Security of Quantum Key Distribution

A dissertation submitted to

SWISS FEDERAL INSTITUTE OF TECHNOLOGY
ZURICH

for the degree of
Doctor of Natural Sciences

presented by

Renato Renner
Dipl. Phys. ETH

September 2005

2.1.4 Distance between states

Intuitively, we say that two states of a physical system are *similar* if any observation of them leads to identical results, except with small probability. For two operators $\rho, \rho' \in \mathcal{P}(\mathcal{H})$ representing the state of a quantum system, this notion of similarity is captured by the L_1-*distance*, i.e., the trace norm[14] $\|\rho - \rho'\|_1$ of the difference between ρ and ρ'.[15] The L_1-distance for operators can be seen as the quantum version of the L_1-distance for probability distributions [16] (or, more generally, nonnegative functions), which is defined by $\|P - P'\|_1 := \sum_z |P(z) - P'(z)|$, for $P, P' \in \mathcal{P}(\mathcal{Z})$. In particular, if ρ and ρ' are operator representations of probability distributions P and P', respectively, then the L_1-distance between ρ and ρ' is equal to the L_1-distance between P and P'.

[14] The *trace norm* $\|S\|_1$ of a hermitian operator S on $\mathcal{H}$ is defined by $\|S\|_1 := \mathrm{tr}(|S|)$.

[15] The L_1-distance between two operators is closely related to the *trace distance*, which is usually defined with an additional factor $\frac{1}{2}$.

[16] The L_1-distance between classical probability distributions is also known as *variational distance* or *statistical distance* (which are often defined with an additional factor $\frac{1}{2}$).

An operator A is *Hermitian* if $A=A^T$. Its eigenvalues are real.

Under the action of a quantum operation, the L_1-distance between two density operators ρ and ρ' cannot increase (cf. Lemma A.2.1). Because any measurement can be seen as a quantum operation, this immediately implies that the distance $\|P - P'\|_1$ between the distributions P and P' obtained from (identical) measurements of two density operators ρ and ρ', respectively, is bounded by $\|\rho - \rho'\|_1$.

The following proposition provides a very simple interpretation of the L_1-distance: If two probability distributions P and P' have L_1-distance at most 2ε, then the two settings described by P and P', respectively, cannot differ with probability more than ε.

Proposition 2.1.1. *Let $P, P' \in \mathcal{P}(\mathcal{X})$ be probability distributions. Then there exists a joint distribution $P_{XX'}$ such that P and P' are the marginals of $P_{XX'}$ (i.e., $P = P_X$, $P' = P_{X'}$) and, for (x, x') chosen according to $P_{XX'}$,*

$$\Pr_{(x,x')}[x \neq x'] \leq \frac{1}{2}\|P - P'\|_1 \ .$$

In particular, if the L_1-distance between two states is bounded by 2ε, then they cannot be distinguished with probability more than ε.

2.2 Universal security of secret keys

Cryptographic primitives (e.g., a secret key or an authentic communication channel) are often used as components within a more complex system. It is thus natural to require that the security of a cryptographic scheme is not compromised when it is employed as part of another system. This requirement is captured by the notion of *universal security*. Roughly speaking, we say that a cryptographic primitive is *universally secure* if it is secure in *any* arbitrary context. For example, the universal security of a secret key S implies that any bit of S remains secret even if some other part of S is given to an adversary.

Universal security definitions are usually based on the idea of characterizing the security of a *real* 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 a *perfect key* U, i.e., a uniformly distributed string which is independent of the adversary's information. As we shall see, such a definition immediately 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.

2.2.1 Standard security definitions are not universal

Unfortunately, many security definitions that are commonly used in quantum cryptography are not universal. For instance, the security of the key S generated by a QKD scheme is typically defined in terms of the mutual information $I(S;W)$ between S and the classical outcome W of a measurement of the adversary's system (see, e.g., [LC99, SP00, NC00, GL03, LCA05] and also the discussion in [BOHL$^+$05] and [RK05]). Formally, S is said to be secure if, for some small ε,

$$\max_W I(S;W) \leq \varepsilon \ , \tag{2.5}$$

where the maximum ranges over all measurements on the adversary's system with output W. Such a definition—although it looks reasonable—does, however, not guarantee that the key S can safely be used in applications. Roughly speaking, the reason for this flaw is that criterion (2.5) does not account for the fact that an adversary might wait with the measurement of her system until she learns parts of the key.

Let us illustrate this potential problem with a concrete example: Assume that we would like to use an n-bit key $S = (S_1, \ldots, S_n)$ as a one-time pad to encrypt an n-bit message $M = (M_1, \ldots, M_n)$.[18] Furthermore, assume that an adversary is interested in the nth bit M_n of the message, but already knows the first $n-1$ bits $M_1, \ldots, M_{n-1}$. Upon observing the ciphertext, the adversary can easily determine[19] the first $n-1$ bits of S. Hence, in order to guarantee the secrecy of the nth message bit M_n, we need to ensure that the adversary still has no information on the nth key bit S_n, even though she already knows all previous key bits $S_1, \ldots, S_{n-1}$. This requirement, however, is not implied by the above definition. Indeed, for any arbitrary $\varepsilon > 0$ and n depending on ε, it is relatively easy to construct examples which satisfy (2.5) whereas an adversary—once she knows the first $n-1$ bits of the key—can determine the nth bit S_n with certainty. For an explicit construction and analysis of such examples, we refer to [Bar05].[20]

[18] That is, the ciphertext $C = (C_1, \ldots, C_n)$ is the bit-wise XOR of S and M, i.e., $C_i = S_i \oplus M_i$.

[19] Note that $S_i = M_i \oplus C_i$.

[20] This phenomenon has also been studied in other contexts (see, e.g., [DHL$^+$04, HLSW04]) where it is called as *locking of classical correlation*.

2.2.2 A universal security definition

Consider a key S distributed according to P_S and let ρ_E^s be the state of the adversary's system given that S takes the value s, for any element s of the *key space* $\mathcal{S}$. According to the discussion in Section 2.1.3, the joint state of the classical key S and the adversary's quantum system can be represented by the density operator

$$\rho_{SE} := \sum_{s \in \mathcal{S}} P_S(s) |s\rangle\langle s| \otimes \rho_E^s \ ,$$

where $\{|s\rangle\}_{s \in \mathcal{S}}$ is an orthonormal basis of some Hilbert space $\mathcal{H}_S$. We say that S is *ε-secure with respect to* $\mathcal{H}_E$ if

$$\frac{1}{2} \left\| \rho_{SE} - \rho_U \otimes \rho_E \right\|_1 \leq \varepsilon \ , \tag{2.6}$$

where $\rho_U = \sum_{s \in \mathcal{S}} \frac{1}{|\mathcal{S}|} |s\rangle\langle s|$ is the fully mixed state on $\mathcal{H}_S$.

The universal security of a key S satisfying this definition follows from a simple argument: Criterion (2.6) guarantees that the *real* situation described by ρ_{SE} is ε-close—with respect to the L_1-distance—to an *ideal* situation where S is replaced by a perfect key U which is uniformly distributed and independent of the state of the system $\mathcal{H}_E$. Moreover, since the L_1-distance cannot increase when applying a quantum operation (cf. Lemma A.2.1), this also holds for any further evolution of the world (where, e.g., the key is used as part of a larger cryptographic system). In fact, it follows from the discussion in Section 2.1.4 that an ε-secure key can be considered *identical* to an *ideal* (perfect) key—except with probability ε.[21] In particular, an ε-secure key is secure within any reasonable framework providing universal composability (e.g., [BOM04] or [Unr04]).[22]

[21] For this statement to hold, it is crucial that the criterion (2.6) is formulated in terms of the L_1-distance (instead of other distance measures such as the fidelity).

