulation protocol [207]. However, the key rate obtained was very pessimistic due to the poor performance of the binary modulation scheme. One expects that the QPSK scheme will have much better performance. It remains an open question to provide a finite key analysis of general discrete modulation beyond binary modulation. As we have extended the underlying numerical method used in this security analysis to finite-key regime [84] recently, we hope to perform the finite key analysis for discrete modulation schemes, especially the protocol studied here. However, there remain technical challenges to solve before such an analysis can be carried out and thus we leave the finite key analysis for future work.

Chapter 7

Asymptotic security analysis of DMCVQKD with homodyne detection

Similarly to the heterodyne scheme discussed in Chapter 6, one can consider a homodyne scheme. For this protocol, we consider only the ideal detector scenario under the photon-number cutoff assumption (see Section 6.5). In principle, the extension to the trusted detector noise can be done in a similar way as the heterodyne scheme. The removal of the photon-number cutoff assumption based on the dimension reduction method requires some work to bound the weight outside the desired subspace. We leave these extensions for future work. This chapter is based on [81].

7.1 Protocol description

(1). *State preparation.*—For each round $k \in [N]$ (where N is sufficiently large), according to the probability distribution $(\frac{p_A}{2}, \frac{p_A}{2}, \frac{1-p_A}{2}, \frac{1-p_A}{2})$, Alice prepares a coherent state $|\psi_k\rangle$ from the set $\{|\alpha\rangle, |-\alpha\rangle, |i\alpha\rangle, |-i\alpha\rangle\}$, where $\alpha \in \mathbb{R}$ is predetermined. Alice sends this state to Bob through an insecure quantum channel.

(2). *Measurement.*—After receiving Alice's state, Bob performs a homodyne measurement on the state. Bob first generates a random bit b_k according to the probability distribution $(p_B, 1-p_B)$. If $b_k = 0$, he measures the q quadrature and if $b_k = 1$, he measures the p quadrature. He obtains the measurement outcome $y_k \in \mathbb{R}$.

(3). *Announcement and sifting.*—After N rounds of first two steps, Alice and Bob communicate via the authenticated classical channel to partition all the rounds $[N]$ into four subsets defined as

$$\begin{aligned} \mathcal{I}_{qq} &= \{k \in [N] : \lvert\psi_k\rangle \in \{\lvert\alpha\rangle, \lvert-\alpha\rangle\}, b_k = 0\}, \\ \mathcal{I}_{qp} &= \{k \in [N] : \lvert\psi_k\rangle \in \{\lvert\alpha\rangle, \lvert-\alpha\rangle\}, b_k = 1\}, \\ \mathcal{I}_{pq} &= \{k \in [N] : \lvert\psi_k\rangle \in \{\lvert i\alpha\rangle, \lvert-i\alpha\rangle\}, b_k = 0\}, \\ \mathcal{I}_{pp} &= \{k \in [N] : \lvert\psi_k\rangle \in \{\lvert i\alpha\rangle, \lvert-i\alpha\rangle\}, b_k = 1\}. \end{aligned} \tag{7.1}$$

Then Alice and Bob randomly select a small test subset $\mathcal{I}_{qq,\text{test}} \subset \mathcal{I}_{qq}$. This selection allows them to define $\mathcal{I}_{\text{key}}$ as the subset of $\mathcal{I}_{qq}$ after removing $\mathcal{I}_{qq,\text{test}}$ and to define $\mathcal{I}_{\text{test}} = \mathcal{I}_{qq,\text{test}} \cup \mathcal{I}_{qp} \cup \mathcal{I}_{pq} \cup \mathcal{I}_{pp}$. Let m denote the size of the index set $\mathcal{I}_{\text{key}}$ and let f be a bijective function from $[m] = \{1, 2, \dots, m\}$ to $\mathcal{I}_{\text{key}}$. After sifting, Alice sets her string $\mathbf{X} = (x_1, x_2, \dots, x_m)$ according to the rule

$$\forall j \in [m],\ x_j = \begin{cases} 0 & \text{if } \lvert\psi_{f(j)}\rangle = \lvert\alpha\rangle, \\ 1 & \text{if } \lvert\psi_{f(j)}\rangle = \lvert-\alpha\rangle. \end{cases} \tag{7.2}$$

(4). *Parameter estimation.*—Alice and Bob perform the parameter estimation by disclosing all the information in the rounds indexed by the test set $\mathcal{I}_{\text{test}}$. To perform such an analysis, they process the data by computing quantities like the first and second moments of q and p quadratures conditioned on each of four states that Alice sends. These quantities allow them to constrain their joint state ρ_{AB}. They then calculate the secret key rate according to the optimization problem in Eq. (7.4). If their analysis shows that no secret keys can be generated, then they abort the protocol. Otherwise, they proceed.

(5). *Reverse reconciliation key map.*—Bob performs a key map to obtain his raw key string. This key map discretizes his measurement outcome y_k to an element in the set $\{0, 1, \perp\}$ for each $k \in \mathcal{I}_{\text{key}}$. For each $j \in [m]$, Bob sets z_j according to the rule

$$z_j = \begin{cases} 0 & \text{if } y_{f(j)} \in [\Delta_c, \infty), \\ 1 & \text{if } y_{f(j)} \in (-\infty, -\Delta_c], \\ \perp & \text{if } y_{f(j)} \in (-\Delta_c, \Delta_c). \end{cases} \tag{7.3}$$

Note that $\Delta_c \geq 0$ is a parameter related to the postselection of data. A protocol without postselection can set $\Delta_c = 0$. At the end of this process, Bob has a string $\mathbf{Z} = (z_1, z_2, \dots, z_m)$. In communication between Alice and Bob, positions with the symbol $\perp$ are deleted from their strings. With a slight abuse of notation, we use $\mathbf{X}, \mathbf{Z}$ to mean the strings after removing the positions related to $\perp$. $\mathbf{Z}$ is called the raw key string.

(6). *Error correction and privacy amplification.*—Bob chooses a suitable error-correction protocol and a suitable privacy-amplification protocol according to the security analysis done in the parameter estimation step and communicates the choices to Alice. Alice and Bob then apply the chosen error-correction protocol and privacy-amplification protocol to generate a secret key.

We remark on the asymmetric roles of these four states and asymmetric choices of quadrature measurements considered here. In this specific setup, Alice and Bob use only signal states $\{|\alpha\rangle, |-\alpha\rangle\}$ and q quadrature measurement data to generate keys and use all other combinations to probe eavesdropping activities. In the asymptotic limit, we can set p_A and p_B arbitrarily close to 1 so that the sifting factor of the protocol is 1 (in the absence of postselection) [163]. However, for a finite number N, it is unlikely that p_A and p_B can be arbitrarily close to 1, since one needs to balance the trade-off between the sifting factor and the accuracy of parameter estimation. In this case, one needs to optimize the choices of p_A and p_B for a given choice of N. The reason that we choose this asymmetric version here is to simplify some numerical calculation and to maximize the sifting factor. We may also consider another variant of this protocol, that is, allowing Alice and Bob to generate keys from both $\mathcal{I}_{qq}$ and $\mathcal{I}_{pp}$. Then for $p_A = p_B = \frac{1}{2}$, the protocol has $\frac{1}{2}$ sifting factor (in the absence of postselection). However, we point out that the essential idea of our security proof in the asymptotic limit is the same for these different variations.

7.2 Key rate optimization

From the homodyne measurement, we can obtain expectation values of the first and second moments of the quadrature operators $\langle\hat{q}\rangle$, $\langle\hat{q}^2\rangle$, $\langle\hat{p}\rangle$, and $\langle\hat{p}^2\rangle$. We can calculate the mean photon number of each conditional state ρ_B^x from the homodyne measurement outcomes, since $\hat{n} = \frac{1}{2}(\hat{q}^2 + \hat{p}^2 - 1) = \hat{a}^\dagger\hat{a}$. Since $\hat{d} = \hat{q}^2 - \hat{p}^2 = \hat{a}^2 + (\hat{a}^\dagger)^2$, we also calculate the expectation value of $\hat{d}$ to constrain ρ_{AB}.

The relevant optimization problem is

$$\begin{aligned}
\text{minimize}\quad & D\big(\mathcal{G}(\rho_{AB})||\mathcal{Z}[\mathcal{G}(\rho_{AB})]\big) \\
\text{subject to}\quad & \\
& \mathrm{Tr}[\rho_{AB}(|x\rangle\langle x|_A \otimes \hat{q})] = p_x\langle\hat{q}\rangle_x, \\
& \mathrm{Tr}[\rho_{AB}(|x\rangle\langle x|_A \otimes \hat{p})] = p_x\langle\hat{p}\rangle_x, \\
& \mathrm{Tr}[\rho_{AB}(|x\rangle\langle x|_A \otimes \hat{n})] = p_x\langle\hat{n}\rangle_x, \\
& \mathrm{Tr}\Big[\rho_{AB}(|x\rangle\langle x|_A \otimes \hat{d})\Big] = p_x\langle\hat{d}\rangle_x, \\
& \mathrm{Tr}[\rho_{AB}] = 1, \\
& \mathrm{Tr}_B[\rho_{AB}] = \sum_{i,j=0}^{3} \sqrt{p_i p_j}\,\langle\varphi_j|\varphi_i\rangle\,|i\rangle\langle j|_A\,, \\
& \rho_{AB} \geq 0,
\end{aligned} \tag{7.4}$$

where $x \in \{0, 1, 2, 3\}$ and $\langle\hat{q}\rangle_x, \langle\hat{p}\rangle_x, \langle\hat{n}\rangle_x$, and $\langle\hat{d}\rangle_x$ denote the corresponding expectation values of operators $\hat{q}, \hat{p}, \hat{n}$, and $\hat{d}$ for the conditional state ρ_B^x, respectively. In Appendix C.2, we discuss how we make these operators finite dimensional under the photon-number cutoff assumption.

We remark that one can add more fine-grained constraints using the POVM description of homodyne measurements or using the interval operators I_0 and I_1, which we define shortly. Additional constraints can only improve the key rate, as they reduce the size of the feasible set **S**. Nevertheless, we observe that this set of constraints already gives us quite tight key rates. We expect that additional constraints will provide only marginal improvements. For the ease of presentation, we choose this set of coarse-grained constraints.

We now specify the maps $\mathcal{G}$ and $\mathcal{Z}$. For the reverse reconciliation, the postprocessing map $\mathcal{G}(\sigma) = K\sigma K^\dagger$ is given by the following Kraus operator:

$$K = \sum_{z=0}^{1} |z\rangle_R \otimes (|0\rangle\langle 0| + |1\rangle\langle 1|)_A \otimes (\sqrt{I_z})_B, \tag{7.5}$$

where I_0 and I_1 are *interval operators* defined in terms of projections onto (improper) eigenstates of q quadrature:

$$I_0 = \int_{\Delta_c}^{\infty} dq\, |q\rangle\langle q|\,, \quad I_1 = \int_{-\infty}^{-\Delta_c} dq\, |q\rangle\langle q|\,. \tag{7.6}$$

In the definition of K, we project Alice's register A onto the subspace spanned by the first two basis states (which are related to the states $|\alpha\rangle$ and $|-\alpha\rangle$) and act on Bob's register by interval operators from the q quadrature measurement, since secret keys are generated only from the rounds where Alice sends $|\alpha\rangle$ or $|-\alpha\rangle$ and Bob performs q quadrature measurements in this protocol. We remark how the postselection is handled in our security proof. Since Δ_c is a postselection parameter, the effect of postselection is reflected in the definition of interval operators which are used in the postprocessing map $\mathcal{G}$. Finally, the pinching quantum channel $\mathcal{Z}$ is described by the projections $Z_0 = |0\rangle\langle 0|_R \otimes \mathbb{1}_{AB}$ and $Z_1 = |1\rangle\langle 1|_R \otimes \mathbb{1}_{AB}$.

7.3 Simulated statistics and calculation of error correction cost

We follow the same simulation method as in Section 6.4. From the homodyne measurement, for each $\alpha_x \in \{\alpha, -\alpha, i\alpha, -i\alpha\}$, the simulated statistics is given as

$$\begin{aligned}
\langle \hat{q} \rangle_x &= \sqrt{2\eta}\,\mathrm{Re}(\alpha_x), \\
\langle \hat{p} \rangle_x &= \sqrt{2\eta}\,\mathrm{Im}(\alpha_x), \\
\langle \hat{n} \rangle_x &= \eta|\alpha_x|^2 + \frac{\eta\xi}{2}, \\
\langle \hat{d} \rangle_x &= \eta[\alpha_x^2 + (\alpha_x^*)^2].
\end{aligned} \tag{7.7}$$

With these values specified, we perform the optimization to bound Eve's information.

Since we simulate the experimental behavior and the cost of error correction is not a part of the optimization, we now present the analytical formula to estimate δ_{EC} from the simulated statistics and numerically evaluate the formula. In this protocol, we use only $|+\alpha\rangle, |-\alpha\rangle$ ($\alpha \in \mathbb{R}$) and the q quadrature measurement to generate keys. After Bob performs his key map, Alice and Bob effectively communicate via a binary channel for the purpose of error correction. From the simulation, the probability distributions of Bob's q quadrature measurement outcomes for conditional states ρ_B^0 and ρ_B^1 are

$$\begin{aligned} P(q|0) &= \frac{1}{\sqrt{\pi(\eta\xi+1)}} e^{\frac{-(q-\sqrt{2\eta}\alpha)^2}{\eta\xi+1}}, \\ P(q|1) &= \frac{1}{\sqrt{\pi(\eta\xi+1)}} e^{\frac{-(q+\sqrt{2\eta}\alpha)^2}{\eta\xi+1}}. \end{aligned} \tag{7.8}$$

Since we allow postselection with the cutoff parameter Δ_c, the sifting probability reads

$$p_{\text{pass}} = 1 - \frac{1}{2}\int_{-\Delta_c}^{\Delta_c} P(q|0)dq - \frac{1}{2}\int_{-\Delta_c}^{\Delta_c} P(q|1)dq. \tag{7.9}$$

The error probability between Alice's and Bob's strings is

$$e = \frac{1}{p_{\text{pass}}}\Big(\frac{1}{2}\int_{-\infty}^{-\Delta_c} P(q|0)dq + \frac{1}{2}\int_{\Delta_c}^{\infty} P(q|1)dq\Big). \tag{7.10}$$

From Eq. (6.21), we can calculate the error correction cost δ_{EC} by the following expression:

$$\begin{aligned} \delta_{\text{EC}} &= (1-\beta_{\text{EC}})H(\mathbf{Z}) + \beta_{\text{EC}} H(\mathbf{Z}|\mathbf{X}) \\ &= (1-\beta_{\text{EC}})H(\mathbf{Z}) + \beta_{\text{EC}} h(e). \end{aligned} \tag{7.11}$$

In this chapter, we use $\beta_{\text{EC}} = 0.95$ in all figures unless mentioned otherwise.

7.4 Simulation results

7.4.1 Loss-only scenario

We first present the results for the loss-only scenario, that is, $\xi = 0$. We plot the key rate versus transmission distance in the loss-only scenario for this protocol in Section 7.4.1. As discussed in Section 6.6, we can also obtain an analytical result for the homodyne scheme to have a direct comparison with our numerical result. We outline this procedure below.

The procedure outlined here is similar to the calculation in Ref. [194]. For this protocol, since $\alpha_x \in \{\alpha, -\alpha\}$, Eve's conditional states $|\epsilon_x\rangle$ are either $\left|\sqrt{1-\eta}\alpha\right\rangle$ or $\left|-\sqrt{1-\eta}\alpha\right\rangle$, which span only a two-dimensional subspace. We can find a two-dimensional representation of $|\epsilon_x\rangle$ as

$$\begin{aligned} |\epsilon_0\rangle &= \left|\sqrt{1-\eta}\alpha\right\rangle = c_0\,|e_0\rangle + c_1\,|e_1\rangle\,, \\ |\epsilon_1\rangle &= \left|-\sqrt{1-\eta}\alpha\right\rangle = c_0\,|e_0\rangle - c_1\,|e_1\rangle\,, \end{aligned} \tag{7.12}$$

where $|e_0\rangle$ and $|e_1\rangle$ are defined as

$$\begin{aligned} |e_0\rangle &= \frac{1}{\sqrt{\cosh[(1-\eta)\alpha^2]}} \sum_{n=0}^{\infty} \frac{(\sqrt{1-\eta}\alpha)^{2n}}{\sqrt{(2n)!}}\,|2n\rangle\,, \\ |e_1\rangle &= \frac{1}{\sqrt{\sinh[(1-\eta)\alpha^2]}} \sum_{n=0}^{\infty} \frac{(\sqrt{1-\eta}\alpha)^{2n+1}}{\sqrt{(2n+1)!}}\,|2n+1\rangle\,, \end{aligned} \tag{7.13}$$

respectively, $c_0 = e^{-\frac{(1-\eta)\alpha^2}{2}}\sqrt{\cosh[(1-\eta)\alpha^2]}$ and $c_1 = e^{-\frac{(1-\eta)\alpha^2}{2}}\sqrt{\sinh[(1-\eta)\alpha^2]}$.

We now directly evaluate the Devetak-Winter formula

$$R^{\infty} = \beta_{\mathrm{EC}} I(\mathbf{X}\!:\!\mathbf{Z}) - \chi(\mathbf{Z}\!:\!E). \tag{7.14}$$

We obtain $I(\mathbf{X}\!:\!\mathbf{Z})$ by a calculation similar to Eq. (7.11). We can directly calculate $\chi(\mathbf{Z}\!:\!E)$ via

$$\chi(\mathbf{Z}\!:\!E) = H(\overline{\rho}_E) - \sum_{j=0}^{1} P(z=j) H(\rho_{E,j}), \tag{7.15}$$

where $H(\sigma) = -\operatorname{Tr}(\sigma \log_2 \sigma)$ is the von Neumann entropy and the relevant states are

$$\begin{aligned} \rho_{E,j} &= \sum_{i=0}^{1} \frac{P(x=i, z=j)}{P(z=j)}\,|\epsilon_i\rangle\!\langle\epsilon_i|\,, \\ \overline{\rho}_E &= \sum_{j=0}^{1} P(z=j)\rho_{E,j}, \end{aligned} \tag{7.16}$$

where $P(x,z)$ is the joint probability distribution of x and z and $P(z)$ is the marginal probability distribution of z. Each of the relevant Eve's states $|\epsilon_x\rangle$ has a two-dimensional matrix representation in the basis $\{|e_0\rangle, |e_1\rangle\}$, and, thus, it is straightforward to directly evaluate the Devetak-Winter formula.

In Section 7.4.1, we observe a similar numerical behavior as in the case of heterodyne detection in Section 6.7.1. We also include the repeaterless secret key capacity bound for the pure-loss channel [64, 65], that is, $R^{\infty} \leq -\log_2(1-\eta)$. With the reconciliation efficiency $\beta_{\mathrm{EC}} = 0.95$, the key rate for this protocol is roughly 1/15 of the secret key capacity bound.

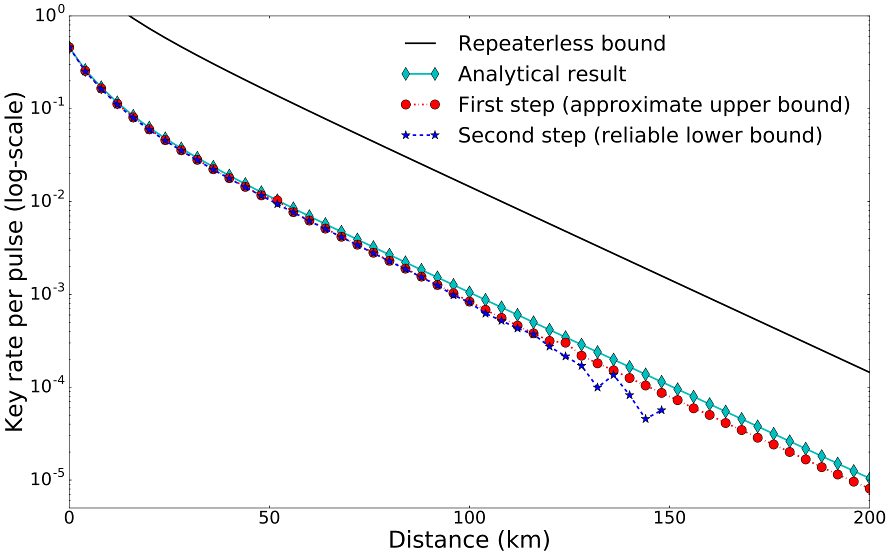
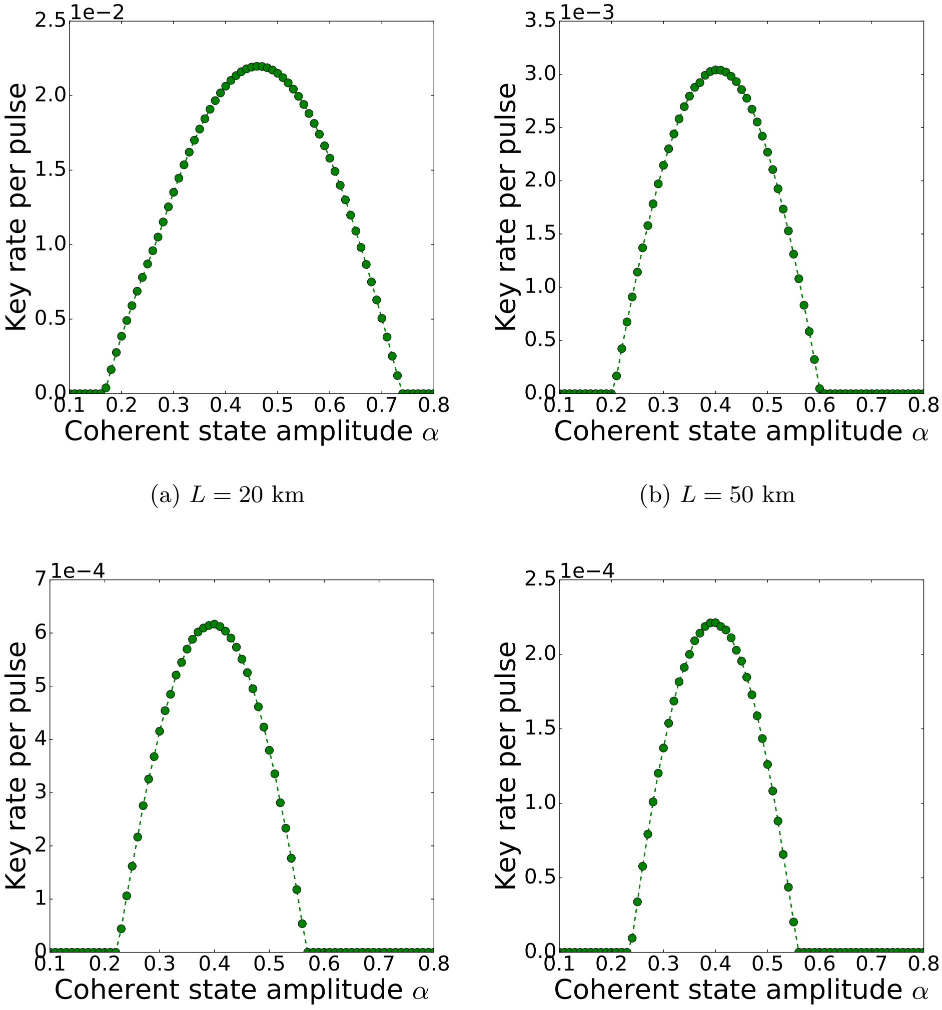


Figure 7.1: Secure key rates versus the transmission distance for the pure-loss channel to demonstrate the numerical behavior of our two-step key rate calculation procedure and to compare with the analytical results (direct evaluation of the Devetak-Winter formula). The transmittance is $\eta = 10^{-0.02L}$ for each distance L in kilometers, and the reconciliation efficiency is $\beta_{\mathrm{EC}} = 0.95$. The curve with circle markers is the approximate upper bound from the first step, and the curve with star markers is the reliable lower bound obtained from the second step. The curve with square markers is the analytical results presented in Section 7.4.1. The solid line with no markers is the repeaterless secret key capacity bound [65]. The coherent state amplitude α is optimized via a coarse-grained search in the interval [0.36, 0.6].

7.4.2 Noisy scenario

We first investigate the optimal choice of coherent state amplitude α in the absence of postselection, that is, $\Delta_c = 0$. In Figure 7.2, we plot the key rate versus the choice of α for a selected set of distances in the case of the excess noise $\xi = 0.01$. The optimal choice of α for each distance $L = 20, 50, 80, 100$ km lies around 0.4, corresponding to a mean photon number of 0.16 from Alice's source. We also see that the optimal choice does not change significantly for different distances. This observation allows us to search in a restricted interval when we optimize α to maximize the key rate for each transmission distance.

In Figure 7.3, we show the secret key rates as a function of the transmission distance for the QPSK protocol with homodyne detection for different choices of excess noise ξ. For this plot, we optimize the coherent state amplitude α by a coarse-grained search in the interval $[0.35, 0.6]$.



(a) $L = 20$ km

(b) $L = 50$ km

(c) $L = 80$ km

(d) $L = 100$ km

Figure 7.2: Secure key rate for the homodyne protocol versus coherent state amplitude α for selected choices of distances (a) $L = 20$ km, (b) $L = 50$ km, (c) $L = 80$ km, and (d) $L = 100$ km with the excess noise $\xi = 0.01$ and reconciliation efficiency $\beta_{\text{EC}} = 0.95$.

As we can see from the plot, we can reach around 200 km with an experimentally feasible value of excess noise, say, $\xi = 0.01$ [30, 34] with the current technology before the key rate becomes insignificant (say, less than 10^{-6} per pulse). To put the number in a more concrete and realistic context, if we consider a system with the repetition rate of 1 GHz and with the detector efficiency 55%, we can obtain 10^3 bits per second at the distance of around 170 km if the total excess noise ξ can be made to be 1% or less.

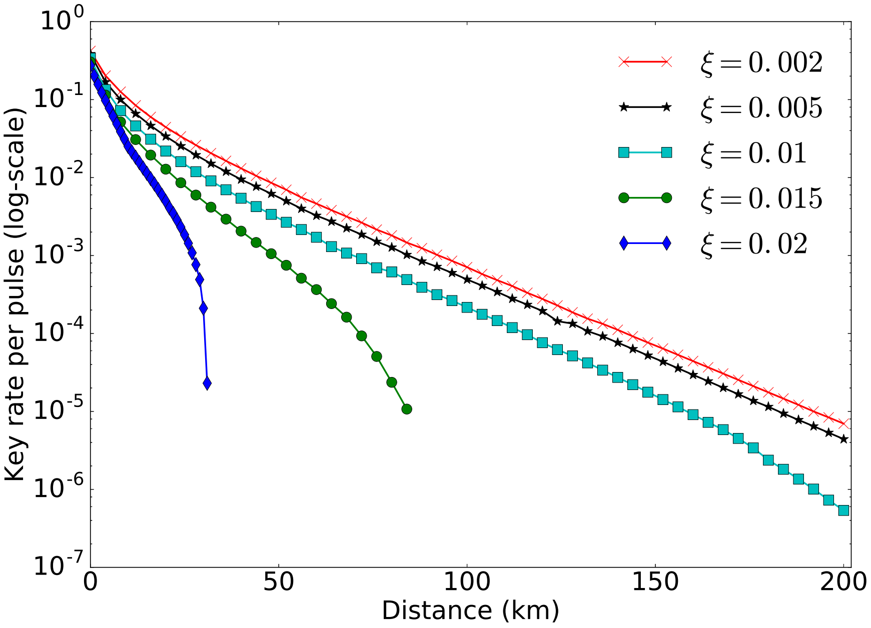
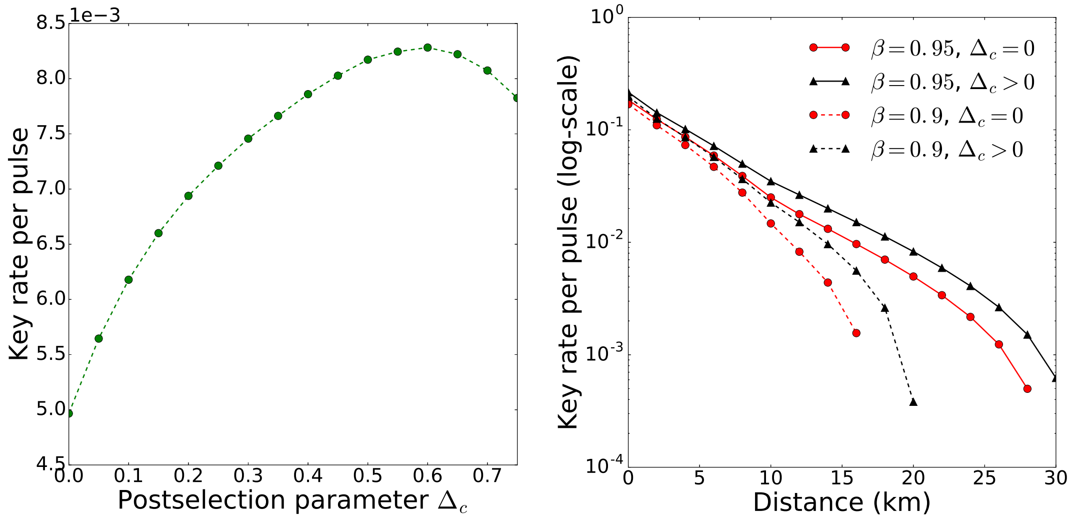


Figure 7.3: Secure key rate versus the transmission distance for the homodyne protocol for different values of the excess noise, from top to bottom, $\xi = 0.002, 0.005, 0.01, 0.015, 0.02$. The coherent state amplitude α is optimized via a coarse-grained search in the interval $[0.35, 0.6]$, the transmittance is $\eta = 10^{-0.02L}$ for each distance L in kilometers and the reconciliation efficiency is $\beta_{\mathrm{EC}} = 0.95$.



(a)

(b)

Figure 7.4: Secure key rate for the homodyne protocol with postselection. The excess noise is $\xi = 0.02$, and the coherent state amplitude is $\alpha = 0.45$. (a) Secure key rate versus the postselection parameter Δ_c at the distance $L = 20$ km with the reconciliation efficiency $\beta_{\text{EC}} = 0.95$. (b) Secure key rate versus the transmission distance with or without postselection for two different values of β_{EC}. Solid lines have $\beta_{\text{EC}} = 0.95$, and dashed lines have $\beta_{\text{EC}} = 0.9$. Lines with (red) circle markers have $\Delta_c = 0$, and lines with (black) triangle markers have Δ_c optimized via a coarse-grained search in the interval $[0.5, 0.7]$.

We also investigate the effects of postselection. The idea of postselection was initially introduced to CVQKD protocols in order to beat the 3 dB limit [192]. The key rate can be potentially improved by discarding very noisy data where Eve has more advantages in determining the raw key than the party (Bob in the case of direct reconciliation and Alice in the case of reverse reconciliation) who needs to match the raw key via the error correction. Intuitively, if we optimize the postselection parameter Δ_c, the key rate can never be lower than the protocol without postselection since one can always set $\Delta_c = 0$ if it is optimal to do so. The important observation here is that our security proof technique allows us to consider postselection with $\Delta_c > 0$ by a simple modification of the postprocessing map $\mathcal{G}$, unlike previous security proofs based on Gaussian optimality. In Figure 7.4, we take the case with an excess noise $\xi = 0.02$ and a fixed coherent state amplitude $\alpha = 0.45$ as an example to illustrate how the postselection strategy can improve the key rate in the reverse reconciliation scheme and to what extent it can help. We first search an optimal value for the postselection parameter Δ_c by a coarse-grained search, and we see, in Figure 7.4a, that the optimal value is around 0.6 at the distance $L = 20$ km. We also obtain similar plots for different choices of distance and find that the optimal value falls roughly in the interval $[0.5, 0.7]$. In Figure 7.4b, we compare the key rate with postselection ($\Delta_c > 0$) to that without postselection ($\Delta_c = 0$) for two values of reconciliation efficiency β_{EC}. In this plot, we optimize the postselection parameter Δ_c via a coarse-grained search in the interval $[0.5, 0.7]$. Since the curves with postselection are above the curves without postselection, we see that the postselection strategy can improve the key rates. We also notice that, for reverse reconciliation schemes, the advantage of postselection also depends on the reconciliation efficiency β_{EC}. The gap between these two scenarios $\Delta_c = 0$ and $\Delta_c > 0$ is smaller when a more efficient code (larger β_{EC}) is used.

7.5 Comparison between the homodyne and heterodyne protocols

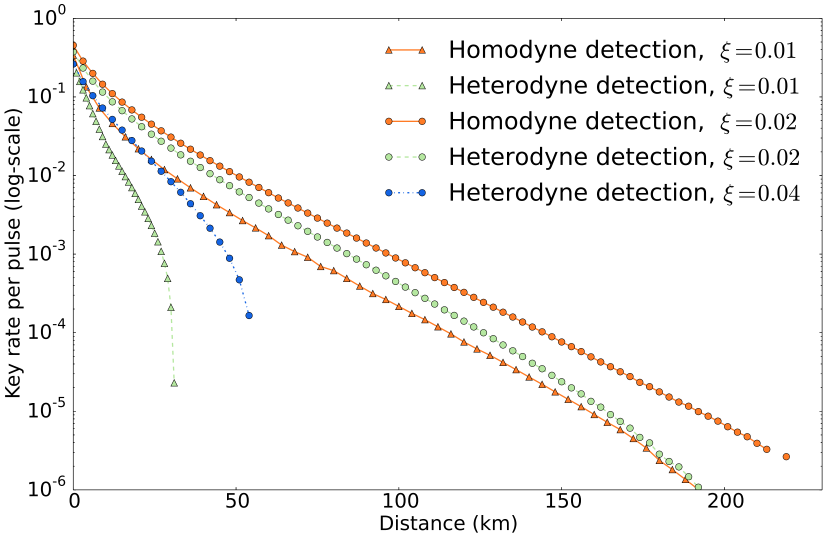


Figure 7.5: Secure key rate versus transmission distance for a direct comparison between the homodyne and heterodyne protocols. It is important to note that these two protocols have different protocol descriptions in terms of probability distribution of signal states and key map in addition to the difference in the detection setup. Curves for the homodyne protocol are plotted with triangle markers (from Figure 7.3); the excess noise is $\xi = 0.01, 0.02$ from top to bottom for curves with triangle markers. Curves for the heterodyne protocol are plotted with circle markers (from Figure 6.6); the excess noise is $\xi = 0.01, 0.02, 0.04$ from top to bottom for curves with circle markers.

