Finite Key Analysis in Quantum Cryptography

Inaugural-Dissertation

zur

Erlangung des Doktorgrades der

Mathematisch-Naturwissenschaftlichen Fakultät

der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Tim Meyer

aus Neustadt am Rübenberge

Oktober 2007

Aus dem Institut für Theoretische Physik, Lehrstuhl III
der Heinrich-Heine Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Referent: Prof. Dr. D. Bruß
Koreferent: Prof. Dr. R. Egger
Tag der mündlichen Prüfung: 31. Oktober 2007

"Die grundlegende Theorie [der Quantenkryptographie] ist [...] schon geklärt. Die Theoriefragen beziehen sich heute hauptsächlich auf die Fragen der Sicherheit, es geht dabei beispielsweise um grundsätzliche Sicherheitsbeweise der Systeme. Was interessanterweise bei unseren Methoden etwas ist, das wir nicht benötigen.

Es gibt verschiedene Möglichkeiten, physisch diese Quantenverbindungen, die dann die Schlüssel austauschen, zur Verschlüsselung zu realisieren. Wir machen das auf eine Art, bei der die Sicherheit offenkundig ist, dazu brauchen wir nicht einmal einen Beweis."

O. Univ.-Prof. Dr. phil. Anton Zeilinger, e&i, Heft 5 2007

Abstract

In view of experimental realization of quantum key distribution schemes, the study of their efficiency becomes as important as the proof of their security. The latter is the subject of most of the theoretical work about quantum key distribution, and many important results such as the proof of unconditional security have been obtained. The efficiency and also the robustness of quantum key distribution protocols against noise can be measured by figures of merit such as the secret key rate (the fraction of input signals that make it into the key) and the threshold quantum bit error rate (the maximal error rate such that one can still create a secret key). It is important to determine these quantities because they tell us whether a certain quantum key distribution scheme can be used at all in a given situation and if so, how many secret key bits it can generate in a given time. However, these figures of merit are usually derived under the "infinite key limit" assumption, that is, one assumes that an infinite number of quantum states are send and that all sub-protocols of the scheme (in particular privacy amplification) are carried out on these infinitely large blocks. Such an assumption usually eases the analysis, but also leads to (potentially) too optimistic values for the quantities in question.

In this thesis, we are explicitly avoiding the infinite key limit for the analysis of the privacy amplification step, which plays the most important role in a quantum key distribution scheme. We still assume that an optimal error correction code is applied and we do not take into account any statistical errors that might occur in the parameter estimation step. In [1], Renner and coworkers derived an explicit formula for the obtainable key rate in terms of Renyi entropies of the quantum states describing Alice's, Bob's, and Eve's systems. This results serves as a starting point for our analysis, and we derive an algorithm that efficiently computes the obtainable key rate for any finite number of input signals, without making any approximations.

As an application, we investigate the so-called "Tomographic Protocol" [2, 3], which is based on the Six-State Protocol [4, 5] and where Alice and Bob can obtain the additional information which quantum state they share after the distribution step of the

protocol. We calculate the obtainable secret key rate under the assumption that the eavesdropper only conducts collective attacks and give a detailed analysis of the dependence of the key rate on various parameters: The number of input signals (the block size), the error rate in the sifted key (the QBER), and the security parameter. Furthermore, we study the influence of multi-photon events which naturally occur in a realistic implementation.

Zusammenfassung

Im Zuge der experimentellen Realisierung von Protokollen zur quantenmechanischen Schlüsselverteilung wird die Analyse ihrer Effizienz genauso wichtig wie der Beweis ihrer Sicherheit. Letzteres ist das Thema der meisten theoretischen Arbeiten auf diesem Gebiet, welche wichtige Ergebnisse lieferten, wie etwa der Beweis der unbedingten Abhörsicherheit. Die Effizienz und die Robustheit eines Protokolls lassen sich durch Gütekriterien wie die Schlüsselrate (der Bruchteil der gesendeten Signale, die den Schlüssel bilden) oder die Schwellen-Quantenfehlerrate (die maximale tolerierbare Fehlerrate, bei der die Schlüsselerzeugung noch möglich ist) definieren. Diese Größen müssen bestimmt werden, um für ein gegebenes Szenario festzustellen, ob ein gewisses Protokoll überhaupt anwendbar ist und wenn ja, wieviele Bits sicherer Schlüssel generiert werden können. Im Allgemeinen jedoch können diese Gütekriterien nur unter der Annahme berechnet werden, dass alle Zwischenschritte des Protokolls — insbesondere der *privacy amplification*-Schritt — mit unendlich vielen Signalen arbeiten. Diese Annahme erleichtert die Analyse zwar, allerdings werden dadurch möglicherweise zu optimitische Werte für die Gütekriterien errechnet.

Aus diesem Grund vermeiden wir in dieser Arbeit die Annahme der unendlich vielen Signale für den *privacy amplification*-Schritt, welcher der wichtigste in einem Schlüsselverteilungsprotokoll ist. Jedoch nehmen wir weiterhin an, dass nur optimale Fehlerkorrekturcodes verwendet werden und wir berücksichtigen auch keine statistischen Fehler, die im Parameter-Abschätzungsschritt auftreten können. In [1] haben Renner et al. eine explizite Formel für die erreichbare Schlüsselrate bzgl. Renyi-Entropien der Quanten-Zustände, die Alices, Bobs und Eves Quanten-System beschreiben, ermittelt. Dieses Ergebnis ist der Ausgangspunkt für unsere Analyse, in der wir einen Algorithmus entwickeln, welcher die erreichbare Schlüsselrate für jegliche Anzahl von Signalen effizient berechnet, ohne auf Näherungen zurückzugreifen.

Als eine Anwendung betrachten wir das sogenannte “Tomographische Protokoll” [2, 3], welches auf dem *Six-State*-Protokoll [4, 5] basiert, und in welchem Alice und Bob zusätzlich bestimmen können, welchen Quantenzustand sie sich nach dem Verteilungss-

chritt des Protokolls teilen. Wir berechnen die erreichbare Schlüsselrate unter der Annahme, dass Eve nur kollektive Attacken durchführt und analysieren detailliert, auf welche Weise die Schlüsselrate von folgenden Parametern abhängt: Die Anzahl der Eingangssignale (die Blocklänge), die Fehlerrate im "gesiebten" Schlüssel (die QBER) und der Sicherheitsparameter. Außerdem untersuchen wir den Einfluß von Mehr-Photonen-Signalen, welche in jeder realistischen Anwendung auftreten.

Contents

Chapter 1

Introduction

Quantum key distribution has taken the step from a theoretician's mind to experimental implementation, becoming a commercial product [6, 7]. During the last basic models which barely described the real world, culminating in the proof of universal composability and security under the most general circumstances. Besides these conceptual proofs it is necessary to investigate the performance of protocols. What rate of secret key bits can be generated by a specific setup? The answer to this question certainly depends on the type of equipment that is used in the actual implementation. A performance measure which only depends on the underlying protocol is given by the number of secret key bits per quantum signal sent. The main subject of this thesis is to compute such a figure of merit (the secret key rate) for a restricted class of quantum key distribution protocols.

1.1 Secret Communication

The ability to secretly communicate has always been of great importance in many aspects of our life: Already in 500 B.C., the Spartans invented a cryptographic device called *scytale* [8]. This is a wooden rod around which one wraps a strap of leather or parchment. Afterwards, the message which is to be encrypted (also called plain text) is written onto the strap such that each letter appears on a new twist. Then the strap is unwrapped, now showing only incoherent letters, and is transported to the receiver who owns a scytale of the same diameter to decode the message. The scytale implements an encryption method known as shift cipher, in which every letter of the plain text gets shifted by a fixed amount. Another example is the Caesar cipher [9], invented by the Roman emperor in 50 B.C., in which letters of the plain text alphabet are replaced by certain other letters. Such a simple scrambling of the plain text renders a message unreadable, at least at first

glance, but 2000 years ago this was apparently enough to scare off adversaries. Ever since then, cryptography (the science of code *making*) and cryptoanalysis (the science of code *breaking*) are two fields constantly feeding each other: Whenever some encryption scheme gets broken, cryptographers are forced to invent a new code, being even harder to break. This in turn encourages code breakers to search for flaws in this new scheme. Eventually, this mutual outperforming has led to the situation we are facing today: The search for an encryption method which is *inherently* secure, thus impossible to break.

RSA

One of the most successful cryptosystems used today is the RSA system, named after its inventors Rivest, Shamir, and Adleman [10]. It is an asymmetric cryptosystem, employing two keys, a private and a public one. The public key (as the name suggests) is announced to everybody who might be willing to communicate secretly with the holder of the private key, which is kept secret. Mathematically, the scheme is based on so-called (trap door) one-way functions [11], which are easy to compute (using the public key), but hard to invert, unless one possesses some kind of "trap door" information, the private key. In this way it is guaranteed that, only having access to the public key, everybody can encrypt messages, but lacking the private key, one cannot decrypt them. However, the fact that one-way functions are "hard" to invert is merely a matter of observation rather than a mathematical statement. Being hard to calculate in this sense means that there has not yet been found any algorithm solving the task in polynomial time. By a reasonable choice of the size of some input parameters (the key length), one can ensure that using any known algorithm, computing the plain text from the cipher text while only knowing the public key becomes unfeasible, as the time needed for these algorithms to finish can be made arbitrarily large.

Still, there are two problems threatening the applicability of RSA: First, it might happen that an algorithm is found, which can invert one-way functions in a polynomial time. Although such a discovery seems unlikely, is has not yet been ruled out by a rigorous mathematical proof. Second, and possibly more severe, as computer power is increasing exponentially all times [12], brute force methods becoming more and more feasible. For instance, older implementation of RSA, using a built-in key length which appeared to providing enough security a decade ago can potentially be broken these days.[1] Also with the advent of the quantum computer, which might be superior to

[1] In 1977, it was supposed to take about $40 \cdot 10^{15}$ years (one million times the age of the universe) to break a 425-bit key. In 1994, 1600 computers "only" needed eight months, and nowadays, a single desktop PC could to the same job. It is recommended today to use at least 2048-bit keys [13].

classical algorithms, it is not clear how long it takes until inverting one-way functions becomes feasible.

Unconditional Security

To provide secure communication which is not suffering the flaw of potentially becoming insecure at a certain time in the future, a new type of cryptographic application is needed. We call a scheme *unconditionally secure* if its security can be mathematically proven and if it is not based on any assumptions about the adversary's abilities, such as being limited in computer power, memory, or time. Fortunately, such a scheme exists:[2] The *Vernam cipher* [14] is used to encrypt messages given in binary notation using a key consisting of random bits which is as long as the message and shared between the parties which wish to communicate. Encryption is performed by calculating bitwise addition modulo two of the plain text and the key, and decryption by adding the cipher text to the key. It has been shown by Shannon [15] that such a cryptographic scheme is unconditionally secure if the key is completely random and only used once (and is of course unknown to the adversary), thus the name "one-time pad". The proof of its security is quite intuitive, since the result $m_i \oplus k_i$ of the addition of a plain text bit m_i to a key bit k_i is completely random if the key bit is completely random. Therefore, the cipher text bit $m_i \oplus k_i$ does not contain any information about the plain text m_i, and consequently the scheme is perfectly secure.

The catch of the one-time pad is the following: Since the key is a random bit string of the same length as the message, which needs to be generated from scratch for each message to be sent, one faces the problem of distributing large amounts of data. Moreover, the problem of keeping this key secret remains. In former times, code books where employed, where pages with used codes were torn out. For state-of-the-art applications, we need to find a reliable and efficient key distribution scheme.

1.2 Quantum Key Distribution

Quantum Key Distribution (QKD) aims exactly at providing such a distribution scheme for random keys. To do so, a QKD protocol makes use of a quantum channel connecting the honest parties, traditionally called Alice and Bob. Through this channel, they can send quantum systems as they see fit. In a real implementation, this quantum channel will usually be an optical fiber guiding photons, but our analysis will not assume any

[2] Actually, exactly one such scheme has been found yet.

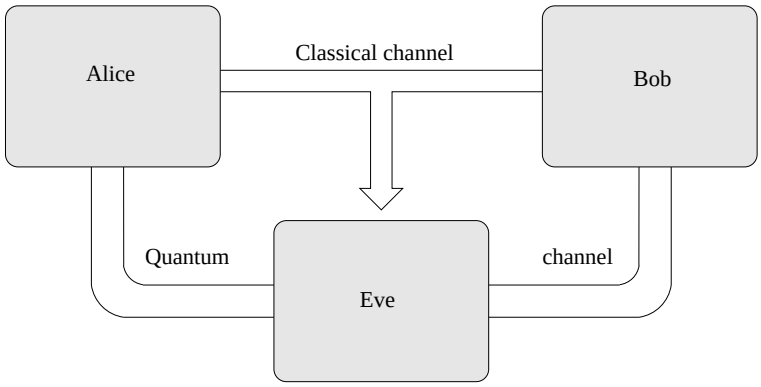

Figure 1.1: The protagonists in the description of a quantum key distribution protocol. The honest parties (Alice and Bob) are connected by an insecure quantum channel which is under the adversary's (Eve's) control. They also have access to a public but authenticated classical channel. Eve can listen to all communication via this channel, but she can neither alter messages sent by Alice and Bob nor create new messages with spurious sender.

special type of channel or quantum system.[3] However, we assume that the quantum channel is fully controlled by the adversary, personified by Eve. This means that in particular Alice and Bob assume that Eve replaces whatever quantum channel originally connected them by anything she wishes. Additionally to the quantum channel, Alice and Bob can use an authenticated public classical channel (see also Fig. 1.2). The adversary can listen to all communication performed via this channel, but she cannot insert new messages pretending to be Alice and Bob. Such an authenticated channel can be created by exploiting a short pre-shared key held by the honest parties [16].

BB84

The prime example of how a random key can be distributed by using an insecure quantum channel and an authenticated classical channel is provided by the so-called BB84 protocol, named after its inventors Bennett and Brassard, who proposed it in 1984 [17]. In the BB84 protocol (like in most other schemes), Alice chooses some (random) data and creates according quantum states which are sent to Bob, who performs a measurement on them, yielding again classical data. More concretely, Alice creates a random bit $x \in \{0, 1\}$, chooses at random the basis $+$ or $\times$, and prepares a quantum state

[3] An exception is the analysis of multi-photon events in attenuated laser pulses, which is a common implementation of QKD. This topic is treated in Ch. 6.

$|x_+\rangle = |x\rangle$ in case of the +-basis or $|x_\times\rangle = (|0\rangle + (-1)^x|1\rangle)/\sqrt{2}$ in case of the $\times$-basis. Thus one out of four possible states is sent through the quantum channel. Bob on his side also chooses one of the two bases + and $\times$ at random and performs a measurement on the arriving quantum state with respect to that basis. His measurement outcome is some random number $y \in \{0, 1\}$, possibly (hopefully) correlated with the bit x Alice chose. This procedure is repeated many times, creating a string of bits on Alice's and Bob's side. Assuming a perfect channel (in particular, no eavesdropping), Bob will get the quantum states sent by Alice undisturbed and we observe that a measurement in an incompatible basis results in a random bit y, uncorrelated with x. On the other hand, when Bob measures in the same basis as Alice chose, the outcome y will be equal to x. In the next step, Alice and Bob announce the bases they used for preparing/measuring each quantum state via the public channel. They determine the non-matching events and discard all corresponding bits. If there was no eavesdropping or other noise, they are now left with an identical, random bit string.

A simple strategy to break this protocol seems to be to simply copy the quantum states while they are traveling from Alice to Bob. In a classical world, there is nothing that the honest parties could do about that. In a quantum world however, things look different (and more promising for the security of key distribution).

The No-Cloning Theorem

It can be viewed as the fundamental concept, rendering quantum key distribution possible, that non-orthogonal quantum states cannot be copied (or cloned) perfectly. This is the statement of the No-Cloning Theorem [18], which can be proven in a simple way: Suppose there exists some unitary operation[4] U with $U|\psi_1\rangle|0\rangle = |\psi_1\rangle|\psi_1\rangle$ and $U|\psi_2\rangle|0\rangle = |\psi_2\rangle|\psi_2\rangle$, where $|\psi_{1,2}\rangle$ are two states which are to be cloned and $|0\rangle$ is some arbitrary input state. By taking the scalar products of the left- and right-hand sides of these equations, it follows that $\langle\psi_1|\psi_2\rangle = |\langle\psi_1|\psi_2\rangle|^2$, which implies that $|\psi_1\rangle$ and $|\psi_2\rangle$ are either identical (the trivial case) or orthogonal. For the BB84 protocol this implies that the adversary cannot perfectly copy the states sent from Alice to Bob, since they are taken from a set containing non-orthogonal states.

A more general strategy for Eve would be to perform a similar unitary operation on the input states $|\psi_i\rangle$ and some probe state $|0\rangle$, $U|\psi_i\rangle|0\rangle = |\psi_i'\rangle|\phi_i\rangle$ and then try to distinguish the output probe states $|\phi_\rangle$. By the same argument as above, one can show that $\langle\psi_1'|\psi_2'\rangle\langle\phi_1|\phi_2\rangle$ is constant, which means that whenever one wants to have the output states $|\phi_i\rangle$ to be more orthogonal, the input states $|\psi_i\rangle$ are getting more

[4] A unitary operation is the most general way to describe the evolution of a pure quantum state.

identical. This means that more distinguishable states probe states come at the cost of more disturbance of the input states. In this way the eavesdropper runs into the danger of being detected by Alice and Bob, who can monitor the error rate in their data.

The Role of Entanglement

In 1991, Ekert proposed a QKD protocol not based on sending classical information encoded into certain quantum systems, but rather on entanglement [19]. The idea of this protocol is to exploit the correlations one obtains when performing local measurements on entangled quantum states. In the original work, Alice and Bob aim at distributing the singlet state $|\psi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$ and then measuring a spin component along a direction chosen at random from a set of three possible directions. The distribution and measurements are repeated many times, and afterwards Alice and Bob reveal the measurement directions they chose. Those are selected such that there exists a common direction for Alice and Bob, yielding an (anti)correlated string of measurement outcomes and moreover, using the expectation value for the other directions, one can evaluate some CHSH inequality [20] (see also [21]). In this way, Alice and Bob can verify whether the quantum state upon which they performed their measurement was entangled. Ideally, they would find that they shared the state $|\psi^-\rangle$, which ensures that the key they draw from the measurement outcomes is perfectly secure, since the state $|\psi^-\rangle$ is pure and cannot be correlated with anything else.

Entanglement theory provides some powerful tools for the understanding of such "entanglement-based" protocols: Most notably, one can show that whenever Alice and Bob share a separable state, no secret key can created [22]. If the quantum systems are qubits, entanglement is even a sufficient condition, which means that entanglement is equivalent to the possibility of secret key extraction [23]. Surprisingly, it is not enough to only share the entanglement, Alice and Bob even need to be able to verify it from their measurement data, if they aim at creating the key using these measurements [24]. In general, the quantum system Alice and Bob distribute between them is some mixed, partly entangled state due to imperfect fibers and/or eavesdropping. Entanglement distillation [25] (see also [26] and references therein) can be used to turn a number of these non-maximally entangled state into fewer pure ones from which the key can be drawn by measurements. It has been shown [27, 28] that the distillation can equivalently be performed by a proper encoding and decoding of the quantum states using CSS codes [29, 30]. In this way, one can show that the BB84 and the Ekert protocol are actually equivalent.

An important consequence of this equivalence is that many QKD protocols can be

reformulated "entanglement-based", which enables us to utilize entanglement theory to quantify many features of the protocol, most notably its security.

Classical Post-processing

As far as we presented quantum key distribution by now, it aims at generating classical, correlated data by measuring a non-local quantum state. Except for the idealized case if which Alice and Bob are connected by a noiseless fiber and in particular, they are not eavesdropped, the classical data is perfectly correlated and unknown to any other party. In reality, the keys will never be perfect, and additional routines need to be expended. Altogether, these routines are called "classical post-processing", indicating that this is a collection of classical algorithms which are performed on classical data. After a "parameter estimation" step, in which the number of errors in the key is appraised, errors correction is performed, which leaves Alice and Bob with perfectly correlated data. Still, it cannot be safely used as a key, since the eavesdropper might have some information about it by attacking the quantum states which Alice and Bob measured. This knowledge can be made arbitrarily small by a procedure called "privacy amplification" [31, 32, 33], in which a certain function is applied to the data, outputting a shorter but more private key. Privacy amplification is of great importance, because it can be applied in any scenario in which the honest parties only share an error-free raw key.

1.3 Efficiency of Quantum Key Distribution

During the last decade, many experiments implementing a quantum key distribution protocol using photons were performed (for an overview of these experiments, refer to [34, 13] and references therein). The optical setup has already been miniaturized to fit into handy boxes, which are commercially sold [7, 6]. They implement the BB84 protocol, using photons as carriers for the quantum information. Photons are particularly suited because they can be easily and cheaply prepared using by lasers and optical elements, they travel through already available optical fibers (e.g. telecom fibers) or free space and they can be easily detected by common photo detectors. Still, this implementation suffers a limitation: The detection rate, i.e. the fraction of signals sent by Alice that get detected by Bob, is quite low, usually of the order 10^{-3}. There a two reasons for this: First, the signals suffer attenuation due to absorbance in the channel[5] or unintentional reflections in optical elements. Second, Bob's detector are not perfect, i.e., only a fraction

[5]Single-mode fibers at 1300 nm and 1550 nm have a loss rate of 0.35 dB/km and 0.2 dB/km, respectively. Note that 0.2 dB/km already results in 99% loss after 100 km.

of the arriving photons will get detected.[6] Also the repetition rate, which is the number of signals Alice's source can prepare and send to Bob per time slot, cannot be made arbitrarily large because of the down-time of Bob's detectors. This is the duration after a detection event in which due to technical limitations, no signal can be detected.

We have already mentioned that if the key generated by the QKD protocol is to be used in a one-time pad, it has to be as long as the message. This means that the user of a QKD device will typically be interested in large keys to be able to encrypt his or her message, which results in the demand for an *efficient* quantum key distribution scheme in the following sense: Given a number n' of quantum states that are sent through the quantum channel, what length ℓ (or rate ℓ/n') of the secret key can we expect the protocol to output? The answer to this question is the central result of this work.

1.4 Summary of the Main Results

This work focuses on the efficiency of the privacy amplification step, which is a building block of any quantum key distribution protocol in order to reduce the eavesdropper's knowledge about the key. More concretely, we are investigating privacy amplification by two-universal hashing [35, 36], in which Alice and Bob pick a certain random function and apply it to the classical data X and Y which they hold, respectively. The output of the hash function is in general much shorter than the input, but one can show that the privacy can be increased by any arbitrary amount [37]. The most important input to our work is a result derived by Renner and König in [37]: It provides the maximal possible length ℓ of the output (the secret key), fulfilling a certain security requirement quantified by a parameter ε, in terms of entropies S_2^ε, S_0^ε, and H_0^ε of the global quantum system describing the classical and quantum data held by all parties ρ_{XE}:

$$\ell = S_2^\varepsilon(\rho_{XE}) - S_0^\varepsilon(\rho_E) - H_0^\varepsilon(\mathbf{X}|\mathbf{Y}) + 2\log(\varepsilon). \tag{1.1}$$

This result is remarkable because it tells us how "much" privacy amplification one has to invest (i.e., how much one has to shrink the input data) in order to obtain a secret key of desired security ε. An important parameter is the size of the input data, which is usually given by a string of bits of length n (which we will call "block size"). This allows us to quantify the efficiency of the parameter estimation step by the secret rate $r = \ell/n$, which is also a function of the desired security (measured by ε). The block size appears implicitly in the above formula as the dimension of the density matrices ρ_{XE} and ρ_E and of the probability distribution $P_{\mathbf{X}|\mathbf{Y}}$.

[6] Typical InGaAs/InP detectors used for 1300 nm photons have a detection efficiency of 15%. For 1550 nm, it is about 5-10%.

Privacy Amplification with a Finite Number of Signals

We are motivated by the fact that as of today, the rate r of the privacy amplification step has only be calculated for the limiting case of $n \to \infty$ and $\varepsilon \to 0$, i.e., infinity block size and perfect security [1] (see also [38]). This is because of the complicated form of the formula for the key length ℓ, in particular since it is a function of the so-called "smooth Renyi entropies" S_α^ε, defined as

$$S_\alpha^\varepsilon(\rho) := \frac{1}{1-\alpha} \inf_{\sigma \in \mathcal{B}^\varepsilon(\rho)} \log \operatorname{tr} \sigma^\alpha, \tag{1.2}$$

where the infimum is taken over all density matrices σ (taken from arbitrarily large Hilbert spaces) which have a distance of at most ε to ρ. We we reformulate this definition to involve only an optimization over a finite set of numbers and eventually provide a simple and efficient algorithm that calculates smooth Renyi entropies of arbitrary density matrices in a time proportional to n. As a corollary, we are able to calculate the length of the secret key generated by privacy amplification for any given block size n and security parameter ε.

Obtainable Secret Key Rates

As an application, we investigate a special variant of the Six-State Protocol [4, 5] which we call "Tomographic Protocol" [2, 3], as its main peculiarity is that Alice and Bob can find out in the parameter estimation step which quantum state they shared. This is possible because the measurements performed in the entanglement-based version of the protocol allow for state tomography [39]. Since the knowledge of the quantum state describing Alice's, Bob's, and Eve's systems is all what is needed for the calculation of the key length ℓ, we can calculate it for this special kind of protocol. Alternatively, one can argue that our result is also applicable for *all* usual QKD protocols under the restriction that the eavesdropper only performs a certain symmetric attacks because in the tomographic protocol, Alice and Bob verify this symmetry and abort the protocol if it is broken.

We will show that the obtainable secret key rate for the Tomographic Protocol strongly depends on the block size n for $n \lesssim 10^4$. At $n \approx 10^4$, the rate reaches about 83% of the asymptotic value for $n \to \infty$ and approaches this value as n increases. From this result one can read off what reasonable block sizes one should choose in the privacy amplification protocol in order to obtain a desired efficiency of the protocol. Moreover, we investigate the dependence of the key rate on the security of the key, measured by the parameter ε. It has the intuitive interpretation that the key is *perfectly* secure except

with probability ε. There is currently no general understanding what is a reasonable range for this parameter. Our results show that, remarkably, up to $\varepsilon \approx 10^{-28}$, one can still generate a secret key for a block size of $n = 10^5$ and a common error rate. We also show that for increasing block size, the key rate becomes less dependent on the security parameter.

Our analysis of the Tomographic Protocol is also valid for arbitrary dimensions d which determine the alphabet size Alice and Bob use for the raw key and also the dimension of the quantum system which are employed. Without taking into account the fraction of the raw key that gets discarded in the sifting step, it turns out that larger dimensions are always favorable, in the sense that they yield the largest key rates. Considering also that for a d-dimensional variant, roughly a number of $n' = (d+1)n$ signals need to be sent in order to obtain a block size of n, we find that the "efficient key rate" ℓ/n' still increases for increasing dimension if the error rate in the sifted key is high. For low error rates however, we find the reverse result, namely that low dimensions yield optimal effective key rates. Interestingly, for each error rate there exists a particular dimension for which the effective key rate becomes maximal.

Multi-photon Events

Every implementation of a QKD protocol that is based on photons as a carrier of quantum information has to deal with multi-photon events which means that two or more photons with the same information encoded are sent through the channel. This enables the eavesdropper to split off one photon, store it, and measure it in the correct basis after Alice and Bob announced these in the sifting step (the so-called photon number-splitting attack). Fortunately, such an attack can be countered by privacy amplification, and we show that the obtainable key rate decreases in order to remove the additional knowledge Eve obtains due to the multi-photon events. It turns out that for a fraction η of single-photon pulses among all non-empty pulses, the key rate is ηr, with r denoting the key rate for the ideal case, i.e. $\eta = 1$. Similarly, we find that when Alice and Bob estimate an error rate Q from their measurement data, one needs to consider the larger value Q/η for the calculation of the key rate. Finally, we show how one can achieve a better estimate by including decoy pulses [40, 41] into the scheme.

1.5 Outline of This Work

This work aims at both providing concise introduction into the theory of quantum key distribution and presenting a novel and important result in the direction of re-

alistic implementations of these concepts. We are motivated by the fact that there exist some ambiguities and misconceptions, in particular about the equivalence between entanglement-based and prepare-and-measure schemes (cf. Sec. 3.2.3) and about the role of purifications (cf. Sec. 4.4). The first part of this thesis (Ch. 3) will deal with these issues while developing the theory of QKD. The main emphasis however lies in the analysis of privacy amplification by two-universal hashing (cf. Ch. 5). Our motivation comes from the lacking analysis of quantum key distribution with a finite number of signals, which plays an important role in experimental realizations. An application of this analysis is provided by the Tomographic Protocol (cf. Ch. 6).

After this introductory chapter, in Ch. 2 we introduce some basic information-theoretic and quantum mechanical concepts and notation. In Ch. 3, we present the general structure of quantum key distribution protocols. We will not focus on any particular protocol, but leave the introduction completely general: The protocol is divided into two parts, a quantum part (Sec. 3.2) and classical part (Sec. 3.3). In the quantum part, we explain how quantum states are distributed among Alice and Bob which yield the raw key upon measurement. Two different classes of protocols, classified based on how the quantum states are distributed, are presented in this section: Prepare-and-measure schemes in Sec. 3.2.1, and entanglement-based schemes in Sec. 3.2.2. In Sec. 3.2.3 we show that these two types are actually equivalent, in the sense that each protocol can be formulated in the other way. An implication of this result on the analysis of eavesdropping attacks is presented in Sec. 3.2.4. The classical part of the QKD protocol is again split up, treating the different sub-protocols that are carried out in this step: Measurements (Sec. 3.3.1), parameter estimation (Sec. 3.3.2), pre-processing (Sec. 3.3.3), and information reconciliation (Sec. 3.3.4).

In Ch. 4, we introduce the notion of security against the background of quantum key distribution. We start by reviewing the definition of security in an information-theoretic sense (Sec. 4.1) and introduce the concept of ε-security in Sec. 4.2. This finally enables us to present possible strategies of the eavesdropper in Sec. 4.3. In Sec. 4.4, we take a little excursion and study the possibility of creating purifications by physical processes. This topic is related to QKD, since purifications naturally appear when we describe eavesdropping attacks.

In Ch. 5, we derive the central result of this work. It treats the privacy amplification protocol, which is the final step in the classical part of any quantum key distribution protocol. The first section (Sec. 5.1) in this chapter provides an introduction, focusing in particular on classical privacy amplification. Sec. 5.2 reviews the main technical results derived in [37, 1], which are taken as a starting point of our own analysis. We

show that in order to derive the achievable key rate of QKD protocol, one needs to calculate so-called smooth Renyi entropies, which involves finding the extremum of a certain function over the space of density matrices. Sec. 5.3 is devoted to the analysis of these entropies and contains most of the technical results of this thesis. After presenting the definition and general properties in Sec. 5.3.1, we focus on three particular Renyi entropies that appear in the formula for the achievable key rate. In Sec. 5.3.2, we derive some important simplifications which ease the analysis significantly, allowing us to construct simple algorithms to compute the entropies for arbitrary density matrices in Sec. 5.3.3. Finally, in Sec. 5.3.4 we derive a special additivity property for the particular Renyi entropies.

