Leftover Hashing Against Quantum Side Information

Marco Tomamichel, Christian Schaffner, Adam Smith, and Renato Renner

***Abstract*—The Leftover Hash Lemma states that the output of a two-universal hash function applied to an input with sufficiently high entropy is almost uniformly random. In its standard formulation, the lemma refers to a notion of randomness that is (usually implicitly) defined with respect to classical side information. Here, a strictly more general version of the Leftover Hash Lemma that is valid even if side information is represented by the state of a quantum system is shown. Our result applies to almost two-universal families of hash functions. The generalized Leftover Hash Lemma has applications in cryptography, e.g., for key agreement in the presence of an adversary who is not restricted to classical information processing.**

***Index Terms*—Leftover hash lemma, quantum information, smooth entropies.**

I. Introduction

CONSIDER a random variable X that is partially known to an agent, that is, the agent possesses side information E correlated to X. One may ask whether it is possible to extract from X a part Z that is completely unknown to the agent, i.e., uniform conditioned on E. If yes, what is the maximum size of Z? And how is Z computed?

The *Leftover Hash Lemma* answers these questions. It states that extraction of uniform randomness Z is possible whenever the agent's uncertainty about X is sufficiently large. More precisely, the number ℓ of extractable bits is approximately equal to the *min-entropy of* X *conditioned on* E, denoted $H_{\min}(\mathrm{X}|\mathrm{E})$ (see Section I-B for a definition and properties). Furthermore, Z can be computed as the output of a function f selected at random from a suitably chosen family of functions $\mathcal{F}$, called *two-universal family of hash functions* (see Section I-A for a definition). Remarkably, the family can be chosen without knowing the actual probability distribution of X and only depends on the alphabet $\mathcal{X}$ of X and the number of bits to be extracted, ℓ.

More specifically, the Leftover Hash Lemma states that, on average over the choices of f from $\mathcal{F}$, the distribution of the output $Z := f(X)$ is at most Δ-far from uniform conditioned on E,[1] where

$$\Delta = \frac{1}{2}\sqrt{2^{\ell - H_{\min}(\mathrm{X}|\mathrm{E})}}. \tag{1}$$

The lemma immediately implies that for a *fixed* joint distribution of X and E, there is a *fixed* function f that extracts almost uniform randomness. In fact, for any $\Delta > 0$, there exists a function f that produces[2]

$$\ell = \left\lfloor H_{\min}(\mathrm{X}|\mathrm{E}) - 2\,\log\frac{1}{\Delta} + 2 \right\rfloor \tag{2}$$

bits that are Δ-close to a bit string that is both uniform and independent of E.

The Leftover Hash Lemma plays an important role in a variety of applications in computer science and cryptography (see, e.g., [1] for an overview). A prominent example is *privacy amplification*, i.e., the task of transforming a weakly secret key (about which an adversary may have partial knowledge E), into a highly secret key (that is uniform and independent of the adversary's information E). It was in this context that the use of two-universal hashing for randomness distillation was first proposed [2]. Originally, the analysis was, however, restricted to situations where X is uniform and E is bounded in size. Later, versions of the Leftover Hash Lemma similar to (1) have been proved independently in [3] and [4]. The term *leftover hashing* was coined in [5], where its use for recycling the randomness in randomized algorithms and for the construction of pseudo-random number generators is discussed (see also [3], [6]).

A. Almost Two-Universal Hashing

The notion of two-universal hashing was introduced by Carter and Wegman [7]. A family $\mathcal{F}$ of functions from $\mathcal{X}$ to $\mathcal{Z}$ is said to be *two-universal* if, for any pair of distinct inputs x and x', and for f chosen at random from $\mathcal{F}$, the probability of a *collision* $f(x) = f(x')$ is not larger than $\delta := 1/|\mathcal{Z}|$. Note that this value for the collision probability corresponds to the one obtained by choosing $\mathcal{F}$ as the family of *all* functions with domain $\mathcal{X}$ and range $\mathcal{Z}$.

Later, the concept of two-universal hashing was generalized to arbitrary collision probabilities δ [8]. Namely, a family of functions $\mathcal{F}$ from $\mathcal{X}$ to $\mathcal{Z}$ is called δ*-almost two-universal* if

$$\Pr_{f\in\mathcal{F}}[f(x) = f(x')] \le \delta \tag{3}$$

Manuscript received July 19, 2010; revised February 03, 2011; accepted March 16, 2011. Date of current version July 29, 2011. M. Tomamichel and R. Renner were supported by the Swiss National Science Foundation under Grant 200021-119868. C. Schaffner was supported by a NWO VICI project. A. Smith was supported by the U.S. National Science Foundation under Grants 0747294 and 0729171.

M. Tomamichel and R. Renner are with the Institute for Theoretical Physics, ETH Zurich, 8093 Zurich, Switzerland (e-mail: marcoto@phys.ethz.ch; renner@phys.ethz.ch).

C. Schaffner is with the Centrum Wiskunde & Informatica (CWI), 1090 GB Amsterdam, The Netherlands.

A. Smith is with the Pennsylvania State University, University Park, PA 16802 USA.

Communicated by P. Hayden, Associate Editor for Quantum Information Theory.

Digital Object Identifier 10.1109/TIT.2011.2158473

[1] The distance from uniform Δ measures the statistical distance of the probability distribution of X given E to a uniform distribution. See Section III for a formal definition.

[2] We use $\log$ to denote the binary logarithm.

for any $x \neq x'$. A two-universal family as above simply corresponds to the special case $\delta = 1/|\mathcal{Z}|$.

The classical Leftover Hash Lemma (1) can be generalized to δ-almost two-universal hash functions [1]. More precisely, when extracting an ℓ-bit string from data X, its distance from uniform conditioned on E is bounded by

$$\Delta = \frac{1}{2}\sqrt{(2^{\ell}\delta - 1) + 2^{\ell - H_{\min}(\mathrm{X}|\mathrm{E})}}. \tag{4}$$

The relaxation of the collision probability condition in the definition of δ-almost two-universal families of hash functions (3) allows for smaller families $\mathcal{F}$, thus reducing the amount of randomness needed to choose a function $f \in \mathcal{F}$. This, in turn, allows for the construction of randomness extractors that require shorter random seeds (cf. Section IV).

B. *Quantum Side Information*

A majority of the original work on universal hashing is based entirely on probability theory and side information is, therefore (often implicitly), assumed to be represented by a *classical* system E (modeled as a random variable).[3] In fact, since hashing is an entirely "classical" process (a simple mapping from a random variable X to another random variable Z), one may expect that the physical nature of the side information is irrelevant and that a purely classical treatment is sufficient. This is, however, not necessarily the case. For example, the output of certain extractor functions may be partially known if side information about their input is stored in a *quantum* device of a certain size, while the same output is almost uniform conditioned on any side information stored in a *classical* system of the same size (see [9] for a concrete example and [10] for a more general discussion).[4]

Here, we follow a line of research started in [11]–[13] and study randomness extraction in the presence of quantum side information E. (This, of course, includes situations where E is partially or fully classical.) More specifically, our goal is to establish a generalized version of (4) which holds if the system E is quantum-mechanical. In order to state the result, we first need to define the *min-entropy* as well as of the notion of *uniformity* in a quantum setting.

The definition of *uniformity* in the context of quantum side information E is rather straightforward. Let Z be a classical random variable which takes any value $z \in \mathcal{Z}$ with probability p_z and let E be a quantum system whose state conditioned on $Z = z$ is given by a density operator $\rho_{\mathrm{E}}^{[z]}$ on $\mathcal{H}_{\mathrm{E}}$. This situation is compactly described by the *classical-quantum* (CQ) state

$$\rho_{\mathrm{ZE}} := \sum_{z \in \mathcal{Z}} p_z \, |z\rangle\langle z|_{\mathrm{Z}} \otimes \rho_{\mathrm{E}}^{[z]} \tag{5}$$

defined on the product space $\mathcal{H}_{\mathrm{Z}} \otimes \mathcal{H}_{\mathrm{E}}$, where $\mathcal{H}_{\mathrm{Z}}$ is a Hilbert space with orthonormal basis $\{|z\rangle_{\mathrm{Z}}\}_{z \in \mathcal{Z}}$. We say that *$Z$ is uniform conditioned on E* if ρ_{ZE} has product form $\omega_{\mathrm{Z}} \otimes \rho_{\mathrm{E}}$, where $\omega_{\mathrm{Z}} := \mathbb{1}_{\mathrm{Z}}/|\mathcal{Z}|$ is the maximally mixed state on $\mathcal{H}_{\mathrm{Z}}$. More generally, we say that Z is *Δ-close to uniform conditioned on E* if there exists a state σ_{E} on E for which the trace distance between ρ_{ZE} and $\omega_{\mathrm{Z}} \otimes \sigma_{\mathrm{E}}$ is at most Δ (see Section III for a formal definition). The trace distance is a natural choice of metric because it corresponds to the *distinguishing advantage*.[5] Furthermore, in the purely classical case, the trace distance reduces to the statistical distance.

Next, we generalize the notion of min-entropy to situations involving quantum side information. Before we do this, note that the classical min-entropy[6] has an operational interpretation as the average guessing probability of X given E, namely

$$H_{\min}(\mathrm{X}|\mathrm{E}) = -\log p_{\mathrm{guess}}(\mathrm{X}|\mathrm{E}). \tag{6}$$

Here, $p_{\mathrm{guess}}(X|E)$ denotes the probability of correctly guessing the value of X using the optimal strategy with access to E. The optimal strategy in the classical case is to guess, for each value of e of E, the X with the highest conditional probability $P_{X|E=e}$. The guessing probability is thus

$$p_{\mathrm{guess}}(\mathrm{X}|\mathrm{E}) = \sum_{e} P_E(e) \, \max_{x} P_{X|E=e}(x).$$

A generalization of the min-entropy to situations where E may be a quantum system was first proposed in [12] (see Section II for a formal definition). As shown in [16], the operational interpretation (6) naturally extends to this more general case. In other words, the min-entropy, $H_{\min}(\mathrm{X}|\mathrm{E})$, is a measure for the probability of guessing X using an optimal strategy with access to the quantum system E.

Next, we argue that the min-entropy, $H_{\min}(\mathrm{X}|\mathrm{E})$, accurately characterizes the total amount of randomness contained in X, i.e., the number of uniformly random bits that can be extracted using an optimal extraction strategy. For this purpose, let ρ_{XE} be fixed and assume that f is a function that maps X to a string $Z = f(X)$ of length ℓ that is uniform conditioned on the side information E. Then, obviously, the probability of guessing Z correctly given E is equal to $2^{-\ell}$ and, by virtue of (6), we find that

$$H_{\min}(\mathrm{Z}|\mathrm{E}) = \ell. \tag{7}$$

Furthermore, the probability of guessing $Z = f(X)$ correctly cannot be smaller than the probability of guessing X correctly. This fact can again be expressed in terms of min-entropies

$$H_{\min}(\mathrm{Z}|\mathrm{E}) \leq H_{\min}(\mathrm{X}|\mathrm{E}) \tag{8}$$

[3] If the side information E is classical, the Leftover Hash Lemma can be formulated without the need to introduce E explicitly (see, e.g., [3]). Instead, one may simply interpret all probability distributions as being conditioned on a fixed value of the side information.

[4] Note that there is no sensible notion of a conditional probability distribution where the conditioning is on the state of a *quantum* (as opposed to a *classical*) system. An implicit treatment of side information E, where one considers all probability distributions to be conditioned on a specific value of E, as explained in the previous footnote, is, therefore, not possible in the general case.

[5] Let p_{succ} be the maximum probability that a distinguisher, presented with a random choice of either the state ρ or the state σ, can correctly guess which of the two he has seen. The *distinguishing advantage* is then defined as the advantage compared to a random guess, which is given by $p_{\mathrm{succ}} - \frac{1}{2} = \frac{1}{4}\|\rho - \sigma\|_1$ (see, e.g., [14])

[6] There are several conventions for defining conditional min-entropy, even for classical random variables. The notion we use is sometimes called *average* conditional min-entropy [15].

i.e., the min-entropy can only decrease under the action of a function. Combining (7) and (8) immediately yields

$$\ell \leq H_{\min}(\mathrm{X}|\mathrm{E}). \tag{9}$$

We conclude that the number ℓ of uniform bits (relative to E) that can be extracted from data X is upper bounded by the min-entropy of X conditioned on E. This result may be seen as a converse of (2).

So far, the claim (9) is restricted to the extraction of *perfectly uniform* randomness. In order to extend this concept to the more general case of approximately uniform randomness, we need to introduce the notion of *smooth* min-entropy. Roughly speaking, for any $\varepsilon \geq 0$, the *ε-smooth min-entropy of X given E*, denoted $H_{\min}^{\varepsilon}(\mathrm{X}|\mathrm{E})$, is defined as the maximum value of $H_{\min}(\mathrm{X}|\mathrm{E})$ evaluated for all density operators $\tilde{\rho}$ that are ε-close to ρ in terms of the purified distance (see Section II for a formal definition).

The above argument leading to (9) can be generalized in a straightforward manner to smooth min-entropy, and results in the bound

$$\ell \leq H_{\min}^{2\sqrt{\Delta}}(\mathrm{X}|\mathrm{E})$$

for the maximum number ℓ of extractable bits that are Δ-close to uniform conditioned on E. Crucially, our extended version of the Leftover Hash Lemma implies that this bound can be reached, up to additive terms of order $\log(1/\Delta)$ (see Theorem 6 and Theorem 7). We, thus, conclude that the min-entropy of X conditioned on E, in particular its "smoothed" version, is an accurate measure for the amount of uniform randomness (conditioned on E) that can be extracted from X.