In this section, we compare the heterodyne detection protocol studied in Chapter 6 and the homodyne detection protocol studied in this chapter. Before we compare results, it is important to notice that these two protocols have different protocol descriptions in addition to the difference in the detection setup. In particular, the homodyne protocol here uses a similar idea of the efficient BB84 protocol [163] and uses only two out of four states for the key generation, while the heterodyne protocol uses all four states and the probability distribution for choosing each signal is uniform. With this in mind, we compare these two protocols and discuss issues. We see that the key rate for the heterodyne protocol is much higher than the homodyne protocol when the excess noise is large. For a direct comparison, we replot key rates of both protocols for the values of excess noise $\xi = 0.01$ and 0.02 from Figure 7.3 and Figure 6.6 in Figure 7.5. We observe that the heterodyne protocol achieves much higher key rates and reaches longer distances than the homodyne protocol studied in this chapter for the same amount of excess noise. In this figure, we also plot the key rate of the heterodyne protocol with the excess noise $\xi = 0.04$ for a direct comparison to the key rate of the homodyne protocol with the excess noise $\xi = 0.02$. We see that the heterodyne protocol behaves similarly as the homodyne protocol with half of the excess noise for those values of excess noise considered here. This result is unsatisfying. We believe this discrepancy comes from our treatment of homodyne protocols. Since we use only two states to generate keys and two other states purely for the testing purpose, our treatment for this QPSK scheme is more like a binary modulation scheme, which has worse performances. We note that a recent work [208] based on our method claims significantly higher key rate for this homodyne scheme by applying a different key map and using all four states. While we have not verified the validity of their results, it is not completely surprising that a different key map can significantly improve the key rate and bring the performance of the homodyne protocol close to the heterodyne protocol. We aim at trying to reproduce their results in a near future.

7.6 Future directions

We would like first to reproduce the results in a recent work [208] and verify that a change of key map can significantly improve the performance of this protocol. We also would like to extend the analysis to the trusted detector noise scenario and also without the photon-number cutoff assumption. Then we want to perform the finite key analysis for this protocol as well.

Chapter 8

Finite key analysis via the entropy accumulation theorem

Finite key analysis is important for practical QKD implementations. As we discussed in Chapter 3, there are several existing methods to calculate the finite key rate for general QKD protocols. We have extended the numerical framework [144, 146] to the finite key regime (Section 4.3) [84], which can be easily combined with the quantum de Finetti theorem (Section 3.5) or the postselection technique (Section 3.6) to prove the security against the general attacks. Another interesting technique that we mentioned in Section 3.7 is the entropy accumulation theorem. It has been successfully applied to DIQKD protocols and produced tight key rates [43]. So far, the entropy accumulation theorem [97, 98] has not been applied to device-dependent QKD protocols beyond a simple application to single-photon entanglement-based BB84 studied in the first EAT paper [97]. There are two major obstacles to overcome in applying EAT to device-dependent QKD. The first one is to guarantee the Markov condition is satisfied by the protocol. This condition in effect guarantees the protocol is well-behaved in what side-information is leaked to Eve during the protocol. This is an important aspect of applying the EAT to device-dependent QKD protocols, as device-dependent protocols often make use of announcing more information about each signal than DI protocols. The second challenge is to have a general method to construct the required min-tradeoff function. This challenge has been addressed for the DI scenario in a recent work [44], but has not been addressed in the device-dependent scenario in which we have more structure of which to take advantage. In this chapter, we apply numerical methods described in Section 4.4 to construct min-tradeoff functions for device-dependent QKD protocols satisfying the above restrictions. The purpose of this chapter is to use simple examples to illustrate principles and demonstrate that our method can produce good key rates for various protocols.

Our work here provides a first step along the research direction of applying the EAT to general device-dependent QKD protocols. Currently, we limit ourselves to entanglement-based protocols. To handle prepare-and-measure protocols, one needs to be able to incorporate the source-replacement scheme [110] into the EAT subprotocol and to add a promise that Alice's reduced density operator is not affected by Eve's attacks. There remain technical challenges to do so. We leave it for future work. Another restriction is that in all examples considered here, all public announcements are based on seeded randomness so that the Markov chain condition (Definition 2.2.39) required in the EAT is trivially satisfied. In our forthcoming paper [83], we will report more results regarding the Markov chain condition. We will show that when Alice's and Bob's POVMs related to announcements and results of the key map satisfy a simple block-diagonal structure, then the protocol satisfies the Markov chain condition. We note that this block-diagonal structure is satisfied for many practical discrete-variable QKD protocols and it is a sufficient (probably not necessary) condition.

8.1 EAT subprotocol

To prove the security of QKD protocol, we use the EAT subprotocol to calculate the smooth min-entropy (or the sandwiched Rényi entropy) rate. A QKD protocol can consist of the EAT subprotocol followed by error correction and privacy amplification subprotocols. The EAT subprotocol is described in Protocol 8.1.

Protocol 8.1 Device-Dependent Entropy Accumulation Subprotocol

Inputs:

$\{M_a^A\}, \{M_b^B\}$	Alice and Bob's measurement devices (POVMs)
$\mathbf{K}$	Subset of Alice and Bob's public announcements kept in the general sifting step
$N \in \mathbb{N}_+$	Number of rounds
$\gamma \in (0,1)$	Probability of testing
$\mathcal{Q}$	Set of acceptable frequency distributions over $\mathcal{X}$

Protocol:

0. **State Transmission:** Eve distributes the N states, which may be entangled in an arbitrary manner, such that the total state is of the form $\rho_{Q_{A,1}^N Q_{B,1}^N E}$.

1. For $i \in [N]$, do Steps 2-7:

2. **Measurements**: Alice and Bob implement their local POVMs $\{M_a^A\}_a$, $\{M_b^B\}_b$ to measure their respective halves of the state and record their outcomes.

3. **Data Partition and Announcement:** Alice partitions her data into public register $\widetilde{A}_i$, and private register $\overline{A}_i$. Likewise, Bob partitions his data into public register $\widetilde{B}_i$ and private register $\overline{B}_i$. Then they announce their public data.

4. **Rounds Labeling**: Alice randomly chooses $T_i \in \{0,1\}$ according to some function $f_T : \widetilde{A} \times \widetilde{B} \to \{0,1\}$, which only depends on the side information that is randomly seeded. This is the division of data into testing and non-testing sets.
5. **General Sifting**: If $(\widetilde{A}_i, \widetilde{B}_i, T_i) \in \mathbf{K}^C \times \{0\}$ where $\mathbf{K}^C$ is the complement of the set $\mathbf{K}$, Alice sets the $\overline{A}_i =\perp$. Denote

$$\mathcal{S} = \{i \in [N] : (\widetilde{A}_i, \widetilde{B}_i, T_i) \in \mathbf{K}^C \times \{0\}\}$$

and let $\overline{A}_{\mathcal{S}}$ denote the collection of registers that correspond to discarded events.
6. **Statistical Tests**:

   - If $T_i = 1$, Alice announces $\overline{A}_i$ publicly. Using this information, Bob generates X_i using a deterministic function t such that $X_i = t(\overline{A}_i, \overline{B}_i, \widetilde{A}_i, \widetilde{B}_i)$.
   - If $T_i = 0$, Bob sets $X_i =\perp$ and sets $\overline{B}_i =\perp$.

   Denote $\mathcal{T} = \{i \in [N] : T_i = 1\}$ and denote registers that Alice announces by $\overline{A}_{\mathcal{T}}$.
7. **Key Map**: If $(\widetilde{A}_i, \widetilde{B}_i, T_i) \in \mathbf{K} \times \{0\}$, Alice updates $\overline{A}_i := f_{\text{KM}}(\overline{A}_i, \widetilde{A}_i, \widetilde{B}_i)$ where f_{KM} is the key map.
8. **Parameter Estimation**: Alice and Bob abort the protocol if $\mathsf{freq}(X_1^N) \notin \mathcal{Q}$.

We remark that in the EAT subprotocol, for the security proof purpose, we set Bob's private register $\overline{B}_i$ to be the $\perp$ symbol when the round is used for key generation. This procedure is not executed in a real protocol.

8.2 Key length

We may use the EAT (Section 3.7) to bound the smooth min-entropy and then apply the leftover hashing lemma (Theorem 3.3.2) to bound the key length. Interestingly, we note a recent work [99] that proves a version of EAT to bound the sandwiched Rényi entropy of order $\alpha \in (1, 2]$ and also a version of leftover hashing lemma based on these sandwiched Rényi entropy directly to bound the key length. We present two theorems about the key length based on these two results.

We state conversions from registers in EAT statements in Section 3.7 to the QKD scenario as described in the EAT subprotocol in Protocol 8.1:

$$\begin{aligned} S_i &\leftrightarrow \overline{A}_i\overline{B}_i \\ P_i &\leftrightarrow \widetilde{A}_i\widetilde{B}_iT_i \\ X_i &\leftrightarrow X_i \\ Q_i &\leftrightarrow A_iB_i \\ R_i &\leftrightarrow R_i \\ E &\leftrightarrow E \ . \end{aligned}$$

With these conversions, we state the smooth min-entropy rate of the entropy accumulation subprotocol (Protocol 8.1).

Theorem 8.2.1. Consider the entropy accumulation protocol defined in Protocol 8.1 where the announcement structure satisfies the Markov chain condition. Let $\Omega = \mathcal{Q}$ and ρ be the output of the protocol. Let h such that $f(\boldsymbol{q}) \geq h$ for all $\boldsymbol{q} \in \mathcal{Q}$ where f is the min-tradeoff function. Then for any $\varepsilon_{\mathrm{EA}}, \bar{\varepsilon} \in (0,1)$, either the protocol aborts with probability greater than $1 - \varepsilon_{\mathrm{EA}}$, or

$$H^{\bar{\varepsilon}}_{\min}(\overline{A}_1^N\overline{B}_1^N|\widetilde{A}_1^N\widetilde{B}_1^NT_1^NE)_{\rho_{|\Omega}} > Nh - c\sqrt{N} - c' \ , \tag{8.1}$$

where c and c' are given by Eqs. (3.68) and (3.69), and we replace $\rho[\Omega]$ by $\varepsilon_{\mathrm{EA}}$.

Moreover, the bound can be further optimized by using Eq. (3.63) and optimizing over α. That is,

$$H^{\bar{\varepsilon}}_{\min}(\overline{A}_1^N\overline{B}_1^N|\widetilde{A}_1^N\widetilde{B}_1^NT_1^NE)_{\rho_{|\Omega}} > Nh - \delta_{\mathrm{EAT}} \ , \tag{8.2}$$

where

$$\delta_{\mathrm{EAT}} = \min_{\alpha\in(1,2)}\left[N\frac{(\alpha-1)\ln 2}{2}V^2 + \frac{1}{\alpha-1}\log_2\frac{2}{\bar{\varepsilon}^2\varepsilon_{\mathrm{EA}}^2} + N(\alpha-1)^2K_\alpha\right] \tag{8.3}$$

with V defined in Eq. (3.64) and K_α defined in Eq. (3.66) with the classical systems $S_i = \overline{A}_i\overline{B}_i$.

As we simply check that the requirements of EAT are satisfied in Protocol 8.1, the proof is trivial.

Remark 8.2.2. The statement of the theorem requires that either the protocol aborts with probability greater than $1 - \varepsilon_{\mathrm{EA}}$ or else the bound holds. In the language of Renner's Ph.D. thesis [89], this says either ρ is $\varepsilon_{\mathrm{EA}}$-securely filtered or the bound holds. This explains why parameter estimation goes away and how it is replaced with $\varepsilon_{\mathrm{EA}}$ in the statement of ε-security.

This leads to the following theorem for the key length if we use the second version of EAT (Theorem 3.7.4) [98].

Theorem 8.2.3. Consider any entanglement-based QKD protocol which has the testing structure as in the EAT subprotocol (Protocol 8.1). Let γ be the probability of testing. Let $h \leq f(\boldsymbol{q})$ for all $\boldsymbol{q} \in \mathcal{Q}$ where f is a min-tradeoff function. The QKD protocol is ε-secure for key length

$$\ell \leq Nh - \text{leak}_{\text{EC}} - N\gamma \log_2(|\mathcal{A}||\mathcal{B}|) - \sqrt{N}\delta - \kappa - 2\log_2(2/\varepsilon_{\text{PA}}) \tag{8.4}$$

where $\mathcal{A}$, $\mathcal{B}$ are the alphabets of private outcomes Alice and Bob announce in the statistical tests excluding the symbol $\perp$, and

$$\begin{aligned} \delta &= 2\log_2(1+2d_S)\sqrt{1-2\log_2(\bar{\varepsilon}/4\cdot(\varepsilon_{\text{EA}}+\varepsilon_{\text{EC}}))} + \tilde{c} \\ \kappa &= 2\log_2\left(1-\sqrt{1-\bar{\varepsilon}^2/16}\right) + \tilde{c}'\ , \end{aligned}$$

where $\tilde{c}$ is c in Eq. (3.68) with $\varepsilon = \bar{\varepsilon}/4$, $\tilde{c}'$ is c' in Eq. (3.69) with $\varepsilon = \bar{\varepsilon}/4$, and d_S is the maximum dimension of $S_i = \overline{A}_i\overline{B}_i$. In particular, $\varepsilon \leq \varepsilon_{\text{EA}}+\bar{\varepsilon}+\varepsilon_{\text{EC}}+\varepsilon_{\text{PA}}$. In other words, either the protocol aborts with a probability greater than $1-\varepsilon_{\text{EA}}-\varepsilon_{\text{EC}}$ or the key is $(\varepsilon_{\text{EC}}+\bar{\varepsilon}+\varepsilon_{\text{PA}})$-correct-and-secure.

Moreover, this result can be further optimized for each specific block-length by optimizing α in Theorem 3.7.4 instead of the fixed α value that leads to c and c'. That is,

$$\begin{aligned} \ell \leq & Nh - \text{leak}_{\text{EC}} - N\gamma\log_2(|\mathcal{A}||\mathcal{B}|) - 2\log_2(2/\varepsilon_{\text{PA}}) - \delta_{\text{EAT}} \\ & - 2\sqrt{N}\log_2(1+2d_S)\sqrt{1-2\log_2(\bar{\varepsilon}/4\cdot(\varepsilon_{\text{EA}}+\varepsilon_{\text{EC}}))} - 2\log_2\left(1-\sqrt{1-\bar{\varepsilon}^2/16}\right) \end{aligned} \tag{8.5}$$

where δ_{EAT} is given in Eq. (8.3) with $\bar{\varepsilon}$ replaced by $\bar{\varepsilon}/4$.

Proof. See Appendix D.2. □

If we use the result for the sandwiched Rényi entropy and the corresponding leftover hashing lemma [99], one can show the following:

Theorem 8.2.4. Consider any entanglement-based QKD protocol which has the testing structure as in the EAT subprotocol (Protocol 8.1). Let γ be the probability of testing. Let $h \leq f(\boldsymbol{q})$ for all $\boldsymbol{q} \in \mathcal{Q}$ where f is a min-tradeoff function. Let $\beta \in (1,2), \delta \in (1/2,1)$, and $\alpha = \frac{-\beta+\delta}{-1+2\delta-\beta\delta}$. The QKD protocol is ε-secure for key length

$$\begin{aligned}\ell \leq\ & Nh - \mathrm{leak}_{\mathrm{EC}} - N\gamma \log_2(|\mathcal{A}||\mathcal{B}|) - N\frac{(\beta-1)\ln 2}{2}V^2 - \frac{\beta}{\beta-1}\log_2\frac{1}{\varepsilon_{\mathrm{EA}}} - N(\beta-1)^2 K_\beta \\ & - \frac{\delta}{1-\delta}\log_2(\frac{1}{\varepsilon_{\mathrm{EA}}+\varepsilon_{\mathrm{EC}}}) - \frac{\alpha}{\alpha-1}\log_2(\frac{1}{\varepsilon_{\mathrm{sec}}}) + 1\end{aligned} \tag{8.6}$$

where $\mathcal{A}$, $\mathcal{B}$ are the alphabets of private outcomes Alice and Bob announce in the statistical tests excluding the symbol $\perp$, d_S is the maximum dimension of $S_i = \overline{A}_i\overline{B}_i$, and

$$\begin{aligned}V &= \sqrt{\mathsf{Var}(f)+2} + \log_2(2d_S^2+1), \\ K_\beta &= \frac{1}{6(2-\beta)^3\ln 2}2^{(\beta-1)[\log_2 d_S + \mathsf{Max}(f) - \mathsf{Min}_\Sigma(f)]}\ln^3\left(2^{\log_2 d_S + \mathsf{Max}(f) - \mathsf{Min}_\Sigma(f)} + e^2\right).\end{aligned} \tag{8.7}$$

In particular, $\varepsilon \leq \varepsilon_{\mathrm{EA}} + \varepsilon_{\mathrm{EC}} + \varepsilon_{\mathrm{sec}}$. In other words, either the protocol aborts with a probability greater than $1 - \varepsilon_{\mathrm{EA}} - \varepsilon_{\mathrm{EC}}$ or the key is $(\varepsilon_{\mathrm{EC}} + \varepsilon_{\mathrm{sec}})$-correct-and-secure.

This proof is similar to the proof of Theorem 8.2.3 by adopting results from Ref. [99] instead of Theorem 3.7.4 [98]. We direct readers to [83] for a full derivation. As we see later, this result gives tighter key rate for small block sizes.

Remark 8.2.5. By comparing Theorem 8.2.4 with Theorem 8.2.3, we note that $\varepsilon_{\mathrm{PA}}$ and $\bar{\varepsilon}$ are not introduced in Theorem 8.2.4. Instead they are replaced by a single security parameter $\varepsilon_{\mathrm{sec}}$. The reason that $\varepsilon_{\mathrm{PA}}$ is used previously is explained after Theorem 3.3.2, the leftover hashing lemma. There are two contributing terms to the secrecy trace distance bound in Eq. (3.8). One of the terms is the smoothing parameter $\bar{\varepsilon}$ as we use the smooth min-entropy to relate the length of secret keys with the trace distance bound. We then set the other term as $\varepsilon_{\mathrm{PA}}$ so that $\varepsilon_{\mathrm{sec}} = \bar{\varepsilon} + \varepsilon_{\mathrm{PA}}$. This separation follows the convention in literature. In contrast, Theorem 8.2.4 uses a version of leftover hashing lemma based on the sandwiched Rényi entropy [99]. Without the smoothing parameter $\bar{\varepsilon}$, we do not divide $\varepsilon_{\mathrm{sec}}$ to more terms here.

8.3 Overview of examples

One simple way to guarantee the Markov chain condition [see Eq. (3.55)] is to use seeded randomness for all public information. That is, each register in the compound register P_i is generated by a random number. All examples in this chapter use the seeded randomness so that the Markov chain condition is trivially satisfied.

We apply the two numerical algorithms described in Section 4.4 to calculate finite-key rates. For the construction of min-tradeoff function, we use the crossover min-tradeoff function (Definition 3.7.5) since it gives higher key rates than the standard min-tradeoff function (Definition 3.7.2). We briefly compare these two algorithms in the qubit BB84 protocol example (Section 8.4) and the six-state four-state protocol example (Section 8.5) [209]. We also use these two examples to compare key rates obtained from Theorems 8.2.3 and 8.2.4.

For all examples considered here, we choose the overall security parameter to be $\varepsilon = 10^{-8}$. For simplicity of calculation, we set all the contributors to the security parameter ε as $\bar{\varepsilon} = \varepsilon_{\text{EA}} = \varepsilon_{\text{EC}} = \varepsilon_{\text{PA}} = 2.5 \times 10^{-9}$ when applying Theorem 8.2.3. In the case of Theorem 8.2.4, we set $\varepsilon_{\text{sec}} = 5 \times 10^{-9}$ while keeping $\varepsilon_{\text{EA}} = \varepsilon_{\text{EC}} = 2.5 \times 10^{-9}$. We note that in principle these parameters can be further optimized although we do not expect the results would be changed significantly by such an optimization. To illustrate concepts, we consider the idea of the unique acceptance for simplicity, that is, $\mathcal{Q} = \{\boldsymbol{F}\}$ for some frequency distribution $\boldsymbol{F} \in \mathbb{P}(\mathcal{X})$. Our implementation allows for the consideration of a more general acceptance set $\mathcal{Q}$ in the form of Eq. (3.21). If the acceptance threshold t in the definition of $\mathcal{Q}$ is small, results in that case would be very similar to results reported here with $t = 0$.

8.4 BB84 with qubit systems

We consider a simple entanglement-based BB84 protocol with qubit systems. In this example, we assume that Alice and Bob each receive a qubit in each round of the protocol and they perform Z-basis or X-basis measurements. We note that this assumption is not practical given current state-of-the-art in experimental implementations. To guarantee this assumption, one may require that the source emits single photons, detectors are photon-number resolving and there is no photon loss in the process or alternatively there is a heralding process. Nevertheless, this example is still interesting to discuss in detail for theoretical purposes.

8.4.1 Protocol details and simulation method

We consider one particular setup for this protocol: (1) Alice chooses the Z basis with a probability p_z and the X basis with a probability $1-p_z$, Bob chooses to measure the Z basis with a probability p_z and X basis with a probability $1 - p_z$. (2) When they both choose Z basis, the round is used for key generation; When they both choose X basis, the round is used for parameter estimation (testing). When they choose different bases, the round is discarded. (3) They perform parameter estimation before error correction. (4) For parameter estimation, we use the phase error POVM $\{E_X, \mathbb{1} - E_X\}$ where E_X is the X-basis error operator. This corresponds to statistics $\{e_x, 1 - e_x\}$ where e_x is the X-basis error rate.

We note that the testing probability γ is given by the probability that both Alice and Bob choose the X-basis, that is, $\gamma = (1-p_z)^2$. The sifting factor for the key rate is p_z^2. We consider an efficient version of BB84 [163] by allowing p_z to be close to 1. This also corresponds to an infrequent testing.

In our simulation, we use the simple depolarizing channel model to model noise. The simulated state is

$$\rho^{\text{sim}} = (1 - \frac{3}{2}Q) \left|\Phi^+\middle\rangle\!\middle\langle\Phi^+\right| + \frac{Q}{2}(\left|\Phi^-\middle\rangle\!\middle\langle\Phi^-\right| + \left|\Psi^+\middle\rangle\!\middle\langle\Psi^+\right| + \left|\Psi^-\middle\rangle\!\middle\langle\Psi^-\right|), \tag{8.8}$$

where $|\Phi^+\rangle, |\Phi^-\rangle, |\Psi^+\rangle$ and $|\Psi^-\rangle$ are Bell states, and Q is the error rate.

8.4.2 Results

For Algorithm 4.2, we optimize the min-tradeoff functions by choosing different $\boldsymbol{q_0} = (Q, 1-Q)$ where Q is searched over the interval $[0.005, 0.07]$ with a step size 0.005. For each min-tradeoff function generated from a particular value of Q, we calculate the key rate, which is the key length ℓ divided by the total number of signals N. We then choose the maximum key rate among all possible choices of min-tradeoff functions generated in this way.

For Algorithm 4.3, we use the interior-point method from MATLAB fmincon function for the first step and then use CVX for the linearized dual problem [Eq. (4.88)]. We note that we can also use the Frank-Wolfe algorithm in a similar way as the asymptotic key rate calculation (Algorithm 4.1).

For both algorithms, we optimize p_z by optimizing $\gamma = (1-p_z)^2 = 10^{-k}$ where k is chosen from the interval $[2, 4]$ with a step size of 0.1. For block size larger than or equal to 10^10, we allow p_z to be more closer to 1 by searching k in the interval $[3, 7]$ with a step size of 0.2. We note that we optimize the choice of α in the statement of Theorem 8.2.3 when we use this theorem, and similarly we optimize both β and δ in the statement of Theorem 8.2.4.

In Figure 8.1, we compare the key rates obtained from these two algorithms with Theorem 8.2.4. Interestingly, both algorithms give similar results while Algorithm 4.3 seems to be slightly better in terms of the smallest number of signals for nonzero key rates. Intuitively, we expect Algorithm 4.3 to behave better as it takes into account some second-order correction terms, while Algorithm 4.2 only looks for the min-tradeoff function that gives the highest leading-order term. As we perform an optimization of the choice of min-tradeoff function by different initial $\boldsymbol{q_0}$'s for Algorithm 4.2, we observe that the optimal finite key rate from Algorithm 4.2 is often given by a min-tradeoff function that does not give the highest value for the leading-order term. For this protocol example, this optimization of $\boldsymbol{q_0}$ seems effective even though we search only a single parameter space instead of the full parameter space $\mathbb{P}(\mathcal{X})$ for the initial point $\boldsymbol{q_0}$. Due to the optimization of $\boldsymbol{q_0}$, the running time of Algorithm 4.2 is much slower than that of Algorithm 4.3.

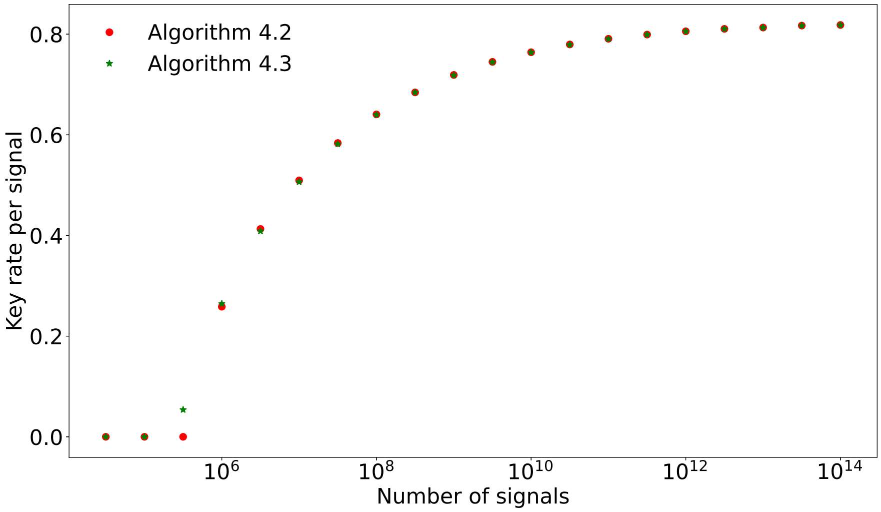


Figure 8.1: Key rate versus the number of signals for the qubit BB84 protocol with the error rate $Q = 0.01$ and with different algorithms for the generation of min-tradeoff functions. The red circle marker corresponds to Algorithm 4.2 while the green star marker corresponds to Algorithm 4.3. Second-order correction terms are based on Theorem 8.2.4. Other protocol parameters are optimized as described in the main text.

In Figure 8.2, we compare key rates given by Theorems 8.2.3 and 8.2.4 when we use Algorithm 4.3. One can see that Theorem 8.2.4 gives better key rates. This confirms our expectation that the EAT based on the sandwiched α-Rényi entropy is tighter than the EAT based on the smooth min-entropy for lower-order correction terms.

8.5 Six-state four-state protocol

In the free space implementation of QKD protocols with polarization encoding, there is naturally one axis that is stable against turbulence while other axes are slowly drifting. The idea of reference-frame-independent [210] was motivated to address this issue and it was shown that such a protocol can be robust to slow drifts. We consider the six-state four-state protocol [209] which has the reference-frame-independent feature and is simpler in implementation.

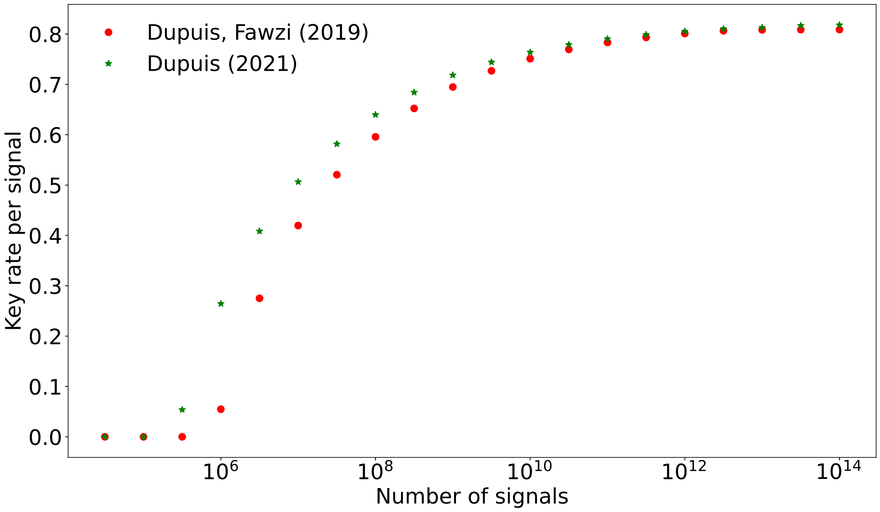


Figure 8.2: Key rate versus the number of signals for the qubit BB84 protocol with the error rate $Q = 0.01$ and with different second-order correction terms. The red circle marker corresponds to Theorem 8.2.3 with the second-order correction terms based on Ref. [98] while the green star marker corresponds to Theorem 8.2.4 with the second-order terms from Ref. [99]. Other protocol parameters are optimized as described in the main text.

8.5.1 Protocol details and simulation

We analyze the entanglement-based version of the six-state four-state protocol [209] and assume that Alice and Bob each receive a qubit in each round for simplicity. In this protocol, Alice measures the state in one of the X, Y and Z bases according to the probability distribution $((1 - p_z)/2, (1 - p_z)/2, p_z)$, while Bob measures in one of X and Z bases with the probability distribution $(1 - p_z, p_z)$. Similarly to the previous qubit BB84 example, when both Alice and Bob choose the Z-basis, this round is used for key generation. When Alice chooses X or Y basis and Bob chooses X basis, this round is used for parameter estimation. We consider an efficient version by allowing p_z to be close to 1.

For the simulation, we assume the Z basis is free of misalignment. The misalignment happens in the X-Y plane of the Bloch sphere. Thus, on top of the qubit depolarizing channel, we also apply a unitary rotation along the Z axis to Bob's qubit in order to model the misalignment. We choose the angle of rotation to be 11° in the simulation. In our security analysis, we use the error rates in the XX and YX bases as constraints.

8.5.2 Results

For Algorithm 4.2, we optimize the min-tradeoff functions by choosing different $\boldsymbol{q_0}$'s. Each $\boldsymbol{q_0}$ is created by choosing a different depolarizing probability Q, which is searched over the interval $[0.005, 0.07]$ with a step size 0.005. We fix the rotation angle along the Z-axis to be 11°. For each min-tradeoff function generated from a particular value of Q, we calculate the key rate and then choose the maximum key rate among them. For Algorithm 4.3, we use same procedure as for the qubit BB84 example in Section 8.4 including the optimization over the choice of p_z.

In Figure 8.3, we compare key rates for Theorems 8.2.3 and 8.2.4 and observe the same behavior as the qubit BB84 example in Figure 8.1. This behavior is not surprising since the sandwiched Rényi entropy was used in the middle step of the proof of Theorems 3.7.3 and 3.7.4 in Refs. [97, 98] before converting to the smooth min-entropy by an additional inequality. Bypassing the smooth min-entropy is expected to give tighter key rates [99].

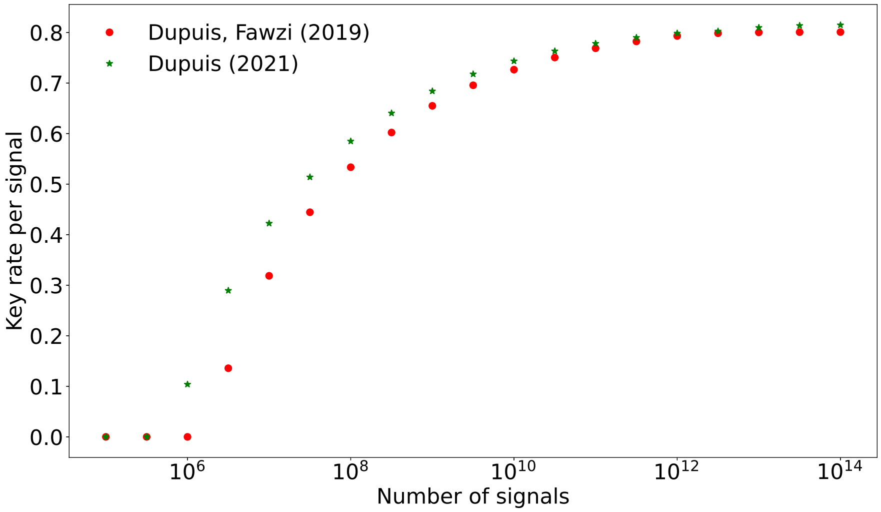


Figure 8.3: Comparison of two key rate expressions: key rate versus the number of signals for the six-state four-state protocol [209] with depolarizing probability $Q = 0.01$. The red circle marker corresponds to Theorem 8.2.3 [98], the green star marker corresponds to Theorem 8.2.4 [99]. Other protocol parameters are optimized as described in the main text.

In Figure 8.4, we compare Algorithms 4.2 and 4.3 for the six-state four-state protocol [209]. Similar to the behavior of the qubit BB84 example, we observe both algorithms give similar results while Algorithm 4.3 gives higher key rates for small block sizes. We note that each tradeoff function given by Algorithm 4.2 does not take into account second-order correction terms during the construction. Instead, we optimize the second-order correction terms afterwards by considering a bunch of different min-tradeoff functions when using this algorithm. The latter optimization is not done in an efficient way since we apply only a simple coarse-grained search over a single parameter instead of a more sophisticated algorithm over the full parameter space of $\mathbb{P}(\mathcal{X})$. In contrast, each min-tradeoff function generated by Algorithm 4.3 uses some information about second-order correction terms. We do not need to perform an additional optimization of different min-tradeoff functions for each fixed set of protocol parameters. However, as explained in Section 4.4.2, we do not use the optimal second-order correction terms when we construct min-tradeoff functions by Algorithm 4.3. After we obtain a fixed min-tradeoff function, we then optimize the second-order correction terms by optimizing free parameters in Theorem 8.2.3 or Theorem 8.2.4. Thus, in principle, Algorithm 4.3 can be further improved. In terms of running time, Algorithm 4.3 takes less time as there is no need to optimize $\boldsymbol{q_0}$.

