ULB Faculté des Sciences

Secure quantum communication with untrusted and lossy devices

Thesis presented by Michele MASINI
in fulfilment of the requirements of the PhD Degree in Physics (”Doctorat en Sciences”)
Année académique 2023-2024

Supervisor : Professor Stefano PIRONIO
Laboratoire d’Information Quantique (LIQ)

Thesis jury :

Serge MASSAR (Université libre de Bruxelles, Chair)
Nicolas CERF (Université libre de Bruxelles, Secretary)
Marco AVESANI (Università degli studi di Padova)
Ramona WOLF (Universität Siegen)

Abstract

Quantum Key Distribution (QKD) enables provably secure communication, but its security depends on accurately characterising quantum devices. Since this is a challenging task, any inaccuracies in characterising the devices can pose security risks. Device Independent (DI) QKD protocols address this issue by making minimal assumptions about the devices used, and we refer to them as untrusted devices.

This thesis begins with an introduction to key concepts in quantum information and quantum optics that are relevant to quantum communication with untrusted devices, along with examples of quantum cryptographic protocols and tasks that can be performed using these untrusted devices.

We then explore the limitations of DI QKD protocols and other protocols that use untrusted quantum devices. We introduce a new mathematical framework to study the necessary conditions for untrusted devices in the presence of photon losses and other common types of noise in quantum communication protocols. We find that protocols using untrusted devices require high detection efficiencies, which refer to the ability of the experimental setup to reliably detect quantum states. This significantly limits the distance over which DI QKD can be implemented.

We present frameworks for analysing the security of DI QKD and apply these methods to the most common protocols of this type. We introduce a general analytical framework applicable to the study of the security of DI QKD protocols with devices producing only two inputs and two outputs. Subsequently, we outline two numerical methods that can be used to analyse setups with any number of inputs and outputs.

Finally, we propose a hybrid solution that combines DI QKD with standard QKD. This approach involves partially trusting one side of the communication protocol while not trusting the other side. It is particularly suitable for secure communication between a server and its clients. Our findings show that this hybrid technology can overcome the distance limitations of DI QKD and achieve security over distances of the same order as standard QKD protocols.

Résumé

La distribution quantique de clés (QKD) permet une communication dont la sécurité peut être prouvée, mais qui dépend d'une caractérisation précise des appareils quantiques. Or cette caractérisation est en pratique complexe et toute inexactitude peut poser des risques de sécurité. Les protocoles de QKD indépendants des dispositifs (DI QKD) abordent ce problème en faisant des hypothèses minimales sur les dispositifs utilisés.

Cette thèse débute par une introduction aux concepts clés de l'information quantique et de l'optique quantique qui sont pertinents pour la communication quantique avec des dispositifs non fiables, ainsi que par des exemples de tâches et de protocoles cryptographiques quantiques pouvant être réalisés avec ces dispositifs non fiables.

Nous étudions ensuite les limitations fondamentales des protocoles de DI QKD, ainsi que d'autres protocoles utilisant des dispositifs quantiques non fiables. Nous introduisons un nouveau cadre mathématique pour étudier, en présence de pertes et de bruit, les conditions nécessaires pour que des dispositifs non fiables puissent néanmoins produire un avantage quantique dans des protocoles de communication quantique. Nous montrons que des efficacités de détection élevées sont nécessaires, ce qui limite de manière significative la distance sur laquelle la DI QKD peut être utilisée.

Nous abordons ensuite la question de l'analyse de la sécurité de la DI QKD. Nous introduisons un cadre analytique général applicable à l'étude de la sécurité des protocoles de DI QKD avec des dispositifs avec seulement deux entrées et deux sorties et appliquons cette approche aux protocoles les plus courants de ce type. Nous décrivons ensuite deux méthodes numériques qui peuvent être utilisées pour analyser des configurations avec un nombre quelconque d'entrées et de sorties.

Finalement, nous proposons une solution hybride qui combine la DI QKD avec la QKD standard. Cette approche consiste à faire partiellement confiance à un côté du protocole de communication tout en ne faisant pas confiance à l'autre côté. Elle est particulièrement adaptée à la communication sécurisée entre un serveur et ses clients. Nos résultats montrent que cette technologie hybride peut surmonter les limitations de distance de la DI QKD et atteindre une sécurité sur des distances de même ordre que les protocoles de QKD standard.

Acknowledgments

This thesis represents a collection of the works I have done and the lessons I have learned throughout my doctorate, which began at the end of 2020. This was a particularly complicated and uncertain time due to the second wave of the COVID-19 pandemic. Despite these challenges and the lockdown that did not allow me to move, my supervisor, Stefano Pironio, gave me the opportunity to start a Ph.D. while working remotely from my parents' home in Italy. Throughout this period and the entirety of my Ph.D., Stefano was always there to discuss with me and enhance my understanding of the field. I am deeply grateful to him for this. I was finally able to move to Brussels in June 2021, thanks to the spread of vaccines.

During the initial six months of doctorate conducted remotely and the subsequent year of my doctorate, I could also count on Erik Woodhead. Erik mentored me almost as a second supervisor during the first part of my doctorate and was there to answer all my questions and doubts.

After moving to Brussels I was given the opportunity to engage with the quantum information community through visits and conferences. During these experiences I met many brilliant and enlightening people. I would like to mention and thank Peter Brown, Pavel Sekatski, Roope Uola, Marco Avesani, Davide Scalcon, and Armin Tavakoli.

Over the nearly four years of my doctoral journey, the research group I was a part of has experienced many changes. It started as a small group consisting only of myself, Stefano, Erik, and Jef Pauwels. After my second year, the group saw an expansion due to the arrival of Edwin Lobo, Abhishek Mishra, Chirag Srivastava, Shubhayan Sarkar, and later, Moisés Bermejo Morán. I am grateful to them for the stimulating discussions and their support during discouraging moments.

I would also like to say a few words about my experience in Brussels. Brussels is an extremely international city, where I had the opportunity to meet people from all over Europe and beyond. It was really hard to get bored in this environment. The friends that I made here during my doctorate taught me a lot through their diverse perspectives and their sensitivity. Among many of them, I am really thankful to Harriet Ní Chinnéide, María Asiain Belloso, Giuseppe Novella, Livia Milana, and Antonio Contini. During the last few years in Brussels, I also rediscovered my passion for music and I would like to thank my teachers Matteo Costarelli and Andrea Virtuoso.

Lastly, I would like to thank my family for allowing me to pursue my interests and study abroad. Without their guidance and support I would have never reached this point.

List of publications

1. M. Masini, S. Pironio, and E. Woodhead. Simple and practical DIQKD security analysis via BB84-type uncertainty relations and Pauli correlation constraints. *Quantum* 6:843, Oct 2022.

2. M. Masini, and S. Sarkar. One-sided DI-QKD secure against coherent attacks over long distances. *arXiv preprint arXiv:2403.11850*, Mar 2024.

3. M. Masini, M. Ioannou, N. Brunner, S. Pironio, and P. Sekatski. Joint-measurability and quantum communication with untrusted devices. *arXiv preprint arXiv:2403.14785*, Mar 2024.

Contents

1 Introduction

The aim of this thesis is to present the works I carried out during my doctorate [1–3]. It also provides an introduction to the subject of Device Independent Quantum Key Distribution [4, 5].

Quantum Key Distribution (QKD) [6, 7] is a technique that was proposed in the 1980s [8], and, for the first time in history, allows provably secure communication to be achieved without relying on trusted carriers. In particular, QKD is able to solve the problem of the key distribution, which was the missing piece to achieve *information theoretic security*, i.e., security against adversaries with unlimited computing resources.

Before QKD the only known way to achieve information theoretic security was to rely on trusted carriers distributing a secret key to the locations where one needed to communicate securely. Having a secret shared key allows two parties to encrypt and decrypt messages. If (and only if) the shared key is random, equally long to the message, and was not used before, then one can prove information theoretic security [9]. This can be done, for instance, using the One Time Pad (OTP) which involves adding modulo-two each bit composing the key with each bit used to represent the message. To decrypt the message, one just has to subtract modulo two the key from the encrypted message.

Currently, cryptosystems tend to rely on a different notion of security which is called *computational security*. According to this notion, a system is said to be secure when it cannot be cracked in "reasonable" time. There are two types of algorithms which are currently used to encrypt communication: symmetric key algorithms and public key algorithms. A symmetric key algorithm uses a single key for both encryption and decryption of secret messages. An example of this class of algorithms is the previously mentioned OTP, which is impractical as it requires the distribution of extremely long keys. For this reason, the most used algorithm of this type is the Advanced Encryption Standard (AES) which is the encryption scheme established by the U.S. National Institute of Standards and Technology [10]. In this algorithm the key is not as long as the message and thus it does not guarantee information theoretic security. Instead, it guarantees computational security and it involves a typically 128 bit key from which 10 other keys are derived and are used in a process involving 10 encryption rounds.

The problem of distributing secret keys is usually solved through public key cryptography. In this scheme, a public key known to everyone is used to encrypt a secret message (in this case the secret key that we want to distribute to perform symmetric key encryption), but the decryption can only be performed using a private key which is known solely by the receiver of the message. Every known public key algorithm aims for computational security but they are never information theoretically secure. These algorithms rely on mathematical problems denoted as one-way functions. These are functions that are easy to compute for any input, but very hard to invert given the image of a random input. More specifically, finding the input with the best currently known algorithms requires a computational time which increases exponentially with the dimension of the input. An example of such function is the multiplication of prime numbers. It is easy to find the product between two large primes, but once we have the product, retrieving the prime factors is believed to be extremely hard. This one-way function is at the base of the RSA, the most-used public key encryption algorithm. This protocol was invented by Ronald Rivest, Adi Shamir and Leonard Adleman in the late 1970s [11] and it is today still used in a large number of applications.

In 1994, Peter Shor discovered a quantum algorithm able to efficiently solve prime factorization (and another one-way function called discrete logarithm) [12]. This means that, if a large-scale quantum computer will be built, most of the public key cryptosystems currently in use will be under threat. As it became clear in the last years that quantum computers are becoming a concrete possibility, other public key algorithms believed to be resistant against quantum computers have started to be studied extensively and initiated undergoing a standardization process [13]. This field of studies is called post-quantum cryptography. One of the most successful classes of post-quantum algorithms are based on lattice problems as, e.g., the shortest vector problem. This problem consists of finding the non-zero vector with the smallest length in an n-dimensional lattice given a basis of vectors in a Euclidean space.

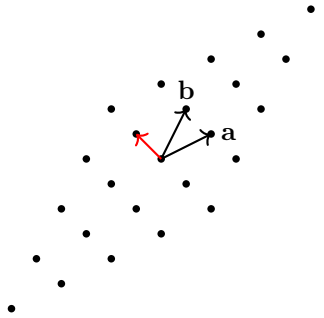

Figure 1.1: Shortest vector problem in a two-dimensional lattice. Given the basis of vectors $\mathbf{a}$ and $\mathbf{b}$, the shortest non-zero vector is the red one.

While this field of studies is extremely promising, it is not immune to sudden technological advances. Great increases in computational power and the discovery of classical or quantum algorithms able to hack these systems efficiently cannot be prevented. Moreover, there are certain types of communication which are required to stay secret for several years or decades (as the identity of undercover agents, DNA data, or health data). Computational security cannot ensure long-term security as one may record encrypted messages and wait for technological advances that will allow for their decryption.

Quantum Key Distribution can offer information theoretic security and this means that security is guaranteed even in the long-term. QKD bases its security on the postulates of Quantum Mechanics. Intuitively, we can say that it relies on the fact that the measurement of a quantum system causes an irreversible disturbance to it. For this reason, if one tries to eavesdrop the key distribution protocol, it is possible to detect the disturbance, abort the protocol, and restart from scratch. Let us remark that QKD is based on a set of assumptions besides the laws of quantum mechanics and, when they are not met, it is still possible to hack it. Anyway, the act of hacking the protocol needs to be performed while the key distribution is happening and once the secret keys are successfully distributed there is no way to retrieve them. This technology does not exist only in theory or in physics laboratories, but since the early 21st century it is available from a number of companies [14–18].

The security of QKD systems in use today relies on a set of assumptions. For example, it is assumed that the source of quantum states prepares states of a specific form, shares them through an untrusted quantum channel, and that the measuring devices perform precise measurements on the received states. In these scenarios, the state preparation and measuring devices are considered *trusted*. This trust implies that these devices are assumed to produce the states and perform the measurements exactly as prescribed by the protocol. These assumptions are used in the proofs of security of the protocols and thus they need to be fulfilled to guarantee that the

distributed keys remain secret.

However, verifying these assumptions about preparation and measurement devices is not straightforward. For instance, it is challenging to confirm whether measurements performed by Bob are restricted to a two-dimensional Hilbert space, and it is not possible to conclude only from the input-output probability distributions whether the fair sampling assumption discussed in Section 1.2.3 holds. Typically, these assumptions derive from a simplified mathematical model of the devices, which does not fully capture the complexities of the real devices in use. Ideally, one could attempt to model the devices with more sophisticated theories, but this would again require trust in the accuracy of these new models relative to the actual workings of the devices.

This discrepancy between the mathematical models used in security proofs and the real-world implementation of QKD systems presents not only theoretical concerns. Soon after the advent of commercial QKD devices, practical vulnerabilities have been exposed. Notably, V. Makarov and his research team exploited these discrepancies to experimentally demonstrate that these commercial devices could indeed be hacked [19–22].

Two main solutions have been explored to address this problem. The first involves patching the commercial devices. This method consists of integrating modules into the devices that block specific types of known attacks. Although this strategy does not resolve the underlying issue of discrepancies between the security proof assumptions and the actual implementation of the protocol—since any new patches would also need to be included in the security proofs—it does prevent the exploitation of known vulnerabilities. However, as new types of attacks are discovered, additional patches must be developed and implemented.

The second solution, which is the focus of this thesis, seeks to establish a more robust set of assumptions for our security proofs. An approach that has been studied extensively consists of making the minimal possible set of assumptions. This approach is denoted as Device Independent (DI) QKD [4, 5] and it consists of a QKD protocol in which the security is proven without trusting the internal workings of some critical parts of the devices used for the protocol, and in particular the quantum elements which are difficult to characterise and prone to errors. Specifically, it involves an untrusted source generating quantum states in an unknown form, which are then distributed to Alice and Bob's untrusted measuring devices. These devices are treated as black boxes, they receive inputs and produce outputs without needing their internal mechanisms to follow a specific mathematical model. Instead, they are only required to satisfy a few reasonable assumptions, which will be discussed in detail in Chapter 3. By analysing the joint statistical distribution of the inputs and outputs from Alice and Bob's devices, we can demonstrate information-theoretic security.

A visual representation of a Device Independent QKD protocol is given in Fig. 1.2. From now on, we will use the colour black to represent untrusted sources or devices and the colour white to represent the trusted ones. Moreover, we will also use the colour gray to draw elements that are only partially trusted.

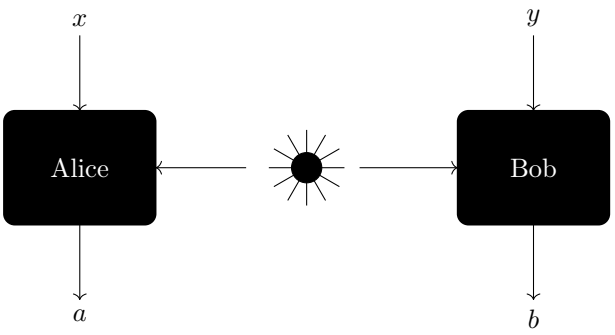

Figure 1.2: Schematic representation of a DI QKD setup. An untrusted source distributes quantum states to Alice and Bob, whose devices are represented as black boxes receiving inputs x, y and producing outputs a, b.

The first three proofs of principle experiments demonstrating the feasibility of DI QKD were performed soon after I started my doctorate, in 2022 [23–25]. However, these experiments were limited to very short distances, of the order of hundreds of meters. As we will explain throughout this thesis, achieving long distances in DI QKD is extremely challenging with the technologies available today.

Finally, note that, while QKD can offer information theoretic security which is a huge selling point, this is not the only aspect to consider when choosing a cryptosystem. QKD is an extremely expansive solution as it requires a completely new infrastructure and, when combined with the OTP, it is extremely slow. Most of the commercial devices offering QKD share a secret key which is used to perform AES instead of the OTP.

QKD will most likely be used, at least in the short and medium term, only for niche applications and its DI version remains at the moment only a research field. That said, it is fundamental to keep studying diversified cryptographic solutions. Diversification allows one to more easily replace a technology with the other existing ones in case things go bad. If we stake everything on a single technology and that technology is found to be flawed, then it will take a very long time to replace it. This is what happened recently to Germany, which staked everything on a single reliable technology to produce energy, relying on a single supplier, and is now facing an economic stall after being forced to stop buying from that supplier[1]. This is not an ideal scenario in the case of cryptography which is used for financial transactions, communication of sensitive data, computer passwords and so on. Using different cryptographic solutions for different security requirements is probably the best way to diversify the use of this technology.

In this thesis, besides presenting the achievements that I obtained together with my collaborators, I will outline some of the latest theoretical achievements in the field of DI QKD.

In this first chapter, we will cover the key concepts needed for the next parts of this thesis. In particular, we will focus on the task of communicating quantum information and we will examine some of its applications. A basic understanding of quantum information, quantum optics and probability theory is recommended to fully grasp the discussions that follow. This chapter aims to provide an introduction rather than an exhaustive review.

In Chapter 2, we will describe a general framework for characterizing the admissible levels of photon loss and noise in a wide range of scenarios and protocols with untrusted measurement devices. Photon loss represents a major challenge for the implementation of quantum communication protocols with untrusted devices. This chapter is predominantly composed of the first part of [3].

[1] See `https://www.bbc.com/news/world-europe-68361717`

Chapter 3 will present the most recent results allowing one to prove the security of DI QKD protocols. We will clarify what assumptions are needed to prove security, how security is defined in mathematical terms, and the different types of attacks that have been defined on QKD protocols. Finally, we will present an existing framework that allows one to prove security against the most general class of attacks on DI QKD. We will conclude by providing expressions that are able to describe when QKD protocols can output an identical and secret couple of keys.

In Chapter 4, we will show how the results of Chapter 2 can be improved in the context of partially DI QKD protocols and DI QKD protocols. This will lead us to compute upper bounds on the key rate of many protocols introduced in the literature, thereby identifying levels of loss and noise where protocols cannot be proven secure. This chapter is largely derived from the second part of [3],

Chapter 5 will introduce three different methods to compute lower bounds on the key rate of DI QKD protocols and thus prove its security. The main focus of the chapter will be a semi-analytic approach introduced in [1] that offers a computationally efficient method to calculate lower bounds on the key rates of a relatively large class of protocols. This section incorporates substantial text from the article in question. We will also discuss two numerical methods that can be used to analyse the security of larger classes of DI QKD protocols.

Finally, in Chapter 6, we will explore a hybrid approach between standard QKD and DI QKD where only one party's quantum device is partially trusted, while the second is untrusted. This chapter, based on [2], will allow us to prove security over distances that are similar to the ones achieved in standard QKD protocols.

1.1 Quantum information

We start by introducing the concept of quantum information and by describing how it can be quantified. A complete discussion of the topics presented here can be found in [26].

In classical information theory, information can be measured in terms of bits. A bit can be represented by a physical system that can take two possible values, usually 0 and 1. When we are "ignorant" about the value of a bit, we say that learning that value gives us one bit of information.

The quantum analogue of a bit, the quantum bit or *qubit*, is represented by a physical two-level quantum system. Our system is now described by the laws of quantum mechanics. We have that a qubit can be prepared in the first level, which we will denote as $|0\rangle$, in the second level, which we will denote as $|1\rangle$ and is a vector orthogonal to $|0\rangle$, or in a superposition of the two, which we can denote as

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle\,, \qquad \alpha, \beta \in \mathbb{C}, \tag{1.1}$$

where $|\alpha|^2 + |\beta|^2 = 1$. Also in this case, it is a natural question to ask what is the amount of quantum information that we can learn from a given qubit. Contrarily to the classical case, the information that we learn depends on the basis that we choose when we measure our system. For instance, if we know that the qubit was prepared in $|0\rangle$, and we measure the $X = |0\rangle\langle 1| + |1\rangle\langle 0|$ observable, then, before performing the measurement, we still are ignorant about the outcome. Before answering this question let us go back for a moment to the classical case.

Classically, we say that a system contains one bit of information when the system can be 0 or 1 with the same probability. If we know that our system always takes the value 0, then reading its value is not surprising at all and we say that it contains zero bits of information as we do not learn anything from it. In general, a two level classical system with probabilities p, $1-p$ of being in 0, 1 respectively contains

$$h(p) = -p\log_2(p) - (1-p)\log_2(1-p) \tag{1.2}$$

bits of information. Here, the function $h(x)$ is called the *binary entropy*. In general, for a random variable K with n realisations $k_1, \ldots, k_n$ each having probability $p(1), \ldots, p(n)$, the information content can be described by the *Shannon entropy*

$$S(K) = -\sum_{x=1}^{n} p(x) \log_2 p(x), \tag{1.3}$$

where the information content is a measure of how surprising it is to learn the possible outcomes of a random variable. More specifically, we can interpret this quantity as an average of how surprising it is to learn K as we are considering the expectation value of $\log_2 p(x)$ over the distribution $p(x)$. In the context of cryptography, it is also useful to define the *min-entropy* [27] which can be interpreted as the measure of how surprising it is to learn K in the worse case scenario. We can write this quantity as

$$S_{\min}(K) = -\log_2 \left(\max_{x=1,\ldots,n} p(x) \right), \tag{1.4}$$

and it is always a lower bound on the Shannon entropy.

The concept of quantum information is defined analogously, but it contains an important difference. As we pointed out before, the surprise in learning the outcome of the measurement of a quantum state, depends on the basis that we choose to perform the measurement. For this reason, we describe the quantum information content as a measure of how surprising it is to learn the outcome of the measurement of a quantum system, optimized by selecting the measurement that minimizes the average surprise or, equivalently, the measurement that disturbs the least the original state of the system. Such measurement basis is given exactly by the eigenbasis of the state of the system. Therefore, the content of quantum information of a state described by a density matrix ρ can be quantified by

$$H(\rho) = -\operatorname{Tr}(\rho \log_2 \rho) = -\sum_{x=1}^{n} \lambda_i \log_2 \lambda_i, \tag{1.5}$$

where λ_i are the eigenvalues of ρ. This measure of information is called the *Von Neumann entropy*. We will often use the notation $H(A)$ to indicate the Von Neumann entropy of the state of the system A. Similarly to the classical case, we can define the quantum version of the min-entropy as

$$H_{\min}(\rho) = -\max_{0 \leq \Pi \leq \mathbb{1}} \log_2 \operatorname{Tr}(\Pi \rho). \tag{1.6}$$

A further quantity that we will use throughout this thesis is the so-called conditional Von Neumann entropy. This measure comes in handy when we work with a state ρ_{AB} that describes the joint state of two subsystems A and B. In this case, the *conditional Von Neumann entropy* of the system A given B is

$$H(A|B)_{\rho_{AB}} = H(AB)_{\rho_{AB}} - H(B)_{\rho_B} = -\operatorname{Tr}(\rho_{AB} \log_2 \rho_{AB} - \rho_B \log_2 \rho_B), \tag{1.7}$$

and it quantifies the remaining information content of the state $\rho_A = \operatorname{Tr}_B(\rho_{AB})$ given that we have access to the state $\rho_B = \operatorname{Tr}_A(\rho_{AB})$.

1.2 Encoding and transmission of quantum information

Quantum information is usually encoded into the various degrees of freedom of photons for the purpose of being transmitted. Photonic degrees of freedom offer many possibilities for encoding quantum information.

The most straightforward physical property of photons where we can encode a qubit is the polarization. Here, a qubit can be prepared in a two-mode state. The first mode represents the presence or absence of a photon in the horizontal polarization (H) and the second mode represents the presence or absence of a photon in the vertical polarization (V). Let $|n\rangle$ represent a photon number state and $|n\rangle_H |m\rangle_V = |n,m\rangle_{H,V}$ a two mode photon number state with n photons polarized horizontally and m photons polarized vertically. A qubit can be expressed as

$$|\psi\rangle = \alpha\,|0,1\rangle_{H,V} + \beta\,|1,0\rangle_{H,V}\,. \tag{1.8}$$

Defining $|0\rangle \coloneqq |0,1\rangle_{H,V}$ and $|1\rangle \coloneqq |1,0\rangle_{H,V}$, we can write our qubit in the form of eq. (1.1). Expanding this concept, a d-dimensional state can be created by placing photons in different modes and creating a superposition as follows

$$|\psi\rangle = \alpha_1\,|0,\ldots,0,1\rangle + \alpha_2\,|0,\ldots,0,1,0\rangle + \cdots + \alpha_d\,|1,0,\ldots,0\rangle\,. \tag{1.9}$$

Other qubit encoding techniques include time-bin encoding, where we send photons in different time windows to represent different qubit states or frequency encoding, where we encode the qubit states in different photon frequencies.

Additionally, photons can be used to transmit infinite-dimensional quantum states by encoding the information in the optical phase space. A notable example is the coherent state. We can write this state in terms of the photon number states as

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}}\,|n\rangle\,, \tag{1.10}$$

where $\alpha \in \mathbb{C}$. In this configuration, the information can be encoded into the Q and P quadratures of the phase space (or into combinations of the two). Coherent states have a Gaussian distribution in the phase space. The expected values for Q and P are the real and imaginary parts of α and the variance in both quadratures is $\frac{1}{2}$ (Fig. 1.3). Quadratures can be measured through homodyne detection.

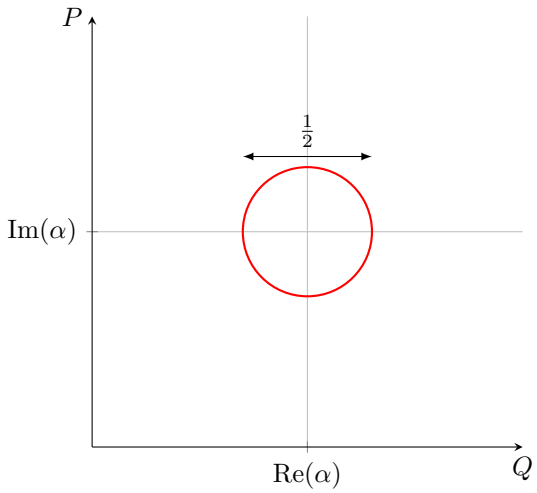

Figure 1.3: A coherent state $|\alpha\rangle$ in the phase space.

Coherent states are a particular type of Gaussian states. Gaussian states are a class of quantum states characterized by a Gaussian Wigner function, which is a quasiprobability distribution in

the phase space. Additionally, channels that map Gaussian states to Gaussian states are denoted as Gaussian channels. Gaussian states and channels encompass a wide range of operations that are relatively easy to work with and can be realized in a laboratory. This framework is detailed in sources such as [28, 29].

Photons are usually transmitted using optical fibers or through free space and, during the transmission, the information that we encoded in their degrees of freedom is susceptible to corruption. Additionally, the instruments used to prepare and to measure the photons often have imperfections that further contribute to the degradation of the transmitted information. In the following, we will describe some of the models that are used to describe and simulate this corruption.

1.2.1 Dephasing and white noise

We begin with the *dephasing noise* model. In this model, the original state partially loses its off-diagonal terms. Let ρ be the ideal quantum state prepared by a source, a dephasing channel Γ_v acts on ρ as

$$\Gamma_v(\rho) = v\rho + (1-v)\mathrm{diag}(\rho), \tag{1.11}$$

where $\mathrm{diag}(\rho) = \sum_i \rho_{ii} |i\rangle\langle i|$ returns the diagonal part of ρ in a given basis $|i\rangle$ and v is denoted as *visibility*. Similarly, the *white noise* model dilutes the original state by mixing it with a state that introduces uniform noise across each diagonal element. The white noise channel acts on ρ as

$$W_v(\rho) = v\rho + (1-v)\frac{\mathbb{1}}{d}, \tag{1.12}$$

where d is the dimension of ρ.

1.2.2 Photon losses

Losses represent another common source of corruption of the information carried by photons. Photons tend to be lost during transmission or during detection. Optical fibers and other optical components can be described as beam splitters of transmittances η_{fiber} and η_{opt} respectively, acting on the incoming state. Similarly, imperfect photon detectors can be described by a beam splitter (with transmittance η_{det}) followed by the ideal detector. Overall, the effect of optical fibers, optical components, and imperfect detectors can be described as a single beam splitter of transmittance $\eta = \eta_{\mathrm{fiber}} \cdot \eta_{\mathrm{opt}} \cdot \eta_{\mathrm{det}}$, referred to as *detection efficiency*. The state ρ is combined at this beam splitter with a vacuum state $|0\rangle$ and one of the two output ports is traced out as in Fig. 1.4.

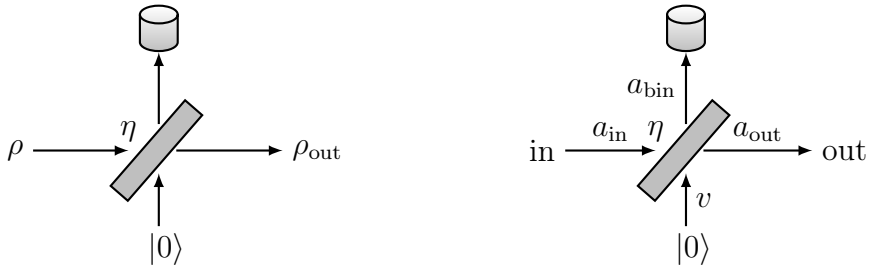

Figure 1.4: Model for photon losses. On the left hand side, the figure represents the model in terms of states, while on the right hand side, it is represented in terms of ladder operators a_{in}, v, a_{out}, a_{bin}.

The ladder operators of the input (a_{in}, v) and output (a_{out}, a_{bin}) ports of a beam splitter with transmittance η follow the relation

$$\begin{pmatrix} a_{\text{out}} \\ a_{\text{bin}} \end{pmatrix} = \begin{pmatrix} \sqrt{\eta} & \sqrt{1-\eta} \\ -\sqrt{1-\eta} & \sqrt{\eta} \end{pmatrix} \begin{pmatrix} a_{\text{in}} \\ v \end{pmatrix}. \tag{1.13}$$

The action of a beam splitter U_η on a photon $|1\rangle_{\text{in}}$ (with the vacuum state at the second input port) is

$$\begin{aligned} U_\eta |1\rangle_{\text{in}} |0\rangle = U_\eta a_{\text{in}}^\dagger |0\rangle_{\text{in}} |0\rangle &= (\sqrt{\eta} a_{\text{out}}^\dagger + \sqrt{1-\eta} a_{\text{bin}}^\dagger) |0\rangle_{\text{out}} |0\rangle_{\text{bin}} \\ &= \sqrt{\eta} |1\rangle_{\text{out}} |0\rangle_{\text{bin}} + \sqrt{1-\eta} |0\rangle_{\text{out}} |1\rangle_{\text{bin}}. \end{aligned} \tag{1.14}$$

The impact of losses on a state of the type (1.9) can be described by appending a vacuum state to each mode of the initial state and applying the transformation given by eq. (1.14) to each pair of modes

$$\begin{aligned} U_\eta^{\otimes n} |\psi\rangle_{\text{in}} |0,\dots,0\rangle = \alpha_1 \big(\sqrt{\eta} |1,0,\dots,0\rangle_{\text{out}} |0,\dots,0\rangle_{\text{bin}} - \sqrt{1-\eta} |0,\dots,0\rangle_{\text{out}} |1,0,\dots,0\rangle_{\text{bin}}\big) + \dots \\ + \alpha_n \big(\sqrt{\eta} |0,\dots,0,1\rangle_{\text{out}} |0,\dots,0\rangle_{\text{bin}} - \sqrt{1-\eta} |0,\dots,0\rangle_{\text{out}} |0,\dots,0,1\rangle_{\text{bin}}\big) \\ = \sqrt{\eta} |\psi\rangle_{\text{out}} |0,\dots,0\rangle_{\text{bin}} + \sqrt{1-\eta} |0,\dots,0\rangle_{\text{out}} |\psi\rangle_{\text{bin}}. \end{aligned}$$

Finally, by tracing out the environment modes, we obtain the following output state

$$\rho_{\text{out}} = \text{Tr}_E \left(U_\eta^{\otimes n} |\psi\rangle_{\text{in}} |0,\dots,0\rangle \langle\psi|_{\text{in}} \langle 0,\dots,0| U_\eta^{\otimes n\,\dagger} \right) \tag{1.15}$$

$$= \eta \, \psi_{\text{out}} + (1-\eta) |0,\dots,0\rangle\langle 0,\dots,0|_{\text{out}}. \tag{1.16}$$

Therefore, when we measure the resulting state, the probability of obtaining the same outcome as in the lossless case is η. Conversely, with a probability of $1-\eta$, we measure the vacuum state. The latter case is denoted as *no-detection event* or *no-click event*.

For a coherent state, the effect of losses is

$$U_\eta |\alpha\rangle_{\text{in}} |0\rangle = |\sqrt{\eta}\,\alpha\rangle_{\text{out}} \left| \sqrt{1-\eta}\,\alpha \right\rangle_{\text{bin}} \tag{1.17}$$

resulting in an attenuated coherent state $\left|\sqrt{\eta}\,\alpha\right\rangle$, if we trace out the environment.

1.2.3 Fair sampling assumption

As discussed in the previous section, no-detection events are one possible outcome in photonic experiments with discrete variables. These outcomes can be handled in two ways: either by including them in the analysis of our experiment or by considering only the subset of events where a photon was detected, discarding the rest. However, in the latter case, the subset may be biased and not representative of the entire set of events. The fair sampling assumption consists of deciding that we can safely discard the no-click events without introducing bias. The name *fair sampling assumption*, which refers to the fact that the subset of considered events is a fair sample of the entire set of events, is mainly used in the context of Bell tests, a subject that will be discussed later in this chapter.

To prevent bias in the considered subset, we adopt a mathematical model to describe photon-detectors. In the type of experiments considered in this thesis, measuring devices are required to perform multiple different measurements during multiple rounds, and the choice of measurement

is performed according to a randomly provided input, that we will here denote as x. In the following, we describe a model which is detailed in [30].

Let us assume that our measuring device is composed of a filter and an ideal detector (Fig. 1.5). The filter outputs a flag $D = \checkmark, \perp$ based on whether our device accepts the round or not. The filter is composed of two parts acting independently on the classical and quantum input. The part acting on the quantum input accepts the round only when the photon was detected. When they both accept the round, we say that our device produces a click. If our device does not click, the round of the protocol gets discarded.

If the filter outputs $D = \checkmark$, our device performs a lossless measurement based on the random input x and yields an outcome a. It is crucial in this model that the choice of the measurement does not influence the output of that part of the filter processing the quantum input. In other words, the fact that the detector clicks or not is independent of the choice x of the measurement.

Modelling our measuring device in this way is equivalent to performing an ideal measurement with unit efficiency and outcomes a.

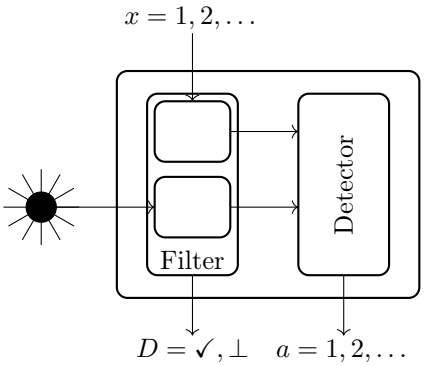

Figure 1.5: A model describing a measuring device under the fair sampling assumption. A filter returns an output $D = \checkmark, \perp$ which is determined independently by the input x and the state ρ. The detector performs a lossless measurement only if the filter returned $D = \checkmark$.

1.2.4 Dark counts

While performing a fair sampling assumption allows us to get rid of the effect of losses, introducing a high number of discarded measurements introduces another problematic effect known as *dark counts*. Dark counts are essentially random clicks generated by detectors resulting from thermal fluctuations inside the detectors. These are of critical importance particularly when operating at low detection efficiencies, as their frequency becomes noteworthy over the fraction of rounds that have not been discarded under the fair sampling assumption.

In the following, we present a model for computing the measurement operators taking into account photon loss and dark counts in two scenarios: one where we discard no-detection events, and one where we do not. For simplicity, we will describe the case of a measuring device with only one input. As previously mentioned, the fair sampling assumption requires multiple measurements. Adapting this model to multiple inputs is straightforward: it involves evolving the measurement operators for each setting accordingly.

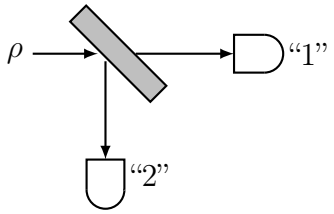

Figure 1.6: Schematic representation of an ideal polarization measurement in the absence of losses and dark counts. A polarizing beam splitter divides the incoming photonic state into two paths. At the two output ports of the beam splitter, we place two detectors and each of them corresponds to one of the two possible outcomes. Outcomes "1" and "2" correspond to two orthogonal polarizations of the photon.

In Fig. 1.6, we represent an ideal measurement of the polarization of a photon. Let us initially address photon loss. When a photon arrives at the measurement device (with probability η), it encounters a polarizing beam splitter, which then directs it to one of two detectors based on its polarization state. If the photon is lost (with probability $1-\eta$), then neither detector clicks.

Let us now evaluate the POVMs describing the aforementioned scenario. We will take into account the ideal POVMs M_1 and M_2 of the photon being detected by the first and second detectors, respectively. The measurement operators evolve as

$$M_a^\eta = \begin{cases} \eta M_a & \text{if a detector clicks} \\ (1-\eta)\mathbb{1} & \text{if neither detector clicks.} \end{cases} \tag{1.18}$$

Note that measuring the POVMs in eq. (1.18) on an ideal qubit state is equivalent[2] to evolving a qubit state as in eq. (1.16) and performing the measurement

$$\tilde{M}_a = \begin{cases} |0,1\rangle\langle 0,1| & \text{for outcome "1",} \\ |1,0\rangle\langle 1,0| & \text{for outcome "2",} \\ |0,0\rangle\langle 0,0| & \text{for the no-click outcome.} \end{cases} \tag{1.19}$$

Let us now assume that our detectors exhibit also a dark count behavior, with each of the two possible outcomes having a probability of p_d of registering a spurious click in every round of the protocol. Our setup does not allow us to differentiate whether the detection was due to a genuine photon detection, or a dark count.

Let us delineate the four possible detection outcomes.

1. **Single click in the first detector:** This outcome indicates that only the detector "1" clicked. It could arise from the detection of a photon with a specific polarization or a dark count in the first detector.

2. **Single click in the second detector:** This outcome indicates that only the detector "2" clicked.

3. **Double click:** It occurs when both detectors register a click simultaneously. This can arise from both detectors experiencing dark counts, or one detecting a photon while the other registers a dark count.

4. **No click:** When neither detector registers a click, suggesting the photon was lost and no dark counts occurred.

[2]By equivalent here we mean that it returns the same probability distribution.

Let us now evaluate the POVMs related to each of these four events. We will take into account the efficiency η of the photon's arrival at the detectors, and the probability p_d of a dark count occurring in each detector.

Assuming the photon arrives at the measurement device, four possible events can occur. We list them in table 1.1.

Event	**POVM**	**Outcome Type**
Photon at first detector, no dark count at second	$\eta \cdot (1-p_d) \cdot M_1$	Single click in the first
Photon at second detector, no dark count at first	$\eta \cdot (1-p_d) \cdot M_2$	Single click in the second
Photon at first detector, dark count at second	$\eta \cdot p_d \cdot M_1$	Double click
Photon at second detector, dark count at first	$\eta \cdot p_d \cdot M_2$	Double click

Table 1.1: POVMs of detection events with photon arrival, considering dark counts and categorized by outcome type.

For the scenario where the photon does not arrive, we have four cases which we list in table 1.2.

Event	**POVM**	**Outcome Type**
No dark count in either detector	$(1-\eta) \cdot (1-p_d)^2 \cdot \mathbb{1}$	No click
Dark count in the first detector only	$(1-\eta) \cdot p_d \cdot (1-p_d) \cdot \mathbb{1}$	Single click in the first
Dark count in the second detector only	$(1-\eta) \cdot (1-p_d) \cdot p_d \cdot \mathbb{1}$	Single click in the second
Dark counts in both detectors	$(1-\eta) \cdot p_d^2 \cdot \mathbb{1}$	Double click

Table 1.2: POVMs of detection events without photon arrival, categorized by outcome type.

To compute the measurement operators without making the fair sampling assumption, given the ideal measurement operators N_1 for outcome "1" and N_2 for outcome "2", we need to consider the probabilities of each event as previously discussed. We obtain in this way four measurement operators $N_b(\eta_B, p_d)$ corresponding to the four outcome types.

$$N_1(\eta, p_d) = \eta(1-p_d)N_1 + (1-\eta)p_d(1-p_d)\mathbb{1}, \tag{1.20}$$

$$N_2(\eta, p_d) = \eta(1-p_d)N_2 + (1-\eta)p_d(1-p_d)\mathbb{1}, \tag{1.21}$$

$$N_{\mathrm{Dc}}(\eta, p_d) = \left(p_d\eta + (1-\eta)p_d^2\right)\mathbb{1}, \tag{1.22}$$

$$N_{\mathrm{Nc}}(\eta, p_d) = (1-\eta)(1-p_d)^2\mathbb{1}. \tag{1.23}$$

For simplicity, we can group no clicks and double clicks into a single outcome, obtaining in this way

$$N_{\varnothing}(\eta, p_d) = \left(p_d\eta + (1-\eta)(p_d^2 + (1-p_d)^2)\right)\mathbb{1}. \tag{1.24}$$

Let us finally focus on the measurements under the fair sampling. Our ideal POVMs are denoted as M_a. Given the assumption that the probability of a click in our detector is independent of the basis choice, we are allowed to discard events where both our detectors do not click or where they both click. As we discard these events, we will need to renormalise our POVMs dividing them by the probability of having only one click. Consequently, the measurement operators $M_a(\eta, p)$ can be expressed in terms of the ideal ones M_1 as

$$M_1(\eta, p_d) = \frac{\eta(1-p_d)M_1 + (1-\eta)p_d(1-p_d)\mathbb{1}}{1 - p_d\eta - (1-\eta)(p_d^2 + (1-p_d)^2)}, \tag{1.25}$$

and $M_2(\eta, p_d) = \mathbb{1} - M_1(\eta, p_d)$.

1.2.5 Thermal noise

In earlier discussions, our focus was primarily on noise models suitable for finite-dimensional systems. However, for infinite-dimensional quantum systems and continuous variable quantum communication protocols, a more representative model of photonic state degradation involves thermal noise. Similarly to the photon loss model, which mixes the input state with a vacuum state at a beam splitter, the thermal noise model mixes it with a thermal state. Thermal states can be expressed in terms of the photon number basis as

$$\tau_\nu = \frac{1}{1+\nu}\sum_{n=0}^{\infty}\left(\frac{\nu}{1+\nu}\right)^n |n\rangle\langle n|\,, \tag{1.26}$$

where ν is the mean photon number. These states, similarly to the coherent ones, have a Gaussian distribution in the phase space, are centred at the origin, and have a variance of $\frac{1}{2}+\nu$. For $\nu \to 0$ the thermal state converges to a vacuum state, reverting this model to the photon loss model. The thermal state can equivalently be expressed in the basis of coherent states $|\beta\rangle$ as

$$\tau_\nu = \frac{1}{\pi\nu}\int_{\mathbb{C}} d^2\beta e^{-|\beta|^2/\nu}\,|\beta\rangle\langle\beta|\,. \tag{1.27}$$

This noise model is typically described in terms of the transmittance η of the beam splitter and of the *excess noise*

$$\epsilon = \frac{2\nu(1-\eta)}{\eta}. \tag{1.28}$$

Let us conclude by discussing the evolution of coherent states under thermal noise. From a phase space perspective, the smaller the transmittance, the more the input state moves to the origin of the phase space and the more the excess noise, the greater the variance of the output state. In particular, a coherent state $|\alpha\rangle$ evolves under this model to

$$\rho_{\text{out}} = \int d^2\beta \frac{2}{\epsilon\eta\pi}\exp\left(-\frac{2|\beta-\sqrt{\eta}\alpha|^2}{\epsilon\eta}\right)|\beta\rangle\langle\beta|\,, \tag{1.29}$$

which is a Gaussian state centred in $\sqrt{\eta}\alpha$ and with variance $\frac{1+\epsilon\eta}{2}$ (Fig. 1.7).

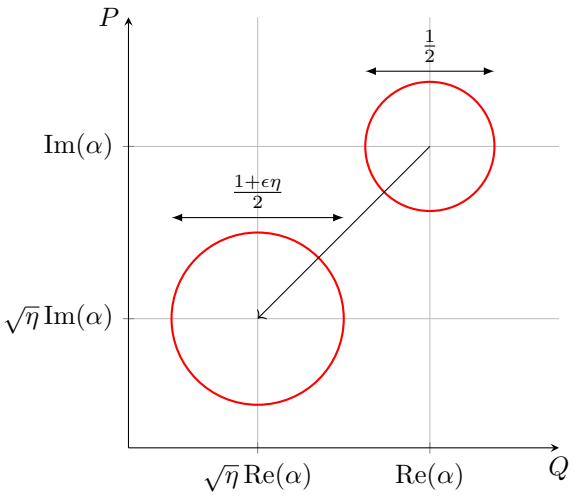

Figure 1.7: A coherent state $|\alpha\rangle$ evolving in the phase space under a thermal noise channel with transmittance η and excess noise ϵ.

1.3 Quantum Key Distribution

Quantum Key Distribution is a technique that allows for the first time to achieve secure communication in a provable manner by distributing a pair of secret keys to two distant parties.

Let us describe some of the most famous QKD protocols in their simplest form. As it is typically done in cryptography, we will describe our scheme using two fictitious parties, Alice and Bob, willing to establish secure communication.

1.3.1 BB84 protocol

The BB84 was introduced in 1984 by C. H. Bennett and G. Brassard and it was the first QKD protocol to be invented [8]. The protocol is described below.

1. **State Preparation:** Alice randomly prepares one of four quantum states: $|0\rangle$, $|1\rangle$, $|+\rangle$, or $|-\rangle$, where
$$|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \quad |-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \tag{1.30}$$
Alice keeps track of what state she prepared and what basis she used (Z for $|0\rangle$ and $|1\rangle$ or X otherwise).

2. **Transmission:** Alice sends the prepared state to Bob through an untrusted quantum channel, susceptible to interception by an eavesdropper, which we will call Eve.

3. **Measurement:** Upon receiving the state, Bob randomly chooses to measure using one of two observables:
$$Z = |0\rangle\langle 0| - |1\rangle\langle 1|, \quad \text{or} \quad X = |0\rangle\langle 1| + |1\rangle\langle 0|. \tag{1.31}$$
Bob notes the observable used and the outcome. Alice and Bob repeat the first three steps for n rounds.

4. **Public Communication:** Alice and Bob perform classical communication through a public but authenticated channel. This means that Eve might be listening but she cannot pretend to be one of the two honest parties. During this step, Alice informs Bob about the basis (Z or X) used for the state preparation at each round, without revealing the specific state.

5. **Basis Reconciliation:** Bob tells Alice whether his measurements matched the basis used by Alice. They discard any rounds where the bases did not match. When they used the same basis, Bob's outcome and Alice's prepared state should match.

6. **Error Checking:** They reveal a subset of their outcomes to detect any interference by Eve. If too many errors are found, they abort the protocol.

7. **Key Generation:** If the error rate is acceptable, Alice and Bob use the remaining, unrevealed, not-discarded outcomes to generate a shared secret key.

It is easy to verify that, if Eve has been measuring the states sent by Alice, then the outcomes of Alice and Bob will not match perfectly as her measurements would alter the outcomes of Bob.

At the end of the protocol, if Eve did not interfere and the protocol proceeded without errors, the two honest parties will share two perfectly matching strings composed by the outcomes of the retained measurements of Z and X.

In real experiments, it is extremely unlikely that the protocol will run without errors. As discussed in Section 1.2, quantum information tends to arrive at the receiver in a corrupted form. Consequently, after completing the steps above, Alice and Bob usually end up with a pair of imperfect keys, denoted as *raw keys*, that do not exactly match and may have been partially eavesdropped by Eve. Therefore, QKD security proofs establish bounds on Eve's potential correlation with the secret key and the correlation between Alice and Bob. Should the security criteria be met, Alice and Bob can employ *information reconciliation* and *privacy amplification* algorithms. These algorithms reduce the length of the raw key in exchange for producing a pair of perfectly matching and perfectly private keys, up to a reasonably low failing probability. These keys can be used to perform secure communication in a symmetric key algorithm.

1.3.2 B92 protocol

A few years later, following his work on the BB84 protocol, Bennett introduced a second QKD protocol using similar principles. This is known as the B92 protocol [31]. Similarly to before, Alice and Bob need an untrusted quantum channel and an authenticated classical channel for this protocol. The steps are outlined below.

1. **State Preparation and Transmission:** Alice sends one of two states to Bob: $|0\rangle$ or $|+\rangle$.
2. **Measurement and Public Communication:** Bob measures the incoming quantum state using the Z or X observables, as in BB84. He publicly communicates that he discards the round if the measurement outcome corresponds to the eigenvalue $+1$, since it doesn't reveal the state Alice sent.
3. **States Discrimination:**
   - If the outcome is -1 and Bob used Z, the state must have been $|+\rangle$.
   - If the outcome is -1 and Bob used X, the state must have been $|0\rangle$.

   Alice records a '0' for $|0\rangle$ and a '1' for $|+\rangle$. Bob records a '0' for retained measurements where he used X and a '1' for Z.
4. **Error Checking:** As with BB84, they publicly reveal a subset of their outcomes to check for eavesdropping. If discrepancies are found indicating Eve's interference, they abort the protocol.
5. **Key Generation:** If no interference is detected, the retained and unrevealed outcomes form the basis of a secure shared key.

As before, if Eve did not interfere and the protocol proceeded without errors, it is easy to verify that the bits stored by Alice and Bob will match perfectly. In a real protocol, we establish a bound on Eve's information and the correlation between Alice and Bob and decide whether to run information reconciliation and privacy amplification to distil perfect keys or whether to abort the protocol.

1.3.3 Entanglement-based protocols

Previously, we discussed the BB84 and B92 protocols in their prepare-and-measure format, where Alice prepares a state each round, sends it to Bob, and Bob then measures the received state. However, it is possible to achieve the same statistical results using entanglement-based methods, where a source prepares entangled states that are distributed to both Alice and Bob for measurement.

Entanglement in BB84: For the BB84 protocol, an equivalent entanglement-based approach involves the use of the maximally entangled state

$$|\phi\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}. \tag{1.32}$$

Here, Alice measures one of the subsystems using the observables Z and X with equal probability. This setup ensures that Bob's subsystem will collapse into one of the four states $|0\rangle$, $|1\rangle$, $|+\rangle$, or $|-\rangle$, exactly as in the prepare-and-measure scenario.

Entanglement in B92: For the B92 protocol, the source should prepare the state

$$|\phi\rangle = \frac{|0\rangle\,|0\rangle + |1\rangle\,|+\rangle}{\sqrt{2}}. \tag{1.33}$$

Alice then measures the first subsystem in the Z basis. Consequently, the second subsystem, which is sent to Bob, will end up in one of the two states $|0\rangle$ or $|+\rangle$, just as it would in the prepare-and-measure version.

1.4 Untrusted devices

In the previous section, we described QKD protocols where Alice and Bob trust their preparation and measuring devices. As we explained in the introduction, it became soon evident that, because of such trust, this technology contained some loopholes that could be exploited to hack and eavesdrop on the communication protocol [19–22]. For this reason, in parallel with the development of QKD, researchers began to explore a new technology that could have guaranteed higher levels of security and would have prevented the loopholes in question to be exploited. This new technology is called Device Independent QKD. The DI approach, initially developed in the context of Bell tests, has applications across various areas of quantum information, which we will now shortly outline.

1.4.1 Bell test

In the first decades of the 20th century, the scientific community started debating around the nature of randomness in quantum mechanics, particularly whether it could be explained through hidden variables that might predict outcomes deemed random by quantum theory.

In 1964, J. S. Bell proposed a thought experiment designed in a way that it could falsify any local hidden variables theory [32]. We will describe a relatively simple example of a *Bell experiment*. For an extensive review on Bell non-locality see for example [33]. According to local hidden variable theories, the results in setups like the one shown in Fig. 1.2 can be described in terms of a probability distribution

$$p(a,b|x,y) = \int d\lambda\, q(\lambda) p(a|x,\lambda) p(b|y,\lambda), \tag{1.34}$$

where a hidden variable λ, occurring with probability $q(\lambda)$, supposedly accounts for every statistical correlation between Alice and Bob's devices.

Assuming that the values of a and b are either $+1$ or -1, and that x and y can be 1 or 2, we define the quantity

$$S = \langle a_1 b_1 \rangle + \langle a_1 b_2 \rangle + \langle a_2 b_1 \rangle - \langle a_2 b_2 \rangle, \tag{1.35}$$

where $\langle a_x b_y \rangle = \sum_{a,b} a \cdot b \cdot p(a,b|x,y)$. It can be shown that for distributions like the one in equation (1.34), the value of S is constrained by the inequality $S \leq 2$, known as the CHSH inequality [34].

However, quantum theory allows for scenarios where this inequality is violated. For example, using a maximally entangled state with observables

$$A_1 = Z, \quad A_2 = X, \quad B_1 = \frac{Z+X}{\sqrt{2}}, \quad B_2 = \frac{Z-X}{\sqrt{2}}, \tag{1.36}$$

we find

$$S = \langle A_1 \otimes (B_1 + B_2) \rangle + \langle A_2 \otimes (B_1 - B_2) \rangle = \frac{1}{\sqrt{2}} (2\langle Z \otimes Z \rangle + 2\langle X \otimes X \rangle) = 2\sqrt{2}, \tag{1.37}$$

which violates the CHSH inequality.

This example provides a specific strategy to achieve $S = 2\sqrt{2}$. However, to experimentally demonstrate the violation of the CHSH inequality, it is not necessary to make assumptions on the specific form of states and measurements. It is sufficient to look at the frequencies of the outcomes for given inputs and compute the probability distribution $p(a,b|x,y)$, which can then be used to calculate S.

The first empirical demonstration of the violation of this inequality was conducted by S. J. Freedman and J. F. Clauser in 1972 [35]. However, this experiment, like others conducted over the subsequent decades, was not entirely conclusive. Specifically, they all suffered from *loopholes*, which are technical flaws in experimental implementations that leave room for alternative explanations involving hidden variables.

One significant loophole relevant to this thesis is the *detection loophole*. This issue arises when experiments discard no-click outcomes. These kinds of experiments do not allow to falsify all hidden variable models. In principle, we cannot exclude that nature is causing the discarding of rounds that would otherwise not support a violation of Bell's inequality. To falsify every possible hidden variable model, it is necessary to perform an experiment without employing the fair sampling assumption discussed in Section 1.2.3. This means that we need to retain all outcomes, because the no-click ones might be dependent on the inputs that have been generated. Importantly, the detection loophole is not only a theoretical concern in Bell tests, it can also lead to practical security vulnerabilities when detectors do not accurately match their theoretical descriptions. For instance, attacks on the first commercial QKD systems [19–22] exploited these discrepancies. Attackers manipulated the detectors to control detection efficiencies based on the measurement settings, effectively reopening the detection loophole. By exploiting the fact that the detectors' real-world behaviour deviated from idealized models, they were able to bias which detection events occurred, potentially compromising the security of the system. This issue will be described with a simple example in the next section and it will be one of the main focuses of the next chapter.