C. Main Result

Our main result is a generalization of the Leftover Hash Lemma for δ-almost two-universal families of hash functions which is valid in the presence of quantum side information. While the statement is new for general δ-almost two-universal hash functions, the special case of two-universal hashing was proved previously by one of us [12].[7]

Lemma 1 (Generalized Leftover Hash Lemma): Let X be a random variable, let E be a quantum system, and let $\mathcal{F}$ be a δ-almost two-universal family of hash functions from $\mathcal{X}$ to $\{0,1\}^{\ell}$. Then, on average over the choices of f from $\mathcal{F}$, $Z := f(X)$ is Δ-close to uniform conditioned on E, where

$$\Delta = \sqrt[4]{(2^{\ell}\delta - 1)^2 + 9 \cdot 2^{\ell - H_{\min}(\mathrm{X}|\mathrm{E})}}. \tag{10}$$

Furthermore, if $\delta \leq 2^{-\ell}$, i.e., if $\mathcal{F}$ is two-universal, then

$$\Delta = \frac{1}{2}\sqrt{2^{\ell - H_{\min}(\mathrm{X}|\mathrm{E})}}. \tag{11}$$

Note that inserting $\delta = 2^{-\ell}$ into the first expression for Δ yields a looser bound than (11). The latter, therefore, requires a separate proof. In the technical part below, the two claims are formulated more generally for the *smooth* min-entropy (Theorem 6 and Theorem 7).

D. Applications and Related Work

Quantum versions of the Leftover Hash Lemma [12] for two-universal families of hash functions have been used in the context of privacy amplification against a quantum adversary [13], [10]. This application has gained prominence with the rise of quantum cryptography and quantum key distribution in particular. There, the side information E is gathered during a key agreement process between two parties by an eavesdropper who is not necessarily limited to classical information processing. The quantum generalization of the Leftover Hash Lemma is then used to bound the amount of secret key that can be distilled by the two parties.

The restriction to two-universal families of hash functions leads to the need for a random seed of length $\Theta(n)$, where n is the length in bits of the original partially secret string. This seed is used to choose f from a two-universal family $\mathcal{F}$. The main result of this paper, Lemma 1, and a suitable construction of a δ-almost two-universal family of hash functions (see Section IV, Theorem 10) allow for a shorter seed of length proportional to ℓ, $\log \frac{n}{\ell}$ and $\log \frac{1}{\Delta}$. The length of secret key that can be extracted with this method is only reduced by an additive term proportional to $\log \frac{1}{\Delta}$ compared to the extractor using two-universal hashing. (See (26) and (27).) Furthermore, the generalized Leftover Hash Lemma allows for an extension of existing cryptographic security proofs to δ-almost two-universal families of hash functions and may lead to a speed-up in practical implementations.[8]

Recently, the problem of randomness extraction with quantum side information has generated renewed interest. For example, XORing a classical source about which an adversary holds quantum information with a δ-biased mask (as in [20]) results in a uniformly distributed string even when conditioned on quantum side information [21][9].

However, to achieve even shorter seed lengths, more advanced techniques such as Trevisan's extractor [23] have been studied in [24]–[26]. In [25], it is shown that a seed of length $O(\mathrm{polylog}\, n)$ is sufficient to generate a key of length $\ell \approx H_{\min}(\mathrm{X}) - \log \dim\mathcal{H}_{\mathrm{E}}$, where $\dim\mathcal{H}_{\mathrm{E}}$ is a measure of the size of the adversary's quantum memory. In [26], the result was extended to the formalism of conditional min-entropies. They attain a key length of $\ell \approx H_{\min}^{\varepsilon}(\mathrm{X}|\mathrm{E})$ minus erms logarithmic in the output size. This construction can be concatenated with our extractor using almost two-universal hashing (Theorem 10) to extract the remaining randomness. The combined key length, $\ell \approx H_{\min}^{\varepsilon}(\mathrm{X}|\mathrm{E})$, is then almost optimal, as we argue in (9) and the seed still grows as $\mathrm{polylog}\, n$. (See the discussion in [26] for more details.)

Moreover, our result should be used instead of the classical Leftover Hash Lemma whenever randomness is extracted in a context governed by the laws of quantum physics. For example, consider a device that needs a seed that is random conditioned

[7] We reprove the result here for a slightly adapted definition of the *smooth* min-entropy that has now become standard. (See, e.g., [17].)

[8] See, e.g., [18] and [19], where a practical implementation of privacy amplification is discussed in Section V.

[9] See also [22] for a generalization of this work to the fully quantum setting.

on its internal state. In this case the use of the *classical* Leftover Hash Lemma instead of its quantum version, Lemma 1, corresponds to the implicit and potentially unjustified assumption that the device is entirely governed by classical mechanics.

E. Organization of the Paper

In Section II, we discuss various aspects of the smooth entropy framework, which will be needed for our proof. We then give the proof of the General Leftover Hash Lemma (Lemma 1) in Section III. More precisely, we provide statements of the Leftover Hash Lemma for two-universal and δ-almost two-universal hashing in terms of the smooth min-entropy (Theorems 6 and 7). Finally, in Section IV, we combine known constructions of δ-almost two-universal hash functions and discuss their use for randomness extraction with shorter random seeds. Appendix B may be of independent interest because it establishes a relation between the smooth min- and max-entropies (as defined above and used in [16], [27], [17]) and certain related entropic quantities used in earlier work (e.g., in [12]) and to prove the main result of this paper.

II. Smooth Entropies

Let $\mathcal{H}$ be a finite-dimensional Hilbert space. We use $\mathcal{L}(\mathcal{H})$, $\mathcal{L}^{\dagger}(\mathcal{H})$ and $\mathcal{P}(\mathcal{H})$ to denote the set of linear, Hermitian and positive semi-definite operators on $\mathcal{H}$, respectively. We define the set of normalized quantum states by $\mathcal{S}_{=}(\mathcal{H}) := \{\rho \in \mathcal{P}(\mathcal{H}) : \operatorname{tr}\rho = 1\}$ and the set of sub-normalized states by $\mathcal{S}_{\leq}(\mathcal{H}) := \{\rho \in \mathcal{P}(\mathcal{H}) : 0 < \operatorname{tr}\rho \leq 1\}$. Given a pure state $|\phi\rangle \in \mathcal{H}$, we use $\phi = |\phi\rangle\langle\phi|$ to denote the corresponding projector in $\mathcal{P}(\mathcal{H})$. The inverse of a Hermitian operator is taken on its support (generalized inverse). Given a bipartite Hilbert space $\mathcal{H}_{\mathrm{AB}} := \mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}}$ and a state $\rho_{\mathrm{AB}} \in \mathcal{S}_{\leq}(\mathcal{H}_{\mathrm{AB}})$, we denote by ρ_{A} and ρ_{B} its marginals $\rho_{\mathrm{A}} = \operatorname{tr}_{\mathrm{B}}\rho_{\mathrm{AB}}$ and $\rho_{\mathrm{B}} = \operatorname{tr}_{\mathrm{A}}\rho_{\mathrm{AB}}$.

The *trace distance* between states ρ and τ is given by $\frac{1}{2}\|\rho - \tau\|_1 = \frac{1}{2}\operatorname{tr}|\rho - \tau|$. We also employ the *purified distance*, P, as a metric on $\mathcal{S}_{\leq}(\mathcal{H})$ [17]. It is an upper bound on the trace distance and defined in terms of the *generalized fidelity*, $\bar{F}$, as

$$P(\rho,\tau) := \sqrt{1 - \bar{F}(\rho,\tau)^2}, \quad \text{where}$$
$$\bar{F}(\rho,\tau) := \operatorname{tr}|\sqrt{\rho}\sqrt{\tau}| + \sqrt{(1-\operatorname{tr}\rho)(1-\operatorname{tr}\tau)}.$$

We will need that the purified distance is a monotone under trace nonincreasing completely positive maps (CPMs). Let $\mathcal{E}$ be a trace nonincreasing CPM, then [17]

$$P(\rho,\tau) \geq P(\mathcal{E}(\rho), \mathcal{E}(\tau)). \tag{12}$$

Note that the projections $\rho \mapsto \Pi\rho\Pi$ for any projector Π is a trace nonincreasing CPM. We define the ε*-ball* of states close to $\rho \in \mathcal{S}_{\leq}(\mathcal{H})$ as

$$\mathcal{B}^{\varepsilon}(\rho) := \{\tilde{\rho} \in \mathcal{S}_{\leq}(\mathcal{H}) : P(\rho,\tilde{\rho}) \leq \varepsilon\}.$$

We define the smooth min-entropy [12].

Definition 1: Let $\varepsilon \geq 0$ and $\rho_{\mathrm{AB}} \in \mathcal{S}_{\leq}(\mathcal{H}_{\mathrm{AB}})$. The *min-entropy of A conditioned on B* is given by

$$H_{\min}(\mathrm{A}|\mathrm{B})_{\rho} := \max_{\sigma_{\mathrm{B}} \in \mathcal{S}_{=}(\mathcal{H}_{\mathrm{B}})} \sup\{\lambda \in \mathbb{R} : \rho_{\mathrm{AB}} \leq 2^{-\lambda}\mathbb{1}_{\mathrm{A}} \otimes \sigma_{\mathrm{B}}\}.$$

Furthermore, the *smooth min-entropy of A conditioned on B* is defined as

$$H_{\min}^{\varepsilon}(\mathrm{A}|\mathrm{B})_{\rho} := \max_{\tilde{\rho}_{\mathrm{AB}} \in \mathcal{B}^{\varepsilon}(\rho_{\mathrm{AB}})} H_{\min}(\mathrm{A}|\mathrm{B})_{\tilde{\rho}}.$$

The conditional min-entropy is a measure of the uncertainty about the state of a system A given quantum side information B. In particular, if the system A describes a classical random variable (i.e., if the state is CQ, cf. (5)), the min-entropy can be interpreted as a guessing probability.[10] For general quantum states, the smooth min-entropy satisfies data-processing inequalities. For example, if a CPM is applied to the B system or if a projective measurement is conducted on the A system, the smooth min-entropy of A given B is guaranteed not to decrease.[11]

The following lemma makes clear that the min-entropy smoothing of a state will not destroy its CQ structure.

Lemma 2: Let ρ_{XB} be a CQ state. Then, there exists a CQ state $\tilde{\rho}_{\mathrm{XB}} \in \mathcal{B}^{\varepsilon}(\rho_{\mathrm{XB}})$ that optimizes $H_{\min}^{\varepsilon}(\mathrm{X}|\mathrm{B})_{\rho} = H_{\min}(\mathrm{X}|\mathrm{B})_{\tilde{\rho}}$.

Proof: Let $\bar{\rho}_{\mathrm{XB}}$ be a state in $\mathcal{B}^{\varepsilon}(\rho_{\mathrm{XB}})$. We can create a CQ state $\tilde{\rho}_{\mathrm{XB}}$ by measuring X. This operation will not increase the distance (cf. [17], Lemma 7), i.e., $P(\tilde{\rho}_{\mathrm{XB}}, \rho_{\mathrm{XB}}) \leq \varepsilon$, while it cannot decrease min-entropy (cf. [17], Theorem 19). ■

Finally, we will need a quantum generalization of the collision entropy (Rényi-entropy of order 2).

Definition 2: Let $\rho_{\mathrm{AB}} \in \mathcal{S}_{\leq}(\mathcal{H}_{\mathrm{AB}})$ and $\sigma_{\mathrm{B}} \in \mathcal{P}(\mathcal{H}_{\mathrm{B}})$, then the *collision entropy of A conditioned on B* of a state ρ_{AB} given σ_{B} is $-\log\Gamma_{\mathrm{C}}(\rho_{\mathrm{AB}}|\sigma_{\mathrm{B}})$, where

$$\Gamma_{\mathrm{C}}(\rho_{\mathrm{AB}}|\sigma_{\mathrm{B}}) := \operatorname{tr}\left(\rho_{\mathrm{AB}}\left(\mathbb{1}_{\mathrm{A}} \otimes \sigma_{\mathrm{B}}^{-1/2}\right)\right)^2.$$

We will use the fact that the collision entropy provides an upper bound on the min-entropy. Equation (14) constitutes one of the main technical contributions of this work.

Lemma 3: Let $\rho_{\mathrm{XB}} \in \mathcal{S}_{\leq}(\mathcal{H}_{\mathrm{XB}})$ be a CQ state and $\eta > 0$. Then, there exists a state $\sigma_{\mathrm{B}} \in \mathcal{S}_{=}(\mathcal{H}_{\mathrm{B}})$ such that

$$\Gamma_{\mathrm{C}}(\rho_{\mathrm{XB}}|\sigma_{\mathrm{B}}) \leq 2^{-H_{\min}(\mathrm{X}|\mathrm{B})_{\rho}}. \tag{13}$$

Moreover, there exists a CQ state $\bar{\rho}_{\mathrm{XB}} \in \mathcal{B}^{\eta}(\rho_{\mathrm{XB}})$ such that

$$\Gamma_{\mathrm{C}}(\bar{\rho}_{\mathrm{XB}}|\bar{\rho}_{\mathrm{B}}) \leq \left(\frac{2}{\eta^2} + \frac{1}{\operatorname{tr}\rho_{\mathrm{XB}}}\right) 2^{-H_{\min}(\mathrm{X}|\mathrm{B})_{\rho}}. \tag{14}$$

Proof: To prove the first statement, we observe that, by the definition of the min-entropy, there exists a state $\sigma_{\mathrm{B}} \in \mathcal{S}_{=}(\mathcal{H}_{\mathrm{B}})$ s.t. $\rho_{\mathrm{XB}} \leq 2^{-H_{\min}(\mathrm{X}|\mathrm{B})_{\rho}}\mathbb{1}_{\mathrm{X}} \otimes \sigma_{\mathrm{B}}$ and, thus

$$(\mathbb{1}_{\mathrm{X}} \otimes \sigma_{\mathrm{B}})^{-\frac{1}{2}}\rho_{\mathrm{XB}}(\mathbb{1}_{\mathrm{X}} \otimes \sigma_{\mathrm{B}})^{-\frac{1}{2}} \leq 2^{-H_{\min}(\mathrm{X}|\mathrm{B})_{\rho}}\mathbb{1}_{\mathrm{XB}}.$$

The statement then follows after we multiply both sides by ρ_{XB}, take the trace and use that $\operatorname{tr}\rho_{\mathrm{XB}} \leq 1$.