We introduce a special quantum key distribution protocol, the "Tomographic Protocol" in Ch. 6 and apply our analysis of the privacy amplification procedure to this special case. The basic idea of the protocol is presented in Sec. 6.1 and we show how it fits into the framework developed in Ch. 3. Of particular importance is Sec. 6.2 in which we compute the obtainable key length for the Tomographic Protocol as a function of the parameters that Alice and Bob choose and which they measure in the parameter estimation step. We make use of the results found in Ch. 5. The dependence of the key rate on the various parameters (in particular the block size n and the security parameter ε) is shown in Sec. 6.3, for the idealized case of a single-photon realization of the protocol. This restriction is dropped in Sec. 6.4, where we take into account that inevitably multiple copies of the signal states are generated in any experiment.

We conclude in Ch. 7. In the appendix, we comment on the numerical methods employed to obtain the results of the preceding chapters.

Chapter 2

Preliminaries

This chapter is devoted to the introduction of the concepts and notation that is used throughout this thesis. In the forthcoming chapters, we will employ elements from both classical probability theory (such as probability distributions and entropies) and quantum information theory (such as density matrices and entanglement). These two fields are not completely separated from each other, rather, many concepts that were introduced in one area were carried over to other, mainly in the direction from the classical to the quantum world. The organization of this chapter is as follows: In Sec. 2.1, we will introduce the concepts and basic notation from classical probability theory that are important for the understanding of this thesis. Very specific definitions, that only appear in a certain section and which are of no importance for the global scope will be introduced in their respective section. Sec. 2.2 presents the notation and some special concepts from quantum information theory. Whenever possible, we will point out direct connections between classical and quantum concepts.

2.1 Classical World

In this section, we present some basic concepts of classical probability theory. We will often use the concept of a random variable. (Very) formally, it is defined as follows:

Definition 2.1.1. *Let (Ω, P) be a discrete probability space, i.e., Ω is some finite or countably infinity set, and P is a probability distribution on Ω, that is, some map $P : \Omega \to [0,1]$ with $\sum_{\omega\in\Omega} P(\omega) = 1$. A random variable X with range $\mathcal{X}$ is a function $X : \Omega \to \mathcal{X}$.*

We will always use the convention that a capital letter X denotes the random variable, a calligraphic letter $\mathcal{X}$ denote its range, that is, X takes values $x = X(\omega) \in \mathcal{X}$. The

probability $P(\omega)$ of this event ω will equivalently be denoted by $\mathrm{Prob}[X = x] \equiv P_X(x)$. Random variables that take vectors as values, e.g. $\mathcal{X} = \{0,1\}^n$, will be denoted by bold letters $\mathbf{X}$. The cardinality of a random variable, i.e., the size of its range, is given by $|\mathcal{X}|$. For two random variables X and Y, we denote the combined probability of X taking the value x and Y taking the value y by $P_{XY}(x,y)$, whereas we denote the corresponding conditional probability by $P_{X|Y}(x,y)$.

Next, we introduce a measure of the similarity of two probability distributions:

Definition 2.1.2. *Let P and Q be two probability distributions over the same range $\mathcal{X}$. Then the* variational distance *between P and Q is given by*

$$\|P - Q\| = \frac{1}{2} \sum_{x \in \mathcal{X}} |P(x) - Q(x)|. \tag{2.1}$$

This definition can easily be generalized to the case where P and Q are defined over different ranges by setting $P(x) = 0$ for all x which are not in the range of P, and similarly for Q. The variational distance is a metric on the set of probability distributions with range $\mathcal{X}$. As such, it fulfills the triangle inequality and $\|P - Q\| = 0$ if and only if P and Q are identical. The distance $\|P - Q\|$ of two probability distributions can be interpreted as the probability that two random variables X and X', described by a joint probability distribution $P_{XX'}$ with $P = P_X$ and $Q = P_{X'}$, take different values: $\|P - Q\| = \mathrm{Prob}[X \neq X']$.

We are often interested in quantifying the information that one random variable X contains about another one Y. This is done by the *mutual information* $I(X : Y) = H(X) - H(X|Y)$, where $H(X)$ is the usual Shannon entropy, $H(X) = -\sum_{x \in \mathcal{X}} P_X(x) \log P_X(x)$, and $H(X|Y) = H(X,Y) - H(Y)$ is the conditional Shannon entropy, with $H(X,Y) = -\sum_{x,y} P_{XY}(x,y) \log P_{XY}(x,y)$. The base of the logarithm is arbitrary; usually, one takes it to be two, which means that the entropy is measured in bits. Note that the mutual information is a symmetric quantity, i.e. $I(X : Y) = I(Y : X)$.

2.2 Quantum World

Quantum mechanical systems are described by positive semidefinite operators ρ with trace one. In the following, we use "positive operator" as a synonym for "positive semidefinite operator". We also adopt the usual habit and call ρ a *state* even if it is not pure. The set of all positive operators acting on a Hilbert space $\mathcal{H}$ will be denoted by $\mathcal{P}(\mathcal{H})$, and the set of all such operators having trace one by $\mathcal{B}(\mathcal{H}) = \{\sigma \in \mathcal{P}(\mathcal{H}) :$

$\operatorname{tr}\sigma = 1\}$. For quantum states, we can also introduce a distance measure, similarly to the classical case of Def. 2.1.2:

Definition 2.2.1. *Let $\rho, \sigma \in \mathcal{B}(\mathcal{H})$ be two density operators. Then the* trace distance *between ρ and σ is given by*

$$\|\rho - \sigma\| = \frac{1}{2} \operatorname{tr} |\rho - \sigma|, \tag{2.2}$$

where $|A| = \sqrt{AA^\dagger}$.

Like the variational distance, the trace distance defines a metric on $\mathcal{B}(\mathcal{H})$. Measurements on quantum states $\rho \in \mathcal{B}(\mathcal{H})$ are defined by *positive operator valued measurements* (POVMs) [42], which are a set $\mathcal{M} = \{M_i\} \subset \mathcal{P}(\mathcal{H})$ of positive operators summing up to the identity, $\sum_i M_i = \mathbb{1}_{\mathcal{H}}$. As the measurement outcome i is obtained with probability $\operatorname{tr}(\rho M_i)$, we can construct a probability distribution $P^\rho_{\mathcal{M}}$ describing the statistics of all possible measurement outcomes with $P^\rho_{\mathcal{M}}(i) = \operatorname{tr}(\rho M_i)$. One can show that the variational distance is a lower bound on the trace distance of two quantum states when the same POVM $\mathcal{M}$ is applied, $\|\rho - \sigma\| \geq \|P^\rho_{\mathcal{M}} - P^\sigma_{\mathcal{M}}\|$. Equality is obtained for "classical states" ρ_X, which are the quantum states representing a classical random variable X with range $\mathcal{X}$ and associated probability distribution P_X, i.e. $\rho_X = \sum_{x\in\mathcal{X}} P_X(x)|x\rangle\langle x| \in \mathcal{B}(\mathcal{H})$, where $\mathcal{H}$ is some $|\mathcal{X}|$-dimensional Hilbert space with basis $\{|x\rangle\}_{x\in\mathcal{X}}$. For those classical states, we have $\|\rho_X - \rho_{X'}\| = \|P_X - P_{X'}\|$.

We will often encounter the situation where classical data, described by some random variable X, is correlated with a quantum system. The quantum state describing the combined system is called "cq-state" ("classical-quantum-state"):

Definition 2.2.2. *Let X be a random variable with range $\mathcal{X}$ and probability distribution P_X, and let $\rho^x_E \in \mathcal{B}(\mathcal{H}_E)$ be a quantum state that depends on the value x of X. Then the joint state of the system is given by the so-called cq-state*

$$\rho_{XE} = \sum_{x\in\mathcal{X}} P_X(x)|x\rangle\langle x| \otimes \rho^x_E, \tag{2.3}$$

with $\rho_{XE} \in \mathcal{B}(\mathcal{H} \otimes \mathcal{H}_E)$, and $\mathcal{H}$ is some $|\mathcal{X}|$-dimensional Hilbert space with basis $\{|x\rangle\}_{x\in\mathcal{X}}$.

This definition is straightforwardly generalized to, say, ccq-states, where a quantum state $\rho^{xx'}_E$ is correlated with two random variables X and X'. We will also encounter cq-states $\rho_{\mathbf{X}E}$ where the classical part is described by a random variable $\mathbf{X}$ taking vectors as values, and the quantum part $\rho^{\mathbf{x}}_E$ may depend on all values $\mathbf{x}$. States of this form naturally appear in the analysis of the security of QKD, where classical data (the key) is correlated with a quantum system held by the eavesdropper.

The evolution of a quantum system is most generally described by a completely positive (CP) map $\Lambda : P(\mathcal{H}) \to P(\mathcal{H})$. Such a map has the property that any extension to larger Hilbert spaces maps positive matrices to positive matrices, i.e. $[\mathrm{id}_{\mathcal{H}'} \otimes \Lambda](\rho) \geq 0$, for any $\mathcal{H}'$ and $\rho \in \mathcal{P}(\mathcal{H}' \otimes \mathcal{H})$. For all deterministic processes, this map is additionally trace-preserving, $\mathrm{tr}\,\Lambda(\rho) = \mathrm{tr}\,\rho$. Any CP map can be written in the so-called Kraus decomposition [43] or alternatively, and more convenient for our analysis, in the following way [44]: $\Lambda(\rho) = \mathrm{tr}_A(U\rho \otimes |0\rangle\langle 0|_B U^\dagger)$. That is, some auxiliary system B in some (without loss of generality) pure state $|0\rangle$ is appended to ρ, then some unitary operation U is performed on the combined system, and the part A (which may be different from B) is traced out.

Chapter 3

Quantum Key Distribution Protocols

3.1 Composition of a QKD Protocol

The goal of all quantum key distribution protocols is to provide the honest parties, Alice and Bob with random, correlated, and private classical data, the key. To achieve this, they have a quantum channel at their disposal, which is however to be assumed completely under the control of the adversary, Eve. This means that whatever quantum state Alice or Bob send through the channel, the output can be completely arbitrary, the only restriction is consistency with quantum mechanics.[1] In addition to the quantum channel, the honest parties can make use of a public, classical channel, which is assumed to be authentic, that is, everybody (in particular Eve) can listen to all communication over the channel, but she cannot alter or forge messages.[2]

In the most general sense, the secret key is generated from classically created random data (e.g. coin-flipping) and/or outcomes of measurements of quantum states. Every QKD protocol can be divided into two parts: A quantum part, in which quantum mechanical systems are distributed between Alice and Bob and upon which some measurements are carried out, and a classical part, in which the classical data generated in the first part is transformed into a secret key by means of so-called "post-processing".[3] Post-processing is a collection of purely classical algorithms such as error correction and

[1] In Ch. 4.3, we give a detailed classification of possible eavesdropping attacks and the resulting structure of the quantum states.

[2] One can show that authenticity can be created from some short secret key that Alice and Bob share beforehand [16].

[3] We will only consider one-way classical post-processing, which will be described in detail below.

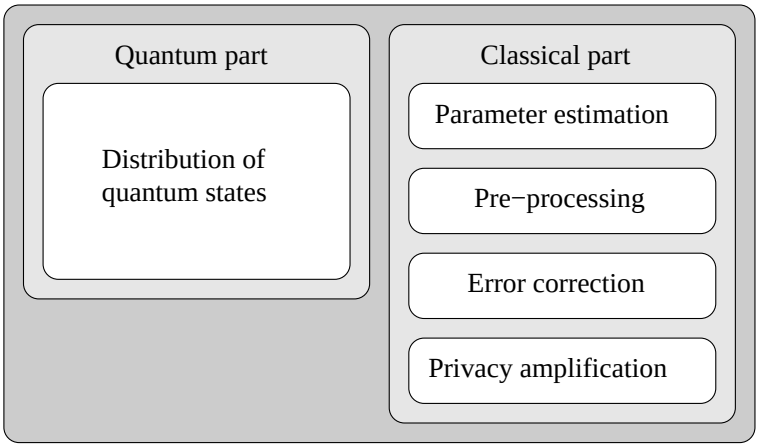

Figure 3.1: Composition of a quantum key distribution protocol.

privacy amplification [31]. The composition is visualized in Fig. 3.1.

The quantum part of the QKD protocol will be treated in Sec. 3.2. We will see that one has to discriminate two different ways of how the quantum states are distributed between Alice and Bob and how the honest parties obtain the classical data. One class of protocols, presented in detail in Sec. 3.2.1, is called "prepare-and-measure" schemes. Common members of this class are the BB84 [17], the Six-State [4, 5], and the B92 Protocol [45]. All these protocols have in common that Alice at first picks some random data (usually a bit, for instance by tossing a coin), prepares some corresponding quantum state and sends it to Bob. Bob on his side performs a measurement on the quantum state that he receives. In general, this state is different from what Alice sent, because of noise in the quantum channel and/or Eve's tampering with the signal. Nonetheless, Bob's measurement outcome will in general be to some extent correlated with Alice's data. The main twist of the prepare-and-measure protocols is that for each bit value that Alice chooses, she picks one state from a *set* of possible quantum states. This redundancy guarantees that the information encoded in the quantum state (the data Alice that chose) cannot be easily retrieved, as this would require the quantum states to be distinguishable, a property which is ruled out by a proper choice of the set of possible states.

The second class of protocols into which we can divide the distribution part is called "entanglement-based schemes". These protocols, which are presented in Sec. 3.2.2 are based on the idea of distributing a maximally entangled quantum state between Alice and Bob. This state might for instance be the Bell state $|\phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$, for which Alice and Bob obtain perfectly correlated outcomes when measuring their

respective particles in the x or z-basis (the eigenbasis of the Pauli operators σ_x and σ_z, respectively). Moreover, the outcomes will be completely random and since the state $|\phi^+\rangle$ is pure, it cannot be correlated with anything else, implying the privacy of the key.

At first glance, the prepare-and-measure and entanglement-based scheme look very different, mostly because for the former class, no entanglement is present during the whole distribution step, although it is an essential ingredient for the latter class. In Sec. 3.2.3 however we will show that these two schemes are equivalent under the assumption that only single copies of the quantum states are sent in the prepare-and-measure version. A detailed analysis of the case where we cannot rely on this equivalence is deferred to Sec. 6.4; until then, we always use the idealized single-copy assumption.

After Bob (and also Alice for an entanglement-based protocol) measured his quantum state and both parties are left with classical data, the classical part of the QKD protocol commences (see Sec. 3.3). It consists of a so-called "parameter estimation" phase (see Sec 3.3.2), in which Alice and Bob try to deduce some features of the quantum state the Bob received (or both shared, for the entanglement-based version), and thus also about the eavesdropping attack. In this step they might also decide to cancel the protocol completely, if it turns out that their data contains too many errors, possibly due to a strong interaction of the eavesdropper. At this point, the classical data may also be processed in order to possibly enhance the robustness of the protocol. This "pre-processing" step, covered in Sec. 3.3.3 is only listed here for completeness but will be ignored in this work. The next step in the post-processing is the "information reconciliation" step described in Sec. 3.3.4. The goal of this step is to make Alice and Bob's data equal. It usually consists of a sifting step, in which Alice and Bob discard all events in which Bob used an incompatible measurement basis (for the details about encodings and measurements, see Sec. 3.2) and/or an error correction step in which all remaining errors are corrected to leave Alice and Bob with equal data. The last step of the classical part, and most importantly for our analysis, is called "privacy amplification" (see Sec. 3.3.5 and Ch. 5). The aim of this step is to remove all residual knowledge which the eavesdropper might have about the key. This is done by applying a so-called hash function to some large block of the key, outputting a shorter key which can be made more private by a proper choice of the function.

In the remaining chapter, we devote a subsection to each step of the QKD protocol.

3.2 Quantum Part/Distribution of Quantum States

In this section we present a general description of the distribution step of a QKD protocol. The next two subsections treat the prepare-and-measure variant (see Sec. 3.2.1) and the entanglement-based variant (see Sec. 3.2.2) separately. Their equivalence is proven in the third subsection, Sec. 3.2.3.

Although most QKD implementations only encode bits into quantum states, we will generalize our analysis to arbitrary alphabets of size d. Note that in the end, after the classical post-processing (cf. Sec. 3.3), we will always end up with a key that only consists of *bits*. Let $\{|\phi_x^j\rangle\}$, with $x = 0, 1, \ldots, d-1$ and $j = 1, 2, \ldots, r$ be a family a pure states such that for each j, the $|\phi_x^j\rangle$ are linearly independent. Furthermore, let $\mathcal{M}^j = \{M_x^j, M_?^j\}$ denote some set of POVMs such that $\mathcal{M}^j$ unambiguously discriminates the $|\phi_x^j\rangle$, i.e. $\langle\phi_x^j|M_y^j|\phi_x^j\rangle \sim \delta_{xy}$ for all j. Again for completeness, we include the possibility of inconclusive outcomes, represented by $M_?^j$, in the case where the $|\phi_x^j\rangle$ are not mutually orthogonal. However, the inconclusive outcomes will not play any special role in our analysis. We say that the index j labels the *encoding* of the $d$it x and the states $\{|\phi_x^j\rangle\}$ are called *signal states*. Finally, let P_J and P_K be some probability distributions on the set of encodings $\{1, 2, \ldots, r\}$. The probability distribution P_J will be used by Alice to choose an encoding for each signal, and likewise P_K will be used by Bob to choose a POVM $\mathcal{M}^k$. For the sake of completeness, define a third probability distribution P_X on $\{0, 1, \ldots, d-1\}$ for the choice of the $d$it x that is to be encoded. Although in our analysis we will only consider uniform distributions, i.e. $P_X(x) = 1/d$ for all x, for the sake of generality and in order to unify the notation, we allow also for arbitrary probability distributions P_X.

The idea behind the introduction of additional redundancy by having r different encodings (that is, dr different quantum states are used to encode only one $d$it) is that it becomes more difficult to identify a certain state taken from a set of different states as the size of the set increases. However, if the encoding j is known, $\mathcal{M}^j$ is constructed such that this task is feasible.[4]

3.2.1 Prepare-and-measure Schemes

Most QKD protocols fall into this category, for instance the BB84 [17], Six-State [4], and the B92 [45] Protocol. They all have in common that Alice encodes some classical

[4] In a real experiment, even the implementation of a measurement that discriminates only two (non-orthogonal) states might not be highly efficient [46]. However, in our analysis we will ignore such practical problems.

information (e.g., a bit) into a set of quantum states that are sent to Bob. Bob on his part will measure the quantum system and obtains some classical measurement result. We will now discuss the details of this procedure.

> Alice chooses n numbers $x_i \in \{0, 1, \ldots, d-1\}$ and n numbers $j_i \in \{1, 2, \ldots, r\}$ according to P_X and P_J, respectively, prepares the state $|\phi_{x_1}^{j_1}\rangle \otimes \cdots \otimes |\phi_{x_n}^{j_n}\rangle$, and sends it to Bob.

Since Alice keeps her choice of all $d$its and encodings in mind, the combined state describing her classical data and the prepared quantum system is given by

$$\rho_{JAB}^n = \left[\sum_{x=0}^{d-1}\sum_{j=1}^{r} P_J(j)P_X(x)|j\rangle\langle j| \otimes |x\rangle\langle x| \otimes |\phi_x^j\rangle\langle\phi_x^j|\right]^{\otimes n} \tag{3.1}$$

Consider for a moment that Alice and Bob are connected by a noiseless channel (i.e., there is no eavesdropper), thus Bob receives $\left[\sum_{x=0}^{d-1}\sum_{j=1}^{r} P_J(j)P_X(x)|\phi_x^j\rangle\langle\phi_x^j|\right]^{\otimes n}$ undisturbed. In order to "decode" the x's, he chooses n numbers $k_i \in \{1, 2, \ldots, r\}$ according to P_K and performs the POVM $\mathcal{M}^{k_1} \otimes \cdots \otimes \mathcal{M}^{k_n}$ on this state obtaining the result $\mathbf{y} \in \{0, 1, \ldots, d-1, ?\}^{\times n}$. That is, for each signal arriving, he picks a encoding k_i according to the probability distribution P_K and performs a measurement, given by $\mathcal{M}^{k_i}$ on the quantum state. He adds the outcome y_i to a list which forms his "raw key". Again, there might be inconclusive measurement outcomes depending on the specific protocol.[5]

To fill this sketch of a protocol with life, let us consider a common version of the BB84 protocol: In the BB84 protocol, we have $r = 2$ different encodings for a bit ($d = 2$), namely the eigenstates of the σ_z and σ_x Pauli operators, i.e. $|\phi_0^1\rangle = |0\rangle$, $|\phi_1^1\rangle = |1\rangle$, $|\phi_0^2\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, and $|\phi_1^2\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$. Since for each encoding, the two different "codewords" are orthogonal, the two POVMs employed by Bob are given by $\mathcal{M}^j = \{|\phi_0^j\rangle\langle\phi_0^j|, |\phi_1^j\rangle\langle\phi_1^j|\}$, $j = 1, 2$, without inconclusive outcomes. The bits 0 and 1 will be encoded with equal probability, therefore we choose P_X to be a flat distribution. On the other hand, it turns out that it is preferable that Alice almost always uses the same encoding to increase the efficiency of the protocol [47]. Bob on his side takes P_K to be flat again, that is, he randomly chooses $\mathcal{M}^1$ or $\mathcal{M}^2$ for his measurement.

[5] Note that we do not take into account any inconclusive outcomes which originate from experimental imperfections such as dark counts, stray light, detector dead time, or losses.

3.2.2 Entanglement-based Schemes

In contrast to the class of protocols described in the previous section, entanglement-based protocols aim at distributing a maximally entangled state between Alice and Bob. The correlation inherent in this state and its detachment from the environment enable Alice and Bob to create a common secret key. This is achieved in the following way: Suppose Alice and Bob share the maximally entangled state $|\phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ and measure their respective system in the computational basis. They will both obtain either the bit 0 or 1 with probability 1/2. Moreover, since $|\phi^+\rangle$ is pure, it can easily be shown that no other (classical or quantum) system can contain any information about this bit [28, 27]. This leaves one with the problem of how to distribute the state $|\phi^+\rangle$, since it has to be prepared locally and therefore it has to pass through the quantum channel which potentially can be attacked by the eavesdropper. The original idea [19, 27] was to let Alice prepare n copies of $|\phi^+\rangle$, send the second half to Bob, resulting in a state $\rho^{\otimes n}$ of n non-maximally entangled states ρ. Instead of performing classical privacy amplification on their data, Alice and Bob now run an entanglement distillation protocol [25, 48, 49] on all these copies and obtain a number $m < n$ of maximally entangled states. Unfortunately, such a protocol is hard to implement because it requires quantum memory. However, it has been shown [28] that the entanglement distillation can equivalently be performed using certain quantum error-correcting codes (CSS codes [29, 30]). We do not need to go into the details here, since we only consider classical privacy amplification.

The distribution of quantum states using an entanglement-based protocols works as follows:

> Alice chooses n numbers $j_i \in \{1, 2, \ldots, r\}$ according to P_J, prepares the state $\sum_{x=0}^{d-1} \sqrt{P_X(x)}|x\rangle|\phi_x^{j_1}\rangle \otimes \cdots \otimes \sum_{x=0}^{d-1} \sqrt{P_X(x)}|x\rangle|\phi_x^{j_n}\rangle$, and sends the second half (the $|\phi_x^{j_i}\rangle$ part for all i) to Bob.

The only difference to the prepare-and-measure scheme is that each signal is in a coherent superposition of $|x\rangle|\phi_x^j\rangle$ for all x. The encoding j of each signal is chosen classically, thus the combined state describing Alice's data and the quantum systems sent to Bob is given by

$$\rho_{JAB} = \left[\sum_{j=1}^{r} P_J(j)|j\rangle\langle j| \otimes \left(\sum_{x=0}^{d-1} \sqrt{P_X(x)}|x\rangle|\phi_x^j\rangle\right)\left(\sum_{x=0}^{d-1} \sqrt{P_X(x)}\langle x|\langle\phi_x^j|\right)\right]^{\otimes n} \tag{3.2}$$

The "encoding" is performed by Alice measuring each of her quantum states in the computational basis $\{|x\rangle\}$, resulting in a measurement outcome x_i with probability

$P_X(x_i)$. The decoding step works in the same way as for the prepare-and-measure version: Bob chooses n numbers $k_i \in \{1, 2, \ldots, r\}$ according to P_K and performs the POVM $\mathcal{M}^{k_1} \otimes \cdots \otimes \mathcal{M}^{k_n}$ on this state. He obtains the result $\mathbf{y} \in \{0, 1, \ldots, d-1, ?\}^{\times n}$, which again might contain inconclusive measurement outcomes.

3.2.3 Equivalence of Entanglement-based and Prepare-and-measure Schemes

It is easy to see that the schemes presented in previous two subsections are equivalent, that is, they provide Alice and Bob with the same correlations and are indistinguishable.

We see that any prepare-and-measure scheme can be recast entanglement-based by introducing a purifying system in the state (3.1). We say that a quantum state $\rho \in \mathcal{B}(\mathcal{H})$ is purified [50, 51] by a state $\sigma \in \mathcal{B}(\mathcal{H}_{\text{aux}})$ (or by a system $\mathcal{H}_{\text{aux}}$) if there exists a pure state $|\Psi\rangle \in \mathcal{H} \otimes \mathcal{H}_{\text{aux}}$ such that ρ and σ are both marginal states of this pure state, i.e. $\rho = \text{tr}_{\mathcal{H}_{\text{aux}}} |\Psi\rangle\langle\Psi|$ and $\sigma = \text{tr}_{\mathcal{H}} |\Psi\rangle\langle\Psi|$ (see also Sec. 4.4). Here, we choose the system R such that it purifies the mixture $\sum_{x=0}^{d-1} P_X(x)|x\rangle\langle x| \otimes |\phi_x^j\rangle\langle\phi_x^j|$ and is neither controlled by Alice and Bob nor by Eve:

$$\rho_{JABR} = \left[\sum_{j=1}^{r} P_J(j)|j\rangle\langle j| \otimes \left(\sum_{x=0}^{d-1} \sqrt{P_X(x)}|x\rangle|\phi_x^j\rangle|x\rangle_R \right) \left(\sum_{x=0}^{d-1} \sqrt{P_X(x)}\langle x|\langle\phi_x^j|\langle x|_R \right) \right]^{\otimes n} \tag{3.3}$$

This state is equal to the state (3.2) from Alice's, Bob's, and Eve's point of view, thus the prepare-and-measure scheme is contained in the class of entanglement-based schemes.

Conversely, consider the state (3.2) describing an entanglement-based scheme. If Alice measures the system A in the $\{|x\rangle\}$-basis without revealing the result, we arrive at the state (3.1). This shows that any entanglement-based scheme is also contained in the class of prepare-and-measure schemes, which proofs their equivalence.

In certain cases however one would like to go further and describe a protocol as being entirely based on distributing entangled states. To this end, define the so-called encoding operators

$$A^j = \sum_{x=0}^{d-1} |\phi_x^j\rangle\langle x|. \tag{3.4}$$

For simplicity, let us assume that the $|\phi_x^j\rangle$ are mutually orthogonal for each encoding j, implying that the A^j are unitary.

Using these operators, we see that the signal states can be prepared by the action

of A^j:

$$\sum_{x=0}^{d-1} \sqrt{P_X(x)}|x\rangle|\phi_x^j\rangle = \mathbb{1} \otimes A^j \sum_{x=0}^{d-1} \sqrt{P_X(x)}|x\rangle|x\rangle =: \mathbb{1} \otimes A^j|\tilde{\phi}^+\rangle, \tag{3.5}$$

where we have defined $|\tilde{\phi}^+\rangle = \sum_{x=0}^{d-1} \sqrt{P_X(x)}|x\rangle|x\rangle$, which, for $P_X(x) = 1/d$ is equal to the maximally entangled state in d dimensions. In this way, the preparation (encoding) can be absorbed into the operator A^j. Suppose that we can find some operators $\tilde{A}^j$ such that $\mathbb{1} \otimes A^j|\tilde{\phi}^+\rangle = \tilde{A}^j \otimes \mathbb{1}|\tilde{\phi}^+\rangle$ holds. This means that Alice can perform a modified encoding operation $\tilde{A}^j$ on her half of the state $|\tilde{\phi}^+\rangle$ instead of applying A^j to Bob's part. In particular, the distributed state is now simply $|\tilde{\phi}^+\rangle$, independent of the encoding and thus it is the same for each protocol. It is easy to see that we can always [1] choose $\tilde{A}^j := A^{j\mathrm{T}}$ if A^j does not map the states $|x\rangle$ from $\mathbb{C}^d$ into a higher-dimensional Hilbert space. The important case where this happens is when Alice sends (probably unintentionally) more than one copy of the signal states, that is $A^j = \sum_{x=0}^{d-1} |\phi_x^j\rangle^{\otimes n}\langle x|$ for some $n > 1$. The analysis of these multi-photon events is more involved and will be discussed for the example of the "Tomographic Protocol" in Sec. 6.4. Until then, we will stick to case where $n = 1$ and where the encoding maps states from $\mathbb{C}^d$ to $\mathbb{C}^{d'}$ with $d' \leq d$.

3.2.4 Eavesdropping

An important consequence of this modified encoding approach in which for any protocol only one particular state $|\tilde{\phi}^+\rangle = \sum_{x=0}^{d-1} \sqrt{P_X(x)}|x\rangle|x\rangle$ has to be distributed is the simple analysis of the eavesdropping attack: We will see in Ch. 4 that Eve's attack can be fully described by a unitary operation that she applies on the quantum state sent from Alice to Bob and a "probe system" that she attaches to it. This means that the total state shared between Alice, Bob, and Eve right after her attack is given by

$$|\Psi\rangle_{ABE} = \mathbb{1}_A \otimes U_{BE}|\tilde{\phi}^+\rangle_{AB}|0\rangle_E, \tag{3.6}$$

where U_{BE} is the unitary operation the eavesdropper applies on the system B and her probe system E, which is in some arbitrary initial state $|0\rangle$. Eve keeps a subsystem of BE and forwards the remaining part to Bob. Without loss of generality, we can assume that she keeps her original probe E and sends B on to Bob, because the unitary operation U_{BE} can in particular contain a swapping operation of arbitrary subspaces. This means that Bob will receive the state $\rho_B = \mathrm{tr}_{AE}(U_{BE}|\tilde{\phi}^+\rangle\langle\tilde{\phi}^+| \otimes |0\rangle\langle 0|_E U_{BE}^\dagger)$, whereas the knowledge of the eavesdropper about the key at this point is also fully characterized by the quantum state held by her, which is

$$\rho_E = \mathrm{tr}_{AB}(U_{BE}|\tilde{\phi}^+\rangle\langle\tilde{\phi}^+| \otimes |0\rangle\langle 0|_E U_{BE}^\dagger). \tag{3.7}$$

This state is independent of the particular encoding and thus, independent of the specific protocol.

To conclude this section, we have shown that the entanglement-based and prepare-and-measure type of the preparation of the quantum system are equivalent, in the sense that the total states (3.1) describing the distributed quantum states together with any classical data the honest parties hold are the same from Alice's, Bob's, and Eve's point of view. Additionally, for a restricted set of encodings, where the signals states live in a Hilbert space of dimension not greater than d, where d is defined by the state $|\tilde{\phi}^+\rangle = \sum_{x=0}^{d-1} \sqrt{P_X(x)}|x\rangle|x\rangle$, the quantum state held by the adversary is independent of the encoding, as shown by (3.7). In particular, this holds for the BB84, B92, and Ekert protocol. This is important because one step of the security analysis, classifying the state held by the adversary, is therefore common for all such protocols (cf. Sec.4.3).

