

Trevisan's extractor in the presence of quantum side information

Anindya De^{*1}, Christopher Portmann^{†2}, Thomas Vidick^{‡1}, and Renato Renner^{§3}

¹Computer Science Division, University of California, Berkeley, CA, USA.

²Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8552, Japan.

³Institute for Theoretical Physics, ETH Zurich, 8093 Zurich, Switzerland.

December 30, 2009

Abstract

Randomness extraction involves the processing of purely classical information and is therefore usually studied in the framework of classical probability theory. However, such a classical treatment is generally too restrictive for applications, where side information about the values taken by classical random variables may be represented by the state of a quantum system. This is particularly relevant in the context of cryptography, where an adversary may make use of quantum devices. Here, we build upon prior work by Ta-Shma and by De and Vidick to show that the well known construction paradigm for extractors proposed by Trevisan is sound in the presence of quantum side information.

1 Introduction

Randomness extraction is the art of generating (almost) uniform randomness Z from an only weakly random source of data X . More precisely, a *randomness extractor* (or, simply *extractor*) is a function that takes as input X together with a uniformly distributed (and usually short) random value Y , called *seed*, and maps this to Z . One then requires that Z is (almost) uniformly distributed whenever the min-entropy of X is larger than some threshold k , i.e.,

$$H_{\min}(X) \geq k \implies Z := \text{Ext}(X, Y) \text{ almost uniform.} \quad (1)$$

We note here that the min-entropy is directly related to the probability of correctly guessing the value of X (using an optimal strategy), i.e., $2^{-H_{\min}(X)} =$

^{*}anindya@cs.berkeley.edu

[†]portmann.c.aa@m.titech.ac.jp

[‡]vidick@cs.berkeley.edu

[§]renner@phys.ethz.ch

$\max_x P_X(x)$. Hence, the criterion can be interpreted operationally: if the probability of guessing the input of the extractor, X , is sufficiently low then its output is almost uniform.

Obviously, notions of randomness such as the *guessing probability* or the *uniformity* of a random variable always depend on the side information relative to which they are defined. We may make this explicit in our formulation of Criterion (1). Denoting by E all side information with respect to which the extractor output should be uniform, the criterion reads

$$H_{\min}(X|E) \geq k \implies Z := \text{Ext}(X, Y) \text{ almost uniform conditioned on } E, \quad (2)$$

where $H_{\min}(X|E)$ is the conditional min-entropy. Following the remark above, $H_{\min}(X|E)$ can be interpreted operationally, i.e., $2^{-H_{\min}(X|E)}$ is the maximum probability of correctly guessing X , given access to side information E . (For definitions of min-entropy we refer to Section 2.2. For definitions of extractors as well as technical versions of the criteria, we refer to Section 3.1.)

Criterion (2) looks like an innocent reformulation of Criterion (1). However, crucially, their relationship depends on the physical nature of the side information E , i.e., whether E is represented by the state of a classical or a quantum system. In the case of purely classical side information, E may be modeled as a random variable and it is known that the two Criteria (2) and (1) are essentially equivalent (see Lemma 3.3 for a precise statement). But in the general case where E is a quantum system, Criterion (2) is strictly stronger than (1): it was shown in [1] that there exist extractors that fulfill (1) but not (2) (see also [2] for a discussion). On the positive side however, it is known [3] that the two criteria are essentially equivalent for extractors with one-bit output.

Since our world is inherently non-classical, one may now argue that (2) rather than (the weaker) Criterion (1) should be taken as the relevant criterion for the definition of extractors. Indeed, it turns out that for various applications of extractors, Criterion (2) is necessary. For example, in the context of cryptography, one typically uses extractors to generate secret keys, i.e., randomness that is uniform from an adversary's point of view (see the literature on privacy amplification, e.g., [4]). Since an adversary may store information E in a quantum system, Criterion (1) is not sufficient to imply secrecy.

In the standard literature on randomness extraction, constructions of extractors are usually shown to fulfill Criterion (1), for certain values of the threshold k (see [5] as well as [6] for an overview). However, only a few constructions have been shown to fulfill Criterion (2) (with arbitrary quantum side information E). Among them is two-universal hashing [7, 8] as well as constructions based on the sample-and-hash approach [2].

Recently, Ta-Shma [9] studied Trevisan's construction of extractors [10] in the presence of quantum side information. Although his proof requires the output length to be much smaller than the min-entropy of the original data, the result was a breakthrough because it, for the first time, implied the existence of "quantum-proof" extractors requiring only short seeds (logarithmic in the output length). Later, two of the present authors [11] were able to improve the performance in terms of the output length, using again a construction based on Trevisan's paradigm. Their result was proved in the context of the bounded-storage model, in which the adversary's power is quantified by the number of (qu)bits he is allowed to use to store his side information E . More precisely, they

showed that for an input X and quantum side information E , the output could be as long as $H_{\min}(X) - H_0(E)$, where $H_0(E)$ denotes the number of quantum bits required to store E . While this expression can (in general) be arbitrarily smaller than $H_{\min}(X|E)$ (see Criterion (2)), the construction is optimal if X is uniform on its support and the quantum state of E is pure conditioned on X .

In this work, we show that the performance of Trevisan's extractor is almost optimal in the presence of quantum side information. More precisely, the output length of the extractor can be close to the conditional min-entropy $H_{\min}(X|E)$ (see Corollary 5.3 for the exact parameters), which is all the randomness that can possibly be extracted (see [7] for a discussion of the optimality of general randomness-extracting functions).

Our main result is a little more general than this. It has been observed, by, e.g., Lu and Vadhan [12, 13], that Trevisan's extractor [10] (and variations of it, such as [14]) is a concatenation of the outputs of a one-bit extractor with different pseudo-random seeds. Since this one-bit extractor (and to some extent the construction of the pseudo-random seeds) can be freely changed without needing to modify the proof, this results in a generic scheme (defined in Section 4.1, Definition 4.2), which we prove to be secure with roughly the same parameters whether quantum side information is present or not (Section 4.2, Theorem 4.4).

Our argument follows in spirit the idea of De and Vidick [11]. Technically, the proof is essentially a concatenation of the two following previously-known results.

- In the first part of the original proof of Trevisan [10], it is shown that the ability to distinguish the extractor output from uniform implies the ability to distinguish the output of the underlying one-bit extractor from uniform (a list-decodable code in Trevisan's original scheme). Ta-Shma has shown that this claim is still true in the context of quantum side information [9].
- This reduces the problem to proving that the one-bit extractor used in the construction is secure in the presence of quantum side information. However, because for one-bit extractors, the general Criterion (2) is essentially equivalent to the classical one (1), as shown by König and Terhal [3], the claim follows from the classical results with roughly the same parameters.

This paper is structured as follows. We first define the technical tools we will need in Section 2, in particular the (conditional) min-entropy. In Section 3 we give formal definitions of extractors and discuss briefly how much randomness can be extracted from a given source. Section 4 contains the description of Trevisan's extractor construction paradigm and a proof that it is still sound in the presence of quantum side information. Then in Section 5 we plug in various one-bit extractors and pseudo-random seed constructions, resulting in, amongst others, a construction which is nearly optimal in the amount of randomness extracted and uses a seed of size poly-logarithmic in the input (which is identical to the best known bound in the classical case [14] for Trevisan's extractor). Finally, in Section 6, we mention a few classical results which modify and improve Trevisan's extractor, but for which the correctness in the presence of quantum side information does not seem to follow immediately from this work.

2 Technical preliminaries

2.1 Notation

We write $[N]$ for the set of integers $\{1, \dots, N\}$. If $x \in \{0, 1\}^n$ is a string of length n , $i \in [n]$ an integer, and $S \subseteq [n]$ a set of integers, we write x_i for the i^{th} bit of x , and x_S for the string formed by the bits of x at the positions given by the elements of S .

\mathcal{H} will always denote a finite-dimensional Hilbert space. We denote by $\mathcal{P}(\mathcal{H})$ the set of positive semi-definite operators on \mathcal{H} . We define the set of normalized quantum states $\mathcal{S}(\mathcal{H}) := \{\mathcal{P}(\mathcal{H}) : \text{tr } \rho = 1\}$ and the set of sub-normalized quantum states $\mathcal{S}_{\leq}(\mathcal{H}) := \{\mathcal{P}(\mathcal{H}) : \text{tr } \rho \leq 1\}$.

We write $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ for a bipartite quantum system and $\rho_{AB} \in \mathcal{P}(\mathcal{H}_{AB})$ for a bipartite quantum state. $\rho_A = \text{tr}_B(\rho_{AB})$ and $\rho_B = \text{tr}_A(\rho_{AB})$ denote the corresponding reduced density operators.