[10] See discussion in Section I and [16] for details.

[11] See [17] for precise statements and proofs.

To prove the second statement, we will make use of properties of the alternative entropic quantities discussed in Appendix B. In particular, we use Lemma 19 to define $\bar{\rho}_{\mathrm{XB}} \in \mathcal{B}^{\eta}(\rho_{\mathrm{XB}})$ s.t.

$$\hat{H}_{\min}(\mathrm{X}|\mathrm{B})_{\bar{\rho}} \geq H_{\min}(\mathrm{X}|\mathrm{B})_{\rho} - \log\left(\frac{2}{\eta^2} + \frac{1}{\operatorname{tr}\rho_{\mathrm{XB}}}\right).$$

In particular, we can choose $\bar{\rho}_{\mathrm{XB}}$ CQ.[12] We now apply the same argument as in the proof of (13) to show that

$$\Gamma_{\mathrm{C}}(\bar{\rho}_{\mathrm{XB}}|\bar{\rho}_{\mathrm{B}}) \leq 2^{-\hat{H}_{\min}(\mathrm{X}|\mathrm{B})_{\bar{\rho}}} \leq 2^{-H_{\min}(\mathrm{X}|\mathrm{B})_{\rho}+\log\left(\frac{2}{\eta^2}+\frac{1}{\operatorname{tr}\rho_{\mathrm{XB}}}\right)}$$

which concludes the proof. ■

III. Proof of the Leftover Hash Lemma

In this section we give bounds on the distance from uniform of the quantum state after hashing with two-universal and δ-almost two-universal functions (Theorems 6 and 7, respectively). The proof of the Leftover Hashing Lemma (Lemma 1) then follows.

We consider the scenario where the random variable X is picked from a set $\mathcal{X}$ and E is a quantum system whose state may depend on X. The situation is described by a CQ state of the form

$$\rho_{\mathrm{XE}} = \sum_x |x\rangle\langle x|_{\mathrm{X}} \otimes \rho_{\mathrm{E}}^{[x]} \tag{15}$$

where the probability of x occurring is the trace of the sub-normalized state $\rho_{\mathrm{E}}^{[x]}$ and $\rho_{\mathrm{E}} = \sum_x \rho_{\mathrm{E}}^{[x]}$. After applying a function $f : \mathcal{X} \to \{0,1\}^{\ell}$ chosen at random from a family of hash functions $\mathcal{F}$, the resulting CQ state is given by

$$\rho_{\mathrm{FZE}} = \sum_f \sum_z p_f |f\rangle\langle f|_{\mathrm{F}} \otimes |z\rangle\langle z|_{\mathrm{Z}} \otimes \rho_{\mathrm{E}}^{[f,z]} \tag{16}$$

where $z \in \{0,1\}^{\ell}$, $p_f = 1/|\mathcal{F}|$ and

$$\rho_{\mathrm{E}}^{[f,z]} := \sum_{x, f(x)=z} \rho_{\mathrm{E}}^{[x]}. \tag{17}$$

Formally, randomness extraction can be modelled as a trace-preserving CPM, $\mathcal{A}$, from $\mathcal{H}_{\mathrm{FX}} \to \mathcal{H}_{\mathrm{FZ}}$ that maps $\rho_{\mathrm{F}} \otimes \rho_{\mathrm{XE}} \mapsto (\mathcal{A} \otimes \mathcal{I}_{\mathrm{E}})(\rho_{\mathrm{F}} \otimes \rho_{\mathrm{XE}}) = \rho_{\mathrm{FZE}}$.

First, we extend the definition of the distance from uniform to sub-normalized states for technical reasons.[13]

Definition 3: Let $\rho_{\mathrm{AB}} \in \mathcal{S}_{\leq}(\mathcal{H}_{\mathrm{AE}})$, then we define the *distance from uniform of A conditioned on B* as

$$D_u(\mathrm{A}|\mathrm{B})_{\rho} := \min_{\sigma_{\mathrm{B}}} \frac{1}{2}\|\rho_{\mathrm{AB}} - \omega_{\mathrm{A}} \otimes \sigma_{\mathrm{B}}\|_1 \tag{18}$$

where $\omega_{\mathrm{A}} := \mathbb{1}_{\mathrm{A}}/\dim\mathcal{H}_{\mathrm{A}}$ and the minimum is taken over all $\sigma_{\mathrm{B}} \in \mathcal{S}_{\leq}(\mathcal{H}_{\mathrm{B}})$ satisfying $\operatorname{tr}\sigma_{\mathrm{B}} = \operatorname{tr}\rho_{\mathrm{B}}$.

In the following, we will give upper bounds on the quantity $D_u(\mathrm{Z}|\mathrm{FE})$, where $Z = f(X)$ is the random variable after two-universal or δ-almost two-universal hashing, respectively. Note that we consider the distance from uniform conditioned on the choice of hash function F as well as well as the side information E. This describes the situation where the chosen hash function (the value of F) is published after its use, i.e., the strong extractor regime.

The distance from uniform conditioned on E averaged over the choice of f (as used in the introduction) is given by

$$\sum_f p_f D_u(\mathrm{Z}|\mathrm{E})_{\rho^{[f]}}, \quad \text{where } \rho_{\mathrm{ZE}}^{[f]} := \sum_z |z\rangle\langle z|_{\mathrm{Z}} \otimes \rho_{\mathrm{E}}^{[f,z]}.$$

It is upper bounded by $D_u(\mathrm{Z}|\mathrm{FE})$, i.e.,

$$\begin{aligned}\sum_f p_f D_u(\mathrm{Z}|\mathrm{E})_{\rho^{[f]}} &\leq \frac{1}{2}\sum_f p_f \left\|\rho_{\mathrm{ZE}}^{[f]} - \omega_{\mathrm{Z}} \otimes \sigma_{\mathrm{E}}\right\|_1 \\ &= D_u(\mathrm{Z}|\mathrm{FE})_{\rho}\end{aligned} \tag{19}$$

where σ_{E} optimizes (18) for $D_u(\mathrm{Z}|\mathrm{FE})_{\rho}$. Hence, an upper bound on $D_u(\mathrm{Z}|\mathrm{FE})$ trivially implies an upper bound on the average distance as well.

As a first step, we bound the distance from uniform in terms of the collision entropy.

Lemma 4: Let $\rho_{\mathrm{AB}} \in \mathcal{S}_{\leq}(\mathcal{H}_{\mathrm{AB}})$ and $\tau_{\mathrm{B}} \in \mathcal{S}_{\leq}(\mathcal{H}_{\mathrm{B}})$ with $\operatorname{supp}\{\tau_{\mathrm{B}}\} \supseteq \operatorname{supp}\{\rho_{\mathrm{B}}\}$, then

$$D_u(\mathrm{A}|\mathrm{B})_{\rho} \leq \frac{1}{2}\sqrt{d_{\mathrm{A}}\Gamma_{\mathrm{C}}(\rho_{\mathrm{AB}}|\tau_{\mathrm{B}}) - \operatorname{tr}\left(\rho_{\mathrm{B}}\tau_{\mathrm{B}}^{-1/2}\rho_{\mathrm{B}}\tau_{\mathrm{B}}^{-1/2}\right)}.$$

Proof: We use a Hölder-type inequality for linear operators and unitarily invariant norms (see [28] for a proof). Let A, B and C be linear operators and $r, s, t > 0$ such that $\frac{1}{r} + \frac{1}{s} + \frac{1}{t} = 1$, then $\|ABC\|_1 \leq \||A|^r\|_1^{1/r}\||B|^s\|_1^{1/s}\||C|^t\|_1^{1/t}$.

We apply the inequality with parameters $r = t = 4$, $s = 2$, $A = C = \mathbb{1}_{\mathrm{A}} \otimes \tau_{\mathrm{B}}^{1/4}$ and $B = (\mathbb{1}_{\mathrm{A}} \otimes \tau_{\mathrm{B}}^{-1/4})(\rho_{\mathrm{AB}} - \omega_{\mathrm{A}} \otimes \rho_{\mathrm{B}})(\mathbb{1}_{\mathrm{A}} \otimes \tau_{\mathrm{B}}^{-1/4})$. This leads to

$$\begin{aligned}2\,D_u(\mathrm{A}|\mathrm{B})_{\rho} &\leq \|\rho_{\mathrm{AB}} - \omega_{\mathrm{A}} \otimes \rho_{\mathrm{B}}\|_1 \\ &= \|ABC\|_1 \leq \|A^4\|_1^{1/4}\|B^2\|_1^{1/2}\|C^4\|_1^{1/4} \\ &\leq \sqrt{d_A \operatorname{tr}\left((\rho_{\mathrm{AB}} - \omega_{\mathrm{A}} \otimes \rho_{\mathrm{B}})\left(\mathbb{1}_{\mathrm{A}} \otimes \tau_{\mathrm{B}}^{-1/2}\right)\right)^2}.\end{aligned}$$

We simplify the expression on the r.h.s. further using

$$\begin{aligned}&\operatorname{tr}\left((\rho_{\mathrm{AB}} - \omega_{\mathrm{A}} \otimes \rho_{\mathrm{B}})\left(\mathbb{1}_{\mathrm{A}} \otimes \tau_{\mathrm{B}}^{-1/2}\right)\right)^2 \\ &= \operatorname{tr}\left(\rho_{\mathrm{AB}}\left(\mathbb{1}_{\mathrm{A}} \otimes \tau_{\mathrm{B}}^{-1/2}\right)\right)^2 + \operatorname{tr}\left((\omega_{\mathrm{A}} \otimes \rho_{\mathrm{B}})\left(\mathbb{1}_{\mathrm{A}} \otimes \tau_{\mathrm{B}}^{-1/2}\right)\right)^2 \\ &\quad - 2\operatorname{tr}\left(\rho_{\mathrm{AB}}\left(\mathbb{1}_{\mathrm{A}} \otimes \tau_{\mathrm{B}}^{-1/2}\right)(\omega_{\mathrm{A}} \otimes \rho_{\mathrm{B}})\left(\mathbb{1}_{\mathrm{A}} \otimes \tau_{\mathrm{B}}^{-1/2}\right)\right) \\ &= \Gamma_{\mathrm{C}}(\rho_{\mathrm{AB}}|\tau_{\mathrm{B}}) - \frac{1}{d_{\mathrm{A}}}\operatorname{tr}\left(\rho_{\mathrm{B}}\tau_{\mathrm{B}}^{-1/2}\rho_{\mathrm{B}}\tau_{\mathrm{B}}^{-1/2}\right)\end{aligned}$$

which concludes the proof. ■

The above bound can be simplified by setting $\tau_{\mathrm{B}} = \rho_{\mathrm{B}}$

$$D_u(\mathrm{A}|\mathrm{B})_{\rho} \leq \frac{1}{2}\sqrt{d_{\mathrm{A}}\Gamma_{\mathrm{C}}(\rho_{\mathrm{AB}}|\rho_{\mathrm{B}}) - \operatorname{tr}\rho_{\mathrm{B}}}. \tag{20}$$

[12] Similar to Lemma 2, measuring $\bar{\rho}_{\mathrm{XB}}$ on the X system can not decrease the alternative min-entropy while the distance to ρ_{XB} can not increase.

[13] Note that sub-normalized states have to be considered due to our definition of the smoothing of the min-entropy.

The following lemma yields a bound on the collision entropy of the output of the hash function in terms of the collision entropy of the input.

Lemma 5: Let $\mathcal{F}$ be δ-almost two-universal, let ρ_{XE} and ρ_{FZE} be defined as in (15) and (16), respectively, and let $\tau_{\mathrm{E}} \in \mathcal{S}_{=}(\mathcal{H}_{\mathrm{E}})$. Then,

$$\Gamma_{\mathrm{C}}(\rho_{\mathrm{FZE}}|\rho_{\mathrm{F}} \otimes \tau_{\mathrm{E}}) \le \Gamma_{\mathrm{C}}(\rho_{\mathrm{XE}}|\tau_{\mathrm{E}}) + \delta \,\mathrm{tr}\, \left(\rho_{\mathrm{E}}\tau_{\mathrm{E}}^{-1/2}\rho_{\mathrm{E}}\tau_{\mathrm{E}}^{-1/2}\right).$$

Proof: The collision entropy on the l.h.s. can be rewritten as an expectation value over F, that is

$$\begin{aligned}
&\Gamma_{\mathrm{C}}(\rho_{\mathrm{FZE}}|\rho_{\mathrm{F}} \otimes \tau_{\mathrm{E}}) \\
&\quad = \sum_f \mathrm{tr}\, \left(\rho_{\mathrm{FZE}}(p_f \mathbb{1}_{\mathrm{FZ}} \otimes \tau_{\mathrm{E}})^{-1/2}\rho_{\mathrm{ZEF}}(p_f \mathbb{1}_{\mathrm{FZ}} \otimes \tau_{\mathrm{E}})^{-1/2}\right) \\
&\quad = \sum_f p_f \sum_z \mathrm{tr}\, \left(|f\rangle\langle f|_{\mathrm{F}} \otimes |z\rangle\langle z|_{\mathrm{Z}} \otimes \rho_{\mathrm{E}}^{[f,z]}\tau_{\mathrm{E}}^{-1/2}\rho_{\mathrm{E}}^{[f,z]}\tau_{\mathrm{E}}^{-1/2}\right) \\
&\quad = \mathop{\mathbb{E}}_{F\in\mathcal{F}} \left[\sum_z \mathrm{tr}\, \left(\rho_{\mathrm{E}}^{[F,z]}\tau_{\mathrm{E}}^{-1/2}\rho_{\mathrm{E}}^{[F,z]}\tau_{\mathrm{E}}^{-1/2}\right)\right] \\
&\quad = \sum_{x,x'} \mathop{\mathbb{E}}_{F\in\mathcal{F}} \left[\sum_z \delta_{F(x)=z}\delta_{F(x')=z}\right] \mathrm{tr}\, \left(\rho_{\mathrm{E}}^{[x]}\tau_{\mathrm{E}}^{-1/2}\rho_{\mathrm{E}}^{[x']}\tau_{\mathrm{E}}^{-1/2}\right).
\end{aligned}$$