[22] These frameworks are usually based on the so-called *simulatability paradigm*. That is, a real cryptosystem is said to be *as secure as* an ideal cryptosystem if any attack to the real scheme can be *simulated* by an attack to the ideal scheme (see also [MRH04]). It is easy to see that our security criterion is compatible with this paradigm: Consider a (real) key agreement protocol and assume that, for any possible attack of the adversary, the final key satisfies (2.6). The adversary's quantum state after the attack is then almost independent of the key, that is, the adversary could simulate virtually all her information without even interacting with the cryptosystem. The real key agreement protocol is thus *as secure as* an ideal key agreement scheme which, by definition, does not leak any information at all.

The security of a key according to (2.6) also implies security with respect to most of the standard security definitions in quantum cryptography. For example, if S is ε-secure with respect to $\mathcal{H}_E$ then the mutual information between S and the outcome of any measurement applied to the adversary's system is small (whereas the converse is often not true, as discussed above). In particular, if the adversary is purely classical, (2.6) reduces to a classical security definition which has been proposed in the context of information-theoretically secure key agreement (see, e,g., [DM04]).

Chapter 3

(Smooth) Min- and Max-Entropy

Entropy measures are indispensable tools in classical and quantum information theory. They quantify *randomness*, that is, the uncertainty that an observer has on the state of a (quantum) physical system. In this chapter, we introduce two entropic quantities, called *smooth min-entropy* and *smooth max-entropy*. As we shall see, these are useful to characterize randomness with respect to fundamental information-theoretic tasks such as the extraction of uniform randomness or data compression.[1] Moreover, smooth min- and max-entropies have natural properties which are similar to those known from the von Neumann entropy and its classical special case, the Shannon entropy[2] (Sections 3.1 and 3.2). In fact, for product states, smooth min- and max-entropy are asymptotically equal to the von Neumann entropy (Section 3.3).

[1] *Randomness extraction* is actually privacy amplification and is the topic of Chapter 5. *Data compression* is closely related to information reconciliation which is treated in Section 6.3.

[2] The *Shannon entropy* of a probability distribution P is defined by $H(P) := -\sum_x P(x) \log P(x)$, where log denotes the binary logarithm. Similarly, the *von Neumann entropy* of a density operator ρ is $H(\rho) := -\mathrm{tr}(\rho \log \rho)$.

Smooth min- and max-entropies are actually families of entropy measures parameterized by some nonnegative real number ε, called *smoothness*. In applications, the smoothness is related to the error probability of certain information-theoretic tasks and is thus typically chosen to be small. We first consider the "non-smooth" special case where $\varepsilon = 0$ (Section 3.1). This is the basis for the general definition where the smoothness ε is arbitrary (Section 3.2).

We will introduce a *conditional* version of smooth min- and max-entropy. It is defined for bipartite operators ρ_{AB} on $\mathcal{H}_A \otimes \mathcal{H}_B$ and measures the uncertainty on the state of the subsystem $\mathcal{H}_A$ given access to the subsystem $\mathcal{H}_B$. Unlike the *conditional von Neumann entropy* $H(A|B) := H(\rho_{AB}) - H(\rho_B)$, however, it cannot be written as a difference between two "unconditional" entropy measures.

To illustrate our definition of (conditional) min- and max-entropy, let us, as an analogy, consider an alternative formulation of the conditional von Neumann entropy $H(A|B)$. Let

$$H(\rho_{AB}|\sigma_B) := -\mathrm{tr}\big(\rho_{AB}(\log \rho_{AB} - \log \mathrm{id}_A \otimes \sigma_B)\big) \ , \tag{3.1}$$

for some state σ_B on $\mathcal{H}_B$. This quantity can be rewritten as

$$H(\rho_{AB}|\sigma_B) = H(\rho_{AB}) - H(\rho_B) - D(\rho_B\|\sigma_B) \ ,$$

where $D(\rho_B\|\sigma_B)$ is the relative entropy[3] of ρ_B to σ_B. Because $D(\rho_B\|\sigma_B)$ cannot be negative, this expression takes its maximum for $\sigma_B = \rho_B$, in which case it is equal to $H(A|B)$. We thus have

$$H(A|B) = \sup_{\sigma_B} H(\rho_{AB}|\sigma_B) \ , \tag{3.2}$$

where the supremum ranges over all density operators σ_B on $\mathcal{H}_B$.

The definitions of (smooth) min- and max-entropies are inspired by this approach. We first introduce a quantity which corresponds to (3.1) (cf. Definitions 3.1.1 and 3.2.1) and then define our entropy measures by a formula of the form (3.2) (Definitions 3.1.2 and 3.2.2).

[3]The *relative entropy* $D(\rho\|\sigma)$ is defined by $D(\rho\|\sigma) := \mathrm{tr}(\rho\log\rho) - \mathrm{tr}(\rho\log\sigma)$.

3.1 Min- and max-entropy

This section introduce a "non-smooth" version of min- and max-entropy. It is the basis for the considerations in Section 3.2, where these entropy measures are generalized. The focus is on min-entropy, which is used extensively in the remaining part of the thesis. However, most of the properties derived in the following also hold for max-entropy.

3.1.1 Definition of min- and max-entropy

Definition 3.1.1. Let $\rho_{AB} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$. The *min-entropy of ρ_{AB} relative to σ_B* is

$$H_{\min}(\rho_{AB}|\sigma_B) := -\log \lambda$$

where λ is the minimum real number such that $\lambda \cdot \mathrm{id}_A \otimes \sigma_B - \rho_{AB}$ is non-negative. The *max-entropy of ρ_{AB} relative to σ_B* is

$$H_{\max}(\rho_{AB}|\sigma_B) := \log \mathrm{tr}\big((\mathrm{id}_A \otimes \sigma_B)\rho^0_{AB}\big)$$

where ρ^0_{AB} denotes the projector onto the support of ρ_{AB}.

Definition 3.1.2. Let $\rho_{AB} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B)$. The *min-entropy* and the *max-entropy of* ρ_{AB} *given* $\mathcal{H}_B$ are

$$H_{\min}(\rho_{AB}|B) := \sup_{\sigma_B} H_{\min}(\rho_{AB}|\sigma_B)$$
$$H_{\max}(\rho_{AB}|B) := \sup_{\sigma_B} H_{\max}(\rho_{AB}|\sigma_B) \ ,$$

respectively, where the supremum ranges over all $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$ with $\mathrm{tr}(\sigma_B) = 1$.

Remark 3.1.3. It follows from Lemma B.5.3 that the min-entropy of ρ_{AB} relative to σ_B, for σ_B invertible, can be written as

$$H_{\min}(\rho_{AB}|\sigma_B) = -\log \lambda_{\max}\big((\mathrm{id}_A \otimes \sigma_B^{-1/2})\rho_{AB}(\mathrm{id}_A \otimes \sigma_B^{-1/2})\big) \ ,$$

where $\lambda_{\max}(\cdot)$ denotes the maximum eigenvalue of the argument.