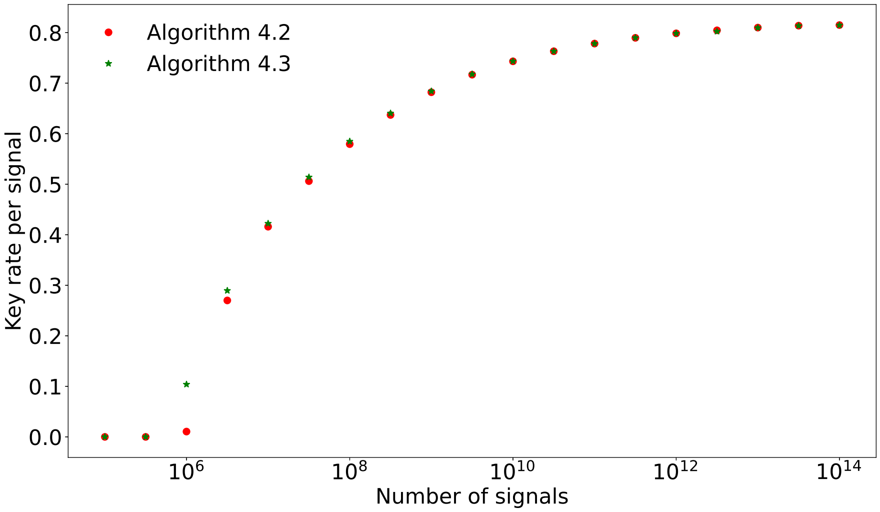


Figure 8.4: Comparison of two algorithms: key rate versus the number of signals for the six-state four-state protocol [209] with the depolarizing probability $Q = 0.01$. The red circle marker corresponds to Algorithm 4.2, the green star marker corresponds to Algorithm 4.3. Second-order correction terms are from Theorem 8.2.4. Other protocol parameters are optimized as described in the main text.

8.6 High-dimensional 2-MUB

Another interesting family of protocols that can be analyzed by our numerical method in the EAT framework are the high-dimensional analogues of the BB84 protocol. In BB84, two mutually unbiased bases (MUBs) are used. While two MUBs exist for any dimension, we consider 2-MUB protocols in prime dimensions only in this section. Our restriction here is mainly due to the particular simulation method adopted from [211], not because of the limitations in our proof technique.

8.6.1 Protocol details and simulation

Recall that the Weyl operators are defined as

$$U_{jk} = \sum_{s=0}^{d-1} \omega^{sk} |s+j\rangle\langle s| \tag{8.9}$$

for $j, k \in \{0, 1, \dots, d-1\}$ where $\omega = e^{2\pi i/d}$ is a dth root of unity. We note that U_{01} is the generalized Pauli-Z matrix and U_{10} is the generalized Pauli-X matrix. We define the qudit version of Z and X operators as $Z := U_{01}$ and $X := U_{10}$. The generalized Bell states are

$$|\Phi_{jk}\rangle = \frac{1}{\sqrt{d}} \sum_{s=0}^{d-1} \omega^{sk} |s, s+j\rangle = \mathbb{1} \otimes U_{jk} |\Phi_{00}\rangle . \tag{8.10}$$

In the 2-MUB protocol, Alice measures in the eigenbasis of either U_{01} or U_{10}. Bob similarly measure in the eigenbasis of either U_{01}^* or U_{10}^*. In the classical phase, Alice and Bob announce their basis choices and discard rounds with mismatched bases. We allow asymmetric basis choice, i.e., with a probability p_z to choose Z basis and a probability $1 - p_z$ to choose X basis. The eigenbasis of the operator $Z := U_{01}$ is the computational basis $\{|s\rangle : 0 \le s \le d-1\}$. The eigenbasis of the operator $X := U_{10}$ is $\{\left|\psi_j^X\right\rangle : 0 \le j \le d-1\}$ where

$$\left|\psi_j^X\right\rangle := \sum_{s=0}^{d-1} \frac{\omega^{-js}}{\sqrt{d}} |s\rangle . \tag{8.11}$$

In this example, we also consider an efficient protocol by setting p_z to be close to 1. For simplicity of calculation, the testing rounds are those when they both choose X basis and the key generation rounds are rounds when they both choose Z basis.

It is interesting to note that the state $|\Phi_{00}\rangle = \frac{1}{\sqrt{d}} \sum_s |s, s\rangle$ is invariant under any $U_{jk} \otimes U_{jk}^*$. In an honest implementation, the source is supposed to prepare the state $|\Phi_{00}\rangle$, and to distribute one half to Alice and the other half to Bob. If the channel is ideal, then they are supposed to obtain perfectly correlated results just like the qubit case.

We follow the simulation in [211] by considering the following observation for the error vector in each basis U_{jk}, which is based on the natural generalization of the qubit depolarizing channel:

$$\boldsymbol{q_{jk}}(Q) := \{1 - Q, Q/(d-1), \dots, Q/(d-1)\}. \tag{8.12}$$

The simulation state ρ_{AB}^{sim} used to generate full statistics is

$$\rho_{AB}^{\text{sim}} = \sum_{j,k} \lambda_{jk} \left|\Phi_{jk}\right\rangle\!\left\langle\Phi_{jk}\right| , \tag{8.13}$$

where

$$\lambda_{jk} = \frac{1}{d}\Big(\sum_{s} q_{1s}^{(sj-k \bmod d)} + q_{01}^{(j)} - 1\Big) \tag{8.14}$$

with $q_{jk}^{(i)}$ being the ith entry of the vector $\boldsymbol{q_{jk}}$.

In our security analysis, we use the X basis error rate and X basis correlation as constraints.

8.6.2 Results

We apply Algorithm 4.3 with the Frank-Wolfe method to calculate the key rate for this protocol example. We optimize the probability of choosing Z-basis in the same way as in the qubit BB84 example. In Figure 8.5, we compare key rates for d-dimensional 2-MUB protocols with $d = 2, 3, 5$ and 7 using Theorem 8.2.4. We set the depolarizing probability Q to be 0.01 for all curves. This example demonstrates that our method can work for high-dimensional protocols.

We remark that Ref. [211] claims their analysis is against coherent attacks. However, they only performed an analysis against collective attacks and did not use any technique like quantum de Finetti representation theorem or postselection technique. Instead, they argued that by following the same arguments from [124, 125] for qubit protocols based on their high symmetries, their analysis is also for coherent attacks for qudit protocols. Unfortunately, these papers [124, 125] only claim asymptotically collective attacks are optimal. They did not provide detailed analysis for the finite-key scenario. Thus, we believe one still needs to lift the analysis in [211] to coherent attacks via either quantum de Finetti representation or postselection technique.

8.7 Future directions

We would like to apply our method to study more complicated protocol examples. In particular, we would like to investigate an optical implementation of the entanglement-based BB84 since it involves announcements about detection which are not based on seeded randomness. Therefore, we need to check the Markov chain condition more explicitly. We will report this result in our forthcoming paper [83]. We also hope to include more direct comparisons with other proof methods. When the dimension is large, we expect the key rate obtained from the EAT is better than the one from the postselection technique as the postselection technique has a strong dependence on the dimension while the EAT does not. Another interesting direction is to extend our method

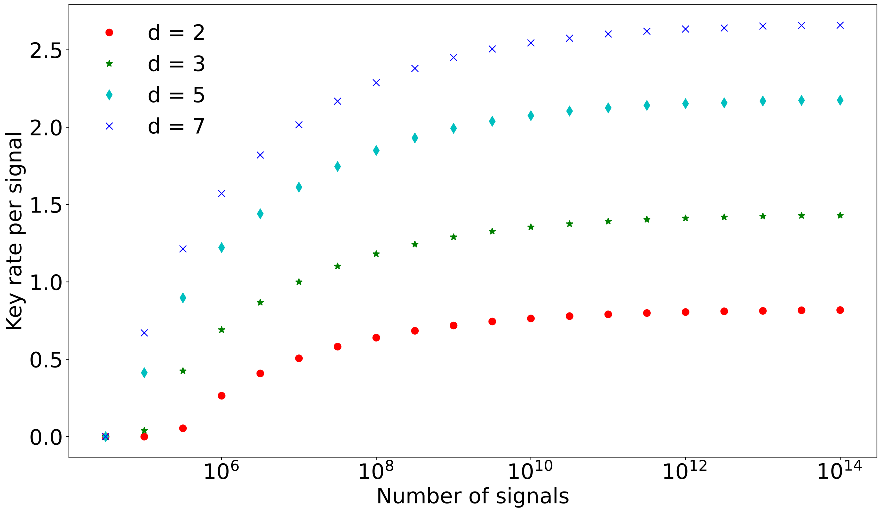


Figure 8.5: 2-MUB protocols with qudits for various values of dimension d with the depolarizing probability $Q = 0.01$. The red circle marker corresponds to $d = 2$ (BB84), the green star marker corresponds to $d = 3$, the cyan diamond marker corresponds to $d = 5$ and the blue cross marker is for $d = 7$. Second-order correction terms are from Theorem 8.2.4. Other protocol parameters are optimized as explained in the main text.

to prepare-and-measure protocols, which we currently cannot directly handle the ρ_A constraint. Similar to the discussion about the postselection technique in Section 3.6, one may ignore this promise or introduce an additional test for ρ_A. However, in either way, we do not expect good key rates. We need to solve technical challenges before we can directly handle this constraint. We leave it for future work.

Chapter 9

Concluding remarks

In this thesis, we focus on security proof methods in Renner's framework [89]. One advantage of this framework is that security proofs can be modular. In this framework, one may first prove the security against i.i.d. collective attacks which are more restricted, and then lift the proof to general attacks by several existing techniques: quantum de Finetti representation theorem [89, 111], postselection technique [112, 120] or entropy accumulation theorem [97–99]. Under the assumption of i.i.d. collective attacks, one can greatly simplify the proof by studying a single-round system as the N-round system is assumed to be an i.i.d. state. Both asymptotic [144, 146] and finite-size key rate [84] calculation problems have been formulated as convex optimization problems. We discuss a numerical framework to calculate key rates. It includes numerical methods to calculate both asymptotic and finite-size key rates. We also review several existing tools that help extend the ranges of protocols that one can study using these numerical methods. This numerical framework is highly versatile and can produce tight key rates. Another advantage of using this numerical framework is that one may perform the security analysis of multiple side channels simultaneously. This can be done by having suitable models of devices that include those side channels and then deriving suitable mathematical descriptions for those devices. Numerical methods also have their limitations due to the scaling of the computational resources with the problem size. We hope more analytical tools and advanced numerical algorithms can be developed to enhance the power of this numerical framework for key rate calculation in the future.

As applications, we calculate asymptotic key rates of the phase-matching MDIQKD protocol [71], a variant of TF-QKD [62, 63] and of the discrete-modulated continuous-variable QKD [81, 82]. While our specific phase-matching MDI protocol is idealized, we confirm the $\sqrt{\eta}$ loss scaling of this protocol. A realistic version of this protocol with finite test states has been shown to overcome the repeaterless bound as well in Ref. [93]. For the DMCVQKD protocols, we apply the numerical method to calculate asymptotic key rates in both untrusted (ideal) [81] and trusted detector noise scenarios [82]. Our initial works [81, 82] use an additional assumption - photon-number cutoff assumption. This assumption allows us to truncate an infinite-dimensional Hilbert space, which makes the numerical calculation possible. Our initial work verifies the validity of this assumption numerically, meaning that the results do not change significantly with the choice of cutoff when the cutoff photon number is sufficiently large (i.e. 10 in our simulations for the ideal detector scenario) compared to the mean photon number of the states received by Bob. While this assumption gives numerically plausible key rates, it lacks rigorous justifications that one desires for a security proof. In our recent work [85], we remove this assumption by a dimension reduction method. We are able to show that key rates obtained without the photon-number cutoff assumption do not deviate much from those obtained previously under the photon-number cutoff assumption. All of these works focus on the asymptotic key rates. It is important for future works to study finite-size effects of those protocols.

In terms of finite-key analysis, we discuss two numerical methods [83, 84]. We present preliminary results to illustrate the numerical method [83] with the entropy accumulation theorem by simple protocol examples. We hope to investigate more complicated examples in the near future.

References

[1] P. W. Shor, *Algorithms for Quantum Computation: Discrete Logarithms and Factoring*, in *Proceedings 35th annual symposium on foundations of computer science* (IEEE, 1994) pp. 124–134.

[2] S. Wiesner, *Conjugate coding*, ACM Sigact News **15**, 78 (1983).

[3] C. H. Bennett and G. Brassard, *Quantum Cryptography: Public Key Distribution and Coin Tossing*, in *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing* (IEEE, New York, 1984) pp. 175–179.

[4] A. K. Ekert, *Quantum Cryptography based on Bell's Theorem*, Phys. Rev. Lett. **67**, 661 (1991).

[5] C. H. Bennett, G. Brassard, and N. D. Mermin, *Quantum Cryptography without Bell's Theorem*, Phys. Rev. Lett. **68**, 557 (1992).

[6] H. Buhrman, R. Cleve, J. Watrous, and R. De Wolf, *Quantum Fingerprinting*, Phys. Rev. Lett. **87**, 167902 (2001).

[7] J. M. Arrazola and N. Lütkenhaus, *Quantum Fingerprinting with Coherent States and a Constant Mean Number of Photons*, Phys. Rev. A **89**, 062305 (2014).

[8] F. Xu, J. M. Arrazola, K. Wei, W. Wang, P. Palacios-Avila, C. Feng, S. Sajeed, N. Lütkenhaus, and H.-K. Lo, *Experimental Quantum Fingerprinting with Weak Coherent Pulses*, Nat. commun. **6**, 8735 (2015).

[9] J.-Y. Guan, F. Xu, H.-L. Yin, Y. Li, W.-J. Zhang, S.-J. Chen, X.-Y. Yang, L. Li, L.-X. You, T.-Y. Chen, Z. Wang, Q. Zhang, and J.-W. Pan, *Observation of Quantum Fingerprinting Beating the Classical Limit*, Phys. Rev. Lett. **116**, 240502 (2016).

[10] B. Lovitz and N. Lütkenhaus, *Families of Quantum Fingerprinting Protocols*, Phys. Rev. A **97**, 032340 (2018).

[11] P. Arrighi and L. Salvail, *Blind Quantum Computation*, International Journal of Quantum Information **4**, 883 (2006).

[12] A. Broadbent, J. Fitzsimons, and E. Kashefi, *Universal Blind Quantum Computation*, in *Proceedings of the 50th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2009)* (IEEE, 2009) pp. 517–526, arXiv:0807.4154 .

[13] S. Barz, E. Kashefi, A. Broadbent, J. F. Fitzsimons, A. Zeilinger, and P. Walther, *Demonstration of Blind Quantum Computing*, Science **335**, 303 (2012).

[14] V. Giovannetti, L. Maccone, T. Morimae, and T. G. Rudolph, *Efficient Universal Blind Quantum Computation*, Phys. Rev. Lett. **111**, 230501 (2013).

[15] A. Mantri, C. A. Pérez-Delgado, and J. F. Fitzsimons, *Optimal Blind Quantum Computation*, Phys.l Rev. Lett. **111**, 230502 (2013).

[16] D. Gottesman and I. Chuang, *Quantum Digital Signatures*, arXiv:quant-ph/0105032 (2001).

[17] P. Wallden, V. Dunjko, A. Kent, and E. Andersson, *Quantum Digital Signatures with Quantum-Key-Distribution Components*, Phys. Rev. A **91**, 042304 (2015).

[18] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, *Quantum Cryptography*, Rev. Mod. Phys. **74**, 145 (2002).

[19] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, *The Security of Practical Quantum Key Distribution*, Rev. Mod. Phys. **81**, 1301 (2009).

[20] E. Diamanti and A. Leverrier, *Distributing Secret Keys with Quantum Continuous Variables: Principle, Security and Implementations*, Entropy **17**, 6072 (2015).

[21] E. Diamanti, H.-K. Lo, B. Qi, and Z. Yuan, *Practical Challenges in Quantum Key Distribution*, npj Quantum Inf **2**, 16025 (2016).

[22] F. Xu, X. Ma, Q. Zhang, H.-K. Lo, and J.-W. Pan, *Secure Quantum Key Distribution with Realistic Devices*, Rev. Mod. Phys. **92**, 025002 (2020).

[23] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani, J. Pereira, M. Razavi, J. S. Shaari, M. Tomamichel, V. C. Usenko, G. Vallone, P. Villoresi, and P. Wallden, *Advances in Quantum Cryptography*, Adv. Opt. Photon. **12**, 1012 (2020).

[24] F. Grosshans and P. Grangier, *Continuous Variable Quantum Cryptography Using Coherent States*, Phys. Rev. Lett. **88**, 057902 (2002).

[25] F. Grosshans, G. Van Assche, J. Wenger, R. Brouri, N. J. Cerf, and P. Grangier, *Quantum Key Distribution Using Gaussian-Modulated Coherent States*, Nature **421**, 238 (2003).

[26] C. Weedbrook, A. M. Lance, W. P. Bowen, T. Symul, T. C. Ralph, and P. K. Lam, *Quantum Cryptography without Switching*, Phys. Rev. Lett. **93**, 170504 (2004).

[27] A. Boaron, G. Boso, D. Rusca, C. Vulliez, C. Autebert, M. Caloz, M. Perrenoud, G. Gras, F. Bussières, M.-J. Li, *et al.*, *Secure Quantum Key Distribution over 421 km of Optical Fiber*, Phys. Rev. Lett. **121**, 190502 (2018).

[28] Y. Zhang, Z. Chen, S. Pirandola, X. Wang, C. Zhou, B. Chu, Y. Zhao, B. Xu, S. Yu, and H. Guo, *Long-Distance Continuous-Variable Quantum Key Distribution over 202.81 km of Fiber*, Phys. Rev. Lett. **125**, 010502 (2020).

[29] J. Lodewyck, M. Bloch, R. García-Patrón, S. Fossier, E. Karpov, E. Diamanti, T. Debuisschert, N. J. Cerf, R. Tualle-Brouri, S. W. McLaughlin, and P. Grangier, *Quantum Key Distribution over 25 km with an All-Fiber Continuous-Variable System*, Phys. Rev. A **76**, 042305 (2007).

[30] P. Jouguet, S. Kunz-Jacques, A. Leverrier, P. Grangier, and E. Diamanti, *Experimental Demonstration of Long-Distance Continuous-Variable Quantum Key Distribution*, Nat. Photonics **7**, 378 (2013).

[31] B. Qi, P. Lougovski, R. Pooser, W. Grice, and M. Bobrek, *Generating the Local Oscillator "Locally" in Continuous-Variable Quantum Key Distribution based on Coherent Detection*, Phys. Rev. X **5**, 041009 (2015).

[32] D. B. S. Soh, C. Brif, P. J. Coles, N. Lütkenhaus, R. M. Camacho, J. Urayama, and M. Sarovar, *Self-Referenced Continuous-Variable Quantum Key Distribution Protocol*, Phys. Rev. X **5**, 041010 (2015).

[33] D. Huang, P. Huang, D. Lin, C. Wang, and G. Zeng, *High-Speed Continuous-Variable Quantum Key Distribution without Sending a Local Oscillator*, Opt. Lett. **40**, 3695 (2015).

[34] D. Huang, P. Huang, D. Lin, and G. Zeng, *Long-Distance Continuous-Variable Quantum Key Distribution by Controlling Excess Noise*, Sci. Rep. **6**, 19201 (2016).

[35] C. Wittmann, J. Fürst, C. Wiechers, D. Elser, H. Häseler, N. Lütkenhaus, and G. Leuchs, *Witnessing Effective Entanglement over a 2 km Fiber Channel*, Opt. Express **18**, 4499 (2010).

[36] X.-Y. Wang, Z.-L. Bai, S.-F. Wang, Y.-M. Li, and K.-C. Peng, *Four-state Modulation Continuous Variable Quantum Key Distribution over a 30-km Fiber and Analysis of Excess Noise*, Chin. Phys. Lett. **30**, 010305 (2013).

[37] B. Heim, C. Peuntinger, N. Killoran, I. Khan, C. Wittmann, C. Marquardt, and G. Leuchs, *Atmospheric Continuous-Variable Quantum Communication*, New J. Phys. **16**, 113018 (2014).

[38] T. Hirano, T. Ichikawa, T. Matsubara, M. Ono, Y. Oguri, R. Namiki, K. Kasai, R. Matsumoto, and T. Tsurumaru, *Implementation of Continuous-Variable Quantum Key Distribution with Discrete Modulation*, Quantum Sci. Technol. **2**, 024010 (2017).

[39] F. Laudenbach, B. Schrenk, C. Pacher, M. Hentschel, C.-H. F. Fung, F. Karinou, A. Poppe, M. Peev, and H. Hübel, *Pilot-Assisted Intradyne Reception for High-Speed Continuous-Variable Quantum Key Distribution with True Local Oscillator*, Quantum **3**, 193 (2019).

[40] J. Barrett, L. Hardy, and A. Kent, *No Signaling and Quantum Key Distribution*, Phys. Rev. Lett. **95**, 010503 (2005).

[41] A. Acín, N. Gisin, and L. Masanes, *From Bell's Theorem to Secure Quantum Key Distribution*, Phys. Rev. Lett. **97**, 120405 (2006).

[42] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, *Device-Independent Security of Quantum Cryptography against Collective Attacks*, Phys. Rev. Lett. **98**, 230501 (2007).

[43] R. Arnon-Friedman, F. Dupuis, O. Fawzi, R. Renner, and T. Vidick, *Practical Device-Independent Quantum Cryptography via Entropy Accumulation*, Nat. Commun. **9**, 459 (2018).

[44] P. Brown, H. Fawzi, and O. Fawzi, *Computing conditional entropies for quantum correlations*, Nat. Commun. **12**, 575 (2021).

[45] V. Makarov, A. Anisimov, and J. Skaar, *Effects of Detector Efficiency Mismatch on Security of Quantum Cryptosystems*, Phys. Rev. A **74**, 022313 (2006).

[46] B. Qi, C.-H. F. Fung, H.-K. Lo, and X. Ma, *Time-Shift Attack in Practical Quantum Cryptosystems*, Quantum Inf. Comput **7**, 73 (2007), arXiv:quant-ph/0512080 .

[47] Y. Zhao, C.-H. F. Fung, B. Qi, C. Chen, and H.-K. Lo, *Quantum Hacking: Experimental Demonstration of Time-Shift Attack Against Practical Quantum-Key-Distribution Systems*, Phys. Rev. A **78**, 042333 (2008).

[48] V. Makarov, *Controlling Passively Quenched Single Photon Detectors by Bright Light*, New J. Phys. **11**, 065003 (2009).

[49] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov, *Hacking Commercial Quantum Cryptography Systems by Tailored Bright Illumination*, Nat. Photonics **4**, 686 (2010).

[50] L. Lydersen, M. K. Akhlaghi, A. H. Majedi, J. Skaar, and V. Makarov, *Controlling a Superconducting Nanowire Single-Photon Detector using Tailored Bright Illumination*, New J. Phys. **13**, 113042 (2011).

[51] I. Gerhardt, Q. Liu, A. Lamas-Linares, J. Skaar, C. Kurtsiefer, and V. Makarov, *Full-Field Implementation of a Perfect Eavesdropper on a Quantum Cryptography System*, Nat. Commun. **2**, 349 (2011).

[52] I. Gerhardt, Q. Liu, A. Lamas-Linares, J. Skaar, V. Scarani, V. Makarov, and C. Kurtsiefer, *Experimentally Faking the Violation of Bell's Inequalities*, Phys. Rev. Lett. **107**, 170404 (2011).

[53] H. Weier, H. Krauss, M. Rau, M. Fürst, S. Nauerth, and H. Weinfurter, *Quantum Eavesdropping without Interception: An Attack Exploiting the Dead Time of Single-Photon Detectors*, New J. Phys. **13**, 073024 (2011).

[54] S. Sajeed, P. Chaiwongkhot, J.-P. Bourgoin, T. Jennewein, N. Lütkenhaus, and V. Makarov, *Security Loophole in Free-Space Quantum Key Distribution due to Spatial-Mode Detector-Efficiency Mismatch*, Phys. Rev. A **91**, 062301 (2015).

[55] V. Makarov, J.-P. Bourgoin, P. Chaiwongkhot, M. Gagné, T. Jennewein, S. Kaiser, R. Kashyap, M. Legré, C. Minshull, and S. Sajeed, *Creation of Backdoors in Quantum Communications via Laser Damage*, Phys. Rev. A **94**, 030302 (2016).

[56] P. Chaiwongkhot, K. B. Kuntz, Y. Zhang, A. Huang, J.-P. Bourgoin, S. Sajeed, N. Lütkenhaus, T. Jennewein, and V. Makarov, *Eavesdropper's Ability to Attack a Free-Space Quantum-Key-Distribution Receiver in Atmospheric Turbulence*, Phys. Rev. A **99**, 062315 (2019).

[57] H.-K. Lo, M. Curty, and B. Qi, *Measurement-Device-Independent Quantum Key Distribution*, Phys. Rev. Lett. **108**, 130503 (2012).

[58] S. L. Braunstein and S. Pirandola, *Side-Channel-Free Quantum Key Distribution*, Phys. Rev. Lett. **108**, 130502 (2012).

[59] Y. Liu, T.-Y. Chen, L.-J. Wang, H. Liang, G.-L. Shentu, J. Wang, K. Cui, H.-L. Yin, N.-L. Liu, L. Li, X. Ma, J. S. Pelc, M. M. Fejer, C.-Z. Peng, Q. Zhang, and J.-W. Pan, *Experimental Measurement-Device-Independent Quantum Key Distribution*, Phys. Rev. Lett. **111**, 130502 (2013).

[60] Y.-L. Tang, H.-L. Yin, S.-J. Chen, Y. Liu, W.-J. Zhang, X. Jiang, L. Zhang, J. Wang, L.-X. You, J.-Y. Guan, D.-X. Yang, Z. Wang, H. Liang, Z. Zhang, N. Zhou, X. Ma, T.-Y. Chen, Q. Zhang, and J.-W. Pan, *Measurement-Device-Independent Quantum Key Distribution over 200 km*, Phys. Rev. Lett. **113**, 190501 (2014).

[61] Y. Cao, Y.-H. Li, K.-X. Yang, Y.-F. Jiang, S.-L. Li, X.-L. Hu, M. Abulizi, C.-L. Li, W. Zhang, Q.-C. Sun, W.-Y. Liu, X. Jiang, S.-K. Liao, J.-G. Ren, H. Li, L. You, Z. Wang, J. Yin, C.-Y. Lu, X.-B. Wang, Q. Zhang, C.-Z. Peng, and J.-W. Pan, *Long-Distance Free-Space Measurement-Device-Independent Quantum Key Distribution*, Phys. Rev. Lett. **125**, 260503 (2020).

[62] M. Lucamarini, Z. L. Yuan, J. F. Dynes, and A. J. Shields, *Overcoming the Rate-Distance Limit of Quantum Key Distribution without Quantum Repeaters*, Nature **557**, 400 (2018).

[63] X. Ma, P. Zeng, and H. Zhou, *Phase-Matching Quantum Key Distribution*, Phys. Rev. X **8**, 031043 (2018).

[64] M. Takeoka, S. Guha, and M. M. Wilde, *Fundamental Rate-Loss Tradeoff for Optical Quantum Key Distribution*, Nat. Commun. **5**, 5235 (2014).

[65] S. Pirandola, R. Laurenza, C. Ottaviani, and L. Banchi, *Fundamental Limits of Repeaterless Quantum Communications*, Nat. Commun. **8**, 15043 (2017).

[66] H.-K. Lo, X. Ma, and K. Chen, *Decoy State Quantum Key Distribution*, Phys. Rev. Lett. **94**, 230504 (2005).

[67] X.-B. Wang, *Beating the Photon-Number-Splitting Attack in Practical Quantum Cryptography*, Phys. Rev. Lett. **94**, 230503 (2005).

[68] X. Ma, B. Qi, Y. Zhao, and H.-K. Lo, *Practical Decoy State for Quantum Key Distribution*, Phys. Rev. A **72**, 012326 (2005).

[69] K. Tamaki, H.-K. Lo, W. Wang, and M. Lucamarini, *Information Theoretic Security of Quantum Key Distribution Overcoming the Repeaterless Secret Key Capacity Bound*, arXiv:1805.05511 (2018).

[70] X.-B. Wang, Z.-W. Yu, and X.-L. Hu, *Twin-field Quantum Key Distribution with Large Misalignment Error*, Phys. Rev. A **98**, 062323 (2018).

[71] J. Lin and N. Lütkenhaus, *Simple Security Analysis of Phase-Matching Measurement-Device-Independent Quantum Key Distribution*, Phys. Rev. A **98**, 042332 (2018).

[72] C. Cui, Z.-Q. Yin, R. Wang, W. Chen, S. Wang, G.-C. Guo, and Z.-F. Han, *Twin-Field Quantum Key Distribution without Phase Post-Selection*, Phys. Rev. Applied **11**, 034053 (2019).

[73] M. Curty, K. Azuma, and H.-K. Lo, *Simple Security Proof of Twin-Field Type Quantum Key Distribution Protocol*, npj. Quantum Inf. **5**, 64 (2019).

[74] L. Salvail, M. Peev, E. Diamanti, R. Alleaume, N. Lütkenhaus, and T. Laenger, *Security of Trusted Repeater Quantum Key Distribution Networks*, Journal of Computer Security **18**, 61 (2010), arXiv:0904.4072 .

[75] S.-K. Liao, W.-Q. Cai, W.-Y. Liu, L. Zhang, Y. Li, J.-G. Ren, J. Yin, Q. Shen, Y. Cao, Z.-P. Li, F.-Z. Li, X.-W. Chen, L.-H. Sun, J.-J. Jia, J.-C. Wu, X.-J. Jiang, J.-F. Wang, Y.-M. Huang, Q. Wang, Y.-L. Zhou, L. Deng, T. Xi, L. Ma, T. Hu, Q. Zhang, Y.-A. Chen, N.-L. Liu, X.-B. Wang, Z.-C. Zhu, C.-Y. Lu, C.-Z. P. Rong Shu, J.-Y. Wang, and J.-W. Pan, *Satellite-to-Ground Quantum Key Distribution*, Nature **549**, 43 (2017).

[76] R. Bedington, J. M. Arrazola, and A. Ling, *Progress in Satellite Quantum Key Distribution*, npj Quantum Inf. **3**, 1 (2017).

[77] P. Sibson, C. Erven, M. Godfrey, S. Miki, T. Yamashita, M. Fujiwara, M. Sasaki, H. Terai, M. G. Tanner, C. M. Natarajan, R. H. Hadfield, J. L. O'Brien, and M. G. Thompson, *Chip-Based Quantum Key Distribution*, Nat. Commun. **8**, 13984 (2017).

[78] G. Zhang, J. Y. Haw, H. Cai, F. Xu, S. Assad, J. F. Fitzsimons, X. Zhou, Y. Zhang, S. Yu, J. Wu, *et al.*, *An Integrated Silicon Photonic Chip Platform for Continuous-Variable Quantum Key Distribution*, Nat. Photonics **13**, 839 (2019).

[79] L. Cao, W. Luo, Y. Wang, J. Zou, R. Yan, H. Cai, Y. Zhang, X. Hu, C. Jiang, W. Fan, *et al.*, *Chip-Based Measurement-Device-Independent Quantum Key Distribution Using Integrated Silicon Photonic Systems*, Phys. Rev. Applied **14**, 011001 (2020).

[80] K. Wei, W. Li, H. Tan, Y. Li, H. Min, W.-J. Zhang, H. Li, L. You, Z. Wang, X. Jiang, *et al.*, *High-speed Measurement-Device-Independent Quantum Key Distribution with Integrated Silicon Photonics*, Phys. Rev. X **10**, 031030 (2020).

[81] J. Lin, T. Upadhyaya, and N. Lütkenhaus, *Asymptotic Security Analysis of Discrete-Modulated Continuous-Variable Quantum Key Distribution*, Phys. Rev. X **9**, 041064 (2019).

[82] J. Lin and N. Lütkenhaus, *Trusted Detector Noise Analysis for Discrete Modulation Schemes of Continuous-Variable Quantum Key Distribution*, Phys. Rev. Applied **14**, 064030 (2020).

[83] T. van Himbeeck, J. Lin, I. George, K. Fang, and N. Lütkenhaus, *Finite-Key Analysis of Quantum Key Distribution Using Entropy Accumulation* (2021), (in preparation).

[84] I. George, J. Lin, and N. Lütkenhaus, *Numerical Calculations of the Finite Key Rate for General Quantum Key Distribution Protocols*, Phys. Rev. Research **3**, 013274 (2021).

[85] T. Upadhyaya, T. van Himbeeck, J. Lin, and N. Lütkenhaus, *Dimension Reduction in Quantum Key Distribution for Continuous- and Discrete-Variable Protocols*, PRX Quantum **2**, 020325 (2021).

[86] Y. Zhang, P. J. Coles, A. Winick, J. Lin, and N. Lütkenhaus, *Security Proof of Practical Quantum Key Distribution with Detection-Efficiency Mismatch*, Phys. Rev. Research **3**, 013076 (2021).

[87] H. Hu, J. Im, J. Lin, N. Lütkenhaus, and H. Wolkowicz, *Robust Interior Point Method for Quantum Key Distribution Rate Computation*, arXiv:2104.03847 (2021).

[88] L. C. G. Govia, D. Bunandar, J. Lin, D. Englund, N. Lütkenhaus, and H. Krovi, *Clifford-Group-Restricted Eavesdroppers in Quantum Key Distribution*, Phys. Rev. A **101**, 062318 (2020).

[89] R. Renner, *Security of Quantum Key Distribution*, Ph.D. thesis, ETH Zürich, Zürich, Switzerland (2005), arXiv:quant-ph/0512258 .

[90] N. J. Beaudry, T. Moroder, and N. Lütkenhaus, *Squashing Models for Optical Measurements in Quantum Communication*, Phys. Rev. Lett. **101**, 093601 (2008).

