Trevisan's Extractor in the Presence of Quantum Side Information

Citation	De, Anindya et al. "Trevisan's Extractor in the Presence of Quantum Side Information." SIAM Journal on Computing 41.4 (2012): 915–940. © 2012, Society for Industrial and Applied Mathematics
As Published	http://dx.doi.org/10.1137/100813683
Publisher	Society for Industrial and Applied Mathematics
Version	Final published version
Citable link	http://hdl.handle.net/1721.1/77888
Terms of Use	Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.

SIAM J. COMPUT.
Vol. 41, No. 4, pp. 915–940

TREVISAN'S EXTRACTOR IN THE PRESENCE OF QUANTUM SIDE INFORMATION[*]

ANINDYA DE[†], CHRISTOPHER PORTMANN[‡], THOMAS VIDICK[§], AND RENATO RENNER[‡]

Abstract. Randomness extraction involves the processing of purely classical information and is therefore usually studied with 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 show that the well-known construction paradigm for extractors proposed by Trevisan is sound in the presence of quantum side information. We exploit the modularity of this paradigm to give several concrete extractor constructions, which, e.g., extract all the conditional (smooth) min-entropy of the source using a seed of length polylogarithmic in the input, or only require the seed to be weakly random.

Key words. randomness extractors, quantum information, quantum cryptography, smooth min-entropy

AMS subject classifications. 68Q01, 81P45, 81P94, 94A17

DOI. 10.1137/100813683

1. Introduction.

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

$$H_{\min}(X) \geq k \implies Z := \text{Ext}(X, Y) \text{ statistically close to uniform.} \tag{1.1}$$

The min-entropy of a random variable X is directly related to the probability of correctly guessing the value of X using an optimal strategy: $2^{-H_{\min}(X)} = \max_x P_X(x)$. Hence criterion (1.1) can be interpreted operationally: if the maximum probability of successfully guessing the input of the extractor, X, is sufficiently low, then its output is statistically close to uniform.

[*] Received by the editors November 3, 2010; accepted for publication (in revised form) June 5, 2012; published electronically August 21, 2012. A preliminary version of this work [5] appeared at STOC '10.

http://www.siam.org/journals/sicomp/41-4/81368.html

[†] Computer Science Division, University of California, Berkeley, CA 94720 (anindya@cs.berkeley.edu). This author's work was supported by the Berkeley fellowship for graduate study and by NSF-CCF-1017403.

[‡] Institute for Theoretical Physics, ETH Zurich, 8093 Zurich, Switzerland, and Group of Applied Physics, University of Geneva, 1211 Geneva, Switzerland (chportma@phys.ethz.ch, renner@phys.ethz.ch). The second author's work was supported in part by the Vienna Science and Technology Fund (WWTF) through project ICT10-067 (HiPANQ). These authors' work was supported by the Swiss National Science Foundation (via grants 200021-119868 and 200020-135048, and the National Centre of Competence in Research "Quantum Science and Technology") and by the European Research Council—ERC (grant 258932).

[§] Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139 (vidick@csail.mit.edu). This author's work was supported by the National Science Foundation under grant 0844626.

The randomness of a value X always depends on the information one has about it; this is called *side information* in what follows. In cryptography, for instance, a key is supposed to be uniformly random from the point of view of an adversary, who may have access to messages exchanged by the honest parties, which we would therefore consider as side information. Here, extractors are typically used for *privacy amplification* [4, 3], i.e., to turn a partially secure raw key (about which the adversary may have nontrivial information) into a perfectly secure key. We thus demand that the extractor output be uniform with respect to the side information held by the adversary. Another example is *randomness recycling* in a computation, which can be done using extractors [10]. The aim is that the recycled randomness be independent of the outputs of previous computations, which are therefore considered as side information.

In the following, we make side information explicit and denote it by E. The notions of randomness that we are going to use, such as the *guessing probability*, *min-entropy*, or the *uniformity* of a random variable, must then be defined with respect to E. We can naturally reformulate criterion (1.1) as

$$H_{\min}(X|E) \geq k \implies Z := \mathrm{Ext}(X,Y) \text{ statistically close to uniform conditioned on } E, \tag{1.2}$$

where $H_{\min}(X|E)$ is the conditional min-entropy, formally defined in section 2.2. This conditioning naturally extends the operational interpretation of the min-entropy to scenarios with explicit side information; i.e., $2^{-H_{\min}(X|E)}$ is the maximum probability of correctly guessing X, given access to side information E [13].

Interestingly, the relationship between the two criteria (1.1) and (1.2) 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 are essentially equivalent (see Lemma 3.3 for a precise statement). But in the general case where E is a quantum system, criterion (1.2) is *strictly stronger* than (1.1): it was shown in [6] that there exist extractors that fulfill (1.1) but for which (1.2) fails (see also [12] for a discussion).

Since our world is inherently nonclassical, it is of particular importance that (1.2) rather than the weaker criterion (1.1) be taken as the relevant criterion for the definition of extractors. In cryptography, for instance, there is generally nothing that prevents an adversary from holding quantum side information. In fact, even if a cryptographic scheme is purely classical, an adversary may acquire information using a nonclassical attack strategy. Hence, when using extractors for privacy amplification, (1.1) does not generally imply security. A similar situation may arise in the context of randomness recycling. If we run a (simulation of) a quantum system E using randomness X, approximately $H_{\min}(X|E)$ bits of X can be reused. If we now, in an attempt to recycle the randomness, apply a function Ext which fulfills (1.1) but not (1.2), the output Z may still be correlated to the system E.

It is known that the conditional min-entropy accurately characterizes the maximum amount of uniform randomness that can be extracted from X while being independent from E. (More precisely, the *smooth conditional min-entropy*, an entropy measure derived from $H_{\min}(X|E)$ by maximizing the latter over all states in an ε-neighborhood, is an upper bound on the amount of uniform randomness that can be extracted; see section 2.2 and [22] for details.) In other words, the characterization of extractors in terms of $H_{\min}(X|E)$ is essentially optimal, and one may thus argue that criterion (1.2) is indeed the correct definition for randomness extraction (see also [22, 12, 14]). In this work, we follow this line of argument and call an extractor *quantum-proof* if it satisfies (1.2) (see section 3.1).

We note that there have been alternative proposals in the literature for defining extractors in the context of quantum side information, which, however, do not satisfy the above optimality condition. One prominent example is the bounded storage model (see section 5.3), where the (quantum) side information E is characterized by the number of qubits, $H_0(E)$, required to store it. In this model, the entropy $H_{\min}(X|E)$ of a source X conditioned on E is lower-bounded by $H_{\min}(X) - H_0(E)$. However, this characterization of side information is strictly weaker than that using $H_{\min}(X|E)$: there are sources X and nontrivial side information E such that $H_{\min}(X) - H_0(E) \ll H_{\min}(X|E)$.[1] In particular, even if an extractor can provably extract $H_{\min}(X) - H_0(E)$ bits of uniform (with respect to E) randomness from a source X, we do not know whether the same extractor can attain the optimal $H_{\min}(X|E)$ bits. Note also that the same considerations apply to the purely classical case. In fact, no recent work defines classical extractors for randomness sources with side information stored in bounded classical memories.[2]

Finally, we remark that the increased generality attained by the notion of quantum-proof extractors used here is crucial for applications. For example, in quantum key distribution, where extractors are used for privacy amplification [22], it is generally impossible to bound the adversary's memory size.