If $\mathcal{H}_B$ is the trivial space $\mathbb{C}$, we simply write $H_{\min}(\rho_A)$ and $H_{\max}(\rho_A)$ to denote the min- and the max-entropy of ρ_A, respectively. In particular,

$$H_{\min}(\rho_A) = -\log \lambda_{\max}(\rho_A)$$
$$H_{\max}(\rho_A) = \log \mathrm{rank}(\rho_A) \ .$$

The classical analogue

The above definitions can be specialized canonically to classical probability distributions.[4] More precisely, for $P_{XY} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ and $Q_Y \in \mathcal{P}(\mathcal{Y})$, we have

$$H_{\min}(P_{XY}|Q_Y) := H_{\min}(\rho_{XY}|\sigma_Y)$$
$$H_{\max}(P_{XY}|Q_Y) := H_{\max}(\rho_{XY}|\sigma_Y)$$

where ρ_{XY} and σ_Y are the operator representations of P_{XY} and Q_Y, respectively (cf. Section 2.1.3).

Remark 3.1.4. Let $P_{XY} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ and $Q_Y \in \mathcal{P}(\mathcal{Y})$. Then[5]

$$H_{\min}(P_{XY}|Q_Y) = -\log \max_{y \in \mathrm{supp}(Q_Y)} \max_{x \in \mathcal{X}} \frac{P_{XY}(x,y)}{Q_Y(y)}$$
$$H_{\max}(P_{XY}|Q_Y) = \log \sum_{y \in \mathcal{Y}} Q_Y(y) \cdot \left|\mathrm{supp}(P_X^y)\right| ,$$

where P_X^y denotes the function $P_X^y : x \mapsto P_{XY}(x,y)$. In particular,

$$H_{\max}(P_{XY}|Y) = \log \max_{y \in \mathcal{Y}} \left|\mathrm{supp}(P_X^y)\right| .$$

[4] Similarly, the Shannon entropy can be seen as the classical special case of the von Neumann entropy.

[5] The *support* of a nonnegative function $f \in \mathcal{P}(\mathcal{X})$, denoted $\mathrm{supp}(f)$, is the set of values $x \in \mathcal{X}$ such that $f(x) > 0$.

3.1.2 Basic properties of min- and max-entropy

Min-entropy cannot be larger than max-entropy

The following lemma gives a relation between min- and max-entropy. It implies that, for a density operator ρ_{AB}, the min-entropy cannot be larger than the max-entropy.

Lemma 3.1.5. *Let $\rho_{AB} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$. Then*

$$H_{\min}(\rho_{AB}|\sigma_B) + \log \operatorname{tr}(\rho_{AB}) \leq H_{\max}(\rho_{AB}|\sigma_B) \ .$$

Additivity of min- and max-entropy

The von Neumann entropy of a state which consists of two independent parts is equal to the sum of the entropies of each part, i.e., $H(\rho_A \otimes \rho_{A'}) = H(\rho_A) + H(\rho_{A'})$. This also holds for min- and max-entropy.

Lemma 3.1.6. *Let $\rho_{AB} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B)$, $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$ and, similarly, $\rho_{A'B'} \in \mathcal{P}(\mathcal{H}_{A'} \otimes \mathcal{H}_{B'})$, $\sigma_{B'} \in \mathcal{P}(\mathcal{H}_{B'})$. Then*

$$H_{\min}(\rho_{AB} \otimes \rho_{A'B'}|\sigma_B \otimes \sigma_{B'}) = H_{\min}(\rho_{AB}|\sigma_B) + H_{\min}(\rho_{A'B'}|\sigma_{B'})$$
$$H_{\max}(\rho_{AB} \otimes \rho_{A'B'}|\sigma_B \otimes \sigma_{B'}) = H_{\max}(\rho_{AB}|\sigma_B) + H_{\max}(\rho_{A'B'}|\sigma_{B'}) \ .$$

Strong subadditivity

The von Neumann entropy is subadditive, i.e., $H(A|BC) \leq H(A|B)$, which means that the entropy cannot increase when conditioning on an additional subsystem. This property can be generalized to min- and max-entropy.

Lemma 3.1.7. *Let $\rho_{ABC} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ and $\sigma_{BC} \in \mathcal{P}(\mathcal{H}_B \otimes \mathcal{H}_C)$. Then*

$$H_{\min}(\rho_{ABC}|\sigma_{BC}) \leq H_{\min}(\rho_{AB}|\sigma_B)$$
$$H_{\max}(\rho_{ABC}|\sigma_{BC}) \leq H_{\max}(\rho_{AB}|\sigma_B) \ .$$

Conditioning on classical information

The min- and max-entropies of states which are partially classical can be expressed in terms of the min- and max-entropies of the corresponding conditional operators (see Section 2.1.3).

Lemma 3.1.8. *Let $\rho_{ABZ} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_Z)$ and $\sigma_{BZ} \in \mathcal{P}(\mathcal{H}_B \otimes \mathcal{H}_Z)$ be classical with respect to an orthonormal basis $\{|z\rangle\}_{z\in\mathcal{Z}}$ of $\mathcal{H}_Z$, and let ρ_{AB}^z and σ_B^z be the corresponding (non-normalized) conditional operators. Then*

$$H_{\min}(\rho_{ABZ}|\sigma_{BZ}) = \inf_{z\in\mathcal{Z}} H_{\min}(\rho_{AB}^z|\sigma_B^z)$$
$$H_{\max}(\rho_{ABZ}|\sigma_{BZ}) = \log \sum_{z\in\mathcal{Z}} 2^{H_{\max}(\rho_{AB}^z|\sigma_B^z)} \ .$$

Classical subsystems have nonnegative min-entropy

Similarly to the conditional von Neumann entropy, the min- and max-entropies of entangled systems can generally be negative. This is, however, not the case for the entropy of a classical subsystem. Lemma 3.1.9 below implies that

$$H_{\min}(\rho_{XC}|\rho_C) \geq 0 \ ,$$

for any density operator ρ_{XC} which is classical on the first subsystem[6]. By Lemma 3.1.5, the same holds for max-entropy.

Lemma 3.1.9. *Let $\rho_{XBC} \in \mathcal{P}(\mathcal{H}_X \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ be classical on $\mathcal{H}_X$ and let $\sigma_C \in \mathcal{P}(\mathcal{H}_C)$. Then*

$$H_{\min}(\rho_{XBC}|\sigma_C) \geq H_{\min}(\rho_{BC}|\sigma_C) \ .$$

[6]To see this, let $\mathcal{H}_B$ be the trivial space $\mathbb{C}$ and set $\sigma_C = \rho_C$.

3.1.3 Chain rules for min-entropy

The chain rule for the von Neumann entropy reads $H(AB|C) = H(A|BC) + H(B|C)$. In particular, since $H(B|C)$ cannot be larger than $H(B)$, we have $H(AB|C) \leq H(A|BC) + H(B)$. The following lemma implies that a similar statement holds for min-entropy, namely,

$$H_{\min}(\rho_{ABC}|C) \leq H_{\min}(\rho_{ABC}|BC) + H_{\max}(\rho_B) \ .$$

Lemma 3.1.10. *Let $\rho_{ABC} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$, $\sigma_C \in \mathcal{P}(\mathcal{H}_C)$, and let $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$ be the fully mixed state on the support of ρ_B. Then*

$$H_{\min}(\rho_{ABC}|\sigma_C) = H_{\min}(\rho_{ABC}|\sigma_B \otimes \sigma_C) + H_{\max}(\rho_B) \ .$$

Data processing

Let A, Y, and C be random variables such that $A \leftrightarrow Y \leftrightarrow C$ is a *Markov chain*, i.e., the conditional probability distributions $P_{AC|Y=y}$ have product form $P_{A|Y=y} \times P_{C|Y=y}$. The uncertainty on A given Y is then equal to the uncertainty on A given Y and C, that is, in terms of Shannon entropy, $H(A|Y) = H(A|YC)$. Hence, by the chain rule, we get the equality $H(AY|C) = H(Y|C) + H(A|Y)$.

