Universally Composable Privacy Amplification Against Quantum Adversaries

Renato Renner and Robert König

Swiss Federal Institute of Technology (ETH), Zürich, Switzerland
renner@inf.ethz.ch rkoenig@inf.ethz.ch

Abstract. Privacy amplification is the art of shrinking a partially secret string Z to a highly secret key S. We show that, even if an adversary holds quantum information about the initial string Z, the key S obtained by two-universal hashing is secure, according to a universally composable security definition. Additionally, we give an asymptotically optimal lower bound on the length of the extractable key S in terms of the adversary's (quantum) knowledge about Z. Our result has applications in quantum cryptography. In particular, it implies that many of the known quantum key distribution protocols are universally composable.

1 Introduction

1.1 Privacy Amplification

Consider two parties having access to a common string Z about which an adversary might have some partial information. *Privacy amplification*, introduced by Bennett, Brassard, and Robert [10], is the art of transforming this partially secure string Z into a highly secret key S by public discussion. A good technique is to compute S as the output of a publicly chosen two-universal hash function[1] F applied to Z. Indeed, it has been shown [10, 21, 9] that, if the adversary holds purely classical information W about Z, this method yields a secure key S and, additionally, is asymptotically optimal with respect to the length of S. For instance, if both the initial string Z and the adversary's knowledge W consist of many independent and identically distributed parts, the number of extractable key bits roughly equals the conditional Shannon entropy $H(Z|W)$.

The analysis of privacy amplification can be extended to a situation where the adversary holds quantum instead of only classical information about Z. This generalizes the classical setting in a non-trivial way. In particular, the adversary might store her quantum information until she learns the hash function F (which is publicly chosen) and then perform a measurement depending on F. This might allow her to obtain more information about the function output (i.e., the resulting key S) than if she had measured her state at the beginning (independently of F).

[1] See Section 2.1 for a definition of two-universal functions.

1.2 Universal Composability

Cryptographic primitives (such as a secret key or an authentic communication channel) are often used as components within a larger system (e.g., a system for secure message transmission usually makes use of a secret key for encryption). It is thus natural to require that the security of these components is not compromised when they are used in any (arbitrarily complex) scheme. This requirement is captured by the notion of universal composability. Roughly speaking, a cryptographic primitive is said to provide *universally composable security* if it is secure in *any* arbitrary context. For instance, the universally composable security of a secret key S guarantees that any bit of S remains secret even if some other part of S is given to an adversary.

In the past few years, composable security has attracted a lot of interest and led to important new definitions and proofs (see, e.g., the framework of Canetti [11] or Pfitzmann and Waidner [27]). Recently, Ben-Or and Mayers [5, 6] and Unruh [30] have generalized the notion of universal composability to the quantum world. Universally composable security definitions are usually based on the idea of characterizing the security of a cryptographic scheme by its distance to an ideal system which (by definition) is perfectly secure. For instance, a secret key S is said to be secure if it is close to an independent and almost uniformly distributed string U. This implies that any cryptosystem which is proven secure when using a perfect key U remains secure when U is replaced by the (real) key S.

Unfortunately, most of the existing security definitions in quantum cryptography do not provide universal composability. For instance, the security of the key S generated by a quantum key distribution (QKD) scheme is usually defined by the requirement that the mutual information between S and the classical outcome W obtained from an arbitrary measurement of the adversary's quantum system be small (for a formal definition, see, e.g., [26] or [18]). This, however, does not necessarily imply composability. Indeed, an adversary might wait with the measurement of her quantum state until she learns some of the bits of S, which possibly allows her to obtain information about the remaining bits (cf. Section 3).

1.3 Contributions

We address the problem of privacy amplification in a setting where an adversary holds quantum information. We show that, by two-universal hashing, one can obtain a key S which is secure according to a universally composable security definition. This means that, in any context, S is virtually as secure as a perfect key, i.e., a uniformly distributed string U which is completely independent of the adversary's knowledge. This has implications in quantum cryptography. In particular, since the security of many of the known QKD protocols such as BB84 [8] or B92 [7] can be proven based on the security of privacy amplification (cf. [13] and [22], or [23]), it follows immediately from our results that these protocols provide universal composability (cf. Section 4.5).

Our main technical result (Section 4) is an easily computable lower bound on the length of the extractable key S in terms of (smooth) Rényi entropy (see Section 2.4 for a definition of smooth Rényi entropy). The bound is asymptotically tight if the initial information Z as well as the adversary's (quantum) knowledge consist of n independent pieces, for n approaching infinity (Section 4.4).

1.4 Related Work

The problem of privacy amplification against quantum adversaries has first been studied for the case where the adversary can only store a certain limited number of qubits. Based on a result on communication complexity [1], Ben-Or [2] argued that it is possible to extract at least one secret bit from a uniformly distributed string Z, if Z is sufficiently longer than the size of the adversary's storage device. In [22], it is shown that two-universal hashing allows for the extraction of a secure key S whose length roughly equals the difference between the entropy of the original string Z and the number of qubits stored by the adversary. The security definition used in [22] does, however, not provide universal composability. Simultaneously, Devetak and Winter [15] gave a full analysis of privacy amplification for the special case where the initial string Z as well as the adversary's information consist of many independent pieces. Interestingly, their result can be reproduced from our general bound (Section 4.4).

Ben-Or, Horodecki, Leung, Mayers, and Oppenheim [3, 4] were the first to address the problem of universal composability in the context of QKD. Our security definition (cf. Definition 3 in Section 3) is essentially equivalent to the definitions proposed in [3, 4], which are based on the framework developed in [6]. More precisely, if S is ε-secure according to our definition, it satisfies the security definition of [3] for some parameter ε' depending on ε. It is thus an immediate consequence of the results in [3] that our security definition provides universal composability with respect to the framework of [6].

2 Preliminaries

2.1 Random Functions and Two-Universal Functions

A *random function* F from $\mathcal{X}$ to $\mathcal{Y}$ is a random variable taking values from the set of functions with domain $\mathcal{X}$ and range $\mathcal{Y}$. F is called a *two-universal* (random) function if $\Pr_{f \leftarrow P_F}[f(x) = f(x')] \leq \frac{1}{|\mathcal{Y}|}$, for any distinct $x, x' \in \mathcal{X}$.[2] In particular, F is two-universal if, for any distinct $x, x' \in \mathcal{X}$, the random variables $F(x)$ and $F(x')$ are independent and uniformly distributed. For instance, the random function distributed uniformly over the set of all functions from $\mathcal{X}$ to $\mathcal{Y}$ is two-universal. Examples of two-universal functions requiring less randomness can, e.g., be found in [12] and [31].

[2] In the literature, two-universality is usually defined for families $\mathcal{F}$ of functions: A family $\mathcal{F}$ is called *two-universal* if the random function F with uniform distribution over $\mathcal{F}$ is two-universal.

2.2 Density Operators and Random States

Let $\mathcal{H}$ be a Hilbert space. We denote by $\mathcal{P}(\mathcal{H})$ the set of non-negative (hermitian) operators ρ on $\mathcal{H}$ with $\mathrm{tr}(\rho) \leq 1$, and call its elements *density operators*. We say that $\rho \in \mathcal{P}(\mathcal{H})$ is *normalized* if $\mathrm{tr}(\rho) = 1$. A normalized density operator $\rho \in \mathcal{P}(\mathcal{H})$ is called *pure* if it has rank 1, i.e., $\rho = P_{|\phi\rangle}$ for some vector $|\phi\rangle \in \mathcal{H}$ (where $P_{|\phi\rangle}$ denotes the projector along $|\phi\rangle$).