To avoid this loophole, experiments must account for all outcomes, including those in photonic implementations where no photons are detected. No-detection outcomes, for instance, can be mapped to another outcome category, effectively modifying the observables O used by Alice and Bob to $\eta O + (1-\eta)\mathbb{1}$. For a maximally entangled state, this adjustment gives:

$$S = \eta^2 2\sqrt{2} + (1-\eta)^2 2, \tag{1.38}$$

which exceeds the value of 2 only if $\eta \gtrsim 82.8\%$. For partially entangled states with an angle that maximizes the value of S for each η, the threshold detection efficiency required is $\eta > 2/3$.

Achieving such high detection efficiencies, while addressing other known loopholes, demanded extensive experimental advancements over many years. In 2015, three different research groups

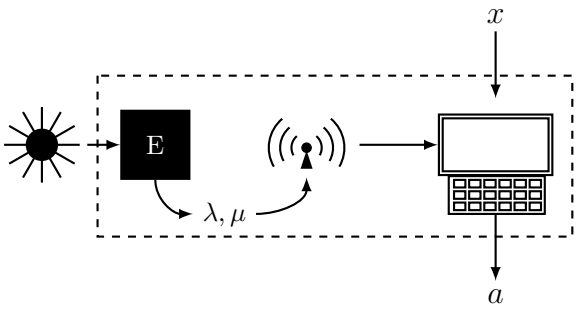

Figure 1.8: An eavesdropper performs a measurement E on the incoming state, obtains outcomes λ and μ, and processes the results.

in Delft, Vienna, and Boulder independently conducted the first loophole-free Bell tests [36–38]. The achievements of all those years of research led to the 2022 Nobel prize in Physics which was awarded to A. Aspect, J. F. Clauser, and A. Zeilinger "for experiments with entangled photons, establishing the violation of Bell inequalities and pioneering quantum information science".

1.4.2 Discarding no-clicks on untrusted devices

When the internal workings of measuring devices are not well characterised, discarding no-click events can introduce a bias in the subset of retained events. This bias can be exploited by a third party to infer the outcomes of the device or, in the case of Bell tests, allow for an explanation of the experiment through hidden variables. Let us illustrate this with an example.

Specifically, we will demonstrate that when the detection efficiency η is $\frac{1}{2}$, discarding no-detection events can create the false impression that a two-measurement experiment is conclusive. Without discarding these events, the experiment can be explained in two ways: either both measurements are randomly being performed, or an alternative explanation exists where only one measurement is performed and the results are manipulated to mimic the two-measurement scenario. When no-detection events are discarded, it may misleadingly appear that the experiment is conclusively demonstrating the two-measurement scenario, even though it remains inconclusive.

Consider a measuring device that performs measurements of the Z and X observables with equal probability. Before discarding no-clicks, the measurements can be represented with the POVMs

$$M_{1|1}^{(\eta)} = \eta\,|0\rangle\langle 0|\,, \quad M_{2|1}^{(\eta)} = \eta\,|1\rangle\langle 1|\,, \quad M_{1|2}^{(\eta)} = \eta\,|+\rangle\langle +|\,, \quad M_{2|2}^{(\eta)} = \eta\,|-\rangle\langle -|\,, \quad M_{\varnothing|x}^{(\eta)} = (1-\eta)\mathbb{1}. \tag{1.39}$$

We will demonstrate that an eavesdropper can perform a single measurement on the incoming state and process the outcomes to reproduce the same statistics as in Eq. (1.39). This scenario is depicted in Fig. 1.8. The eavesdropper's single measurement is a two-outcome measurement described by the POVMs

$$E_{1,1} = \frac{1}{2}\,|0\rangle\langle 0|\,, \quad E_{2,1} = \frac{1}{2}\,|1\rangle\langle 1|\,, \quad E_{1,2} = \frac{1}{2}\,|+\rangle\langle +|\,, \quad E_{2,2} = \frac{1}{2}\,|-\rangle\langle -|\,. \tag{1.40}$$

We will label the first outcome with the variable λ and the second with μ. The measuring device is built in a way that it outputs λ only when the outcome μ matches the device's input x, and

it triggers a no-click event otherwise. In terms of conditional probabilities we have

$$p(a|x,\lambda,\mu) = \delta_{\mu,x}\delta_{\lambda,a}, \quad \text{for } a = 1,2, \tag{1.41}$$
$$p(\varnothing|x,\lambda,\mu) = 1 - \delta_{\mu,x}. \tag{1.42}$$

After this post-processing of the outcomes, the measurement operators become

$$\sum_{\mu,\lambda=1}^{2} p(a|x,\lambda,\mu)E_{\lambda,\mu} = E_{a,x} = M_{a|x}^{(1/2)}, \quad \text{for } a = 1,2, \tag{1.43}$$
$$\sum_{\mu,\lambda=1}^{2} p(\varnothing|x,\lambda,\mu)E_{\lambda,\mu} = \frac{1}{2}\mathbb{1} = M_{\varnothing|x}^{(1/2)}. \tag{1.44}$$

This demonstrates that in an experiment where $\eta = \frac{1}{2}$, we cannot rule out the presence of Eve or the validity of hidden variables theories unless our device is well characterised in the way explained in Section 1.2.3. In Chapter 2, we will see how this strategy can be adapted to more general scenarios.

1.4.3 Steering

The concept of steering captures a scenario in which Bob performs a measurement on his subsystem, influencing the state of Alice's subsystem to a specific post-measurement state, as depicted in Fig. 1.9. In this context, Bob is said to *steer* Alice's state by his choice of measurement. It is crucial to note that this phenomenon does not allow for instantaneous communication because Bob cannot control the outcomes of his measurements directly. Different outcomes will result in different post-measurement states for Alice.

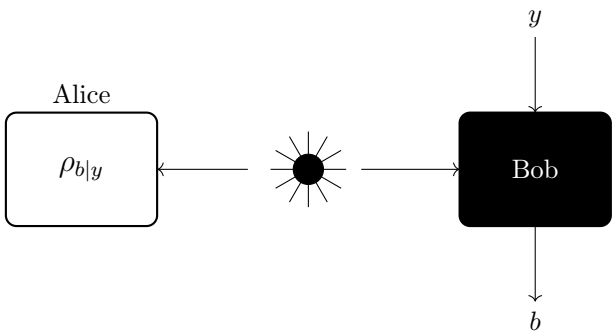

Figure 1.9: A quantum steering experiment. Bob performs a measurement selected by an input y and obtains an outcome b, leaving Alice's subsystem in the state $\rho_{b|y}$.

If Alice's state can be described only by local properties and is independent of Bob's actions, the state ρ_{AB} that describes both Alice and Bob's systems is said unsteerable. This is the case when Alice's state can be modelled using a statistical ensemble of states σ_λ, where σ_λ represents the state of Alice's subsystem corresponding to a hidden variable λ. The hidden variable λ occurs with probability $q(\lambda)$, and Alice's overall state can be written as

$$\rho_{b|y} = \int d\lambda\, q(\lambda)p(a|x,\lambda)\sigma_\lambda. \tag{1.45}$$

In this setting, to determine if Alice's state is unsteerable, it is necessary to perform tomography on her system, which requires trusting her measurement devices. Conversely, trust in Bob's measurement devices is not required as only the knowledge of his inputs and outputs is needed to evaluate the steerability of Alice's states. Therefore, the steering scenario is usually denoted as a one-sided device-independent scenario. For a comprehensive review on steering, see [39].

1.4.4 Device Independent QKD

Finally, we discuss the case of Device Independent Quantum Key Distribution, which is the primary focus of this thesis. As we outlined in Section 1.4.1, the results of the loophole-free Bell tests have shown that nature cannot be described in terms of hidden variables, meaning that the outcomes of some measurements are intrinsically random and cannot be predicted. Because such measurement outputs are unpredictable, Bell tests can be exploited to generate secret keys that are unknown to any third party. This is the intuition at the base of DI QKD.

In DI QKD, Alice and Bob aim to distribute a secret key without relying on specific mathematical models to describe their measurement devices or the quantum states generated. This section provides an intuitive example of how DI QKD can be implemented, while a comprehensive discussion of the assumptions and security proofs will be detailed in Chapters 3-5.

Consider a setup similar to the one shown in Fig. 1.2, where a protocol is executed over many rounds. In each round, Alice performs one of two measurements, A_1 or A_2, while Bob chooses from three measurements, B_1, B_2, or B_3 and they both store their outcomes. At the conclusion of the protocol, Alice and Bob disclose the inputs and outputs for a fraction of these measurements and use them to compute the CHSH parameter S, as defined in eq. (1.35). If the computed value of S reaches $2\sqrt{2}$, it implies [40, 41] that the quantum state they shared was maximally entangled and their measurements were, up to local rotations, given by

$$A_1 = Z, \quad A_2 = X, \quad B_1 = \frac{Z+X}{\sqrt{2}}, \quad B_2 = \frac{Z-X}{\sqrt{2}}. \tag{1.46}$$

Since they share a maximally entangled state, the monogamy of entanglement [42] ensures that no third party can be correlated with Alice and Bob. If Bob's measurement B_3 correlates with Alice's measurement A_1 (e.g., if $A_1 = Z$ and $B_3 = Z$), they can derive a secret key from the measurement outcomes corresponding to the inputs $x = 1$ and $y = 3$. A similar version of this protocol was first proposed by A. Ekert in 1991 [43].

This protocol allows for the exclusion of eavesdropping by Eve without dependence on the specific forms of the measurements or states involved. However, in practical settings, achieving $S = 2\sqrt{2}$ is highly unlikely due to noise, photon losses, and other imperfections in implementation, leading to a lower value of S. The methods for extracting a secret key under these imperfect conditions will be elaborated in Chapter 5.

Lastly, if Alice and Bob are located in the same laboratory, they can also use the outcomes of their measurements to generate randomness in a device-independent manner, ensuring that the resultant random strings are not correlated with any external entity. This process, known as *DI randomness expansion*, requires an initial amount of randomness to determine the measurement settings. The final random string combines this initial seed (the part that was not revealed to compute S) with the new random bits generated from the measurement outcomes.

1.4.5 Semi-Device Independent QKD

We have seen that Quantum Key Distribution protocols rely on ensuring that quantum devices behave as expected. If this is not done correctly we potentially open security loopholes in

the protocol. In contrast, Device Independent Quantum Key Distribution protocols do not require such precise characterization of the devices. However, DIQKD protocols are currently strongly limited in distance. This follows from the fact that they are much more sensitive than device-dependent protocols to photon losses, which in optical fibers scale exponentially with the distance. To address these challenges, one possible approach is to incorporate well-chosen additional assumptions and ensure their validity. Many approaches have been proposed along these lines. Let us shortly describe some of the most famous ones.

The first semi-DI protocol we will introduce is Measurement-Device Independent QKD (MDI-QKD) [44, 45]. As depicted in Fig. 1.10, in this scenario two trusted parties send quantum states to a central untrusted measuring device controlled by a referee. The two sources prepare one of the four BB84 states $|0\rangle$, $|1\rangle$, $|+\rangle$, or $|-\rangle$ according to random inputs x and y. The untrusted referee performs a Bell State Measurement on the two incoming states and publicly announces the outcome c. This allows Alice and Bob to discard rounds where they did not end up correctly correlated. At the end of the protocol, Alice and Bob communicate the basis used to prepare their states and extract secret keys.

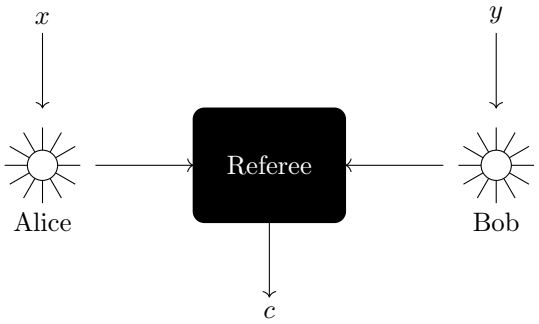

Figure 1.10: A schematic representation of a Measurement-DI QKD protocol. Two trusted sources send quantum states to a central untrusted measuring device. We use the colour black to fill untrusted elements and white for the trusted ones.

A second option are the so called one-sided device-independent (1SDI) protocols, where only one of the two parties in the communication protocol is trusted. This setting can either be of the prepare-and-measure type [46, 47] (Fig. 1.11), where we trust the source, or of the entanglement-based type [48–55], (Fig. 1.11) where we trust a measuring device.

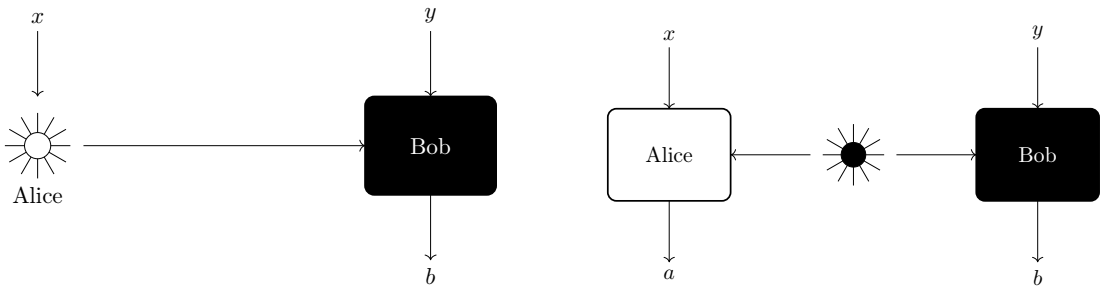

Figure 1.11: A schematic representation of 1SDI QKD protocols: a) in the prepare-and-measure setting where the source is trusted and b) in the entanglement-based setting where one measuring device is trusted.

The prepare-and-measure setting of [46, 47] involves semi-trusted sources with an assump-

tion on the overlap between prepared states. We will explain these protocols in Section 4.3.2. Other assumptions can also be considered, such as restricted distrust in state preparation [56] or constraints on the guessing probability by third parties regarding the prepared states [57]. Generally, anyway, entanglement-based 1SDI settings have an advantage over prepare-and-measure ones since they allow achieving longer distances by placing the source close to the untrusted side, which is more sensitive to losses. In a prepare-and-measure scenario, losses must be accounted for across the entire path from the source to the measuring device. For this reason, in Chapter 6, we will study in detail an entanglement-based 1SDI scenario.

Finally, it is also possible to prove the security of semi-DI scenarios based on dimension bounds, as demonstrated in [58–61].

2 Joint measurability and quantum communication with untrusted devices

In this chapter, we describe a general framework for characterizing the admissible levels of loss and noise in a wide range of scenarios and protocols with untrusted devices. In particular, we present general bounds that apply to prepare-and-measure protocols for the semi-DI approach, as well as to Bell tests for DI protocols. A key step in our work is to establish a general connection between quantum protocols with untrusted devices and the fundamental notions of channel extendibiliy and joint-measurability, which capture essential aspects of the communication and measurement of quantum information. The content of this chapter is reproduced from the first half of [3].

2.1 Introduction

Photon loss, which results from unavoidable absorption and scattering in optical channels, as well as limited detector efficiency, represents one of the key challenges for the implementation of long-distance quantum communication. This issue becomes even more critical in experiments and applications where quantum states are transmitted through untrusted channels and measured using untrusted apparatuses. This is common in quantum cryptography, most notably within the DI [4, 5] or semi-device-independent (SDI) [58, 59] approaches, but also in fundamental tests of quantum physics such as quantum Bell nonlocality [32, 33] and steering [39, 62].

In these situations, certifying quantum properties such as entanglement or randomness, and achieving quantum advantages, such as information-theoretic security, becomes impossible when photon losses exceed a certain threshold. The underlying reason is that the untrusted channels and measurement devices could potentially implement processes in which losses are not merely passive phenomena, but are actively influenced by the experimenter's chosen measurement settings. The most famous example of this is arguably the so-called "detection loophole" [63], that plagued experimental tests of quantum nonlocality for decades, before experiments could finally close it, and, eventually led to the celebrated loophole-free Bell experiments. From a more practical point of view, losses also open the door to potential attacks from eavesdroppers on quantum cryptographic systems, such as blinding attacks [19–22]. In the latter, Eve, the attacker, can remotely manipulate the device of the receiver, such that conclusive detection events will be recorded only for a subset of measurements determined by Eve.

The level of admissible losses is generally highly dependent on the specific quantum communication protocol that is considered. Moreover, this tolerance also depends on the detailed aspects of the experimental setup, including the specific states that are intended to be transmitted through the channel and the expected measurements to be performed on them. An intense effort has been devoted to characterizing critical loss thresholds. This started in the context of Bell experiments (see e.g. [63–66] and [33, 67] for reviews) as well as in steering test (see e.g. [68–70]). In turn, critical efficiencies for DI and SDI protocols have also been discussed (see e.g. [71, 72]). While all these works have played a significant role towards the first loophole-free Bell tests [36–38] and proof-of-principle implementations of DI protocols [23–25], it is fair to say that our understanding is currently limited to few specific cases.

In this chapter, we develop a general framework for characterizing critical losses in quantum communication setups. This allows us to establish upper-bounds on admissible losses that are applicable to a wide range of experiments and protocols involving untrusted measurement devices. The key point of this chapter is to establish a connection with the notion of "joint-measurability", which captures the incompatibility of quantum measurements [73–75].

Our starting point is to introduce the concept of a "channel-measurement unit" (CMU), which represents a fundamental component of quantum communication protocols with untrusted devices. Specifically, the CMU consists of an untrusted communication channel C followed by an untrusted measuring device M, as depicted in Fig. 2.1. The CMU serves as the basic building block for a broad range of setups. The simplest examples are SDI prepare-and-measure scenarios [76, 77], as in Fig. 2.2.a), and bipartite Bell experiment where a pair of entangled photons are sent to two separated CMUs (see Fig. 2.2.b.). These setups serve as a basis for SDI QKD protocols [58, 59, 72]) and DI protocols [4, 23] respectively. More generally, the CMU can also be seen as a part of more complex scenarios, e.g. involving quantum networks or quantum circuits (see Fig. 2.2.c).

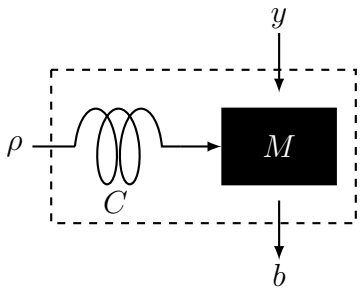

Figure 2.1: A channel-measurement unit (CMU) is composed of an untrusted channel C and an untrusted measurement device M. The CMU receives as input a quantum state ρ that is transmitted through the channel C and then measured by M. The classical input y denotes the choice of measurement (performed e.g. by the honest user), while its output is denoted b.

We aim to determine the admissible level of loss in the CMU before it loses its ability to exhibit 'quantumness', i.e. when it becomes classically simulable. We provide several natural definitions for this, which are shown to be all equivalent to the condition that the effective POVMs implemented by the CMU, representing the combined effect of the channel and the measurement device, are jointly measurable. An equivalent condition, based on the channel extendibility of the CMU, is also particularly convenient to work with. Using these definitions, we provide several upper bounds on the loss threshold. Notably, these bounds typically depend only on the number of measurements implemented by the CMU, and on the noise level, but not on the details of the channel and/or the measurements.

More generally, our work motivates the investigation of joint measurability for sets of measurements that are not only noisy, the case most existing works have focused on [78–82], but that also feature losses, which has been considered only in a limited number of previous works [72, 83–85].

We consider two different scenarios for taking losses into account in protocols with untrusted devices. The first consists of attributing a specific measurement outcome (an inconclusive 'no-click' outcome) to events where the photons are lost. We refer to this as the *no-click scenario*, corresponding e.g. to experiments involving single-photon detectors. The second consists of general optical scenarios where measurements may always produce a conclusive outcome, as, e.g., in continuous variable setups based on homodyne (or heterodyne) measurements. In this

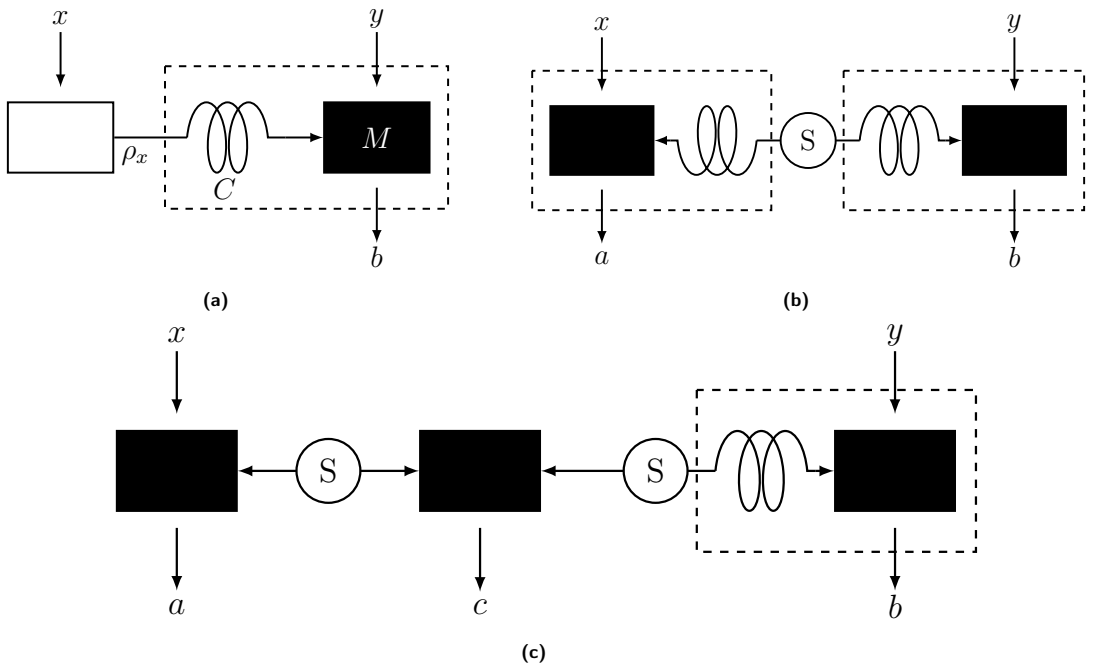

Figure 2.2: The CMU is a basic component of many experiments with untrusted devices, such as (a) a simple prepare-and-measure scenario, (b) A Bell test, (c) a quantum network featuring several sources and measurements

case, we consider the standard model of loss for the optical channel described in Section 1.2.5, corresponding to mixing the input system with a thermal state through a beam splitter of limited transmissivity. We refer to this as the *thermal-noise channel scenario*.

In the no-click scenario, we introduce general bounds on critical loss thresholds, which, in their simplest form, recover or improve on certain known bounds [64, 71]. In particular, we prove a bound on N-extendibility of channels combining white noise and loss. In the thermal-noise channel scenario, we point out that it is possible to apply directly existing bounds for channel extendibility [86] and joint-measurability [87], but we also provide in addition explicit strategies for achieving these bounds.

We will proceed as follows. In Section 2.2, we introduce the CMU, we make it clear that the concept of joint-measurability is appropriate for defining at what point a CMU loses its utility in a quantum protocol with untrusted devices, and we relate it to other properties of the CMU, such as channel extendibility. In Section 2.3, we derive upper-bounds in the no-click scenario and in Section 2.4 in the thermal-noise channel scenario.

2.2 Joint-measurability of a CMU

A CMU consists of a quantum channel C followed by a measurement device M, as in Fig. 2.1. The channel is described by CPTP map Φ, while the measurement is given by a set of POVMs $M_y = \{M_{b|y}\}_b$, where $M_{b|y} \geq 0$ and $\sum_b M_{b|y} = \mathbb{1}$ for all y, b. Note that we use the notation M_y to denote the POVMs and $M_{b|y}$ their elements.

From a black-box perspective, the CMU is characterized by the probabilities $P(b|y,\rho)$ of obtaining a measurement outcome b, given the measurement input y and the incoming quantum

state ρ. Formally we have that

$$P(b|y,\rho) = \mathrm{Tr}\left[\Phi(\rho)\, M_{b|y}\right]. \tag{2.1}$$

The CMU can also be described through a set of effective POVMs with elements $M^{\Phi}_{b|y} = \Phi^*\left(M_{b|y}\right)$, where Φ^* is the dual channel of Φ. These POVMs describe the combined effect of the channel and of the measurements. The probabilities $P(b|y,\rho)$ can then also be written as

$$P(b|y,\rho) = \mathrm{Tr}\left[\rho\, M^{\Phi}_{b|y}\right]. \tag{2.2}$$

We would like to determine the admissible level of loss in the CMU, regardless of its usage in a larger experiment, before it loses its ability to exhibit 'quantumness'. Since most experiments and protocols with untrusted devices require the distribution of entanglement and/or incompatible measurements, we could define this threshold using one of the two following natural conditions:

(1) The CMU can be replaced by an equivalent CMU for which the channel Φ is entanglement-breaking.

(2) The CMU can be replaced by an equivalent CMU for which the N measurements M_y are compatible, i.e., jointly measurable.

By 'equivalent' CMU, we mean one that may differ in implementation from the original CMU, but yields the same output probabilities $P(b|y,\rho)$, since from an operational and blackbox perspective, only these output probabilities are relevant.

As an example of the above definition, consider a standard bipartite nonlocality test. If either Alice's or Bob's CMU satisfies condition (1) or (2), it becomes clear that the entire experiment admits a local hidden-variable theory and no violation of a Bell inequality is possible. Similar conclusions can be drawn for steering experiments. Moreover, even protocols that do not directly rely on entanglement may require condition (1) to hold. For example, in a prepare-and-measure quantum key distribution (QKD) scheme where the channel connecting Alice to Bob is entanglement-breaking, no key rate can be extracted [88].

More generally, the purpose of a CMU is to distribute quantum information from the entry of the channel to a remote measurement device where one of several measurements should be performed. Then an alternative, and seemingly more stringent, way to define the threshold at which a CMU becomes ineffective is the following:

(3) The CMU can be replaced by an equivalent CMU for which the channel Φ is a quantum-classical channel and the measurements M_y represent classical measurements on the output of this channel.

In a quantum experiment involving a CMU that satisfies condition (3), the quantum part of the experiment can be truncated just before the CMU, since all information processing that happens in the CMU is purely classical, see Fig. 2.3. This criterion is used, for instance, in [89] to determine when a routed Bell experiment can establish only short-range quantum correlations. Furthermore, in any cryptographic protocol where the CMU channel is a public channel and satisfies condition (3), an eavesdropper could have a perfect copy of the information sent in that channel since it is purely classical.

Although conditions (1), (2), (3) may appear distinct, they are actually all equivalent in an untrusted scenario, and equivalent to the following condition:

(4) The set of effective POVMs $M^{\Phi}_{b|y}$ of the CMU are jointly measurable.

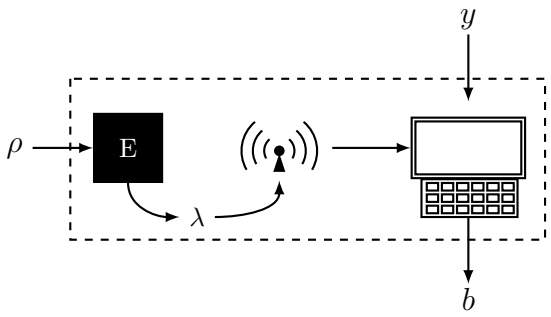

Figure 2.3: A CMU that satisfies conditions (3) and (4) can be implemented through a (single) quantum measurement E_λ at the entrance of the channel, whose classical output λ is then transmitted and processed to generate the output b depending on y.

The set of effective POVMs $M^\Phi_{b|y}$ is said to be jointly measurable [75] if there exists a *parent POVM* E with elements E_λ, and probability distributions $p(b|y,\lambda)$ such that

$$M^\Phi_{b|y} = \int_\lambda d\lambda\, p(b|y,\lambda) E_\lambda\,. \tag{2.3}$$

Operationally, and as illustrated in Fig. 2.3, this means that the CMU can in principle be implemented via the following strategy. First, at the entrance of the channel, upon receiving the quantum state ρ, one measures the parent POVM E; note that this is a single (fixed) quantum measurement, hence independent of the input y. The outcome λ of the parent POVM is then transmitted through a classical channel to the remote measurement device, where a simple classical device generates the output b using the probability distribution $p(b|y,\lambda)$, which depends on y and on the classical outcome λ.

From the above operational interpretation of (4), it is immediate that conditions (3) and (4) are equivalent. It is also clear that conditions (2) and (4) are equivalent, since if the measurements M_y are jointly measurable, the same holds for the effective measurements M^Φ_y, and vice versa. Finally, conditions (1) and (4) are also equivalent, as shown in [90]. Indeed, Φ is entanglement-breaking if $\Phi(\rho) = \sum_\lambda \sigma_\lambda \operatorname{Tr}[E_\lambda \rho]$. But then $P(b|y,\rho) = \sum_\lambda \operatorname{Tr}[E_\lambda\rho] \operatorname{Tr}\left[\sigma_\lambda M_{b|y}\right] = \operatorname{Tr}\left[\rho M^\Phi_{b|y}\right]$, for $M^\Phi_{b|y} = \sum_\lambda \operatorname{Tr}\left[\sigma_\lambda M_{b|y}\right] E_\lambda$ which is of the form (2.3) if we define $P(b|y,\lambda) = \operatorname{Tr}\left[\sigma_\lambda M_{b|y}\right]$. This argument can be extended to infinite-dimensional quantum systems; in this case, an entanglement breaking channel is of the form $\Phi(\rho) = \int \mathrm{d}\lambda\, \sigma_\lambda \operatorname{Tr}\rho E_\lambda$ where the POVM is in general not atomic [91].

Note that condition (4) is not only sufficient for guaranteeing that a CMU is "useless", but also a necessary one; indeed, any set of measurements that is incompatible (in the sense of not being jointly measurable) can be used to demonstrate the effect of quantum steering [78, 79].

In the following, we will thus take condition (4) as our definition of when a CMU does not exhibit quantumness. We say that a CMU is *jointly-measurable* (*JM*) whenever condition (4) holds.

Finally, note that there are other possible conditions than (1) on the channel of a CMU that are equivalent to (4). For instance, a channel Φ is said to be *incompatibility-breaking* if the set of measurements $\{\Phi^*(M_y)\}$ are compatible for any M_y [80]. It is said to be *N-extendable* if there exists another channel $\Phi_{1\to N}$, called the parent channel, which outputs N quantum system $S_1 \ldots, S_N$ such that

$$\operatorname{Tr}_{S_\mu|\mu\neq k}\, \Phi_{1\to N}(\rho) = \Phi(\rho) \tag{2.4}$$

for any output system S_k and any input state ρ. These notions give rise to the following additional conditions.

(5) The CMU can be replaced by an equivalent CMU for which the channel Φ is incompatibility-breaking.

(6) The CMU can be replaced by an equivalent CMU for which the channel Φ is N-extendable.

The entanglement-breaking, incompatibility-breaking, and N-extendibility of a channel are in principle distinct properties. While every entanglement-breaking channel is N-extendable, and any N-extendable channel is incompatibility breaking, there exist incompatible channels that are not N-extendable, and N-extendable channels that are not entanglement-breaking. However, in the context of an untrusted CMU, the conditions (1), (5), and (6) are all equivalent to (4). This is because they are not statements about the properties of the channel in the actual realization of the CMU, but they are statements about an operationally equivalent CMU that reproduces the *combined* effect of the *channels* and of the *measurements* of the original CMU.

Specifically, since a classical channel is both incompatibility-breaking and N-extendable , we clearly have (3)⇒(5) and (3)⇒(6), and thus (4)⇒(5) and (4)⇒(6). It is also clear that (5)⇒(4). On the other hand, if (6) holds, i.e., the channel Φ is N-extendable, the effective POVMs $\Phi^*(M_{b|y})$ are jointly-measurable because they can be obtained by applying the channel $\Phi_{1\to N}$ on the input state, and then performing each of the N different measurement M_y for $y = 1, \dots N$ on each of the N output systems, as illustrated in Fig. 2.4. Thus (6)⇒(4). We will make extensive use of this equivalence between the *JM* property of the CMU and N-extendibility in the remainder of the paper.

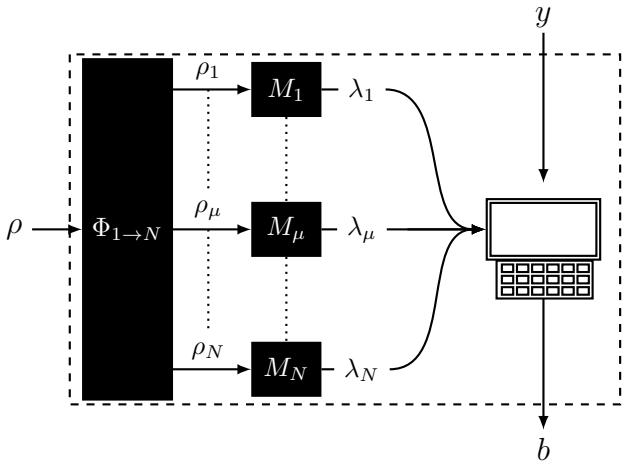

Figure 2.4: A CMU satisfying condition (6) can be implemented by splitting the incoming system in N systems, measuring each of the output systems with one of the N different measurements, and sending their classical outcomes to Bob.

2.3 No-click CMUs

We first analyse a large class of CMUs taking as input quantum states ρ in a finite d-dimensional Hilbert space $\mathcal{H}$ and for which the N measurements $y = 1, \dots, N$ yield a finite, discrete set of $L+1$ outcomes $b = 1, \dots, L, \varnothing$, where the last outcome $\varnothing$ is a 'no-click' outcome happening when

the measurement device fails to register an actual outcome, e.g. because the photon is lost. In addition to the loss, the measurements can also be affected by another type of noise.

In general, such a CMU is described by effective POVMs of the form

$$M_{b|y}^{\eta,\Phi} = \begin{cases} \eta\, \Phi^*\big(M_{b|y}\big) & \text{if } b \neq \varnothing \\ (1-\eta)\mathbb{1} & \text{if } b = \varnothing\,, \end{cases} \tag{2.5}$$

where $M_{b|y}$ describe the ideal POVMs, η is the *detection efficiency*, and the CPTP map $\Phi : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ models the additional noise channel (Φ^* is its adjoint). Similarly to the case described in Section 1.2.4, this CMU can be understood as a concatenation of the channel Φ and the *loss channel*

$$\Lambda_\eta(\rho) = \eta\rho + (1-\eta)\,|\varnothing\rangle\langle\varnothing| \tag{2.6}$$

acting on the state ρ before the measurements $\tilde{M}_y$ with elements

$$\tilde{M}_{b|y} = \begin{cases} M_{b|y} & \text{if } b \neq \varnothing \\ |\varnothing\rangle\langle\varnothing| & \text{if } b = \varnothing \end{cases} \tag{2.7}$$

are performed. Here, $|\varnothing\rangle$ is an additional state orthogonal to $\mathcal{H}$ describing the loss of a photon, and (2.7) extends the ideal POVMs so as to produce the no-click outcome whenever the state $|\varnothing\rangle$ is received. This allows one to write $M_{b|y}^{\eta,\Phi} = (\Lambda_\eta \circ \Phi)^*(\tilde{M}_{b|y})$.

For most of our applications, we will consider the noise channel Φ to be the *white-noise* channel

$$W_v(\rho) = v\,\rho + (1-v)\frac{\mathbb{1}}{d}\,, \tag{2.8}$$

with *visibility* v. In this case, the effective POVMs of the CMU read

$$M_{b|y}^{\eta,v} = \begin{cases} \eta\, v\, M_{b|y} + \eta\,(1-v) q_{b|y}\,\mathbb{1} & \text{if } b \neq \varnothing \\ (1-\eta)\mathbb{1} & \text{if } b = \varnothing\,, \end{cases} \tag{2.9}$$

where $q_{b|y} = \mathrm{Tr}\left(M_{b|y}\right)/d$.

Given the ideal POVMs $M_{b|y}$, the noise channel Φ and the efficiency η, the question of whether the effective POVMs (2.5) are *JM* can be formulated as a semidefinite program (SDP) [75, 92]. Furthermore, the threshold efficiency η_* below which they are *JM* can readily be determined by solving a single SDP since the above effective POVMs depend linearly on η.

In the following, we are interested in providing bounds on the critical values of η that hold for a large class of measurements M_y and noise models Φ.

2.3.1 Bounds on η for arbitrary sets of measurements implied by channel extendibility

In this section, we first collect some results on the N-extendibility of some families of channels. As discussed in the introduction, we then use them to imply that any CMU featuring such a channel and composed of N POVMs is *JM*. We start by introducing two general parametric families of channels.

A channel R_q with $q \in [0,1]$ is called *replacement* if it is of the form

$$R_q(\rho) = q\,\rho + (1-q)\sigma, \tag{2.10}$$

where σ is any fixed state (possibly orthogonal to $\mathcal{H}$). The loss channel Λ_η in Eq. (2.6) and the white-noise channel W_v in Eq. (2.8) are instances of replacement channels.

A channel Γ_v with $v \in [0,1]$ is called *convex noise* if it is of the form

$$\Gamma_v(\rho) = v\rho + (1-v)\Gamma_0(\rho), \tag{2.11}$$

for some Γ_0. An equivalent definition is to impose the properties $\Gamma_1 = \mathrm{id}$ and $p\,\Gamma_{v_1} + (1-p)\,\Gamma_{v_2} = \Gamma_{\bar{v}}$, where $\bar{v} = p\,v_1 + (1-p)v_2$. It is easy to see that all the replacement channels are convex noise channels. However, there are other prominent examples of convex noise channels, e.g. the dephasing channel

$$\Gamma_v(\rho) = v\rho + (1-v)\mathrm{diag}\,(\rho)\,, \tag{2.12}$$

where $\mathrm{diag}(\rho) = \sum_i \rho_{ii}\,|i\rangle\langle i|$ is the diagonal part of ρ in the computational basis.

With these definitions, we now formulate a few results on N-extendibility. The first one holds for all replacement channels.

Lemma 1 ([80])**.** *A replacement channel $R_q(\rho)$ as in Eq. (2.10) is N-extendable if $q \leq \frac{1}{N}$.*

Proof. For $q \leq \frac{1}{N}$ the channel

$$\begin{aligned} R_{1\to N}(\rho) = {} & q\,\rho \otimes \sigma^{\otimes(N-1)} + q\,\sigma \otimes \rho \otimes \sigma^{\otimes(N-2)} \\ & + \ldots + q\,\sigma^{\otimes(N-1)} \otimes \rho + (1-qN)\sigma^{\otimes N}\,. \end{aligned} \tag{2.13}$$

is well defined and is the parent of R_q. □

This directly implies that the loss channel Λ_η in Eq. (2.6) is N-extendable for $\eta \leq \frac{1}{N}$. A similar implication holds for the white-noise channel W_v in Eq. (2.8), however for this channel the following result gives a tighter bound.

Lemma 2 ([93])**.** *The white-noise channel W_v in Eq. (2.8) is N-extendable if*

$$v \leq \frac{N+d}{N(d+1)}. \tag{2.14}$$

Proof. This follows from results on universal quantum cloning. An optimal one-to-N universal cloning machine is a channel $\Phi_{1\to N}^{\mathrm{UCM}}$ from one to N qudits whose output resembles N copies of the input state as closely as possible for all possible input states [93, 94]. Furthermore, for each of the output systems, the marginal channels of Eq. (2.4) induced by the optimal universal cloning machine are precisely given by a white-noise channel W_v (implied by the symmetry of the task) with visibility saturating the inequality (2.14) [93]. More noise can always be added if v is lower (see Lemma 3 below). □

Now, following our original motivation we proceed to study the N-extendibility of concatenated channels. The next result is a very simple but useful observation.

Lemma 3. *Consider an N-extendable channel Φ and another arbitrary channel Γ. The concatenated channels*

$$\Gamma \circ \Phi \quad \textit{and} \quad \Phi \circ \Gamma \tag{2.15}$$

are N-extendable.

Proof. The parent channels are given by $\Gamma^{\otimes N} \circ \Phi_{1\to N}$ and $\Phi_{1\to N} \circ \Gamma$ respectively. □

Now consider a concatenation of channels with known properties.

Lemma 4. *Consider a convex noise channel Γ_v (Eq. 2.11) with $v \geq v_*$ and a replacement channel R_q (Eq. 2.10) with $q \geq q_*$, where above the threshold values the channels Γ_v and R_q are N-extendable. The concatenated channel*

$$\Phi_{q,v} = R_q \circ \Gamma_v \tag{2.16}$$

is N-extendable for

$$q \leq \frac{1 - v_*}{(1 - v) + (v - v_*)/q_*} \,. \tag{2.17}$$

Proof. First, let us write the concatenated channel explicitly as

$$\begin{aligned} \Phi_{q,v}(\rho) &= R_q\big(v\rho + (1-v)\Gamma_0(\rho)\big) \\ &= q\, v\, \rho + q(1-v)\Gamma_0(\rho) + (1-q)\sigma. \end{aligned} \tag{2.18}$$

By assumption, there exist parent channels $\Gamma_{1\to N}$ and $R_{1\to N}$ that reproduce the marginal channels Γ_{v_*} and R_{q_*} when tracing out $N-1$ systems. Let us now consider a mixture thereof $p\,\Gamma_{1\to N} + (1-p)\,R_{1\to N}$ for some parameter $p \in [0,1]$. It defines a parent channel which establishes the N-extendibility of $\mathcal{E}_p(\rho) = \mathrm{Tr}_{N-1}(p\,\Gamma_{1\to N} + (1-p)R_{1\to N})(\rho)$ given by

$$\begin{aligned} \mathcal{E}_p(\rho) &= p\,\Gamma_{v_*}(\rho) + (1-p)R_{q_*}(\rho) \\ &= (p\,v_* + (1-p)q_*)\rho + p(1-v_*)\Gamma_0(\rho) \\ &+ (1-p)(1-q_*)\sigma. \end{aligned}$$

Comparing with Eq. (2.18) we conclude that $\Phi_{q,v}$ is identical to $\mathcal{E}_p$ if

$$\begin{aligned} 1 - q &= (1-p)(1-q_*) \\ q(1-v) &= p(1-v_*)\,. \end{aligned} \tag{2.19}$$

Solving the first equation for p gives $1 - p = \frac{1-q}{1-q_*}$. Plugging this in the second equation implies

$$q = \frac{1 - v_*}{(1 - v) + (v - v_*)/q_*} \,. \tag{2.20}$$

The channel $\Phi_{q,v}$ is thus N-extendable when this equality is satisfied. To see that any channel with a lower transmission $q' \leq q$ is also N-extendable note that it can be decomposed as $\Phi_{q',v} = R_{q'/q} \circ \Phi_{q,v}$ and apply Lemma 3. □

Let us now come back to our channels of interest and apply the above Lemmas to obtain sufficient conditions for the CMUs with effective POVMs (2.5) and (2.9) to become *JM*. First, consider the concatenation of any channel Φ with the loss channel Λ_η and apply Lemmas 1 and 3 to obtain the following bound.

Upper bound 1 ([71])**.** *The concatenation of the loss channel with any noise channel Φ*

$$\Lambda_\eta \circ \Phi(\rho) = \eta\,\Phi(\rho) + (1-\eta)\,|\varnothing\rangle\langle\varnothing| \tag{2.21}$$

is N-extendable, and thus the corresponding CMU with effective POVMs (2.5) is JM, if

$$\eta \leq \frac{1}{N} \,. \tag{2.22}$$

This is essentially a reformulation from the channel perspective of a result already obtained in [64] in the context of Bell experiments, in [71] for general one-sided-device-independent protocols, and in [80] where channels of the form $R_\eta \circ \Gamma(\rho) = \eta\,\Gamma(\rho) + (1-\eta)\sigma$ were shown to be N-incompatibility breaking for $\eta \leq \frac{1}{N}$.

Next, we also assume that the noise channel $\Phi = \Phi_v$ is convex and apply Lemmas 1 and 4 to get.

Upper bound 2. *The channel*

$$\Lambda_\eta \circ \Phi_v(\rho) = \eta\,\Phi_v(\rho) + (1-\eta)\,|\varnothing\rangle\langle\varnothing|\,, \tag{2.23}$$

corresponding to the concatenation of the loss channel and any convex noise channel Φ_v with $v \geq v_*$ where Φ_{v_*} is N-extendable, and thus the corresponding CMU is JM, if

$$\eta \leq \frac{1-v_*}{(1-v)+N(v-v_*)}\,. \tag{2.24}$$

Remarkably, this bound is useful even if we only know that the channel Φ_v becomes N-extendable at $v = 0$. Direct application of (2.24) shows that in this case $\Lambda_\eta \circ \Phi_v$ is N-extendable for

$$\eta \leq \frac{1}{(1-v)+Nv}. \tag{2.25}$$

This bound applies for example when $\Phi_v = \Gamma_v$ is the dephasing channel in Eq. (2.12).

Similarly, if Φ_v is a replacement channel itself (or becomes N-extendable for $v \leq \frac{1}{N}$ for another reason) the bound (2.24) implies that $\Lambda_\eta \circ \Phi_v$ is N-extendable for

$$\eta \leq \frac{1}{vN}. \tag{2.26}$$

Finally, we give special treatment to the case where $\Phi_v = W_v$ is the white-noise channel, as it will be important for applications. Combining the Upper bound 2 with Lemma 2 (giving $v_* = \frac{N+d}{N(d+1)}$) we get our last bound in this section.

Upper bound 3. *The concatenation of loss and white-noise*

$$\Lambda_\eta \circ W_v(\rho) = \eta\,v\,\rho + \eta(1-v)\frac{\mathbb{1}}{d} + (1-\eta)\,|\varnothing\rangle\langle\varnothing|\,, \tag{2.27}$$

is N-extendable, and thus the corresponding CMU with effective POVMs (2.9) *is JM, if*

$$\eta \leq \frac{d}{N\Big(v(d+1)-1\Big)}\,. \tag{2.28}$$

Note that we can make this bound dimension-independent by taking the worst-case value of the right-hand side of (2.28), $\eta \leq \min_{d\geq 1} d/\big(N(v(d+1)-1)\big) = \frac{1}{Nv}$, which coincides with the more general bound (2.26).

The fundamental trade-off between the visibility v and the detection efficiency η provided by the bound (2.28) is illustrated in Fig. 2.5 in the case of qubits.

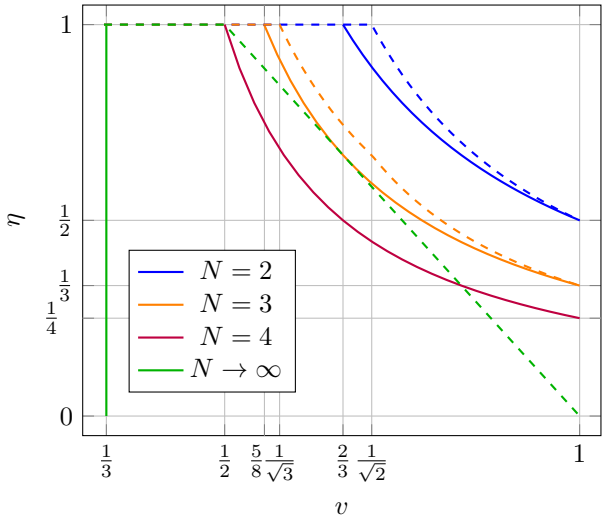

Figure 2.5: The full curves depict the lines below and on the left of which the channel $\Phi_{\eta,v}$ of the no-click white-noise qubit CMU is N-extendable according to Upper bound 3 for $N = 2, 3, 4$ and according to Lemma 2 for $N = \infty$. The dashed curves correspond to the additional assumption that the ideal measurements are binary and are determined according to Upper bound 6 (see Section 2.3.2) for $N = 2, 3, 4$ and the bound (2.52) (see Section 2.3.2) for $N = \infty$. Note that for the case $N = 4$ the two lines coincide.

2.3.2 Bounds on η for binary qubit measurements

The bounds obtained so far on η depend on generic parameters such as the number of measurements N, the level of noise v, or the dimension d of the underlying Hilbert space in the case of white noise. And as they follow from the N-extendibility of the CMU channel, they hold for any possible measurements M_y implemented by the CMU. We now take additionally into account information about the specific measurements M_y that define the CMU, focusing on the particular case of qubits.

First note the following variation of Upper bound 2.

Upper bound 2'. *Consider a no-click CMU defined in terms of a convex noise channel Φ_v with $v \geq v_*$ and ideal measurements M_y such that the effective POVMs $\Phi^*_{v_*}(M_y)$ are jointly measurable. Then the effective POVMs $\Phi^*_{\eta,v}(M_y)$ defining the no-click CMU in Eq.* (2.5) *are JM for*

$$\eta \leq \frac{1 - v_*}{(1 - v) + N(v - v_*)} \,. \tag{2.29}$$

This can be proven identically to the Lemma 4 by comparing the POVMs

$$\begin{aligned} \Phi^*_{\eta,v}(M_{b|y}) = \eta v M_{b|y} + \eta(1 - v)\Phi^*_{1,0}(M_{b|y}) \\ + (1 - \eta)\Phi^*_{0,1}(M_{b|y}) \end{aligned} \tag{2.30}$$

with the family of POVMs

$$p\,\Phi^*_{1,v_*}(M_{b|y}) + (1 - p)\Phi^*_{\frac{1}{N},1}(M_{b|y}), \tag{2.31}$$

which are *JM* by construction.

From now on we set $d = 2$ for the rest of the section. A binary qubit measurements is composed of two POVM elements $\{E_+, E_-\}$ of the form

$$E_\pm = \frac{1}{2}\big((1 \pm \gamma)\mathbb{1} \pm \boldsymbol{m} \cdot \boldsymbol{\sigma}\big), \tag{2.32}$$

with $1 - |\gamma| \leq \|\boldsymbol{m}\|$ guaranteed by positivity. If $\gamma = 0$ it is said to be unbiased, otherwise it is biased. A biased measurement can be decomposed as a mixture

$$\{E_+, E_-\} = \|\boldsymbol{m}\|\{M_+, M_-\} + (1 - \|\boldsymbol{m}\|)\{R_+, R_-\}, \tag{2.33}$$

of a PVM (unbiased)

$$M_\pm = \frac{1}{2}(\mathbb{1} \pm \hat{\boldsymbol{m}} \cdot \boldsymbol{\sigma}) \tag{2.34}$$

with a unit vector $\hat{\boldsymbol{m}} = \frac{\boldsymbol{m}}{\|\boldsymbol{m}\|}$, and a "dummy measurement" $R_\pm = q_\pm \mathbb{1}$ with $q_\pm = \frac{1 \pm \gamma - \|m\|}{2(1-\|m\|)}$ whose output is independent of the measured state. This observation leads to the following result.

Lemma 5. *Consider a channel Φ and a set of binary measurements $E_{\pm|y} = \frac{1}{2}\big((1 \pm \gamma_y)\mathbb{1} \pm \boldsymbol{m}_y \cdot \boldsymbol{\sigma}\big)$. The measurements set $\Phi^*(E_{\pm|y})$ is JM if the measurements $\Phi^*(M_{\pm|y})$ with $M_{\pm|y} = \frac{1}{2}\big(\mathbb{1} \pm \hat{\boldsymbol{m}}_y \cdot \boldsymbol{\sigma}\big)$ and $\hat{\boldsymbol{m}}_y = \frac{\boldsymbol{m}_y}{\|\boldsymbol{m}_y\|}$ are JM.*

Proof. By Eq. (2.33) the measurements $\Phi^*(E_{\pm|y})$ can be decomposed as

$$\Phi^*(E_{\pm|y}) = \|\boldsymbol{m}_y\|\Phi^*(M_{\pm|y}) + (1 - \|\boldsymbol{m}_y\|)R_{\pm|y}, \tag{2.35}$$

since $R_{\pm|y} = \Phi^*(R_{\pm|y}) = q_{\pm|y}\mathbb{1}$ (the adjoint of a CPTP map is unital).

Since the measurements $\Phi^*(M_{\pm|y})$ are *JM* by assumption there exists a parent POVM E_λ simulating them. Then the measurements $\Phi^*(E_{\pm|y})$ are simulated by the following strategy. Perform the parent POVM E_λ. Upon receiving the setting y simulate $\Phi^*(M_{\pm|y})$ with probability $\|\boldsymbol{m}_y\|$, or with probability $1 - \|\boldsymbol{m}_y\|$ sample a random output from $q_{\pm|y}$. □

This result implies that when discussing the *JM* of sets of *ideal* binary measurements subject to noise and loss, biased measurements can be ignored. That is, if we show that the set $\Phi^*_{\eta,v}(M_{\pm|y})$ with

$$M_{\pm|y} = \frac{1}{2}(\mathbb{1} \pm \hat{\boldsymbol{m}}_y \cdot \boldsymbol{\sigma}), \tag{2.36}$$

and $\|\hat{\boldsymbol{m}}_y\| = 1$ is *JM* it automatically implies that any set of (biased) measurements $\Phi^*_{\eta,v}(E_{\pm|y})$ with $\boldsymbol{m}_y$ parallel to $\hat{\boldsymbol{m}}_y$ is also *JM*. We will thus restrict our consideration to N measurements of the form (2.36) in the rest of the section. Furthermore, we now set the channel Φ_v to be the white-noise channel W_v mapping PVMs to

$$W_v^*(M_{\pm|y}) = \frac{1}{2}(\mathbb{1} \pm v\,\hat{\boldsymbol{m}}_y \cdot \boldsymbol{\sigma}), \tag{2.37}$$

which are also unbiased.

The question of the joint measurability of unbiased qubit measurements has been studied extensively, see e.g. section III.A in the review [75].

In particular, from [95] it follows that any set of $N = 2$ measurements of the form (2.37) is *JM* if and only if

$$v \leq v_*(\hat{\boldsymbol{m}}_1, \hat{\boldsymbol{m}}_2) := \frac{2}{\|\hat{\boldsymbol{m}}_1 + \hat{\boldsymbol{m}}_2\| + \|\hat{\boldsymbol{m}}_1 - \hat{\boldsymbol{m}}_2\|}. \tag{2.38}$$

For $N = 3$, the measurements are *JM* if and only if [96, 97]

$$v \leq v_*(\hat{\boldsymbol{m}}_1, \hat{\boldsymbol{m}}_2, \hat{\boldsymbol{m}}_3) := \frac{4}{\sum_{k=0}^{3} \|\boldsymbol{t}_k - \boldsymbol{t}_{FT}\|}, \tag{2.39}$$

where $\boldsymbol{t}_0 = \hat{\boldsymbol{m}}_1 + \hat{\boldsymbol{m}}_2 + \hat{\boldsymbol{m}}_3$, $\boldsymbol{t}_k = 2\hat{\boldsymbol{m}}_k - \boldsymbol{t}_0$ for $k \geq 1$, and $\boldsymbol{t}_{FT}$ is the so-called Fermat-Toricelli vector of the set $\{\hat{\boldsymbol{m}}_1, \hat{\boldsymbol{m}}_2, \hat{\boldsymbol{m}}_3\}$, i.e. $\boldsymbol{t}_{FT} = \text{argmin}_{\boldsymbol{t}} \sum_{k=1}^{3} \|\hat{\boldsymbol{m}}_k - \boldsymbol{t}\|$

For any set of N unbiased qubit measurements, we now show the following result.

Lemma 6. *A set of N unbiased qubit measurements $M_{\pm|y} = \frac{1}{2}(\mathbb{1} \pm \boldsymbol{m}_y \cdot \boldsymbol{\sigma})$ specified by the vectors $\boldsymbol{m}_1, \ldots, \boldsymbol{m}_N$ with $\|\boldsymbol{m}_y\| \leq 1$ is JM if*

$$\sum_{\boldsymbol{a}} \| \sum_{k=1}^{N} (-1)^{a_k} \boldsymbol{m}_k \| \leq 2^N \tag{2.40}$$

where $\boldsymbol{a} = (a_1, \ldots, a_N) \in \{0,1\}^N$ runs through the 2^N bitstrings. (See also the relaxed condition in Eq. (2.49)).