3.3 Classical Part

In this part of the key distribution the classical strings which Alice and Bob obtain upon measuring the quantum state distributed in the quantum part, will be made equal and secure. The states given in Eq. (3.1) (for the prepare-and-measure scheme) and Eq. (3.2) (for the entanglement-based scheme) describe the situation before the signal states reach Bob. Since they pass through a quantum channel about which Alice and Bob have only partial knowledge (they might know some of its basic properties such as attenuation in the absence of an eavesdropper), the state shared by Alice and Bob after the transmission might be of a complicated structure. In general, it is only possible for Alice and Bob to obtain some partial knowledge about this state, because there are many states that are compatible with a given measurement statistics. In Ch. 6 we present a special protocol in which Alice and Bob can in principle exactly infer which quantum state they shared prior to the measurement. In this section, we present the classical part of the protocol by only assuming that Alice and Bob share some n-partite state ρ_{AB}^n.

3.3.1 Measurement

We already indicated in Sec. 3.2.1 and 3.2.2 how Bob decodes the $d$it string chosen by Alice: For the ith signal, he chooses an encoding k_i according to the probability distribution P_K, applies the POVM $\mathcal{M}^{k_i}$ and obtains the outcome $y_i \in \{0, 1, \ldots, d-1, ?\}$, where "?" denotes an inconclusive answer.

Denote by $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ the $d$it string Alice has chosen in the preparation step

(or which she has measured in an entanglement-based scheme). Since each x_i was chosen according to the probability distribution P_X, we can introduce a random variable $\mathbf{X}$ with range of all n $d$it-strings $\{0,1,\ldots,d-1\}^{\times n}$ and an associated probability distribution $P_{\mathbf{X}}$ with $P_{\mathbf{X}}(\mathbf{x}) = P_X(x_1)\cdots P_X(x_n)$. Likewise, let $\mathbf{y}$ with $y_i \in \{0,1,\ldots,d-1,?\}$ denote the n outcomes of Bob's application of $\mathcal{M}^k$. Finally, we can introduce random variables $\mathbf{Y}$, $\mathbf{J}$, and $\mathbf{K}$ for Bob's measurement outcomes and Alice's and Bob's choice of the encoding, respectively.

3.3.2 Parameter Estimation

All data Alice and Bob gathered is described by the random variables $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{J}$, and $\mathbf{K}$. By comparing (part of) this data, Alice and Bob can try to infer some characteristics of the eavesdropping attack. In particular, they have to be able to conclude whether it is at least *possible* to create a secret key from those classical data. An important quantity that can be estimated in this step is the quantum bit error rate (QBER), which is the fraction of signals i for which Alice and Bob chose the same encoding, i.e. $j_i = k_i$, but where Bob got a measurement outcome different from what Alice prepared, i.e. $x_i \neq y_i$. The calculation of secret key rates (cf. Sec. 6.2) also yields results that are dependent on the parameters Alice and Bob estimate in this step.

3.3.3 Pre-processing

It might be advantageous for Alice and Bob not to generate a secret key directly from $\mathbf{X}$ and $\mathbf{Y}$, but to have Alice calculate two new random variables $\mathbf{U}$ and $\mathbf{V}$ obtained by some conditional probability distributions $P_{\mathbf{U}|\mathbf{X}}$ and $P_{\mathbf{V}|\mathbf{U}}$. Alice keeps $\mathbf{U}$ and sends $\mathbf{V}$ to Bob. (Bob will calculate a guess for $\mathbf{U}$ using $\mathbf{V}$, the error correction information, and his data $\mathbf{Y}$, see below.) In our analysis, we will neglect this step, thus choosing $\mathbf{U} \equiv \mathbf{X}$ and $\mathbf{V}$ to be trivial (uniform). Nevertheless, it turned out that the performance of a protocol can actually be improved by choosing $\mathbf{U}$ to be a "noisy" version of $\mathbf{X}$. On the other hand, the variable $\mathbf{V}$ does not seem to play any role [1].

3.3.4 Information Reconciliation

Up to this point, the classical strings $\mathbf{x}$ and $\mathbf{y}$ Alice and Bob hold are not identical, because in general, they originate from measuring some quantum state which is disturbed due to Eve's interaction. In the information reconciliation step, the strings are made equal. To achieve this, Alice sends error correction information — quantified by another random variable $\mathbf{W}$ — to Bob, who uses $\mathbf{Y}$, $\mathbf{V}$, $\mathbf{W}$ to calculate a guess for $\mathbf{U}$. (In our

case, $\mathbf{U} \equiv \mathbf{X}$, since we neglect the pre-processing.) We assume that after this step, Alice and Bob hold the same random variable $\mathbf{U}$, at least with probability $1-\varepsilon$. Usually, in this step $\mathbf{W}$ contains at least the information about $\mathbf{J}$, that is, which encodings Alice used for each signal. A simple so-called "sifting" method for Bob is then to simply discard all signals for which he used a different encoding. There exist also more sophisticated sifting strategies, for instance in the SARG protocol [52], in which more signals get discarded, but Eve also has less knowledge about the remaining ones. In addition to the sifting, also all other errors that were introduced by Eve or the noisy channel are corrected.

Since we neglect the pre-processing step, we will from now on denote the $d$it string shared by Alice and Bob (from which the key will be created) by $\mathbf{X}$ and the error correction information sent by Alice by $\mathbf{W}$, which is also known by Eve, because it is sent over the public channel. It has been shown in [53] that the minimal amount of information W which allows Bob to guess X with probability of at least $1-\varepsilon$, knowing only Y and W, is given by $H_0^\varepsilon(X|Y)$. This quantity, called conditional smooth Renyi entropy, will be discussed in detail in Sec. 5.3 (see also Sec. 5.2).

3.3.5 Privacy Amplification.

Although at this point Alice and Bob share a common $d$it string $\mathbf{x}$, it is not very reasonable to directly use it as a secret key, since the eavesdropper might have too much information about it. In the privacy amplification step, Alice and Bob shrink the length of the key $\mathbf{x}$ and at the same time reduce the information that Eve might have about it, thereby generating a secret key. Since privacy amplification is an important step which will be the starting point of the analysis of of secret key rates, we review this sub-protocol in more detail in Sec. 5.2. We also review the security analysis of privacy amplification and present an expression for an achievable secret key length, as found in [1].

Chapter 4

Security of Quantum Key Distribution

In this section, we present methods that enable us to quantify the secrecy of a key created with a QKD protocol described in the previous chapter. The main peculiarity is that we want to quantify the knowledge that an adversary holding a quantum system might have about a classical key. For a better understanding of this issue, we start with a brief review of a classical security definition in Sec. 4.1 and show that generalizing the classical definition for the quantum case does not yield a satisfactory security definition. In the main section of this chapter, Sec. 4.2, we will see how this issue can be settled by a completely different approach, involving the distance of quantum states.

4.1 Security in the Classical World

Before focusing on the problem of how to define the security of a QKD protocol, we take one step back and consider the classical case. Suppose Alice and Bob run some (classical) key distribution protocol such that, at the end, they hold two strings S_A and S_B, respectively. We now ask ourselves what qualifies S_A and S_B as secret key? Certainly, the answer to that question depends on "how secure" we want the key to be. Usually, we are interested in a security definition which is not based on any assumptions such as limited computing power or memory, or restricting the possible actions of the eavesdropper. Rather, we are looking for a statement that qualifies a key as *information-theoretically secure*: [54]

Definition 4.1.1. *Let S_A, S_B, and Z be random variables and $\varepsilon \geq 0$. Then (S_A, S_B) is an* ε-secure *key pair with respect to Z if there exists some random variable S with range*

$\mathcal{S}$ such that

$$\begin{aligned} \mathrm{Prob}[S_A = S_B = S] &\geq 1 - \varepsilon && (4.1)\\ H(S) &= \log|\mathcal{S}| && (4.2)\\ I(S:Z) &\leq \varepsilon. && (4.3) \end{aligned}$$

This definition has the following intuitive interpretation: The strings S_A and S_B are secure with respect to an adversary having access to the information contained in[1] Z, if they are almost (except with probability ε) equal to a uniform string S about which the eavesdropper has almost no information. Uniformity is quantified by the Shannon information $H(S)$, and the knowledge of the eavesdropper about Z is given by the mutual information $I(S:Z)$. It is important to allow for small deviations from the case of a perfectly secure key ($\varepsilon = 0$), since perfect security is impossible to obtain with a probabilistic protocol running on a finite time scale. Def. 4.1.1 already provides us with the strongest security definition appropriate for any classical application, thus it seems reasonable to take it as a starting point for our quantum version.

Let us now think about how we can translate our consideration from the last subsection to the quantum world: Clearly, at the end of any (quantum) protocol, Alice and Bob are left with classical data, the key. However, we have to consider the case that the adversary conducts an attack that provides her with a quantum system which contains some information about the key (we will see in the next section how she can do this). In this case, a natural generalization of Eq. (4.3) would be to bound Eve's accessible information, which is the maximal amount of information she can obtain about the key when choosing the optimal measurement on her quantum state. However, it has been shown that such an approach is problematic, due to an undesirable feature of the accessible information: Although the adversary's accessible information about the key S is negligibly small, it might happen that S is completely insecure in certain applications [55]. This shows that the security definition Def. 4.1.1 and also its generalization is not appropriate in the quantum case. In the next section, we introduce a security definition which accurately describes secret keys generated with QKD protocols.

4.2 Security in the Quantum World

As we have already mentioned, in quantum key distribution we need to deal with the case in which the adversary potentially holds a quantum system correlated with the

[1] For instance, Z might describe all public communication the adversary learns.

classical data to which the privacy amplification is to be applied. Moreover, we need to keep in mind that a QKD protocol is a probabilistic procedure that needs to be described at least partly by classical probability theory. For instance, we have to be able to include in the right way cases where the eavesdropper is extraordinarily lucky when it comes to making decisions such as the choice of measurement bases. In every protocol it might happen that Eve by accident chooses all the time the right basis in an intercept-resend attack, but obviously only with exceedingly small probability. A second major issue which one has to deal with when looking for reasonable security definitions for QKD is *universal composability*: Certainly, the key generated by a QKD scheme will be used in some other cryptographic application (e.g. a one-time pad); quite rarely, Alice and Bob will be satisfied with shared randomness. Therefore, we need to consider a QKD protocol *composed* with another cryptographic primitive. In this section, we will introduce a security definition that takes this into account.

In order to describe such universal composability, we need a formalism introduced in [56], the so-called "cq-states" defined in Def. 2.2.2: Consider a protocol together with an arbitrary eavesdropping strategy that can output keys from some set $\mathcal{X}$. For instance, if the protocol consists of n rounds each creating a single bit, $\mathcal{X}$ might be the set of all n-bit strings. Suppose a key $x \in \mathcal{X}$ is generated with probability p_x, and in this case the eavesdropper holds a quantum system ρ_E^x. Introducing a classical random variable X with range $\mathcal{X}$ and probability distribution $P_X(x) := p_x$, the state

$$\rho_{XE} = \sum_{x \in \mathcal{X}} P_X(x) |x\rangle\langle x| \otimes \rho_E^x, \tag{4.4}$$

describes all possible outcomes of the QKD protocol, together with the corresponding states of the adversary. Using this formalism, we can introduce the following definition of a secret key [37, 55, 57, 58]:

Definition 4.2.1. *Let $\rho_{S_A S_B E} \in \mathcal{B}(\mathcal{H}_S \otimes \mathcal{H}_S \otimes \mathcal{H}_E)$ be a ccq-state for some random variables S_A and S_B with range $\mathcal{S}$ and let $\varepsilon \geq 0$. We say that (S_A, S_B) is an ε-secure key pair, if*

$$\|\rho_{S_A S_B E} - \rho_{\text{uniform}} \otimes \rho_E\| \leq \varepsilon, \tag{4.5}$$

where $\rho_{\text{uniform}} = \sum_{s \in \mathcal{S}} |s\rangle\langle s| \otimes |s\rangle\langle s| / |\mathcal{S}|$ is the state describing an identical, uniformly distributed key pair, and $\rho_E = \text{tr}_{S_A, S_B}(\rho_{S_A S_B E})$.

The interpretation of the state $\rho_{S_A S_B E}$ is that Alice and Bob hold (not necessarily identical) bit strings described by the random variables S_A and S_B, respectively, and Eve holds a quantum system ρ_E that might be correlated with them. The ideal case, in which Alice's and Bob's strings are identical and uniformly distributed, together with

Eve being completely uncorrelated with them, is given by $\rho_{\text{uniform}} \otimes \rho_E$. A distance of ε of the state $\rho_{S_A S_B E}$ to this ideal scenario means that the key pair (S_A, S_B) behaves as an ideal key, except with probability ε.

4.3 Classification of Eavesdropping Strategies

In this section, we will give a classification of eavesdropping strategies based on the complexity of the eavesdropper's interaction with the quantum states and measurement. Formally, the set of all possible eavesdropping strategies can be divided into three classes, "individual", "collective", and "coherent" attacks. Individual attacks are the simplest ones, corresponding to an eavesdropper with little power. Coherent attacks are potentially the most powerful, assuming an eavesdropper with unlimited technological power and resources, only being limited by the laws of nature. More concretely, these classes of strategies are defined by how Eve interacts with the quantum signals that are sent from Alice to Bob and how she processes the information she gathers in this way. The most general way to describe how information about a quantum system ρ_A is extracted is the following: Attach an ancilla system in a predefined state $|0\rangle\langle 0|_E$ to ρ_A and perform a unitary operation U on the composite system $\rho_A \otimes |0\rangle\langle 0|_E$. Then do a measurement on the ancilla system $\rho_E := \text{tr}_A(U^\dagger \rho_A \otimes |0\rangle\langle 0|_E U)$. The measurement is given by a POVM $\mathcal{M} = \{M_j\}$ which yields outcome j with probability $\text{tr}(M_j \rho)$, when measuring a state ρ. We denote the classical probability distribution which is obtained in this way by $P^\rho_{\mathcal{M}}$, i.e. $P^\rho_{\mathcal{M}}(j) = \text{tr}(M_j \rho)$.

Consider the case where Alice sends n quantum systems $\rho_A^1, \ldots, \rho_A^n$ to Bob. An *individual* attack is an attack where Eve attaches an ancilla system $|0\rangle\langle 0|_E$ to each state ρ_A^i, applies the same unitary operation U and measures her part of all the composite systems individually using a POVM $\mathcal{M}^1$ for each system. *Collective* attacks are more general concerning the measurement, as they allow the eavesdropper to measure all ancilla systems collectively, using a POVM $\mathcal{M}^n$ acting on all of her probes simultaneously. The most general attack is the *coherent* attack, in which it is assumed that Eve attaches one "large" ancilla system to the total state $\rho_A^1 \otimes \cdots \otimes \rho_A^n$ and then performs a *global* unitary transformation U_n and measurement. More formally, the probability distribution that the eavesdropper obtains for each class of attacks is given by:

$$\text{Individual:} \quad P^{\rho_E^1}_{\mathcal{M}^1} \cdots P^{\rho_E^n}_{\mathcal{M}^1}, \quad \rho_E^i = \text{tr}_A(U^\dagger \rho_A^i \otimes |0\rangle\langle 0|_E U), \tag{4.6}$$

$$\text{Collective:} \quad P^{\rho_E^1 \otimes \cdots \otimes \rho_E^n}_{\mathcal{M}^n}, \quad \rho_E^i = \text{tr}_A(U^\dagger \rho_A^i \otimes |0\rangle\langle 0|_E U), \tag{4.7}$$

$$\text{Coherent:} \quad P^{\rho_E}_{\mathcal{M}^n}, \quad \rho_E = \text{tr}_A(U_n^\dagger (\rho_A^1 \otimes \cdots \otimes \rho_A^n) \otimes |0\rangle\langle 0|_E U_n), \tag{4.8}$$

Note that in this classification we always included a final measurement on the eavesdropper's system. This assumption has to some extent only a historical legitimation, since before the introduction of universal composability, the mutual information was used to quantify the eavesdropper's knowledge about the key. Thus, one had to classify how the adversary extract classical information out of her quantum state, which naturally leads to the distinction we have presented so far.

In the previous section we have seen that if we demand universal composable security, we have to consider the case where the adversary keeps a quantum system beyond the end of the QKD protocol. Since the distinction between individual and collective attacks is only on the level of measurements, for our security analysis which does not consider measurements on the adversary's system, these attacks are actually equivalent. The only two possibilities we can now discriminate is whether Eve's attaches a probe to each signal individually (the collective attack) or globally to all signals (the coherent attack).

4.4 The Role of Purifications

This section covers a simple mathematical concept in quantum information theory, purifications. Note that there exists two different notions of the word "purification": One of them describes the process of performing a global operation on several identical copies of input states, such that the purity of some of them is increased. One example of this process in entanglement distillation. When we are talking about a purification of some state $\rho \in \mathcal{B}(\mathcal{H})$, we mean a pure state $|\Psi\rangle \in \mathcal{H} \otimes \mathcal{H}_{\text{aux}}$ from a higher-dimensional Hilbert space such that $\text{tr}_{\text{aux}} |\Psi\rangle\langle\Psi| = \rho$. In subsection 4.4.1, we explain how this concept can be used to find the quantum state describing the system of Alice, Bob, and Eve. Subsection 4.4.2 deals with physical processes that create purifications. This section partly goes beyond the concepts of QKD, but we believe that the results are interesting on their own.

4.4.1 Purifications in QKD

In chapter 5 dealing with privacy amplification, it is shown that it is important for the analysis of the protocol to know the overall distributed state ρ_{ABE} describing Alice's, Bob's and Eve's system after the eavesdropping attack, but before the measurements. We have already derived in Sec. 3.2.3 that this state is given by

$$|\Psi\rangle_{ABE} = U_{BE}(\mathbb{1} \otimes A^j |\tilde{\phi}^+\rangle)|0\rangle_E, \tag{4.9}$$

where the encoding operators A^j account for the signal states used in the protocol and $|\tilde{\phi}^+\rangle = \sum_{x=0}^{d-1} \sqrt{P_X(x)}|x\rangle|x\rangle$. Eq. (4.9) is valid for all three classes of eavesdropping strategies presented in the previous section, but for simplicity, we investigate it for collective attacks. Also recall that for the idealized scenario in which in fact single copies of the signal states $|\phi_x^j\rangle$ are prepared, the preparation can be performed by the action of some modified encoding operators $\tilde{A}^j$ acting on *Alice's* system,

$$|\Psi\rangle_{ABE} = \tilde{A}^j \otimes U_{BE}|\tilde{\phi}^+\rangle|0\rangle_E, \tag{4.10}$$

thus the attack of the eavesdropper is performed on (one half of) the state $|\tilde{\phi}^+\rangle$, independently of the actual encoding. However, Eq. (4.10) is not of great help, because it still contains an arbitrary (unknown) unitary operation U_{BE}. Fortunately, one can easily characterize Eve's state simply in terms of the state shared between Alice and Bob after the distribution, ρ_{AB}. This is because if Eve applies a unitary operation on the pure state $|\tilde{\phi}^+\rangle$ sent by Alice, $U_{BE}|\tilde{\phi}^+\rangle|0\rangle_E = |\Psi\rangle_{ABE}$, the overall state $|\Psi\rangle_{ABE}$ will remain pure and fulfills $\mathrm{tr}_E\,|\Psi\rangle\langle\Psi|_E = \rho_{AB}$, i.e., Eve holds the purifying system of ρ_{AB}.

This approach is particularly suited for the analysis of the tomographic protocol (cf. Ch. 6) as in this protocol it is (at least in principle) possible for Alice and Bob to determine the state ρ_{AB} they share after the distribution. As we will see, for the calculation of the achievable key rate it is necessary to know the quantum state Eve holds, which then can be easily calculated by assuming that Eve holds the purifying system of ρ_{AB}.

4.4.2 On the (Im)possibility of Physical Purification

The question that arises in this context is: "How realistic is the assumption that Eve can create a purification of a given quantum system?" Being realistic means that there should exist a physical process, mapping the input state to its purification. We started investigating this topic in view of its application to QKD, as it is presented in the previous section. Specifically, we ask whether it is possible to purify a number of *unknown* quantum states by a single physical operation. However, it turned out that in QKD, this question can be answered by "no".

For any state $\rho \in \mathcal{B}(\mathcal{H})$, from the spectral decomposition $\rho = \sum_i p_i|\lambda_i\rangle\langle\lambda_i|$, with $p_i \geq 0$ and $\sum_i p_i = 1$, a purification is given by $|\Psi\rangle = \sum_i \sqrt{p_i}|\lambda_i\rangle|\chi_i\rangle$, where the $|\chi_i\rangle$ are mutually orthogonal states in an auxiliary Hilbert space $\mathcal{H}_{\mathrm{aux}}$. Such a state is unique up to unitary operation on $\mathcal{H}_{\mathrm{aux}}$ [51]. Since we are interested in physical processes, we look for completely positive and trace preserving maps Λ (which we will call "purifier") fulfilling $\mathrm{tr}\,\Lambda(\rho)^2 = 1$ (purity) and $\mathrm{tr}_{\mathrm{aux}}\,\Lambda(\rho) = \rho$ ("faithfullness").

The first question that we may ask is: "Does some physical map Λ exist that maps *any* input state onto its purification?" The answer to this question is given by the following [59]

Theorem 4.4.1. *(i) Every purifier that fulfills the purity condition for all input states is a constant map.*

(ii) Every purifier that fulfills the faithfullness condition for all input states does not increase the purity of any state.

Proof. We proof (i) by contradiction. Suppose there exists some purifier Λ such that $\Lambda(\rho)^2 = \Lambda(\rho)$ for all $\rho \in \mathcal{B}(\mathcal{H})$ with at least $\Lambda(\rho_1) \neq \Lambda(\rho_2)$ for some $\rho_1, \rho_2 \in \mathcal{B}(\mathcal{H})$. But then for the state $\rho_3 = (\rho_1 + \rho_2)/2$, by linearity of Λ, requiring $\Lambda(\rho_3) = (\Lambda(\rho_1) + \Lambda(\rho_2))/2$ to be pure also implies that $\Lambda(\rho_1) = \Lambda(\rho_2)$.

To show (ii), note that since for pure states $|\phi\rangle$, $\mathrm{tr}_{\mathrm{aux}}\, \Lambda(|\phi\rangle\langle\phi|) = |\phi\rangle\langle\phi|$, we have that $\Lambda(|\phi\rangle\langle\phi|) = |\phi\rangle\langle\phi| \otimes \sigma_\phi$, for some $\sigma_\phi \in \mathcal{B}(\mathcal{H}_{\mathrm{aux}})$. Considering the spectral decomposition of an arbitrary state ρ, $\rho = \sum_i p_i |\lambda_i\rangle\langle\lambda_i|$, due to linearity we find for the purity of the output: $\mathrm{tr}\, \Lambda(\rho)^2 = \mathrm{tr}(\sum_i p_i \Lambda(|\lambda_i\rangle\langle\lambda_i|))^2 = \mathrm{tr}(\sum_i p_i |\lambda_i\rangle\langle\lambda_i| \otimes \sigma_{\lambda_i})^2 = \sum_i p_i^2 \,\mathrm{tr}\, \sigma_{\lambda_i}^2 \leq \sum_i p_i^2 = \mathrm{tr}\, \rho^2$, i.e., the purity of ρ is not increased. □

This theorem shows that it is impossible to find a physical map that creates the purification of its input state whenever the input state is arbitrary (or unknown). Moreover, it tells us that only one of the two properties (purity and faithfullness) is already sufficient to rule out the existence. Clearly, the assumptions that Λ purifies *all* possible input states poses some great restrictions.

The next task therefore is to find sets $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ of states that can be purified. We call a set $\mathcal{M}$ of states *essentially pure* if for all states $\rho_i \in \mathcal{M}$, there exist some states σ_{B} and σ_{aux}, some unitary operation U (independent of ρ_i), and a pure state $|\phi_i\rangle \in \mathcal{H}_{\mathrm{A}}$ such that $\rho_i \otimes \sigma_{\mathrm{aux}} = U|\phi_i\rangle\langle\phi_i| \otimes \sigma_{\mathrm{B}} U^\dagger$. Note that the splitting between the Hilbert spaces is different for the left and right hand side: $\mathcal{H} \otimes \mathcal{H}_{\mathrm{aux}} \simeq \mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}}$. It is easy to see that essentially pure states can be purified: First note that any completely positive map can be characterized by appending a pure state, performing a global unitary operation and tracing out part of the combined system. Consider a completely positive map Λ generating pure states and suppose there exists some map Λ^{-1} such that $\Lambda^{-1}(\Lambda(\rho)) = \rho$. If we write $\Lambda^{-1}(\rho) = \mathrm{tr}_{\mathrm{aux}}\, U\rho \otimes |\psi\rangle\langle\psi| U^\dagger$, then $\tilde{\Lambda}(\rho) := U\Lambda(\rho) \otimes |\psi\rangle\langle\psi| U^\dagger$ still yields a pure state, because adding a pure state and performing unitary operations does not change the purity. Moreover, the map $\tilde{\Lambda}$ is a purifier, since $\mathrm{tr}_{\mathrm{aux}}\, \tilde{\Lambda}(\rho) = \rho$. This means that whenever we can find a reversible process Λ^{-1} for a map Λ that outputs pure states, we can construct a purifier. Applying this result to essentially pure states, we see that

we have a map $\Lambda(\rho_i) = \mathrm{tr}_\mathrm{B}\, U\rho_i \otimes \sigma_\mathrm{aux} U^\dagger$ that maps each state $\rho_i \in \mathcal{M}$ to a pure state $|\phi_i\rangle$. Additionally, there exists a reverse map $\Lambda^{-1}(|\phi_i\rangle\langle\phi_i|) = \mathrm{tr}_\mathrm{aux}\, U^\dagger|\phi_i\rangle\langle\phi_i| \otimes \sigma_\mathrm{B} U$ which implies that there exists a purifier for the set $\mathcal{M}$.

Not so obvious, essentially pure state are the only states that can be collectively purified by a single physical operation. This is formalized by the following [59]

Theorem 4.4.2. *Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$. Then the following statements are equivalent:*

(i) $\mathcal{M}$ is a set of essentially pure states.

(ii) A purifier for $\mathcal{M}$ exists.

(iii) There exists a completely positive and trace preserving map Λ such that $\mathrm{tr}\,\Lambda(\rho)^2 = 1$ for all $\rho \in \mathcal{M}$ and $\|\rho - \rho'\| = \|\Lambda(\rho) - \Lambda(\rho')\|$ for all $\rho, \rho' \in \mathcal{M}$.

Proof. For the proof, see [59]. □

Although this theorem classifies all states that can be purified completely, it is in general not easy to check whether a set of states is essentially pure.

Let us close our excursion about physical purification with the remark that, although the impetus for the study of this topic was the description of eavesdropping attacks in QKD, the concept presented in this section is not directly applicable: For the entanglement-based scheme (cf. Sec. 3.2.2), Eve seeks to create a purification of the state ρ_{AB} that gets distributed between Alice and Bob. Since Alice sends half of the state $|\tilde{\phi}^+\rangle_{AB} = \sum_{x=0}^{d-1} \sqrt{P_X(x)}|x\rangle|x\rangle$, Eve can only act on this half of the state: $|\tilde{\phi}^+\rangle \to \mathbb{1}_A \otimes U_{BE}|\tilde{\phi}^+\rangle_{AB}|0\rangle_E =: |\Psi\rangle_{ABE}$. Thus obviously, after her attack, the overall state $|\Psi\rangle_{ABE}$ is pure and she holds the purifying system $\rho_E = \mathrm{tr}_{AB}\,|\Psi\rangle\langle\Psi|_{ABE}$ of the distributed state $\rho_{AB} = \mathrm{tr}_E\,|\Psi\rangle\langle\Psi|_{ABE}$. On the other hand, for the prepare-and-measure scheme (cf. Sec. 3.2.1), Alice sends one of the r signals states $|\phi_j^x\rangle$ to Bob. But here, Eve is not interested in obtaining a purification of each single state. Rather, she would like to have the purifying system of the state describing all possible signals, Eq. (3.1). But we already have shown in Sec. 3.2.3 that this state is equivalent to the entanglement-based version, therefore we know that Eve can actually obtain the purifying system.

Chapter 5

Privacy Amplification

This chapter deals with the privacy amplification (PA) protocol Alice and Bob apply after the information reconciliation step (cf. Sec. 3.3.4). We start in Sec. 5.1 by explaining the basic ideas of privacy amplification from a purely classical point of view. Although in QKD, PA remains a classical protocol, i.e. it is some algorithm which is applied to classical data. Problems arise when we need to consider the case where the adversary holds a quantum system which might be correlated with this data. In Sec. 5.2, we will deal with this issue. The main result which we present in this section was derived in [37, 1]; it is an expression for the maximally achievable key length one can obtained by privacy amplification, as a function of entropies of the quantum states the adversary holds. These entropies are discussed in detail in Sec. 5.3, because they play a crucial role in the calculation of the key length. This section will be a cornerstone for the analysis of the "Tomographic Protocol" in Ch. 6.

5.1 Introduction

The aim of privacy amplification is to turn a string of (e.g.) bits held by the honest parties and about which the adversary might have some knowledge into a secret one. This is in general done by applying a certain function which shrinks the length of the string, but on the other hand outputs a key about which Eve has less information. In this section, we present a classical scenario as an example of how this can be achieved. We denote the string that Alice and Bob hold by random variables $\mathbf{X}$ and $\mathbf{Y}$, respectively, with range $\mathcal{X} = \mathcal{Y} = \{0,1\}^n$. Eve's knowledge is quantified by a random variable $\mathbf{Z}$ with the same range, and their correlation is given by a tripartite probability distribution $P_{\mathbf{XYZ}}$. For simplicity, we assume that the data is generated by independent and identically performed experiments, that is, the probability distribution $P_{\mathbf{XYZ}}$ factorizes with respect

X	Y	Z	P_{XYZ}
0	0	0	1/6
0	0	1	1/6
0	1	0	1/6
1	0	1	1/6
1	1	0	1/6
1	1	1	1/6

Figure 5.1: Example for correlations between three random variables X, Y, and Z. As a side note, these correlations can be obtained when measuring a certain bound-entangled state given in [60], in the computational basis.

to the n realizations, $P_{\mathbf{XYz}}(\mathbf{x}, \mathbf{y}, \mathbf{z}) = P_{XYZ}(x_1, y_1, z_1) \cdots P_{XYZ}(x_n, y_n, z_n)$. This means that we only need to consider the probability distribution P_{XYZ} of a single bit, which, as an example, is given in Fig. 5.1. Alice and Bob now apply a simple privacy amplification protocol devised in [31]: They agree on a block length L and divide their n-bit strings $\mathbf{x}$ and $\mathbf{y}$ into blocks of length L. For each block i, Alice randomly choses a "key bit" a_i and adds it modulo 2 to each bit in the block: $A_i := (x_1 \oplus a_i, x_2 \oplus a_i, \ldots, x_L \oplus a_i)$. All the blocks A_i are then announced over the public channel.