We will be concerned with settings involving both classical and quantum information. More precisely, we will consider a situation where the state $\rho_x \in \mathcal{P}(\mathcal{H})$ of a quantum system depends on the value x of a classical random variable X with range $\mathcal{X}$. Note that ρ_X is then itself a random variable with range $\mathcal{P}(\mathcal{H})$. In the following, we call such a random variable with range $\mathcal{P}(\mathcal{H})$ a *random state* on $\mathcal{H}$, and denote it by a bold symbol $\boldsymbol{\rho}$. We say that the random state $\boldsymbol{\rho}$ is *normalized* if $\mathrm{tr}(\boldsymbol{\rho}) \equiv 1$.

It is often convenient to represent classical information as a state of a quantum system. Let $\mathcal{X}$ be a set and let $\mathcal{H}$ be a Hilbert space with orthonormal basis $\{|x\rangle\}_{x\in\mathcal{X}}$. The *state representation* of $x \in \mathcal{X}$, denoted $\{x\}$, is defined as the projector along $|x\rangle$, i.e., $\{x\} := P_{|x\rangle}$. In particular, for a random variable X on $\mathcal{X}$, $\{X\}$ is a random state on $\mathcal{H}$.

Consider a quantum system described by a random state $\boldsymbol{\rho}$ on $\mathcal{H}$, i.e., if the random variable $\boldsymbol{\rho}$ takes the value ρ, then the system is in state ρ. For an observer which is ignorant of the value of the random variable $\boldsymbol{\rho}$, the system is described by the density operator $[\boldsymbol{\rho}]$ defined as the expectation value of $\boldsymbol{\rho}$,

$$[\boldsymbol{\rho}] := \mathop{\mathrm{E}}_{\rho \leftarrow P_{\boldsymbol{\rho}}}[\rho] = \sum_{\rho\in\mathcal{P}(\mathcal{H})} P_{\boldsymbol{\rho}}(\rho)\rho \ ,$$

where $P_{\boldsymbol{\rho}}$ is the probability distribution of $\boldsymbol{\rho}$. More generally, for any event $\mathcal{E}$, we define

$$[\boldsymbol{\rho}|\mathcal{E}] := \mathop{\mathrm{E}}_{\rho \leftarrow P_{\boldsymbol{\rho}|\mathcal{E}}}[\rho] \ .$$

Let X be a random variable and let $\boldsymbol{\rho}$ be a random state. The random state $\{X\} \otimes \boldsymbol{\rho}$ then describes a system consisting of both a state representation of X and a quantum subsystem which is in state $\rho_x := [\boldsymbol{\rho}|X = x]$ whenever X takes the value x. The density operator $[\{X\} \otimes \boldsymbol{\rho}]$ of the overall system is thus given by

$$[\{X\} \otimes \boldsymbol{\rho}] = \mathop{\mathrm{E}}_{x\leftarrow P_X}\big[P_{|x\rangle} \otimes \rho_x\big] = \sum_{x\in\mathcal{X}} P_X(x) P_{|x\rangle} \otimes \rho_x \ . \tag{1}$$

In particular, $[\{X\} \otimes \boldsymbol{\rho}] = [\{X\}] \otimes [\boldsymbol{\rho}]$ if and only if X is independent of $\boldsymbol{\rho}$.

2.3 Distance Measures and Non-Uniformity

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

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

The variational distance between 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 1. *Let P and Q be two probability distributions. Then there exists a joint probability distribution $P_{XX'}$ such that $P_X = P$, $P_{X'} = Q$, and*

$$\Pr_{(x,x')\leftarrow P_{XX'}}[x \neq x'] = \delta(P, Q) .$$

The *trace distance* between two density operators ρ and σ on the same Hilbert space $\mathcal{H}$ is defined as

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

The trace distance is a metric on the set of density operators $\mathcal{P}(\mathcal{H})$. We say that ρ is *ε-close* to σ if $\delta(\rho, \sigma) \leq \varepsilon$, and denote by $\mathcal{B}^{\varepsilon}(\rho)$ the set of density operators which are ε-close to ρ, i.e., $\mathcal{B}^{\varepsilon}(\rho) = \{\sigma \in \mathcal{P}(\mathcal{H}) : \delta(\rho, \sigma) \leq \varepsilon\}$.

The trace distance is subadditive with respect to the tensor product, i.e., for any $\rho, \sigma \in \mathcal{P}(\mathcal{H})$ and $\rho', \sigma' \in \mathcal{P}(\mathcal{H}')$,

$$\delta(\rho \otimes \rho', \sigma \otimes \sigma') \leq \delta(\rho, \sigma) + \delta(\rho', \sigma') , \tag{2}$$

with equality if $\rho' = \sigma'$ is normalized,

$$\delta(\rho \otimes \rho', \sigma \otimes \rho') = \delta(\rho, \sigma) . \tag{3}$$

Moreover, $\delta(\cdot, \cdot)$ cannot increase when the same quantum operation $\mathcal{E}$ is applied to both arguments, i.e.,

$$\delta(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \leq \delta(\rho, \sigma) . \tag{4}$$

The variational distance can be seen as a (classical) special case of the trace distance. Let X and Y be random variables. Then the variational distance between the probability distributions of X and Y equals the trace distance between the state representations $[\{X\}]$ and $[\{Y\}]$, i.e.,

$$\delta(P_X, P_Y) = \delta\big([\{X\}], [\{Y\}]\big) .$$

In particular, it follows directly from (4) that the trace distance between two normalized density operators ρ and σ is an upper bound for the variational distance between the probability distributions P_X and P_Y of the outcomes when applying the same measurement to ρ and σ, respectively, i.e.,

$$\delta(P_X, P_Y) \leq \delta(\rho, \sigma) . \tag{5}$$

The trace distance between two density operators involving a state representation of the same classical random variable X can be written as the expectation of the trace distance between the density operators conditioned on X.

Lemma 2. *Let X be a random variable and let $\boldsymbol{\rho}$ and $\boldsymbol{\sigma}$ be random states. Then*

$$\delta\big([\{X\} \otimes \boldsymbol{\rho}], [\{X\} \otimes \boldsymbol{\sigma}]\big) = \mathop{\mathrm{E}}_{x \leftarrow P_X}\big[\delta(\rho_x, \sigma_x)\big]$$

where $\rho_x := [\boldsymbol{\rho}|X = x]$ and $\sigma_x := [\boldsymbol{\sigma}|X = x]$.

Proof. Using (1) and the orthogonality of the vectors $|x\rangle$, we can write

$$\begin{aligned}
\delta\big([\{X\} \otimes \boldsymbol{\rho}], [\{X\} \otimes \boldsymbol{\sigma}]\big) &= \frac{1}{2}\mathrm{tr}\Big(\Big|\mathop{\mathrm{E}}_{x \leftarrow P_X}\big[P_{|x\rangle} \otimes (\rho_x - \sigma_x)\big]\Big|\Big) \\
&= \frac{1}{2}\mathrm{tr}\Big(\mathop{\mathrm{E}}_{x \leftarrow P_X}\Big[\big|P_{|x\rangle} \otimes (\rho_x - \sigma_x)\big|\Big]\Big) \ .
\end{aligned}$$

The assertion then follows from the linearity of the trace and the fact that $\mathrm{tr}(|P_{|x\rangle} \otimes (\rho_x - \sigma_x)|) = \mathrm{tr}(|\rho_x - \sigma_x|)$. □

In Section 3, we will see that a natural measure for characterizing the secrecy of a key is its trace distance to a uniform distribution. This motivates the following definition.

Definition 1. *Let X be a random variable with range $\mathcal{X}$ and let $\boldsymbol{\rho}$ be a random state. The* non-uniformity *of X* given *$\boldsymbol{\rho}$ is defined by*

$$d(X|\boldsymbol{\rho}) := \delta\big([\{X\} \otimes \boldsymbol{\rho}], [\{U\}] \otimes [\boldsymbol{\rho}]\big)$$

where U is a random variable uniformly distributed on $\mathcal{X}$.

