Randomness Extraction via δ-Biased Masking in the Presence of a Quantum Attacker

Serge Fehr[⋆] and Christian Schaffner[⋆⋆]

CWI[⋆⋆⋆] Amsterdam, The Netherlands
{S.Fehr,C.Schaffner}@cwi.nl

Abstract. Randomness extraction is of fundamental importance for information-theoretic cryptography. It allows to transform a raw key about which an attacker has some limited knowledge into a fully secure random key, on which the attacker has essentially no information. Up to date, only very few randomness-extraction techniques are known to work against an attacker holding quantum information on the raw key. This is very much in contrast to the classical (non-quantum) setting, which is much better understood and for which a vast amount of different techniques are known and proven to work.

We prove a new randomness-extraction technique, which is known to work in the classical setting, to be secure against a quantum attacker as well. Randomness extraction is done by xor'ing a so-called δ-biased mask to the raw key. Our result allows to extend the classical applications of this extractor to the quantum setting. We discuss the following two applications. We show how to encrypt a long message with a short key, information-theoretically secure against a quantum attacker, provided that the attacker has enough quantum uncertainty on the message. This generalizes the concept of entropically-secure encryption to the case of a quantum attacker. As second application, we show how to do error-correction without leaking partial information to a quantum attacker. Such a technique is useful in settings where the raw key may contain errors, since standard error-correction techniques may provide the attacker with information on, say, a secret key that was used to obtain the raw key.

1 Introduction

Randomness extraction allows to transform a raw key X about which an attacker has some limited knowledge into a fully secure random key S. It is required that the attacker has essentially no information on the resulting random key S, no

[⋆] Supported by a Veni grant from the Dutch Organization for Scientific Research (NWO).

[⋆⋆] Supported by the EU projects SECOQC and QAP IST 015848 and a NWO Vici grant 2004-2009.

[⋆⋆⋆] Centrum voor Wiskunde en Informatica, the national research institute for mathematics and computer science in the Netherlands.

matter what kind of information he has about the raw key X, as long as his uncertainty on X is lower bounded in terms of a suitable entropy measure. One distinguishes between extractors which use a private seed (preferably as small as possible) [29], and, what is nowadays called *strong* extractors, which only use public coins [15, 21]. In the context of cryptography, the latter kind of randomness extraction is also known as privacy amplification [5]. Randomness-extraction techniques play an important role in various areas of theoretical computer science. In cryptography, they are at the core of many constructions in information-theoretic cryptography, but they also proved to be useful in the computational setting. As such, there is a huge amount of literature on randomness extraction, and there exist various techniques which are optimized with respect to different needs; we refer to Shaltiel's survey [26] for an informative overview on classical and recent results.

Most of these techniques, however, are only guaranteed to work in a non-quantum setting, where information is formalized by means of classical information theory. In a quantum setting, where the attacker's information is given by a quantum state, our current understanding is much more deflating. Renner and König [23] have shown that privacy amplification via universal$_2$ hashing is secure against quantum adversaries. And, König and Terhal [18] showed security against quantum attackers for certain extractors, namely for one-bit-output strong extractors, as well as for strong extractors which work by extracting bit wise via one-bit-output strong extractors. Concurrent to our work, Smith has shown recently that Renner and König's result generalizes to *almost*-universal hashing, i.e., that Srinivasan-Zuckerman extractors remain secure against quantum adversaries [27]. On the negative side, Gavinsky *et al.* recently showed that there exist (strong) extractors that are secure against classical attackers, but which become completely insecure against quantum attackers [13]. Hence, it is not only a matter of lack of proof, but in fact classical extractors may turn insecure when considering *quantum* attackers.

We prove a new randomness-extraction technique to be secure against a quantum attacker. It is based on the concept of *small-biased spaces*, see e.g. [20]. Concretely, randomness extraction is done by xor'ing the raw key $X \in \{0,1\}^n$ with a δ-biased mask $A \in \{0,1\}^n$, chosen privately according to some specific distribution, where the distribution may be chosen publicly from some family of distributions. Roughly, A (or actually the family of distributions) is δ-biased, if any non-trivial parity of A can only be guessed with advantage δ. We prove that if A is δ-biased, then the bit-wise xor $X \oplus A$ is ε-close to random and independent of the attacker's quantum state with $\varepsilon = \delta \cdot 2^{(n-t)/2}$, where t is the attacker's quantum collision-entropy in X. Thus, writing $\delta = 2^{-\kappa}$, the extracted key $X \oplus A$ is essentially random as long as 2κ is significantly larger than $n - t$. Note that in its generic form, this randomness extractor uses public coins, namely the choice of the distribution, *and* a private seed, the sampling of A according to the chosen distribution. Specific instantiations though, may lead to standard extractors with no public coins (as in Section 5), or to a strong extractor with no private seed (as in Section 6). The proof of the new randomness-extraction result combines quantum-information-theoretic techniques developed by Renner [22, 23] and techniques from Fourier analysis, similar to though slightly more involved than those used in [2].

We would like to point out that the particular extractor we consider, δ-biased masking, is well known to be secure against *non*-quantum attackers. Indeed, classical security was shown by Dodis and Smith, who also suggested useful applications [11, 12]. Thus, our main contribution is the *security analysis* in the presence of a *quantum* attacker. Our positive result not only contributes to the general problem of the security of extractors against quantum attacks, but it is particularly useful in combination with the classical applications of δ-biased masking where it leads to interesting new results in the quantum setting. We discuss these applications and the arising new results below.

The first application is entropically secure encryption [25, 12]. An encryption scheme is entropically secure if the ciphertext gives essentially no information away on the plaintext (in an information-theoretic sense), provided that the attacker's a priori information on the plaintext is limited. Entropic security allows to overcome Shannon's pessimistic result on the size of the key for information-theoretically secure encryption, in that a key of size essentially $\ell \approx n - t$ suffices to encrypt a plaintext of size n which has t bits of entropy given the attacker's a priori information. This key size was known to suffice for a non-quantum adversary [25, 12]. By our analysis, this result carries over to the setting where we allow the attacker to store information as quantum states: a key of size essentially $\ell \approx n - t$ suffices to encrypt a plaintext of size n which has t bits of (min- or collision-) entropy given the attacker's quantum information about the plaintext.

Note that entropic security in a quantum setting was also considered explicitly in [8] and implicitly for the task of approximate quantum encryption [2, 16, 10]. However, all these results are on encrypting a *quantum* message into a quantum ciphertext on which the attacker has limited *classical* information (or none at all), whereas we consider encrypting a *classical* message into a classical ciphertext on which the attacker has limited *quantum* information. Thus, our result in quantum entropic security is in that sense orthogonal. As a matter of fact, the results in [2, 16, 10, 8] about randomizing quantum states can also be appreciated as extracting "quantum randomness" from a quantum state on which the attacker has limited *classical* information. Again, this is orthogonal to our randomness-extraction result which allows to extract classical randomness from a *classical* string on which the attacker has limited *quantum* information. In independent recent work, Desrosiers and Dupuis showed that one can combine techniques to get the best out of both: they showed that δ-biased masking (as used in [2]) allows to extract "quantum randomness" from a *quantum* state on which the attacker has limited *quantum* information. This in particular implies our result.

