arXiv:quant-ph/0402131 v2 4 Mar 2004

A Generic Security Proof for Quantum Key Distribution

Matthias Christandl * Renato Renner † Artur Ekert * ‡

March 4, 2004

Abstract

Quantum key distribution allows two parties, traditionally known as Alice and Bob, to establish a secure random cryptographic key if, firstly, they have access to a quantum communication channel, and secondly, they can exchange classical public messages which can be monitored but not altered by an eavesdropper, Eve. Quantum key distribution provides perfect security because, unlike its classical counterpart, it relies on the laws of physics rather than on ensuring that successful eavesdropping would require excessive computational effort. However, security proofs of quantum key distribution are not trivial and are usually restricted in their applicability to specific protocols. In contrast, we present a general and conceptually simple proof which can be applied to a number of different protocols. It relies on the fact that a cryptographic procedure called privacy amplification is equally secure when an adversary's memory for data storage is quantum rather than classical [1].

1 Introduction

The potential power of quantum phenomena to protect information was first adumbrated by Wiesner who, in the early 1970's, introduced the concept of quantum conjugate coding [2]. He showed how to store or transmit two messages by encoding them in two conjugate observables, such as linear and

*Centre for Quantum Computation, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, United Kingdom

†Computer Science Department, ETH Zürich, CH-8092 Zürich, Switzerland

‡Department of Physics, National University of Singapore, Singapore 117 542, Singapore

circular polarization of light, so that either but not both of may be received and decoded. He illustrated his idea with a design of unforgeable bank notes. Building upon this work, Bennett and Brassard proposed a quantum key distribution scheme, known as BB84 or the four state protocol, in which Alice repeatedly sends to Bob one of four prescribed states of a qubit, and Bob measures them in one of two conjugate bases [3]. Independently and initially unaware of the earlier work, Ekert developed a different approach to quantum cryptography based on quantum entanglement. He proposed a key distribution protocol, known as E91, in which entangled pairs of qubits are distributed to Alice and Bob, who then extract key bits by measuring their qubits in prescribed bases [4]. A particularly nice feature of E91, for the purpose of security analysis, is that Eve herself is allowed to prepare and deliver all the qubit pairs that Alice and Bob will subsequently use to generate the key.

Many variations on quantum key distribution have been subsequently proposed and we will mention some of them later on. They can be roughly divided into "prepare and measure" protocols, such as BB84 and B92 [5], and "entanglement based" protocols, such as E91. Many interesting techniques for manipulating quantum entanglement have been discovered in the last few years. Thus it is often convenient to cast some of the "prepare and measure" protocols in terms of the "entanglement based" ones.

1.1 Security Proofs

All good quantum key distribution protocols must be operable in the presence of noise that may or may not result from eavesdropping. The protocols must specify for which values of measurable parameters Alice and Bob can establish a secret key and provide a physically implementable procedure which generates such a key. The design of the procedure must take into account that an eavesdropper may have access to unlimited quantum computing power. On Alice and Bob's side, the procedure should rely on simple and easily implementable operations. For example, good protocols should not assume that Alice and Bob have quantum computers, or any sophisticated quantum technology, apart from the ability to transmit over a quantum channel.

The search for operational security criteria led to early studies of quantum eavesdropping [6, 7] and finally to the first proof of the security of key distribution [8]. The original proof showed that the E91 and all entanglement based key distributions are indeed secure and noise-tolerant against an adversary with unlimited computing power as long as Alice and Bob can

implement quantum privacy amplification. Quantum privacy amplification allows one to establish a secure key over any distance, e.g. using entanglement swapping [9] in a chain of quantum repeaters [10, 11]. However, this procedure, which distills pure entangled states from corrupted mixed states of two qubits, requires a small scale quantum computation. Subsequent proofs by Inamori [12] and Ben-Or [13] showed that Alice and Bob can also distill a secret key from partially entangled particles using only classical error correction and classical privacy amplification [14, 15].

Quantum privacy amplification was also used by Lo and Chau to prove the security of the BB84 protocol over an arbitrary distance [16]. A concurrent and independent proof by Mayers showed that the protocol can be secure without Alice and Bob having to rely on the use of quantum computers [17]. The same conclusion, but using different techniques, was subsequently reached by Biham *et al.* [18]. Although the two proofs did not require quantum privacy amplification they were rather complex. A nice fusion of quantum privacy amplification and error correction was proposed by Shor and Preskill who formulated a relatively simple proof of the security of the BB84 protocol based on virtual quantum error correction [19]. They showed that a protocol which employs quantum error-correcting codes to prevent Eve from becoming entangled with qubits that are used to generate the key reduces to the BB84 augmented by classical error correction and classical privacy amplification. This proof has been further extended by Gottesman and Lo [20] to cover the case of two-way public communication in BB84 which allows a higher bit error rate, and by Tamaki *et al.*[21] to prove the security of the B92 protocol. More recently another simple proof of the security of BB84, which employs results from quantum communication complexity, has been provided by Ben-Or [13].

1.2 Do we need another Security Proof?

Most popular quantum key distribution schemes have been analyzed in terms of their security criteria and there is a pretty good understanding of the limitations of the techniques involved e.g. those due to imperfect sources or detectors. The schemes vary but every single one of them must involve either quantum or classical privacy amplification as an inherent part of the secure key distillation protocol.

Classical privacy amplification, originally proposed by Bennett, Brassard and Robert, was restricted to the case in which Eve acquires *classical, deterministic* information about the raw key [14]. The applicability of the method was then extended by Bennett, Brassard, Crépeau and Maurer to

cover scenarios where Eve's information is *classical* and *probabilistic* [15]. We use the recent result by König, Maurer and Renner on the power of quantum memory [1] in a quantum cryptographic context. It can be viewed as a further generalization of classical privacy amplification to cases in which Eve's information about the key is *quantum*. Of course, the privacy amplification is useless unless we can derive an upper bound on the amount of quantum information available to Eve. We show how to do this for common quantum key distribution protocols. Taken together these results give a very general and powerful technique for assessing security of a wide class of quantum key distribution protocols.

1.3 Scenario

In our scenario Eve has a technological advantage over Alice and Bob. She can distribute qubits to Alice and Bob, she can entangle the qubits with an ancilla that she controls, she can have access to unlimited quantum computational power, and she can monitor all the public communication between Alice and Bob in which they reveal their measurement choices and exchange further information in order to correct errors in their shared key and to amplify its privacy. In contrast Alice and Bob can only perform measurements on individual qubits and communicate classically over a public channel. We will assess the security in the case of a noisy quantum channel without losses.

Alice and Bob go through prescribed stages of quantum key distribution and at some point they end up with perfectly correlated binary strings about which Eve has some information, namely all information communicated in public together with all information contained in her ancilla. The ancilla is a quantum entity which Eve may measure at the very end of the key distribution protocol. Hence its information content has to be expressed in qubits rather than bits. Classical privacy amplification allows Alice and Bob to increase the privacy of the shared string as long as they can estimate the amount of classical information that leaked to Eve [14]. For any shared string of n bits upon which Eve has some r bits of information the procedure outputs a binary string of length s shorter than $n-r$ and such that Eve has virtually no information about the new string. The snag is that Eve, who can delay her measurement of the ancilla, has r qubits rather than r bits of information about the n–bit string. However, in this particular context, it does not matter, as shown in [1]. We show how Alice and Bob are able to estimate the quantum information content of r qubits in Eve's ancilla in a generic quantum key distribution protocol.

2 Outline of the Main Result

It is convenient for the purpose of this outline to start with a generic scenario in which Eve distributes quantum particles to Alice and Bob. Without any loss of generality we assume that Eve starts with a tripartite pure state describing a batch of particles delivered to Alice, a batch of particles delivered to Bob, and an ancilla which is retained by Eve.

When Alice and Bob receive their respective particles they perform measurements following a quantum key distribution protocol, which they agreed to in advance. For example, they may measure every single particle choosing randomly from a prescribed set of different measurements. They also communicate in public and agree which outcomes of the measurements are to be discarded and which will be used for the key generation.

At this point Alice and Bob have partially correlated n bit strings labeled, respectively, as X and Y. Eve knows the protocol and holds an ancilla which was entangled with the qubits prior to Alice's and Bob's measurements. After the measurements the ancilla is in a quantum state which, in general, depends on X and Y and is described by some density operator ρ^E. The initial public communication must allow Alice and Bob to estimate the degree of the correlation between X and Y and to derive an upper bound on the quantum information content of the ancilla in state ρ^E. This is not trivial as we do not assume that the pairs of qubits are independent and identically distributed (i.i.d); they can be entangled between themselves and the ancilla in an arbitrary way.

We solve the problem in its full generality. However, in this section we present a rough outline based on the i.i.d case. This, we hope, will serve as a gentle introduction to the more technical sections that follow.

Let Alice and Bob be given n realizations of i.i.d random variables X and Y respectively. Let the degree of correlations be quantified by the mutual information $I(X;Y)$ and let the quantum information content of the ancilla be no more than r qubits. The strings of Alice and Bob can be made identical with high probability by a procedure called information reconciliation. Alice has to communicate in public approximately $nH(X|Y)$ (the conditional entropy of X given Y) bits about her string so that Bob, who holds n realizations of Y, can guess Alice's string correctly.

Thus, after the information reconciliation, Eve's information about Alice's string consists of $nH(X|Y)$ classical bits and r qubits. Without any loss of generality we can assume that Eve's information is contained in $nH(X|Y) + r$ qubits. Eve can wait and perform her measurement on the ancilla whenever she sees fit. However, no matter which observable she mea-

sures after the classical privacy amplification she is not better off than she would be if she had $nH(X|Y)+r$ classical bits of information about X prior to the privacy amplification. This follows from the recent work by König, Maurer, and Renner on the power of quantum memory [1]. We will elaborate on this in more detail in section 3 and section 4. Thus the length of the secret key after the privacy amplification is $nH(X)-nH(X|Y)-r = nI(X;Y)-r$, i.e. the key can be established when $nI(X;Y) > r$.

In the main part of the paper we will show how the estimation of r works in general. In order to illustrate the idea behind this estimation, let us consider the particular case of independent and identically distributed pairs of quantum states. Each pair that Eve delivers to Alice and Bob comes from a tripartite pure state $|\Psi\rangle$ such that $\rho = \mathrm{tr}_E|\Psi\rangle\langle\Psi|$ is the density operator of each pair of quantum states and $\rho^E = \mathrm{tr}_{AB}|\Psi\rangle\langle\Psi|$ is the density operator of a part of the ancilla. The state of the ancilla in an n-fold tensor product of the form $\rho^E = \rho^e \otimes \cdots \otimes \rho^e$. In this particular case we can use the quantum coding results [22, 23] to estimate r in the limit of large n; $r = nS(\rho^e) = nS(\rho)$ qubits, where $S(\rho)$ is the von Neumann entropy of ρ; $S(\rho) = -\mathrm{tr}(\rho \log \rho)$.

In the qubit case, the mutual information can be written as $I(X;Y) = n(1-h(\epsilon))$, where $h(\epsilon) = -\epsilon \log \epsilon - (1-\epsilon)\log(1-\epsilon)$ is the binary entropy function and ϵ is the average bit error rate. The threshold error rate can be then established from the condition

$$1 - h(\epsilon) \geq S(\rho). \tag{1}$$

A key distribution protocol should allow Alice and Bob to estimate the purity of the pairs of quantum states in terms of the von Neumann entropy $S(\rho)$. If not they need to maximize $S(\rho)$ over all possible density operators ρ which are consistent with the estimated bit error rate ϵ.

Moreover the key rate R is

$$R = H(X) - H(X|Y) - \max_{\hat{\rho}\in\mathcal{R}} S(\hat{\rho}) \ , \tag{2}$$

(see section 4.4.3)

The argument above relies on the extension of the applicability of classical privacy amplification to the cases where Eve has partial quantum rather than partial classical information about the key. This follows from a more general observation that encoding classical information into qubits rather than bits, although never worse, does not offer any significant advantage in some scenarios; ours being one of them. Amazingly enough potential advantages of quantum encoding were already pointed out by Wiesner in his

seminal paper on conjugate coding [2]. Subsequently Ambainis, Nayak, Ta-Shma and Vazirani [24] considered a scenario where one has to use partial information about a binary string X to answer a random binary question about X. One might think that storing partial information about X in a quantum rather than classical memory has a natural advantage because one can delay a measurement on the quantum memory until after the question has been asked. This gives an extra freedom of choosing the most appropriate measurement. However, Ambainis *et al.* [24] and Nayak [25] showed that if information about an n–bit string X is stored in r qubits and one is asked about a particular bit of X then in order to err with probability less than ϵ one needs $r > n\ (1 - h(\epsilon))$. Thus, asymptotically, in this particular case, quantum storage does not offer any advantage. König, Maurer and Renner [1] show that there is no advantage even if one is asked more general, non-binary, questions about X. This made it possible to make the connection to privacy amplification.

In the following we provide a detailed and reasonably self-contained description of the new security proof. In section 3 we introduce the relevant concepts and methods of probability theory and quantum mechanics. The main results are presented in section 4. This is followed by applications of our security criteria to selected quantum key distribution protocols (section 5).

3 Preliminaries

3.1 Notation

Let a be a subset of a set $\mathcal{I}$. The *characteristic function* χ_a of a on $\mathcal{I}$ is the function from $\mathcal{I}$ to $\{0, 1\}$ defined by $\chi(i) = 1$ if and only if $i \in a$.