We have used (17) to substitute for $\rho_{\mathrm{E}}^{[F,z]}$ in the last step. The expectation value can be evaluated using the defining property (3) of δ-almost two-universal families. We get

$$\mathop{\mathbb{E}}_{F\in\mathcal{F}} \left[\sum_z \delta_{F(x)=z}\delta_{F(x')=z}\right] \le \delta$$

if $x \neq x'$ and 1 otherwise. We use this relation and the fact that the trace terms are positive to bound

$$\begin{aligned}
&\Gamma_{\mathrm{C}}(\rho_{\mathrm{FZE}}|\rho_{\mathrm{F}} \otimes \tau_{\mathrm{E}}) \\
&\quad \le \sum_x \mathrm{tr}\left(\rho_{\mathrm{E}}^{[x]}\tau_{\mathrm{E}}^{-1/2}\rho_{\mathrm{E}}^{[x]}\tau_{\mathrm{E}}^{-1/2}\right) + \delta \sum_{x\neq x'} \mathrm{tr}\left(\rho_{\mathrm{E}}^{[x]}\tau_{\mathrm{E}}^{-1/2}\rho_{\mathrm{E}}^{[x']}\tau_{\mathrm{E}}^{-1/2}\right).
\end{aligned}$$

We now complete the second sum with the terms where $x = x'$ to get the statement of the lemma. ■

If we set $\tau_{\mathrm{E}} = \rho_{\mathrm{E}}$, the result can be simplified further

$$\Gamma_{\mathrm{C}}(\rho_{\mathrm{FZE}}|\rho_{\mathrm{F}} \otimes \rho_{\mathrm{E}}) \le \Gamma_{\mathrm{C}}(\rho_{\mathrm{XE}}|\rho_{\mathrm{E}}) + \delta \,\mathrm{tr}\, \rho_{\mathrm{E}}. \qquad (21)$$

We are now ready to give a bound on the distance from uniform $D_u(\mathrm{Z}|\mathrm{FE})$ after privacy amplification with two-universal and δ-almost two-universal families of hash functions. For two-universal hashing, we get the following bound. (See also [12], where the same result was shown for a slightly different definition of the smooth min-entropy.)

Theorem 6: Let $\mathcal{F}$ be two-universal and let ρ_{XE} and ρ_{ZEF} be defined as in (15) and (16), respectively. Then, for any $\varepsilon \ge 0$

$$D_u(\mathrm{Z}|\mathrm{FE})_\rho \le \varepsilon + \frac{1}{2}\sqrt{2^{\ell - H_{\min}^{\varepsilon}(\mathrm{X}|\mathrm{E})_\rho}}.$$

Proof: We use Lemma 4 to bound $D_u(\mathrm{Z}|\mathrm{FE})_\rho$. In particular, we set $\tau_{\mathrm{FE}} := \rho_{\mathrm{F}} \otimes \tau_{\mathrm{E}}$ to get

$$\begin{aligned}
2D_u(\mathrm{Z}|\mathrm{FE})_\rho &\le \sqrt{2^\ell \Gamma_{\mathrm{C}}(\rho_{\mathrm{ZFE}}|\tau_{\mathrm{FE}}) - \mathrm{tr}\, \left(\rho_{\mathrm{E}}\tau_{\mathrm{E}}^{-1/2}\rho_{\mathrm{E}}\tau_{\mathrm{E}}^{-1/2}\right)} \\
&\le \sqrt{2^\ell \Gamma_{\mathrm{C}}(\rho_{\mathrm{XE}}|\tau_{\mathrm{E}})}
\end{aligned}$$

where we have used Lemma 5 and that $\mathcal{F}$ is two-universal ($\delta \le 2^{-\ell}$) in the last step. The r.h.s. can be expressed in terms of a min-entropy using (13). With an appropriate choice of τ_{E}, we have

$$2D_u(\mathrm{Z}|\mathrm{FE})_\rho \le \sqrt{2^{\ell - H_{\min}(\mathrm{X}|\mathrm{E})_\rho}}. \qquad (22)$$

We have now shown the theorem for the case $\varepsilon = 0$.

Finally, the bound can be expressed in terms of a smooth min-entropy. Let $\tilde{\rho}_{\mathrm{XE}} \in \mathcal{B}^{\varepsilon}(\rho_{\mathrm{XE}})$ be the CQ state (cf. Lemma 2) that optimizes the smooth min-entropy $H_{\min}^{\varepsilon}(\mathrm{X}|\mathrm{E})_\rho = H_{\min}(\mathrm{X}|\mathrm{E})_{\tilde{\rho}}$. We define $\tilde{\rho}_{\mathrm{FZE}} := (\mathcal{A} \otimes \mathcal{I}_{\mathrm{E}})(\rho_{\mathrm{F}} \otimes \tilde{\rho}_{\mathrm{XE}})$ and note that privacy amplification can only decrease the purified distance (12), i.e.,

$$\frac{1}{2}\|\rho_{\mathrm{FZE}} - \tilde{\rho}_{\mathrm{FZE}}\|_1 \le P(\rho_{\mathrm{FZE}}, \tilde{\rho}_{\mathrm{FZE}}) \le P(\rho_{\mathrm{XE}}, \tilde{\rho}_{\mathrm{XE}}) \le \varepsilon.$$

Moreover, let $\tilde{\sigma}_{\mathrm{FE}}$ be the state that minimizes the distance from uniform $D_u(\mathrm{Z}|\mathrm{FE})_{\tilde{\rho}}$. Then

$$\begin{aligned}
2D_u(\mathrm{Z}|\mathrm{FE})_\rho &\le \|\rho_{\mathrm{FZE}} - \omega_{\mathrm{Z}} \otimes \tilde{\sigma}_{\mathrm{FE}}\|_1 \\
&\le \|\rho_{\mathrm{FZE}} - \tilde{\rho}_{\mathrm{FZE}}\|_1 + \|\tilde{\rho}_{\mathrm{FZE}} - \omega_{\mathrm{Z}} \otimes \tilde{\sigma}_{\mathrm{FE}}\|_1 \\
&\le 2\varepsilon + 2D_u(\mathrm{Z}|\mathrm{FE})_{\tilde{\rho}}.
\end{aligned}$$

We now apply (22) for $\tilde{\rho}_{\mathrm{FZE}}$ (instead of ρ_{FZE}) to get

$$\begin{aligned}
D_u(\mathrm{Z}|\mathrm{FE})_\rho &\le \varepsilon + \frac{1}{2}\sqrt{2^{\ell - H_{\min}(\mathrm{X}|\mathrm{E})_{\tilde{\rho}}}} \\
&= \varepsilon + \frac{1}{2}\sqrt{2^{\ell - H_{\min}^{\varepsilon}(\mathrm{X}|\mathrm{E})_\rho}}
\end{aligned}$$

which concludes the proof. ■

Next, we consider the case of δ-almost two-universal hashing.

Theorem 7: Let $\mathcal{F}$ be δ-almost two-universal and let ρ_{XE} and ρ_{ZEF} be defined as in (15) and (16), respectively. Then, for any $\varepsilon \ge 0$ and $\eta > 0$

$$D_u(\mathrm{Z}|\mathrm{FE})_\rho \le \varepsilon + \eta + \frac{1}{2}\sqrt{(2^\ell\delta - 1) + \left(\frac{2}{\eta^2} + \frac{1}{1-\varepsilon}\right)2^{\ell - H_{\min}^{\varepsilon}(\mathrm{X}|\mathrm{E})_\rho}}.$$

Proof: We use Lemma 4 as in (20) to bound $D_u(\mathrm{Z}|\mathrm{FE})_\rho$. For normalized ρ_{ZFE}, we find

$$\begin{aligned}
2D_u(\mathrm{Z}|\mathrm{FE})_\rho &\le \sqrt{2^\ell \Gamma_{\mathrm{C}}(\rho_{\mathrm{FZE}}|\rho_{\mathrm{F}} \otimes \rho_{\mathrm{E}}) - 1} \\
&\le \sqrt{2^\ell \Gamma_{\mathrm{C}}(\rho_{\mathrm{XE}}|\rho_{\mathrm{E}}) + (2^\ell\delta - 1)}
\end{aligned}$$

where we used Lemma 5 as stated in (21).

The smoothing of the above equation is achieved using the same arguments as in the proof of Theorem 6. However, this time we need to include an additional smoothing parameter $\eta > 0$ in order to be able to apply (14).

Let $\tilde{\rho}_{\mathrm{XE}} \in \mathcal{B}^{\varepsilon}(\rho_{\mathrm{XE}})$ be a CQ state (cf. Lemma 2) that optimizes the smooth min-entropy $H^{\varepsilon}_{\min}(\mathrm{X}|\mathrm{E})_{\rho} = H_{\min}(\mathrm{X}|\mathrm{E})_{\tilde{\rho}}$ and let $\bar{\rho}_{\mathrm{XE}} \in \mathcal{B}^{\eta}(\tilde{\rho}_{\mathrm{XE}})$ be the CQ state (cf. Lemma 2) that satisfies

$$\begin{aligned}\Gamma_{\mathrm{C}}(\bar{\rho}_{\mathrm{XE}}|\bar{\rho}_{\mathrm{E}}) &\le \left(\frac{2}{\eta^2}+\frac{1}{\operatorname{tr}\tilde{\rho}_{\mathrm{XE}}}\right)2^{-H_{\min}(\mathrm{X}|\mathrm{E})_{\tilde{\rho}}}\\ &\le \left(\frac{2}{\eta^2}+\frac{1}{1-\varepsilon}\right)2^{-H^{\varepsilon}_{\min}(\mathrm{X}|\mathrm{E})_{\rho}}. \qquad (23)\end{aligned}$$

Then, $\bar{\rho}_{\mathrm{XE}} \in \mathcal{B}^{\varepsilon+\eta}(\rho_{\mathrm{XE}})$ holds due to the triangle inequality of the purified distance. Moreover, we define the state after randomness extraction, $\bar{\rho}_{\mathrm{FZE}} := (\mathcal{A}\otimes\mathcal{I}_{\mathrm{E}})(\rho_{\mathrm{F}}\otimes\bar{\rho}_{\mathrm{XE}})$. Following the arguments laid out in the proof of Theorem 6, we have

$$\begin{aligned}D_u(\mathrm{Z}|\mathrm{FE})_{\rho} &\le \varepsilon+\eta+D_u(\mathrm{Z}|\mathrm{FE})_{\bar{\rho}}\\ &\le \varepsilon+\eta+\frac{1}{2}\sqrt{2^{\ell}\Gamma_{\mathrm{C}}(\bar{\rho}_{\mathrm{XE}}|\bar{\rho}_{\mathrm{E}})+(2^{\ell}\delta-1)}.\end{aligned}$$

This can be bounded using (23), which concludes the proof. ■

The proof of the Leftover Hash Lemma stated in the introduction (Lemma 1) follows when we set $\varepsilon = 0$ and $\eta = \frac{\Delta}{2}$ in Theorem 6 and Theorem 7. To see this, note that the statements of both theorems can be expressed in terms of the distance from uniform averaged over the choice of hash function, Δ, using (19).

IV. Explicit Constructions With Shorter Seeds

Here, we combine known constructions of two-universal and δ-almost two-universal hash functions and discuss their use for randomness extraction with shorter random seeds. We restrict ourselves to the scenario where X is an n-bit string $x \in \{0,1\}^n$. The challenge is typically to optimize the following parameters:

a) the error described by the distance from uniform, Δ, which should be small;

b) the length of the extracted key, ℓ, which one wants to make as large as possible (close to $H^{\varepsilon}_{\min}(\mathrm{X}|\mathrm{E})$);

c) the length of the random seed, $s := \log|\mathcal{F}|$, needed to choose f, which one wants to keep small.

The latter point is important in practical implementations of privacy amplification, for example in quantum key distribution (QKD), where the choice of f has to be communicated between two parties.

We will first review the explicit constructions of (δ-almost) two-universal hash functions used in this section. In [7], Carter and Wegman proposed several constructions of two-universal function families, trying to minimize the size of $\mathcal{F}$. An example of a two-universal set of hash functions with $|\mathcal{F}| = 2^n$ is the set $\mathcal{F} = \{f_\alpha\}_{\alpha\in\{0,1\}^n}$ consisting of elements

$$\begin{array}{cccccc} f_\alpha & : & \{0,1\}^n & \longrightarrow & \{0,1\}^{\ell} & \\ & & x & \longmapsto & x\cdot\alpha \mod 2^{\ell} & \end{array} \qquad (24)$$

where $x\cdot\alpha$ denotes the multiplication in the field $\mathrm{GF}(2^n)$. The fact that $\mathcal{F}$ is two-universal can be readily verified by considering the difference $f_\alpha(x) - f_\alpha(x') = (x-x')\cdot\alpha \bmod 2^{\ell}$ and noting that the mapping $\alpha \mapsto (x-x')\cdot\alpha$ is a bijection if $x - x' \neq 0$.