If a classical random variable X takes the value $x \in \mathcal{X}$ with probability p_x , it can be represented by the state $\rho_X = \sum_{x \in \mathcal{X}} p_x |x\rangle\langle x|$, where $\{|x\rangle\}_{x \in \mathcal{X}}$ is an orthonormal basis of a Hilbert space \mathcal{H}_X . If the classical system X is part of a composite system XB , any state of that composite system can be written as $\rho_{XB} = \sum_{x \in \mathcal{X}} p_x |x\rangle\langle x| \otimes \rho_B^x$.

$\|\cdot\|_{\text{tr}}$ denotes the trace-norm and is defined by $\|A\|_{\text{tr}} := \text{tr } \sqrt{A^\dagger A}$.

2.2 Min-entropy

To measure how much randomness a source contains and can be extracted, we will use the *smooth conditional min-entropy*. This entropy measure was first defined by Renner [7], and represents the optimal measure for randomness extraction in the sense that it is always possible to extract that amount of (almost) uniform randomness from a source, but never more.

Definition 2.1 (conditional min-entropy). Let $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$, then the *min-entropy* of A conditioned on B is defined as

$$H_{\min}(A|B)_\rho := \max_{\sigma_B \in \mathcal{S}(\mathcal{H}_B)} \{\lambda \in \mathbb{R} : 2^{-\lambda} \mathbf{1}_A \otimes \sigma_B \geq \rho_{AB}\}.$$

We will often drop the subscript ρ when there is no doubt about which state is meant.

This definition has a simple operational interpretation when the system A is classical, which is the case we consider. König et al. [15] showed that for a state $\rho_{XB} = \sum_{x \in \mathcal{X}} p_x |x\rangle\langle x| \otimes \rho_B^x$ classical on X ,

$$H_{\min}(X|B)_\rho = -\log p_{\text{guess}}(X|B)_\rho, \quad (3)$$

where $p_{\text{guess}}(X|B)$ is the probability of guessing X given B , namely the maximum over all POVMs $\{E_B^x\}_{x \in \mathcal{X}}$ on B of

$$p_{\text{guess}}(X|B)_\rho := \max_{\{E_B^x\}_{x \in \mathcal{X}}} \left(\sum_{x \in \mathcal{X}} p_x \text{tr}(E_B^x \rho_B^x) \right).$$

If the system B is empty, then the min-entropy of X reduces to the standard definition, $H_{\min}(X) = -\log \max_{x \in \mathcal{X}} p_x$ (sometimes written $H_\infty(X)$). In this

case the connection to the guessing probability is particularly obvious: when no side information is available, the best guess we can make is simply the value $x \in \mathcal{X}$ with highest probability.

We will need the following “chain-rule type” statement about the min-entropy.

Lemma 2.2 ([3, Lemma 1]). *Consider a state ρ_{XYZB} , where X, Y and Z are classical, $\rho_{XY} = \rho_X \otimes \rho_Y$ and $YZ \leftrightarrow X \leftrightarrow B$ form a Markov chain, that is $\rho_{XYZB} = \sum_{x,y,z} P_{XYZ}(x, y, z) |x, y, z\rangle\langle x, y, z| \otimes \rho_B^x$, then*

$$H_{\min}(X|YZB) \geq H_{\min}(X|B) - H_0(Z),$$

where $H_0(Z) = \log \text{rank}(\rho_Z)$.

As hinted at the beginning of this section, the min-entropy is not quite optimal, in the sense that it is sometimes possible to extract more randomness. However, the *smooth* min-entropy is optimal. This information measure consists in maximizing the min-entropy over all states ε -close to the actual state ρ_{XB} of the system considered. Thus by introducing an extra error ε , we have a state with more entropy. (See Section 3.2 for more details.)

As distance metric on the set of sub-normalized states, we will use the *purified distance*, as defined in [16].

Definition 2.3 (purified distance). Let $\rho, \sigma \in \mathcal{S}_{\leq}(\mathcal{H})$. We define the *purified distance* between ρ and σ as

$$P(\rho, \sigma) := \sqrt{1 - \bar{F}(\rho, \sigma)^2},$$

where $\bar{F}(\cdot, \cdot)$ denotes the generalized fidelity

$$\begin{aligned} \bar{F}(\rho, \sigma) &= \left\| \sqrt{\rho \oplus (1 - \text{tr } \rho)} \sqrt{\sigma \oplus (1 - \text{tr } \sigma)} \right\|_{\text{tr}} \\ &= \left\| \sqrt{\rho} \sqrt{\sigma} \right\|_{\text{tr}} + \sqrt{(1 - \text{tr } \rho)(1 - \text{tr } \sigma)}. \end{aligned}$$

Note that the purified distance is larger than the trace distance (for a proof, we refer to [16]), namely for any two states $\rho, \sigma \in \mathcal{S}_{\leq}(\mathcal{H})$,

$$P(\rho, \sigma) \geq \frac{1}{2} \|\rho - \sigma\|_{\text{tr}}. \quad (4)$$

This allows us to define an ε -ball of states.

Definition 2.4. Let $\varepsilon \geq 0$ and $\rho \in \mathcal{S}_{\leq}(\mathcal{H})$ with $\sqrt{\text{tr } \rho} > \varepsilon$. Then, we define an ε -ball in \mathcal{H} around ρ as

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

We can now define the smooth conditional min-entropy.

Definition 2.5 (smooth min-entropy). Let $\varepsilon \geq 0$ and $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$, then the ε -*smooth min-entropy* of A conditioned on B is defined as

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

3 Extractors

3.1 Extractors, adversaries, and privacy amplification

An extractor $\text{Ext} : \{0, 1\}^n \times \{0, 1\}^d \rightarrow \{0, 1\}^m$ is a function which takes a weak source of randomness X and a uniformly random, short seed Y , and produces some output $\text{Ext}(X, Y)$, which is nearly uniform. The extractor is said to be strong, if the output is approximately independent from the seed.

Definition 3.1 (strong extractor). A function $\text{Ext} : \{0, 1\}^n \times \{0, 1\}^d \rightarrow \{0, 1\}^m$ is a (k, ε) -strong extractor, if for all distributions X with $H_{\min}(X) \geq k$ and a uniform seed Y , we have¹

$$\frac{1}{2} \left\| \rho_{\text{Ext}(X, Y)Y} - \rho_U \otimes \rho_Y \right\|_{\text{tr}} \leq \varepsilon,$$

where ρ_U is the fully mixed state on a system of dimension 2^m .

Using the connection between min-entropy and guessing probability (Eq. (3)), a (k, ε) -strong extractor can be seen as a function which guarantees that if the probability of guessing X correctly is not too high (less than 2^{-k}), then it can produce a random variable which is approximately uniform and independent from the seed Y .

As discussed in the introduction, we consider here a more general situation involving side information, denoted E , which may be represented by the state of a quantum system. We then want to find some function Ext such that, if the probability of guessing X given E is not too high, Ext can produce a random variable $\text{Ext}(X, Y)$ which is approximately uniform and independent from the seed Y and the side information E . Equivalently, one may think of a *privacy amplification* scenario [17, 4], where E is the information available to an adversary and where the goal is to turn weakly secret data X into a *secret* key $\text{Ext}(X, Y)$, where the seed Y is assumed to be public. (In typical key agreement protocols, the key is chosen by the legitimate parties and exchanged over public channels.)

The following definition covers the general situation where the side information E may be represented quantum-mechanically. The case of purely classical side information is then formulated as a restriction on the nature of E .

Definition 3.2 (strong extractor against adversaries). A function $\text{Ext} : \{0, 1\}^n \times \{0, 1\}^d \rightarrow \{0, 1\}^m$ is a (k, ε) -strong extractor against quantum adversaries, if for all states ρ_{XE} classical on X with $H_{\min}(X|E)_\rho \geq k$, and for a uniform seed Y , we have

$$\frac{1}{2} \left\| \rho_{\text{Ext}(X, Y)YE} - \rho_U \otimes \rho_Y \otimes \rho_E \right\|_{\text{tr}} \leq \varepsilon,$$

where ρ_U is the fully mixed state on a system of dimension 2^m .

The function Ext is a (k, ε) -strong extractor against classical adversaries if the same holds with the system E restricted to classical states.