Let $\mathbf{z} = (z_1, \ldots, z_n)$ be an n-tuple and $a \subseteq \{1, \ldots, n\}$ a set of indices. Then $\mathbf{z}_a$ denotes the $|a|$-tuple containing all z_i with $i \in a$. For two n-tuples $\mathbf{z} = (z_1, \ldots, z_n)$ and $\mathbf{z}' = (z'_1, \ldots, z'_n)$ of real values, $\mathbf{z}$ is said to be *majorized* by $\mathbf{z}'$, denoted $\mathbf{z} \prec \mathbf{z}'$, if for any $k \in \{1, \ldots, n\}$

$$\sum_{i \in a_k} z_i \leq \sum_{j \in b_k} z'_j$$

where a_k and b_k are the sets containing the indices of the k largest elements of $\mathbf{z}$ and $\mathbf{z}'$, respectively. A real valued function f on the set of real n-tuples is said to be *Schur-convex* if

$$\mathbf{z} \prec \mathbf{z}' \implies f(\mathbf{z}) \leq f(\mathbf{z}')$$

for any $\mathbf{z}$ and $\mathbf{z}'$.

For a function f on $\mathcal{Z}$, we denote by $f^{\max}$ and $f^{\min}$ the functions on the power set of $\mathcal{Z}$ defined by

$$f^{\max}(\mathcal{W}) = \max_{z \in \mathcal{W}} f(z) \quad \text{and} \quad f^{\min}(\mathcal{W}) = \min_{z \in \mathcal{W}} f(z) \ ,$$

for any $\mathcal{W} \subseteq \mathcal{Z}$.

Let $\delta : \mathcal{Z} \times \mathcal{Z} \to \mathbb{R}^+$ be a metric on a set $\mathcal{Z}$. The *ε-environment of an element $z \in \mathcal{Z}$* is defined by

$$\mathcal{B}^{\varepsilon}(z) := \{z' \in \mathcal{Z} : \delta(z, z') \leq \varepsilon\} \ .$$

Similarly, the *ε-environment of a subset $\mathcal{W} \subseteq \mathcal{Z}$* is the union of all ε-environments of elements of $\mathcal{W}$, i.e.,

$$\mathcal{B}^{\varepsilon}(\mathcal{W}) := \bigcup_{z \in \mathcal{W}} \mathcal{B}^{\varepsilon}(z) \ .$$

3.2 Elements of Classical Probability and Information Theory

The goal of this subsection is to introduce some concepts of probability and information theory that we will use for the proofs of our main results. For a more complete overview, we refer to the standard literature (e.g., [26]).

In the following, we use capital letters (Z) for random variables, calligraphic letters ($\mathcal{Z}$) for their range, and small letters (z) for the elements of their range. The probability distribution of a random variable Z is denoted by P_Z. The *expectation over Z* of a function f of $\mathcal{Z}$ is given by $E_Z[f(Z)] := \sum_{z \in \mathcal{Z}} P_Z(z) f(z)$. A random variable or probability distribution is called *binary* if it has range $\mathcal{Z} = \{0, 1\}$. We write P_p^{bin} for the binary probability distribution with $P_p^{\text{bin}}(1) = p$.

An n-tuple $(Z_1, \ldots, Z_n)$ of random variables with the same range $\mathcal{Z}$ is called *exchangeable* if, for all permutations π on $\{1, \ldots, n\}$,

$$P_{Z_1 \cdots Z_n} = P_{Z_{\pi(1)} \cdots Z_{\pi(n)}}$$

It is easy to see that, for any n-tuple $(Z_1, \ldots, Z_n)$ of random variables with range $\mathcal{Z}$, the n-tuple

$$(Z'_1, \ldots, Z'_n) := (Z_{\Pi(1)}, \ldots, Z_{\Pi(n)})$$

obtained by permuting the indices according to a random permutation Π on $\{1, \ldots, n\}$ is exchangeable.

The *variational distance* between two probability distributions P and Q over the same range $\mathcal{Z}$ is defined by

$$\delta(P,Q) := \frac{1}{2}\sum_{z\in\mathcal{Z}} |P(z) - Q(z)| \ .$$

The variational distance δ is a metric on the set of probability distributions with range $\mathcal{Z}$. In particular, $\delta(P,Q) = 0$ if and only if $P = Q$, it is symmetric, and it satisfies the triangle inequality. For random variables Z and Z', we also write $\delta(Z,Z')$ instead of $\delta(P_Z, P_{Z'})$. The variational distance between two probability distributions P and Q can be interpreted as the probability that two random experiments described by P and Q, respectively, are different. This is formalized by the following lemma.

Lemma 3.1. *Let P and Q be two probability distributions. Then there exists a pair of random variables Z and Z' with joint probability distribution $P_{ZZ'}$ such that $P_Z = P$, $P_{Z'} = Q$, and*

$$\mathrm{Prob}[Z \neq Z'] = \delta(P,Q) \ .$$

It is easy to see that the variational distance between Z and Z' can not increase when applying the same function f on both Z and Z', i.e.,

$$\delta(Z,Z') \geq \delta(f(Z), f(Z')) \ . \tag{3}$$

Let $[Z,W]$ and $[Z',W']$ be two pairs of random variables, and let $P_{Z|W}(\cdot,w) := P_{Z|W=w}$ and $P_{Z'|W'}(\cdot,w) := P_{Z'|W'=w}$ be the probability distribution of Z and Z' conditioned on $W = w$ and $W' = w$, respectively. Using the triangle inequality, it can be shown that

$$\left|\delta(P_{ZW}, P_{Z',W'}) - E_W[\delta(P_{Z|W}(\cdot,W), P_{Z'|W'}(\cdot,W))]\right| \leq \delta(P_W, P_{W'}) \ . \tag{4}$$

Combining this with (3) for the function $f : (z,w) \mapsto w$ leads to

$$E_W[\delta(P_{Z|W}(\cdot,W), P_{Z'|W'}(\cdot,W))] \leq 2\delta(P_{ZW}, P_{Z'W'}) \ , \tag{5}$$

and, similarly, for $f : (z,w) \mapsto z$,

$$\delta(P_Z, P_{Z'}) \leq E_W[\delta(P_{Z|W}(\cdot,W), P_{Z'|W'}(\cdot,W))] + \delta(P_W, P_{W'}) \ . \tag{6}$$

Let P be a probability distribution over $\mathcal{Z}$. The *non-uniformity* of P,

$$d(P) := \delta(P,U) \ ,$$

is defined as the variational distance of P from the uniform distribution U over $\mathcal{Z}$. For a random variable Z with probability distribution P_Z, we also write $d(Z)$ instead of $d(P_Z)$. Similarly, for two random variables Z and W, the *expected non-uniformity* of Z given W is defined by

$$d(Z|W) := E[d(P_{Z|W}(\cdot, W))] \ .$$

Definition 3.2. Let $\mathbf{z} := (z_1, \ldots, z_n)$ be an n-tuple of elements from a set $\mathcal{Z}$. The *frequency distribution* $Q_{\mathbf{z}}$ of $\mathbf{z}$ is the real valued function on $\mathcal{Z}$ defined by

$$Q_{\mathbf{z}}(z) := \frac{|\{i : z_i = z\}|}{n}$$

for $z \in \mathcal{Z}$.

It is easy to see that the frequency $Q_{\mathbf{z}}$ is a probability distribution on $\mathcal{Z}$, i.e., $Q_{\mathbf{z}}(z) \in [0, 1]$ and $\sum_{z \in \mathbf{Z}} Q_{\mathbf{z}}(z) = 1$.

Definition 3.3. The *probability range* of an n-tuple $\mathbf{Z} = (Z_1, \ldots, Z_n)$ of random variables with range $\mathcal{Z}$ is the smallest convex set $\mathcal{P}$ of probability distributions on $\mathcal{Z}$ such that

$$P_{Z_k|Z_1=z_1,\ldots,Z_{k-1}=z_{k-1}} \in \mathcal{P}$$

for all $k \in \{1, \ldots, n\}$ and $z_1, \ldots, z_{n-1} \in \mathcal{Z}$.

The following result of [27] states that the frequency distribution of a sequence of random variables is with high probability contained in an ε-environment of its probability range.

Lemma 3.4. *Let $\mathbf{Z} = (Z_1, \ldots, Z_n)$ be an n-tuple of random variables with alphabet $\mathcal{Z}$ of size $|\mathcal{Z}| = q$ and let $\mathcal{P}$ be the probability range of $\mathbf{Z}$. Then, for any $\varepsilon > 0$,*

$$\text{Prob}[Q_{\mathbf{Z}} \in \mathcal{B}^{\varepsilon}(\mathcal{P})] \geq 1 - 2^q e^{-n\varepsilon^2/2} \ .$$

We will make use of different entropy measures to characterize random variables or, more precisely, their probability distributions. Let P be a probability distribution with range $\mathcal{Z}$, support $\mathcal{Z}^+ := \{z \in \mathcal{Z} : P(z) > 0\}$, and maximum probability $p_{\max}(P) := \max_{z \in \mathcal{Z}} P(z)$. Then, the *Rényi entropy of order* α, for $\alpha \in \mathbb{R}^+ \cup \{\infty\}$,[1] is defined by[2]

$$H_\alpha(P) := \frac{1}{1-\alpha} \log\big(\sum_{z \in \mathcal{Z}} P(z)^\alpha\big) \ .$$

[1] For $\alpha \in \{0, 1, \infty\}$, $H_\alpha(P)$ is defined by the limit value $\lim_{\beta \to \alpha} H_\beta(P)$.

[2] All logarithms in this paper are binary.

It turns out that, for $\alpha = 1$, $H_1(P)$ corresponds to the *Shannon entropy* $H(P) = -\sum_{z\in\mathcal{Z}^+} P(z)\log(P(z))$. Moreover, for $\alpha = \infty$, we have $H_\infty(P) = -\log(p_{\max})$, which is also called *min-entropy*, and, for $\alpha = 0$, $H_0(P) = \log(|\mathcal{Z}^+|)$. For a random variable Z with probability distribution P_Z, we also write $H(Z)$ instead of $H(P_Z)$, and, more generally, for an event $\mathcal{E}$, $H(Z|\mathcal{E})$ instead of $H(P_{Z|\mathcal{E}})$.

The Rényi entropy of order α of a random variable Z conditioned on another random variable W is given by

$$H_\alpha(Z|W) := \min_{w\in\mathcal{W}} H_\alpha(Z|W=w) \quad (\text{for } \alpha > 1)$$

and

$$H_\alpha(Z|W) := \max_{w\in\mathcal{W}} H_\alpha(Z|W=w) \quad (\text{for } \alpha < 1)\ .$$

We will often be interested in the entropy of a probability distribution which is close to a given distribution P. This is formalized by the notion of *smooth Rényi entropy* introduced in [28].

Definition 3.5. Let $\varepsilon \geq 0$ and $\alpha \in \mathcal{R}^+ \cup \{\infty\}$. The *$\varepsilon$-smooth Rényi entropy or order α* of a probability distribution P is defined by

$$H_\alpha^\varepsilon(P) := H_\alpha^{\max}(\mathcal{B}^\varepsilon(P)) \quad (\text{for } \alpha > 1)$$

and

$$H_\alpha^\varepsilon(P) := H_\alpha^{\min}(\mathcal{B}^\varepsilon(P)) \quad (\text{for } \alpha < 1)\ .$$

Similarly, the notion of conditional Rényi entropy can be generalized to smooth Rényi entropy. In particular, for $\alpha = \infty$, we have

$$H_\infty^\varepsilon(Z|W) := \max_{P_{Z'W'}:\delta(P_{Z'W'},P_{ZW})\leq\varepsilon} H_\infty(Z'|W')\ .$$

The following lemma is an immediate consequence of the above definition for $\alpha = 0$.

Lemma 3.6. *Let Z be a random variable with range $\mathcal{Z}$ and let $\mathcal{W}$ be a subset of $\mathcal{Z}$. Then, for any $\varepsilon \geq 0$,*

$$\text{Prob}[Z \in \mathcal{W}] \geq 1-\varepsilon \quad \Longrightarrow \quad H_0^\varepsilon(P_Z) \leq \log|\mathcal{W}|\ .$$

For $\alpha = 0$, the (smooth) Rényi entropy is sub-additive.

Lemma 3.7. *Let Z and W be two random variables. Then, for any $\varepsilon, \varepsilon' > 0$,*

$$H_0^{\varepsilon+\varepsilon'}(ZW) \leq H_0^\varepsilon(Z) + H_0^{\varepsilon'}(W)\ .$$

The min-entropy of a random variable Z when conditioning on another random variable W cannot decrease more than the Rényi entropy of order zero of W.

Lemma 3.8. *Let Z and W be random variables. Then, for any $\varepsilon, \varepsilon', \varepsilon'' \in \mathbb{R}^+$,*

$$H_\infty^{\varepsilon+\varepsilon'+\varepsilon''}(Z|W) \geq H_\infty^{\varepsilon}(ZW) - H_0^{\varepsilon'}(W) - \log(\frac{1}{\varepsilon''}) \ .$$

Lemma 3.9. *Let $\mathbf{Z}$ be an exchangeable n-tuple of random variables with range $\mathcal{Z}$. Then*

$$H_\infty(\mathbf{Z}|Q_{\mathbf{Z}} = Q) \geq nH(Q) - |\mathcal{Z}|(\log(n) + 1) \ .$$

Proof. By the definition of exchangeability, $P_{\mathbf{Z}|Q_{\mathbf{Z}}=Q}$ is the uniform distribution over the set of all n-tuples $\mathbf{z}$ with $Q_{\mathbf{z}} = Q$. It is easy to see that there are

$$N_Q := \frac{n!}{\prod_{z\in\mathcal{Z}}(nQ(z))!}$$

such tuples, i.e., we have $H_\infty(\mathbf{Z}|Q_{\mathbf{Z}} = Q) = -\log p_{\max}(P_{\mathbf{Z}|Q_{\mathbf{Z}}=Q}) = \log(N_Q)$. The assertion then follows from a straightforward calculation using Stirling's approximation

$$\sqrt{2\pi}m^{m+\frac{1}{2}}e^{-m} \leq m! \leq m^{m+\frac{1}{2}}e^{-m+1} \ ,$$

for any $m \in \mathbb{N}$. □

The notion of typical sets is widely used in information theory. Note that the following definition slightly differs from the one given in [26].