1.1. Related results.

In the standard literature on randomness extraction, constructions of extractors are usually shown to fulfill criterion (1.1) for certain values of the threshold k (see [38] as well as [23] for an overview). However, only a few constructions have been shown to fulfill (1.2) with arbitrary quantum side information E. Among them is two-universal hashing [22, 32], constructions based on the sample-and-hash approach [12], as well as all extractors with one-bit output [14].

Recently, Ta-Shma [28] studied Trevisan's construction of extractors [33] in the bounded quantum storage model. The result was a breakthrough, implying, for the first time, the existence of quantum-proof extractors requiring only short seeds (logarithmic in the input length). Unfortunately, Ta-Shma's result is proved in the bounded quantum storage model. More precisely, he requires the output length to be much smaller than the min-entropy of the original data: it scales as $(H_{\min}(X)/H_0(E))^{1/c}$, where $c > 1$ is a constant.

Subsequent to this work, Ben-Aroya and Ta-Shma [2] showed how two versions of Trevisan's extractor, shown to be quantum-proof in this paper, can be combined to extract a constant fraction of the min-entropy of an n-bit source with a seed of length $O(\log n)$, when $H_{\min}(X|E) > n/2$. This is better than the straightforward application of Trevisan's extractor analyzed here, which requires $O(\log^2 n)$ bits of seed for the same output size (but works for any $H_{\min}(X|E)$).

1.2. Our contribution.

In this work, we show that the performance of Trevisan's extractor does not suffer in the presence of quantum side information. This

[1] This can easily be seen by considering the following example. Let X be uniformly distributed on $\{0,1\}^n$, and let E be X with each bit flipped with constant probability $\varepsilon < 1/2$. Then $H_{\min}(X|E) = \Theta(n)$, but $H_{\min}(X) - H_0(E) = 0$.

[2] Restricting the class of randomness sources further by bounding their min-entropy can have advantages. For example, if we consider only bit-fixing sources, or sources generated by a random walk on a Markov chain, then the extractor can be deterministic. (See [23] for a brief overview of restricted families of sources studied in the literature.) There is, however, no known advantage (e.g., in terms of seed length) in considering only input sources with side information stored in a memory of bounded size, whether it is classical or quantum.

TABLE 1.1

Plugging various weak designs and one-bit extractors into Trevisan's construction, we obtain these concrete extractors. Here n is the input length, α and γ are arbitrary constants such that $0 < \gamma < \alpha \leq 1$, and $\frac{1}{2} < \beta < 1$ is a specific constant. For succinctness we take the error to be $\varepsilon = \mathrm{poly}(1/n)$ and give the output length m up to a term in $O(\log n)$. We refer to the corresponding corollaries in section 5 for the exact seed and output lengths.

	Min-entropy	Output length	Seed length	Note
Corollary 5.2	Any k	$m = k$	$d = O(\log^3 n)$	Optimized output length
Corollary 5.3	$k = n^{\alpha}$	$m = n^{\alpha-\gamma}$	$d = O(\log n)$	Optimized seed length
Corollary 5.4	$k = \alpha n$	$m = (\alpha - \gamma)n$	$d = O(\log^2 n)$	Local extractor
Corollary 5.5	$k = n^{\alpha}$	$m = n^{\alpha-\gamma}$	$d = O(\log n)$	Seed with min-entropy βd

improves on the best previously known result [28] in two major ways. First, we prove our results in the most general model, where the min-entropy of the source is measured relative to quantum side information (criterion (1.2)). Second, we show that the output length of the extractor can be close to the optimal conditional min-entropy $H_{\min}(X|E)$ (see Corollary 5.2 for the exact parameters).[3] This provides the first proof of soundness for an extractor with polylogarithmic seed meeting (1.2) in the presence of arbitrary quantum side information.

More generally, we show that a whole class of extractors is quantum-proof. It has been observed, by, e.g., Lu [15] and Vadhan [34], that Trevisan's extractor [33] (and variations of it, such as [21]) can be seen as a concatenation of the outputs of a one-bit extractor with different pseudorandom seeds. Since the proof of the extractor property is independent of the type of the underlying one-bit extractor (and to some extent the construction of the pseudorandom seeds), our result is valid for a generic scheme (defined in section 4.1, Definition 4.2). We find that the performance of this generic scheme in the context of quantum side information (section 4.2.1, Theorem 4.6) is roughly equivalent to the (known) case of purely classical side information [21].

In practical situations where quantum-proof extractors are used, e.g., privacy amplification in quantum key distribution [22], the players do not necessarily have access to a uniform source of randomness. We therefore separately analyze the situation where the seed is only weakly random and show that Trevisan's extractor is quantum-proof in that setting as well (section 4.2.2, Theorem 4.7).

By "plugging" various one-bit extractors and pseudorandom seeds into the generic scheme, we obtain different final constructions, optimized for different needs, e.g., maximizing the output length, minimizing the seed, or using a nonuniform seed. In Table 1.1 we give a brief overview of the final constructions proposed.

1.3. Proof technique. The proof proceeds by contradiction. We first assume that a player holding the side information E can distinguish the output from uniform with probability greater than ε. We then show that such a player can reconstruct the input X with high probability, which means that X must have low min-entropy ($H_{\min}(X|E) < k$). Taking the contrapositive proves that the extractor is sound.

Trevisan [33] originally proved the soundness of his extractor this way. His construction starts by encoding the source X using a list-decodable code C. The output of the extractor then consists of certain bits of $C(X)$, which are specified by the seed and a construction called a (weak) design [17, 21]. (See section 4.1 for a precise de-

[3] In the conference version of this paper [5], the authors showed that a similar result could be obtained in the more restricted bounded-storage model.

scription of Trevisan's extractor.) His proof can then be broken down into two steps. He first shows that a player who can distinguish the output from uniform can guess a random bit of $C(X)$. In the second step, he shows that such a player can reconstruct X.

Proving the soundness of Trevisan's extractor in the quantum min-entropy framework requires some important changes. In order to better explain these new elements, it will be useful to first give a brief overview of the main steps that go into Ta-Shma's proof [28]. For the sake of contradiction, assume that there is a test T which performs a measurement on the side information E in order to distinguish the output from uniform with advantage ε. Using a standard hybrid argument along with properties of the (weak) design, one can then construct a new test T' (using a little extra classical advice about X), which predicts a random bit of $C(X)$ with probability $\frac{1}{2}+\frac{\varepsilon}{m}$, where m is the number of output bits. Further, T' makes exactly *one* query to T.

The proof in [28] proceeds by showing how from such a test one can construct another test T'' which predicts any bit of X with probability 0.99 and queries T' at most $q=(m/\varepsilon)^c$ times ($c=15$ for the code in [28]). This gives a random access code (RAC) [1] for X; however, since it requires q queries to the side information E, the no-cloning theorem forces us to see it as querying a single system of length $qH_0(E)$ (recall that Ta-Shma's result was proved in the bounded storage model, where one bounds the information provided by E by its number of qubits $H_0(E)$). Finally, using a new bound on the dimension of RACs [28], one finds that $H_{\min}(X) \gtrsim m^c H_0(E)$; hence $m \lesssim (H_{\min}(X)/H_0(E))^{1/c}$, where for simplicity we have taken the error ε to be a constant.