Proof. Define the following set of 2^N vectors

$$\boldsymbol{w}_{\boldsymbol{a}} = (-1)^{a_1} \boldsymbol{m}_1 + (-1)^{a_2} \boldsymbol{m}_2 + \cdots + (-1)^{a_N} \boldsymbol{m}_N \tag{2.41}$$

labeled by the bitstrings $\boldsymbol{a} = (a_1, \ldots, a_N) \in \{0,1\}^N$ and denoted their normalized versions as $\hat{\boldsymbol{w}}_{\boldsymbol{a}} = \frac{\boldsymbol{w}_{\boldsymbol{a}}}{\|\boldsymbol{w}_{\boldsymbol{a}}\|}$, which define 2^N projective measurements. Using the probability distribution $p(\boldsymbol{a}) = \frac{\|\boldsymbol{w}_{\boldsymbol{a}}\|}{\sum_{\boldsymbol{a}} \|\boldsymbol{w}_{\boldsymbol{a}}\|}$ define the parent POVM composed of 2^{N+1} elements

$$E_{\pm,\boldsymbol{a}} = \frac{p(\boldsymbol{a})}{2}(\mathbb{1} \pm \hat{\boldsymbol{w}}_{\boldsymbol{a}} \cdot \boldsymbol{\sigma}) = \frac{1}{2}\left(p(\boldsymbol{a})\mathbb{1} \pm \frac{\boldsymbol{w}_{\boldsymbol{a}} \cdot \boldsymbol{\sigma}}{\sum_{\boldsymbol{a}} \|\boldsymbol{w}_{\boldsymbol{a}}\|} \right). \tag{2.42}$$

With the post-processing

$$p(b|y, \pm, \boldsymbol{a}) = \begin{cases} 1 & \pm 1 = \text{sign}(b\,(-1)^{a_y}) \\ 0 & \text{otherwise} \end{cases} \tag{2.43}$$

for $b = \pm 1$, the parent POVMs can simulate all measurements of the form

$$\tilde{M}_{b|y} = \sum_{\pm,\boldsymbol{a}} E_{\pm,\boldsymbol{a}}\, p(b|y, \pm, \boldsymbol{a}) = \sum_{\boldsymbol{a}} E_{\text{sign}(b(-1)^{a_y}),\boldsymbol{a}} \tag{2.44}$$

$$= \frac{1}{2}\left(\mathbb{1} + b\frac{2^N}{\sum_{\boldsymbol{a}} \|\boldsymbol{w}_{\boldsymbol{a}}\|} \boldsymbol{m}_k \cdot \boldsymbol{\sigma} \right). \tag{2.45}$$

These measurements are identical to $M_{b|y}$ if $\frac{2^N}{\sum_{\boldsymbol{a}} \|\boldsymbol{w}_{\boldsymbol{a}}\|} = 1$, in which case the proof is done. If this ratio is even larger $\frac{2^N}{\sum_{\boldsymbol{a}} \|\boldsymbol{w}_{\boldsymbol{a}}\|} \geq 1$, one can mix some randomness into $\tilde{M}_{b|y}$ to complete the proof. □

Applying this Lemma to the measurements $W_v^*(M_{\pm|y})$ in Eq. (2.36) with $\boldsymbol{m}_y = v\,\hat{\boldsymbol{m}}_y$ gives the follwing condition for their joint measurability

$$v \leq v_\star(\hat{\boldsymbol{m}}_1, \ldots \hat{\boldsymbol{m}}_N) := \frac{2^N}{\sum_{\boldsymbol{a}} \| \sum_{k=1}^{N} (-1)^{a_k} \hat{\boldsymbol{m}}_k \|} \tag{2.46}$$

Note that Eq. (2.38) is a particular case of Eq. (2.46) for $N = 2$, while it is unclear to us if this is also the case for Eq. (2.39). An attentive reader also noticed that for general N we used $v_\star$ instead of v_* to indicate that the bound is potentially not tight.

The threshold visibilities for two, three and N measurements in Eqs. (2.38,2.39,2.46) can be used together with the Upper bound 2' to obtain sufficient conditions on the joint measurability of the corresponding CMUs $M^{\eta,v}_{b|y}$. The resulting bounds would depend on the specific choice of binary the qubit measurements described by the vectors $\hat{\boldsymbol{m}}_1, \ldots, \hat{\boldsymbol{m}}_N$.

Nevertheless, it is also possible to derive a simple bound for any set of N binary qubit measurements by noting that the rhs of Eq. (2.46) satisfies

$$v_*(\hat{\boldsymbol{m}}_1, \ldots, \hat{\boldsymbol{m}}_N) \geq \frac{1}{\sqrt{N}}. \tag{2.47}$$

To see this first use the concavity of the square root to obtain $\sum_{\boldsymbol{a}} \frac{1}{2^N} \|\boldsymbol{w}_{\boldsymbol{a}}\| = \sum_{\boldsymbol{a}} \frac{1}{2^N} \sqrt{\|\boldsymbol{w}_{\boldsymbol{a}}\|^2} \leq \sqrt{\frac{1}{2^N} \sum_{\boldsymbol{a}} \|\boldsymbol{w}_{\boldsymbol{a}}\|^2}$. Then in the last term simply compute

$$\sum_{\boldsymbol{a}} \| \sum_{k=1}^{N} (-1)^{a_k} \hat{\boldsymbol{m}}_k \|^2 = \sum_{\boldsymbol{a}} \sum_{kj=1}^{N} (-1)^{a_k + a_j} \hat{\boldsymbol{m}}_k \cdot \hat{\boldsymbol{m}}_j = \sum_{\boldsymbol{a}} \sum_{k=1}^{N} \hat{\boldsymbol{m}}_k \cdot \hat{\boldsymbol{m}}_k = 2^N N, \tag{2.48}$$

which implies $\sqrt{\sum_{\boldsymbol{a}} \frac{1}{2^N} \| \sum_{k=1}^{N} (-1)^{a_y} \hat{\boldsymbol{m}}_k \|^2} = \sqrt{\frac{1}{2^N} 2^N N} = \sqrt{N}$ and the inequality (2.47). This bound is, however, only saturable for $N = 2$ and 3 by the measurements corresponding to MUBs. Remarkably the same arguments imply $\frac{1}{2^N} \sum_{\boldsymbol{a}} \| \sum_{k=1}^{N} (-1)^{a_k} \boldsymbol{m}_k \| \leq \sqrt{\sum_k \|\boldsymbol{m}_k\|^2}$, hence

$$\sum_{k=1}^{N} \|\boldsymbol{m}_k\|^2 \leq 1 \tag{2.49}$$

can be used as a relaxation of the condition (2.40) in Lemma 6.

To conclude this section we combine the Upper bound 2', the Lemmas 5 with 6, and the bound (2.47) to obtain the following simple sufficient condition for joint measurability of our CMUs.

Upper bound 6. *The no-click CMU implementing N binary qubit measurements (2.9,2.36) with white-noise visibility v and detection efficiency η is JM if*

$$\eta \leq \frac{1}{\sqrt{N}\left((\sqrt{N}+1)v - 1\right)}. \tag{2.50}$$

It is easy to see that for $N \geq 4$ this condition does not bring any improvement over the general upper bound (2.28), which reduces in the case $d = 2$ to

$$\eta \leq \frac{2}{N\Big(3v - 1\Big)}. \tag{2.51}$$

In contrast, for $N = 2$ and 3, it improves on this bound, as illustrated in Fig. 2.5. We can also demonstrate numerically that in these cases the bound (2.50) is tight. Indeed, selecting σ_z and σ_x as the two measurement directions in the case $N = 2$, and additionally σ_y in the case $N = 3$, we can determine, for any fixed v, the maximal value of η such that the resulting effective POVMs are *JM* by solving a single SDP, as noted below Eq. (2.9). We observed that for all the tested values v, the maximal transmission η found by the SDP matches with the upper bound of Eq. (2.50) up to numerical precision.

2.3.3 Bound on η independent of N

Finally, let us mention the asymptotic case where the number of measurements N is unbounded. In [85] a necessary and sufficient condition for the joint measurability of all d-dimensional projective measurements subject to white noise and loss was derived. In the case of qubits, this condition takes a simple form

$$\eta \leq 2(1-v)\,, \tag{2.52}$$

depicted in Fig. 2.5. This bound improves over the upper bound (2.51) whenever $(1-v)(3v-1) > \frac{1}{N}$. By the Lemma 5 this result is promoted to all binary measurements.

For all POVMs, a sufficient condition is given by $\eta \leq (1-v)^{d-1}$ [84]. When N is large these bounds can be tighter and used instead of the bound (2.28).

2.4 Thermal-noise channel scenario

The bounds obtained in the previous section are based on the assumption that the measurements of the CMU have a finite number of discrete outcomes and that a photon loss can be modeled as a no-click outcome. However, some common quantum optics measurements are not of that form. Quadrature measurements are an example of measurements with an infinite, continuous set of outcomes, and which always produce an outcome, even when the photon is lost.

In this section, we thus make no hypothesis at all on the measurement device and we assume that losses in the channel are modeled, as is usual in quantum optics, through a thermal-noise (or attenuator) channel [28, 29] characterized through the *transmittance* η and the *excess noise* ϵ. As discussed in Section 1.2.5, the thermal-noise channel (Fig. 2.6) can be viewed as combining on a beam-splitter of transmittance η, the incoming state ρ and a thermal state τ_ν with mean photon number $\nu = (\eta\epsilon)/(2(1-\eta))$. The resulting CPTP map is

$$\Phi_{\eta,\epsilon}(\rho) = \mathrm{Tr}_2\left(U_\eta \rho \otimes \tau_\nu U_\eta^\dagger\right)\,, \tag{2.53}$$

where

$$\tau_\nu = \frac{1}{\pi\nu}\int_{\mathbb{C}} d^2\beta e^{-|\beta|^2/\nu}\,|\beta\rangle\langle\beta|\,, \tag{2.54}$$

is a thermal state written in the basis of coherent states $|\beta\rangle$, and

$$U_\eta = e^{\arccos(\sqrt{\eta})(\hat{a}^\dagger\hat{b} - \hat{a}\hat{b}^\dagger)}, \tag{2.55}$$

represent the action of the beam-splitter, where $\hat{a}$ and $\hat{b}$ are annihilation operators acting on the input state ρ and the thermal state τ_ν, respectively.

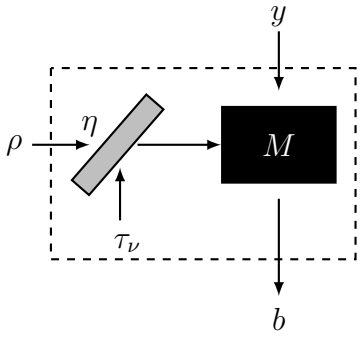

Figure 2.6: CMU where the channel is a thermal-noise channel which can be viewed as combining ρ with a thermal state τ_ν at a beam-splitter of transmittance η.

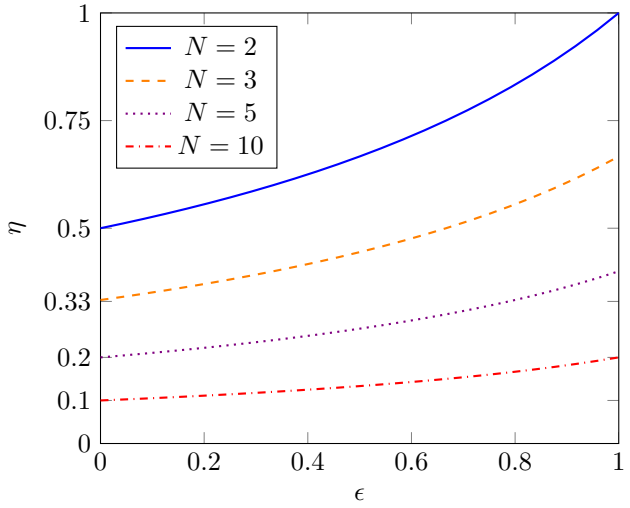

Figure 2.7: The curves depict the lines below which the channel $\Phi_{\eta,\epsilon}$ of the thermal-noise CMU is jointly measurable according to Upper bound 7 for $N = 2, 3, 5, 10$. The blue line provides also the region below which any number of Gaussian measurements is jointly measurable as we state with Upper bound 8.

Upper bound 7. *The channel $\Phi_{\eta,\epsilon}$ of a thermal-noise CMU is N-extendable, and thus the corresponding CMU is JM irrespective of the measurements M_y, if*

$$\eta \leq \frac{1}{N(1-\epsilon/2)}\,. \tag{2.56}$$

Proof. Consider the channel $\Phi_{1\to N} = \Phi_{\mathrm{BS}_N} \circ \Phi_{\mathrm{Amp}}$ corresponding to first amplifying the input state ρ with a gain $G > 1$ and then splitting it using a symmetric N-mode beam-splitter, as illustrated in Fig. 2.8. We now show that if we trace out $N-1$ output modes at the exit of this channel, we get the thermal-noise channel $\Phi_{\eta,\epsilon}$ by choosing appropriately G.

The thermal-noise channel, the amplifier, and the beam-splitter are Gaussian channels, and their action can be fully described in terms of two real matrices, the scaling and noise matrices X and Y [29]. These matrices capture the evolution of the characteristic function of the system, regardless of whether the initial state is Gaussian or not. For the thermal-noise channel, the amplifier, and the channel corresponding to tracing out all but one ouptput mode of a symmetric N-mode beam-splitter, these matrices are respectively

$$X_{\mathrm{Th}} = \sqrt{\eta}\mathbb{1}, \qquad Y_{\mathrm{Th}} = (1-\eta+\epsilon\eta)\mathbb{1}\,, \tag{2.57}$$

$$X_{\mathrm{Amp}} = \sqrt{G}\mathbb{1}_2, \qquad Y_{\mathrm{Amp}} = (G-1)\,\mathbb{1}_2, \tag{2.58}$$

and

$$X_{\mathrm{BS}} = \sqrt{\frac{1}{N}}\mathbb{1}_2, \qquad Y_{\mathrm{BS}} = \frac{N-1}{N}\mathbb{1}_2,. \tag{2.59}$$

In the Gaussian systems framework, the channel obtained by applying two channels in succession, each described by the matrices $X_{A,B}$ and $Y_{A,B}$, is described by the matrices

$$X_{B\circ A} = X_B X_A, \tag{2.60}$$

$$Y_{B\circ A} = X_B^T Y_A X_B + Y_B. \tag{2.61}$$

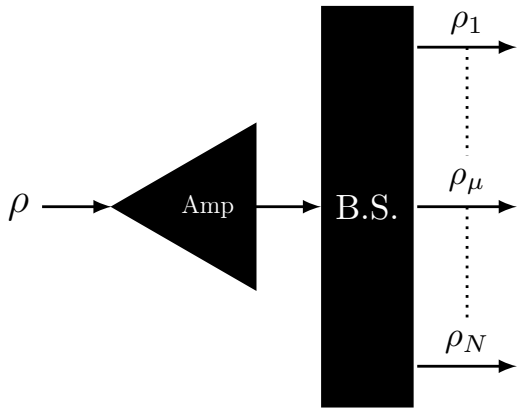

Figure 2.8: Strategy for establishing the N-extendibility of the thermal-noise-channel. The initial state ρ is amplified and split into N modes using a beam splitter.

We thus have

$$X_{\mathrm{BS}\circ\mathrm{Amp}} = X_{\mathrm{BS}} X_{\mathrm{Amp}} = \sqrt{\frac{G}{N}}\mathbb{1}_2\,, \tag{2.62}$$
$$Y_{\mathrm{BS}\circ\mathrm{Amp}} = X_{\mathrm{BS}}^T Y_{\mathrm{Amp}} X_{\mathrm{BS}} + Y_{\mathrm{BS}} = \frac{G-N-2}{N}\mathbb{1}_2\,.$$

These matrices are equal to the target matrices X_{Th} and Y_{Th} for $G = 1/(1-\epsilon/2)$ and

$$\eta = \frac{1}{N(1-\epsilon/2)}. \tag{2.63}$$

Of course, lower values of η are also possible by simply attenuating the overall channel output, resulting in the bound (2.56). □

We note that the above result was also obtained in [86], where a necessary and sufficient condition for the N-extendibility of single-mode Gaussian channels is presented. The above proof, however, provides an explicit strategy for establishing the N-extendibility of the thermal-noise channel.

While we considered a single-mode thermal-noise channel $\Phi_{\eta,\epsilon}$, the argument extends to n-mode scenarios, where the global noise channel is the product $\bar{\Phi}_{\eta,\epsilon} = \bigotimes_{i=1} \Phi_{(\eta,\epsilon)}$. It is easy to see that this channel $\bar{\Phi}_{\eta,\epsilon}$ is N-extendable if $\Phi_{\eta,\epsilon}$ is. Therefore the Upper bound 7 also holds in this scenario.

The bound (2.56) is completely generic as it depends only on the number of measurements N and the excess noise ϵ, and holds for any possible measurements M_y implemented by the CMU. A much more stringent bound can be obtained if we assume that the measurements M_y are Gaussian measurements, which can be implemented by first applying a Gaussian channel to the state and then performing homodyne detection [98].

Upper bound 8 ([87])**.** *A CMU consisting of a thermal-noise channel followed by Gaussian measurements is JM if*

$$\eta \leq \frac{1}{2-\epsilon}\,. \tag{2.64}$$

Remakably, the above bound on the tranmsissivity η is independent of the total number N of measurements performed. It follows from the fact that the thermal-noise channel is incompatibility breaking for Gaussian measurements as shown in [87]. We provide an explicit strategy to obtain this bound in the following in the case of pure homodyne measurements.

Proof. Homodyne measurements corresponding to the observables

$$\hat{X}(\theta) = \cos\theta\hat{X} + \sin\theta\hat{P}, \tag{2.65}$$

specified by the angle θ and where $\hat{X}$ and $\hat{P}$ are the position and momentum quadrature operators

$$\hat{X} = \frac{\hat{a} + \hat{a}^\dagger}{\sqrt{2}}, \quad \text{and} \quad \hat{P} = \frac{\hat{a} - \hat{a}^\dagger}{i\sqrt{2}}. \tag{2.66}$$

We now show that when the measurements M_y in a thermal-noise CMU correspond to the above observables (with the angle $\theta(y)$ depending on y), the resulting effective POVMs are JM independently of the total number N of such measurements.

In a thermal-noise model, the annihilation operators follow the input-output relation

$$\hat{a}_{\text{out}} = \sqrt{\eta}\hat{a}_{\text{in}} + \sqrt{1-\eta}a_{\tau_\nu}, \tag{2.67}$$

where a_{τ_ν} is the annihilation operator acting on the thermal state that is combined with the input state in our noise model. We can use the latter relation to express the effective CMU quadratures that are implemented by the CMU as

$$\hat{X}_{\text{out}}(\theta) = \sqrt{\eta}\hat{X}_{\text{in}}(\theta) + \sqrt{1-\eta}\hat{X}_{\tau_\nu}(\theta). \tag{2.68}$$

We now introduce a single parent POVM that simulates the results of the measurements of such quadratures for any value of θ. This parent POVM consists in performing first a heterodyne measurement on the input state, i.e, we split the initial mode in two modes using a 50/50 beam splitter and measuring the X quadrature of one output mode and the P quadrature of the other output mode. The obtained classical outcomes, x and p, are then sent to the measurement device, where given the input angle θ, it ouputs

$$x_\theta = G(\cos\theta x + \sin\theta p) + \Delta, \tag{2.69}$$

where G is a constant real number, while Δ is a random real number with centered Gaussian distribution of variance σ^2.

In this equivalent CMU, the input-output relations for the annihilation operators after the 50/50 beam splitter are

$$\hat{a}_1 = \frac{\hat{a}_{\text{in}} + \hat{v}}{\sqrt{2}}, \quad \text{and} \quad \hat{a}_2 = \frac{\hat{a}_{\text{in}} - \hat{v}}{\sqrt{2}}, \tag{2.70}$$

where $\hat{v}$ is the annihilation operator of the vacuum state entering the dark port of our beam splitter and $\hat{a}_1$ and $\hat{a}_2$ are the annihilation operators of the two output modes. The quadrature $\hat{X}$ measured in the first ouptut mode and the quadrature $\hat{P}$ measureed in the second output mode can then be written as

$$\hat{X}_1 = \frac{\hat{X}_{\text{in}} + \hat{X}_v}{\sqrt{2}}, \quad \text{and} \quad \hat{P}_2 = \frac{\hat{P}_{\text{in}} - \hat{P}_v}{\sqrt{2}}. \tag{2.71}$$

The probability distribution of the final output x_θ of the CMU after the post-processing (2.69) is then fully characterized by the moments $\langle x_\theta^n \rangle$. This strategy simulates the original CMU if these

moments coincides with the moments $\langle \hat{X}(\theta)^n \rangle_{\rm th}$ of the distribution of the effective observables (2.68). We now determine such moments.

We first compute the target moments.

$$\langle \hat{X}(\theta)^n \rangle_{\rm th} = \left\langle \left(\sqrt{\eta} \hat{X}_{\rm in} \theta) + \sqrt{1-\eta} \hat{X}_{\tau_\nu}(\theta) \right)^n \right\rangle \tag{2.72}$$

$$= \sum_{k=0}^{n} \binom{n}{k} \sqrt{\eta}^{n-k} \langle \hat{X}_{\rm in}(\theta)^{n-k} \rangle \sqrt{1-\eta}^{k} \langle \hat{X}_{\tau_\nu}(\theta)^k \rangle \tag{2.73}$$

$$= \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} \sqrt{\eta}^{n-2k} \langle \hat{X}_{\rm in}(\theta)^{n-2k} \rangle \\ (1-\eta)^k \langle \hat{X}_{\tau_\nu}(\theta)^{2k} \rangle \tag{2.74}$$

$$= \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} \sqrt{\eta}^{n-2k} \langle \hat{X}_{\rm in}(\theta)^{n-2k} \rangle \\ (2k-1)!! \left(\frac{1}{2}(1-\eta+\epsilon\eta) \right)^k , \tag{2.75}$$

where we used the fact that the odd moments of the thermal state in Eq. (2.54) are zero, while the even ones are

$$(2k-1)!! \left(\frac{1}{2} + \frac{\epsilon\eta}{2(1-\eta)} \right)^k . \tag{2.76}$$

We now determine the moments following from the *JM* strategy.

$$\langle x_\theta^n \rangle = \langle \Big(G(\cos\theta \hat{X}_1 + \sin\theta \hat{P}_2) + \Delta \Big)^n \rangle \tag{2.77}$$

$$= \left\langle \left(\frac{G}{\sqrt{2}} (\hat{X}_{\rm in}(\theta) + \hat{X}_v(-\theta)) + \Delta \right)^n \right\rangle \tag{2.78}$$

$$= \sum_{k=0}^{n} \binom{n}{k} \left(\frac{G}{\sqrt{2}} \right)^{n-k} \langle X_{\rm in}(\theta)^{n-k} \rangle \\ \left\langle \left(\frac{G}{\sqrt{2}} \hat{X}_v(-\theta) + \Delta \right)^k \right\rangle \tag{2.79}$$

$$= \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} \left(\frac{G}{\sqrt{2}} \right)^{n-2k} \langle X_{\rm in}(\theta)^{n-2k} \rangle \\ \left\langle \left(\frac{G}{\sqrt{2}} \hat{X}_v(-\theta) + \Delta \right)^{2k} \right\rangle, \tag{2.80}$$

where in the last line we used the fact that the sum of two random variables with Gaussian

central distribution has zero odd moments. Now, we have

$$\left\langle \left(\frac{G}{\sqrt{2}} \hat{X}_v(-\theta) + \Delta \right)^{2k} \right\rangle =$$

$$= \sum_{l=0}^{2k} \binom{2k}{l} \left(\frac{G}{\sqrt{2}} \right)^{2k-l} \langle \hat{X}_v(-\theta)^{2k-l} \rangle \langle \Delta^l \rangle \tag{2.81}$$

$$= \sum_{l=0}^{k} \binom{2k}{2l} \left(\frac{G}{\sqrt{2}} \right)^{2(k-l)} \left(\frac{1}{2} \right)^{k-l} \sigma^{2l} \tag{2.82}$$

$$(2(k-l)-1)!!(2l-1)!!, \tag{2.83}$$

where we used again that the odd moments are zero and that the even ones are

$$\langle \hat{X}_v(-\theta)^{2(k-l)} \rangle = \left(\frac{1}{2} \right)^{k-l} (2(k-l)-1)!!, \tag{2.84}$$

$$\langle \Delta^l \rangle = \sigma^{2l} (2l-1)!!. \tag{2.85}$$

Using $(2n)! = (2n)!!(2n-1)!!$ and $(2n)!! = 2^n n!$, we get

$$\binom{2k}{2l} (2(k-l)-1)!!(2l-1)!! = \binom{k}{l} (2k-1)!!, \tag{2.86}$$

and thus

$$\left\langle \left(\frac{G}{\sqrt{2}} \hat{X}_v(-\theta) + \Delta \right)^{2k} \right\rangle = (2k-1)!! \left(\frac{G^2}{4} + \sigma^2 \right)^k. \tag{2.87}$$

We finally obtain that

$$\langle x_\theta^n \rangle = \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} \left(\frac{G}{\sqrt{2}} \right)^{n-2k} \langle X_{\text{in}}(\theta)^{n-2k} \rangle$$

$$(2k-1)!! \left(\frac{G^2}{4} + \sigma^2 \right)^k. \tag{2.88}$$

It thus follows that $\langle \hat{X}(\theta)^n \rangle_{\text{th}} = \langle x_\theta^n \rangle$ for all n if

$$G = \sqrt{2\eta}, \quad \text{and} \quad \sigma^2 = \frac{1}{2}(1 - 2\eta + \epsilon\eta). \tag{2.89}$$

Since $\sigma^2 \geq 0$, we conclude that this strategy works as long as (2.64) holds.

□

2.5 Discussion

We have developed a general framework for characterizing the impact of losses in quantum communication setups with untrusted measurement devices. A key element in our work is the concept of a CMU, which should be viewed as a set of effective measurements, characterizing both the imperfections of the channel and the subsequent measurement apparatus. We investigated

the joint measurability of this set, establishing a direct connection between this fundamental notion and quantum communication.

Our work motivates the investigation of the (in)compatibility of sets of POVM under the combined effect of noise and loss, of which we discussed several cases. In particular, we have derived a sufficient condition for the compatibility of sets of N unbiased qubit measurements (Lemma 6). We have also seen that results on channel extendibility are very useful in this context, as they imply compatibility of CMUs solely based on the noise channel and the number of measurements that compose it. We derived a general result on the extendibility of concatenated noise channels. Our results can be applied to any protocol or experimental scenario with untrusted measurement devices, including DI and SDI protocols, providing insight into the admissible trade-off between noise and loss. Obtaining more results in this direction would be desirable.

3 Security and key rate

Device Independent Quantum Key Distribution relies on specific assumptions and on the principles of quantum mechanics. In this chapter, we describe these assumptions, and we define security for both DI QKD and semi-Device Independent protocols. We explore various attack types and methods to prove security against some of them, particularly the most general type. Key rate formulas are discussed, highlighting the relationship between eavesdropping detection and the final length of the secret key. The chapter focuses on key rate formulas in the asymptotic case, where the number of protocol rounds tends to infinity, and describes the generalised Entropy Accumulation Theorem which allows to derive security against the most general attacks.

3.1 Introduction

The security of a quantum cryptographic protocol relies on the properties of quantum mechanics and an additional set of assumptions. A DI QKD protocol minimises these additional assumptions, while semi-DI QKD protocols add a few carefully chosen assumptions that seek to be more manageable compared to those in standard QKD protocols. In standard QKD, we need a complete characterization of all operations, including for example the dimension of the Hilbert space. In semi-DI QKD, the assumptions can be chosen to avoid such complex characterizations of the devices used.

Unlike classical algorithms, which rely on the belief in the computational hardness of certain mathematical problems and the assumption that potential adversaries lack the computational power to solve them, DI QKD protocols adopt a different approach consisting only on securing the location of measuring devices and trusting the classical computations performed during the protocols. This approach makes it possible to achieve information-theoretical security through principles rooted in quantum mechanics.

Proving the security of a quantum cryptographic protocol starting from a minimal set of assumptions is a non-trivial task. A crucial step in this process is obtaining a formula for the *key rate*, defined as the amount of secret key bits per round of the protocol. The key rate formulas discussed in this chapter express that the less an eavesdropper could have accessed information about the raw key, the longer the final key can be. Conversely, the less correlation there is between Alice and Bob, the shorter the final key can be. If it is possible for an eavesdropper to have accessed "too much" information about the key, then the key rate formulas tell us that the length of the final key is zero and we need to abort the protocol.

In a real protocol, only a finite number of rounds are performed, introducing statistical fluctuations on the parameters estimated during the protocols and reducing the performance of the error correction algorithms. These fluctuations need to be taken into account in the security analysis of protocols, ensuring that the protocol remains secure despite these finite-size effects. Analyses of this type are typically conducted when the security of protocols is assessed. References such as [99, 100] provide detailed insights into this aspect. However, in this chapter, we will focus mainly on the asymptotic case, where we assume that the number of rounds of the protocol tends to infinity.

In the following, we will explain how key rate formulas can be derived and what expressions are used to quantify the information that Eve might have obtained, and the amount of correlation

between Alice and Bob. We will begin by outlining the assumptions used to prove security in Section 3.2. In Section 3.3, we will discuss the definition of security in QKD, followed by a description of the different types of attacks in Section 3.4, leading up to the most general class of attacks that can be performed within our assumptions. In this section, we will also express a formula for the key rate that holds when eavesdroppers are restricted to a less general class of attacks, denoted as collective attacks. Finally, in Section 3.5 we will employ the Entropy Accumulation Theorem and we will sketch the derivation of the asymptotic key rate for protocols secure against the most general class of attacks. By the end of the chapter, we should be able to understand how to express the key rate in different scenarios and what we mean by security in the context of DI QKD. The following subjects are discussed in [5, 99, 100].

3.2 Assumptions in Device Independent QKD protocols

Just as in mathematical or physical theories, where proofs are established based on axioms or postulates, the proofs of security in cryptography also start from assumptions that we believe true. These assumptions must be met to ensure the security of a protocol. If they are not satisfied, it might be possible to break the cryptographic protocols. It is crucial for those implementing and using the protocol to understand these assumptions and ensure they are fulfilled.

Let us list the assumptions necessary for proving the security of a DI QKD protocol. We will discuss the definition of security in more detail later in this chapter. A DI QKD protocol is considered secure under the following conditions.

- The theory of quantum mechanics is correct.
- Alice and Bob's locations and their classical devices are secure. This means that no information can leak from their laboratories and that the classical units behave as established by the protocol. For instance, there is no quantum cryptographic protocol that could prevent an undercover agent from copying the secret key after it has been generated and sending it to his allies, and, similarly, compromised classical devices could transmit the secret key to the eavesdroppers without the quantum protocol failing.
- Alice and Bob can rely on a secure random number generator. There is no eavesdropper that can decide the random numbers generated or read them before they are revealed.
- Alice and Bob can communicate through a public authenticated classical channel. This means that eavesdroppers might be listening their classical communication but they cannot pretend to be one of the honest parties.

Assuming the correctness of quantum mechanics is necessary because, in principle, a theory with greater predictive power could allow an eavesdropper to better predict the outcomes of Alice and Bob's measurements. Typically, one would also assume that quantum mechanics is complete—that is, no extension of quantum mechanics can predict measurement outcomes more accurately than quantum mechanics itself. However, in this specific case, this additional assumption is not required. By assuming only the correctness of quantum mechanics and that the inputs are generated by a secure random number generator, we can guarantee that Eve cannot obtain a more accurate description of Alice and Bob's outcomes than what quantum mechanics provides. This result was demonstrated in [101].

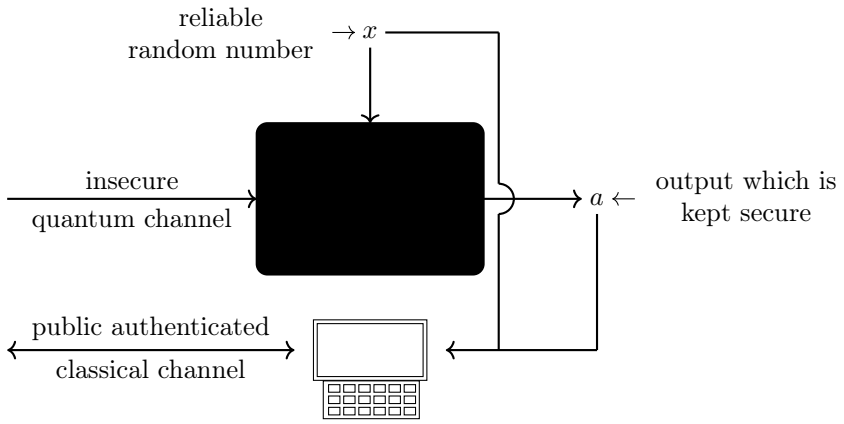

Figure 3.1: A quantum device represented as a black box receiving inputs x and producing outputs a. The generated inputs and outputs are sent to a trusted computing unit that can perform authenticated public communication.

In traditional QKD protocols, such as the well-known BB84, there is an additional assumption that the quantum devices used by Alice and Bob are accurately characterized and consistently perform as expected, similarly to what we assumed about the classical ones. For instance, in the traditional version of the BB84 protocol, Alice prepares specific two-dimensional states $|0\rangle$, $|1\rangle$, $|+\rangle$, $|-\rangle$, and Bob measures corresponding two-dimensional observables Z and X. However, as we pointed out in Chapter 1, characterizing devices is extremely challenging and when we do not succeed, we open loopholes that can be exploited to eavesdrop the communication.

In contrast, Device Independence refers to the fact that the devices of Alice and Bob are not characterized as in a standard QKD protocol. In DI QKD the devices used are modelled as black boxes which receive inputs and produce outputs (Fig. 3.1). The devices do not need to be characterized in the sense that we do not need to have a specific model describing the behaviour of the measurement apparatus. However, untrusted does not mean that the entire device could be provided by a malicious third party, otherwise it could for instance be designed to initially generate the secret key securely, but then transmit it back to the eavesdropper. It is important that every unit responsible for performing classical computation steps must be trusted. This includes protocols for authentication on the classical channel, privacy amplification, information reconciliation algorithms, and so on. This makes sense, as classical devices are much less susceptible to noise and errors compared to quantum devices, making their characterization more straightforward. Indeed, hacking the classical device can compromise any type of QKD protocol, but this risk is inherent to all cryptographic protocols, as a malicious actor could always install a virus on the end user's system to share the generated key with the eavesdropper once it has been generated. Therefore, implementing robust cybersecurity measures and securing the classical infrastructure are essential to prevent such attacks.

For semi-DI QKD protocols, additional assumptions can be introduced in a way that they make the protocol more practical while still being secure. These assumptions can vary. For example, in Section 4.3.2, we will explain that we can consider fixed overlaps between the states prepared. Another approach is to have characterized sources and untrusted measuring devices, as in Measurement-Device-Independent QKD. Alternatively, we can assume that only one side in the protocol is trusted, which can be useful for asymmetric tasks such as communication between a server and clients, as we will discuss in Chapter 6.

Finally, let us remark that, as in the standard QKD scenario, in DI and semi-DI frameworks the quantum channel allowing Alice and Bob to exchange quantum information is not trusted,

and Eve is allowed to intercept and modify any signal passing through it.

From this point onward in the chapter, we will assume that the secret key is generated at Alice's device, while the string produced by Bob is used to reconstruct Alice's key through information reconciliation. In this scenario, the protocol involves one-way public communication from Alice to Bob. The initial strings obtained are referred to as raw keys, and they are composed by realisations of the random variables K_A and K_B. These random variables consist of the outcomes of measurements during the rounds designated as *key generation rounds*. In the remaining rounds of the protocol, known as *parameter estimation rounds*, Alice and Bob estimate quantities that will be used to bound the information possessed by Eve. During these rounds, they store the symbols $K_A = \perp$ and $K_B = \perp$ in their raw keys to indicate that such rounds should be excluded from the extraction of the final secret key.

3.3 Security definition

In QKD, as with most existing technologies, a protocol carries a possibility of failure, though we require that the probability of this happening is small. A QKD protocol can fail in three ways: it can either abort even without an eavesdropper's interference, fail to protect the secret key from Eve, or fail to produce matching keys for Alice and Bob. Security proofs assign a maximum probability to each of these failure events.

1. **Completeness**: For the first eventuality, we require that an *honest implementation* of the protocol has a probability ϵ_{comp} of aborting which has to be low. An honest implementation refers to a simulated implementation of the protocol based on a mathematical model of the devices. The models include a realistic amount of noise and imperfections in the devices. The completeness condition ensures that the protocol has a low probability of aborting when everything functions as expected.

2. **Secrecy**: For the second case, we require that the QKD protocol is ϵ_s-secret, meaning that the probability of an eavesdropper gaining sufficient information to correctly guess the key is less than ϵ_s. Specifically, let A be Alice's register, E be any register possessed by Eve, $p(\text{accept})$ be the probability that the protocol does not abort, and l be the length of the final key. We require that the output state τ_{AE} of our DI QKD protocol shared by Alice and Eve satisfies

$$p(\text{accept})\frac{1}{2}\left|\left|\tau_{AE} - \frac{1}{2^l}\sum_{k\in\{0,1\}^l} |k\rangle\langle k|_A \otimes \tau_E\right|\right|_1 \leq \epsilon_{\text{s}}. \tag{3.1}$$

   Here, τ_{AE} represent normalized states conditioned on the protocol not aborting and $\tau_E = \text{Tr}_A(\tau_{AE})$.

3. **Correctness**: Lastly, we require that the protocol is ϵ_c-correct, indicating that the probability of the protocol not aborting and, at the same time, Alice's and Bob's keys not matching is at most ϵ_c.

A QKD protocol, before performing information reconciliation and privacy amplification, outputs a raw key of length n which is composed by the realisations of the random variables $K_A^n = K_{A,1}\dots K_{A,n}$ of each key generation round. This raw key may contain imperfections due to noise or potential eavesdropping and thus, not meet the secrecy and correctness conditions.

First, the keys produced by Alice and Bob might not perfectly match. This issue is addressed by employing an information reconciliation algorithm, which reveals a fraction leak_{EC} of Alice's raw key to allow Bob to reconstruct Alice's key (or vice versa). We can estimate the cost of information reconciliation in the case of a honest implementation of our protocol, where the random variables K_A^n are identical and independently distributed. We will refer to one of the random variables as K_A since they all have the same probability distribution. As shown in [99], the fraction of bits leaked in a protocol with $n \to \infty$ is

$$\lim_{n\to\infty} \frac{\text{leak}_{EC}}{n} \leq H(K_A|K_B, X, Y), \tag{3.2}$$

where X and Y are the random variables representing the inputs of Alice and Bob's devices. From now on we will neglect the conditionings on X and Y. Information reconciliation can be implemented in such a way that the probability of failure is at most ϵ_c, thereby meeting the correctness requirement.

As for the secrecy requirement, we can invoke the generalised Leftover Hash Lemma [102]. This lemma enables us to meet the secrecy condition by applying a two-universal hash function $f(x)$ to Alice's raw key. This function ensures that for any two inputs x and y, the probability that the outputs $f(x)$ and $f(y)$ are the same is no more than $1/m$, where m is the number of possible inputs. This guarantees that even if Eve knows part of the raw key, the output she obtains after applying the hash function is very likely to be different from Alice's.

Let us now define the conditional smooth min entropy which is

$$H_{\min}^{\epsilon_s}(K_A^n|E)_{\rho_{AE}} = \max_{\sigma \text{ s.t. } 0\leq \text{Tr}\,\sigma\leq 1 \text{ and } P(\rho_{AE},\sigma)\leq\epsilon_s} H_{\min}(K_A^n|E)_\sigma, \tag{3.3}$$

where the min-entropy was defined in Section 1.1, while its conditional version will be discussed in Section 5.4. The conditional smooth min entropy quantifies the minimum amount of randomness remaining in Alice's key K_A^n when conditioned on any information E available to Eve. Specifically, it is defined as the maximum value that the function $H_{\min}(K_A^n|E)$ can take over a set of states σ that are ϵ_s-close to the original state ρ_{AE} in terms of purified distance

$$P(\rho, \sigma) = \sqrt{1 - F^2(\rho, \sigma)}, \tag{3.4}$$

where

$$F(\rho, \sigma) = \text{Tr}\left|\sqrt{\rho}\sqrt{\sigma}\right| + \sqrt{(1 - \text{Tr}(\rho))(1 - \text{Tr}(\sigma))}. \tag{3.5}$$

The generalised Leftover Hash Lemma states that the secrecy condition can be met as long as the length l of the final key, after the hash function has been applied, is less than or equal to the conditional smooth min-entropy of Alice's raw key, hence

$$l \leq H_{\min}^{\epsilon_s}(K_A^n|E)_{\rho_{AE}} - \text{leak}_{EC}, \tag{3.6}$$

where we subtracted the term leak_{EC} because the hash function is applied after information reconciliation.

In conclusion, it is desirable to choose l to be as close as possible to the right-hand side of equation (3.6), while still ensuring that it is smaller. In the following, we will see how to simplify the computation of this quantity.

3.4 Attacks

As we pointed out in Section 3.2, a DI quantum key distribution protocol is designed to be secure against any attack within the defined set of assumptions. These protocols are considered secure

against *coherent* or *general attacks*. Historically, however, three increasingly general classes of attacks on quantum cryptographic protocols have been defined (Fig. 3.2). Additionally, there is a category of attacks known as no-signalling attacks. These fall outside the assumptions of quantum mechanics but still respect all other assumptions we made, and they include the constraint that the speed of light cannot be exceeded.

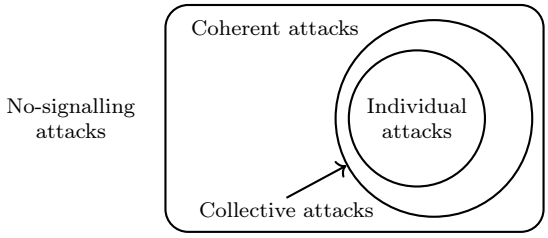

Figure 3.2: Classes of attacks

Let us present the three classes of attacks allowed by quantum mechanics. We will start from the most specific and proceed to the most general ones.

3.4.1 Individual attacks

The first class that was discussed in the literature [103–106] is the one of individual attacks (Fig. 3.3).

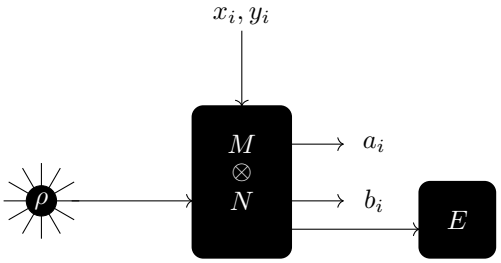

Figure 3.3: An individual attack performed by Eve. At each round she prepares a state ρ and she stores a subsystem of it. At the end of the protocol she measures individually each subsystem stored.

A protocol performed by Alice and Bob proceeds in many rounds. At each round (labelled as $i = 1, \ldots, n$), the two parties receive a quantum state, they produce random numbers $x_i,\ y_i$, they measure the state received, and they obtain outcomes $a_i,\ b_i$. For individual and collective attacks, we will assume that the state representing the systems of Alice, Bob and Eve during the entire protocol is of the form

$$\rho_{\text{protocol}} = \rho \otimes \rho \otimes \cdots \otimes \rho = \rho^{\otimes n}, \tag{3.7}$$

where n is the number of rounds of the protocol, and the measurements performed by Alice and Bob are of the form

$$M_{\text{protocol}} = M \otimes M \otimes \cdots \otimes M \qquad \text{and} \qquad N_{\text{protocol}} = N \otimes N \otimes \cdots \otimes N. \tag{3.8}$$

This means that the outcomes generated at each round of the protocol are random variables which are identical and independently distributed (i.i.d.) of the ones at the previous rounds.

In an individual attack, Eve performs an identical attack in every round of the protocol. At each round, she generates the initial state ρ, she sends one part to Alice, one part to Bob, and conserves one subsystem for herself. She is allowed to store her subsystem in a quantum memory and measure it at the end of the protocol, after she learned the information leaked during public communication, but she is not allowed to perform a joint measurement of all systems she kept.

Moreover, as we are in a DI setting, Eve is allowed to tamper the measuring devices of Alice and Bob. In particular, Eve is allowed to decide what is the operation performed inside their device once the device is provided with a certain quantum state and a classical input (x_i, y_i). The outcome produced, anyway, will remain private.

3.4.2 Collective attacks

The second class that we describe was introduced in [107, 108] and is denoted as collective attacks (Fig. 3.4).

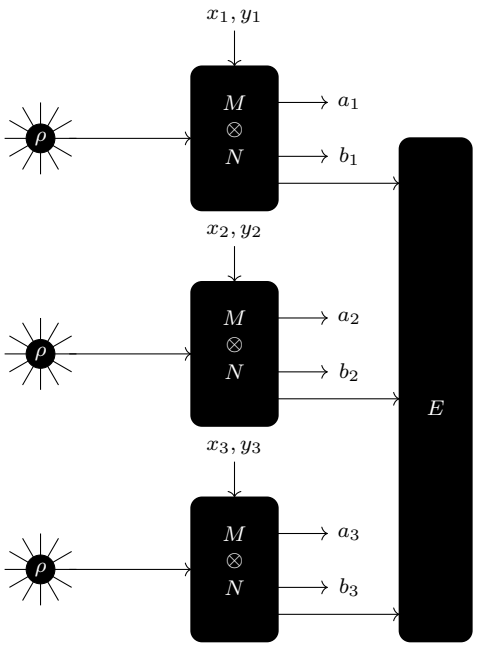

Figure 3.4: A collective attack performed by Eve. For simplicity, we represent a protocol consisting of three rounds. At each round a state ρ is prepared and all the operations performed do not depend on the ones at the previous rounds. Eve is allowed to perform a joint measurement E on all the subsystems that she stores at each round of the protocol.

As before, for this class of attack, we assume that the outcomes at each round of the protocol are i.i.d. random variables. This allows us to drop the label i indicating the round when we refer to the random variables of a single round. The only difference with the previous section is that Eve is allowed to perform a joint measurement of all the subsystems that she conserved at each round of the protocol.

When Eve is limited to collective attacks, the quantum Asymptotic Equipartition Property [109] can be used to compute the conditional smooth min-entropy of eq. (3.3) using the conditional

von Neumann entropies of individual protocol rounds. We have

$$H_{\min}^{\epsilon_s}(K_A^n|E)_{\rho^{\otimes n}} \geq n\left(H(K_A|E)_\rho - \frac{c(\epsilon_s)}{\sqrt{n}}\right), \tag{3.9}$$

where $c(\epsilon_s)$ is an error term dependent on ϵ_s, and K_A is the random variable that we obtain from a single key generation round. Note that during test rounds the outcomes and inputs are revealed, and we get $H(K_A|E, \text{test round}) = 0$, while for all other n rounds the entropies are equal since K_A are i.i.d. Finally, we obtain the asymptotic key rate which is

$$r_\infty \geq \lim_{n\to\infty} \frac{l}{n} = H(K_A|E) - H(K_A|K_B). \tag{3.10}$$

The asymptotic key rate for collective attacks is often referred to as Devetak-Winter rate as it was first proven by I. Devetak and A. Winter in [110] in the context of fully characterized devices. However, as we just described, this formula holds for all protocols where only collective attacks are allowed. In the previous formulas, we expressed the conditonal Von Neumann entropy in terms of the random variables K whose realisations are stored in registers. Alternatively, indicating as A and B the registers of Alice and Bob used to store the raw keys, we can write the key rate formula (3.10) in terms of their registers. In this way, we obtain the most common expression of this formula, that is

$$r_\infty \geq H(A|E) - H(A|B), \tag{3.11}$$

or, in case the information reconciliation is performed by Alice to retrieve Bob's key, we have

$$r_\infty \geq H(B|E) - H(B|A). \tag{3.12}$$

3.4.3 Coherent attacks

The classes of attacks that we just discussed, are characterized by the assumption that at each round the behaviour of the quantum devices is independent of what happens before. However, in principle, it is not completely safe to assume that the outcomes of Alice and Bob are not correlated with the outcomes from previous rounds. For example, their measuring device might be warmed up by the arrival of a certain type of quantum states and, because of its temperature, it might be more likely to perform some specific operation during the following round. This piece of information can be exploited by Eve to better perform her attack.

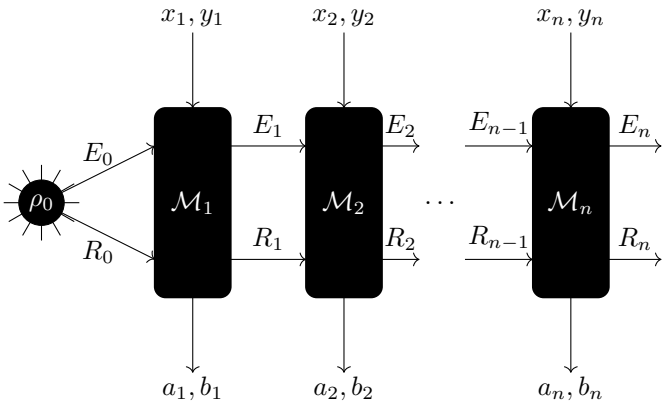

Figure 3.5: A coherent attack performed by Eve. At each step of the protocol the memory of Alice's and Bob's device R_i and Eve's system E_i are updated.

In the most general case, the attack of Eve consists of preparing an initial state ρ_0 at the beginning of the protocol and providing with a subsystem R_0 of it the devices of Alice and Bob, while keeping a second subsystem E_0 for her. The devices produce at each round random inputs x_i and y_i, perform operations $\mathcal{M}_i$ decided by Eve, and return outputs a_i and b_i. At each round, the internal memory of the devices R_i and the system of Eve E_i are updated.

This scenario precludes the use of the quantum Asymptotic Equipartition Property to provide a lower bound on the conditional smooth min-entropy because the variables $K_{A,i}$ across the n rounds of the protocol are not independent of each other. In the next section, we will provide an instrument to handle this more general and complex case.

3.5 Entropy Accumulation

The Entropy Accumulation Theorem (EAT) is a result which was first proven in 2016 [111]. This theorem indicates that, for protocols meeting certain criteria, the entropy of the random variables $K_{A,i}$ conditioned on Eve's information, accumulates in each round in a way that is similar to the case where $K_{A,i}$ are i.i.d. In this section, we will present a more recent result which is more general than the initial EAT and it is denoted as generalised Entropy Accumulation Theorem [112].

The generalised EAT aims to establish a bound on the conditional min-entropy in the most general context of a sequential process like that depicted in Fig. 3.5. The state describing the system in this scenario evolves at each round through the maps $\mathcal{M}_i : R_{i-1}E_{i-1} \to K_{A,i}R_iE_i$, ultimately becoming $\mathcal{M}_n \circ \cdots \circ \mathcal{M}_1(\rho^0_{R_0E_0})$. The generalised EAT applies only when the channels $\mathcal{M}_i$ satisfy a *no-signalling condition*, which requires that for each $\mathcal{M}_i$, there exists a channel $\mathcal{R}_i : E_{i-1} \to E_i$ such that

$$\mathrm{Tr}_{K_{A,i}R_i} \circ \mathcal{M}_i = \mathcal{R}_i \circ \mathrm{Tr}_{R_{i-1}} . \tag{3.13}$$

This condition ensures that at each step i, the side information in E_{i+1} does not reveal any additional information about the previous outputs $K_{A,1} \dots K_{A,i}$ that was not already present in E_i.

According to the Generalized EAT, if our DI QKD protocol respects the no-signalling condition, the conditional min-entropy can be bounded as

$$H^{\epsilon_s}_{\min}(K^n_A|E_n)_{\mathcal{M}_n \circ \cdots \circ \mathcal{M}_1(\rho^0_{R_0E_0})} \geq \sum_{i=1}^{n} \inf_{\omega} H(K_{A,i}|E_i\tilde{E}_{i-1})_{\mathcal{M}_i \otimes \mathrm{id}_{\tilde{E}_{i-1}}(\omega)} - O(\sqrt{n}), \tag{3.14}$$

where ω represents a quantum state lying on the purification of the systems $R_{i-1}E_{i-1}$, denoted as $\tilde{E}_{i-1}$.

The bound in eq. (3.14) can be used by finding a convex function f of the observed probability distributions $p(a, b|x, y)$ or of some witness as the CHSH parameter S introduced in Section 1.4. This function, known as the *min-tradeoff function*, has to lower bound the term $\inf_\omega H(K_{A,i}|E_i\tilde{E}_{i-1})_{\mathcal{M}_i \otimes \mathrm{id}_{\tilde{E}_{i-1}}(\omega)}$ independently of the round of the protocol, in a way that the infimum is taken over the states ω that are compatible with the observed witness or probability distribution. This allows us to lower bound the asymptotic key rate in the same way as we did for collective attacks. A simple example of a min-tradeoff function can be found at the beginning of Chapter 5 (eq. (5.1)). During this chapter, we will provide different ways to obtain min-tradeoff functions.

Finally, it is important to remark the limitations imposed by the no-signalling condition on the types of protocols that can be analysed using the generalised EAT. A notable example is protocols that employ *post-selection* techniques, which are particularly advantageous in scenarios

characterized by photon losses. In such protocols, Alice and Bob differentiate the rounds used for secret key generation from those used for estimating the probability distribution $p(a, b|x, y)$. During parameter estimation rounds, all outcomes, including non-detection events, are retained. Conversely, during key generation rounds, only rounds with photon detections are considered. This approach differs from a fair sampling assumption, where no-detection events are uniformly discarded. In this case, a loophole-free Bell test is performed to bound the conditional min-entropy, but the secret key is extracted solely from the subset of detected events.

Protocols involving post-selection can be analysed assuming that Eve is limited to collective attacks under the quantum Asymptotic Equipartition Property. However, they are incompatible with the generalized EAT framework (and also with the standard EAT) due to their violation of the no-signalling condition. Generally, in DI protocols, the outcomes at each round may depend on those from previous rounds due to potential design flaws. Announcements made at round i might be dependent on the outcome at round $i-1$. This implies that by announcing the discarding of a round due to non-detection of a photon, we inadvertently disclose side information about the outcomes of previous rounds, violating the no-signalling condition. This is not only a theoretical concern arising from the assumptions of the EAT framework, it was indeed shown in [113, 114] that for protocols employing post-selection, coherent attacks are stronger than collective attacks.

4 Upper bounds on DI and semi-DI QKD key rates based on joint measurability

Finding lower bounds on the conditional Von Neumann entropy in DI or semi-DI QKD protocols can be a very complicated problem as we will see in the following chapter. In this chapter we will use the results of Chapter 2 to compute upper bounds on the asymptotic key rate of DI or semi-DI QKD protocols. This will allow us to easily establish when protocols have no chance to be proven secure within certain noise regimes. To do this, we will introduce the new notion of partial joint-measurability, which arises naturally in this context. The content of this chapter is reproduced from the second part of [3].

4.1 Joint measurability and DI QKD

The approach outlined in Chapter 2 allows one to put bounds on the performance of any QKD protocol involving a CMU, in particular within the DI or semi-DI (SDI) approaches. Consider a protocol in which one of the parties, say Bob, generates their copy of the secret key by performing an untrusted measurement on a quantum state received from an untrusted channel, i.e., the key is obtained by post-processing the input y and output b of a CMU. Then no key can be extracted from the protocol if the effective POVMs of the CMU are JM. This follows from the fact that in any CMU that is JM, Alice and Bob cannot certify the presence of entanglement, which is a necessary condition for the security of QKD [88].