The same equality also holds for quantum states ρ_{AYC} on $\mathcal{H}_A \otimes \mathcal{H}_Y \otimes \mathcal{H}_C$ which are classical on $\mathcal{H}_Y$ and where, analogously to the Markov condition, the conditional density operators $\bar{\rho}^y_{AC}$ have product form, i.e., $\bar{\rho}^y_{AC} = \bar{\rho}^y_A \otimes \bar{\rho}^y_C$. The following lemma generalizes this statement to min-entropy.

Lemma 3.1.11. *Let $\rho_{AYC} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_Y \otimes \mathcal{H}_C)$ be classical with respect to an orthonormal basis $\{|y\rangle\}_{y\in\mathcal{Y}}$ of $\mathcal{H}_Y$ such that the corresponding conditional operators ρ^y_{AC}, for any $y \in \mathcal{Y}$, have product form and let $\sigma_C \in \mathcal{P}(\mathcal{H}_C)$. Then*

$$H_{\min}(\rho_{AYC}|\sigma_C) \geq H_{\min}(\rho_{YC}|\sigma_C) + H_{\min}(\rho_{AY}|\rho_Y) \ .$$

3.1.4 Quantum operations can only increase min-entropy

The min-entropy can only increase when applying quantum operations. Because the partial trace is a quantum operation, this general statement also implies the first assertion of Lemma 3.1.7 (strong subadditivity).

Lemma 3.1.12. *Let $\rho_{AB} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B)$, $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$, $\tilde{\sigma}_{B'} \in \mathcal{P}(\mathcal{H}_{B'})$ and let $\mathcal{E}$ be a CPM from $\mathcal{H}_A \otimes \mathcal{H}_B$ to $\mathcal{H}_{A'} \otimes \mathcal{H}_{B'}$ such that $\mathrm{id}_{A'} \otimes \tilde{\sigma}_{B'} - \mathcal{E}(\mathrm{id}_A \otimes \sigma_B)$ is nonnegative. Then, for $\tilde{\rho}_{A'B'} := \mathcal{E}(\rho_{AB})$,*

$$H_{\min}(\tilde{\rho}_{A'B'}|\tilde{\sigma}_{B'}) \geq H_{\min}(\rho_{AB}|\sigma_B) \ .$$

3.1.5 Min-entropy of superpositions

Let $\{|x\rangle\}_{x\in\mathcal{X}}$ be an orthonormal basis on $\mathcal{H}_X$, let $\{|\psi^x\rangle\}_{x\in\mathcal{X}}$ be a family of vectors on $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_E$, and define

$$\rho_{ABE} := |\psi\rangle\langle\psi| \qquad \text{where } |\psi\rangle := \sum_{x\in\mathcal{X}} |\psi^x\rangle \tag{3.4}$$

$$\tilde{\rho}_{ABEX} := \sum_{x\in\mathcal{X}} |\psi^x\rangle\langle\psi^x| \otimes |x\rangle\langle x| \ . \tag{3.5}$$

Note that, if the states $|\psi^x\rangle$ are orthogonal then $\tilde{\rho}_{ABEX}$ can be seen as the state resulting from an orthogonal measurement of ρ_{ABE} with respect to the projectors along $|\psi^x\rangle$. While ρ_{ABE} is a *superposition* (linear combination) of vectors $|\psi^x\rangle$, $\tilde{\rho}_{ABE}$ is a *mixture* of vectors $|\psi^x\rangle$. The following lemma gives a lower bound on the min-entropy of ρ_{ABE} in terms of the min-entropy of $\tilde{\rho}_{ABE}$.

Lemma 3.1.13. *Let ρ_{ABE} and $\tilde{\rho}_{ABEX}$ be defined by* (3.4) *and* (3.5)*, respectively, and let $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$. Then*

$$H_{\min}(\rho_{AB}|\sigma_B) \geq H_{\min}(\tilde{\rho}_{AB}|\sigma_B) - H_{\max}(\tilde{\rho}_X) \ .$$

Lemma 3.1.14. *Let ρ_{ABE}, $\tilde{\rho}_{ABEX}$ be defined by* (3.4) *and* (3.5)*, respectively, and let $\sigma_{BX} \in \mathcal{P}(\mathcal{H}_B \otimes \mathcal{H}_{\mathcal{X}})$. Then*

$$H_{\min}(\rho_{AB}|\sigma_B) \geq H_{\min}(\tilde{\rho}_{ABX}|\sigma_{BX}) - H_{\max}(\tilde{\rho}_X) \ .$$

3.2 Smooth min- and max-entropy

The min-entropy and the max-entropy, as defined in the previous section, are discontinuous in the sense that a slight modification of the system's state might have a large impact on its entropy. To illustrate this, consider for example a classical random variable X on the set $\{0, \ldots, n-1\}$ which takes the values 0 and 1 with probability almost one half, i.e., $P_X(0) = P_X(1) = \frac{1-\varepsilon}{2}$, for some small $\varepsilon > 0$, whereas the other values have equal probabilities, i.e., $P_X(x) = \frac{\varepsilon}{n-2}$, for all $x > 1$. Then, by the definition of the max-entropy, $H_{\max}(P_X) = \log n$. On the other hand, if we slightly change the probability distribution P_X to some probability distribution $\bar{P}_X$ such that $\bar{P}_X(x) = 0$, for all $x > 1$, then $H_{\max}(\bar{P}_X) = 1$. In particular, for n large, $H_{\max}(P_X) \gg H_{\max}(\bar{P}_X)$, while $\|P_X - \bar{P}_X\|_1 \leq \varepsilon$.

We will see later (cf. Section 6.3) that the max-entropy $H_{\max}(P_X)$ can be interpreted as the minimum number of bits needed to encode X in such a way that its value can be recovered from the encoding without errors. The above example is consistent with this interpretation. Indeed, while we need at least $\log n$ bits to store a value X distributed according to P_X, one single bit is sufficient to store a value distributed according to $\bar{P}_X$. However, for most applications, we allow some small error probability. For example, we might want to encode X in such a way that its value can be recovered with probability $1 - \varepsilon$. Obviously, in this case, one single bit is sufficient to store X even if it is distributed according to P_X.

The example illustrates that, given some probability distribution P_X, one might be interested in the maximum (or minimum) entropy of any distribution $\bar{P}_X$ which is close to P_X. This idea is captured by the notion of smooth min- and max-entropy.

3.2.1 Definition of smooth min- and max-entropy

The definition of smooth min- and max-entropy is based on the "non-smooth" version (Definition 3.1.1).