Definition 3.10. Let $\mathcal{Z}$ be a set, $n \in \mathbb{N}$, and $r \geq 0$. The *r-typical set over $\mathcal{Z}^n$* is defined as

$$\mathcal{T}_{\mathcal{Z}}^n(r) := \{\mathbf{z} \in \mathcal{Z}^n : H(Q_{\mathbf{z}}) \leq r\} \ .$$

Lemma 3.11. *For any set $\mathcal{Z}$ of size $|\mathcal{Z}| = q$, $n, \in \mathbb{N}$, and $r \geq 0$,*

$$|\mathcal{T}_{\mathcal{Z}}^n(r)| \leq 2^{nr}n^{q-1} \ .$$

Proof. Let $\mathcal{Q} := \{Q_{\mathbf{z}} : \mathbf{z} \in \mathcal{Z}^n\}$ be the set of frequency distributions of n-tuples over $\mathcal{Z}$ and, for any $\hat{Q} \in \mathcal{Q}$, let $\mathcal{S}(\hat{Q}) := \{\mathbf{z} \in \mathcal{Z}^n : Q_{\mathbf{z}} = \hat{Q}\}$ be the set of n-tuples $\mathbf{z} = (z_1, \ldots, z_n)$ with frequency distribution $\hat{Q}$. We first show that, for any $\hat{Q} \in \mathcal{Q}$,

$$|\mathcal{S}(\hat{Q})| \leq 2^{nH(\hat{Q})} \ . \tag{7}$$

Let $\mathbf{Z} = (Z_1, \ldots, Z_n)$ be an n-tuple of independent random variables Z_i distributed according to $\hat{Q}$. Since, for any n-tuple $\mathbf{z}$ in $\mathcal{S}(\hat{Q})$, each symbol z occurs $n\hat{Q}(z)$ times in $\mathbf{z}$, we find

$$1 \geq \text{Prob}[\mathbf{Z} \in \mathcal{S}(\hat{Q})] = |\mathcal{S}(\hat{Q})| \prod_{z \in \mathcal{Z}} \hat{Q}(z)^{n\hat{Q}(z)} = |\mathcal{S}(\hat{Q})| 2^{\sum_{z \in \mathcal{Z}} n\hat{Q}(z) \log(\hat{Q}(z))}$$

which implies (7). The assertion of the lemma then follows from

$$|\mathcal{T}^n_{\mathcal{Z}}(r)| = \sum_{\hat{Q} \in \mathcal{Q}:\, H(\hat{Q}) \leq r} |\mathcal{S}(\hat{Q})| \leq \sum_{\hat{Q} \in \mathcal{Q}:\, H(\hat{Q}) \leq r} 2^{nH(\hat{Q})} \leq |\mathcal{Q}|\, 2^{nr}$$

and the observation that $|\mathcal{Q}| \leq n^{q-1}$. □

Definition 3.12. Let $p \in [0,1]$ and let $\mathcal{I}$ be a set. A *p-random selection* A on $\mathcal{I}$ is a random variable describing the subset obtained by independently picking each element of $\mathcal{I}$ with probability p, i.e., for any $a \subseteq \mathcal{I}$,

$$P_A(a) = \prod_{i \in \mathcal{I}} P_p^{\text{bin}}(\chi_a(i))$$

where χ_a be the characteristic function of a on $\mathcal{I}$.

A random function G from $\mathcal{X}$ to $\mathcal{Y}$ is called *two-universal* if $\text{Prob}[G(x) = G(x')] \leq \frac{1}{|\mathcal{Y}|}$ holds for any distinct $x, x' \in \mathcal{X}$. In particular, G is two-universal if, for any distinct $x, x' \in \mathcal{X}$, the random variables $G(x)$ and $G(x')$ are independent and uniformly distributed. For instance, the uniform random function from a set $\mathcal{X}$ to a set $\mathcal{Y}$ is two-universal.[3]

3.3 Elements of Quantum Theory

In this section, we introduce some basic concepts of quantum theory which we will use. For a more complete overview we refer to the standard literature (e.g., [31]).

Let $\mathcal{H}$ be a Hilbert space of dimension d. We denote by $\mathcal{S}(\mathcal{H})$ the set of *density operators* on $\mathcal{H}$, i.e., $\mathcal{S}(\mathcal{H})$ is the set of positive operators ρ on $\mathcal{H}$ with $\text{tr}(\rho) = 1$. For any $\rho \in \mathcal{S}(\mathcal{H})$, let $\boldsymbol{\lambda}(\rho)$ be the d-tuple of eigenvalues of ρ

[3] In the literature, two-universality is usually defined for families $\mathcal{G}$ of functions: A family $\mathcal{G}$ is called two-universal if the random function G with uniform distribution over $\mathcal{G}$ is two-universal. Non-trivial examples of two-universal families $\mathcal{G}$ of functions can, e.g., be found in [29] and [30].

(e.g., in decreasing order). The *trace distance* between two density operators ρ and σ on the same Hilbert space $\mathcal{H}$ is defined by

$$\delta(\rho, \sigma) := \frac{1}{2}\mathrm{tr}(|\rho - \sigma|) \ .$$

We will use several well-known properties of the trace distance (for proofs, see e.g. [31]).

The trace distance can be seen as a generalization of the variational distance to density operators. Many of the properties of the variational distance thus also hold for the trace distance. In particular, the trace distance is a metric on $\mathcal{S}(\mathcal{H})$.

Moreover, for two probability distributions P and Q over $\mathcal{W}$ and two families of density operators $\{\rho_w\}_{w\in\mathcal{W}}$ and $\{\sigma_w\}_{w\in\mathcal{W}}$,

$$\delta(\sum_{w\in\mathcal{W}} P(w)\rho_w, \sum_{w\in\mathcal{W}} Q(w)\sigma_w) \leq \sum_{w\in\mathcal{W}} P(w)\,\delta(\rho_w, \sigma_w) + \delta(P, Q) \ . \quad (8)$$

This inequality can be seen as the quantum analogue of (6). The trace distance between two *pure states* $\rho = |\phi\rangle\langle\phi|$ and $\sigma = |\psi\rangle\langle\psi|$ can easily be computed explicitly,

$$\delta(\rho, \sigma) = \sqrt{1 - |\langle\phi|\psi\rangle|^2} = \sqrt{1 - \mathrm{tr}(\rho\sigma)} \ . \quad (9)$$

Let $\mathcal{F}$ be a *positive operator valued measure (POVM)* on a Hilbert space $\mathcal{H}$, i.e., $\mathcal{F} = \{F_z\}_{z\in\mathcal{Z}}$ is a family of positive operators on $\mathcal{H}$ such that $\sum_{z\in\mathcal{Z}} F_z = \mathrm{id}$. We say that $\mathcal{F}$ is *orthogonal* if there exists an orthonormal basis $\{|z\rangle\}_{z\in\mathcal{Z}}$ of $\mathcal{H}$ such that $F_z = |z\rangle\langle z|$, for any $z \in \mathcal{Z}$.

Definition 3.13. Let $\mathcal{F} = \{F_z\}_{z\in\mathcal{Z}}$ be a POVM on a Hilbert space $\mathcal{H}$. The *measurement mapping* $\gamma_\mathcal{F}$ of $\mathcal{F}$ is the function mapping each density operator $\rho \in \mathcal{S}(\mathcal{H})$ to the probability distribution $P = \gamma_\mathcal{F}(\rho)$ on $\mathcal{Z}$ defined by $P(z) := \mathrm{tr}(F_z\rho)$. The *probability range* $\mathcal{P}_\mathcal{F}$ of $\mathcal{F}$ is the range of $\gamma_\mathcal{F}$, i.e., $\mathcal{P}_\mathcal{F} := \gamma_\mathcal{F}(\mathcal{S}(\mathcal{H}))$.

For a POVM $\mathcal{F} = \{F_z\}_{z\in\mathcal{Z}}$ on a Hilbert space $\mathcal{H}$ and a probability distribution P on $\mathcal{Z}$, we write $\gamma_\mathcal{F}^{-1}(P)$ to denote the set of density operators ρ on $\mathcal{H}$ such that $\gamma_\mathcal{F}(\rho) = P$. More generally, for a set $\mathcal{P}$ of probability distributions, $\gamma_\mathcal{F}^{-1}(\mathcal{P}) := \bigcup_{P\in\mathcal{P}} \gamma_\mathcal{F}^{-1}(P)$ is the set of density operators ρ with $\gamma_\mathcal{F}(\rho) \in \mathcal{P}$

The trace distance between two density operators ρ and σ turns out to be an upper bound for the variational distance between the probability distributions of the outcomes of the same measurement $\mathcal{F}$ applied to ρ and σ.

Lemma 3.14. *Let $\mathcal{F}$ be a POVM on a Hilbert space $\mathcal{H}$ and let $\rho, \sigma \in \mathcal{S}(\mathcal{H})$. Then*

$$\delta(\gamma_{\mathcal{F}}(\rho), \gamma_{\mathcal{F}}(\sigma)) \leq \delta(\rho, \sigma) \ .$$

The probability distribution resulting from an orthogonal measurement of a quantum state ρ is in a certain sense less ordered than the eigenvalues of ρ. This is formalized by the following lemma. A proof can, for instance, be found in [32] (see also [33]).

Lemma 3.15 (Schur's majorization theorem). *Let $\mathcal{F} = \{F_1, \ldots, F_d\}$ be an orthogonal measurement on a d-dimensional Hilbert space. Then, for any density operator $\rho \in \mathcal{S}(\mathcal{H})$,*

$$\mathbf{p} \prec \boldsymbol{\lambda}(\rho)$$

where $\mathbf{p} = (\gamma_{\mathcal{F}}(\rho)(1), \ldots, \gamma_{\mathcal{F}}(\rho)(d))$ are the probabilities of the outcomes when measuring ρ with respect to $\mathcal{F}$.

Let $\mathcal{H} \otimes \mathcal{H}'$ be a bipartite Hilbert space, let $\rho \in \mathcal{S}(\mathcal{H} \otimes \mathcal{H}')$, and let Z be the outcome of a measurement of ρ with respect to a POVM $\mathcal{F} = \{F_z\}_{z \in \mathcal{Z}}$ on (a subspace of) $\mathcal{H}'$. The density operator on $\mathcal{H}$ resulting from conditioning ρ on the measurement outcome $Z = z$, denoted $\rho^{\mathcal{H}}_{z \leftarrow \mathcal{F}}$, is given by

$$\rho^{\mathcal{H}}_{z \leftarrow \mathcal{F}} := \frac{1}{c} \mathrm{tr}_{\mathcal{H}'}(\mathrm{id}_{\mathcal{H}} \otimes F_z \, \rho)$$

where $c := \mathrm{tr}(\mathrm{id}_{\mathcal{H}} \otimes F_z \, \rho)$ is a normalization constant and where $\mathrm{tr}_{\mathcal{H}'}$ denotes the partial trace over the subspace $\mathcal{H}'$.

Let $\mathcal{H}^{\otimes n} := \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$ be the product of n identical factor spaces $\mathcal{H}_i = \mathcal{H}$. The following definition can be seen as a quantum version of Definition 3.3.

Definition 3.16. The *density range* of a density operator $\rho \in \mathcal{H}^{\otimes n}$ is the smallest convex subset $\mathcal{P}$ of $\mathcal{S}(\mathcal{H})$ such that for any $k \in \{1, \ldots, n\}$, for any POVM $\mathcal{F}^{k-1} = \{F_z\}_{z \in \mathcal{Z}}$ on $\mathcal{H}^{k-1} := \bigotimes_{i=1}^{k-1} \mathcal{H}_i$, and for any $z \in \mathcal{Z}$, the density operator $\rho^{\mathcal{H}_k}_{z \leftarrow \mathcal{F}^{k-1}}$ is contained in $\mathcal{P}$.

Let $\mathcal{E}$ be a *quantum operation* $\mathcal{E}$ on a Hilbert space $\mathcal{H}$, i.e., $\mathcal{E} = \{E_z\}_{z \in \mathcal{Z}}$ is a family of linear operators E_z on $\mathcal{H}$ such that $\sum_{z \in \mathcal{Z}} E_z^\dagger E_z = \mathrm{id}$. Then, the density operator $\sigma = \mathcal{E}(\rho)$ *resulting from applying* $\mathcal{E}$ to a density operator ρ is given by

$$\sigma := \sum_{z \in \mathcal{Z}} E_z \rho E_z^\dagger \ .$$

Lemma 3.17. *Let ρ be a density operator on $\mathcal{H}$ and let $\sigma := \mathcal{E}(\rho)$ be the density operator resulting from applying a quantum operation $\mathcal{E} = \{E_z\}_{z\in\mathcal{Z}}$ to ρ. Then*

$$\delta(\rho,\sigma) \leq \sqrt{1 - \sum_{z\in\mathcal{Z}} |\mathrm{tr}(E_z\rho)|^2} \ .$$

Proof. We first show that the assertion of the lemma holds if $\rho = |\phi\rangle\langle\phi|$ is a pure state. For $z \in \mathcal{Z}$, let $p_z := \mathrm{tr}(E_z\rho E_z^\dagger)$, $|\psi_z\rangle := \frac{1}{\sqrt{p_z}} E_z|\phi\rangle$, and $\sigma_z := |\psi_z\rangle\langle\psi_z|$. Note that $p_z \in [0,1]$, $\sum_z p_z = 1$, and

$$\sigma = \sum_{z\in\mathcal{Z}} p_z\sigma_z \ .$$

We can thus apply (8) yielding

$$\delta(\rho,\sigma) \leq \sum_{z\in\mathcal{Z}} p_z\delta(\rho,\sigma_z) \ . \tag{10}$$

Since, $\rho = |\phi\rangle\langle\phi|$ and $\sigma_z = |\psi_z\rangle\langle\psi_z|$ are pure states, it follows from (9) that

$$\delta(\rho,\sigma_z) = \sqrt{1 - |\langle\phi|\psi_z\rangle|^2} = \sqrt{1 - \frac{1}{p_z}|\mathrm{tr}(E_z\rho)|^2} \ .$$

Combining this with (10), we find

$$\delta(\rho,\sigma) \leq \sum_{z\in\mathcal{Z}} p_z \sqrt{1 - \frac{1}{p_z}|\mathrm{tr}(E_z\rho)|^2} \leq \sqrt{\sum_{z\in\mathcal{Z}} p_z\big(1 - \frac{1}{p_z}|\mathrm{tr}(E_z\rho)|^2\big)}$$

where the second inequality follows from the concavity of the square root and Jensen's inequality. This concludes the proof of the lemma for pure states ρ.