The second application is in the context of private error-correction. Consider a situation where the raw key X is obtained by Alice and Bob with the help of some (short) common secret key K, and where the attacker Eve, who does not know K, has high entropy on X. Assume that, due to noise, Bob's version of the raw key X' is slightly different from Alice's version X. Such a situation may for instance occur in the bounded-storage model or in a quantum-key-distribution setting. Since Alice and Bob have different versions of the raw key, they first need to correct the errors before they can extract (by means of randomness extraction) a secure key S from X. However, since X and X' depend on K, standard techniques for correcting the errors between X and X' leak information not only on X but also on K to Eve, which prohibits that Alice and Bob can re-use K in a future session. In the case of a non-quantum attacker, Dodis and Smith showed how to do error-correction in such a setting without leaking information on K to Eve [11], and thus that K can be safely re-used an unlimited number of times. We show how our randomness-extraction result gives rise to a similar way of doing error correction without leaking information on K, even if Eve holds her partial information on X in a quantum state. Such a private-error-correction technique is a useful tool in various information-theoretic settings with a quantum adversary. Very specifically, this technique has already been used as essential ingredient to derive new results in the bounded-(quantum)-storage model and in quantum key distribution [7].

The paper is organized as follows. We start with some quantum-information-theoretic notation and definitions. The new randomness-extraction result is presented in Section 3 and proven in Section 4. The two applications discussed are given in Sections 5 and 6.

2 Preliminaries

2.1 Notation and Terminology

A *quantum system* is described by a complex Hilbert space $\mathcal{H}_A$ (in this paper always of finite dimension). The *state* of the system is given by a *density matrix*: a positive semi-definite operator ρ_A on $\mathcal{H}_A$ with trace $\mathrm{tr}(\rho_A) = 1$. We write $\mathcal{P}(\mathcal{H}_A)$ for the set of all positive semi-definite operators on $\mathcal{H}_A$, and we call $\rho_A \in \mathcal{P}(\mathcal{H}_A)$ *normalized* if it has trace 1, i.e., if it is a density matrix. For a density matrix $\rho_{AB} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B)$ of a composite quantum system $\mathcal{H}_A \otimes \mathcal{H}_B$, we write $\rho_B = \mathrm{tr}_A(\rho_{AB})$ for the state obtained by tracing out system $\mathcal{H}_A$. A density matrix $\rho_{XB} \in \mathcal{P}(\mathcal{H}_X \otimes \mathcal{H}_B)$ is called *classical on* $\mathcal{H}_X$ *with* $X \in \mathcal{X}$, if it is of the form $\rho_{XB} = \sum_x P_X(x)|x\rangle\langle x| \otimes \rho_B^x$ with normalized $\rho_B^x \in \mathcal{P}(\mathcal{H}_B)$, where $\{|x\rangle\}_{x\in\mathcal{X}}$ forms an orthonormal basis of $\mathcal{H}_X$. Such a density matrix ρ_{XB} which is classical on $\mathcal{H}_X$ can be viewed as a random variable X with distribution P_X together with a family $\{\rho_B^x\}_{x\in\mathcal{X}}$ of *conditional density matrices*, such that the state of $\mathcal{H}_B$ is given by ρ_B^x if and only if X takes on the value x. We can introduce a new random variable Y which is obtained by "processing" X, i.e., by extending the distribution P_X to a consistent joint distribution P_{XY}. Doing so then naturally defines the density matrix $\rho_{XYB} = \sum_{x,y} P_{XY}(x,y)|x\rangle\langle x|\otimes|y\rangle\langle y|\otimes \rho_B^x$, and thus also the density matrix $\rho_{YB} = \mathrm{tr}_X(\rho_{XYB}) = \sum_y P_Y(y)|y\rangle\langle y| \otimes \big(\sum_x P_{X|Y}(x|y)\rho_B^x\big)$. If the meaning is clear from the context, we tend to slightly abuse notation and write the latter also as $\rho_{YB} = \sum_y P_Y(y)|y\rangle\langle y| \otimes \rho_B^y$, i.e., understand ρ_B^y as $\sum_x P_{X|Y}(x|y)\rho_B^x$. Throughout, we write $\mathbb{1}$ for the identity matrix of appropriate dimension.

2.2 Distance and Entropy Measures for Quantum States

We recall some definitions from [22]. Let $\rho_{XB} \in \mathcal{P}(\mathcal{H}_X \otimes \mathcal{H}_B)$. Although the following definitions make sense (and are defined in [22]) for arbitrary ρ_{XB}, we may assume ρ_{XB} to be normalized[1] and to be classical on $\mathcal{H}_X$.

Definition 2.1. *The* L_1-distance from uniform *of ρ_{XB} given B is defined by*

$$d(\rho_{XB}|B) := \|\rho_{XB} - \rho_U \otimes \rho_B\|_1 = \operatorname{tr}\big(|\rho_{XB} - \rho_U \otimes \rho_B|\big)$$

where $\rho_U := \frac{1}{\dim(\mathcal{H}_X)}\mathbb{1}$ is the fully mixed state on $\mathcal{H}_X$ and $|A| := \sqrt{A^\dagger A}$ is the positive square root of $A^\dagger A$ (where $A^\dagger$ is the complex-conjugate transpose of A).

If ρ_{XB} is classical on $\mathcal{H}_X$, then $d(\rho_{XB}|B) = 0$ if and only if X is uniformly distributed and ρ_B^x does not depend on x, which in particular implies that no information on X can be learned by observing system $\mathcal{H}_B$. Furthermore, if $d(\rho_{XB}|B) \leq \varepsilon$ then the real system ρ_{XB} "behaves" as the ideal system $\rho_U \otimes \rho_B$ except with probability ε in that for any evolution of the system no observer can distinguish the real from the ideal one with advantage greater than ε [23].