Note that $d(X|\boldsymbol{\rho}) = 0$ if and only if X is uniformly distributed and independent of $\boldsymbol{\rho}$.

2.4 (Smooth) Rényi Entropy

Let $\rho \in \mathcal{P}(\mathcal{H})$ be a density operator and let $\alpha \in [0, \infty]$. The *Rényi entropy of order α of ρ* is defined by[3]

$$S_\alpha(\rho) := \frac{1}{1-\alpha} \log\big(\mathrm{tr}(\rho^\alpha)\big)$$

with the convention $S_\alpha(\rho) := \lim_{\beta \to \alpha} S_\beta(\rho)$ for $\alpha \in \{0, 1, \infty\}$.[4] In particular, for $\alpha = 0$, $S_0(\rho) = \log\big(\mathrm{rank}(\rho)\big)$ and, for $\alpha = \infty$, $S_\infty(\rho) = -\log\big(\lambda_{\max}(\rho)\big)$ where $\lambda_{\max}(\rho)$ denotes the maximum eigenvalue of ρ. Note that, for a classical random variable X, the Rényi entropy $S_\alpha([\{X\}])$ of the state representation of X corresponds to the (classical) Rényi entropy $H_\alpha(X)$ of X [29].

The notion of *ε-smooth* Rényi entropy H_α^ε has been introduced in [28] for the classical case, and can be seen as a generalization of (conventional) Rényi

[3] All logarithms in this paper are binary.

[4] Note that, for this definition, the density operator ρ must not necessarily be normalized.

entropy H_α (see Appendix C for a definition). Smooth Rényi entropy is useful for characterizing basic properties of random variables such as the amount of extractable randomness or the minimum encoding length. Moreover, it has natural properties similar to Shannon entropy.

Definition 2 below generalizes classical smooth Rényi entropy H_α^ε to density operators. This quantum version of smooth Rényi entropy will be useful to state our main results.

Definition 2. *Let $\rho \in \mathcal{P}(\mathcal{H})$ and let $\varepsilon \geq 0$. The* ε-smooth Rényi entropy of order α of ρ *is defined by*[5]

$$S_\alpha^\varepsilon(\rho) := \frac{1}{1-\alpha} \log\Bigl(\inf_{\sigma \in \mathcal{B}^{\varepsilon/2}(\rho)} \bigl(\mathrm{tr}(\sigma^\alpha)\bigr) \Bigr) ,$$

for $\alpha \in (0,1) \cup (1,\infty)$, and $S_\alpha^\varepsilon(\rho) := \lim_{\beta \to \alpha} S_\beta^\varepsilon(\rho)$, for $\alpha \in \{0,\infty\}$.

The classical definition of smooth Rényi entropy can be seen as a special case of Definition 2. In particular, the smooth Rényi entropy $H_\alpha^\varepsilon(X)$ of a classical random variable X is equal to the smooth Rényi entropy $S_\alpha^\varepsilon([\{X\}])$ of the state representation of X. On the other hand, the smooth Rényi entropy of a density operator ρ can be expressed in terms of the classical smooth Rényi entropy of its eigenvalues. Formally,

$$S_\alpha^\varepsilon(\rho) = H_\alpha^\varepsilon(P) , \tag{6}$$

where P is the (not necessarily normalized) probability distribution defined by the eigenvalues $\lambda_1, \ldots, \lambda_d$ of ρ, i.e., $P(i) = \lambda_i$, for $i \in \{1, \ldots, d\}$.

It is important to note that equation (6) provides an efficient method for *computing* the smooth Rényi entropy $S_\alpha^\varepsilon(\rho)$ of a given density operator ρ. In particular, since the smooth Rényi entropy $H_\alpha^\varepsilon(P)$ of a classical probability distribution P can be calculated in a simple way (see Appendix C), it is also easy to compute $S_\alpha^\varepsilon(\rho)$ if the eigenvalues of ρ are known.

The following lemma is a direct generalization of the corresponding statement for classical smooth Rényi entropy (see Lemma 15 in Appendix C) saying that the smooth Rényi entropy $H_\alpha^\varepsilon(Z^n)$ of a random variable Z^n consisting of many independent and identically distributed pieces asymptotically equals its Shannon entropy $H(Z^n)$.

Lemma 3. *Let ρ be a normalized density operator. Then, for any $\alpha \in [0,\infty]$,*

$$\lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{1}{n} S_\alpha^\varepsilon(\rho^{\otimes n}) = S(\rho) ,$$

where $S(\rho)$ denotes the von Neumann entropy of ρ.

---

[5] Recall that $\mathcal{B}^{\varepsilon/2}(\rho)$ denotes the set of non-negative operators $\sigma \in \mathcal{P}(\mathcal{H})$ such that $\delta(\sigma,\rho) \leq \frac{\varepsilon}{2}$, i.e., $\mathrm{tr}(|\sigma - \rho|) \leq \varepsilon$.

3 Secret Keys and Composability

A very intuitive way of defining the security of a *real* cryptographic protocol is to compare it with an *ideal* functionality. The ideal functionality of a secret key S is simply an independent and uniformly distributed random variable U (in particular, U is fully independent of the adversary's information). This motivates the following definition.

Definition 3. *Let S be a random variable, let $\boldsymbol{\rho}$ be a random state, and let $\varepsilon \geq 0$. S is said to be* ε-secure with respect to $\boldsymbol{\rho}$ *if $d(S|\boldsymbol{\rho}) \leq \varepsilon$.*

Consider a situation where S is used as a secret key and where the adversary's information is given by a random state $\boldsymbol{\rho}$. If S is ε-secure with respect to $\boldsymbol{\rho}$ then it is guaranteed that this situation is ε-close—with respect to the trace distance—to an ideal setting where S is replaced by a perfect key U which is uniformly distributed and independent of $\boldsymbol{\rho}$. Since the trace distance does not increase when appending an additional quantum system (cf. (2) or (3)) or when applying any arbitrary quantum operation (cf. (4)), this also holds for any further evolution of the system. In particular, it follows from (5) and Lemma 1 that the real and the ideal setting can be considered to be identical with probability at least $1-\varepsilon$.

Note that our security definition can be seen as a natural generalization of classical security definitions based on the variational distance (which is the classical analogue of the trace distance). Indeed, if the adversary's knowledge is purely classical, Definition 3 is equivalent to a security definition as it is, e.g., used in [17].

The security of a key S according to Definition 3 implies that S is also secure according to many of the widely used security definitions in quantum cryptography. One of the most popular security requirements for a key S with respect to an adversary holding information $\boldsymbol{\rho}$ is that S be almost independent of the classical outcome W resulting from any arbitrary measurement of $\boldsymbol{\rho}$.[6] Obviously, if a key S is ε-secure with respect to $\boldsymbol{\rho}$ (according to our definition), the probability distribution P_{SW} is ε-close (with respect to the variational distance) to a product distribution. Note, however, that the converse is not true: Even if S and W are almost independent for any measurement of $\boldsymbol{\rho}$, the quantum state $\boldsymbol{\rho}$ might still strongly depend on S.

Indeed, security definitions which are formulated in terms of the adversary's measurement results W do not necessarily provide universal composability: If it is only known that a key S is almost independent of the classical outcome W obtained from measuring the quantum state $\boldsymbol{\rho}$—for any measurement strategy chosen independently of S—, one cannot necessarily use S in any arbitrary cryptosystem, e.g., as a one-time pad. Consider for instance a cryptographic application where S consists of two parts S_1 and S_2, and where S is used in such a way that an adversary learns S_1. Hence, the adversary can let the measurement of her quantum system depend on the specific value of S_1. This might provide

[6] See, e.g., [26], and the references therein.

her with more information about S_2 than if she had chosen her measurement independently of S_1.[7]

4 Main Result

4.1 Theorem and Proof

Consider a situation where an adversary holds quantum information $\boldsymbol{\rho}$ about a classical string Z. Additionally, let S be a key of length s computed by applying a (publicly chosen) two-universal function F to Z, that is, $S := F(Z)$. Theorem 1 below states that, if the length s is chosen to be sufficiently smaller than $\bar{s} := S_2([\{Z\} \otimes \boldsymbol{\rho}]) - S_0([\boldsymbol{\rho}])$, then the key S is ε-secure with respect to $\boldsymbol{\rho}$ (for ε decreasing exponentially fast in the difference $\bar{s} - s$). In other words, a two-universal function F can be used to turn a partially secure string Z into a highly secure key S of length roughly $\bar{s}$. In Section 4.3, we will discuss this application in more detail.