Our proof improves upon Ta-Shma's through two major changes. First, we model the side information E explicitly, instead of viewing it as an oracle which one queries. Indeed, the measurement performed by the test T' to predict the bits of $C(X)$ will be different from the measurement performed by T'' to reconstruct X, and this cannot be captured by the "oracle side-information" model of Ta-Shma. We thus show (in section 4.2, Proposition 4.4) that if the output of the extractor can be distinguished from uniform with probability $\frac{1}{2}+\varepsilon$ by a player holding the side information E, then the bits of $C(X)$ can be guessed with probability $\frac{1}{2}+\frac{\varepsilon}{m}$ by a player holding E and some extra small classical information G.

Second, we depart from the reconstruction paradigm at the heart of the second half of the proof of both Trevisan's and Ta-Shma's results. Instead of explicitly defining the measurement and computation necessary to reconstruct X, we use the fact that for any list-decodable code $C:\{0,1\}^n \to \{0,1\}^{\bar{n}}$ the function

$$\begin{aligned} C' : \{0,1\}^n \times [\bar{n}] &\to \{0,1\}, \\ (x,i) &\mapsto C(x)_i \end{aligned}$$

is a one-bit extractor according to (1.1) (see Appendix D for more details). It was, however, proved by König and Terhal [14] that in the one-bit setting the more general criterion (1.2) is essentially equivalent to the usual criterion (1.1). This result lets us conclude directly that the input X must have low min-entropy relative to the quantum side information E.

This proof structure results in a very modular extractor construction paradigm, which allows arbitrary one-bit extractors and pseudorandom seeds to be plugged in, producing many different final constructions, some of which are given in Table 1.1 and detailed in section 5.

1.4. Organization of the paper. We first define the necessary technical tools in section 2, in particular the conditional min-entropy. In section 3 we give formal definitions of extractors and discuss how much randomness can be extracted from a given source. Section 4 contains the description of Trevisan's extractor construction paradigm and our main result: a proof that this construction paradigm is sound in the presence of quantum side information in the cases of both uniform and weakly random seeds. Then in section 5 we plug into Trevisan's construction various one-bit extractors and pseudorandom seed constructions, resulting in various different extractors. For example, section 5.1 contains a construction which is nearly optimal in the amount of randomness extracted (which is identical to the best known bound in the classical case [21] for Trevisan's extractor), and section 5.4 gives an extractor which is still sound if there is a small linear entropy loss in the seed. Finally, in section 6, we give a brief outlook on further work. In particular, we mention a few classical results which modify and improve Trevisan's extractor, but for which the soundness in the presence of quantum side information does not seem to follow immediately from this work.

The appendix contains many technical sections and lemmas which are not essential for understanding Trevisan's extractor but are nonetheless an important part of the construction and proof. Appendix A develops a bit more the general theory of extractors: it contains two subsections which, respectively, define extractors for weakly random seeds and show how to compose extractors to obtain more randomness from the same source. In Appendix B we state several technical lemmas: min-entropy chain rules and the details of the reduction from Trevisan's construction to the underlying one-bit extractor. Appendix C contains previously known constructions for one-bit extractors and weak designs which we use in this work and plug into Trevisan's paradigm. Finally, in Appendix D we give a proof that list-decodable codes are one-bit extractors.

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 ith bit of x, and x_S for the string formed by the bits of x at the positions given by the elements of S.

$\mathcal{H}$ always denotes a finite-dimensional Hilbert space. We denote by $\mathcal{P}(\mathcal{H})$ the set of positive semidefinite operators on $\mathcal{H}$. We define the set of normalized quantum states $\mathcal{S}(\mathcal{H}) := \{\rho \in \mathcal{P}(\mathcal{H}) : \operatorname{tr}\rho = 1\}$ and the set of subnormalized quantum states $\mathcal{S}_{\leq}(\mathcal{H}) := \{\rho \in \mathcal{P}(\mathcal{H}) : \operatorname{tr}\rho \leq 1\}$.

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. $\rho_A = \operatorname{tr}_B(\rho_{AB})$ and $\rho_B = \operatorname{tr}_A(\rho_{AB})$ 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 $\rho_X = \sum_{x \in X} p_x |x\rangle\langle x|$, 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 $\rho_{XB} = \sum_{x \in \mathcal{X}} p_x |x\rangle\langle x| \otimes \rho_B^x$.

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

2.2. Min-entropy. To measure how much randomness a source contains and can be extracted, we need to use the *smooth conditional min-entropy.* This entropy measure was first defined by Renner [22] 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. Before defining this notion, we first state a *nonsmooth* version.

DEFINITION 2.1 (conditional min-entropy [22]). *Let $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$. The* min-entropy *of A conditioned on B is defined as*

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

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

This definition has a simple operational interpretation when the first system is classical, which is the case we consider. König, Renner, and Schaffner [13] 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,

$$H_{\min}(X|B)_\rho = -\log p_{\text{guess}}(X|B)_\rho, \tag{2.1}$$

where $p_{\text{guess}}(X|B)$ is the maximum probability of guessing X given B, namely,

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

where the maximum is taken over all positive operator-valued measure (POVMs) $\{E_B^x\}_{x\in\mathcal{X}}$ on B. If the system B is empty, then the min-entropy of X reduces to the Renyi entropy of order infinity, $H_{\min}(X) = -\log \max_{x\in\mathcal{X}} p_x$ (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.

The *smooth* min-entropy then consists of maximizing the min-entropy over all subnormalized 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 potentially much more entropy. (See section 3.2 for more details.)

DEFINITION 2.2 (smooth min-entropy [22, 31]). *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*

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

where $\mathcal{B}^\varepsilon(\rho_{AB}) \subseteq \mathcal{S}_{\leq}(\mathcal{H}_{AB})$ is a ball of subnormalized states of radius ε around ρ_{AB}.[4]

3. Extractors.

3.1. Extractors, side information, and privacy amplification.

An extractor $\text{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ 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 almost uniform. The extractor is said to be strong if the output is approximately independent of the seed.