Definition 2.2. *The* collision-entropy *and the* min-entropy *of ρ_{XB}* relative to *a normalized and invertible $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$ are defined by*

$$\begin{aligned}
\mathrm{H}_2(\rho_{XB}|\sigma_B) &:= -\log \operatorname{tr}\left(\left((\mathbb{1} \otimes \sigma_B^{-1/4})\, \rho_{XB}\, (\mathbb{1} \otimes \sigma_B^{-1/4})\right)^2\right) \\
&= -\log \sum_x P_X(x)^2 \operatorname{tr}\left(\left(\sigma_B^{-1/4}\, \rho_B^x\, \sigma_B^{-1/4}\right)^2\right) \qquad \textit{and} \\
\mathrm{H}_\infty(\rho_{XB}|\sigma_B) &:= -\log \lambda_{max}\Big((\mathbb{1} \otimes \sigma_B^{-1/2})\, \rho_{XB}\, (\mathbb{1} \otimes \sigma_B^{-1/2})\Big) \\
&= -\log \max_x \lambda_{max}\Big(P_X(x)\, \sigma_B^{-1/2}\, \rho_B^x\, \sigma_B^{-1/2}\Big),
\end{aligned}$$

respectively, where $\lambda_{max}(\cdot)$ denotes the largest eigenvalue of the argument. The collision-entropy *and the* min-entropy *of ρ_{XB}* given $\mathcal{H}_B$ *are defined by*

$$\mathrm{H}_2(\rho_{XB}|B) := \sup_{\sigma_B} \mathrm{H}_2(\rho_{XB}|\sigma_B) \qquad \textit{and} \qquad \mathrm{H}_\infty(\rho_{XB}|B) := \sup_{\sigma_B} \mathrm{H}_\infty(\rho_{XB}|\sigma_B)$$

respectively, where the supremum ranges over all normalized $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$.

[1] For a non-normalized ρ_{XB}, there is a normalizing $1/\operatorname{tr}(\rho_{XB})$-factor in the definition of collision-entropy. Also note that $\operatorname{tr}(\sigma^{-1/2}\rho\sigma^{-1/2}) = \operatorname{tr}(\rho\sigma^{-1})$ for any invertible σ.

Note that without loss of generality, the supremum over σ_B can be restricted to the set of normalized *and invertible* states σ_B which is dense in the set of normalized states in $\mathcal{P}(\mathcal{H}_B)$. Note furthermore that it is not clear, neither in the classical nor in the quantum case, what the "right" way to define conditional collision- or min-entropy is, and as a matter of fact, it depends on the context which version serves best. An alternative way to define the collision- and min-entropy of ρ_{XB} given $\mathcal{H}_B$ would be as $\tilde{\mathrm{H}}_2(\rho_{XB}|B) := \mathrm{H}_2(\rho_{XB}|\rho_B)$ and $\tilde{\mathrm{H}}_\infty(\rho_{XB}|B) := \mathrm{H}_\infty(\rho_{XB}|\rho_B)$. For a density matrix ρ_{XY} that is classical on $\mathcal{H}_X$ and $\mathcal{H}_Y$, it is easy to see that $\tilde{\mathrm{H}}_2(\rho_{XY}|Y) = -\log \sum_y P_Y(y) \sum_x P_{X|Y}(x|y)^2$, i.e., the negative logarithm of the average conditional collision probability, and $\tilde{\mathrm{H}}_\infty(\rho_{XY}|Y) = -\log \max_{x,y} P_{X|Y}(x|y)$, i.e., the negative logarithm of the maximal conditional guessing probability. These notions of classical conditional collision- and min-entropy are commonly used in the literature, explicitly (see e.g. [24, 6]) or implicitly (as e.g. in [5]). We stick to Definition 2.2 because it leads to stronger results, in that asking $\mathrm{H}_2(\rho_{XB}|B)$ to be large is a weaker requirement than asking $\tilde{\mathrm{H}}_2(\rho_{XB}|B)$ to be large, as obviously $\mathrm{H}_2(\rho_{XB}|B) \geq \tilde{\mathrm{H}}_2(\rho_{XB}|B)$, and similarly for the min-entropy.

3 The New Randomness-Extraction Result

We start by recalling the definition of a δ-biased random variable and of a δ-biased family of random variables [20, 11].

Definition 3.1. *The* bias *of a random variable A, with respect to $\alpha \in \{0,1\}^n$, is defined as*

$$\mathrm{bias}_\alpha(A) := \sum_a P_A(a)(-1)^{\alpha \cdot a} = 2\big(P[\alpha \cdot A = 1] - \tfrac{1}{2}\big)\,,$$

and A is called δ-biased *if* $\mathrm{bias}_\alpha(A) \leq \delta$ *for all non-zero* $\alpha \in \{0,1\}^n$*. A family of random variables* $\{A_i\}_{i\in\mathcal{I}}$ *over* $\{0,1\}^n$ *is called* δ-biased *if, for all* $\alpha \neq 0$,

$$\sqrt{\mathbb{E}_{i\leftarrow\mathcal{I}}\big[\mathrm{bias}_\alpha(A_i)^2\big]} \leq \delta$$

where the expectation is over a i chosen uniformly at random from $\mathcal{I}$.

Note that by Jensen's inequality, $\mathbb{E}_{i\leftarrow\mathcal{I}}[\mathrm{bias}_\alpha(A_i)] \leq \delta$ for all non-zero α is a necessary (but not sufficient) condition for $\{A_i\}_{i\in\mathcal{I}}$ to be δ-biased. In case though the family consists of only one member, then it is δ-biased if and only if its only member is.

Our main theorem states that if $\{A_i\}_{i\in\mathcal{I}}$ is δ-biased for a small δ, and if an adversary's conditional entropy $\mathrm{H}_2(\rho_{XB}|B)$ on a string $X \in \{0,1\}^n$ is large enough, then masking X with A_i for a random but known i gives an essentially random string.

Theorem 3.2. *Let the density matrix $\rho_{XB} \in \mathcal{P}(\mathcal{H}_X \otimes \mathcal{H}_B)$ be classical on $\mathcal{H}_X$ with $X \in \{0,1\}^n$. Let $\{A_i\}_{i\in\mathcal{I}}$ be a δ-biased family of random variables over $\{0,1\}^n$, and let I be uniformly and independently distributed over $\mathcal{I}$. Then*

$$d\big(\rho_{(A_I \oplus X)BI}\big|BI\big) \leq \delta \cdot 2^{-\frac{1}{2}(\mathrm{H}_2(\rho_{XB}|B)-n)}.$$

By the inequalities

$$\mathrm{H}_\infty(X) - \log\dim(\mathcal{H}_B) \leq \mathrm{H}_\infty(\rho_{XB}|B) \leq \mathrm{H}_2(\rho_{XB}|B)\,,$$

proven in [22], Theorem 3.2 may also be expressed in terms of conditional min-entropy $\mathrm{H}_\infty(\rho_{XB}|B)$ or in terms of classical min-entropy of X minus the size of the quantum state (i.e. the number of qubits). If B is the "empty" quantum state, i.e., $\log\dim(\mathcal{H}_B) = 0$, then Theorem 3.2 coincides with Lemma 4 of [11]. Theorem 3.2 also holds, with a corresponding normalization factor, for non-normalized operators, from which it follows that it can also be expressed in terms of the *smooth* conditional min-entropy $\mathrm{H}^\varepsilon_\infty(\rho_{XB}|B)$, as defined in [22], as $d(\rho_{(A_I\oplus X)BI}|BI) \leq 2\varepsilon + \delta \cdot 2^{-\frac{1}{2}(\mathrm{H}^\varepsilon_\infty(\rho_{XB}|B)-n)}$.