Definition 3.2.1. Let $\rho_{AB} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B)$, $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$, and $\varepsilon \geq 0$. The *ε-smooth min-entropy* and the *ε-smooth max-entropy of ρ_{AB} relative to σ_B* are

$$H_{\min}^{\varepsilon}(\rho_{AB}|\sigma_B) := \sup_{\bar{\rho}_{AB}} H_{\min}(\bar{\rho}_{AB}|\sigma_B)$$
$$H_{\max}^{\varepsilon}(\rho_{AB}|\sigma_B) := \inf_{\bar{\rho}_{AB}} H_{\max}(\bar{\rho}_{AB}|\sigma_B) \ ,$$

where the supremum and infimum ranges over the set $\mathcal{B}^{\varepsilon}(\rho_{AB})$ of all operators $\bar{\rho}_{AB} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B)$ such that $\|\bar{\rho}_{AB} - \rho_{AB}\|_1 \leq \mathrm{tr}(\rho_{AB}) \cdot \varepsilon$ and $\mathrm{tr}(\bar{\rho}_{AB}) \leq \mathrm{tr}(\rho_{AB})$.

Definition 3.2.2. Let $\rho_{AB} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and let $\varepsilon \geq 0$. The *ε-smooth min-entropy* and the *ε-smooth max-entropy of ρ_{AB} given $\mathcal{H}_B$* are

$$H_{\min}^{\varepsilon}(\rho_{AB}|B) := \sup_{\sigma_B} H_{\min}^{\varepsilon}(\rho_{AB}|\sigma_B)$$
$$H_{\max}^{\varepsilon}(\rho_{AB}|B) := \sup_{\sigma_B} H_{\max}^{\varepsilon}(\rho_{AB}|\sigma_B) \ ,$$

where the supremum ranges over all $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$ with $\mathrm{tr}(\sigma_B) = 1$.

Note that, similar to the description in Section 3.1, these definitions can be specialized to classical probability distributions.

3.2.2 Basic properties of smooth min-entropy

Superadditivity

The following is a generalization of (one direction of) Lemma 3.1.6 to smooth min-entropy.

Lemma 3.2.6. *Let $\rho_{AB} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B)$, $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$ and, similarly, $\rho_{A'B'} \in \mathcal{P}(\mathcal{H}_{A'} \otimes \mathcal{H}_{B'})$, $\sigma_{B'} \in \mathcal{P}(\mathcal{H}_{B'})$, and let $\varepsilon, \varepsilon' \geq 0$. Then*

$$H_{\min}^{\varepsilon+\varepsilon'}(\rho_{AB} \otimes \rho_{A'B'}|\sigma_B \otimes \sigma_{B'}) \geq H_{\min}^{\varepsilon}(\rho_{AB}|\sigma_B) + H_{\min}^{\varepsilon'}(\rho_{A'B'}|\sigma_{B'}) \ .$$

Strong subadditivity

The following statement is a generalization of Lemma 3.1.7 to smooth min-entropy.

Lemma 3.2.7. *Let $\rho_{ABC} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$, $\sigma_{BC} \in \mathcal{P}(\mathcal{H}_B \otimes \mathcal{H}_C)$, and let $\varepsilon \geq 0$. Then*

$$H_{\min}^{\varepsilon}(\rho_{ABC}|\sigma_{BC}) \leq H_{\min}^{\varepsilon}(\rho_{AB}|\sigma_B) \ .$$

Conditioning on classical information

The following lemma generalizes (one direction of) Lemma 3.1.8 to smooth min-entropy.

Lemma 3.2.8. *Let $\rho_{ABZ} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_Z)$ and $\sigma_{BZ} \in \mathcal{P}(\mathcal{H}_B \otimes \mathcal{H}_Z)$ be classical with respect to an orthonormal basis $\{|z\rangle\}_{z\in\mathcal{Z}}$ of $\mathcal{H}_Z$, let ρ_{AB}^z and σ_B^z be the corresponding (non-normalized) conditional operators, and let $\varepsilon \geq 0$. Then*

$$H_{\min}^{\varepsilon}(\rho_{ABZ}|\sigma_{BZ}) \geq \inf_{z\in\mathcal{Z}} H_{\min}^{\varepsilon}(\rho_{AB}^z|\sigma_B^z) \ .$$

3.2.3 Chain rules for smooth min-entropy

The following lemma generalizes (one direction of) Lemma 3.1.10 to smooth min-entropy.

Lemma 3.2.9. *Let $\rho_{ABC} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$, $\sigma_C \in \mathcal{P}(\mathcal{H}_C)$, let $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$ be the fully mixed state on the support of ρ_B, and let $\varepsilon \geq 0$. Then*

$$H_{\min}^{\varepsilon}(\rho_{ABC}|\sigma_C) \leq H_{\min}^{\varepsilon}(\rho_{ABC}|\sigma_B \otimes \sigma_C) + H_{\max}(\rho_B) \ .$$

Data processing

The following lemma is a generalization of Lemma 3.1.11 to smooth min-entropy.

Lemma 3.2.10. *Let $\rho_{AYC} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_Y \otimes \mathcal{H}_C)$ be classical with respect to an orthonormal basis $\{|y\rangle\}_{y\in\mathcal{Y}}$ of $\mathcal{H}_Y$ such that the corresponding conditional operators ρ_{AC}^{y}, for any $y \in \mathcal{Y}$, have product form, let $\sigma_C \in \mathcal{P}(\mathcal{H}_C)$, and let $\varepsilon \geq 0$. Then*

$$H_{\min}^{\varepsilon}(\rho_{AYC}|\sigma_C) \geq H_{\min}^{\varepsilon}(\rho_{YC}|\sigma_C) + H_{\min}(\rho_{AY}|\rho_Y) \ .$$

3.2.4 Smooth min-entropy of superpositions

The following statement generalizes Lemma 3.1.14.

Lemma 3.2.11. *Let ρ_{ABE}, $\tilde{\rho}_{ABEX}$ be defined by (3.4) and (3.5), respectively, for mutually orthogonal vectors $|\psi^x\rangle$, let $\sigma_{BX} \in \mathcal{P}(\mathcal{H}_B \otimes \mathcal{H}_{\mathcal{X}})$, and let $\varepsilon \geq 0$. Then*

$$H_{\min}^{\varepsilon}(\rho_{AB}|\sigma_B) \geq H_{\min}^{\tilde{\varepsilon}}(\tilde{\rho}_{ABX}|\sigma_{BX}) - H_{\max}(\tilde{\rho}_X) \ ,$$

where $\tilde{\varepsilon} = \frac{\varepsilon^2}{6|\mathcal{X}|}$.

3.2.5 Smooth min-entropy calculus

Theorem 3.2.12. *Let $\varepsilon, \varepsilon' \geq 0$. Then the following inequalities hold:*

- *(Super-)additivity:*

$$H_{\min}^{\varepsilon+\varepsilon'}(\rho_{AB} \otimes \rho_{A'B'}|BB') \geq H_{\min}^{\varepsilon}(\rho_{AB}|B) + H_{\min}^{\varepsilon'}(\rho_{A'B'}|B') \ , \quad (3.18)$$

  for $\rho_{AB} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and $\rho_{A'B'} \in \mathcal{P}(\mathcal{H}_{A'} \otimes \mathcal{H}_{B'})$.

- *Strong subadditivity:*

$$H_{\min}^{\varepsilon}(\rho_{ABC}|BC) \leq H_{\min}^{\varepsilon}(\rho_{AB}|B) \ , \quad (3.19)$$

  for $\rho_{ABC} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$.

- *Conditioning on classical information:*

$$H_{\min}^{\varepsilon}(\rho_{ABZ}|BZ) \geq \inf_{z \in \mathcal{Z}} H_{\min}^{\varepsilon}(\bar{\rho}_{AB}^{z}|B) \ , \quad (3.20)$$

  for $\rho_{ABZ} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_Z)$ normalized and classical on $\mathcal{H}_Z$, and for normalized conditional operators $\bar{\rho}_{AB}^{z}$.