To verify that the assertion of the lemma also holds for mixed states ρ, write ρ as a convex combination of pure states ρ_w, i.e., $\rho = \sum_{w\in\mathcal{W}} q_w\rho_w$ for appropriate $q_w \in [0,1]$ with $\sum_{w\in\mathcal{W}} q_w = 1$, and let $\mathcal{E}(\rho_w)$ be the state resulting from applying the quantum operation $\mathcal{E}$ on ρ_w. Then, since $\sigma = \mathcal{E}(\rho) = \sum_{w\in\mathcal{W}} q_w\mathcal{E}(\rho_w)$, inequality (8) yields

$$\delta(\rho,\sigma) \leq \sum_{w\in\mathcal{W}} q_w\delta(\rho_w,\mathcal{E}(\rho_w)) \leq \sum_{w\in\mathcal{W}} q_w \sqrt{1 - \sum_{z\in\mathcal{Z}} |\mathrm{tr}(E_z\rho_w)|^2}$$

where the last inequality follows from the statement of the lemma applied to the pure states ρ_w. Using again Jensen's inequality, we obtain

$$\delta(\rho,\sigma) \leq \sqrt{1 - \sum_{z\in\mathcal{Z}}\sum_{w\in\mathcal{W}} q_w|\mathrm{tr}(E_z\rho_w)|^2} \leq \sqrt{1 - \sum_{z\in\mathcal{Z}} |\mathrm{tr}(E_z \sum_{w\in\mathcal{W}} q_w\rho_w)|^2}$$

which concludes the proof. □

We will now use Lemma 3.17 to derive a lower bound for the variational distance between two probability distributions in terms of the trace distance between two corresponding density operators. This is in a certain sense the converse of Lemma 3.14.

Lemma 3.18. *Let $\mathcal{F} = \{F_z\}_{z\in\mathcal{Z}}$ be an orthogonal POVM on a Hilbert space $\mathcal{H}$, let $\rho \in \mathcal{S}(\mathcal{H})$, let $P := \gamma_{\mathcal{F}}(\rho)$, and let Q be a probability distribution on $\mathcal{Z}$. Then there exists $\sigma \in \mathcal{S}(\mathcal{H})$ such that*

$$Q = \gamma_{\mathcal{F}}(\sigma) \tag{11}$$

and

$$\delta(\rho, \sigma) \leq \sqrt{2\delta(P, Q)} \ . \tag{12}$$

In particular,

$$\mathcal{B}^{\varepsilon}(\gamma_{\mathcal{F}}(\rho)) \subseteq \gamma_{\mathcal{F}}(\mathcal{B}^{\sqrt{2\varepsilon}}(\rho)) \ . \tag{13}$$

Proof. From Lemma 3.1, there exist random variables Z and Z' distributed according to P and Q, respectively, such that

$$\text{Prob}[Z \neq Z'] = \delta := \delta(P, Q) \ .$$

Let $\mathcal{W} := \{(z, z') \in \mathcal{Z} \times \mathcal{Z} : z \neq z'\}$, let $p_{z,z'} := P_{Z'|Z}(z', z)$, and let $\{|z\rangle\}_{z\in\mathcal{Z}}$ be an orthonormal basis of $\mathcal{H}$ such that $F_z = |z\rangle\langle z|$. Let

$$E_0 := \sum_{z\in\mathcal{Z}} \sqrt{p_{z,z}}|z\rangle\langle z| \ ,$$

and, for $(z, z') \in \mathcal{W}$,

$$E_{z,z'} := \sqrt{p_{z,z'}}|z'\rangle\langle z| \ ,$$

be linear operators on $\mathcal{H}$. It is easy to verify that the family $\mathcal{E} = \{E_0\} \cup \{E_{z,z'}\}_{(z,z')\in\mathcal{W}}$ is a quantum operation, i.e.,

$$E_0^{\dagger}E_0 \ + \sum_{(z,z')\in\mathcal{W}} E_{z,z'}^{\dagger}E_{z,z'} = \text{id} \ .$$

Let

$$\sigma := \mathcal{E}(\rho) = E_0\rho E_0^{\dagger} \ + \sum_{(z,z')\in\mathcal{W}} E_{z,z'}\rho E_{z,z'}^{\dagger}$$

be the quantum state resulting from applying $\mathcal{E}$ to ρ. It then follows from a straightforward calculation that

$$\gamma_{\mathcal{F}}(\sigma) = P_{Z'}$$

which implies (11) since $P_{Z'} = Q$. To show that also (12) holds, we use Lemma 3.17 yielding

$$\delta(\rho,\sigma) \leq \sqrt{1 - |\mathrm{tr}(E_0\rho)|^2 - \sum_{(z,z')\in\mathcal{W}} |\mathrm{tr}(E_{z,z'}\rho)|^2} \leq \sqrt{1 - |\mathrm{tr}(E_0\rho)|^2} \ .$$

Since $P_Z = \gamma_{\mathcal{F}}(\rho)$, we have

$$\mathrm{tr}(E_0\rho) = \sum_{z\in\mathcal{Z}} \sqrt{p_{z,z}}\gamma_{\mathcal{F}}(\rho)(z) \geq \sum_{z\in\mathcal{Z}} p_{z,z}P_Z(z) = \mathrm{Prob}[Z = Z'] = 1 - \delta \ ,$$

and thus

$$\delta(\rho,\sigma) \leq \sqrt{1-(1-\delta)^2} = \sqrt{2\delta - \delta^2} \leq \sqrt{2\delta} \ .$$

□

The entropy of a quantum state can be defined in terms of the entropy of a classical probability distribution. Let ρ be a density operator on a d-dimensional Hilbert space $\mathcal{H}$ and let $(\lambda_1, \ldots, \lambda_d) := \boldsymbol{\lambda}(\rho)$ be the d eigenvalues of ρ. Note that there exists an orthonormal basis $\{|1\rangle, \ldots |d\rangle\}$ of ρ (namely the eigenbasis) such that $\lambda_i = P(i)$ where $P := \gamma_{\mathcal{F}}(\rho)$ is the probability distribution of a measurement of ρ with respect to the POVM $\mathcal{F} = \{|1\rangle\langle 1|, \ldots, |d\rangle\langle d|\}$. In particular $\boldsymbol{\lambda}(\rho)$ can be interpreted as a probability distribution on $\{1, \ldots, d\}$.

The *Rényi entropy* (of order α) of a density operator ρ is defined by the Rényi entropy of $\boldsymbol{\lambda}(\rho)$, i.e., $S_\alpha(\rho) := H_\alpha(\boldsymbol{\lambda}(\rho))$, for $\alpha \in \mathbb{R}^+ \cup \{\infty\}$. In particular, for $\alpha = 1$, $S(\rho) := S_1(\rho)$ is the *von Neumann entropy* of ρ. Note that, for $\alpha = 0$,

$$S_0(\rho) = \log(\mathrm{rank}(\rho)) \ .$$

The smooth Rényi entropy for density operators can be defined by generalizing the classical Definition 3.5.

Definition 3.19. Let $\varepsilon \geq 0$ and $\alpha \in \mathcal{R}^+ \cup \{\infty\}$. The *$\varepsilon$-smooth Rényi entropy of order α* of a density operator ρ is defined by

$$S_\alpha(\rho) := S_\alpha^{\max}(\mathcal{B}^\varepsilon(\rho)) \quad (\text{for } \alpha > 1) \quad \text{and} \quad S_\alpha(\rho) := S_\alpha^{\min}(\mathcal{B}^\varepsilon(\rho)) \quad (\text{for } \alpha < 1) \ .$$

The following lemma is a direct consequence of Lemma 3.15 and the fact that the entropy functions $-H_\alpha$ are Schur-convex.

Lemma 3.20. *Let $\mathcal{F}$ be an orthogonal POVM on a d-dimensional Hilbert space. Then, for any density operator $\rho \in \mathcal{S}(\mathcal{H})$ and any $\alpha \in \mathbb{R}^+ \cup \{\infty\}$,*

$$S_\alpha(\rho) \leq H_\alpha(\gamma_{\mathcal{F}}(\rho)) \ .$$

We often will use this result for the case $\alpha = 1$. To simplify the notation, let

$$S_{\mathcal{F}}(\rho) := H(\gamma_{\mathcal{F}}(\rho)). \tag{14}$$

be the Shannon entropy of the outcomes when measuring a density operator $\rho \in \mathcal{S}(\mathcal{H})$ with respect to a POVM $\mathcal{F}$. If $\mathcal{F}$ corresponds to a measurement in an eigenbasis of ρ, we obviously have $S(\rho) = S_{\mathcal{F}}(\rho)$, and thus, from Lemma 3.20,

$$S(\rho) = \min_{\mathcal{F}} S_{\mathcal{F}}(\rho) \tag{15}$$

where the minimum is taken over all orthogonal POVMs $\mathcal{F}$ in $\mathcal{H}$.

The following lemma is an extension of Lemma 3.20 to smooth Rényi entropy.

Lemma 3.21. *Let $\mathcal{F}$ be an orthogonal POVM on a Hilbert space $\mathcal{H}$. Then, for any density operator $\rho \in \mathcal{S}(\mathcal{H})$. Then, for any density operator ρ, $\alpha < 1$, and $\varepsilon \geq 0$,*

$$S_\alpha^{\sqrt{2\varepsilon}}(\rho) \leq H_\alpha^\varepsilon(\gamma_{\mathcal{F}}(\rho)) \ .$$

Proof. From Lemma 3.20, we have

$$S_\alpha^{\sqrt{2\varepsilon}}(\rho) = \inf_{\sigma \in \mathcal{B}^{\sqrt{2\varepsilon}}(\rho)} S_\alpha(\sigma) \leq \inf_{\sigma \in \mathcal{B}^{\sqrt{2\varepsilon}}(\rho)} H_\alpha(\gamma_{\mathcal{F}}(\sigma)) \ .$$

The assertion then follows from Lemma 3.18,

$$\inf_{\sigma \in \mathcal{B}^{\sqrt{2\varepsilon}}(\rho)} H_\alpha(\gamma_{\mathcal{F}}(\sigma)) \leq \inf_{Q \in \mathcal{B}^\varepsilon(\gamma_{\mathcal{F}}(\rho))} H_\alpha(Q) = H_\alpha^\varepsilon(\gamma_{\mathcal{F}}(\rho)) \ .$$

□

4 Main Result

This section contains the main result of the paper, namely, an explicit expression for the rate of secure quantum key distribution (cf. equation (22)).

In the first part, we derive Lemma 4.1 which says that the frequency distribution obtained when measuring the subsystems of an n-partite quantum state with respect to a certain POVM $\bar{\mathcal{F}}$ can be estimated from the results obtained by applying another POVM $\mathcal{F}$ on a few randomly chosen subsystems. This is then used to show Lemma 4.2 which gives an upper bound for the Rényi entropy of order 0 of the outcomes when applying the POVM $\mathcal{F}$ given only the outcomes of the measurements with respect to $\bar{\mathcal{F}}$ on a few (randomly chosen) subsystems. The result is then applied to bound the size (rank) of the n-partite quantum system given the outcomes of a measurement on a few subsystems (Corollary 4.3).

In Sections 4.2 and 4.3 we review information reconciliation and the security of privacy amplification in the presence of a quantum adversary, respectively. These are main ingredients of the post–processing stage.

In Section 4.4, we introduce the generic quantum key distribution protocol and prove its security by combining the above mentioned results with the information reconciliation and privacy amplification to obtain our main result, i.e., the secret key rate (22).

4.1 Parameter Estimation

Let $\mathcal{H}$ be a Hilbert space, let $\rho \in \mathcal{S}(\mathcal{H}^{\otimes n})$, let $a \subseteq \{1, \ldots, n\}$, and let $\mathcal{F} = \{F_z\}_{z \in \mathcal{Z}}$ be a POVM on $\mathcal{H}$. Then $\Gamma^a_{\mathcal{F}}(\rho)$ denotes the $|a|$-tuple $\hat{\mathbf{Z}}$ of outcomes resulting from applying $\mathcal{F}$ to ρ on $\mathcal{H}_a$, where $\mathcal{H}_a$ is the tensor product of the factor spaces $\mathcal{H}_i$, for $i \in a$.

Lemma 4.1. *Let $\rho \in \mathcal{S}(\mathcal{H}^{\otimes n})$ be an n-partite state with density range $\mathcal{R} \subseteq \mathcal{S}(\mathcal{H})$, let $\mathcal{F} = \{F_z\}_{z \in \mathcal{Z}}$ and $\bar{\mathcal{F}} = \{\bar{F}_z\}_{z \in \bar{\mathcal{Z}}}$ be two POVMs on $\mathcal{H}$, and let A be a p-random selection on $\{1, \ldots, n\}$. Let $\mathbf{Z}_A := \Gamma^A_{\mathcal{F}}(\rho)$ and $\mathbf{Z}_{\bar{A}} := \Gamma^{\bar{A}}_{\bar{\mathcal{F}}}(\rho)$ be the outcomes when measuring ρ in $\mathcal{H}_A$ and $\mathcal{H}_{\bar{A}}$ with respect to $\mathcal{F}$ and $\bar{\mathcal{F}}$, respectively. Then, for any $\varepsilon > 0$,*

$$\mathrm{Prob}\big[\exists \hat{\rho} \in \mathcal{R} : \; p\,\delta(Q_{\mathbf{Z}_A}, \gamma_{\mathcal{F}}(\hat{\rho})) + (1-p)\,\delta(Q_{\mathbf{Z}_{\bar{A}}}, \gamma_{\bar{\mathcal{F}}}(\hat{\rho})) \leq \varepsilon\big] \geq 1 - \mu$$

where $\mu := 2^{|\mathcal{Z}|+|\bar{\mathcal{Z}}|} e^{-\frac{n\varepsilon^2}{8}}$.