4 The Proof

We start by pointing out some elementary observations regarding the Fourier transform over the hypercube. In particular, we can extend the Convolution theorem and Parseval's identity to the case of matrix-valued functions. Further properties of the Fourier transform (with a different normalization) of matrix-valued functions over the hypercube have recently been established by Ben-Aron, Regev and de Wolf [4]. In Section 4.2, we introduce and recall a couple of properties of the L_2-distance from uniform. The actual proof of Theorem 3.2 is given in Section 4.3.

4.1 Fourier Transform and Convolution

For some fixed positive integer d, consider the complex vector space $\mathcal{MF}$ of all functions $M : \{0,1\}^n \to \mathbb{C}^{d\times d}$. The *convolution* of two such matrix-valued functions $M, N \in \mathcal{MF}$ is the matrix-valued function

$$M * N : x \mapsto \sum_y M(y)N(x-y)$$

and the *Fourier transform* of a matrix-valued function $M \in \mathcal{MF}$ is the matrix-valued function

$$\mathfrak{F}(M) : \alpha \mapsto 2^{-n/2} \sum_x (-1)^{\alpha\cdot x} M(x)$$

where $\alpha \cdot x$ denotes the standard inner product modulo 2. Note that if X is a random variable with distribution P_X and M is the matrix-valued function $x \mapsto P_X(x) \cdot \mathbb{1}$, then

$$\mathfrak{F}(M)(\alpha) = 2^{-n/2} \cdot \mathrm{bias}_\alpha(X) \cdot \mathbb{1}\,.$$

The *Euclidean* or L_2-norm of a matrix-valued function $M \in \mathcal{MF}$ is given by

$$\|M\|_2 := \sqrt{\operatorname{tr}\Big(\sum_x M(x)^\dagger M(x)\Big)}$$

where $M(x)^\dagger$ denotes the complex-conjugate transpose of the matrix $M(x)$.[2]

The following two properties known as Convolution Theorem and Parseval's Theorem are straightforward to prove (see Appendix A).

Lemma 4.1. *For all $M, N \in \mathcal{MF}$:*

$$\mathfrak{F}(M * N) = 2^{n/2} \cdot \mathfrak{F}(M) \cdot \mathfrak{F}(N) \qquad \text{and} \qquad \|\mathfrak{F}(M)\|_2 = \|M\|_2 \,.$$

4.2 The L_2-Distance from Uniform

The following lemmas together with their proofs can be found in [22]. Again, we restrict ourselves to the case where ρ_{XB} and σ_B are normalized and ρ_{XB} is classical on X, whereas the claims hold (partly) more generally.

Definition 4.2. *Let $\rho_{XB} \in \mathcal{P}(\mathcal{H}_X \otimes \mathcal{H}_B)$ and $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$. Then the conditional L_2-distance from uniform of ρ_{XB} relative to σ_B is*

$$d_2(\rho_{XB}|\sigma_B) := \operatorname{tr}\left(\left((\mathbb{1} \otimes \sigma_B^{-1/4})(\rho_{XB} - \rho_U \otimes \rho_B)(\mathbb{1} \otimes \sigma_B^{-1/4})\right)^2\right),$$

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

Lemma 4.3. *Let $\rho_{XB} \in \mathcal{P}(\mathcal{H}_X \otimes \mathcal{H}_B)$. Then, for any normalized $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$,*

$$d(\rho_{XB}|B) \leq \sqrt{\dim(\mathcal{H}_X)}\sqrt{d_2(\rho_{XB}|\sigma_B)}.$$

Lemma 4.4. *Let $\rho_{XB} \in \mathcal{P}(\mathcal{H}_X \otimes \mathcal{H}_B)$ be classical on $\mathcal{H}_X$ with $X \in \mathcal{X}$, and let ρ_B^x be the corresponding normalized conditional operators. Then, for any $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$*

$$d_2(\rho_{XB}|\sigma_B) = \sum_x \operatorname{tr}\left((\sigma_B^{-1/4} P_X(x)\rho_B^x \sigma_B^{-1/4})^2\right) - \frac{1}{|\mathcal{X}|}\operatorname{tr}\left((\sigma_B^{-1/4}\rho_B\sigma_B^{-1/4})^2\right).$$

4.3 Proof Theorem 3.2

Write $D_i = A_i \oplus X$ and $D_I = A_I \oplus X$. Since $\rho_{D_I BI} = \frac{1}{|\mathcal{I}|}\sum_i \rho^i_{D_I B} \otimes |i\rangle\langle i| = \frac{1}{|\mathcal{I}|}\sum_i \rho_{D_i B} \otimes |i\rangle\langle i|$, and similar for ρ_{BI}, it follows that the L_1-distance from uniform can be written as an expectation over the random choice of i from $\mathcal{I}$. Indeed

$$d(\rho_{D_I BI}|BI) = \frac{1}{|\mathcal{I}|}\operatorname{tr}\Big(\Big|\sum_i (\rho_{D_i B} - \rho_U \otimes \rho_B) \otimes |i\rangle\langle i|\Big|\Big)$$

[2] We will only deal with Hermitian matrices $M(x)$ where $\|M\|_2 = \sqrt{\operatorname{tr}\left(\sum_x M(x)^2\right)}$.

$$= \frac{1}{|\mathcal{I}|} \sum_i \operatorname{tr}\big(\big|\rho_{D_iB} - \rho_U \otimes \rho_B\big|\big) = \frac{1}{|\mathcal{I}|} \sum_i d(\rho_{D_iB}|B) = \mathbb{E}_{i \leftarrow \mathcal{I}}\big[d(\rho_{D_iB}|B)\big] \, .$$

where the second equality follows from the block-diagonal form of the matrix. With Lemma 4.3, the term in the expectation can be bounded in terms of the L_2-distance from uniform, that is, for any normalized $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$,

$$d(\rho_{D_IBI}|BI) \leq \sqrt{2^n}\, \mathbb{E}_{i \leftarrow \mathcal{I}}\Big[\sqrt{d_2(\rho_{D_iB}|\sigma_B)}\Big] \leq 2^{n/2} \sqrt{\mathbb{E}_{i \leftarrow \mathcal{I}}\big[d_2(\rho_{D_iB}|\sigma_B)\big]}$$

where the second inequality is Jensen's inequality. By Lemma 4.4, we have for the L_2-distance

$$\begin{aligned} & d_2(\rho_{D_iB}|\sigma_B) \\ & = \operatorname{tr}\left(\sum_d (\sigma_B^{-1/4}\, P_{D_i}(d)\rho_B^d\, \sigma_B^{-1/4})^2\right) - \frac{1}{2^n} \operatorname{tr}\left((\sigma_B^{-1/4}\, \rho_B\, \sigma_B^{-1/4})^2\right) \, . \end{aligned} \tag{1}$$