This result can also easily be directly proven. Let $B' = BY$ where B and Y are the random variables describing the input and output of Bob's measuring device, let A denote Alice's registers used in the QKD protocol and $\mathcal{E}$ be Eve's classical register (potentially obtained by measuring a quantum system she holds). In this chapter, Eve's attacks are limited to the individual class as it is not required to consider the most general class of attacks when looking for necessary conditions on quantum protocols. In particular, we will assume that at each round of the protocol Eve performs a measurement and stores the outcome in a classical register. It is known that for arbitrary post-processing of A and B', the key rate that can be extracted is upper-bounded by the intrinsic information [115]

$$I(A : B' \downarrow \mathcal{E}) \equiv \min_{\mathcal{E}\to\mathcal{F}} I(A : B'|\mathcal{F}), \tag{4.1}$$

where $\mathcal{F}$ runs through all post-processing of Eve's register $\mathcal{E}$ and $I(A : B'|\mathcal{F}) = \sum_f p(f) I(A : B'|\mathcal{F} = f)$ is the conditional mutual information

$$I(A : B) = H(A) + H(B) - H(AB). \tag{4.2}$$

Now consider the case where Bob's CMU is JM. Then, as Eve controls the public channel of the CMU, we can assume that she implements herself the parent POVM E, obtains the output Λ, keeps a copy for her, and sends the other copy to Bob's apparatus that uses it to determine B given Y. Hence, the probability distribution conditional on Eve's register $\mathcal{E} = \Lambda$ factorizes

$$p(A = a, B' = by|\mathcal{E} = \lambda) = p(a|\lambda)p(y)p(b|y,\lambda) \tag{4.3}$$

and $I(A : B' \downarrow \mathcal{E}) \leq I(A : B'|\mathcal{E}) = 0$, proving that no key rate can be extracted.

We conclude from the above discussion that in any DI or SDI QKD protocol where Bob's device is untrusted and his inputs and outputs are used to establish the final key, no key can be extracted if the losses and noise satisfy the various bounds presented above. However, these bounds can be further improved by using the concept of partial-joint-measurability, which we introduce in the next section.

4.2 Partial-joint measurability

In many QKD protocols, the key on Bob's side is obtained by post-processing the input y and output b only of a subset $\mathcal{K} \subset \{1, \ldots, N\}$ of cardinality $K = |\mathcal{K}| < N$ of all possible measurement inputs, while other measurements are solely used for parameter estimation. In this situation, we can weaken the notion of *JM* through the following definition.

We say that the POVMs M_y for $y \in \{1, \ldots, N\}$ are $\mathcal{K}$-JM, where $\mathcal{K}$ is subset of $\{1, \ldots, N\}$, if there exists *i)*, a quantum instrument $I = \{I_\lambda\}_\lambda$ with classical measurement outcomes λ, *ii)* probability distributions $p(b|y, \lambda)$, and *iii)* measurement $M_{y,\lambda}$ such that, for any state ρ, the probabilities $p(b|y, \rho) = \mathrm{Tr}\left[\rho M_{b|y}\right]$ can be written as

$$P(b|y, \rho) = \begin{cases} \sum_\lambda p(b|y, \lambda) \,\mathrm{Tr}\left[I_\lambda(\rho)\right] & \text{if } y \in \mathcal{K} \\ \sum_\lambda \mathrm{Tr}\left[I_\lambda(\rho) M_{b|y,\lambda}\right] & \text{if } y \notin \mathcal{K}. \end{cases} \tag{4.4}$$

Operationally, this means that instead of carrying out the original POVMs M_y on the state ρ, one can instead first perform the parent instrument I on the state ρ, obtaining a classical output λ and a quantum output $I_\lambda(\rho)$. If $y \in \mathcal{K}$, then one only uses the classical output λ to generate the outcome b with probability $p(b|y, \lambda)$, as in the case of full *JM*. If $y \notin \mathcal{K}$, then one generates b by carrying out on the state $I_\lambda(\rho)$ a measurement defined by the operators $\{M_{b|y,\lambda}\}$, which depend on y, but also on λ.

Since the relations (4.4) should hold for any input state ρ, they can be equivalently formulated as

$$M_{b|y} = \begin{cases} \sum_\lambda p(b|y, \lambda) I^*_\lambda(\mathbb{1}) & \text{if } y \in \mathcal{K} \\ \sum_\lambda I^*_\lambda(M_{b|y,\lambda}) & \text{if } y \notin \mathcal{K} \end{cases}, \tag{4.5}$$

where I^*_λ denotes the adjoint of I_λ and $I^*_\lambda(\mathbb{1}) = E_\lambda$ defines a parent POVM as in the case of full *JM*.

In the context of QKD applications, when a CMU is $\mathcal{K}$-JM, we can assume that the parent instrument I_λ is performed by Eve, which can store a copy of its classical output λ. If Bob's register $B' = BY$ used to distill the final key only depends on the inputs $y \in \mathcal{K}$ (while the other inputs may still be used for parameter estimation) the intrinsic information is zero

$$I(A : B' \downarrow \mathcal{E}) = 0\,. \tag{4.6}$$

This is because, analogously to the discussion leading to Eq. (4.3), the probability distribution of Alice's and Bob's random variable factorizes $p(A, B'|\mathcal{E} = \lambda) = p(A|\mathcal{E} = \lambda)p(B'|\mathcal{E} = \lambda)$ when conditioned on λ, which is hold by Eve. Hence, if the CMU gives rise to $\mathcal{K}$-JM measurements, no key can be distilled from the outputs of any measurements in $\mathcal{K}$.

Note that in the case where $K = N - 1$, partial joint-measurability becomes equivalent to full joint measurability. This is because there is then a single $y \notin \mathcal{K}$ and the corresponding measurement can be directly performed right after the parent instrument.

In the case of the no-click scenario, the following result relates the full joint-measurability of a subset $\mathcal{K}$ of the measurements to the partial joint-measurability of the full set of measurements.

Upper bound 9. *Consider a no-click CMU with an arbitrary number N of inputs and assume that for a subset $\mathcal{K} \subseteq \{1, \ldots, N\}$, the measurements $\{\Phi^*_\eta(M_y)\}_{y\in\mathcal{K}}$ are JM for a detection efficiency η. Then the full CMU is $\mathcal{K}$-JM for detection efficiency $\eta' \leq \eta/(1+\eta)$.*

Proof. By assumption, the set $\{\Phi^*_\eta(M_y)\}_{y\in\mathcal{K}}$ is *JM*, which means that there exists a parent instrument $\mathcal{E}$ that simulates it. Define then the following strategy. With probability q, perform the parent POVM $\mathcal{E}$ and follow the corresponding simulation scheme if $y \in \mathcal{K}$, while output no-click $\varnothing$ if $y \notin \mathcal{K}$. With probability $1-q$, leave the system untouched, output no-click $\varnothing$ if $y \in \mathcal{K}$, and perform the measurement $\Phi^*(M_y)$ if $y \notin \mathcal{K}$. On average this strategy realizes the POVMs with elements

$$\tilde{M}_{b|y} = \begin{cases} q\eta\Phi^*(M_{b|y}) & \text{if } b \neq \varnothing \\ (q(1-\eta) + (1-q))\,\mathbb{1} & \text{if } b = \varnothing \end{cases} \tag{4.7}$$

if $y \in \mathcal{K}$, and

$$\tilde{M}_{b|y} = \begin{cases} (1-q)\Phi^*(M_{b|y}) & \text{if } b \neq \varnothing \\ q\mathbb{1} & \text{if } b = \varnothing \end{cases} \tag{4.8}$$

if $y \notin \mathcal{K}$. These measurements have the form of the no-click CMU (2.5) if $q\eta = (1-q) = \eta'$, which implies $\eta' = \eta/(1+\eta)$. □

As an illustration, if we apply the above construction to the bound (2.50), we obtain that a white-noise no-click CMU implementing the qubit measurements (2.36) is $\mathcal{K}$-JM if

$$\eta \leq \frac{1}{v(K+\sqrt{K})}\,, \tag{4.9}$$

which improves the original bound when $K \leq N-1$.

Note that by construction, the condition of partial joint measurability is weaker than that of joint measurability. The bound (4.9) allows us to illustrate this difference explicitly. Indeed, we mentioned below the bound (2.50) that for $N = 3$, it is tight. In particular, the three MUBs measurements $\sigma_x, \sigma_y, \sigma_z$ are incompatible if $\eta > 2/(9v-3)$. However, from (4.9) these three measurements are 1-*JM* up to $\eta \leq 1/(2v)$. For instance, for $v = 1$, the three MUB measurements are incompatible but 1-*JM* in the region $\frac{1}{3} < \eta \leq \frac{1}{2}$.

The construction above can also be applied to, e.g, the generic white-noise no-click bound (2.28). However in this case a better improvement can be obtained using the following relation between partial joint measurability and channel extendability.

Lemma 7. *If a channel Φ is $(K+1)$-extendable, then the effective POVMs $\Phi^*(M_y)$ are $\mathcal{K}$-JM for any subset $\mathcal{K} \subseteq \{1, \ldots, N\}$ of K measurements.*

To see this it is sufficient to apply, as illustrated in Fig. 4.1, the extended channel $\Phi_{1\to K+1}$, perform the intended measurements M_y for y in $\mathcal{K}$ on the first K systems and leave the last one untouched. This defines a quantum instrument that satisfies all the required properties.

The above observation can be applied to the bounds (2.28) and (2.56) and implies that the white-noise no-click CMU is $\mathcal{K}$-JM if

$$\eta \leq \frac{d}{(K+1)\,(v(d+1)-1)}\,, \tag{4.10}$$

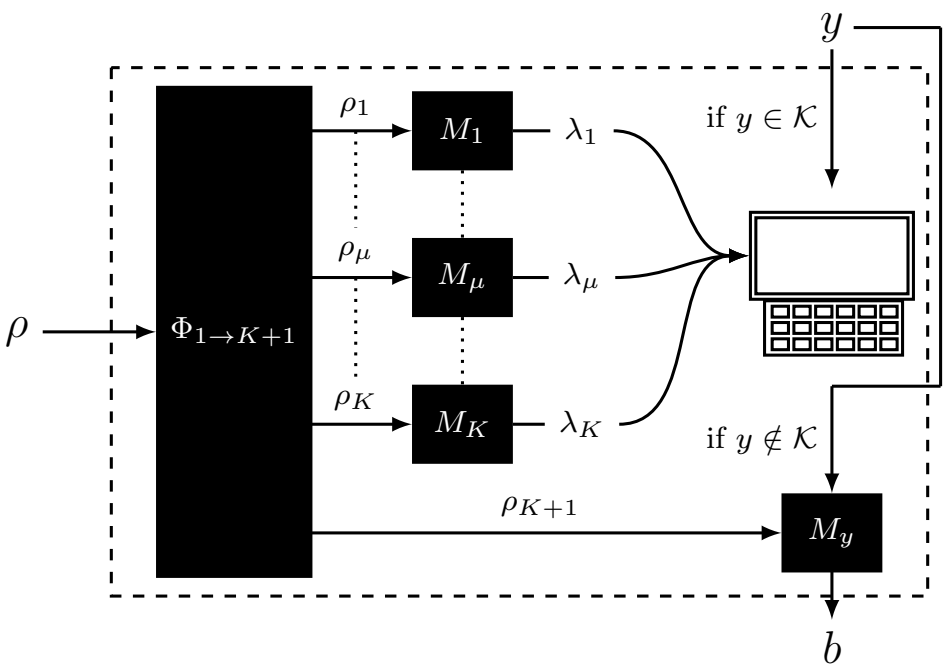

Figure 4.1: Consider a CMU that admits an implementation with a channel Φ that is $K+1$-extendable. Perform each of the K measurements M_y for $y \in \mathcal{K}$ in the first K extensions, resulting in outcomes $\lambda_1, \ldots, \lambda_K$, and forward the $K+1$th extension to the measurement device. Then by simply outputting λ_y if $y \in \mathcal{K}$ and performing the measurement M_y on the $K+1$th extension if $y \notin \mathcal{K}$ defines a CMU with effective POVMs that are $\mathcal{K}$-JM.

and that the thermal-noise CMU is $\mathcal{K}$-JM if

$$\eta \leq \frac{1}{(K+1)(1-\epsilon/2)}\,, \tag{4.11}$$

which both improve (or equal) the original bounds when $K \leq N-1$.

In the context of DI QKD or semi-DI QKD where the key is generated from a subset $\mathcal{K}$ of the measurements, the bounds (4.9), (4.9), and (4.11) can be used to strengthen the critical levels of losses and noises beyond which no key can be generated whenever $K \leq N-1$. For instance, in the thermal-noise channel scenario, no key can be generated if $\eta < 1/(2-\epsilon)$ whenever the key is generated from a single measurement, independently of the total number N of measurements used in the protocol, including those used for parameter estimation. While in the white-noise no-click scenario, we deduce from (4.10) that no key can be generated for $K = 1$ if $\eta \leq 1/(3v-1)$, corresponding, e.g., to $\eta \leq 0.5235$ for $v = 0.97$.

4.3 Convex combination attacks for public communication from Bob to Alice

As we explained above in any QKD protocol involving Alice and Bob and where Bob's quantum device is an untrusted CMU, no security can be guaranteed if the CMU is (partially)-*JM*, since Eve can then have full information about Bob's output, i.e., Alice and Bob are decorrelated

conditioned on Eve's knowledge. This is true independently of the protocol used to distil the shared key from Alice and Bob's generated data. However, as Eve's *JM* attack targets Bob's device, one would intuitively expect protocol involving one-way public communication from Bob to Alice for key distillation to be more sensitive to such attacks. Indeed, for such protocols, the maximum achievable key rate is given by the Devetak-Winter formula [110, 116]

$$r_\infty = H(B|E) - H(B|A) \tag{4.12}$$

where $H(X|Y)$ is the von Neumann entropy of X conditioned on Y. This key rate is zero as long as Eve's uncertainty $H(B|E)$ on Bob's symbols is lower than Alice's uncertainty $H(B|A)$, thus Eve's does not necessarily need to have full information about Bob's raw key for the protocol to be insecure.

This leads us to consider a class of *convex combination attacks* directly inspired by a similar strategy proposed in [117] in the context of DI QKD[1]. In such attacks, Eve employs a mixture of two strategies: either, with probability p, she implements a lossy and noisy version of the expected measurements performed by Bob, in which case she exploits the fact that this implementation is (partially) *JM* to get full information about Bob's outcome. Or with probability $1-p$, she replaces Bob's CMU with an ideal, loss-free, and noise-free CMU, in which case she has no information about Bob's outcome. This mixed strategy allows her to reproduce the actual CMU used by Bob beyond the admissible level of losses and noises derived in the previous sections, while at the same time providing her with enough information to make $r_\infty = H(B|E) - H(B|A) = 0$.

We consider such attacks only in the no-click scenario. Indeed, they cannot straightforwardly be applied to CMUs with a thermal-noise channel, since such channels are not closed under convex combinations. Specifically, we show the following in Section 4.5.

Lemma 8. *A thermal-noise channel $\Phi_{\eta,\epsilon}$ that is not N-extendable cannot be decomposed as*

$$\Phi_{\eta,\epsilon} = p\,\Phi_G + (1-p)\Phi' \qquad (0 < p < 1) \tag{4.13}$$

where Φ_G is a N-extendable Gaussian channel and Φ' is an arbitrary channel.

We prove this observation by showing that the Q-function of the channel $\Phi' = (\Phi_{\eta,\epsilon} - p\,\Phi_G)/(1-p)$ (for a coherent state input) can not be positive for such a decomposition.

In the following, we thus focus on QKD protocols with one-way communication from Bob to Alice and involving a white-noise no-click CMU represented by effective POVMs $\{M_y^{\eta,v}\}$ of the form (2.9). We assume that this set of POVMs is neither *JM*, nor $\mathcal{K}$-JM, where $\mathcal{K}$ is the set of inputs used to establish the raw key, since otherwise we have already shown that no key can be extracted.

Consider now an attack, where, with probability p, Eve uses a strategy S_* and, with probability $1-p$, a strategy S_1. In the former case, Bob's untrusted measurement device implements the POVMs $M_y^{\eta^*,v^*}$, where η^* and v^* are chosen such that these POVMs are $\mathcal{K} - JM$ so that Eve has full information about Bob's outcome, i.e., $H(B|E,S_*) = 0$. In the latter case, Bob's measurement device implements the ideal POVMs $M_y^{1,1}$, which may prevent Eve from having

[1] In the case of DI QKD, our attacks improve on the ones of [117] because they take into account not only losses but also white-noise. However, we can further improve these attacks by targeting both CMUs of a DI QKD protocol, see next subsection. We also remark that it is important to consider separately the case where a no-click outcome is kept as a distinct outcome in Alice and Bob's post-processing and the case where it is discarded. In [117] only the former case is analysed. An example of a DI QKD protocol where binning of the no-click outcome leads to a positive key rate for a detection efficiency that is lower than the threshold computed in [117] can be found in [1].

any non-trivial information about his outcome. Conservatively, we can upper-bound Eve's information by $H(B|E,S_1) \leq H_{1,1}(B)$, where $H_{1,1}(B)$ is the entropy of Bob's outcome, averaged over the inputs $y \in \mathcal{K}$, when Bob performs ideal POVMs $M_y^{1,1}$.

Altogether Eve's information on Bob's outcome is then given by

$$\begin{aligned} H(B|E) &= pH(B|E,S_*) + (1-p)H(B|E,S_1) \\ &\leq (1-p)H_{1,1}(B)\,, \end{aligned} \tag{4.14}$$

hence the key rate is upper-bounded by

$$r_\infty \leq (1-p)H_{1,1}(B) - H_{\eta,v}(B|A)\,, \tag{4.15}$$

where p remains to be determined, and where $H_{\eta,v}(B|A)$ is computed from the actual correlations between Alice and Bob's outcomes, characterized by an efficiency η and visibility v.

This attack must reproduce on average Bob's actual POVMs $M_{b|y}^{\eta,v}$. Thus, it must satisfy

$$M_{b|y}^{\eta,v} = pM_{b|y}^{\eta_*,v_*} + (1-p)M_{b|y}^{1,1}, \tag{4.16}$$

which is equivalent to the following system of equations

$$\begin{aligned} p(1-\eta_*\, v_*) &= 1-\eta v \\ p\,(1-\eta_*) &= 1-\eta\,. \end{aligned} \tag{4.17}$$

We have seen in the previous sections, that the POVMs $M_y^{\eta^*,v^*}$ are $\mathcal{K}$-JM if η^* satisfies a bound $\eta^* \leq \hat{\eta}_{N,K,d}(v^*)$ which depends on the visibility v^*, the number of measurements N or K, the Hilbert space dimension d, and other properties of the measurements. Eve will pick the transmission value η^* which saturates this bound

$$\eta_* = \hat{\eta}_{N,K,d}(v_*). \tag{4.18}$$

Inserting this value for η_* in Eq. (4.17), we can then determine p (and v_*) as a function of Bob's actual CMU parameters η and v, and the obtained value of p can be used in (4.15) to get a bound on the key rate. We now illustrate this using several of the bounds obtained in the previous sections.

Arbitrary measurements in dimension d. Using the bound (4.10) and setting $\eta_* = \frac{d}{(K+1)(v_*(d+1)-1)}$ to the value saturating this bound, by solving Eq. (4.17) we find that $p = p'_{K,N,d}$ with

$$\begin{aligned} p'_{K,N,d}(\eta,v) =& \frac{\eta(1-v)+d(1-\eta v)}{d} \\ &\times \max\left\{\frac{K+1}{K}, \frac{N}{N-1}\right\}. \end{aligned} \tag{4.19}$$

Here we used the fact that if all the N measurements are used to extract the key one simply replaces $K+1$ with N. An interesting special case is the one where the key is extracted from an extremely large number of measurements. A bound on the key rate independent of N and K can be obtained from Eq. (4.19), by taking the limit of large K, N. We find

$$\lim_{K,N\to\infty} p'_{K,N,d}(\eta,v) = 1 - \frac{\eta(v(d+1)-1)}{d}, \tag{4.20}$$

and thus by Eq. (4.15) the key rate $r_\infty = 0$ is zero if $\frac{\eta(v(d+1)-1)}{d} \leq \frac{H_{\eta,v}(B|A)}{H_{1,1}(B)}$.

A few binary measurements in dimension 2. In the case $d = 2$, the value (4.19) based on the generic bound (4.10) can be used. However, in the specific case of qubit measurement of the form (2.36) and $K = N = 2$ or $K = N = 3$, it is more advantageous to use the bound (2.51). Setting η_* to saturate this bound we find that p is given by

$$p''_N(\eta, v) = \frac{N(1-\eta v) + \eta\sqrt{N}(1-v)}{N-1}\,. \tag{4.21}$$

Note that for any specific set of N qubit measurements, one can of course use a tighter *JM* condition based on Upper bound 2' and (2.46) to derive a better attack.

All binary measurements in dimension 2. Finally, as discussed in Section 2.3.3 the set of all qubit PVMs subject to noise and loss become *JM* for $\eta \leq 2(1-v)$. Setting $\eta_* = 2(1-v_*)$ in 4.17 we find p to be equal to

$$p'''(\eta, v) = 1 - \eta v + \sqrt{\eta(1-v)(2-\eta(1+v))}\,. \tag{4.22}$$

For the case of a qubit ($d = 2$) with projective measurements, which is the most studied in the literature, we have thus obtained three different expressions for p leading to different bounds on the key rate in Eq. (4.15). In the following table, we compare these expressions and indicate which one gives a better attack depending on the values N and K.

K \ N	2	3	≥ 4
1	Eq. (4.21)	Eq. (4.19)	Eq. (4.19)
2	Eq. (4.21)	Eq. (4.21)	Eq. (4.19)
3	x	Eq. (4.21)	Eqs. (4.19) or (4.22)
≥ 4	x	x	Eqs. (4.19) or (4.22)
$\to \infty$	x	x	Eq. (4.22)

Table 4.1: In the case of N binary measurements (ideally) on a qubit, K of which are used for key extraction, the table gives the optimal attack parameter p among the expressions $p'_{K,N,2}(\eta, v), p''_N(\eta, v)$ or $p'''(\eta, v)$ in Eqs. (4.19 ,4.21, 4.22) derived by using different JM bounds discussed in the previous section. In addition to the table, we find that for $v = 1$ the maximal value is always $p = (1-\eta)\max\left\{\frac{K+1}{K}, \frac{N}{N-1}\right\}$.

4.3.1 Application to BB84/CHSH-type protocols

We now illustrate the above approach on a family of qubit protocols where the key on Bob's side is based on the output $b \in \{0, 1\}$ of a single measurement, which we assume in the honest implementation to be a σ_z measurement. We also assume that in the key generation rounds, the state entering Bob's CMU is the σ_z eigenstate $|0\rangle$ or $|1\rangle$, depending on Alice's key variable $a = 0$ or $a = 1$. These are the only features of the protocol that we need to apply a bound on the key rate using our approach.

Protocols satisfying the above conditions are the one-sided-DI version of the original BB84 protocol, where the source is fully characterized [52], semi-DI versions of it, where Alice's device is only trusted to prepare qubit states [60], the semi-DI CHSH prepare-and-measure protocol

of [59], entanglement-based protocols such as the one-sided-DI protocols based on steering of [2, 50], fully DI CHSH-based protocols [4], or their routed version [118].

As $K = 1$ and $d = 2$, we have from (4.19), $p = 2 + \eta(1 - 3v)$, while $H_{1,1}(B) = 1$. To compute $H_\eta(B|A)$, we can consider two possibilities. Either non-detection events are treated as distinct outcomes, i.e., $b \in \{0, 1, \varnothing\}$. Or they are grouped with one of the other possible outcomes. In the first case, we have $H_{\eta,v}(B|A) \geq \eta\, h\left(\frac{1+v}{2}\right) + h(\eta)$, so that the key rate is upper-bounded as

$$r_\infty \leq \eta(3v - 1) - 1 - \eta\, h\left(\frac{1+v}{2}\right) - h(\eta)\,. \tag{4.23}$$

In the case where Bob's bin his no-click outcome with one of the other outcomes, we have $H(B|A) \geq \frac{1}{2}\left(h\left(\frac{\eta}{2}(1+v)\right) + h\left(\frac{\eta}{2}(1-v)\right)\right)$ and the key rate is upper-bounded as

$$r_\infty \leq \eta(3v - 1) - 1 - \frac{1}{2}\left(h\left(\frac{\eta}{2}(1+v)\right) + h\left(\frac{\eta}{2}(1-v)\right)\right)\,.$$

As an illustration, for a visibility $v = 1$, we find thresholds of $\eta = 83.0\%$ in the case where Bob does not bin and $\eta = 71.6\%$ in the case where he bins, and for a visibility $v = 0.97$, we find thresholds of $\eta = 87\%$ and $\eta = 76\%$, respectively.

4.3.2 Application to receiver-device-independent protocols

As a second application, we consider the class of so-called receiver device-independent (RDI) protocols [46, 47], which can be seen as an extension of the B92 protocol [31] to the case of many states and measurements. These are prepare-and-measure protocols as outlined in Fig. 2.2(a). Here, the measuring device is completely untrusted, while the preparing device is assumed to produce N pure states whose Gram matrix is fixed. We will consider the version of the protocol where Alice prepares qubit states

$$|\psi_x\rangle = \cos(\theta/2)\,|0\rangle + e^{\frac{i2\pi}{N}x}\sin(\theta/2)\,|1\rangle\,, \tag{4.24}$$

for some choice of $\theta \in [0, \frac{\pi}{2}]$, analysed in [46] in detail. Bob's ideal measurements are then $M_{b|y} = \left|\psi_y^\perp\right\rangle\!\left\langle\psi_y^\perp\right|, |\psi_y\rangle\langle\psi_y|$ for $b = 0, 1$ respectively, with the setting y chosen at random. After his measurement, Bob tells Alice to reject the round if $b \neq 0$. This allows him to unambiguously exclude the value $x = y$, since $\langle\psi_y|\psi_y^\perp\rangle = 0$, in all the rounds which are not rejected. Next, Alice reveals an unordered pair of values (x, x') where x is the value she encoded in the state and x' is a different value chosen at random. This pair is only accepted by Bob if $y \in \{x, x'\}$. Since there are only two possible values left, and Bob excludes one of them, Alice and Bob are left with a perfectly correlated bit after the sifting. This gives the idea of the protocol. We now explicitly compute the correlations observed by Alice and Bob.

The probability to observe the outcome $b = 0$ depends on the state $|\psi_x\rangle$ sent by Alice and is given by

$$p^{(\eta,v)}(0|x, y) = \eta\, v\left|\langle\psi_x|\psi_y^\perp\rangle\right|^2 + \frac{\eta(1-v)}{2} \tag{4.25}$$

where $|\langle\psi_x|\psi_y^\perp\rangle|^2 = \sin^2(\theta)\sin^2(\frac{\pi(x-y)}{N})$. A round is successful if $b = 0$ and $x = y$ (for any x') or $x' = y$ (for some $x \neq y$), leading to

$$p(\text{succ}|y) = \frac{1}{N}p(0|y, y) + \frac{1}{N(N-1)}\sum_{x\neq y} p(0|x, y). \tag{4.26}$$

Combining with Eq. (4.25) we find that the success probability is independent of y and given by

$$p^{(\eta,v)}(\text{succ}) = \frac{\eta v \sin^2(\theta)}{2(N-1)} + \frac{\eta(1-v)}{N} \tag{4.27}$$

Eve's strategy consists of replacing the CMU with a jointly measurable one, given by the parameters (η_*, v_*), with probability $p = p(\text{attack})$, or with the ideal one with probability $1-p$. In fact, the probability that the round is successful depends on her choice. The conditional probability of the attack reads

$$\bar{p} = p(\text{attack}|\text{succ}) = \frac{p^{(\eta_*,v_*)}(\text{succ})\, p}{p^{(\eta,v)}(\text{succ})}, \tag{4.28}$$

and implies $H(B|E,\text{succ}) = (1-\bar{p})$. With the help of Eqs. (4.17) we see that

$$p^{(\eta_*,v_*)}(\text{succ})\, p = \frac{p\eta_* v_* \sin^2(\theta)}{2(N-1)} + \frac{p\eta_*(1-v_*)}{N} \tag{4.29}$$

$$= p^{(\eta,v)}(\text{succ}) - \frac{(1-p)\sin^2(\theta)}{2(N-1)} \tag{4.30}$$

and thus $\bar{p} = 1 - (1-p)\frac{\frac{\sin^2(\theta)}{2(N-1)}}{p^{(\eta,v)}(\text{succ})}$.

When Eve replaces, with probability p, the CMU with a *JM* one, her parent POVM predicts, according to the strategies detailed in Section 2.3 that a no-click outcome $b = \{0,1\}$ will be obtained for Bob for only one of his possible measurements y. This implies that when Bob announces that he obtained the value $b = 0$, Eve can infer which input y was used by Bob and hence her entropy about Bob's raw key is 0. Analogously to Eq. (4.14), we thus have $H(B|E,\text{succ}) \leq (1-\bar{p})H_{1,1}(B|\text{succ}) \leq 1-\bar{p}$ or

$$H(B|E,\text{succ}) \leq (1-p)\frac{\frac{\sin^2(\theta)}{2(N-1)}}{p^{(\eta,v)}(\text{succ})}. \tag{4.31}$$

While this formula is affected by post-selection, we see that the optimal strategy for Eve is still the one maximizing p such that the noisy CMU is jointly measurable and we can use the formulas derived in the last section and summarized in Table. 4.1 (in our case $K = N$ as all of Bob's measurements are used for the key).

Finally, to compute the key rate, we also need the expression of $H(B|A,\text{succ})$. For a successful round where Alice prepared the state x and also revealed x', Bob must have chosen the measurements with $y = x$ or $y = x'$. For Alice the likelihood of these events is

$$p(y = x|A,\text{succ}) = \frac{p^{(\eta,v)}(0|x,x)}{p^{(\eta,v)}(0|x,x) + p^{(\eta,v)}(0|x,x')} \tag{4.32}$$

$$p(y = n|A,\text{succ}) = \frac{p^{(\eta,v)}(0|x,x')}{p^{(\eta,v)}(0|x,x) + p^{(\eta,v)}(0|x,x')} \tag{4.33}$$

The entropy of this distribution (as well as the success probability) depends on the pair $\{x,x'\}$. The entropy is minimal if $\sin^2(\pi\frac{x-x'}{N}) = 1$ in which case it is equal to the rhs of

$$H(B|A,\text{succ}) \geq h\left(\frac{1-v}{2(1-v\cos^2(\theta))}\right). \tag{4.34}$$

This gives a general upper bound on the key rate with one-way error correction

$$\begin{aligned} r &\leq p(\text{succ})\Big(H(B|E,\text{succ}) - H(B|A,\text{succ})\Big) \\ &\leq (1-p)\frac{\sin^2(\theta)}{2(N-1)} - p^{(\eta,v)}(\text{succ})h\left(\frac{1-v}{2(1-v\cos^2(\theta))}\right) \end{aligned} \tag{4.35}$$

To apply the bound we now take the limit of many measurements $N \to \infty$, where it was shown [47] that the key rate remains positive for arbitrary losses provided that the visibility is perfect $v = 1$ and $N > \frac{1}{\eta}$. Since the ideal measurements of Bob are qubit PVMs we can use the Eq. (4.22) for the strategy of Eve that maximizes $(1-p)$. We find that for $\eta = 10\%, 1\%, 0.1\%$ the key rate is zero if $v \leq 95\%, 99.5\%, 99.95\%$ respectively.

With the same approach tighter bounds can be derived on the RDI protocol by computing the entropy $H(B|A,\text{succ})$ explicitly, and deriving a tighter *JM* condition for the specific set of the CMUs stemming from the measurements $M_{b|y} = \{|\psi_y\rangle\langle\psi_y|, |\psi_y^\perp\rangle\langle\psi_y^\perp|\}$ specified by Eq. (4.24).

4.4 Convex combination attacks for DI QKD protocols

The convex combination attacks introduced in the previous section are asymmetric as they apply to generic QKD protocols where Bob holds a CMU, but where Alice may hold, e.g., a trusted preparation device. Fully DI QKD protocols, however, are symmetric in the sense that both parties hold a CMU. Exploiting this feature, we analyse in this section improved convex combination attacks for DI QKD protocols that lead to more stringent constraints on the key rate. Furthermore, again due to the symmetry of the protocol, they can be applied to situations where the public communication for key distillation goes from Bob to Alice, as in the previous section, or from Alice to Bob. For simplicity, we consider explicitly only the former case. But we note that in many cases the measurements and post-processing performed by Alice and Bob in DI QKD are essentially identical (up to e.g. local unitaries), implying that the same bounds on the key rate would be obtained by considering the public communication in the other direction.

Our convex combination attacks, in their simplest version, are a mixture of strategies where Eve replaces either Alice's or Bob's CMU with a full *JM* for a portion of the time, and allowing the devices to operate ideally for the remainder. In the former case, the correlations between Alice and Bob are local, while they are non-local in the latter case. This class of attacks thus fits in the same framework as those proposed in [119, 120] involving a convex combination of local and non-local correlations. Such local/non-local mixtures are obtained in [119, 120] through analytical or numerical (linear programming) results tailored to specific quantum correlations. However, the analytical results are limited to highly specific cases and linear programming methods become increasingly challenging when dealing with a large number of measurements and outputs. In contrast, the approach based on joint measurability is generic, as it depends only on rough properties of the quantum protocol, and it can provide bounds even in scenarios with many measurements. The downside is that it may lead to bounds on the key rate less stringent than the ones of [119, 120] when applied to the specific correlations for which their methods are tailored.

Note, however, that more generally, we will consider attacks where Eve replaces Bob's CMU with a $\mathcal{K}$-*JM* one with some probability. This gives rise to a class of attacks different from the one considered in [119, 120], since correlations arising from a $\mathcal{K}$-*JM* POVMs are not necessarily local, yet, are sufficient for Eve to have full information about Bob's key-generating outcomes. This may compensate for the fact that our joint measurability bounds are not tailored to specific correlations.

Let us now describe our attacks in more detail. Consider a generic DI QKD scenario consisting of a central source distributing bipartite states ρ, a CMU for Alice, and a CMU for Bob, as depicted in Fig. 2.2.b. For concreteness, we assume that both Alice and Bob have a white-noise no-click CMU characterized by the effective POVMs (2.9) with detector efficiency η and white-noise visibility v. Therefore, together they implement the joint effective POVMs $M_{a|x}^{\eta,v} \otimes M_{b|y}^{\eta,v}$ on the incoming state ρ. We further denote by N_A the number of measurements on Alice's side, and by N_B the overall number of measurements on Bob's side. Furthermore, let us assume that the key is only extracted from a subset of Bob's measurements with the inputs $y \in \mathcal{K}_B$.

Eve's attack is based on the following convex decomposition of Alice and Bob's joint POVMs

$$\begin{aligned} M_{a|x}^{\eta,v} \otimes M_{b|y}^{\eta,v} = p_A M_{a|x}^{\eta_A,v_A} \otimes M_{b|y}^{1,1} + p_B M_{a|x}^{1,1} \otimes M_{b|y}^{\eta_B,v_B} \\ + q M_{a|x}^{\eta',0} \otimes M_{b|y}^{\eta',0} \\ + (1 - p_A - p_B - q) M_{a|x}^{1,1} \otimes M_{b|y}^{1,1} , \end{aligned} \tag{4.36}$$

where $\eta' = \eta(1-v)/(1-\eta v)$, $q = (1-\eta v)^2$, and η_A, v_A, η_B, v_B, p_A, and p_B are free parameters that Eve can chose. Using the explicit form for the white-noise no-click CMU POVMs (2.9), it is readily verified that this decomposition is valid provided that these parameters satisfy the following system of equations

$$2\eta v(1-\eta v) = p_A(1-\eta_A v_A) + p_B(1-\eta_B v_B) \tag{4.37}$$

$$\eta^2(1-v)v = \eta_A p_A\,(1-v_A) = \eta_B p_B\,(1-v_B) \tag{4.38}$$

$$\eta v(1-\eta) = p_A\,(1-\eta_A) = p_B\,(1-\eta_B)\ . \tag{4.39}$$

The first constraint (4.37) is linearly redundant and can be omitted. The remaining ones are equivalent to

$$\eta v(1-\eta v) = p_A\,(1-\eta_A v_A) = p_B\,(1-\eta_B v_B) \tag{4.40}$$

$$\eta v(1-\eta) = p_A\,(1-\eta_A) = p_B\,(1-\eta_B)\ . \tag{4.41}$$

Here, one sees that the tuples of parameters (p_A, η_A, v_A) and (p_B, η_A, v_A) are subject to independent constraints, leaving one free parameter for each tuple.

The convex decomposition (4.36) can be exploited by Eve as follows. With probability $1 - p_A - p_B - q$, Alice's and Bob's CMUs behave ideally as they implement the ideal POVMs $M_{a|x}^{1,1} \otimes M_{b|y}^{1,1}$. In this case, we assume that Eve has no extra information about Bob's outcomes, i.e. her conditional entropy is only bounded by $H(B|E) \leq H_{1,1}(B)$, where $H_{1,1}(B)$ the entropy of Bob's outcomes averaged over the inputs $y \in \mathcal{K}_B$ in the ideal (noise-free and loss-free) situation. With probability q, the CMUs of Alice and Bob implement the POVMs $M_{a|x}^{\eta',0} \otimes M_{b|y}^{\eta',0}$, which are maximally noisy, i.e., the visibility is zero. In this case, the correlations generated by Alice and Bob are local. Since Eve can further decompose these correlations as a mixture of deterministic correlations, she has full information about Bob's outcomes, i.e., $H(B|E) = 0$. With probability p_A, Alice's CMU implements the POVMs $M_{a|x}^{\eta_A,v_A}$. If we choose η_A and v_A such that these POVMs are (full) JM, then, again, Alice and Bob's correlations are local, and thus $H(B|E) = 0$. Finally, with probability p_B, Bob's CMU implements the POVMs $M_{b|y}^{\eta_B,v_B}$. If we now choose η_B and v_B such that these POVMs are $\mathcal{K}_B$-JM, Eve has again full information about Bob's outcomes for the inputs $y \in \mathcal{K}_B$, i.e., $H(B|E) = 0$, since she can implement the parent instrument and predict all these outcomes. Note that in this last case, though, the correlations between Alice and Bob, which involve also the inputs $y \notin \mathcal{K}$, are not necessarily local. Summarizing, by our

attack the key rate is upper-bounded as

$$\begin{aligned} &r_\infty \leq tH_{1,1}(B) - H_{\eta,v}(B|A)\,, \\ &\quad \text{with} \\ &t = 1 - q - p_A - p_B = 2\eta v - \eta^2 v^2 - p_A - p_B\,. \end{aligned} \tag{4.42}$$

Hence to find the tightest upper-bound, we must maximize the probabilities p_A and p_B subject to the constraints that the POVMs $M_x^{\eta_A,v_A}$ are full *JM* and the POVMs $M_y^{\eta_B,v_B}$ are $\mathcal{K}$-*JM*. For this, it is optimal for Eve to set the efficiency $\eta_A = \hat{\eta}_{N_A,d}(v_A)$ saturating one of the bounds for *JM* that we have derived in the previous sections and similarly setting $\eta_B = \hat{\eta}_{N_B,K_B,d}(v_B)$ saturating one of the bounds for $\mathcal{K}_B$-*JM*. Then, the optimal values of p_A and p_B are given by the solutions of the Eqs. (4.40,4.41), i.e

$$\begin{aligned} \frac{p_A}{\eta v}\left(1 - \hat{\eta}_{N_A,d}(v_A)\, v_A\right) &= \frac{p_B}{\eta v}\left(1 - \hat{\eta}_{N_B,K_B,d}(v_B) v_B\right) \\ &= 1 - \eta v \\ \frac{p_A}{\eta v}\left(1 - \hat{\eta}_{N_A,d}(v_A)\right) &= \frac{p_B}{\eta v}\left(1 - \hat{\eta}_{N_B,K_B,d}(v_B)\right) \\ &= 1 - \eta. \end{aligned} \tag{4.43}$$

Remarkably, these equations become equivalent to those in (4.17) by the following parameter change $(p_{A(B)}, \eta_{A(B)}, v_{A(B)}) \to (\eta v\, p, \eta_*, v_*)$. It follows that the expressions $\eta\, v\, p'_{K,N,d}(\eta, v), \eta\, v\, p''_N(\eta, v)$ or $\eta\, v\, p'''(\eta, v)$ in Eqs. (4.19 ,4.21, 4.22) can be readily used to find $p_{A(B)}$ in different cases.

In particular, our bound (4.42) guarantees that the key rate is zero if one can choose the values p_A and p_B such that

$$p_A + p_B \geq 2\eta v - \eta^2 v^2 - \frac{H_{\eta,v}(B|A)}{H_{1,1}(B)} \implies r_\infty = 0. \tag{4.44}$$

Noticing that the term $\frac{H_{\eta,v}(B|A)}{H_{1,1}(B)}$ is always positive one obtains a weaker but correlation-independent condition for zero key rate

$$p_A + p_B \geq 2\eta v - \eta^2 v^2 \implies r_\infty = 0. \tag{4.45}$$

We now illustrate this bound on the three specific attack strategies already discussed in Sec. 4.3.

Arbitrary measurements in dimension *d*. Using the bounds (2.28) and (4.10) we find that

$$p_A = \eta\, v\, p'_{N_A-1,N_A,d}(\eta, v) \tag{4.46}$$

$$p_B = \eta\, v\, p'_{K_B,N_B,d}(\eta, v)\,, \tag{4.47}$$

with the function $p'_{K,N,d}(\eta, v)$ defined in equation (4.19). In particular, no key can be extracted (Eq. 4.45), whenever

$$\eta \leq \frac{(T_{N_A,K_B,N_B} - 2)\, d}{(T_{N_A,K_B,N_B} - 1)\, d\, v - T_{N_A,K_B,N_B}(1 - v)},$$

where we introduced $T_{N_A,K_B,N_B} = \frac{N_A}{N_A-1} + \min\left\{\frac{K_B+1}{K_B}, \frac{N_B}{N_B-1}\right\}$ to shorten the equation. This bound becomes

$$\eta \leq 1 - \frac{1}{T_{N_A,K_B,N_B} - 1} \tag{4.48}$$

for $v = 1$, and

$$v \leq \frac{T_{N_A,K_B,N_B} + d(T_{N_A,K_B,N_B} - 2)}{T_{N_A,K_B,N_B} + d(T_{N_A,K_B,N_B} - 1)} \tag{4.49}$$

for $\eta = 1$.

A few binary measurements on a qubit. When Alice and Bob use N_A, N_B (respectively) binary qubit measurements we can use the bound in Eq. (2.50) to obtain

$$p_{A(B)} = \eta\, v\, p''_{N_{A(B)}}(\eta, v), \tag{4.50}$$

with the function $p''_N(\eta, v)$ defined in Eq. (4.21). In particular, for $N_A = N_B = N$ we find that no key can be extracted (Eq. (4.45)) whenever

$$\eta \leq \frac{2(N+1)v + 4(1-v)\sqrt{N}}{(N-1)^2 v^2 + 4N(2v-1)},$$

which becomes $v \leq \frac{2}{\sqrt{N}+1}$ for $\eta = 1$.

All binary measurements on a qubit. Finally, in the same situation but with an unrestricted number of measurements for Alice and Bob, we can use the bound of Eq. (2.52) valid for the set of all qubit PVMs to obtain

$$p_{A(B)} = \eta\, v\, p'''(\eta, v), \tag{4.51}$$

with $p'''(\eta, v)$ defined in Eq. (4.22). In particular, we find that no key can be extracted (Eq. (4.45)) if

$$\eta \leq \frac{8(1-v)}{4 - 3v^2},$$

which becomes $v \leq \frac{2}{3}$ for $\eta = 1$.

Note that different bounds can be used on the sides of Alice and Bob. Furthermore, the attack we presented can be easily improved with additional knowledge about the measurements performed by the parties, as in this case one can use tighter (partial)-JM bounds specific to a given CMU.

Recall also that in the most common case where Alice and Bob perform qubit PVMs, the comparison of the three bounds $p'_{K,N,2}(\eta, v)$, $p''_N(\eta, v)$ and $p'''(\eta, v)$ was given in Table 4.1. It can be readily used to find the optimal attack of Eve in different cases. With its help, let us now explore the performance of popular DI QKD protocols in noisy conditions and determine the maximum achievable key rate. The challenge here is to compute the terms $H_{\eta,v}(B|A)$ and $H_{1,1}(B)$.

4.4.1 Application to qubit protocols without binning

Consider the protocols where Alice and Bob share a two-qubit state

$$|\Psi_\theta\rangle = \cos(\theta)\,|00\rangle + \sin(\theta)\,|11\rangle\,. \tag{4.52}$$

When the state is not maximally entangled Alice and Bob can only establish perfect correlation when they both measure in the Z basis. In this case, we thus assume that the key is only extracted from this measurement and set $K_B = 1$. The entropy of the key-generating outcome is then given by

$$H_{1,1}(B) = h\left(\cos^2(\theta)\right) \tag{4.53}$$

for the ideal implementation. The conditional entropy term $H_{\eta,v}(B|A)$ is always larger than (equal for optimally aligned measurements)

$$H_{\eta,v}(B|A) = \eta(1-\eta) + h(\eta) + \eta^2\, h\left(\frac{1+v^2}{2}\right), \tag{4.54}$$

where $h(x)$ is the binary entropy.

Plugging these expressions into Eq. (4.42) gives

$$r_\infty \leq h\big(\cos^2(\theta)\big)(2\eta v - p_A - p_B - \eta^2 v^2) - H_{\eta,v}(B|A). \tag{4.55}$$

One notices that $h(\cos^2(\theta))$ is the only term depending on the parameter θ, and it is straightforward to see that the optimal protocol (where the key rate bound is the less stringent) consists of preparing the maximally entangled state. We can thus set $\theta = \frac{\pi}{4}$ for which $H_{1,1}(B) = 1$. To find the best bound it remains to maximize $p_A + p_B$ with the guidance of Table 4.1 and the functions $p'_{K,N,d}(\eta, v), p''_N(\eta, v)$ or $p'''(\eta, v)$ defined in Eqs. (4.19 ,4.21, 4.22). To do so we consider three situations.

(i) $(N_A, N_B, K_B) = (3, 2, 1)$ is the setting of the DI QKD CHSH protocol introduced in [4, 121]. Here the optimal bound reads

$$r_\infty \leq\ \eta v\big(2 - p''_3(\eta, v) - p''_2(\eta, v)\big) - \eta^2 v^2 - H_{\eta,v}(B|A). \tag{4.56}$$

(ii) $(N_A, N_B, K_B) = (\infty, \infty, 1)$ corresponds to the setting where only one measurement is used for key generation, but the number of test measurements is not restricted. Here the bounds

$$r_\infty \leq\ \eta v\big(2 - p'_{1,\infty,2}(\eta, v) - p'''(\eta, v)\big) - \eta^2 v^2 - H_{\eta,v}(B|A), \tag{4.57}$$

are to be used depending on the values of η and v.

(iii) $(N_A, N_B, K_B) = (\infty, \infty, \infty)$ is the scenario with no restriction on the number of measurements. This is a meaningful scenario as for the maximally entangled state Alice and Bob can have perfect correlation for any number of measurements. Here the bound

$$r_\infty \leq\ \eta v\big(2 - 2\, p'''(\eta, v)\big) - \eta^2 v^2 - H_{\eta,v}(B|A) \tag{4.58}$$

is optimal.

In Fig. 4.2 we plot the lines in the (η, v) plane below which the bounds of Eqs. (4.56-4.58) guarantee a zero key rate. The threshold values of η and v (for $v, \eta = 1$ respectively) are reported in Table. 4.2

Setting (N_A, N_B, K_B)	Threshold η (at $v = 1$)	Threshold v (at $\eta = 1$)
(3,2,1)	88.3 %	89.8 %
(∞,∞,1)	87.4 %	88.8 %
(∞,∞,∞)	85.3 %	87.1 %

Table 4.2: Threshold efficiency η and visibility v for varions number of measurements (N_A, N_B, K_B). This applies to DI QKD protocols with the maximally entangled two-qubit state, dichotomic measurements (ideally), no-binning of the no-click outcome, and one-way public communication from Bob to Alice.

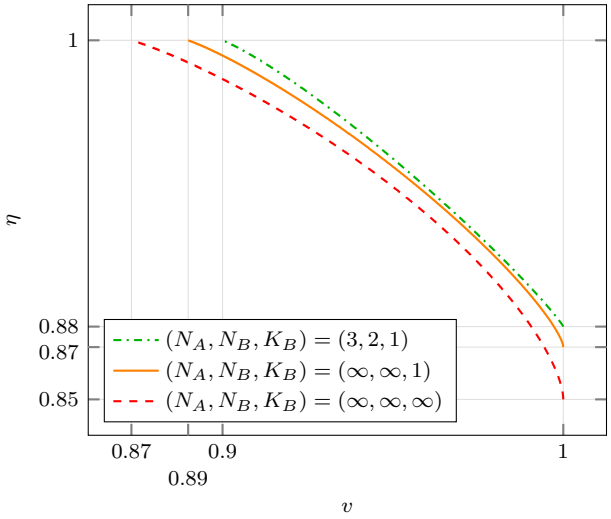

Figure 4.2: For the parameters (η, v) below the lines the key rate r_∞ is guaranteed to be zero. The lines show when the right-hand sides of the corresponding equation fall to zero. (From top to bottom) for Eq. (4.56) : *(i)* $(N_A, N_B, K_B) = (3, 2, 1)$ dash-dotted line , Eq. (4.57) : *(ii)* $(N_A, N_B, K_B) = (\infty, \infty, 1)$ full line, and Eq. (4.58) : *(iii)* $(N_A, N_B, K_B) = (\infty, \infty, \infty)$ dashed line. This applies to DI QKD protocols based on two-qubit states, dichotomic measurements (ideally), no-binning of the no-click outcome, and one-way public communication from Bob to Alice.

4.4.2 Application to qubit protocols with binning

A common strategy in the DI QKD protocols is to bin non-click outcomes of Bob with the most likely among the other two outcomes. After such binning the key rate bound in the Eq. (4.55) remains valid upon replacing the conditional entropy term with

$$H^{\theta}_{\eta,v}(B|A) = (1-\eta)\, h\left(\frac{\eta}{2}(1 - v\cos(2\theta))\right) + \frac{\eta}{2}(1 - v\cos(2\theta))\; h\left(\eta\left(1 - \frac{1-v^2}{2(1-v\cos(2\theta))}\right)\right) + \frac{\eta}{2}(1 + v\cos(2\theta))\; h\left(\eta\frac{1-v^2}{2(1+v\cos(2\theta))}\right). \quad (4.59)$$

Setting (N_A, N_B, K_B)	Threshold η (at $v = 1$)	Threshold v (at $\eta = 1$)
(3,2,1)	72.7 %	89.8 %
(∞,∞,1)	68.3 %	88.8 %
(∞,∞,∞)	74.2 %	87.1 %

Table 4.3: Threshold efficiency η and visibility v for different number of measurements (N_A, N_B, K_B). This applies to DI QKD protocols with the partially entangled two-qubit state, dichotomic measurements (ideally), binning of the no-click outcomes, and one-way public communication from Bob to Alice.

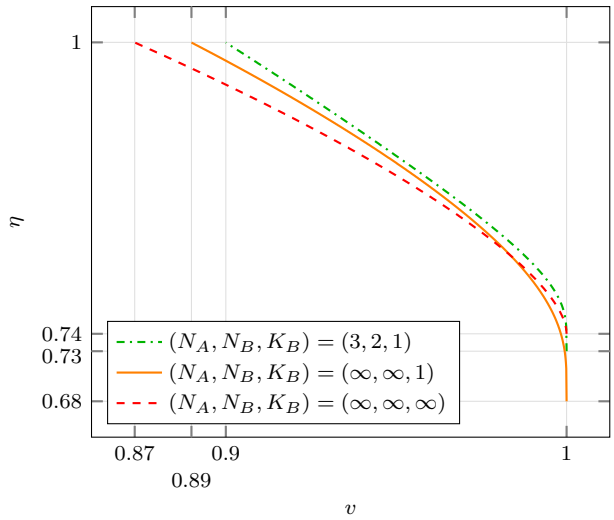

Figure 4.3: For the parameters (η, v) below the lines the key rate r_∞ was found to be zero after a maximization over θ. The lines show when the right-hand sides of the corresponding equation fall to zero. (From top to bottom) for Eq. (4.60) : *(i')* $(N_A, N_B, K_B) = (3, 2, 1)$ dash-dotted line , and Eq. (4.61) : *(ii')* $(N_A, N_B, K_B) = (\infty, \infty, 1)$ full line. For the curve, Eq. (4.62) : *(iii')* $(N_A, N_B, K_B) = (\infty, \infty, \infty)$ dashed line, no maximization over θ is needed as it only makes sense to consider the maximally entangled state $\theta = \frac{\pi}{4}$. This applies to DI QKD protocols based on two-qubit states, dichotomic measurements (ideally), binning of the no-click outcomes, and one-way public communication from Bob to Alice.

Note that in general this reduces Alice's entropy, i.e., $H^{\theta}_{\eta,v}(B|A)$ here above is lower or equal to $H_{\eta,v}(B|A)$ in Eq. (4.54) since the binning channel applied by Bob is two-to-one. It follows that the key rate bounds that we obtain using binning are less stringent than without. With the new expression for $H^{\theta}_{\eta,v}(B|A)$ let us consider the following settings.

(i') For $(N_A, N_B, K_B) = (3, 2, 1)$ the optimal bound reads

$$r_\infty \leq h\big(\cos^2(\theta)\big)\big(\eta v(2 - p''_3(\eta, v) - p''_2(\eta, v)) - \eta^2 v^2\big) - H^{\theta}_{\eta,v}(B|A). \tag{4.60}$$

(ii') For $(N_A, N_B, K_B) = (\infty, \infty, 1)$ the optimal bound reads

$$r_\infty \leq h\big(\cos^2(\theta)\big)\big(\eta v(2 - p'_{1,\infty,2}(\eta, v) - p'''(\eta, v)) - \eta^2 v^2\big) - H^{\theta}_{\eta,v}(B|A) \tag{4.61}$$

(iii') For $(N_A, N_B, K_B) = (\infty, \infty, \infty)$ we only consider the case of the maximally entangled state $\theta = \frac{\pi}{4}$, as it is the only one where Alice and Bob can have perfect correlations in more than one measurement basis. We then get the optimal bound

$$r_\infty \leq \eta v\big(2 - 2\,p'''(\eta, v)\big) - \eta^2 v^2 - H^{\frac{\pi}{4}}_{\eta,v}(B|A). \tag{4.62}$$

To derive state-independent upper-bounds one needs to maximize the right-hand side of the above expressions with respect to θ. Here, both quantities $H_{1,1}(B)$ and $H^{\theta}_{\eta,v}(B|A)$ have a nontrivial dependence on the angle θ, and this maximization has to be done for all η and v. In Fig. 4.3, we display the thresholds that we found after maximizing the key rate over the angle θ in a heuristical way. Specifically, we used the Nelder–Mead technique [122]. In Table. 4.3 we give the corresponding threshold values.

4.5 Proof of Lemma 8

In this section, we prove Lemma 8 from Section 4.3. But first, we recall some elements of the formalism of Gaussian channels, states, and their phase-space representation.

The Wigner function of a single-mode Gaussian state with the covariance matrix V and displacement vector $\boldsymbol{\mu}$ is given by

$$W(\alpha) = \frac{1}{\sqrt{\det(\pi V/2)}} \exp\left(-2(\boldsymbol{\alpha}-\boldsymbol{\mu})^T V^{-1}(\boldsymbol{\alpha}-\boldsymbol{\mu})\right). \tag{4.63}$$

with $\boldsymbol{\alpha} = \binom{\mathrm{Re}\,\alpha}{\mathrm{Im}\,\alpha}$. By definition, the pair $V, \boldsymbol{\mu}$ thus uniquely specifies a Gaussian state. In particular, for a coherent state $|\gamma\rangle$ we have $\boldsymbol{\mu} = \boldsymbol{\gamma} = \binom{\mathrm{Re}\,\gamma}{\mathrm{Im}\,\gamma}$ and $V = \mathbb{1}$.