Theorem 1. *Let Z be a random variable with range $\mathcal{Z}$, let $\boldsymbol{\rho}$ be a random state, and let F be a two-universal function from $\mathcal{Z}$ to $\mathcal{S} = \{0,1\}^s$ which is independent of Z and $\boldsymbol{\rho}$. Then*

$$d(F(Z)|\{F\} \otimes \boldsymbol{\rho}) \leq \frac{1}{2} 2^{-\frac{1}{2}(S_2([\{Z\}\otimes\boldsymbol{\rho}]) - S_0([\boldsymbol{\rho}]) - s)} .$$

Let us state some technical lemmas to be used for the proof of Theorem 1.

Lemma 4. *Let Z be a random variable with range $\mathcal{Z}$, let $\boldsymbol{\rho}$ be a random state, and let F be a random function on $\mathcal{Z}$ which is independent of Z and $\boldsymbol{\rho}$. Then*

$$d(F(Z)|\{F\} \otimes \boldsymbol{\rho}) = \mathop{\mathrm{E}}_{f \leftarrow P_F}[d(f(Z)|\boldsymbol{\rho})] .$$

Proof. Let U be a random variable uniformly distributed on the range of F and independent of F and $\boldsymbol{\rho}$. Then

$$d(F(Z)|\{F\} \otimes \boldsymbol{\rho}) = \delta\left([\{F(Z)\} \otimes \boldsymbol{\rho} \otimes \{F\}], [\{U\} \otimes \boldsymbol{\rho} \otimes \{F\}]\right) .$$

Now, applying Lemma 2 to the random states $\{F(Z)\} \otimes \boldsymbol{\rho}$ and $\{U\} \otimes \boldsymbol{\rho}$ gives the desired result since

$$\begin{aligned} [\{F(Z)\} \otimes \boldsymbol{\rho}|F = f] &= [\{f(Z)\} \otimes \boldsymbol{\rho}] \\ [\{U\} \otimes \boldsymbol{\rho}|F = f] &= [\{U\}] \otimes [\boldsymbol{\rho}] , \end{aligned}$$

which holds because F is independent of Z, $\boldsymbol{\rho}$, and U. □

[7] The effect of side information on the maximum classical correlation that can be obtained by measurements has been studied in different contexts [16, 19]. A simple example which demonstrates that classical information is indeed helpful for choosing a "good" measurement is as follows: Let S_1 and S_2 be random bits and let ρ be the state of a two-dimensional quantum system obtained by encoding the bit S_2 using either the rectilinear basis (if $S_1 = 0$) or the diagonal basis (if $S_1 = 1$). Clearly, if S_1 is known, S_2 can easily be determined by applying the appropriate measurement to ρ. On the other hand, the probability of correctly guessing S_2 from the outcome of any measurement chosen independently of S_1 is bounded away from 1.

The following lemmas can most easily be formalized in terms of the square of the Hilbert-Schmidt distance. For two density operators ρ and σ, let

$$\Delta(\rho, \sigma) := \operatorname{tr}\bigl((\rho - \sigma)^2\bigr) .$$

Moreover, for a random variable X and a random state $\boldsymbol{\rho}$, we define

$$D(X|\boldsymbol{\rho}) := \Delta([\{X\} \otimes \boldsymbol{\rho}], [\{U\}] \otimes [\boldsymbol{\rho}])$$

where U is a random variable uniformly distributed on $\mathcal{X}$.

Lemma 5. *Let ρ and σ be two density operators on $\mathcal{H}$. Then*

$$\delta(\rho, \sigma) \leq \frac{1}{2}\sqrt{\operatorname{rank}(\rho - \sigma) \cdot \Delta(\rho, \sigma)} .$$

Proof. The assertion follows directly from Lemma 11 (cf. Appendix A) and the definition of the distance measures $\delta(\cdot, \cdot)$ and $\Delta(\cdot, \cdot)$. □

Lemma 6. *Let X be a random variable with range $\mathcal{X}$ and let $\boldsymbol{\rho}$ be a random state. Then*

$$d(X|\boldsymbol{\rho}) \leq \frac{1}{2} 2^{\frac{H_0(X) + S_0([\boldsymbol{\rho}])}{2}} \sqrt{D(X|\boldsymbol{\rho})} .$$

Proof. Note that the rank of $[\{X\} \otimes \boldsymbol{\rho}] - [\{U\}] \otimes [\boldsymbol{\rho}]$ is bounded by $2^{H_0(X) + S_0([\boldsymbol{\rho}])}$. The assertion thus follows as an immediate consequence of the definitions and Lemma 5. □

Lemma 7. *Let X be a random variable with range $\mathcal{X}$ and let $\boldsymbol{\rho}$ be a random state. Then*

$$D(X|\boldsymbol{\rho}) = \operatorname{tr}\Bigl(\Bigl(\sum_{x \in \mathcal{X}} P_X(x)^2 \rho_x^2\Bigr) - \frac{1}{|\mathcal{X}|}[\boldsymbol{\rho}]^2\Bigr)$$

where $\rho_x := [\boldsymbol{\rho}|X = x]$, for any $x \in \mathcal{X}$.

Proof. From (1) and the fact that $\operatorname{tr}(P_{|x\rangle} P_{|x'\rangle}) = \delta_{x,x'}$ (where $\delta_{x,x'}$ is the Kronecker delta which equals 1 if $x = x'$ and 0 otherwise), we find

$$\begin{aligned}
D(X|\boldsymbol{\rho}) &= \operatorname{tr}\Bigl(\Bigl(\sum_{x \in \mathcal{X}} P_X(x) P_{|x\rangle} \otimes \rho_x - \frac{1}{|\mathcal{X}|} \sum_{x \in \mathcal{X}} P_{|x\rangle} \otimes [\boldsymbol{\rho}]\Bigr)^2\Bigr) \\
&= \operatorname{tr}\Bigl(\sum_{x \in \mathcal{X}} \Bigl(P_X(x)\rho_x - \frac{1}{|\mathcal{X}|}[\boldsymbol{\rho}]\Bigr)^2\Bigr) \\
&= \operatorname{tr}\Bigl(\sum_{x \in \mathcal{X}} P_X(x)^2 \rho_x^2 - \frac{2}{|\mathcal{X}|}[\boldsymbol{\rho}] \sum_{x \in \mathcal{X}} P_X(x)\rho_x + \frac{1}{|\mathcal{X}|}[\boldsymbol{\rho}]^2\Bigr) .
\end{aligned}$$

Inserting the identity

$$[\boldsymbol{\rho}] = \sum_{x \in \mathcal{X}} P_X(x)\rho_x$$

concludes the proof. □

Lemma 8. *Let Z be a random variable, let $\boldsymbol{\rho}$ be a random state, and let F be a two-universal function on $\mathcal{Z}$ chosen independently of Z and $\boldsymbol{\rho}$. Then*

$$\mathop{\mathrm{E}}_{f \leftarrow P_F}\big[D(f(Z)|\boldsymbol{\rho})\big] \leq 2^{-S_2([\{Z\}\otimes\boldsymbol{\rho}])} \ .$$