Note that

$$\begin{aligned} P_{D_i}(d)\rho_B^d &= P_{D_i}(d) \sum_x P_{X|D_i}(x|d)\rho_B^x = \sum_x P_{XD_i}(x,d)\rho_B^x \\ &= \sum_x P_{XA_i}(x, d \oplus x)\rho_B^x = \sum_x P_X(x) P_{A_i}(d \oplus x)\rho_B^x \end{aligned}$$

so that the first term on the right-hand side of (1) can be written as

$$\begin{aligned} & \operatorname{tr}\left(\sum_d (\sigma_B^{-1/4}\, P_{D_i}(d)\rho_B^d\, \sigma_B^{-1/4})^2\right) \\ & = \operatorname{tr}\left(\sum_d \left(\sum_x P_X(x)\sigma_B^{-1/4}\, \rho_B^x\, \sigma_B^{-1/4} P_{A_i}(d \oplus x)\right)^2\right) \, . \end{aligned}$$

The crucial observation now is that the term that is squared on the right side is the convolution of the two matrix-valued functions $M : x \mapsto P_X(x)\sigma_B^{-1/4}\rho_B^x\sigma_B^{-1/4}$ and $N : x \mapsto P_{A_i}(x)\mathbb{1}$, and the whole expression equals $\|\!|M * N|\!\|_2^2$. Thus, using Lemma 4.1 we get

$$\begin{aligned} \operatorname{tr}\left(\sum_d (\sigma_B^{-1/4}\, P_{D_i}(d)\rho_B^d\, \sigma_B^{-1/4})^2\right) &= \|\!|M * N|\!\|_2^2 = \|\!|\mathfrak{F}(M * N)|\!\|_2^2 \\ &= \|\!|2^{n/2} \cdot \mathfrak{F}(M) \cdot \mathfrak{F}(N)|\!\|_2^2 = 2^n \operatorname{tr}\left(\sum_\alpha \big(\mathfrak{F}(M)(\alpha)\mathfrak{F}(N)(\alpha)\big)^2\right) \\ &= \frac{1}{2^n} \operatorname{tr}\left((\sigma_B^{-1/4}\, \rho_B\, \sigma_B^{-1/4})^2\right) + \operatorname{tr}\left(\sum_{\alpha \neq 0} \mathfrak{F}(M)(\alpha)^2 \operatorname{bias}_\alpha(A_i)^2\right) \, , \end{aligned} \tag{2}$$

where the last equality uses

$$\mathfrak{F}(M)(0) = 2^{-n/2} \sum_x P_X(x) \sigma_B^{-1/4}\, \rho_B^x\, \sigma_B^{-1/4} = 2^{-n/2} \sigma_B^{-1/4}\, \rho_B\, \sigma_B^{-1/4}$$

as well as

$$\mathfrak{F}(N)(0) = 2^{-n/2} \sum_x P_{A_i}(x) \mathbb{1} = 2^{-n/2} \mathbb{1} \quad \text{and} \quad \mathfrak{F}(N)(\alpha) = 2^{-n/2} \cdot \mathrm{bias}_\alpha(A_i) \mathbb{1}\,.$$

Substituting (2) into (1) gives

$$d_2(\rho_{D_i B}|\sigma_B) = \mathrm{tr}\left(\sum_{\alpha \neq 0} \mathfrak{F}(M)(\alpha)^2\, \mathrm{bias}_\alpha(A_i)^2 \right)\,.$$

Using the linearity of the expectation and trace, and using the bound on the expected square-bias, we get

$$\begin{aligned}
\mathbb{E}_{i \leftarrow \mathcal{I}}\big[d_2(\rho_{D_i B}|\sigma_B)\big] &\le \delta^2\, \mathrm{tr}\Big(\sum_{\alpha \neq 0} \mathfrak{F}(M)(\alpha)^2 \Big) \le \delta^2\, \mathrm{tr}\Big(\sum_{\alpha} \mathfrak{F}(M)(\alpha)^2 \Big) \\
&= \delta^2 |\!|\!|\mathfrak{F}(M)|\!|\!|_2^2 = \delta^2 |\!|\!|M|\!|\!|_2^2 = \delta^2 \sum_x \mathrm{tr}\left(P_X(x)^2 (\sigma_B^{-1/4}\, \rho_B^x\, \sigma_B^{-1/4})^2 \right) \\
&= \delta^2 2^{-\mathrm{H}_2(\rho_{XB}|\sigma_B)}\,,
\end{aligned}$$

where the second inequality follows because of

$$\mathrm{tr}\big(\mathfrak{F}(M)(0)^2\big) = 2^{-n}\, \mathrm{tr}\big((\sigma_B^{-1/4}\, \rho_B\, \sigma_B^{-1/4})^2\big) \ge 0\,.$$

Therefore,

$$d(\rho_{D_I BI}|BI) \le 2^{n/2} \sqrt{\mathbb{E}_{i \leftarrow \mathcal{I}}\big[d_2(\rho_{D_i B}|\sigma_B)\big]} \le \delta \cdot 2^{-\frac{1}{2}(\mathrm{H}_2(\rho_{XB}|\sigma_B) - n)}$$

and the assertion follows from the definition of $H_2(\rho_{XB}|B)$ because σ_B was arbitrary. □

5 Application I: Entropic Security

Entropic security is a relaxed but still meaningful security definition for (information-theoretically secure) encryption that allows to circumvent Shannon's pessimistic result, which states that any perfectly secure encryption scheme requires a key at least as long as the message to be encrypted. Entropic security was introduced by Russell and Wang [25], and later more intensively investigated by Dodis and Smith [12]. Based on our result, and in combination with techniques from [12], we show how to achieve entropic security against quantum adversaries. We would like to stress that in contrast to perfect security e.g. when using the one-time-pad, entropic security does *not* a priori protect against a quantum adversary.

Informally, entropic security is defined as follows. An encryption scheme is entropically secure if no adversary can obtain any information on the message M from its ciphertext C (in addition to what she can learn from scratch), provided the message M has enough uncertainty from the adversary's point of view. The impossibility of obtaining any information on M is formalized by requiring that any adversary that can compute $f(M)$ for some function f when given C, can also compute $f(M)$ *without* C (with similar success probability). A different formulation, which is named *indistinguishability*, is to require that there exists a random variable C', independent of M, such that C and C' are essentially identically distributed. It is shown in [12], and in [8] for the case of a *quantum* message, that the two notions are equivalent if the adversary's information on M is classical. In recent work, Desrosiers and Dupuis proved this equivalence to hold also for an adversary with quantum information [9].

The adversary's uncertainty on M is formalized, for a *classical* (i.e. non-quantum) adversary, by the *min-entropy* $\mathrm{H}_\infty(M|V=v)$ (or, alternatively, the collision-entropy) of M, conditioned on the value v the adversary's view V takes on. We formalize this uncertainty for a quantum adversary in terms of the quantum version of conditional min- or actually collision-entropy, as introduced in Section 2.2.