Proof. Let $\mathcal{G}$ be the POVM on $\mathcal{H}$ obtained by combining $\mathcal{F}$ and $\bar{\mathcal{F}}$ with probability p and $1-p$, respectively, i.e., $\mathcal{G} := \{G_{(z,r)}\}_{(z,r) \in (\mathcal{Z} \cup \bar{\mathcal{Z}}) \times \{0,1\}}$ with $G_{(z,1)} := pF_z$ and $G_{(z,0)} := (1-p)\bar{F}_z$. Let $\mathbf{W} := \Gamma_{\mathcal{G}}(\rho)$ be the n-tuple of outcomes (Z_i, R_i) when measuring ρ with respect to $\mathcal{G}$. The random variables occurring in the lemma can then equivalently be defined by $A := \{i : R_i = 1\}$, $\mathbf{Z}_A := (Z_1, \ldots, Z_n)_A$, and $\mathbf{Z}_{\bar{A}} := (Z_1, \ldots, Z_n)_{\bar{A}}$.

The probability range of the n-tuple $\mathbf{W}$ is contained in $\mathcal{P} := \gamma_{\mathcal{G}}(\mathcal{R})$. We can thus apply Lemma 3.4 for $\bar{\varepsilon} := \varepsilon/2$ leading to

$$\mathrm{Prob}[Q_{\mathbf{W}} \in \mathcal{B}^{\bar{\varepsilon}}(\mathcal{P})] \geq 1 - \mu \ .$$

Let $\mathbf{z} \in (\mathcal{Z} \cup \bar{\mathcal{Z}})^n$ and $a \subseteq \{1, \ldots, n\}$ such that the n-tuple $\mathbf{w}$ of values $w_i = (z_i, \chi_a(i))$ satisfies $Q_{\mathbf{w}} \in \mathcal{B}^{\bar{\varepsilon}}(\mathcal{P})$, i.e.,

$$\delta(Q_{\mathbf{w}}, \gamma_{\mathcal{G}}(\hat{\rho})) \leq \varepsilon/2$$

for some $\hat{\rho} \in \mathcal{R}$. It remains to be shown that this implies

$$p\,\delta(Q_{\mathbf{z}_a}, \gamma_{\mathcal{F}}(\hat{\rho})) + (1-p)\,\delta(Q_{\mathbf{z}_{\bar{a}}}, \gamma_{\bar{\mathcal{F}}}(\hat{\rho})) \leq \varepsilon \ . \tag{16}$$

Let (Z, R) and (Z', R') be two pairs of random variables distributed according to $\gamma_{\mathcal{G}}(\hat{\rho})$ and $Q_{\mathbf{w}}$, respectively. It follows from the construction of the POVM $\mathcal{G}$ that $P_R = P_p^{\mathrm{bin}}$, $P_{Z|R=1} = \gamma_{\mathcal{F}}(\hat{\rho})$, and $P_{Z|R=0} = \gamma_{\bar{\mathcal{F}}}(\hat{\rho})$. Moreover, by the definition of the frequency distribution, $P_{Z'|R'=1} = Q_{\mathbf{z}_a}$ and $P_{Z'|R'=0} = Q_{\mathbf{z}_{\bar{a}}}$. Hence, using (5),

$$\begin{aligned} p\,\delta(\gamma_{\mathcal{F}}(\hat{\rho}), Q_{\mathbf{z}_a}) + (1-p)\,\delta(\gamma_{\bar{\mathcal{F}}}(\hat{\rho}), Q_{\mathbf{z}_{\bar{a}}}) &= E_R[P_{Z|R}(\cdot, R), P_{Z'|R'}(\cdot, R)] \\ &\leq 2\delta(P_{ZR}, P_{Z'R'}) = 2\delta(\gamma_{\mathcal{G}}(\hat{\rho}), Q_{\mathbf{w}}) \leq \varepsilon \ . \end{aligned}$$

which implies (16) and thus concludes the proof. □

Lemma 4.2. *Let $\rho \in \mathcal{S}(\mathcal{H}^{\otimes n})$ be an n-partite state with density range $\mathcal{R} \subseteq \mathcal{S}(\mathcal{H})$, let $\mathcal{F} = \{F_z\}_{z \in \mathcal{Z}}$ and $\bar{\mathcal{F}} = \{\bar{F}_z\}_{z \in \bar{\mathcal{Z}}}$ be two POVMs on $\mathcal{H}$, let A be a p-random selection on $\{1, \ldots, n\}$, and let $\mathbf{Z}_A := \Gamma^A_{\mathcal{F}}(\rho)$, $\mathbf{Z}_{\bar{A}} := \Gamma^{\bar{A}}_{\bar{\mathcal{F}}}(\rho)$. Then, for any $\varepsilon > 0$, there exists a real valued function μ with*

$$E[\mu(Q_{\mathbf{Z}_A}, A)] \leq 2^{|\mathcal{Z}|+|\mathcal{Z}'|} e^{-\frac{n\varepsilon^2}{2}}$$

such that, for any probability distribution $\hat{Q}$ on $\mathcal{Z}$ and any $a \subseteq \{1, \ldots, n\}$,

$$H_0^{\mu(\hat{Q},a)}(\mathbf{Z}_{\bar{A}} | Q_{\mathbf{Z}_A} = \hat{Q}, A = a) \leq |\bar{a}|\, H^{\max}(\mathcal{B}(\hat{Q})) + \log(|\bar{a}|)(|\bar{\mathcal{Z}}| - 1) \ ,$$

where $\mathcal{B}(\hat{Q}) := \mathcal{B}_{\varepsilon/(1-p)}(\gamma_{\bar{\mathcal{F}}}(\mathcal{R} \cap \gamma_{\mathcal{F}}^{-1}(\mathcal{B}_{\varepsilon/p}(\hat{Q}))))$.

Proof. Let $\mathcal{W}$ be the set of pairs $(\mathbf{z}, a)$ consisting of an n-tuple $\mathbf{z}$ of elements from $\mathcal{Z} \cup \bar{\mathcal{Z}}$ and a subset $a \subseteq \{1, \ldots, n\}$ such that there exists a density operator $\hat{\rho} \in \mathcal{R}$ satisfying

$$p\,\delta(Q_{\mathbf{z}_a}, \gamma_{\mathcal{F}}(\hat{\rho})) + (1-p)\,\delta(Q_{\mathbf{z}_{\bar{a}}}, \gamma_{\bar{\mathcal{F}}}(\hat{\rho})) \leq \varepsilon \ .$$

For any probability distribution $\hat{Q}$ on $\mathcal{Z}$ and any $a \subseteq \{1, \ldots, n\}$, let

$$\mathcal{C}(\hat{Q}, a) := \{\mathbf{z}_{\bar{a}} : (\mathbf{z}, a) \in \mathcal{W} \text{ and } Q_{\mathbf{z}_a} = \hat{Q}\} \ .$$

We first show that

$$\log(|\mathcal{C}(\hat{Q}, a)|) \leq |\bar{a}|\, H^{\max}(\mathcal{B}(\hat{Q})) + \log(|\bar{a}|)(|\mathcal{Z}| - 1) \ , \tag{17}$$

for any $\hat{Q}$ and a. It follows from the definition of the set $\mathcal{C}(\hat{Q}, a)$ that for any $\mathbf{z}' \in \mathcal{C}(\hat{Q}, a)$ there exists $\hat{\rho} \in \mathcal{R}$ such that $\delta(\hat{Q}, \gamma_{\mathcal{F}}(\hat{\rho})) \leq \varepsilon/p$ and $\delta(Q_{\mathbf{z}'}, \gamma_{\mathcal{G}}(\hat{\rho})) \leq \varepsilon/(1-p)$. which directly implies $Q_{\mathbf{z}'} \in \mathcal{B}(\hat{Q})$ and thus

$$H(Q_{\mathbf{z}'}) \leq r := H^{\max}(\mathcal{B})(\hat{Q})) \ .$$

Hence, by Definition 3.10, $\mathbf{z}'$ is contained in the r-typical set $\mathcal{T}^k_{\hat{\mathcal{Z}}}(r)$ for $k := |\bar{a}|$. By Lemma 3.11 the size of $\mathcal{T}^k_{\hat{\mathcal{Z}}}(r)$ can not be larger than $2^{kr}|a|^{|\mathcal{Z}|-1}$, from which (17) follows.

Lemma 4.1 gives a lower bound for the probability that $(\mathbf{Z}, A)$ is contained in $\mathcal{W}$,

$$\text{Prob}[(\mathbf{Z}, A) \in \mathcal{W}] \geq 1 - 2^{|\mathcal{Z}|+|\bar{\mathcal{Z}}|} e^{-\frac{np^2}{8}} \ .$$

Let the function μ be defined by

$$\mu(\hat{Q}, a) := 1 - \text{Prob}[\mathbf{Z}_{\bar{A}} \in \mathcal{C}(Q_{\mathbf{Z}_A}, A) | Q_{\mathbf{Z}_A} = \hat{Q}, A = a] \ .$$

Then, since $\text{Prob}[\mathbf{Z}_{\bar{A}} \in \mathcal{C}(Q_{\mathbf{Z}_A}, A)|] \geq \text{Prob}[(\mathbf{Z}, A) \in \mathcal{W}]$, we obtain

$$E[\mu(Q_{\mathbf{Z}_A}, A)] = 1 - \text{Prob}[\mathbf{Z}_{\bar{A}} \in \mathcal{C}(Q_{\mathbf{Z}_A}, A)|] \leq 2^{2q} e^{-\frac{n\varepsilon^2}{8}} \ .$$

On the other hand, from Lemma 3.6,

$$H_0^{\mu(\hat{Q},a)}(\mathbf{Z}_{\bar{A}} | \mathbf{Z}_A = \hat{Q}, A = a) \leq \log(|\mathcal{C}(\hat{Q}, a)|)$$

for any $\hat{Q}$ and a. Combining this with (17) concludes the proof. □

Corollary 4.3. *Let $\rho \in \mathcal{S}(\mathcal{H}^{\otimes n})$ be an n-partite state with density range $\mathcal{R} \subseteq \mathcal{S}(\mathcal{H})$, let $\mathcal{F} = \{F_z\}_{z \in \mathcal{Z}}$ be a POVM on $\mathcal{H}$, let $\bar{\mathcal{F}}$ be an orthogonal POVM on $\mathcal{H}$, and let A be a p-random selection on $\{1, \ldots, n\}$. Let $\mathbf{Z}_A := \Gamma^A_{\mathcal{F}}(\rho)$ be the outcomes when measuring ρ in $\mathcal{H}_A$ with respect to $\mathcal{F}$ and let $\rho_{\bar{A}}$ be the remaining quantum state in $\mathcal{H}_{\bar{A}}$. Then, for any $\varepsilon > 0$, there exists a real valued function μ with*

$$E[\mu(Q_{\mathbf{Z}_A}, A)] \leq 2^{\frac{\dim(\mathcal{H})+|\mathcal{Z}|}{2}} e^{-\frac{n\varepsilon^2}{16}}$$

such that, for any probability distribution $\hat{Q}$ on $\mathcal{Z}$ and any $a \subseteq \{1, \ldots, n\}$,

$$S_0^{\mu(\hat{Q},a)}(\rho_{\bar{A}}|Q_{\mathbf{Z}_A} = \hat{Q}, A = a) \leq |\bar{a}|\, H^{\max}(\mathcal{B}(\hat{Q})) + \log(|\bar{a}|)(\dim(\mathcal{H}) - 1) \ ,$$

where $\mathcal{B}(\hat{Q}) := \mathcal{B}_{\varepsilon/(1-p)}(\gamma_{\bar{\mathcal{F}}}(\mathcal{R} \cap \gamma_{\mathcal{F}}^{-1}(\mathcal{B}_{\varepsilon/p}(\hat{Q}))))$.