¹A more standard classical notation would be $\frac{1}{2} |\text{Ext}(X, Y) \circ Y - U \circ Y| \leq \varepsilon$, where the distance metric is the variational distance. However, since classical random variables can be represented by quantum states diagonal in the computational basis, and the variational distance is a special case of the trace distance, we use the quantum notation for compatibility with the rest of this work.

It turns out that if the adversary holding E is restricted to classical information about X , then this definition is essentially equivalent to the conventional Definition 3.1.

Lemma 3.3 ([3, Proposition 1]). *Any (k, ε) -strong extractor is a $(k + \log 1/\varepsilon, 2\varepsilon)$ -strong extractor against classical adversaries.*

However, if the adversary has quantum information, this does not necessarily hold. Gavinsky et al. [1] give an example of a (k, ε) -strong extractor, which does not work in the presence of a quantum adversary, even when $H_{\min}(X|E)$ is much larger than k .

3.2 Extracting more randomness

Radhakrishnan and Ta-Shma [18] have shown that a (k, ε) -strong extractor $\text{Ext} : \{0, 1\}^n \times \{0, 1\}^d \rightarrow \{0, 1\}^m$ will necessarily have $m \leq k - 2 \log 1/\varepsilon + O(1)$. However, in some situations we can extract much more randomness than the min-entropy. For example, let X be distributed on $\{0, 1\}^n$ with $\Pr[X = x_0] = 1/n$ and for all $x \neq x_0$, $\Pr[X = x] = \frac{n-1}{n(2^n-1)}$. We have $H_{\min}(X) = \log n$, so using a $(\log n, 1/n)$ -strong extractor we could obtain at most $\log n$ bits of randomness. But X is already $1/n$ -close to uniform, since $\frac{1}{2} \|\rho_X - \rho_U\|_{\text{tr}} \leq \frac{1}{n}$. So we already have n bits of nearly uniform randomness, exponentially more than by using a $(\log n, 1/n)$ -strong extractor.

In the case of an extractor against quantum adversaries, similar examples can be found, e.g., in [16, Remark 22]. We therefore need a different entropy measure to characterize how much randomness can be extracted. For this we will use the *smooth min-entropy* (Definition 2.5), which is essentially optimal [7, Section 5.6].

Lemma 3.4. *If $\text{Ext} : \{0, 1\}^n \times \{0, 1\}^d \rightarrow \{0, 1\}^m$ is a (k, ε) -strong extractor against (classical or quantum) adversaries, then for any state ρ_{XE} with $H_{\min}^{\varepsilon'}(X|E)_\rho \geq k$,*

$$\frac{1}{2} \|\rho_{\text{Ext}(X,Y)YE} - \rho_U \otimes \rho_Y \otimes \rho_E\|_{\text{tr}} \leq \varepsilon + 2\varepsilon'.$$

This lemma implies that an extractor which is known to extract m bits from any source such that $H_{\min}(X|E) \geq k$ can in fact extract the same number of bits, albeit with a slightly larger error, from sources which only satisfy $H_{\min}^{\varepsilon'}(X|E) \geq k$, a much weaker requirement in some cases.

Proof. Let $\tilde{\rho}_{XE}$ be the state ε' -close to ρ_{XE} for which $H_{\min}(X|E)_{\tilde{\rho}}$ reaches its maximum. Then

$$\begin{aligned} & \frac{1}{2} \|\rho_{\text{Ext}(X,Y)YE} - \rho_U \otimes \rho_Y \otimes \rho_E\|_{\text{tr}} \\ & \leq \frac{1}{2} \|\rho_{\text{Ext}(X,Y)YE} - \tilde{\rho}_{\text{Ext}(X,Y)YE}\|_{\text{tr}} + \frac{1}{2} \|\tilde{\rho}_{\text{Ext}(X,Y)YE} - \rho_U \otimes \rho_Y \otimes \tilde{\rho}_E\|_{\text{tr}} \\ & \quad + \frac{1}{2} \|\rho_U \otimes \rho_Y \otimes \tilde{\rho}_E - \rho_U \otimes \rho_Y \otimes \rho_E\|_{\text{tr}} \\ & \leq \frac{1}{2} \|\tilde{\rho}_{\text{Ext}(X,Y)YE} - \rho_U \otimes \rho_Y \otimes \tilde{\rho}_E\|_{\text{tr}} + \|\rho_{XE} - \tilde{\rho}_{XE}\|_{\text{tr}} \\ & \leq \varepsilon + 2\varepsilon'. \end{aligned}$$

In the second inequality above we used (twice) the fact that a trace-preserving quantum operation can only decrease the trace distance. And in the last line we used Eq. (4), which says that the purified distance is larger than the trace distance. \square

3.3 Entropy loss

Since a (k, ε) -strong extractor can be applied to any ε source with smooth min-entropy $H_{\min}^{\varepsilon'}(X|E) \geq k$, we can measure the entropy loss of the extractor – namely how much entropy was not extracted – with

$$\Delta := k - m,$$

where m is the size of the output.

As already quoted in the previous section, Radhakrishnan and Ta-Shma [18] show that we necessarily have $m \leq k - 2 \log 1/\varepsilon + O(1)$. Hence a scheme is considered to have optimal entropy loss if $\Delta = 2 \log 1/\varepsilon + O(1)$.

4 Constructing m -bit extractors from 1-bit extractors and weak designs

In this section we show how to construct an m -bit extractor against quantum adversaries from any 1-bit strong extractor.

This can be seen as a derandomization of a result by König and Terhal [3], who also extract m -bits against quantum adversaries by concatenating m times a 1-bit extractor. They however choose a different seed for each bit, thus having a seed of total length $d = mt$, where t is the length of the seed of the 1-bit extractor. In the case of classical adversaries, this derandomization was done by Trevisan [10], who shows how to concatenate m times a 1-bit extractor using only $d = \text{poly}(t, \log m)$ bits of seed.² We combine the weak designs from Raz et al. [14], which they use to improve Trevisan’s extractor, and a previous observation by two of the authors [11], that since 1-bit extractors were proven secure against quantum adversaries by König and Terhal [3], Trevisan’s extractor is also secure against quantum adversaries.

This results in a generic scheme, which can use any weak design and 1-bit strong extractor. We define it in Section 4.1, then prove bounds on the min-entropy and error in Section 4.2.

4.1 Description of the generic scheme

In order to shorten the seed while still outputting m bits, we treat the seed as a string of length $d < mt$, from which the m overlapping blocks of t bits needed to specify the different seeds will be extracted.

Let $y \in \{0, 1\}^d$ be the total seed. To specify the seeds for each application of the 1-bit extractor we need m sets $S_1, \dots, S_m \subset [d]$ of size $|S_i| = t$ for all

²Trevisan’s original paper does not explicitly define his extractor as a pseudo-random concatenation of a 1-bit extractors. It has however been noted in, e.g., [12, 13], that this is basically what Trevisan’s extractor does.

i. The seeds for the different runs of the 1-bit extractor are then given by y_{S_i} , namely the bits of y at the positions specified by the elements of S_i .

The seeds for the different outputs of the 1-bit extractor must however be nearly independent, so we will minimize the overlap between the sets, namely $|S_i \cap S_j|$. For this we use the *weak designs* introduced by Raz et al. [14].³

Definition 4.1 (weak design, [14, Definition 5]). A family of sets $S_1, \dots, S_m \subset [d]$ is a *weak (t, r) -design* if

1. For all i , $|S_i| = t$.
2. For all i , $\sum_{j=1}^{i-1} 2^{|S_j \cap S_i|} \leq rm$.

We can now describe the extractor construction.

Definition 4.2. For a one-bit extractor $C : \{0, 1\}^n \times \{0, 1\}^t \rightarrow \{0, 1\}$, which uses a seed of length t , and for a weak (t, r) -design $S_1, \dots, S_m \subset [d]$, we define the m -bit extractor $\text{Ext}_C : \{0, 1\}^n \times \{0, 1\}^d \rightarrow \{0, 1\}^m$ as

$$\text{Ext}_C(x, y) := C(x, y_{S_1}) \cdots C(x, y_{S_m}).$$

y_{S_i} is the string composed of the bits of y at the positions specified by the elements in S_i .

Remark 4.3. The length of the seed of the extractor Ext_C is d , one of the parameters of the weak design, which in turn depends on t , the size of the seed of the 1-bit extractor C . In Section 5 we will give concrete instantiations of weak designs and 1-bit extractors, achieving various entropy losses and seed sizes. The size of the seed will always be $d = \text{poly}(\log n)$, if the error is $\varepsilon = \text{poly}(1/n)$. For example, to achieve a near optimal entropy loss (Section 5.1), we need $d = O(t^2 \log m)$ and $t = O(\log n)$, hence $d = O(\log^3 n)$.