Definition 5.1. *We call a (possibly randomized) encryption scheme* $E : \mathcal{K} \times \mathcal{M} \to \mathcal{C}$ (t, ε)-quantum-indistinguishable *if there exists a random variable* C' *over* $\mathcal{C}$ *such that for any normalized* $\rho_{MB} \in \mathcal{P}(\mathcal{H}_M \otimes \mathcal{H}_B)$ *which is classical on* $\mathcal{H}_M$ *with* $M \in \mathcal{M}$ *and* $\mathrm{H}_2(\rho_{MB}|B) \geq t$*, we have that*

$$\left\| \rho_{E(K,M)B} - \rho_{C'} \otimes \rho_B \right\|_1 \leq \varepsilon \,,$$

where K *is uniformly and independently distributed over* $\mathcal{K}$*.*

Note that in case of an "empty" state B, our definition coincides with the indistinguishability definition from [12] (except that we express it in collision- rather than min-entropy).

Theorem 3.2, with $\mathcal{I} = \{i_\circ\}$ and $A_{i_\circ} = K$, immediately gives a generic construction for a quantum-indistinguishable encryption scheme (with C' being uniformly distributed). Independently, this result was also obtained in [9].

Theorem 5.2. *Let* $\mathcal{K} \subseteq \{0,1\}^n$ *be such that the uniform distribution* K *over* $\mathcal{K}$ *is* δ*-biased. Then the encryption scheme* $E : \mathcal{K} \times \{0,1\}^n \to \{0,1\}^n$ *with* $E(k,m) = k \oplus m$ *is* (t, ε)*-quantum-indistinguishable with* $\varepsilon = \delta \cdot 2^{\frac{n-t}{2}}$*.*

Alon *et al.* [1] showed how to construct subsets $\mathcal{K} \subseteq \{0,1\}^n$ of size $|\mathcal{K}| = O(n^2/\delta^2)$ such that the uniform distribution K over $\mathcal{K}$ is δ-biased and elements in $\mathcal{K}$ can be efficiently sampled. With the help of this construction, we get the following result, which generalizes the bound on the key-size obtained in [12] to the quantum setting.

Corollary 5.3. *For any* $\varepsilon \geq 0$ *and* $0 \leq t \leq n$*, there exists a* (t, ε)*-quantum-indistinguishable encryption scheme encrypting* n*-bit messages with key length* $\ell = \log|\mathcal{K}| = n - t + 2\log(n) + 2\log(\frac{1}{\varepsilon}) + O(1)$.

In the language of extractors, defining a (t, ε)-*quantum extractor* in the natural way as follows, Corollary 5.3 translates to Corollary 5.5 below.

Definition 5.4. *A function* $E : \mathcal{J} \times \mathcal{X} \to \{0,1\}^m$ *is called a* (t, ε)-weak quantum extractor *if* $d(\rho_{E(J,X)B}|B) \leq \varepsilon$*, and a* (t, ε)-strong quantum extractor *if* $d(\rho_{E(J,X)JB}|JB) \leq \varepsilon$ *for any normalized* $\rho_{XB} \in \mathcal{P}(\mathcal{H}_X \otimes \mathcal{H}_B)$ *which is classical on* $\mathcal{H}_X$ *with* $X \in \mathcal{X}$ *and* $\mathrm{H}_2(\rho_{XB}|B) \geq t$*, and where* J *is uniformly and independently distributed over* $\mathcal{J}$.

Corollary 5.5. *For any* $\varepsilon \geq 0$ *and* $0 \leq t \leq n$*, there exists a* (t, ε)*-weak quantum extractor with* n*-bit output and seed length* $\ell = \log|\mathcal{K}| = n - t + 2\log(n) + 2\log(\frac{1}{\varepsilon}) + O(1)$.

6 Application II: Private Error Correction

Consider the following scenario. Two parties, Alice and Bob, share a common secret key K. Furthermore, we assume a "random source" which can be queried by Alice and Bob so that on identical queries it produces identical outputs. In particular, when Alice and Bob both query the source on input K, they both obtain the same "raw key" $X \in \{0,1\}^n$. We also give an adversary Eve access to the source. She can obtain some (partial) information on the source and store it possibly in a quantum state ρ_Z. However, we assume she has some uncertainty about X, because due to her ignorance of K, she is unable to extract "the right" information from the source. Such an assumption of course needs to be justified in a specific implementation. Specifically, we require that $\mathrm{H}_\infty(\rho_{XKZ}|KZ)$ is lower bounded, i.e., Eve has uncertainty in X even if at some later point she learns K but only the source has disappeared in the meantime.

Such a scenario for instance arises in the bounded-storage model [19, 3] (though with classical Eve), when K is used to determine which bits of the long randomizer Alice and Bob should read to obtain X, or in a quantum setting when Alice sends n qubits to Bob and K influences the basis in which Alice prepares them respectively Bob measures them.

In this setting, it is well-known how to transform by public (authenticated) communication the weakly-secure raw key X into a fully secure key S: Alice and Bob do privacy amplification, as shown in [14, 5] in case of a classical Eve, respectively as in [23, 22] in case of a quantum Eve. Indeed, under the above assumptions on the entropy of X, privacy amplification guarantees that the resulting key S looks essentially random for Eve even given K. This guarantee implies that S can be used, say, as a one-time-pad encryption key, but it also implies that if Eve learns S, she still has essentially no information on K, and thus K can be safely re-used for the generation of a new key S.

Consider now a more realistic scenario, where due to noise or imperfect measurements Alice's string X and Bob's string X' are close but not exactly equal. There are standard techniques to do error correction (without giving Eve too much information on X): Alice and Bob agree on a suitable error-correcting code $\mathcal{C}$, Alice samples a random codeword C from $\mathcal{C}$ and sends $Y = X \oplus C$ to Bob, who can recover X by decoding $C' = Y \oplus X'$ to the nearest codeword C and compute $X = Y \oplus C$. Or equivalently, in case of a linear code, Alice can send the syndrome of X to Bob, which allows Bob to recover X in a similar manner. If Eve's entropy in X is significantly larger than the size of the syndrome, then one can argue that privacy amplification still works and the resulting key S is still (close to) random given Eve's information (including the syndrome) and K. Thus, S is still a secure key. However, since X depends on K, and the syndrome of X depends on X, the syndrome of X may give information on K to Eve, which makes it insecure to re-use K. A common approach to deal with this problem is to use part of S as the key K in the next session. Such an approach not only creates a lot of inconvenience for Alice and Bob in that they now have to be stateful and synchronized, but in many cases Eve can prevent Alice and Bob from agreeing on a secure key S (for instance by blocking the last message) while nevertheless learning information on K, and thus Eve can still cause Alice and Bob to run out of key material.

In [11], Dodis and Smith addressed this problem and proposed an elegant solution in case of a classical Eve. They constructed a family of codes which not only allow to efficiently correct errors, but at the same time also serve as randomness extractors. More precisely, they show that for every $0 < \lambda < 1$, there exists a family $\{\mathcal{C}_j\}_{j \in \mathcal{J}}$ of binary linear codes of length n, which allows to efficiently correct a constant fraction of errors, and which is δ-biased for $\delta < 2^{-\lambda n/2}$. The latter is to be understood that the family $\{C_j\}_{j \in \mathcal{J}}$ of random variables, where C_j is uniformly distributed over $\mathcal{C}_j$, is δ-biased for $\delta < 2^{-\lambda n/2}$. Applying Lemma 4 of [11] (the classical version of Theorem 3.2) implies that $C_j \oplus X$ is close to random for any X with large enough entropy, given j. Similarly, applying our Theorem 3.2 implies the following.