Bob, learning A_i, adds his own bits modulo 2, i.e. he computes for each block $B_i = (x_1 \oplus a_i \oplus y_1, x_2 \oplus a_i \oplus y_2, \ldots, x_L \oplus a_i \oplus y_L)$ and accepts the ith block if all bit values in B_i are the same, $x_j \oplus y_j \oplus a_i := b_i$, for all $1 \leq j \leq L$. If Bob accepts the block i, which happens when his bits y_j are either all correlated or all anti-correlated with Alice's bits x_j, both add a_i and b_j, respectively, to the list of newly created key bits. The probability that Alice's and Bob's key bit coincide, given that Bob accepts a block, is given by

$$\text{Prob}[a_i = b_i | \text{Bob accepts block } i] = \frac{\left(\frac{2}{3}\right)^L}{\left(\frac{2}{3}\right)^L + \left(\frac{1}{3}\right)^L}, \tag{5.1}$$

which tends to 1 for large L.

Eve's best strategy is to do the same as Bob does, i.e., to subtract her own block from Alice's block A_i in order to compute her guess e_i for the bit a_i. To calculate Eve's rate of success, we look again at the probability distribution P_{XYZ} given in Fig. 5.1: Given Bob accepts block, the corresponding random variables $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ must be distributed either according to line 3 and 4 or according to line 1, 2, 5, and 6, without any mixtures between those groups. In the first case, Eve will always guess Alice's bit value correctly, whereas in the second case, she will succeed only with probability 1/2, since her variable

is completely uncorrelated from Alice's in that situation. Thus,

$$\text{Prob}[a_i = e_i|\text{Bob accepts block } i] = \frac{\left(\frac{1}{3}\right)^L + \left(\frac{2}{3}\right)^L \frac{1}{2}}{\left(\frac{2}{3}\right)^L + \left(\frac{1}{3}\right)^L}, \tag{5.2}$$

which tends to 1/2 for large L. This means that in the limit of large block length, Alice and Bob can generate a new pair of strings (a_i, b_i), which are almost perfectly correlated, yet completely unknown to Eve.

This simple privacy amplification scheme may no be very efficient, since possibly very large block lengths L are needed to remove Eve's correlations with the key. We will focus on another method, the so-called two-universal hashing. Roughly speaking, Alice randomly chooses a function f from a certain set of functions (see Def. 5.1.1 below) and computes $f(\mathbf{x})$. Then she announces f via the public channel to Bob, who on his side computes $f(\mathbf{y})$. The key pair is then given by $(f(\mathbf{x}), f(\mathbf{y}))$, which can be shown to be more private than the original data $(\mathbf{x}, \mathbf{y})$ before. We will formalize this for the relevant quantum mechanical case in Sec. 5.2.

Definition 5.1.1. *A family of two-universal functions is a set $\mathcal{F}$ of functions $F : \mathcal{X} \to \mathcal{Y}$ together with a probability distribution P_F, such that*

$$\text{Prob}[f(x) = f(x')] \leq \frac{1}{|\mathcal{Y}|}, \quad \forall x \neq x' \in \mathcal{X}, \tag{5.3}$$

where $f \in \mathcal{F}$ is randomly chosen according to P_F.

We are interested in *hash* functions, i.e., functions where the cardinality of the image is (much) smaller than that of the domain. In particular, we will consider functions mapping their input to a bit string, that is, $F : \mathcal{X} \to \{0,1\}^\ell$, with $\ell \leq |\mathcal{X}|$. In the following, we will call ℓ the *key length*.

5.2 Privacy Amplification in the Quantum World

The main peculiarity with which we need to deal is the fact that the eavesdropper holds a quantum system which is correlated with Alice's and Bob's data. This implies that we cannot argue in the way we did in the previous section, and we have also explained in Sec. 4.1 that it is not enough to consider arbitrary measurements for Eve. It is important that we consider that the adversary holds a quantum system throughout the protocol. In this section, we review a result showing that privacy amplification by two-universal hashing can generate unconditionally secure keys. At first, we focus on the case where Alice's and Bob's data are already identical. Afterwards, we will loosen this assumption by including the error correction step into the analysis.

Assume that Alice and Bob apply the privacy amplification protocol at a step in which Alice's and Bob's classical data is already identical. Therefore, we only need to consider a single random variable X which is correlated with a quantum system E under Eve's control, as described by a cq-state ρ_{XE}. Alice chooses a hash function f from the set F according to P_F (which may be taken to be uniform), computes $f(x)$, and communicates f to Bob. After this step, they hold a classical bit string $f(x)$, where x is chosen according to P_X and f according to P_F. Since the choice of f is broadcasted over the classical channel, we have to add this information to Eve's pool of knowledge. We denote the quantum state describing this situation together with the system E that was correlated with x by $\rho_{F(X)EF}$. Formally, it is given by

$$\rho_{F(X)EF} = \sum_{f\in\mathcal{F}} P_F(f)\rho_{f(X)E} \otimes |f\rangle\langle f|, \tag{5.4}$$

where $\rho_{f(X)E}$ is the joint state where a particular function f is applied to X:

$$\rho_{f(X)E} = \sum_{x\in\mathcal{X}} |f(x)\rangle\langle f(x)| \otimes \rho_E^x \tag{5.5}$$

Note that in (5.4), both the system E and F are under Eve's control.

The following theorem bounds the distance of a secret key generated by two-universal hashing from an ideal key:

Theorem 5.2.1. *Let $\rho_{XE} \in \mathcal{B}(\mathcal{H}_X \otimes \mathcal{H}_E)$ be a cq-state for some random variable X with range $\mathcal{X}$ and let $\{F : \mathcal{X} \to \{0,1\}^{\tilde{\ell}}\}$ be a family of two-universal hash functions. Furthermore, denote by $\rho_{F(X)EF}$ the state after the two-universal hashing. Then*

$$\|\rho_{F(X)EF} - \rho_U \otimes \rho_{EF}\| \leq \frac{1}{2} 2^{-\frac{1}{2}(S_2(\rho_{XE}) - S_0(\rho_E) - \tilde{\ell})}, \tag{5.6}$$

where $\rho_U = 1/2^{\tilde{\ell}} \sum_{x\in\{0,1\}^{\tilde{\ell}}} |x\rangle\langle x|$ is a uniform key and $\rho_{EF} = \mathrm{tr}_{F(X)}\, \rho_{F(X)EF}$ and $\rho_E = \mathrm{tr}_X\, \rho_{XE}$.

Proof. For the proof, see [37, 58]. □

The quantities S_2 and S_0, occurring in (5.6) (and $S_2^{\varepsilon'}$ and $S_0^{\varepsilon'}$ that appear below in (5.7)) are called (smooth) Renyi entropies. We devote Sec. 5.3 to their detailed analysis. It turns out that the bound in (5.6) can be generalized and potentially improved by exploiting the triangle inequality of the trace distance. The simple calculation can be found in [58], yielding

$$\|\rho_{F(X)EF} - \rho_U \otimes \rho_{EF}\| \leq \frac{1}{2} 2^{-\frac{1}{2}(S_2^{\varepsilon'}(\rho_{XE}) - S_0^{\varepsilon'}(\rho_E) - \tilde{\ell})} + 2\varepsilon', \tag{5.7}$$

for any $\varepsilon' \geq 0$.

When studying QKD protocols, one is usually interested in the achievable key length for a given scenario. Moreover, one might require that the generated key fulfills some security requirements, for instance, it is ε-secure for some predefined ε. Thus, one would like to have the left-hand side of (5.7) to be upper bounded by ε, i.e.

$$\frac{1}{2}2^{-\frac{1}{2}(S_2^{\varepsilon'}(\rho_{XE})-S_0^{\varepsilon'}(\rho_E)-\tilde{\ell})} + 2\varepsilon' \stackrel{!}{\leq} \varepsilon. \tag{5.8}$$

This relation holds if the secret key length $\tilde{\ell}$ is bounded by

$$\tilde{\ell} \leq S_2^{\varepsilon'}(\rho_{XE}) - S_0^{\varepsilon'}(\rho_E) + 2\log(2\varepsilon - 4\varepsilon'). \tag{5.9}$$

In this equation, we have one free parameter $0 \leq \varepsilon' < \varepsilon/2$, with ε fixed by the security requirement. In [37], ε' was put to $\varepsilon/4$, for no apparent reason. To keep our notation consistent with [1], we start with a slightly modified version of (5.9), which was used in [1], however no proof was given:

$$\tilde{\ell} \leq S_2^{\varepsilon'}(\rho_{XE}) - S_0^{\varepsilon'}(\rho_E) + 2\log(\varepsilon). \tag{5.10}$$

with $\varepsilon' = (\varepsilon/8)^2$ (again for no apparent reason).

Relation (5.10) has the following interpretation: If Alice and Bob managed to distribute some random string X about which the adversary has some knowledge (in the form of a correlated quantum state), such that the total system is described by the state ρ_{XE}, then, performing a PA protocol based on two-universal hashing, it is impossible to turn X into ε-secret of length $\tilde{\ell}$ larger than the right hand side of Eq. (5.10). Note that this result is independent of how Alice and Bob actually obtained X, in particular, which distribution protocol they chose (cf. Sec. 3.2). Another very important feature of relation (5.10) is that the bound on the key length $\tilde{\ell}$ is tight, i.e. for an optimal choice of the PA protocol, equality can be achieved [61]. Since the proof of this statement is even constructive [61], we will from now on assume that Alice and Bob use an optimal PA protocol, thus, we will take

$$\tilde{\ell} = S_2^{\varepsilon'}(\rho_{XE}) - S_0^{\varepsilon'}(\rho_E) + 2\log(\varepsilon), \tag{5.11}$$

where $\tilde{\ell}$ is the maximally achievable key length, as a starting point for our further analysis (cf. also [1]).

Note that up to now, we have assumed that Alice and Bob both already hold the same string $\mathbf{X}$. However to achieve this, they have to correct their original strings $\mathbf{X}$ and $\mathbf{Y}$ (cf. Sec. 3.3.4). This involves sending some error correction information $\mathbf{W}$ over the

public channel, which means that some information is leaking to Eve. One can show [1] that, using an optimal error correction protocol[1], the length of $\mathbf{W}$ is given by $H_0^{\varepsilon'}(\mathbf{X}|\mathbf{Y})$. Here, $H_0^{\varepsilon'}$ is another smooth Renyi entropy, which is defined below in Sec. 5.3. This amount has to be subtracted from (5.11), leading to [1]

$$\ell = S_2^{\varepsilon'}(\rho_{XE}) - S_0^{\varepsilon'}(\rho_E) - H_0^{\varepsilon}(\mathbf{X}|\mathbf{Y}) + 2\log(\varepsilon). \tag{5.12}$$

Finally, we have obtained an expression for the obtainable secret key length ℓ for the case where Alice and Bob start with some classical data $\mathbf{X}$ and $\mathbf{Y}$, respectively, and where the correlation between the data $\mathbf{X}$ and the eavesdropper is described by the quantum state ρ_{XE}.

5.3 Smooth Renyi Entropies

In this section, we derive technical results concerning smooth Renyi entropies. Part of these result can be found in [62]. We start with the general definition and show in Sec. 5.3.2 how it can be reformulated in way that allows us to explicitly calculate the entropies for arbitrary input states. The main focus will lie on the analysis of S_2^{ε}, S_0^{ε}, and H_0^{ε}, which are the entropies occurring in the formula for the key length, Eq. (5.12). For these cases, we derive in Sec. 5.3.3 a solution which can be calculated by a simple algorithm. This allows us to calculate the obtainable key rate for the "Tomographic Protocol" (cf. Ch. 6). Finally, in Sec. 5.3.4, we derive some additivity properties of the Renyi entropies S_2^{ε} and S_0^{ε}, which play an important role in the analysis of multi-photon events in Sec. 6.4.

5.3.1 General Properties

Before we give the definition of the smooth Renyi entropies, we have to introduce some notation: Recall that we denote by $\mathcal{P}(\mathbb{C}^d)$ the set of positive operators acting on $\mathbb{C}^d$ and by $\mathcal{B}(\mathbb{C}^d) := \{\sigma \in \mathcal{P}(\mathbb{C}^d) : \operatorname{tr}\sigma = 1\}$ the set of all density matrices acting on $\mathbb{C}^d$. For $\rho_d \in \mathcal{B}(\mathbb{C}^d)$, define $e_{d'}(\rho_d) := \rho_d \oplus \operatorname{diag}\underbrace{(0,\ldots,0)}_{d'-d}$ to be an embedding of ρ_d into a Hilbert space of dimension d'. Finally, let $\mathcal{B}^{\varepsilon}(\rho_d) := \{\sigma_{d'} \in \mathcal{B}(\mathbb{C}^{d'}) : d' \geq d, \|\sigma_{d'} - e_{d'}(\rho_d)\| \leq \varepsilon\}$ be the set of density operators of dimension greater than or equal to d which are ε-close to ρ_d. In this chapter, we will often index a density matrix with the dimension of the Hilbert space on which it is acting.

[1] Note that realistic error correction protocols need more overhead. Their analysis however is much more involved and a topic of current research on its own.

The quantum version of smooth Renyi entropies is defined in the following way [1, 37]

Definition 5.3.1. *Let $\rho_d \in \mathcal{B}(\mathbb{C}^d)$, $\alpha \in [0,\infty]$, and $\varepsilon \in [0,\infty)$. The* ε-smooth Renyi entropy of order α *of ρ_d is defined as*

$$S_\alpha^\varepsilon(\rho_d) := \frac{1}{1-\alpha} \inf_{\sigma_{d'} \in \mathcal{B}^\varepsilon(\rho_d)} \log \operatorname{tr} \sigma_{d'}^\alpha. \tag{5.13}$$

For the definition of the classical conditional smooth Renyi entropy, we use the notation that for some random variable X with range $\mathcal{X}$ and associated probability distribution P_X, we denote by $\mathcal{B}^\varepsilon(P_X) := \{Q_X : Q_X(x) \leq P_X(x)\, \forall x \in \mathcal{X}, \sum_{x\in\mathcal{X}} Q_X(x) \geq 1-\varepsilon\}$ the set of probability distributions Q_X which are ε-close to P_X.

Definition 5.3.2. *Let X and Y be two random variables with range $\mathcal{X}$ and $\mathcal{Y}$, respectively, and let P_{XY} be some joint probability distribution. Furthermore, let $\alpha \in [0,\infty]$, and $\varepsilon \in [0,\infty)$. The* conditional ε-smooth Renyi entropy of order α *of X given Y is defined as*

$$H_\alpha^\varepsilon(X|Y) := \frac{1}{1-\alpha} \inf_{Q_{XY} \in \mathcal{B}^\varepsilon(P_{XY})} \max_{y\in\mathcal{Y}} \log \sum_{x\in\mathcal{X}} Q_{X|Y=y}(x)^\alpha. \tag{5.14}$$

We are in particular interested in the entropies occurring in Eq. (5.12), that is, S_2^ε, S_0^ε, and H_0^ε. For these cases, the two definitions from above reduce to the following form:

$$S_2^\varepsilon(\rho_d) = -\inf_{\sigma_{d'} \in \mathcal{B}^\varepsilon(\rho_d)} \log \operatorname{tr} \sigma_{d'}^2 \tag{5.15}$$

$$S_0^\varepsilon(\rho_d) = \inf_{\sigma_{d'} \in \mathcal{B}^\varepsilon(\rho_d)} \log \operatorname{rank} \sigma_{d'} \tag{5.16}$$

$$H_0^\varepsilon(X|Y) = \min_{\mathcal{A}:\mathrm{Prob}[\mathcal{A}]\geq 1-\varepsilon} \left(\max_{y\in\mathcal{Y}} \log |\{x \in \mathcal{X} : P_{X\mathcal{A}|Y=y}(x) > 0\}| \right), \tag{5.17}$$

where the minimum in (5.17) ranges over all events $\mathcal{A}$ occurring with probability of at least $1-\varepsilon$.

The explicit calculation of these entropies will be done in Sec. 5.3.3. However, for the case of the quantum Renyi entropies S_2^ε and S_0^ε, we will first translate the minimization over the set $\mathcal{B}^\varepsilon(\rho_d)$ to a minimization over real vectors, which eases the analysis significantly. The derivation of this transformation is the subject of the following section.

5.3.2 Simplifications for S_2^ε and S_0^ε

The main result of this section is that since the Renyi entropies only depend on the eigenvalues of the states taken from the set $\mathcal{B}^\varepsilon(\rho_d)$, the minimization can equivalently

be performed over the set of spectra that can be found in $\mathcal{B}^\varepsilon(\rho_d)$. This statement is proven below in Lemma 5.3.1.

Recall the definition of the set of operators which are ε-close to a given one, $\mathcal{B}^\varepsilon(\rho_d) := \{\sigma_{d'} \in \mathcal{B}(\mathbb{C}^{d'}) : d' \geq d, \|\sigma_{d'} - e_{d'}(\rho_d)\| \leq \varepsilon\}$. Now define $\mathcal{D}^\varepsilon(\rho_d) := \{\sigma_{d'} \in \mathcal{B}^\varepsilon(\rho_d) : [\sigma_{d'}, e_{d'}(\rho_d)] = 0\}$ to be the set of all such operators which additionally commute with ρ_d. Finally, let $\boldsymbol{\lambda}(\rho_d)$ be the ordered spectrum of ρ_d, i.e. $\boldsymbol{\lambda}(\rho_d) = (\lambda_1, \ldots, \lambda_d) \in \mathbb{R}^d$ in ascending order and denote by $\|\boldsymbol{\lambda} - \boldsymbol{\lambda}'\| = 1/2 \sum_i |\lambda_i - \lambda'_i|$ the distance of the vectors $\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}'$.

Lemma 5.3.1. *The two sets $\Lambda^\varepsilon_{\mathcal{B}}(\rho_d) = \{\boldsymbol{\lambda}(\sigma_{d'}) : \sigma_{d'} \in \mathcal{B}^\varepsilon(\rho_d)\}$, and $\Lambda^\varepsilon_{\mathcal{D}}(\rho_d) = \{\boldsymbol{\lambda}(\sigma_{d'}) : \sigma_{d'} \in \mathcal{D}^\varepsilon(\rho_d)\}$, defined as the sets of spectra that correspond to the sets of density matrices $\mathcal{B}^\varepsilon(\rho_d)$ and $\mathcal{D}^\varepsilon(\rho_d)$, respectively, are identical.*

Proof. Since $\mathcal{D}^\varepsilon(\rho_d) \subset \mathcal{B}^\varepsilon(\rho_d)$, it follows directly that $\Lambda^\varepsilon_{\mathcal{D}}(\rho_d) \subset \Lambda^\varepsilon_{\mathcal{B}}(\rho_d)$. To show the other inclusion, let $\boldsymbol{\mu} \in \Lambda^\varepsilon_{\mathcal{B}}(\rho_d)$. This means that there exists some $\sigma_{d'} \in \mathcal{B}^\varepsilon(\rho_d)$ such that $\boldsymbol{\mu} = \boldsymbol{\lambda}(\sigma_{d'})$ and $\|\sigma_{d'} - e_{d'}(\rho_d)\| \leq \varepsilon$. From the spectral decomposition $e_{d'}(\rho_d) = \sum_{i=1}^{d'} \nu_i |i\rangle\langle i|$ define $\tilde{\sigma}_{d'} := \sum_{i=1}^{d'} \mu_i |i\rangle\langle i|$, i.e. $[\tilde{\sigma}_{d'}, e_{d'}(\rho_d)] = 0$. We then have that

$$\begin{aligned} \|\tilde{\sigma}_{d'} - e_{d'}(\rho_d)\| &= \|\boldsymbol{\lambda}(\tilde{\sigma}_{d'}) - \boldsymbol{\lambda}(e_{d'}(\rho_d))\| && (5.18)\\ &= \|\boldsymbol{\lambda}(\sigma_{d'}) - \boldsymbol{\lambda}(e_{d'}(\rho_d))\| && (5.19)\\ &\leq \|\sigma_{d'} - e_{d'}(\rho_d)\| && (5.20)\\ &\leq \varepsilon, && (5.21) \end{aligned}$$

which implies that $\tilde{\sigma}_{d'} \in \mathcal{D}^\varepsilon(\rho_d)$, and since $\boldsymbol{\mu} = \boldsymbol{\lambda}(\tilde{\sigma}_{d'})$, it follows that $\boldsymbol{\mu} \in \Lambda^\varepsilon_{\mathcal{D}}(\rho_d)$. □

We will now reformulate the definition of $\mathcal{D}^\varepsilon(\rho_d)$. Since for all $\sigma_{d'} \in \mathcal{D}^\varepsilon(\rho_d)$, we have that $[\sigma_{d'}, e_{d'}(\rho_d)] = 0$, it follows that $\sigma_{d'} = \sigma_d \oplus \sigma_{d'-d}$ with $\sigma_d, \sigma_{d'-d} \in \mathcal{P}(\mathbb{C}^d)$ and $\operatorname{tr}\sigma_d, \operatorname{tr}\sigma_{d'-d} \leq 1$. Thus,

$$\begin{aligned} \|\sigma_{d'} - e_{d'}(\rho_d)\| &= \|\sigma_d - \rho_d\| + \|\sigma_{d'-d}\| && (5.22)\\ &= \|\sigma_d - \rho_d\| + \frac{1}{2}\operatorname{tr}\sigma_{d'-d} && (5.23)\\ &= \|\sigma_d - \rho_d\| + \frac{1}{2}(1 - \operatorname{tr}\sigma_d). && (5.24) \end{aligned}$$

We can therefore write

$$\begin{aligned} \mathcal{D}^\varepsilon(\rho_d) = \{&\sigma_d \oplus \sigma_{d'-d} \in \mathcal{B}(\mathbb{C}^{d'}) : && (5.25)\\ &\operatorname{tr}\sigma_d, \operatorname{tr}\sigma_{d'-d} \leq 1, [\sigma_d, \rho_d] = 0, \|\sigma_d - \rho_d\| \leq \varepsilon - \frac{1}{2}(1 - \operatorname{tr}\sigma_d)\}. && (5.26) \end{aligned}$$

Turning to the smooth Renyi entropies $S_\alpha^\varepsilon(\rho_d)$, since they only depend on the eigenvalues of ρ_d, we can use the above lemma to replace the infimum over the set $B^\varepsilon(\rho_d)$ by the infimum over $D^\varepsilon(\rho_d)$:

$$\begin{aligned}
S_\alpha^\varepsilon(\rho_d) &:= \frac{1}{1-\alpha} \inf_{\sigma_{d'} \in \mathcal{B}^\varepsilon(\rho_d)} \log \operatorname{tr} \sigma_{d'}^\alpha && (5.27)\\
&= \frac{1}{1-\alpha} \inf_{\sigma_{d'} \in \mathcal{D}^\varepsilon(\rho_d)} \log \operatorname{tr} \sigma_{d'}^\alpha && (5.28)\\
&= \frac{1}{1-\alpha} \inf_{\sigma_d \oplus \sigma_{d'-d} \in \mathcal{D}^\varepsilon(\rho_d)} \log \operatorname{tr} (\sigma_d \oplus \sigma_{d'-d})^\alpha && (5.29)\\
&= \frac{1}{1-\alpha} \inf_{\sigma_d \oplus \sigma_{d'-d} \in \mathcal{D}^\varepsilon(\rho_d)} \log (\operatorname{tr} \sigma_d^\alpha + \operatorname{tr} \sigma_{d'-d}^\alpha) && (5.30)\\
&= \frac{1}{1-\alpha} \inf_{\sigma_d \in \bar{\mathcal{D}}^\varepsilon(\rho_d), d' \geq d} \log \left[\operatorname{tr} \sigma_d^\alpha + (d'-d) \left(\frac{1-\operatorname{tr} \sigma_d}{d'-d} \right)^\alpha \right] && (5.31)\\
&= \frac{1}{1-\alpha} \inf_{\sigma_d \in \bar{\mathcal{D}}^\varepsilon(\rho_d), d' \geq d} \log \left[\operatorname{tr} \sigma_d^\alpha + \frac{1}{(d'-d)^{\alpha-1}} (1-\operatorname{tr} \sigma_d)^\alpha \right] && (5.32)\\
&= \frac{1}{1-\alpha} \inf_{\sigma_d \in \bar{\mathcal{D}}^\varepsilon(\rho_d)} \log \operatorname{tr} \sigma_d^\alpha, \quad \alpha \neq 1 && (5.33)
\end{aligned}$$

The crucial step in this derivation is from (5.30) to (5.31): Note that $\sigma_{d'-d}$ is only restricted by $\operatorname{tr} \sigma_{d'-d} = 1 - \operatorname{tr} \sigma_d$, otherwise, it is completely arbitrary (it only has to be positive). Thus the infimum over $\operatorname{tr} \sigma_{d'-d}^\alpha$ can be evaluated separately, and it is obtained for a uniform distribution of the eigenvalues, i.e. each eigenvalue is $\operatorname{tr} \sigma_{d'-d}/(d'-d) = (1-\operatorname{tr} \sigma_d)/(d'-d)$. The only freedom one still has is to choose the dimension d' and σ_d, which is now taken from the set

$$\bar{\mathcal{D}}^\varepsilon(\rho_d) := \{\sigma_d \in \mathcal{P}(\mathbb{C}^d) : \operatorname{tr} \sigma_d \leq 1, [\sigma_d, \rho_d] = 0, \|\sigma_d - \rho_d\| \leq \varepsilon - \frac{1}{2}(1 - \operatorname{tr} \sigma_d)\}, \quad (5.34)$$

which is loosely speaking the "remaining" part of the set $\mathcal{D}^\varepsilon(\rho_d)$ when forgetting about $\sigma_{d'-d}$. Concerning the step from (5.32) to (5.33), note that the second term in the logarithm is always non-negative, so for $\alpha > 1$ the infimum is obtained for $d' \to \infty$, whereas for $\alpha < 1$ it is obtained for $d' = d$, which leads to (5.33) in both cases. In particular, for $\alpha < 1$, the final infimum has to be taken over all $\sigma_d \in \bar{\mathcal{D}}^\varepsilon(\rho_d)$ with $\operatorname{tr} \sigma_d = 1$. Since for $\alpha < 1$, the minimization is performed over the set $\bar{\mathcal{D}}^\varepsilon(\rho_d)$ containing states with arbitrary trace, we can split up the infimum by fixing the trace to be t and finally minimizing over t, which leads to

$$S_\alpha^\varepsilon(\rho_d) = \begin{cases} \frac{1}{1-\alpha} \inf_{\sigma_d \in \bar{\mathcal{D}}_1^\varepsilon(\rho_d)} \log \operatorname{tr} \sigma_d^\alpha & \text{for } \alpha < 1 \\ \frac{1}{1-\alpha} \inf_{1-\varepsilon \leq t \leq 1} \inf_{\sigma_d \in \bar{\mathcal{D}}_t^\varepsilon(\rho_d)} \log \operatorname{tr} \sigma_d^\alpha & \text{for } \alpha > 1 \end{cases}, \quad (5.35)$$

where we have defined the set $\bar{\mathcal{D}}_t^\varepsilon(\rho_d)$ to be the set $\bar{\mathcal{D}}^\varepsilon(\rho_d)$ where all density matrices have trace t:

$$\bar{\mathcal{D}}_t^\varepsilon(\rho_d) := \{\sigma_d \in \mathcal{P}(\mathbb{C}^d) : \operatorname{tr}\sigma_d = t, [\sigma_d, \rho_d] = 0, \|\sigma_d - \rho_d\| \leq \varepsilon - \frac{1}{2}(1-t)\} \tag{5.36}$$

The range of the parameter t is obtained by using the following lower bound on the trace distance of the operators σ_d and ρ_d which occurs in (5.36):

$$\begin{aligned} \varepsilon - \frac{1}{2}(1-t) &\geq \|\sigma_d - \rho_d\| && (5.37)\\ &= \frac{1}{2}\operatorname{tr}|\sigma_d - \rho_d| && (5.38)\\ &\geq \frac{1}{2}|\operatorname{tr}(\sigma_d - \rho_d)| && (5.39)\\ &= \frac{1}{2}(1-t) && (5.40)\\ \Leftrightarrow t &\geq 1-\varepsilon, && (5.41) \end{aligned}$$

where we have used that σ_d and ρ_d commute.

The calculations can finally be further simplified by observing that the function which is to be minimized, $\log \operatorname{tr}\sigma_d^\alpha$, only depends on the eigenvalues of σ_d. Since the infimum in (5.33) only ranges over all σ_d taken from a set $\bar{\mathcal{D}}^\varepsilon(\rho)$ of operators which commute with the input state ρ, we can entirely focus on the eigenvalues of ρ (this is the essence of Lem. 5.3.1). Therefore, we can write

$$S_\alpha^\varepsilon(\rho_d) = \begin{cases} \frac{1}{1-\alpha} \inf\limits_{\boldsymbol{\mu}\in\Lambda^\varepsilon_{\bar{\mathcal{D}}_1}(\rho_d)} \log \sum\limits_{i=1}^d \mu_i^\alpha & \text{for } \alpha < 1 \\ \frac{1}{1-\alpha} \inf\limits_{1-\varepsilon\leq t\leq 1} \inf\limits_{\boldsymbol{\mu}\in\Lambda^\varepsilon_{\bar{\mathcal{D}}_t}(\rho_d)} \log \sum\limits_{i=1}^d \mu_i^\alpha & \text{for } \alpha > 1 \end{cases} \tag{5.42}$$

where

$$\Lambda^\varepsilon_{\bar{\mathcal{D}}_t}(\rho_d) = \{\boldsymbol{\mu} \in \mathbb{R}^d : \sum_{i=1}^d \mu_i = t, \|\boldsymbol{\mu} - \boldsymbol{\lambda}(\rho_d)\| \leq \varepsilon - \frac{1}{2}(1-t)\}. \tag{5.43}$$

To conclude, we have transformed the minimization over a subset of a Hilbert space to a minimization over a set of real vectors subjected to simple constraints. In the next subsection, we will evaluate Eq. (5.42) for two special cases (namely $\alpha = 0$ and $\alpha = 2$) which play an important role in the privacy amplification step (cf. Sec. 5.2).

5.3.3 Explicit Calculation of S_2^ε, S_0^ε, and H_0^ε

The aim of this section is to construct an algorithm that computes the Renyi entropies S_2^ε and S_0^ε. Due to the minimization involved, it is not possible to present a closed expression, since $S_\alpha^\varepsilon(\rho)$ strongly depends on the explicit distribution of the eigenvalues of ρ. However, we found a strategy that allows us compute these entropies efficiently.

Calculation of $S_0^\varepsilon(\rho)$

For the case of $\alpha = 0$, from Eq. (5.42) we find that the smooth Renyi entropy takes a very simple form:

$$S_0^\varepsilon(\rho) = \inf_{\boldsymbol{\mu}\in\Lambda^\varepsilon_{\mathcal{D}_1}(\rho)} \log \sum_{i=1}^{d} \mu_i^0 \tag{5.44}$$

$$= \inf_{\boldsymbol{\mu}\in\Lambda^\varepsilon_{\mathcal{D}_1}(\rho)} \log |\{\mu_i > 0\}|, \tag{5.45}$$

where the set $\Lambda^\varepsilon_{\mathcal{D}_1}(\rho)$ is defined in (5.43). Note that for $\alpha < 1$, there is no optimization over t, rather, we only need to consider the case $t = 1$. Calculating $S_0^\varepsilon(\rho)$ is thus equivalent to finding the solution $\mathbf{x}$ of the following optimization problem:

$$|\{x_i > 0\}| \to \min \tag{5.46}$$

$$\sum_{i=1}^{n} x_i = 1 \tag{5.47}$$

$$\sum_{i=1}^{n} |x_i - y_i| \leq 2\varepsilon, \tag{5.48}$$

where the vector $\mathbf{y}$ corresponds to the vector of eigenvalues of ρ_d (which is given). The solution to this optimization problem is very simple: Find the largest number k such that the sum of the smallest k values y_i is smaller or equal to ε. We define the solution $\mathbf{x}$ by starting from $\mathbf{y}$ (which certainly satisfies (5.47) and (5.48)) and modify it in the following way: Put the smallest k values y_i to zero and rise the largest y_i by ε such that both (5.47) and (5.48) are still satisfied. The solution $\mathbf{x}$ obtained in this way fulfills $|\{x_i > 0\}| = |\{y_i > 0\}| - k$.