Proof. Since the POVM $\bar{\mathcal{F}} = \{\bar{F}_z\}_{z \in \bar{\mathcal{Z}}}$ is orthogonal, we have $|\bar{\mathcal{Z}}| = \dim(\mathcal{H})$. According to Lemma 4.2, there exists a function $\bar{\mu}$ satisfying

$$E[\bar{\mu}(Q_{\mathbf{Z}_A}, A)] \leq 2^{\dim(\mathcal{H})+|\mathcal{Z}|} e^{-\frac{n\varepsilon^2}{8}}$$

such that

$$H_0^{\bar{\mu}(\hat{Q},a)}(Q_{\mathbf{Z}_{\bar{A}}}|Q_{\mathbf{Z}_A} = \hat{Q}, A = a) \leq |\bar{a}|\, H^{\max}(\mathcal{B}(\hat{Q})) + \log(|\bar{a}|)(\dim(\mathcal{H}) - 1) \tag{18}$$

holds. Let the function μ be defined by $\mu(\hat{Q}, a) := \sqrt{2\bar{\mu}(\hat{Q}, a)}$. Using Jensen's inequality, we obtain

$$E[\mu(Q_{\mathbf{Z}_A}, A)] \leq \sqrt{2E[\bar{\mu}(Q_{\mathbf{Z}_A}, A)]} \leq 2^{\frac{\dim(\mathcal{H})+|\mathcal{Z}|}{2}} e^{-\frac{n\varepsilon^2}{16}} \ .$$

On the other hand, since $\bar{\mathcal{F}}$ is orthogonal, Lemma 3.21 implies that

$$S_0^{\mu(\hat{Q},a)}(\rho_{\bar{A}}|Q_{\mathbf{Z}_A} = \hat{Q}, A = a) \leq H_0^{\bar{\mu}(\hat{Q},a)}(\mathbf{Z}_{\bar{A}}|Q_{\mathbf{Z}_A} = \hat{Q}, A = a)$$

for any $\hat{Q}$ and a which, together with (18), concludes the proof. □

Corollary 4.4. *Let $\mathcal{Z}$ be a set and let A be a p-random selection on $\{1, \ldots, n\}$. Then, for any n-tuple $\mathbf{Z}$ of random variables with range $\mathcal{Z}$ and any $\varepsilon > 0$, there exists a real valued function μ with*

$$E[\mu(Q_{\mathbf{Z}_A}, A)] \leq 2^{2|\mathcal{Z}|} e^{-\frac{n\varepsilon^2}{2}}$$

such that, for any probability distribution $\hat{Q}$ on $\mathcal{Z}$ and any $a \subseteq \{1, \ldots, n\}$,

$$H_0^{\mu(\hat{Q},a)}(\mathbf{Z}_{\bar{A}}|Q_{\mathbf{Z}_A} = \hat{Q}, A = a) \leq |\bar{a}|\, H^{\max}(\mathcal{B}^{\varepsilon/p(1-p)}(\hat{Q})) + \log(|\bar{a}|)(|\mathcal{Z}| - 1) \ .$$

Corollary 4.5. *Let $\mathbf{X}$ and $\mathbf{Y}$ be n-tuples of random variables with range $\mathcal{X}$ and $\mathcal{Y}$, respectively, and let A be a p-random selection on $\{1, \ldots, n\}$. Then, for any $\mathbf{y} \in \mathcal{Y}^n$ and $\varepsilon > 0$, there exists a real valued function μ with*

$$E[\mu(Q_{\mathbf{X}_A|\mathbf{Y}_A}, A)] \leq |\mathcal{Y}| 2^{2|\mathcal{X}|} e^{-\frac{n\varepsilon^3}{2}}$$

such that, for any channel $\hat{Q}$ from $\mathcal{Y}$ to $\mathcal{X}$ and for any set $a \subseteq \{1,\ldots,n\}$,

$$H_0^{\mu(\hat{Q},a)}(\mathbf{X}_{\bar{A}}|Q_{\mathbf{X}_A|\mathbf{Y}_A} = \hat{Q}, \mathbf{Y} = \mathbf{y}, A = a) \leq n\,(r+\varepsilon|\mathcal{Y}|\log(|\mathcal{X}|))+\log(n)(|\mathcal{X}|-1)\ .$$

where

$$r := \sum_{y\in\mathcal{Y}} Q_{\mathbf{Y}}(y) H^{\max}(\mathcal{B}^{\varepsilon/p(1-p)}(\hat{Q}(\cdot|y)))\ .$$

Proof. Let $\mathcal{Y}'$ be the subset of $\mathcal{Y}$ containing all values y such that $Q_{\mathbf{y}}(y) \geq \varepsilon n$. For any $y \in \mathcal{Y}$, let

$$a_y := \{i : \mathbf{y}_i = y\}\ ,$$

and, for any $y \in \mathcal{Y}'$, let μ_y be the function defined by Corollary 4.4 applied to the tuple $\mathbf{X}_{a_y}$. In particular, we have, for $y \in \mathcal{Y}'$, any probability distribution $\hat{Q}'$ on $\mathcal{X}$, and any $a \subseteq \{1,\ldots,n\}$,

$$h_y := H_0^{\mu_y(\hat{Q}',a\cap a_y)}(\mathbf{X}_{\bar{A}\cap a_y}|Q_{\mathbf{X}_{A\cap a_y}} = \hat{Q}', A = a) \leq |a_y|\, r_y + \log(n)(|\mathcal{X}|-1)\ .$$

where $r_y := H^{\max}(\mathcal{B}^{\varepsilon/p(1-p)}(\hat{Q}))$. On the other hand, for $y \in \mathcal{Y} - \mathcal{Y}'$, let

$$h_y := H_0(\mathbf{X}_{\bar{A}\cap a_y}|Q_{\mathbf{X}_{A\cap a_y}} = \hat{Q}', A = a) \leq |a_y|\log(|\mathcal{X}|) \leq n\varepsilon\log(|\mathcal{X}|)$$

Applying Lemma 3.7 yields

$$H_0^{\mu(\hat{Q},a)}(\mathbf{X}_{\bar{A}}|Q_{\mathbf{X}_A|\mathbf{Y}_A} = \hat{Q}, \mathbf{Y} = \mathbf{y}, A = a) \leq \sum_{y\in\mathcal{Y}} h_y$$

for

$$\mu(\hat{Q}, a) := \sum_{y\in\mathcal{Y}'} \mu_y(\hat{Q}(\cdot|y), a \cap a_y)$$

from which the assertion follows. □

4.2 Information Reconciliation

Lemma 4.6. *Let Z be a random variable with $H_0^{\varepsilon}(Z) \leq r$ and let F be a two-universal hash function from $\mathcal{Z}$ to $\{0,1\}^s$. Then there exists a guessing function g such that*

$$\text{Prob}[g(F, F(Z)) = Z] \geq 1 - 2^{-(s-r)} + \varepsilon$$

For a proof, see e.g. [34].

4.3 Privacy Amplification Against Quantum Adversaries

We will use the following theorem proven in [1].

Theorem 4.7. *Let Z be a random variable with $H_\infty^\varepsilon(Z) \geq n$ and let $\rho \in \mathcal{S}(\mathcal{H})$ be a density operator with $S_0^{\varepsilon'}(\rho) \leq r$ which depends on X. Let F be a two-universal hash function from $\mathcal{Z}$ to $\{0,1\}^s$ and let $W := \Gamma_{\mathcal{G}}(\rho)$ be the outcome of a measurement of ρ with respect to an arbitrary POVM $\mathcal{G}$ which might depend on F. Then*

$$d(F(Z)|WF) \leq \frac{3}{4} 2^{-\frac{n-r-s}{2}} + \varepsilon + \varepsilon' \ .$$

4.4 A Generic Quantum Key Distribution Protocol

In this section, we will describe the generic protocol and apply the results from the previous sections to prove its security. To enhance the readability of this exposition, we will restrict our attention to the asymptotic behavior of the relevant quantities. The exact statements about eventual constants may be taken directly from the lemmas that we refer to.

Let ρ be a density operator on $(\mathcal{H}_A \otimes \mathcal{H}_B)^{\otimes n}$. Let $\mathcal{F}$ and $\mathcal{G}$ be two POVMs on $\mathcal{H}_A$ and let $\mathcal{F}'$ and $\mathcal{G}'$ be two POVMs on $\mathcal{H}_B$. Let T and T' be two p-random selections on $\{1, \ldots, n\}$. For any $i \in \{1, \ldots, n\}$, let X_i be the outcome of a measurement of the subsystem $(\mathcal{H}_A)_i$ with respect to $\mathcal{F}$, if $i \in T$, or with respect to $\mathcal{G}$, otherwise. Similarly, let Y_i be the outcome of a measurement of $(\mathcal{H}_B)_i$ with respect to $\mathcal{F}'$, if $i \in T'$, or with respect to $\mathcal{G}'$, otherwise. Let S be a p-random selection on $\overline{T_A}$.

For the following asymptotic analysis, we assume that $p = \Theta(n^{-\alpha})$ for some $\alpha \in (0,1)$. In particular, pn grows less than linearly in n.

4.4.1 Parameter Estimation

The goal of this protocol phase is to estimate the parameters used for the subsequent information reconciliation and privacy amplification phase. In particular, Alice and Bob have to determine the minimum length r of the error correcting information needed and the maximum length s of the final key such that it is guaranteed to be secure.

Alice		**Bob**
S, T, $\mathbf{X} = (X_1, \ldots, X_n)$		T', $\mathbf{Y} = (Y_1, \ldots, Y_n)$
	$\xrightarrow{S, T, \mathbf{X}_{S \cup T}}$	$P_{XY} := \hat{Q}(\mathbf{X}_{S \cap \overline{T'}}, \mathbf{Y}_{S \cap \overline{T'}})$, $\mathcal{R} := \mathcal{R}(\mathbf{X}_{T \cap T'}, \mathbf{Y}_{T \cap T'})$
		$r := \lceil nH(X\|Y) \rceil$
		$t := \lfloor nH(X) \rfloor$
		$u := \lceil n \max_{\hat{\rho} \in \mathcal{R}} S(\hat{\rho}) \rceil$
	$\xleftarrow{r, s, T'}$	$s := t - r - u$
$\mathbf{X}' := \mathbf{X}_{\overline{S \cup T \cup T'}}$		$\mathbf{Y}' := \mathbf{X}_{\overline{S \cup T \cup T'}}$

The functions $\hat{Q}$ and $\mathcal{R}$ are defined as follows. Let $\mathbf{x} = (x_1, \ldots, x_k)$ and $\mathbf{y} = (y_1, \ldots, y_k)$ be two k-tuples. Then $\hat{Q}(\mathbf{x}, \mathbf{y})$ is the frequency distribution $Q_{\mathbf{z}}$ of the k-tuple $\mathbf{z} = ((x_1, y_1), \ldots, (x_k, y_k))$. Similarly, $\mathcal{R} := \mathcal{R}(\mathbf{x}, \mathbf{y})$ is the set of density operators on $\mathcal{H}_A \otimes \mathcal{H}_B$ such that the outcomes of a measurement of any $\hat{\rho} \in \mathcal{R}$ with respect to $\mathcal{F} \otimes \mathcal{F}'$ are distributed according to $\hat{Q}(\mathbf{x}, \mathbf{y})$.

Note that $T \cap T'$ is a p^2-random selection on $\{1, \ldots, n\}$ and that $S \cap \overline{T'}$ is a $p^2(1-p)$-random selection on $\{1, \ldots, n\}$. Corollary 4.4 implies that

$$H_0^\varepsilon(\mathbf{X}'|\mathbf{Y} = \mathbf{y}, C) \leq H_0^\varepsilon(\mathbf{X}_{\overline{S \cap \overline{T'}}}) = nH(X|Y) + o(n) \tag{19}$$

holds for ε exponentially small in n. Similarly, Lemma 3.9 implies

$$H_\infty^\varepsilon(\mathbf{X}'|C) = H_\infty^\varepsilon(\mathbf{X}_{\overline{T \cap T'}}) + o(n) = nH(X) + o(n) \ . \tag{20}$$

4.4.2 Information Reconciliation

Let $n' := n - |S \cup T \cup T'|$ be the length of the tuples $\mathbf{X}'$ and $\mathbf{Y}'$, and let, for some $r' \leq n'$, $\mathcal{H}_{r'} := \mathcal{H}(\mathcal{X}^{n'} \to \{0,1\}^{r'})$ be the set of two-universal hash functions mapping $\mathbf{X}'$ to r' bits.

Alice		**Bob**
$F \in_R \mathcal{H}_{r'}$	$\xrightarrow{F, F(\mathbf{X}')}$	$\bar{\mathbf{X}}'$: guess $\mathbf{X}'$ from $\mathbf{Y}'$

It follows from Lemma 4.6 and (19) that for some $r' = r + o(n) \leq nH(X|Y) + o(n)$, $\mathbf{X}' = \bar{\mathbf{X}}'$ holds except with probability exponentially small

in n. Moreover, since F is independent of $\mathbf{X}'$ and since $H_0(F(X')) = r'$, Lemma 3.8 together with (20) implies

$$H_\infty^{\varepsilon'}(\mathbf{X}'|C, C') \geq H_\infty^{\varepsilon}(\mathbf{X}'|C) - r' + o(n) = nH(X) - r' + o(n) \tag{21}$$

where $C' := (F, F(\mathbf{X}'))$ are the messages sent by Alice during the information reconciliation protocol.

4.4.3 Privacy Amplification

Let $\mathcal{P}$ be the set of permutations of the n' elements of $\mathbf{X}'$ and for some $s' \leq n'$ let $\mathcal{H}_{s'} := \mathcal{H}(\mathcal{X}^{n'} \to \{0,1\}^{s'})$ be a two-universal hash function mapping $\mathbf{X}'$ to s' bits.

Alice		**Bob**
$P \in \mathcal{P}, G \in_R \mathcal{H}_{s'}$	$\xrightarrow{P, G}$	
$S := G(\mathbf{X}')$		$S' := G(P(\bar{\mathbf{X}}'))$

Since $\mathbf{X}' = \mathbf{X}''$ holds except with probability exponentially small in n, we have $S = S'$.

It follows from Corollary 4.3 and (21) that for some $s' = s + o(n)$,

$$H_\infty^{\varepsilon'}(\mathbf{X}'|C, C') - H_0^{\varepsilon}(\rho) - s' \geq nH(X) - r' - n \max_{\hat{\rho} \in \mathcal{R}} S(\hat{\rho}) - s' + o(n)$$

is smaller than zero. Theorem 4.7 thus implies that the knowledge of Eve about the key S is negligible.

Note that the length s' of the final key is $t - r - u + o(n)$. The rate R of this generic protocol is thus given by

$$R = I(X;Y) - \max_{\hat{\rho} \in \mathcal{R}} S(\hat{\rho}) \ . \tag{22}$$

Note further that we can carry out the same analysis for the difference $H_\infty^{\varepsilon'}(\mathbf{X}'|C, W) - H_0^{\varepsilon}(\rho|W)$ where we condition on additional information W. This might improve the rate R with a clever choice of the information W as we will see in the next section.

5 Examples

In the following we will illustrate our result by calculating the secret key rate and the tolerable error rates for common quantum key distribution protocols.