A single mode Gaussian channel Φ maps Gaussian states to Gaussian states. It can be represented by matrices X, Y and a vector $\boldsymbol{\delta}$, which transform the Wigner function as follows [28, 86]

$$\boldsymbol{\mu} \mapsto X\boldsymbol{\mu} + \boldsymbol{\delta} \tag{4.64}$$

$$V \mapsto XVX^T + Y \tag{4.65}$$

Therefore, if a Gaussian channel acts on a coherent state, the resulting Wigner function of state $\Phi(|\gamma\rangle\langle\gamma|)$ is Gaussian and characterized by the pair $V' = XX^T + Y$ and $\boldsymbol{\mu}' = X\boldsymbol{\gamma} + \boldsymbol{\delta}$.

Finally, let us compute the Q-function corresponding to this state. It is well known that it is the convolution of the Wigner function with a Gaussian

$$Q(\alpha) = \frac{2}{\pi}\int \mathrm{d}^2\beta\, W(\beta) e^{-2|\alpha-\beta|^2}. \tag{4.66}$$

Hence the Q-function is also a Gaussian with the displacement vector $\boldsymbol{\mu}_Q = \boldsymbol{\mu}'$ and covariance matrix $V_Q = V' + \mathbb{1}$ (note that a convolution of two distributions describes a random variable given by the sum of the random variables described by the distributions). Thus if a Gaussian channel (X, Y, δ) acts on a coherent state $|\gamma\rangle$, the Q-function of the finals state is Gaussian with

$$\boldsymbol{\mu}_Q = X\boldsymbol{\gamma} + \boldsymbol{\delta} \tag{4.67}$$

$$V_Q = XX^T + Y + \mathbb{1}. \tag{4.68}$$

Now let us come back to the Lemma that we want to prove.

Lemma 8. *A thermal-noise channel $\Phi_{\eta,\epsilon}$ that is not N-extendable cannot be decomposed as*

$$\Phi_{\eta,\epsilon} = p\,\Phi_G + (1-p)\Phi' \qquad (0 < p < 1) \tag{4.69}$$

where Φ_G is a N-extendable Gaussian channel and Φ' is an arbitrary channel.

Proof. To prove this result we use the following Lemma, which is demonstrated below.

Lemma 9. *The thermal-noise channel $\Phi_{\eta,\epsilon}$ can admit a decomposition*

$$\Phi_{\eta,\epsilon} = p\,\Phi_G + (1-p)\Phi' \tag{4.70}$$

with a Gaussian channel Φ_G characterized by the matrices X and Y (see e.g.[28]), a positive trace preserving map Φ', and some $p > 0$, only if

$$X = \sqrt{\eta}\mathbb{1} \qquad Y \leq (1 - \eta + \eta\varepsilon)\mathbb{1}. \tag{4.71}$$

Let us now take a thermal channel $\Phi_{\eta,\epsilon}$ which is not N-extendable, that is with

$$\eta(1-\epsilon/2) > \frac{1}{N} \tag{4.72}$$

accordingly to [86] (note that on the level of the channel, this is a necessary and sufficient condition). Consider any Gaussian channel Φ_G appearing in the decomposition of Eq. (4.70). On the one hand we know that it is N-extendable if and only if [86]

$$\sqrt{\det Y} \geq 1 - \frac{1}{N} + \left|\det X - \frac{1}{N}\right|. \tag{4.73}$$

On the other, Lemma 1 implies that $\sqrt{\det Y} \leq (1 - \eta + \eta\epsilon)$ and $\det X = \eta$ must hold. By using these identities together with the bound in Eq. (4.72), it is easy to see that the condition of Eq. (4.73) can never be satisfied. Hence, the decomposition in Eq. (4.70) is impossible. □

Proof of Lemma 9. First note that the case $\eta = 0$ is trivial, we thus assume $\eta > 0$.

Consider the map $\tilde{\Phi} = (1-p)\Phi' = \Phi_{\eta,\epsilon} - p\Phi_G$, which must be at least positive for the decomposition to make physical sense. Let us act with $\tilde{\Phi}$ on a coherent state $|\gamma\rangle$, the Q-function of final state $\tilde{\Phi}(|\gamma\rangle\langle\gamma|)$ is given by

$$\tilde{Q}_\gamma(\alpha) = Q^{\eta,\epsilon}_\gamma(\alpha) - p\,Q^G_\gamma(\alpha), \tag{4.74}$$

where $Q^{\eta,\epsilon}_\gamma$ and Q^G_γ are the Q-functions associated to the states $\Phi_{\eta,\epsilon}(|\gamma\rangle\langle\gamma|)$ and $\Phi_G(|\gamma\rangle\langle\gamma|)$. The Q-function of a state is proportional to the probability density of the output of the heterodyne measurement. Therefore, if $\tilde{\Phi}$ is a positive channel, the function $\tilde{Q}_\gamma(\alpha)$ must be positive for all α and γ, i.e. we must have

$$Q^{\eta,\epsilon}_\gamma(\alpha) \geq p\,Q^G_\gamma(\alpha) \tag{4.75}$$

for some nonzero p and all α and γ. This is equivalent to requiring that the ratio $\frac{Q^G_\gamma(\alpha)}{Q^{\eta,\epsilon}_\gamma(\alpha)}$ is bounded. Since both distributions are Gaussian this is equivalent to the following bound

$$(\boldsymbol{\alpha} - \boldsymbol{\mu}_{th})^T V_{th}^{-1}(\boldsymbol{\alpha} - \boldsymbol{\mu}_{th})^T - (\boldsymbol{\alpha} - \boldsymbol{\mu}_G)^T V_G^{-1}(\boldsymbol{\alpha} - \boldsymbol{\mu}_G) < \infty, \tag{4.76}$$

where for the thermal channel $\Phi_{\eta,\epsilon}$ we have $\boldsymbol{\mu}_{th} = \sqrt{\eta}\boldsymbol{\gamma}$ and $V_{th} = (2 + \epsilon\eta)\mathbb{1}$, while for a general Gaussian channel $V_G = XX^T + Y + \mathbb{1}$ and $\boldsymbol{\mu}_G = X\boldsymbol{\gamma} + \boldsymbol{\delta}$ accordingly to Eqs. (4.67,4.68). This condition is only nontrivial in the limit where $|\alpha|, |\gamma|$ or both go to infinity. In this limit, $\boldsymbol{\delta}$ plays no role, and we can thus ignore it.

Plugging the values of $\boldsymbol{\mu}_{th}$, V_{th} and $\boldsymbol{\mu}_G$ in Eq. (4.76) we obtain

$$\frac{|\alpha - \sqrt{\eta}\gamma|^2}{2 + \eta\epsilon} - (\boldsymbol{\alpha} - X\boldsymbol{\gamma})^T (V_G)^{-1}(\boldsymbol{\alpha} - X\boldsymbol{\gamma}) \leq \infty. \tag{4.77}$$

This is a bilinear form in the varibles $\boldsymbol{\xi} = (\mathrm{Re}\alpha, \mathrm{Im}\alpha, \mathrm{Re}\gamma, \mathrm{Im}\gamma)^T$ and can be expressed as

$$\boldsymbol{\xi}^T M \boldsymbol{\xi} \leq \infty \tag{4.78}$$

$$M = \left(\begin{array}{c|c} \frac{\mathbb{1}}{2+\eta\epsilon} - V_G^{-1} & -\frac{\sqrt{\eta}\mathbb{1}}{2+\eta\epsilon} + X^T V_G^{-1} \\ \hline -\frac{\sqrt{\eta}\mathbb{1}}{2+\eta\epsilon} + V_G^{-1} X & \frac{\eta\mathbb{1}}{2+\eta\epsilon} + X^T V_G^{-1} X \end{array} \right), \tag{4.79}$$

or simply $M \leq 0$. The positivity of this matrix M remains unchanged if we multiply it by the diagonal matrix

$$\mathrm{diag}(\sqrt{2+\eta\epsilon}, \sqrt{2+\eta\epsilon}, -\frac{\sqrt{2+\eta\epsilon}}{\sqrt{\eta}}, -\frac{\sqrt{2+\eta\epsilon}}{\sqrt{\eta)}}) \tag{4.80}$$

from the left and from the right. This allows us to rewrite the condition as

$$\left(\begin{array}{c|c} \mathbb{1} - W^{-1} & \mathbb{1} - Z^T W^{-1} \\ \hline \mathbb{1} - W^{-1} Z & \mathbb{1} - Z^T W^{-1} Z \end{array} \right) \leq 0 \tag{4.81}$$

or

$$\mathbb{1}_4 \leq \left(\begin{array}{c|c} W^{-1} & Z^T W^{-1} \\ \hline W^{-1} Z & Z^T W^{-1} Z \end{array} \right) \tag{4.82}$$

for $W^{-1} = (2+\eta\epsilon) V_G^{-1}$ and $Z = X/\sqrt{\eta}$.

We now consider two cases. First, assume that the matrix X (and hence Z) is invertible. Multiplying the Eq. (4.82) with $\mathrm{diag}(\mathbb{1}|Z^{-1})$ from the right and $\mathrm{diag}(\mathbb{1}|(Z^{-1})^T)$ from the left does not change the positivity of a matrix, and allows us to cast the inequality in the form

$$\left(\begin{array}{c|c} \mathbb{1} & (Z^{-1})^T \\ \hline Z^{-1} & (Z^{-1})^T Z \end{array} \right) \leq \left(\begin{array}{c|c} W^{-1} & W^{-1} \\ \hline W^{-1} & W^{-1} \end{array} \right). \tag{4.83}$$

The right hand side is block-diagonal $\binom{1}{1}\binom{1}{1}^T \otimes W^{-1}$. And since the inequality must hold in both blocks, we get

$$(\mathbb{1} + {Z^{-1}}^T)(\mathbb{1} + Z^{-1}) \leq 4W^{-1} \tag{4.84}$$

$$(\mathbb{1} - {Z^{-1}}^T)(\mathbb{1} - Z^{-1}) \leq 0. \tag{4.85}$$

The second inequality implies $\mathbb{1} - Z^{-1} = 0$, or $X = \sqrt{\eta}\mathbb{1}$. This reduces the fist one to $W^{-1} \geq \mathbb{1}$, or $W \leq \mathbb{1}$. Recalling that $W = V_G/(2+\eta\epsilon)$ with $V_G = XX^T + Y + 1 = \eta\mathbb{1} + Y + 1$ we obtain

$$\frac{(1+\eta)\mathbb{1} + Y}{2+\eta\epsilon} \leq \mathbb{1}, \tag{4.86}$$

leading to

$$Y \leq (1 - \eta + \eta\epsilon)\mathbb{1}. \tag{4.87}$$

This proves the Lemma for an invertible X.

Next, consider the remaining case where X is not invertible. There is thus a normalized vector $\boldsymbol{v}$ such that $X\boldsymbol{v} = Z\boldsymbol{v} = 0$. Multiplying the Eq (4.82) with $(0,0|\boldsymbol{v})$ from the right and $(0,0|\boldsymbol{v})^T$ from the left we find

$$1 \leq 0. \tag{4.88}$$

X must thus be invertible, which concludes the proof. □

Note that the Lemma is most probably also true with an if and only if statement. The if direction could be shown by an explicit decomposition, noting that the thermal-noise channel can be viewed as a loss channel followed by a random displacement with a normally distributed complex amplitude.

4.6 Discussion

We applied the tools presented in Chapter 2 to derive upper bounds on the key rates of DI or semi-DI QKD protocols. In the context of QKD, the notion of joint measurability of a CMU leads to necessary criteria for extracting a secure key. The analysis of QKD protocols also led us to introduce the concept of partial joint measurability. We showed that this notion differs from standard joint measurability, in the sense that there exist sets of measurements that are partially jointly measurable but not (fully) jointly measurable. It would be interesting to investigate this concept further, in particular how to cast partial joint measurability as a semi-definite program. We note that this concept as we defined it differs from the notions discussed in previous works, where sets of POVMs that are incompatible can feature subsets of POVMs that are jointly measurable [75, 123, 124]. In contrast, our definition is closely related to the notion of no-exclusivity introduced recently for quantum instruments [125], and to the notion of weak-compatibility introduced earlier in the same context [126]. In particular, the notion of partial joint-measurability that we introduced for a channel followed by a set of POVMs can be generalised to an analogous notion of partial no-exclusivity for a channel followed by a set of instruments. Lemma 7 then can naturally be extended to this setting.

Finally, the intuition behind the notion of partial joint measurability can be used to improve the class of attacks on DI QKD considered in [119, 120]. These attacks are based on decompositions of the correlations observed by Alice and Bob $p(a,b|x,y) = q\, p_L(a,b|x,y) + (1-1)p_{NL}(a,b|x,y)$ into a mixture of local $p_L(a,b|x,y)$ and non-local $p_{NL}(a,b|x,y)$ correlations. In fact, the locality of correlation $p_L(a,b|x,y)$ is sufficient but not necessary for Eve to predict the key. Since she only needs to predict the outcomes of the key-generating measurements, we introduced here attacks where $p(a,b|x,y)$ is decomposed into a mixture of non-local quantum correlations and partly-*JM* $p_{JM}(a,b|x,y)$ correlations. The partly-*JM* correlations can be associated with a quantum-classical state, where only the classical register is used to simulate the outputs of the key-generating measurements, whereas for the other measurements the quantum register is used, hence can display non-locality.

5 Lower bounds on the key rate of fully DI QKD protocols

In this chapter, we will provide an overview of methods used to provide lower bounds on the asymptotic key rate of DI QKD protocols and, thus, prove the security of DI QKD protocols. There are two main types of approaches used to carry out this task, the analytical approaches and the numerical ones. We will focus mainly on the analytical case which was studied in [1]. After introducing our analytical approach, we will apply it to compute the key rates of different versions of DI QKD protocols as a function of different noise parameters. At the end of this chapter, we will also outline two types of numerical methods to compute lower bounds which will be used in the following chapters. The content of this chapter up to Section 5.3 included is reproduced from [1].

5.1 Introduction

Realizing a working DI QKD protocol has long presented a significant challenge both to theorists, due to the mathematical difficulty of devising practical and rigorous security proofs, and to experimental researchers, due to the difficulty of distributing entangled quantum systems with low noise and high detection rates over long distances. Recent advances paved the way to three successful proof-of-principle experiments demonstrating the feasibility of this technology [23–25]. However, there is still a long way from these proof-of-principle experiments to practical DI QKD implementations, with the necessity to improve the distance and the rate at which the keys are distributed.

As we explained in Chapter 3, one major theoretical advance introduced a few years ago is the entropy accumulation theorem [127], and the related technique of quantum probability estimation [128], which reduces proving the unconditional security of a generic DI QKD protocol in the finite-key regime to the problem of obtaining a lower bound (referred to as *min-tradeoff function*) on the conditional von Neumann entropy $H(K_A|E)$ of Alice's raw key variable K_A conditioned on an eavesdropper's possible quantum side information E, as a function of the expected value of a Bell expression or of the input-output probability distribution of our measurements. For instance the security of the simplest DI QKD protocol based on the CHSH inequality follows from the following lower bound on the conditional von Neumann entropy of Alice's measurement outcome A_1

$$H(A_1|E) \geq 1 - \phi\big(\sqrt{S^2/4 - 1}\big)\,, \tag{5.1}$$

where $\phi(x) = h\big(\frac{1}{2} + \frac{1}{2}x\big)$, $h(x)$ is the binary entropy, and $S = \langle A_1B_1\rangle + \langle A_1B_2\rangle + \langle A_2B_1\rangle - \langle A_2B_2\rangle$ is the expected value of the CHSH Bell expression [4].

The basic CHSH protocol based on the above lower bound is, however, not optimal in a number of respects. There has thus been in the last few years a search for ways to bound the conditional entropy for more general DI QKD protocols, either focusing on the 2-input/2-output setting [100, 129, 130], or finding numerical methods to tackle the problem in a more general way [131, 132]. Despite these efforts, bounding the entropy can be a numerically-intensive problem, with one recent approach [100] notably requiring thousands of processor core-hours of computing time to numerically bound the average entropy for a two-basis variant [130] of the CHSH-based

DI QKD protocol. This has significant drawbacks, reducing confidence in the results (as they are harder for others to reproduce), increasing the difficulty to optimize over parameters in simulations, and generally increasing the time and computing resources necessary just to calculate a key rate.

In this chapter, we will proceed by introducing three approaches for computing the key rates. First, we present in Section 5.2 a semi-analytic approach to bound the conditional entropy in the 2-input/2-output device-independent setting that is conceptually and technically relatively simple. This is the approach introduced in [1]. It is a generalization of the approach in [133] that was used to derive an analytical bound on the conditional entropy for a family of asymmetric CHSH inequalities. As we explain here, the main conceptual steps of this security analysis are not specific to the protocol considered in [133] but can actually be easily adapted to other 2-input/2-output device-independent protocols. Second, we present in Section 5.4 a popular numerical approach, which consists of bounding the probability for Eve to guess correctly the secret key. This approach is extremely versatile and not limited to 2-input/2-output settings, but it is not able to provide tight lower bounds on the key rate. Finally, we will describe in Section 5.5 a very recent numerical method introduced in [132] which is extremely versatile and can provide the tightest lower bounds on the key rate which can be currently computed.

5.2 Semi-analytic DI QKD security analysis via BB84-type uncertainty relations and Pauli correlation constraints

Let us start presenting our semi-analytic approach. As usual in the 2-input/2-output scenario, the starting point is to use Jordan's lemma to reduce the analysis to convex combinations of qubit strategies. From there, our approach is based on three steps. First, as in a standard qubit QKD protocol like BB84, we bound the conditional entropy of Alice's key generating measurement, say, A_1 through an uncertainty relation involving the correlations $\langle \bar{A}_1 \otimes B \rangle$ between an *orthogonal* measurement $\bar{A}_1$ on Alice's subsystem and a binary observable B on Bob's system. In a device-independent setting, though, and contrarily to, e.g., BB84, we cannot have direct access to the correlations $\langle \bar{A}_1 \otimes B \rangle$ as we cannot assume that Alice's measurement devices perform measurements in two orthogonal bases A_1, $\bar{A}_1$. The second step is then to establish a device-independent qubit constraint on $\langle \bar{A}_1 \otimes B \rangle$ which is based on correlations between Alice and Bob that are actually observed in the protocol, e.g., the CHSH expectation value or some other Bell expression. Combining the first and second step, we obtain a bound on the conditional entropy which is device-independent, apart from the assumptions that Alice and Bob are measuring qubits. The third step then involves a convexity analysis: either the resulting bound happens to be convex or, if this is not the case, we convexify it. In this way, we get a lower bound that is valid for convex combination of qubit strategies, and thus by Jordan's lemma, for arbitrary, dimension-free strategies.

We illustrate this approach in detail on two variants of the CHSH-based DI QKD protocol: the two-basis variant [130] and a new variant that incorporates, in addition to the CHSH value, information about the bias in the key generating measurement A_1. This last feature is particularly relevant for photonic implementations of DI QKD where no-click outcomes $\emptyset$ are mapped to a given key bit value, say $\emptyset \mapsto +1$, resulting in highly biased outcomes. The bounds that we obtain are optimal or close to optimal and significantly simpler technically and less computationally demanding than other approaches. We show in particular that a qubit DI QKD protocol can tolerate detector efficiencies as low as 80.26%.

In the following, we provide a high-level description of our approach to bounding the conditional entropy in 2-input/2-output scenarios.

5.2.1 Description of the approach

We start by specifying the class of problems that we aim to solve. We consider a tripartite setup involving a state ρ_{ABE} shared among Alice, Bob, and the eavesdropper Eve. We assume that Alice can measure one of two ± 1-valued observables A_1 or A_2 on her system, and similarly Bob can measure one of two ± 1-valued observables B_1 or B_2. We refer to the tuple $\mathcal{Q} \equiv (\rho_{ABE}, A_1, A_2, B_1, B_2)$ as a *strategy*.

A strategy $\mathcal{Q}$ can be seen as describing a single round of a multi-round DI QKD protocol. The measurements by Alice and Bob serve two purposes: generating some random variable K_A on Alice's side (which will constitute Alice's copy of the *raw key* in the DI QKD protocol) and establishing some correlations between Alice and Bob (which will be estimated in a *parameter estimation* step of the DI QKD protocol). Any strategy $\mathcal{Q}$ implies some tradeoff between how random K_A is to Eve and how correlated Alice's and Bob's measurement outcomes are. This tradeoff can be formalized as follows.

Eve's information on the raw key K_A. Let us assume that Alice uses the following general procedure to generate a random key value K_A: she first selects a measurement choice $X = 1, 2$ according to a probability distribution μ_X, she measures the corresponding observable A_1 or A_2, she gets the classical output $A = \pm 1$, and finally she applies to A a (possibly stochastic) map $\$_x : \{\pm 1\} \to \mathcal{K}_A : A \mapsto K_A$ to obtain a value K_A in some finite alphabet $\mathcal{K}_A$. A measure of how random K_A is to Eve, given knowledge of the measurement choice X, is the conditional von Neumann entropy

$$H(K_A|XE) = H(\rho_{K_A XE}) - H(\rho_{XE}) \tag{5.2}$$

where $H(\rho) = -\operatorname{Tr}(\rho \log_2(\rho))$ is the Von Neumann entropy and $\rho_{XE} = \operatorname{Tr}_{K_A}(\rho_{K_A XE})$ where

$$\rho_{K_A XE} = \sum_{k_A, x} \mu(x) \left|k_A, x\right\rangle \left\langle k_A, x\right| \otimes \rho_E^{k_A, x} \tag{5.3}$$

is the classical-quantum state describing the correlations between K_A, X, and E. In the above expression, the reduced states of Eve are given by

$$\rho_E^{k_A, x} = \sum_{a=\pm 1} p_x(k_A|a) \operatorname{Tr}_{AB}\left[\rho_{ABE} \, \frac{\mathbb{1} + aA_x}{2} \otimes \mathbb{1}_B \otimes \mathbb{1}_E\right] \tag{5.4}$$

where $p_x(k|a)$ are the transition probabilities of the map $\$_x$.

In this paper, we will often be interested in the case where K_A is simply obtained as the outcome of one of Alice's measurement, e.g., A_1 (i.e., there is no random input choice X and no classical preprocessing). By a slight abuse of notation, we write A_1 both for the random variable denoting the measurement outcome of A_1 and for the measurement A_1 itself. We thus write in such cases $K_A = A_1$ and $H(K_A|XE) = H(A_1|E)$. We will also consider noisy preprocessing [116, 134], where Alice's raw key bit K_A is again the outcome of the measurement A_1, but with probability q she flips it and with probability $1 - q$ she keeps it as it is. We write $K_A = A_1^q$ for the corresponding random variable and thus $H(K_A|XE) = H(A_1^q|E)$ for the conditional entropy. Finally, the last case we will consider is one where K_A is obtained by choosing the observables A_1 and A_2 with probabilities p and $\bar{p} = 1 - p$, respectively, and applying noisy preprocessing with flip probability q to the measurement output. We then write $K_A = A_X^q$ and $H(K_A|XE) = H(A_X^q|XE)$.

Alice-Bob correlations. In a device-independent setting, the correlations between Alice and Bob can be characterized through *Bell linear functionals*, which are linear functions of 1-body and 2-body correlators. In the 2-input/2-output scenario, 1-body and 2-body correlators can all be written in the common form

$$\langle A_x \otimes B_y \rangle = \mathrm{Tr}\big[\rho_{AB}\, A_x \otimes B_y\big] \quad \text{for } x = 0, 1, 2 \tag{5.5}$$

if we define $A_0 = \mathbb{1}_A$ and $B_0 = \mathbb{1}_B$. A Bell linear functional S is then specified by 9 real coefficients $\{S_{xy}\}_{x,y=0,1,2}$ $(x, y = 0, 1, 2)$ and its value on a given set of correlators $\{\langle A_x \otimes B_y \rangle\}$ is given by

$$S = \sum_{x,y=0}^{2} S_{xy} \langle A_x \otimes B_y \rangle \,. \tag{5.6}$$

We refer to S as a *Bell expectation*. We will particularly be interested in the following in the CHSH functional

$$S = \langle A_1 B_1 \rangle + \langle A_1 B_2 \rangle + \langle A_2 B_1 \rangle - \langle A_2 B_2 \rangle \,. \tag{5.7}$$

Tradeoff between Eve's information on the raw key and Alice-Bob correlations. Assume that a procedure for generating a raw key value (as specified by a measurement probability distribution μ_X and preprocessing maps $\$_x$) and a series of $m \geq 1$ Bell expectation values $\boldsymbol{S} = (S_1, \ldots, S_m)$[1] are fixed. Our objective is to establish a lower bound

$$H(K_A|XE) \geq f(\boldsymbol{S}) \tag{5.8}$$

that is device independent, in the sense that it is satisfied by every quantum strategy $\mathcal{Q}$. For technical reasons, we require f to be a convex function of its arguments[2].

Relation to the security of DI QKD protocols. In a typical DI QKD protocol, Alice's and Bob's devices are successively used for n rounds. Some of the rounds are used to generate raw key values K_A on Alice's side and K_B on Bob's side. Some of the rounds are used to gather statistical data to decide, based on whether one or several Bell statistics are above some thresholds, if the protocol should be aborted or if it can proceed. In the latter case, error correction and privacy amplification are applied to the final raw key string. Following the application of the entropy accumulation theorem [127], the security of such a generic multi-round protocol can be reduced to deriving a tradeoff bound (5.8), which can be understood as characterizing the behavior of a single round[3] in expectation. In particular a tradeoff bound allows one to compute the key rate in the finite-key regime and in the asymptotic one, where it simply reduces to the Devetak-Winter formula [110]

$$r = H(K_A|XE) - H(K_A|K_B) \,, \tag{5.9}$$

where $H(K_A|K_B)$ is the conditional Shannon entropy of the classical random variables K_A and K_B.

[1] This can range from a single Bell functional, such as CHSH, to the entire set of correlators $\{\langle A_x \otimes B_y \rangle\}$, or anything in between.

[2] This is required for application of the entropy accumulation theorem, and follows naturally when reducing the analysis to qubits. Furthermore, if f defines a bound on $H(K|XE)$ that is tight, it must necessarily be convex by concavity of the conditional entropy and because any convex mixture of two strategies defines a valid strategy.

[3] The raw key generation procedure and the set of Bell statistics to be used in the single-round bound (5.8) should obviously coincide with those of the multi-round protocol.

5.2.2 Reduction to qubits

The lower bounds (5.8) we aim to derive must be proven valid for any quantum strategy $\mathcal{Q} = (\rho_{ABE}, A_1, A_2, B_1, B_2)$, defined a priori on Hilbert spaces of arbitrary dimension. However, because the strategies we consider involve only two binary measurements for Alice and for Bob, it is well-known that it is sufficient, thanks to Jordan's lemma, to consider pure qubit strategies [135].

More specifically, suppose that we have derived a lower bound $H(K_A|XE) \geq f(\boldsymbol{S})$, that is valid for any strategy $\mathcal{Q} = (|\Psi\rangle_{ABE}, A_1, A_2, B_1, B_2)$ where *i)* Alice's and Bob's systems are two-dimensional, *ii)* $|\Psi\rangle_{ABE}$ is a pure state, *iii)* A_1, A_2, B_1, B_2 are qubit, non-degenerate Pauli observables constrained to the Z–X plane on the Bloch sphere, and where *iv)* the function f is convex. Then this lower bound is valid for arbitrary strategies. For details, see for instance [133].

Note that the "2-input/2-output" restriction, which allows to make this qubit simplification, only applies to Alice's measurements and to those measurements of Bob that are involved in the definition of the Bell functionals $\boldsymbol{S}$, as these are the only measurements involved in the relation (5.8). The raw key generation procedure on Bob's side leading to the raw key value K_B can, however, involve further measurement choices with more outputs, see examples in the Section 5.2.6.

We now assume the above simplification and present our approach to deriving tradeoff bounds, which follows three technical steps described in the next three subsections.

5.2.3 BB84-type uncertainty relations

The first non-trivial step in our approach is *device-dependent* and consists in deriving a qubit uncertainty relation akin to those used in the analysis of the standard entanglement-based BB84 protocol and variants of it. Let us illustrate this on several examples. In the following, $\phi(x) = h\big(\frac{1}{2} + \frac{1}{2}x\big)$, where $h(x)$ is the binary entropy.

Consider first the simple situation where Alice's raw key bit $K_A = A_1$ is simply obtained as the outcome of the measurement A_1, i.e., there is no random input choice X and no classical preprocessing. We then have the following bound.

Entropy bound 1. *(BB84)*

$$H(A_1|E) \geq 1 - \phi\big(\big|\langle \bar{A}_1 \otimes B\rangle\big|\big)\,, \tag{5.10}$$

where $\bar{A}_1$ is a Pauli observable orthogonal to A_1 on the Bloch sphere and B any given ± 1-valued observable on Bob's subsystem.

This bound is simply a reexpression of the one-sided device-independent entropy bound $H(\mathrm{Z}|E) \geq 1 - \phi\big(|\langle \mathrm{X} \otimes B\rangle|\big)$ for the BB84 protocol [136] that relates the information Eve has about the outcome of a Z measurement by how much Bob is correlated to the complementary X measurement. The bound (5.10) directly follows from the fact that A_1 and $\bar{A}_1$ are Pauli operators, which we can identify with the Z and X operators.

As a second example, let us add noisy preprocessing [116, 134] to the raw key procedure: Alice's raw key bit $K_A = A_1^q$ is again the outcome of the measurement A_1, but with probability q she flips it and with probability $1 - q$ she keeps it as it is.

Entropy bound 2. *(BB84 bound with noisy preprocessing)*

$$H(A_1^q|E) \geq f_q\big(\big|\langle \bar{A}_1 \otimes B\rangle\big|\big)\,, \tag{5.11}$$

where

$$f_q(x) = 1 + \phi\Big(\sqrt{(1-2q)^2 + 4q(1-q)x^2}\Big) - \phi(x)\,, \tag{5.12}$$

and $\bar{A}_1$ is a Pauli observable orthogonal to A_1 on the Bloch sphere and B any given ±1-valued observable on Bob's subsystem.

This again follows by identifying A_1 and $\bar{A}_1$ with the Z and X operators and reusing a one-sided device-independent bound known for BB84 with noisy preprocessing [133, 137].

The two above bounds were used in [133] to analyse the security of a family of CHSH-based DI QKD protocols. But more generally, it is also possible to obtain other bounds, such as the two ones below, which we will apply to other variants of CHSH-based DI QKD protocols in Section 5.2.6.

Entropy bound 3. *(BB84 with noisy preprocessing and bias)*

$$H(A_1^q|E) \geq g_q\big(|\langle A_1\rangle|, \big|\langle \bar{A}_1 \otimes B\rangle\big|\big)\,, \tag{5.13}$$

where

$$g_q(z,x) = \phi\big(\tfrac{1}{2}(R_+ + R_-)\big) + \phi\big(\tfrac{1}{2}(R_+ - R_-)\big) - \phi\big(\sqrt{z^2+x^2}\big)\,, \tag{5.14}$$

with

$$R_\pm = \sqrt{(1-2q\pm z)^2 + 4q(1-q)x^2}\,, \tag{5.15}$$

and $\bar{A}_1$ is a Pauli observable orthogonal to A_1 on the Bloch sphere and B any given ±1-valued observable on Bob's subsystem.

This bound represents a refinement of the bound 2, as it depends not only on $\langle \bar{A}_1 \otimes B\rangle$, but also on the value of the 1-body correlator $\langle A_1\rangle$ measuring how much Alice's raw output is biased.

Our last example is one where Alice's raw key bit $K_A = A_X^q$ is obtained by choosing the observables A_1 and A_2 with probability p and $\bar{p} = 1-p$, respectively, and applying noisy preprocessing with flip probability q to the measurement output. The conditional entropy is then

$$H(A_X^q|XE) = pH(A_1^q|E) + \bar{p}H(A_2^q|E)\,, \tag{5.16}$$

and one has the following bound.

Entropy bound 4. *(Two-basis bound)*

$$H(A_X^q|XE) \geq f_q\Big(\sqrt{p\langle \bar{A}_1 \otimes B\rangle^2 + \bar{p}\langle \bar{A}_2 \otimes B'\rangle^2}\Big) \tag{5.17}$$

where $\bar{A}_1$ and $\bar{A}_2$ are observables orthogonal to A_1, A_2, respectively and $f_q(x)$ is the function defined in (5.12)*.*

The above bounds are essentially similar to those used in the analysis of standard entanglement-based QKD. They are valid for arbitrary entangled states $|\Psi\rangle_{ABE}$ where Alice's and Bob's systems are two dimensional and are expressed in terms of correlators $\langle A \otimes B\rangle$ between Alice and Bob that involve (contrarily to the device-independent case) *specific, fixed* observables, such as $\bar{A}_1$ on Alice's side. As such they can be derived using existing techniques.

We remark that all of these bounds can be derived from bound 3, which we derive in detail in Section 5.3.1. In particular, bound 2 is a special a case of bound 3 evaluated with $\langle A_1\rangle = 0$[4], while bound 1 is obtained by further setting $q = 0$. Bound 4 follows from bounding both contributions to the average entropy separately using bound 2,

$$\begin{aligned} H(A_X^q|XE) &= pH_q(A_1|E) + \bar{p}H_q(A_2|E) \\ &\geq pf_q\big(\big|\langle \bar{A}_1 \otimes B\rangle\big|\big) + \bar{p}f_q\big(\big|\langle \bar{A}_2 \otimes B'\rangle\big|\big)\,, \end{aligned} \tag{5.18}$$

[4]The resulting bound holds independently of the actual value of $\langle A_1\rangle$ thanks to the monotonicity property discussed below: if we make in bound 3 the replacement $|\langle A_1\rangle| \mapsto 0$ we obtain a bound that remains valid.

and then using that the function $x \mapsto f_q(\sqrt{x})$ is convex (see Appendix B of [133] for a proof of this property).

Importantly, we also show in Section 5.3.1 that all the above bounds satisfy a type of monotonicity property. We say that a bound $H(K_A|XE) \geq f(x)$ is *monotone* in x if the bound $H(K_A|XE) \geq f(x_-)$ holds for all $x_- \leq x$ and similarly in the multivariate case for each variable independently, e.g., $H(K_A|XE) \geq f(x, y)$ is *monotone* in x and y if the bound $H(K_A|XE) \geq f(x_-, y_-)$ hold for all $x_- \leq x$ and $y_- \leq y$. Note that this monotonicity property is weaker than monotonicity of the function f itself: if the function f is monotonically increasing then the bound $H(K_A|XE) \geq f(x)$ is monotone, but the converse does not necessarily hold.

Monotonicity property. *The entropy bounds* (5.10) *and* (5.11) *are monotone in* $\left|\langle \bar{A}_1 \otimes B\rangle\right|$, *the bound* (5.13) *is monotone in* $|\langle A_1\rangle|$ *and* $\left|\langle \bar{A}_1 \otimes B\rangle\right|$, *and the bound* (5.17) *is monotone in* $p\langle \bar{A}_1 \otimes B\rangle^2 + \bar{p}\langle \bar{A}_2 \otimes B'\rangle^2$.

The monotonicity of the bound (5.13) is established in Section 5.3.1 from which the monotonicity of the other bounds follows[5]. This property will be important in Section 5.2.4 as it allows replacing in the entropy bounds the correlators on which they depend in the right-hand side by a lower bound on these correlators and in Section 5.2.5 where it allows the systematic computation of a convex envelope based on a discrete set of points.

5.2.4 Pauli correlation constraints

The bounds on the conditional entropy $H(K_A|XE)$ that we have given in the previous subsection are expressed in terms of correlators involving observables which are not necessarily accessible through the devices, e.g., the correlator $\langle \bar{A}_1 \otimes B\rangle$ involving the observable $\bar{A}_1$. The second step of our approach consists in deriving a constraint on these correlators in terms of correlators involving only the observables A_1, A_2, B_1, B_2 *actually measured* by the devices.

For instance, it is a straightforward exercise, see [133], to show the following bound.

Correlation bound 1. *Correlation bound 1 (CHSH)*

$$\left|\langle \bar{A}_1 \otimes B\rangle\right| \geq \sqrt{S^2/4 - 1}\,, \tag{5.19}$$

where $S = \langle A_1B_1\rangle + \langle A_1B_2\rangle + \langle A_2B_1\rangle - \langle A_2B_2\rangle$ *is the expected value of the CHSH statistic and* $B \propto B_1 - B_2$.

More generally, one can also consider a family of asymmetric versions of the CHSH statistic for which the following bounds are shown in [133].

Correlation bound 2. *Correlation bound 2 (asymmetric CHSH) Let* $S_\alpha = \alpha\langle A_1B_1\rangle + \alpha\langle A_1B_2\rangle + \langle A_2B_1\rangle - \langle A_2B_2\rangle$ *be a variant of CHSH depending on a given parameter* $\alpha \in \mathbb{R}$. *Then for some appropriate choice of a* ± 1*-valued observable* B,

$$\left|\langle \bar{A}_1 \otimes B\rangle\right| \geq E_\alpha(S_\alpha)\,, \tag{5.20}$$

where

$$E_\alpha(S_\alpha) = \sqrt{{S_\alpha}^2/4 - \alpha^2} \tag{5.21}$$

if $|\alpha| \geq 1$ *or* $|S_\alpha| \geq 2\sqrt{1 + \alpha^2 - \alpha^4}$ *and*

$$E_\alpha(S_\alpha) = \sqrt{1 - \left(1 - \tfrac{1}{|\alpha|}\sqrt{(1-\alpha^2)({S_\alpha}^2/4 - 1)}\right)^2} \tag{5.22}$$

otherwise.

[5] In the case of bounds (5.10), (5.11), (5.17), it also follows from the stronger property that the function $f_q(x)$ is monotonically increasing in x, as shown in Appendix B. of [133].

The correlation bounds (5.19) and (5.20) can be derived analytically. But more generically, one can derive numerical lower bounds on polynomial functions of arbitrary qubit correlators, such as $\langle \bar{A}_1 \otimes B\rangle$ or $\langle \bar{A}_2 \otimes B'\rangle$, in terms of Bell functionals involving only the accessible correlators $\langle A_x \otimes B_y\rangle$ $(x, y = 0, 1, 2)$, using the Lasserre hierarchy of semidefinite programming relaxations for polynomial optimization [138, 139]. This can be done by parameterizing explicitly all qubit operators in the Z–X plane.

We illustrate this general idea on the specific problem of deriving a lower bound for the expression

$$p\langle \bar{A}_1 \otimes B\rangle^2 + \bar{p}\langle \bar{A}_2 \otimes B'\rangle^2 \tag{5.23}$$

appearing on the right-hand side of (5.17) in terms of the CHSH expectation value S.

We first recall that we can use any ± 1-valued observables B and B' in (5.17). Taking these to be of the form

$$B^{(\prime)} = \cos\big(\varphi_B^{(\prime)}\big)\mathrm{Z} + \sin\big(\varphi_B^{(\prime)}\big)\mathrm{X} \tag{5.24}$$

and then choosing the angles φ_B and φ'_B that maximize (5.23) we obtain

$$\begin{aligned} &p\langle \bar{A}_1 \otimes B\rangle^2 + \bar{p}\langle \bar{A}_2 \otimes B'\rangle^2 \\ &\quad = p\big(\langle \bar{A}_1 \otimes \mathrm{Z}\rangle^2 + \langle \bar{A}_1 \otimes \mathrm{X}\rangle^2\big) \\ &\qquad + \bar{p}\big(\langle \bar{A}_2 \otimes \mathrm{Z}\rangle^2 + \langle \bar{A}_2 \otimes \mathrm{X}\rangle^2\big)\,. \end{aligned} \tag{5.25}$$

We then choose Alice's basis such that

$$A_1 = \cos\big(\tfrac{\varphi_A}{2}\big)\mathrm{Z} - \sin\big(\tfrac{\varphi_A}{2}\big)\mathrm{X}\,, \tag{5.26}$$

$$A_2 = \cos\big(\tfrac{\varphi_A}{2}\big)\mathrm{Z} + \sin\big(\tfrac{\varphi_A}{2}\big)\mathrm{X} \tag{5.27}$$

and the complementary operators are

$$\bar{A}_1 = \sin\big(\tfrac{\varphi_A}{2}\big)\mathrm{Z} + \cos\big(\tfrac{\varphi_A}{2}\big)\mathrm{X}\,, \tag{5.28}$$

$$\bar{A}_2 = -\sin\big(\tfrac{\varphi_A}{2}\big)\mathrm{Z} + \cos\big(\tfrac{\varphi_A}{2}\big)\mathrm{X} \tag{5.29}$$

for some unknown angle φ_A. Using these in the above expression we obtain, explicitly,

$$\begin{aligned} &p\langle \bar{A}_1 \otimes B\rangle^2 + \bar{p}\langle \bar{A}_2 \otimes B'\rangle^2 \\ &\quad = \sin\big(\tfrac{\varphi_A}{2}\big)^2 ({E_{\mathrm{zz}}}^2 + {E_{\mathrm{zx}}}^2) + \cos\big(\tfrac{\varphi_A}{2}\big)^2 ({E_{\mathrm{xz}}}^2 + {E_{\mathrm{xx}}}^2) \\ &\qquad + 2(2p-1)\sin\big(\tfrac{\varphi_A}{2}\big)\cos\big(\tfrac{\varphi_A}{2}\big)(E_{\mathrm{zz}}E_{\mathrm{xz}} + E_{\mathrm{zx}}E_{\mathrm{xx}})\,, \end{aligned} \tag{5.30}$$

where we note the expectation values of products of Pauli operators $E_{\mathrm{xx}} = \langle \mathrm{X} \otimes \mathrm{X}\rangle$ and similarly for E_{xz}, E_{zx}, and E_{zz}.

We wish to constrain (5.30) for a given value of the CHSH expectation value which, in the choice of basis made above, takes the form

$$\begin{aligned} S &= \langle (A_1 + A_2) \otimes B_1\rangle + \langle (A_1 - A_2) \otimes B_2\rangle \\ &= 2\cos\big(\tfrac{\varphi_A}{2}\big)\langle \mathrm{Z} \otimes B_1\rangle - 2\sin\big(\tfrac{\varphi_A}{2}\big)\langle \mathrm{X} \otimes B_2\rangle\,. \end{aligned} \tag{5.31}$$

Maximizing the second line over (nondegenerate) ± 1-valued observables B_1 and B_2 in the Z–X plane gives

$$\begin{aligned} S/2 \leq\ &\big|\cos\big(\tfrac{\varphi_A}{2}\big)\big|\sqrt{{E_{\mathrm{zz}}}^2 + {E_{\mathrm{zx}}}^2} \\ &+ \big|\sin\big(\tfrac{\varphi_A}{2}\big)\big|\sqrt{{E_{\mathrm{xz}}}^2 + {E_{\mathrm{xx}}}^2}\,, \end{aligned} \tag{5.32}$$

which can be read as a constraint on the unknown angle φ_A and Pauli correlations E_{xx}, E_{xz}, E_{zx}, and E_{zz} appearing in (5.30).

To complete the problem, we finally remark that E_{xx}, E_{xz}, E_{zx}, and E_{zz} can be interpreted as expectations of products of the Z and X Pauli operators for some underlying state only if they satisfy

$$E_{\mathrm{zz}}{}^2 + E_{\mathrm{zx}}{}^2 \leq 1\,, \tag{5.33}$$
$$E_{\mathrm{xz}}{}^2 + E_{\mathrm{xx}}{}^2 \leq 1\,, \tag{5.34}$$

and

$$\begin{aligned}&\big(1 - E_{\mathrm{zz}}{}^2 - E_{\mathrm{zx}}{}^2\big)\big(1 - E_{\mathrm{xz}}{}^2 - E_{\mathrm{xx}}{}^2\big)\\ &\quad\geq \big(E_{\mathrm{zz}}E_{\mathrm{xz}} + E_{\mathrm{zx}}E_{\mathrm{xx}}\big)^2\end{aligned} \tag{5.35}$$

as shown in Section 4.3 of [133].

To get a valid lower bound on (5.40), it is thus sufficient to minimize the left-hand side of (5.30) given the constraints (5.32)–(5.35). The problem can be simplified by introducing the new variables

$$E_{\mathrm{zz}} = \lambda\cos(z)\,, \qquad E_{\mathrm{zx}} = \lambda\sin(z)\,, \tag{5.36}$$
$$E_{\mathrm{xz}} = \mu\cos(x)\,, \qquad E_{\mathrm{xx}} = \mu\sin(x)\,, \tag{5.37}$$
$$s = \sin\tfrac{\varphi_A}{2}\,, \qquad c = \cos\tfrac{\varphi_A}{2}\,, \tag{5.38}$$
$$\Delta = \cos(x - z)\,. \tag{5.39}$$

Using the trigonometric identity $\cos\tfrac{\varphi_A}{2}^2 + \sin\tfrac{\varphi_A}{2}^2 = 1$ and that we can drop the absolute values from (5.32) without substantially changing the problem, we arrive at the following.

Correlation bound 3. *Correlation bound 3 (two-basis) There exist ±1-valued qubit operators B and B' acting on Bob's subsystem such that*

$$p\langle\bar{A}_1\otimes B\rangle^2 + \bar{p}\langle\bar{A}_2\otimes B'\rangle^2 \geq E_p(S)^2\,, \tag{5.40}$$

where $E_p(S)^2$ is the solution to the minimization problem

$$\begin{aligned} E_p(S)^2 = \min \quad & s^2\lambda^2 + c^2\mu^2 + 2(2p-1)sc\lambda\mu\Delta\\ \text{s.t.} \quad & c\lambda + s\mu \geq S/2\\ & \lambda^2 \leq 1\\ & \mu^2 \leq 1\\ & (1-\lambda^2)(1-\mu^2) \geq \lambda^2\mu^2\Delta^2\\ & c^2 + s^2 = 1\\ & \Delta^2 \leq 1 \end{aligned} \tag{5.41}$$

in the five variables $\lambda, \mu, c, s, \Delta \in \mathbb{R}$.

As the above is a polynomial optimization problem, it can be reduced to a sequence of semidefinite programs using the Lasserre hierarchy [138, 139]. Importantly, every SDP relaxation at a given order in the hierarchy provides a valid lower bound to the optimization problem and consequently a valid lower bound of the form (5.40). At level 3 of the Lasserre hierarchy, the problem takes less than a second to solve and appears to already give the optimal solution.

In the case in which $p = 1/2$, the above problem can actually be solved analytically, as shown in Section 5.3.2. The result in that case is

$$E_{\frac{1}{2}}(S)^2 = \frac{1 + {x_*}^2}{1 - x_*} + \frac{S^2}{4}\frac{1 + x_*}{1 - x_*} - \frac{S}{\sqrt{2}}\frac{(1 + x_*)^{3/2}}{1 - x_*}\,, \tag{5.42}$$

where the variable x_* is the solution of

$$4x(2 - x) + 2(S^2 + 2) + S(x - 5)\sqrt{2(1 + x)} = 0 \tag{5.43}$$

in the range

$$-\frac{S}{4}\sqrt{8 - S^2} \leq x \leq \frac{S}{4}\sqrt{8 - S^2}\,. \tag{5.44}$$

Eq. (5.43) can be rearranged to a root-finding problem for a degree 4 polynomial in x and can thus be solved analytically, though the solution is quite lengthy and we do not explicitly report it here.

5.2.5 Convexity and fully device-independent bounds

Combining the above correlation bounds and the entropy bounds of the previous section, one obtains bounds on the conditional entropy that are device independent modulo the qubit reduction. For instance, using the CHSH correlation bound (5.19) in the BB84 entropy bound (5.10), where the substitution of (5.19) in (5.10) is possible thanks to the monotonicity property of the BB84 entropy bound, we recover the CHSH entropy bound

$$H(A_1|E) \geq 1 - \phi\big(\sqrt{S^2/4 - 1}\big) \tag{5.45}$$

given in the introduction and originally derived in [4]. Using (5.20) in the BB84 bound with noisy preprocessing (5.11), one obtains the more general qubit bound

$$H(A_1^q|E) \geq f_q\big(E_\alpha(S_\alpha)\big) \tag{5.46}$$

derived in [133].

But other combinations are also possible, such as the two original following ones, which we are going to consider in more detail in Section 5.2.6.

The first, which gives a bound on the entropy in terms of $\langle A_1 \rangle$ in addition to CHSH, is simply obtained by combining (5.19) and (5.13):

$$H(A_1^q|E) \geq g_q\big(|\langle A_1 \rangle|, \sqrt{S^2/4 - 1}\big)\,. \tag{5.47}$$

For the second, let $\tilde{E}_p(S)^2$ denote any lower bound to $E_p(S)^2$ obtained by solving analytically or numerically the polynomial optimization problem (5.41) or any of its relaxations in the Lasserre hierarchy. Then using such a bound in (5.17), we obtain

$$H(A_X^q|XE) \geq f_q\big(\tilde{E}_p(S)\big) \tag{5.48}$$

with $\tilde{E}_p(S) \equiv \sqrt{\tilde{E}_p(S)^2}$.

Convexity analysis

Regardless of the combination used, the result is a bound on the conditional entropy valid for two-qubit systems, which can only be extended to give a fully device-independent bound, valid in arbitrary dimension, if it is convex. The third and final step thus consists of a convexity analysis.

If we obtain a qubit bound on the conditional entropy with a reasonably simple analytic expression then it may be feasible to study its properties directly. Either we simply prove it is convex, as can be done for (5.45), or more generally as was done in [133] for (5.46) for $|\alpha| \geq 1$. Or we analytically establish that it is not convex and determine its convex envelope, as was done in [133] for (5.46) for $|\alpha| < 1$.

More generally, however, the qubit bound may be obtained numerically or it may be analytic but of a form that does not easily lend itself to an analytic convexity analysis, as is the case for the bounds (5.47) and (5.48). In such cases, we need a way of constructing a convex lower bound on whatever qubit bound we obtain.

Convex lower bounds through linear programming

A simple solution that we can use, provided our entropy bounds satisfy the monotonicity property introduced in Section 5.2.3, is based on a discretization of the qubit bound, similar to the approach used in [130]. In the following, let us generically write the bound valid for two-qubit systems as

$$H(K_A|XE) \geq f(\boldsymbol{S})\,, \tag{5.49}$$

where $f\colon \mathcal{D} \to \mathbb{R}$ is a function, defined on some domain $\mathcal{D}$, that we either know analytically or can compute numerically, of one or more Bell expectation values $\boldsymbol{S} = (S_1, S_2, \ldots, S_n) \in \mathcal{D}$.

Let us introduce a covering $\mathcal{K} = \{K\}$ of the domain $\mathcal{D}$ by polytopes $\{K\}$, such that every $\boldsymbol{S} \in \mathcal{D}$ is contained in at least one of the polytopes K. In practice, we would typically use a grid partition in terms of hyperrectangles where each point (outside of vertices and shared edges) is contained in only one hyperrectangle K (but this is not strictly necessary for the method to work).

Let us suppose, furthermore, that for every K we have a way of identifying a value $f[K]$ that we can use as a lower qubit bound on the conditional entropy valid for the entire polytope, i.e., such that

$$H(K_A|XE) \geq f[K]\,, \qquad \forall \boldsymbol{S} \in K\,. \tag{5.50}$$

We can then define a discretized qubit bound,

$$H(K_A|XE) \geq f_{\mathcal{K}}(\boldsymbol{S}) \tag{5.51}$$

where $f_{\mathcal{K}}$ is defined as

$$f_{\mathcal{K}}(\boldsymbol{S}) = \min_{K \ni \boldsymbol{S}} f[K]\,, \tag{5.52}$$

where the minimization is taken over all polytopes K that contain $\boldsymbol{S}$. This, in particular, associates unique values $f_{\mathcal{K}}(\boldsymbol{S}_j)$ to the vertices $\boldsymbol{S}_j$ of the polytopes. The convex envelope of the discretized function $f_{\mathcal{K}}$, finally, is readily given by the solution to the following linear program-

ming problem,

$$\begin{aligned} \bar{f}_{\mathcal{K}}(\boldsymbol{S}) = \ & \text{minimize} \sum_j \theta_j \, f_{\mathcal{K}}(\boldsymbol{S}_j) \\ & \text{subject to} \sum_j \theta_j \boldsymbol{S}_j = \boldsymbol{S} \\ & \sum_j \theta_j = 1 \\ & \theta_j \geq 0 \,, \end{aligned} \tag{5.53}$$

where the $\boldsymbol{S}_j$ are the combined vertices of all the polytopes K in $\mathcal{K}$. We thus obtain a bound

$$H(K_A|XE) \geq \bar{f}_{\mathcal{K}}(\boldsymbol{S}) \tag{5.54}$$

on the conditional entropy that is convex and extends to the fully device-independent setting.

We have not explained, however, how one can identify in (5.50) the lower-bound values $f[K]$ for each polytope K, which is crucial to define a discretized qubit bound. This can be done if the bound (5.49) is monotone in $|\boldsymbol{S}| = (|S_1|, |S_2|, \ldots, |S_n|)$, i.e., if the bound still holds if we replace in (5.49) any of the n Bell expectation values S_i by a value s_i that is smaller in absolute value, $|s_i| \leq |S_i|$. This is in particular the case for all the bounds (5.45)–(5.48) presented above since they are obtained by combining the monotone entropy bounds of Section 5.2.3 with the monotonically increasing correlation bounds of Section 5.2.4. Using this monotonicity property, we can now simply divide the domain $\mathcal{D}$ into hyperrectangles K and use as the lower-bound value $f[K]$ for each hyperrectangle K, the value of the qubit bound evaluated at the corner that is closest to the origin.

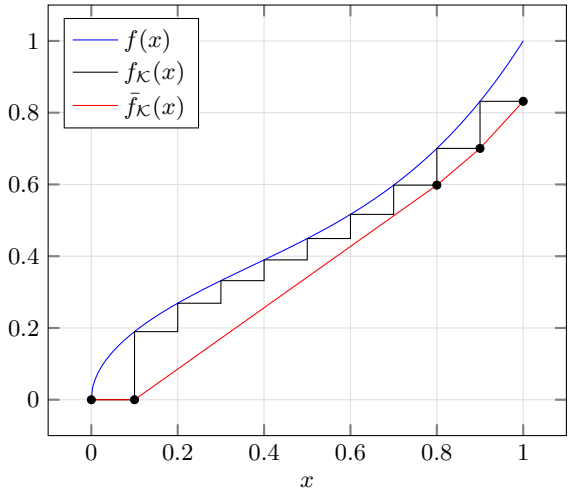

Figure 5.1: Convex lower bound $\bar{f}_{\mathcal{K}}$ of a function f constructed on n (in the figure $n = 10$) equally spaced subdivisions of its domain, i.e., the polytopes K are here n consecutive line segments between $x = 0$ and $x = 1$. We actually used this method on the qubit bound (5.48), but the function $f_q(\tilde{E}_p(S))$ is too close to convex to make a visually interesting example. The construction is thus illustrated on the figure for the visibly non-convex function $f(x) = 0.6\sqrt{x} + 0.4x^4$.

Finally, in the special case that we are working with a qubit entropy bound $H(K_A|XE) \geq f(S)$ of a single variable S, we remark that one can avoid the linear program and compute $f_{\mathcal{K}}(S)$ very

rapidly essentially by eliminating the redundant vertices and interpolating between the remaining ones, as illustrated in Figure 5.1. This can be done in linear time in the number of vertices [140, 141]. We in particular applied this technique to the two-basis bound (5.48) to compute the key-rate bounds obtained in Section 5.2.6 below.

Certifying an affine tradeoff bound

While we can always use the above approach when we have a qubit entropy bound satisfying the monotonicity property, it is not always necessary to solve the linear programming problem to obtain a valid convex lower bound on the conditional entropy. An alternative approach, which would ultimately lend itself to more direct use in the entropy accumulation theorem, is to certify a linear or affine lower bound on the entropy.