Proof. Let us define $\rho_z := [\boldsymbol{\rho}|Z=z]$ for every $z \in \mathcal{Z}$ and let $\mathcal{S}$ be the range of F. With Lemma 7, we obtain

$$\mathop{\mathrm{E}}_{f \leftarrow P_F}\big[D(f(Z)|\boldsymbol{\rho})\big] = \mathrm{tr}\left(\mathop{\mathrm{E}}_{f \leftarrow P_F}\left[\sum_{s\in\mathcal{S}} P_{f(Z)}(s)^2[\boldsymbol{\rho}|f(Z)=s]^2\right]\right) - \frac{1}{|\mathcal{S}|}\mathrm{tr}([\boldsymbol{\rho}]^2) \ , \tag{7}$$

where we have used the linearity of the trace. Note that

$$P_{f(Z)}(s)\cdot[\boldsymbol{\rho}|f(Z)=s] \ = \sum_{z\in f^{-1}(\{s\})} P_Z(z)\rho_z \ .$$

Using this identity and rearranging the summation order, we get

$$\sum_{s\in\mathcal{S}} P_{f(Z)}(s)^2[\boldsymbol{\rho}|f(Z)=s]^2 = \sum_{z,z'\in\mathcal{Z}} P_Z(z)P_Z(z')\rho_z\rho_{z'}\delta_{f(z),f(z')} \ .$$

Taking the expectation value over the random choice of F then gives

$$\mathop{\mathrm{E}}_{f \leftarrow P_F}\left[\sum_{s\in\mathcal{S}} P_{f(Z)}(s)^2[\boldsymbol{\rho}|f(Z)=s]^2\right] = \sum_{z,z'\in\mathcal{Z}} P_Z(z)P_Z(z')\rho_z\rho_{z'} \mathop{\mathrm{Pr}}_{f \leftarrow P_F}[f(z)=f(z')] \ .$$

Similarly, we obtain

$$[\boldsymbol{\rho}]^2 = \sum_{z,z'\in\mathcal{Z}} P_Z(z)P_Z(z')\rho_z\rho_{z'} \ .$$

Inserting this into (7), we get

$$\mathop{\mathrm{E}}_{f \leftarrow P_F}[D(f(Z)|\boldsymbol{\rho})] = \sum_{z,z'\in\mathcal{Z}} P_Z(z)P_Z(z')\left(\mathop{\mathrm{Pr}}_{f \leftarrow P_F}[f(z)=f(z')] - \frac{1}{|\mathcal{S}|}\right)\mathrm{tr}(\rho_z\rho_{z'}) \ .$$

As we assumed that F is two-universal, all summands with $z \neq z'$ are not larger than zero and we are left with

$$\mathop{\mathrm{E}}_{f \leftarrow P_F}[D(f(Z)|\boldsymbol{\rho})] \leq \sum_{z\in\mathcal{Z}} P_Z(z)^2\mathrm{tr}(\rho_z^2) = \mathrm{tr}\big([\{Z\}\otimes\boldsymbol{\rho}]^2\big)$$

from which the assertion follows by the definition of the Rényi entropy S_2. □

Proof (Theorem 1). Using Lemma 4 and Lemma 6, we get

$$\begin{aligned}
d(F(Z)|\{F\}\otimes\boldsymbol{\rho}) &= \mathop{\mathrm{E}}_{f \leftarrow P_F}[d(f(Z)|\boldsymbol{\rho})] \\
&\leq \frac{1}{2}2^{\frac{s+S_0([\boldsymbol{\rho}])}{2}} \mathop{\mathrm{E}}_{f \leftarrow P_F}[\sqrt{D(f(Z)|\boldsymbol{\rho})}] \\
&\leq \frac{1}{2}2^{\frac{s+S_0([\boldsymbol{\rho}])}{2}} \sqrt{\mathop{\mathrm{E}}_{f \leftarrow P_F}[D(f(Z)|\boldsymbol{\rho})]} \ ,
\end{aligned}$$

where the last inequality follows from Jensen's inequality and the convexity of the square root. Applying Lemma 8 concludes the proof. □

4.2 A Bound in Terms of Smooth Rényi Entropy

The goal of this section is to reformulate Theorem 1 in terms of smooth Rényi entropy (cf. Corollary 1 below). Since, e.g., $S_0([\boldsymbol{\rho}])$ is generally larger than $S_0^\varepsilon([\boldsymbol{\rho}])$, this gives a better bound on the length of the extractable key. Indeed, for the situation where Z and $\boldsymbol{\rho}$ are obtained from many repetitions of the same random experiment, the bound in terms of smooth Rényi entropy is asymptotically optimal (cf. Section 4.4), which is not true if conventional Rényi entropy is used instead.

The following derivation is based on the idea that, for any normalized density operator ρ with smooth Rényi entropy $S_\alpha^\varepsilon(\rho)$, there exists a (not necessarily normalized) density operator ρ' which is ε-close to ρ such that the (conventional) Rényi entropy of ρ', $S_\alpha(\rho')$, is equal to $S_\alpha^\varepsilon(\rho)$.[8]

Lemma 9. *Let X be a random variable and let $\boldsymbol{\rho}$ be a normalized random state. Then, for any $\varepsilon \geq 0$, there exists a random variable X' and a random state $\boldsymbol{\rho}'$ with $\delta([\{X'\} \otimes \boldsymbol{\rho}'], [\{X\} \otimes \boldsymbol{\rho}]) \leq 2\sqrt{\varepsilon}$ such that, for any $\alpha > 1$,*

$$S_\alpha([\{X'\} \otimes \boldsymbol{\rho}']) - S_0([\boldsymbol{\rho}']) \geq S_\alpha^\varepsilon([\{X\} \otimes \boldsymbol{\rho}]) - S_0^\varepsilon([\boldsymbol{\rho}]) \ .$$

Proof. Let P be the projector onto the minimum subspace which corresponds to eigenvalues of $[\boldsymbol{\rho}]$ with total weight (at least) $1 - \varepsilon$, i.e.,

$$\mathrm{tr}(P[\boldsymbol{\rho}]P^\dagger) \geq 1 - \varepsilon \ . \tag{8}$$

It is easy to verify that $\log\bigl(\mathrm{rank}(P)\bigr) = S_0^\varepsilon([\boldsymbol{\rho}])$. Similarly, there exists a random variable X' and a random state $\boldsymbol{\sigma}$ with $\mathrm{tr}([\boldsymbol{\sigma}]) \leq \mathrm{tr}([\boldsymbol{\rho}]) = 1$ such that

$$S_\alpha([\{X'\} \otimes \boldsymbol{\sigma}]) = S_\alpha^\varepsilon([\{X\} \otimes \boldsymbol{\rho}])$$

and

$$\delta([\{X'\} \otimes \boldsymbol{\sigma}], [\{X\} \otimes \boldsymbol{\rho}]) \leq \frac{\varepsilon}{2} \ . \tag{9}$$

Let $\boldsymbol{\rho}'$ be the random state defined by $\boldsymbol{\rho}' := P\boldsymbol{\sigma}P^\dagger$. Then,

$$S_0([\boldsymbol{\rho}']) \leq \log\bigl(\mathrm{rank}(P)\bigr) = S_0^\varepsilon([\boldsymbol{\rho}])$$

and by Lemma 14 (see Appendix B), since $[\{X'\} \otimes \boldsymbol{\rho}']$ is the projection of $[\{X'\} \otimes \boldsymbol{\sigma}]$ (with respect to the projection operation $(\mathrm{id} \otimes P)$),

$$S_\alpha([\{X'\} \otimes \boldsymbol{\rho}']) \geq S_\alpha([\{X'\} \otimes \boldsymbol{\sigma}]) = S_\alpha^\varepsilon([\{X\} \otimes \boldsymbol{\rho}]) \ .$$

It thus remains to be shown that

$$\delta([\{X'\} \otimes \boldsymbol{\rho}'], [\{X\} \otimes \boldsymbol{\rho}]) \leq 2\sqrt{\varepsilon} \ . \tag{10}$$

[8] Note that $S_\alpha(\rho')$ is also defined for density operators ρ' with $\mathrm{tr}(\rho') < 1$.