To be more specific, let $\rho \in \mathcal{B}(\mathbb{C}^d)$ some density matrix having eigenvalues λ_i with degeneracy n_i, for $1 \leq i \leq m$, i.e. $\sum_{i=1}^m n_i\lambda_i = 1$ and $\sum_{i=1}^m n_i = d$. We assume that the λ_i are given in ascending order. Define

$$s^-(r) := \sum_{i=1}^{r} n_i\lambda_i, \tag{5.49}$$

for $0 \leq r \leq m$, which is the sum of r smallest different eigenvalues. (For $r = 0$, the sum is taken to be zero.) Moreover, let

$$b^- := \max\{r : s^-(r) \leq \varepsilon\} \tag{5.50}$$

be the largest number r such that the sum of the r smallest different eigenvalues is smaller than ε. Then the number k we are looking for is thus given by

$$k = \sum_{i=1}^{b^-} n_i + \left\lfloor \frac{\varepsilon - s^-(b^-)}{\lambda_{b^-+1}} \right\rfloor, \tag{5.51}$$

where $\lfloor x \rfloor$ denotes the largest integer smaller than or equal to x. Finally, we have found that

$$S_0^\varepsilon(\rho) = \log(d-k). \tag{5.52}$$

Calculation of $H_0^\varepsilon(X|Y)$

The calculation of this classical Renyi entropy is similar to the calculation of $S_0^\varepsilon(\rho)$ performed in the previous section. First note that we can rewrite its definition,

$$H_0^\varepsilon(X|Y) = \min_{\mathcal{A}:\mathrm{Prob}[\mathcal{A}]\geq 1-\varepsilon} \max_{y\in\mathcal{Y}} \log|\mathcal{P}_{\mathcal{A}y}|, \tag{5.53}$$

introducing the set $\mathcal{P}_{\mathcal{A}y} := \{x \in \mathcal{X} : P_{X\mathcal{A}|Y=y}(x) > 0\}$. In the analysis of the Tomographic Protocol (see Sec. 6.1) we will encounter the case where the set $\mathcal{P}_{\mathcal{A}y}$ is actually independent of y. This means that we can drop the minimization over y and simply consider the set $\mathcal{P}_{\mathcal{A}y_0}$, for some arbitrarily chosen y_0. Thus the only restriction on the size of the $\mathcal{P}_{\mathcal{A}y_0}$ is given by $\mathcal{A}$. The minimization over all such events occurring with probability of at least $1-\varepsilon$ can be solved in the following way: It is easy to see that all such events are necessarily of the form $[X = x_1] \vee [X = x_2] \vee \cdots \vee [X = x_k]$ with $\sum_{i=1}^k P_X(x_i) \geq 1-\varepsilon$. Since we are searching for the smallest set $\mathcal{P}_{\mathcal{A}y_0}$, we are interested in those events which are most restrictive, i.e. which have k as small as possible. This means we need to find the smallest number k such that the sum of k largest probabilities occurring in $P_{X|Y=y_0}$ is greater than or equal to $1-\varepsilon$. We then have found that $H_0^\varepsilon(X|Y) = \log k$.

To explicitly calculate k, consider some probability distribution $P_{X|Y=y_0}$ having m entries p_i with multiplicity n_i, for $1 \leq i \leq m$, i.e. $\sum_{i=1}^m n_i p_i = 1$. We again assume that the p_i are given in ascending order. In analogy to the calculation of $S_0^\varepsilon(\rho)$, define

$$s^+(r) := \sum_{i=1}^{r} n_{m-i+1} p_{m-i+1}, \tag{5.54}$$

which is the sum of r largest different probabilities. Also define

$$b^+ := \min\{r : s^+(r) \geq 1-\varepsilon\} \tag{5.55}$$

to be the smallest number r such that sum of the largest r probabilities is greater than or equal to $1-\varepsilon$. The number k is then given by

$$k = \sum_{i=1}^{b^+} n_{m-i+1} + \left\lfloor \frac{s^+(b^+) - (1-\varepsilon)}{p_{m-b^+}} \right\rfloor, \tag{5.56}$$

which finally leads to

$$H_0^\varepsilon(X|Y) = \log k. \tag{5.57}$$

Note that this result only holds if the number of non-zero entries in $P_{X|Y=y}$ does not depend on y. (However, this is the only case we are interested in.) If this is not the case, one would first need to find the corresponding y which maximizes the number of non-zero entries, which corresponds to evaluating the maximum in (5.53).

Calculation of $S_2^\varepsilon(\rho)$

The calculation of $S_2^\varepsilon(\rho)$ is more involved, mainly because we have an additional parameter t which corresponds to the trace of the density matrices over which we have to minimize (in contrast to the entropies S_0^ε and H_0^ε). From Eq. (5.42), we find that

$$S_2^\varepsilon(\rho) = -\inf_{1-\varepsilon\le t\le 1} \inf_{\boldsymbol{\mu}\in\Lambda^\varepsilon_{\mathcal{D}_t}(\rho)} \log \sum_{i=1}^d \mu_i^2, \tag{5.58}$$

where the set $\Lambda^\varepsilon_{\mathcal{D}_t}(\rho)$ is defined by (5.43). We defer the minimization over t until the very end. Therefore, for a fixed t, we are facing an optimization problem of the following form:

$$\text{(o1)} \qquad \sum_{i=1}^n x_i^2 \to \min \tag{5.59}$$

$$\text{(c1)} \qquad \sum_{i=1}^n x_i = t \tag{5.60}$$

$$\text{(c2)} \qquad \sum_{i=1}^n |x_i - y_i| \le 2\varepsilon - (1-t) \tag{5.61}$$

Here, the y_i are the eigenvalues of ρ_d (which are given).

We will construct the solution of this optimization problem successively by a series of small steps. First note that we can demand that the components in the vector $\mathbf{y}$ are given in ascending order, i.e. $y_i \le y_j$ for all $i \le j$. It is then easy to see that any solution $\mathbf{x}$ can be ordered in the same way:

Lemma 5.3.2. *Let $\mathbf{x}$ be a solution of (o1,c1,c2). Suppose there exist two indices $k \le l$ such that $y_k \le y_l$ and $x_k \ge x_l$. Then $\tilde{\mathbf{x}} = (x_1, \ldots, x_l, \ldots, x_k, \ldots, x_n)$, i.e., the original solution with x_k and x_l interchanged, is also a solution of (o1,c1,c2).*

Proof. Obviously, we have that $\sum_{i=1}^n \tilde{x}_i^2 = \sum_{i=1}^n x_i^2$ and $\sum_{i=1}^n \tilde{x}_i = \sum_{i=1}^n x_i$, that is, $\tilde{\mathbf{x}}$ fulfills (o1) and (c1). It remains to show that also (c2) holds, which can be done by verifying $|x_k - y_k| + |x_l - y_l| \ge |x_l - y_k| + |x_k - y_l|$ using the triangle inequality and considering all (six) possible relations between x_k, x_l, y_k, and y_l separately. □

By successive application of Lemma 5.3.2 we see that any non-monotonically increasing solution $\mathbf{x}$ can be turned into a monotonically increasing one $\tilde{\mathbf{x}}$. The next lemma reveals something about the structure of the solution.

Lemma 5.3.3. *Let $\mathbf{x}$ be a solution of (o1,c1,c2), where the x_i are in ascending order. Then we can transform $\mathbf{x}$ in the following two ways without changing its optimality:*

1. *(a) Let $J \subset [1,n]$ be the set[2] of indices such that $x_j \leq y_j$ for all $j \in J$. Let $j_{\min}$ be the smallest index in J such that $x_{j_{\min}} < y_{j_{\min}}$ and $j_{\max}$ the largest index in J. Define*

$$\tilde{x}_{j_{\min}} := \min(y_{j_{\min}}, (x_{j_{\min}} + x_{j_{\max}})/2) \tag{5.62}$$

$$\tilde{x}_{j_{\max}} := \max(x_{j_{\max}} - x_{j_{\min}} + y_{j_{\max}}, (x_{j_{\min}} + x_{j_{\max}})/2), \tag{5.63}$$

   and $\tilde{x}_j = x_j$ for all $j \neq j_{\min}, j_{\max}$.

   (b) Permute the $\{\tilde{x}_j\}_{j\in J}$ such that they are again in ascending order.

2. *(a) Let $K \subset [1,n]$ be the set of indices such that $x_k \geq y_k$ for all $k \in K$. Let $k_{\min}$ be the smallest index in K and $k_{\max}$ be the largest index in K such that $x_{k_{\max}} > y_{k_{\max}}$. Define*

$$\tilde{x}_{k_{\max}} := \max(y_{k_{\max}}, (x_{k_{\min}} + x_{k_{\max}})/2) \tag{5.64}$$

$$\tilde{x}_{k_{\min}} := \min(x_{k_{\min}} + x_{k_{\max}} - y_{k_{\max}}, (x_{k_{\min}} + x_{k_{\max}})/2), \tag{5.65}$$

   and $\tilde{x}_k = x_k$ for all $k \neq k_{\min}, k_{\max}$.

   (b) Permute the $\{\tilde{x}_k\}_{k\in K}$ such that they are again in ascending order.

Proof. We proof the assertion only for the first transformation, the proof for the second one is completely analogous. Due to Lemma 5.3.2, step (b) preserves the optimality of the solution. It remains to consider step (a), i.e., we have to show that $\tilde{\mathbf{x}}$ is a solution of (o1,c1,c2). By construction, we have that $\tilde{x}_{j_{\min}} - x_{j_{\min}} = -\tilde{x}_{j_{\max}} + x_{j_{\max}} =: \delta$, which implies that $\tilde{\mathbf{x}}$ fulfills (c1) and (c2). To show (o1), we directly calculate

$$\sum_{j\in J} \tilde{x}_j^2 = \sum_{j\neq j_{\min}, j_{\max}} x_j^2 + (x_{j_{\min}} + \delta)^2 + (x_{j_{\max}} - \delta)^2 \tag{5.66}$$

$$= \sum_{j\in J} x_j^2 + 2\delta^2 + 2\delta(\underbrace{x_{j_{\min}} - x_{j_{\max}}}_{=-2\delta}) \tag{5.67}$$

$$\leq \sum_{j\in J} x_j^2. \tag{5.68}$$

□

[2]It is easy to see that the set J is always non-empty unless $\varepsilon = 0$ and $t = 1$.

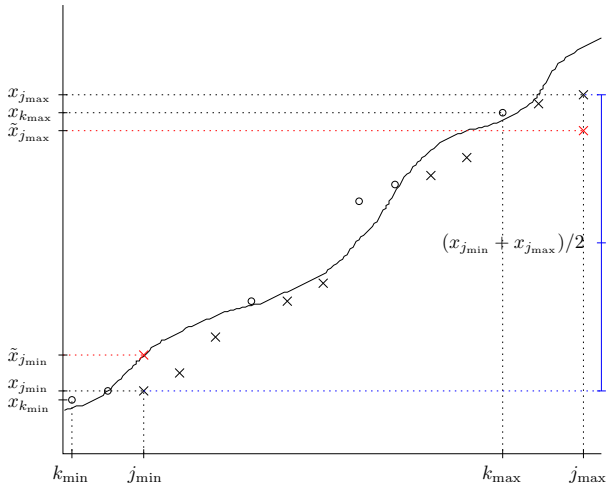

Figure 5.2: (Color online.) Visualization of the notation given in Lemma 5.3.3. For better clarity, we have plotted the y_i as a continuous function. All x_j with $j \in J$ are depicted as crosses, whereas the x_k with $k \in K$ are depicted as circles. We have indicated one step of transformation 1 (a), where we have $x_{j_{\min}} \to \tilde{x}_{j_{\min}} = y_{j_{\min}}$ and $x_{j_{\max}} \to \tilde{x}_{j_{\max}} = x_{j_{\max}} - x_{j_{\min}} + y_{j_{\max}}$, that is, the minimum in Eq. (5.62) and the maximum in Eq. (5.63) are attained for the first argument.

To illustrate the meaning of the above lemma, we give the following geometric interpretation: Transformation 1 rises the smallest x_i while simultaneously lowering the largest one. Their final value is either their arithmetic mean if this is smaller than y_i. If it is larger, then x_i will only be raised up to y_i, and the largest x will only be lowered by the same smaller amount (see also Fig. 5.2).

We can iteratively apply Lemma 5.3.3 to any given solution $\mathbf{x}$ of (o1,c1,c2) until both transformations do not change the solution anymore. At this point, $\mathbf{x}$ will be of the following form: There might exist some is with $x_i = y_i$, for all j, j' such that $x_j < y_j$ and $x_{j'} < y_{j'}$ we have that $x_j = x_{j'} =: x_{\max}$, and for all k, k' such that $x_k > y_k$ and $x_{k'} > y_{k'}$ we have that $x_k = x_{k'} =: x_{\min}$. This means that all x_i are either equal to y_i or take one out of two constant values, depending on whether $x_i < y_i$ or $x_i > y_i$ (see also Fig. 5.3). It remains to find these constants and the set of indices for which $x_i < y_i$ and $x_i > y_i$, which will be done in the remaining part of this section:

Lemma 5.3.4. *Let* $\mathbf{x}$ *be a solution of (o1,c1,c2). Then* $\mathbf{x}$ *satisfies (c2) tightly.*

Proof. Suppose $\mathbf{x}$ is a solution of (o1,c1,c2) with $\sum_{i=1}^{n} |x_i - y_i| =: \delta < 2\varepsilon - (1-t)$.

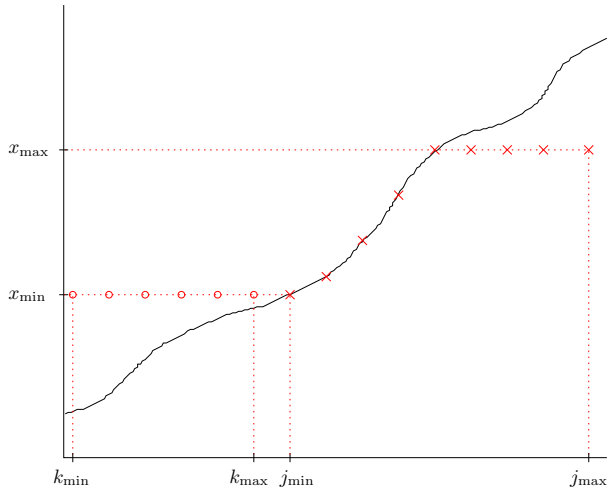

Figure 5.3: (Color online.) Visualization of a solution of (o1,c1,c2) after iteratively applying Lemma 5.3.3 until the transformations no longer change the result. At this point, the solution $\{x_i\}$ (depicted as crosses and circles) is of the following form: For $k_{\min} \leq k \leq k_{\max}$, we have $\tilde{x}_k = x_{\min}$, for $j_{\min} \leq j \leq j_{\max}$, we have $\tilde{x}_j = x_{\max}$, and $\tilde{x}_i = y_i$ otherwise. Again we have plotted the y_i as a continuous function for better clarity.

Define $J^+(x) = \{i \in [1,n] : x_i \geq x\}$ and $J^-(x) = \{i \in [1,n] : x_i \leq x\}$, and furthermore

$$s^+(x) = \sum_{j \in J^+(x)} (x_i - x), \tag{5.69}$$

$$s^-(x) = \sum_{j \in J^-(x)} (x - x_i) \tag{5.70}$$

Find $x_{\max}$ and $x_{\min}$ such that[3] $s^+(x_{\max}) = \delta/2 = s^-(x_{\min})$. We can then construct a new solution $\tilde{\mathbf{x}}$ in the following way: Define $\tilde{x}_j := x_{\min}$ for $j \in J^-(x_{\min})$, $\tilde{x}_j := x_{\max}$ for $j \in J^+(x_{\max})$, and $\tilde{x}_i = x_i$ otherwise. It is easy to see that $\tilde{\mathbf{x}}$ satisfies (c1) and it is constructed such that $\sum_{i=1}^n |\tilde{x}_i - y_i| = \varepsilon$, i.e., it satisfies (c2) tightly. The proof that $\tilde{\mathbf{x}}$ fulfills condition (o1) follows the same lines as in Lemma 5.3.3. □

Define $x_i =: y_i + \Delta x_i$ and suppose there exists some m, $1 \leq m \leq n$ such that $x_i \geq y_i$ for $1 \leq i \leq m-1$, $x_i \leq y_i$ for $m \leq i \leq n$, and $x_i = y_i$ otherwise. Since $\mathbf{x}$ with $x_i \geq y_i$ for all i is certainly not a solution of (o1,c1,c2)[4], such an m clearly exists; at least we

[3] It is easy to see that both $x_{\max}$ and $x_{\min}$ are well-defined since $s^+(x)$ and $s^-(x)$ are continuous functions.

[4] Except for the degenerate case with $\varepsilon = 0$ and $t = 1$.

can have $m = 1$ which means that $x_i \leq y_i$ for all i. We can now rewrite (c2) and (c1), using Lemma 5.3.4 to turn the inequality in (c2) into an equality:

$$2\varepsilon - (1-t) = \sum_{i=1}^{m-1} \Delta x_i - \sum_{i=m}^{n} \Delta x_i \tag{5.71}$$

$$t = \sum_{i=1}^{n} x_i = 1 + \sum_{i=1}^{n} \Delta x_i \tag{5.72}$$

Inserting (5.72) into (5.71) yields the following two constraints:

$$\sum_{i=1}^{m-1} \Delta x_i = \varepsilon - (1-t) \tag{5.73}$$

$$-\sum_{i=m}^{n} \Delta x_i = \varepsilon \tag{5.74}$$

Using these two equations, we can compute $x_{\min}$ and $x_{\max}$ by cutting the largest and smallest values of the given vector $\mathbf{y}$ until (5.73) and (5.74) are saturated.

We are left with evaluating the final minimum over t, which has the range $1 - \varepsilon \leq t \leq 1$. From Eqs. (5.73) and (5.74), which are equivalent to the original constraints (c1,c2), we see directly that the optimal choice to minimize (o1) is to choose $(1-t)$ as large as possible, i.e., $t = 1 - \varepsilon$.

Let us recapitulate the results so far: We have shown that the solution $\mathbf{x}$ of the optimization problem (o1,c1,c2) given by Eqs. (5.59), (5.60), and (5.61) is of the following form: There exists some index $i_{\max} \in [1, n]$ with $x_i = x_{\max} < y_i$ for all $i \geq i_{\max}$ and $x_i = y_i$ for all $i < i_{\max}$. The value $x_{\max}$ can be found by exploiting Eq. (5.74): One cuts as many of the largest values y_i as possible down to $x_{\max}$ until the sum of these changes equals ε. We are left with finding a way to calculate the value $x_{\max}$. To this end let $\rho \in \mathcal{B}(\mathbb{C}^d)$ some density matrix having eigenvalues λ_i with degeneracy n_i, for $1 \leq i \leq m$. We assume that the λ_i are given in ascending order. Define

$$s^+(r) := \sum_{i=1}^{r} n_{m-i+1}(\lambda_{m-i+1} - \lambda_{m-r}), \tag{5.75}$$

for $0 \leq r \leq m-1$, to be the sum of the differences of the largest r eigenvalues. Moreover, let

$$b^+ := \max\{r : s^+(r) \leq \varepsilon\}, \tag{5.76}$$

which implies that

$$x_{\max} = \lambda_{m-b^+} - \frac{\varepsilon - s^+(b^+)}{\sum_{i=0}^{b^+} n_{m-i}}. \tag{5.77}$$

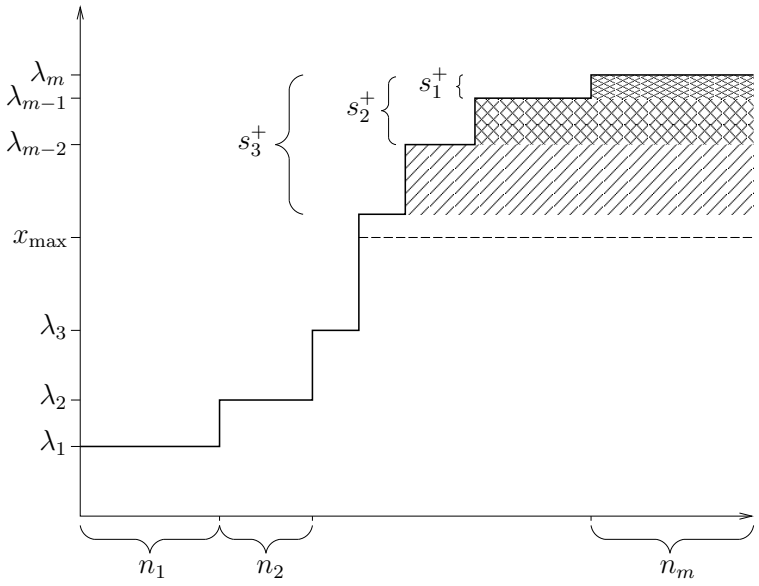

Figure 5.4: Visualization of the definition of $s^+(r)$ given in Eq. (5.75), $x_{\max}$, given by Eq. (5.77), and the eigenvalues λ_i and their multiplicity n_i. In this example, we have $b^- = 1$ and $b^+ = 3$.

For better clarity, see also Fig. 5.4. The entropy is then finally given by

$$S_2^\varepsilon(\rho) = -\log\left(\sum_{i=1}^{b^+-1} n_i\lambda_i^2 + \sum_{i=b^+}^{m} n_i x_{\max}^2\right). \tag{5.78}$$

5.3.4 Additivity

In this section, we calculate the quantum Renyi entropies S_0^ε and S_2^ε for density matrices which are of a special form. We will encounter this type of density matrices when we are dealing with multi-photon events in Sec. 6.4. The aim of this section is to provide a simple expression resembling an additivity property for the Renyi entropies of these states. They are of the form

$$\sigma = \rho \otimes (\mathbb{1}_r/r \oplus \operatorname{diag}\underbrace{(0,\ldots,0)}_{s})^{\otimes m}, \tag{5.79}$$

where $\rho \in \mathcal{B}(\mathbb{C}^d)$ and $\mathbb{1}_r/r \in \mathcal{B}(\mathbb{C}^r)$. Denote by λ_i and n_i, $i = 1, 2, \ldots, n$ the eigenvalues of ρ and their multiplicity, respectively, where the eigenvalues are in ascending order such that $\lambda_1 = 0$ and $\lambda_i > 0$ for $i > 1$. Then the eigenvalues of σ are given by $\bar{\lambda}_i = r^{-m}\lambda_i$ for

all i, and the multiplicities are given by $\bar{n}_1 = r^m n_1 + d[(r+s)^m - r^m)]$ and $\bar{n}_i = r^m n_i$ for $i > 1$.

Approximate additivity of S_0^ε

Recall that, due to Eq. (5.52) and the proceeding discussion, $S_0^\varepsilon(\rho) = \log(\dim \rho - k)$, where $k = \sum_{i=1}^{b^-} n_i + \lfloor(\varepsilon - s^-(b^-))/\lambda_{b^-+1}\rfloor$ with $b^- = \max\{r : s^-(r) \leq \varepsilon\}$ and $s^-(r) = \sum_{i=1}^{r} n_i\lambda_i$. Likewise, denote by $\bar{k}$, $\bar{b}^-$, and $\bar{s}(r)$ all those quantities referring to σ given by (5.79). We see that

$$\bar{s}^-(r) = \sum_{i=1}^{r} \bar{n}_i\bar{\lambda}_i = \sum_{i=2}^{r} n_i\lambda_i + (r^m n_0 + d[(r+s)^m - r^m]) \cdot 0 = s^-(r), \tag{5.80}$$

and therefore also $\bar{b}^- = b^-$. However, the expression for $\bar{k}$ is more complicated due to the Gauss brackets involved. We introduce the quantity δ which amounts for the rounding:

$$k = \sum_{i=1}^{b^-} n_{i-1} + \left\lfloor \frac{\varepsilon - s^-(b^-)}{\lambda_{b^-+1}} \right\rfloor = \sum_{i=1}^{b^-} n_{i-1} + \frac{\varepsilon - s^-(b^-)}{\lambda_{b^-+1}} - \delta \tag{5.81}$$

and similar for $\bar{k}$ with $\bar{\delta}$:

$$\bar{k} = \sum_{i=1}^{b^-} \bar{n}_{i-1} + \left\lfloor \frac{\varepsilon - s^-(b^-)}{\bar{\lambda}_{b^-+1}} \right\rfloor = \sum_{i=1}^{b^-} \bar{n}_{i-1} + \frac{\varepsilon - s^-(b^-)}{\bar{\lambda}_{b^-+1}} - \bar{\delta} \tag{5.82}$$

Now we can express $\bar{k}$ in terms of k:

$$\begin{aligned}
\bar{k} &= \sum_{i=1}^{b^-} \bar{n}_{i-i} + \left\lfloor \frac{\varepsilon - s^-(b^-)}{\bar{\lambda}_{b^-+1}} \right\rfloor && (5.83)\\
&= \sum_{i=2}^{r} r^m n_{i-1} + r^m n_0 + d[(r+s)^m - r^m] + \left\lfloor \frac{\varepsilon - s^-(b^-)}{\bar{\lambda}_{b^-+1}} \right\rfloor && (5.84)\\
&= \sum_{i=2}^{r} r^m n_{i-1} + r^m n_0 + d[(r+s)^m - r^m] + \frac{\varepsilon - s^-(b^-)}{r^{-m}\lambda_{b^-+1}} - \bar{\delta} && (5.85)\\
&= r^m k + r^m\delta - \delta + d[(r+s)^m - r^m]. && (5.86)
\end{aligned}$$

Since $\dim \sigma = (r+s)^m \dim \rho = d(r+s)^m$, we finally obtain for the entropy:

$$\begin{aligned}
S_0^\varepsilon(\sigma) &= \log\left(d(r+s)^m - r^m k - d[(r+s)^m - r^m] - r^m\delta - \bar{\delta}\right) && (5.87)\\
&= \log r^m + \log\left(d - k - \delta + r^{-m}\bar{\delta}\right) && (5.88)\\
&= \log r^m + \log\left((d-k)\left[1 + \frac{r^{-m}\bar{\delta} - \delta}{d-k}\right]\right) && (5.89)\\
&= \log r^m + S_0^\varepsilon(\rho) + \log\left(1 + \frac{r^{-m}\bar{\delta} - \delta}{d-k}\right), && (5.90)
\end{aligned}$$

where the last term on the right hand side is only exponentially small in m:

$$\begin{aligned} \log\left(1+\frac{r^{-m}\bar{\delta}-\delta}{d-k}\right) &\leq \frac{r^{-m}\bar{\delta}-\delta}{d-k} && (5.91)\\ &\leq \frac{1}{r^m(d-k)} && (5.92)\\ &\leq r^{-m} \quad \text{for } \varepsilon<1, && (5.93)\end{aligned}$$

where we have used that $\bar{\delta}<1$ and $\delta<1$. To conclude, for two density matrices σ and ρ related by (5.79), we have found that

$$S_0^\varepsilon(\sigma) \approx S_0^\varepsilon(\rho) + \log r^m, \tag{5.94}$$

where the approximation is up to r^{-m}.

Exact additivity of S_2^ε

The general solution for S_2^ε is given by Eq. (5.78): $S_2^\varepsilon(\rho) = -\log(\sum_{i=1}^{b^+-1} n_i\lambda_i^2 + \sum_{i=b^+}^{d} n_i x_{\max}^2)$, with $x_{\max} = \lambda_{d-b^+} - (\varepsilon - s^+(b^+))/\sum_{i=0}^{b^+} n_{d-i}$, where $s^+(r) := \sum_{i=1}^{r} n_{d-i+1}(\lambda_{d-i+1} - \lambda_{d-r})$ and $b^+ := \max\{r : s^+(r) \leq \varepsilon\}$. Denote by $\bar{x}_{\max}$, $\bar{s}^+(r)$, and $\bar{b}^+$ all these quantities referring to σ. We now calculate

$$\begin{aligned} \bar{s}^+(r) &= \sum_{i=1}^{r} \bar{n}_{d-i+1}(\bar{\lambda}_{d-i+1} - \bar{\lambda}_{d-r}) && (5.95)\\ &= \sum_{i=1}^{r} r^m n_{d-i+1}(r^{-m}\lambda_{d-i+1} - r^{-m}\lambda_{d-r}) && (5.96)\\ &= s^+(r) && (5.97)\end{aligned}$$

which also directly implies that $\bar{b}^+ = b^+$ and $\bar{x}_{\max} = r^{-m}x_{\max}$. Finally, we find for the entropy:

$$\begin{aligned} S_2^\varepsilon(\sigma) &= -\log\left(\sum_{i=1}^{\bar{b}^+-1} \bar{n}_i\bar{\lambda}_i^2 + \sum_{i=\bar{b}^+}^{d} \bar{n}_i\bar{x}_{\max}^2\right) && (5.98)\\ &= -\log\left(r^{-m}\sum_{i=1}^{b^+-1} n_i\lambda_i^2 + r^{-m}\sum_{i=b^+}^{d} n_i x_{\max}^2\right) && (5.99)\\ &= S_2^\varepsilon(\rho) + \log r^m, && (5.100)\end{aligned}$$

which is very similar to the result obtained for S_0^ε, Eq. (5.94).

To conclude this section, we have shown that for density matrices of the special form (5.79), the smooth Renyi entropies S_0^ε and S_2^ε behave (for the former case at least approximately) additive.

Chapter 6

Finite Key Analysis for the Tomographic Protocol

In this chapter we present the main results of this thesis: An application of the analysis of privacy amplification to a specific quantum key distribution protocol, the so-called "Tomographic Protocol". Part of the results presented here can be found in [62]. The emphasis lies on analyzing the success of two-universal hashing for *finite* block sizes n. We will not deal with any finite-size effects that may emerge from statistical issues such as estimating an error rate in the parameter estimation step.[1] Additionally we assume that the adversary performs a collective eavesdropping attack (cf. Ch 4.3).[2] Our goal is to calculate the maximally achievable key rate of this protocol, that is the fraction of signals (after the sifting step) that make it into the final key. This number is of interest because of two reasons: First, it can be viewed as a measure of the efficiency of the privacy amplification step, and second Alice and Bob can use this information to determine how "much" privacy amplification they need to apply, because the key rate is equal to the factor by which the raw key needs to be shrunken in order to have the desired security.

Besides the dependence on the block size n, we present a detailed analysis of the dependence of the key rate on a number of different parameters: We show that for a security parameter ε (which is the probability that the key is insecure) of to 10^{-28}, one can still create a secret key. Moreover, we will see that the dependence on ε becomes less pronounced as the block size n increases, which means that one can create keys with

[1] We will also assume that optimal error correction protocols are employed.

[2] Is has been shown in [63] that the analysis of collective attacks is sufficient, since the most general attacks, coherent attacks, are not more powerful, at least in the limiting case of $n \to \infty$. Note however that we are explicitly avoiding this approximation.

higher security by increasing the block size. We are also able to compare variants of the Tomographic Protocol using different alphabet sizes (and quantum systems of corresponding dimension) and show for instance that for higher error rates, low dimensions are always favorable.