DEFINITION 3.1 (strong extractor [18]). *A function* $\text{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *is a* (k,ε)-strong extractor with uniform seed *if for all distributions X with*

[4] The distance measure used in this definition is the *purified distance* [31], $P(\rho,\sigma) := \sqrt{1 - F(\rho,\sigma)^2}$, where $F(\cdot,\cdot)$ is the fidelity. The only property of the purified distance that we need in this work is that it upper-bounds the trace distance, i.e., $P(\rho,\sigma) \geq \frac{1}{2}\|\rho - \sigma\|_{\text{tr}}$. We refer the reader to [31] for a formal definition of the purified distance (and fidelity) on subnormalized states and a discussion of its advantages.

min-entropy $H_{\min}(X) \geq k$ *and a uniform seed* Y *we have*[5]

$$\frac{1}{2}\|\rho_{\mathrm{Ext}(X,Y)Y} - \rho_{U_m} \otimes \rho_Y\|_{\mathrm{tr}} \leq \varepsilon,$$

where ρ_{U_m} *is the fully mixed state on a system of dimension* 2^m*.*

Using the connection between min-entropy and guessing probability ((2.1)), a (k,ε)-strong extractor can be seen as a function which guarantees that if the guessing probability of X is not too high ($p_{\mathrm{guess}}(X) \leq 2^{-k}$), then it produces 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 by E, which may be represented by the state of a quantum system. A function Ext is then an extractor if, when the probability of guessing X *given* E is not too high, Ext can produce a random variable $\mathrm{Ext}(X,Y)$ which is approximately uniform and independent from the seed Y and the side information E. Equivalently, one may think of a *privacy amplification* scenario [4, 3], where E is the information available to an adversary and where the goal is to turn weakly secret data X into a *secret* key $\mathrm{Ext}(X,Y)$, where the seed Y is assumed to be public. (In typical key agreement protocols, the seed 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 (quantum-proof strong extractor [12, section 2.6]). *A function* Ext : $\{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *is a* quantum-proof *(or simply* quantum*)* (k,ε)-strong extractor with uniform seed *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*[6]

$$\frac{1}{2}\|\rho_{\mathrm{Ext}(X,Y)YE} - \rho_{U_m} \otimes \rho_Y \otimes \rho_E\|_{\mathrm{tr}} \leq \varepsilon, \tag{3.1}$$

where ρ_{U_m} *is the fully mixed state on a system of dimension* 2^m*.*

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

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

Lemma 3.3 (see [12, section 2.5] and [14, Proposition 1]). *Any* (k,ε)*-strong extractor is a classical-proof* $(k + \log 1/\varepsilon, 2\varepsilon)$*-strong extractor.*

However, if the system E is quantum, this does not necessarily hold. Gavinsky et al. [6] give an example of a (k,ε)-strong extractor that breaks down in the presence of quantum side information, even when $H_{\min}(X|E)$ is significantly larger than k.

Remark 3.4. In this section we defined extractors with a uniform seed, as this is the most common way of defining them. Instead one could use a seed which is only weakly random, but require it to have a min-entropy larger than a given threshold,

[5] A more standard classical notation would be $\frac{1}{2}\|\mathrm{Ext}(X,Y) \circ Y - U_m \circ Y\| \leq \varepsilon$, 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 trace distance reduces to the variational distance, we use the quantum notation for compatibility with the rest of this work.

[6] The authors of [32] substitute $\exists \sigma_{YE}$ s.t. $\frac{1}{2}\|\rho_{\mathrm{Ext}(X,Y)YE} - \rho_{U_m} \otimes \sigma_{YE}\|_{\mathrm{tr}} \leq \varepsilon$ for (3.1). This results in a weaker definition which does not offer the same composability guarantees. In particular, Lemma A.4 does not hold with the same parameters when extractors are defined as in [32].

$H_{\min}(Y) \geq s$. The seed must still be independent from the input and the side information. Since having access to a uniform seed is often an unrealistic assumption, it is much more useful for practical applications to define and prove the soundness of extractors with a weakly random seed. We redefine extractors formally this way in Appendix A.1, and show in section 4.2.2 that Trevisan's extractor is still quantum-proof in this setting.

All the considerations of this section, in particular Lemma 3.3 and the gap between classical and quantum side-information, also apply if the seed is only weakly random. In the following, when we talk about a strong extractor without specifying the nature of the seed, we are referring to both uniform seeded and weakly random seeded extractors.

3.2. Extracting more randomness. Radhakrishnan and Ta-Shma [19] have shown that a (k, ε)-strong extractor $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ will necessarily have

$$m \leq k - 2\log 1/\varepsilon + O(1). \tag{3.2}$$

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 $\frac{1}{2}\|\rho_X - \rho_{U_n}\|_{\mathrm{tr}} \leq \frac{1}{n}$. So we already have n bits of nearly uniform randomness, exponentially more than the min-entropy suggests.

In the case of quantum extractors, similar examples can be found, e.g., in [31, Remark 22]. However, an upper bound on the extractable randomness can be obtained by replacing the min-entropy by the *smooth* min-entropy (Definition 2.2). More precisely, the total number of ε-uniform bits that can be extracted in the presence of side information E can never exceed $H^{\varepsilon}_{\min}(X|E)$ [22, section 5.6].

Conversely, the next 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 satisfy only $H^{\varepsilon'}_{\min}(X|E) \geq k$, a much weaker requirement in some cases.

LEMMA 3.5. *If* $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *is a quantum-proof* (k, ε)*-strong extractor, then for any state* ρ_{XE} *and any* $\varepsilon' > 0$ *with* $H^{\varepsilon'}_{\min}(X|E)_\rho \geq k$,

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

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

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

In the second inequality above we twice used the fact that a trace-preserving quantum operation can only decrease the trace distance. And in the last line we used the fact that the purified distance—used in the smooth min-entropy definition (Definition 2.2)—upper-bounds the trace distance. □

Remark 3.6. Since a (k, ε)-strong extractor can be applied to any 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. From (3.2) we know that an extractor has optimal entropy loss if $\Delta = 2 \log 1/\varepsilon + O(1)$.

4. Constructing m-bit extractors from one-bit extractors and weak designs. In this section we prove our main result: we show that Trevisan's extractor paradigm [33]—which shows how to construct an m-bit extractor from any (classical) one-bit strong extractor—is sound in the presence of quantum side information.

This construction paradigm can be seen as a derandomization of the simple concatenation of the outputs of a one-bit extractor applied m times to the same input with different (independent) seeds. The construction with independent seeds needs a total seed of length $d = mt$, where t is the length of the seed of the one-bit extractor. Trevisan [33] shows how to do this using only $d = \text{poly}(t, \log m)$ bits of seed and proves that it is sound when no side information is present.[7] We combine a combinatorial construction called weak designs by Raz, Reingold, and Vadhan [21], which they use to improve Trevisan's extractor, and a previous observation by two of the authors [5] that, since one-bit extractors were shown to be quantum-proof by König and Terhal [14], Trevisan's extractor is also quantum-proof.

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

4.1. Description of Trevisan's construction. In order to shorten the seed while still outputting m bits, in Trevisan's extractor construction paradigm the seed is treated as a string of length $d < mt$, which is then split into m overlapping blocks of t bits, each of which is used as a (different) seed for the one-bit extractor. Let $y \in \{0,1\}^d$ be the total seed. To specify the seeds for each application of the one-bit extractor we need m sets $S_1, \ldots, S_m \subset [d]$ of size $|S_i| = t$ for all i. The seeds for the different runs of the one-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 one-bit extractor must, however, be nearly independent. To achieve this, Nisan and Wigderson [17] proposed minimizing the overlap $|S_i \cap S_j|$ between the sets, and Trevisan used this idea in his original work [33]. Raz, Reingold, and Vadhan [21] improved this, showing that it is sufficient for these sets to meet the conditions of a *weak design.*[8]

DEFINITION 4.1 (weak design [21, Definition 5]). *A family of sets $S_1, \ldots, S_m \subset [d]$ is a* weak (t, r)-design *if the following hold:*

[7] Trevisan's original paper does not explicitly define his extractor as a pseudorandom concatenation of a one-bit extractor. It has, however, been noted in, e.g., [15, 34], that this is basically what Trevisan's extractor does.