With δ-almost two-universal families, a larger value of δ typically allows for a smaller set $\mathcal{F}$. This is nicely illustrated by the following well-known construction based on polynomials. Let $\mathbb{F}$ be an arbitrary field and let r be a positive integer. We define the family $\mathcal{F} = \{f_\alpha\}_{\alpha\in\mathbb{F}}$ of functions

$$\begin{array}{ccccc} f_\alpha & : & \mathbb{F}^r & \longrightarrow & \mathbb{F} \\ & & (x_1,\ldots,x_r) & \longmapsto & \sum_{i=1}^{r} x_i\alpha^{r-i}. \end{array} \qquad (25)$$

Using the fact that a polynomial of degree $r-1$ can only have $r-1$ zeros, it is easy to verify that $\mathcal{F}$ is δ-almost two-universal, for $\delta = (r-1)/|\mathbb{F}|$.

Another method to construct δ-almost two-universal families of hash functions is to concatenate two such families. We will use the following lemma by Stinson (see Theorem 5.4 in [8]).

Lemma 8: Let $\mathcal{F}_1$ be δ_1-almost two-universal from $\{0,1\}^n$ to $\{0,1\}^k$ and let $\mathcal{F}_2$ be δ_2-almost two-universal from $\{0,1\}^k$ to $\{0,1\}^{\ell}$. Then, the family $\{f_2\circ f_1 : f_1\in\mathcal{F}_1, f_2\in\mathcal{F}_2\}$ of all concatenated hash functions is $(\delta_1+\delta_2)$-almost two-universal.

Combining the general results on δ-almost two-universal hashing of Section III with the explicit constructions described above, we obtain the following statements.

If we do not care about the seed length s, we may choose a two-universal family of hash functions and recover a result by Renner [12]:

Theorem 9: For any $\varepsilon \ge 0$, there exists a family of hash functions from $\{0,1\}^n$ to $\{0,1\}^{\ell}$ satisfying

$$s = n \text{ and } \Delta = \varepsilon + \frac{1}{2}\sqrt{2^{\ell - H^{\varepsilon}_{\min}(\mathrm{X}|\mathrm{E})_{\rho}}}.$$

Proof: We apply Theorem 6 using the two-universal family constructed in (24), which yields $s = \log|\mathcal{F}| = n$. ■

We now show that we can choose a family of hash functions such that s is proportional to the key length ℓ instead of the input string length n.

Theorem 10: For any $\varepsilon \ge 0$, $\eta > 0$ and $\mu > 0$, there exists a family of hash functions from $\{0,1\}^n$ to $\{0,1\}^{\ell}$ satisfying[14]

$$s = 2\left\lfloor \ell + \log\left(\frac{n}{\ell}\right) + \log\left(\frac{1}{\mu^2}\right) - 1\right\rfloor \quad \text{and}$$

$$\Delta = \varepsilon + \eta + \mu + \frac{1}{2}\sqrt{\left(\frac{2}{\eta^2}+\frac{1}{1-\varepsilon}\right)2^{\ell - H^{\varepsilon}_{\min}(\mathrm{X}|\mathrm{E})_{\rho}}}.$$

Proof: As in the classical approach of [29], we concatenate two hash functions using Lemma 8 with some clever choice of the parameters. For the first function, we set $k = \lfloor \ell + \log(n/\ell) + \log(1/\mu^2) - 1\rfloor$ and use the field $\mathbb{F} = \mathrm{GF}(2^k)$ in the polynomial-based hash construction from (25). Interpreting the n-bit strings as $r = \lceil n/k\rceil$ blocks of k bits, the first hash function maps from $\{0,1\}^n$ to $\{0,1\}^k$ and requires a k-bit seed. Then, regular two-universal hashing from (24) with a seed length of again k bits is used to map from $\{0,1\}^k$ to $\{0,1\}^{\ell}$. The two seed lengths add up to $s = 2k = 2\lfloor \ell + \log(n/\ell) + \log(1/\mu^2) - 1\rfloor$.

Polynomial-based hashing achieves a δ_1 of at most

$$\frac{r-1}{2^k} \le \frac{n}{k\,2^k} \le \frac{4\,\ell\,\mu^2}{k\,2^{\ell}} \le \frac{4\mu^2}{2^{\ell}}$$

[14] The variable μ allows us to trade off key length against seed length.

by the choice of r and the fact that $k \geq \ell + \log(n/\ell) + \log(1/\varepsilon^2) - 2$. Together with $\delta_2 \leq 2^{-\ell}$ from the two-universal hashing, we get from Lemma 8 that this construction yields a $\delta_1 + \delta_2 \leq \frac{1+4\mu^2}{2^\ell}$-almost two-universal family of hash functions. Inserting this expression for δ into Theorem 7 yields

$$\Delta \leq \varepsilon + \eta + \frac{1}{2}\sqrt{\left(\frac{2}{\eta^2} + \frac{1}{1-\varepsilon}\right) 2^{\ell - H_{\min}^{\varepsilon}(\mathrm{X}|\mathrm{E})_\rho} + 4\mu^2}.$$

The theorem then follows as an upper bound to this expression. ■

To illustrate the difference between the two constructions above, we keep the distance from uniform, Δ, fixed and consider the min-entropy $H_{\min}^{\varepsilon}(\mathrm{X}|\mathrm{E})$ for $\varepsilon = \Delta/2$. The constructions allow us to extract ℓ bits of randomness using s bits of random seed. For the two-universal construction in Theorem 9, we find

$$\ell = \left\lfloor H_{\min}^{\Delta/2}(\mathrm{X}|\mathrm{E}) - 2\log\frac{1}{\Delta}\right\rfloor \quad \text{and} \quad s = n. \tag{26}$$

For the almost two-universal construction in Theorem 10, using $\eta = \mu = \Delta/8$ and $\varepsilon = \Delta/2$, we find

$$\begin{aligned} \ell &= \left\lfloor H_{\min}^{\Delta/2}(\mathrm{X}|\mathrm{E}) - 4\log\frac{1}{\Delta} - 10\right\rfloor \\ s &= 2\left\lfloor \ell + \log\frac{n}{\ell} + 2\log\frac{1}{\Delta} + 5\right\rfloor. \end{aligned} \tag{27}$$

Therefore, using the almost two-universal construction reduces the required random seed s from n to something proportional to ℓ, $\log\frac{n}{\ell}$ and $\log\frac{1}{\Delta}$ while keeping the extractable randomness ℓ unchanged up to terms in $\log\frac{1}{\Delta}$.

Appendix A
Technical Results

The first lemma is an application of Uhlmann's theorem [30] to the purified distance[15] (see [17] for a proof).

Lemma 11: Let $\rho, \tau \in \mathcal{S}_{\leq}(\mathcal{H})$, $\mathcal{H}' \cong \mathcal{H}$ and $\varphi \in \mathcal{H} \otimes \mathcal{H}'$ be a purification of ρ. Then, there exists a purification $\vartheta \in \mathcal{H} \otimes \mathcal{H}'$ of τ with $P(\rho, \tau) = P(\varphi, \vartheta)$.

Corollary 12: Let $\rho, \tau \in \mathcal{S}_{\leq}(\mathcal{H})$ and $\bar{\rho} \in \mathcal{S}_{\leq}(\mathcal{H} \otimes \mathcal{H}')$ be an extension of ρ. Then, there exists an extension $\bar{\tau} \in \mathcal{S}_{\leq}(\mathcal{H} \otimes \mathcal{H}')$ of τ with $P(\rho, \tau) = P(\bar{\rho}, \bar{\tau})$.

In the following, we apply this result to an ε-ball of pure states, $\mathcal{B}_{\mathrm{p}}^{\varepsilon}(\rho) := \{\tilde{\rho} \in \mathcal{B}^{\varepsilon}(\rho) : \mathrm{rank}\,\tilde{\rho} = 1\}$.

Corollary 13: Let $\rho \in \mathcal{S}_{\leq}(\mathcal{H})$ and $\varphi \in \mathcal{H} \otimes \mathcal{H}'$ be a purification of ρ. Then

$$\mathcal{B}^{\varepsilon}(\rho) \supseteq \left\{\tilde{\rho} \in \mathcal{S}_{\leq}(\mathcal{H}) : \exists\, \tilde{\phi} \in \mathcal{B}_{\mathrm{p}}^{\varepsilon}(\varphi) \text{ s.t. } \tilde{\rho} = \mathrm{tr}_{\mathcal{H}'}\tilde{\phi}\right\}$$

and equality holds if the Hilbert space dimensions satisfy $\dim\mathcal{H}' \geq \dim\mathcal{H}$.

We will use the following property of pure bipartite states.

Lemma 14: Let $\phi_{\mathrm{AB}} \in \mathcal{P}(\mathcal{H}_{\mathrm{AB}})$ be pure, $\rho_{\mathrm{A}} = \mathrm{tr}_{\mathrm{B}}\,\phi_{\mathrm{AB}}$, $\rho_{\mathrm{B}} = \mathrm{tr}_{\mathrm{A}}\,\phi_{\mathrm{AB}}$ and let $X \in \mathcal{L}(\mathcal{H}_{\mathrm{A}})$ be an operator with support and image in $\mathrm{supp}\{\rho_{\mathrm{A}}\}$. Then

$$(X \otimes \mathbb{1}_{\mathrm{B}})|\phi\rangle_{\mathrm{AB}} = \left(\mathbb{1}_{\mathrm{A}} \otimes \left(\rho_{\mathrm{B}}^{1/2} X^T \rho_{\mathrm{B}}^{-1/2}\right)\right)|\phi\rangle_{\mathrm{AB}}$$

and the transpose is taken in the Schmidt basis of ϕ_{AB}.

Proof: We introduce the Schmidt decomposition $|\phi\rangle_{\mathrm{AB}} = \sum_i \sqrt{\lambda_i}\,|i\rangle_{\mathrm{A}} \otimes |i\rangle_{\mathrm{B}}$. Clearly, $(\mathbb{1}_{\mathrm{A}} \otimes \rho_{\mathrm{B}}^{-1/2})|\phi\rangle_{\mathrm{AB}} = \sum_i |i\rangle_{\mathrm{A}} \otimes |i\rangle_{\mathrm{B}} =: |\gamma\rangle_{\mathrm{AB}}$ is the (unnormalized) fully entangled state on the support of ρ_{A} and ρ_{B}. It is easy to verify that $(X \otimes \mathbb{1}_{\mathrm{B}})|\gamma\rangle_{\mathrm{AB}} = (\mathbb{1}_{\mathrm{A}} \otimes X^T)|\gamma\rangle_{\mathrm{AB}}$, where the transposed matrix is given by $X^T = \sum_{i,j} \langle i|X|j\rangle_{\mathrm{A}} |j\rangle\langle i|_{\mathrm{B}}$. ■

Corollary 15: Let $\phi_{\mathrm{AB}} \in \mathcal{P}(\mathcal{H}_{\mathrm{AB}})$ be pure, $\rho_{\mathrm{A}} = \mathrm{tr}_{\mathrm{B}}\,\phi_{\mathrm{AB}}$, $\rho_{\mathrm{B}} = \mathrm{tr}_{\mathrm{A}}\phi_{\mathrm{AB}}$ and $f : \mathbb{R}^+ \to \mathbb{R}$ a real-valued function, then

$$(f(\rho_{\mathrm{A}}) \otimes \mathbb{1}_{\mathrm{B}})|\phi\rangle_{\mathrm{AB}} = (\mathbb{1}_{\mathrm{A}} \otimes f(\rho_{\mathrm{B}}))|\phi\rangle_{\mathrm{AB}}.$$

Proof: Observe that $\rho_{\mathrm{B}} = \rho_{\mathrm{A}}^T$ and apply Lemma 14. ■

We define the notion of a *dual projector* with regard to a pure state using the following corollary:

Corollary 16: Let $|\phi\rangle_{\mathrm{AB}} \in \mathcal{H}_{\mathrm{AB}}$ be pure, $\rho_{\mathrm{A}} = \mathrm{tr}_{\mathrm{B}}\,\phi_{\mathrm{AB}}$, $\rho_{\mathrm{B}} = \mathrm{tr}_{\mathrm{A}}\,\phi_{\mathrm{AB}}$ and let $\Pi_{\mathrm{A}} \in \mathcal{P}(\mathcal{H}_{\mathrm{A}})$ be a projector in $\mathrm{supp}\{\rho_{\mathrm{A}}\}$. Then, there exists a dual projector Π_{B} on $\mathcal{H}_{\mathrm{B}}$ such that

$$\left(\Pi_{\mathrm{A}} \otimes \rho_{\mathrm{B}}^{-1/2}\right)|\phi\rangle_{\mathrm{AB}} = \left(\rho_{\mathrm{A}}^{-1/2} \otimes \Pi_{\mathrm{B}}\right)|\phi\rangle_{\mathrm{AB}}.$$

The next Lemma gives a bound on the purified distance of a state ρ and a projected state $\Pi\rho\Pi$.

Lemma 17: Let $\rho \in \mathcal{S}_{\leq}(\mathcal{H})$ and Π a projector on $\mathcal{H}$, then

$$P(\rho, \Pi\rho\Pi) \leq \sqrt{2\,\mathrm{tr}(\Pi^{\perp}\rho) - \mathrm{tr}(\Pi^{\perp}\rho)^2}$$

where $\Pi^{\perp} = \mathbb{1} - \Pi$ is the complement of Π on $\mathcal{H}$.

Proof: The generalized fidelity between the two states can be bounded using $\mathrm{tr}(\Pi\rho) \leq \mathrm{tr}(\rho)$. We have

$$\bar{F}(\rho, \Pi\rho\Pi) \geq \mathrm{tr}(\Pi\rho) + 1 - \mathrm{tr}\,\rho = 1 - \mathrm{tr}(\Pi^{\perp}\rho).$$

The desired bound on the purified distance follows from its definition. ■

Appendix B
Alternative Entropic Quantities

Here, we discuss two alternative entropic quantities, $\hat{H}_{\min}^{\varepsilon}(\mathrm{A}|\mathrm{B})$ and $\hat{H}_{\max}^{\varepsilon}(\mathrm{A}|\mathrm{B})$, and show that they are equivalent (up to terms in $\log\varepsilon$) to the smooth min-entropy and smooth max-entropy, respectively. Some of the technical results of this Appendix will be used to give a bound on the collision entropy in terms of the smooth min-entropy (cf. Lemma 2).