[91] T. Tsurumaru and K. Tamaki, *Security Proof for Quantum-Key-Distribution Systems with Threshold Detectors*, Phys. Rev. A **78**, 032302 (2008).

[92] O. Gittsovich, N. J. Beaudry, V. Narasimhachar, R. R. Alvarez, T. Moroder, and N. Lütkenhaus, *Squashing Models for Detectors and Applications to Quantum Key Distribution Protocols*, Phys. Rev. A **89**, 012325 (2014).

[93] I. W. Primaatmaja, E. Lavie, K. T. Goh, C. Wang, and C. C. W. Lim, *Versatile Security Analysis of Measurement-Device-Independent Quantum Key Distribution*, Phys. Rev. A **99**, 062332 (2019).

[94] J. Watrous, *The Theory of Quantum Information* (Cambridge University Press, Cambridge, UK, 2018).

[95] M. Tomamichel, *Quantum Information Processing with Finite Resources: Mathematical Foundations*, Springer Briefs in Mathematical Physics, Vol. 5 (Springer Verlag, Berlin, 2018).

[96] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, Cambridge, UK, 2010).

[97] F. Dupuis, O. Fawzi, and R. Renner, *Entropy Accumulation*, Commun. Math. Phys. **379**, 867 (2020).

[98] F. Dupuis and O. Fawzi, *Entropy Accumulation with Improved Second-Order*, IEEE Trans. Inf. Theory **65**, 7596 (2019).

[99] F. Dupuis, *Privacy Amplification and Decoupling without Smoothing*, arXiv:2105.05342 (2021).

[100] C. C. Gerry and P. L. Knight, *Introductory Quantum Optics* (Cambridge University Press, Cambridge, UK, 2005).

[101] U. Leonhardt, *Essential Quantum Optics: from Quantum Measurements to Black Holes* (Cambridge University Press, Cambridge, UK, 2010).

[102] S. M. Barnett and P. L. Knight, *Thermofield Analysis of Squeezing and Statistical Mixtures in Quantum Optics*, J. Opt. Soc. Am. B **2**, 467 (1985).

[103] K. E. Cahill and R. J. Glauber, *Ordered Expansions in Boson Amplitude Operators*, Phys. Rev. **177**, 1857 (1969).

[104] K. E. Cahill and R. J. Glauber, *Density Operators and Quasiprobability Distributions*, Phys. Rev. **177**, 1882 (1969).

[105] M. S. Kim, F. A. M. de Oliveira, and P. L. Knight, *Properties of Squeezed Number States and Squeezed Thermal States*, Phys. Rev. A **40**, 2494 (1989).

[106] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge university press, Cambridge, UK, 2004).

[107] J. Borwein and A. S. Lewis, *Convex Analysis and Nonlinear Optimization: Theory and Examples* (Springer-Verlag, New York, USA, 2006).

[108] F. Grosshans, N. J. Cerf, J. Wenger, R. Tualle-Brouri, and P. Grangier, *Virtual Entanglement and Reconciliation Protocols for Quantum Cryptography with Continuous Variables*, Quantum Inf. Comput. **3**, 535 (2003).

[109] M. Curty, M. Lewenstein, and N. Lütkenhaus, *Entanglement as Precondition for Secure Quantum Key Distribution*, Phys. Rev. Lett. **92**, 217903 (2004).

[110] A. Ferenczi and N. Lütkenhaus, *Symmetries in Quantum Key Distribution and the Connection between Optimal Attacks and Optimal Cloning*, Phys. Rev. A **85**, 052310 (2012).

[111] R. Renner and J. I. Cirac, *de Finetti Representation Theorem for Infinite-Dimensional Quantum Systems and Applications to Quantum Cryptography*, Phys. Rev. Lett. **102**, 110504 (2009).

[112] M. Christandl, R. König, and R. Renner, *Postselection Technique for Quantum Channels with Applications to Quantum Cryptography*, Phys. Rev. Lett. **102**, 020504 (2009).

[113] M. Ben-Or, M. Horodecki, D. W. Leung, D. Mayers, and J. Oppenheim, *The Universal Composable Security of Quantum Key Distribution*, in *Theory of Cryptography Conference* (Springer, Berlin, Heidelberg, 2005) pp. 386–406.

[114] R. Renner and R. König, *Universally Composable Privacy Amplification against Quantum Adversaries*, in *Theory of Cryptography Conference* (Springer, Berlin, Heidelberg, 2005) pp. 407–425.

[115] J. Müller-Quade and R. Renner, *Composability in Quantum Cryptography*, New J. Phys. **11**, 085006 (2009).

[116] R. König, R. Renner, A. Bariska, and U. Maurer, *Small Accessible Quantum Information Does Not Imply Security*, Phys. Rev. Lett. **98**, 140502 (2007).

[117] V. Scarani and R. Renner, *Security Bounds for Quantum Cryptography with Finite Resources*, in *Theory of Quantum Computation, Communication, and Cryptography*, edited by Y. Kawano and M. Mosca (Springer, Berlin, Heidelberg, 2008) pp. 85–95, arXiv:0806.0120v1 .

[118] R. G. Gallager, *Low Density Parity Check Codes*, Ph.D. thesis, MIT, Cambridge, Massachusetts, USA (1963).

[119] E. Arikan, *Channel Polarization: A Method for Constructing Capacity-Achieving Codes*, in *2008 IEEE International Symposium on Information Theory* (2008) pp. 1173–1177.

[120] R. Renner, *Simplifying Information-Theoretic Arguments by Post-Selection*, Quantum Cryptography and Computing **26**, 66 (2010).

[121] H. Chernoff *et al.*, *A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on the Sum of Observations*, The Annals of Mathematical Statistics **23**, 493 (1952).

[122] W. Hoeffding, *Probability Inequalities for Sums of Bounded Random Variables*, Journal of the American Statistical Association **58**, 13 (1963).

[123] R. J. Serfling, *Probability Inequalities for the Sum in Sampling without Replacement*, The Annals of Statistics **2**, 39 (1974).

[124] B. Kraus, N. Gisin, and R. Renner, *Lower and Upper Bounds on the Secret-Key Rate for Quantum Key Distribution Protocols Using One-Way Classical Communication*, Phys. Rev. Lett. **95**, 080501 (2005).

[125] R. Renner and N. G. B. Kraus, *Information-Theoretic Security Proof for Quantum-Key-Distribution Protocols*, Phys. Rev. A **72**, 012332 (2005).

[126] V. Scarani and R. Renner, *Quantum Cryptography with Finite Resources: Unconditional Security Bound for Discrete-Variable Protocols with One-Way Postprocessing*, Phys. Rev. Lett. **100**, 200501 (2008).

[127] R. Y. Q. Cai and V. Scarani, *Finite-key Analysis for Practical Implementations of Quantum Key Distribution*, New J. Phys. **11**, 045024 (2009).

[128] M. Tomamichel, R. Colbeck, and R. Renner, *A Fully Quantum Asymptotic Equipartition Property*, IEEE Trans. Inf. Theory **55**, 5840 (2009).

[129] M. Tomamichel, *A Framework for Non-Asymptotic Quantum Infomration Theory*, Ph.D. thesis, ETH Zürich, Zürich, Switzerland (2012), arXiv:1203.2142 .

[130] K. Zyczkowski and H.-J. Sommers, *Induced Measures in the Space of Mixed Quantum sSates*, Journal of Physics A: Mathematical and General **34**, 7111 (2001).

[131] D. Mayers, *Unconditional Security in Quantum Cryptography*, J. ACM **48**, 351 (2001).

[132] P. W. Shor and J. Preskill, *Simple Security Proof of Security of the BB84 Quantum Key Distribution Protocol*, Phys. Rev. Lett. **85**, 441 (2000).

[133] M. Koashi, *Simple Security Proof of Quantum Key Distribution Based on Complementarity*, New J. Phys. **11**, 045018 (2009).

[134] M. Hayashi, *Upper Bounds of Eavesdropper's Performances in Finite-Length Code with the Decoy Method*, Phys. Rev. A **76**, 012329 (2007).

[135] H.-K. Lo, *Method for Decoupling Error Correction from Privacy Amplification*, New J. Phys. **5**, 36 (2003).

[136] D. Gottesman, H.-K. Lo, N. Lütkenhaus, and J. Preskill, *Security of quantum key distribution with imperfect devices*, in *International Symposium on Information Theory, 2004. ISIT 2004. Proceedings.* (IEEE, 2004) p. 136.

[137] T. Tsurumaru, *Leftover Hashing from Quantum Error Correction: Unifying the Two Approaches to the Security Proof of Quantum Key Distribution*, IEEE Trans. Inf. Theory **66**, 3465 (2020).

[138] P. J. Coles, M. Berta, M. Tomamichel, and S. Wehner, *Entropic Uncertainty Relations and Their Applications*, Rev. Mod. Phys. **89**, 015002 (2017).

[139] J. M. Renes and J.-C. Boileau, *Conjectured Strong Complementary Information Tradeoff*, Phys. Rev. Lett. **103**, 020402 (2009).

[140] M. Berta, M. Christandl, R. Colbeck, J. M. Renes, and R. Renner, *The Uncertainty Principle in the Presence of Quantum Memory*, Nat. Phys. **6**, 659 (2010).

[141] P. J. Coles, L. Yu, V. Gheorghiu, and R. B. Griffiths, *Information-Theoretic Treatment of Tripartite Systems and Quantum Channels*, Phys. Rev. A **83**, 062338 (2011).

[142] M. Tomamichel and R. Renner, *Uncertainty Relation for Smooth Entropies*, Phys. Rev. Lett. **106**, 110506 (2011).

[143] F. Furrer, T. Franz, M. Berta, A. Leverrier, V. B. Scholz, M. Tomamichel, and R. F. Werner, *Continuous Variable Quantum Key Distribution: Finite-Key Analysis of Composable Security Against Coherent Attacks*, Phys. Rev. Lett. **109**, 100502 (2012).

[144] A. Winick, N. Lütkenhaus, and P. J. Coles, *Reliable Numerical Key Rates for Quantum Key Distribution*, Quantum **2**, 77 (2018).

[145] I. Devetak and A. Winter, *Distillation of Secret Key and Entanglement from Quantum States*, Proc. R. Soc. A **461**, 207 (2005).

[146] P. J. Coles, E. M. Metodiev, and N. Lütkenhaus, *Numerical Approach for Unstructured Quantum Key Distribution*, Nat. Commun. **7**, 11712 (2016).

[147] P. J. Coles, *Unification of Different Views of Decoherence and Discord*, Phys. Rev. A **85**, 042103 (2012).

[148] M. Frank and P. Wolfe, *An Algorithm for Quadratic Programming*, Nav. Res. Logist. Q. **3**, 95 (1956).

[149] M. Grant and S. Boyd, *CVX: Matlab Software for Disciplined Convex Programming, version 2.1*, `http://cvxr.com/cvx` (2018).

[150] M. Grant and S. Boyd, *Graph Implementations for Nonsmooth Convex Programs*, in *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, edited by V. Blondel, S. Boyd, and H. Kimura (Springer-Verlag, Berlin, 2008) pp. 95–110.

[151] K. C. Toh, M. J. Todd, and R. H. Tutuncu, *SDPT3 — a Matlab Software Package for Semidefinite Programming*, Optim. Methods Software **11**, 545 (1999).

[152] R. H. Tutuncu, K. C. Toh, and M. J. Todd, *Solving Semidefinite-Quadratic-Linear Programs Using SDPT3*, Math. Program. Ser. B **95**, 189 (2003).

[153] MOSEK ApS, *The MOSEK Optimization Toolbox for MATLAB Manual. Version 8.0.0.60* (2016).

[154] W. Wang and H.-K. Lo, *Machine Learning for Optimal Parameter Prediction in Quantum Key Distribution*, Phys. Rev. A **100**, 062334 (2019).

[155] M.-G. Zhou, Z.-P. Liu, W.-B. Liu, C.-L. Li, J.-L. Bai, Y.-R. Xue, Y. Fu, H.-L. Yin, and Z.-B. Chen, *Machine Learning for Secure Key Rate in Continuous-Variable Quantum Key Distribution*, arXiv:2108.02578 (2021).

[156] T. Tsurumaru, *Squash Operator and Symmetry*, Phys. Rev. A **81**, 012328 (2010).

[157] A. Winter, *Tight Uniform Continuity Bounds for Quantum Entropies: Conditional Entropy, Relative Entropy Distance and Energy Constraints*, Commun. Math. Phys. **347**, 291 (2016).

[158] N. Killoran and N. Lütkenhaus, *Strong Quantitative Benchmarking of Quantum Optical Devices*, Phys. Rev. A **83**, 052320 (2011).

[159] H.-J. Briegel, W. Dür, J. I. Cirac, and P. Zoller, *Quantum Repeaters: The Role of Imperfect Local Operations in Quantum Communication*, Phys. Rev. Lett. **81**, 5932 (1998).

[160] D. Luong, L. Jiang, J. Kim, and N. Lütkenhaus, *Overcoming Lossy Channel Bounds Using a Single Quantum Repeater Node*, Appl. Phys. B **122**, 96 (2016).

[161] K. Tamaki, H.-K. Lo, C.-H. F. Fung, and B. Qi, *Phase Encoding Schemes for Measurement-Device-Independent Quantum Key Distribution with Basis-Dependent Flaw*, Phys. Rev. A **85**, 042307 (2012).

[162] A. Ferenczi, *Security Proof Methods for Quantum Key Distribution Protocols*, Ph.D. thesis, University of Waterloo (2013).

[163] H.-K. Lo, H. F. Chau, and M. Ardehali, *Efficient Quantum Key Distribution Scheme and a Proof of its Unconditional Security*, J. Cryptol. **18**, 133 (2005).

[164] N. Mukunda and E. C. G. Sudarshan, *New Light on the Optical Equivalence Theorem and a New Type of Discrete Diagonal Coherent State Representation*, Pramana **10**, 227 (1978).

[165] J. K. Sharma, C. L. Mehta, N. Mukunda, and E. C. G. Sudarshan, *Representation and Properties of Para-Bose Oscialltor Operators. II. Coherent States and the Minimum Uncertainty States*, J. Math. Phys. **22**, 78 (1981).

[166] M. Minder, M. Pittaluga, G. Roberts, M. Lucamarini, J. Dynes, Z. Yuan, and A. Shields, *Experimental Quantum Key Distribution beyond the Repeaterless Secret Key Capacity*, Nat. Photonics **13**, 334 (2019).

[167] S. Wang, D.-Y. He, Z.-Q. Yin, F.-Y. Lu, C.-H. Cui, W. Chen, Z. Zhou, G.-C. Guo, and Z.-F. Han, *Beating the Fundamental Rate-Distance Limit in a Proof-of-Principle Quantum Key Distribution System*, Physical Review X **9**, 021046 (2019).

[168] X. Zhong, J. Hu, M. Curty, L. Qian, and H.-K. Lo, *Proof-of-Principle Experimental Demonstration of Twin-Field Type Quantum Key Distribution*, Phys. Rev. Lett. **123**, 100506 (2019).

[169] Y. Liu, Z.-W. Yu, W. Zhang, J.-Y. Guan, J.-P. Chen, C. Zhang, X.-L. Hu, H. Li, C. Jiang, J. Lin, T.-Y. Chen, L. You, Z. Wang, X.-B. Wang, Q. Zhang, and J.-W. Pan, *Experimental Twin-Field Quantum Key Distribution through Sending or Not Sending*, Phys. Rev. Lett. **123**, 100505 (2019).

[170] K. Maeda, T. Sasaki, and M. Koashi, *Repeaterless Quantum Key Distribution with Efficient Finite-Key Analysis Overcoming the Rate-Distance Limit*, Nat. Commun. **10**, 3140 (2019).

[171] G. Currás-Lorenzo, Á. Navarrete, K. Azuma, G. Kato, M. Curty, and M. Razavi, *Tight Finite-Key Security for Twin-Field Quantum Key Distribution*, npj Quantum Inf. **7**, 22 (2021).

[172] F. Grasselli, Á. Navarrete, and M. Curty, *Asymmetric Twin-Field Quantum Key Distribution*, New J. Phys. **21**, 113032 (2019).

[173] W. Wang and H.-K. Lo, *Simple Method for Asymmetric Twin-Field Quantum Key Distribution*, New J. Phys. **22**, 013020 (2020).

[174] J.-P. Chen, C. Zhang, Y. Liu, C. Jiang, W. Zhang, X.-L. Hu, J.-Y. Guan, Z.-W. Yu, H. Xu, J. Lin, M.-J. Li, H. Chen, H. Li, L. You, Z. Wang, X.-B. Wang, Q. Zhang, and J.-W. Pan, *Sending-or-Not-Sending with Independent Lasers: Secure Twin-Field Quantum Key Distribution over 509 km*, Phys. Rev. Lett. **124**, 070501 (2020).

[175] X.-T. Fang, P. Zeng, H. Liu, M. Zou, W. Wu, Y.-L. Tang, Y.-J. Sheng, Y. Xiang, W. Zhang, H. Li, Z. Wang, L. You, H. C. Ming-Jun Li, Y.-A. Chen, Q. Zhang, C.-Z. Peng, X. M. andTeng Yun Chen, and J.-W. Pan, *Implementation of Quantum Key Distribution Surpassing the Linear Rate-Transmittance Bound*, Nat. Photonics **14**, 422 (2020).

[176] P. Zeng, W. Wu, and X. Ma, *Symmetry-Protected Privacy: Beating the Rate-Distance Linear Bound over a Noisy Channel*, Phys. Rev. Applied **13**, 064013 (2020).

[177] C.-M. Zhang, Y.-W. Xu, R. Wang, and Q. Wang, *Twin-Field Quantum Key Distribution with Discrete-Phase-Randomized Sources*, Phys. Rev. Applied **14**, 064070 (2020).

[178] C. Jiang, Z.-W. Yu, X.-L. Hu, and X.-B. Wang, *Sending-or-Not-Sending Twin-Field Quantum Key Distribution with Discrete-Phase-Randomized Weak Coherent States*, Phys. Rev. Research **2**, 043304 (2020).

[179] G. Currás-Lorenzo, L. Wooltorton, and M. Razavi, *Twin-Field Quantum Key Distribution with Fully Discrete Phase Randomization*, Phys. Rev. Applied **15**, 014016 (2021).

[180] Z. Cao, Z. Zhang, H.-K. Lo, and X. Ma, *Discrete-Phase-Randomized Coherent State Source and its Application in Quantum Key Distribution*, New J. Phys. **17**, 053014 (2015).

[181] F. Grasselli, H. Kampermann, and D. Bruß, *Conference Key Agreement with Single-Photon Interference*, New J. Phys. **21**, 123002 (2019).

[182] M. Navascués, F. Grosshans, and A. Acín, *Optimality of Gaussian Attacks in Continuous-Variable Quantum Cryptography*, Phys. Rev. Lett. **97**, 190502 (2006).

[183] R. García-Patrón and N. J. Cerf, *Unconditional Optimality of Gaussian Attacks against Continuous-Variable Quantum Key Distribution*, Phys. Rev. Lett. **97**, 190503 (2006).

[184] P. Jouguet, S. Kunz-Jacques, E. Diamanti, and A. Leverrier, *Analysis of Imperfections in Practical Continuous-Variable Quantum Key Distribution*, Phys. Rev. A **86**, 032309 (2012).

[185] E. Kaur, S. Guha, and M. M. Wilde, *Asymptotic Security of Discrete-Modulation Protocols for Continuous-Variable Quantum Key Distribution*, Phys. Rev. A **103**, 012412 (2021).

[186] H. Häseler and N. Lütkenhaus, *Quantum Benchmarks for the Storage or Transmission of Quantum Light from Minimal Resources*, Phys. Rev. A **81**, 060306(R) (2010).

[187] Y.-B. Zhao, M. Heid, J. Rigas, and N. Lütkenhaus, *Asymptotic Security of Binary Modulated Continuous-Variable Quantum Key Distribution under Collective Attacks*, Phys. Rev. A **79**, 012307 (2009).

[188] K. Brádler and C. Weedbrook, *Security Proof of Continuous-Variable Quantum Key Distribution using Three Coherent States*, Phys. Rev. A **97**, 022310 (2018).

[189] A. Leverrier and P. Grangier, *Unconditional Security Proof of Long-Distance Continuous-Variable Quantum Key Distribution with Discrete Modulation*, Phys. Rev. Lett. **102**, 180504 (2009).

[190] P. Papanastasiou, C. Lupo, C. Weedbrook, and S. Pirandola, *Quantum Key Distribution with Phase-Encoded Coherent States: Asymptotic Security Analysis in Thermal-Loss Channels*, Phys. Rev. A **98**, 012340 (2018).

[191] S. Ghorai, P. Grangier, E. Diamanti, and A. Leverrier, *Asymptotic Security of Continuous-Variable Quantum Key Distribution with a Discrete Modulation*, Phys. Rev. X **9**, 021059 (2019).

[192] C. Silberhorn, T. C. Ralph, N. Lütkenhaus, and G. Leuchs, *Continuous Variable Quantum Cryptography: Beating the 3 dB Loss Limit*, Phys. Rev. Lett. **89**, 167901 (2002).

[193] M. Heid and N. Lütkenhaus, *Security of Coherent-State Quantum Cryptography in the Presence of Gaussian Noise*, Phys. Rev. A **76**, 022313 (2007).

[194] M. Heid and N. Lütkenhaus, *Efficiency of Coherent-State Quantum Cryptography in the Presence of Loss: Influence of Realistic Error Correction*, Phys. Rev. A **73**, 052316 (2006).

[195] S. Fossier, E. Diamanti, T. Debuisschert, R. Tualle-Brouri, and P. Grangier, *Improvement of Continuous-Variable Quantum Key Distribution Systems by Using Optical Preamplifiers*, J. Phys. B **42**, 114014 (2009).

[196] V. C. Usenko and R. Filip, *Trusted Noise in Continuous-Variable Quantum Key Distribution: a Threat and a Defense*, Entropy **18**, 20 (2016).

[197] F. Laudenbach and C. Pacher, *Analysis of the Trusted-Device Scenario in Continuous-Variable Quantum Key Distribution*, Adv. Quantum Technol. **2**, 1900055 (2019).

[198] R. Namiki, A. Kitagawa, and T. Hirano, *Secret Key Rate of a Continuous-Variable Quantum-Key-Distribution Scheme when the Detection Process is Inaccessible to Eavesdroppers*, Phys. Rev. A **98**, 042319 (2018).

[199] B. R. Mollow and R. J. Glauber, *Quantum Theory of Parameteric Amplification. I*, Phys. Rev. **160**, 1076 (1967).

[200] W. Vogel and D.-G. Welsch, *Quantum Optics* (Wiley, New York, USA, 2006).

[201] R. Renner, *Symmetry of Large Physical Systems Implies Independence of Subsystems*, Nat. Phys. **3**, 645 (2007).

[202] M. E. Shirokov, *Tight Uniform Continuity Bounds for the Quantum Conditional Mutual Information, for the Holevo Quantity, and for Capacities of Quantum Channels*, J. Math. Phys. (N. Y.) **58**, 102202 (2017).

[203] M. Dušek, M. Jahma, and N. Lütkenhaus, *Unambiguous State Discrimination in Quantum Cryptography with Weak Coherent States*, Phys. Rev. A **62**, 022306 (2000).

[204] C. Pacher, P. Grabenweger, J. Martinez-Mateo, and V. Martin, *An Information Reconciliation Protocol for Secret-Key Agreement with Small Leakage*, in *2015 IEEE International Symposium on Information Theory (ISIT)* (IEEE, 2015) pp. 730–734.

[205] F. Kanitschar and C. Pacher, *Postselection Strategies for Continuous-Variable Quantum Key Distribution Protocols with Quadrature Phase-Shift Keying Modulation*, arXiv:2104.09454 (2021).

[206] A. Denys, P. Brown, and A. Leverrier, *Explicit Asymptotic Secret Key Rate of Continuous-Variable Quantum Key Distribution with an Arbitrary Modulation*, Quantum **5**, 540 (2021).

[207] T. Matsuura, K. Maeda, T. Sasaki, and M. Koashi, *Finite-Size Security of Continuous-Variable Quantum Key Distribution with Digital Signal Processing*, Nat. Commun. **12**, 252 (2021).

[208] W.-B. Liu, C.-L. Li, Y.-M. Xie, C.-X. Weng, J. Gu, X.-Y. Cao, Y.-S. Lu, B.-H. Li, H.-L. Yin, and Z.-B. Chen, *Homodyne Detection Quadrature Phase Shift Keying Continuous-Variable Quantum Key Distribution with High Excess Noise*, arXiv:2104.11152 (2021).

[209] R. Tannous, Z. Ye, J. Jin, K. B. Kuntz, N. Lütkenhaus, and T. Jennewein, *Demonstration of a 6 State-4 State Reference Frame Independent Channel for Quantum Key Distribution*, Appl. Phys. Lett. **115**, 211103 (2019).

[210] A. Laing, V. Scarani, J. G. Rarity, and J. L. O'Brien, *Reference-Frame-Independent Quantum Key Distribution*, Phys. Rev. A **82**, 012304 (2010).

[211] L. Sheridan and V. Scarani, *Security Proof for Quantum Key Distribution Using Qudit Systems*, Phys. Rev. A **82**, 030301(R) (2010).

[212] P. Marian and T. A. Marian, *Squeezed States with Thermal Noise. I. Photon-Number Statistics*, Phys. Rev. A **47**, 4474 (1993).

[213] M. Milicevic, C. Feng, L. M. Zhang, and P. G. Gulak, *Quasi-Cyclic Multi-Edge LDPC Codes for Long-Distance Quantum Cryptography*, npj Quantum Inf. **4**, 21 (2018).

[214] P. Kok and S. L. Braunstein, *Postselected versus Nonpostselected Quantum Teleportation using Parametric Down-Conversion*, Phys. Rev. A **61**, 042304 (2000).

[215] M. M. Wilde, *Quantum Information Theory* (Cambridge University Press, Cambridge, UK, 2013).

[216] S. Barnett and P. M. Radmore, *Methods in Theoretical Quantum Optics*, Vol. 15 (Oxford University, New York, USA, 2002).

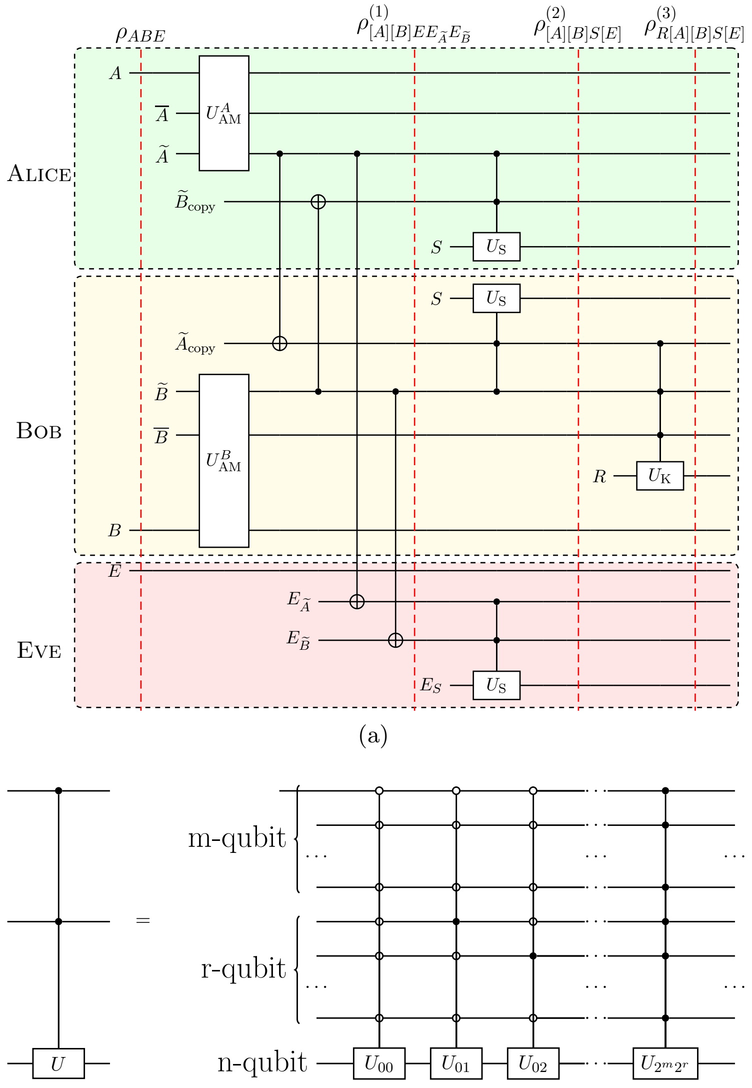
APPENDICES

Appendix A

Simplification of postprocessing map

In this appendix, we present a derivation of a general postprocessing framework for the numerical method in Ref. [144]. Our derivation reproduces the form of postprocessing map $\mathcal{G}$ presented in Ref. [144] when Eve's systems are traced out. This formulation is an alternative formulation from the one presented in Section 4.1.2. This formulation allows us to see directly why we can use the quantum relative entropy as the objective function of the key rate optimization problem. Moreover, based on this derivation, we can make several observations to simplify the postprocessing map $\mathcal{G}$ in special cases, which leads to a reduction of the dimensions required in the numerical analysis.

We begin with some definitions to set up the notations. When we write an operator on composite registers, we omit the identity operator on the unspecified registers and may reorder registers for ease of writing. Moreover, the relevant unspecified registers depend on the context. Let $\mathcal{X}$ denote the set of Alice's measurement outcomes. With Alice's set of announcements S^A, we partition the set $\mathcal{X}$ into subsets $\mathcal{X}_a$ for $a \in S^A$ such that $\mathcal{X} = \bigcup_{a \in S^A} \mathcal{X}_a$. Similarly, we partition Bob's measurement outcomes $\mathcal{Y}$ as $\mathcal{Y} = \bigcup_{b \in S^B} \mathcal{Y}_b$ by using his set of announcements S^B. To simplify our notation, we assume without the loss of generality that $|\mathcal{X}_a| = \omega_A$ for all $a \in S^A$ and $|\mathcal{Y}_b| = \omega_B$ for all $b \in S^B$ for some numbers ω_A and ω_B independent of a and b. This assumption can be easily satisfied by a clever way of bookkeeping measurement outcomes. Then, we define a family of maps $f_a : \Omega^A := \{1, 2, \ldots, \omega_A\} \to \mathcal{X}_a$ and a family of maps $f_b : \Omega^B := \{1, 2, \ldots, \omega_B\} \to \mathcal{Y}_b$ such that f_a's and f_b's are bijective. Finally, we label Alice's POVM P^A as $P^A := \{P^A_x : x \in \mathcal{X}\} = \{P^A_{[a, f_a(\alpha)]} : a \in S^A, \alpha \in \Omega^A\}$ and Bob's POVM $P^B := \{P^B_y : y \in \mathcal{Y}\} = \{P^B_{[b, f_b(\beta)]} : b \in S^B, \beta \in \Omega^B\}$.



(a)

(b)

Figure A.1: (a) A schematic circuit diagram for the relevant postprocessing steps: announcement (with measurement), sifting, and key map. Three dashed boxes separate Alice's, Bob's, and Eve's domains. The initial pure state ρ_{ABE} is evolved by an isometry at each step which introduces additional registers and applies a unitary operation on relevant registers. $U^A_{\mathrm{AM}}, U^B_{\mathrm{AM}}, U_{\mathrm{S}}$ and U_{K} are unitary operations related to the announcement and measurement, sifting, and key map. (b) An explanation for the controlled unitary operation used in (a). Here, m, r, and n are some sufficiently large integers so that we have a representation of the basis elements for each register in the computational basis of qubits. See the text in Appendix A for more explanations.

A.1 A full model for the relevant postprocessing steps

We give a schematic circuit diagram in Figure A.1 to describe the announcement (including measurement), sifting and key map steps. This diagram covers the scenarios related to this work, and it works for protocols with one round of announcements and with reverse reconciliation. It is not difficult to draw a similar diagram in other scenarios, including the direct reconciliation schemes. Under collective attacks, Alice and Bob share a bipartite quantum state ρ_{AB} after each transmission of a quantum signal. In the worst-case scenario, Eve holds a purification ρ_{ABE} of ρ_{AB}. The initial state in the circuit diagram is ρ_{ABE}. At each step, the state is evolved by an isometry; that is, we introduce some local registers and evolve the state by a local unitary. We also keep track of the information leakage during the classical communication. If some information is publicly available during the classical communication in the protocol, then each party holds a copy of the relevant registers. One can recover the classical communication information by measuring a local copy of these registers in the computational basis. Now, we discuss each of the three steps in detail.

The isometries related to the announcement and measurement step are denoted by W_A for Alice and W_B for Bob. In particular, W_A first introduces two registers $\widetilde{A}$ and $\overline{A}$ and then applies a unitary operator U^{A}_{AM} to implement Alice's POVM P^A in a coherent fashion, where announcements are stored in the register $\widetilde{A}$ and measurement outcomes are in the register $\overline{A}$. Like Alice's isometry W_A, Bob's isometry W_B implements his POVM P^B with the announcement register $\widetilde{B}$ and measurement outcome register $\overline{B}$ using a local unitary U^{B}_{AM}. Since the announcement information is available to everyone, Eve obtains a copy (denoted by $E_{\widetilde{A}}$) of $\widetilde{A}$ and a copy (denoted by $E_{\widetilde{B}}$) of $\widetilde{B}$. The coherent version of copying is represented by the controlled NOT operation. Also, Alice and Bob each have a copy of the other party's announcement register, denoted by $\widetilde{B}_{\text{copy}}$ and $\widetilde{A}_{\text{copy}}$, respectively. For convenience of writing, we use $\widetilde{A}$ and $\widetilde{B}$ to refer to both Alice's and Bob's copies of $\widetilde{A}$ and $\widetilde{B}$. As we see later, we can actually combine the register $\widetilde{A}$ with $\widetilde{A}_{\text{copy}}$ and combine $\widetilde{B}$ with $\widetilde{B}_{\text{copy}}$ in the key rate calculation. In this diagram, the state after the announcement and measurement step is $\rho^{(1)}_{[A][B]EE_{\widetilde{A}}E_{\widetilde{B}}} = (W_A \otimes W_B)\rho_{ABE}(W_A \otimes W_B)^{\dagger}$, where for ease of writing, we reorder the registers and use a shorthand notation for collections of registers: $[A]$ for registers $A\widetilde{A}\overline{A}$ and $[B]$ for registers $B\widetilde{B}\overline{B}$.