[8] The 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 to use the version of [8], since it simplifies the notation without changing the design constructions.

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 Trevisan's generic extractor construction.

Definition 4.2 (Trevisan's extractor [33]). *For a one-bit extractor $C : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}$, which uses a (not necessarily uniform) seed of length t, and for a weak (t, r)-design $S_1, \dots, S_m \subset [d]$, we define the m-bit extractor $\mathrm{Ext}_C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ as*

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

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 one-bit extractor C. In section 5 we will give concrete instantiations of weak designs and one-bit extractors, achieving various entropy losses and seed sizes. The size of the seed will always be $d = \mathrm{poly}(\log n)$ if the error is $\varepsilon = \mathrm{poly}(1/n)$. 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 quantum-proof strong extractor. The first step follows the structure of the classical proof [33, 21]. We show that a player holding the side information and who can distinguish the output of the extractor Ext_C from uniform can—given a little extra information—distinguish the output of the underlying one-bit extractor C from uniform. This is summed up in the following proposition.

Proposition 4.4. *Let X be a classical random variable correlated to some quantum system E; let Y be a (not necessarily uniform) seed, independent from XE; and let*

$$\|\rho_{\mathrm{Ext}_C(X,Y)YE} - \rho_{U_m} \otimes \rho_Y \otimes \rho_E\|_{\mathrm{tr}} > \varepsilon, \tag{4.1}$$

where Ext_C is the extractor from Definition 4.2. Then there exists a fixed partition of the seed Y into two substrings, V and W, and a classical random variable G such that G has size $H_0(G) \leq rm$, where r is one of the parameters of the weak design (Definition 4.1), $V \leftrightarrow W \leftrightarrow G$ form a Markov chain,[9] *and*

$$\|\rho_{C(X,V)VWGE} - \rho_{U_1} \otimes \rho_{VWGE}\|_{\mathrm{tr}} > \frac{\varepsilon}{m}. \tag{4.2}$$

We provide a proof of Proposition 4.4 in Appendix B.2, where it is restated as Proposition B.5.[10]

For readers familiar with Trevisan's scheme [33, 21], we briefly sketch the correspondence between the variables of Proposition 4.4 and quantities analyzed in Trevisan's construction. Trevisan's proof proceeds by assuming by contradiction that there exists a player, holding E, who can distinguish between the output of the extractor and the uniform distribution (4.1). Part of the seed is then fixed (this corresponds to W in the above statement), and some classical advice is taken (this corresponds

[9] Three random variables are said to form a Markov chain $X \leftrightarrow Y \leftrightarrow Z$ if for all x, y, z we have $P_{Z|YX}(z|y,x) = P_{Z|Y}(z|y)$ or equivalently $P_{ZX|Y}(z,x|y) = P_{Z|Y}(z|y)P_{X|Y}(x|y)$.

[10] Note that Ta-Shma [28] has already implicitly proved that this proposition must hold in the presence of quantum side information, by arguing that the side information can be viewed as an oracle. The present statement is a strict generalization of that reasoning, which allows conditional min-entropy as well as nonuniform seeds to be used.

to G in the above statement) to construct another player who can distinguish a specific bit of the output from uniform. But since a specific bit of Trevisan's extractor is just the underlying one-bit extractor applied to a substring of the seed (V in the above statement), this new player (who holds WGE) can distinguish the output of the one-bit extractor from uniform (see (4.2)).

In the classical case Proposition 4.4 would be sufficient to prove the soundness of Trevisan's scheme, since it shows that if a player can distinguish Ext_C from uniform, then he can distinguish C from uniform given a few extra advice bits, which contradicts the assumption that C is an extractor.[11] But since our assumption is that the underlying one-bit extractor is only classical-proof, we still need to show that the quantum player who can distinguish $C(X, V)$ from uniform is not more powerful than a classical player, and so if he can distinguish the output of C from uniform, so can a classical player. This has already been done by König and Terhal [14], who show that one-bit extractors are quantum-proof.

THEOREM 4.5 (see [14, Theorem III.1]). *Let $C : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}$ be a (k, ε)-strong extractor. Then C is a quantum-proof $(k + \log 1/\varepsilon, 3\sqrt{\varepsilon})$-strong extractor.*[12]

We now need to put Proposition 4.4 and Theorem 4.5 together to prove that Trevisan's extractor is quantum-proof. The cases of uniform and weak random seeds differ somewhat in the details. We therefore give two separate proofs for these two cases in sections 4.2.1 and 4.2.2.

4.2.1. Uniform seed. We show that Trevisan's extractor is a quantum-proof strong extractor with uniform seed with the following parameters.

THEOREM 4.6. *Let $C : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}$ be a (k, ε)-strong extractor with uniform seed and let $S_1, \dots, S_m \subset [d]$ be a weak (t, r)-design. Then the extractor given in Definition 4.2, $\text{Ext}_C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$, is a quantum-proof $(k + rm + \log 1/\varepsilon, 3m\sqrt{\varepsilon})$-strong extractor.*

Proof. In Proposition 4.4, if the seed Y is uniform, then V is independent from W and hence, by the Markov chain property, from G as well, so (4.2) can be rewritten as

$$\|\rho_{C(X,V)VWGE} - \rho_{U_1} \otimes \rho_V \otimes \rho_{WGE}\|_{\text{tr}} > \frac{\varepsilon}{m},$$

which corresponds to the exact criterion of the definition of a quantum-proof extractor.

Let C be a (k, ε)-strong extractor with uniform seed, and assume that a player holds a system E such that

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

Then by Proposition 4.4 and because Y is uniform, we know that there exists a classical system G with $H_0(G) \leq rm$, and a partition of Y in V and W, such that

$$\|\rho_{C(X,V)VWGE} - \rho_{U_1} \otimes \rho_V \otimes \rho_{WGE}\|_{\text{tr}} > 3\sqrt{\varepsilon}. \tag{4.3}$$

Since C is a (k, ε)-strong extractor, we know from Theorem 4.5 that we must have $H_{\min}(X|WGE) < k + \log 1/\varepsilon$ for (4.3) to hold. Hence by Lemma B.3, $H_{\min}(X|E) = H_{\min}(X|WE) \leq H_{\min}(X|WGE) + H_0(G) < k + rm + \log 1/\varepsilon$. ▯

[11] In the classical case, [33, 21] still show that a player who can distinguish $C(X, V)$ from uniform can reconstruct X with high probability. But this is nothing other than proving that C is an extractor.

[12] This result holds whether the seed is uniform or not.

4.2.2. Weak random seed.

We show with the following parameters that Trevisan's extractor is a quantum-proof strong extractor with weak random seed.