Here, let us suppose we believe that the conditional entropy respects an affine lower bound

$$H(K_A|XE) \geq \beta + \boldsymbol{\alpha} \cdot \boldsymbol{S} - \varepsilon\,, \tag{5.55}$$

that we wish to certify up to some precision ε. Such a bound may be obtained, for example, by computing at a particular point the tangent of a function $\bar{f}(\boldsymbol{S})$ that we believe to be the convex hull of a known qubit bound $f(\boldsymbol{S})$. As above, we introduce a covering $\mathcal{K} = \{K\}$ of the domain $\mathcal{D}$ with polytopes K and assume for every K a lower bound $f[K]$ on the conditional entropy, as defined in (5.50). We also define

$$\begin{aligned} \alpha[K] &= \max_{\boldsymbol{S} \in K} \boldsymbol{\alpha} \cdot \boldsymbol{S} \\ &= \max_{\boldsymbol{S} \in \mathrm{Vert}(K)} \boldsymbol{\alpha} \cdot \boldsymbol{S} \end{aligned} \tag{5.56}$$

where $\mathrm{Vert}(K)$ are the vertices of K. To check that (5.55) holds, we then only need to verify that

$$\beta + \alpha[K] - f[K] \leq \varepsilon \tag{5.57}$$

holds for all polytopes K in the covering $\mathcal{K}$, which is now a finite problem. Alternatively, we can compute the maximal value over $\mathcal{K}$ of $\beta + \alpha[K] - f[K]$ to determine the best possible precision ε we can achieve given our covering choice.

An important difference with the linear programming approach above is that we do not necessarily have to decide on a covering $\mathcal{K}$ in advance. In fact, this is often very wasteful as, to obtain a good bound with a small tolerance, we would typically find we need a fine discretization of the domain only close to where the bound coincides with its tangent. Finding a suitable discretization can then be done naturally, and in practice often very rapidly, by starting by testing (5.57) for the polytopes K in an initially coarse covering (which could consist of just one polytope containing the entire domain) and then, for each K for which the test fails, subdividing K into smaller polytopes and recursively applying the test to each of those (see illustration in Figure 5.2).

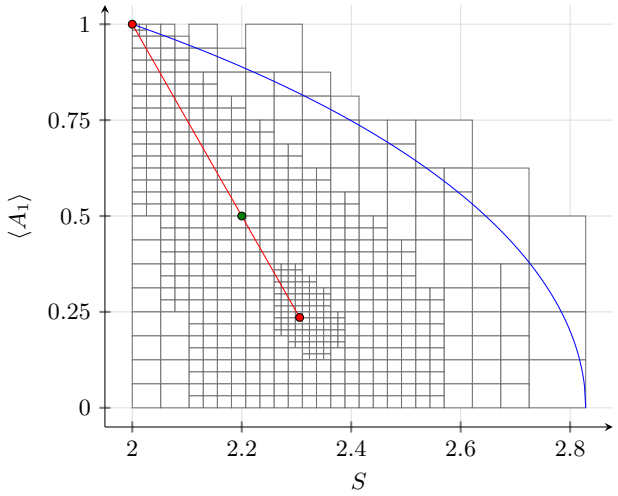

Figure 5.2: Certification of an affine lower entropy bound based on the qubit bound (5.47) depending on the CHSH expectation value S and the one-body correlator $\langle A_1 \rangle$. The blue curve represents the boundary of the domain $\mathcal{D} \subset [0,1] \times [2, 2\sqrt{2}]$ where the values of $(\langle A_1 \rangle, S)$ are consistent with quantum theory. We conjecture that the convex envelope of the function $\tilde{g}_q(\langle A_1 \rangle, S) = g_q\big(|\langle A_1 \rangle|, \sqrt{S^2/4 - 1}\big)$ in $\mathcal{D}$ is obtained by taking a convex decomposition of the point $(1, 2)$ and a point on the line from $(1, 2)$ to $(\langle A_1 \rangle, S)$. The figure illustrates such a convex decomposition (red points) for the point $(0.5, 2.2)$ (green point). From this, we can compute a candidate affine function (5.55) that optimally certifies the entropy of the point $(0.5, 2.2)$. Setting a value for ε, we then run a recursive algorithm to find a rectangle covering, depicted in the figure, that certifies the candidate affine function. We chose a value $\varepsilon = 0.025$ such that the resultant covering is coarse enough that it can be visualized, but much smaller values, e.g., $\varepsilon \approx 10^{-8}$ or less can readily be used.

Application to the bound (5.47) including the bias $\langle A_1 \rangle$. We used this recursive certification method, coupled with a guess on the optimal linear tradeoff functions, for the qubit bound (5.47) which depends on the two variables $\langle A_1 \rangle$ and S. The function $\tilde{g}_q(\langle A_1 \rangle, S) \equiv g_q\big(|\langle A_1 \rangle|, \sqrt{S^2/4 - 1}\big)$ defining this bound is not convex as its Hessian matrix is not positive semidefinite everywhere. It appears, though, to be convex in each of the parameters $\langle A_1 \rangle$ and S individually, and more generally in any direction passing through the positive orthant in the plane $\langle A_1 \rangle$–S. This implies that the convex envelope of $\tilde{g}_q(\langle A_1 \rangle, S)$ can be constructed by considering at most convex combinations of *two* points in the plane, instead of three points as follows by Carathéodory's theorem. Indeed, any non-trivial convex combination of three points in the plane $\langle A_1 \rangle$–S would have at least two of those points joined by a segment aligned in the direction of the positive orthant. But since the function is convex in that direction, one can advantageously replace the two points by a mixture of those.

Furthermore, if we are interested in computing a valid entropy bound for a point with $\langle A_1 \rangle$ positive, it is sufficient to consider convex combinations in the domain $\mathcal{D} \subset [0,1] \times [2, 2\sqrt{2}]$ of the plane $\langle A_1 \rangle$–S, i.e., points with negative values of $\langle A_1 \rangle$ can be neglected. To see this, consider a convex combination

$$(\langle A_1 \rangle, S) = t\,\big(\langle A_1 \rangle', S'\big) + (1-t)\,\big(\langle A_1 \rangle'', S''\big) \tag{5.58}$$

where $\langle A_1 \rangle' < 0$ is negative for the point $(\langle A_1 \rangle, S)$ yielding a corresponding value for the entropy function

$$t\,\tilde{g}_q\big(\langle A_1 \rangle', S'\big) + (1-t)\,\tilde{g}_q\big(\langle A_1 \rangle'', S''\big) \tag{5.59}$$

that is a valid lower bound for $H(A_1^q|E)$. Replace now this convex strategy by the (valid) convex

combination

$$(\langle A_1\rangle, S) = t\,(0, S') + (1-t)\left(\frac{\langle A_1\rangle}{1-t},\, S''\right). \tag{5.60}$$

The corresponding value for the entropy function is

$$t\,\tilde{g}_q(0, S') + (1-t)\,\tilde{g}_q\left(\frac{\langle A_1\rangle}{1-t},\, S''\right), \tag{5.61}$$

which is still a valid lower bound for $H(A_1^q|E)$ because of the monotonicity property of the bound and the fact that $\frac{\langle A_1\rangle}{1-t} \leq \langle A_1\rangle''$ (since $\langle A_1\rangle' < 0$).

Finally, we numerically observed that the convex envelope of $\tilde{g}_q(\langle A_1\rangle, S)$ in the domain $[0,1]\times[2, 2\sqrt{2}]$ was always obtained by taking a convex decomposition of two particular points: the point $(1,2)$ and a point on the line from $(1,2)$ to $(\langle A_1\rangle, S)$. This observation gives a conjecture for the convex envelope of the qubit bound (5.47), from which candidate linear tradeoff functions of the form (5.55) can readily be computed as tangents to this envelope. We can then attempt to certify that such candidates are indeed proper tradeoff functions through a rectangle covering and the recursive procedure described above, as illustrated in Figure 5.2. We can in principle perform such certification to arbitrary precision ε, though, in practice, we may be limited by the number of rectangles required to reach a very small ε and by the limited precision of hardware floating-point arithmetic on typical computers. The key rates and results presented in Section 5.2.6 have been computed using this procedure. From our results, it appears that our conjecture on the convex envelope of $\tilde{g}_q(\langle A_1\rangle, S)$ is correct as we are always able to certify the resultant linear tradeoff functions up to a precision of the order of $\varepsilon \approx 10^{-6}$ or better.

5.2.6 Applications

Here, we apply our method to bound the asymptotic one-way key rate, given by the Devetak-Winter rate

$$r = H(K_A|XE) - H(K_A|K_B)\,, \tag{5.62}$$

for DI QKD in two situations of interest: white noise, where we assume that Alice and Bob share an attenuated version,

$$\rho = v\phi^+ + (1-v)\mathbb{1}/4\,, \tag{5.63}$$

depending on some visibility v, of the ideal maximally-entangled state

$$\left|\phi^+\right\rangle = \frac{1}{\sqrt{2}}\big(|00\rangle + |11\rangle\big)\,, \tag{5.64}$$

and limited detection efficiency, where we assume that Alice's and Bob's devices return one of the expected outcomes ± 1 with a probability η less than one.

The qubit bound (5.45) (which is already convex) was used in [4] to compute the key rate of the standard CHSH DI QKD protocol and the convexification of (5.46) was used in [133] to generalize the analysis in terms of the asymmetric CHSH expressions S_α and incorporating noisy preprocessing. We will now illustrate the use of the two other qubit bounds (5.47) and (5.48) given in the preceding section, in subsections 5.2.6 and 5.2.6, respectively.

In [133], the asymmetric CHSH expressions were chosen for parameter estimation because they retain the same symmetries as the version of the DI QKD protocol where only one of Alice's measurements, A_1, is used to generate the key and they can be used to derive the optimal one-way key rate for that protocol with respect to white noise. There is no analogous connection between the asymmetric CHSH expressions and losses and, in fact, the lowest threshold, $\eta \approx 82.57\%$, on

the global detection efficiency reported in [133] was obtained using CHSH (the special case of S_α with $\alpha = 1$).

In the following, we reanalyse these correlation models using different setups. In particular, as [133] already does an optimal analysis for white noise using one measurement basis for key generation and with noisy preprocessing, the only remaining way to improve the noise robustness is to use a different protocol. For that case, we apply our approach to a variant of the protocol based on CHSH, proposed recently in [130], in which both of Alice's measurements A_1 and A_2 are used to generate the key. For losses, by contrast, as remarked in [133] the analysis performed there was likely not optimal as the treatment of losses introduced biases in the probabilities of Alice's and Bob's measurement outcomes, while the analytic bound on the entropy used there was optimized for the case that Alice's outcomes are obtained equiprobably. For losses, therefore, we concentrate on bounding the key rate using the expectation value $\langle A_1 \rangle$ of Alice's key-generation measurement in addition to the Bell violation.

White noise analysis for the two-basis protocol

In the two-basis protocol of [130], Alice and Bob ideally share a maximally-entangled state $|\phi^+\rangle$ and have devices that, for Alice, ideally perform the two measurements

$$A_1 = \mathrm{Z}\,, \qquad\qquad A_2 = \mathrm{X}\,, \tag{5.65}$$

and, for Bob, the four measurements

$$B_1 = \frac{\mathrm{Z}+\mathrm{X}}{\sqrt{2}}\,, \qquad\qquad B_3 = \mathrm{Z}\,, \tag{5.66}$$

$$B_2 = \frac{\mathrm{Z}-\mathrm{X}}{\sqrt{2}}\,, \qquad\qquad B_4 = \mathrm{X}\,. \tag{5.67}$$

This ideal realization is designed so that the measurements A_1, A_2, B_1, and B_2 yield a maximal violation of the CHSH Bell inequality while Bob's measurements B_3 and B_4 yield outcomes that are perfectly correlated with Alice's when she measures, respectively, A_1 and A_2, i.e., $\langle A_1 B_3 \rangle = \langle A_2 B_4 \rangle = 1$.

In the protocol, Alice and Bob use rounds where Bob measures B_1 or B_2 to estimate CHSH; they use a small fraction of the rounds where Bob measures B_3 and B_4 to estimate how correlated the outcomes are with A_1 and A_2, and use the results of the remaining rounds where Alice and Bob measured A_1 and B_3 or A_2 and B_4 as their raw key. We also assume in the following that Alice flips her outcomes in the key generation rounds (i.e., applies noisy preprocessing) with some probability q.

Let us suppose that Alice uses the measurements A_1 and A_2 with probabilities p' and $\bar{p}' = 1-p'$ and that Bob uses the measurements B_3 and B_4 with the same relative probabilities. Then, out of the rounds not used for parameter estimation, the asymptotic key rate, taking into account the effect of sifting[6], is

$$\begin{aligned} r &= p'^2\, r_{13} + \bar{p}'^2\, r_{24} \\ &= (p'^2 + \bar{p}'^2)(p\, r_{13} + \bar{p}\, r_{24})\,, \end{aligned} \tag{5.68}$$

where

$$r_{xy} = H(A_x^q|E) - H(A_x^q|B_y) \tag{5.69}$$

[6] In particular, the key rate is attenuated by the probability $p'^2 + \bar{p}'^2$ that Alice and Bob use matching bases. It has been pointed out in [100] that this can be avoided, but this requires the parties to either possess quantum memories or to use a very long preshared key to coordinate the measurement choices.

and we introduced $p = p'^2/(p'^2 + \bar{p}'^2)$ and $\bar{p} = 1 - p$ in the second line. Here, $H(A_1^q|B_3)$ and $H(A_2^q|B_4)$ depend only on the correlations between Alice's and Bob's measurement outcomes, which they know from parameter estimation. Assuming Alice and Bob perform the ideal measurements on an attenuated state (5.63), the entropies of Alice's outcomes conditioned on Bob are

$$H(A_1^q|B_3) = H(A_2^q|B_4) = h\big(q + \delta(1-2q)\big)\,, \tag{5.70}$$

where the channel error rate δ is related to the visibility v in (5.63) by $v = 1 - 2\delta$, while the CHSH expectation value is

$$S = 2\sqrt{2}(1-2\delta)\,. \tag{5.71}$$

Bounding the key rate thus amounts to establishing a lower bound on the weighted average conditional entropy

$$pH(A_1^q|E) + \bar{p}H(A_2^q|E) = H(A_X^q|XE) \tag{5.72}$$

depending on the CHSH violation. A valid *qubit* bound in terms of the CHSH expectation value S is given by (5.48), from which a valid, fully device-independent, convex lower bound can be obtained using the techniques discussed in Section 5.2.5.

We can thus express the bound we obtain on the key rate, via CHSH, in terms of δ using our approach as

$$r \geq (p'^2 + \bar{p}'^2)\Big[\tilde{f}_q\big(2\sqrt{2}(1-2\delta)\big) - h\big(q + \delta(1-2q)\big)\Big]\,, \tag{5.73}$$

where $\tilde{f}_q(S)$ is the convex lower bound we obtain for the entropy, evaluated at $S = 2\sqrt{2}(1-2\delta)$.

We remark here that we could, in principle, bound the average entropy in terms of any correlation Bell inequality. We use only the CHSH expectation value here both for simplicity and because, in the most interesting case where the bases are used equiprobably (i.e., $p = 1/2$), we can infer from the symmetries of the protocol that CHSH is already the optimal measure of nonlocality for white noise (see Section 5.3.3 for details).

The key rate we obtain using our approach for $p = 0.5$ and $p = 0.75$ are illustrated, and compared with the known analytical bounds for $p = 1$, without noisy preprocessing (i.e., $q = 0$) and with the optimal amount of noisy preprocessing applied in Figures 5.3 and 5.4. The threshold noise rates up to which we obtain a positive key rate are reported for different values of q in Table 5.1. For $q = 0$ and q close to $1/2$, the results essentially rigorously confirm the thresholds of 8.36% and 9.24% that were anticipated could be obtained in the conclusion of [133]. For $0 < p < 1/2$, similar to [130], we did not see any improvement to the key rate; the highest rate appeared to always be obtained with either $p = 1$ or $p = 1/2$, depending on the value of S. However, as it may not be realistic to be sure that the measurements are used *exactly* equiprobably in a real implementation, we note that it is important to be able to bound the entropy for values of p that may deviate a little from 0.5. The key rate is in fact very robust against deviations of p from 0.5, as can be seen comparing the results for $p = 0.5$ and $p = 0.75$ in Figures 5.3 and 5.4.

The best threshold of 9.24% obtained for q close to $1/2$ using our method is close to the best threshold of 9.33% recently reported in [100] and obtained for $q = 0.3$, although the method we have used allows the key rate to be bounded much more rapidly[7]. Without noisy preprocessing, the threshold of 8.36% we obtain is slightly better than the threshold around 8.24% found in [130] and the same as the threshold that would be obtained using the "conjectured alternative proof"

[7] Ref. [100] reports requiring ~ 5000 processor-core hours to obtain a numerical bound on the average conditional entropy. For comparison, using our method we could generate a plot of the conditional entropy with 500 points in a minute or two on a regular laptop using the Lasserre hierarchy or almost instantaneously using the analytic method for $p = 1/2$ described in Section 5.3.2.

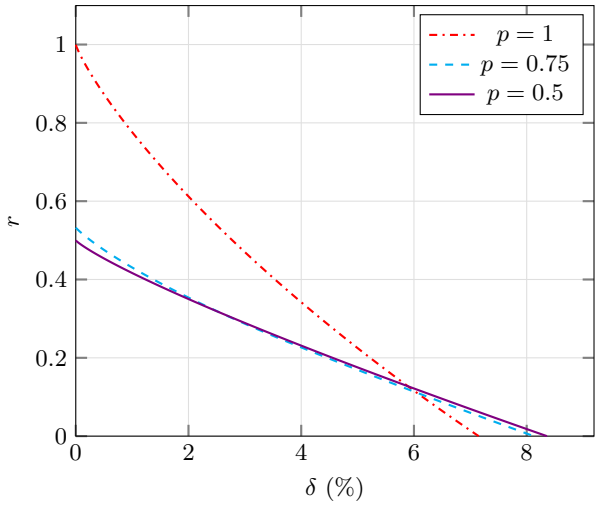

Figure 5.3: Lower bound on the Devetak-Winter rate as a function of the channel error rate δ, assuming $q = 0$.

p	$q = 0$	$q = 0.2$	$q = 0.3$	$q = 0.49$	$q \to 1/2$
1	7.1492	7.9503	8.0321	8.0848	8.0848
0.5	8.3599	9.1130	9.1923	9.2434	9.2435

Table 5.1: Threshold error rates (%) obtained for different probabilities p of measuring A_1 after sifting non-matching basis.

(after taking the convex envelope of the result) proposed in SectionI.H of the supplementary information to the same paper[8].

We provide an indication of how close the key-rate bound we obtain in the case $p = 1/2$ is to being optimal by comparing with a specific strategy, which was already identified as a likely candidate for the optimal collective attack for $q = 0$ in [133], and described in Section 5.3.4. This attack yields the following value for the average entropy

$$\tfrac{1}{2}H(A_1^q|E) + \tfrac{1}{2}H(A_2^q|E) = \bar{f}_q\big(S/\sqrt{8}\big)\,, \tag{5.74}$$

where

$$\bar{f}_q(x) = \begin{cases} f_q(x) & \text{if } x \geq x_* \\ h(q) + f_q'(x_*)(x - 1/\sqrt{2}) & \text{if } x \leq x_* \end{cases} \tag{5.75}$$

with x_* (dependent on q) such that

$$h(q) + f_q'(x_*)(x - 1/\sqrt{2}) = f_q(x_*)\,, \tag{5.76}$$

and where $f_q(x)$ is defined in Eq. (5.12).

The results of numerical tests done without noisy preprocessing in [133] and [143] strongly suggest that (5.74) actually gives the optimal bound on the average entropy for $q = 0$. Additional tests we did for this work did not find a counterexample for $q \neq 0$. But even without a proof of

[8] This is not a coincidence. The section in question proposes to bound the key rate using a lower bound on the conditional entropy in terms of the fidelity of Eve's marginal states. This is very closely related to the BB84 bound [142] and, in fact, all of the lower bounds we derive on the correlation terms $\big|\langle \bar{A}_x \otimes B\rangle\big|$ appearing in the BB84 bounds we use are also (typically tight) lower bounds on the fidelity of Eve's marginals following the qubit reduction.

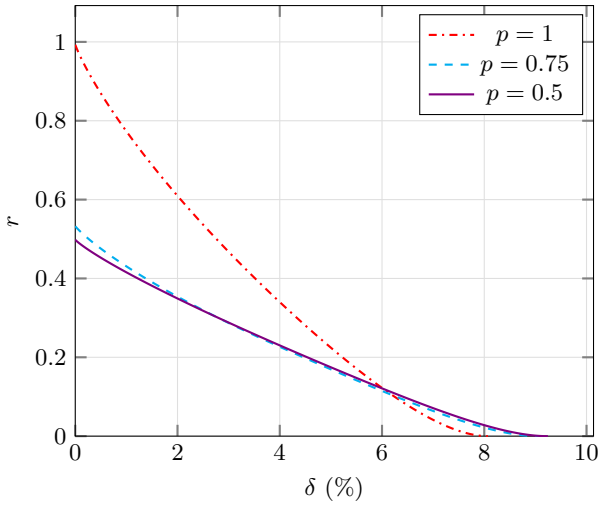

Figure 5.4: Lower bound on the Devetak-Winter rate as a function of the channel error rate δ, using an optimal noisy preprocessing.

optimality, as (5.74) is obtained with a known collective attack it gives an upper bound on the one-way asymptotic key rate with noisy preprocessing. A comparison of the key rates, optimized over q, using our numerical lower bound (already given in Figure 5.4) and using (5.74) is given in Figure 5.5 and shows the two to be very close. The threshold error rate obtained using (5.74) ranges from $\delta \approx 8.4447\%$ for $q = 0$ up to $\delta \approx 9.4756\%$ for $q \to 1/2$, and is compared with the threshold obtained using our numerical method in Figure 5.6.

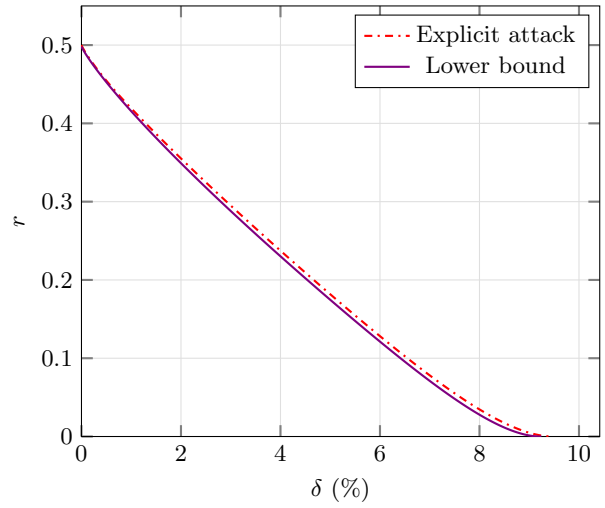

Figure 5.5: Comparison between the conjectured optimal attack and the lower bound on the Devetak-Winter rate as a function of the channel error rate δ, using an optimal noisy preprocessing.

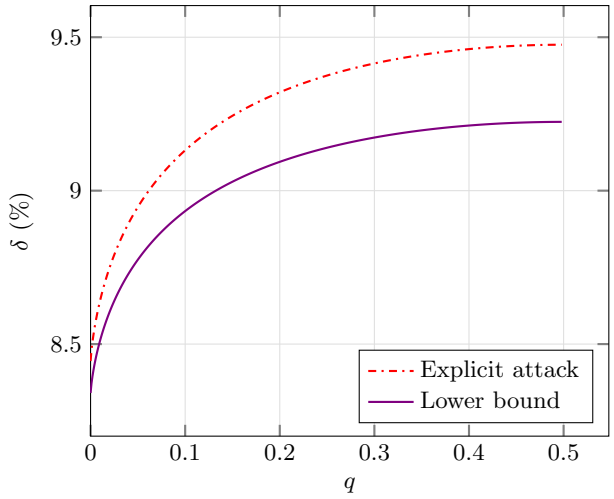

Figure 5.6: Thresholds for the channel error rate as a function of the noisy preprocessing computed using the conjectured optimal attack and our lower bound on the conditional entropy.

More refined loss analysis exploiting bias

Here, we consider a setup where we suppose that the main imperfection is that Alice's and Bob's devices have a detection efficiency that is less than perfect, i.e., we suppose that, in each protocol round, each of their devices outputs one of the regular outcomes ± 1 with probability η and outputs nothing, or a "nondetection" outcome $\emptyset$, with probability $1-\eta$. In order to use our approach, which strictly applies to protocols in which the measurements in the Bell test have binary outcomes, we map nondetection events resulting from the measurements A_1, A_2, B_1, and B_2 used to perform the Bell test to $+1$.

In this case we consider the usual, single-basis, version of the DI QKD protocol, but with different states and measurements. Similar to the Eberhard scheme [63], we suppose that Alice and Bob (ideally) share a partially-entangled two-qubit state

$$|\psi_\theta\rangle = \cos\left(\tfrac{\theta}{2}\right)|00\rangle + \sin\left(\tfrac{\theta}{2}\right)|11\rangle \,, \tag{5.77}$$

and that Alice and Bob (ideally) perform, respectively, two and three measurements

$$A_x = \cos(\varphi_{A,x})\mathrm{Z} + \sin(\varphi_{A,x})\mathrm{X}\,, \qquad x = 1,2 \tag{5.78}$$

$$B_y = \cos(\varphi_{B,y})\mathrm{Z} + \sin(\varphi_{B,y})\mathrm{X}\,, \qquad y = 1,2,3\,, \tag{5.79}$$

determined by angles $\varphi_{A,x}$ and $\varphi_{B,y}$ that we will optimize over when bounding the key rate[9]. Alice and Bob use the measurements A_1, A_2, B_1, and B_2 to estimate the CHSH expectation value and use A_1 and B_3 to generate the key.

As we are only considering the usual single-basis version of the protocol, the asymptotic key rate is

$$r = H(A_1^q|E) - H(A_1^q|B_3) \tag{5.80}$$

where the Shannon entropy of Alice's outcome conditioned on Bob,

$$H(A_1^q|B_3) = -\sum_{a,b} p(a,b)\log_2\big(p(a|b)\big)\,, \tag{5.81}$$

[9]Note that this is a slight generalization with respect to [133], which fixed A_1 and B_3 to Z.

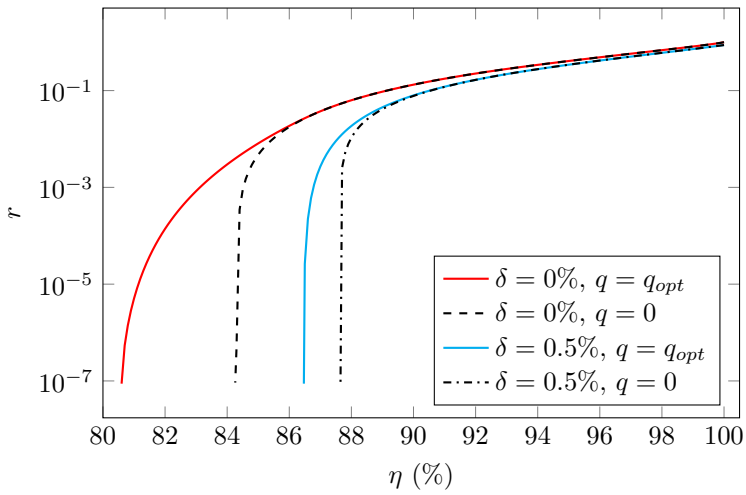

Figure 5.7: Key rate as a function of the detection efficiency with no channel error rate and with a little error rate.

depends on the joint probability $p(a, b)$ that Alice obtains the outcome $a \in \{+1, -1\}$ from measuring A_1 after mapping nondetection events to $+1$ and flipping the result with probability q, and Bob obtains the outcome $b \in \{+1, -1, \emptyset\}$ from measuring B_3 and possibly obtaining the loss outcome $\emptyset$ with probability $1 - \eta$.

To bound the key rate we need to bound $H(A_1^q|E)$. As mentioned above, mapping nondetection events deterministically to $+1$ and deliberately using a partially-entangled state bias Alice's and Bob's measurements to giving one of the outcomes more frequently than the other. We can exploit this by taking into account the expectation value $\langle A_1 \rangle$ of Alice's key generation measurement, in addition to the CHSH expectation value S, to derive a better lower bound on the entropy.

The expectation value $\langle A_1 \rangle$ can be taken into account using the qubit bound (5.47) and the convexification procedure discussed at the end of Section 5.2.5 and illustrated in Figure 5.2. Using this approach, we optimized the key rate numerically over the angles φ_{A_j}, φ_{B_k}, and θ. The optimized key rates, both assuming no noise and a white noise rate of $\delta = 0.5\%$ are represented both for $q = 0$ and with optimized q in Figure 5.7.

As one can see in the figure, the highest key rate is very small for a significant range of global detection efficiencies close to the threshold as a result of being obtained for values of q close to $1/2$ and very weakly entangled states. Due to this, the threshold detector efficiency above which a positive key rate can be certified is very sensitive and, for example, significantly worsened by the addition of even a small amount of depolarizing noise. To illustrate this, we plot the threshold global detection efficiency as a function of the error rate δ in Figure 5.8, where a comparison is provided with the earlier results of [133] using the analytic entropy bound for the asymmetric CHSH expressions.

Table 5.2 gives the thresholds on the detection efficiency that we find using our approach for different values of q assuming no additional noise. We include in the table both the thresholds for which we can certify a positive key rate and the ones obtained using our conjecture regarding the convex envelope of the qubit bound. The small discrepancy between the two values, particularly for larger values of q, is due to the difficulty of numerically certifying the key rate accurately when the key rate becomes very small (the key rate for the last column of Table 5.2 is of $O(10^{-12})$). Indeed to certify the entropy to a very high precision using a discretized qubit bound requires using a very dense covering, which at some point becomes too time-consuming computationally.

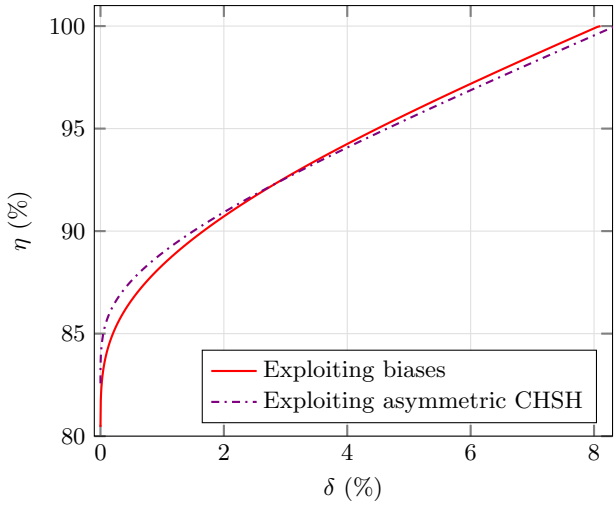

Figure 5.8: Threshold detection efficiency η as a function of the channel error rate δ.

	$q = 0$	$q = 0.2$	$q = 0.3$	$q = 0.49$
Certified	84.2149	80.4642	80.3411	80.2593
Conjectured	84.2147	80.4362	80.3046	80.2283

Table 5.2: Threshold detection efficiencies (%) for different probabilities q of flipping Alice's outcome assuming no channel noise. For q between 0.49 and 0.5, we did not observe an improvement of the threshold up to the precision reported in the table.

This issue however only affects the certification of extremely small asymptotic key rates, such as the long tail observed in Figure 5.7, which are probably too low to be of practical value and likely to be dwarfed by the difference made by even small amounts of noise or corrections due to finite-key effects. To illustrate this, in Table 5.3 we report the detection efficiency thresholds in the presence of a channel noise rate of $\delta = 0.5$. In this case, the thresholds using the conjectured convex envelope and those that can be properly certified are the same up to the precision to which we report the results.

Finally, we remark that the qubit bound (5.47) is tight in $\langle A_1 \rangle$ and S for all q as there is an explicit attack, described in Section 5.3.5, that saturates it. This means that our conjecture regarding the convex envelope of the qubit bound represents a valid attack yielding upper bounds on the key rate (as it corresponds to an explicit mixture of two-qubit strategies). This means that the certified bounds that we report in Table 5.3 are, up to the precision we use, optimal in terms of $\langle A_1 \rangle$ and S, and that the second line of Table 5.2 corresponds to the minimal detection thresholds one can hope to attain using only information about $\langle A_1 \rangle$ and S.

	$q = 0$	$q = 0.2$	$q = 0.3$	$q = 0.49$
$\delta = 0.5\%$	87.6017	86.5842	86.5013	86.4490

Table 5.3: Certified threshold detection efficiencies (%) obtained for different probabilities q of flipping Alice's outcome and with $\delta = 0.5\%$ of channel error rate. We do not observe a difference with the conjectured case up to the precision reported in the table.

5.3 Technical Details

5.3.1 Derivation of BB84 bound with bias

The BB84 entropy bound (5.13) is a generalization of the two bounds (5.10) and (5.11), which give the special cases of (5.13) with $\langle A_1 \rangle = 0$ and both with $\langle A_1 \rangle = 0$ and no noisy preprocessing ($q = 0$). It can be derived, in a way that also confirms the monotonicity property, essentially by modifying the symmetrization step in the derivation done in Section4.2 of the paper [133]. We do this in detail here.

As in the derivation of [133], we suppose that Alice, Bob, and Eve share a pure tripartite state

$$|\psi\rangle_{ABE} = |0\rangle_A |\psi_0\rangle_{BE} + |1\rangle_A |\psi_1\rangle_{BE} \,, \tag{5.82}$$

where $|0\rangle$ and $|1\rangle$ are the eigenstates of A_1, which we identify here with Z, and $|\psi_0\rangle$ and $|\psi_1\rangle$ are arbitrary (and not necessarily orthogonal) states shared by Bob and Eve normalized so that

$$\langle\psi_0|\psi_0\rangle + \langle\psi_1|\psi_1\rangle = 1 \,. \tag{5.83}$$

After Alice measures $A_1 = \mathrm{Z}$ and flips the outcome with probability q, the correlations between Alice and Eve are described by the classical-quantum state

$$\tau_{AE} = [0]_A \otimes (\bar{q}\psi_0^E + q\psi_1^E) + [1]_A \otimes (q\psi_0^E + \bar{q}\psi_1^E) \,, \tag{5.84}$$

where $\bar{q} = 1 - q$ and $\psi_a^E = \mathrm{Tr}_B[\psi_a]$ are the partial traces of the states $|\psi_a\rangle$ accessible to Eve.

Now, since renaming the outcomes does not change the entropy, the conditional entropy $H(\mathrm{Z}|E) = H(\mathrm{Z}E) - H(E)$ computed on the above state is the same as the conditional entropy computed on

$$\tau'_{AE} = [1]_A \otimes (\bar{q}\psi_0^E + q\psi_1^E) + [0]_A \otimes (q\psi_0^E + \bar{q}\psi_1^E) \,, \tag{5.85}$$

which is the same state as above except that we have swapped $[0]_A$ and $[1]_A$. They in addition have the same entropy as a partly symmetrized state,

$$\bar{\tau}_{AEF} = \bar{p}\,\tau_{AE} \otimes [0]_F + p\,\tau'_{AE} \otimes [1]_F \,, \tag{5.86}$$

for any probability p and $\bar{p} = 1 - p$, since

$$H(\mathrm{Z}|EF)_{\bar{\tau}} = \bar{p}\,H(\mathrm{Z}|E)_\tau + p\,H(\mathrm{Z}|E)_{\tau'} = H(\mathrm{Z}|E)_\tau \,. \tag{5.87}$$

The above state, written out explicitly, is

$$\begin{aligned}\bar{\tau}_{AEF} = [0]_A \otimes \Big[&\bar{p}(\bar{q}\psi_0^E + q\psi_1^E) \otimes [0]_F \\ &+ p(q\psi_0^E + \bar{q}\psi_1^E) \otimes [1]_F\Big] \\ + [1]_A \otimes \Big[&\bar{p}(q\psi_0^E + \bar{q}\psi_1^E) \otimes [0]_F \\ &+ p(\bar{q}\psi_0^E + q\psi_1^E) \otimes [1]_F\Big] \,.\end{aligned} \tag{5.88}$$

We rewrite this as

$$\bar{\tau}_{AEF} = [0]_A \otimes (\bar{q}\sigma_= + q\sigma_{\neq}) + [1]_A \otimes (q\sigma_= + \bar{q}\sigma_{\neq}) \tag{5.89}$$

with the (unnormalized) states

$$\sigma_= = \bar{p}\,\psi_0^E \otimes [0]_F + p\,\psi_1^E \otimes [1]_F \,, \tag{5.90}$$

$$\sigma_{\neq} = \bar{p}\,\psi_1^E \otimes [0]_F + p\,\psi_0^E \otimes [1]_F \,. \tag{5.91}$$

The state can be obtained as the marginal of an extended one,

$$\begin{aligned}\bar{\tau}_{ABEE'FF'} = [0]_A \otimes (\bar{q}\chi_{=} + q\chi_{\neq}) \\ + [1]_A \otimes (q\chi_{=} + \bar{q}\chi_{\neq}) ,\end{aligned} \tag{5.92}$$

where $|\chi_{=}\rangle , |\chi_{\neq}\rangle \in \mathcal{H}_B \otimes \mathcal{H}_E \otimes \mathcal{H}_{E'} \otimes \mathcal{H}_F \otimes \mathcal{H}_{F'}$ are unnormalized pure states

$$|\chi_{=}\rangle = \sqrt{\bar{p}}\,|\psi_0\rangle\,|\phi_0\rangle\,|00\rangle + \sqrt{p}\,|\psi_1'\rangle\,|\phi_1\rangle\,|11\rangle \ , \tag{5.93}$$

$$|\chi_{\neq}\rangle = \sqrt{\bar{p}}\,|\psi_1'\rangle\,|\phi_1\rangle\,|00\rangle + \sqrt{p}\,|\psi_0\rangle\,|\phi_0\rangle\,|11\rangle \ , \tag{5.94}$$

in which

$$|\psi_1'\rangle = e^{i\varphi} B \otimes \mathbb{1}_E\,|\psi_1\rangle \in \mathcal{H}_B \otimes \mathcal{H}_E , \tag{5.95}$$

where B is a Hermitian unitary operator (thus satisfying $B^2 = \mathbb{1}_B$) acting on $\mathcal{H}_B$ and φ is a phase chosen such that $\langle\psi_0|\psi_1'\rangle$ is real and nonnegative, and

$$|\phi_0\rangle , |\phi_1\rangle \in \mathcal{H}_{E'} \tag{5.96}$$

are normalized states chosen to have some nonnegative real overlap $\langle\phi_0|\phi_1\rangle = \lambda_{\mathrm{X}} \in [0,1]$.

Using that the conditional entropy cannot increase if we extend the Hilbert space being conditioned on, direct calculation of the conditional entropy on the state (5.92) gives

$$\begin{aligned} H(\mathrm{Z}|E)_\tau &= H(\mathrm{Z}|EF)_{\bar{\tau}} \\ &\geq H(\mathrm{Z}|BEE'FF')_{\bar{\tau}} \\ &= S(\bar{\tau}_{ABEE'FF'}) - S(\chi_{=} + \chi_{\neq}) \\ &= H(\boldsymbol{\lambda}) - \phi\Big(\sqrt{Z'^2 + X'^2}\Big) , \end{aligned} \tag{5.97}$$

where

$$Z' = \|\chi_{=}\| - \|\chi_{\neq}\| \equiv \langle\chi_{=}|\chi_{=}\rangle - \langle\chi_{\neq}|\chi_{\neq}\rangle \ , \tag{5.98}$$

$$X' = 2|\langle\chi_{=}|\chi_{\neq}\rangle| , \tag{5.99}$$

and $H(\boldsymbol{\lambda}) = -\sum_{jk} \lambda_{jk} \log_2(\lambda_{jk})$ is the Shannon entropy associated to the four eigenvalues of (5.92),

$$\lambda_{11} = \frac{1}{4}\Big[1 + QZ' + \sqrt{R'^2 + 2QZ'}\Big] , \tag{5.100}$$

$$\lambda_{12} = \frac{1}{4}\Big[1 - QZ' + \sqrt{R'^2 - 2QZ'}\Big] , \tag{5.101}$$

$$\lambda_{21} = \frac{1}{4}\Big[1 - QZ' - \sqrt{R'^2 - 2QZ'}\Big] , \tag{5.102}$$

$$\lambda_{22} = \frac{1}{4}\Big[1 + QZ' - \sqrt{R'^2 + 2QZ'}\Big] , \tag{5.103}$$

where Q is related to the amount of noisy preprocessing applied by

$$Q = \bar{q} - q = 1 - 2q \tag{5.104}$$

and

$$R' = \sqrt{Z'^2 + Q^2 + (1 - Q^2)X'^2} . \tag{5.105}$$

We can factorize the four eigenvalues above as $\lambda_{jk} = p_j p'_k$ with

$$p_1 = \frac{1}{2} + \frac{1}{4}\big(R'_+ + R'_-\big)\,, \tag{5.106}$$

$$p_2 = \frac{1}{2} - \frac{1}{4}\big(R'_+ + R'_-\big)\,, \tag{5.107}$$

$$p'_1 = \frac{1}{2} + \frac{1}{4}\big(R'_+ - R'_-\big)\,, \tag{5.108}$$

$$p'_2 = \frac{1}{2} - \frac{1}{4}\big(R'_+ - R'_-\big)\,, \tag{5.109}$$

and

$$R'_\pm = \sqrt{R'^2 \pm 2QZ'}\,, \tag{5.110}$$

so that $H(\boldsymbol{\lambda}) = H(\boldsymbol{p}) + H(\boldsymbol{p}')$. This allows us to express the qubit entropy bound more concisely as

$$H(\mathrm{Z}|E) \geq g_q(Z', X') \tag{5.111}$$

with

$$\begin{aligned} g_q(Z', X') = \phi\big(\tfrac{1}{2}(R'_+ + R'_-)\big) + \phi\big(\tfrac{1}{2}(R'_+ - R'_-)\big) \\ - \phi\big(\sqrt{Z'^2 + X'^2}\big) \end{aligned} \tag{5.112}$$

and

$$R'_\pm = \sqrt{(Q \pm Z')^2 + (1 - Q^2)X'^2}\,. \tag{5.113}$$

At this point, we have recovered the form of the function g_q defined in Section 5.2.1. To complete the derivation note that, from the definitions of $|\chi_=\rangle$ and $|\chi_\neq\rangle$ we have

$$\begin{aligned} Z' &= \|\chi_0\| - \|\chi_1\| \\ &= \bar{p}\|\psi_0\| + p\|\psi_1\| - \bar{p}\|\psi_1\| - p\|\psi_0\| \\ &= \lambda_\mathrm{Z}\big(\|\psi_0\| - \|\psi_1\|\big) \\ &= \lambda_\mathrm{Z}\langle A_1\rangle\,, \end{aligned} \tag{5.114}$$

where $\lambda_\mathrm{Z} \in [-1, 1]$ is related to the symmetrization-step probability by $\lambda_\mathrm{Z} = \bar{p} - p$, and that

$$\begin{aligned} \langle\chi_=|\chi_\neq\rangle &= \bar{p}\,\langle\psi_0|\psi'_1\rangle\,\langle\phi_0|\phi_1\rangle + p\,\langle\psi'_1|\psi_0\rangle\,\langle\phi_1|\phi_0\rangle \\ &= \lambda_\mathrm{X}\, e^{i\varphi}\,\langle\psi_0|\, B \otimes \mathbb{1}_E\,|\psi_1\rangle \\ &= \lambda_\mathrm{X}\,\big|\mathrm{Re}\big[\langle\psi_0|\, B \otimes \mathbb{1}_E\,|\psi_1\rangle\big]\big|\,, \end{aligned} \tag{5.115}$$

where we recall that we set $\langle\phi_0|\phi_1\rangle = \lambda_\mathrm{X} \in [0, 1]$, while

$$\langle \mathrm{X} \otimes B\rangle = 2\,\mathrm{Re}\big[\langle\phi_0|\, B \otimes \mathbb{1}_E\,|\phi_1\rangle\big]\,, \tag{5.116}$$

so that

$$2\,\langle\chi_=|\chi_\neq\rangle = \lambda_\mathrm{X}\,|\langle \mathrm{X} \otimes B\rangle|\,. \tag{5.117}$$

Putting all this together and recalling that we identify A_1 with Z, and can choose $\bar{A}_1 = \mathrm{X}$, means that we finally get

$$H(A_1|E) \geq g_q\big(\lambda_\mathrm{Z}\langle A_1\rangle,\, \lambda_\mathrm{X}\big|\langle\bar{A}_1 \otimes B\rangle\big|\big) \tag{5.118}$$

for all $-1 \leq \lambda_\mathrm{Z}, \lambda_\mathrm{X} \leq 1$ (as the derivation we have given applies for any values of the symmetrization probability p and overlap $\langle\phi_0|\phi_1\rangle$ we may wish to use). This confirms that the inequality

$$H(A_1|E) \geq g_q(Z, X) \tag{5.119}$$

holds for any (real) numbers satisfying

$$|Z| \leq |\langle A_1\rangle| \qquad \text{and} \qquad |X| \leq \big|\langle\bar{A}_1 \otimes B\rangle\big|\,. \tag{5.120}$$

5.3.2 Analytic solution for p = 1/2

Here we derive in detail the average entropy bound for the two-basis protocol in the case that Alice's measurements are used equiprobably. When $p = 1/2$, the minimization problem (5.41) in Section 5.2.4 simplifies to

$$\begin{aligned} \text{minimize} \quad & f(\lambda,\mu,\varphi_A) = \sin\left(\tfrac{\varphi_A}{2}\right)^2\lambda^2 + \cos\left(\tfrac{\varphi_A}{2}\right)^2\mu^2 \\ \text{subject to} \quad & \left|\cos\left(\tfrac{\varphi_A}{2}\right)\right||\lambda| + \left|\sin\left(\tfrac{\varphi_A}{2}\right)\right||\mu| \geq S/2 \\ & \lambda^2 \leq 1 \\ & \mu^2 \leq 1\,, \end{aligned} \tag{5.121}$$

where we have reintroduced the angle φ_A from earlier in the section explicitly and used that the single constraint involving the variable Δ becomes irrelevant. As we stated in Section 5.2.4 and show here, the above problem can be solved analytically subject to finding the root of a degree four polynomial.

In the following, we will assume that $S > 2$, since the solution to the classical case $S = 2$ is trivially ${E_{\frac{1}{2}}}^2 = 0$.

First, we note that, as our problem is invariant under the transformations $\lambda \mapsto -\lambda$ and $\mu \mapsto -\mu$ and that, for $S > 2$, the points $\mu = 0$ or $\lambda = 0$ do not satisfy the first constraint

$$\left|\cos\left(\tfrac{\varphi_A}{2}\right)\right||\lambda| + \left|\sin\left(\tfrac{\varphi_A}{2}\right)\right||\mu| \geq S/2\,, \tag{5.122}$$

we can replace the constraints $\lambda^2 \leq 1$ and $\mu^2 \leq 1$ with $0 < \lambda \leq 1$ and $0 < \mu \leq 1$.

Moreover, the problem is also invariant under the transformation $\varphi_A \mapsto 2\pi - \varphi_A$, meaning that for all solutions such that $\varphi_A \in [0,\pi]$, there exists an equivalent solution in $[\pi, 2\pi]$. Thus, we can restrict the domain of φ_A to be $0 < \varphi_A < \pi$, where we excluded the boundaries since the cases $\varphi_A = 0, \pi$ are not in agreement with $S > 2$.

The function that we need to minimize can be rewritten as

$$f(\lambda,\mu,\varphi_A) = \frac{\lambda^2}{2}\big(1 - \cos(\varphi_A)\big) + \frac{\mu^2}{2}\big(1 + \cos(\varphi_A)\big)\,. \tag{5.123}$$

Let us look for a minimum for our function by checking where its derivatives are zero. We start with

$$\frac{d}{d\mu} f(\lambda,\mu,\varphi_A) = \mu\big(1 + \cos(\varphi_A)\big)\,. \tag{5.124}$$

Here, $\frac{d}{d\mu} f(\lambda,\mu,\varphi_A) = 0$ if and only if $\mu = 0$ or $\varphi_A = \pi$. These points are not part of the restricted domain that we are considering. We conclude that the minimum must be at the boundaries of our domain. From now on, we will analyse this case.

Case 1: We consider the boundary $\lambda = 1$. We have

$$f(1,\mu,\varphi_A) = \frac{1+\mu^2}{2} + \frac{\cos(\varphi_A)}{2}(\mu^2 - 1) \tag{5.125}$$

and

$$\frac{d}{d\mu} f(1,\mu,\varphi_A) = \mu\big(1 + \cos(\varphi_A)\big)\,, \tag{5.126}$$

thus $\frac{d}{d\mu} f(\lambda,\mu,\varphi_A) = 0$ if and only if $\mu = 0$ or $\varphi_A = \pi$. Such solutions are not in the domain.

Case 2: We consider the boundary $\mu = 1$. Analogously, we obtain non-feasible solutions.

Case 3: We consider the boundary $\cos\left(\frac{\varphi_A}{2}\right)\lambda + \sin\left(\frac{\varphi_A}{2}\right)\mu = S/2$. This region is the one in which

$$\begin{aligned}\mu_* &= \lambda\frac{\sin(\varphi_A)}{\cos(\varphi_A)-1} - S\frac{\sin\left(\frac{\varphi_A}{2}\right)}{\cos(\varphi_A)-1} \\ &= \lambda\frac{\sqrt{1-x^2}}{x-1} - \frac{S}{\sqrt{2}}\frac{\sqrt{1-x}}{x-1}\,,\end{aligned} \tag{5.127}$$

where we made the change of variable $x = \cos(\varphi_A)$. The domain of x is $-1 < x < 1$.

We have

$$f(\lambda, \mu_*, x) = \lambda^2\frac{1+x^2}{1-x} - \lambda\frac{S(1+x)^{3/2}}{\sqrt{2}(1-x)} + \frac{S^2(1+x)}{4(1-x)} \tag{5.128}$$

and

$$\frac{d}{d\lambda}f(\lambda, \mu_*, x) = 2\lambda\frac{1+x^2}{1-x} - \frac{S(1+x)^{3/2}}{\sqrt{2}(1-x)}\,. \tag{5.129}$$

Now, recalling that we assumed $x \neq 1$, we have that $\frac{d}{d\lambda}f(\lambda, \mu_*, x) = 0$ iff

$$\lambda = \frac{S(x+1)^{3/2}}{2\sqrt{2}(x^2+1)} = \lambda_*\,. \tag{5.130}$$

Thus,

$$f(\lambda_*, \mu_*, x) = \frac{S^2}{8}\frac{1-x^2}{1+x^2}\,, \tag{5.131}$$

which is a concave function of x, meaning the minimum is at the intersection between boundaries.

Case 3+1: We intersect the boundary of case 3 with $\lambda = 1$. We get

$$\mu_* = \frac{\sqrt{1-x^2}}{x-1} - \frac{S}{\sqrt{2}}\frac{\sqrt{1-x}}{x-1}\,. \tag{5.132}$$

Here, requiring $\mu_* \leq 1$, we obtain the condition

$$-\frac{S}{4}\sqrt{8-S^2} \leq x \leq \frac{S}{4}\sqrt{8-S^2}\,. \tag{5.133}$$

We have

$$f(1, \mu_*, x) = \frac{x^2+1}{1-x} - \frac{S(x+1)^{3/2}}{\sqrt{2}(1-x)} + \frac{S^2(x+1)}{4(1-x)} \tag{5.134}$$

and

$$\begin{aligned}&\frac{d}{dx}f(1, \mu_*, x) = \\ &\quad\frac{4x(2-x) + 2(S^2+2) + S(x-5)\sqrt{2(1+x)}}{4(x-1)^2}\,,\end{aligned} \tag{5.135}$$

hence, since $x \neq 1$, $\frac{d}{dx}f(1, \mu_*, x) = 0$ iff

$$4x(2-x) + 2(S^2+2) + S(x-5)\sqrt{2(1+x)} = 0\,. \tag{5.136}$$

Case 3+2: We intersect the boundary of case 3 with $\mu = 1$. Here, one can check that we obtain the same result as in case 3+1.

Case 1+2: We consider $\lambda = \mu = 1$. With this choice we have $E_{\frac{1}{2}}{}^2 = 1\ \forall\varphi_A$. This region of parameters does not contain in general the absolute minimum.

We conclude that the solution to the optimization problem must be the one of case 3+1 (or equivalently 3+2). If there is more than one solution to Eq. (5.136) satisfying the constraints (5.133), then we take the smallest one.

We used Mathematica to find the roots of Eq. (5.136) analytically. Moreover, imposing the constraints (5.133) and $S > 2$, we found a single solution. We used the resulting expression for the computations for $p = 1/2$ done in Section 5.2.6.

5.3.3 Optimality of CHSH for the two-basis protocol

In the case that the bases are used equiprobably, i.e., $p = 1/2$, the symmetries of the two-basis DI QKD protocol studied in Section 5.2.6 imply that the CHSH Bell expectation value alone already gives the optimal bound on the average conditional entropy

$$H(A_x|XE) \propto \tfrac{1}{2}H(A_1|E) + \tfrac{1}{2}H(A_2|E) \tag{5.137}$$

for the optimal CHSH-violating correlations attenuated by white noise. The reason for this is that, given any quantum strategy giving a particular value of the average entropy and CHSH expectation value, one can construct a new symmetrized strategy giving the same entropy and CHSH expectation value.

To see this, let us suppose we have a particular quantum strategy $\mathcal{Q} = (\rho_{ABE}, A_1, A_2, B_1, B_2)$. We note first that both conditional entropies $H(A_x|E)$ and the CHSH expectation value $S = \langle A_1B_1\rangle + \langle A_1B_2\rangle + \langle A_2B_1\rangle - \langle A_2B_2\rangle$ are unchanged if we flip all the measurements, i.e., do $A_x \mapsto -A_x$ and $B_y \mapsto -B_y$. By randomly and equiprobably using these two strategies we can force Alice's and Bob's local outcomes to become equiprobable. This corresponds to using a new strategy $\mathcal{Q}' = (\rho'_{ABE}, A'_1, A'_2, B'_1, B'_2)$ with

$$A'_x = A_x \oplus -A_x\,, \tag{5.138}$$
$$B'_y = B_y \oplus -B_y\,, \tag{5.139}$$
$$\rho'_{ABE} = \tfrac{1}{2}\rho_{ABE} \oplus \tfrac{1}{2}\rho_{ABE}\,, \tag{5.140}$$

for which the CHSH expectation value and the values of the entropies are unchanged, but for which $\langle A'_x\rangle = \langle B'_y\rangle = 0$.