First, note that conditional entropies can be defined from relative entropies, as is well-known for the case of the von Neumann

[15] The main advantage of the purified distance over the trace distance is that we can always find extensions and purifications without increasing the distance.

entropy. Let ρ_{AB} be a bipartite quantum state. Then, the conditional von Neumann entropy of A given B is

$$\begin{aligned} H(\mathrm{A}|\mathrm{B})_\rho &:= H(\rho_{\mathrm{AB}}) - H(\rho_{\mathrm{B}}) \\ &= -D(\rho_{\mathrm{AB}} \,\|\, \mathbb{1}_{\mathrm{A}} \otimes \rho_{\mathrm{B}}) \qquad (28) \\ &= -\min_{\sigma_{\mathrm{B}} \in \mathcal{S}_=(\mathcal{H}_{\mathrm{B}})} D(\rho_{\mathrm{AB}} \,\|\, \mathbb{1}_{\mathrm{A}} \otimes \sigma_{\mathrm{B}}) \qquad (29) \end{aligned}$$

where we used Klein's inequality [31], [14] in the last step. The relative entropy is defined as $D(\rho \,\|\, \tau) := \mathrm{tr}(\rho(\log\rho - \log\tau))$ and $H(\rho) := -\mathrm{tr}(\rho\log\rho)$.

We will now define the smooth min-entropy and an alternative to the smooth entropy as first introduced in [12]. The definition of two versions of the min-entropy is parallel to the case of the von Neumann entropy above; however, the two identities (28) and (29) now lead to different definitions. We follow [32] and first introduce the *max relative entropy*. For two positive operators $\rho \in \mathcal{S}_\leq(\mathcal{H})$ and $\tau \in \mathcal{P}(\mathcal{H})$ we define

$$D_{\max}(\rho \,\|\, \tau) := \inf\{\lambda \in \mathbb{R} : \rho \leq 2^\lambda \tau\}.$$

Definition 4: Let $\varepsilon \geq 0$ and $\rho_{\mathrm{AB}} \in \mathcal{S}_\leq(\mathcal{H}_{\mathrm{AB}})$. The min-entropy and the *alternative min-entropy* of A conditioned on B are given by

$$\begin{aligned} H_{\min}(\mathrm{A}|\mathrm{B})\rho &= \max_{\sigma_{\mathrm{B}} \in \mathcal{S}_=(\mathcal{H}_{\mathrm{B}})} -D_{\max}(\rho_{\mathrm{AB}} \,\|\, \mathbb{1}_{\mathrm{A}} \otimes \sigma_{\mathrm{B}}) \quad \text{and} \\ \hat{H}_{\min}(\mathrm{A}|\mathrm{B})_\rho &:= -D_{\max}(\rho_{\mathrm{AB}} \,\|\, \mathbb{1}_{\mathrm{A}} \otimes \rho_{\mathrm{B}}) \end{aligned}$$

respectively. Furthermore, the *alternative smooth min-entropy* of A conditioned on B is defined as

$$\hat{H}^\varepsilon_{\min}(\mathrm{A}|\mathrm{B})_\rho := \max_{\tilde{\rho}_{\mathrm{AB}} \in \mathcal{B}^\varepsilon(\rho_{\mathrm{AB}})} \hat{H}_{\min}(\mathrm{A}|\mathrm{B})_{\tilde{\rho}}.$$

The *smooth max-entropies* can be defined as duals of the smooth min-entropies.

Definition 5: Let $\varepsilon \geq 0$ and $\rho_{\mathrm{AB}} \in \mathcal{S}_\leq(\mathcal{H}_{\mathrm{AB}})$, then we define the *smooth max-entropy* and the *alternative smooth max-entropy* of A conditioned on B as

$$\begin{aligned} H^\varepsilon_{\max}(\mathrm{A}|\mathrm{B})_\rho &:= -H^\varepsilon_{\min}(\mathrm{A}|\mathrm{C})_\rho \quad \text{and} \\ \hat{H}^\varepsilon_{\max}(\mathrm{A}|\mathrm{B})_\rho &:= -\hat{H}^\varepsilon_{\min}(\mathrm{A}|\mathrm{C})_\rho \end{aligned}$$

where $\rho_{\mathrm{ABC}} \in \mathcal{S}_\leq(\mathcal{H}_{\mathrm{ABC}})$ is any purification of ρ_{AB}.

The max-entropies are well-defined since the min-entropies are invariant under local isometries on the C system (cf. [17] and Lemma 21) and, thus, independent of the chosen purification. The nonsmooth max-entropies $H_{\max}(\mathrm{A}|\mathrm{B})_\rho$ and $\hat{H}_{\max}(\mathrm{A}|\mathrm{B})_\rho$ are defined as the limit $\varepsilon \to 0$ of the corresponding smooth quantities. The alternative max-entropy is discussed in Appendix C, where it is shown that (cf. also [33])

$$\hat{H}_{\max}(\mathrm{A}|\mathrm{B})_\rho = \max_{\sigma_{\mathrm{B}} \in \mathcal{S}_=(\mathcal{H}_{\mathrm{B}})} \log\mathrm{tr}(\Pi_{\rho_{\mathrm{AB}}}(\mathbb{1}_{\mathrm{A}} \otimes \sigma_{\mathrm{B}})) \qquad (30)$$

where $\Pi_{\rho_{\mathrm{AB}}}$ is the projector onto the support of ρ_{AB}. Furthermore, we find that

$$\hat{H}^\varepsilon_{\max}(\mathrm{A}|\mathrm{B})_\rho = \inf_{\mathcal{H}_{\mathrm{B}'} \supseteq \mathcal{H}_{\mathrm{B}}} \; \min_{\tilde{\rho}_{\mathrm{AB}'} \in \mathcal{B}^\varepsilon(\rho_{\mathrm{AB}'})} \hat{H}_{\max}(\mathrm{A}|\mathrm{B}')_{\tilde{\rho}} \qquad (31)$$

where the infimum is taken over all embeddings $\rho_{\mathrm{AB}'}$ of ρ_{AB} into $\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}'}$. In fact, it is sufficient to consider an embedding into a space of size $\dim\mathcal{H}_{\mathrm{B}'} = \mathrm{rank}\,\{\rho_{\mathrm{AB}}\} \cdot \dim\mathcal{H}_{\mathrm{A}}$.

The first definition of the smooth max-entropy, $H^\varepsilon_{\max}(\mathrm{A}|\mathrm{B})$, is used in [16], [27] and is found to have many interesting properties, e.g. it satisfies a data-processing inequality [17]. The alternative definition, $\hat{H}^\varepsilon_{\max}(\mathrm{A}|\mathrm{B})$, was first introduced in [12] and is used to quantitatively characterize various information theoretic tasks (cf., e.g., [32], [34], [35]). Here, we find that the two smooth min-entropies and the two smooth max-entropies are pairwise equivalent up to terms in $\log\varepsilon$. Namely, the following lemma holds:

Lemma 18: Let $\varepsilon > 0$, $\varepsilon' \geq 0$ and $\rho_{\mathrm{AB}} \in \mathcal{S}_=(\mathcal{H}_{\mathrm{AB}})$, then

$$H^{\varepsilon'}_{\min}(\mathrm{A}|\mathrm{B})_\rho - \log c \leq \hat{H}^{\varepsilon+\varepsilon'}_{\min}(\mathrm{A}|\mathrm{B})_\rho \leq H^{\varepsilon+\varepsilon'}_{\min}(\mathrm{A}|\mathrm{B})_\rho$$

where $c = 2/\varepsilon^2 + 1/(1-\varepsilon')$.

The equivalence of the max-entropies follows by their definition as duals, i.e., we have

$$H^{\varepsilon'}_{\max}(\mathrm{A}|\mathrm{B})_\rho + \log c \geq \hat{H}^{\varepsilon+\varepsilon'}_{\max}(\mathrm{A}|\mathrm{B})_\rho \geq H^{\varepsilon+\varepsilon'}_{\max}(\mathrm{A}|\mathrm{B})_\rho.$$

The proof of Lemma 18 is based on the following result, where, for convenience of exposition, we introduce the generalized conditional min-entropy

$$h_{\min}(\mathrm{A}|\mathrm{B})_{\rho|\sigma} := -D_{\max}(\rho_{\mathrm{AB}} \,\|\, \mathbb{1}_{\mathrm{A}} \otimes \sigma_{\mathrm{B}}).$$

Lemma 19: Let $\varepsilon > 0$ and $\rho_{\mathrm{ABC}} \in \mathcal{S}_\leq(\mathcal{H}_{\mathrm{ABC}})$ be pure. Then, there exists a projector Π_{AC} on $\mathcal{H}_{\mathrm{AC}}$ and a state $\tilde{\rho}_{\mathrm{ABC}} = \Pi_{\mathrm{AC}}\,\rho_{\mathrm{ABC}}\,\Pi_{\mathrm{AC}}$ such that $\tilde{\rho}_{\mathrm{ABC}} \in \mathcal{B}^\varepsilon_{\mathrm{p}}(\rho_{\mathrm{ABC}})$ and

$$h_{\min}(\mathrm{A}|\mathrm{B})_{\tilde{\rho}|\rho} \geq H_{\min}(\mathrm{A}|\mathrm{B})\rho - \log\frac{2}{\varepsilon^2}.$$

Furthermore, there exists a state $\bar{\rho}_{\mathrm{AB}} \in \mathcal{S}_\leq(\mathcal{H}_{\mathrm{AB}})$ that satisfies $\bar{\rho}_{\mathrm{AB}} \in \mathcal{B}^\varepsilon(\rho_{\mathrm{AB}})$ and

$$\hat{H}_{\min}(\mathrm{A}|\mathrm{B})_{\bar{\rho}} \geq H_{\min}(\mathrm{A}|\mathrm{B})\rho - \log\left(\frac{2}{\varepsilon^2} + \frac{1}{\mathrm{tr}\,\rho_{\mathrm{AB}}}\right).$$

Proof: The proof is structured as follows: First, we give a lower bound on the entropy $h_{\min}(\mathrm{A}|\mathrm{B})_{\tilde{\rho}|\rho}$ in terms of $H_{\min}(\mathrm{A}|\mathrm{B})\rho$ and a projector Π_{B} that is the dual projector (cf. Corollary 16) of Π_{AC} with regard to ρ_{ABC}. We then find a lower bound on the purified distance between ρ_{ABC} and $\tilde{\rho}_{\mathrm{ABC}}$ in terms of Π_{B} and define Π_{B} (and, thus, Π_{AC}) such that this distance does not exceed ε.

Let λ and σ_{B} be the pair that optimizes the min-entropy $H_{\min}(\mathrm{A}|\mathrm{B})\rho$, i.e., $H_{\min}(\mathrm{A}|\mathrm{B})\rho = h_{\min}(\mathrm{A}|\mathrm{B})_{\rho|\sigma} = -\log\lambda$.

We have $\tilde\rho_{B} \le \rho_{B}$ by definition of $\tilde\rho_{ABC}$. Hence, $h_{\min}(A|B)_{\tilde\rho|\rho}$ is finite and can be written as

$$2^{-h_{\min}(A|B)_{\tilde\rho|\rho}} = \left\|{}^{-1/2}_{\rho_B}\tilde\rho_{AB}\rho_B^{-1/2}\right\|_\infty$$

where $\|X\|_\infty$ denotes the maximum eigenvalue of X. We bound this expression using the dual projector Π_B of Π_{AC} with regard to ρ_{ABC} and the fact that $\rho_{AB} \le \lambda \mathbb{1}_A \otimes \sigma_B$ by definition of λ and σ_B

$$\begin{aligned}\text{rhs.} &= \left\|\mathrm{tr}_C\left(\left(\Pi_{AC}\otimes\rho_B^{-1/2}\right)\rho_{ABC}\left(\Pi_{AC}\otimes\rho_B^{-1/2}\right)\right)\right\|_\infty\\ &= \left\|\Pi_B\,\rho_B^{-1/2}\rho_{AB}\,\rho_B^{-1/2}\Pi_B\right\|_\infty\\ &\le \lambda\,\left\|\mathbb{1}_A\otimes\Pi_B\,\rho_B^{-1/2}\sigma_B\,\rho_B^{-1/2}\Pi_B\right\|_\infty\\ &= \lambda\|\Pi_B\Gamma_B\Pi_B\|_\infty\end{aligned}$$

where, in the last step, we introduced the Hermitian operator $\Gamma_B := \rho_B^{-1/2}\sigma_B\,\rho_B^{-1/2}$. Taking the logarithm on both sides leads to

$$h_{\min}(A|B)_{\tilde\rho|\rho} \ge H_{\min}(A|B)\rho - \log\|\Pi_B\Gamma_B\Pi_B\|_\infty. \tag{32}$$

We use Lemma 17 to bound the distance between ρ_{ABC} and $\tilde\rho_{ABC}$, namely

$$P(\rho_{ABC},\tilde\rho_{ABC}) \le \sqrt{2\,\mathrm{tr}(\Pi_{AC}^{\perp}\rho_{ABC})} = \sqrt{2\,\mathrm{tr}(\Pi_B^{\perp}\rho_B)}$$

where the last equality can be verified using Corollary 16. Clearly, the optimal choice of Π_B will cut off the largest eigenvalues of Γ_B in (32) while keeping the states ρ_{ABC} and $\tilde\rho_{ABC}$ close. We, thus, define Π_B to be the minimum rank projector onto the smallest eigenvalues of Γ_B such that $\mathrm{tr}(\Pi_B\rho_B) \ge \mathrm{tr}\,\rho_B - \varepsilon^2/2$ or, equivalently, $\mathrm{tr}(\Pi_B^{\perp}\rho_B) \le \varepsilon^2/2$. This definition immediately implies that ρ_{ABC} and $\tilde\rho_{ABC}$ are ε-close and it remains to find an upper bound on $\|\Pi_B\Gamma_B\Pi_B\|_\infty$.