The first part of our analysis will deal with the ideal case in which Alice sends a single copy of each signal state to Bob, as she would do if she possessed a single-photon source. We will first derive all results for this scenario and then show in the second part how one can include also the case where Alice sends (accidentally) more than one copy of each state into the analysis. This is the situation one naturally encounters in most experimental implementations.

This chapter is organized as follows: The specifics of the Tomographic Protocol are presented in Sec. 6.1 and in Sec. 6.2 we derive the maximally achievable key rates for this protocol, using the results from Ch. 5, at first for the idealized case where in the quantum part of the protocol only single copies of the signal states are prepared (Sec. 6.3). This assumption will be loosened in Sec. 6.4, where we generalize the analysis to incorporate multi-photon events.

6.1 Description of the Protocol

The Tomographic Protocol was originally introduced in [2] (see also [3]). It is based on the Six-State Protocol [4] and generalizations thereof to higher dimensions. The main twist of the Tomographic Protocol is its special strategy to do parameter estimation: Since the measurements involved in the six state protocol form a tomographically complete POVM, it is (at least in principle) possible to reconstruct the density matrix of the measured state. Alice and Bob then use this knowledge to rigorously abort the protocol if they learn that the quantum state they shared is not compatible with some channel model (the depolarizing channel) they are expecting. In this way, they force the adversary to launch an attack that provides Alice and Bob with the same quantum state as they would obtain if they were only connected by a depolarizing channel, i.e. Eve has to perform a symmetric attack.

Such a symmetric attack has the advantage that the analysis of the privacy amplification step becomes simple, since the density matrices involved have only a few different eigenvalues (due to the high symmetry).

In more detail, the protocol works as follows: Alice and Bob use an alphabet of size d and consequently d-dimensional quantum systems for the encoding of the information. Some of the results we derive only hold for the case of qubits, i.e. $d = 2$, but when not

mentioned explicitly, any dimension is valid. We assume that the privacy amplification protocol outputs a string of bits, a binary key. As already mentioned, Alice chooses the encodings (cf. Sec. 3.2) such that they form a set of mutually unbiased bases [64, 42], that is, we have $d+1$ different encodings (or "bases") for the signal states: For each $d$it value $x \in 0, 1, \ldots, d-1$ and encoding $j = 1, 2, \ldots, d+1$, we denote the signal state by $|\phi_j^x\rangle$. They have the following properties: $\langle\phi_j^x|\phi_j^y\rangle = \delta_{xy}$ (for each encoding, they form a basis) and $\langle\phi_i^x|\phi_j^y\rangle = 1/\sqrt{d}$ for $i \neq j$ (the basis vectors from different bases are "mutually unbiased"). It was shown in [65, 66] that such a set of bases exists at least for the case of prime dimensions. The measurement performed by Bob (cf. Sec. 3.3.1) is given by the POVM $\mathcal{M}^k = \{|\phi_k^0\rangle\langle\phi_k^0|, |\phi_k^1\rangle\langle\phi_k^1|, \ldots, |\phi_k^{d-1}\rangle\langle\phi_k^{d-1}|\}$ for some randomly chosen $k = 1, 2, \ldots, d+1$. This measurement is said to be tomographically complete [39], which means that from the measurement statistics obtained $P^{\rho}_{\mathcal{M}^k}$, it is possible to infer the measured state ρ. We assume that Alice chooses each $d$it x with equal probability $1/d$, and likewise, Alice and Bob choose each encoding and measurement, respectively, with equal probability $1/(d+1)$.

We are now turning to the entanglement-based version (cf. Sec. 3.2.2 and 3.2.3) of this protocol in which Alice prepares a maximally entangled state in d dimensions, $|\phi_d^+\rangle = \sum_{x=0}^{d-1} |xx\rangle/\sqrt{d}$ and sends the second half of it to Bob. Due to Eve's interaction, it gets disturbed, and we denote the quantum state they share at this point by ρ_{AB}. So far we have described the entanglement-based version as Alice applying the modified encoding operators

$$A^{j\mathrm{T}} = \sum_{x=0}^{d-1} |x\rangle\langle\phi_x^j| \tag{6.1}$$

to her part of the state ρ_{AB} followed by a measurement in the computational basis $\{|x\rangle\}$ by Alice. Since the signal state $|\phi_x^j\rangle$ form a basis for each j, the application of $A^{j\mathrm{T}}$ corresponds to a basis rotation of system A. Instead of performing the rotation and measuring in the computational basis, Alice can equivalently directly measure in the $\{|\phi_x^j\rangle\}$-basis. This means that now Alice and Bob both perform the same tomographically complete measurement on their respective subsystems. We denote the outcome of this measurement by random variables X and Y, respectively. By classical communication, Alice and Bob can now check on which state ρ_{AB} they actually performed their measurements.

When we are talking about state tomography, we are usually working in a scenario where the same measurements are applied to identical copies of quantum states. In particular, the number of copies involved can be very large [67], which is an important issue for the performance of a QKD protocol. However, we will take the idea that

Alice and Bob have the possibility to reconstruct their shared quantum state as nothing more than an impetus for an assumption of the eavesdropping attack, namely that it is symmetric. We will not actually assume that Alice and Bob could do an efficient state tomography on the state ρ_{AB}, rather, Eve is restricted to an attack that leads to a specific form for ρ_{AB} which Alice and Bob could *in principle* check to probe for an eavesdropper.

At this point, we have to introduce the multi signal picture. So far we have considered measurements on the states ρ_{AB} corresponding to one signal sent from Alice to Bob, and one $d$it X and Y of data. Consider now the case where Alice sends n' signals to Bob and that the eavesdropper launches a collective attack (cf. Sec. 4.3), that is, for each half of $|\phi_d^+\rangle$ Alice sends to Bob, Eve performs the same unitary transformation. Moreover, we are assuming that Eve holds a purifying system ρ_E of each state ρ_{AB} shared by Alice and Bob after her attack (cf. Sec. 4.4.1), i.e. there exists some pure state $|\Psi\rangle_{ABE}$ with $\rho_E = \mathrm{tr}_{AB}\,|\Psi\rangle\langle\Psi|$ and $\rho_{AB} = \mathrm{tr}_E\,|\Psi\rangle\langle\Psi|$ for each signal. This implies that if Alice sends n' times half of the state $|\phi_d^+\rangle$ to Bob, they will share the state $\rho_{AB}^{\otimes n'}$ in the end, where Eve holds the purifying system of each state in the tensor product. Because of the tensor structure of the quantum state $\rho_{AB}^{\otimes n'}$, for most of the analysis it is sufficient to consider the single states ρ_{AB}. Whenever we are referring to a quantum state depending on a string of $d$its described by a random variable $\mathbf{X}$, we will use the notation $\rho_{\mathbf{X}E}$ to avoid confusion.

In the parameter estimation step (cf. Sec. 3.3.2), Alice and Bob use part of the data they obtained by the tomographic measurements to verify that the quantum state ρ_{AB} they shared prior to their measurement is of the form

$$\rho_{AB}^{\mathrm{dep}}(\beta_0, \beta_1) := (\beta_0 - \beta_1)|\phi_d^+\rangle\langle\phi_d^+| + \frac{\beta_1}{d}\mathbb{1} \otimes \mathbb{1}, \tag{6.2}$$

where we have adopted the notation of [2]. If it is not of this form, they abort the protocol. The state $\rho_{AB}^{\mathrm{dep}}(\beta_0, \beta_1)$ is the result of $|\phi_d^+\rangle$ passing through a depolarizing channel [67]. The two parameters are not independent, but fulfill the normalization condition $\beta_0 + (d-1)\beta_1 = 1$. When measuring $\rho_{AB}^{\mathrm{dep}}(\beta_0, \beta_1)$ in the same basis, the probability of Alice and Bob obtaining the same outcome is β_0, whereas the probability of obtaining two particular different outcomes is β_1. After Alice and Bob have verified that they shared the state (6.2), they discard all instances where they chose a different encoding for their measurement or where they got an inconclusive measurement outcome. They will be left with classical data $\mathbf{X}$ and $\mathbf{Y}$, respectively, which are $d$it strings of length $n < n'$. The error rate in this sifted key (for $d = 2$ this is the QBER) is given by $1 - \beta_0 = (d-1)\beta_1$ (strictly speaking, only in the limit of $n \to \infty$). We assume that

$0 \leq \beta_1 < 1/d < \beta_0 \leq 1$, since $\beta_0 = \beta_1 = 1/d$ corresponds to case of no correlations in the state ρ_{AB}.

We skip the pre-processing step (cf. Sec. 3.3.3) and turn directly to information reconciliation and privacy amplification protocols, which will be described in more detail in the next section.

6.2 Privacy Amplification for Finite Block Size

In this section, we analyze the success of the privacy amplification protocol described in Sec. 5.2, using the notion of ε-security (cf. Sec. 4.2). There are two free parameters that can be chosen at will by Alice and Bob, namely the block size n on which the PA protocol works and the security parameter ε that provides a measure of how secure the final key will be.

At this point of the QKD protocol, Alice and Bob hold classical data described by random variables $\mathbf{X}$ and $\mathbf{Y}$, respectively, with range $\mathcal{X} = \mathcal{Y} = \{0, 1, \ldots, d-1\}^n$. Their correlation can be easily obtained from Eq. (6.2) and it is given by the probability distribution

$$P_{\mathbf{XY}}(\mathbf{x}, \mathbf{y}) = \prod_{i=1}^{n} \beta_1 + \delta_{x_i y_i}(\beta_0 - \beta_1). \tag{6.3}$$

Since $\mathbf{X}$ and $\mathbf{Y}$ originate from measuring a tensor product state $\rho_{AB}^{\mathrm{dep}}(\beta_0, \beta_1)^{\otimes n}$, this probability distribution also factors: $P_{\mathbf{XY}}(\mathbf{x}, \mathbf{y}) = P_{XY}(x_1, y_1) \cdots P_{XY}(x_n, y_n)$, with

$$P_{XY}(x, y) = \beta_1 + \delta_{xy}(\beta_0 - \beta_1). \tag{6.4}$$

As already mentioned in Sec. 3.3.4 and 5.2, we assume that Alice and Bob employ an optimal error correction protocol which needs to communicate a string of length $H_0^{\varepsilon}(\mathbf{X}|\mathbf{Y})$ (in the limit of $n \to \infty$) over the public channel. After the error correction, Bob will be able to guess the string $\mathbf{X}$ from his data $\mathbf{Y}$ and the error correction information with probability of at least $1 - \varepsilon$.

In order to obtain the information the eavesdropper has about the key, we need to consider the purification of the state (6.2), which was shown in [3] to be

$$|\Psi\rangle_{ABE} = \sqrt{\frac{\beta_0}{d}} \sum_{x=0}^{d-1} |xx\rangle_{AB} |E_{xx}\rangle_E + \sqrt{\frac{\beta_1}{d}} \sum_{x \neq y} |xy\rangle_{AB} |E_{xy}\rangle_E, \tag{6.5}$$

where Eve's states $|E_{xy}\rangle$ are orthogonal to all other states for $x \neq y$ and $\langle E_{xx}|E_{yy}\rangle = 1 - \beta_1/\beta_0$ for $x \neq y$. From the purification (6.5), we can easily obtain the state

$$\rho_E^{xy} = {}_{AB}\langle xy|\Psi\rangle_{ABE} \tag{6.6}$$

held by the eavesdropper, when Alice and Bob obtain the measurement result x and y, respectively.

Finally, consider all possible measurement outcomes x and y for a single signal, which occur with probability $P_{XY}(x,y)$. The ccq-state describing these outcomes together with the adversary's quantum state ρ_E^{xy} is given by

$$\rho_{XYE} = \sum_{x,y} P_{XY}(x,y)|x\rangle\langle x| \otimes |y\rangle\langle y| \otimes \rho_E^{xy}. \tag{6.7}$$

Since after the error correction, Bob has corrected his $d$it y to Alice's $d$it x, we are only interested in Eve's correlation with the final key $d$it x. Therefore, we need to consider the state $\rho_{XE} = \mathrm{tr}_Y \rho_{XYE}$, which is given by

$$\rho_{XE} = \sum_{x,y} P_{XY}(x,y)|x\rangle\langle x| \otimes \rho_E^{xy}. \tag{6.8}$$

Considering all n signals, because of the tensor product structure of the shared state, the final cq-state $\rho_{\mathbf{X}E}$ describing Eve's correlation with the whole n-$d$it key is also of a simple tensor structure:

$$\rho_{\mathbf{X}E} = \left[\sum_{x,y} P_{XY}(x,y)|x\rangle\langle x| \otimes \rho_E^{xy}\right]^{\otimes n} \tag{6.9}$$

We can now employ the formula for the secret key length which was derived in Sec. 5.2:

$$\ell = S_2^{\varepsilon'}(\rho_{\mathbf{X}E}) - S_0^{\varepsilon'}(\rho_E) - H_0^{\varepsilon}(\mathbf{X}|\mathbf{Y}) + 2\log(\varepsilon). \tag{6.10}$$

where $\varepsilon' = (\varepsilon/8)^2$ and ε is the security parameter that quantifies how secure the final key will be (cf. Sec. 4.2). The secret key *rate* is given by ℓ/n, where n is the length of the string $\mathbf{X}$ on which the PA protocol is working. Note that this length is much shorter than the initial number of signals sent from Alice to Bob, which was denoted by n', since during the parameter estimation and sifting step, a large number of signals need to be discarded. For instance, even if no parameter estimation would be carried out, n would only be of the order of $n'/(d+1)$, since the protocol uses $d+1$ different encodings. But since we are only interested in the PA step, we take n to quantify the input resources, whereas ℓ quantifies the length of the output.

In Sec. 5.3.3, we have derived an explicit formula for the Renyi entropies occurring in Eq. (6.10) in terms of eigenvalues and probabilities of $\rho_{\mathbf{X}E}$, $\rho_E = \mathrm{tr}_{\mathbf{X}} \rho_{\mathbf{X}E}$, and $P_{\mathbf{X}|\mathbf{Y}=\mathbf{y}_0}$ for $\mathbf{y}_0 \in \{0,1,\ldots,d-1\}^n$ chosen arbitrarily. The eigenvalues of $\rho_{\mathbf{X}E}$ and ρ_E can be

easily obtained using Eqs. (6.9) and (6.6), yielding

$$\rho_{\mathbf{X}E}: \qquad \lambda_0 = 0, \qquad n_0 = d^{3n} - d^{2n} \tag{6.11}$$

$$\lambda_{i+1} = \tfrac{1}{d^n}\beta_0^i\beta_1^{n-i}, \qquad n_{i+1} = d^n\binom{n}{i}(d-1)^{n-i} \tag{6.12}$$

$$\rho_E: \quad \lambda_i = \left(\beta_0 - \beta_1 + \tfrac{\beta_1}{d}\right)^i\left(\tfrac{\beta_1}{d}\right)^{n-i}, \quad n_i = \binom{n}{i}(d-1)^{n-i}, \tag{6.13}$$

for $0 \leq i \leq n$. The probabilities occurring in the conditioned distribution $P_{\mathbf{X}|\mathbf{Y}=\mathbf{y}_0}$ can be calculated from Eq. (6.3), and they are given by

$$P_{\mathbf{X}|\mathbf{Y}=\mathbf{y}_0}: \quad p_i = \beta_0^i\beta_1^{n-i}, \quad n_i = \binom{n}{i}(d-1)^{n-i}, \tag{6.14}$$

for $0 \leq i \leq n$. The results obtained by calculating the right hand side of Eq. (6.10) are given in the next section.

6.3 Results for Single-Copy Signal States

Let us recall the main features of the special protocol we are investigating, the security assumptions we made, and which parameters we have to consider: The results presented in this section are only valid for collective eavesdropping attacks and a specific QKD protocol using d-dimensional quantum systems, in which the honest parties have verified in the parameter estimation step that prior to their measurements, they share n' copies of the state (6.2). This state has one free parameter, the error rate in the sifted key $1 - \beta_0$. (For qubits, $1 - \beta_0$ equals the QBER Q.) The PA protocol is carried out on n out of n' signals; the rest gets discarded. The number n can be adjusted by Alice and Bob by simply choosing a different number of input signals n', and we will call n the block size in the PA protocol. Another free parameter besides d and n is the security parameter ε, which Alice and Bob can choose at will to have the PA protocol output a more secure or less secure key, according to the security definition 4.2.1. Moreover, the results presented in this section only hold for the idealized case in which Alice sends only a single copy of each signal state to Bob, which corresponds to the case of employing a single-photon source (the generalization to multi-photon events will be the topic of the next section).

We will compare our results with those found in [1], where the authors calculated the right hand side of Eq. (6.10) for the case of infinite block sizes n, i.e. $r_\infty = \lim_{n\to\infty} \ell/n$. These results also hold for the more general case of coherent eavesdropping attacks. In the first part of this section, we restrict ourselves to the qubit case, because it is relevant

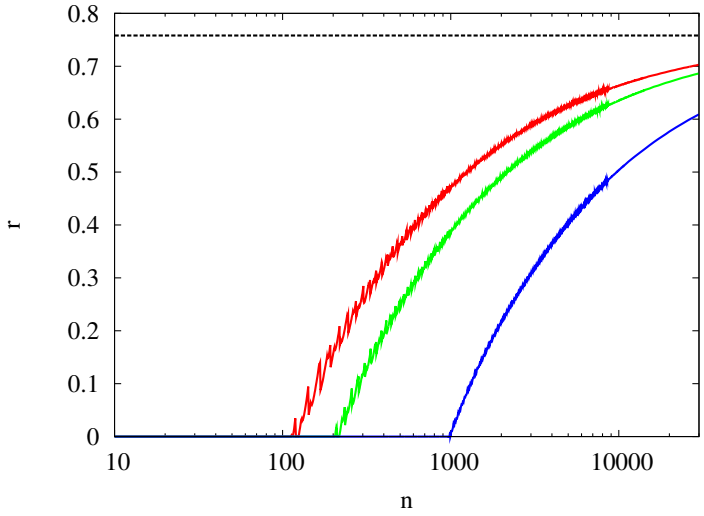

Figure 6.1: (Color online.) Achievable key rate r versus block length n for a QBER of $Q = 0.02$ and three different (arbitrarily chosen) values of the security parameter $\varepsilon = 0.1, 0.01, 10^{-10}$ (from top to bottom). The dashed line at $r_\infty = 0.758059$ is the asymptotic result for $n \to \infty$, obtained in [1]. The dimension of the quantum systems sent is taken to be $d = 2$.

for photon implementations. Different dimensions d of the quantum systems are covered in the second part of this section.

Dependence on n, Q, and ε

Fig. 6.1 shows a plot of the achievable key rate $r := \ell/n$ for the qubit case ($d = 2$) and a quantum bit error rate (QBER $Q = 1 - \beta_0$) of $Q = 0.02$ (a value that is also found in current experimental realizations). The security parameter ε, which measured the optimality of the key (cf. Def. 4.2.1) is chosen somewhat arbitrarily between 0.5 and 0.01 to illustrate the effect of this parameter on the key rate. A detailed discussion on the parameter ε can be found below. The achievable key rate apparently converges towards the asymptotic value r_∞ very slowly, that is, on a logarithmic scale in n, r still only grows sub-logarithmic, as one can see from a fitting. For very small block sizes (e.g. of the order of 10^2 to 10^3), we find a considerable derivation from the asymptotic value. However, for a moderate block size of the order of 10^4, we already obtain over 84% of r_∞, even for $\varepsilon = 0.01$. The quantitative dependence of the key rate on the security parameter is two-fold: First, for a smaller ε (i.e. higher security) the obtainable key rate gets smaller, as one would suspect. Second, the larger the block size, the smaller

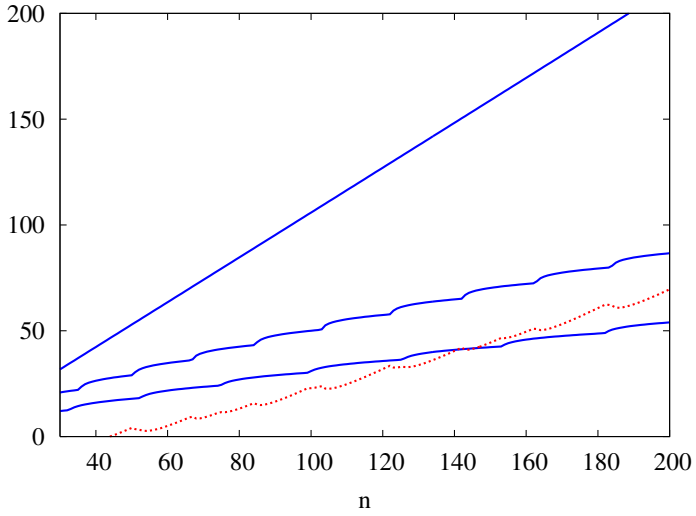

Figure 6.2: (Color online.) The three entropies constituting the achievable key length ℓ (dotted line), $S_2^{\varepsilon'}$, $S_0^{\varepsilon'}$, and $H_0^{\varepsilon'}$ (from top to bottom), depicted by solid lines. The QBER is chosen to be $Q = 0.02$ and $\varepsilon = 0.5$, with $\varepsilon' = (\varepsilon/8)^2$.

is the dependence on ε. This observation can be made from Fig. 6.1 by realizing that the curves for different ε approach each other. Such a result was also found analytically in [1] for the limit $n \to \infty$.

A prominent feature of the results presented in Fig. 6.1 are the "oscillations" of the achievable key rate, the amplitude of which decreases as n increases. Analytically, the oscillations arise from the structure of ℓ given in Eq. (5.12), being the difference of the three monotonic functions $S_2^{\varepsilon'}$, $S_0^{\varepsilon'}$, and $H_0^{\varepsilon'}$ where the last two are smoothened versions (see Fig. 6.2) of a non-continuous function. In the limit $n \to \infty$, the non-continuities disappear, leading to a monotonic key rate. Up to now, we can give no physical explanation for the non-monotonicity, besides the fact that our formula is just an achievable key rate and we disregarded the classical pre-processing step in our analysis, thus the key rate might also increase in some cases.

An important figure of merit for a QKD protocol is its *threshold QBER* (or tolerable error rate), which is the largest quantum bit error rate for which is the key rate is still non-zero. For the Six-State Protocol without pre-processing, it was found that a non-vanishing key rate is obtained whenever $Q \leq 0.126$ [68]. Using degenerate codes, which can be interpreted as some form of pre-processing [1], this bound can be slightly improved to $Q \leq 0.127$ [68]. For the "noisy" pre-processing, i.e. where Alice chooses $\mathbf{U}$ to be a noisy version of $\mathbf{X}$ (cf. Sec. 3.3.3), the bound can be further improved to

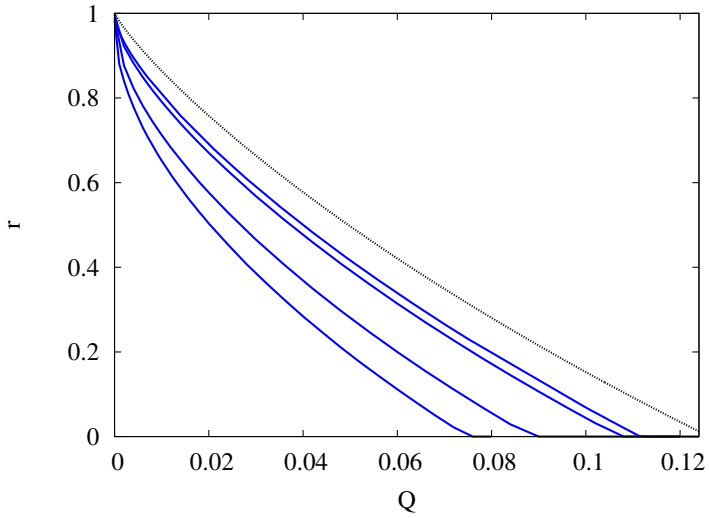

Figure 6.3: (Color online.) Key rate r versus QBER Q, for $d = 2$ and different block sizes n and security parameters ε: $n = 20,000$, $\varepsilon = 10^{-1}$; $n = 20,000$, $\varepsilon = 10^{-2}$, $n = 20,000$, $\varepsilon = 10^{-10}$, and $n = 10,000$, $\varepsilon = 10^{-10}$ (solid lines from top to bottom). The dotted line is the asymptotic result r_∞ found by [1].

$Q \leq 0.141$ [1]. In order to display the effect of finite block sizes n on the threshold QBER, in Fig. 6.3 we plot the achievable key rate r as a function of the QBER Q, for different security parameters ε.

Up to this point, we have not specified a reasonable range for the security parameter. Recall that a key being ε-secure means that it is perfectly secure except with probability ε. Suppose that the tomographic protocol is implemented in a QKD solution, i.e., a pair of black boxes, connected by a fiber, supplying the user with secret keys. The user might be interested in the long term application in which the apparatus generates secret keys of an accumulated length of N bits, for instance, $N = 10^{12}$. These keys are of length n, which implies that N/n keys are to be generated. Among all the N/n keys, we might require the probability of any of these keys being corrupted to be "very small", i.e. $N\varepsilon/n \ll 1$. This means that for block sizes of the order of $n = 10^5$, we need to consider security parameters that fulfill $\varepsilon \ll 10^{-7}$. Since as of today, apparently no rigorous theory exists about what range of ε has to be considered, we investigate a large range of ε and the limit of $\varepsilon \to 0$:

For $\varepsilon = 0$, all smooth Renyi entropies appearing in Eq. (5.12) reduce to the conventional Renyi entropies which are additive, i.e. $S^0_\alpha(\rho^{\otimes n}) = nS^0_\alpha(\rho)$ and similar for H^ε_0. We can therefore calculate all entropies and thus also the secret key length ℓ analytically

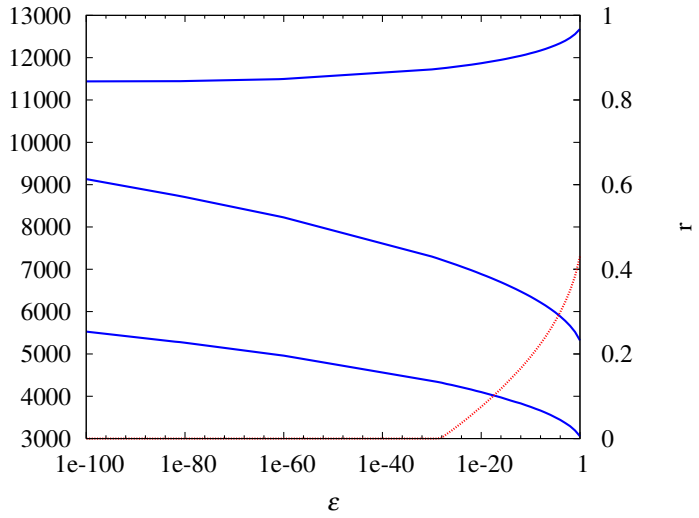

Figure 6.4: (Color online.) Plot of the entropies constituting the secret key length. The solid lines correspond to (from top to bottom): $S_2^{\varepsilon'}(\rho_{\mathbf{X}E})$, $S_0^{\varepsilon'}(\rho_E)$, and $H_0^{\varepsilon}(\mathbf{X}|\mathbf{Y})$, for $n = 10^4$, $d = 2$, and a QBER of $Q = 0.05$. The dotted line (plotted versus the right y-axis) corresponds to the secret key rate r.

in this case. In Fig. 6.4, we present the Renyi entropies and the secret key length for $\varepsilon \geq 10^{-100}$ for a block size of $n = 10^4$ and a QBER of $Q = 0.05$. For these values, we analytically find $S_2^0(\rho_{\mathbf{X}E}) = 1.14401n$, $S_0^0(\rho_E) = 2n$, and $H_0^0(\mathbf{X}|\mathbf{Y}) = n$, which shows in comparison with Fig. 6.4 that all entropies except S_2 approach their asymptotic value for $\varepsilon = 0$ rather slowly. (For $\varepsilon = 10^{-100}$ and $n = 10^4$, we find that $S_2^{10^{-100}}(\rho_{\mathbf{X}E}) = 114401$, $S_0^{10^{-100}}(\rho_E) = 9130.37$, and $H_0^{10^{-100}}(\mathbf{X}|\mathbf{Y}) = 5527.24$.)

Dependence on d

So far, we have only presented explicit results for the case where Alice and Bob use two-dimensional quantum systems in the distribution step. The Tomographic Protocol however is applicable for quantum systems of any (prime) dimension and our derivation of the results so far was completely general, without any assumptions about the dimensionality. The implementation of a QKD protocol using three-dimensional quantum states has already been demonstrated in [69]. As future implementations may even use higher-dimensional systems, it is worth to investigate the dependency of the privacy amplification protocol on the dimensionality, because it might reveal whether some dimension is preferable in terms of efficiency.

In order to make it possible to compare the results for different dimensions, we use

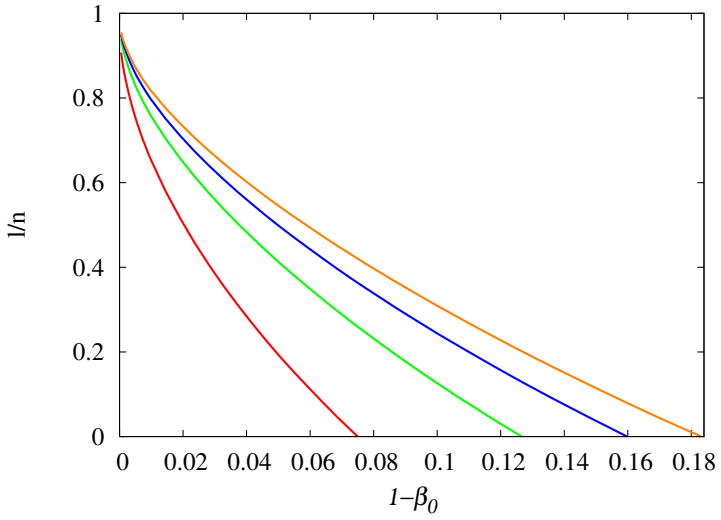

Figure 6.5: (Color online.) Key rate ℓ/n, measured in $d$its, for different dimensions of the quantum systems (from top to bottom: $d = 5, 4, 3, 2$) therefore also the alphabet size of the key, plotted versus the error rate per $d$it in the sifted key $1 - \beta_0$. The block size is $n = 10000$, and the security parameter is $\varepsilon = 10^{-10}$.

three approaches: First, we only look at the privacy amplification step, that is, we calculate the achievable key length ℓ measured in $d$its, for an n $d$it-string as input. The parameter $1 - \beta_0$ quantifies the fraction of erroneous $d$its in the sifted key (for $d = 2$, this is called "quantum bit error rate" (QBER) Q), thus it is well-suited for comparison of different dimensions. The result is depicted in Fig. 6.5, revealing that the privacy amplification step indeed becomes more efficient for higher-dimensional quantum states. Due to the complexity of the explicit formula (5.11) for the key length, we cannot tell whether this result comes form two-universal hashing being potentially more efficient for larger dimension, or the adversary being more restricted in her action and knowledge about Alice's and Bob's data.