THEOREM 4.7. *Let $C : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}$ be a (k, ε)-strong extractor with an s-bit seed—i.e., the seed needs at least s bits of min-entropy—and $S_1, \dots, S_m \subset [d]$ a weak (t, r)-design. Then the extractor given in Definition* 4.2*,* $\mathrm{Ext}_C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$*, is a quantum-proof $(k + rm + \log 1/\varepsilon, 6m\sqrt{\varepsilon})$-strong extractor for any seed with min-entropy $d - (t - s - \log \frac{1}{3\sqrt{\varepsilon}})$.*

The main difference between this proof and that of Theorem 4.6 is that, since the seed Y is not uniform in Proposition 4.4, the substring W of the seed not used by the one-bit extractor C is correlated to the seed V of C and acts as classical side information about the seed. To handle this, we show in Lemma A.3 that with probability $1 - \varepsilon$ over the values of W, V still contains a lot of min-entropy, roughly $s' - d'$, where d' is the length of W and s' is the min-entropy of Y. And hence a player holding WGE can distinguish the output of C from uniform, even though the seed has enough min-entropy.

Proof. Let C be a (k, ε)-strong extractor with s bits of min-entropy in the seed, and assume that a player holds a system E such that

$$\|\rho_{\mathrm{Ext}_C(X,Y)YE} - \rho_{U_m} \otimes \rho_Y \otimes \rho_E\|_{\mathrm{tr}} > 6m\sqrt{\varepsilon}.$$

Then by Proposition 4.4 we have

$$\|\rho_{C(X,V)VWGE} - \rho_{U_1} \otimes \rho_{VWGE}\|_{\mathrm{tr}} > 6\sqrt{\varepsilon}. \tag{4.4}$$

Since this player has classical side information W about the seed V, we need an extra step to handle it. Lemma A.3 tells us that from (4.4) and because, by Theorem 4.5, C is a quantum $(k+\log 1/\varepsilon, 3\sqrt{\varepsilon})$-strong extractor, we must have either, for some w, $H_{\min}(X|GEW = w) < k + \log 1/\varepsilon$, and hence

$$H_{\min}(X|E) = H_{\min}(X|EW = w)$$
$$\leq H_{\min}(X|GEW = w) + H_0(G) < k + rm + \log \frac{1}{\varepsilon},$$

or $H_{\min}(V|W) < s + \log \frac{1}{3\sqrt{\varepsilon}}$, from which we obtain, using Lemma B.1,

$$H_{\min}(Y) \leq H_{\min}(V|W) + H_0(W) < s + \log \frac{1}{3\sqrt{\varepsilon}} + d - t. \qquad \square$$

5. Concrete constructions.

Depending on what goal has been set—e.g., maximize the output, minimize the seed length—different one-bit extractors and weak designs will be needed. In this section we give a few examples of what can be done, by taking various classical extractors and designs, and plugging them into Theorem 4.6 (or Theorem 4.7), to obtain bounds on the seed size and entropy loss in the presence of quantum side information.

The results are usually given using the O-notation. This is always meant with respect to all the free variables; e.g., $O(1)$ is a constant independent of the input length n, the output length m, and the error ε. Likewise, $o(1)$ goes to 0 for both n and m large.

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, Reingold, and Vadhan [21], so we use the same one-bit extractor and weak design as they used.

In section 5.2 we give a scheme which uses a seed of length $d = O(\log n)$ but can extract only part of the entropy. This is also based on [21] in the classical case.

In section 5.3 we combine an extractor and design which are locally computable (from Vadhan [34] and Hartman and Raz [8], respectively) to produce a quantum m-bit extractor such that each bit of the output depends on only $O(\log(m/\varepsilon))$ bits of the input.

And finally in section 5.4 we use a one-bit extractor from Raz [20], which requires only a weakly random seed, resulting in a quantum m-bit extractor, which also works with a weakly random seed.

These constructions are summarized in Table 1.1.

5.1. Near-optimal entropy loss. To achieve a near-optimal entropy loss we need to combine a one-bit extractor with near-optimal entropy loss and a weak $(t, 1)$-design. We use the same extractor and design as Raz, Reingold, and Vadhan [21] to do so, namely Lemma C.1 for the design and Proposition C.5 for the one-bit extractor. Plugging this into Theorem 4.6, we get a quantum extractor with parameters similar to those of [21].

Corollary 5.1. *Let C be the extractor from Proposition* C.5 *with error $\varepsilon' = \frac{\varepsilon^2}{9m^2}$, and let us use the weak design from Lemma* C.1. *Then Trevisan's extractor* $\mathrm{Ext}_C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *is a quantum-proof* $(m + 8\log m + 8\log 1/\varepsilon + O(1), \varepsilon)$*-strong extractor with uniform seed, with $d = O(\log^2(n/\varepsilon)\log m)$.*

For $\varepsilon = \mathrm{poly}(1/n)$ the seed has length $d = O(\log^3 n)$. The entropy loss is $\Delta = 8\log m + 8\log 1/\varepsilon + O(1)$, which means that the input still has that much randomness left in it (conditioned on the output). We can extract a bit more by now applying a second extractor to the input. For this we will use the extractor by Tomamichel et al. [32], which is a quantum (k', ε')-strong extractor[13] with seed length $d' = O(m' + \log n' + \log 1/\varepsilon')$ and entropy loss $\Delta' = 4\log 1/\varepsilon' + O(1)$, where n' and m' are the input and output string lengths. Since we will use it for $m' = 8\log m + 4\log 1/\varepsilon' + O(1)$, we immediately get the following corollary from Lemma A.4.

Corollary 5.2. *By applying the extractors from Corollary* 5.1 *and* [32, Theorem 10] *in succession, we get a new function,* $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$*, which is a quantum-proof* $(m + 4\log 1/\varepsilon + O(1), \varepsilon)$*-strong extractor with uniform seed of length* $d = O(\log^2(n/\varepsilon)\log m)$.

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

The entropy loss is $\Delta = 4\log 1/\varepsilon + O(1)$, which is only a factor 2 times larger than the optimal entropy loss. By Lemma 3.5 this extractor can produce $m = H_{\min}^{\varepsilon'}(X|E) - 4\log 1/\varepsilon - O(1)$ bits of randomness with an error $\varepsilon + 2\varepsilon'$.

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 one-bit extractor. Since t cannot be less than $\Omega(\log n)$ [19], 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, e.g., Lemma C.2, at the cost of extracting less randomness from the source.

For the one-bit extractor we use the same one in the previous section, Proposition C.5. Plugging this into Theorem 4.6, we get a quantum extractor with logarithmic seed length.

[13] The authors of [32] define quantum-proof extractors a little differently than we do (see footnote 6), but it is not hard to see that their result holds with the same parameters, as the differences are absorbed in the O-notation.

COROLLARY 5.3. *If for any constant $0 < \alpha \leq 1$ the source has min-entropy $H_{\min}(X|E) = n^{\alpha}$, and the desired error is $\varepsilon = \mathrm{poly}(1/n)$, then using the extractor C from Proposition* C.5 *with error $\varepsilon' = \frac{\varepsilon^2}{9m^2}$ and the weak design from Lemma* C.2 *with $r = n^{\gamma}$ for any $0 < \gamma < \alpha$, we have that Trevisan's extractor* $\mathrm{Ext}_C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *is a quantum-proof $(n^{\gamma}m + 8\log m + 8\log 1/\varepsilon + O(1), \varepsilon)$-strong extractor with uniform seed, with $d = O\big(\frac{1}{\gamma}\log n\big)$.*

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 3.5 this can be increased to $m = H_{\min}^{\varepsilon'}(X|E)^{1-\frac{\gamma}{\alpha}} - o(1)$ with an error of $\varepsilon + 2\varepsilon'$.