The sifting step partitions the set of announcement events $S^A \times S^B$ as $S^A \times S^B = \mathbf{K} \cup \mathbf{D}$, where $\mathbf{K}$ is the set of announcement events to be kept and $\mathbf{D}$ is the set of announcements to be discarded. The sifting isometry (denoted by V_S) introduces a register S to store the result of sifting ("keep" or "discard") and performs a unitary operator U_{S} on the local copies of registers $\widetilde{A}$ and $\widetilde{B}$ to compute the sifting decision. In a common scenario, each party can implement this unitary U_{S} from the description of a protocol. If it is not from the protocol description and additional classical communication is needed, then, after a party implements this unitary operation, other parties obtain a copy of the register S. For simplicity, we use S to refer to both Alice's and Bob's copies of this register. In the diagram, the state after this sifting step is $\rho^{(2)}_{[A][B]S[E]} = V_S \rho^{(1)}_{[A][B]EE_{\widetilde{A}}E_{\widetilde{B}}} V_S^{\dagger}$, where we use a shorthand notation $[E]$ to refer to Eve's collection of registers $EE_{\widetilde{A}}E_{\widetilde{B}}E_S$.

The key map isometry V_K introduces a register R and applies a local unitary U_{K} to compute the key map g and to store the result in the register R. This key map g takes the announcement $(a, b) \in S^A \times S^B$ and Alice's measurement outcome $f_a(\alpha)$ in the case of direct reconciliation or Bob's measurement outcome $f_b(\beta)$ in the case of reverse reconciliation as inputs and outputs a value in $\{0, 1, \ldots, N-1\}$ where N is the number of key symbols. For the purpose of derivation, we include an additional key symbol $\perp$ to this set and map all discarded events to it. We see later that we can eventually remove the symbol $\perp$ from the set of key symbols. In the diagram, the state after the key map state is $\rho^{(3)}_{R[A][B]S[E]} = V_K \rho^{(2)}_{[A][B]S[E]} V_K^{\dagger}$.

To give explicit expressions for these isometries W_A, W_B, V_S, and V_K, we first define K_a^A and K_b^B as

$$\begin{aligned} K_a^A &= \sum_{\alpha \in \Omega^A} \sqrt{P^A_{[a, f_a(\alpha)]}} \otimes |a\rangle_{\widetilde{A}} \otimes |\alpha\rangle_{\overline{A}}\,, \\ K_b^B &= \sum_{\beta \in \Omega^B} \sqrt{P^B_{[b, f_b(\beta)]}} \otimes |b\rangle_{\widetilde{B}} \otimes |\beta\rangle_{\overline{B}}\,, \end{aligned} \tag{A.1}$$

where $\{|a\rangle : a \in S^A\}$ and $\{|b\rangle : b \in S^B\}$ are orthonormal bases for registers $\widetilde{A}$ ($E_{\widetilde{A}}$) and $\widetilde{B}$ ($E_{\widetilde{B}}$) and $\{|\alpha\rangle : \alpha \in \Omega^A\}$ and $\{|\beta\rangle : \beta \in \Omega^B\}$ are orthonormal bases for registers $\overline{A}$ and $\overline{B}$, respectively. (Note that $|\alpha\rangle$ here is not a coherent state discussed in the main text.) We remark that K_a^A and K_b^B are the same as defined in Eqs. (40) and (41) of Ref. [144] if we write $f_a(\alpha)$ as α_a and $f_b(\beta)$ as β_b. Then W_A, W_B and V_S are defined, respectively, as

$$\begin{aligned} W_A &= \sum_{a \in S^A} K_a^A \otimes |a\rangle_{E_{\widetilde{A}}}\,, \\ W_B &= \sum_{b \in S^B} K_b^B \otimes |b\rangle_{E_{\widetilde{B}}}\,, \\ V_S &= \Pi \otimes |\mathbf{K}\rangle_S \otimes |\mathbf{K}\rangle_{E_S} + (\mathbb{1}_{\widetilde{A}\widetilde{B}} - \Pi) \otimes |\mathbf{D}\rangle_S \otimes |\mathbf{D}\rangle_{E_S}\,, \end{aligned} \tag{A.2}$$

where $\Pi = \sum_{(a,b)\in\mathbf{K}} |a\rangle\langle a|_{\widetilde{A}} \otimes |b\rangle\langle b|_{\widetilde{B}}$ and $\{|\mathbf{K}\rangle, |\mathbf{D}\rangle\}$ is an orthonormal basis for the register S (E_S). To write out the key map isometry V_K, we take the reverse reconciliation schemes as an example, and it is straightforward to write out V_K in the case of direct reconciliation schemes by using Alice's measurement outcome $f_a(\alpha)$ instead of Bob's outcome $f_b(\beta)$. We first define an (partial) isometry V on the subspace that Π projects onto and then write out V_K:

$$\begin{aligned} V &= \sum_{\substack{(a,b)\in\mathbf{K} \\ \beta \in \Omega^B}} |g(a, b, f_b(\beta))\rangle_R \otimes |a\rangle\langle a|_{\widetilde{A}} \otimes |b\rangle\langle b|_{\widetilde{B}} \otimes |\beta\rangle\langle\beta|_{\overline{B}}\,, \\ V_K &= V\Pi + |\perp\rangle_R \otimes (\mathbb{1}_{\widetilde{A}\widetilde{B}} - \Pi). \end{aligned} \tag{A.3}$$

We remark that the final state $\rho^{(3)}_{R[A][B]S[E]} = V_K V_S (W_A \otimes W_B)\rho_{ABE}(W_A \otimes W_B)^\dagger V_S^\dagger V_K^\dagger$ is a pure state since ρ_{ABE} is a pure state, and we apply only isometries to it.

A.2 Removing the dependence on Eve's registers

To access the key information, we use the projective measurement $\{Z_j = |j\rangle\langle j|_R : j \in \{0, 1, \dots, N-1, \perp\}\}$. Since the final state $\rho^{(3)}_{R[A][B]S[E]}$ is pure, we apply Theorem 4.1.1 [147, Theorem 1] to rewrite conditional entropy $H(\mathbf{Z}|[E])$ as

$$H(\mathbf{Z}|[E]) = D(\rho^{(3)}_{R[A][B]S} || \sum_j Z_j \rho^{(3)}_{R[A][B]S} Z_j), \tag{A.4}$$

where $\rho^{(3)}_{R[A][B]S} = \mathrm{Tr}_{[E]}(\rho^{(3)}_{R[A][B]S[E]})$. We then define an announcement map $\mathcal{A}$ for an input state σ as $\mathcal{A}(\sigma) = \sum_{a \in S^A} \sum_{b \in S^B} (K^A_a \otimes K^B_b)\sigma(K^A_a \otimes K^B_b)^\dagger$ and rewrite $\rho^{(3)}_{R[A][B]S}$ as

$$\begin{aligned} \rho^{(3)}_{R[A][B]S} =& \mathrm{Tr}_{[E]}(\rho^{(3)}_{R[A][B]S[E]}) \\ =& p_{\text{pass}} \rho^{\mathbf{K}}_{R[A][B]} \otimes |\mathbf{K}\rangle\langle\mathbf{K}|_S \\ &+ (1 - p_{\text{pass}})\rho^{\mathbf{D}}_{R[A][B]} \otimes |\mathbf{D}\rangle\langle\mathbf{D}|_S , \end{aligned} \tag{A.5}$$

where $p_{\text{pass}} = \mathrm{Tr}\big(V\Pi\mathcal{A}(\rho_{AB})\Pi V^\dagger\big) = \mathrm{Tr}(\mathcal{A}(\rho_{AB})\Pi)$ is the same sifting probability defined in the main text and

$$\begin{aligned} \rho^{\mathbf{K}}_{R[A][B]} &= \frac{V\Pi\mathcal{A}(\rho_{AB})\Pi V^\dagger}{p_{\text{pass}}}, \\ \rho^{\mathbf{D}}_{R[A][B]} &= \frac{|\perp\rangle\langle\perp|_R \otimes (\mathbb{1}_{\widetilde{A}\widetilde{B}} - \Pi)\mathcal{A}(\rho_{AB})(\mathbb{1}_{\widetilde{A}\widetilde{B}} - \Pi)}{1 - p_{\text{pass}}}. \end{aligned} \tag{A.6}$$

To show that the symbol $\perp$ has no contribution to the key rate, we use the following lemma (see Ref. [215]).

Lemma A.2.1. For quantum-classical states ρ_{QX} and σ_{QX} defined as $\rho_{QX} = \sum_x p(x)\rho^x_Q \otimes |x\rangle\langle x|_X$, $\sigma_{QX} = \sum_x q(x)\sigma^x_Q \otimes |x\rangle\langle x|_X$, where p and q are probability distributions over a finite alphabet $\mathcal{X}$ and ρ^x_Q and σ^x_Q are density operators for all $x \in \mathcal{X}$, the quantum relative entropy is $D(\rho_{QX}||\sigma_{QX}) = \sum_x p(x)D(\rho^x_Q||\sigma^x_Q) + D(p||q)$.

Applying the lemma to the state $\rho^{(3)}_{R[A][B]S}$ with the classical register S gives us

$$\begin{aligned}
&D\big(\rho^{(3)}_{R[A][B]S}||\mathcal{Z}(\rho^{(3)}_{R[A][B]S})\big)\\
&= p_{\text{pass}}D\big(\rho^{\mathbf{K}}_{R[A][B]}||\mathcal{Z}(\rho^{\mathbf{K}}_{R[A][B]})\big)\\
&\quad + (1-p_{\text{pass}})D(\rho^{\mathbf{D}}_{R[A][B]}||\rho^{\mathbf{D}}_{R[A][B]})\\
&= p_{\text{pass}}D\big(\rho^{\mathbf{K}}_{R[A][B]}||\mathcal{Z}(\rho^{\mathbf{K}}_{R[A][B]})\big)\\
&= D\big(\mathcal{G}(\rho_{AB})||\mathcal{Z}[\mathcal{G}(\rho_{AB})]\big),
\end{aligned} \tag{A.7}$$

where we define $\mathcal{G}(\rho_{AB}) = V\Pi\mathcal{A}(\rho_{AB})\Pi V^\dagger$, which is the same as in Ref. [144].

Finally, we remark that, on the subspace where Π projects, the symbol $\perp$ does not show up anymore. Thus, we can modify $\{Z_j\}$ to remove the symbol $\perp$ in the end. This modification gives back to the definition of $\mathcal{Z}$ shown in Ref. [144].

A.3 Simplifying the postprocessing map

We now provide several remarks to explain how we can simplify the map $\mathcal{G}$ while making sure that such a simplification does not change our calculated key rates. Our discussion here takes the reverse reconciliation schemes as an example. It is straightforward to adapt the arguments to the direct reconciliation schemes.

We first make a remark about the registers $\widetilde{A}_{\text{copy}}$ and $\widetilde{B}_{\text{copy}}$ that are hidden in our notation $\widetilde{A}$ and $\widetilde{B}$. After tracing out Eve's registers $E_{\widetilde{A}}$ and $E_{\widetilde{B}}$, Alice's register $\widetilde{A}$ and Bob's copy $\widetilde{A}_{\text{copy}}$ are both classical registers, and, likewise, Bob's register $\widetilde{B}$ and Alice's copy $\widetilde{B}_{\text{copy}}$ are classical. Since each of the sifting and key map steps is done locally via a controlled unitary whose target is the register S or R alone, we can pull out two copies of registers $\widetilde{A}$ and $\widetilde{B}$ to write the final state in the form of the quantum-classical state to which the previous lemma applies. If we look at the block diagonal structure of $\mathcal{G}(\rho_{AB})$ with respect to two copies of the register $\widetilde{A}$, we see directly that the state with a single copy of the register $\widetilde{A}$ is just embedded in a larger space with two copies of the register $\widetilde{A}$. This result means the eigenvalues of the state are unaffected by removing one copy of the register $\widetilde{A}$. A similar argument works for two copies of $\widetilde{B}$. Moreover, from the previous lemma, we see immediately that we can calculate the key rate from individual announcements if we write the key map g as $g[a,b,f_b(\beta)] =: g_{ab}(\beta)$ for a collection of functions g_{ab}, one for each $(a,b) \in \mathbf{K}$. In this case, for each $(a,b) \in \mathbf{K}$, we define an isometry $V_{ab} = \sum_{\beta\in\Omega^B} |g_{ab}(\beta)\rangle_R \otimes |\beta\rangle\langle\beta|_{\overline{B}}$ and a completely positive map $\mathcal{G}_{ab}$ for an input state σ as $\mathcal{G}_{ab}(\sigma) = V_{ab}(\widetilde{K}^A_a \otimes \widetilde{K}^B_b)\sigma(\widetilde{K}^A_a \otimes \widetilde{K}^B_b)^\dagger V^\dagger_{ab}$, where we define $\widetilde{K}_a^A$ such that $K_a^A = \widetilde{K}_a^A \otimes |a\rangle_{\widetilde{A}}$ and $\widetilde{K}_b^B$ such that $K_b^B = \widetilde{K}_b^B \otimes |b\rangle_{\widetilde{B}}$. Then,

$$\begin{aligned} &D\big(\mathcal{G}(\rho_{AB})||\mathcal{Z}[\mathcal{G}(\rho_{AB})]\big) \\ =& \sum_{(a,b)\in\mathbf{K}} D\big(\mathcal{G}_{ab}(\rho_{AB})||\mathcal{Z}[\mathcal{G}_{ab}(\rho_{AB})]\big). \end{aligned} \tag{A.8}$$

Besides the lemma, our objective function has another important property. Since the quantum relative entropy is invariant under an isometry, if an isometry can commute with $\mathcal{G}$ and $\mathcal{Z}$ maps, then our objective function is also invariant under this isometry. In other words, we can add or remove an isometry W (that acts only on Alice's and Bob's registers) in the final expression of Eq. (A.7) if W commutes with $\mathcal{G}$ and $\mathcal{Z}$ maps.

From this property of our objective function, if those functions g_{ab}'s are the identity function, then we see that each isometry V_{ab} simply copies the register $\overline{B}$ and stores this copy to the register R. Adding this copy is a local isometry and renaming the register $\overline{B}$ by the name R is a unitary. Thus, we can combine the registers $\overline{B}$ and R and retain the name of R.

Also, based on this property of our objective function, we now discuss when we can omit the appearance of the register $\overline{A}$. As depicted in Figure A.2, the announcement and measurement step can be decomposed into two steps. First, Alice performs a coarse-grained measurement (with associated unitary U_{A}^A in this figure) to make announcements. Second, conditioned on her own announcement result, Alice performs a refined measurement (with a controlled unitary U_{M}^A in the figure) if needed to obtain the fine-grained measurement outcomes. As the refined measurement is done in Alice's local registers, to which Eve has no access, Alice's refined measurement can be described by a local isometry. This local isometry can be performed after Eve obtains a copy of the announcement registers. In the reverse reconciliation scheme, since the key map isometry V does not depend on the register $\overline{A}$, the isometry that describes Alice's refined measurement then commutes with $\mathcal{G}$ and $\mathcal{Z}$ maps. This result means that, from the key rate calculation perspective, we can drop this isometry due to the property of our objective function. More precisely, if we define $P_a^A = \sum_{\alpha\in\Omega^A} P_{[a,f_a(\alpha)]}^A$ for each $a \in S^A$ and define $K_a^{A'} = \sqrt{P_a^A}\otimes|a\rangle_{\widetilde{A}}$, then we can replace the map $\mathcal{G}$ by the map $\mathcal{G}'$ defined as

$$\mathcal{G}'(\sigma) = V\Pi\Big(\sum_{\substack{a\in S^A \\ b\in S^B}} (K_a^{A'}\otimes K_b^B)\sigma(K_a^{A'}\otimes K_b^B)^\dagger\Big)\Pi V^\dagger \tag{A.9}$$

for an input state σ. (A similar replacement can be done for $\mathcal{G}_{ab}$.) A similar argument can be applied to the direct reconciliation schemes by interchanging the roles of Alice's and Bob's registers to show that we can omit the register $\overline{B}$ for direct reconciliation.

Along the same line of argument, we remark that the refined measurement conditioned on the announcements can be a coarse-grained instead of the fine-grained measurement if the key map

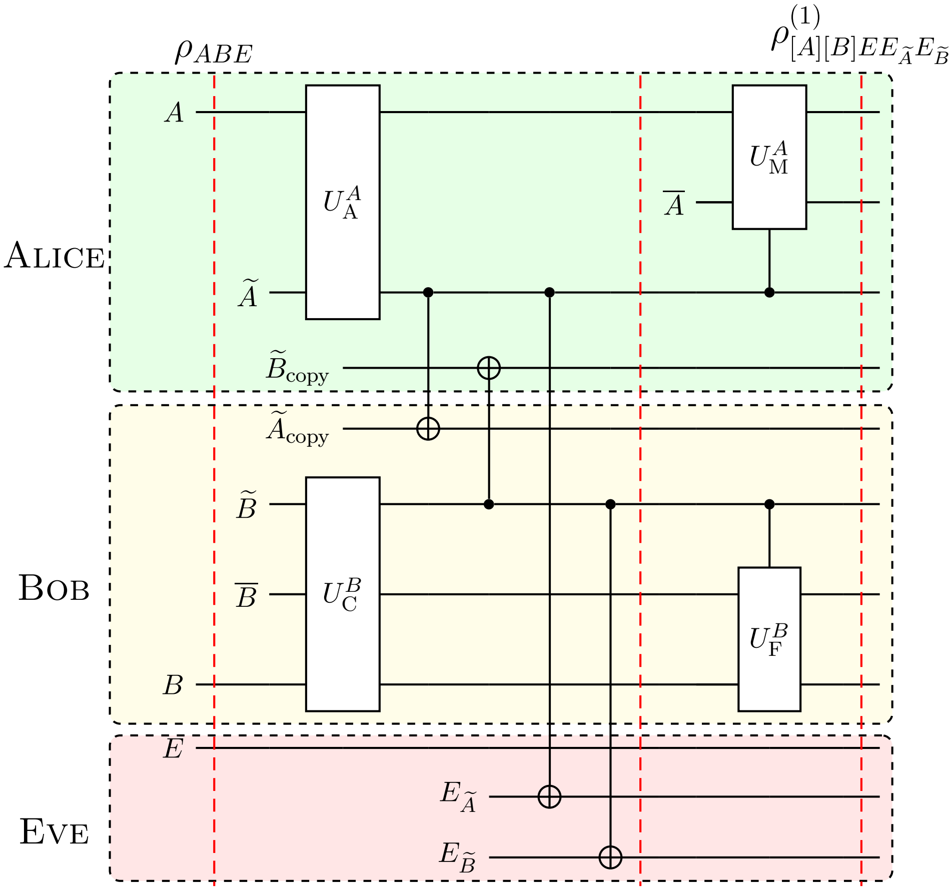


Figure A.2: An alternative description of the announcement step for the reverse reconciliation schemes. This step can be decomposed into two steps. At the first step, Alice performs only a coarse-grained measurement with a unitary U^A_{A} to obtain announcement results, and Bob also performs a coarse-grained measurement with a unitary U^B_{C} to obtain announcement outcomes and coarse-grained measurement information. At the second step, they choose to perform optional refined measurements (U^A_{M} and U^B_{F}) conditioned on the announcements (and previous coarse-grained measurement information for Bob). They can postpone the refined measurements after giving Eve announcement results and in some cases choose not to perform the refined measurements.

uses only the coarse-grained information. This process is also described in Figure A.2. Bob first applies an isometry (with associated unitary U_{C}^{B} in the figure) to implement the coarse-grained measurement which gives the same coarse-grained information needed for the key map g and then applies an additional isometry (with associated unitary U_{F}^{B} in the figure) to obtain fine-grained measurement outcomes. As the sifting step depends only on the announcement, we can move Bob's refined measurement after the sifting step (not shown in this figure). Since the key map uses only the coarse-grained information, the key map isometry effectively undoes the unitary U_{F}^{B}. Therefore, we can take the POVM related to the coarse-grained measurement when we write out Kraus operators K_b^B in Eq. (A.1).

Appendix B

Technical details of PM-MDI QKD

B.1 A representation of Eve's POVM in the subspace $\mathcal{S}$

In this appendix, we will describe how to find a representation of Eve's POVM $\{F^\gamma\}$ in an orthonormal basis of the subspace spanned by the signal states, which we previously denoted as $\text{span}(\mathcal{S})$. When we discuss about the two-mode coherent states $|\alpha_A, \alpha_B\rangle$ prepared by Alice and Bob, for the ease of notation, we write $\boldsymbol{\alpha} = (\alpha_A, \alpha_B)$ and $|\boldsymbol{\alpha}\rangle = |\alpha_A, \alpha_B\rangle$.

If we are given $\langle\boldsymbol{\alpha}| F^\gamma |\boldsymbol{\beta}\rangle$ for every $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathcal{S}$, then the procedure described here allows us to find a four-dimensional representation of F^γ in the subspace $\text{span}(\mathcal{S})$ and helps us evaluate the von Neumann entropy of Eve's conditional states more straightforwardly. We remark here that the values of $\langle\boldsymbol{\alpha}| F^\gamma |\boldsymbol{\beta}\rangle$ can be determined by test states in the test mode of our protocol, as discussed in Section 5.2.4. For our simulations in Appendix B.2, we also provide a simulation method to obtain $\langle\boldsymbol{\alpha}| F^\gamma |\boldsymbol{\beta}\rangle$.

Before we proceed, we want to emphasize that the set $\mathcal{S}$ defined in Eq. (5.2) is a basis for the subspace $\text{span}(\mathcal{S})$ and we will use this particular ordering of basis elements in the later discussion.

B.1.1 Orthonormal basis decomposition

Since two coherent states $\{\left|+\sqrt{\mu}\right\rangle, \left|-\sqrt{\mu}\right\rangle\}$ span a two-dimensional space, we start with a canonical two-dimensional description of $\{\left|+\sqrt{\mu}\right\rangle, \left|-\sqrt{\mu}\right\rangle\}$.

$$\begin{aligned} |+\sqrt{\mu}\rangle &= c_0 |e_0\rangle + c_1 |e_1\rangle \\ |-\sqrt{\mu}\rangle &= c_0 |e_0\rangle - c_1 |e_1\rangle, \end{aligned} \tag{B.1}$$

where $\{|e_0\rangle, |e_1\rangle\}$ is an orthonormal basis, $|c_0|^2 + |c_1|^2 = 1$ and $|c_0|^2 - |c_1|^2 = \langle +\sqrt{\mu}|-\sqrt{\mu}\rangle$. Without loss of generality, we choose c_0 and c_1 to be real numbers by absorbing the complex phases into the definitions of $|e_0\rangle, |e_1\rangle$. We remark here that the explicit expressions for $|e_0\rangle$ and $|e_1\rangle$ are irrelevant for our discussion, but a canonical choice of this basis written in the Fock state basis is

$$\begin{aligned} |e_0\rangle &= \frac{1}{\sqrt{\cosh(\mu)}} \sum_{n=0}^{\infty} \frac{\sqrt{\mu}^{2n}}{\sqrt{(2n)!}} |2n\rangle , \\ |e_1\rangle &= \frac{1}{\sqrt{\sinh(\mu)}} \sum_{n=0}^{\infty} \frac{\sqrt{\mu}^{2n+1}}{\sqrt{(2n+1)!}} |2n+1\rangle , \end{aligned} \tag{B.2}$$

and with this choice of basis, $c_0 = e^{-\frac{\mu}{2}}\sqrt{\cosh(\mu)}$ and $c_1 = e^{-\frac{\mu}{2}}\sqrt{\sinh(\mu)}$.

We then obtain a basis for span($\mathcal{S}$) as $\mathcal{B} = \{|e_0, e_0\rangle, |e_1, e_1\rangle, |e_0, e_1\rangle, |e_1, e_0\rangle\}$. We remark that this particular ordering of basis elements is useful for the presentation and later allows us to see the block diagonal structures of some particular POVM elements more straightforwardly.

We now write out signal states in the set $\mathcal{S}$ as column vectors in this basis $\mathcal{B}$:

$$\begin{aligned} |+\sqrt{\mu}, +\sqrt{\mu}\rangle &= \begin{pmatrix} c_0^2 \\ c_1^2 \\ c_0c_1 \\ c_0c_1 \end{pmatrix}, \quad |-\sqrt{\mu}, -\sqrt{\mu}\rangle = \begin{pmatrix} c_0^2 \\ c_1^2 \\ -c_0c_1 \\ -c_0c_1 \end{pmatrix}, \\ |+\sqrt{\mu}, -\sqrt{\mu}\rangle &= \begin{pmatrix} c_0^2 \\ -c_1^2 \\ -c_0c_1 \\ c_0c_1 \end{pmatrix}, \quad |-\sqrt{\mu}, +\sqrt{\mu}\rangle = \begin{pmatrix} c_0^2 \\ -c_1^2 \\ c_0c_1 \\ -c_0c_1 \end{pmatrix} \end{aligned} \tag{B.3}$$

Once we write out F^γ in the basis $\mathcal{B}$, we can then find Eve's conditional states in the basis $\mathcal{B}$ by appropriate multiplications. Then the evaluation of von Neumann entropy of conditional states is straightforward since finding eigenvalues of 4 × 4 matrices is computationally simple.

B.1.2 Change of basis matrix

Suppose we have determined $\langle \boldsymbol{\alpha}| F^\gamma |\boldsymbol{\beta}\rangle$, where $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathcal{S}$. Now we want to write out F^γ in the basis $\mathcal{B}$. This can be done by a change of basis matrix.

We can write $|\boldsymbol{e}_m\rangle \in \mathcal{B}$ in the basis $\mathcal{S}$:

$$|\boldsymbol{e}_m\rangle = \sum_n A_{nm} |\boldsymbol{\alpha}_n\rangle , \tag{B.4}$$

where A_{nm} is the (n, m) entry of the desired change of basis matrix A and $|\boldsymbol{\alpha}_n\rangle \in \mathcal{S}$.

Similarly,

$$\langle \boldsymbol{e}_m| = \sum_n \bar{A}_{nm} \langle \boldsymbol{\alpha}_n| = \sum_n (A^\dagger)_{mn} \langle \boldsymbol{\alpha}_n| \,, \tag{B.5}$$

where $\bar{A}_{nm}$ is the complex conjugate of A_{nm} and $A^\dagger$ is the Hermitian conjugate of A.

Combining previous two equations, we have

$$\langle \boldsymbol{e}_m| F^\gamma |\boldsymbol{e}_n\rangle = \sum_{i,j} (A^\dagger)_{mj} \langle \boldsymbol{\alpha}_j| F^\gamma |\boldsymbol{\alpha}_i\rangle A_{in}. \tag{B.6}$$

In the ordering of $\mathcal{S}$ and $\mathcal{B}$, this change of basis matrix A can be expressed as

$$A = \begin{pmatrix} \frac{1}{4c_0^2} & \frac{1}{4c_1^2} & \frac{1}{4c_0c_1} & \frac{1}{4c_0c_1} \\ \frac{1}{4c_0^2} & \frac{1}{4c_1^2} & -\frac{1}{4c_0c_1} & -\frac{1}{4c_0c_1} \\ \frac{1}{4c_0^2} & -\frac{1}{4c_1^2} & -\frac{1}{4c_0c_1} & \frac{1}{4c_0c_1} \\ \frac{1}{4c_0^2} & -\frac{1}{4c_1^2} & \frac{1}{4c_0c_1} & -\frac{1}{4c_0c_1} \end{pmatrix}, \tag{B.7}$$

where c_0 and c_1 are defined from Eq. (B.1).

B.2 Simulation

In this section, we explain how to obtain values of $\langle \boldsymbol{\alpha}| F^\gamma |\boldsymbol{\beta}\rangle$ for $|\boldsymbol{\alpha}\rangle, |\boldsymbol{\beta}\rangle \in \mathcal{S}$ through simulations. After knowing $\langle \boldsymbol{\alpha}| F^\gamma |\boldsymbol{\beta}\rangle$, we can then use the results from Appendix B.1, in particular, Eq. (B.6), to express F^γ in the basis $\mathcal{B}$ and then proceed with the evaluation of key rate.

To avoid the confusion between our simulation method and experimental execution of the protocol, we remark on how to obtain $\langle \boldsymbol{\alpha}| F^\gamma |\boldsymbol{\beta}\rangle$ in the actual implementation of the protocol. In the parameter estimation step, the values of $\langle \boldsymbol{\alpha}| F^\gamma |\boldsymbol{\alpha}\rangle$ for $\boldsymbol{\alpha} \in \mathcal{S}$ are directly obtained from observed correlation. The values of $\langle \boldsymbol{\alpha}| F^\gamma |\boldsymbol{\beta}\rangle$ for $\boldsymbol{\alpha} \neq \boldsymbol{\beta}$ can be calculated by observed correlation of test states, as explained in Section 5.2.4.

For our simulation, we propagate the input states through our simulated model of imperfections and then apply the POVM of detectors to the final states arriving at the detectors to calculate $\langle \boldsymbol{\alpha}| F^\gamma |\boldsymbol{\beta}\rangle$.

B.2.1 Eve's POVM associated with loss-only scenario

We now consider the loss-only scenario. Here, we simulate the quantum channel as a lossy channel and we consider the normal situation where Charlie (Eve) is honest and performs the

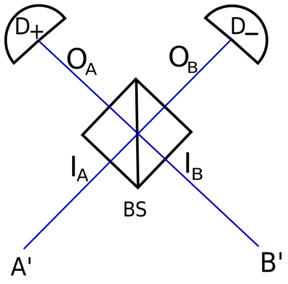


Figure B.1: Explanation of labels for input and output modes of the beam splitter.

measurements shown in Figure 5.1a. Namely, we calculate the POVM F corresponding to the real setup with ideal devices at the central node. Our protocol can verify via test states in the test mode that this is the actual Eve's POVM in the loss-only scenario. As mentioned in the main text, for the purpose of presentation, we consider the symmetric setup. For a total single-photon transmitivity η between Alice and Bob, each path has a transmissivity $\sqrt{\eta}$.

In this section, we will label Eve's POVM F^{γ} associated with the loss-only scenario by adding the subscript "loss". As shown in Figure B.1, we will label the input modes of the beam splitter at the central node as I_A and I_B and the output modes as O_A and O_B, where O_A reaches the detector D_+ and O_B reaches the detector D_-.

We describe how to obtain $\langle \boldsymbol{\alpha} | F^{\gamma}_{\text{loss}} | \boldsymbol{\beta} \rangle$. First, Alice and Bob prepare coherent states $|\alpha_A\rangle_{A'}$ and $|\alpha_B\rangle_{B'}$ in the registers A' and B', respectively, and send them to Charlie. After the lossy channel, the state becomes $\left|\sqrt{\sqrt{\eta}}\alpha_A, \sqrt{\sqrt{\eta}}\alpha_B\right\rangle_{I_AI_B}$, and Eve has the state $\left|\sqrt{1-\sqrt{\eta}}\alpha_A, \sqrt{1-\sqrt{\eta}}\alpha_B\right\rangle_{E_AE_B}$ at her disposal. Then after the beam splitter, the state becomes $\left|\frac{\sqrt{\sqrt{\eta}}\alpha_A+\sqrt{\sqrt{\eta}}\alpha_B}{\sqrt{2}}, \frac{\sqrt{\sqrt{\eta}}\alpha_A-\sqrt{\sqrt{\eta}}\alpha_B}{\sqrt{2}}\right\rangle_{O_AO_B}$. We now apply the POVM of the detectors to this state.

The ideal detectors are described by the following POVM:

$$\begin{aligned}
\Pi^{+}_{\text{ideal}} &= (\mathbb{1}_{O_A} - |0\rangle\langle 0|_{O_A}) \otimes |0\rangle\langle 0|_{O_B}\,, \\
\Pi^{-}_{\text{ideal}} &= |0\rangle\langle 0|_{O_A} \otimes (\mathbb{1}_{O_B} - |0\rangle\langle 0|_{O_B}), \\
\Pi^{?}_{\text{ideal}} &= |0\rangle\langle 0|_{O_A} \otimes |0\rangle\langle 0|_{O_B}\,, \\
\Pi^{d}_{\text{ideal}} &= (\mathbb{1}_{O_A} - |0\rangle\langle 0|_{O_A}) \otimes (\mathbb{1}_{O_B} - |0\rangle\langle 0|_{O_B}),
\end{aligned} \tag{B.8}$$

where $|0\rangle$ is the vacuum state.