- *Chain rule:*

$$H_{\min}^{\varepsilon}(\rho_{ABC}|C) \leq H_{\min}^{\varepsilon}(\rho_{ABC}|BC) + H_{\max}(\rho_B) \ , \quad (3.21)$$

  for $\rho_{ABC} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$.

- *Data processing:*

$$H_{\min}^{\varepsilon}(\rho_{AYC}|C) \geq H_{\min}^{\varepsilon}(\rho_{YC}|C) + H_{\min}(\rho_{AY}|\rho_Y) \ , \quad (3.22)$$

  for $\rho_{AYC} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_Y \otimes \mathcal{H}_C)$ classical on $\mathcal{H}_Y$ such that the conditional operators ρ_{AC}^{y} have product form.

3.3 Smooth min- and max-entropy of products

In this section, we show that the smooth min- and max-entropies of product states are asymptotically equal to the von Neumann entropy. In a first step, we consider a purely classical situation, i.e., we prove that the smooth min- and max-entropies of a sequence of independent and identically distributed random variables can be expressed in terms of Shannon entropy (which is the classical analogue of the von Neumann entropy). Then, in a second step, we generalize this statement to quantum states (Section 3.3.2).

3.3.1 The classical case

The proof of the main result of this section (Theorem 3.3.4) is based on a Chernoff style bound (Theorem 3.3.3) which is actually a variant of the asymptotic equipartition property (AEP) known from information theory (see, e.g., [CT91]). It states that, with high probability, the negative logarithm of the probability of an n-tuple of values chosen according to a product distribution P^n is close to the Shannon entropy of P^n.

Typical sequences and their probabilities

Lemma 3.3.1. *Let $P_{XY} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ be a probability distribution. Then, for any $t \in \mathbb{R}$ with $|t| \leq \frac{1}{\log(|\mathcal{X}|+3)}$,*

$$\log \mathop{\mathbb{E}}_{x,y} \left[P_{X|Y}(x,y)^{-t} \right] \leq tH(X|Y) + \frac{1}{2} t^2 \log(|\mathcal{X}| + 3)^2 \ ,$$

where the expectation is taken over pairs (x, y) chosen according to P_{XY}.

Lemma 3.3.2. *Let $P_{XY} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ be a probability distribution and let γ be the function on $\mathcal{X} \times \mathcal{Y}$ defined by*

$$\gamma(x, y) := -\log P_{X|Y}(x, y) - H(X|Y) \ .$$

Then, for any $t \in \mathbb{R}$ with $|t| \leq \frac{1}{\log(|\mathcal{X}|+3)}$,

$$\mathop{\mathbb{E}}_{x,y}\left[2^{t\gamma(x,y)}\right] \leq 2^{\frac{1}{2}t^2 \log(|\mathcal{X}|+3)^2} \ .$$

Theorem 3.3.3. *Let $P_{XY} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ be a probability distribution and let $n \in \mathbb{N}$. Then, for any $\delta \in [0, \log|\mathcal{X}|]$ and $(\mathbf{x}, \mathbf{y})$ chosen according to $P_{X^nY^n} := (P_{XY})^n$,*

$$\Pr_{\mathbf{x},\mathbf{y}}\left[-\log P_{X^n|Y^n}(\mathbf{x}, \mathbf{y}) \geq n\big(H(X|Y) + \delta\big)\right] \leq 2^{-\frac{n\delta^2}{2\log(|\mathcal{X}|+3)^2}} \ ,$$

and, similarly,

$$\Pr_{\mathbf{x},\mathbf{y}}\left[-\log P_{X^n|Y^n}(\mathbf{x}, \mathbf{y}) \leq n\big(H(X|Y) - \delta\big)\right] \leq 2^{-\frac{n\delta^2}{2\log(|\mathcal{X}|+3)^2}} \ .$$

Asymptotic equality of smooth entropy and Shannon entropy

Theorem 3.3.4. *Let $P_{XY} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ be a probability distribution and let $n \in \mathbb{N}$. Then, for any $\varepsilon \geq 0$ and $P_{X^nY^n} := (P_{XY})^n$,*

$$\frac{1}{n} H_{\max}^{\varepsilon}(P_{X^nY^n}|P_{Y^n}) \leq H(X|Y) + \delta$$
$$\frac{1}{n} H_{\min}^{\varepsilon}(P_{X^nY^n}|P_{Y^n}) \geq H(X|Y) - \delta \ ,$$

where $\delta := \log(|\mathcal{X}| + 3)\sqrt{\frac{2\log(1/\varepsilon)}{n}}$.

Because the min-entropy $H_{\min}(P_{X^nY^n}|P_{Y^n})$ cannot be larger than the max-entropy $H_{\max}(P_{X^nY^n}|P_{Y^n})$ (cf. Lemma 3.1.5), Theorem 3.3.4 implies that

$$\frac{1}{n} H_{\min}^{\varepsilon}(P_{X^nY^n}|P_{Y^n}) \approx \frac{1}{n} H_{\max}^{\varepsilon}(P_{X^nY^n}|P_{Y^n}) \approx \frac{1}{n} H(X^n|Y^n) \ , \qquad (3.29)$$

where asymptotically, for increasing n, the approximation becomes an equality.

Remark 3.3.5. It is easy to see that Theorem 3.3.4 can be generalized to probability distributions $P_{X^nY^n}$ which are the product of not necessarily identical distributions $P_{X_iY_i}$. That is, for any distribution of the form $P_{X^nY^n} = \prod_{i=1}^{n} P_{X_iY_i}$, the approximation (3.29) still holds.

3.3.2 The quantum case

The following theorem and its corollary can be seen as a quantum version of Theorem 3.3.4 for smooth min-entropy (where the Shannon entropy is replaced by the von Neumann entropy). The proof essentially follows the same line as the classical argument described above. A similar argument shows that the statement also holds for smooth max-entropy.

Theorem 3.3.6. *Let $\rho_{AB} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B)$, $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$ be density operators, and let $n \in \mathbb{N}$. Then, for any $\varepsilon \geq 0$,*

$$\frac{1}{n} H_{\min}^{\varepsilon}(\rho_{AB}^{\otimes n} | \sigma_B^{\otimes n}) \geq H(\rho_{AB}) - H(\rho_B) - D(\rho_B \| \sigma_B) - \delta \ ,$$

where $\delta := 2 \log\big(\mathrm{rank}(\rho_A) + \mathrm{tr}(\rho_{AB}^2 (\mathrm{id}_A \otimes \sigma_B^{-1})) + 2\big) \sqrt{\frac{\log(1/\varepsilon)}{n} + 1}$.

The following corollary specializes Theorem 3.3.6 to the case where the first part of the state $\rho_{AB} = \rho_{XB}$ is classical and where $\sigma_B = \rho_B$.

Corollary 3.3.7. *Let $\rho_{XB} \in \mathcal{P}(\mathcal{H}_X \otimes \mathcal{H}_B)$ be a density operator which is classical on $\mathcal{H}_X$. Then, for any $\varepsilon \geq 0$,*

$$\frac{1}{n} H_{\min}^{\varepsilon}(\rho_{XB}^{\otimes n} | \rho_B^{\otimes n}) \geq H(\rho_{XB}) - H(\rho_B) - \delta \ ,$$

where $\delta := \big(2 H_{\max}(\rho_X) + 3\big) \sqrt{\frac{\log(1/\varepsilon)}{n} + 1}$.