4.2 Analysis

We now prove that the extractor defined in the previous section is a strong extractor against quantum adversaries with following parameters.

Theorem 4.4. *Let $C : \{0, 1\}^n \times \{0, 1\}^t \rightarrow \{0, 1\}$ be a (k, ε) -strong extractor and $S_1, \dots, S_m \subset [d]$ a weak (t, r) -design. Then the extractor given in Definition 4.2, $\text{Ext}_C : \{0, 1\}^n \times \{0, 1\}^d \rightarrow \{0, 1\}^m$, is a $(k + rm + \log 1/\varepsilon, 3m\sqrt{\varepsilon})$ -strong extractor against quantum adversaries.*

The proof follows the same structure as the security proof of Trevisan's extractor [10, 14]. We first show that an adversary who can distinguish the output of the extractor Ext_C from uniform can guess the bits of $C(X, \cdot)$ on average given a little extra information. This is summed up in the following proposition:

³The second condition of the weak design was originally defined as $\sum_{j=1}^{i-1} 2^{|S_j \cap S_i|} \leq r(m-1)$. We prefer use the version of [19], since it simplifies the notation without changing the design constructions.

Proposition 4.5. If an adversary holding the system E can distinguish the output of Ext_C from uniform with probability greater than ε , then he can distinguish the output of the one-bit extractor C from uniform with probability greater than ε/m , given E and a classical system GW , namely

$$\left\| \rho_{C(X,V)VEGW} - \rho_U \otimes \rho_V \otimes \rho_{EGW} \right\|_{\text{tr}} > \frac{\varepsilon}{m}.$$

The system G has size $H_0(G) \leq rm$, where r is one of the parameters of the weak design (Definition 4.1). W is independent from X , namely $\rho_{XW} = \rho_X \otimes \rho_W$. And $GW \leftrightarrow X \leftrightarrow E$ form a Markov chain, that is $\rho_{XGW} = \sum_{x,g,w} P_{XGW}(x,g,w) |x,g,w\rangle\langle x,g,w| \otimes \rho_E^x$.

This proposition has been (implicitly) proved in [9], in which the (memory of the) adversary is viewed as an oracle. For completeness, and in order to have the proposition stated in exactly the form we need it, we provide a separate proof of Proposition 4.5 in Appendix B as Lemma B.2.

The second part of the security proof of Trevisan’s extractor consists in showing that an adversary who can distinguish the output of C from uniform can reconstruct X . But this is nothing else than proving that C is an extractor against (in our case, quantum) adversaries, since by definition of an extractor, an adversary who can distinguish the output from uniform has low conditional min-entropy about the input, i.e., can correctly guess / reconstruct X with high probability. By our extractor construction, C is already a 1-bit extractor, so all we need is the result by König and Terhal [3] which says that 1-bit extractors are secure against quantum adversaries.

Theorem 4.6 ([3, Theorem III.1]). *Let $C : \{0,1\}^n \times \{0,1\}^t \rightarrow \{0,1\}$ be a (k, ε) -strong extractor. Then C is a $(k + \log 1/\varepsilon, 3\sqrt{\varepsilon})$ -strong extractor against quantum adversaries.*

We are now ready to prove the main theorem.

Proof of Theorem 4.4. We proceed by contradiction, assuming that an adversary holds a system E such that $H_{\min}(X|E) \geq k + rm + \log 1/\varepsilon$ and

$$\left\| \rho_{\text{Ext}_C(X,Y)YE} - \rho_U \otimes \rho_Y \otimes \rho_E \right\|_{\text{tr}} > 3m\sqrt{\varepsilon}.$$

Then by Proposition 4.5 we know that there exist classical “advice” systems G and W such that,

$$\left\| \rho_{C(X,V)VEGW} - \rho_U \otimes \rho_V \otimes \rho_{EGW} \right\|_{\text{tr}} > 3\sqrt{\varepsilon},$$

where $H_0(G) \leq rm$, W is independent from X and $GW \leftrightarrow X \leftrightarrow E$ form a Markov chain.

By the assumption on C and Theorem 4.6, this means that $H_{\min}(X|EGW) < k + \log 1/\varepsilon$, hence by Lemma 2.2, $H_{\min}(X|E) < k + H_0(G) + \log 1/\varepsilon \leq k + rm + \log 1/\varepsilon$. A contradiction. \square

5 Concrete constructions

Depending on what goal has been set – e.g., maximize the output, minimize the seed length – different 1-bit extractors and weak designs will be needed.

In this section we give a few examples of what can be done, by taking various extractors and designs used against classical adversaries, and plugging them into Theorem 4.4, to obtain bounds on the seed size and entropy loss against quantum adversaries.

We first consider the problem of extracting all the min-entropy of the source in Section 5.1. This was achieved in the classical case by Raz et al. [14], so we use the same 1-bit extractor and weak design as them.

In Section 5.2 we give a scheme which uses a seed of length $d = O(\log n)$, but can only extract part of the entropy. This is also based on Raz et al. [14] in the classical case.

And finally in Section 5.3 we use an extractor and design which are locally computable (from Vadhan [13] and Hartman and Raz [19] respectively), to produce an m -bit extractor against quantum adversaries, such that each bit of the output depends on only $O(m \log(m/\varepsilon))$ bits of the input.

5.1 Near optimal entropy loss

To achieve a near optimal entropy loss we need to combine a 1-bit extractor with near optimal entropy loss and a weak $(t, 1)$ -design. We use the same extractor and design as Raz et al. [14] to do so.

Lemma 5.1 ([14, Lemma 17]⁴). *For every $t, m \in \mathbb{N}$ there exists a weak $(t, 1)$ -design $S_1, \dots, S_m \subset [d]$ such that $d = t \lceil \frac{t}{\ln 2} \rceil \lceil \log 4m \rceil = O(t^2 \log m)$. Moreover, such a design can be found in time $\text{poly}(m, d)$ and space $\text{poly}(m)$.*

As 1-bit extractor, Raz et al. [14] (and Trevisan [10] too) used the bits of a list-decodable code. We give the parameters here as Proposition 5.2 and refer to Appendix C for details on the construction and proof.

Proposition 5.2. For any $\varepsilon > 0$ and $n \in \mathbb{N}$ there exists a (k, ε) -strong extractor $\text{Ext}_{n, \varepsilon} : \{0, 1\}^n \times \{0, 1\}^d \rightarrow \{0, 1\}$ with $d = O(\log(n/\varepsilon))$ and $k = 3 \log 1/\varepsilon$.

Plugging this into Theorem 4.4 we get an extractor against quantum adversaries with parameters similar to Raz et al. [14].

Corollary 5.3. *Let $C_{n, \delta} : \{0, 1\}^n \times \{0, 1\}^t \rightarrow \{0, 1\}$ be the extractor from Proposition 5.2 with $\delta = \frac{\varepsilon^2}{9m^2}$ and let $S_1, \dots, S_m \subset [d]$ be the weak $(t, 1)$ -design from Lemma 5.1. Then*

$$\begin{aligned} \text{Ext} : \{0, 1\}^n \times \{0, 1\}^d &\rightarrow \{0, 1\}^m \\ (x, y) &\mapsto C(x, y_{S_1}) \cdots C(x, y_{S_m}) \end{aligned}$$

is a $(m + 8 \log m + 8 \log 1/\varepsilon + O(1), \varepsilon)$ -strong extractor against quantum adversaries, with $d = O(\log^2(n/\varepsilon) \log m)$.

For $\varepsilon = \text{poly}(1/n)$ the seed has length $d = O(\log^3 n)$.

The entropy loss is $\Delta = 8 \log m + 8 \log 1/\varepsilon + O(1)$. For $\varepsilon = O(1/n)$ it is then $\Delta = O(\log 1/\varepsilon)$ which is optimal up to a constant factor. By Lemma 3.4 this extractor can produce $m = H_{\min}^\varepsilon(X|E) - O(\log 1/\varepsilon)$ bits of randomness with an error 3ε .

⁴Hartman and Raz [19] give a more efficient construction of this lemma, namely in time $\text{poly}(\log m, t)$ and space $\text{poly}(\log m + \log t)$, with the extra minor restriction that $m > t^{\log t}$.

5.2 Seed of logarithmic size

The weak design used in Section 5.1 requires the seed to be of size $d = \Theta(t^2 \log m)$, where t is the size of the seed of the 1-bit extractor. Since t cannot be less than $\log n$, a scheme using this design will always have $d = \Omega(\log^2 n \log m)$. If we want to use a seed of size $d = O(\log n)$ we need a different weak design.