Let Π_B' be the projector onto the largest remaining eigenvalue in $\Pi_B\Gamma_B\Pi_B$ and note that Π_B' and $\Pi_B^{\perp}$ commute with Γ_B. Then

$$\|\Pi_B\Gamma_B\Pi_B\|_\infty = \mathrm{tr}(\Pi_B'\Gamma_B) = \min_{\mu_B}\frac{\mathrm{tr}(\mu_B(\Pi_B^{\perp}+\Pi_B)\Gamma_B)}{\mathrm{tr}(\mu_B)}$$

where μ_B is minimized over all positive operators in the support of $\Pi_B^{\perp}+\Pi_B'$. Fixing instead $\mu_B = (\Pi_B^{\perp}+\Pi_B')\rho_B(\Pi_B^{\perp}+\Pi_B')$, we find

$$\begin{aligned}\|\Pi_B\Gamma_B\Pi_B\|_\infty &\le \frac{\mathrm{tr}\left(\Gamma_B^{1/2}\rho_B\Gamma_B^{1/2}(\Pi_B^{\perp}+\Pi_B')\right)}{\mathrm{tr}((\Pi_B^{\perp}+\Pi_B')\rho_B)}\\ &\le \frac{\mathrm{tr}\left(\Gamma_B^{1/2}\rho_B\Gamma_B^{1/2}\right)}{\mathrm{tr}((\Pi_B^{\perp}+\Pi_B')\rho_B)} \le \frac{2}{\varepsilon^2}.\end{aligned}$$

In the last step we used that $\mathrm{tr}(\rho_B^{1/2}\Gamma_B\rho_B^{1/2}) = \mathrm{tr}(\sigma_B) = 1$ and that $\mathrm{tr}((\Pi_B^{\perp}+\Pi_B')\rho_B) \ge \frac{\varepsilon^2}{2}$ by definition of $\Pi_B^{\perp}$. We have now established the first statement.

To prove the second statement, we introduce an operator $\Delta_B := \rho_B - \tilde\rho_B \ge 0$. The state $\bar\rho_{AB} = \tilde\rho_{AB} + \mathbb{1}_A/d_A\otimes\Delta_B$, where $d_A = \dim\mathcal{H}_A$, satisfies $\bar\rho_B = \rho_B$. We now show that the state $\bar\rho_{AB}$ is ε-close to ρ_{AB}. The inequality $\tilde\rho_{AB} \le \bar\rho_{AB}$ implies $\|_{\sqrt{\tilde\rho_{AB}}\sqrt{\rho_{AB}}}\|_1 \le \|\sqrt{\bar\rho_{AB}}\sqrt{\rho_{AB}}\|_1$ and, thus

$$\begin{aligned}\bar F(\rho_{AB},\bar\rho_{AB}) &\ge F(\tilde\rho_{AB},\rho_{AB}) + 1 - \mathrm{tr}\,\rho_{AB}\\ &\ge F(\tilde\rho_{ABC},\rho_{ABC}) + 1 - \mathrm{tr}\,\rho_{AB}\\ &= 1 - \mathrm{tr}(\Pi_{AC}^{\perp}\rho_{AC}) \ge 1 - \varepsilon^2/2\end{aligned}$$

where we used the monotonicity of the fidelity $F(\rho,\tau) := \|\sqrt{\rho}\sqrt{\tau}\|_1$ under the partial trace. Thus, $P(\bar\rho_{AB},\rho_{AB}) \le \varepsilon$.

We use that $\bar\rho_B = \rho_B$ and $\bar\rho_{AB} \le \tilde\rho_{AB} + \mathbb{1}_A/d_A\otimes\rho_B$ to find a lower bound on $\hat H_{\min}(A|B)_{\bar\rho} = h_{\min}(A|B)_{\bar\rho|\rho}$

$$\begin{aligned}2^{-\hat H_{\min}(A|B)_{\bar\rho}} &= \left\|\rho_B^{-1/2}\bar\rho_{AB}\,\rho_B^{-1/2}\right\|_\infty\\ &\le \left\|\rho_B^{-1/2}\tilde\rho_{AB}\,\rho_B^{-1/2} + \frac{1}{d_A}\mathbb{1}_{AB}\right\|_\infty\\ &\le \lambda\,\frac{2}{\varepsilon^2} + \frac{1}{d_A}.\end{aligned}$$

We have $\lambda \ge \mathrm{tr}\,\rho_{AB}/d_A$ (Lemma 20 in [17]) and, thus

$$\hat H_{\min}(A|B)_{\bar\rho} \ge H_{\min}(A|B)\rho - \log\left(\frac{2}{\varepsilon^2} + \frac{1}{\mathrm{tr}\,\rho_{AB}}\right).$$

This concludes the proof of the second statement. ■

Furthermore, the alternative smooth min-entropy is a lower bound on the smooth min-entropy by definition.

Lemma 20: Let $\rho_{AB} \in \mathcal{S}_{\le}(\mathcal{H}_{AB})$, then

$$\hat H_{\min}(A|B)_\rho \le H_{\min}(A|B)\rho - \log\frac{1}{\mathrm{tr}\,\rho_{AB}}.$$

We are now ready to prove Lemma 18. Namely, we show that, for $\varepsilon > 0$, $\varepsilon' \ge 0$ and $\rho_{AB} \in \mathcal{S}_{\le}(\mathcal{H}_{AB})$, it holds that

$$H_{\min}^{\varepsilon'}(A|B)_\rho - \log c \le \hat H_{\min}^{\varepsilon+\varepsilon'}(A|B)_\rho \le H_{\min}^{\varepsilon+\varepsilon'}(A|B)_\rho$$

where $c = 2/\varepsilon^2 + 1/(\mathrm{tr}\,\rho_{AB} - \varepsilon')$.

Proof of Lemma 18: Let $\tilde\rho_{AB} \in \mathcal{B}^{\varepsilon'}(\rho_{AB})$ be the state that maximizes $H_{\min}^{\varepsilon'}(A|B)_\rho$. Clearly, $\mathrm{tr}\,\tilde\rho_{AB} \ge \mathrm{tr}\,\rho_{AB} - \varepsilon'$. Moreover, Lemma 19 and the triangle inequality of the purified distance imply that there exists a state $\bar\rho_{AB} \in \mathcal{B}^{\varepsilon+\varepsilon'}(\rho_{AB})$ that satisfies

$$\hat H_{\min}^{\varepsilon+\varepsilon'}(A|B)_\rho \ge \hat H_{\min}(A|B)_{\bar\rho} \ge H_{\min}^{\varepsilon'}(A|B)_\rho - \log c$$

which concludes the proof of the first inequality. The second inequality follows by applying Lemma 20 to the state that maximizes $\hat H_{\min}^{\varepsilon+\varepsilon'}(A|B)_\rho$. ■

APPENDIX C
DUALITY RELATION FOR ALTERNATIVE SMOOTH ENTROPIES

Here, we find that the alternative smooth min-entropy of A conditioned on B is invariant under local isometries on the B system. Since all purifications are equivalent up to isometries on the purifying system, this allows the definition of the alternative max-entropy as its dual (see Definition 5). Moreover, the max-entropy is invariant under local isometries on the B system as a direct consequence. Note that the alternative smooth min- and max-entropies are in general not invariant under isometries on the A system, i.e., they depend on the dimension of the Hilbert space $\mathcal{H}_{\mathrm{A}}$.

Lemma 21: Let $\varepsilon \geq 0$ and $\rho_{\mathrm{AB}} \in \mathcal{S}_{\leq}(\mathcal{H}_{\mathrm{AB}})$. Moreover, let $U : \mathcal{H}_{\mathrm{B}} \to \mathcal{H}_{\mathrm{D}}$ be an isometry with $\tau_{\mathrm{AD}} := (\mathbb{1}_{\mathrm{A}} \otimes U)\rho_{\mathrm{AB}}(\mathbb{1}_{\mathrm{A}} \otimes U^\dagger)$. Then

$$\begin{aligned}&\hat{H}^{\varepsilon}_{\min}(\mathrm{A}|\mathrm{B})_\rho = \hat{H}^{\varepsilon}_{\min}(\mathrm{A}|\mathrm{D})_\tau \quad \text{and}\\&\hat{H}^{\varepsilon}_{\max}(\mathrm{A}|\mathrm{B})_\rho = \hat{H}^{\varepsilon}_{\max}(\mathrm{A}|\mathrm{D})_\tau.\end{aligned}$$

Proof: Let $\tilde{\rho}_{\mathrm{AB}} \in \mathcal{B}^{\varepsilon}(\rho_{\mathrm{AB}})$ be the state that maximizes the alternative min-entropy of A conditioned on B and let λ be defined with $\hat{H}^{\varepsilon}_{\min}(\mathrm{A}|\mathrm{B})_\rho = -\log \lambda$. Then $\tilde{\rho}_{\mathrm{AB}} \leq \lambda \mathbb{1}_{\mathrm{A}} \otimes \tilde{\rho}_{\mathrm{B}}$, which implies

$$\underbrace{(\mathbb{1}_{\mathrm{A}} \otimes U)\tilde{\rho}_{\mathrm{AB}}(\mathbb{1}_{\mathrm{A}} \otimes U^\dagger)}_{=:\tilde{\tau}_{\mathrm{AD}}} \leq \lambda \mathbb{1}_{\mathrm{A}} \otimes (U\tilde{\rho}_{\mathrm{B}}U^\dagger).$$

Hence, $\tilde{\tau}_{\mathrm{AD}} \leq \lambda \mathbb{1}_{\mathrm{A}} \otimes \tilde{\tau}_{\mathrm{D}}$. Moreover, $\tilde{\tau}_{\mathrm{AD}} \in \mathcal{B}^{\varepsilon}(\tau_{\mathrm{AD}})$ due to (12), which implies $\hat{H}^{\varepsilon}_{\min}(\mathrm{A}|\mathrm{D})_\rho \geq \hat{H}^{\varepsilon}_{\min}(\mathrm{A}|\mathrm{B})_\rho$. The same argument in reverse can be applied to get $\hat{H}^{\varepsilon}_{\min}(\mathrm{A}|\mathrm{B})_\rho \geq \hat{H}^{\varepsilon}_{\min}(\mathrm{A}|\mathrm{D})_\tau$.

The invariance under isometry of the dual quantity follows by definition. Namely, let ρ_{ABE} be any purification of ρ_{AB}, then

$$\begin{aligned}\hat{H}^{\varepsilon}_{\max}(\mathrm{A}|\mathrm{B})_\rho &= -\hat{H}^{\varepsilon}_{\min}(\mathrm{A}|\mathrm{E})_\rho\\&= -\hat{H}^{\varepsilon}_{\min}(\mathrm{A}|\mathrm{E})_\tau = \hat{H}^{\varepsilon}_{\max}(\mathrm{A}|\mathrm{D})_\tau\end{aligned}$$

where $\tau_{\mathrm{ADE}} := (\mathbb{1}_{\mathrm{A}} \otimes U \otimes \mathbb{1}_{\mathrm{E}})\rho_{\mathrm{ABE}}(\mathbb{1}_{\mathrm{A}} \otimes U^\dagger \otimes \mathbb{1}_{\mathrm{E}})$ is a purification of τ_{AD}. ■

Next, we derive expression (30) for the alternative nonsmooth and smooth max-entropies. The result for the nonsmooth entropy was first shown in [33] and an alternative proof is provided here for completeness.

Lemma 22: Let $\rho_{\mathrm{AB}} \in \mathcal{S}_{\leq}(\mathcal{H}_{\mathrm{AB}})$, then

$$\hat{H}_{\max}(\mathrm{A}|\mathrm{B})_\rho = \max_{\sigma_{\mathrm{B}} \in \mathcal{S}_{=}(\mathcal{H}_{\mathrm{B}})} \log \mathrm{tr}(\Pi_{\rho_{\mathrm{AB}}}(\mathbb{1}_{\mathrm{A}} \otimes \sigma_{\mathrm{B}})$$

Proof: Let ρ_{ABC} be a purification of ρ_{AB}. Then, $\tau_{\mathrm{ABC}} := (\mathbb{1}_{\mathrm{AB}} \otimes \rho_{\mathrm{C}}^{-1/2})\rho_{\mathrm{ABC}}\,(\mathbb{1}_{\mathrm{AB}} \otimes \rho_{\mathrm{C}}^{-1/2})$ has marginal $\tau_{\mathrm{AB}} = \Pi_{\rho_{\mathrm{AB}}}$ due to Lemma 14. This allows us to write

$$\begin{aligned}2^{\hat{H}_{\max}(\mathrm{A}|\mathrm{B})_\rho} &= 2^{-\hat{H}_{\min}(\mathrm{A}|\mathrm{C})_\rho} = \|\tau_{\mathrm{AC}}\|_\infty = \|\tau_{\mathrm{B}}\|_\infty\\&= \max_{\sigma_{\mathrm{B}}} \mathrm{tr}(\sigma_{\mathrm{B}}\tau_{\mathrm{B}}) = \max_{\sigma_{\mathrm{B}}} \mathrm{tr}(\Pi_{\rho_{\mathrm{AB}}}(\mathbb{1}_{\mathrm{A}} \otimes \sigma_{\mathrm{B}}))\end{aligned}$$

where the maximization is over all $\sigma_{\mathrm{B}} \in \mathcal{S}_{=}(\mathcal{H}_{\mathrm{B}})$. ■

The alternative smooth max-entropy can be seen as an optimization of the nonsmooth quantity over an ε-ball of states, where the ball is embedded in a sufficiently large Hilbert space. We show that (31) holds.