Since the trace distance cannot increase when applying the projection P (cf. (4)), we obtain from (9)

$$\mathrm{tr}\big(\big|P[\boldsymbol{\sigma}]P^\dagger - P[\boldsymbol{\rho}]P^\dagger\big|\big) = 2\delta(P[\boldsymbol{\sigma}]P^\dagger, P[\boldsymbol{\rho}]P^\dagger) \leq 2\delta([\boldsymbol{\sigma}],[\boldsymbol{\rho}]) \leq \varepsilon \ .$$

Hence, with (8),

$$\mathrm{tr}([\boldsymbol{\rho}']) = \mathrm{tr}([P\boldsymbol{\sigma}P^\dagger]) \geq \mathrm{tr}([P\boldsymbol{\rho}P^\dagger]) - \mathrm{tr}\big(\big|P[\boldsymbol{\rho}]P^\dagger - P[\boldsymbol{\sigma}]P^\dagger\big|\big) \geq 1 - 2\varepsilon$$

and thus, from Lemma 12 (cf. Appendix A),

$$\delta([\{X'\}\otimes\boldsymbol{\rho}'],[\{X'\}\otimes\boldsymbol{\sigma}]) \leq \sqrt{\mathrm{tr}([\boldsymbol{\sigma}])(\mathrm{tr}([\boldsymbol{\sigma}]) - \mathrm{tr}([\boldsymbol{\rho}']))} \leq \sqrt{1-\mathrm{tr}([\boldsymbol{\rho}'])} \leq \sqrt{2\varepsilon} \ .$$

Using once again (9) and applying the triangle inequality for the trace distance implies (10) and thus concludes the proof. □

Using Lemma 9, the following corollary of Theorem 1 follows directly from the triangle inequality for the trace distance.

Corollary 1. *Let Z be a random variable with range $\mathcal{Z}$, let $\boldsymbol{\rho}$ be a normalized random state, let F be a two-universal function from $\mathcal{Z}$ to $\mathcal{S} = \{0,1\}^s$ which is independent of Z and $\boldsymbol{\rho}$, and let $\varepsilon \geq 0$. Then*

$$d(F(Z)|\{F\}\otimes\boldsymbol{\rho}) \leq \frac{1}{2}2^{-\frac{1}{2}(S_2^\varepsilon([\{Z\}\otimes\boldsymbol{\rho}]) - S_0^\varepsilon([\boldsymbol{\rho}]) - s)} + 4\sqrt{\varepsilon} \ .$$

Note that the smooth Rényi entropies occurring in the bound of Corollary 1 can easily be computed from the eigenvalues of the density operators $[\{Z\}\otimes\boldsymbol{\rho}] = \sum_z P_Z(z)P_{|z\rangle}\otimes\rho_z$ and $[\boldsymbol{\rho}] = \sum_z P_Z(z)\rho_z$, where $\rho_z = [\boldsymbol{\rho}|Z=z]$ (cf. Section 2.4).

4.3 Privacy Amplification Against Quantum Adversaries

We now apply the results of the previous section to show that privacy amplification by two-universal hashing is secure (with respect to the universally composable security definition of Section 3) against an adversary holding quantum information. Consider two distant parties which are connected by an authentic, but otherwise fully insecure classical communication channel. Additionally, they have access to a common random string Z about which an adversary has some partial information represented by the state $\boldsymbol{\rho}$ of a quantum system. The two legitimate parties can apply the following simple *privacy amplification protocol* to obtain a secure key S of length s. Let F be a two-universal random function from the range of Z to $\{0,1\}^s$. First, one of the parties randomly chooses an instance of F and announces his choice to the other party using the public communication channel. Then, both parties compute $S = F(Z)$.

Note that, during the execution of this protocol, the adversary might learn F. The final key S must thus be secure with respect to both $\{F\}$ and $\boldsymbol{\rho}$. It is an immediate consequence of Corollary 1 that, for any $\varepsilon \geq 0$, the key S generated by the described privacy amplification protocol is ε-secure with respect to $\boldsymbol{\rho}\otimes\{F\}$ if its length s is not larger than

$$s_\varepsilon = S_2^{\bar\varepsilon}([\{Z\}\otimes\boldsymbol{\rho}]) - S_0^{\bar\varepsilon}([\boldsymbol{\rho}]) - 2\log(1/\varepsilon) \ , \qquad (11)$$

where $\bar\varepsilon = (\varepsilon/8)^2$.

4.4 Asymptotic Optimality

We now show that the bound (11) is asymptotically optimal, i.e., that the right hand side of (11) is (in an asymptotic sense) also an upper bound for the number of key bits that can be extracted by any protocol. Consider a setting where both the initial information $Z^{(n)}$ as well as the adversary's state $\boldsymbol{\rho}^{(n)}$ consist of n independent pieces, for $n \in \mathbb{N}$. Formally, let $Z^{(n)} = (Z_1, \ldots, Z_n)$ and $\boldsymbol{\rho}^{(n)} = \boldsymbol{\rho}_1 \otimes \cdots \otimes \boldsymbol{\rho}_n$ where the pairs $(Z_i, \boldsymbol{\rho}_i)$ are independent and identically distributed. Let $s(n)$ be the length of the key $S^{(n)}$ that can be extracted from $Z^{(n)}$ by an optimal privacy amplification protocol. Using Lemma 3, we conclude from (11) that

$$s(n) \geq H(Z^{(n)}|\boldsymbol{\rho}^{(n)}) + o(n) \tag{12}$$

where, for any Z and $\boldsymbol{\rho}$, $H(Z|\boldsymbol{\rho})$ is defined in terms the von Neumann entropy $S(\cdot)$ by

$$H(Z|\boldsymbol{\rho}) := S([\{Z\} \otimes \boldsymbol{\rho}]) - S([\boldsymbol{\rho}]) \ .$$

To derive an upper bound for $s(n)$, consider an arbitrary privacy amplification protocol for generating a key $S^{(n)}$ from $Z^{(n)}$. Let $C^{(n)}$ be the whole communication exchanged over the public channel during the execution of the protocol, and let $f_{C^{(n)}}$ be the function depending on $C^{(n)}$ which describes how the final key $S^{(n)}$ is computed from $Z^{(n)}$, that is, $S^{(n)} = f_{C^{(n)}}(Z^{(n)})$.

It is a direct consequence of Definition 3 that the von Neumann entropy of an ε-secure key $S^{(n)}$ virtually cannot be smaller than its length $s(n)$, i.e.,

$$s(n) \leq H(f_{C^{(n)}}(Z^{(n)})|\boldsymbol{\rho}^{(n)} \otimes \{C^{(n)}\}) + o(n) \ . \tag{13}$$

Using some well-known properties of the von Neumann entropy, it is easy to see that the quantity $H(Z|\boldsymbol{\rho})$ can only decrease when applying any function f to its first argument or when introducing an additional random variable in the second argument. We thus have

$$H(f_{C^{(n)}}(Z^{(n)})|\boldsymbol{\rho}^{(n)} \otimes \{C^{(n)}\}) \leq H(Z^{(n)}|\boldsymbol{\rho}^{(n)} \otimes \{C^{(n)}\}) \leq H(Z^{(n)}|\boldsymbol{\rho}^{(n)}) \ . \tag{14}$$

Hence, combining (12), (13), and (14), we obtain an expression for the maximum number $s(n)$ of extractable key bits,

$$s(n) = H(Z^{(n)}|\boldsymbol{\rho}^{(n)}) + o(n) \ .$$

In particular, the maximum rate $R := \lim_{n\to\infty} \frac{s(n)}{n}$ at which secret key bits can be generated—from independent realizations of Z about which the adversary has information given by $\boldsymbol{\rho}$—is

$$R = S([\{Z\} \otimes \boldsymbol{\rho}]) - S([\boldsymbol{\rho}]) = H(Z|\boldsymbol{\rho}) \ . \tag{15}$$

This exactly corresponds to the expression for the secret key rate obtained by Devetak and Winter [15].