Theorem 6.1. *For every $0 < \lambda < 1$ there exists a family $\{\mathcal{C}_j\}_{j \in \mathcal{J}}$ of binary linear codes of length n which allows to efficiently correct a constant fraction of errors, and such that for any density matrix $\rho_{XB} \in \mathcal{P}(\mathcal{H}_X \otimes \mathcal{H}_B)$ which is classical on $\mathcal{H}_X$ with $X \in \{0,1\}^n$ and $\mathrm{H}_2(\rho_{XB}|B) \geq t$, it holds that*

$$d\big(\rho_{(C_J \oplus X)BJ}\big|BJ\big) \leq 2^{-\frac{t-(1-\lambda)n}{2}} ,$$

where J is uniformly distributed over $\mathcal{J}$ and C_J is uniformly distributed over $\mathcal{C}_J$.

Using a random code from such a family of codes allows to do error correction in the noisy setting described above without leaking information on K to Eve: By the chain rule [22, Sect. 3.1.3], the assumed lower bound on $\mathrm{H}_\infty(\rho_{XKZ}|KZ)$ implies a lower bound on $\mathrm{H}_\infty(\rho_{XSKZG}|SKZG)$ (essentially the original bound minus the bit length of S), where G is the randomly chosen universal hash function used to extract S from X. Combining systems S, K, Z and G into system B, Theorem 6.1 implies that $\rho_{(C_J \oplus X)SKZGJ} \approx \frac{1}{2^n}\mathbb{1} \otimes \rho_{SKZGJ}$. From standard privacy amplification follows that $\rho_{SKZGJ} \approx \frac{1}{2^\ell}\mathbb{1} \otimes \rho_{KZGJ}$. Using the independence of K, G, J (from Z and from each other), we obtain $\rho_{(C_J \oplus X)SKZGJ} \approx$

$\frac{1}{2^n}\mathbb{1} \otimes \frac{1}{2^\ell}\mathbb{1} \otimes \rho_K \otimes \rho_Z \otimes \rho_G \otimes \rho_J$. This in particular implies that S is a secure key (even when K is given to Eve) and that K is still "fresh" and can be safely re-used (even when S is additionally given to Eve).

Specifically, our private-error-correction techniques allow to add robustness against noise to the bounded-storage model in the presence of a quantum attacker as considered in [17], without the need for updating the common secret key. The results of [17] guarantee that the min-entropy of the sampled substring is lower bounded given the attacker's quantum information and hence, security follows as outlined above. Furthermore, in [7] the above private-error-correction technique is an essential ingredient to add robustness against noise but also to protect against man-in-the-middle attacks in new quantum-identification and quantum-key-distribution schemes in the bounded-quantum-storage model.

In the language of extractors, we get the following result for arbitrary, not necessarily efficiently decodable, binary linear codes.

Corollary 6.2. *Let $\{\mathcal{C}_j\}_{j\in\mathcal{J}}$ be a δ-biased family of binary linear $[n,k,d]_2$-codes. For any $j \in \mathcal{J}$, let G_j be a generator matrix for the code $\mathcal{C}_j$ and let H_j be a corresponding parity-check matrix. Then $E : \mathcal{J} \times \{0,1\}^n \to \{0,1\}^{n-k}$, $(j,x) \mapsto H_j x$ is a (t,ε)-strong quantum extractor with $\varepsilon = \delta \cdot 2^{\frac{1}{2}(n-t)}$.*

This result gives rise to new privacy-amplification techniques, beyond using universal$_2$ hashing as in [23] or one-bit extractors as in [18]. Note that using arguments from [11], it is easy to see that the condition that $\{\mathcal{C}_j\}_{j\in\mathcal{J}}$ is δ-biased and thus the syndrome function H_j is a good strong extractor, is equivalent to requiring that $\{G_j\}_{j\in\mathcal{J}}$ seen as family of (encoding) functions is δ^2-almost universal$_2$ [30, 28].

For a family of binary linear codes $\{\mathcal{C}_j\}_{j\in\mathcal{J}}$, another equivalent condition for δ-bias of $\{\mathcal{C}_j\}_{j\in\mathcal{J}}$ is to require that for all non-zero α, $\Pr_{j\in\mathcal{J}}[\alpha \in \mathcal{C}_j^\perp] \leq \delta^2$, i.e. that the probability that α is in the dual code of $\mathcal{C}_j$ is upper bounded by δ^2 [11]. It follows that the family size $|\mathcal{J}|$ has to be exponential in n to achieve an exponentially small bias δ and therefore, the seed length $\log|\mathcal{J}|$ of the strong extractor will be linear in n as for the case of two-universal hashing.

7 Conclusion

We proposed a new technique for randomness extraction in the presence of a quantum attacker. This is interesting in its own right, as up to date only very few extractors are known to be secure against quantum adversaries, much in contrast to the classical non-quantum case. The new randomness-extraction technique has various cryptographic applications like entropically secure encryption, in the classical bounded-storage model and the bounded-quantum-storage model, and in quantum key distribution. Furthermore, because of the wide range of applications of classical extractors not only in cryptography but also in other areas of theoretical computer science, we feel that our new randomness-extraction technique will prove to be useful in other contexts as well.

Acknowledgments

We would like to thank Ivan Damgård, Renato Renner, and Louis Salvail for helpful discussions and the anonymous referees for useful comments.