Lemma 5.4 ([14, Lemma 15]). *For every $t, m \in \mathbb{N}$ and $r > 1$, there exists a weak (t, r) -design $S_1, \dots, S_m \subset [d]$ such that $d = t \lceil \frac{t}{\ln r} \rceil = O\left(\frac{t^2}{\log r}\right)$. Moreover, such a design can be found in time $\text{poly}(m, d)$ and space $\text{poly}(m)$.*

For the 1-bit extractor we can use the same as in the previous section, Proposition 5.2.

Plugging this into Theorem 4.4 with $\log r = \Theta(t)$, we get an extractor against quantum adversaries with logarithmic seed length.

Corollary 5.5. *If for any constant $0 < \alpha \leq 1$, the source has min-entropy $H_{\min}(X|E) = n^\alpha$, and the desired error is $\varepsilon = \text{poly}(1/n)$, then using the extractor $C_{n,\delta} : \{0, 1\}^n \times \{0, 1\}^t \rightarrow \{0, 1\}$ from Proposition 5.2 with $\delta = \frac{\varepsilon^2}{9m^2}$ and the weak (t, r) -design $S_1, \dots, S_m \subset [d]$ from Lemma 5.4 with $r = n^\gamma$ for any $0 < \gamma < \alpha$, we have that*

$$\begin{aligned} \text{Ext} : \{0, 1\}^n \times \{0, 1\}^d &\rightarrow \{0, 1\}^m \\ (x, y) &\mapsto C(x, y_{S_1}) \cdots C(x, y_{S_m}) \end{aligned}$$

is a $(n^\gamma m + 8 \log m + 8 \log 1/\varepsilon + O(1), \varepsilon)$ -strong extractor against quantum adversaries, with $d = O\left(\frac{1}{\gamma} \log n\right)$.

Choosing γ to be a constant results in a seed of length $d = O(\log n)$. The output length is $m = n^{\alpha-\gamma} - o(1) = H_{\min}(X|E)^{1-\frac{\gamma}{\alpha}} - o(1)$. By Lemma 3.4 this can be increased to $m = H_{\min}^\varepsilon(X|E)^{1-\frac{\gamma}{\alpha}} - o(1)$ with an error of 3ε .

5.3 Locally computable extractor

Another interesting feature of extractors is to be *local*, that is, the m -bit output depends only a small subset of the n input bits. This is useful in, e.g., the bounded storage model [20, 12, 13], where we assume a huge source of random bits, say n , are available, and the adversary's storage is bounded by νn for some constant $\nu < 1$. Legitimate parties are also assumed to have bounded workspace for computation. In particular, for the model to be meaningful, the bound is stricter than that on the adversary. So to extract a secret key from the large source of randomness, they need an extractor which only reads $\ell \ll n$ bits.

Definition 5.6 (ℓ -local extractor). An extractor $\text{Ext} : \{0, 1\}^n \times \{0, 1\}^d \rightarrow \{0, 1\}^m$ is ℓ -locally computable (or ℓ -local), if for every $r \in \{0, 1\}^d$, the function $x \mapsto \text{Ext}(x, r)$ depends on only ℓ bits of its input, where the bit locations are determined by r .

Lu [12] modified Trevisan's scheme [10, 14] to use a local list-decodable code as 1-bit extractor. Vadhan [13] proposes another construction for local extractors, which is optimal up to constant factors. Both these constructions

have similar parameters in the case of 1-bit extractors.⁵ We state the parameters of Vadhan’s construction here and Lu’s constructions in Appendix C.

Lemma 5.7 ([13, Theorem 8.5]). *For any $\varepsilon > \exp(-n/2^{O(\log^* n)})$, $n \in \mathbb{N}$ and constant $0 < \gamma < 1$, there exists an explicit ℓ -local (k, ε) -strong extractor $\text{Ext}_{n, \varepsilon, \gamma} : \{0, 1\}^n \times \{0, 1\}^d \rightarrow \{0, 1\}$ with $d = O(\log(n/\varepsilon))$, $k = \gamma n$ and $\ell = O(\log 1/\varepsilon)$.*

Since we assume that the available memory is limited, we also want the construction of the weak design to be particularly efficient. For this we can use a construction by Hartman and Raz [19].

Lemma 5.8 ([19, Theorem 3]). *For every $m, t \in \mathbb{N}$, such that $m = \Omega(t^{\log t})$ and constant $r > 1$, there exists an explicit weak (t, r) -design $S_1, \dots, S_m \subset [d]$, where $d = O(t^2)$. Such a design can be found in time $\text{poly}(\log m, t)$ and space $\text{poly}(\log m + \log t)$.*

Remark 5.9. For the extractor from Lemma 5.7 and an error $\varepsilon = \text{poly}(1/n)$, this design requires $m = \Omega((\log n)^{\log \log n})$. If we are interested in a smaller m , say $m = \text{poly}(\log n)$, then we can use the weak design from Lemma 5.4 with $r = n^\gamma$. This construction would require time and space $\text{poly}(\log n) = \text{poly}(\log 1/\varepsilon)$. The resulting seed would have length only $O(\log n)$ instead of $O(\log^2 n)$.

Plugging this into Theorem 4.4 we get a local extractor against quantum adversaries.

Corollary 5.10. *If for any constant $0 < \alpha \leq 1$, the source has min-entropy $H_{\min}(X|E) = \alpha n$, then using the weak (t, r) -design $S_1, \dots, S_m \subset [d]$ from Lemma 5.8 and the extractor $C_{n, \delta, \gamma} : \{0, 1\}^n \times \{0, 1\}^t \rightarrow \{0, 1\}$ from Lemma 5.7 with $\delta = \frac{\varepsilon^2}{9m^2}$ and any constant $\gamma < \alpha$, we have that*

$$\begin{aligned} \text{Ext} : \{0, 1\}^n \times \{0, 1\}^d &\rightarrow \{0, 1\}^m \\ (x, y) &\mapsto C(x, y_{S_1}) \cdots C(x, y_{S_m}) \end{aligned}$$

is an ℓ -local $(\gamma n + rm + 2 \log m + 2 \log 1/\varepsilon + O(1), \varepsilon)$ -strong extractor against quantum adversaries, with $d = O(\log^2(n/\varepsilon))$ and $\ell = O(m \log(m/\varepsilon))$. Furthermore, each bit of the output depends on only $O(\log(m/\varepsilon))$ bits of the input.

With these parameters the extractor can produce up to $m = (\alpha - \gamma)n/r - O(\log 1/\varepsilon) = (H_{\min}(X|E) - \gamma n)/r - O(\log 1/\varepsilon)$ bits of randomness, when $\varepsilon = \text{poly}(1/n)$. By Lemma 3.4 this can be increased to $m = (H_{\min}^\varepsilon(X|E) - \gamma n)/r - O(\log 1/\varepsilon)$ with an error of 3ε .

6 Other variations of Trevisan’s scheme

There exist many results modifying and improving Trevisan’s extractor. Some of them still follow the “design and 1-bit extractor” pattern, such as the work of Raz et al. [14] and Lu [12], which correspond to modifications of the design and 1-bit extractor respectively. Hence our work implies that these are immediately secure against quantum adversaries with roughly the same parameters.

⁵If the extractor is used to extract m -bits, then Vadhan’s scheme reads less input bits and uses a shorter seed than Lu’s.

Other results such as [14, 21, 22] replace the binary list-decoding codes with multivariate codes over a field F . The connection to 1-bit extractors is not clear anymore, and the security against quantum adversaries not guaranteed.

Raz et al. extract a little more randomness than we do in Section 5.1. They reduce the entropy loss from $\Delta = O(\log 1/\varepsilon)$ to $\Delta = 2 \log 1/\varepsilon + O(1)$ by composing (in the sense described in Appendix A) the scheme of Section 5.1 with an extractor by Srinivasan and Zuckerman [23], which has an entropy loss of $\Delta = 2 \log 1/\varepsilon + O(1)$ and seed length $O(k + \log n)$. Since $k = O(\log n)$, the total seed remains poly-logarithmic in n . The same approach could be useful in the context of a quantum adversary using Lemma A.1. While it is unknown whether the extractor of Srinivasan and Zuckerman is sound against quantum adversaries, one may try to replace it by alternative constructions such as δ -almost two-universal hashing (see [8]).