Next, we use that the average entropy and CHSH both remain unchanged under the two transformations

$$T_1 : \begin{cases} A_1 \mapsto A_1 \\ A_2 \mapsto -A_2 \\ B_1 \mapsto B_2 \\ B_2 \mapsto B_1 \end{cases} \qquad T_2 : \begin{cases} A_1 \mapsto A_2 \\ A_2 \mapsto A_1 \\ B_1 \mapsto B_1 \\ B_2 \mapsto -B_2 \end{cases}, \tag{5.141}$$

as well as their composition $T_2 \circ T_1$. By randomly using the strategy $\mathcal{Q}'$ with neither, either one,

or both transformations applied, we construct a new strategy $\mathcal{Q}'' = (\rho''_{ABE}, A''_1, A''_2, B''_1, B''_2)$ with

$$A''_1 = A'_1 \oplus A'_1 \oplus A'_2 \oplus A'_2\,, \tag{5.142}$$
$$A''_2 = A'_2 \oplus -A'_2 \oplus A'_1 \oplus -A'_1\,, \tag{5.143}$$
$$B''_1 = B'_1 \oplus B'_2 \oplus B'_1 \oplus -B'_2\,, \tag{5.144}$$
$$B''_2 = B'_2 \oplus B'_1 \oplus -B'_2 \oplus B'_1\,, \tag{5.145}$$
$$\rho''_{ABE} = \tfrac{1}{4}\rho'_{ABE} \oplus \tfrac{1}{4}\rho'_{ABE} \oplus \tfrac{1}{4}\rho'_{ABE} \oplus \tfrac{1}{4}\rho'_{ABE}\,, \tag{5.146}$$

for which

$$\langle A''_1 B''_1\rangle = \langle A''_1 B''_2\rangle = \langle A''_2 B''_1\rangle = -\langle A''_2 B''_2\rangle = S/4\,. \tag{5.147}$$

As, given any strategy $\mathcal{Q}$, we can in this way always construct a strategy $\mathcal{Q}''$ with the same average entropy and CHSH expectation value, but satisfying $\langle A''_x\rangle = \langle B''_y\rangle = 0$ and $\langle A''_1 B''_1\rangle = \langle A''_1 B''_2\rangle = \langle A''_2 B''_1\rangle = -\langle A''_2 B''_2\rangle$, we can infer that these constraints, if they are satisfied for real correlations, do not contain any information other than the CHSH expectation value that can be used to improve the entropy bound.

5.3.4 Explicit attack for the two-basis protocol

We describe here an explicit attack for the two-basis protocol in the case $p = 1/2$, which we conjecture to be optimal.

Suppose that Alice, Bob, and Eve share the optimal symmetric BB84 attack state

$$\begin{aligned}|\Psi\rangle_{ABE} = \frac{1}{2}\Big[&(1+E)\,\big|\phi^+\big\rangle_{AB}\,|{+}{+}\rangle_E \\ &+ \sqrt{1-E^2}\,\big|\phi^-\big\rangle_{AB}\,|{+}{-}\rangle_E \\ &+ \sqrt{1-E^2}\,\big|\psi^+\big\rangle_{AB}\,|{-}{+}\rangle_E \\ &+ (1-E)\,\big|\psi^-\big\rangle_{AB}\,|{-}{-}\rangle_E\ \Big]\,,\end{aligned} \tag{5.148}$$

where $|\phi^\pm\rangle$ and $|\psi^\pm\rangle$ are the four Bell states, depending on some number $0 \le E \le 1$. Its marginal once Eve is traced out is

$$\Psi_{AB} = \frac{1}{4}\Big[\mathbb{1}\otimes\mathbb{1} + E\,\mathrm{X}\otimes\mathrm{X} - E^2\,\mathrm{Y}\otimes\mathrm{Y} + E\,\mathrm{Z}\otimes\mathrm{Z}\Big]\,. \tag{5.149}$$

By measuring $A_1 = \mathrm{Z}$, $A_2 = \mathrm{X}$, and $B_{1,2} = (\mathrm{Z}\pm\mathrm{X})/\sqrt{2}$, the highest possible CHSH expectation value of $S = 2\sqrt{2}E$ with this state is obtained. Direct computation of the conditional entropies after Alice measures Z and X on this state gives

$$\tfrac{1}{2}H(A^q_1|E) + \tfrac{1}{2}H(A^q_2|E) = f_q\big(S/\sqrt{8}\big) \tag{5.150}$$

where f_q is the same BB84 bound with noisy preprocessing used earlier and given by Eq. (5.12). This is too high to be the optimal bound on the average entropy for all S, as the correct bound must attain $h(q)$ at $S = 2$. But we can construct a plausible strategy by taking a convex mixture (similar to the construction in Section 2 of [133]) of the strategy just described with a deterministic one giving $\big(H(A^q_X|XE),\, S\big) = (h(q),\, 2)$. This gives

$$\tfrac{1}{2}H(A^q_1|E) + \tfrac{1}{2}H(A^q_2|E) = \bar{f}_q\big(S/\sqrt{8}\big)\,, \tag{5.151}$$

where

$$\bar{f}_q(x) = \begin{cases} f_q(x) & \text{if } x \ge x_* \\ h(q) + f'_q(x_*)(x - 1/\sqrt{2}) & \text{if } x \le x_* \end{cases} \tag{5.152}$$

with x_* (dependent on q) such that

$$h(q) + f_q'(x_*)(x - 1/\sqrt{2}) = f_q(x_*)\,. \tag{5.153}$$

5.3.5 Explicit attack saturating the qubit entropy bound with bias (5.47)

One can verify that the qubit bound (5.47) is attained with measurements and an initial state of the form

$$A_1 = \mathrm{Z}\,, \tag{5.154}$$
$$A_2 = \mathrm{X}\,, \tag{5.155}$$
$$B_1 = \cos\left(\tfrac{\varphi_B}{2}\right)\mathrm{Z} + \sin\left(\tfrac{\varphi_B}{2}\right)\mathrm{Z}\,, \tag{5.156}$$
$$B_2 = \cos\left(\tfrac{\varphi_B}{2}\right)\mathrm{Z} - \sin\left(\tfrac{\varphi_B}{2}\right)\mathrm{Z}\,, \tag{5.157}$$

and

$$|\Psi\rangle_{ABE} = \cos\left(\tfrac{\theta}{2}\right)|00\rangle_{AB}\,|\psi_0\rangle_E + \sin\left(\tfrac{\theta}{2}\right)|11\rangle_{AB}\,|\psi_1\rangle_E \tag{5.158}$$

with

$$\cos(\theta) = \langle A_1\rangle\,, \tag{5.159}$$
$$\sin(\theta)\,\langle\psi_0|\psi_1\rangle = \sqrt{S^2/4 - 1}\,, \tag{5.160}$$
$$\cos\left(\tfrac{\varphi_B}{2}\right) = 2/S\,, \tag{5.161}$$
$$\sin\left(\tfrac{\varphi_B}{2}\right) = \sqrt{1 - 4/S^2}\,. \tag{5.162}$$

Note that, because $\cos(\theta)^2 + \sin(\theta)^2|\langle\psi_0|\psi_1\rangle|^2 \leq 1$, (5.159) and (5.160) are only consistent with each other if

$$\langle A_1\rangle^2 + S^2/4 \leq 2\,, \tag{5.163}$$

but this is a known boundary of the quantum set [144, 145].

5.4 Numerical DI QKD security analysis via Eve's guessing probability

The semi-analytic method described earlier in this chapter is limited to protocols where each party has only two inputs and two outputs. Additionally, even for two input-two output scenarios, if we want to compute $H(K_A|E)$ by constraining Eve's operations using the full probability distribution $p(a,b|x,y)$ rather than simpler witnesses like S, the method becomes significantly more complicated. This is because it requires identifying qubit bounds on the entropy based on the entire probability distribution. Even if such bounds were found, we must also verify their convexity with respect to each variable in the bound. Due to these complexities, it is beneficial to introduce numerical methods that can handle more general scenarios.

A very popular way of bounding the conditional entropy of the random variable K_A given Eve's information is the evaluation of the so called *min entropy*. This method is extremely versatile in the sense that it can be used to compute the key rate of any protocol, not only in a 2-input/2output scenario. Moreover, it has the advantage of providing a clear physical picture of the reason why it works, as we will see in this section. Anyway, the bounds that it provides tend to be very far from being tight. In this section, we will outline explain this approach.

Let's consider Eve's task of correctly guessing the random variable K_A representing Alice's measurements outcomes. Eve aims to generate an outcome e that matches Alice's outcome. If

Alice performs measurement x with probabilities $\mu(x)$ and reveals x afterward, Eve's guessing probability can be expressed as

$$P_G(K_A|E) = \sum_{a,x} \mu(x)p(a|x)p(e=a|a), \tag{5.164}$$

where $p(e=a|a)$ denotes the probability of Eve obtaining the same outcome as Alice when Alice gets outcome a. To achieve this, Eve holds a portion of the global state ρ_{AE} and she performs a measurement $\{E_e\}_e$ on her subsystem. The number of outcomes of Eve needs to be the same as the number of the possible combinations of a, x, hence we will denote her measurement as $\{E_{a,x}\}_{a,x}$. At this point, we can write

$$p(a|x)p(e=a|a) = \mathrm{Tr}\left(\rho_{AE} M_{a|x} \otimes E_{a,x}\right). \tag{5.165}$$

Consider the scenario where Alice and Bob observe input-output statistics $p(a,b|x,y)$ during an experiment, while Eve wants to correctly guess Alice's outcomes. The maximum guessing probability $P_G^{\max}(K_A|E)$ for Eve, consistent with this data, is determined by solving the following optimization problem:

$$\begin{aligned}
&\text{maximize} && \sum_{a,x} \mu(x) \langle\psi| M_{a|x} \otimes \mathbb{1}_B \otimes E_{a,x} |\psi\rangle && (5.166)\\
&\text{subject to} && \langle\psi| M_{a|x} \otimes N_{b|y} |\psi\rangle = p(a,b|x,y) && (5.167)\\
& && M_{a|x} \geq 0, \quad N_{b|y} \geq 0, \quad E_{a,x} \geq 0, && (5.168)\\
& && \sum_a M_{a|x} = \mathbb{1}, \quad \sum_b N_{b|y} = \mathbb{1}, && (5.169)\\
& && \sum_{a,x} E_{a,x} = \mathbb{1}, \quad \langle\psi| \mathbb{1} |\psi\rangle = 1. && (5.170)
\end{aligned}$$

This optimization problem, a non-commuting polynomial optimization problem, can be relaxed to a converging hierarchy of SDPs using the NPA hierarchy [146, 147]. Consequently, upper bounds on Eve's guessing probability can be computed, independent of the specific forms of Alice's and Bob's measurements and the quantum state they hold.

These upper bounds can be used to compute a bound on the conditional Von Neumann entropy of the variable K_A given Eve's information. Let us assume that the random variable K_A is given by the two outcomes of the measurements of Alice, then we define the conditional min entropy of K_A given Eve as

$$H_{\min}(K_A|E) = \min_{\{E_{a,x}\}} \left(-\log_2 \left(\sum_{a,x} \mu(x) \,\mathrm{Tr}\left(\rho_{AE} M_{a|x} \otimes E_{a,x}\right) \right) \right). \tag{5.171}$$

For a given state ρ_{AE} and measurements $M_{a|x}$, the Von Neumann entropy is bounded by the min entropy

$$H(K_A|E)_{\rho_{AE}, M_{a|x}} \geq H_{\min}(K_A|E)_{\rho_{AE}, M_{a|x}}. \tag{5.172}$$

This means that we can compute a device-independent bound on the Von Neumann as

$$H(K_A|E) \geq \min_{\rho_{AE}, M_{a|x}} H_{\min}(K_A|E)_{\rho_{AE}, M_{a|x}} = -\log_2 P_G^{\max}(K_A|E). \tag{5.173}$$

The intuition behind this bound is that the more accurately Eve can guess Alice's outcomes used to construct the secret key, the lower the secret key rate will be.

5.5 Numerical DI QKD security analysis via Gauss-Radau quadrature

More recently, Peter Brown, Hamza Fawzi and Omar Fawzi [132] introduced a very effective method to bound numerically the conditional Von Neumann entropy and they showed how this can be used in the context of DI QKD. This approach, similarly to the one based on the min entropy, allows to compute the key rate for protocols featuring any number of inputs and outputs. Notably, it yields remarkably tight results, surpassing the approach from the previous section. In addition, this method allows one to consider the entire set of probability distributions to constraint the lower bound on the key rate, hence yielding even better results than the semi-analytic approach from Section 5.2. Unfortunately, understanding the underlying physical motivation behind this method can be extremely challenging. For ease of reference, we will refer to this approach as "BFF" throughout this thesis.

In the following, we will outline the optimization problem required to establish lower bounds on the conditional Von Neumann entropy of Alice's random variable K_A given Eve's information. The optimization problem in question is a non-commuting polynomial optimization problem. As in the previous section, employing the NPA hierarchy enables the derivation of a converging hierarchy of lower bounds through solving semidefinite programs. Additionally, we will explore a speed-up for computing lower bounds by relaxing one of the constraints within the problem formulation. Delving deeply into the specifics of this approach would go beyond the scope of this thesis. More detailed explanations can be found in the original paper [132].

Let us assume that the random variable K_A represents the outcomes of a specific measurement x^* of Alice. Let $t_1, \ldots, t_m$ and $w_1, \ldots, w_m$ be the nodes and weights of an m-point Gauss-Radau quadrature on $(0, 1]$ where we fix $t_m = 1$. These coefficients can be computed efficiently in terms of Legendre polynomials [148]. Moreover, let

$$\alpha_i = \frac{3}{2} \min\left\{\frac{1}{t_i}, \frac{1}{1-t_i}\right\}. \tag{5.174}$$

We have that

$$H(K_A|E) \geq \sum_{i=1}^{m-1} \frac{w_i}{t_i \ln(2)} + \sum_{i=1}^{m-1} \frac{w_i}{t_i \ln(2)} O_i, \tag{5.175}$$

where O_i are the solutions to the following polynomial optimization problems

$$\begin{aligned}
\text{minimize} \quad & \sum_a \langle\psi| M_{a|1}(Z_{a,i} + Z_{a,i}^* + (1-t_i) Z_{a,i}^* Z_{a,i}) + t_i Z_{a,i} Z_{a,i}^* |\psi\rangle && (5.176)\\
\text{subject to} \quad & \langle\psi| M_{a|x} N_{b|y} |\psi\rangle = p(a,b|x,y), && (5.177)\\
& \sum_a M_{a|x} = \sum_a N_{b|y} = \mathbb{1}, && (5.178)\\
& M_{a|x} \geq 0, \quad N_{b|y} \geq 0, && (5.179)\\
& Z_{a,i}^* Z_{a,i} \leq \alpha_i, \quad Z_{a,i} Z_{a,i}^* \leq \alpha_i, && (5.180)\\
& [M_{a|x}, N_{b|y}] = [M_{a|x}, Z_{a,i}^{(*)}] = [Z_{a,i}^{(*)}, N_{b|y}] = 0. && (5.181)
\end{aligned}$$

It is worth noting that the problem introduced here isn't the original formulation presented in [132]; rather, it aligns with one of the speed-ups outlined in Remark 2.6. The original problem, while comprehensive, is computationally intensive, making it impractical for the scope of this thesis, especially in the context of DI QKD scenarios.

Each of the $m-1$ problems introduced above requires the introduction of $2a$ localizing matrices to enforce constraints (5.180). Although these constraints yield tighter results, they significantly decelerate the computation of the lower bounds. Consequently, it proves advantageous to consider a simplified problem formulation:

$$\begin{aligned}
\text{minimize} \quad & \sum_a \langle\psi| M_{a|1}(Z_{a,i} + Z_{a,i}^* + (1-t_i)Z_{a,i}^* Z_{a,i}) + t_i Z_{a,i} Z_{a,i}^* |\psi\rangle && (5.182)\\
\text{subject to} \quad & \langle\psi| M_{a|x} N_{b|y} |\psi\rangle = p(a,b|x,y), && (5.183)\\
& \sum_a M_{a|x} = \sum_a N_{b|y} = \mathbb{1}, && (5.184)\\
& M_{a|x} \geq 0, \quad N_{b|y} \geq 0, && (5.185)\\
& \langle\psi| Z_{a,i}^* Z_{a,i} |\psi\rangle \leq \alpha_i, \quad \langle\psi| Z_{a,i} Z_{a,i}^* |\psi\rangle \leq \alpha_i, && (5.186)\\
& [M_{a|x}, N_{b|y}] = [M_{a|x}, Z_{a,i}^{(*)}] = [Z_{a,i}^{(*)}, N_{b|y}] = 0. && (5.187)
\end{aligned}$$

This formulation is a lower bound on the problem introduced in eqs. (5.176-5.181), thereby allowing the computation of a lower bound on the conditional entropy $H(K_A|E)$. Furthermore, it tends to yield results more fastly, offering a practical advantage in computational efficiency.

Finally, we implemented the approach of this section in order to compare its performances with the semi-analytic bounds introduced at the beginning of the chapter. We include in Fig. 5.9 a plot of the key rate as a function of the detection efficiency for a protocol like the one analysed in Fig. 5.7 without noisy pre-processing (hence we fixed $q = 0$). In the analytical case we bounded the key rate exploiting S and the bias $\langle A_1 \rangle$, while in the numerical case we used the entire input-output probability distribution to constraint the SDP from eqs. (5.176-5.181). In both cases, we maximized the key rate over the angle of the partially entangled state shared between Alice and Bob and the angles of the measurements performed on the ZX plane. The level of the NPA hierarchy that we used was $2 + AB + AZ + ABZ$ and we fixed $m = 16$.

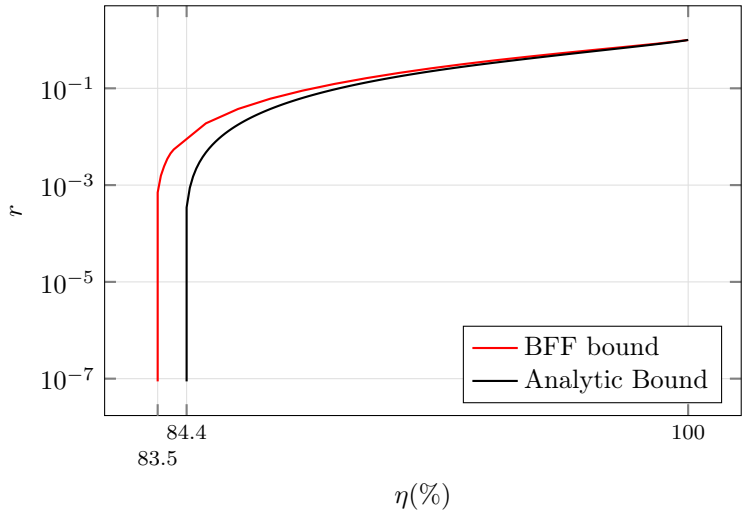

Figure 5.9: Plot of the key rate lower bounds obtained with the methods of [1] and of [132].

5.6 Discussion

In the first part of this chapter, building on [133], we have introduced a flexible approach to derive practical and fully device-independent bounds on the key rate for DI QKD in the 2-input/2-output setting. We have illustrated it on to the two-basis variant of the CHSH DI QKD protocol as well as to undertake a more optimized analysis of the single-basis variant when the main anticipated experimental imperfection is losses. Contrarily to [133], we used numerical methods to solve part of the problem in both cases and obtain optimal or close to optimal bounds on the conditional entropy within a very low amount of computation time. The results may be used to derive bounds on the key rate in the asymptotic limit or in the finite-key regime via the entropy accumulation theorem. They may also be useful as a point of comparison with different numerical approaches used to bound the conditional entropy in the device-independent setting.

When considering losses we found that the global detection efficiency can be brought under 80.26%. This is notably below the detection efficiency of 87.49% attained in the recent experimental demonstration of device-independent quantum key distribution based on a photonic setup [25]. As we remarked in the previous section, however, our threshold is attained using a very weakly entangled state and increases significantly if any realistic amount of noise is added to the model we studied. (Separately, a finite-key analysis would likely have the same effect.)

While writing the manuscript [1], the promising numerical method of Section 5.5 to bound the conditional entropy in general DI scenarios was proposed [132]. Our detection threshold, derived using only the expectation value $\langle A_1 \rangle$ of Alice's key-generation measurement in addition to CHSH, is slightly lower than the threshold of 80.5% reported in [132] using full statistics. This is not a limitation of the method of [132], but rather a matter of using a suboptimal state and measurement implementation parameters in that work. Indeed, running their method on the correlations achieving the threshold of 80.2593% in Table 5.2, the authors of [132] confirmed to us that they also find a positive key rate [149] (though, again, using full statistics instead of only $\langle A_1 \rangle$ and S). This illustrates the interest of having complementary methods. While [132] can in principle be used to tackle very general problems, our method specializing on the 2-input/2-output scenario allows us to rapidly explore the parameter space to find a good implementation. Moreover, there exist scenarios in which our analysis can provide slightly better bounds compared to the numerical method as one can observe from [132, Figure 6b].

A recent result [150] obtained lower bounds on the key rate for the finite-size case without the use of the entropy accumulation theorem in the two-input/two-output scenario. It might be interesting to investigate whether our results involving different parameters to bound the conditional Von Neumann entropy can be used in combination with their technique.

Finally, although we discussed in detail two specific examples illustrating our approach to bounding the conditional von Neumann entropy, we point out that other bounds can be derived. For instance, we could combine the BB84-type bound (5.13) using bias with the correlation bound (5.20) in terms of the asymmetric CHSH expectations. As suggested by Figure 5.8, this should slightly improve the analysis presented here (are least for larger amounts of noise δ). One could also, much more generally, use numerical techniques [151] to derive device-dependent bounds on the conditional von Neumann entropy that are more stringent and combine them with correlation bounds involving full-statistics obtained through relaxations of the Lasserre hierarchy. Our method can also in principle be applied to the n-partite setting, e.g., to derive entropy bounds based on Mermin-type Bell inequalities [152, 153].

The code used to obtain the semi-analytical results in this paper is available on GitHub [154] as well as the one for the numerical method of Section 5.5 [155].

6 One-sided DI QKD protocols

As we pointed out in the previous chapters, it is desirable to find quantum key distribution protocols that are based on realistic assumptions on the devices as well as implementable over long distances. In this chapter, we consider a one-sided DI QKD scheme with two measurements per party and show that it is secure against coherent attacks up to detection efficiencies greater than 50.1% specifically on the untrusted side. This is almost the theoretical limit achievable for protocols with two untrusted measurements. Interestingly, we also show that, by placing the source of states close to the untrusted side, our protocol is secure over distances comparable to standard QKD protocols. The content of this chapter is reproduced from [2].

6.1 Introduction

As we discussed in Section 1.4.5, well-chosen additional assumptions can simultaneously address the device characterization challenges typical of QKD protocols and the distance limitations inherent to DIQKD protocols. In this chapter, we will focus on an entanglement-based 1SDI scenario.

In our analysis, we assume that the source and one party's device (Bob's) are completely untrusted, while the second party's device (Alice's) is trusted. In particular, we assume that the trusted side measures two anti-commuting observables with a discrete number of outcomes and that the probability for a detector to click is independent of the basis choice (a.k.a. the fair-sampling assumption in the context of Bell experiments). This implies that rounds where she does not detect a photon can safely be discarded. As we do not make any assumption on the Hilbert space dimension or on the specific form of the measurements, we will say from now on that Alice's device is semi-trusted. This framework, where only one party is untrusted while the second is (semi) trusted, is usually referred to as quantum steering [39]. From a real-world perspective, this scenario resembles a situation where a user needs to establish secure communication with a server. In this context, the device belonging to the server can be considered semi-trusted, as it is affiliated with a company that possesses the necessary resources and expertise to regularly test and validate their devices.

The robustness against photon losses of various forms of 1SDI QKD protocols with discrete variables have previously been explored in [46, 47, 50, 51, 53]. However, all these protocols are based on different forms of post-selection, where the secret key is derived only from specific measurement outcomes, discarding data from other rounds. Notably, in [46, 47, 51, 53], which are B92-like protocols, rounds yielding a particular outcome are discarded entirely. As we pointed out in Section 3.5, the use of protocols where certain outcomes coming from untrusted devices are discarded does not allow to prove security against the most general class of attacks in a device-independent scenario using the recently introduced Entropy Accumulation Theorem (EAT) framework of [111, 112, 127]. Moreover, it was shown in [113, 114] that protocols relying on post-selection are indeed sensitive to coherent attacks. As a result, the work by Branciard et al. [50], as mentioned by the same authors, is based on a security proof applicable against coherent attacks, only under the constraint that the devices do not retain memory from previous rounds, thus precluding consideration of fully untrusted parties. This limitation extends to subsequent

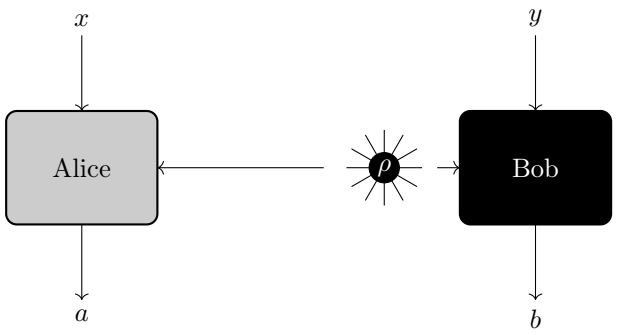

Figure 6.1: The scenario under consideration consists of an untrusted source preparing states ρ and distributing them to both an untrusted measuring device (the black box), positioned near to the source, and a semi-trusted measuring device (the gray box), situated far from the source.

works such as [51, 53]. More recently, Ioannou et al. [46, 47] studied a 1SDI prepare-and-measure setting, described in Section 4.3.2, where the security was proven only against collective attacks. Notably, their class of protocols allows for the demonstration of security at detection efficiencies that can be relatively low. However, this setting, since it is of the prepare-and-measure type, does not allow to achieve secure communication over distances comparable to standard QKD protocols.

In this work, we explore a generalised entanglement-based version of the well-known BB84 protocol [8] and we demonstrate that, in the simple case where Bob represents the undetected quantum states as separate outcomes, our 1SDI QKD scheme is secure as long as Bob's measurement has a detection efficiency greater than 50.1% which is well within the current experimental limits. Interestingly, this threshold roughly corresponds to the one where any two-untrusted-measurements protocol becomes insecure, as shown with Upper bound 1 in Chapter 2. Furthermore, as we do not perform any type of post-selection, one can analyse our protocols using the Entropy Accumulation Theorem, allowing one to prove that our scheme is secure against coherent attacks. Finally, by placing a source of entangled photons close to Bob's detector, we estimate that our 1SDI QKD protocol can be effectively implemented over distances of roughly 247 km, assuming that Bob's detectors and optical components have a total detection efficiency of 90%, that the visibility of the state prepared is of 99%, and that the detectors have a dark count rate of $p_d = 10^{-6}$. All of these values have been achieved in a recent experiment on DI QKD [25].

6.2 1SDI scenario

We will begin by outlining our setting, the additional assumptions we make beyond the standard assumptions made in DI QKD [5], and the protocol considered in this work.

Within this work, we have two spatially separated observers, Alice and Bob, who share a quantum state ρ (Fig. 6.1). They perform measurements on their respective portions of the state corresponding to inputs x, y and obtain outputs a, b for Alice, Bob respectively. Alice's and Bob's measurements are represented as $M_x = \{M_{a|x}\}$ and $N_y = \{N_{b|y}\}$, respectively where $M_{a|x}$, $N_{b|y}$ are the measurement operators that are positive semi-definite and sum up to identity. The experiment's observed correlations are characterized by a set of joint probability distributions or correlations, denoted as $\vec{p} = \{p(a, b|x, y)\}$. The observables of Alice and Bob are represented using the measurement operators as $A_x = M_{1|x} - M_{2|x}$ and $B_y = N_{1|y} - N_{2|y}$. The observables

A_1, A_2 corresponding to Alice's measurements are trusted to obey the relation:

$$\{A_1, A_2\} = 0. \tag{6.1}$$

In contrast, Bob's measurement device is untrusted.

Let us remark here that, to account for no-detection events on Alice's side, we will model her semi-trusted device using the assumption explained in Section 1.2.3. In this model, a filter returns an output $D = \checkmark, \perp$, determined independently by the input x and the state ρ. When $D = \checkmark$, the detector performs an ideal lossless measurement. When we obtain a $D = \perp$, meaning that the quantum state was not detected, we skip the round. In contrast, if Bob fails to detect an incoming state, he stores a separate outcome as $b = \varnothing$.

One should be careful to note that the scenario we consider — where an untrusted source is placed close to Bob's untrusted measuring device, both inside Bob's laboratory — cannot be described merely in terms of an untrusted source alone in Bob's lab. If we were to consider only an untrusted source in Bob's laboratory, that source could simply produce infinite copies of the same quantum state, send one to Alice, and the rest to Eve. Eve could then perform full tomography using the infinite copies she receives, gaining complete knowledge of the state sent by Bob and compromising the security of the protocol.

The key difference between these two scenarios lies in the sequence and dependence of state generation and measurement. In the entanglement-based scenario, the quantum state is first generated independently of Bob's measurement choice y, and only afterwards does Bob perform a measurement based on y. This means the state generated is independent of y. Bob's measurement choice cannot influence the state at Alice's location due to the no-signalling condition at Bob's side. If Bob's measurement could instantaneously affect the state at Alice's lab, it would allow for faster-than-light communication, violating causality. In contrast, in the prepare-and-measure scenario, the state prepared by Bob is directly dependent on his random input y. The outcomes obtained by Alice can indeed depend on the specific state that Bob sends. However, this dependence is permissible because the state must physically travel from Bob to Alice, and any influence is limited by the finite speed at which information can propagate.

Let us now describe the key distribution protocol.

1SDI QKD protocol:

- An untrusted source prepares an entangled state ρ which is distributed to Alice and Bob.
- The two parties choose the measurements settings $x = 1, 2$ and $y = 1, 2$ according to a known distribution $p(x)$ and $p(y)$.
- They obtain outcomes $a = 1, 2$, $D = \checkmark, \perp$, and $b = 1, 2, \varnothing$. The measurement settings and D are announced publicly. If $D = \perp$, then the round is discarded.
- Based on a random bit T, Alice and Bob decide if the round will be used to generate the secret key ($T = k$) in which case they store their outcomes, or to assess the security of the protocol ($T = t$) by revealing publicly their outcomes to estimate the probability distribution $\vec{p}$.
- They use $\vec{p}$ to compute the key rate and decide whether to abort or not the protocol.
- Finally, they perform error correction and privacy amplification.

An example of an honest implementation of the protocol is the case where Alice and Bob share a partially entangled state and perform the measurements of the observables $A_1 = B_1 = Z$ and $A_2 = B_2 = X$ and extract the secret key from the outcomes of the measurement of Z. The particular case where the state prepared is a maximally entangled state corresponds to an entanglement-based version of the BB84 protocol.

In a typical DI QKD scenario, the source of quantum states is placed in the middle between Alice and Bob and distributes two subsystems of an entangled state to the two parties. In this work, we consider a scenario (Fig. 6.1) where the untrusted source is positioned near Bob's laboratory, a configuration previously explored in various contexts, e.g., [50, 156–159]. This arrangement minimizes losses during the transmission of the state to Bob's device, which is the one particularly sensitive to losses. The performance of Bob's device is primarily affected by detector inefficiency. On Alice's side, losses are less relevant as rounds with non-detection are simply discarded.

6.3 Key rate lower bounds

The 1SDI QKD protocol described above satisfies the no-signalling condition of the generalised Entropy Accumulation Theorem [112]. This allows us to assess their security in the asymptotic case by verifying the positivity of the Devetak-Winter rate [110, 116]. When the secret key is extracted from Alice's outcomes of the observables A_1 and Bob uses the outcomes of B_1 to reconstruct her string, the asymptotic key rate is given by

$$r_\infty \geq H(A_1|E) - H(A_1|B_1). \tag{6.2}$$

Here, we decided to extract the key by using one-way public communication from Alice to Bob. We make this choice because, as Alice's device is semi-trusted, the convex-combination attacks as introduced in Section 4.3 do not apply to her device. Therefore, with this choice, we can achieve better key rates.

In eq. (6.2), $H(A_1|E)$ (or $H(A_1|B_1)$) represents the conditional Von Neumann entropy of the outcomes of the measurement of A_1 given Eve's information (or given the measurement outcomes

of the observable B_1). As one can observe from (6.2), to find the asymptotic key rate, one needs to obtain the values of the conditional Von Neumann entropies $H(A_1|E)$ and $H(A_1|B_1)$.

As far as $H(A_1|B_1)$ is concerned, it can be estimated as

$$H(A_1|B_1) = \sum_{a,b} p(a,b|1,1) \log_2 \frac{p(a,b|1,1)}{p(b|1)}. \tag{6.3}$$

The above quantity can be directly inferred from the input-output statistics obtained through measurements of A_1 and B_1.

Let us now focus on $H(A_1|E)$. Here, we describe how to obtain a lower bound of $H(A_1|E)$ using two different techniques. The first one is analytical and based on the approach described in Section 5.2, while the second one is numerical and based on the method outlined in Section 5.5. Here, we provide a sketch of both techniques while more detailed descriptions are provided in Sections 6.6.1-6.6.2.

Let us start with the analytical case. As shown in Section 5.2.3 with Entropy bound 3 (for $q = 0$), for measurements of Alice and Bob acting on two-dimensional spaces, $H(A_1|E)$ can be lower bounded as

$$H(A_1|E) \geq \phi\left(\langle A_1 \rangle\right) - \phi\left(\sqrt{\langle A_1 \rangle^2 + \langle \bar{A}_1 \otimes B \rangle^2}\right). \tag{6.4}$$

Here, $\phi(x) = h\left(\frac{1}{2}(1+x)\right)$, where $h(x)$ represents the binary entropy function. In this context, $\bar{A}_1$ is any observable that anti-commutes with the one employed for key extraction on Alice's side, while B is a unitary operator, that is, $B^\dagger B = \mathbb{1}$. As described earlier, Alice's observables are trusted to be anti-commuting. Consequently, we can substitute $\bar{A}_1 = A_2$ which allows us to easily bound the quantity $\langle \bar{A}_1 \otimes B \rangle$ by estimating the correlator $\langle A_2 \otimes B_2 \rangle$. At this point, as we are in a two-input-two-output scenario, we can use Jordan's lemma [160] to make our bound valid for measurements of any dimension. This lemma asserts that all statistics in the two-input-two-output scenario can be derived by employing convex combinations of strategies involving Alice and Bob using qubits. In essence, if a lower bound is determined by assuming a qubit strategy, it must be convexified to account for the possibility of obtaining a lower bound through the mixture of strategies. This ensures that one cannot obtain further lower bounds by convex mixing of two different strategies. We show in Section 6.6.2 that the function appearing on the right-hand side of eq. (6.4) is convex in $\langle A_1 \rangle$ and $\langle A_2 \otimes B_2 \rangle$. Hence, we can say that our bound holds for states and measurements of any dimension.

The application of Jordan's lemma necessitates the restriction that the statistics employed to bound the adversary's information are derived only from two inputs per party with only two possible outcomes. Therefore, for the specific rounds where we estimate $\langle A_2 \otimes B_2 \rangle$, we need to modify the protocol by deterministically mapping Bob's non-detection outcomes ($\varnothing$) to one of his other possible outcomes for all the rounds used for parameter estimation.

Alternatively, to compute $H(A_1|E)$, we can adopt the BFF approach of Section 5.5 focused on numerical techniques based on the NPA (Navascués-Pironio-Acín) hierarchy of semi-definite programs (SDP) [146, 147]. As we described before, the quantity $H(A_1|E)$ can be lower bounded by the solution to a noncommutative polynomial optimization problem. In our case, we include into the optimization problem a further polynomial constraint (given by eq (6.1)) that will force Alice's measurements to anti-commute (Section 6.6.1).

The latter analysis can be used to analyse more general scenarios including both the case where Bob maps his non-detection outcomes into one of the other outcomes and the case where he keeps it as a separate outcome. Moreover, it is possible to constraint the information of Eve not only by taking into account the correlators $\langle A_2 \otimes B_2 \rangle$ and $\langle A_1 \rangle$, but using the full probability table $p(a,b|x,y)$.

6.4 Reference experiment

In order to simulate the performance of an experiment, we will assume an honest implementation of the protocol where we consider a reference state prepared by the source and reference measurements performed by both parties that generate the probability distribution $\vec{p}_{\text{ref}}$. For our purpose, we assume that the source prepares the two-qubit entangled states of the form

$$|\psi(\theta)\rangle = \cos\theta\,|00\rangle + \sin\theta\,|11\rangle\,. \tag{6.5}$$

Alice measures the Pauli operators $A_1 = Z$ and $A_2 = X$. Since she is semi-trusted, she can discard the non-detection events due to which her measurements will not be affected by detection losses. For Bob's measurements, we include the detection losses using the factor η in the ideal measurements as

$$N_{1|y}(\eta) = \eta N_{1|y} + (1-\eta)\mathbb{1}, \tag{6.6}$$
$$N_{2|y}(\eta) = \eta N_{2|y}, \tag{6.7}$$

where $N_{b|y}$ are the projective measurements that Bob performs in the ideal case, and η represents the transmittance or detection efficiency. To obtain the optimal key rates we will maximize the lower bound to the key rate over the parameter θ of the state $|\psi(\theta)\rangle$ (6.5) for every value of η taken into account. For the ideal measurements, we choose $B_1 = Z$ as this will make Bob's measurements used for the secret key correlated to Alice's ones, and $B_2 = X$ as this will maximize the correlator $\langle A_2 \otimes B_2\rangle$ and thus also the value of the conditional entropy in eq. (6.4) (we expect the case where we use BFF to be analogous).

We now plot the lower bound to the key rates in Fig. 6.2 obtained using the above-mentioned techniques. Firstly, we display the bound of [50]. This bound was obtained for a protocol that includes post-selection on the untrusted party and the key rate was estimated using a different technique which allowed to prove security only for the case of memoryless devices. Notice that, as we show in Section 6.6.2, the bound of [50] can be obtained also under the same assumptions considered in this work without performing post-selection on Bob's side and, thus, allowing us to prove security without assuming memoryless devices. In particular, the conditional entropy of Alice can be bounded by

$$H(A_1|E) \geq 1 - \phi(\langle A_2 \otimes B_2\rangle). \tag{6.8}$$

Hence, using the measurements described above and partially entangled states (eq. (6.5)), the key rate becomes

$$r_\infty \geq 1 - h\left(\frac{1}{2}(1+\eta\sin(2\theta))\right) - (1-\eta)h(\cos^2(\theta)), \tag{6.9}$$

and its maximum is at $\theta = \pi/4$ for all η, meaning that the optimal state in this case is always a maximally entangled state.

Secondly, we plot the bound obtained using eq. (6.4). In this case, we obtained an improvement that comes from the use of partially entangled states together with the inclusion of the information on the bias $\langle A_1\rangle$ in the bound.

Additionally, we include in Fig. 6.2 the key rates for three different scenarios computed using the BFF technique. Firstly, we have the case where Bob maps his non-detection events into one of his other outcomes, and we only constrained the mean value $\langle A_2 \otimes B_2\rangle$. In this case, displayed with a black dashed line in Fig. 6.2, we recovered a lower bound on the key rate which is almost identical to the one identified in [50]. We maximized the key rate over the angle of the partially entangled state for all values of η, but, as in the analytical case, the optimal angle was always $\theta \simeq \pi/4$, i.e., a maximally entangled state. Using the same technique, we could recover also the

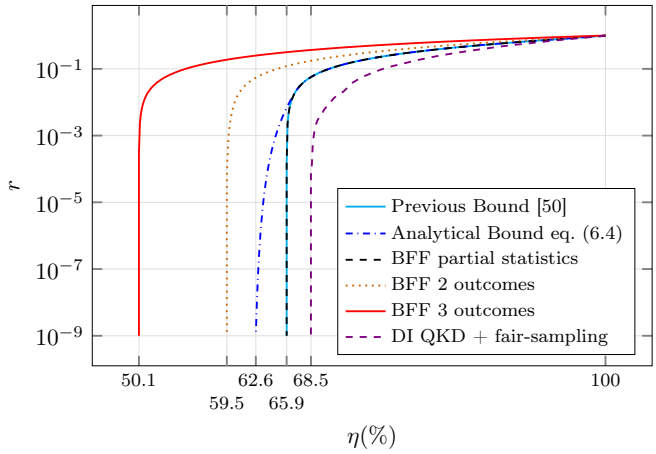

Figure 6.2: Plot of the key rate as a function of Bob's detection efficiency computed using different techniques as explained in the main text.

bound of eq. (6.4) by constraining the mean values $\langle A_2 \otimes B_2 \rangle$ and $\langle A_1 \rangle$ and maximizing over the partially entangled states. Anyway, for this case, the closer we got to the threshold the more we had to increase the level of the NPA hierarchy and the number of Gauss-Radau coefficients (see Section 5.5 for the definition), and this was possible up to slightly less than 64%, where the key rate was of the order of 10^{-5} and the precision of the SDP solver used was not enough to provide reliable results. We did not encounter this problem in the first case as the key rate remains above 10^{-4} up to $\eta = 65.9\%$.

Secondly, we analysed the same protocol using the full probability distribution $\vec{p}_{\text{ref}}$ as a constraint on our problem. This procedure led to an improvement of the key rate which was positive up to a detection efficiency of 59.5% as we can see from the orange dotted line. The use of optimal angles θ led to a small improvement compared to a scenario where we use full statistics and binning of Bob's third outcome, but with maximally entangled states.

Lastly, we computed the key rate using full statistics in the scenario where Bob keeps undetected photons as a separate outcome (red line). For this last case, the optimal θ was again the one of a maximally entangled state for all η. Here, we could show the positivity of the key rate up to 50.1%. Notice that at $\eta = 50\%$ Bob's measurements become jointly measurable, and thus no secret key can be extracted under our assumptions [3, 71].

Finally, for the sake of comparison, we include in Fig. 6.2 the plot of an analogous setting for the case of a 1SDI scenario where we make only a fair sampling assumption on Alice's side (purple line). In this case, Alice performs two measurements, and we do not assume anti-commutativity. Bob measures three different observables each with three possible outcomes. The ideal observables of Alice and Bob are

$$A_x = \cos(\alpha_x)Z + \sin(\alpha_x)X, \tag{6.10}$$
$$B_y = \cos(\beta_y)Z + \sin(\beta_y)X. \tag{6.11}$$

We apply the same noise model described before, hence, Bob's measurement operators are evolved according to eqs. (6.6-6.7), while Alice's ones remain the ideal ones due to the fair-sampling assumption. Here, Alice has only two possible outcomes as she discards the undetected photons, while Bob keeps his three outcomes separate. The secret key is extracted from the outcomes of A_1, while Bob uses the outcomes of B_3 to guess Alice's ones. Moreover, we use a technique denoted

as *noisy pre-processing* [161], where Alice randomly flips her outcomes during key generation rounds with a probability q. We did not use this technique for the previous case as our best result allowed us to obtain a 50.1% threshold which is already very close to optimal. We finally compute the key rates using BFF using the full probability distribution as a constraint. We optimize the key rate heuristically at each value of η over the parameters θ, α_1, α_2, β_1, β_2, β_3, q and obtain a threshold of 68.5%.

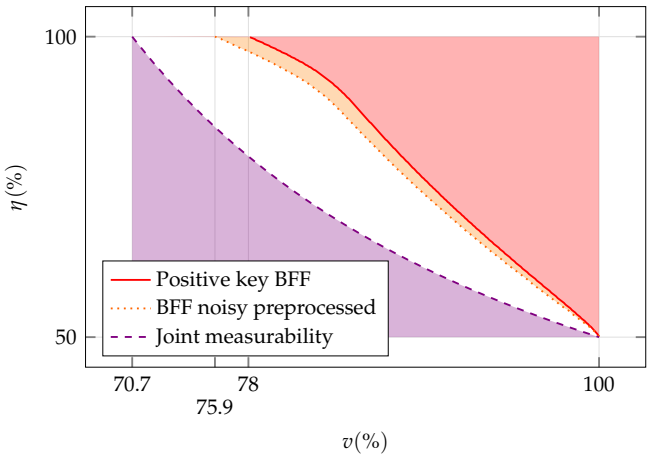

Figure 6.3: Regions where the key rate becomes zero as a function of visibility and detection efficiency. The region where the BFF 3 outcome protocol with partially entangled states is secure is above the red line, the case where we include also noisy preprocessing is above the orange dotted line, while the purple line denotes the boundary below which Bob's observables are jointly measurable, rendering any protocol insecure in that region.

We will now study the robustness of our technique in the presence of depolarizing noise acting on the initial state together with photon losses. The reference statistics are obtained by adding depolarizing noise to the state generated by the source as

$$\rho = v\,|\psi(\theta)\rangle\,\langle\psi(\theta)| + \frac{1-v}{4}\mathbb{1}, \tag{6.12}$$

where v is denoted as visibility. We show in Fig. 6.3 the region in the space of parameters η and v where the key rate is positive. This analysis is conducted in two scenarios. One where we optimize the key rate only over the angle of the partially entangled state and another where we additionally perform optimal noisy pre-processing. The key rate was computed using BFF with full statistics and keeping Bob's undetected photons as a separate outcome. We noticed that for visibilities smaller than one, the use of partially entangled states (and of noisy pre-processing) leads to an improvement also for the BFF case where the three outcomes of Bob are kept separate. Moreover, we include in the plot a curve from [3] which displays the zone where the evolution of X and Z measurements of Bob under depolarizing noise and losses are jointly measurable. When Bob's measurements are jointly measurable, the correlations between Alice and Bob are unsteerable. As a result, no secure key can be extracted in this region regardless of the initial state prepared, and regardless of the measurements performed by the semi-trusted party. One can see in Fig. 6.3 that, for non-unit values of the visibility, a region exists wherein the key rate is non-positive, and the noisy versions of X and Z observables are not jointly measurable. This suggests that refining our protocol or improving its analysis might enhance its resilience against such noise. Specifically, exploring variations of our protocol involving secret key extraction from

both Alice's measurements and using the BFF technique with higher levels of the NPA hierarchy may prove insightful. However, recall that the joint measurability of Bob's measurements is only a necessary condition for the protocol to be insecure. For non-unit values of visibility, stricter conditions may be necessary.

6.5 Theoretical distance estimate of implementability

Finally, in order to estimate the distance at which our protocols can be implemented, we will deploy a model where Alice's and Bob's devices observe dark counts. We describe our model in Section 1.2.4. It is crucial to note that the noise model we are about to use is regarded as part of the untrusted quantum channel that evolves the untrusted state ρ. Therefore, as it is typically assumed in standard QKD protocols, dark counts are not included in the model of Alice's semi-trusted device, ensuring that her measurements remain anti-commuting. For convenience, we will evolve Alice's ideal POVMs to compute the probability distribution resulting from our model as this is equivalent to evolving the state ρ.

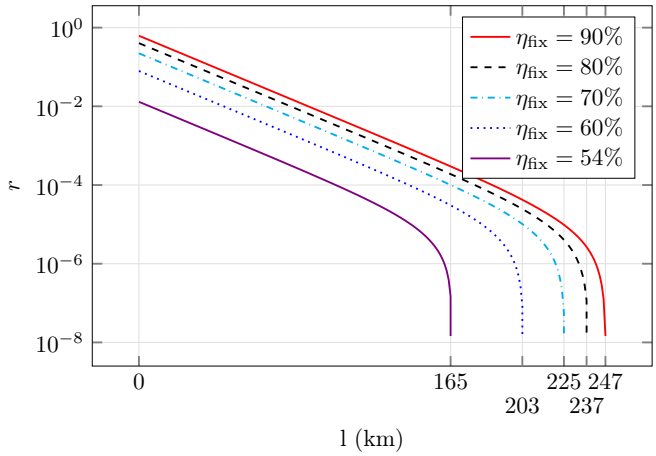

Figure 6.4: Plot of the key rate as a function of the distance in a model that includes photon losses, depolarizing noise, and dark counts in the case where we assume that Alice's measurements anti-commute.

Let us provide Bob's measurement operators. We choose to categorize all the events where Bob experiences zero or double clicks as a separate outcome $\varnothing$. As explained in Section 1.2.4, his POVMs evolve as

$$N_{1,2|y}(\eta_B, p_d) = \eta_B(1-p_d)N_{1,2|y} + (1-\eta_B)p_d(1-p_d)\mathbb{1}, \tag{6.13}$$

$$N_{\varnothing|y}(\eta_B, p_d) = \big(p_d\eta_B + (1-\eta_B)(p_d^2 + (1-p_d)^2)\big)\mathbb{1}, \tag{6.14}$$

where $N_{b|y}$ are the projective measurements that Bob performs in the ideal case.

On Alice's side, we choose to discard all events where her detectors do not click or where they both click. As we discard these events, we will need to renormalize her POVMs dividing them by the probability of having only one click. We show in Section 1.2.4 that the measurement operators of Alice $M_{1|x}(\eta_A, p)$ can be expressed in terms of the ideal ones $M_{1|x}$ as

$$M_{1|x}(\eta_A, p_d) = \frac{\eta_A(1-p_d)M_{1|x} + (1-\eta_A)p_d(1-p_d)\mathbb{1}}{1 - p_d\eta_A - (1-\eta_A)(p_d^2 + (1-p_d)^2)}, \tag{6.15}$$

and $M_{2|x}(\eta_A, p_d) = \mathbb{1} - M_{1|x}(\eta_A, p_d)$.

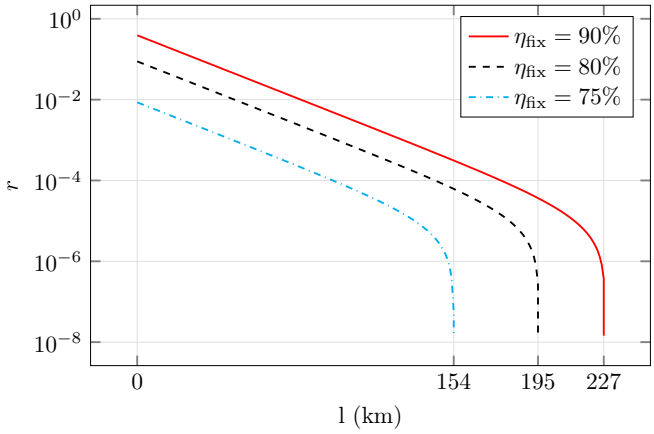

Figure 6.5: Plot of the key rate as a function of the distance in a model that includes photon losses, depolarizing noise, and dark counts in the case where we do not assume that Alice's measurements anti-commute.

For the reference statistics, we will take into account the values of visibility and dark counts of the most advanced experiment we are aware of, which employs an entanglement-based scheme [25]. Specifically, we take $v = 0.99$ and $p_d = 10^{-6}$. As far as the detection efficiency is concerned, we will take into account a fixed component η_{fix} due to the efficiency of the detectors and of the other optical components, and a variable component that accounts for optical fiber losses based on its length l, given by $\eta_{\text{fiber}} = 10^{-\alpha \cdot l/10}$. Here, we will use the typical value for telecom wavelength of $\alpha = 0.2\, dB/km$ (see e.g. [71, 159]). Since the source is placed very close to Bob's laboratory, we can ignore the losses arising due to the fiber. Consequently, his efficiency will be $\eta_B = \eta_{\text{fix}}$, while Alice's one is $\eta_A(l) = \eta_{\text{fix}} \cdot 10^{-\alpha \cdot l/10}$. Given these parameters of the setup, in Fig. 6.4 we plot the lower bound to the asymptotic key rate as a function of the distance. To show the number of key bits generated in each actual round of the protocol (and not only in the ones that have been retained), the key rates in Fig. 6.4 have been multiplied by the probability of retaining a round which is $1 - p_d\eta_A - (1-\eta_A)(p_d^2 + (1-p_d)^2)$. To compute the key rates, we used the BFF method where Bob's third outcomes are kept separate. We provide the plot for different values of η_{fix}. For each specific η_{fix} and l value, we optimized the key rate by adjusting the angle of the partially entangled state and the amount of noisy pre-processing. Let us remark that the optimization for each value of l resulted only in a few kilometers of improvement compared to the case where we optimize only for $l = 0$ and apply the same parameters for each l. This is due to the fact that the key rate drops quickly to zero when η_A and p_d reach a similar order of magnitude.

Finally, we include in Fig. 6.5 a plot where we estimate the achievable distance in the case of a protocol where we do not assume that Alice's observables anti-commute and Bob makes three measurements as explained before. Similarly to the previous case, we computed the key rates using the BFF method where Bob's third outcomes are kept separate, and we fixed $v = 0.99$ and $p_d = 10^{-6}$. In this case, for each specific η_{fix} value, we optimized the key rate by adjusting the parameters θ, α_1, α_2, β_1, β_2, β_3, q for a distance of $l = 0$, and then applied the same parameters to compute the key rate across all other distances. This was due to the fact that computing a key rate in this setting requires a higher level of the NPA hierarchy and optimizing over the parameters in question for each value of l would require a too large computational time. We can see that we can obtain similar distances as in the previous case, but Bob's device needs

greater detection efficiencies. In particular, for $\eta_{\text{fix}} \lesssim 70\%$ this type of protocol is not secure, while the one where we assume that Alice's observables anticommute can still be proven secure. This requirement can be inconvenient in a client-server scenario where Bob is a user with limited resources.

6.6 Technical Details

6.6.1 Brown-Fawzi-Fawzi bounds

In this section, we provide the non-commutative polynomial optimization problem that we used to lower bound the conditional Von Neumann entropy to verify the security of our protocols. Our lower bound is a simple extension of the one described in Section 5.5 to a steering scenario.

Let $t_1, \ldots, t_m$, $w_1, \ldots, w_m$, $\alpha_1, \ldots \alpha_{m-1}$ be defined as before. We have that

$$H(A_1|E) \geq \sum_{i=1}^{m-1} \frac{w_i}{t_i \ln(2)} + \sum_{i=1}^{m-1} \frac{w_i}{t_i \ln(2)} O_i, \tag{6.16}$$

where O_i are the solutions to the following polynomial optimization problems

$$\begin{aligned}
\inf \quad & \sum_a \langle\psi| M_{a|1}(Z_{a,i} + Z_{a,i}^* + (1-t_i) Z_{a,i}^* Z_{a,i}) + t_i Z_{a,i} Z_{a,i}^* |\psi\rangle && (6.17)\\
\text{s.t.} \quad & \langle\psi| M_{a|x} N_{b|y} |\psi\rangle = p(a,b|x,y), && (6.18)\\
& \sum_a M_{a|x} = \sum_a N_{b|y} = \mathbb{1}, && (6.19)\\
& M_{a|x} \geq 0, \quad N_{b|y} \geq 0, && (6.20)\\
& Z_{a,i}^* Z_{a,i} \leq \alpha_i, \quad Z_{a,i} Z_{a,i}^* \leq \alpha_i, && (6.21)\\
& [M_{a|x}, N_{b|y}] = [M_{a|x}, Z_{a,i}^{(*)}] = [Z_{a,i}^{(*)}, N_{b|y}] = 0, && (6.22)\\
& (M_{1|1} - M_{2|1})(M_{1|2} - M_{2|2}) + (M_{1|2} - M_{2|2})(M_{1|1} - M_{2|1}) = 0. && (6.23)
\end{aligned}$$

In the last constraint (eq. (6.23)) of the optimization problem, we introduced the anti-commutation constraint. Note that we expressed w.l.o.g. $A_x = M_{1|x} - M_{2|x}$. The probabilities $p(a,b|x,y)$ are the ones observed during the experiment or, in our case, computed from the simulations.

In the case where we perform noisy preprocessing, the measurement operators $M_{a|1}$ in the objective function need to be replaced with

$$\begin{aligned}
M_{1|1} &\to (1-q) M_{1|1} + q M_{2|1}, && (6.24)\\
M_{2|1} &\to (1-q) M_{2|1} + q M_{1|1}. && (6.25)
\end{aligned}$$

The latter problem is a non-commutative polynomial optimization problem and it can be relaxed to a hierarchy of semi-definite programs [146, 147]. The lower bound that we obtain from the solution of this problem is increasingly tighter as a function of the level of the hierarchy and as a function of the number of points of the Gauss-Radau quadrature. For the computations performed in this work in the case where Alice's observables are assumed to anti-commute, we kept the level of the localizing matrices of the NPA hierarchy fixed to 1 and used $m = 15$ in the Gauss Radau quadrature. The level of the principal moment matrix was slightly bigger than one in order to make sure that all the terms that appear in the localizing moment matrices were also present in the principal one. Additionally, our analysis revealed that for the problems examined in this article, incorporating localizing matrices representing the constraints $Z_{a,i}^* Z_{a,i} \leq \alpha_i$ and

$Z_{a,i}Z_{a,i}^* \leq \alpha_i$ did not improve the lower bounds on the conditional Von Neumann entropy. Consequently, we opted for simpler constraints $\langle\psi| Z_{a,i}^* Z_{a,i} |\psi\rangle \leq \alpha_i$ and $\langle\psi| Z_{a,i} Z_{a,i}^* |\psi\rangle \leq \alpha_i$, which still establish a valid lower bound on the conditional Von Neumann entropy.