5.3. Locally computable extractor.

5.3. Locally computable extractor. Another interesting feature of extractors is *locality*; that is, the m-bit output depends on only a small subset of the n input bits. This is useful in, e.g., the bounded storage model (see [16, 15, 34] for the case of a classical adversary, and [12] for a general quantum treatment), where we assume that a huge source of random bits, say n, are available, and the adversary's storage is bounded by αn for some constant $\alpha < 1$. 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 reads only $\ell \ll n$ bits. An extractor with such a property is called ℓ-local. We will use a construction of an ℓ-local extractor by Vadhan [34], stated in Lemma C.7.

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 [8], given in Lemma C.3. Plugging this into Theorem 4.6, we get a quantum local extractor.

COROLLARY 5.4. *If for any constant $0 < \alpha \leq 1$ the source has min-entropy $H_{\min}(X|E) = \alpha n$, then using the weak design from Lemma* C.3 *for any constant $r > 1$, and the extractor C from Lemma* C.7 *with error $\epsilon' = \frac{\varepsilon^2}{9m^2}$ and any constant $\gamma < \alpha$, we have that Trevisan's extractor* $\mathrm{Ext}_C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *is a quantum-proof ℓ-local $(\gamma n + rm + 2\log m + 2\log 1/\varepsilon + O(1), \varepsilon)$-strong extractor with uniform seed, 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, with an error of $\varepsilon = \mathrm{poly}(1/n)$. By Lemma 3.5 this can be increased to $m = (H_{\min}^{\varepsilon'}(X|E) - \gamma n)/r - O(\log 1/\varepsilon)$ with an error of $\varepsilon + 2\varepsilon'$.

5.4. Weak random seed. Extractors with weak random seeds typically require the seed to have a min-entropy linear in its length. Theorem 4.7 says that the difference between the length and the min-entropy of the seed needed in Trevisan's extractor is roughly the same as the difference between the length and min-entropy of the seed of the underlying one-bit extractor. So we will describe in detail how to modify the construction from section 5.2 to use a weakly random seed. As that extractor uses a seed of length $O(\log n)$, this new construction allows us to preserve the linear loss in the min-entropy of the seed. Any other version of Trevisan's extractor can be modified in the same way to use a weakly random seed, albeit with weaker parameters.

For this we need a one-bit extractor which uses a weakly random seed. We will use a result by Raz [20] (Lemma C.8), which allows us to construct the extractor from Corollary C.9. Plugging this and the weak design from Lemma C.2 into Theorem 4.7, we get the following extractor with weak random seed.

Corollary 5.5. *Let $\alpha > 0$ be a constant such that the source has min-entropy $H_{\min}(X|E) = n^{\alpha}$ and the desired error is $\varepsilon = \mathrm{poly}(1/n)$. Using the extractor C from Corollary* C.9 *with error $\varepsilon' = \frac{\varepsilon^2}{9m^2}$ and the weak design from Lemma* C.2 *with $r = n^{\gamma}$ for any $0 < \gamma < \alpha$, we have that Trevisan's extractor $\mathrm{Ext}_C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a quantum-proof $(n^{\gamma}m + 8\log m + 8\log 1/\varepsilon + O(1), \varepsilon)$-strong extractor with an s-bit weak random seed, where the seed has length $d = O\big(\frac{1}{\beta^2\gamma}\log n\big)$ and min-entropy $s = \big(1 - \frac{1/2-\beta}{c}\big)d$ for some constant c.*[14]

Choosing β and γ to be constants results in a seed of length $d = O(\log n)$ with a possible entropy loss linear in d. The output length is the same as in section 5.2, $m = n^{\alpha-\gamma} - o(1) = H_{\min}(X|E)^{1-\frac{\gamma}{\alpha}} - o(1)$.

If we are interested in extracting all the min-entropy of the source, we can combine Lemma C.8 with the extractor from section 5.1. This results in a new extractor with seed length $d = O(\log^3 n)$ and seed min-entropy $s = d - O(\sqrt[3]{d})$.

6. Outlook.

There exist many results modifying and improving Trevisan's extractor. We briefly describe a few of them here, and refer the reader to [23] for a more extensive review.

Some of these constructions still follow the "design and one-bit extractor" pattern —hence our work implies that they are immediately quantum-proof with roughly the same parameters—e.g., the work of Raz, Reingold, and Vadham [21] and Lu [15], which were mentioned in section 5 and correspond to modifications of the design and one-bit extractor, respectively. Other results such as [21, 30, 24] replace the binary list-decoding codes with multivariate codes over a field F. Raz, Reingold, and Vadham [21] use this technique to reduce the dependence of the seed on the error from $O(\log^2 1/\varepsilon)$ to $O(\log 1/\varepsilon)$. Ta-Shma, Zuckerman, and Safra [30] and Shaltiel and Umans [24] reduce the size of the seed to $d \leq 2\log n$ in several constructions with different parameters for the min-entropy. In these constructions the connection to one-bit extractors is not clear anymore, and it is therefore not guaranteed that these extractors are quantum-proof.

Raz, Reingold, and Vadham [21] extract a little more randomness than we do in section 5.1. They achieve this by composing (in the sense described in Appendix A.2) the scheme of Corollary 5.1 with an extractor by Srinivasan and Zuckerman [26], which has an optimal entropy loss of $\Delta = 2\log 1/\varepsilon + O(1)$. In the presence of quantum side information this extractor has been proven to have an entropy loss of $\Delta = 4\log 1/\varepsilon + O(1)$ in [32]; hence our slightly weaker result in Corollary 5.2, which can possibly be improved.

Impagliazzo, Shaltiel, and Wigderson [9] and then Ta-Shma, Umans, and Zuckerman [29] modify Trevisan's extractor to work for a subpolynomial entropy source, still using a seed of size $d = O(\log n)$. The latter group [29] achieves a construction which can extract all the min-entropy k of the source with such a seed length for some $k = o(n)$. While it is unclear whether these modifications preserve the "design and one-bit extractor" structure, it is an interesting open problem to analyze them in the context of quantum side information.

Another research direction consists of making these constructions practically implementable. Whether the extractor is used for privacy amplification [4, 3], randomness recycling [10], or for generating true randomness [36], the extractor has to have a running time which makes it useful. This does not seem to be the case of Trevisan's

[14] If we work out the exact constant, we find that $c \approx d/t \approx \frac{8(1+4a)}{\beta\gamma\ln 2}$ for $\varepsilon = n^{-a}$.

construction [25]. An important open problem is thus to find variations which are practical to execute.

It is also of great interest to study quantum-proof two-source extractors, that is, extractors which can be applied to two independent sources, each of which is correlated to independent quantum side information. This has so far been studied only by Kasher and Kempe [11], and we refer to their work for more details and open problems.