In the case of a logarithmic seed length, Impagliazzo et al. [24] and Ta-Shma et al. [25] modify Trevisan's extractor to work for a sub-polynomial entropy source, still using a seed of size $d = O(\log n)$. While it is unclear whether these modifications preserve the “design and 1-bit extractor” structure, it is an interesting open problem to analyze them in the context of quantum side information.

A Composing extractors

If an extractor does not have optimal entropy loss, a useful approach to extract more entropy is to apply a second extractor to the original input, trying to extract the randomness that remains when the output of the first extractor is known. This was first proposed in the classical case by Wigderson and Zuckerman [26], and improved by Raz et al. [14]. König and Terhal [3] gave the first quantum version for composing m times 1-bit extractors against quantum adversaries. We slightly generalize the result of König and Terhal [3] to the composition of arbitrary extractors against quantum adversaries.

Lemma A.1. *Let $\text{Ext}_1 : \{0, 1\}^n \times \{0, 1\}^{d_1} \rightarrow \{0, 1\}^{m_1}$ and $\text{Ext}_2 : \{0, 1\}^n \times \{0, 1\}^{d_2} \rightarrow \{0, 1\}^{m_2}$ be (k, ε_1) - and $(k - m_1, \varepsilon_2)$ -strong extractors against (classical or quantum) adversaries respectively. Then the composition of the two, namely*

$$\begin{aligned} \text{Ext}_3 : \{0, 1\}^n \times \{0, 1\}^{d_1} \times \{0, 1\}^{d_2} &\rightarrow \{0, 1\}^{m_1} \times \{0, 1\}^{m_2} \\ (x, y_1, y_2) &\mapsto (\text{Ext}_1(x, y_1), \text{Ext}_2(x, y_2)), \end{aligned}$$

is a $(k, \varepsilon_1 + \varepsilon_2)$ -strong extractor against (classical or quantum) adversaries.

Proof. We need to show that for any state ρ_{XE} with $H_{\min}(X|E) \geq k$,

$$\frac{1}{2} \left\| \rho_{\text{Ext}_1(X, Y_1) \text{Ext}_2(X, Y_2) Y_1 Y_2 E} - \rho_{U_1} \otimes \rho_{U_2} \otimes \rho_{Y_1} \otimes \rho_{Y_2} \otimes \rho_E \right\|_{\text{tr}} \leq \varepsilon_1 + \varepsilon_2. \quad (5)$$

The left-hand side of Eq. (5) can be upper-bounded by

$$\begin{aligned} &\frac{1}{2} \left\| \rho_{\text{Ext}_1(X, Y_1) Y_1 E} \otimes \rho_{U_2} \otimes \rho_{Y_2} - \rho_{U_1} \otimes \rho_{Y_1} \otimes \rho_E \otimes \rho_{U_2} \otimes \rho_{Y_2} \right\|_{\text{tr}} \\ &+ \frac{1}{2} \left\| \rho_{\text{Ext}_2(X, Y_2) Y_2 \text{Ext}_1(X, Y_1) Y_1 E} - \rho_{U_2} \otimes \rho_{Y_2} \otimes \rho_{\text{Ext}_1(X, Y_1) Y_1 E} \right\|_{\text{tr}}. \quad (6) \end{aligned}$$

By the definition of Ext_1 the first term in Eq. (6) is upper-bounded by ε_1 . For the second term we have $\rho_{XY_1} = \rho_X \otimes \rho_{Y_1}$ and $\text{Ext}_1(X, Y_1)Y_1 \leftrightarrow X \leftrightarrow E$ form a Markov chain, so we can apply Lemma 2.2 and get

$$H_{\min}(X | \text{Ext}_1(X, Y_1)Y_1 E) \geq H_{\min}(X | E) - H_0(\text{Ext}_1(X, Y_1)) \geq k - m_1.$$

By the definition of Ext_2 the second term in Eq. (6) can then be upper-bounded by ε_2 . \square

B Security reduction

To show that an adversary who can distinguish the output of Ext_C (defined in Definition 4.2 on page 9) from uniform can also guess the output of the extractor C , we first show that such an adversary can guess one of the bits of the output of Ext_C given some extra classical information. This is a quantum version of a result by Yao [27].

Lemma B.1. *If a player holds a quantum state ρ_B correlated with a classical random variable Z on m -bit strings, such that he can distinguish Z from uniform with probability greater than ε , then there exists a bit $i \in [m]$ such that when given the previous $i-1$ bits of Z , he can distinguish the i^{th} bit of Z from uniform with probability greater than $\frac{\varepsilon}{m}$, i.e., if $\|\rho_{ZB} - \rho_U \otimes \rho_B\|_{\text{tr}} > \varepsilon$, then there exists an $i \in [m]$ such that⁶*

$$\left\| \sum_{\substack{z \in \mathcal{Z} \\ z_i=0}} p_z |z_{[i-1]} \rangle \langle z_{[i-1]}| \otimes \rho_B^z - \sum_{\substack{z \in \mathcal{Z} \\ z_i=1}} p_z |z_{[i-1]} \rangle \langle z_{[i-1]}| \otimes \rho_B^z \right\|_{\text{tr}} > \frac{\varepsilon}{m}. \quad (7)$$

Proof. The proof uses a hybrid argument. Let

$$\sigma_i = \sum_{\substack{z \in \mathcal{Z} \\ r \in \{0,1\}^m}} \frac{p_z}{2^m} |z_{[i], r_{\{i+1, \dots, m\}}} \rangle \langle z_{[i], r_{\{i+1, \dots, m\}}} | \otimes \rho_B^z.$$

Then

$$\begin{aligned} \varepsilon &< \|\rho_{ZB} - \rho_U \otimes \rho_B\|_{\text{tr}} \\ &= \|\sigma_m - \sigma_0\|_{\text{tr}} \\ &\leq \sum_{i=1}^m \|\sigma_i - \sigma_{i-1}\|_{\text{tr}} \\ &\leq m \max_i \|\sigma_i - \sigma_{i-1}\|_{\text{tr}}. \end{aligned}$$

By rearranging $\|\sigma_i - \sigma_{i-1}\|_{\text{tr}}$ we get the lhs of Eq. (7). \square

⁶To simplify the notation, the statement of this lemma uses the fact that for any *binary* random variable X and quantum system Q , the following equality holds:

$$\|\rho_{XQ} - \rho_U \otimes \rho_Q\|_{\text{tr}} = \|p_0 \rho_Q^0 - p_1 \rho_Q^1\|_{\text{tr}}.$$

We now need to bound the size of this extra information, the “previous $i - 1$ bits”, and show that when averaging over all the seeds of Ext_C , we average over all the seeds of C , which means that guessing a bit of the output of Ext_C corresponds to distinguishing the output of C from uniform.

Lemma B.2. *If an adversary holding the system E can distinguish the output of Ext_C from uniform with probability greater than ε , then he can distinguish the output of the one-bit extractor C from uniform with probability greater than ε/m , given E and a classical system GW , namely*

$$\left\| \rho_{C(X,V)VEGW} - \rho_U \otimes \rho_V \otimes \rho_{EGW} \right\|_{\text{tr}} > \frac{\varepsilon}{m}. \quad (8)$$

The system G has size $H_0(G) \leq rm$, where r is one of the parameters of the weak design (Definition 4.1). W is independent from X , namely $\rho_{XW} = \rho_X \otimes \rho_W$. And $GW \leftrightarrow X \leftrightarrow E$ form a Markov chain, that is $\rho_{XGW} = \sum_{x,g,w} P_{XGW}(x,g,w) |x,g,w\rangle\langle x,g,w| \otimes \rho_E^x$.