5.1 BB84 (The Four–State Protocol)

The BB84 quantum key distribution protocol [3] belongs to the class of so-called *prepare and measure* protocols. In this protocol, Alice chooses randomly, with probability $(1-p)$, the first out of a set of two conjugate bases of a qubit, the second basis is chosen with probability p. She then prepares one of the orthogonal basis states, each chosen with equal probability, and sends the quantum state to Bob. [4]

The BB84 protocol can be regarded as an entanglement based protocol and is in this version known as BBM92 [36]. The preparation stage on Alice's side is then given by a measurement on one half of an entangled quantum state whose second part is sent off to Bob. The relevant quantum state ρ is a two qubit state, $\rho \in S(\mathcal{C}^2 \otimes \mathcal{C}^2)$ and we denote measurement basis one by $\{|0\rangle, |1\rangle\}$ and basis two by $\{|+\rangle, |-\rangle\}$, where $|\pm\rangle := \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$. It is understood that e.g. $\langle 01|\rho|01\rangle$ corresponds to the probability that Alice obtains outcome 0 and Bob outcome 1 when both choose to measure in the first basis. We further identify $|+\rangle$ with outcome 0 and $|-\rangle$ with outcome 1.

After the phase, where the transmission of the quantum states and the measurements have been finished, both parties publicly announce the bases in which they conducted their measurements. They discard the cases in which they did not measure in the same basis. On a small subset of the remaining data, they compare a small part of the string to obtain an estimate of the error rate. Let us assume that the error rate for measurements in both basis are the same and equal to ϵ. If this is not the case, Alice and Bob can always randomly flip some of the bits of the set with the lower error rate in order to make the error probabilities of both sets equal.

The entropy of Alice's string X equals $H(X) = 1$ and the conditional entropy of X given Y is given by $H(X|Y) = h(\epsilon)$. The von Neumann entropy of ρ can be estimated as follows. Note that for all projective measurements on ρ with outcome given by a random variable Z, $H(Z) \geq S(\rho)$. Using Alice's and Bob's data, we want to construct the data that a Bell measurement, saved in the random variable Z had resulted in. Let us define the Bell states

$$|\psi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle) \quad \text{and} \quad |\phi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle)$$

[4] The original proposal by Bennett and Brassard fixes $p = \frac{1}{2}$. A more efficient protocol, which achieves twice the key generation rate of the original proposal, can be obtained by choosing the two bases with different probabilities [35]. We choose $p = \Theta(n^{-\alpha})$, for $\alpha \in (0, 1)$ and n the number of transmitted qubits (see also section 4.4) .

and express the probabilities of Z, denoted by λ_i, in terms of the probabilities of measurements in basis one and basis two:

$$\begin{aligned}
\lambda_1 &:= \langle\psi^+|\rho|\psi^+\rangle = \langle++|\rho|++\rangle + \langle--|\rho|--\rangle - \langle\phi^+|\rho|\phi^+\rangle && (23)\\
\lambda_2 &:= \langle\psi^-|\rho|\psi^-\rangle = \langle+-|\rho|+-\rangle + \langle-+|\rho|-+\rangle - \langle\phi^-|\rho|\phi^-\rangle && (24)\\
\lambda_3 &:= \langle\phi^+|\rho|\phi^+\rangle = \langle01|\rho|01\rangle + \langle10|\rho|10\rangle - \langle\phi^-|\rho|\phi^-\rangle && (25)\\
\lambda_4 &:= \langle\phi^-|\rho|\phi^-\rangle\ . && (26)
\end{aligned}$$

The symmetric error probability ϵ yields

$$\begin{aligned}
\langle00|\rho|00\rangle + \langle11|\rho|11\rangle &= 1-\epsilon\\
\langle++|\rho|++\rangle + \langle--|\rho|--\rangle &= 1-\epsilon\\
\langle+-|\rho|+-\rangle + \langle-+|\rho|-+\rangle &= \epsilon\\
\langle01|\rho|01\rangle + \langle10|\rho|10\rangle &= \epsilon
\end{aligned}$$

and can be inserted into eqs. (23)-(26). We obtain

$$\begin{aligned}
\lambda_3 &= \epsilon - \lambda_4 && (27)\\
\lambda_2 &= \epsilon - \lambda_4 && (28)\\
\lambda_1 &= 1-\epsilon-\lambda_3 = 1-2\epsilon+\lambda_4\ . && (29)
\end{aligned}$$

It remains to find the value of the free parameter $\lambda_4 \in [0,\epsilon]$ such that $H(Z)$ is maximized. It can easily be shown that this is the case for $\lambda_4 = \epsilon^2$ with $H(Z) = 2h(\epsilon)$.

The rate R of the protocol according to eq. (22) is given by

$$R = H(X) - H(X|Y) - H(Z) = 1 - 3h(\epsilon)$$

The security threshold is the highest value of ϵ such that the rate R is positive and is henceforth the solution to the equation $1 - 3h(\epsilon) = 0$. We obtain $\epsilon \approx 0.061$ which corresponds to a 6.1% bit error rate. Conversely, there exists a quantum state ρ for which this rate is achieved and it is given by the mixture

$$\rho = \lambda_1|\psi^+\rangle\langle\psi^+| + \lambda_2|\psi^-\rangle\langle\psi^-| + \lambda_3|\phi^+\rangle\langle\phi^+| + \lambda_4|\phi^-\rangle\langle\phi^-| \qquad (30)$$

Making use of the remark at the end of section 4.4.3, we can improve this security threshold. To do so, we introduce a random variable $W = X \oplus Y$,

which contains the information about the error positions. The min entropy of the string X does not decrease, whereas the size of the quantum data does, thus improving the key rate R. This can be seen as follows: given the fact that Alice and Bob measured in bases number one/two and that an error/no error and has occurred, the quantum system can be divided into 4 subsystems. The subsystems in the case of one error/no error contain a fraction of $\frac{\epsilon}{2}$ and $\frac{1-\epsilon}{2}$ of the total number of qubits, respectively. For each of the systems the entropy can be estimated separately. If no error occurred we obtain $h(\frac{1-2\epsilon+\lambda_4}{1-\epsilon})$ and if an error occurred we get $h(\frac{\epsilon-\lambda_4}{\epsilon})$. Averaging over the four systems gives

$$(1-\epsilon)h\left(\frac{1-2\epsilon+\lambda_4}{1-\epsilon}\right)+\epsilon h\left(\frac{\epsilon-\lambda_4}{\epsilon}\right)=H(Z)-h(\epsilon)$$

The key rate for BB84 is thus given by $R=1-2h(\epsilon)$ and the security threshold $\epsilon\approx 0.1100$ is the solution to the equation $1-2h(\epsilon)=0$. The same rate has previously been obtained by Shor and Preskill [19].

5.2 The Six–State Protocol

The six–state protocol [37, 38] is similar to the BB84 protocol, but makes use of a third basis on either side. This additional basis is defined as $\{\frac{1}{\sqrt{2}}(|0\rangle+i|1\rangle),\frac{1}{\sqrt{2}}(|0\rangle-i|1\rangle)\}$ and conjugate to the other bases [5]. This protocol admits higher symmetry, since the six states that are sent are symmetrically distributed on the Bloch sphere. Similarly to the derivation of eqs. (27)-(29), we easily derive the following additional constraint on the eigenvalues

$$\lambda_3=\epsilon-\lambda_2$$

which results in $\lambda_1=1-3/2\epsilon$ and $\lambda_i=\epsilon/2$ for $i\in\{2,3,4\}$ corresponding to a security threshold of 6.8% with corresponding state

$$\begin{aligned}\rho &= \lambda_1|\psi^+\rangle\langle\psi^+|+\lambda_2|\psi^-\rangle\langle\psi^-|+\lambda_3|\phi^+\rangle\langle\phi^+|+\lambda_4|\phi^-\rangle\langle\phi^-| &(31)\\ &= (1-2\epsilon)|\psi^+\rangle\langle\psi^+|+2\epsilon\frac{\mathbb{1}}{4} &(32)\end{aligned}$$

Another way to derive this result uses the average fidelity $\bar{F}$ of a qubit quantum channel. It can be shown [39] to be equal to the average fidelity of the six states. Here, the fidelity of each state equals $1-\epsilon$ and therefore

[5] Note that we can choose bases two and three e.g. with probability $\frac{p}{2}$ each

$\bar{F} = 1 - \epsilon$. $\bar{F}$ and the entanglement fidelity F_e are related by the formula $F_e = \frac{3\bar{F}-1}{2}$[40] which also leads to the 6.8% by use of the quantum Fano inequality.

The given bounds for BB84 and the six–state protocol on the maximal entropy of ρ are optimal, since Eve can simply prepare the state eq.(30). Even if we consider Alice preparing the particles and sending them off to Bob, we cannot achieve a better bound. This is because the state eq. (30) can be effected by Eve with the following strategy: apply the Pauli matrix σ_z with probability λ_2 (σ_y with λ_3 and σ_x with λ_4) on the sent quantum state and with probability λ_1 take no action.

By conditioning on the random variable $W = X \oplus Y$, however, we can improve the security threshold in a similar manner as we did in the BB84 analysis. This leads to a value of $\epsilon \approx 0.1262$ for the six–state protocol, which coincides with the result of an earlier calculation by Lo [41] based on a result by Bennett *et al.* [42].

5.3 B92

In 1992, Bennett [5] suggested a protocol for quantum key distribution that belongs to the class of prepare and measure protocols differs, however, significantly from BB84 and the six–state protocol. In the specification of the protocol, known as B92, Alice sends one of two non–orthogonal quantum states, which we will denote by $|u_\pm\rangle$, to Bob. He chooses randomly to measure in one of two von Neumann measurements. The first measurement consists of the vectors $\{|u_-\rangle, |\tilde{u}_-\rangle\}$, where $|\tilde{u}_-\rangle$ is orthogonal to $|u_-\rangle$. Similarly, the second measurement is given by $\{|u_+\rangle, |\tilde{u}_+\rangle\}$ with $|\tilde{u}_+\rangle$ orthogonal to $|u_+\rangle$. Bob announces acceptance if he obtains outcomes corresponding to $|\tilde{u}_\pm\rangle$, otherwise both parties discard the values that they recorded.

Alice records the bit value 0/1 if she sends $|u_+\rangle/|u_-\rangle$ and Bob jots down the value 0/1 if he obtains $|\tilde{u}_-\rangle/|\tilde{u}_+\rangle$. We will assume throughout the analysis that Alice sends each quantum state with equal probability and Bob chooses randomly and with equal probability between his two measurements.

Note that in the case of perfect transmission, the strings, conditioned upon acceptance are identical and randomly distributed. We will now proceed to show how one can apply our generic security proof to this specific protocol in the presence of noise. To do so, we need to estimate the expressions in (22) where $\mathcal{R}$ is the sets of possible quantum states conditioned on the event that Alice and Bob accept. As in the analysis of BB84 and the six–state protocol, we will condition on an additional random variable, which equals the XOR of Alice's and Bob's bits after acceptance.

For the following analysis, let p_{xy} for $x, y \in \{0,1\}$, be the probability that Alice and Bob accept a particle and that they have the bit values x and y, respectively. We can without loss of generality assume that $p_{00} = p_{11}$ and $p_{01} = p_{10}$ (Alice and Bob can simply abort the protocol if this is not the case).

Let $\{|0\rangle, |1\rangle\}$ be an orthonormal basis and write

$$|u_\pm\rangle = \beta|0\rangle \pm \alpha|1\rangle$$
$$|\tilde{u}_\pm\rangle = \alpha|0\rangle \mp \beta|1\rangle$$

with $\alpha \in (0, \frac{1}{\sqrt{2}})$ and $\beta = \sqrt{1-\alpha^2}$. The interaction of the transmitted quantum states with the environment or a possible eavesdropper, Eve, is given by

$$|u_\pm\rangle|e\rangle \mapsto |\Psi_\pm\rangle := \sqrt{1-\delta}|u_\pm\rangle|e_\pm\rangle + \sqrt{\delta}|\tilde{u}_\pm\rangle|\tilde{e}_\pm\rangle \ , \tag{33}$$

where $\delta = 4p_{01} = 4p_{10}$. (Note that the factor 4 results from the random choices of Alice and Bob.)