Chapter 5

Privacy Amplification

A fundamental problem in cryptography is to distill a secret key from only partially secret data, on which an adversary might have information encoded into the state of a quantum system. In this chapter, we propose a general solution to this problem, which is called *privacy amplification*: We show that the key computed as the output of a hash function (chosen at random from a two-universal[1] family of functions) is secure under the sole condition that its length is smaller than the adversary's uncertainty on the input, measured in terms of (smooth) min-entropy.

We start with the derivation of various technical results (Sections 5.1–5.4). These are used for the proof of the main statement, which is first formulated in terms of min-entropy (Section 5.5) and then generalized to *smooth* min-entropy (Section 5.6).

[1] See Section 5.4 for a definition.

5.1 Bounding the norm of hermitian operators

In this section, we derive an upper bound on the trace norm for hermitian operators (Lemma 5.1.3). The bound only involves matrix multiplications, which makes it easy to evaluate.

Lemma 5.1.1. *Let S and T be hermitian operators on $\mathcal{H}$. Then*

$$\mathrm{tr}(ST) \leq \sqrt{\mathrm{tr}(S^2)\mathrm{tr}(T^2)} \ .$$

Lemma 5.1.2. *Let S be a hermitian operator on $\mathcal{H}$ and let σ be a nonnegative operator on $\mathcal{H}$. Then*

$$\mathrm{tr}|\sqrt{\sigma}S\sqrt{\sigma}| \leq \sqrt{\mathrm{tr}(S^2)\mathrm{tr}(\sigma^2)} \ .$$

Lemma 5.1.3. *Let S be a hermitian operator on $\mathcal{H}$ and let σ be a nonnegative operator on $\mathcal{H}$. Then*

$$\|S\|_1 \leq \sqrt{\mathrm{tr}(\sigma)\mathrm{tr}(S\sigma^{-1/2}S\sigma^{-1/2})} \ .$$

5.2 Distance from uniform

According to the discussion on universal security in Section 2.2.2, the security of a key is defined with respect to its L_1-distance from a perfect key which is uniformly distributed and independent of the adversary's state (see (2.6)). This motivates the following definition.

Definition 5.2.1. Let $\rho_{AB} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Then the *$L_1$-distance from uniform of ρ_{AB} given B* is

$$d(\rho_{AB}|B) := \left\|\rho_{AB} - \rho_U \otimes \rho_B\right\|_1 ,$$

where $\rho_U := \frac{1}{\dim(\mathcal{H}_A)}\mathrm{id}_A$ is the fully mixed state on $\mathcal{H}_A$.

For an operator ρ_{XZ} defined by a classical probability distribution P_{XZ}, $d(\rho_{XZ}|Z)$ is the expectation (over z chosen according to P_Z) of the L_1-distance between the conditional distribution $P_{X|Z=z}$ and the uniform distribution. This property is generalized by the following lemma.

Lemma 5.2.2. *Let ρ_{ABZ} be classical with respect to an orthonormal basis $\{|z\rangle\}_{z\in\mathcal{Z}}$ of $\mathcal{H}_Z$ and let ρ^z_{AB}, for $z \in \mathcal{Z}$, be the corresponding (non-normalized) conditional operators. Then*

$$d(\rho_{ABZ}|BZ) = \sum_{z\in\mathcal{Z}} d(\rho^z_{AB}|B) \ .$$

To derive our result on the security of privacy amplification, it is convenient to consider an alternative measure for the distance from uniform. Let $\rho_{AB} \in \mathcal{P}(\mathcal{H}_{AB})$ and $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$. The *(conditional) L_2-distance from uniform of ρ_{AB} relative to σ_B* is defined by

$$d_2(\rho_{AB}|\sigma_B) := \mathrm{tr}\Big(\big((\rho_{AB} - \rho_U \otimes \rho_B)(\mathrm{id}_A \otimes \sigma_B^{-1/2})\big)^2\Big) \ ,$$

where ρ_U is the fully mixed state on $\mathcal{H}_A$. Note that $d_2(\rho_{AB}|\sigma_B)$ can equivalently be written as

$$d_2(\rho_{AB}|\sigma_B) = \mathrm{tr}\Big(\big((\mathrm{id}_A \otimes \sigma_B^{-1/4})(\rho_{AB} - \rho_U \otimes \rho_B)(\mathrm{id}_A \otimes \sigma_B^{-1/4})\big)^2\Big) \ , \quad (5.3)$$

which proves that $d_2(\rho_{AB}|\sigma_B)$ cannot be negative.

The L_2-distance from uniform can be used to bound the L_1-distance from uniform.

Lemma 5.2.3. *Let $\rho_{AB} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Then, for any $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$,*

$$d(\rho_{AB}|B) \leq \sqrt{\dim(\mathcal{H}_A)\mathrm{tr}(\sigma_B)d_2(\rho_{AB}|\sigma_B)} \ .$$

The following lemma provides an expression for the L_2-distance from uniform for the case where the first subsystem is classical.

Lemma 5.2.4. *Let $\rho_{XB} \in \mathcal{P}(\mathcal{H}_X \otimes \mathcal{H}_B)$ be classical with respect to an orthonormal basis $\{|x\rangle\}_{x\in\mathcal{X}}$ of $\mathcal{H}_X$, let ρ_B^x, for $x \in \mathcal{X}$, be the corresponding (non-normalized) conditional operators, and let $\sigma \in \mathcal{P}(\mathcal{H}_B)$. Then*

$$d_2(\rho_{XB}|\sigma_B) = \sum \mathrm{tr}\big((\sigma_B^{-1/4}\rho_B^x\sigma_B^{-1/4})^2\big) - \frac{1}{|\mathcal{X}|}\mathrm{tr}\big((\sigma_B^{-1/4}\rho_B\sigma_B^{-1/4})^2\big) \ .$$

5.3 Collision entropy

Definition 5.3.1 below can be seen as a generalization of the well-known classical (conditional) collision entropy to quantum states.

Definition 5.3.1. Let $\rho_{AB} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$. Then the *collision entropy of ρ_{AB} relative to σ_B* is

$$H_2(\rho_{AB}|\sigma_B) := -\log \frac{1}{\mathrm{tr}(\rho_{AB})} \mathrm{tr}\Big(\big(\rho_{AB}(\mathrm{id}_A \otimes \sigma_B^{-1/2})\big)^2 \Big) \ .$$

Remark 5.3.2. It follows immediately from Lemma B.5.3 that

$$H_{\min}(\rho_{AB}|\sigma_B) \leq H_2(\rho_{AB}|\sigma_B) \ .$$

Remark 5.3.3. If $\rho_{XB} \in \mathcal{P}(\mathcal{H}_X \otimes \mathcal{H}_B)$ is classical with respect to an orthonormal basis $\{|x\rangle\}_{x \in \mathcal{X}}$ of $\mathcal{H}_X$ such that the (non-normalized) conditional operators ρ_B^x on $\mathcal{H}_B$, for $x \in \mathcal{X}$, are orthogonal then

$$2^{-H_2(\rho_{XB}|\sigma_B)} = \frac{1}{\mathrm{tr}(\rho_{XB})} \sum_x \mathrm{tr}\big((\sigma_B^{-1/4} \rho_B^x \sigma_B^{-1/4})^2\big) \ .$$

5.4 Two-universal hashing

Definition 5.4.1. Let $\mathcal{F}$ be a family of functions from $\mathcal{X}$ to $\mathcal{Z}$ and let P_F be a probability distribution on $\mathcal{F}$. The pair $(\mathcal{F}, P_F)$ is called *two-universal* if $\Pr_f[f(x) = f(x')] \leq \frac{1}{|\mathcal{Z}|}$, for any distinct $x, x' \in \mathcal{X}$ and f chosen at random from $\mathcal{F}$ according to the distribution P_F.