In the purely classical case, i.e., if the adversary's information is given by a classical random variable W, expression (15) reduces to

$$R = H(ZW) - H(W) = H(Z|W) ,$$

which is a well known result of Csiszár and Körner [14] (see also [24]).[9]

4.5 Applications to QKD

Theorem 1 has interesting implications for quantum key distribution (QKD). Recently, a generic protocol for QKD has been presented and proven secure against general attacks [13] (see also [23]). Moreover, it has been shown that many of the known protocols, such as BB84 or B92, are special instances of this generic protocol, i.e., their security directly follows from the security of the generic QKD protocol. Since the result in [13] is based on the security of privacy amplification, the strong type of security implied by Theorem 1 immediately carries over to this generic QKD protocol. In particular, the secret keys generated by the BB84 and the B92 protocol satisfy Definition 3 and thus provide universal composability.

5 Acknowledgment

The authors would like to thank Ueli Maurer for many inspiring discussions, and Dominic Mayers as well as anonymous referees for very useful comments. This project was partially supported by the Swiss National Science Foundation, project No. 200020-103847/1.

A Some Useful Identities

Lemma 10 (Schur's inequality). *Let A be a linear operator on a d-dimensional Hilbert space $\mathcal{H}$ and let $\lambda_1, \ldots, \lambda_d$ be its eigenvalues. Then*

$$\sum_{i=1}^{d} |\lambda_i|^2 \leq \mathrm{tr}(AA^\dagger) ,$$

with equality if and only if A is normal (i.e., $AA^\dagger = A^\dagger A$).

Proof. See, e.g., [20].

[9] In the setting of [14], the two parties are connected by a channel which leaks partial information to an adversary. As shown in [24], the result of [14] also applies if the two parties are connected by a completely public channel, but start with some common information Z about which an adversary has partial knowledge W.

Lemma 11. *Let A be a normal operator with rank r. Then*

$$\mathrm{tr}|A| \leq \sqrt{r}\sqrt{\mathrm{tr}(AA^\dagger)}\ .$$

Proof. Let $\lambda_1, \ldots, \lambda_r$ be the r nonzero eigenvalues of A. Since the square root is concave, we can apply Jensen's inequality leading to

$$\mathrm{tr}|A| = \sum_{i=1}^{r} |\lambda_i| = \sum_{i=1}^{r} \sqrt{|\lambda_i|^2} \leq \sqrt{r}\sqrt{\sum_{i=1}^{r} |\lambda_i|^2}\ .$$

The assertion then follows from Schur's inequality (Lemma 10). □

Lemma 12. *Let $\rho \in \mathcal{P}(\mathcal{H})$ and let P be a projection on $\mathcal{H}$, i.e., $P \circ P = P$. Then, for $\rho' := P\rho P^\dagger$,*

$$\delta(\rho, \rho') \leq \sqrt{\mathrm{tr}(\rho)(\mathrm{tr}(\rho) - \mathrm{tr}(\rho'))}\ .$$

Proof. We first show that the assertion holds for normalized pure states $\rho = P_{|\phi\rangle}$. Since P is a projection, there exist $a, b \in \mathbb{R}$ with $a^2 + b^2 = 1$ and two orthogonal vectors $|\alpha\rangle$, $|\beta\rangle$ with $P|\alpha\rangle = |\alpha\rangle$ and $P|\beta\rangle = 0$ such that $|\phi\rangle = a|\alpha\rangle + b|\beta\rangle$. In particular, $\rho' = a^2 P_{|\alpha\rangle}$. It then follows by a straightforward calculation that

$$\delta(\rho, \rho') = \delta(P_{a|\alpha\rangle + b|\beta\rangle}, a^2 P_{|\alpha\rangle}) \leq b = \sqrt{1 - \mathrm{tr}(\rho')}\ .$$

To prove the assertion for general density operators $\rho \in \mathcal{P}(\mathcal{H})$, let

$$\rho = \sum_{i \in \mathcal{I}} p_i \rho_i$$

where, for any $i \in \mathcal{I}$, $p_i \geq 0$ and ρ_i is a normalized pure state. In particular, $\sum_{i \in \mathcal{I}} p_i = \mathrm{tr}(\rho)$. By linearity, we have

$$\rho' = \sum_{i \in \mathcal{I}} p_i \rho'_i\ ,$$

where $\rho'_i := P\rho_i P^\dagger$. Hence, using the convexity of the trace distance,

$$\delta(\rho, \rho') \leq \sum_{i \in \mathcal{I}} p_i \delta(\rho_i, \rho'_i) \leq \sum_{i \in \mathcal{I}} p_i \sqrt{1 - \mathrm{tr}(\rho'_i)}\ .$$

The assertion then follows from Jensen's inequality. □

B Rényi Entropy and Quantum Operations

The following lemma states that the Rényi entropy of a density operator ρ can only increase when applying a quantum operation $\mathcal{E}$ on ρ.

Lemma 13. *Let $\mathcal{E} : \rho \mapsto \sum_i E_i \rho E_i^\dagger$ be a doubly stochastic quantum operation on $\mathcal{H}$, i.e., E_i are linear operators on $\mathcal{H}$ satisfying $\sum_i E_i^\dagger E_i = \mathrm{id}$ and $\sum_i E_i E_i^\dagger = \mathrm{id}$. Then, for any $\rho \in \mathcal{P}(\mathcal{H})$ and $\alpha \in [0, \infty]$,*

$$S_\alpha(\mathcal{E}(\rho)) \geq S_\alpha(\rho) \ .$$

Proof. See, e.g., [25] (Theorem 5.1 together with Theorem 4.2, applied to the function S_α).

Lemma 13 can be used to show that, for $\alpha > 1$, the Rényi entropy of a density operator ρ can only increase when applying a projector P to ρ.

Lemma 14. *Let $\rho \in \mathcal{P}(\mathcal{H})$ and let P be a projection on $\mathcal{H}$, i.e., $P \circ P = P$. Then, for $\alpha > 1$,*

$$S_\alpha(P \rho P^\dagger) \geq S_\alpha(\rho) \ .$$

Proof. Consider the quantum operation $\mathcal{E}$ defined by

$$\mathcal{E} : \quad \rho \longmapsto P \rho P^\dagger + (\mathrm{id} - P) \rho (\mathrm{id} - P)^\dagger \ .$$

It is easy to verify that $\mathcal{E}$ is doubly stochastic. Hence, from Lemma 13,

$$S_\alpha(\rho' + \rho'') \geq S_\alpha(\rho) \ ,$$

where $\rho' := P \rho P^\dagger$ and $\rho'' := (\mathrm{id} - P) \rho (\mathrm{id} - P)^\dagger$. The assertion then follows from the fact that, because ρ' and ρ'' are orthogonal,

$$\mathrm{tr}\big((\rho' + \rho'')^\alpha\big) \geq \mathrm{tr}\big((\rho')^\alpha\big) \ ,$$

and the definition of S_α.

C Smooth Rényi Entropy of Classical Distributions

Smooth Rényi entropy has been introduced in [28] as a generalization of Rényi entropy. For any set $\mathcal{Z}$, let $\bar{\mathcal{P}}(\mathcal{Z})$ be the set of non-negative functions P on $\mathcal{Z}$ such that $\sum_{z \in \mathcal{Z}} P(z) \leq 1$, i.e., $\bar{\mathcal{P}}(\mathcal{Z})$ contains all (not necessarily normalized) probability distributions on $\mathcal{Z}$. For any $P \in \bar{\mathcal{P}}(\mathcal{Z})$, let $\mathcal{B}^\varepsilon(P)$ be the set of functions $Q \in \bar{\mathcal{P}}(\mathcal{Z})$ such that $\delta(P, Q) := \frac{1}{2} \sum_z |P(z) - Q(z)| \leq \varepsilon$.