Then, for $|\boldsymbol{\alpha}\rangle = |\alpha_A, \alpha_B\rangle$ and $|\boldsymbol{\beta}\rangle = |\beta_A, \beta_B\rangle$, we evaluate $\langle \boldsymbol{\alpha} | F^{\gamma}_{\text{loss}} | \boldsymbol{\beta} \rangle$ as

$$
\begin{aligned}
&\langle \boldsymbol{\alpha} | \, F^{\gamma}_{\text{loss}} \, | \boldsymbol{\beta} \rangle \\
&= \left\langle \frac{\eta^{\frac{1}{4}}(\alpha_A + \alpha_B)}{\sqrt{2}}, \frac{\eta^{\frac{1}{4}}(\alpha_A - \alpha_B)}{\sqrt{2}} \right| \Pi^{\gamma}_{\text{ideal}} \left| \frac{\eta^{\frac{1}{4}}(\beta_A + \beta_B)}{\sqrt{2}}, \frac{\eta^{\frac{1}{4}}(\beta_A - \beta_B)}{\sqrt{2}} \right\rangle \\
&\quad \times \left\langle \sqrt{1-\sqrt{\eta}}\alpha_A \middle| \sqrt{1-\sqrt{\eta}}\beta_A \right\rangle_{E_A} \left\langle \sqrt{1-\sqrt{\eta}}\alpha_B \middle| \sqrt{1-\sqrt{\eta}}\beta_B \right\rangle_{E_B} .
\end{aligned}
\tag{B.9}
$$

Now we have obtained values for $\langle \boldsymbol{\alpha} | \, F^{\gamma}_{\text{loss}} \, | \boldsymbol{\beta} \rangle$ from simulations, and we can then proceed to write F^{γ}_{loss} in the basis $\mathcal{B}$ by using Eq. (B.6).

$$
\begin{aligned}
F^{+}_{\text{loss}} &= (1-\xi^2) \begin{pmatrix} \frac{1-\xi^2\Omega^2}{8c_0^4} & \frac{1-\xi^2\Omega^2}{8c_0^2c_1^2} & 0 & 0 \\ \frac{1-\xi^2\Omega^2}{8c_0^2c_1^2} & \frac{1-\xi^2\Omega^2}{8c_1^4} & 0 & 0 \\ 0 & 0 & \frac{1+\xi^2\Omega^2}{8c_0^2c_1^2} & \frac{1+\xi^2\Omega^2}{8c_0^2c_1^2} \\ 0 & 0 & \frac{1+\xi^2\Omega^2}{8c_0^2c_1^2} & \frac{1+\xi^2\Omega^2}{8c_0^2c_1^2} \end{pmatrix} \\
F^{-}_{\text{loss}} &= (1-\xi^2) \begin{pmatrix} \frac{1-\xi^2\Omega^2}{8c_0^4} & \frac{-1+\xi^2\Omega^2}{8c_0^2c_1^2} & 0 & 0 \\ \frac{-1+\xi^2\Omega^2}{8c_0^2c_1^2} & \frac{1-\xi^2\Omega^2}{8c_1^4} & 0 & 0 \\ 0 & 0 & \frac{1+\xi^2\Omega^2}{8c_0^2c_1^2} & \frac{-1-\xi^2\Omega^2}{8c_0^2c_1^2} \\ 0 & 0 & \frac{-1-\xi^2\Omega^2}{8c_0^2c_1^2} & \frac{1+\xi^2\Omega^2}{8c_0^2c_1^2} \end{pmatrix} \\
F^{?}_{\text{loss}} &= \xi^2 \begin{pmatrix} \frac{(1+\Omega)^2}{4c_0^4} & 0 & 0 & 0 \\ 0 & \frac{(1-\Omega)^2}{4c_1^4} & 0 & 0 \\ 0 & 0 & \frac{1-\Omega^2}{4c_0^2c_1^2} & 0 \\ 0 & 0 & 0 & \frac{1-\Omega^2}{4c_0^2c_1^2} \end{pmatrix} \\
F^{d}_{\text{loss}} &= 0,
\end{aligned}
\tag{B.10}
$$

where for the ease of representation, we define $\Omega = e^{-2(1-\sqrt{\eta})\mu}$, and $\xi = e^{-\sqrt{\eta}\mu}$. Also, c_0 and c_1 are defined from the decomposition in Eq. (B.1). By noting that $2c_0^2 = 1+\xi^2\Omega$ and $2c_1^2 = 1-\xi^2\Omega$, we can easily check that $F^{+}_{\text{loss}} + F^{-}_{\text{loss}} + F^{?}_{\text{loss}} + F^{d}_{\text{loss}} = \mathbb{1}$.

B.2.2 Models for imperfections

In this section, we consider realistic imperfections in the experimental setup. Specifically, we consider the mode mismatch, phase mismatch, dark counts of detectors and the inefficiency of detectors. In the following, we describe the physical models for those imperfections.

Mode mismatch

We consider the mode mismatch with a simulation parameter V. The model of the mode mismatch is that for a coherent state $|\alpha\rangle_1$ in a mode 1, due to the mode mismatch, the state $\left|\sqrt{V}\alpha\right\rangle_1$ remains in the mode 1 and $\left|\sqrt{1-V}\alpha\right\rangle_2$ is in the mode 2, which is distinct from the mode 1. To derive values for $\langle\boldsymbol{\alpha}| F^{\gamma} |\boldsymbol{\beta}\rangle$, we will propagate the input states, similar to the loss-only case. In the place of I_A, I_B, O_A, O_B used in the discussion of loss-only scenario, we will replace them by $I_{A1}, I_{B1}, O_{A1}, O_{B1}$ for the initial mode and $I_{A2}, I_{B2}, O_{A2}, O_{B2}$ for the additional mode.

Suppose Alice sends $|\alpha_A\rangle_{A'}$ and Bob sends $|\alpha_B\rangle_{B'}$. Since only the relative mode mismatch between Alice's mode and Bob's mode matters, without loss of generality, we leave the state in the register A' untouched when it reaches I_{A1}, that is, we have $|\alpha_A\rangle_{I_{A1}}$. Due to the mode mismatch, the state $|\alpha_B\rangle_{B'}$ becomes $\left|\sqrt{V}\alpha_B\right\rangle_{I_{B1}} \left|\sqrt{1-V}\alpha_B\right\rangle_{I_{B2}}$. Correspondingly, we have the vacuum state in the mode I_{A2}. To summarize, the state arriving at the beam splitter of Charlie's station due to mode mismatch is $|\alpha_A, 0\rangle_{I_{A1}I_{A2}} \left|\sqrt{V}\alpha_B, \sqrt{1-V}\alpha_B\right\rangle_{I_{B1}I_{B2}}$. The mode I_{A1} interferes with the mode I_{B1} and the mode I_{A2} interferes with the mode I_{B2} independently. We remark here that this parameter V can be made close to 1 with experimentally available compensation systems, for example, see Ref. [59] in the setting of MDI protocols. In particular, $V \geq 95\%$ is readily achievable.

Phase mismatch

Ideally, Alice and Bob should prepare coherent states with the same global phase. We consider the situation where there is a phase mismatch between Alice's signal state and Bob's signal state. For the ideal input state $|\alpha_A, \alpha_B\rangle_{A'B'}$, due to the phase mismatch, the state becomes $\left|\alpha_A, \alpha_B e^{i\delta}\right\rangle_{I_A I_B}$ for some δ. We expect that with an experimentally feasible phase compensation system, the value of δ is typically small. For instance, the continuous-variable QKD experiment in Ref. [32] reports a value less than $\frac{\pi}{60}$.

Detector dark count

We now consider dark counts of detectors. For simplicity of our presentation, we model two detectors to have the same dark count probability p_d. It is also straightforward to model the case where two detectors have different dark count probabilities.

The effect of dark counts can be taken into consideration by modifying the POVM for detectors. Eq. (B.8)) gives the POVM associated with the ideal detectors. When the detectors have dark counts, the associated POVM associated is modified as below

$$\begin{aligned}
\Pi^{+}_{\text{dark}} &= (\mathbb{1}_{O_A} - |0\rangle\langle 0|_{O_A}) \otimes (1-p_d)\, |0\rangle\langle 0|_{O_B} \\
&\quad + p_d\, |0\rangle\langle 0|_{O_A} \otimes (1-p_d)\, |0\rangle\langle 0|_{O_B} \\
&= (1-p_d)\Pi^{+}_{\text{ideal}} + (1-p_d)p_d\Pi^{?}_{\text{ideal}} \\
\Pi^{-}_{\text{dark}} &= (1-p_d)\, |0\rangle\langle 0|_{O_A} \otimes (\mathbb{1}_{O_B} - |0\rangle\langle 0|_{O_B}) \\
&\quad + (1-p_d)\, |0\rangle\langle 0|_{O_A} \otimes p_d\, |0\rangle\langle 0|_{O_B} \\
&= (1-p_d)\Pi^{-}_{\text{ideal}} + (1-p_d)p_d\Pi^{?}_{\text{ideal}} \\
\Pi^{?}_{\text{dark}} &= (1-p_d)\, |0\rangle\langle 0|_{O_A} \otimes (1-p_d)\, |0\rangle\langle 0|_{O_B} \\
&= (1-p_d)^2\Pi^{?}_{\text{ideal}} \\
\Pi^{d}_{\text{dark}} &= (\mathbb{1}_{O_A} - |0\rangle\langle 0|_{O_A}) \otimes p_d\, |0\rangle\langle 0|_{O_B} \\
&\quad + p_d\, |0\rangle\langle 0|_{O_A} \otimes (\mathbb{1}_{O_B} - |0\rangle\langle 0|_{O_B}) + p_d\, |0\rangle\langle 0|_{O_A} \otimes p_d\, |0\rangle\langle 0|_{O_B} \\
&\quad + (\mathbb{1}_{O_A} - |0\rangle\langle 0|_{O_A}) \otimes (\mathbb{1}_{O_B} - |0\rangle\langle 0|_{O_B}) \\
&= p_d\Pi^{+}_{\text{ideal}} + p_d\Pi^{-}_{\text{ideal}} + p_d^2\Pi^{?}_{\text{ideal}} + \Pi^{d}_{\text{ideal}}.
\end{aligned} \tag{B.11}$$

Since in our simulation, we propagate the input states through the physical models of imperfections to derive the final states before the detectors and then use the POVM of detectors to derive the values of $\langle \boldsymbol{\alpha}| F^{\gamma} |\boldsymbol{\beta}\rangle$, the expression of $\langle \boldsymbol{\alpha}| F^{\gamma} |\boldsymbol{\beta}\rangle$ will have a similar structure as the loss-only case shown in Eq. (B.9). From this observation, we know that once we obtained the expression of Eve's POVM element F^{γ}, when we have considered all other imperfections except dark counts, we can then derive the POVM elements including dark counts of detectors by probabilistic mixtures of F^{γ}, following the same relation as between $\Pi^{\gamma}_{\text{dark}}$ and $\Pi^{\gamma}_{\text{ideal}}$.

To illustrate the idea, we give a simple example where we consider the physical channel to be a lossy channel and we want to include dark counts of detectors in our simulation. Since we have derived F^{γ}_{loss} in Eq. (B.10), we can derive the expressions of F^{γ}_{dark} corresponding to this simulation as follows:

$$\begin{aligned}
F^{+}_{\text{dark}} &= (1-p_d)F^{+}_{\text{loss}} + (1-p_d)p_dF^{?}_{\text{loss}} \\
F^{-}_{\text{dark}} &= (1-p_d)F^{-}_{\text{loss}} + (1-p_d)p_dF^{?}_{\text{loss}} \\
F^{?}_{\text{dark}} &= (1-p_d)^2F^{?}_{\text{loss}} \\
F^{d}_{\text{dark}} &= p_dF^{+}_{\text{loss}} + p_dF^{-}_{\text{loss}} + p_d^2F^{?}_{\text{loss}} + F^{d}_{\text{loss}}.
\end{aligned} \tag{B.12}$$

Detector's efficiency

We take into account that any practical single-photon detector has a limited efficiency. In our simulation method, we can easily modify the POVM of detectors as in Eq. (B.8), to include the efficiency of each detector separately. However, for simplicity of our presentation, we assume that both detectors have the same efficiency η_d so that we can combine the detector's efficiency and the channel transmittance by redefining the total transmissivity η. Let η_t refer to the single-photon transmission probability of the quantum channel between Alice and Bob. Since both detectors have the same efficiency, we can redefine the total transmissivity $\eta = \eta_t \eta_d^2$ and use this value of η in the simulation. Then we can still use the POVM of detectors with the perfect efficiency in our simulation.

B.2.3 Eve's POVM with those imperfections

As discussed in the previous section, we now take into consideration the mode mismatch with a simulation parameter V, and the phase mismatch with a simulation parameter δ. We consider both detectors have the same detector efficiency η_d and the same dark count probability p_d.

We will first derive Eve's POVM $F^{\gamma}_{\text{mismatch}}$ when we consider both the mode mismatch and the phase mismatch. Then we derive Eve's POVM $F^{\gamma}_{\text{model}}$ when we include dark counts of detectors as well. Finally, the effects of detector efficiency is taken into consideration by a redefinition of η.

For an input coherent state $|\alpha_A, \alpha_B\rangle_{A'B'}$, the state after the lossy channels and models for the mode mismatch and phase mismatch becomes $\left|\sqrt{\sqrt{\eta}}\alpha_A, \sqrt{\sqrt{\eta}}\sqrt{V}\alpha_B e^{i\delta}\right\rangle_{I_{A1}I_{B1}}$ $\otimes \left|0, \sqrt{\sqrt{\eta}}\sqrt{1-V}\alpha_B e^{i\delta}\right\rangle_{I_{A2}I_{B2}}$ and Eve has $\left|\sqrt{1-\sqrt{\eta}}\alpha_A, \sqrt{1-\sqrt{\eta}}\alpha_B\right\rangle_{E_A E_B}$ at her disposal. Then, the state arriving at the detectors is $\left|\frac{\sqrt{\sqrt{\eta}}\alpha_A + \sqrt{\sqrt{\eta}}\sqrt{V}\alpha_B e^{i\delta}}{\sqrt{2}}, \frac{\sqrt{\sqrt{\eta}}\alpha_A - \sqrt{\sqrt{\eta}}\sqrt{V}\alpha_B e^{i\delta}}{\sqrt{2}}\right\rangle_{O_{A1}O_{B1}}$ $\otimes \left|\frac{\sqrt{\sqrt{\eta}}\sqrt{1-V}\alpha_B e^{i\delta}}{\sqrt{2}}, -\frac{\sqrt{\sqrt{\eta}}\sqrt{1-V}\alpha_B e^{i\delta}}{\sqrt{2}}\right\rangle_{O_{A2}O_{B2}}$.

We now introduce the POVM of the ideal detectors when there are two independent modes entering the detectors due to mode mismatch.

$$\begin{aligned}
\Pi^{+}_{\text{mismatch}} &= (\mathbb{1}_{O_{A1}O_{A2}} - |00\rangle\langle 00|_{O_{A1}O_{A2}}) \otimes |00\rangle\langle 00|_{O_{B1}O_{B2}}\,, \\
\Pi^{-}_{\text{mismatch}} &= |00\rangle\langle 00|_{O_{A1}O_{A2}} \otimes (\mathbb{1}_{O_{B1}O_{B2}} - |00\rangle\langle 00|_{O_{B1}O_{B2}}), \\
\Pi^{?}_{\text{mismatch}} &= |00\rangle\langle 00|_{O_{A1}O_{A2}} \otimes |00\rangle\langle 00|_{O_{B1}O_{B2}}\,, \\
\Pi^{d}_{\text{mismatch}} &= (\mathbb{1}_{O_{A1}O_{A2}} - |00\rangle\langle 00|_{O_{A1}O_{A2}}) \otimes (\mathbb{1}_{O_{B1}O_{B2}} - |00\rangle\langle 00|_{O_{B1}O_{B2}}).
\end{aligned} \tag{B.13}$$

Then, for $|\boldsymbol{\alpha}\rangle = |\alpha_A, \alpha_B\rangle$ and $|\boldsymbol{\beta}\rangle = |\beta_A, \beta_B\rangle$, we first define

$$\begin{aligned}|\tilde{\alpha}_{\text{final}}\rangle &= \left|\frac{\eta^{1/4}(\alpha_A + \sqrt{V}\alpha_B e^{i\delta})}{\sqrt{2}}, \frac{\eta^{1/4}\sqrt{1-V}\alpha_B e^{i\delta}}{\sqrt{2}}, \frac{\eta^{1/4}(\alpha_A - \sqrt{V}\alpha_B e^{i\delta})}{\sqrt{2}}, -\frac{\eta^{1/4}\sqrt{1-V}\alpha_B e^{i\delta}}{\sqrt{2}}\right\rangle \\ \left|\tilde{\beta}_{\text{final}}\right\rangle &= \left|\frac{\eta^{1/4}(\beta_A + \sqrt{V}\beta_B e^{i\delta})}{\sqrt{2}}, \frac{\eta^{1/4}\sqrt{1-V}\beta_B e^{i\delta}}{\sqrt{2}}, \frac{\eta^{1/4}(\beta_A - \sqrt{V}\beta_B e^{i\delta})}{\sqrt{2}}, -\frac{\eta^{1/4}\sqrt{1-V}\beta_B e^{i\delta}}{\sqrt{2}}\right\rangle\end{aligned} \tag{B.14}$$

We evaluate $\langle \boldsymbol{\alpha}| F^{\gamma}_{\text{mismatch}} |\boldsymbol{\beta}\rangle$ as

$$\langle \boldsymbol{\alpha}| F^{\gamma}_{\text{mismatch}} |\boldsymbol{\beta}\rangle = \langle \tilde{\alpha}_{\text{final}}| \Pi^{\gamma}_{\text{mismatch}} \left|\tilde{\beta}_{\text{final}}\right\rangle \left\langle \sqrt{1-\sqrt{\eta}}\alpha_A \middle| \sqrt{1-\sqrt{\eta}}\beta_A \right\rangle_{E_A} \left\langle \sqrt{1-\sqrt{\eta}}\alpha_B \middle| \sqrt{1-\sqrt{\eta}}\beta_B \right\rangle_{E_B}, \tag{B.15}$$

where $\eta = \eta_t \eta_d^2$ and $\eta_t = 10^{-\frac{0.2L}{10}}$ for a distance L in km.

Now, we write down $F^{\gamma}_{\text{mismatch}}$ in the basis $\mathcal{B}$. For the ease of representation, we define $\xi = e^{-\sqrt{\eta}\mu}$ and $\Omega = e^{-2(1-\sqrt{\eta})\mu}$ as before. We then define

$$\begin{aligned}
a &= (1-\xi^{(1+\sqrt{V}\cos\delta)})\xi^{(1-\sqrt{V}\cos\delta)}, \\
b &= (\xi^{2(1+\sqrt{V}\cos\delta)} - \xi^{(1+\sqrt{V}\cos\delta)})\xi^{(1-\sqrt{V}\cos\delta)}\Omega^2, \\
c &= (\xi^{1+i\sqrt{V}\sin\delta} - \xi)\xi\Omega, \\
d &= (\xi^{1-i\sqrt{V}\sin\delta} - \xi)\xi\Omega, \\
o &= (1-\xi^{(1-\sqrt{V}\cos\delta)})\xi^{(1+\sqrt{V}\cos\delta)}, \\
p &= (\xi^{2(1-\sqrt{V}\cos\delta)} - \xi^{(1-\sqrt{V}\cos\delta)})\xi^{(1+\sqrt{V}\cos\delta)}\Omega^2, \\
q &= (\xi^{1+i\sqrt{V}\sin\delta} - \xi)(\xi^{1-i\sqrt{V}\sin\delta} - \xi)\Omega, \\
m &= (1-\xi^{(1+\sqrt{V}\cos\delta)})(1-\xi^{(1-\sqrt{V}\cos\delta)}), \\
n &= (\xi^{2(1+\sqrt{V}\cos\delta)} - \xi^{(1+\sqrt{V}\cos\delta)})(\xi^{2(1-\sqrt{V}\cos\delta)} - \xi^{(1-\sqrt{V}\cos\delta)})\Omega^2.
\end{aligned} \tag{B.16}$$

And, c_0 and c_1 are again defined from the decomposition in Eq. (B.1).

$$F^{+}_{\text{mismatch}} = \begin{pmatrix} \frac{a+b+2c+2d+o+p}{8c_0^4} & \frac{a+b-o-p}{8c_0^2c_1^2} & 0 & 0 \\ \frac{a+b-o-p}{8c_0^2c_1^2} & \frac{a+b-2c-2d+o+p}{8c_1^4} & 0 & 0 \\ 0 & 0 & \frac{a-b+o-p}{8c_0^2c_1^2} & \frac{a-b+2c-2d-o+p}{8c_0^2c_1^2} \\ 0 & 0 & \frac{a-b-2c+2d-o+p}{8c_0^2c_1^2} & \frac{a-b+o-p}{8c_0^2c_1^2} \end{pmatrix} \tag{B.17}$$

$$F^{-}_{\text{mismatch}} = \begin{pmatrix} \frac{a+b+2c+2d+o+p}{8c_0^4} & -\frac{a+b-o-p}{8c_0^2c_1^2} & 0 & 0 \\ -\frac{a+b-o-p}{8c_0^2c_1^2} & \frac{a+b-2c-2d+o+p}{8c_1^4} & 0 & 0 \\ 0 & 0 & \frac{a-b+o-p}{8c_0^2c_1^2} & -\frac{a-b+2c-2d-o+p}{8c_0^2c_1^2} \\ 0 & 0 & -\frac{a-b-2c+2d-o+p}{8c_0^2c_1^2} & \frac{a-b+o-p}{8c_0^2c_1^2} \end{pmatrix} \tag{B.18}$$

$$F^{?}_{\text{mismatch}} = \xi^2 \begin{pmatrix} \frac{(1+\Omega)^2}{4c_0^4} & 0 & 0 & 0 \\ 0 & \frac{(1-\Omega)^2}{4c_1^4} & 0 & 0 \\ 0 & 0 & \frac{1-\Omega^2}{4c_0^2c_1^2} & 0 \\ 0 & 0 & 0 & \frac{1-\Omega^2}{4c_0^2c_1^2} \end{pmatrix} \tag{B.19}$$

$$F^{d}_{\text{mismatch}} = \begin{pmatrix} \frac{m+n+2q}{4c_0^4} & 0 & 0 & 0 \\ 0 & \frac{m+n-2q}{4c_1^4} & 0 & 0 \\ 0 & 0 & \frac{m-n}{4c_0^2c_1^2} & 0 \\ 0 & 0 & 0 & \frac{m-n}{4c_0^2c_1^2} \end{pmatrix}. \tag{B.20}$$

Finally, Eve's effective POVM corresponding to the mode mismatch, phase mismatch and dark counts of detectors is given as below

$$\begin{aligned} F^{+}_{\text{model}} &= (1-p_d)F^{+}_{\text{mismatch}} + (1-p_d)p_d F^{?}_{\text{mismatch}} \\ F^{-}_{\text{model}} &= (1-p_d)F^{-}_{\text{mismatch}} + (1-p_d)p_d F^{?}_{\text{mismatch}} \\ F^{?}_{\text{model}} &= (1-p_d)^2 F^{?}_{\text{mismatch}} \\ F^{d}_{\text{model}} &= p_d F^{+}_{\text{mismatch}} + p_d F^{-}_{\text{mismatch}} + p_d^2 F^{?}_{\text{mismatch}} + F^{d}_{\text{mismatch}}. \end{aligned} \tag{B.21}$$

Appendix C

Technical details of DMCVQKD

C.1 Mathematical description of a noisy heterodyne detector

C.1.1 Derivation

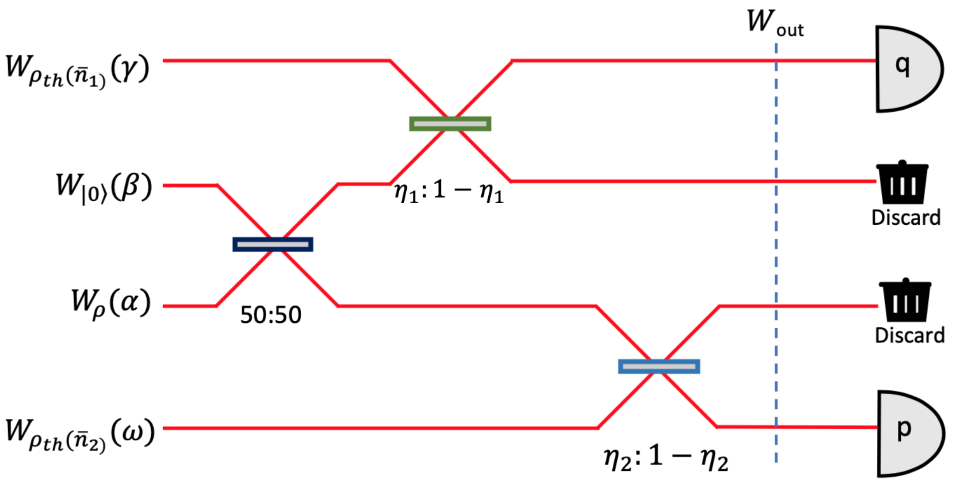


Figure C.1: A concise but equivalent view of the noisy heterodyne detector model depicted in Figure 6.2. Input modes are labeled in terms of Wigner functions.

As the physical model of a noisy heterodyne detector is presented in Figure 6.2, our goal here is to find the corresponding POVM elements that correctly produce the probability density function $P(y)$ of obtaining an outcome $y \in \mathbb{C}$ for an arbitrary input state ρ to the detector. In our trusted noise model, the homodyne detector for the q quadrature measurement has its detector efficiency η_1 and electronic noise ν_1 which is related to a thermal state of the mean photon number $\bar{n}_1 = \frac{\nu_1}{2(1-\eta_1)}$. Similarly, the homodyne detector for the p quadrature measurement has its detector efficiency η_2 and electronic noise ν_2 which corresponds to a thermal state with the mean photon number $\bar{n}_2 = \frac{v_2}{2(1-\eta_2)}$. Figure C.1 shows a compact but equivalent representation of Figure 6.2 with Wigner functions associated to input modes. In this setup, for an output state $W_{\text{out}}(\alpha, \beta, \gamma, \omega)$ at the step labeled in Figure C.1, we measure the q quadrature of the mode α and p quadrature of the mode β with two ideal homodyne detectors, and discard the rest modes γ and ω. The Wigner function of an ideal homodyne detector for the q quadrature measurement that produces a measurement outcome $\text{Re}(y)$ is $W_{H_{\text{Re}(y)}}(\alpha) = \frac{1}{\sqrt{2}\pi}\delta(\text{Re}(\alpha) - \frac{\text{Re}(y)}{\sqrt{2}})$ where δ is the Dirac delta function and similarly, the one for the p quadrature measurement with a measurement outcome $\text{Im}(y)$ is $W_{H_{\text{Im}(y)}}(\alpha) = \frac{1}{\sqrt{2}\pi}\delta(\text{Im}(\alpha) - \frac{\text{Im}(y)}{\sqrt{2}})$. The factors of $\sqrt{2}$ are included such that we can rederive the ideal heterodyne detector POVM $\{E_y : y \in \mathbb{C}\}$ in the limit of unity detector efficiency and zero electronic noise. To discard modes γ and ω that are not measured, we perform the integration over variables γ and ω.

For any input state ρ to the detector, one can in principle obtain the underlying probability density function $P(y) = \text{Tr}(\rho G_y)$ for every measurement outcome $y \in \mathbb{C}$. As the correct POVM element G_y needs to produce the observed probability density function $P(y) = \text{Tr}(\rho G_y)$, this requirement in terms of Wigner functions becomes $P(y) = \pi \int d^2\alpha W_\rho(\alpha) W_{G_y}(\alpha)$, where W_ρ is the Wigner function of the input state ρ and W_{G_y} is the Wigner function of the operator G_y, by the overlap formula in Eq. (2.75). In Figure C.1, we know the mathematical description of measurements on the right, but the description of the state W_{out} is unknown. On the other hand, we want to find the description of the measurement directly acting on the input state and the Wigner function description of the input state and those of ancillary modes on the left are either assumed to be given or known. To connect these known descriptions on the two sides of this diagram to find the desired Wigner function of the POVM element G_y that acts on the input state directly, we start from the right-hand side of this diagram with an unknown four-mode state W_{out} and the known measurements on these modes, perform inverse beam splitter transformations from right to left of this diagram and finally obtain W_{G_y} by integrating over variables other than α. By starting with the multi-mode overlap formula for $P(y)$ on the right-hand side of the diagram and performing the process as described, we obtain

$$
\begin{aligned}
P(y) = \pi^4 \int d^2\alpha \int d^2\beta \int d^2\gamma \int d^2\omega \; \frac{1}{\pi^2} W_{\mathrm{out}}(\alpha,\beta,\gamma,\omega) W_{H_{\mathrm{Re}(y)}}(\alpha) W_{H_{\mathrm{Im}(y)}}(\beta) \\
= \pi^2 \int d^2\alpha \; W_{\rho}(\alpha) \int d^2\beta \; W_{|0\rangle}(\beta) \int d^2\gamma \; W_{\rho_{\mathrm{th}}(\bar{n}_1)}(\gamma) W_{H_{\mathrm{Re}(y)}}(\sqrt{\eta_1}\frac{\alpha+\beta}{\sqrt{2}} + \sqrt{1-\eta_1}\gamma) \\
\times \int d^2\omega \; W_{\rho_{\mathrm{th}}(\bar{n}_2)}(\omega) W_{H_{\mathrm{Im}(y)}}(\sqrt{\eta_2}\frac{\alpha-\beta}{\sqrt{2}} + \sqrt{1-\eta_2}\omega).
\end{aligned}
\tag{C.1}
$$

The next step is to substitute the Wigner function of the vacuum state in Eq. (2.77) and that of the thermal state in Eq. (2.78) and then to perform the integrals over variables β, γ and ω. We first integrate over the variable ω. The relevant integral that involves the variable ω is

$$
\begin{aligned}
&\int d^2\omega \; W_{\rho_{\mathrm{th}}(\bar{n}_2)}(\omega) W_{H_{\mathrm{Im}(y)}}(\sqrt{\eta_2}\frac{\alpha-\beta}{\sqrt{2}} + \sqrt{1-\eta_2}\omega) \\
=&\frac{1}{\pi\sqrt{\pi}} \frac{1}{\sqrt{(1-\eta_2)(1+2\bar{n}_2)}} \exp\left(-\frac{\eta_2[\mathrm{Im}(\beta) + \frac{1}{\sqrt{\eta_2}}\mathrm{Im}(y) - \mathrm{Im}(\alpha)]^2}{(1+2\bar{n}_2)(1-\eta_2)}\right).
\end{aligned}
\tag{C.2}
$$

Next, we perform the integral related to the variable γ. Since Eq. (C.2) does not involve the variable γ, we do not need to plug it back to solve the integral that involves the variable γ. This integration shown in Eq. (C.3) is actually similar to the integration that we just did in Eq. (C.2).

$$
\begin{aligned}
&\int d^2\gamma \; W_{\rho_{\mathrm{th}}(\bar{n}_1)}(\gamma) W_{H_{\mathrm{Re}(y)}}(\sqrt{\eta_1}\frac{\alpha+\beta}{\sqrt{2}} + \sqrt{1-\eta_1}\gamma) \\
=&\frac{1}{\pi\sqrt{\pi}} \frac{1}{\sqrt{(1-\eta_1)(1+2\bar{n}_1)}} \exp\left(-\frac{\eta_1\left[\mathrm{Re}(\beta) - \frac{1}{\sqrt{\eta_1}}\mathrm{Re}(y) + \mathrm{Re}(\alpha)\right]^2}{(1+2\bar{n}_1)(1-\eta_1)}\right).
\end{aligned}
\tag{C.3}
$$

Finally, we integrate over the variable β. We now need to substitute results of Eqs. (C.2) and (C.3) back to Eq. (C.1). The prefactor is simplified to be $\frac{1}{\pi^3}\frac{1}{\sqrt{(1-\eta_1)(1+2\bar{n}_1)(1-\eta_2)(1+2\bar{n}_2)}}$. Except this prefactor, we perform the following integral

$$
\begin{aligned}
&\int d^2\beta \; W_{|0\rangle}(\beta) \exp\left(-\frac{\eta_1\left[\mathrm{Re}(\beta) - \frac{1}{\sqrt{\eta_1}}\mathrm{Re}(y) + \mathrm{Re}(\alpha)\right]^2}{(1+2\bar{n}_1)(1-\eta_1)}\right) \exp\left(-\frac{\eta_2\left[(\mathrm{Im}(\beta) + \frac{1}{\sqrt{\eta_2}}\mathrm{Im}(y) - \mathrm{Im}(\alpha)\right]^2}{(1+2\bar{n}_2)(1-\eta_2)}\right) \\
=&2\sqrt{\frac{(1+2\bar{n}_1)(1+2\bar{n}_2)(1-\eta_1)(1-\eta_2)}{(1+(1-\eta_1)+4\bar{n}_1(1-\eta_1))(1+(1-\eta_2)+4\bar{n}_2(1-\eta_2))}} \\
&\times \exp\left(\frac{-2\eta_1\left[\frac{1}{\sqrt{\eta_1}}\mathrm{Re}(y) - \mathrm{Re}(\alpha)\right]^2}{1+(1-\eta_1)+4\bar{n}_1(1-\eta_1)} + \frac{-2\eta_2\left[\frac{1}{\sqrt{\eta_2}}\mathrm{Im}(y) - \mathrm{Im}(\alpha)\right]^2}{1+(1-\eta_2)+4\bar{n}_2(1-\eta_2)}\right).
\end{aligned}
\tag{C.4}
$$

Finally, by putting the prefactor back and expressing the final expression in a format of Gaussian functions, we obtain the following result

$$\begin{aligned} W_{G_y}(\alpha) =& \frac{1}{\sqrt{\eta_1\eta_2}\pi}\frac{2}{\pi}\frac{1}{\sqrt{1+\frac{2(1-\eta_1)(1+2\bar{n}_1)}{\eta_1}}}\frac{1}{\sqrt{1+\frac{2(1-\eta_2)(1+2\bar{n}_2)}{\eta_2}}} \\ &\times \exp\left(\frac{-2[\frac{1}{\sqrt{\eta_1}}\operatorname{Re}(y)-\operatorname{Re}(\alpha)]^2}{1+\frac{2(1-\eta_1)(1+2\bar{n}_1)}{\eta_1}}+\frac{-2[\frac{1}{\sqrt{\eta_2}}\operatorname{Im}(y)-\operatorname{Im}(\alpha)]^2}{1+\frac{2(1-\eta_2)(1+2\bar{n}_2)}{\eta_2}}\right). \end{aligned} \tag{C.5}$$

By substituting in $\bar{n}_1 = \frac{\nu_1}{2(1-\eta_1)}$ and $\bar{n}_2 = \frac{\nu_2}{2(1-\eta_2)}$, we derive Eq. (6.3) after a straightforward simplification.