In accordance with the standard literature on two-universal hashing, we will, for simplicity, assume that P_F is the uniform distribution on $\mathcal{F}$. In particular, the family $\mathcal{F}$ is said to be *two-universal* if $(\mathcal{F}, P_F)$, for P_F uniform, is two-universal. It is, however, easy to see that all statements proven below also hold with respect to the general definition where P_F is arbitrary.

We will use the following lemma on the existence of two-universal function families.

Lemma 5.4.2. *Let $0 \leq \ell \leq n$. Then there exists a two-universal family of hash functions from $\{0,1\}^n$ to $\{0,1\}^\ell$.*

Consider an operator ρ_{XB} which is classical with respect to an orthonormal basis $\{|x\rangle\}_{x\in\mathcal{X}}$ of $\mathcal{H}_X$ and assume that f is a function from $\mathcal{X}$ to $\mathcal{Z}$. The density operator describing the classical function output together with the quantum system $\mathcal{H}_B$ is then given by

$$\rho_{f(X)B} := \sum_{z\in\mathcal{Z}} |z\rangle\langle z| \otimes \rho_B^z \qquad \text{for } \rho_B^z := \sum_{x\in f^{-1}(z)} \rho_B^x \ , \tag{5.4}$$

where $\{|z\rangle\}_{z\in\mathcal{Z}}$ is an orthonormal basis of $\mathcal{H}_Z$.

Assume now that the function f is randomly chosen from a family of functions $\mathcal{F}$ according to a probability distribution P_F. The function output $f(x)$, the state of the quantum system, and the choice of the function f is then described by the operator

$$\rho_{F(X)BF} := \sum_{f\in\mathcal{F}} P_F(f)\rho_{f(X)B} \otimes |f\rangle\langle f| \tag{5.5}$$

on $\mathcal{H}_Z \otimes \mathcal{H}_B \otimes \mathcal{H}_F$, where $\mathcal{H}_F$ is a Hilbert space with orthonormal basis $\{|f\rangle\}_{f\in\mathcal{F}}$.

The following lemma provides an upper bound on the expected L_2-distance from uniform of a key computed by two-universal hashing.

Lemma 5.4.3. *Let $\rho_{XB} \in \mathcal{P}(\mathcal{H}_X \otimes \mathcal{H}_B)$ be classical on $\mathcal{H}_X$, let $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$, and let $\mathcal{F}$ be a two-universal family of hash functions from $\mathcal{X}$ to $\mathcal{Z}$. Then*

$$\mathop{\mathbb{E}}_{f}\Big[d_2(\rho_{f(X)B}|\sigma_B)\Big] \leq \mathrm{tr}(\rho_{XB})2^{-H_2(\rho_{XB}|\sigma_B)},$$

for $\rho_{f(X)B} \in \mathcal{P}(\mathcal{H}_Z \otimes \mathcal{H}_B)$ defined by (5.4) and f chosen uniformly from $\mathcal{F}$.

5.5 Security of privacy amplification

We are now ready to state our main result on privacy amplification in the context of quantum adversaries. Let X be a string and assume that an adversary controls a quantum system $\mathcal{H}_B$ whose state is correlated with X. Theorem 5.5.1 provides a bound on the security of a key $f(X)$ computed from X by two-universal hashing. The bound only depends on the uncertainty of the adversary on X, measured in terms of collision entropy, min-entropy (cf. Corollary 5.5.2), or smooth min-entropy (Corollary 5.6.1), where the latter is (nearly) optimal (see Section 5.6).

Theorem 5.5.1. *Let $\rho_{XB} \in \mathcal{P}(\mathcal{H}_X \otimes \mathcal{H}_B)$ be classical with respect to an orthonormal basis $\{|x\rangle\}_{x\in\mathcal{X}}$ of $\mathcal{H}_X$, let $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$, and let $\mathcal{F}$ be a two-universal family of hash function from $\mathcal{X}$ to $\{0,1\}^\ell$. Then*

$$d(\rho_{F(X)BF}|BF) \leq \sqrt{\mathrm{tr}(\rho_{XB}) \cdot \mathrm{tr}(\sigma_B)} \cdot 2^{-\frac{1}{2}(H_2(\rho_{XB}|\sigma_B)-\ell)} ,$$

for $\rho_{F(X)BF} \in \mathcal{P}(\mathcal{H}_Z \otimes \mathcal{H}_B \otimes \mathcal{H}_F)$ defined by (5.5).

Corollary 5.5.2. *Let $\rho_{XB} \in \mathcal{P}(\mathcal{H}_X \otimes \mathcal{H}_B)$ be classical with respect to an orthonormal basis $\{|x\rangle\}_{x\in\mathcal{X}}$ of $\mathcal{H}_X$ and let $\mathcal{F}$ be a two-universal family of hash functions from $\mathcal{X}$ to $\{0,1\}^\ell$. Then*

$$d(\rho_{F(X)BF}|BF) \leq \sqrt{\mathrm{tr}(\rho_{XB})} \cdot 2^{-\frac{1}{2}(H_{\min}(\rho_{XB}|B)-\ell)} ,$$

for $\rho_{F(X)BF} \in \mathcal{P}(\mathcal{H}_Z \otimes \mathcal{H}_B \otimes \mathcal{H}_F)$ defined by (5.5).

5.6 Characterization using smooth min-entropy

The characterization of privacy amplification in terms of the collision entropy or min-entropy is not optimal.[2] Because of Remark 5.3.2, the same problem arises if we replace the collision entropy by the min-entropy (as in Corollary 5.5.2). However, as we shall see, the statement of Theorem 5.5.1 still holds if the uncertainty is measured in terms of *smooth* min-entropy. That is, the key generated from X by two-universal hashing is secure if its length is slightly smaller than roughly $H_{\min}^{\varepsilon}(\rho_{XB}|B)$, where ρ_{XB} is the joint state of the initial string X and the adversary's knowledge. This is essentially optimal, i.e., $H_{\min}^{\varepsilon}(\rho_{XB}|B)$ is also an upper bound on the maximum number of key bits that can be generated from X.

[2] This also holds for the classical result, as observed in [BBCM95]. In fact, depending on the probability distribution P_X of the initial string X, it might be possible to extract a key whose length exceeds the collision entropy of P_X.

Corollary 5.6.1. *Let $\rho_{XB} \in \mathcal{P}(\mathcal{H}_X \otimes \mathcal{H}_B)$ be a density operator which is classical with respect to an orthonormal basis $\{|x\rangle\}_{x\in\mathcal{X}}$ of $\mathcal{H}_X$, let $\mathcal{F}$ be a two-universal family of hash functions from $\mathcal{X}$ to $\{0,1\}^\ell$, and let $\varepsilon \geq 0$. Then*

$$d(\rho_{F(X)BF}|BF) \leq 2\varepsilon + 2^{-\frac{1}{2}(H_{\min}^{\varepsilon}(\rho_{XB}|B)-\ell)} ,$$

for $\rho_{F(X)BF} \in \mathcal{P}(\mathcal{H}_Z \otimes \mathcal{H}_B \otimes \mathcal{H}_F)$ defined by (5.5).

Contents