This results does not take into account that also the other parts of the QKD protocol depend on the dimension d: Most importantly, since Alice and Bob use $d + 1$ different encodings in the Tomographic Protocol, the length of the sifted key n is on average a factor $1/(d+1)$ shorter than the number n' of signals sent. Thus, for higher-dimensional quantum systems, a larger fraction of the raw key gets discarded during the sifting (at least it is used for the tomography). To accommodate for this fact, in the second approach, we take $n' = (d + 1)n$ rather than n to quantify the number of input signals to the privacy amplification step and call the quantity ℓ/n' "effective key rate". From

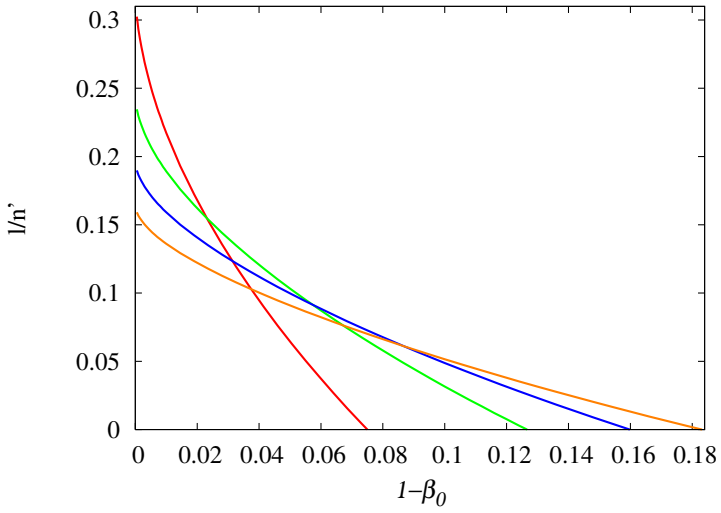

Figure 6.6: (Color online.) Effective key rate ℓ/n', measured in $d$its, for different dimensions of the quantum systems (left side from top to bottom: $d = 2, 3, 4, 5$), plotted versus the error rate per $d$it in the sifted key, $1 - \beta_0$. The block size is again $n = 10000$, with $n' = (d + 1)n$, and the security parameter $\varepsilon = 10^{-10}$.

Fig. 6.6 we see that due to the great overhead needed for higher dimensions, large dimensions do no longer provide the largest key rate for all error rates, as it was the case when we did not take into account the loss in the sifting step. Rather, for low error rates, we find the reverse result, that is, lower dimensions give the largest effective key rates. An interesting observation that can be made from Fig. 6.6 is that for each error rate, there exists an optimal dimension d for which the effective key rate becomes maximal. This is remarkable because of the great number of signals that get discarded for large dimensions, however the effective key rate is still larger than for small dimensions. Since in this approach only the key rate gets scaled, we also observe that the threshold error rate, i.e., the error for which the key rate vanishes, remains unchanged.

The third approach additionally takes into account that the creation of higher-dimensional quantum system uses more resources, quantified by the dimension of all quantum system combined. Let us clarify this by an example: Suppose that Alice would like to send an amount of ten bits of information to Bob. If we do not care about potential experimental problems which one might encounter, she could achieve this by either sending five qubits, or equally well two five-dimensional quantum states. Here, the sum of the dimensions of all quantum systems would be ten. We will now ask the question: Given a certain amount $\tilde{n}$ of quantum information to be sent in the distri-

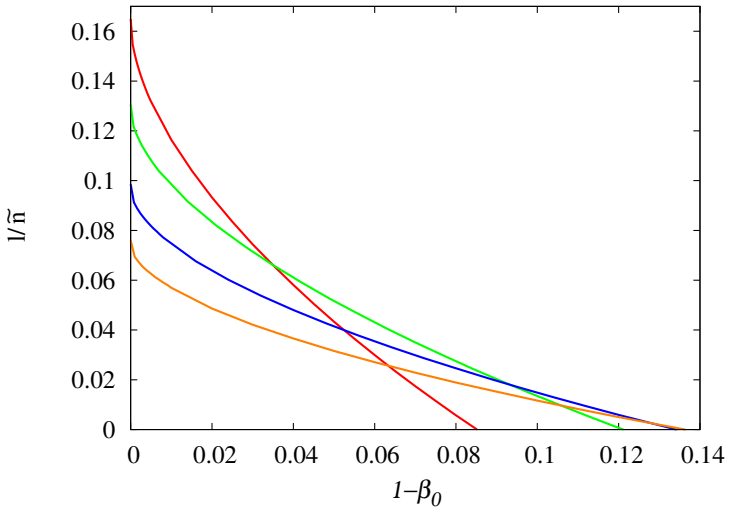

Figure 6.7: (Color online.) Effective key rate $\ell/\tilde{n}$, measured in bits, for different dimensions of the quantum systems (left side from top to bottom: $d = 2, 3, 4, 5$), plotted versus the error rate per $d$it in the sifted key, $1 - \beta_0$. We have fixed $\tilde{n} = 10^5$, which is the total dimension of all quantum systems sent, $\tilde{n} = dn' = d(d+1)n$, and the security parameter is taken to be $\varepsilon = 10^{-10}$.

bution phase of the protocol, what dimension d of the quantum systems is preferable? In a way, we are looking for the optimal factorization of $\tilde{n}$ into d times n'. We are also taking in account the total number n' of systems that need to be sent such that after the sifting, n signals contribute to the secret key. This means that we are fixing $\tilde{n} = dn' = d(d+1)n$ and investigate the dependence of the achievable key rate $\ell/\tilde{n}$ as a function of d, for a fixed $\tilde{n}$. Since in this way the key rate is related to the quantity $\tilde{n}$ that already incorporates the dimension d, we have to measure it in bits. Although this protocol takes the total dimension of all quantum systems sent into account, from Fig. 6.7 we see that qualitatively, we find the same behaviour as for the case where we only took the sifting into account: For smaller error rates, lower dimensions yield larger effective key rates, whereas for larger error rates it is the opposite. In all results presented here, we have chosen the block size $n = 10000$ and the security parameter $\varepsilon = 10^{-10}$ arbitrarily (but still reasonable in view of a realistic implementation). It turns out however that a change of these parameters does not change the qualitative behaviour of the key rate.

To conclude the analysis of the influence of the dimension d (thus, the alphabet size that Alice and Bob use), we have shown that in the tomographic protocol, privacy

amplification is more efficient if greater dimensions are used. We cannot tell whether this is a consequence of two-universal hashing itself or because of the special restrictions the eavesdropper encounters due to the tomography. If we take into account that actually $(d+1)n$ signals need to be sent in the protocol because of the sifting and define the key rate with respect to this number, it turns out that for lower error rates, lower dimensions become favorable, whereas for higher error rates, higher dimensions yield larger key rates. Moreover, for each error rate there exists a certain dimension which yields the optimal key rate. Lastly, defining the error rate with respect to the total dimension $d(d+1)n$ of all signals sent (the "resources" needed to establish the key), we also find that smaller dimensions are favorable for lower error rates and larger dimensions for larger error rates.

6.4 Inclusion of Multi-Photon Events

In this chapter, we turn towards experimental realizations of QKD protocols and the security problems that naturally arise in this context. Since most of the experiments performed so far only employ two-dimensional quantum systems, we restrict our analysis to the qubit case ($d = 2$). A common experimental implementation is to use the polarization degree of freedom of photons as a qubit. Our discussion of the Tomographic Protocol so far was idealized in the sense that we assumed that Alice sends one of the signal states $\{|\phi_x^j\rangle\}$ to Bob. This corresponds to the case where Alice prepares exactly *one* photon in the corresponding polarization state and sends it to Bob, which is an unrealistic assumptions because single photon sources do not exist. In reality, Alice uses an attenuated laser beam with very low intensity such that the average photon number in each pulse is very low. The light pulse emitted by a laser is a coherent state

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{m=0}^{\infty} \frac{\alpha^m}{\sqrt{m!}} |m\rangle, \tag{6.15}$$

which is parameterized by some complex number α. Here, $|m\rangle$ is a Fock state. If the phase $\arg \alpha$ of this coherent state is randomized, we obtain the following state:

$$\rho = \int \frac{\mathrm{d} \arg \alpha}{2\pi} |\alpha\rangle\langle\alpha| = \sum_{m=0}^{\infty} P(m) |m\rangle\langle m|, \tag{6.16}$$

which is a mixture of all Fock states wighted with a Poissonian distribution,

$$P(m) = e^{-|\alpha|^2} \frac{|\alpha|^{2m}}{m!}. \tag{6.17}$$

The average photon number of such a pulse is given by $\bar{m} = |\alpha|^2$. Common QKD implementations (using these so-called "weak coherent pulses") employ attenuated lasers where the average photon number of each pulse is of the order of 0.1. From Eq. (6.17), we can compute that such pulses are mostly empty, $P(0) \approx 0.905$, contain one photon with probability $P(1) \approx 0.09$ and more than one photon with probability $\sum_{m=2}^{\infty} P(m) \approx 0.005$.

Using such a weak coherent pulse-implementation, it follows that Alice does not send the signal states $|\phi_x^j\rangle$ to Bob, but rather an unknown number of copies, i.e. $|\phi_x^j\rangle\langle\phi_x^j|^{\otimes m}$ where the number of copies is distributed according to the Poissonian statistic (6.17). In this situation, the following eavesdropping attack (called "photon number splitting" (PNS) attack), which cannot be covered by our analysis so far, becomes imminent: Consider that Alice and Bob are connected by a lossy fiber and that Bob's photo detectors cannot resolve the photon number (i.e. they can only tell "there was no photon" or "there were some photons"), which is an assumption that holds in any real-world implementation. For each pulse sent from Alice to Bob, Eve measures its photon number via a non-demolition measurement[3] [70]. If it is more than one, she splits off one photon, stores it, and sends the remaining photons to Bob through a lossless fiber. If the pulse contains no photons, she forwards the vacuum pulse to Bob. If the photon number is one, she either completely blocks the pulse or performs some single-state attack as described in Sec. 4.3. She chooses one of these possibilities and the disturbance caused by the single-state attack in a way that the photon statistics that Bob expects to receive (he assumes a lossy channel having some yield γ, or attenuation $1-\gamma$) are maintained. It is easy to see that such a PNS attack always exists if the yield γ is smaller than the probability of emitting a multi-photon pulse [40]. In this attack, it is impossible for Alice and Bob to detect the eavesdropper, because their estimation of the channel yield is what they expect from a lossy fiber, but Eve has a copy of each signal state of the multi-photon events. If the QKD protocol employs a simple information reconciliation step, in which after the distribution, Alice and Bob talk about the encodings used, Eve can measure all her copies in the corresponding basis and obtains the same classical data as Alice for signals that were encoded into a multi-photon pulse.

However, under these circumstances, the QKD protocol is not completely insecure. As long as Alice and Bob can find a proper description of the total system, i.e. the ccq-state $\rho_{\mathbf{X}E}$ given by Eq. (6.9), the obtainable secret key length (6.10) can be calculated also for the weak coherent pulse implementation. The result will certainly be a smaller key length ℓ than which is obtained for the single-photon case, which tells Alice and

[3] Using such an operation Eve introduces no errors and thus cannot be detected.

Bob that they need to do "more" privacy amplification, thus sacrificing more raw key. In the next section, we show how we can find the ccq-state for a weak coherent pulse implementation of the Tomographic Protocol, under the assumptions of a collective eavesdropping attack (which is assumed throughout this chapter).

6.4.1 Concept

For the single-photon case, recall that we can express the preparation of the signal states $\{|\phi_x^j\rangle\}$ in the entanglement-based picture (cf. Sec. 3.2.2), which implies that for an encoding j chosen by Alice, she prepares

$$|\psi^j\rangle_{AB} = \sum_{x=0}^{1} \sqrt{P_X(x)}|x\rangle_A|\phi_x^j\rangle_B. \tag{6.18}$$

It is sufficient to consider some fixed encoding j, because we assume that Alice chooses them with equal probability and all signals in which Bob chooses a different one are discarded in the sifting step. In the previous section we have seen that in a weak coherent pulse implementation, Alice does not prepare the signal states $|\phi_x^j\rangle$, but rather a classical mixture of $|\phi_x^j\rangle\langle\phi_x^j|^{\otimes m}$ where m is distributed according to a Poissonian distribution[4] $P(m)$. We can incorporate this in Eq. (6.18) by introducing an auxiliary system R, which is neither controlled by Alice and Bob nor by Eve and which holds the photon number m, thus purifying the mixture $\sum_{m=0}^{\infty} P(m)|\phi_x^j\rangle\langle\phi_x^j|^{\otimes m}$:

$$|\psi^j\rangle_{ABR} = \sum_{m=0}^{\infty}\sum_{x=0}^{1} \sqrt{P(m)}\sqrt{P_X(x)}|x\rangle_A|\phi_x^j\rangle_B^{\otimes m}|m\rangle_R. \tag{6.19}$$

As we have seen in Sec. 4.3, the eavesdropper's attack is fully described by the attachment of a probe system $|0\rangle_E$ to the signal states underway to Bob and some unitary operation U_{BE} acting on these two systems, which implies that the state after Eve's attack is given by

$$|\psi^j\rangle_{ABER} = \sum_{m=0}^{\infty}\sum_{x=0}^{1} \sqrt{P(m)}\sqrt{P_X(x)}|x\rangle_A U_{BE}|\phi_x^j\rangle_B^{\otimes m}|0\rangle_E|m\rangle_R. \tag{6.20}$$

Since it is possible for Eve to find out the number of photons m in each signal without introducing any noise, she can actually choose a different unitary operation $U_{BE}^{(m)}$ depending on the photon number. In particular, for $m > 1$, she can apply a swap operation

[4] The Poissonian distribution is a special case that occurs for the weak coherent pulse implementation. Our analysis is valid for arbitrary probability distributions $P(m)$.

which supplies her with one of the signal states, and some operation $U_{BE}^{(1)}$ for the pulses containing exactly one photon, as already mentioned in the previous chapter.

Note that this eavesdropping attack is no longer a strictly collective attack, because we allow the eavesdropper to apply a different unitary operation $U_{BE}^{(m)}$ to each signal state depending on its photon number. However, this strategy is the most simple generalization of a collective attack to the multi-photon case, since $U_{BE}^{(m)}$ *only* depends on the photon number m, and we do not allow Eve to choose an arbitrary unitary operation for each signal. All signals containing the same photon number will be attacked in the same way.

Consider the case where Alice employs a source that emits single-photon pulses with a probability of p_{single} and multi-photon pulses with a probability of p_{multi}. During the sifting, Alice and Bob will discard all empty pulses, which makes it reasonable to define $\eta = p_{\text{single}}/(p_{\text{single}} + p_{\text{multi}})$ to be the rate of single-photon pulses among all non-empty pulses. Finally, we can define the number of single-photon pulses to be $n_{\text{s}} = \lfloor \eta n \rfloor$, and the number of multi-photon pulses to be $n_{\text{m}} = \lfloor (1-\eta) n \rfloor$, where n is total number of signals after the sifting.[5]

Let us now investigate the state ρ_{ABE}^{n} describing all n signals that survive the sifting step. If we introduce the operators $U_{BE}^{(m)}$ in Eq. (6.20) and project the system R onto a certain photon number m for each of the n signals, we can separate the total n-signal state between $m = 1$ and $m > 1$:

$$\begin{aligned}\rho_{ABE}^{n} = &\left[\left(\sum_{x=0}^{1} \sqrt{P_X(x)} |x\rangle_A U_{BE}^{(1)} |\phi_x^j\rangle_B |0\rangle_E\right) (\text{h.c.})\right]^{\otimes n_{\text{s}}} \\ &\otimes \left[\sum_{m=2}^{\infty} P(m) \left(\sum_{x=0}^{1} \sqrt{P_X(x)} |x\rangle_A U_{BE}^{(m)} |\phi_x^j\rangle_B^{\otimes m} |0\rangle_E\right) (\text{h.c.})\right]^{\otimes n_{\text{m}}} \qquad (6.21) \\ =: &\ |\Psi^{\text{single}}\rangle\langle\Psi^{\text{single}}|_{ABE}^{\otimes n_{\text{s}}} \otimes \rho_{ABE}^{\text{multi}\otimes n_{\text{m}}}, \qquad (6.22)\end{aligned}$$

where we have defined $|\Psi^{\text{single}}\rangle_{ABE}^{\otimes n_{\text{s}}}$ and $\rho_{ABE}^{n_{\text{m}}}$ for the single and multi-photon part, respectively. Note that Eqs. (6.21) and (6.22) have to be read modulo permutation of the n subsystems, because the order of the pulses is in general different. Also recall that we can fix some encoding j, because we are only looking at the signals which survive the sifting step, and for all those signals Alice and Bob have chosen the same encoding j.

For the multi-photon part $\rho_{ABE}^{\otimes n_{\text{m}}}$, we have already argued that Eve can split one signal off $|\phi_x^j\rangle\langle\phi_x^j|_B^{\otimes m}$ to store it until she learns its encoding j. Then she can measure

[5] We ignore possible rounding errors and assume that $n_{\text{s}} + n_{\text{m}} = n$.

it and obtains the same bit as Bob. Since we assumed that Eve replaces the lossy fiber between her and Bob by a noiseless one, it is sufficient to forward only one of the photons in the pulse to Bob, even if there were more than two. Using this strategy, the state $\rho_{ABE}^{n_{\mathrm{m}}}$ is actually also pure and of the form

$$|\Psi^{\mathrm{multi}}\rangle_{ABEE'}^{\otimes n_{\mathrm{m}}} = \left[\sum_{x=0}^{1} \sqrt{P_X(x)}|x\rangle_A U_{BE}^{(1)'}|\phi_x^j\rangle_B|0\rangle_E|\phi_x^j\rangle_{E'}\right]^{\otimes n_{\mathrm{m}}}, \tag{6.23}$$

where Eve is still free to choose some unitary operation $U_{BE}^{(1)'}$ on her auxiliary system E and the remaining photon that she will send on to Bob, and we have introduced a second system E' under Eve's control which stores the copy of the signal state. Comparing the multi-photon $|\Psi^{\mathrm{multi}}\rangle_{ABEE'}^{\otimes n_{\mathrm{m}}}$ and single-photon part $|\Psi^{\mathrm{single}}\rangle_{ABE}^{\otimes n_{\mathrm{s}}}$, we observe that they are essentially identical, the only difference being that for the multi-photon case, Eve has the additional advantage that she holds the system $|\phi_x^j\rangle_{E'}$. Moreover, the state $|\Psi^{\mathrm{single}}\rangle_{ABE}$ describes exactly the situation discussed in the previous section, which only dealt with the idealized single-photon case. We therefore make the same assumption as before, namely that Eve can obtain the purifying system of the state ρ_{AB} distributed between Alice and Bob, now for both the single and multi-photon part. In this way we can adopt our previous analysis for the state describing all n signals (after the sifting),

$$|\Psi^n\rangle_{ABE} = |\Psi^{\mathrm{single}}\rangle_{ABE}^{\otimes n_{\mathrm{s}}} \otimes |\Psi^{\mathrm{multi}}\rangle_{ABE}^{\otimes n_{\mathrm{m}}}, \tag{6.24}$$

where we only have to keep in mind that Eve also holds a copy of the signal state for each of the multi-photon pulses.

For each (non-empty) signal arriving at Alice and Bob, they measure a tomographically complete POVM, which makes it in principle possible to check for the "effectively" distributed single signal state

$$\rho_{AB}^{\mathrm{eff}} = \mathrm{tr}_E(\eta|\Psi^{\mathrm{single}}\rangle\langle\Psi^{\mathrm{single}}|_{ABE} + (1-\eta)|\Psi^{\mathrm{multi}}\rangle\langle\Psi^{\mathrm{multi}}|_{ABE}) \tag{6.25}$$

to be of a certain form. Since otherwise Alice and Bob will abort the protocol, we assume that Eve prepares the states $|\Psi^{\mathrm{single}}\rangle_{ABE}$ and $|\Psi^{\mathrm{multi}}\rangle_{ABE}$ such that

$$\rho_{AB}^{\mathrm{eff}} = \rho_{AB}^{\mathrm{dep}}(\beta_0, \beta_1) := (\beta_0 - \beta_1)|\phi^+\rangle\langle\phi^+| + \frac{\beta_1}{2}\mathbb{1}, \tag{6.26}$$

that is, the effectively distributed state is a depolarized version of the maximally entangled state $|\phi^+\rangle := \sum_{x=0}^{1}|xx\rangle/\sqrt{2}$, in the same way as in the previous section.

The measurements Alice and Bob perform are carried out on the effectively distributed state (6.25) which is the same as for the idealized single-photon case in the

previous section. Thus their correlation, described by $P_{\mathbf{XY}}$, given by Eq. (6.3), remains unchanged. To find the correlation between Eve and Alice's data $\mathbf{X}$, we need to construct the cq-state corresponding to Eq. (6.9) which we found in the previous section for the case of single-photon events.

In order to calculate the obtainable key length in a scenario including multi-photon events, we need to know the structure of the global state (6.24). But Alice and Bob can only find on which "effective" state (6.25) they performed their measurements. The global state can thus be any state that is compatible with Eq. (6.25) and Eq. (6.26). In the next two subsections, we present two different approaches for the determination of $|\Psi^n\rangle_{ABE}$. In Sec. 6.4.2, we are making the apparent assumption that the single-photon part $\mathrm{tr}_E |\Psi^{\mathrm{single}}\rangle_{ABE}$ and the multi-photon part $\mathrm{tr}_E |\Psi^{\mathrm{multi}}\rangle_{ABE}$ are themselves depolarized states (6.26). Thus also their convex combination via the parameter η in Eq. (6.25) is a depolarized state. Using this ansatz as a "first guess", we are able to calculate the secret key rate for this scenario. However, this approach is not the most general one, because there are many other states $|\Psi^{\mathrm{single}}\rangle_{ABE}$ and $|\Psi^{\mathrm{multi}}\rangle_{ABE}$ which have the desired feature that they lead to the depolarized state in the convex combination (6.25). One of these possibilities could be favorable for Eve, i.e. it could lead to a lower obtainable key rate for Alice and Bob. Therefore, in Sec. 6.4.3, we extend our considerations from Sec. 6.4.2 to a more general case. It will turn out that it is indeed not sufficient to consider the single- and multi-photon part to be the same depolarized state — there exists a more general strategy for Eve that leads to a lower key rate.

6.4.2 Symmetric Splitting

Fixing the form of the effectively distributed state (6.25) to be (6.26) leaves a lot of freedom in the structure of the states $|\Psi^{\mathrm{single}}\rangle_{ABE}$ and $|\Psi^{\mathrm{multi}}\rangle_{ABE}$; only their mixture has to be a depolarized maximally entangled state.

In this section, we restrict our attention to the case where both these states are purifications of the "same" depolarized state, i.e.,

$$\mathrm{tr}_E |\Psi^{\mathrm{single}}\rangle\langle\Psi^{\mathrm{single}}|_{ABE} = \rho_{AB}^{\mathrm{dep}}(\beta_0^{\mathrm{s}}, \beta_1^{\mathrm{s}}), \tag{6.27}$$

$$\mathrm{tr}_E |\Psi^{\mathrm{multi}}\rangle\langle\Psi^{\mathrm{multi}}|_{ABE} = \rho_{AB}^{\mathrm{dep}}(\beta_0^{\mathrm{m}}, \beta_1^{\mathrm{m}}), \tag{6.28}$$

only for different parameters $\beta_{0,1}^{\mathrm{s}}$ and $\beta_{0,1}^{\mathrm{m}}$. This assumption is an obvious choice which fulfills Eq. (6.25), which implies the parameters β_0 and β_1 measured by Alice and Bob

in the parameter estimation step are given by

$$\beta_0 = \eta\beta_0^{\mathrm{s}} + (1-\eta)\beta_0^{\mathrm{m}}, \tag{6.29}$$

$$\beta_1 = \eta\beta_1^{\mathrm{s}} + (1-\eta)\beta_1^{\mathrm{m}}. \tag{6.30}$$

Recall that the parameters β_0 and β_1 are not independent, rather, they are related by $\beta_0 + \beta_1 = 1$.

By Eqs. (6.27) and (6.28), the total state for both the single and the multi-photon part is the same as for the analysis of the idealized single-photon case in Sec. 6.2. For this case, we have already derived the cq-state describing Eve's correlation with Alice's data $\mathbf{X}$, and it is given by (cf. also Eq. (6.9))

$$\rho_{\mathbf{X}E} = \left[\sum_{x,y} P_{XY}(x,y)|x\rangle\langle x| \otimes \rho_E^{xy}\right]^{\otimes n_{\mathrm{s}}}, \tag{6.31}$$

with $P_{XY}(x,y) = \beta_1 + \delta_{xy}(\beta_0 - \beta_1)$ and where

$$\rho_E^{xy} = {}_{AB}\langle xy|\Psi^{\mathrm{single}}\rangle_{ABE}. \tag{6.32}$$

If we assume that Alice and Bob employ a simple sifting strategy in which they announce the encoding used for each pulse, Eve will know all key bits that are generated from the multi-photon part, because she then can measure the stored signals in the correct basis. This means that the cq-state for the multi-photon part is of the simple form $[1/2\sum_{x=0}^{1}|x\rangle\langle x| \otimes |x\rangle\langle x|_E]^{\otimes n_{\mathrm{m}}}$. It follows that for all n signals, the cq-state is given by

$$\sigma_{\mathbf{X}E} = \rho_{\mathbf{X}E} \otimes \left[\frac{1}{2}\sum_{x=0}^{1}|x\rangle\langle x| \otimes |x\rangle\langle x|_E\right]^{\otimes n_{\mathrm{m}}}, \tag{6.33}$$

again modulo permutation of the n subsystems (cf. discussion in Sec. 6.4.1).

What is now the maximally obtainable secret key length that can be extracted from $\mathbf{X}$ and $\mathbf{Y}$ in the realistic scenario described by Eq. (6.33)? One might expect that the only contribution to the secret key comes from the single-photon part,

$$\ell(\sigma_{\mathbf{X}E}) \stackrel{?}{=} \ell(\rho_{\mathbf{X}E}). \tag{6.34}$$

In order to answer this question, we need to evaluate $\ell(\sigma_{\mathbf{X}E})$, i.e. Eq. (6.10). Note that $\sigma_{\mathbf{X}E} = \rho_{\mathbf{X}E} \otimes (\mathbb{1}_r/r \oplus \mathrm{diag}(\underbrace{0,\ldots,0}_{s}))^m$, with $r = 2$, $s = 2^2 - 2$, and $m = n_{\mathrm{m}}$. We have calculated the Renyi entropies appearing in Eq. (6.10) for density matrices of this form in Sec. 5.3.4, and we already noticed that the correlations between $\mathbf{X}$ and $\mathbf{Y}$ remain

unchanged when we are taking multi-photon events into account. Thus we immediately obtain

$$\ell(\sigma_{\mathbf{X}E}) \approx \ell(\rho_{\mathbf{X}E}), \tag{6.35}$$

where the approximation is up to an additive constant of the order of $2^{-n_{\mathrm{m}}}$. This result tells us that all n_{m} multi-photon signals do not contribute to the key length at all, only the single-photon part is important. Due to Eq. (6.27), this part is described by $|\Psi^{\mathrm{single}}\rangle_{ABE}$, the purification of $\rho_{AB}^{\mathrm{dep}}(\beta_0^{\mathrm{s}}, \beta_1^{\mathrm{s}})$, which is the depolarized state given by Eq. (6.26). The amount of secret key that can be extracted from this state depends on the parameter β_0^{s} (or $\beta_1^{\mathrm{s}} = 1 - \beta_0^{\mathrm{s}}$). But Alice and Bob cannot measure this parameter directly, only the convex combination β_0 given by Eq. (6.29). They can therefore only infer a lower bound on the secret key length by making the most pessimistic assumption on β_0^{s} compatible with their measurement, which is $\beta_0^{\mathrm{m}} = 1$, meaning that most of the correlations they measured come from the multi-photon part (which does not contribute). This leads to the lowest possible value for β_0^{s}:

$$\beta_0^{\mathrm{s,min}} = \frac{1}{\eta}\left(\beta_0 - (1-\eta)\right), \tag{6.36}$$

or, in terms of quantum bit error rates,

$$Q^{\mathrm{min}} = \frac{1}{\eta} Q, \tag{6.37}$$

where $Q = 1 - \beta_0$ is the QBER Alice and Bob obtain in the parameter estimation step. From these considerations we also see that if $\beta_0 \leq 1-\eta \Leftrightarrow Q = 1-\beta_0 \geq \eta$, all correlations can be contributed to the multi-photon part, i.e. $\beta_0^{\mathrm{s}} = 0$ and $\ell = 0$. This is a well-known result [40, 41, 71], meaning that if the error rate in the sifted key is larger than the probability of creating a single-photon pulse, no secret key can be generated. Another simple upper bound on the tolerable QBER is obtained by the following consideration: Remember that the secret key is drawn entirely from the single-photon part (6.27), which we assumed to be the depolarized state (6.26). If this state is separable, no secret key can be distilled from it [22, 23]. The condition on the parameters for which this happens is $\beta_0 \leq 2/3$, i.e. $Q \geq 1/3$.

As a recap, if Alice and Bob run the Tomographic Protocol in a weak coherent pulse implementation with a fraction η of single-photon pulses among all n pulses after the sifting and QBER Q, they can expect to generate a secret key of the same length as if they ran the protocol using a single photon implementation with only $n_{\mathrm{s}} = \eta n$ signals and a larger QBER of Q/η. This result is visualized in Fig. 6.8. Note that if Alice and Bob would employ decoy pulses [40, 41], they might be able to learn more about

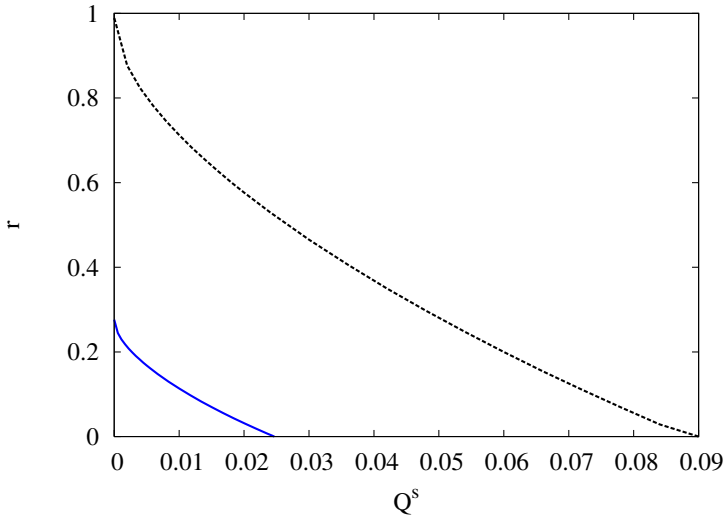

Figure 6.8: (Color online.) Secret key rate r versus the QBER of the single-photon part, $Q^{\mathrm{s}} = 1-\beta_0^{\mathrm{s}}$ for a weak-coherent pulse implementation of the Tomographic Protocol with a fraction $\eta = 0.279$ of single-photon events among all non-empty pulses (solid lines). This corresponds to an average photon number of 0.425 (chosen for comparison with Fig. 6.9). The dotted line corresponds to the single-photon implementation. We have chosen $n = 20,000$ and $\varepsilon = 10^{-10}$.

the parameters β_0^{s} and β_0^{m} (which ultimately describe the total system) than just their convex combination $\beta_0 = \eta\beta_0^{\mathrm{s}} + (1-\eta)\beta_0^{\mathrm{m}}$. This would allow them to estimate a tighter bound on the secret key rate. We address this issue in Sec. 6.4.4.