C.1.2 POVM elements in the general case

As we derive the Wigner function of an arbitrary POVM element G_y corresponding to the detector model in Figure 6.2, we next show that the POVM elements G_y's are projections onto displaced squeezed thermal states up to a scaling factor. To see this, we make the following definitions:

$$\begin{aligned} \lambda_j &:= \frac{(1-\eta_j)(1+2\bar{n}_j)}{\eta_j} = \frac{1-\eta_j+\nu_j}{\eta_j} \quad \text{for } j=1,2, \\ \bar{n}_{het} &:= \frac{\sqrt{(1+2\lambda_1)(1+2\lambda_2)}-1}{2}, \\ \xi_{het} &:= \frac{1}{4}\ln\left(\frac{1+2\lambda_2}{1+2\lambda_1}\right), \\ \alpha_{het} &:= \frac{1}{\sqrt{\eta_1}}\operatorname{Re}(y)+\frac{i}{\sqrt{\eta_2}}\operatorname{Im}(y). \end{aligned} \tag{C.6}$$

With these choices of parameters α_{het}, ξ_{het} and $\bar{n}_{het}$, Eq. (C.5) can be rewritten as

$$W_{G_y}(\gamma) = \frac{1}{\sqrt{\eta_1\eta_2}\pi}\frac{2}{\pi}\frac{1}{1+2\bar{n}_{het}}\exp\left\{-2[\frac{e^{2\xi_{het}}\operatorname{Re}(\gamma-\alpha_{het})^2+e^{-2\xi_{het}}\operatorname{Im}(\gamma-\alpha_{het})^2}{1+2\bar{n}_{het}}]\right\}. \tag{C.7}$$

By comparing the Wigner function of G_y in Eq. (C.5) and the Wigner function of a displaced squeezed thermal state in Eq. (2.81), we can identify $G_y = \frac{1}{\sqrt{\eta_1\eta_2}\pi}\rho_{\text{DSTS}}(\alpha_{het}, \xi_{het}, \bar{n}_{het})$ for the choices of parameters α_{het}, ξ_{het} and $\bar{n}_{het}$ in Eq. (C.6). Therefore, each G_y is a scaled projection onto a displaced squeezed thermal state with the displacement α_{het}, squeezing parameter ξ_{het} and the mean photon number of the initial thermal state before squeezing and displacement $\bar{n}_{het}$.

C.1.3 POVM elements in the simple case

If $\eta_1 = \eta_2 = \eta_d$ and $\nu_1 = \nu_2 = \nu_{\text{el}}$, it is easy to verify that Eq. (C.5) reduces to Eq. (6.3). Then one can identify each POVM element G_y is the projection onto a displaced thermal state with a prefactor $1/(\eta_d \pi)$ in Eq. (6.5). An alternative view is to look at the parameters α_{het}, ξ_{het} and $\bar{n}_{het}$ in Eq. (C.6). In particular, since $\lambda_1 = \lambda_2$, the amount of squeezing ξ_{het} becomes zero. Thus, by neglecting squeezing in the POVM elements of the general case, one can also conclude each POVM element is proportional to the projection onto a displaced thermal state. One can further verify that the displacement is $\alpha_{het} = \frac{y}{\sqrt{\eta_d}}$ and the mean photon number of the initial thermal state $\bar{n}_{het}$ becomes $\frac{1-\eta_d+\nu_{\text{el}}}{\eta_d}$.

C.2 Photon-number basis representation of operators

Let $\mathcal{N}$ denote the photon-number basis up to N_c, that is, $\mathcal{N} = \{|0\rangle, \dots, |N_c\rangle\}$. To analyze DMCVQKD protocols under the photon-number cutoff assumption, that is, $\rho_{AB} = (\mathbb{1}_A \otimes \Pi_{N_c})\rho_{AB}(\mathbb{1}_A \otimes \Pi_{N_c})$ where Π_{N_c} is the projection onto the subspace spanned by the basis $\mathcal{N}$., we need to represent operators as well as observables used in the key rate optimization problems [see Eqs. (6.11), (6.16) and (7.4)] in this photon-number basis $\mathcal{N}$. When these operators are represented in this finite-dimensional Hilbert space, we can then proceed with numerical optimization to calculate the key rate. Since Alice's system is irrelevant for our discussion here, we focus on the conditional states ρ_B^x in the following discussion. For any operator $\hat{O}$ acting on Bob's system, we observe that

$$\begin{aligned} \mathrm{Tr}\left[\rho_B^x \hat{O}\right] &= \mathrm{Tr}\left[\Pi_{N_c}\rho_B^x\Pi_{N_c}\hat{O}\right] \\ &= \mathrm{Tr}\left[(\Pi_{N_c}\rho_B^x\Pi_{N_c})(\Pi_{N_c}\hat{O}\Pi_{N_c})\right]. \end{aligned} \tag{C.8}$$

This observation allows us to define the truncated version of the operator $\hat{O}$ by $\Pi_{N_c}\hat{O}\Pi_{N_c}$.

In Appendix C.2.1, we discuss those operators in the ideal detector scenarios for both the heterodyne protocol (Chapter 6) and the homodyne protocol (Chapter 7). We then focus on the trusted detector noise scenario for the heterodyne protocol in later subsections. In Appendix C.2.2, our discussion in the trusted detector noise scenario is restricted to the simplified scenario that we use in the main text for presenting simulation results; that is, we set $\eta_1 = \eta_2 = \eta_d$ and $\nu_1 = \nu_2 = \nu_{\text{el}}$. Under this scenario, we present formulae that can be efficiently evaluated in Matlab. We then discuss the general case where the imperfections in two homodyne detectors are not necessarily the same in Appendix C.2.3. For the general case, we provide the matrix representation of the POVM elements G_y's and leave the evaluation of region operators and observables for the optimization problem to be done numerically.

C.2.1 Ideal detector scenario

In our optimization problem [see Eqs. (6.11) and (7.4)], the relevant operators are of the forms $\Pi_{N_c}\rho_B^x\Pi_{N_c}$ and $\Pi_{N_c}\hat{O}\Pi_{N_c}$, which have finite-dimensional matrix representations. Specifically, we can find a matrix representation of $\hat{O}$ in the basis $\mathcal{N}$. We start by writing out the annihilation operator $\hat{a}$ in this basis, and then the creation operator $\hat{a}^\dagger$ is just its conjugate transpose. Consequently, other relevant operators $\hat{q}, \hat{p}$, $\hat{n}$, and $\hat{d}$ can be written directly following from their definitions in terms of $\hat{a}$ and $\hat{a}^\dagger$. In this basis,

$$\Pi_{N_c}\hat{a}\Pi_{N_c} = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \sqrt{2} & 0 & \cdots & 0 \\ \vdots & & \ddots & & & \vdots \\ 0 & & \cdots & & 0 & \sqrt{N_c} \\ 0 & & \cdots & & 0 & 0 \end{pmatrix}. \tag{C.9}$$

It is not difficult to write out the interval operators I_0 and I_1 and region operators R_0, R_1, R_2, and R_3 in this basis. To do so, we use the overlap $\langle q|n\rangle$ between a quadrature eigenstate $|q\rangle$ and a photon-number state $|n\rangle$ and the overlap $\left\langle \gamma e^{i\theta} \middle| n \right\rangle$ between a coherent state $\left|\gamma e^{i\theta}\right\rangle$ and a photon-number state $|n\rangle$. With our definition of quadrature operators in Eq. (2.47), the overlaps $\langle q|n\rangle$ and $\left\langle \gamma e^{i\theta} \middle| n \right\rangle$ read [216]

$$\begin{aligned} \langle q|n\rangle &= \frac{1}{\sqrt{\pi^{\frac{1}{2}}2^n(n!)}}\exp\left(-\frac{q^2}{2}\right)H_n(q), \\ \left\langle \gamma e^{i\theta} \middle| n \right\rangle &= e^{-\frac{\gamma^2}{2}}\frac{\gamma^n e^{-in\theta}}{\sqrt{n!}}, \end{aligned} \tag{C.10}$$

respectively, where H_n is the Hermite polynomial of the order of n. We then perform the relevant integrals to obtain a finite-dimensional matrix representation in this basis.

Finally, in the expressions of Kraus operators shown in Eqs. (6.9) and (7.5), we need to take the square root of each of region operators R_0, R_1, R_2, and R_3 or the interval operators I_0 and I_1 in Eq. (7.6). A caution about the ordering of truncation and square root is needed. For example, even though the interval operators are projective on the entire infinite-dimensional space such that the square root of each operator is identical to itself, the truncated version of each interval operator is no longer projective in the finite-dimensional subspace spanned by the basis $\mathcal{N}$. We now explain the proper way to handle this issue. With the photon-number cutoff assumption $\rho = \Pi_{N_c}\rho\Pi_{N_c}$, we see from Eq. (C.8) that, for a POVM element F on the infinite-dimensional space, the corresponding POVM element on this finite-dimensional subspace becomes $\Pi_{N_c}F\Pi_{N_c}$. As we know from Appendix A, the purpose of taking the square root of a POVM element is to realize this POVM measurement in an isometric fashion. Since the relevant POVM element on this finite-dimensional subspace is $\Pi_{N_c} F \Pi_{N_c}$, we need to take the square root of $\Pi_{N_c} F \Pi_{N_c}$. Therefore, we first take the truncation and then take the square root.

C.2.2 Trusted detector noise scenario with the simple case

In this simple case where two homodyne detectors in the heterodyne detection scheme have the same imperfection, the POVM element G_y given in Eq. (6.5) in the photon-number basis is expressed as [199]

$$\langle m| G_y |n\rangle = \frac{1}{\eta_d \pi} \exp\left[-\frac{|y|^2}{\eta_d(1+\bar{n}_d)}\right] \frac{\bar{n}_d^m}{(1+\bar{n}_d)^{n+1}} (\frac{y^*}{\sqrt{\eta_d}})^{n-m} (\frac{m!}{n!})^{1/2} L_m^{(n-m)}\Big(-\frac{|y|^2}{\eta_d \bar{n}_d (1+\bar{n}_d)}\Big), \tag{C.11}$$

where we define $\bar{n}_d = \frac{1-\eta_d+\nu_{\text{el}}}{\eta_d}$ for ease of writing and $L_k^{(j)}(x)$ is the generalized Laguerre polynomial of degree k with a parameter j in the variable x. In particular, the diagonal entries are simplified to be

$$\langle n| G_y |n\rangle = \frac{1}{\eta_d \pi} \exp\left[-\frac{|y|^2}{\eta_d(1+\bar{n}_d)}\right] \frac{\bar{n}_d^n}{(1+\bar{n}_d)^{n+1}} L_n(-\frac{|y|^2}{\eta_d \bar{n}_d (1+\bar{n}_d)}), \tag{C.12}$$

where $L_k(x) = L_k^{(0)}(x)$ is the Laguerre polynomial of degree k in the variable x. For ease of writing later, we define $C_{m,n} = \frac{1}{\pi \eta_d^{(n-m)/2+1}} (\frac{m!}{n!})^{1/2} \frac{\bar{n}_d^m}{(1+\bar{n}_d)^{n+1}}$.

Region operators

Our goal here is to write region operators $R_j = \int_{y \in \mathcal{A}_j} G_y \; d^2y$ in the photon-number basis. For simplicity, we work out the expressions in the absence of postselection. To include the postselection, one may numerically integrate over the discarded region and subtract this result from the expression without postselection since this numerical integration is efficiently computable in MATLAB. We first consider off-diagonal elements (i.e. $m \neq n$). In this case, we plug the expression of $\langle m| G_y |n\rangle$ in Eq. (C.11) into the definition of R_j in Eq. (6.12), write it in the polar coordinate with $y = re^{i\theta}$ and perform the integration over the phase θ to obtain the following expressions.

$$\begin{aligned}\langle m| R_0 |n\rangle &= \int_0^\infty dr\, r \int_{-\pi/4}^{\pi/4} d\theta \;\, \langle m| G_{re^{i\theta}} |n\rangle \\ &= C_{m,n} \frac{2\sin[\frac{m-n}{4}\pi]}{m-n} \int_0^\infty dr \;\, \exp\Big[-\frac{r^2}{\eta_d(1+\bar{n}_d)}\Big] L_m^{(n-m)}\Big(-\frac{r^2}{\eta_d \bar{n}_d (1+\bar{n}_d)}\Big) r^{n-m+1}.\end{aligned} \tag{C.13}$$

$$
\begin{aligned}
\langle m| R_1 |n\rangle &= \int_0^\infty dr\ r \int_{\pi/4}^{3\pi/4} d\theta\ \langle m| G_{re^{i\theta}} |n\rangle \\
&= C_{m,n} \frac{i(1-i^{m-n})e^{i(m-n)\pi/4}}{m-n} \int_0^\infty dr\ \exp\Big[-\frac{r^2}{\eta_d(1+\bar{n}_d)}\Big] L_m^{(n-m)}\Big(-\frac{r^2}{\eta_d \bar{n}_d(1+\bar{n}_d)}\Big) r^{n-m+1}.
\end{aligned} \tag{C.14}
$$

$$
\begin{aligned}
\langle m| R_2 |n\rangle &= \int_0^\infty dr\ r \int_{3\pi/4}^{5\pi/4} d\theta\ \langle m| G_{re^{i\theta}} |n\rangle \\
&= C_{m,n} \frac{2(-1)^{m+n}\sin[\frac{m-n}{4}\pi]}{m-n} \int_0^\infty dr\ \exp\Big[-\frac{r^2}{\eta_d(1+\bar{n}_d)}\Big] L_m^{(n-m)}\Big(-\frac{r^2}{\eta_d \bar{n}_d(1+\bar{n}_d)}\Big) r^{n-m+1}.
\end{aligned} \tag{C.15}
$$

$$
\begin{aligned}
\langle m| R_3 |n\rangle &= \int_0^\infty dr\ r \int_{5\pi/4}^{7\pi/4} d\theta\ \langle m| G_{re^{i\theta}} |n\rangle \\
&= C_{m,n} \frac{i(1-i^{m-n})e^{5i(m-n)\pi/4}}{m-n} \int_0^\infty dr\ \exp\Big[-\frac{r^2}{\eta_d(1+\bar{n}_d)}\Big] L_m^{(n-m)}\Big(-\frac{r^2}{\eta_d \bar{n}_d(1+\bar{n}_d)}\Big) r^{n-m+1}.
\end{aligned} \tag{C.16}
$$

We notice that the integral over r is a common integral in these four equations. We now perform this common integral to obtain the result in terms of Taylor series expansion of a simple function.

$$
\begin{aligned}
&\int_0^\infty dr\ \exp\Big[-\frac{r^2}{\eta_d(1+\bar{n}_d)}\Big] L_m^{(n-m)}\Big(-\frac{r^2}{\eta_d \bar{n}_d(1+\bar{n}_d)}\Big) r^{n-m+1} \\
=\ & \frac{1}{2}[\eta_d(1+\bar{n}_d)]^{\frac{n-m}{2}+1}\Gamma(\frac{n-m}{2}+1) f_m(\bar{n}_d, n-m, \frac{n-m}{2}),
\end{aligned} \tag{C.17}
$$

where Γ is the gamma function and $f_m(a,\alpha,k)$ is defined as the Taylor series coefficients of the function below in the variable t as

$$
(1-t)^{-\alpha+k}(1-(1+\frac{1}{a})t)^{-(k+1)} = \sum_{n=0}^{\infty} f_n(a,\alpha,k)t^n. \tag{C.18}
$$

We note that the Taylor series coefficients here can be quickly found in MATLAB.

Now, we consider the diagonal entries of R_j (i.e. $m = n$). By substituting $y = re^{i\theta}$ in Eq. (C.12), we note that this expression does not depend on θ. Thus, it is easy to see $\langle n| R_0 |n\rangle = \langle n| R_1 |n\rangle = \langle n| R_2 |n\rangle = \langle n| R_3 |n\rangle$. The integration over the phase θ gives a factor of $\frac{\pi}{2}$. We proceed the integration over variable r and obtain

$$\begin{aligned}\langle n| R_j |n\rangle &= \frac{\pi}{2}\frac{1}{\eta_d \pi}\int_0^\infty dr\, r \exp\Big[-\frac{r^2}{\eta(1+\bar{n}_d)}\Big]\frac{\bar{n}_d^n}{(1+\bar{n}_d)^{n+1}} L_n\Big(-\frac{r^2}{\eta_d\bar{n}_d(1+\bar{n}_d)}\Big) \\ &= \frac{1}{4}\frac{\bar{n}_d^n}{(1+\bar{n}_d)^n}(1+\frac{1}{\bar{n}_d})^n = \frac{1}{4}.\end{aligned} \tag{C.19}$$

To include postselection, the common integral in the case $m \neq n$ becomes

$$\begin{aligned}&\int_{\Delta_a}^\infty dr\ \exp\Big[-\frac{r^2}{\eta_d(1+\bar{n}_d)}\Big] L_m^{(n-m)}\Big(-\frac{r^2}{\eta_d\bar{n}_d(1+\bar{n}_d)}\Big) r^{n-m+1} \\ =&\frac{1}{2}[\eta_d(1+\bar{n}_d)]^{\frac{n-m}{2}+1}\Gamma(\frac{n-m}{2}+1) f_m(\bar{n}_d, n-m, \frac{n-m}{2}) \\ &-\int_0^{\Delta_a} dr\ \exp\Big[-\frac{r^2}{\eta_d(1+\bar{n}_d)}\Big] L_m^{(n-m)}\Big(-\frac{r^2}{\eta_d\bar{n}_d(1+\bar{n}_d)}\Big) r^{n-m+1},\end{aligned} \tag{C.20}$$

where the second term is efficiently computable numerically. The case for $m = n$ follows similarly.

First-moment observables

We then proceed to evaluate the matrix elements of $\hat{F}_Q$ and $\hat{F}_P$. In the photon-number basis, the matrix elements are

$$\begin{aligned}\langle m| \hat{F}_Q |n\rangle &= \int \frac{y+y^*}{\sqrt{2}} \langle m| G_y |n\rangle\ d^2y \\ &= \frac{C_{m,n}}{\sqrt{2}}\int_0^\infty dr\ \exp\Big[-\frac{r^2}{\eta_d(1+\bar{n}_d)}\Big] L_m^{(n-m)}\Big(-\frac{r^2}{\eta_d\bar{n}_d(1+\bar{n}_d)}\Big) r^{n-m+2} \\ &\times \int_0^{2\pi} d\theta\, e^{-i(n-m)\theta}(e^{i\theta}+e^{-i\theta}),\end{aligned} \tag{C.21}$$

$$\begin{aligned}\langle m| \hat{F}_P |n\rangle &= \int \frac{i(y^*-y)}{\sqrt{2}} \langle m| G_y |n\rangle\ d^2y \\ &= \frac{iC_{m,n}}{\sqrt{2}}\int_0^\infty dr\ \exp\Big[-\frac{r^2}{\eta_d(1+\bar{n}_d)}\Big] L_m^{(n-m)}\Big(-\frac{r^2}{\eta_d\bar{n}_d(1+\bar{n}_d)}\Big) r^{n-m+2} \\ &\times \int_0^{2\pi} d\theta\, e^{-i(n-m)\theta}(e^{-i\theta}-e^{i\theta}).\end{aligned} \tag{C.22}$$

As $\hat{F}_Q$ is a Hermitian operator, we can first find entries $\langle m| \hat{F}_Q |n\rangle$ for $m \leq n$. Then for $m > n$, we simply set $\langle m| \hat{F}_Q |n\rangle$ to be the complex conjugate of $\langle n| \hat{F}_Q |m\rangle$. From the integration over θ, the nonzero entries for $m \leq n$ are

$$\begin{aligned}\langle m|\,\hat{F}_Q\,|m+1\rangle &= \sqrt{2}\pi C_{m,m+1}\int_0^\infty dr\ \exp\Big[-\frac{r^2}{\eta_d(1+\bar{n}_d)}\Big]L_m^{(1)}\Big(-\frac{r^2}{\eta_d\bar{n}_d(1+\bar{n}_d)}\Big)r^3\\ &= \frac{\pi}{\sqrt{2}}C_{m,m+1}((1+\bar{n}_d)\eta_d)^2 f_m(\bar{n}_d,1,1).\end{aligned} \tag{C.23}$$

By a similar procedure for $\hat{F}_P$, we have

$$\begin{aligned}\langle m|\,\hat{F}_P\,|m+1\rangle &= -\sqrt{2}i\pi C_{m,m+1}\int_0^\infty dr\ \exp\Big[-\frac{r^2}{\eta_d(1+\bar{n}_d)}\Big]L_m^{(1)}\Big(-\frac{r^2}{\eta_d\bar{n}_d(1+\bar{n}_d)}\Big)r^3\\ &= -\frac{i\pi}{\sqrt{2}}C_{m,m+1}((1+\bar{n}_d)\eta_d)^2 f_m(\bar{n}_d,1,1).\end{aligned} \tag{C.24}$$

Second-moment observables

Next, we evaluate the matrix elements of $\hat{S}_Q$ and $\hat{S}_P$. In the photon-number basis, they are

$$\begin{aligned}\langle m|\,\hat{S}_Q\,|n\rangle &= \int (\frac{y+y^*}{\sqrt{2}})^2\,\langle m|\,G_y\,|n\rangle\, d^2y\\ &= \frac{C_{m,n}}{2}\int_0^\infty dr\ \exp\Big[-\frac{r^2}{\eta_d(1+\bar{n}_d)}\Big]L_m^{(n-m)}\Big(-\frac{r^2}{\eta_d\bar{n}_d(1+\bar{n}_d)}\Big)r^{n-m+3}\\ &\times\int_0^{2\pi} d\theta\ e^{-i(n-m)\theta}(e^{i\theta}+e^{-i\theta})^2,\end{aligned} \tag{C.25}$$

$$\begin{aligned}\langle m|\,\hat{S}_P\,|n\rangle &= \int (\frac{i(y^*-y)}{\sqrt{2}})^2\,\langle m|\,G_y\,|n\rangle\, d^2y\\ &= -\frac{C_{m,n}}{2}\int_0^\infty dr\ \exp\Big[-\frac{r^2}{\eta_d(1+\bar{n}_d)}\Big]L_m^{(n-m)}\Big(-\frac{r^2}{\eta_d\bar{n}_d(1+\bar{n}_d)}\Big)r^{n-m+3}\\ &\times\int_0^{2\pi} d\theta\ e^{-i(n-m)\theta}(e^{-i\theta}-e^{i\theta})^2.\end{aligned} \tag{C.26}$$

Again, since $\hat{S}_Q$ and $\hat{S}_P$ are Hermitian operators, we only need to define the upper triangular part and then set the lower triangular part using the Hermitian property. The relevant integrals are simplified to be

$$\begin{aligned}\langle m|\,\hat{S}_Q\,|m\rangle &= 2\pi C_{m,m}\int_0^\infty dr\;\exp\Big[-\frac{r^2}{\eta_d(1+\bar{n}_d)}\Big]L_m\Big(-\frac{r^2}{\eta_d\bar{n}_d(1+\bar{n}_d)}\Big)r^3\\ &= \pi C_{m,m}(\eta_d(1+\bar{n}_d))^2 f_m(\bar{n}_d,0,1),\\ \langle m|\,\hat{S}_Q\,|m+2\rangle &= \pi C_{m,m+2}\int_0^\infty dr\;\exp\Big[-\frac{r^2}{\eta_d(1+\bar{n}_d)}\Big]L_m^{(2)}\Big(-\frac{r^2}{\eta_d\bar{n}_d(1+\bar{n}_d)}\Big)r^5\\ &= \pi C_{m,m+2}(\eta_d(1+\bar{n}_d))^3 f_m(\bar{n}_d,2,2).\end{aligned} \tag{C.27}$$

For $\hat{S}_P$, we have

$$\begin{aligned}\langle m|\,\hat{S}_P\,|m\rangle &= 2\pi C_{m,m}\int_0^\infty dr\;\exp\Big[-\frac{r^2}{\eta_d(1+\bar{n}_d)}\Big]L_m\Big(-\frac{r^2}{\eta_d\bar{n}_d(1+\bar{n}_d)}\Big)r^3\\ &= \pi C_{m,m}(\eta_d(1+\bar{n}_d))^2 f_m(\bar{n}_d,0,1),\\ \langle m|\,\hat{S}_P\,|m+2\rangle &= -\pi C_{m,m+2}\int_0^\infty dr\;\exp\Big[-\frac{r^2}{\eta_d(1+\bar{n}_d)}\Big]L_m^{(2)}\Big(-\frac{r^2}{\eta_d\bar{n}_d(1+\bar{n}_d)}\Big)r^5\\ &= -\pi C_{m,m+2}(\eta_d(1+\bar{n}_d))^3 f_m(\bar{n}_d,2,2).\end{aligned} \tag{C.28}$$

C.2.3 Trusted detector noise scenario in the general case

We consider the general case where two homodyne detectors may have different imperfections. In this case, each POVM element G_y is given in Eq. (C.7). Given the POVM

$$G_y = \frac{1}{\sqrt{\eta_1\eta_2}\pi}\rho_{\text{DSTS}}(\alpha_{het},\xi_{het},\bar{n}_{het}),$$

its matrix elements are given by [212, Eq. (5.2)] with the prefactor $\frac{1}{\sqrt{\eta_1\eta_2}\pi}$ as

$$\begin{aligned}\langle m|\,G_y\,|n\rangle = \frac{1}{\sqrt{\eta_1\eta_2}}\frac{Q(0)}{\sqrt{m!n!}}\sum_{k=0}^{\min(m,n)} k!\binom{m}{k}\binom{n}{k}\tilde{A}^k\Big(\frac{\tilde{B}}{2}\Big)^{\frac{m-k}{2}}\Big(\frac{\tilde{B}^*}{2}\Big)^{\frac{n-k}{2}}\\ \times H_{m-k}((2\tilde{B})^{-\frac{1}{2}}\tilde{C})H_{n-k}((2\tilde{B}^*)^{-\frac{1}{2}}\tilde{C}^*),\end{aligned} \tag{C.29}$$

where H_ℓ is the Hermite polynomial of order ℓ. With simple substitutions, one may verify that these parameters $\tilde{A}$, $\tilde{B}$, $\tilde{C}$ and $Q(0)$ are defined in terms of $\lambda_1, \lambda_2, \alpha_{het}$ in Eq. (C.6) as

$$\begin{aligned}
\tilde{A} &:= 1 - \frac{\lambda_1 + \lambda_2 + 2}{2(\lambda_1 + 1)(\lambda_2 + 1)}, \\
\tilde{B} &:= \frac{-|\lambda_1 - \lambda_2|}{2(\lambda_1 + 1)(\lambda_2 + 1)}, \\
\tilde{C} &:= \frac{\mathrm{Re}(\alpha_{het})}{\max(\lambda_1, \lambda_2) + 1} + i\frac{\mathrm{Im}(\alpha_{het})}{\min(\lambda_1, \lambda_2) + 1}, \\
Q(0) &:= \frac{1}{\pi}\frac{1}{\sqrt{(\lambda_1 + 1)(\lambda_2 + 1)}} \exp\left[-\frac{\mathrm{Re}(\alpha_{het})^2}{\max(\lambda_1, \lambda_2) + 1} - \frac{\mathrm{Im}(\alpha_{het})^2}{\min(\lambda_1, \lambda_2) + 1}\right].
\end{aligned} \tag{C.30}$$

As indicated in [212, Eqs. (5.3) and (5.4)], the choice of square roots of $\tilde{B}$ is as $\tilde{B}^{1/2} = ie^{i(\varphi/2)}\left|\tilde{B}\right|^{1/2}$ and $(\tilde{B}^*)^{1/2} = (\tilde{B}^{1/2})^*$, where $\varphi = 0$ if $\lambda_1 \leq \lambda_2$ and $\varphi = \pi$ if $\lambda_1 > \lambda_2$. We note that $\bar{n}_{het}$ and ξ_{het} are defined in terms of λ_1 and λ_2 and one may rewrite these parameters in terms of $\bar{n}_{het}, \xi_{het}$ and α_{het} to make the matrix elements more explicitly depend on the parameters of the displaced squeezed thermal states.

From the expression of $\langle m| G_y |n\rangle$ in Eq. (C.29), one can apply the definition of region operators R_j's in Eq. (6.12) to find $\langle m| R_j |n\rangle$ by numerical integration. Similarly, from the definitions of first and second-moment observables in Eq. (6.15) in terms of POVM elements G_y's, one can numerically obtain a representation of these operators in the photon-number basis.

Appendix D

Proofs and technical details related to EAT

D.1 Deriving dual problem of Eq. (4.85)

In this section, let $\mathcal{H}$ be a Hilbert space and let $\Phi_{\boldsymbol{\mathcal{M}}}$ be a measurement map given in Eq. (3.17) with associated alphabet $\mathcal{X}$ for outcomes. Let W be the quantum conditional entropy function in Eq. (4.19). Let $m = |\mathcal{X}|$ be the number of outcomes. We define the set

$$\mathcal{F}_{\text{dual}} := \{\boldsymbol{f} \in \mathbb{R}^m : \boldsymbol{f} \cdot \Phi_{\boldsymbol{\mathcal{M}}}(\rho) \leq W(\rho)\ \forall \rho \in \mathrm{D}(\mathcal{H})\}. \tag{D.1}$$

It is also easy to see that $\mathcal{F}_{\text{dual}}$ is a convex set. Moreover, we can write the indicator function (see Eq. (2.96)) for the set $\mathcal{F}_{\text{dual}}$ in a useful way by the following lemma.

Lemma D.1.1. Let $\boldsymbol{f} \in \mathbb{R}^m$. Let $\delta_{\mathcal{F}_{\text{dual}}}(\boldsymbol{f})$ be the indicator function for the set $\mathcal{F}_{\text{dual}}$. Then,

$$\sup_{\rho}\{\langle \boldsymbol{f}, \Phi_{\boldsymbol{\mathcal{M}}}(\rho)\rangle - W(\rho) : \rho \in \mathrm{Pos}(\mathcal{H})\} = \delta_{\mathcal{F}_{\text{dual}}}(\boldsymbol{f}). \tag{D.2}$$

Proof. Let $\alpha := \sup_{\rho}\{\langle \boldsymbol{f}, \Phi_{\boldsymbol{\mathcal{M}}}(\rho)\rangle - W(\rho) : \rho \in \mathrm{Pos}(\mathcal{H})\}$.

If $\boldsymbol{f} \in \mathcal{F}_{\text{dual}}$, then $\langle \boldsymbol{f}, \Phi_{\boldsymbol{\mathcal{M}}}(\rho)\rangle - W(\rho) \leq 0$ for any $\rho \in \mathrm{D}(\mathcal{H})$. As any positive operator can be scaled from $\rho \in \mathrm{D}(\mathcal{H})$ by a non-negative coefficient, it is the case that $\alpha \leq 0$. Thus, $\alpha = 0$ with $\rho = 0$.

If $\boldsymbol{f} \notin \mathcal{F}_{\text{dual}}$, then there exists $\rho_0 \in \mathrm{D}(\mathcal{H})$ such that $\langle \boldsymbol{f}, \Phi_{\boldsymbol{\mathcal{M}}}(\rho_0)\rangle - W(\rho_0) > 0$. Let $\gamma = \langle \boldsymbol{f}, \Phi_{\boldsymbol{\mathcal{M}}}(\rho_0)\rangle - W(\rho_0)$ and $\rho_\beta = \beta\rho_0$ with $\beta \geq 0$. We notice that $W(\beta\rho) = \beta W(\rho)$ for any $\beta \geq 0$. Because $\Phi_{\boldsymbol{\mathcal{M}}}$ is linear, with ρ_β, it is the case that $\alpha > \beta\gamma$. With $\beta \to \infty$, $\alpha \to \infty$. Thus, Eq. (D.2) holds. □

To derive the dual problem of Eq. (4.85), we first compute the Fenchel conjugate of two functions that will become useful.

Lemma D.1.2. For $c_0 \geq 0$ and $c_1 > 0$, let $s : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ be defined as

$$s(x) = \begin{cases} -\sqrt{c_0^2 - |x|^2/c_1^2} & \text{if } |x| \leq c_0 c_1, \\ \infty & \text{otherwise.} \end{cases} \tag{D.3}$$

The Fenchel conjugate function of s is

$$s^*(y) = c_0\sqrt{1 + c_1^2 y^2}. \tag{D.4}$$

Proof. By Definition 2.4.11, the conjugate function of s is

$$\begin{aligned} s^*(y) &= \sup_{x\in\mathbb{R}}\{xy - s(x)\} \\ &= \sup_{x\in\mathbb{R}}\{xy + \sqrt{c_0^2 - |x|^2/c_1^2} : |x| \leq c_0 c_1\} \end{aligned} \tag{D.5}$$

By a simple calculation to optimize over x, it is easy to verify that Eq. (D.4) is indeed the conjugate function of s.[13] □

Lemma D.1.3. For $c_0 \geq 0$ and $c_1 > 0$, let $\mathcal{E}(\boldsymbol{\lambda}) : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ be defined as

$$\mathcal{E}(\boldsymbol{\lambda}) = \begin{cases} s(\|\boldsymbol{\lambda}\|_1/2) & \text{if } \sum_x \boldsymbol{\lambda}(x) = 0 \\ \infty & \text{otherwise,} \end{cases} \tag{D.6}$$

where s is defined in Lemma D.1.2. Then $\mathcal{E}$ is a convex function and its conjugate function $\mathcal{E}^*(\boldsymbol{f})$ is $\mathcal{E}^*(\boldsymbol{f}) = s^*(\max(\boldsymbol{f}) - \min(\boldsymbol{f}))$, that is,

$$\mathcal{E}^*(\boldsymbol{f}) = c_0\sqrt{1 + c_1^2[\max(\boldsymbol{f}) - \min(\boldsymbol{f})]^2}\ . \tag{D.7}$$

Proof. The convexity of $\mathcal{E}$ follows from the convexity of s and the convexity of the set $\{\boldsymbol{\lambda} \in \mathbb{R}^m : \sum_x \boldsymbol{\lambda}(x) = 0\}$. From Definition 2.4.11,

$$\begin{aligned} \mathcal{E}^*(\boldsymbol{f}) &= \sup_{\boldsymbol{\lambda}\in\mathbb{R}^m}\{\langle\boldsymbol{\lambda}, \boldsymbol{f}\rangle - \mathcal{E}(\boldsymbol{\lambda})\} \\ &= \sup_{\boldsymbol{\lambda}\in\mathbb{R}^m}\{\langle\boldsymbol{\lambda}, \boldsymbol{f}\rangle - s(\|\boldsymbol{\lambda}\|_1/2) : \sum_x \boldsymbol{\lambda}(x) = 0\} \\ &= \sup_{\boldsymbol{\lambda}\in\mathbb{R}^m}\{\langle\boldsymbol{\lambda}, \boldsymbol{f}\rangle + \sqrt{c_0^2 - \|\boldsymbol{\lambda}\|_1^2/(4c_1^2)} : \|\boldsymbol{\lambda}\|_1 \leq 2c_0c_1, \sum_x \boldsymbol{\lambda}(x) = 0\} \end{aligned} \tag{D.8}$$

[13] The supremum can be found by various methods. For example, with the help of Mathematica, one can easily find the optimal value as a function of y. Alternatively, one can solve case by case for two cases: $x \geq 0$ and $x < 0$. For each case, one looks for a solution of x where the derivative is zero by simple computation. Then, one verifies that both solutions in these two cases lead to the same optimal value which is indeed the supremum.