Proof. We apply Lemma B.1 to the output Ext_C (Definition 4.2), and get that if $\left\| \rho_{\text{Ext}_C(X,Y)YE} - \rho_U \otimes \rho_Y \otimes \rho_E \right\|_{\text{tr}} > \varepsilon$, then there exists an $i \in [m]$ such that

$$\begin{aligned} & \left\| \sum_{\substack{x,y \\ C(x,y_{S_i})=0}} \frac{p_x}{2^d} |C(x,y_{S_1}) \cdots C(x,y_{S_{i-1}}), y\rangle\langle C(x,y_{S_1}) \cdots C(x,y_{S_{i-1}}), y| \otimes \rho^x \right. \\ & \left. - \sum_{\substack{x,y \\ C(x,y_{S_i})=1}} \frac{p_x}{2^d} |C(x,y_{S_1}) \cdots C(x,y_{S_{i-1}}), y\rangle\langle C(x,y_{S_1}) \cdots C(x,y_{S_{i-1}}), y| \otimes \rho^x \right\|_{\text{tr}} > \frac{\varepsilon}{m}. \end{aligned} \quad (9)$$

We split $y \in \{0,1\}^d$ in two strings of $t = |S_i|$ and $\ell = d - t$ bits, and write $v := y_{S_i}$ and $w := y_{[d] \setminus S_i}$. To simplify the notation, we set $g(w,x,j,v) := C(x,y_{S_j})$. Fix w, x and j , and consider the function $g(w,x,j,\cdot) : \{0,1\}^t \rightarrow \{0,1\}$. This function only depends on $|S_j \cap S_i|$ bits of v . So to describe this function we need a string of at most $2^{|S_j \cap S_i|}$ bits. And to describe $g^{w,x}(\cdot) := g(w,x,1,\cdot) \cdots g(w,x,i-1,\cdot)$, which is the concatenation of the bits of $g(w,x,j,\cdot)$ for $1 \leq j \leq i-1$, we need a string of length at most $\sum_{j=1}^{i-1} 2^{|S_j \cap S_i|}$. So a system G containing a description of $g^{w,x}$ has size at most $H_0(G) \leq \sum_{j=1}^{i-1} 2^{|S_j \cap S_i|}$. We now rewrite Eq. (9) as

$$\begin{aligned} & \left\| \sum_{\substack{x,v,w \\ C(x,v)=0}} \frac{p_x}{2^{t+\ell}} |g^{w,x}(v), v, w\rangle\langle g^{w,x}(v), v, w| \otimes \rho^x \right. \\ & \left. - \sum_{\substack{x,v,w \\ C(x,v)=1}} \frac{p_x}{2^{t+\ell}} |g^{w,x}(v), v, w\rangle\langle g^{w,x}(v), v, w| \otimes \rho^x \right\|_{\text{tr}} > \frac{\varepsilon}{m}. \end{aligned}$$

By providing a complete description of $g^{w,x}$ instead of its value at the point v , we can only increase the trace distance, hence

$$\left\| \sum_{\substack{x,v,w \\ C(x,v)=0}} \frac{p_x}{2^{t+\ell}} |g^{w,x}, v, w\rangle\langle g^{w,x}, v, w| \otimes \rho^x - \sum_{\substack{x,v,w \\ C(x,v)=1}} \frac{p_x}{2^{t+\ell}} |g^{w,x}, v, w\rangle\langle g^{w,x}, v, w| \otimes \rho^x \right\|_{\text{tr}} > \frac{\varepsilon}{m}.$$

By rearranging this a little more we finally get

$$\left\| \rho_{C(X,V)VEGW} - \rho_U \otimes \rho_V \otimes \rho_{EGW} \right\|_{\text{tr}} > \frac{\varepsilon}{m},$$

where G is a classical system of size $H_0(G) \leq \sum_{j=1}^{i-1} 2^{|S_j \cap S_i|}$, and W is a substring of the seed, therefore independent of X . By the definition of weak designs, we have for all $i \in [m]$, $\sum_{j=1}^{i-1} 2^{|S_j \cap S_i|} \leq rm$ for some $r \geq 1$. So $H_0(G) \leq rm$. \square

C List-decodable codes are 1-bit extractors

C.1 Construction

A standard error correcting code guarantees that if the error is small, any string can be uniquely decoded. A list-decodable code guarantees that for a larger (but bounded) error, any string can be decoded to a list of possible messages.

Definition C.1 (list-decodable code). A code $C : \{0, 1\}^n \rightarrow \{0, 1\}^{\bar{n}}$ is said to be (ε, L) -list-decodable if every Hamming ball of relative radius $1/2 - \varepsilon$ in $\{0, 1\}^{\bar{n}}$ contains at most L codewords.

Neither Trevisan [10] nor Raz et al. [14] state it explicitly, but both papers contain an implicit proof that if $C : \{0, 1\}^n \rightarrow \{0, 1\}^{\bar{n}}$ is a (ε, L) -list-decodable code, then

$$\begin{aligned} \text{Ext} : \{0, 1\}^n \times [\bar{n}] &\rightarrow \{0, 1\} \\ (x, y) &\mapsto C(x)_y, \end{aligned}$$

is a $(\log L + \log 1/2\varepsilon, 2\varepsilon)$ -strong extractor. We have rewritten their proof in Section C.2 for completeness.⁷

There exist list-decodable codes with following parameters.

Lemma C.2. *For every $n \in \mathbb{N}$ and $\delta > 0$ there is a code $C_{n,\delta} : \{0, 1\}^n \rightarrow \{0, 1\}^{\bar{n}}$, which is $(\delta, 1/\delta^2)$ -list-decodable, with $\bar{n} = \text{poly}(n, 1/\delta)$. Furthermore, $C_{n,\delta}$ can be evaluated in time $\text{poly}(n, 1/\delta)$ and \bar{n} can be assumed to be a power of 2.*

⁷A slightly more general proof, that *approximate* list-decodable codes are 1-bit extractors can be found in [11, Claim 3.7].

For example, Guruswami et al. [28] combine a Reed-Solomon code with a Hadamard code, obtaining such a list-decodable code with $\bar{n} = O(n/\delta^4)$.

Such codes require all bits of the input x to be read to compute any single bit $C(x)_i$ of the output. If we are interested in so-called *local* codes, we can use a construction by Lu [12].

Lemma C.3 ([12, Corollary 1]). *For every $n \in \mathbb{N}$, $0 < \delta < 1/m$ and constant $0 < \gamma < 1$, there is a code $C_{n,\delta,\gamma} : \{0,1\}^n \rightarrow \{0,1\}^{\bar{n}}$, which is $(\delta, 2^{\gamma n}/\delta^2)$ -list-decodable, with $\bar{n} = \text{poly}(n, 1/\delta)$. Furthermore, for every $i \in [\bar{n}]$, $C_{n,\delta,\gamma}(x)_i$ is the parity of $O(\log(1/m\delta))$ bits of x .*

C.2 Proof

Theorem C.4. *Let $C : \{0,1\}^n \rightarrow \{0,1\}^{\bar{n}}$ be an (ε, L) -list-decodable code. Then the function*

$$C' : \{0,1\}^n \times [\bar{n}] \rightarrow \{0,1\} \\ (x, y) \mapsto C(x)_y,$$

is a $(\log L + \log 1/2\varepsilon, 2\varepsilon)$ -strong extractor.⁸

To prove this theorem we first show that an adversary who can distinguish the bits of $C(X)$ from uniform can construct a string α which is close to $C(X)$ on average (over X). Then using the error correcting properties of the code C , he can reconstruct X . Hence an adversary who can break the extractor must have low min-entropy about X .

Lemma C.5. *Let X and Y be two independent random variables with alphabets $\{0,1\}^n$ and $[n]$ respectively. Let Y be uniformly distributed and X be distributed such that $\frac{1}{2}|X_Y \circ Y - U_1 \circ Y| > \delta$, where U_1 is uniformly distributed on $\{0,1\}$. Then there exists a string $\alpha \in \{0,1\}^n$ with*

$$\Pr \left[d(X, \alpha) \leq \frac{1}{2} - \frac{\delta}{2} \right] > \delta,$$

where $d(\cdot, \cdot)$ is the relative Hamming distance.

Proof. Define $\alpha \in \{0,1\}^n$ to be the concatenation of the most probable bits of X , i.e., $\alpha_y := \arg \max_b P_{X_y}(b)$, where $P_{X_y}(b) = \sum_{x \in \{0,1\}^n} P_X(x)$.

The average relative Hamming distance between X and α is

$$\begin{aligned} \sum_{x \in \{0,1\}^n} P_X(x) d(x, \alpha) &= \frac{1}{n} \sum_{x \in \{0,1\}^n} P_X(x) \sum_{y=1}^n |x_y - \alpha_y| \\ &= \sum_{\substack{x,y \\ x_y \neq \alpha_y}} P_X(x) \\ &= 1 - \frac{1}{n} \sum_{y=1}^n P_X(\alpha_y). \end{aligned}$$

⁸This theorem still holds against a classical adversary with exactly the same parameters.