The evolution in equation eq. (33) is unitary which implies the important constraint

$$\langle\, u_+ \,|\, u_- \,\rangle = \langle\, \Psi_+ \,|\, \Psi_- \,\rangle \ .$$

This constraint reads in its expanded form

$$\begin{aligned} \beta^2 - \alpha^2 = (1-\delta)(\beta^2 - \alpha^2) \langle\, e_+ \,|\, e_- \,\rangle \\ + \sqrt{(1-\delta)\delta}\ 2\alpha\beta \left(\langle\, e_+ \,|\, \tilde{e}_- \,\rangle + \langle\, \tilde{e}_+ \,|\, e_- \,\rangle\right) \\ + \delta(\alpha^2 - \beta^2) \langle\, \tilde{e}_+ \,|\, \tilde{e}_- \,\rangle \ . \end{aligned} \tag{34}$$

Without loss of generality we can take $\langle\, e_+ \,|\, e_- \,\rangle$ to be real. Eve's quantum states, given the outcome was accepted by Bob and that Alice and Bob have the same bit value, are denoted by $|f_\pm\rangle$. In the case of an error and acceptance, we write $|\tilde{f}_\pm\rangle$, where $\pm$ denotes Alice's bit value 0/1. One easily obtains

$$|f_\pm\rangle := \frac{\langle\, \tilde{u}_\mp \,|\, \Psi_\pm \,\rangle}{\sqrt{\gamma}} = \frac{\sqrt{1-\delta}\ 2\alpha\beta|e_\pm\rangle + \sqrt{\delta}(\alpha^2 - \beta^2)|\tilde{e}_\pm\rangle}{\sqrt{\gamma}} \tag{35}$$

$$|\tilde{f}_\pm\rangle := \frac{\langle\, \tilde{u}_\pm \,|\, \Psi_\pm \,\rangle}{\sqrt{\delta}} = |\tilde{e}_\pm\rangle \tag{36}$$

where $\gamma = 4p_{00} = 4p_{11}$ is given by the probability that Alice and Bob have a correct value,

$$\gamma = (1-2\delta)(2\alpha\beta)^2 + 2\sqrt{(1-\delta)\delta}2\alpha\beta(\alpha^2 - \beta^2)\mathrm{Re}\,\langle\, e_\pm \,|\, \tilde{e}_\pm \,\rangle + \delta \ . \tag{37}$$

Eve's density matrices, conditioned on the correctness/ falseness of the accepted bit, are given by

$$\begin{aligned}\sigma &:= \tfrac{1}{2}\left(|f_+\rangle\langle f_+| + |f_-\rangle\langle f_-|\right)\\ \tilde{\sigma} &:= \tfrac{1}{2}\big(|\tilde{f}_+\rangle\langle\tilde{f}_+| + |\tilde{f}_-\rangle\langle\tilde{f}_-|\big) = \tfrac{1}{2}\big(|\tilde{e}_+\rangle\langle\tilde{e}_+| + |\tilde{e}_-\rangle\langle\tilde{e}_-|\big)\end{aligned}$$

and have eigenvalues $\frac{1\pm|\langle f_+\,|\,f_-\rangle|}{2}$ and $\frac{1\pm|\langle \tilde{e}_+\,|\,\tilde{e}_-\rangle|}{2}$, respectively. Every estimate for the scalar products $\langle f_+\,|\,f_-\rangle$ and $\langle \tilde{e}_+\,|\,\tilde{e}_-\rangle$ thus leads to an estimate of the entropy of σ and $\tilde{\sigma}$. $\langle f_+\,|\,f_-\rangle$ takes the form

$$\langle f_+\,|\,f_-\rangle \;=\; \frac{(1-\delta)\,\langle e_+\,|\,e_-\rangle - (\beta^2-\alpha^2)^2}{\gamma}$$

where we made use of eq. (34) to simplify the expression in the nominator.

It thus remains to find an estimate for $\langle e_+\,|\,e_-\rangle$. This will be done by use of the unitarity constraint, eq. (34). In particular, we can choose δ such that $\langle e_+\,|\,e_-\rangle$ is sufficiently close to one. For small $\mathrm{Re}\,\langle e_\pm\,|\,\tilde{e}_\pm\rangle$, we thus derived lower bound on the scalar product $\langle f_+\,|\,f_-\rangle$. Note that δ and γ can be derived from the probabilities p_{xy} which are determined by Alice and Bob. Together with eq. (37), this gives an estimate for $\mathrm{Re}\,\langle e_\pm\,|\,\tilde{e}_\pm\rangle$. Using the trivial bound $S(\tilde{\sigma}) \le 1$, this suffices to find a bound for the rate of B92 according to eq. (22).

As a specific example let us consider the depolarizing channel

$$\rho \to (1-p)\rho + \frac{p}{3}\sum_i \sigma_i\rho\sigma_i\ .$$

It is easy to compute the quantities p_{xy} for this channel. In particular, we obtain

$$p_{01} = p_{10} = p/6$$

and

$$p_{00} = p_{11} = \frac{1}{4}(1-\frac{4}{3}p)(2\alpha\beta)^2 + \frac{2}{3}p$$

i.e., $\delta = \frac{2}{3}p$. Using eq. (37) and $\gamma = 4p_{00} = 4p_{11}$, we have $\mathrm{Re}\,\langle e_\pm\,|\,\tilde{e}_\pm\rangle = 0$. The error rate conditioned on acceptance, is thus given by

$$\epsilon = \frac{\delta}{(1-2\delta)\eta + 2\delta} \qquad \text{with } \eta := (2\alpha\beta)^2\ .$$

From $\mathrm{Re}\,\langle e_+\,|\,\tilde{e}_+\rangle = 0$ follows $\mathrm{Re}\,\langle e_+\,|\,\tilde{e}_-\rangle \le \sqrt{1-|\langle e_+\,|\,e_-\rangle|^2}$ which we insert into eq. (34). We therefore have an estimate of the terms proportional

to $\sqrt{\delta}$. For the third term of the right hand side of eq. (34), we use take the trivial estimate $\mathrm{Re}\,\langle\,\tilde{e}_+\,|\,\tilde{e}_-\,\rangle \geq -1$. Altogether we have

$$\nu\,(1-\langle\,e_+\,|\,e_-\,\rangle)^2 \leq 1-\langle\,e_+\,|\,e_-\,\rangle^2 \quad \text{with } \nu := \frac{(1-\delta)(1-\eta)}{4\delta\eta}\;.$$

The valid solutions of this quadratic expression are given by

$$\langle\,e_+\,|\,e_-\,\rangle \geq \frac{\nu-1}{\nu+1}$$

and directly lead to an estimate for $S(\sigma)$. Using $S(\tilde{\sigma}) \leq 1$ we obtain an estimate for the entropy of the quantum state of Alice and Bob conditioned on the random variable W. The total rate is given by

$$R = \frac{(1-2\delta)\eta+2\delta}{2}\big(1-h(\epsilon)-\epsilon-(1-\epsilon)h(x)\big)$$

where

$$x := \frac{(1-5\delta)(1-\delta)\eta(1-\eta)}{(\delta+(1-2\delta)\eta)((1-\delta)-(1-5\delta)\eta)}$$

The highest security threshold p is obtained for $\alpha \approx 0.38$ and equals $p \approx 0.036$. This is a slight improvement of the previously obtained security threshold $p \approx 0.034$ by Tamaki, Koashi and Imoto [21].

6 Conclusion

In this paper we have presented a security proof for a generic quantum key distribution protocol. The protocol requires only single particle measurements on Alice's and Bob's sides and uses one-way information reconciliation and privacy amplification to extract a secret key from the raw data. In our proof we estimate the amount of classical correlation contained in Alice's and Bob's data and derive a bound on the quantum information, which a possible adversary might have about this data. Subsequently, we apply a recent result by König, Maurer and Renner [1] to ensure the security of the privacy amplification stage.

Special cases of our protocol include entanglement based quantum key distribution, such as E91, and prepare and measure schemes, such as BB84 or the six state protocol. We were able to derive security thresholds of 11.0% bit error rate for BB84 (four–state protocol) and 12.6% for the six–state protocol, previously obtained by Shor and Preskill, and Lo, respectively.

Furthermore we have shown how our technique can be applied to prove the security of B92. In the case of the depolarizing channel this leads to a slight improvement of the security threshold that has been recently obtained by Tamaki, Koashi and Imoto.

7 Acknowledgements

This work was supported in part by a grant from the Cambridge-MIT Institute, A*Star Grant No. 012-104-0040 and the EU under project RESQ (IST-2001-37559). MC was supported by a DAAD Doktorandenstipendium. RR was partially supported by the Swiss National Science Foundation, project No. 20-66716.01.

References

[1] R. König, U. Maurer, and R. Renner. On the power of quantum memory. quant-ph/0305154.

[2] S. Wiesner. Conjugate coding. *Sigact News*, 15(1):78–88, 1983. Originally written c. 1970 but unpublished.

[3] C. H. Bennett and G. Brassard. Quantum cryptography, public key distribution and coin tossing. In *Proceedings of International Conference on Computer Systems and Signal Processing*, page 175, 1984.

[4] A. K. Ekert. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.*, 67(6):661, 1991.

[5] C.H. Bennett. Quantum cryptography using any two nonorthogonal states. *Phys. Rev. Lett.*, 68(21):3121–3124, 1992.

[6] A.K. Ekert and B. Huttner. Eavesdropping techniques in quantum cryptosystems. *Journal of Modern Optics*, 41:2455–2466, 1994. Special Issue on Quantum Communication.

[7] A.C.-C. Yao. Security of quantum protocols against coherent measurements. In *Proceedings of the 27th ACM Symposium on the Theory of Computing*, pages 67–75. ACM Press, 1995.

[8] D. Deutsch, A. Ekert, R. Jozsa, C. Macchiavello, S. Popescu, and A. Sanpera. Quantum privacy amplification and the security of quan-

tum cryptography over noisy channels. *Phys. Rev. Lett.*, 77:2818–2821, 1996. Erratum-ibid. 80 (1998) 2022–2022, quant-ph/9604039.

[9] M. Zukowski, A. Zeilinger, M. Horne, and A.K. Ekert. "Event-ready detectors"; Bell experiment via entanglement swapping. *Physical Review Letters*, 71:4287–4290, 1993.

[10] W. Dür, H.-J. Briegel, J. I. Cirac, and P. Zoller. Quantum repeaters based on entanglement purification. *Phys. Rev. A*, 59:169–181, 1999.

[11] H. Aschauer and H.-J. Briegel. A security proof for quantum cryptography based entirely on entanglement purification. *Physical Review*, A 66:032302, 2002.

[12] H. Inamori. Security of EPR-based quantum key distribution. quant-ph/0008064.

[13] M. Ben-Or. Simple security proof for quantum key distribution. Online presentation available at `http://www.msri.org/publications/ln/msri/2002/qip/ben-or/1/index.html`.

[14] C.H. Bennett, G. Brassard, and J.-M. Robert. Privacy amplification by public discusion. *SIAM Journal on Computing*, 17(2):210–229, 1988.

[15] C.H. Bennett, G. Brassard, C. Crépeau, and U. Maurer. Generalized privacy amplification. *IEEE Transactions on Information Theory*, 41(6):1915–1923, 1995.

[16] H.-K. Lo and H. F. Chau. Unconditional Security of Quantum Key Distribution over Arbitrarily Long Distances. *Science*, 283(5410):2050–2056, 1999.

[17] D. Mayers. Unconditional security in quantum cryptography. quant-ph/9802025, 1998.

[18] E. Biham, M. Boyer, P. O. Boykin, T. Mor, and V. Roychowdhury. A proof of the security of quantum key distribution. In *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing (STOC)*, page 715, New York, 2000. ACM Press. quant-ph/9912053.

[19] P. W. Shor and J. Preskill. Simple proof of security of the BB84 quantum key distribution protocol. *Phys. Rev. Lett.*, 85(2):441–444, 2000. quant-ph/0003004.

[20] D. Gottesman and H.-K. Lo. Proof of security of quantum key distribution with two-way classical communication. *IEEE Trans. Inf. Th.*, 49(2):457–475, 2003. quant-ph/0105121.

[21] K. Tamaki, M. Koashi, and N. Imoto. Unconditionally secure key distribution based on two nonorthogonal states. *Phys. Rev. Lett.*, 90:167904, 2003.

[22] B. Schumacher. Quantum coding. *Physical Review A*, 51:2738–2747, 1995.

[23] M. Horodecki. Limits for compression of quantum information carried by ensembles of mixed states. *Physical Review A*, 57:3364, 1998.

[24] A. Ambainis, A. Nayak, A. Ta-Shma, and U. Vazirani. Dense quantum coding and a lower bound for 1-way quantum automata. In *Proceedings of the 31st ACM Symposium on Theory of Computing*, 1999. quant-ph/9804043.

[25] A. Nayak. Optimal lower bounds for quantum automata and random access codes. In *Proceedings of the 40th Annual Symposium on Foundations of Computer Science*, pages 369–377. quant-ph/9904093.

[26] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley Series in Telecommunications. Wiley, New York, 1991.

[27] T. Holenstein and R. Renner. On the frequency distribution of non-independent random values. Available at `http://www.crypto.ethz.ch/~renner/publications.html`, November 2003.

[28] R. Renner and S. Wolf. Smooth Rényi entropy and applications. Available at `http://www.crypto.ethz.ch/~renner/publications.html`, October 2003.

[29] J. L. Carter and M. N. Wegman. Universal classes of hash functions. *Journal of Computer and System Sciences*, 18:143–154, 1979.

[30] M. N. Wegman and J. L. Carter. New hash functions and their use in authentication and set equality. *Journal of Computer and System Sciences*, 22:265, 1981.

[31] M. A. Nielsen and I. L. Chuang. *Quantum computation and quantum information*. Cambridge University Press, 2000.

[32] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press, 1985.

[33] R. Rajendra. *Matrix analysis*. Graduate Texts in Mathematics. Springer, 1996.

[34] R. Renner and S. Wolf. The exact price for unconditionally secure asymmetric cryptography. In *Advances in Cryptology - EUROCRYPT '04*, Lecture Notes in Computer Science. Springer-Verlag, 2004.

[35] H.-K. Lo, H. F. Chau, and M. Ardehali. Efficient quantum key distribution scheme and proof of its unconditional security. quant-ph/0011056, 2000.

[36] C. H. Bennett, G. Brassard, and N. D. Mermin. Quantum cryptography without Bell's theorem. *Phys. Rev. Lett,*, 68(5):557, 1992.

[37] D. Bruß. Optimal eavesdropping in quantum cryptography with six states. *Phys. Rev. Lett.*, 81:3018–3021, 1998.

[38] H. Bechmann-Pasquinucci and N. Gisin. Incoherent and coherent eavesdropping in the six-state protocol of quantum cryptography. *Phys. Rev. A*, 59(6):4238–4248, 1999. quant-ph/9807041.

[39] A. J. Short K. Banaszek M. D. Bowdrey, D. K. L. Oi and J. A. Jones. Fidelity of single qubit maps. *Physics Letters A*, 294(5-6):258–260, 2002. quant-ph/0201106.

[40] M. Horodecki, P. Horodecki, and R. Horodecki. General teleportation channel, singlet fraction, and quasidistillation. *Phys. Rev. A*, 60(3):1888–1898, 1999.

[41] H.-K. Lo. Proof of unconditional security of six-state quantum key distribution scheme. *Quant. Inf. Comp.*, 2(2):81–92, 2001. quant-ph/0102138.

[42] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters. Mixed state entanglement and quantum error correction. *Phys. Rev. A*, 54(5):3824–3851, 1996. quant-ph/9604024.