6.4.3 Asymmetric Splitting

In this section, we generalize the considerations of the previous section to more general cases where the states in Eq. (6.25) are not necessarily depolarized maximally entangled states. As we have already seen, only the single-photon part is important for generating the secret key, as Eve has full information about the multi-photon pulses. Thus, we will concentrate on $|\Psi^{\mathrm{single}}\rangle_{ABE}$ and choose $|\Psi^{\mathrm{multi}}\rangle_{ABE}$ such that (6.26) holds, for predefined values β_0 (Alice and Bob measure it) and η (determined by the experimental setup).

Our goal will be to show that the asymmetric splitting is advantageous for Eve, in the sense that there exists a larger range (than for the symmetric splitting) of the parameter β_0, detectable for Alice and Bob, for which no key agreement is possible.

We consider the following general[6] ansatz for $|\Psi^{\text{single}}\rangle_{ABE}$:

$$|\Psi^{\text{single}}\rangle_{ABE} = \sqrt{\alpha_1}|\phi^+\rangle|1\rangle + \sqrt{\alpha_2}|\phi^-\rangle|2\rangle + \sqrt{\alpha_3}|\psi^+\rangle|3\rangle + \sqrt{\alpha_4}|\psi^-\rangle|4\rangle, \tag{6.38}$$

with $\sum_{i=1}^{4}\alpha_i = 1$. The marginal state shared by Alice and Bob is thus $\rho_{AB}^{\text{single}} = \alpha_1|\phi^+\rangle\langle\phi^+| + \cdots + \alpha_4|\psi^-\rangle\langle\psi^-|$, which is separable (i.e., no secret key can be extracted) if and only if $\alpha_i \leq 1/2$ for all i.[7] Condition (6.25), evaluated for the marginal states implies that

$$\rho_{AB}^{\text{multi}} = \frac{1}{1-\eta}(\rho_{AB}^{\text{dep}}(\beta_0,\beta_1) - \eta\rho_{AB}^{\text{single}}), \tag{6.39}$$

where we again have assumed that the effectively shared state is the depolarized state (6.26). Inserting the explicit form of $\rho_{AB}^{\text{single}}$ and $\rho_{AB}^{\text{dep}}(\beta_0,\beta_1)$ into (6.39) yields

$$\begin{aligned}\rho_{AB}^{\text{multi}} &= \frac{1}{1-\eta}\left[\left(\frac{3\beta_0-1}{2} - \eta\alpha_1\right)|\phi^+\rangle\langle\phi^+| + \left(\frac{1-\beta_0}{2} - \eta\alpha_2\right)|\phi^-\rangle\langle\phi^-| \right. &(6.40)\\ &\left. + \left(\frac{1-\beta_0}{2} - \eta\alpha_3\right)|\psi^+\rangle\langle\psi^+| + \left(\frac{1-\beta_0}{2} - \eta\alpha_4\right)|\psi^-\rangle\langle\psi^-|\right]. &(6.41)\end{aligned}$$

The only assumption we make on ρ_{AB}^{multi} is its positivity, which yields the following bounds on the parameters α_i:

$$\alpha_1 \leq \frac{3\beta_0-1}{2\eta} \tag{6.42}$$

$$\alpha_i \leq \frac{1-\beta_0}{2\eta} \quad \text{for } i = 2,3,4 \tag{6.43}$$

A simple bound for the threshold QBER is given by the value of Q for which $\rho_{AB}^{\text{single}}$ becomes separable. (However, separability is only a sufficient conditions, i.e., there exist entangled states which also lead to a vanishing key rate.) Thus we look for the range of β_0 which allows $\alpha_i \leq 1/2$ for all i. The only α_i that could possibly be larger than $1/2$ is α_1, since $\alpha_i < 1/6$ for $i = 2,3,4$ is equivalent to $\alpha_1 = 1-\alpha_2-\alpha_3-\alpha_4 > 1/2$. From Eq. (6.43), we see that this happens if $(1-\beta_0)/(2\eta) > 1/6$, i.e. $1-\beta_0 > \eta/3$. In terms of the QBER Q, this reads $Q \geq \eta/3$, which needs to be contrasted to the bound $Q \geq 1/3$ for the symmetric splitting found in the previous section. This means that for the more general splitting of the effectively distributed state (6.25) considered in this section, we can derive a lower tolerable QBER than for the symmetric splitting. This tells us that the symmetric splitting assumption is not well-suited for a security analysis, because there exists a strategy for Eve (the asymmetric splitting) which provides Alice and Bob

[6] This is the most general ansatz because we have shown in Sec. 5.3.2 that the Renyi entropies $S_\alpha^2(\rho)$ only depend on the eigenvalues of ρ.

[7] This can be easily verified by consulting the PPT criterion [72].

with a possibly lower key rate for the same QBER. Thus we need to consider the case where the eavesdropper chooses this attack, which is favorable from her perspective.

To conclude, we have found that for an asymmetric splitting and a measured QBER of $Q \geq \eta/3$, there exists a separable state $\rho_{AB}^{\text{single}}$ that produces the depolarized state in the effective mixture (6.25), i.e., no secret key agreement is possible. For the symmetric splitting, the corresponding bound reads $Q \geq 1/3$, which is too optimistic. This indicates that the symmetric splitting approach is potentially yields a too optimistic key rate, since we have shown that it at least yields a too optimistic threshold QBER. Unfortunately, we cannot calculate the they rate for the asymmetric splitting approach, as the corresponding density matrices contain too many different eigenvalues, resulting in an increasing running time of our algorithms. Therefore, we have to leave this issue unsolved.

6.4.4 Decoy States

The problem that Alice and Bob face whenever they have to deal with multi-photon pulses is that Eve's options are greatly increased. We have seen this in the previous section where the eavesdropper was able to devise a strategy that depends on the photon number of each pulse. Unfortunately for Alice and Bob, using the standard protocol, they cannot get enough information about this new strategy to tightly bound the newly arising parameters, and thus they have to make the most pessimistic assumptions (cf. Eq. 6.36).

The decoy state protocol aims at solving this problem: By employing a second photon source with a different photon number statistic, Alice and Bob can measure more characteristics of their data and can potentially infer the value of all parameters in the protocol.

Suppose that Alice holds two photon sources which we label a and b. Originally [40, 41], the idea was that pulses from one source (the "signal" source) are used in a standard (i.e., BB84 or 6-state) protocol to generate the secret key. "Decoy" pulses are interspersed in between them to probe the eavesdropping strategy. It turned out that both sources can play an equal role in the key generation, making the naming misleading. The two sources need to have exactly the same characteristic such as frequency distribution and polarization, but it is crucial that they have a different photon number distribution. For simplicity, we assume that they are both Poissonian, but with a different mean photon number. We denote by $\eta^{a,b}$ the fraction of single photon pulses among all non-empty pulses emitted by source a and b, respectively, and $1 - \eta^{a,b}$ is the fraction of multi-photon pulses.

We consider the case where Alice decides for each pulse whether to choose source a or b, and that Eve cannot distinguish these pulses. However, she can choose her attack according to the photon number of the pulse. The distributed state is therefore of the form

$$\rho_{AB}^{n} = \rho_{AB}^{\mathrm{dep}}(\beta_0^{\mathrm{s}},\beta_1^{\mathrm{s}})^{\otimes n_{\mathrm{s}}^{a}} \otimes \rho_{AB}^{\mathrm{dep}}(\beta_0^{\mathrm{m}},\beta_1^{\mathrm{m}})^{\otimes n_{\mathrm{m}}^{a}} \otimes \rho_{AB}^{\mathrm{dep}}(\beta_0^{\mathrm{s}},\beta_1^{\mathrm{s}})^{\otimes n_{\mathrm{s}}^{b}} \otimes \rho_{AB}^{\mathrm{dep}}(\beta_0^{\mathrm{m}},\beta_1^{\mathrm{m}})^{\otimes n_{\mathrm{m}}^{b}}, \quad (6.44)$$

where $n_{\mathrm{s}}^{a,b} = \eta^{a,b} n^{a,b}$ are the number of single photon pulses originating from source a and b, respectively, and $n_{\mathrm{m}}^{a,b} = (1-\eta^{a,b})n^{a,b}$ are the corresponding number of multi-photon pulses. The total number of pulses in source a and b are denoted by $n^{a,b}$, with $n^a + n^b = n$. As Alice and Bob can (only) differentiate between pulses from the two different sources, they measure two effectively distributed states:

$$\rho_{AB}^{a} = \frac{n_{\mathrm{s}}^{a}}{n^a}\rho_{AB}^{\mathrm{dep}}(\beta_0^{\mathrm{s}},\beta_1^{\mathrm{s}}) + \frac{n_{\mathrm{m}}^{a}}{n^a}\rho_{AB}^{\mathrm{dep}}(\beta_0^{\mathrm{m}},\beta_1^{\mathrm{m}}) = \rho_{AB}^{\mathrm{dep}}(\beta_0^{a},\beta_1^{a}), \quad (6.45)$$

$$\rho_{AB}^{b} = \frac{n_{\mathrm{s}}^{b}}{n^a}\rho_{AB}^{\mathrm{dep}}(\beta_0^{\mathrm{s}},\beta_1^{\mathrm{s}}) + \frac{n_{\mathrm{m}}^{b}}{n^b}\rho_{AB}^{\mathrm{dep}}(\beta_0^{\mathrm{m}},\beta_1^{\mathrm{m}}) = \rho_{AB}^{\mathrm{dep}}(\beta_0^{b},\beta_1^{b}), \quad (6.46)$$

with

$$\beta_0^{a,b} = \frac{n_{\mathrm{s}}^{a,b}}{n^{a,b}}\beta_0^{\mathrm{s}} + \frac{n_{\mathrm{m}}^{a,b}}{n^{a,b}}\beta_0^{\mathrm{m}} = \eta^{a,b}\beta_0^{\mathrm{s}} + (1-\eta^{a,b})\beta_0^{\mathrm{m}}, \quad (6.47)$$

$$\beta_1^{a,b} = \frac{n_{\mathrm{s}}^{a,b}}{n^{a,b}}\beta_1^{\mathrm{s}} + \frac{n_{\mathrm{m}}^{a,b}}{n^{a,b}}\beta_1^{\mathrm{m}} = \eta^{a,b}\beta_1^{\mathrm{s}} + (1-\eta^{a,b})\beta_1^{\mathrm{m}}. \quad (6.48)$$

Here, $\beta_{0,1}^{a,b}$ are the measurable quantities for Alice and Bob, and the $\eta^{a,b}$ are known because Alice knows the characteristics of her sources. The quantity β_0^{s}, i.e., the correlations present in the single-photon pulses, determine the extractable key length, and need to be determined by Alice and Bob. Since Eqs. (6.47) and (6.48) form a linear system of equations, Alice and Bob can compute β_0^{s} from their data. This is in contrast to the scenario depicted in the previous section, where Alice and Bob were not able to determine the value of these variables exactly, but needed a pessimistic assumption.

As we have shown in Sec. 6.4.2, only the single-photon pulses contribute to the secret key rate. Thus, by introducing the appropriate cq-states for the different terms in Eq. (6.44), which are given by Eq. (6.33), we find again that only the single-photon part, i.e. only n_{s} signals, contribute to the secret key rate. The important parameter determining the extractable key length is the correlation β_0^{s} in the single-photon part, which is given by

$$\beta_0^{\mathrm{s}} = \frac{1}{\eta^{a,b}}(\beta_0^{a,b} - (1-\eta^{a,b})\beta_0^{\mathrm{m}}) \quad (6.49)$$

$$= \frac{(1-\eta^{a})\beta_0^{b} - (1-\eta^{b})\beta_0^{a}}{\eta^{a} - \eta^{b}}, \quad (6.50)$$

which follows from Eqs. (6.47) and (6.48). Here, β_0^{s} and β_0^{m} characterize the eavesdropping strategy, i.e., they are chosen by Eve, and Alice and Bob need to determine them. As discussed above, the employment of two different photon sources provides them with a method to achieve this, in contrast to the original protocol discussed in Section 6.4.1. In the latter case, Eq. (6.49) has to be replaced by Eq. (6.36), where β_0^{m} was pessimistically assumed to be 1. Eq. (6.50) can also be re-expressed in terms of quantum bit error rates, yielding

$$Q^{\mathrm{s}} = \frac{(1-\eta^b)Q^a - (1-\eta^a)Q^b}{\eta^a - \eta^b}, \tag{6.51}$$

where Q^{s} is the QBER in the single-photon part and $Q^{a,b}$ are the measured QBERs for source a and b, respectively.

To conclude, we have seen that using decoy state, Alice and Bob do not need to make a pessimistic assumption of the parameter Q^{s} which determines the error rate in the single-photon signals and thus the obtainable secret key rate. Still, only the single-photon part contributes, and the fraction of single-photon pulses is given by $\eta = \alpha\eta^a + (1-\alpha)\eta^b$, where α determines the fraction of signals originating from source a, which Alice can tune at will. In Fig. 6.9, we compare the obtainable key rates for a single-photon implementation with a realistic weak coherent pulse scheme and with the decoy state method presented in this section. We see that the decoy state method is more robust than the method without decoy states, as the key rate remains non-zero for a larger range of the error rate. Only for very low error rates, the decoy state method becomes unfavorable, as multi-photon pulses are added without the need for estimating Q^{s}, which is almost equal to Q for $Q \to 0$.

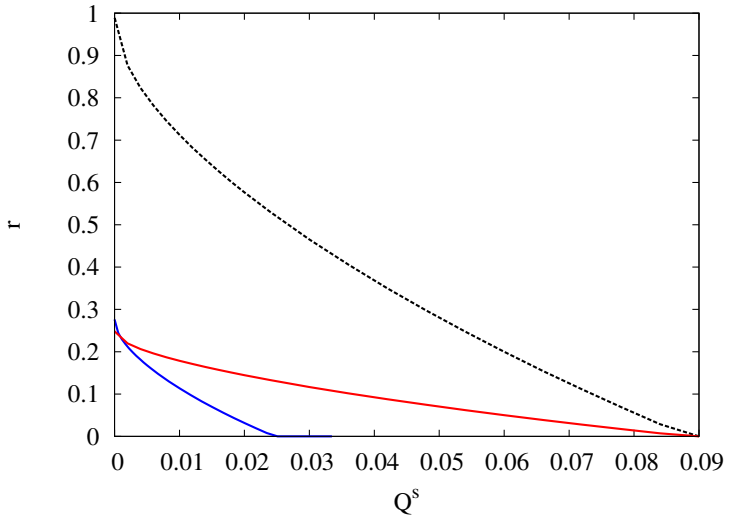

Figure 6.9: (Color online.) Secret key rate r versus the QBER of the single-photon part, $Q^{\mathrm{s}} = 1 - \beta_0^{\mathrm{s}}$ for a single-photon implementation (dashed line), a weak-coherent pulse implementation with a fraction $\eta = 0.279$ of single-photon events (blue), and a decoy state implementation (red). For the decoy states, we have used the same parameters as in [73], that is, $\eta^a = 0.279$, $\eta^b = 0.166$, and a fraction of $\alpha = 3/4$ of signals from source a. We have chosen $n = 20,000$ and $\varepsilon = 10^{-10}$.

Chapter 7

Conclusion

In this thesis, we have developed a novel method to calculate secret key rates for quantum key distribution protocols. It consists of algorithms that allow for the explicit calculation of smooth Renyi entropies, which determine the key rate r by virtue of an expression found by [1] (see Eq. (5.12)). Our notion of secret key rate is defined with respect to the privacy amplification step, i.e., it is the ratio of the output (the number of secret key bits ℓ) and the input (the number of raw key bits n, the block size) of the privacy amplification step. In particular, our results are valid for any finite value of n (however, for numerical limitations, we are restricted to values of $n \lesssim 7 \cdot 10^4$).

As an example application of our algorithms, we investigated the Tomographic Protocol, a variant of the well-known Six-state Protocol. In its analysis, we made two assumptions: First, Alice and Bob can verify in the parameter estimation step that the quantum state they share after the distribution is of a certain form ρ_{AB}^{dep} (see Eq. 6.2). The idea behind this assumption is that Alice and Bob would expect to share this state in the absence of an eavesdropper if they are connected by a depolarizing channel, which is a reasonable error model for a realistic channel. The maximally entangled state $|\phi^+\rangle$ Alice distributes evolves into ρ_{AB}^{dep} under the action of this channel. "Verification" means that we assume that Alice and Bob always share this state, because any instance in which they find a different state will be treated as an indication for an eavesdropping attack and the protocol will be aborted. Alternatively, we can interpret the Tomographic Protocol as the standard Six-state Protocol with an additional restriction on the eavesdropping attack, namely that it is symmetric, i.e. the noise that it introduces is white noise, in the same way as for the depolarizing channel. The second assumption is that the eavesdropper only conducts collective attacks, that is, each signal is attacked in the same way. There are indications that assuming collective attacks is not really a restriction, because Alice and Bob can enforce such a high symmetry by only adding a

small extra step to the protocol [1, 38, 63].

We are able to calculate the secret key rate for the Tomographic Protocol as a function of various parameters: The block size n, the quantum bit error rate (QBER) Q (which is the only free parameter describing the quantum states distributed between Alice and Bob), the dimension of the quantum systems d, and the security parameter ε. For a realistic implementation, it is important to assume a finite block size n (the number of signals that are inputed to the privacy amplification step) and a non-vanishing security parameter (which is the probability that the final key is insecure). We have found that the key rate heavily depends on both n and ε: It converges to the asymptotic value r_∞ for $n \to \infty$ found in [1], reaching approximately 86% of it at $n = 7 \cdot 10^4$ for low error rates ($Q = 0.02$) and a security parameter of $\varepsilon = 10^{-3}$. We believe that typical privacy amplification protocols work with block sizes of the order of 10^5, therefore our results suggest that the finiteness of n does not lead to a dramatic decrease of the secret key rate in a realistic setting. Potentially more severe is the influence of the security parameter on the key rate for a finite block size, as it was shown that r becomes independent of it only in the limit $n \to \infty$ [1]. Indeed our results show that the key rate decreases with decreasing ε. Still, for a realistic error rate of $Q = 0.05$ and a block size of $n = 10^4$, we have found that the key rate remains non-zero even for very small security parameters up to $\varepsilon > 10^{-28}$. This is an important result that shows that unconditionally secure keys can also be created in a realistic scenario. Note that up to now, there is no general understanding of what are reasonable values to choose for ε.

We extended our analysis of the tomographic protocol to include also multi-photon events. These are signals send from Alice containing more than one copy of a signal states, enabling the so-called photon number splitting (PNS) attack. Such events inevitably occur in any realistic application relying on weak coherent pulses. Making the most pessimistic assumptions that the eavesdropper has complete knowledge about all key bits originating form multi-photon pulses, we have shown that the key rate will decrease only by a factor of η (the fraction of single photon events) with respect to an idealized single-photon implementation. However, the estimation of the QBER Q becomes more involved. Our results suggest that the "real" QBER, which is used for the calculation of the key rate, is at least a factor of $1/\eta$ larger than the QBER which is estimated from Alice's and Bob's data.

Our results show that secret key generation by privacy amplification remains feasible when taking into account realistic assumptions, namely: The privacy amplification protocol works with finite blocks of classical data, and the quality of the key is measured by some security parameter which can be set to a desired value.

Appendix A

Numerical methods

In this chapter, we present the numerical methods that were employed to derive the results presented in Ch. 6. The main task in that section was to find the achievable key rate for a specific QKD protocol, which involved calculating smooth Renyi entropies of the quantum states describing Alice's and Eve's systems (and of the probability distribution describing Alice's and Bob's correlations). In Sec. 5.3.3, we derived a simple expression for the Renyi entropies $S_2^\varepsilon(\rho)$, $S_0^\varepsilon(\rho)$, and $H_0^\varepsilon(X|Y)$ in terms of the eigenvalues and probabilities occurring in ρ and $P_{X|Y}$, respectively. Since the numerical evaluation of these expressions are very similar for all entropies, we focus on $S_0^\varepsilon(\rho)$ to present the details. For $\rho \in \mathcal{B}(\mathbb{C}^d)$, recall that

$$S_0^\varepsilon(\rho) = \log(d-k), \tag{A.1}$$

where

$$\begin{aligned} k &= \sum_{i=1}^{b^-} n_{i-1} + \left\lfloor \frac{\varepsilon - s^-(b^-)}{\lambda_{b^-}} \right\rfloor, && \text{(A.2)} \\ b^- &= \max\{r : s^-(r) \leq \varepsilon\} && \text{(A.3)} \\ s^-(r) &= \sum_{i=1}^{r} n_{i-1}\lambda_{i-1}. && \text{(A.4)} \end{aligned}$$

Here, λ_i and n_i with $1 \leq i \leq m$ are the eigenvalues and multiplicities of ρ, respectively, and $0 \leq r, b^- \leq m+1$. The specific form of the eigenvalues in given in Eq. (6.13). The main complication lies in the calculation of b^-, for which we cannot give a closed expression, because we cannot solve (A.4) for r. Therefore, we need to devise an algorithm that checks for all r whether $s^-(r) \leq \varepsilon$ and returns the largest one of these. Unfortunately, it is unfeasible to start from $r = 0$, increasing r by one in each step and

test if $s^-(r) \leq \varepsilon$, because this requires (in the worst case scenario) $m+1$ operations. For the QKD protocol we investigated in Ch. 6, the density matrix ρ is an n-fold tensor product (cf. Eq. (6.9)), i.e. m is usually of the order of 2^n (unless it is pure), where n is the block size of the privacy amplification protocol, which can easily be of the order of 10^5 (cf. Sec. 6.2). Thus, this simple approach for finding r needs a number of operations that scales exponentially in n.

A more desirable approach is to employ a more sophisticated search algorithm to find b^-. This is possible because the elements of the search space, the $s^+(r)$, are ordered. In this work, we implemented a binary search algorithm [74] that cuts the interval containing the solution in half in each step. Such an algorithm has a running time of $\log(m+1)$, if the search space contains $m+1$ elements, which is in our case of the order of $\log 2^n = n$. In this way, the running time stays polynomial in the block size of the PA protocol.

The calculations were performed on standard desktop PCs, equipped with Pentium 4 processors running at 3.2 GHz and 1GB memory. The algorithms were written in C++, using Ginac's arbitrary precision library, cln [75]. This is necessary because we are dealing with very small numbers due to the large size of input density matrices: The smallest eigenvalue is typically of the order of 2^{-n} (cf. Sec. 6.2). Note that all calculations performed do not contain any approximations and are exact up to possible rounding errors. For most of the calculations, it is sufficient to use a precision of 20 significant digits; we verified however for all calculations that a variation of the precision does not lead to change in the result (the key length, which is of the order of n) of more than 10^{-5}.

A typical calculation of the secret key length for, say, a block size of $n = 20,000$ and security parameter $\varepsilon = 10^{-10}$ lasts about 10 min. Since the runtime of the algorithms is only linear in n, there is in principle no limitation for block sizes n to be calculated, as long as one is investing more computer power and memory. With our limited resources, we were able to calculate key rates for block sizes up to $n = 70,000$, which lasts a couple of hours, but already uses up all memory. Moving to vector processors or clusters, this limit could certainly be pushed further, but we do not expect any new insights from calculating secret key rates for block sizes beyond 10^5.

Bibliography

[1] R. Renner, N. Gisin, and B. Kraus, Phys. Rev. A **72**, 012332 (2005).

[2] D. Bruß *et al.*, Phys. Rev. Lett. **91**, 097901 (2003).

[3] Y. C. Liang *et al.*, Phys. Rev. A **68**, 22324 (2003).

[4] D. Bruß, Phys. Rev. Lett. **81**, 3018 (1998).

[5] H. Bechmann-Pasquinucci and N. Gisin, Phys. Rev. A **59**, 4238 (1999).

[6] http://www.magiqtech.com.

[7] http://www.idquantique.com.

[8] T. Kelly, Cryptologia **22**, 244 (1998).

[9] D. Kahn, *The Codebreakers: The Comprehensive History of Secret Communication from Ancient Times to the Internet* (B&T, Charlotte, 1996).

[10] R. Rivest, A. Shamir, and L. Adleman, Communications of the ACM **21**, 120 (1978).

[11] C. Papadimitriou, *Computational Complexity* (Addison Wesley, Amsterdam, 1993).

[12] G. Moore, Electronics **38**, (1965).

[13] M. Dŭsek, N. Lütkenhaus, and M. Hendrych, Progress in Optics **49**, 381 (2006).

[14] G. S. Vernam, Journal of IEEE **55**, 109 (1926).

[15] C. Shannon, Bell System Technical Journal **28**, 656 (1949).

[16] P. Gemmell and N. Naor, Advances in Cryptology–CRYPTO '93 **773**, 355 (1993).

[17] C. H. Bennett and G. Brassard, in *Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing, Bangalora, India* (IEEE, New York, 1985), pp. 175–179.

[18] W. K. Wootters and W. H. Zurek, Nature **299**, 802 (1982).

[19] A. Ekert, Phys. Rev. Lett. **67**, 661 (1991).

[20] J. Clauser, M. Horne, A. Shimony, and R. Holt, Phys. Rev. Lett. **23**, 880 (1969).

[21] J. S. Bell, Physics **1**, 195 (1964).

[22] A. Acín and N. Gisin, Phys. Rev. Lett. **94**, 020501 (2005).

[23] A. Acín, L. Masanes, and N. Gisin, Phys. Rev. Lett. **91**, 167901 (2003).

[24] M. Curty, M. Lewenstein, and N. Lütkenhaus, Phys. Rev. Lett. **92**, 217903 (2004).

[25] C. H. Bennett *et al.*, Phys. Rev. Lett. **76**, 722 (1996).

[26] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, quant-ph/0702225 (2007).

[27] H.-K. Lo and H. F. Chau, Science **283**, 2050 (1999).

[28] P. W. Shor and J. Preskill, Phys. Rev. Lett. **85**, 441 (2000).

[29] A. R. Calderbank and P. W. Shor, Phys. Rev. A **54**, 1098 (1996).

[30] A. M. Steane, Proc. R. Soc. London A **452**, 2551 (1996).

[31] U. M. Maurer, IEEE Transactions on Information Theory **39**, 733 (1993).

[32] C. H. Bennett, G. Brassard, C. Crepeau, and U. M. Maurer, IEEE Trans. Info. Theory **41**, 1915 (1995).

[33] C. H. Bennett, G. Brassard, and J.-M. Robert, SIAM Journal on Computing **17**, 210 (1988).

[34] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Rev. Mod. Phys. **74**, 145 (2002).

[35] J. L. Carter and M. N. Wegman, Journal of Computer and System Sciences **18**, 143 (1979).

[36] M. N. Wegman and J. L. Carter, Journal of Computer and System Sciences **22**, 265 (1981).

[37] R. Renner and R. König, in *Second Theory of Cryptography Conference, TCC 2005*, Vol. 3378 of *LNCS*, edited by J. Kilian (Springer, New York, 2005), pp. 407–425.

[38] R. Renner, N. Gisin, and B. Kraus, Phys. Rev. Lett. **95**, 080501 (2005).

[39] W. K. Wootters and B. D. Fields, Ann. Phys. (N.Y.) **191**, 363 (1989).

[40] W.-Y. Hwang, Phys. Rev. Lett. **91**, 057901 (2003).

[41] H.-K. Lo, X. Ma, and K. Chen, Phys. Rev. Lett. **94**, 230504 (2005).

[42] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer Academic Publishers, Dordrecht, 1995).

[43] K. Kraus, *States, Effects and Operations: Fundamental Notions of Quantum Theory* (Springer, New York, 1983).

[44] F. W. Stinespring, Proc. Amer. Math. Soc. **6**, 211 (1955).

[45] C. Bennett, Phys. Rev. Lett. **68**, 3121 (1992).

[46] C. W. Helstrøm, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).

[47] H.-K. Lo, H. F. Chau, and M. Ardehali, J. of Cryptology **18**, 133 (2005).

[48] D. Deutsch *et al.*, Phys. Rev. Lett. **77**, 2818 (1996).

[49] I. Devetak and A. Winter, Phys. Rev. Lett. **93**, 080501 (2004).

[50] N. Gisin, Helv. Phys. Acta **62**, 363 (1989).

[51] L. P. Hughston, R. Jozsa, and W. K. Wootters, Phys. Lett. A **183**, 14 (1993).

[52] V. Scarani, A. Acín, G. Ribordy, and N. Gisin, Phys. Rev. Lett. **92**, 057901 (2004).

[53] R. Renner and S. Wolf, Lecture Notes in Computer Science **3788**, 199 (2005).

[54] N. Gisin, R. Renner, and S. Wolf, Algorithmica **34**, 389 (2002).

[55] R. König, R. Renner, A. Bariska, and U. Maurer, Phys. Rev. Lett. **98**, 140502 (2007).

[56] I. Devetak and A. Winter, Phys. Rev. A **68**, 042301 (2003).

[57] M. Ben-Or *et al.*, in *Second Theory of Cryptography Conference, TCC 2005*, Vol. 3378 of *LNCS*, edited by J. Kilian (Springer, New York, 2005), pp. 386–406.

[58] R. Renner, Ph.D. thesis, ETH Zürich, 2005.

[59] M. Kleinmann, H. Kampermann, T. Meyer, and D. Bruß, Phys. Rev. A **73**, 062309 (2006).

[60] W. Dür, J. I. Cirac, and R. Tarrach, Phys. Rev. Lett. **83**, 3562 (1999).

[61] R. Renner, (private communication).

[62] T. Meyer, H. Kampermann, M. Kleinmann, and D. Bruss, Phys. Rev. A **74**, 042340 (2006).

[63] R. Renner, quant-ph/0703069 (2007).

[64] J. Schwinger, Proceedings of the National Academy of Sciences of the United States of America **46**, 570 (1960).

[65] S. Bandyopadhyay, P. O. Boykin, V. Roychowdhury, and F. Vatan, Algorithmica **34**, 512 (2002).

[66] A. Klappenecker and M. Rötteler, in *7th International Conference on Finite Fields and Applications* (Springer, New York, 2004), Vol. 2984, pp. 137–144.

[67] M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[68] H.-K. Lo, QIC **1**, 81 (2001).

[69] S. Groeblacher *et al.*, New J. Phys. **8**, 75 (2005).

[70] N. Imoto, H. A. Haus, and Y. Yamamoto, Phys. Rev. A **32**, 2287 (1985).

[71] D. Gottesman, H.-K. Lo, N. Lütkenhaus, and J. Preskill, Quant. Inf. Comput. **5**, 325 (2004).

[72] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996).

[73] Z. L. Yuan, A. W. Sharpe, and A. J. Shields, Appl. Phys. Lett. **90**, 011118 (2007).

[74] D. E. Knuth, *The Art of Computer Programming* (Addison-Wesley Longman, Amsterdam, 1999).

[75] http://www.ginac.de/CLN/.

Acknowledgements

First and foremost, I would like to thank Prof. Dr. Dagmar Bruß for giving me the opportunity to work in her group, for the great freedom I experienced during my work, and for the encouragement I received. Furthermore, I am indebted to the other (also former) group members Dr. Hermann Kampermann, Dr. Razmik Unanyan, Matthias Kleinman, Zahra Shadman, and Patrick Skwara for the pleasant atmosphere, countless discussions, and many advices and enlightments.

Special thanks go to Dr. Renato Renner, Dr. Barbara Kraus, and Prof. Dr. Norbert Lütkenhaus for helpful discussions.

Finally, I would like to thank my parents and my girlfriend Alexandra for their understanding and support during the work on this thesis.