Definition 4. *Let $P \in \bar{\mathcal{P}}(\mathcal{Z})$ and let $\varepsilon \geq 0$. The ε-smooth Rényi entropy $H_\alpha^\varepsilon(P)$ of order α of P is defined by*

$$H_\alpha^\varepsilon(P) := \frac{1}{1 - \alpha} \log \left(\inf_{Q \in \mathcal{B}^{\varepsilon/2}(P)} \Big(\sum_{z \in \mathcal{Z}} Q(z)^\alpha \Big) \right) \ ,$$

for $\alpha \in (0, 1) \cup (1, \infty)$, and $H_\alpha^\varepsilon(P) := \lim_{\beta \to \alpha} H_\beta^\varepsilon(P)$, for $\alpha \in \{0, \infty\}$.

For a random variable Z with probability distribution P_Z, we also write $H_\alpha^\varepsilon(Z)$ instead of $H_\alpha^\varepsilon(P_Z)$.

It turns out that, for $\alpha < 1$, the logarithm on the right hand side of this definition takes its minimum for the function $Q \in \mathcal{B}^{\varepsilon/2}(P)$ which is obtained from P by setting the smallest probabilities to zero. Similarly, for $\alpha > 1$, the minimum is taken for the function Q obtained by cutting the largest probabilities of P. The smooth Rényi entropy $H_\alpha^\varepsilon(P)$ can thus easily be computed from the probabilities $P(z)$, for $z \in \mathcal{Z}$.

Smooth Rényi entropy has many natural properties which are similar to the properties of Shannon entropy. In particular, the smooth Rényi entropy of many independent and uniformly distributed random variables is close to the Shannon entropy.

Lemma 15. *Let $Z_1, \ldots, Z_n$ be independent random variables distributed according to P_Z. Then, for any $\alpha \neq 1$,*

$$\lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{1}{n} H_\alpha^\varepsilon(Z_1 \cdots Z_n) = H(Z) \ .$$

For a discussion of further properties and applications of smooth Rényi entropy, see [28].

References

1. A. Ambainis, L. J. Schulman, A. Ta-Shma, U. Vazirani, and A. Wigderson. The quantum communication complexity of sampling. In *Proceedings of the 39th Annual Symposium on Foundations of Computer Science*, pages 342–351, 1998.
2. M. Ben-Or. Security of BB84 QKD Protocol. Slides available at `http://www.msri.org/publications/ln/msri/2002/quantumintro/ben-or/2/`, 2002.
3. M. Ben-Or, M. Horodecki, D. Leung, D. Mayers, and J. Oppenheim. Composability of QKD. Slides available at `http://www.msri.org/publications/ln/msri/2002/qip/mayers/1/` (Part II), 2002.
4. M. Ben-Or, M. Horodecki, D. Leung, D. Mayers, and J. Oppenheim. The universal composable security of quantum key distribution. In *Proceedings of TCC 2005*, 2005.
5. M. Ben-Or and D. Mayers. Quantum universal composability. Slides available at `http://www.msri.org/publications/ln/msri/2002/quantumcrypto/mayers/1/banner/01.html`, 2002.
6. M. Ben-Or and D. Mayers. General security definition and composability for quantum & classical protocols. Available at `http://arxiv.org/abs/quant-ph/0409062`, 2004.
7. C. H. Bennett. Quantum cryptography using any two nonorthogonal states. *Physical Review Letters*, 68(21):3121–3124, 1992.
8. C. H. Bennett and G. Brassard. Quantum cryptography: Public-key distribution and coin tossing. In *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, pages 175–179, 1984.
9. C. H. Bennett, G. Brassard, C. Crépeau, and U. Maurer. Generalized privacy amplification. *IEEE Transaction on Information Theory*, 41(6):1915–1923, 1995.
10. C. H. Bennett, G. Brassard, and J.-M. Robert. Privacy amplification by public discussion. *SIAM Journal on Computing*, 17(2):210–229, 1988.

11. R. Canetti. Universally composable security: A new paradigm for cryptographic protocols. In *Proceedings of the 42nd IEEE Symposium on Foundations of Computer Science*, pages 136–145, 2001.
12. J. L. Carter and M. N. Wegman. Universal classes of hash functions. *Journal of Computer and System Sciences*, 18:143–154, 1979.
13. M. Christandl, R. Renner, and A. Ekert. A generic security proof for quantum key distribution. Available at `http://arxiv.org/abs/quant-ph/0402131`, February 2004.
14. I. Csiszár and J. Körner. Broadcast channels with confidential messages. *IEEE Transactions on Information Theory*, 24:339–348, 1978.
15. I. Devetak and A. Winter. Distillation of secret key and entanglement from quantum states. Available at `http://arxiv.org/abs/quant-ph/0306078`, June 2003.
16. D. DiVincenzo, M. Horodecki, D. Leung, J. Smolin, and B. Terhal. Locking classical correlation in quantum states. *Physical Review Letters*, 92, 067902, 2004.
17. S. Dziembowski and U. Maurer. Optimal randomizer efficiency in the bounded-storage model. *Journal of Cryptology*, 17(1):5–26, 2004. Conference version appeared in Proc. of STOC '02.
18. D. Gottesman and H.-K. Lo. Proof of security of quantum key distribution with two-way classical communications. *IEEE Transactions on Information Theory*, 49(2):457–475, 2003.
19. P. Hayden, D. Leung, P. W. Shor, and A. Winter. Randomizing quantum states: Constructions and applications *Communications in Mathematical Physics*, 250(2):371–391, 2004.
20. R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press, 1985.
21. R. Impagliazzo, L. A. Levin, and M. Luby. Pseudo-random generation from one-way functions (extended abstract). In *Proceedings of the Twenty-First Annual ACM Symposium on Theory of Computing*, pages 12–24, 1989.
22. R. König, U. Maurer, and R. Renner. On the power of quantum memory. Available at `http://arxiv.org/abs/quant-ph/0305154`, May 2003.
23. B. Kraus, N. Gisin, and R. Renner. Lower and upper bounds on the secret key rate for QKD protocols using one-way classical communication. Available at `http://arxiv.org/abs/quant-ph/0410215`, 2004.
24. U. M. Maurer. Secret key agreement by public discussion from common information. *IEEE Transactions on Information Theory*, 39(3):733–742, 1993.
25. M. A. Nielsen. Majorization and its applications to quantum information theory. Available at `http://www.qinfo.org/talks/1999/06-maj/maj.pdf`, June 1999.
26. M. A. Nielsen and I. L. Chuang. *Quantum computation and quantum information*. Cambridge University Press, 2000.
27. B. Pfitzmann and M. Waidner. Composition and integrity preservation of secure reactive systems. In *7th ACM Conference on Computer and Communications Security*, pages 245–254. ACM Press, 2000.
28. R. Renner and S. Wolf. Smooth Rényi entropy and applications. In *Proceedings of the 2004 IEEE International Symposium on Information Theory*, page 233, 2004.
29. A. Rényi. On measures of entropy and information. In *Proceedings of the 4th Berkeley Symp. on Math. Statistics and Prob.*, volume 1, pages 547–561. Univ. of Calif. Press, 1961.
30. D. Unruh. Simulatable security for quantum protocols. Available at `http://arxiv.org/abs/quant-ph/0409125`, 2004.
31. M. N. Wegman and J. L. Carter. New hash functions and their use in authentication and set equality. *Journal of Computer and System Sciences*, 22:265–279, 1981.