Finally, in the case where we analysed a DI protocol with a fair sampling assumption on Alice's side, we did not include any localizing matrix. This comes from two reasons. First, we did not need to enforce anti-commutativity and, second, as before, we used the simpler constraints $\langle\psi| Z_{a,i}^* Z_{a,i} |\psi\rangle \leq \alpha_i$ and $\langle\psi| Z_{a,i} Z_{a,i}^* |\psi\rangle \leq \alpha_i$. However, in this scenario, achieving satisfactory results required the use of a principal moment matrix at level $2 + ABZ$.

6.6.2 Analytical bounds

In this section, we will prove extensively the analytical lower bound on the key rate described in the main text.

We will bound the quantity $H(A_1|E)$ as a function of input-output statistics derived from measurements of $A_{1,2}$ and $B_{1,2}$ that only have two possible outcomes. In this way, we can use Jordan's lemma [160] and reduce our problem to deriving a convex lower bound on the conditional Von Neumann entropy between Alice and Eve in the case where Alice and Bob's systems are two-dimensional.

If the secret key is extracted from the outcomes of the measurement of the two-outcomes observable A_1 on Alice's side, the conditional Von Neumann entropy of Alice conditioned on Eve can be bounded by (Entropy bound 1 in Section 5.2.3)

$$H(A_1|E) \geq 1 - \phi\left(\langle|\bar{A}_1 \otimes B\rangle|\right), \tag{6.26}$$

where B is any unitary observable on Bob's subspace, $\bar{A}_1$ is an observable on Alice's subspace that anticommutes with A_1, and $\phi(x) = h(1/2(1+x))$ where $h(x)$ is the binary entropy function. Under the same assumptions, a second bound that we can use (Entropy bound 3 with $q = 0$) is

$$H(A_1|E) \geq \phi\left(\langle A_1\rangle\right) - \phi\left(\sqrt{\langle A_1\rangle^2 + \langle\bar{A}_1 \otimes B\rangle^2}\right), \tag{6.27}$$

At this point, we need to find a lower bound on $\langle\bar{A}_1 \otimes B\rangle$ as a function of the input-output statistics of Alice and Bob's devices such that it is independent of what is the form of the measurements performed by Bob.

For the steering scenario, we can choose to bound the correlator directly. By assuming that $\bar{A}_1 = A_2$, and choosing $B = B_2$, we can evaluate $\langle A_2 \otimes B_2\rangle$ from the input-output statistics. In this case, using eq. (6.27), our bound becomes of the type

$$H(A_1|E) \geq \phi\left(\langle A_1\rangle\right) - \phi\left(\sqrt{\langle A_1\rangle^2 + \langle A_2 \otimes B_2\rangle^2}\right). \tag{6.28}$$

Notice that for the case of eq. (6.26) applied to an honest implementation where Alice and Bob measure Pauli operators Z, X on a maximally entangled state undergoing losses and depolarizing noise, we obtain the same lower bound of [50] just by assuming the anti-commutativity of Alice's measurements and making a fair sampling assumption on her side.

Finally, in order to use Jordan's lemma, we will need to prove that the bounds obtained are convex in the variables $x = \langle A_2 \otimes B_2\rangle$ and $z = \langle A_1\rangle$.

The case of eq. (6.26) is straightforward since the binary entropy function is a concave function. Let us now prove the convexity for the case of eq. (6.27).

We start with the direct case. We need to show that, for the function

$$f(z, x) = \phi(z) - \phi\left(\sqrt{z^2 + x^2}\right), \tag{6.29}$$

the hessian matrix is such that

$$\mathbf{H}(f(z,x)) = \begin{pmatrix} \frac{d^2}{dz^2}f(z,x) & \frac{d^2}{sz\,dx}f(z,x) \\ \\ \frac{d^2}{sz\,dx}f(z,x) & \frac{d^2}{dx^2}f(z,x) \end{pmatrix} \geq 0. \tag{6.30}$$

Being a 2×2 matrix, the positivity is ensured by the positivity of the diagonal elements and of the determinant. We have

$$\frac{d^2}{dz^2}f(z,x) = \frac{x^2}{\ln(2)}\left(\frac{x^2-1+2z^2}{(z^2+x^2)(1-x^2-z^2)(1-z^2)} + \frac{\operatorname{arctanh}\left(\sqrt{z^2+x^2}\right)}{\left(x^2+z^2\right)^{3/2}}\right). \tag{6.31}$$

we can easily notice that, since $z^2 \leq 1$, the positivity of eq. (6.31) is ensured by the condition $z^2 + x^2 \leq 1$ which is proven in Section 6.6.3.

The second diagonal element is

$$\frac{d^2}{dx^2}f(z,x) = \frac{1}{\ln(2)}\left(\frac{x^2}{(z^2+x^2)(1-x^2-z^2)} + z^2\frac{\operatorname{arctanh}\left(\sqrt{z^2+x^2}\right)}{\left(x^2+z^2\right)^{3/2}}\right), \tag{6.32}$$

whose positivity is ensured by the same conditions as before. Finally, the determinant is

$$\begin{aligned} &\frac{d^2}{dz^2}f(z,x)\frac{d^2}{dx^2}f(z,x) - \left(\frac{d^2}{dx\,dz}f(z,x)\right)^2 = \\ &\frac{x^2}{(\ln(2))^2(z^2+x^2)^2(1-x^2-z^2)(1-z^2)}\left(\sqrt{z^2+x^2}\operatorname{arctanh}\sqrt{z^2+x^2} - (z^2+x^2)\right), \end{aligned} \tag{6.33}$$

whose positivity is ensured by $z^2 \leq 1$, by $z^2 + x^2 \leq 1$, and by $\operatorname{arctanh}(x) \geq x$.

We conclude that eq. (6.28) is a valid lower bound under the assumptions of anti-commutativity between A_1 and A_2 and under a fair sampling assumption on Alice's side. Notice that Alice's Hilbert space dimension does not need to be bounded.

6.6.3 Domain of the correlations

In this section, we prove the condition

$$\langle A\rangle^2 + \langle \bar{A} \otimes B\rangle^2 \leq 1, \tag{6.34}$$

which is required in Section 6.6.2.

Let us begin by considering a state ρ_{AB} and then purifying it to $|\psi_{ABE}\rangle$ such that $\rho_{AB} = \mathrm{Tr}_E\ \psi_{ABE}$ where E denotes the ancillary system. Now, any general state $|\psi_{ABE}\rangle \in \mathbb{C}^2 \otimes \mathcal{H}_{BE}$ can be written as

$$|\psi_{ABE}\rangle = \sum_{i=0,1} \lambda_i |i\rangle_A |e_i\rangle_{BE} \tag{6.35}$$

where $\lambda_i \geq 0$, $\sum_{i=0,1}\lambda_i^2 = 1$ and $|e_i\rangle_{BE}$ are normalised but in general not orthogonal. Now, evaluating the left-hand side of eq. (6.34) by plugging in the state (6.35), we obtain

$$\langle Z\rangle^2 + \langle X \otimes B\rangle^2 = \left(\lambda_0^2 - \lambda_1^2\right)^2 + 4\lambda_0^2\lambda_1^2\ \left(\mathrm{Re}\langle e_0|B|e_1\rangle\right)^2 \tag{6.36}$$

where we used the fact that B is Hermitian. As B is unitary, we have that $-1 \leq \mathrm{Re}\langle e_0|B|e_1\rangle \leq 1$ using which we arrive at

$$\langle Z\rangle^2 + \langle X \otimes B\rangle^2 \leq \left(\lambda_0^2 - \lambda_1^2\right)^2 + 4\lambda_0^2\lambda_1^2 = \left(\lambda_0^2 + \lambda_1^2\right)^2 = 1. \tag{6.37}$$

This completes the proof.

6.7 Discussion

We applied to a 1SDI scenario the most recent methods to derive lower bounds on the conditional Von Neumann entropy. We provided lower bounds with both analytical and numerical methods and showed that our analytical results can be recovered with the numerical method of [132]. We showed that the first QKD protocol ever introduced, the BB84 protocol, can be proven secure in an entanglement-based 1SDI scenario with up to 50.1% detection efficiency on the untrusted side. This result is almost optimal as no protocol employing an untrusted party performing two measurements can achieve security below a 50% detection efficiency threshold. Furthermore, we observed that, by increasing the number of Gauss-Radau coefficients, we can get closer to this critical threshold.

Detection efficiencies above 50% have now been obtained in a large number of photonic experiments both in the context of Bell tests [37, 38, 162], Quantum Random Number Generation [163–170] and DI QKD [25]. Therefore, our protocol is well within the current experimental limits. It is worth noting that DI QKD protocols can also be proven secure at similar distances as the ones simulated in this chapter but they require more complex schemes [157, 159] and much higher detection efficiencies. In the case of [159], for example, the detection efficiency required to prove security with unit visibility is 96.1%. Notice also that [159] assumes a dark count rate of 10^{-7}, while here we make the more pessimistic assumption of a dark count rate of 10^{-6}.

Finally, while our findings achieve near-optimal thresholds in terms of detection efficiency, it remains an open question whether our result is optimal also in terms of depolarizing noise. As we showed in Fig. 6.3, there is a region where either our protocol could be improved, or there is an attack that goes beyond the joint measurability of Bob's measurements. Let us also point out that our work can be easily extended to the case where the measurements of Alice are not perfectly anticommuting using the BFF method. For instance, one can explore the case of $A_1A_2 + A_2A_1 = \epsilon\mathbb{1}$, where ϵ is small. Another possible extension is to consider multipartite scenarios like Conference Key Agreement.

The code used to compute the key rates numerically can be found in the GitHub folder [155].

7 Conclusions

In this thesis, we delved into the subject of Quantum Key Distribution with untrusted devices. We reviewed some of its most important theoretical aspects, we examined the advantages that it offers in terms of security, and we highlighted the practical challenges relative to its implementation. One of the most demanding experimental requirements for an effective DI QKD realisation is achieving high detection efficiencies over distances that are adequately long for practical use. While current technology does not allow to overcome this requirement, we investigated whether theoretical advancements or modified assumptions could ease this challenge.

In Chapter 2 we studied the minimal requirements for communication channels and measurement devices in Device Independent scenarios. We found that high detection efficiencies are not a necessary condition when a large number of measurements is performed. While this is true in general, we observed that, when the measurements are Gaussian, we cannot reduce the requirements in terms of detection efficiency by increasing the number of measurements. In Chapter 4, we focused on the necessary conditions for DI QKD protocols. We analysed DI QKD setting where a large number of measurements are performed. We could show that the simplest protocols, the ones using maximally entangled states and a large number of measurements in the ZX plane, cannot be proven secure at low detection efficiencies. However, considering more sophisticated protocols, which were not included in our analysis, might show improvements. For instance, noisy preprocessing was not considered in our study.

In Chapter 5, we described some of the most promising approaches to lower bound the conditional Von Neumann entropy in DI QKD protocols and, thus, prove their security. Our best analysis applied on the most sophisticated protocol that we could study, showed that realising such DI protocol in practice requires detection efficiencies above 80% on both Alice's and Bob's sides. This result was confirmed also using the numerical method of Section 5.5 in [132]. This 80% value is extremely challenging to achieve in an experiment because each optical component has a non unit detection efficiency and each of them needs to be multiplied with the efficiency of the detectors and the one of the optical fiber. Even assuming to use ideal detectors and optical components with 100% detection efficiency, since the detection efficiency of optical fibers decreases exponentially with their length[1], the DI QKD protocols analysed cannot be realized over more than 10km of distance.

The study of more sophisticated DI QKD protocols, where a large number of measurements is involved, is extremely complex with the means available at the moment. Analytically, solving this would require a result similar to the Jordan lemma for protocols with more than two inputs and outputs. Numerically, handling many measurements demands immense computational power because the NPA hierarchy requires solving extremely large Semidefinite Programs when a large number of measurement operators are involved.

We observed in Chapter 4 that protocols where only one side's quantum devices are untrusted have less strict requirements. This was for instance the case of the receiver-device independent protocols analysed in Section 4.3.2. Guided by this observation, we explored in Chapter 6 a scenario where one side's quantum device in the communication protocol is partially trusted, while the other side's quantum devices are untrusted. We could show that this scenario can be proven secure against coherent attacks over distances that are similar to the ones achieved in

[1] For telecom wavelength we have that $\eta \simeq 10^{-0.02/l}$, where l is the length of the fiber in kilometers.

standard QKD protocols. Moreover, we observed that our result was quasi-optimal in terms of detection efficiency.

Advancing the field of quantum cryptography could benefit from studying hybrid approaches between Device-Independent and traditional Quantum Key Distribution protocols. Addressing the challenges of characterizing quantum devices in traditional QKD and the experimental complexity of DI QKD is important. Hybrid approaches may offer a solution to both issues simultaneously. Exploring various assumptions and tailoring specific protocols to different security requirements could be particularly beneficial. For instance, the asymmetric case discussed in Chapter 6 is especially suitable for scenarios where a server needs to securely communicate with a client, demonstrating how specific protocols can be designed to meet distinct security needs.

A further scenario worth studying involves protocols with assumptions about the information an eavesdropper might have about the source, a setting introduced in [57], or protocols with imperfect state preparations [56]. My preliminary studies in this direction indicate that QKD under these assumptions can be proven secure. Additionally, there are other assumptions worth exploring, such as limitations on the energy emitted by the source or received by the detector [171], and protocols where Eve is only allowed to perform collective attacks, as the protocol described in Section 4.3.2 or the one of [172].

Bibliography

[1] M. Masini, S. Pironio, and E. Woodhead. Simple and practical DIQKD security analysis via BB84-type uncertainty relations and Pauli correlation constraints. *Quantum*, 6:843, Oct 2022. https://doi.org/10.22331/q-2022-10-20-843.

[2] M. Masini and S. Sarkar. One-sided DI-QKD secure against coherent attacks over long distances. *arXiv preprint arXiv:2403.11850*, Mar 2024. https://doi.org/10.48550/arXiv.2403.11850.

[3] M. Masini, M. Ioannou, N. Brunner, S. Pironio, and P. Sekatski. Joint-measurability and quantum communication with untrusted devices. *arXiv preprint arXiv:2403.14785*, Mar 2024. https://doi.org/10.48550/arXiv.2403.14785.

[4] 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, Jun 2007. https://link.aps.org/doi/10.1103/PhysRevLett.98.230501.

[5] I.W. Primaatmaja, K.T. Goh, E.Y.Z. Tan, J.T.F. Khoo, S. Ghorai, and C.C.W. Lim. Security of device-independent quantum key distribution protocols: a review. *Quantum*, 7:932, Mar 2023. https://doi.org/10.22331/q-2023-03-02-932.

[6] 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.*, 81:1301–1350, Sep 2009. https://link.aps.org/doi/10.1103/RevModPhys.81.1301.

[7] 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, May 2020. https://link.aps.org/doi/10.1103/RevModPhys.92.025002.

[8] C.H. Bennett and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. *Theoretical Computer Science*, 560:7–11, May 2014. https://www.sciencedirect.com/science/article/pii/S0304397514004241.

[9] C.E. Shannon. Communication theory of secrecy systems. *Bell. Syst. Tech. J.*, 28(4): 656–715, 1949. https://doi.org/10.1002/j.1538-7305.1949.tb00928.x.

[10] N.I. of Standards and Technology. Federal information processing standards publication 197. Nov 2001. https://doi.org/10.6028/NIST.FIPS.197-upd1.

[11] R.L. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Commun. ACM*, 21(2):120–126, feb Feb. 1978. https://doi.org/10.1145/359340.359342.

[12] P.W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.*, 26(5):1484–1509, Oct 1997. https://doi.org/10.1137/S0097539795293172.

[13] N.I. of Standards and Technology. Status report on the third round of the nist post-quantum cryptography standardization process. Jul 2022. https://doi.org/10.6028/NIST.IR.8413-upd1.

[14] MagiQ Technologies. https://www.magiqtech.com/solutions/network-security/.

[15] ID Quantique. https://www.idquantique.com/quantum-safe-security/quantum-key-distribution/.

[16] Quintessence Labs. https://www.quintessencelabs.com/.

[17] ThinkQuantum. https://www.thinkquantum.com/tq-technology/.

[18] Quantum Telecommunications Italy. https://www.qticompany.com/products/.

[19] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov. Hacking commercial quantum cryptography systems by tailored bright illumination. *Nat. Photon.*, 4(10):686–689, Aug 2010. https://doi.org/10.1038/nphoton.2010.214.

[20] 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(1):1–6, Jun 2011. https://doi.org/10.1038/ncomms1348.

[21] N. Jain, C. Wittmann, L. Lydersen, C. Wiechers, D. Elser, C. Marquardt, V. Makarov, and G. Leuchs. Device calibration impacts security of quantum key distribution. *Phys. Rev. Lett.*, 107(11):110501, May 2011. https://doi.org/10.1103/PhysRevLett.107.110501.

[22] A.N. Bugge, S. Sauge, A.M.M. Ghazali, J. Skaar, L. Lydersen, and V. Makarov. Laser damage helps the eavesdropper in quantum cryptography. *Phys. Rev. Lett.*, 112(7):070503, Feb 2014. https://doi.org/10.1103/PhysRevLett.112.070503.

[23] D. Nadlinger, P. Drmota, B. Nichol, G. Araneda, D. Main, R. Srinivas, D. Lucas, C. Ballance, K. Ivanov, E.Z. Tan, et al. Experimental quantum key distribution certified by bell's theorem. *Nature*, 607(7920):682–686, Jul 2022. https://doi.org/10.1038/s41586-022-04941-5.

[24] W. Zhang, T. van Leent, K. Redeker, R. Garthoff, R. Schwonnek, F. Fertig, S. Eppelt, W. Rosenfeld, V. Scarani, C.C.W. Lim, et al. A device-independent quantum key distribution system for distant users. *Nature*, 607(7920):687–691, Jul 2022. https://doi.org/10.1038/s41586-022-04891-y.

[25] W.Z. Liu, Y.Z. Zhang, Y.Z. Zhen, M.H. Li, Y. Liu, J. Fan, F. Xu, Q. Zhang, and J.W. Pan. Toward a photonic demonstration of device-independent quantum key distribution. *Phys. Rev. Lett.*, 129(5):050502, Jul 2022. https://doi.org/10.1103/PhysRevLett.129.050502.

[26] M.M. Wilde. *Quantum information theory*. Cambridge university press, 2013.

[27] R. Renner. Security of quantum key distribution. *Int. J. Quantum Inf.*, 6(01):1–127, Feb 2008.

[28] A. Serafini. *Quantum continuous variables: a primer of theoretical methods*. CRC press, 2017.

[29] R. Garcia-Patron Sanchez. Quantum information with optical continuous variables: from bell tests to key distribution.

[30] D. Orsucci, J.D. Bancal, N. Sangouard, and P. Sekatski. How post-selection affects device-independent claims under the fair sampling assumption. *Quantum*, 4:238, Mar 2020. https://doi.org/10.22331/q-2020-03-02-238.

[31] C.H. Bennett. Quantum cryptography using any two nonorthogonal states. *Phys. Rev. Lett.*, 68:3121–3124, May 1992. https://link.aps.org/doi/10.1103/PhysRevLett.68.3121.

[32] J.S. Bell. On the Einstein Podolsky Rosen paradox. *Physics*, 1(3):195–200, Nov 1964. http://cds.cern.ch/record/111654/.

[33] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner. Bell nonlocality. *Rev. Mod. Phys.*, 86:419–478, Apr 2014. https://doi.org/10.1103/RevModPhys.86.419.

[34] J.F. Clauser, M.A. Horne, A. Shimony, and R.A. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, 23:880–884, Oct 1969. https://link.aps.org/doi/10.1103/PhysRevLett.23.880.

[35] S.J. Freedman and J.F. Clauser. Experimental test of local hidden-variable theories. *Phys. Rev. Lett.*, 28:938–941, Apr 1972. https://link.aps.org/doi/10.1103/PhysRevLett.28.938.

[36] B. Hensen, H. Bernien, A.E. Dréau, A. Reiserer, N. Kalb, M.S. Blok, J. Ruitenberg, R.F. Vermeulen, R.N. Schouten, C. Abellán, et al. Loophole-free bell inequality violation using electron spins separated by 1.3 kilometres. *Nature*, 526(7575):682–686, Oct 2015. https://doi.org/10.1038/nature15759.

[37] M. Giustina, M.A.M. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J.A. Larsson, C. Abellán, et al. Significant-loophole-free test of bell's theorem with entangled photons. *Phys. Rev. Lett.*, 115:250401, Dec 2015. https://link.aps.org/doi/10.1103/PhysRevLett.115.250401.

[38] L.K. Shalm, E. Meyer-Scott, B.G. Christensen, P. Bierhorst, M.A. Wayne, M.J. Stevens, T. Gerrits, S. Glancy, D.R. Hamel, M.S. Allman, et al. Strong loophole-free test of local realism. *Phys. Rev. Lett.*, 115:250402, Dec 2015. https://link.aps.org/doi/10.1103/PhysRevLett.115.250402.

[39] R. Uola, A.C.S. Costa, H.C. Nguyen, and O. Gühne. Quantum steering. *Rev. Mod. Phys.*, 92:015001, Mar 2020. https://link.aps.org/doi/10.1103/RevModPhys.92.015001.

[40] D. Mayers and A. Yao. Quantum cryptography with imperfect apparatus. *Proceedings 39th Annual Symposium on Foundations of Computer Science (Cat. No.98CB36280)*, pages 503–509, Sep 1998. https://doi.org/10.48550/arXiv.quant-ph/9809039.

[41] D. Mayers and A. Yao. Self testing quantum apparatus. *Quantum Information & Computation*, 4(4):273–286, Sep 2004. https://doi.org/10.48550/arXiv.quant-ph/0307205.

[42] T.J. Osborne and F. Verstraete. General monogamy inequality for bipartite qubit entanglement. *Phys. Rev. Lett.*, 96(22):220503, Jun 2006. https://doi.org/10.1103/PhysRevLett.96.220503.

[43] A.K. Ekert. Quantum cryptography based on bell's theorem. *Phys. Rev. Lett.*, 67:661–663, Aug 1991. https://link.aps.org/doi/10.1103/PhysRevLett.67.661.

[44] H.K. Lo, M. Curty, and B. Qi. Measurement-device-independent quantum key distribution. *Phys. Rev. Lett.*, 108(13):130503, Mar 2012. https://doi.org/10.1103/PhysRevLett.108.130503.

[45] S.L. Braunstein and S. Pirandola. Side-channel-free quantum key distribution. *Phys. Rev. Lett.*, 108:130502, Mar 2012. https://link.aps.org/doi/10.1103/PhysRevLett.108.130502.

[46] M. Ioannou, P. Sekatski, A.A. Abbott, D. Rosset, J.D. Bancal, and N. Brunner. Receiver-device-independent quantum key distribution protocols. *New J. Phys.*, 24(6):063006, Jun 2022. https://doi.org/10.1088%2F1367-2630%2Fac71bc.

[47] M. Ioannou, M.A. Pereira, D. Rusca, F. Grünenfelder, A. Boaron, M. Perrenoud, A.A. Abbott, P. Sekatski, J.D. Bancal, N. Maring, et al. Receiver-device-independent quantum key distribution. *Quantum*, 6:718, May 2022. https://doi.org/10.22331/q-2022-05-24-718.

[48] D. Mayers. Unconditional security in quantum cryptography. *J. ACM*, 48(3):351–406, May 2001. https://doi.org/10.48550/arXiv.quant-ph/9802025.

[49] M. Tomamichel and R. Renner. Uncertainty relation for smooth entropies. *Phys. Rev. Lett.*, 106:110506, Mar 2011. https://link.aps.org/doi/10.1103/PhysRevLett.106.110506.

[50] C. Branciard, E.G. Cavalcanti, S.P. Walborn, V. Scarani, and H.M. Wiseman. One-sided device-independent quantum key distribution: Security, feasibility, and the connection with steering. *Phys. Rev. A*, 85(1):010301, Jan 2012. https://doi.org/10.1103/PhysRevA.85.010301.

[51] M. Lucamarini, G. Vallone, I. Gianani, P. Mataloni, and G. Di Giuseppe. Device-independent entanglement-based bennett 1992 protocol. *Phys. Rev. A*, 86(3):032325, Sep 2012. https://doi.org/10.1103/PhysRevA.86.032325.

[52] M. Tomamichel, S. Fehr, J. Kaniewski, and S. Wehner. One-sided device-independent qkd and position-based cryptography from monogamy games. In *Advances in Cryptology–EUROCRYPT 2013: 32nd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Athens, Greece, May 26-30, 2013. Proceedings 32*, pages 609–625. Springer, Mar 2013.

[53] G. Vallone, A. Dall´ Arche, M. Tomasin, and P. Villoresi. Loss tolerant device-independent quantum key distribution: a proof of principle. *New J. Phys.*, 16(6):063064, Jun 2014. https://dx.doi.org/10.1088/1367-2630/16/6/063064.

[54] T. Gehring, V. Händchen, J. Duhme, F. Furrer, T. Franz, C. Pacher, R.F. Werner, and R. Schnabel. Implementation of continuous-variable quantum key distribution with composable and one-sided-device-independent security against coherent attacks. *Nat. Commun.*, 6(1):8795, Oct 2015. https://doi.org/10.1038/ncomms9795.

[55] J. Xin, X.M. Lu, X. Li, and G. Li. One-sided device-independent quantum key distribution for two independent parties. *Opt. Express*, 28(8):11439–11450, Apr 2020. https://opg.optica.org/oe/abstract.cfm?URI=oe-28-8-11439.

[56] A. Tavakoli. Semi-device-independent framework based on restricted distrust in prepare-and-measure experiments. *Phys. Rev. Lett.*, 126:210503, May 2021. https://link.aps.org/doi/10.1103/PhysRevLett.126.210503.

[57] A. Tavakoli, E. Zambrini Cruzeiro, E. Woodhead, and S. Pironio. Informationally restricted correlations: a general framework for classical and quantum systems. *Quantum*, 6:620, Jan 2022. https://doi.org/10.22331/q-2022-01-05-620.

[58] M. Pawłowski and N. Brunner. Semi-device-independent security of one-way quantum key distribution. *Phys. Rev. A*, 84(1):010302, Jul 2011. https://doi.org/10.1103/PhysRevA.84.010302.

[59] E. Woodhead and S. Pironio. Secrecy in prepare-and-measure clauser-horne-shimony-holt tests with a qubit bound. *Phys. Rev. Lett.*, 115(15):150501, Oct 2015. https://doi.org/10.1103/PhysRevLett.115.150501.

[60] E. Woodhead. Semi device independence of the bb84 protocol. *New J. Phys.*, 18(5):055010, Jun 2016. https://doi.org/10.1088/1367-2630/18/5/055010.

[61] Z.Q. Yin, C.H.F. Fung, X. Ma, C.M. Zhang, H.W. Li, W. Chen, S. Wang, G.C. Guo, and Z.F. Han. Mismatched-basis statistics enable quantum key distribution with uncharacterized qubit sources. *Phys. Rev. A*, 90:052319, Nov 2014. https://link.aps.org/doi/10.1103/PhysRevA.90.052319.

[62] H.M. Wiseman, S.J. Jones, and A.C. Doherty. Steering, entanglement, nonlocality, and the einstein-podolsky-rosen paradox. *Phys. Rev. Lett.*, 98:140402, Apr 2007. https://link.aps.org/doi/10.1103/PhysRevLett.98.140402.

[63] P.H. Eberhard. Background level and counter efficiencies required for a loophole-free Einstein-Podolsky-Rosen experiment. *Phys. Rev. A*, 47:R747–R750, Feb 1993. https://doi.org/10.1103/PhysRevA.47.R747.

[64] S. Massar and S. Pironio. Violation of local realism vs detection efficiency. *Phys. Rev. A*, 68(6):062109, Dec 2003. http://arxiv.org/abs/quant-ph/0210103.

[65] T. Vértesi, S. Pironio, and N. Brunner. Closing the detection loophole in bell experiments using qudits. *Phys. Rev. Lett.*, 104:060401, Feb 2010. https://link.aps.org/doi/10.1103/PhysRevLett.104.060401.

[66] C. Branciard. Detection loophole in bell experiments: How postselection modifies the requirements to observe nonlocality. *Phys. Rev. A*, 83:032123, Mar 2011. https://link.aps.org/doi/10.1103/PhysRevA.83.032123.

[67] J.Å. Larsson. Loopholes in bell inequality tests of local realism. *J. Phys. A*, 47(42):424003, Oct 2014. https://doi.org/10.1088%2F1751-8113%2F47%2F42%2F424003.

[68] B. Wittmann, S. Ramelow, F. Steinlechner, N.K. Langford, N. Brunner, H.M. Wiseman, R. Ursin, and A. Zeilinger. Loophole-free einstein–podolsky–rosen experiment via quantum steering. *New J. Phys.*, 14(5):053030, May 2012. https://doi.org/10.1088%2F1367-2630%2F14%2F5%2F053030.

[69] A.J. Bennet, D.A. Evans, D.J. Saunders, C. Branciard, E.G. Cavalcanti, H.M. Wiseman, and G.J. Pryde. Arbitrarily loss-tolerant einstein-podolsky-rosen steering allowing a demonstration over 1 km of optical fiber with no detection loophole. *Phys. Rev. X*, 2:031003, Jul 2012. https://link.aps.org/doi/10.1103/PhysRevX.2.031003.

[70] V. Srivastav, N.H. Valencia, W. McCutcheon, S. Leedumrongwatthanakun, S. Designolle, R. Uola, N. Brunner, and M. Malik. Quick quantum steering: Overcoming loss and noise with qudits. *Phys. Rev. X*, 12:041023, Nov 2022. https://link.aps.org/doi/10.1103/PhysRevX.12.041023.

[71] A. Acín, D. Cavalcanti, E. Passaro, S. Pironio, and P. Skrzypczyk. Necessary detection efficiencies for secure quantum key distribution and bound randomness. *Phys. Rev. A*, 93 (1):012319, Jan 2016. https://doi.org/10.1103/PhysRevA.93.012319.

[72] M. Ioannou, P. Sekatski, S. Designolle, B.D.M. Jones, R. Uola, and N. Brunner. Simulability of high-dimensional quantum measurements. *Phys. Rev. Lett.*, 129:190401, Nov 2022. https://link.aps.org/doi/10.1103/PhysRevLett.129.190401.

[73] T. Heinosaari, J. Kiukas, D. Reitzner, and J. Schultz. Incompatibility breaking quantum channels. *J. Phys. A*, 48(43):435301, Oct 2015. https://doi.org/10.1088/1751-8113/48/43/435301.

[74] P. Busch, P. Lahti, J.P. Pellonpää, and K. Ylinen. *Quantum Measurement*. Springer International Publishing, 2016. https://doi.org/10.1007/978-3-319-43389-9.

[75] O. Gühne, E. Haapasalo, T. Kraft, J.P. Pellonpää, and R. Uola. Colloquium: Incompatible measurements in quantum information science. *Rev. Mod. Phys.*, 95(1):011003, Feb 2023. https://doi.org/10.1103/RevModPhys.95.011003.

[76] R. Gallego, N. Brunner, C. Hadley, and A. Acín. Device-independent tests of classical and quantum dimensions. *Phys. Rev. Lett.*, 105(23):230501, Nov 2010. https://doi.org/10.1103/PhysRevLett.105.230501.

[77] Y. Wang, I.W. Primaatmaja, E. Lavie, A. Varvitsiotis, and C.C.W. Lim. Characterising the correlations of prepare-and-measure quantum networks. *npj Quantum Information*, 5 (1), Feb 2019. https://doi.org/10.1038%2Fs41534-019-0133-3.

[78] M.T. Quintino, T. Vértesi, and N. Brunner. Joint measurability, einstein-podolsky-rosen steering, and bell nonlocality. *Phys. Rev. Lett.*, 113:160402, Oct 2014. https://link.aps.org/doi/10.1103/PhysRevLett.113.160402.

[79] R. Uola, T. Moroder, and O. Gühne. Joint Measurability of Generalized Measurements Implies Classicality. *Phys. Rev. Lett.*, 113(16):160403, Oct 2014. https://link.aps.org/doi/10.1103/PhysRevLett.113.160403.

[80] T. Heinosaari, J. Kiukas, and D. Reitzner. Noise robustness of the incompatibility of quantum measurements. *Phys. Rev. A*, 92:022115, Aug 2015. https://link.aps.org/doi/10.1103/PhysRevA.92.022115.

[81] J. Bavaresco, M.T. Quintino, L. Guerini, T.O. Maciel, D. Cavalcanti, and M.T. Cunha. Most incompatible measurements for robust steering tests. *Phys. Rev. A*, 96:022110, Aug 2017. https://link.aps.org/doi/10.1103/PhysRevA.96.022110.

[82] S. Designolle, P. Skrzypczyk, F. Fröwis, and N. Brunner. Quantifying measurement incompatibility of mutually unbiased bases. *Phys. Rev. Lett.*, 122:050402, Feb 2019. https://link.aps.org/doi/10.1103/PhysRevLett.122.050402.

[83] P. Skrzypczyk and D. Cavalcanti. Loss-tolerant einstein-podolsky-rosen steering for arbitrary-dimensional states: Joint measurability and unbounded violations under losses. *Phys. Rev. A*, 92(2):022354, Aug 2015. https://doi.org/10.1103/physreva.92.022354.

[84] P. Sekatski, F. Giraud, R. Uola, and N. Brunner. Unlimited one-way steering. *Phys. Rev. Lett.*, 131(11):110201, Sep 2023. https://doi.org/10.1103/PhysRevLett.131.110201.

[85] P. Sekatski. Compatibility of projective measurements subject to white noise and loss. *Phys. Rev. A*, 109(2):022215, Feb 2024. https://doi.org/10.1103/PhysRevA.109.022215.

[86] L. Lami, S. Khatri, G. Adesso, and M.M. Wilde. Extendibility of bosonic gaussian states. *Phys. Rev. Lett.*, 123(5):050501, Jul 2019. https://doi.org/10.1103/PhysRevLett.123.050501.

[87] S. Rahimi-Keshari, M. Mehboudi, D. De Santis, D. Cavalcanti, and A. Acín. Verification of joint measurability using phase-space quasiprobability distributions. *Phys. Rev. A*, 104 (4):042212, Oct 2021. http://arxiv.org/abs/2012.06853.

[88] M. Curty, M. Lewenstein, and N. Lütkenhaus. Entanglement as a Precondition for Secure Quantum Key Distribution. *Phys. Rev. Lett.*, 92(21):217903, May 2004. https://link.aps.org/doi/10.1103/PhysRevLett.92.217903.

[89] E.P. Lobo, J. Pauwels, and S. Pironio. Certifying long-range quantum correlations through routed bell tests. *Quantum*, 8:1332, May 2024. https://doi.org/10.22331/q-2024-05-02-1332.

[90] M.F. Pusey. Verifying the quantumness of a channel with an untrusted device. *arXiv preprint arxiv:1502.03010*, Mar 2015. http://arxiv.org/abs/1502.03010.

[91] A. Holevo, M. Shirokov, and R. Werner. Separability and entanglement-breaking in infinite dimensions. *arXiv preprint quant-ph/0504204*, 2005. https://doi.org/10.48550/arXiv.quant-ph/0504204.

[92] M.M. Wolf, D. Perez-Garcia, and C. Fernandez. Measurements incompatible in quantum theory cannot be measured jointly in any other no-signaling theory. *Phys. Rev. Lett.*, 103 (23):230402, Dec 2009. https://doi.org/10.1103/physrevlett.103.230402.

[93] R.F. Werner. Optimal cloning of pure states. *Phys. Rev. A*, 58(3):1827, Sep 1998. https://doi.org/10.1103/PhysRevA.58.1827.

[94] M. Keyl and R.F. Werner. Optimal cloning of pure states, testing single clones. *J. Math. Phys.*, 40(7):3283–3299, Jul 1999. https://doi.org/10.1063/1.532887.

[95] P. Busch. Unsharp reality and joint measurements for spin observables. *Phys. Rev. D*, 33: 2253–2261, Apr 1986. https://link.aps.org/doi/10.1103/PhysRevD.33.2253.

[96] R. Pal and S. Ghosh. Approximate joint measurement of qubit observables through an arthur–kelly model. *J. Phys. A*, 44(48):485303, Nov 2011. https://dx.doi.org/10.1088/1751-8113/44/48/485303.

[97] S. Yu and C. Oh. Quantum contextuality and joint measurement of three observables of a qubit. *arXiv preprint arXiv:1312.6470*, Dec 2013. https://doi.org/10.48550/arXiv.1312.6470.

[98] J. Kiukas and J. Schultz. Informationally complete sets of gaussian measurements. *J. Phys. A*, 46(48):485303, Nov 2013. https://doi.org/10.1088/1751-8113/46/48/485303.

[99] R. Arnon-Friedman, R. Renner, and T. Vidick. Simple and tight device-independent security proofs. *SIAM J. Comput.*, 48(1):181–225, Jan 2019. https://doi.org/10.1137/18M1174726.

[100] E.Y.Z. Tan, P. Sekatski, J.D. Bancal, R. Schwonnek, R. Renner, N. Sangouard, and C.C.W. Lim. Improved diqkd protocols with finite-size analysis. *Quantum*, 6:880, Dec 2022. https://doi.org/10.48550/arXiv.2012.08714.

[101] R. Colbeck and R. Renner. No extension of quantum theory can have improved predictive power. *Nat. Commun.*, 2(1):411, Aug 2011.

[102] M. Tomamichel, C. Schaffner, A. Smith, and R. Renner. Leftover hashing against quantum side information. *IEEE Trans. Inf. Theory*, 57(8):5524–5535, Aug 2011. https://doi.org/10.1109/TIT.2011.2158473.

[103] B.A. Slutsky, R. Rao, P.C. Sun, and Y. Fainman. Security of quantum cryptography against individual attacks. *Phys. Rev. A*, 57:2383–2398, Apr 1998. https://link.aps.org/doi/10.1103/PhysRevA.57.2383.

[104] N. Lütkenhaus. Estimates for practical quantum cryptography. *Phys. Rev. A*, 59:3301–3319, May 1999. https://link.aps.org/doi/10.1103/PhysRevA.59.3301.

[105] H. Bechmann-Pasquinucci and N. Gisin. Incoherent and coherent eavesdropping in the six-state protocol of quantum cryptography. *Phys. Rev. A*, 59:4238–4248, Jun 1999. https://link.aps.org/doi/10.1103/PhysRevA.59.4238.

[106] N.J. Cerf, M. Bourennane, A. Karlsson, and N. Gisin. Security of quantum key distribution using d-level systems. *Phys. Rev. Lett.*, 88:127902, Mar 2002. https://link.aps.org/doi/10.1103/PhysRevLett.88.127902.

[107] E. Biham and T. Mor. Security of quantum cryptography against collective attacks. *Phys. Rev. Lett.*, 78:2256–2259, Mar 1997. https://link.aps.org/doi/10.1103/PhysRevLett.78.2256.

[108] E. Biham, M. Boyer, G. Brassard, J. van de Graaf, and T. Mor. Security of quantum key distribution against all collective attacks. *Algorithmica*, 34:372–388, Nov 2002. https://doi.org/10.1007/s00453-002-0973-6.

[109] M. Tomamichel, R. Colbeck, and R. Renner. A fully quantum asymptotic equipartition property. *IEEE Trans. Inf. Theory*, 55(12):5840–5847, Dec 2009. https://doi.org/10.1109/TIT.2009.2032797.

[110] I. Devetak and A. Winter. Distillation of secret key and entanglement from quantum states. *Proc. R. Soc. Lond. A*, 461(2053):207–235, Jan 2005. https://doi.org/10.1098/rspa.2004.1372.

[111] F. Dupuis, O. Fawzi, and R. Renner. Entropy accumulation. *Commun. Math. Phys.*, 379 (3):867–913, Sep 2020. https://doi.org/10.1007/s00220-020-03839-5.

[112] T. Metger, O. Fawzi, D. Sutter, and R. Renner. Generalised entropy accumulation. In *2022 IEEE 63rd An. Symp. on FOCS*, pages 844–850. IEEE, Dec 2022. https://doi.org/10.1109/FOCS54457.2022.00085.

[113] G. de la Torre, J.D. Bancal, S. Pironio, V. Scarani, et al. Randomness in post-selected events. *New J. Phys.*, 18(3):035007, Mar 2016. https://doi.org/10.1088/1367-2630/18/3/035007.

[114] M. Sandfuchs and R. Wolf. Coherent attacks are stronger than collective attacks on diqkd with random postselection. *arXiv preprint arXiv:2306.07364*, Jun 2023. https://doi.org/10.48550/arXiv.2306.07364.

[115] R. Renner and S. Wolf. New bounds in secret-key agreement: The gap between formation and secrecy extraction. In *Advances in Cryptology—EUROCRYPT 2003: International Conference on the Theory and Applications of Cryptographic Techniques, Warsaw, Poland, May 4–8, 2003 Proceedings 22*, pages 562–577. Springer, 2003.

[116] R. Renner, N. Gisin, and B. Kraus. Information-theoretic security proof for quantum-key-distribution protocols. *Phys. Rev. A*, 72:012332, Jul 2005. https://doi.org/10.1103/PhysRevA.72.012332.

[117] J. Kołodyński, A. Máttar, P. Skrzypczyk, E. Woodhead, D. Cavalcanti, K. Banaszek, and A. Acín. Device-independent quantum key distribution with single-photon sources. *Quantum*, 4:260, Apr 2020. https://doi.org/10.22331/q-2020-04-30-260.

[118] T.L. Roy-Deloison, E.P. Lobo, J. Pauwels, and S. Pironio. Device-independent quantum key distribution based on routed bell tests. *arXiv preprint arXiv:2404.01202*, Apr 2024. https://doi.org/10.48550/arXiv.2404.01202.

[119] M. Farkas, M. Balanzó-Juandó, K. Łukanowski, J. Kołodyński, and A. Acín. Bell nonlocality is not sufficient for the security of standard device-independent quantum key distribution protocols. *Phys. Rev. Lett.*, 127(5):050503, Jul 2021. https://doi.org/10.1103/PhysRevLett.127.050503.

[120] K. Łukanowski, M. Balanzó-Juandó, M. Farkas, A. Acín, and J. Kołodyński. Upper bounds on key rates in device-independent quantum key distribution based on convex-combination attacks. *Quantum*, 7:1199, Dec 2023. https://doi.org/10.22331/q-2023-12-06-1199.

[121] A. Acín, S. Massar, and S. Pironio. Efficient quantum key distribution secure against no-signalling eavesdroppers. *New J. Phys.*, 8(8):126, Aug 2006. https://dx.doi.org/10.1088/1367-2630/8/8/126.

[122] J.A. Nelder and R. Mead. A Simplex Method for Function Minimization. *The Computer Journal*, 7(4):308–313, Jan 1965. https://doi.org/10.1093/comjnl/7.4.308.

[123] T. Heinosaari, D. Reitzner, and P. Stano. Notes on joint measurability of quantum observables. *Foundations of Physics*, 38(12):1133–1147, Nov 2008. https://doi.org/10.1007%2Fs10701-008-9256-7.

[124] Y.C. Liang, R.W. Spekkens, and H.M. Wiseman. Specker's parable of the overprotective seer: A road to contextuality, nonlocality and complementarity. *Physics Reports*, 506(1):1–39, May 2011. https://www.sciencedirect.com/science/article/pii/S0370157311001517.

[125] F. Buscemi, K. Kobayashi, S. Minagawa, P. Perinotti, and A. Tosini. Unifying different notions of quantum incompatibility into a strict hierarchy of resource theories of communication. *Quantum*, 7:1035, Jun 2023. https://doi.org/10.22331/q-2023-06-07-1035.

[126] G.M. D'Ariano, P. Perinotti, and A. Tosini. Incompatibility of observables, channels and instruments in information theories. *J. Phys. A*, 55(39):394006, Sep 2022. https://dx.doi.org/10.1088/1751-8121/ac88a7.

[127] R. Arnon-Friedman, F. Dupuis, O. Fawzi, R. Renner, and T. Vidick. Practical device-independent quantum cryptography via entropy accumulation. *Nat. Commun.*, 9:459, Jan 2018. https://doi.org/10.1038/s41467-017-02307-4.

[128] Y. Zhang, H. Fu, and E. Knill. Efficient randomness certification by quantum probability estimation. *Phys. Rev. Res.*, 2:013016, Jan 2020. https://doi.org/10.1103/PhysRevResearch.2.013016.

[129] E.Y.Z. Tan, R. Schwonnek, K.T. Goh, I.W. Primaatmaja, and C.C.W. Lim. Computing secure key rates for quantum cryptography with untrusted devices. *npj Quantum Information*, 7(1):1–6, Oct 2021. https://doi.org/10.1038/s41534-021-00494-z.

[130] R. Schwonnek, K.T. Goh, I.W. Primaatmaja, E.Y.Z. Tan, R. Wolf, V. Scarani, and C.C.W. Lim. Device-independent quantum key distribution with random key basis. *Nat. Commun.*, May 2021. https://doi.org/10.1038/s41467-021-23147-3.

[131] P. Brown, H. Fawzi, and O. Fawzi. Computing conditional entropies for quantum correlations. *Nat. Commun.*, 12:575, Jan 2021. https://doi.org/10.1038/s41467-020-20018-1.

[132] P. Brown, H. Fawzi, and O. Fawzi. Device-independent lower bounds on the conditional von neumann entropy. *arXiv preprint 2106.13692*, Jun 2021. https://doi.org/10.48550/arXiv.2106.13692.

[133] E. Woodhead, A. Acín, and S. Pironio. Device-independent quantum key distribution with asymmetric CHSH inequalities. *Quantum*, 5:443, Apr 2021. https://doi.org/10.22331/q-2021-04-26-443.

[134] O. Kern and J.M. Renes. Improved one-way rates for BB84 and 6-state protocols. *Quantum Inf. Comput.*, 8(8,9):0756–0772, Sep 2008. https://doi.org/10.26421/QIC8.8-9-6.

[135] S. Pironio, A. Acín, N. Brunner, N. Gisin, S. Massar, and V. Scarani. Device-independent quantum key distribution secure against collective attacks. *New J. Phys.*, 11(4):045021, Apr 2009. https://doi.org/10.1088/1367-2630/11/4/045021.

[136] M. Berta, M. Christandl, R. Colbeck, J.M. Renes, and R. Renner. The uncertainty principle in the presence of quantum memory. *Nat. Phys.*, 6:659–662, Jul 2010. https://doi.org/10.1038/nphys1734.

[137] E. Woodhead. Tight asymptotic key rate for the Bennett-Brassard 1984 protocol with local randomization and device imprecisions. *Phys. Rev. A*, 90:022306, Aug 2014. https://doi.org/10.1103/PhysRevA.90.022306.

[138] J.B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Comput.*, 11:796–817, Jan 2001. https://doi.org/10.1137/S1052623400366802.

[139] D. Henrion and J.B. Lasserre. Convergent relaxations of polynomial matrix inequalities and static output feedback. *IEEE Trans. Autom. Control*, 51(2):192–202, Feb 2006. https://doi.org/10.1109/TAC.2005.863494.

[140] D. McCallum and D. Avis. A linear algorithm for finding the convex hull of a simple polygon. *Inf. Process. Lett.*, 9(5):201–206, Dec 1979. https://doi.org/10.1016/0020-0190(79)90069-3.

[141] A.A. Schäffer and C.J. Van Wyk. Convex hulls of piecewise-smooth Jordan curves. *J. Algorithms*, 8(1):66–94, Mar 1987. https://doi.org/10.1016/0196-6774(87)90028-9.

[142] E. Woodhead. Quantum cloning bound and application to quantum key distribution. *Phys. Rev. A*, 88:012331, Jul 2013. https://doi.org/10.1103/PhysRevA.88.012331.

[143] R. Bhavsar, S. Ragy, and R. Colbeck. Calculation and application of various von Neumann entropies in CHSH-based device-independent randomness expansion. Mar 2021. https://doi.org/10.48550/arXiv.2103.07504.

[144] S. Pironio, A. Acín, S. Massar, A. Boyer de La Giroday, D.N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T.A. Manning, and C. Monroe. Random numbers certified by Bell's theorem. *Nature*, 464:1021–1024, Apr 2010. https://doi.org/10.1038/nature09008.

[145] A. Acín, S. Massar, and S. Pironio. Randomness versus nonlocality and entanglement. *Phys. Rev. Lett.*, 108:100402, Mar 2012. https://doi.org/10.1103/PhysRevLett.108.100402.

[146] M. Navascués, S. Pironio, and A. Acín. A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. *New J. Phys.*, 10(7):073013, Jul 2008. https://dx.doi.org/10.1088/1367-2630/10/7/073013.

[147] S. Pironio, M. Navascués, and A. Acín. Convergent relaxations of polynomial optimization problems with noncommuting variables. *SIAM J. Optim.*, 20(5):2157–2180, Jan 2010. https://doi.org/10.1137/090760155.

[148] https://mathworld.wolfram.com/RadauQuadrature.html.

[149] P. Brown. private communication.

[150] X. Zhang, P. Zeng, T. Ye, H.K. Lo, and X. Ma. Quantum complementarity approach to device-independent security. *Phys. Rev. Lett.*, 131(14):140801, Oct 2023. https://doi.org/10.48550/arXiv.2111.13855.

[151] A. Winick, N. Lütkenhaus, and P.J. Coles. Reliable numerical key rates for quantum key distribution. *Quantum*, 2:77, Jul 2018. https://doi.org/10.22331/q-2018-07-26-77.

[152] N.D. Mermin. Extreme quantum entanglement in a superposition of macroscopically distinct states. *Phys. Rev. Lett.*, 65:1838–1840, Oct 1990. https://doi.org/10.1103/PhysRevLett.65.1838.

[153] F. Grasselli, G. Murta, H. Kampermann, and D. Bruß. Entropy bounds for multiparty device-independent cryptography. *PRX Quantum*, 2:010308, Jan 2021. https://doi.org/10.1103/PRXQuantum.2.010308.

[154] https://github.com/MicheleMasini1996/diqkd-2input2output.

[155] https://github.com/MicheleMasini1996/BFF.git.

[156] G. Vallone. Einstein-podolsky-rosen steering: Closing the detection loophole with non-maximally-entangled states and arbitrary low efficiency. *Phys. Rev. A*, 87(2):020101, Feb 2013. https://doi.org/10.1103/PhysRevA.87.020101.

[157] N. Gisin, S. Pironio, and N. Sangouard. Proposal for implementing device-independent quantum key distribution based on a heralded qubit amplifier. *Phys. Rev. Lett.*, 105(7): 070501, Aug 2010. https://doi.org/10.1103/PhysRevLett.105.070501.

[158] S. Wollmann, N. Walk, A.J. Bennet, H.M. Wiseman, and G.J. Pryde. Observation of genuine one-way einstein-podolsky-rosen steering. *Phys. Rev. Lett.*, 116(16):160403, Apr 2016. https://doi.org/10.1103/PhysRevLett.116.160403.

[159] V. Zapatero and M. Curty. Long-distance device-independent quantum key distribution. *Scientific Reports*, 9(1):17749, Nov 2019. https://doi.org/10.1038/s41598-019-53803-0.

[160] C. Jordan. Essai sur la géométrie à n dimensions. *Bull. Soc. Math. Fr.*, 3:103–174, May 1875. http://www.numdam.org/item?id=BSMF_1875__3__103_2.

[161] M. Ho, P. Sekatski, E.Y.Z. Tan, R. Renner, J.D. Bancal, and N. Sangouard. Noisy preprocessing facilitates a photonic realization of device-independent quantum key distribution. *Phys. Rev. Lett.*, 124:230502, Jun 2020. https://doi.org/10.1103/PhysRevLett.124.230502.

[162] M.H. Li, C. Wu, Y. Zhang, W.Z. Liu, B. Bai, Y. Liu, W. Zhang, Q. Zhao, H. Li, Z. Wang, et al. Test of local realism into the past without detection and locality loopholes. *Phys. Rev. Lett.*, 121:080404, Aug 2018. https://link.aps.org/doi/10.1103/PhysRevLett.121.080404.

[163] Y. Liu, X. Yuan, M.H. Li, W. Zhang, Q. Zhao, J. Zhong, Y. Cao, Y.H. Li, L.K. Chen, H. Li, et al. High-speed device-independent quantum random number generation without a detection loophole. *Phys. Rev. Lett.*, 120:010503, Jan 2018. https://link.aps.org/doi/10.1103/PhysRevLett.120.010503.

[164] L. Shen, J. Lee, L.P. Thinh, J.D. Bancal, A. Cerè, A. Lamas-Linares, A. Lita, T. Gerrits, S.W. Nam, V. Scarani, and C. Kurtsiefer. Randomness extraction from bell violation with continuous parametric down-conversion. *Phys. Rev. Lett.*, 121:150402, Oct 2018. https://link.aps.org/doi/10.1103/PhysRevLett.121.150402.

[165] P. Bierhorst, E. Knill, S. Glancy, Y. Zhang, A. Mink, S. Jordan, A. Rommal, Y.K. Liu, B. Christensen, S.W. Nam, et al. Experimentally generated randomness certified by the impossibility of superluminal signals. *Nature*, 556(7700):223–226, Apr 2018. https://doi.org/10.1038/s41586-018-0019-0.

[166] Y. Liu, Q. Zhao, M.H. Li, J.Y. Guan, Y. Zhang, B. Bai, W. Zhang, W.Z. Liu, C. Wu, X. Yuan, et al. Device-independent quantum random-number generation. *Nature*, 562 (7728):548–551, Sep 2018. https://doi.org/10.1038/s41586-018-0559-3.

[167] Y. Zhang, L.K. Shalm, J.C. Bienfang, M.J. Stevens, M.D. Mazurek, S.W. Nam, C. Abellán, W. Amaya, M.W. Mitchell, H. Fu, et al. Experimental low-latency device-independent quantum randomness. *Phys. Rev. Lett.*, 124:010505, Jan 2020. https://link.aps.org/doi/10.1103/PhysRevLett.124.010505.

[168] W.Z. Liu, M.H. Li, S. Ragy, S.R. Zhao, B. Bai, Y. Liu, P.J. Brown, J. Zhang, R. Colbeck, J. Fan, Q. Zhang, and J.W. Pan. Device-independent randomness expansion against quantum side information. *Nat. Phys.*, 17:448–451, Feb 2021. https://doi.org/10.1038/s41567-020-01147-2.

[169] M.H. Li, X. Zhang, W.Z. Liu, S.R. Zhao, B. Bai, Y. Liu, Q. Zhao, Y. Peng, J. Zhang, Y. Zhang, et al. Experimental realization of device-independent quantum randomness expansion. *Phys. Rev. Lett.*, 126:050503, Feb 2021. https://link.aps.org/doi/10.1103/PhysRevLett.126.050503.

[170] L.K. Shalm, Y. Zhang, J.C. Bienfang, C. Schlager, M.J. Stevens, M.D. Mazurek, C. Abellán, W. Amaya, M.W. Mitchell, M.A. Alhejji, et al. Device-independent randomness expansion with entangled photons. *Nat. Phys.*, 17(4):452–456, Jan 2021. https://doi.org/10.1038/s41567-020-01153-4.

[171] T. Van Himbeeck, E. Woodhead, N.J. Cerf, R. García-Patrón, and S. Pironio. Semi-device-independent framework based on natural physical assumptions. *Quantum*, 1:33, Nov 2017. https://doi.org/10.22331/q-2017-11-18-33.

[172] F. Xu, Y.Z. Zhang, Q. Zhang, and J.W. Pan. Device-independent quantum key distribution with random postselection. *Phys. Rev. Lett.*, 128:110506, Mar 2022. https://link.aps.org/doi/10.1103/PhysRevLett.128.110506.