Lemma 23: Let $\varepsilon \geq 0$ and $\rho_{\mathrm{AB}} \in \mathcal{S}_{\leq}(\mathcal{H}_{\mathrm{AB}})$, then

$$\hat{H}^{\varepsilon}_{\min}(\mathrm{A}|\mathrm{B})_\rho = \inf_{\mathcal{H}_{\mathrm{B}'} \supseteq \mathcal{H}_{\mathrm{B}}} \min_{\tilde{\rho}_{\mathrm{AB}'} \in \mathcal{B}^{\varepsilon}(\rho_{\mathrm{AB}'})} \hat{H}_{\min}(\mathrm{A}|\mathrm{B}')_{\tilde{\rho}}$$

where $\rho_{\mathrm{AB}'}$ is the embedding of ρ_{AB} into $\mathcal{H}_{\mathrm{AB}'}$. Furthermore, the infimum is attained for embeddings with $\dim \mathcal{H}_{\mathrm{B}'} \geq \mathrm{dimsupp}\{\rho_{\mathrm{AB}}\} \cdot \dim \mathcal{H}_{\mathrm{A}}$.

Proof: Let ρ_{ABC} be a purification of ρ_{AB} on a Hilbert space $\mathcal{H}_{\mathrm{C}}$ with $\mathrm{dim}\mathcal{H}_{\mathrm{C}} = \mathrm{rank}\,\{\rho_{\mathrm{AB}}\}$. Furthermore, for any $\mathcal{H}_{\mathrm{B}'} \supseteq \mathcal{H}_{\mathrm{B}}$, let $\rho_{\mathrm{AB}'\mathrm{C}'}$ be the embedding of ρ_{ABC} into $\mathcal{H}_{\mathrm{AB}'\mathrm{C}'}$ with $\mathrm{dim}\mathcal{H}_{\mathrm{C}'} = \mathrm{dim}\mathcal{H}_{\mathrm{AB}'}$. We use Corollary 13 twice to upper bound

$$\begin{aligned}\hat{H}^{\varepsilon}_{\max}(\mathrm{A}|\mathrm{B})_\rho &= -\hat{H}^{\varepsilon}_{\min}(\mathrm{A}|\mathrm{C}')_\rho\\&= \min_{\tilde{\rho}_{\mathrm{AC}'} \in \mathcal{B}^{\varepsilon}(\rho_{\mathrm{AC}'})} -\hat{H}_{\min}(\mathrm{A}|\mathrm{C}')_{\tilde{\rho}}\\&\leq \min_{\tilde{\rho}_{\mathrm{AB}'\mathrm{C}'} \in \mathcal{B}^{\varepsilon}_{\mathrm{p}}(\rho_{\mathrm{AB}'\mathrm{C}'})} \hat{H}_{\min}(\mathrm{A}|\mathrm{B}')_{\tilde{\rho}}\\&= \min_{\tilde{\rho}_{\mathrm{AB}'} \in \mathcal{B}^{\varepsilon}(\rho_{\mathrm{AB}'})} \hat{H}_{\min}(\mathrm{A}|\mathrm{B}')_{\tilde{\rho}}.\end{aligned}$$

A lower bound on $\hat{H}^{\varepsilon}_{\min}(\mathrm{A}|\mathrm{B})_\rho$ follows when we require that $\mathrm{dim}\mathcal{H}_{\mathrm{B}'} \geq \mathrm{rank}\,\{\rho_{\mathrm{AB}}\} \cdot \mathrm{dim}\mathcal{H}_{\mathrm{A}} = \mathrm{dim}\mathcal{H}_{\mathrm{AC}}$. Then, $\mathcal{H}_{\mathrm{B}'}$ is large enough to accommodate all purifications of states in $\mathcal{H}_{\mathrm{AC}}$. Using Corollary 13 twice, we find

$$\begin{aligned}\hat{H}^{\varepsilon}_{\max}(\mathrm{A}|\mathrm{B})_\rho &= \min_{\tilde{\rho}_{\mathrm{AC}} \in \mathcal{B}^{\varepsilon}(\varepsilon)\rho_{\mathrm{AC}}} -\hat{H}_{\min}(\mathrm{A}|\mathrm{C})_{\tilde{\rho}}\\&= \min_{\tilde{\rho}_{\mathrm{AB}'\mathrm{C}} \in \mathcal{B}^{\varepsilon}_{\mathrm{p}}(\rho_{\mathrm{AB}'\mathrm{C}})} \hat{H}_{\max}(\mathrm{A}|\mathrm{B}')_{\tilde{\rho}}\\&\geq \min_{\tilde{\rho}_{\mathrm{AB}'} \in \mathcal{B}^{\varepsilon}(\rho_{\mathrm{AB}'})} \hat{H}_{\min}(\mathrm{A}|\mathrm{B}')_{\tilde{\rho}}.\end{aligned}$$

The infimum is, therefore, attained and it is sufficient to consider embeddings with $\mathrm{dim}\mathcal{H}_{\mathrm{B}'} = \dim \mathrm{supp}\{\rho_{\mathrm{AB}}\} \cdot \mathrm{dim}\mathcal{H}_{\mathrm{A}}$. ■

ACKNOWLEDGMENT

We thank Johan Åberg, Roger Colbeck and Christopher Portmann for useful discussions and comments.

REFERENCES

[1] D. R. Stinson, "Universal hash families and the leftover hash lemma, and applications to cryptography and computing," *J. Combin. Math. Combin. Comput.*, vol. 42, pp. 3–31, 2002.

[2] C. H. Bennett, G. Brassard, and J.-M. Robert, "Privacy amplification by public discussion," *SIAM J. Comput.*, vol. 17, no. 2, p. 210, 1988.

[3] R. Impagliazzo, L. A. Levin, and M. Luby, "Pseudo-random generation from one-way functions," in *Proc. ACM STOC*, 1989, pp. 12–24, ACM Press.

[4] C. H. Bennett, G. Brassard, C. Crepeau, and U. M. Maurer, "Generalized privacy amplification," *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pp. 1915–1923, Jun., 1995.

[5] R. Impagliazzo and D. Zuckerman, "How to recycle random bits," in *Proc. IEEE Symp. Found. Comp. Sci.*, 1989, pp. 248–253.

[6] J. Håstad, R. Impagliazzo, L. A. Levin, and M. Luby, "A pseudorandom generator from any one-way function," *SIAM J. Comput.*, vol. 28, no. 4, pp. 1364–1396, 1999.

[7] J. L. Carter and M. N. Wegman, "Universal classes of hash functions," *J. Comp. Syst. Sci.*, vol. 18, no. 2, pp. 143–154, 1979.

[8] D. R. Stinson, "Universal hashing and authentication codes," *Designs, Codes, Cryptogr.*, vol. 4, no. 3, pp. 369–380, Jul. 1994.

[9] D. Gavinsky, J. Kempe, I. Kerenidis, R. Raz, and R. de Wolf, "Exponential separation for one-way quantum communication complexity, with applications to cryptography," in *Proc. ACM STOC*, 2007, pp. 516–525, ACM Press.

[10] R. König and R. Renner, Sampling of Min-Entropy Relative to Quantum Knowledge [Online]. Available: http://arxiv.org/abs/0712.4291

[11] R. König, U. M. Maurer, and R. Renner, "On the power of quantum memory," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2391–2401, Jul. 2005.

[12] R. Renner, "Security of Quantum Key Distribution" Ph.D. dissertation, ETH, Zurich, Switzerland, Dec. 2005 [Online]. Available: http://arxiv.org/abs/quant-ph/0512258

[13] R. Renner and R. König, "Universally composable privacy amplification against quantum adversaries," in *Proc. TCC*, Cambridge, MA, 2005, vol. 3378, pp. 407–425, ser. LNCS, Springer.

[14] M. A. Nielsen and I. Chuang*, Quantum Computation and Quantum Information*. Cambridge, U.K.: Cambridge Univ. Press, 2000.

[15] Y. Dodis, R. Ostrovsky, L. Reyzin, and A. Smith, "Fuzzy extractors: How to generate strong keys from biometrics and other noisy data," *SIAM J. Comput.*, vol. 38, no. 1, pp. 97–139, 2008.

[16] R. König, R. Renner, and C. Schaffner, "The operational meaning of min- and max-entropy," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4337–4347, Sep. 2009.

[17] M. Tomamichel, R. Colbeck, and R. Renner, "Duality between smooth min- and max-entropies," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4674–4681, Sep. 2010.

[18] G. Van Assche*, Quantum Cryptography and Secret-Key Distillation*. Cambridge, U.K.: Cambridge Univ. Press, 2006.

[19] J. Lodewyck, M. Bloch, R. García-Patrón, S. Fossier, E. Karpov, E. Diamanti, T. Debuisschert, N. Cerf, R. Tualle-Brouri, S. McLaughlin, and P. Grangier, "Quantum key distribution over 25 km with an all-fiber continuous-variable system," *Phys. Rev. A*, vol. 76, no. 4, 2007.

[20] Y. Dodis and A. Smith, "Correcting errors without leaking partial information," in *Proc. ACM STOC*, 2005, pp. 654–663, ACM Press.

[21] S. Fehr and C. Schaffner, "Randomness extraction via delta-biased masking in the presence of a quantum attacker," in *Proc. TCC*, New York, 2008, vol. 4948, pp. 465–481, ser, LNCS, Springer.

[22] S. P. Desrosiers and F. Dupuis, "Quantum entropic security and approximate quantum encryption," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3455–3464, Jul. 2010.

[23] L. Trevisan, "Extractors and pseudorandom generators," *J. ACM*, vol. 48, no. 4, pp. 860–879, Jul. 2001.

[24] A. Ta-Shma, "Short seed extractors against quantum storage," in *Proc. ACM STOC*, Aug. 2009, pp. 401–408, ACM Press.

[25] A. De and T. Vidick, "Near-optimal extractors against quantum storage," in *Proc. ACM STOC*, Nov. 2010, pp. 161–170, ACM Press.

[26] A. De, C. Portmann, T. Vidick, and R. Renner, Trevisan's Extractor in the Presence of Quantum Side Information, Dec. 2009 [Online]. Available: http://arxiv.org/abs/0912.5514

[27] M. Tomamichel, R. Colbeck, and R. Renner, "A fully quantum asymptotic equipartition property," *IEEE Trans. Inf. Theory*, vol. 55, no. 12, pp. 5840–5847, Dec. 2009.

[28] R. Bhatia*, Matrix Analysis*, ser. Graduate Texts in Mathematics. New York: Springer, 1997.

[29] A. Srinivasan and D. Zuckerman, "Computing with very weak random sources," *SIAM J. Comput.*, vol. 28, no. 4, pp. 1433–1459, 1999.

[30] A. Uhlmann, "The transition probability for states of star-algebras," *Ann. Phys.*, vol. 497, no. 4, pp. 524–532, 1985.

[31] O. Klein, "Zur quantenmechanischen Begründung des zweiten Hauptsatzes der Wärmelehre," *Z. Phys.*, vol. 72, no. 11–12, pp. 767–775, Nov. 1931.

[32] N. Datta, "Min- and max-relative entropies and a new entanglement monotone," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2816–2826, Jun. 2009.

[33] M. Berta, "Single-Shot Quantum State Merging," M.S. thesis, ETH, Zurich, Switzerland, 2008.

[34] M. Mosonyi and N. Datta, "Generalized relative entropies and the capacity of classical-quantum channels," *J. Math. Phys.*, vol. 50, no. 7, p. 072104, 2009.

[35] F. Buscemi and N. Datta, "The quantum capacity of channels with arbitrarily correlated noise," *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp. 1447–1460, Mar. 2010.

Marco Tomamichel was born on March 13, 1981, in St. Gallen (Switzerland). He studied electrical engineering at ETH Zurich (Switzerland), where he graduated in 2007 with a M.Sc. in electrical engineering and information technology degree.

He is currently working on a Ph.D. thesis with the Institute of Theoretical Physics at ETH Zurich (Switzerland). His research interests include nonasymptotic quantum information theory and its applications.

Christian Schaffner received a diploma degree in mathematics from ETH Zurich (Switzerland) in 2003 and a Ph.D. degree in computer science from Århus University (Århus, Denmark) in 2007.

Currently, he holds a postdoctoral position at Centrum Wiskunde & Informatica (Amsterdam, The Netherlands). His research interests include quantum cryptography, cryptographic protocols, and (quantum) information theory.

Adam Smith received his Ph.D. in electrical engineering and computer science from MIT (Cambridge, MA, USA) in 2004.

He is currently an Associate Professor of Computer Science and Engineering at the Pennsylvania State University (University Park, PA, USA). Previously, he was a visiting scholar at the Weizmann Institute of Science and UCLA. His research interests lie in cryptography, data privacy and their connections to information theory, quantum computing and statistics.

Dr. Smith received a US Presidential Early Career Award for Scientists and Engineers (PECASE) in 2009. He has served as a reviewer for several IEEE journals and conferences and on the program committees of the IEEE conferences on Security and Privacy (2009) and on the Foundations of Computer Science (2011).

Renato Renner was born on December 11, 1974, in Lucerne (Switzerland). He studied physics, first at EPF Lausanne and later at ETH Zurich (Switzerland), where he graduated in 2000 with a Dipl. Phys. degree. He then moved to the Computer Science Department of ETH to work on a thesis in the area of quantum cryptography. He received his Ph.D. degree in 2005.

Between 2005 and 2007, he held a HP research fellowship in the Department for Applied Mathematics and Theoretical Physics at the University of Cambridge (United Kingdom). Since 2007, he is an Assistant Professor in Theoretical Physics at ETH Zurich. His research interests range from quantum information science to foundations of quantum mechanics.

Prof. Renner is a member of the American Physical Society. He has been a member of the technical program committee of ISIT 2008.