Let $\boldsymbol{\lambda}_+$ be defined as $\boldsymbol{\lambda}_+(x) = \max(\boldsymbol{\lambda}(x), 0)$ and $\boldsymbol{\lambda}_-$ be defined as $\boldsymbol{\lambda}_-(x) = \max(-\boldsymbol{\lambda}(x), 0)$. Then, it is clear that $\boldsymbol{\lambda} = \boldsymbol{\lambda}_+ - \boldsymbol{\lambda}_\mathbf{1}$ and $\|\boldsymbol{\lambda}\|_1 = \|\boldsymbol{\lambda}_+\|_1 + \|\boldsymbol{\lambda}_-\|_1$. The condition $\sum_x \boldsymbol{\lambda}(x) = 0$ implies $\|\boldsymbol{\lambda}_+\|_1 = \|\boldsymbol{\lambda}_-\|_1 = \frac{1}{2}\|\boldsymbol{\lambda}\|_1$. Thus,

$$\begin{aligned}
\mathcal{E}^*(\boldsymbol{f}) &= \sup_{\substack{\boldsymbol{\lambda}_+ \in \mathbb{R}^m \\ \boldsymbol{\lambda}_- \in \mathbb{R}^m}} \{\langle \boldsymbol{f}, \boldsymbol{\lambda}_+\rangle - \langle \boldsymbol{f}, \boldsymbol{\lambda}_-\rangle + \sqrt{c_0^2 - \|\boldsymbol{\lambda}_+\|_1^2/c_1^2} : \|\boldsymbol{\lambda}_+\|_1 = \|\boldsymbol{\lambda}_-\|_1 \leq c_0c_1, \boldsymbol{\lambda}_+ \geq 0, \boldsymbol{\lambda}_- \geq 0\} \\
&= \sup_{\boldsymbol{\lambda}_+ \in \mathbb{R}^m, v \in \mathbb{R}} \{\langle \boldsymbol{f}, \boldsymbol{\lambda}_+\rangle - v\min(\boldsymbol{f}) + \sqrt{c_0^2 - v^2/c_1^2} : 0 \leq v = \|\boldsymbol{\lambda}_+\|_1 \leq c_0c_1, \boldsymbol{\lambda}_+ \geq 0\} \\
&= \sup_{v \in \mathbb{R}} \{v[\max(\boldsymbol{f}) - \min(\boldsymbol{f})] + \sqrt{c_0^2 - v^2/c_1^2} : 0 \leq v \leq c_0c_1\} \\
&= \sup_{v \in \mathbb{R}} \{v[\max(\boldsymbol{f}) - \min(\boldsymbol{f})] + \sqrt{c_0^2 - v^2/c_1^2} : |v| \leq c_0c_1\} \\
&= \sup_{v \in \mathbb{R}} \{v[\max(\boldsymbol{f}) - \min(\boldsymbol{f})] - s(v)\} \\
&= s^*(\max(\boldsymbol{f}) - \min(\boldsymbol{f})).
\end{aligned} \tag{D.9}$$

In the second line above, we optimize $\boldsymbol{\lambda}_-$ and the optimal value is achieved when $\boldsymbol{\lambda}_-$ contains a single nonzero entry in the position corresponding to $\min(\boldsymbol{f})$. In the third line, we optimize $\boldsymbol{\lambda}_+$ and the optimal value is achieved when $\boldsymbol{\lambda}_+$ contains a single nonzero entry in the position corresponding to $\max(\boldsymbol{f})$. In the fourth line, we drop the constraint $v \geq 0$ since the optimal value is always achieved when $v \geq 0$ due to the fact $\max(\boldsymbol{f}) - \min(\boldsymbol{f}) \geq 0$. In the fifth line, we use the definition of s. The last line follows from the definition of s^*. □

Proposition D.1.4. Let $c_0 \geq 0$ and $c_1 > 0$ be constants. For the problem

$$\begin{aligned}
h_{\text{primal}}(\boldsymbol{q_0}) := \underset{\rho}{\text{minimize}} \qquad & W(\rho) + \sqrt{c_0^2 - \|\boldsymbol{q}_0 - \Phi_{\boldsymbol{\mathcal{M}}}(\rho)\|_1^2/(4c_1^2)} \\
\text{subject to} \qquad & \operatorname{Tr}(\rho) = 1 \\
& \|\boldsymbol{q}_0 - \Phi_{\boldsymbol{\mathcal{M}}}(\rho)\|_1 \leq 2c_0c_1 \\
& \rho \geq 0,
\end{aligned} \tag{D.10}$$

its Fenchel dual problem is

$$\begin{aligned}
h_{\text{dual}}(\boldsymbol{q_0}) := \underset{\boldsymbol{f}}{\text{maximize}} \qquad & \boldsymbol{q_0} \cdot \boldsymbol{f} - c_0\sqrt{1 + c_1^2[\max(\boldsymbol{f}) - \min(\boldsymbol{f})]^2} \\
\text{subject to} \qquad & \boldsymbol{f} \cdot \Phi_{\boldsymbol{\mathcal{M}}}(\rho) \leq W(\rho) \ \forall \rho \in \mathrm{D}(\mathcal{H}).
\end{aligned} \tag{D.11}$$

Moreover, if there exists $\sigma \in \mathrm{D}(\mathcal{H})$ such that $\Phi_{\boldsymbol{\mathcal{M}}}(\sigma) = \boldsymbol{q_0}$, then $h_{\text{primal}}(\boldsymbol{q_0}) = h_{\text{dual}}(\boldsymbol{q_0})$.

Proof. We introduce the indicator function $\delta_{\mathrm{Pos}(\mathcal{H})}$ for the set $\mathrm{Pos}(\mathcal{H})$ (see Eq. (2.96)). We identify any $\rho \in \mathrm{Pos}(\mathcal{H})$ as an element in $\mathbb{R}^n$ for $n = 4\dim(\mathcal{H})^2$ by representing ρ according to Eq. (2.90) and then applying an appropriate vectorization of ρ. Similarly, we abuse the notation $\Phi_{\mathcal{M}}$ and let it denote the corresponding linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$ where m is the number of POVM elements for this measurement channel. To apply Theorem 2.4.12, we define $f(\rho) = W(\rho) + \delta_{\mathrm{Pos}(\mathcal{H})}(\rho)$, $g(\boldsymbol{\lambda}) = \mathcal{E}(\boldsymbol{\lambda} - \boldsymbol{q}_0)$ and let $\Phi_{\mathcal{M}}$ play the role of A in the theorem. We first calculate f^* and g^*:

$$\begin{aligned}
f^*(Z) &= \sup_{\rho}\{\langle Z, \rho\rangle - f(\rho)\} \\
&= \sup_{\rho}\{\langle Z, \rho\rangle - W(\rho) : \rho \in \mathrm{Pos}(\mathcal{H})\} && \text{(D.12)} \\
g^*(\boldsymbol{f}) &= \sup_{\boldsymbol{\lambda}}\{\langle \boldsymbol{f}, \boldsymbol{\lambda}\rangle - \mathcal{E}(\boldsymbol{\lambda} - \boldsymbol{q_0})\} \\
&= \sup_{\boldsymbol{\lambda'}}\{\langle \boldsymbol{f}, \boldsymbol{\lambda'}\rangle + \langle \boldsymbol{f}, \boldsymbol{q_0}\rangle - \mathcal{E}(\boldsymbol{\lambda'})\} \\
&= \mathcal{E}^*(\boldsymbol{f}) + \langle \boldsymbol{f}, \boldsymbol{q_0}\rangle. && \text{(D.13)}
\end{aligned}$$

Because $\sum_x \boldsymbol{q_0}(x) = 1$ and $\sum_x \Phi_{\mathcal{M}}(\rho)(x) = 1$ for $\mathrm{Tr}(\rho) = 1$, one can use the definition of f and $\mathcal{E}$ to rewrite the problem in Eq. (D.10) as

$$\begin{aligned}
h_{\text{primal}}(\boldsymbol{q_0}) = &\inf_{\rho \in \mathbb{R}^n}\{f(\rho) + \mathcal{E}(\Phi_{\mathcal{M}}(\rho) - \boldsymbol{q_0})\} \\
= &\inf_{\rho \in \mathbb{R}^n}\{f(\rho) + g(\Phi_{\mathcal{M}}(\rho))\}.
\end{aligned} \tag{D.14}$$

We then use the indicator function for the set $\mathcal{F}_{\text{dual}}$ and the conjugate function $\mathcal{E}^*$ of $\mathcal{E}$ to rewrite Eq. (D.11) as

$$\begin{aligned}
h_{\text{dual}}(\boldsymbol{q_0}) &= \sup_{\boldsymbol{f}}\{-\langle \boldsymbol{q_0}, -\boldsymbol{f}\rangle - \mathcal{E}^*(\boldsymbol{f}) - \delta_{\mathcal{F}_{\text{dual}}}(\boldsymbol{f})\} \\
&= \sup_{\boldsymbol{f}}\{-g^*(-\boldsymbol{f}) - \delta_{\mathcal{F}_{\text{dual}}}(\boldsymbol{f})\}.
\end{aligned} \tag{D.15}$$

By Lemma D.1.1, $\delta_{\mathcal{F}_{\text{dual}}}(\boldsymbol{f}) = \sup_{\rho}\{\langle \boldsymbol{f}, \Phi_{\boldsymbol{\mathcal{M}}}(\rho)\rangle - W(\rho) : \rho \in \mathrm{Pos}(\mathcal{H})\}$. Using this substitution, we find

$$\begin{aligned}
h_{\text{dual}}(\boldsymbol{q_0}) = &\sup_{\boldsymbol{f} \in \mathbb{R}^m}\{-\sup_{\rho}\{\langle \boldsymbol{f}, \Phi_{\boldsymbol{\mathcal{M}}}(\rho)\rangle - W(\rho) : \rho \in \mathrm{Pos}(\mathcal{H})\} - g^*(-\boldsymbol{f})\} \\
= &\sup_{\boldsymbol{f} \in \mathbb{R}^m}\{-\sup_{\rho}\{\langle \Phi^{\dagger}_{\boldsymbol{\mathcal{M}}}(\boldsymbol{f}), \rho\rangle - W(\rho) : \rho \in \mathrm{Pos}(\mathcal{H})\} - g^*(-\boldsymbol{f})\} \\
= &\sup_{\boldsymbol{f} \in \mathbb{R}^m}\{-f^*(\Phi^{\dagger}_{\boldsymbol{\mathcal{M}}}(\boldsymbol{f})) - g^*(-\boldsymbol{f})\}.
\end{aligned} \tag{D.16}$$

Therefore, the optimization problem in Eq. (D.11) is the dual problem of the optimization problem in Eq. (D.10).

Since both f and g are convex, to show $h_{\text{primal}}(\boldsymbol{q_0}) = h_{\text{dual}}(\boldsymbol{q_0})$, we just need to verify the condition

$$\operatorname{relint}(\operatorname{dom}(g)) \cap \Phi_{\boldsymbol{\mathcal{M}}} \operatorname{relint}(\operatorname{dom}(f)) \neq \emptyset. \tag{D.17}$$

From the definition of domain [see Eq. (2.89)],

$$\begin{aligned} \operatorname{dom}(g) &= \{\boldsymbol{\lambda} \in \mathbb{R}^m : \sum_x \boldsymbol{\lambda}(x) = 1, \|\boldsymbol{\lambda} - \boldsymbol{q_0}\|_1 \leq 2c_0c_1\}\ , \\ \operatorname{dom}(f) &= \{\rho \in \mathbb{R}^n : \rho \geq 0\}. \end{aligned} \tag{D.18}$$

It is clear that $\boldsymbol{q_0} \in \operatorname{relint}(\operatorname{dom}(g))$. Moreover, there exists some $\epsilon > 0$ such that $(1-\epsilon)\boldsymbol{q_0}+\epsilon\mathbf{1}/m \in \operatorname{relint}(\operatorname{dom}(g))$. We now show that $(1-\epsilon)\boldsymbol{q_0}+\epsilon\mathbf{1}/m \in \Phi_{\boldsymbol{\mathcal{M}}} \operatorname{relint}(\operatorname{dom}(f))$. By the assumption, there exists $\sigma \in \mathrm{D}(\mathcal{H})$ such that $\Phi_{\boldsymbol{\mathcal{M}}}(\sigma) = \boldsymbol{q_0}$. Then for any $\delta > 0$, it is the case that $\mathcal{D}_\delta(\sigma) \in \operatorname{relint}(\operatorname{dom}(f))$. In particular, $\mathcal{D}_\epsilon(\sigma) \in \operatorname{relint}(\operatorname{dom}(f))$. Since $\Phi_{\boldsymbol{\mathcal{M}}}(\mathcal{D}_\epsilon(\sigma)) = (1-\epsilon)\boldsymbol{q_0}+\epsilon\mathbf{1}/m$, it is the case that $(1-\epsilon)\boldsymbol{q_0}+\epsilon\mathbf{1}/m \in \Phi_{\boldsymbol{\mathcal{M}}} \operatorname{relint}(\operatorname{dom}(f))$. Therefore, Eq. (D.17) holds. We conclude that $h_{\text{primal}}(\boldsymbol{q_0}) = h_{\text{dual}}(\boldsymbol{q_0})$. □

Since we establish that $h_{\text{primal}}(\boldsymbol{q_0}) = h_{\text{dual}}(\boldsymbol{q_0})$, we can now solve the primal problem in Eq. (D.10). The difficulty is that the objective function is not differentiable at points where $\Phi_{\boldsymbol{\mathcal{M}}}(\rho) = \boldsymbol{q_0}$. Since many standard solvers handle differentiable functions more effectively, we introduce a slack variable and rewrite the problem to an equivalent one according to the following lemma.

Lemma D.1.5. The following optimization problem

$$\begin{aligned} \tilde{h}_{\text{primal}}(\boldsymbol{q_0}) := \underset{\rho,\boldsymbol{\xi}}{\text{minimize}} \qquad & W(\rho) - \sqrt{c_0^2 - \Big(\sum_j \boldsymbol{\xi}(j)\Big)^2/(4c_1^2)} \\ \text{subject to} \qquad & -\boldsymbol{\xi} \leq \Phi_{\boldsymbol{\mathcal{M}}}(\rho) - \boldsymbol{q_0} \leq \boldsymbol{\xi} \\ & \sum_j \boldsymbol{\xi}(j) \leq 2c_0c_1 \\ & \operatorname{Tr}(\rho) = 1 \\ & \rho \geq 0 \end{aligned} \tag{D.19}$$

has the same optimal value as the problem in Eq. (D.10), that is, $\tilde{h}_{\text{primal}}(\boldsymbol{q_0}) = h_{\text{primal}}(\boldsymbol{q_0})$.

Proof. We first notice that the function $-\sqrt{c_0^2 - t^2/(4c_1^2)}$ is a monotonically non-decreasing function of $0 \leq t \leq 2c_0c_1$. For $\boldsymbol{\xi} \geq 0, \|\boldsymbol{\xi}\|_1 = \sum_j \boldsymbol{\xi}(j)$. The condition $-\boldsymbol{\xi} \leq \Phi_{\boldsymbol{\mathcal{M}}}(\rho) - \boldsymbol{q_0} \leq \boldsymbol{\xi}$ implies that $\boldsymbol{\xi} \geq 0$ and $\|\boldsymbol{\xi}\|_1 \geq \|\boldsymbol{q_0} - \Phi_{\boldsymbol{\mathcal{M}}}(\rho)\|_1$. Thus, it is clear that $\tilde{h}_{\text{primal}}(\boldsymbol{q_0}) \geq h_{\text{primal}}(\boldsymbol{q_0})$. Let $\rho^\star$ be an optimal solution for $h_{\text{primal}}(\boldsymbol{q_0})$ and let $\boldsymbol{\xi}^\star(j) := |\boldsymbol{q_0}(j) - \Phi_{\boldsymbol{\mathcal{M}}}(\rho^\star)(j)|$. Then $\|\boldsymbol{\xi}^\star\|_1 = \|\boldsymbol{q_0} - \Phi_{\boldsymbol{\mathcal{M}}}(\rho^\star)\|_1 \leq 2c_0c_1$. This implies, $(\rho^\star, \boldsymbol{\xi}^\star)$ is a feasible solution for $\tilde{h}_{\text{primal}}(\boldsymbol{q_0})$ that achieves the same objective function value as $h_{\text{primal}}(\boldsymbol{q_0})$. Therefore, $\tilde{h}_{\text{primal}}(\boldsymbol{q_0}) \leq h_{\text{primal}}(\boldsymbol{q_0})$. We conclude that $\tilde{h}_{\text{primal}}(\boldsymbol{q_0}) = h_{\text{primal}}(\boldsymbol{q_0})$. □

D.2 Proof of Theorem 8.2.3

The proof of this theorem relies on the following lemma.

Lemma D.2.1. Consider a CCQ state $\rho_{XPE} = \sum_{(x,p)} |x\rangle\langle x| \otimes |p\rangle\langle p| \otimes \rho_E^{(x,p)}$. Consider another classical register Y which be constructed using a deterministic function on the registers X and P, i.e. $f : X \times P \to Y$. That is to say, there exists an isometry, $V \equiv \sum_{(x,p)} |x\rangle\langle x| \otimes |p\rangle\langle p| \otimes |f(x,p)\rangle_Y$ reconstructing Y from X and P. Then

$$H_{\min}^{\varepsilon}(YX|PE)_{V\rho V^\dagger} = H_{\min}^{\varepsilon}(X|PE)_\rho \ . \tag{D.20}$$

Proof. To prove this lemma, we simply prove

$$H_{\min}^{\varepsilon}(YX|PE)_{V\rho V^\dagger} \geq H_{\min}^{\varepsilon}(X|PE)_\rho \tag{D.21}$$

$$H_{\min}^{\varepsilon}(YX|PE)_{V\rho V^\dagger} \leq H_{\min}^{\varepsilon}(X|PE)_\rho \ . \tag{D.22}$$

We start with Eq. (D.21) as it is straightforward using previous results. Namely, we see:

$$H_{\min}^{\varepsilon}(YX|PE)_{V\rho V^\dagger} \geq H_{\min}^{\varepsilon}(X|PE)_{V\rho V^\dagger} = H_{\min}^{\varepsilon}(X|PE)_\rho \ . \tag{D.23}$$

The first inequality comes from the fact that the entropy of a classical register is non-negative [95, Lemma 6.7]. The equality follows from noting that $\mathrm{Tr}_Y(V\rho V^\dagger) = \rho$. Therefore we can focus on the other direction of the inequality.

By combining [95, Definitions 6.2 and 6.5] we note the definition of the smooth min-entropy terms we are concerned with:

$$H_{\min}^{\varepsilon}(YX|PE)_{V\rho V^\dagger} := \max_{\overline{\rho}\in\mathcal{B}_\varepsilon(V\rho V^\dagger)} \max\left\{\overline{\lambda}\in\mathbb{R} : \mathbb{1}_Y \otimes \mathbb{1}_X \otimes 2^{-\overline{\lambda}}\overline{M}_{PE} \succeq \overline{\rho}_{YXPE}\right\} \tag{D.24}$$

$$H_{\min}^{\varepsilon}(X|PE)_\rho := \max_{\widetilde{\rho}\in\mathcal{B}_\varepsilon(\rho)} \max\left\{\widetilde{\lambda}\in\mathbb{R} : \mathbb{1}_X \otimes 2^{-\widetilde{\lambda}}\widetilde{M}_{PE} \succeq \widetilde{\rho}_{XPE}\right\} \ . \tag{D.25}$$

One can see from these definitions it would be sufficient to show the optimal solution for $H_{\min}^{\varepsilon}(YX|PE)_{V\rho V^\dagger}$ is a feasible solution for $H_{\min}^{\varepsilon}(X|PE)_\rho$ to prove Eq. (D.22). Suppose the optimal solution of $H_{\min}^{\varepsilon}(YX|PE)_{V\rho V^\dagger}$ is taken at $(\overline{\lambda}, \overline{M}_{PE}, \overline{\rho}_{YXPE})$ where $\overline{\rho}_{YXPE}$ is chosen as a CCCQ state by [95, Lemma 6.6]. This gives

$$\mathbb{1}_Y \otimes \mathbb{1}_X \otimes 2^{-\overline{\lambda}}\overline{M}_{PE} \succeq \overline{\rho}_{YXPE} := \sum_{(y,x,p)} |y\rangle\langle y| \otimes |x\rangle\langle x| \otimes |p\rangle\langle p| \otimes \overline{\rho}_E^{(y,x,p)} \ . \tag{D.26}$$

Applying the pinching map $\Delta_P(\cdot) := \sum_p |p\rangle\,\langle p| \cdot |p\rangle\,\langle p|$ on both sides, we get

$$\sum_{(y,x,p)} |y\rangle\langle y| \otimes |x\rangle\langle x| \otimes |p\rangle\langle p| \otimes 2^{-\overline{\lambda}}\overline{M}_E^{(p)} \succeq \sum_{(y,x,p)} |y\rangle\langle y| \otimes |x\rangle\langle x| \otimes |p\rangle\langle p| \otimes \overline{\rho}_E^{(y,x,p)}, \tag{D.27}$$

with $\overline{M}_E^{(p)} := \langle p|\overline{M}_{PE}|p\rangle$. We can further *un-compute* Y by applying $V^\dagger \cdot V$ on both sides,

$$\sum_{(x,p)} |x\rangle\langle x| \otimes |p\rangle\langle p| \otimes 2^{-\overline{\lambda}}\overline{M}_E^{(p)} \succeq \sum_{(x,p)} |x\rangle\langle x| \otimes |p\rangle\langle p| \otimes \overline{\rho}_E^{(f(x,p),x,p)}, \tag{D.28}$$

which is equivalent to

$$\mathbb{1}_X \otimes 2^{-\overline{\lambda}} \sum_p |p\rangle\langle p| \otimes \overline{M}_E^{(p)} \succeq V^\dagger \overline{\rho}_{YXPE} V \tag{D.29}$$

Since $V^\dagger \cdot V$ is a CPTNI map, we know that

$$P(V^\dagger \overline{\rho}_{YXPE} V, \rho_{XPE}) \leq P(\overline{\rho}_{YXPE}, V\rho_{XPE}V^\dagger) \leq \varepsilon, \tag{D.30}$$

where the second equality follows by the assumption of $\overline{\rho}_{YXPE}$. Thus, it is the case that $(\overline{\lambda}, \sum_p |p\rangle\langle p| \otimes \overline{M}_E^{(p)}, V^\dagger \overline{\rho}_{YXPE} V)$ is a feasible solution for $H_{\min}^{\varepsilon}(X|PE)_\rho$, which implies (D.22). □

We now prove the theorem.

Proof of Theorem 8.2.3. Let Ω' be the event that Protocol 8.1 does not abort (i.e. Ω is satisfied) and error correction succeeds (i.e. $Z_1^N = \widehat{Z}_1^N$)[14]. Let Ω'' be the event that the QKD protocol does not abort and error correction succeeds. Let $\tilde{\rho}_{|\Omega''}$ be the output of the QKD protocol conditioned on not aborting and error correction succeeding and $\rho_{|\Omega'}$ be the output of Protocol 8.1 conditioned on not aborting *and* error correction being applied and succeeding.

Fundamentally, we are interested in the entropy of the raw key which excludes the sifted registers or the registers used for parameter estimation. We can denote this set of registers $\mathbf{Z} \equiv \overline{A}_1^N \setminus (\overline{A}_{\mathcal{S}} \cup \overline{A}_{\mathcal{T}})$. However, we note that Alice announces $\overline{A}_{\mathcal{T}}$, making it part of what is conditioned on, and $\overline{A}_{\mathcal{S}}$ is determined entirely by the registers $\widetilde{A}_1^N$, $\widetilde{B}_1^N$, and T_1^N. Therefore, by Lemma D.2.1, we can conclude:

$$H_{\min}^{\bar{\varepsilon}}(\mathbf{Z}|\widetilde{A}_1^N \widetilde{B}_1^N T_1^N \overline{A}_{\mathcal{T}} CE)_{\tilde{\rho}_{|\Omega''}} = H_{\min}^{\bar{\varepsilon}}(\overline{A}_1^N|\widetilde{A}_1^N \widetilde{B}_1^N T_1^N \overline{A}_{\mathcal{T}} CE)_{\tilde{\rho}_{|\Omega''}}. \tag{D.31}$$

We can handle the classical communication cost of error correction by [95, Lemma 6.8]:

$$\begin{aligned} H_{\min}^{\bar{\varepsilon}}(\overline{A}_1^N|\widetilde{A}_1^N \widetilde{B}_1^N T_1^N \overline{A}_{\mathcal{T}} CE)_{\tilde{\rho}_{|\Omega''}} &\geq H_{\min}^{\bar{\varepsilon}}(\overline{A}_1^N|\widetilde{A}_1^N \widetilde{B}_1^N T_1^N \overline{A}_{\mathcal{T}} E)_{\tilde{\rho}_{|\Omega''}} - \mathrm{leak}_{\mathrm{EC}} &\text{(D.32)}\\ &= H_{\min}^{\bar{\varepsilon}}(\overline{A}_1^N|\widetilde{A}^n \widetilde{B}^N T_1^N \overline{A}_{\mathcal{T}} E)_{\rho_{|\Omega'}} - \mathrm{leak}_{\mathrm{EC}} &\text{(D.33)}\end{aligned}$$

[14] The requirement is formally $\overline{A}_1^N = \widehat{A}_1^N$. However, Bob knows Alice's result whenever she does not apply the key map.

where C is the register storing the classical communication due to error correction, and the equality comes from noting that $\rho_{\overline{A}_1^N \widetilde{A}_1^N \widetilde{B}_1^N T_1^N X_1^N E_{|\Omega'}} = \tilde{\rho}_{\overline{A}_1^N \widetilde{A}_1^N \widetilde{B}_1^N T_1^N X_1^N E_{|\Omega''}}$. We stress that Bob's private register, $\overline{B}_1^N$, for both the EAT and the QKD protocol, is excluded in the equality as they differ[15]. We note that $\overline{A}_{\mathcal{T}}$ is conditioned upon as Alice publicly announced it, but X_1^N is not conditioned upon as Bob computes it locally and only announces if they do not abort.

With all of this addressed, we can conclude that we would like to do privacy amplification using the smooth min-entropy term:

$$H^{\bar{\varepsilon}}_{\min}(\overline{A}_1^N | \widetilde{A}_1^N \widetilde{B}_1^N T_1^N \overline{A}_{\mathcal{T}} E)_{\rho_{|\Omega'}} \tag{D.34}$$

This imposes one major difficulty: we need to find a way to relate this to the term $H^{\bar{\varepsilon}}_{\min}(\overline{A}_1^N \overline{B}_1^N | \widetilde{A}^N \widetilde{B}^N T_1^N E)_{\rho_{|\Omega'}}$ on the left-hand side of Eq. (8.1).

Following a similar derivation in the DIQKD setting in [43], and also handling the $\overline{A}_{\mathcal{T}}$ register, we use the following series of claims:

$$\begin{aligned} H^{\bar{\varepsilon}}_{\min}(\overline{A}_1^N | \widetilde{A}_1^N \widetilde{B}_1^N T_1^N \overline{A}_{\mathcal{T}} E)_{\rho_{|\Omega'}} &\geq H^{\bar{\varepsilon}}_{\min}(\overline{A}_1^N | \widetilde{A}_1^N \widetilde{B}_1^N T_1^N \overline{A}_{\mathcal{T}} \overline{B}_1^N E)_{\rho_{|\Omega'}} && \text{(D.35)} \\ &\geq H^{2\bar{\varepsilon}}_{\min}(\overline{A}_1^N \overline{A}_{\mathcal{T}} \overline{B}_1^N | \widetilde{A}_1^N \widetilde{B}_1^N T_1^N E)_{\rho_{|\Omega'}} - H^{\bar{\varepsilon}/4}_{\max}(\overline{A}_{\mathcal{T}} \overline{B}_1^N | \widetilde{A}_1^N \widetilde{B}_1^N T_1^N E)_{\rho_{|\Omega'}} \\ &\qquad - 2\log_2(1 - \sqrt{1 - \bar{\varepsilon}^2/16}) && \text{(D.36)} \\ &\geq H^{2\bar{\varepsilon}}_{\min}(\overline{A}_1^N \overline{B}_1^N | \widetilde{A}_1^N \widetilde{B}_1^N T_1^N E)_{\rho_{|\Omega'}} - H^{\bar{\varepsilon}/4}_{\max}(\overline{A}_{\mathcal{T}} \overline{B}_1^N | T_1^N E)_{\rho_{|\Omega'}} \\ &\qquad - 2\log_2(1 - \sqrt{1 - \bar{\varepsilon}^2/16})\ . && \text{(D.37)} \end{aligned}$$

The first inequality follows from the data-processing inequality for smooth min-entropy as the left hand side may be obtained by tracing off the $\overline{B}_1^N$ registers. The second inequality follows from the chain rules in [95, Eq. (6.59)]. The final equality follows from Lemma D.2.1 for the min-entropy term, since $\overline{A}_{\mathcal{T}}$ may be reconstructed from $\overline{A}_1^N$ and T_1^N, and the data-processing inequality for the max-entropy term.

Now what is left is to bound the max-entropy terms using the original version of EAT (Theorem 3.7.3) [97]. To do this we use the EAT in the form of Theorem 3.7.3 with the replacements $S_i \to \overline{A}_{\mathcal{T},i}\overline{B}_i$, $P_i \to T_i$, and $E \to E$. We note it is allowed to use $\overline{A}_{\mathcal{T}}$ as we may treat it as if it were actually N registers where it is $\perp$ whenever $T_i \neq 0$. The Markov conditions trivially hold since T_i uses seeded randomness. Then using the fact that $\overline{A}_{\mathcal{T}_i}\overline{B}_i = \perp \times \perp$ except when $T_i = 1$ which happens with probability γ, we construct the max-tradeoff function analytically by the

[15] In the EAT subprotocol (Protocol 8.1), we set each $\overline{B}_i$ to be the $\perp$ symbol due to the proof technique. This change is not executed in the QKD protocol.

following observation:

$$\begin{aligned} H(\overline{A}_{\mathcal{T}_i}\overline{B}_i|T_iR')_{\mathcal{M}_i(\omega)} &\leq H(\overline{A}_{\mathcal{T}_i}\overline{B}_i|T_i)_{\mathcal{M}_i(\omega)} \\ &= (1-\gamma)H(\overline{A}_{\mathcal{T}_i}\overline{B}_i|T_i=0)_{\mathcal{M}_i(\omega)} + \gamma H(\overline{A}_{\mathcal{T}_i}\overline{B}_i|T_i=1)_{\mathcal{M}_i(\omega)} \leq \gamma \log_2|\mathcal{A}\times\mathcal{B}|, \end{aligned} \tag{D.38}$$

where $\mathcal{A},\mathcal{B}$ are the alphabets of private outcomes that are announced during parameter estimation. This tells us we can let the max-tradeoff function $f_{\max}(\boldsymbol{q}) = \gamma\log_2|\mathcal{A}\times\mathcal{B}|$ for all $\boldsymbol{q}\in\mathbb{P}(\mathcal{X})$. This also implies $\|\nabla f_{\max}\|_\infty = 0$. Therefore, using this with Theorem 3.7.3, we get

$$H_{\max}^{\bar{\varepsilon}/4}(\overline{A}_{\mathcal{T}}\overline{B}_1^N|T_1^NE)_{\rho_{|\Omega'}} \leq N\gamma\log_2(|\mathcal{A}\times\mathcal{B}|) + 2\sqrt{N}\log_2(1+2d_S)\sqrt{1-2\log_2(\bar{\varepsilon}/4\cdot(\varepsilon_{\mathrm{EA}}+\varepsilon_{\mathrm{EC}}))} \tag{D.39}$$

where we have replaced $\rho[\Omega']$ with $\varepsilon_{\mathrm{EA}}+\varepsilon_{\mathrm{EC}}$.

We now combine Eq. (D.37) and Eq. (D.39) to get the following:

$$\begin{aligned} H_{\min}^{\bar{\varepsilon}}(\overline{A}_1^N|\widetilde{A}_1^N\widetilde{B}_1^NT_1^N\overline{A}_{\mathcal{T}}E)_{\rho_{|\Omega'}} \geq H_{\min}^{2\bar{\varepsilon}}(\overline{A}_1^N\overline{B}_1^N|\widetilde{A}_1^N\widetilde{B}_1^NT_1^NE)_{\rho_{|\Omega'}} &- N\gamma\log_2(|\mathcal{A}\times\mathcal{B}|) \\ &- 2\log_2(1-\sqrt{1-\bar{\varepsilon}^2/16}) \\ &- 2\sqrt{N}\log_2(1+2d_S)\sqrt{1-2\log_2(\bar{\varepsilon}/4\cdot(\varepsilon_{\mathrm{EA}}+\varepsilon_{\mathrm{EC}}))} \end{aligned} \tag{D.40}$$

We now apply Theorem 8.2.1 to the right-hand-side min-entropy term (with appropriate choice of $\bar{\varepsilon}/4$) so that we have that either the input state aborts with a probability greater than $1-\varepsilon_{\mathrm{EA}}-\varepsilon_{\mathrm{EC}}$ or

$$\begin{aligned} H_{\min}^{\bar{\varepsilon}}(\overline{A}_1^N|\widetilde{A}_1^N\widetilde{B}_1^NT_1^N\overline{A}_{\mathcal{T}}E)_{\rho_{|\Omega'}} \geq N\beta - \tilde{c}\sqrt{N} - \tilde{c}' &- N\gamma\log_2(|\mathcal{A}\times\mathcal{B}|) \\ &- 2\log_2(1-\sqrt{1-\bar{\varepsilon}^2/16}) \\ &- 2\sqrt{N}\log_2(1+2d_S)\sqrt{1-2\log_2(\bar{\varepsilon}/4\cdot(\varepsilon_{\mathrm{EA}}+\varepsilon_{\mathrm{EC}}))} \end{aligned} \tag{D.41}$$

Applying the leftover hashing lemma (Theorem 3.3.2) [89] completes the proof. □

Index