References

1. N. Alon, O. Goldreich, J. Håstad, and R. Peralta. Simple constructions of almost k-wise independent random variables. In *31st Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, volume II, pages 544–553, 1990.
2. A. Ambainis and A. Smith. Small pseudo-random families of matrices: Derandomizing approximate quantum encryption. In K. Jansen, S. Khanna, J. D. P. Rolim, and D. Ron, editors, *Approximation Algorithms for Combinatorial Optimization Problems, APPROX 2004, and 8th International Workshop on Randomization and Computation, RANDOM 2004*, volume 3122 of *Lecture Notes in Computer Science*, pages 249–260. Springer, 2004.
3. Y. Aumann, Y. Z. Ding, and M. O. Rabin. Everlasting security in the bounded storage model. *IEEE Transactions on Information Theory*, 48(6):1668–1680, June 2002.
4. A. Ben-Aroya, O. Regev, and R. de Wolf. A hypercontractive inequality for matrix-valued functions with applications to quantum computing. `http://arxiv.org/abs/0705.3806`, 2007.
5. C. H. Bennett, G. Brassard, C. Crépeau, and U. M. Maurer. Generalized privacy amplification. *IEEE Transactions on Information Theory*, 41:1915–1923, Nov. 1995.
6. I. B. Damgård, S. Fehr, L. Salvail, and C. Schaffner. Oblivious transfer and linear functions. In *Advances in Cryptology—CRYPTO '06*, volume 4117 of *Lecture Notes in Computer Science*, pages 427–444. Springer, 2006.
7. I. B. Damgård, S. Fehr, L. Salvail, and C. Schaffner. Secure identification and QKD in the bounded-quantum-storage model. In *Advances in Cryptology—CRYPTO '07*, volume 4622 of *Lecture Notes in Computer Science*, pages 342–359. Springer, 2007.
8. S. P. Desrosiers. Entropic security in quantum cryptography. `http://arxiv.org/abs/quant-ph/0703046`, 2007.
9. S. P. Desrosiers and F. Dupuis. Quantum entropic security and approximate quantum encryption. `http://arxiv.org/abs/0707.0691`, July 5, 2007.
10. P. A. Dickinson and A. Nayak. Approximate randomization of quantum states with fewer bits of key. In *Quantum Computing: Back Action 2006*, volume 864 of *American Institute of Physics Conference Series*, pages 18–36, November 2006. `quant-ph/0611033`.
11. Y. Dodis and A. Smith. Correcting errors without leaking partial information. In *37th Annual ACM Symposium on Theory of Computing (STOC)*, pages 654–663, 2005.
12. Y. Dodis and A. Smith. Entropic security and the encryption of high entropy messages. In *Theory of Cryptography Conference (TCC)*, volume 3378 of *Lecture Notes in Computer Science*, pages 556–577. Springer, 2005.
13. D. Gavinsky, I. Kerenidis, J. Kempe, R. Raz, and R. de Wolf. Exponential separations for one-way quantum communication complexity, with applications to cryptography. In *39th Annual ACM Symposium on Theory of Computing (STOC)*, pages 516–525, 2007. `http://arxiv.org/abs/quant-ph/0611209`.

14. J. Håstad, R. Impagliazzo, L. A. Levin, and M. Luby. A pseudorandom generator from any one-way function. *SIAM Journal on Computing*, 28(4), 1999.
15. R. Impagliazzo, L. A. Levin, and M. Luby. Pseudo-random generation from one-way functions. In *21st Annual ACM Symposium on Theory of Computing (STOC)*, pages 12–24, 1989.
16. I. Kerenidis and D. Nagaj. On the optimality of quantum encryption schemes. *Journal of Mathematical Physics*, 47:092102, 2006. `http://arxiv.org/abs/quant-ph/0509169`.
17. R. König and R. Renner. Sampling of min-entropy relative to quantum knowledge. In *Workshop on Quantum Information Processing (QIP 2008)*, 2007.
18. R. König and B. M. Terhal. The bounded storage model in the presence of a quantum adversary. `http://arxiv.org/abs/quant-ph/0608101`, 2006.
19. U. M. Maurer. A provably-secure strongly-randomized cipher. In *Advances in Cryptology—EUROCRYPT '90*, volume 473 of *Lecture Notes in Computer Science*, pages 361–373. Springer, 1990.
20. J. Naor and M. Naor. Small-bias probability spaces: efficient constructions and applications. In *22nd Annual ACM Symposium on Theory of Computing (STOC)*, pages 213–223, 1990.
21. N. Nisan and D. Zuckerman. More deterministic simulation in logspace. In *25th Annual ACM Symposium on the Theory of Computing (STOC)*, pages 235–244, 1993.
22. R. Renner. *Security of Quantum Key Distribution*. PhD thesis, ETH Zürich (Switzerland), September 2005. `http://arxiv.org/abs/quant-ph/0512258`.
23. R. Renner and R. König. Universally composable privacy amplification against quantum adversaries. In *Theory of Cryptography Conference (TCC)*, volume 3378 of *Lecture Notes in Computer Science*, pages 407–425. Springer, 2005.
24. R. Renner and S. Wolf. Simple and tight bounds for information reconciliation and privacy amplification. In *Advances in Cryptology—ASIACRYPT 2005*, Lecture Notes in Computer Science, pages 199–216. Springer, 2005.
25. A. Russell and H. Wang. How to fool an unbounded adversary with a short key. In *Advances in Cryptology—EUROCRYPT '02*, volume 2332 of *Lecture Notes in Computer Science*, pages 133–148. Springer, 2002.
26. R. Shaltiel. Recent developments in explicit constructions of extractors. *Bulletin of the EATCS*, 77:67–95, 2002.
27. A. Smith, 2007. Private communication.
28. D. R. Stinson. Universal hashing and authentication codes. In *Advances in Cryptology—CRYPTO '91*, volume 576 of *Lecture Notes in Computer Science*, pages 74–85. Springer, 1991.
29. A. Ta-Shma. On extracting randomness from weak random sources. In *28th Annual ACM Symposium on the Theory of Computing (STOC)*, pages 276–285, 1996.
30. M. N. Wegman and L. Carter. New hash functions and their use in authentication and set equality. *J. Comput. Syst. Sci.*, 22(3):265–279, 1981.

A Proof of Lemma 4.1

Concerning the first claim,

$$\mathfrak{F}(M * N)(\alpha) = \frac{1}{2^{n/2}} \sum_x (-1)^{\alpha \cdot x} \sum_y M(y) N(x \oplus y)$$

$$
\begin{aligned}
&= 2^{-n/2} \sum_y (-1)^{\alpha \cdot y} M(y) \sum_x (-1)^{\alpha \cdot (x \oplus y)} N(x \oplus y) \\
&= 2^{-n/2} \sum_y (-1)^{\alpha \cdot y} M(y) \sum_z (-1)^{\alpha \cdot z} N(z) \\
&= 2^{n/2} \cdot \mathfrak{F}(M)(\alpha) \cdot \mathfrak{F}(N)(\alpha) \,.
\end{aligned}
$$

The second claim is argued as follows.

$$
\begin{aligned}
\|\mathfrak{F}(M)\|_2^2 &= \operatorname{tr} \left(\sum_\alpha \mathfrak{F}(M)(\alpha)^\dagger \mathfrak{F}(M)(\alpha) \right) \\
&= 2^{-n} \operatorname{tr} \left(\sum_\alpha \left(\sum_x (-1)^{\alpha \cdot x} M(x) \right)^* \left(\sum_{x'} (-1)^{\alpha \cdot x'} M(x') \right) \right) \\
&= 2^{-n} \operatorname{tr} \left(\sum_{x,x'} M(x)^\dagger M(x') \sum_\alpha (-1)^{\alpha \cdot (x \oplus x')} \right) \\
&= \operatorname{tr} \left(\sum_x M(x)^\dagger M(x) \right) = \|M\|_2^2
\end{aligned}
$$

where the last equality follows from the fact that $\sum_\alpha (-1)^{\alpha \cdot y} = 2^n$ if $y = (0, \ldots, 0)$ and 0 otherwise. □