And since $\frac{1}{2}|X_Y \circ Y - U_1 \circ Y| > \delta$ is equivalent to $\frac{1}{n} \sum_{y=1}^n \max_{b \in \{0,1\}} P_{X_y}(b) > \frac{1}{2} + \delta$, we have

$$\sum_{x \in \{0,1\}^n} P_X(x) d(x, \alpha) < \frac{1}{2} - \delta. \quad (10)$$

We now wish to lower bound the probability that the average Hamming distance is less than $\frac{1}{2} - \frac{\delta}{2}$. Let $B := \{x : d(x, \alpha) \leq \frac{1}{2} - \frac{\delta}{2}\}$ be the set of values $x \in \{0,1\}^n$ meeting this requirement. Then the weight of B , $w(B) := \sum_{x \in B} P_X(x)$, is the quantity we wish to lower bound. It is at its minimum if all $x \in B$ have Hamming distance $d(x, \alpha) = 0$. In which case the average Hamming distance is

$$\sum_{x \in \{0,1\}^n} P_X(x) d(x, \alpha) > (1 - w(B)) \left(\frac{1}{2} - \frac{\delta}{2} \right). \quad (11)$$

Combining Eqs. (10) and (11) we get

$$w(B) > \frac{\delta}{1 - \delta} \geq \delta. \quad \square$$

We are now ready to prove Theorem C.4.

Proof of Theorem C.4. We will show that if it is possible to distinguish $C'(X, Y)$ from uniform with probability at least 2ε , then X must have min-entropy $H_{\min}(X) < \log L + \log 1/2\varepsilon$.

If $\frac{1}{2}|C'(X, Y) \circ Y - U_1 \circ Y \circ E| > 2\varepsilon$, then by Lemma C.5 we know that there exists an $\alpha \in \{0,1\}^n$ such that

$$\Pr \left[d(C(X), \alpha) \leq \frac{1}{2} - \varepsilon \right] > 2\varepsilon,$$

where $d(\cdot, \cdot)$ is the relative Hamming distance.

This means that with probability at least 2ε , X take values x such that $\frac{1}{n}d(C(x), \alpha) \leq \frac{1}{2} - \varepsilon$. So for these values of X , if we choose one of the codewords in the Hamming ball of relative radius $\frac{1}{2} - \varepsilon$ around α uniformly at random as our guess for x , we will have chosen correctly with probability at least $1/L$, since the Hamming ball contains at most L code words. The total probability of guessing X is then at least $2\varepsilon/L$.

Hence by Eq. (3), $H_{\min}(X) < \log L + \log 1/2\varepsilon$. \square

References

- [1] Dmitry Gavinsky, Julia Kempe, Iordanis Kerenidis, Ran Raz, and Ronald de Wolf. Exponential separations for one-way quantum communication complexity, with applications to cryptography. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing, STOC '07*, pages 516–525, New York, NY, USA, 2007. ACM. [doi:10.1145/1250790.1250866, arXiv:quant-ph/0611209].
- [2] Robert König and Renato Renner. Sampling of min-entropy relative to quantum knowledge. eprint, 2007. [arXiv:0712.4291].

- [3] Robert König and Barbara M. Terhal. The bounded-storage model in the presence of a quantum adversary. *IEEE Transactions on Information Theory*, 54(2):749–762, Feb 2008. [doi:10.1109/TIT.2007.913245, arXiv:quant-ph/0608101].
- [4] Charles H. Bennett, Gilles Brassard, Claude Crépeau, and Ueli Maurer. Generalized privacy amplification. *IEEE Transaction on Information Theory*, 41(6):1915–1923, November 1995. [doi:10.1109/18.476316].
- [5] David Zuckerman. General weak random sources. In *Proceedings of the 31st IEEE Symposium on Foundations of Computer Science, FOCS '90*, pages 534–543, Los Alamitos, CA, USA, 1990. IEEE Computer Society. [doi:10.1109/FSCS.1990.89574].
- [6] Ronen Shaltiel. Recent developments in explicit constructions of extractors. *Bulletin of the European Association for Theoretical Computer Science*, 77:67–95, June 2002.
- [7] Renato Renner. *Security of Quantum Key Distribution*. PhD thesis, Swiss Federal Institute of Technology Zurich, September 2005. [arXiv:quant-ph/0512258].
- [8] Marco Tomamichel, Christian Schaffner, Adam Smith, and Renato Renner. Leftover hashing in the presence of quantum side information. To appear, 2010.
- [9] Amnon Ta-Shma. Short seed extractors against quantum storage. In *Proceedings of the 41st annual ACM symposium on Theory of computing, STOC '09*, pages 401–408, New York, NY, USA, 2009. ACM. [doi:10.1145/1536414.1536470, arXiv:0808.1994].
- [10] Luca Trevisan. Extractors and pseudorandom generators. *Journal of the ACM*, 48(4):860–879, 2001. [doi:10.1145/502090.502099].
- [11] Anindya De and Thomas Vidick. Near-optimal extractors against quantum storage. To appear at QIP '10. eprint, 2009. [arXiv:0911.4680].
- [12] Chi-Jen Lu. Encryption against storage-bounded adversaries from on-line strong extractors. *Journal of Cryptology*, 17(1):27–42, 2004. [doi:10.1007/s00145-003-0217-1].
- [13] Salil P. Vadhan. Constructing locally computable extractors and cryptosystems in the bounded-storage model. *Journal of Cryptology*, 17(1):43–77, 2004. [doi:10.1007/s00145-003-0237-x].
- [14] Ran Raz, Omer Reingold, and Salil Vadhan. Extracting all the randomness and reducing the error in Trevisan’s extractors. *Journal of Computer and System Sciences*, 65(1):97–128, 2002. [doi:10.1006/jcss.2002.1824].
- [15] Robert König, Renato Renner, and Christian Schaffner. The operational meaning of min- and max-entropy. *IEEE Transactions on Information Theory*, 55(9):4337–4347, 2009. [arXiv:0807.1338].
- [16] Marco Tomamichel, Roger Colbeck, and Renato Renner. Duality between smooth min- and max-entropies. eprint, 2009. [arXiv:0907.5238].

- [17] Charles H. Bennett, Gilles Brassard, and Jean-Marc Robert. Privacy amplification by public discussion. *SIAM Journal on Computing*, 17(2):210–229, 1988. [doi:10.1137/0217014].
- [18] Jaikumar Radhakrishnan and Amnon Ta-Shma. Bounds for dispersers, extractors, and depth-two superconcentrators. *SIAM Journal on Discrete Mathematics*, 13(1):2–24, 2000. [doi:10.1137/S0895480197329508].
- [19] Tzvika Hartman and Ran Raz. On the distribution of the number of roots of polynomials and explicit weak designs. *Random Structures and Algorithms*, 23(3):235–263, 2003. [doi:10.1002/rsa.10095].
- [20] Ueli M. Maurer. Conditionally-perfect secrecy and a provably-secure randomized cipher. *Journal of Cryptology*, 5(1):53–66, 1992.
- [21] Amnon Ta-Shma, David Zuckerman, and Shmuel Safra. Extractors from reed-muller codes. *Journal of Computer and System Sciences*, 72(5):786–812, 2006. [doi:10.1016/j.jcss.2005.05.010].
- [22] Ronen Shaltiel and Christopher Umans. Simple extractors for all min-entropies and a new pseudorandom generator. *Journal of the ACM*, 52(2):172–216, 2005. [doi:10.1145/1059513.1059516].
- [23] Aravind Srinivasan and David Zuckerman. Computing with very weak random sources. *SIAM Journal on Computing*, 28(4):1433–1459, 1999. [doi:10.1137/S009753979630091X].
- [24] Russell Impagliazzo, Ronen Shaltiel, and Avi Wigderson. Extractors and pseudo-random generators with optimal seed length. In *Proceedings of the thirty-second annual ACM symposium on Theory of computing, STOC '00*, pages 1–10, New York, NY, USA, 2000. ACM. [doi:10.1145/335305.335306].
- [25] Amnon Ta-Shma, Christopher Umans, and David Zuckerman. Lossless condensers, unbalanced expanders, and extractors. In *Proceedings of the thirty-third annual ACM symposium on Theory of computing, STOC '01*, pages 143–152, New York, NY, USA, 2001. ACM. [doi:10.1145/380752.380790].
- [26] Avi Wigderson and David Zuckerman. Expanders that beat the eigenvalue bound: Explicit construction and applications. *Combinatorica*, 19(1):125–138, 1999. [doi:10.1007/s004930050049].
- [27] Andrew Chi-Chih Yao. Theory and applications of trapdoor functions (extended abstract). In *Proceedings of the 23rd IEEE Symposium on Foundations of Computer Science, FOCS '82*, pages 80–91. IEEE, November 1982.
- [28] Venkatesan Guruswami, Johan Håstad, Madhu Sudan, and David Zuckerman. Combinatorial bounds for list decoding. *IEEE Transactions on Information Theory*, 48(5):1021–1034, 2002.