Appendix A. More on extractors.

A.1. Weak random seed. In section 3.1 we defined extractors as functions which take a uniformly random seed. This is the most common way of defining them, but not a necessary condition. Instead we can consider extractors which use a seed which is only weakly random, but with bounded min-entropy. We extend Definition 3.1 this way.

DEFINITION A.1 (strong extractor with weak random seed [20]). *A function* $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *is a* (k, ε)-strong extractor with an s-bit seed *if, for all distributions* X *with* $H_{\min}(X) \geq k$ *and any seed* Y *independent from* X *with* $H_{\min}(Y) \geq s$*, we have*

$$\frac{1}{2}\|\rho_{\mathrm{Ext}(X,Y)Y} - \rho_{U_m} \otimes \rho_Y\|_{\mathrm{tr}} \leq \varepsilon,$$

where ρ_{U_m} *is the fully mixed state on a system of dimension* 2^m*.*

If quantum side information about the input is present in a system E, then, as before, we require the seed and the output to be independent from that side-information.

DEFINITION A.2 (quantum-proof strong extractor with weak random seed). *A function* $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *is a* quantum-proof (k, ε)-strong extractor with an s-bit seed *if for all states* ρ_{XE} *classical on* X *with* $H_{\min}(X|E)_\rho \geq k$*, and for any seed* Y *independent from* XE *with* $H_{\min}(Y) \geq s$*, we have*

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

where ρ_{U_m} *is the fully mixed state on a system of dimension* 2^m*.*

Lemma 3.3 says that any extractor will work with roughly the same parameters when classical side information about the input X is present. The same holds in the case of classical side information Z about the seed Y.

LEMMA A.3. *Let* $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *be a quantum-proof* (k, ε)*-strong extractor with an* s*-bit seed. Then for any classical* X*,* Y*, and* Z*, and quantum* E*, such that* XE *and* Y *are independent,* $Y \leftrightarrow Z \leftrightarrow E$ *form a Markov chain,*[15] $H_{\min}(Y|Z) \geq s + \log 1/\varepsilon$*, and for all* $z \in \mathcal{Z}$*,* $H_{\min}(X|EZ = z) \geq k$*, we have*

$$\frac{1}{2}\|\rho_{\mathrm{Ext}(X,Y)YZE} - \rho_U \otimes \rho_{YZE}\|_{\mathrm{tr}} \leq 2\varepsilon.$$

Proof. For any two classical systems Y and Z, we have

$$2^{-H_{\min}(Y|Z)} = \mathop{\mathbb{E}}_{z \leftarrow Z}\left[2^{-H_{\min}(Y|Z=z)}\right],$$

[15] A state ρ_{XYE}, where X and Y ae classical systems, forms a Markov chain $X \leftrightarrow Y \leftrightarrow E$ if it can be expressed as $\rho_{XYE} = \sum_{x,y} P_{XY}(x,y)|x,y\rangle\langle x,y| \otimes \rho_E^y$.

so by Markov's inequality,

$$\Pr_{z\leftarrow Z}\left[H_{\min}(Y|Z=z) \leq H_{\min}(Y|Z) - \log 1/\varepsilon\right] \leq \varepsilon.$$

And since $Y \leftrightarrow Z \leftrightarrow E$ form a Markov chain, we have for all $z \in \mathcal{Z}$

$$\rho_{YE|Z=z} = \rho_{Y|Z=z} \otimes \rho_{E|Z=z}.$$

Hence

$$\begin{aligned}\frac{1}{2}&\|\rho_{\mathrm{Ext}(X,Y)YEZ} - \rho_U \otimes \rho_{YEZ}\|_{\mathrm{tr}} \\ &= \frac{1}{2}\sum_{z\in\mathcal{Z}} P_Z(z)\|\rho_{\mathrm{Ext}(X,Y)YE|Z=z} - \rho_U \otimes \rho_{YE|Z=z}\|_{\mathrm{tr}} \\ &= \frac{1}{2}\sum_{z\in\mathcal{Z}} P_Z(z)\|\rho_{\mathrm{Ext}(X,Y)YE|Z=z} - \rho_U \otimes \rho_{Y|Z=z} \otimes \rho_{E|Z=z}\|_{\mathrm{tr}} \leq 2\varepsilon. \qquad \square\end{aligned}$$

The case of quantum side information correlated to both the input and the seed is beyond the scope of this work.

A.2. Composing extractors. If an extractor does not have optimal entropy loss, a useful approach to extracting more entropy is to apply a second extractor to the original input, 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 [35] and improved by Raz, Reingold, and Vadham [21]. König and Terhal [14] gave the first quantum version for composing m times quantum one-bit extractors. We slightly generalize the result of König and Terhal [14] to the composition of arbitrary quantum extractors.

LEMMA A.4. *Let* $\mathrm{Ext}_1 : \{0,1\}^n \times \{0,1\}^{d_1} \to \{0,1\}^{m_1}$ *and* $\mathrm{Ext}_2 : \{0,1\}^n \times \{0,1\}^{d_2} \to \{0,1\}^{m_2}$ *be quantum-proof* (k, ε_1)*- and* $(k-m_1, \varepsilon_2)$*-strong extractors. Then the composition of the two, namely*

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

is a quantum-proof $(k, \varepsilon_1 + \varepsilon_2)$*-strong extractor.*

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

$$\text{(A.1)} \qquad \frac{1}{2}\|\rho_{\mathrm{Ext}_1(X,Y_1)\,\mathrm{Ext}_2(X,Y_2)Y_1Y_2E} - \rho_{U_1} \otimes \rho_{U_2} \otimes \rho_{Y_1} \otimes \rho_{Y_2} \otimes \rho_E\|_{\mathrm{tr}} \leq \varepsilon_1 + \varepsilon_2.$$

The left-hand side of (A.1) can be upper-bounded by

$$\text{(A.2)} \qquad \begin{aligned}&\frac{1}{2}\|\rho_{\mathrm{Ext}_1(X,Y_1)Y_1E} \otimes \rho_{U_2} \otimes \rho_{Y_2} - \rho_{U_1} \otimes \rho_{Y_1} \otimes \rho_E \otimes \rho_{U_2} \otimes \rho_{Y_2}\|_{\mathrm{tr}} \\ &\qquad + \frac{1}{2}\|\rho_{\mathrm{Ext}_2(X,Y_2)Y_2\,\mathrm{Ext}_1(X,Y_1)Y_1E} - \rho_{U_2} \otimes \rho_{Y_2} \otimes \rho_{\mathrm{Ext}_1(X,Y_1)Y_1E}\|_{\mathrm{tr}}.\end{aligned}$$

By the definition of Ext_1 the first term in (A.2) is upper-bounded by ε_1. For the second term we use Lemma B.3 and get

$$\begin{aligned}H_{\min}(X|\,\mathrm{Ext}_1(X,Y_1)Y_1E) &\geq H_{\min}(X|Y_1E) - H_0(\mathrm{Ext}_1(X,Y_1)) \\ &= H_{\min}(X|E) - H_0(\mathrm{Ext}_1(X,Y_1)) \geq k - m_1.\end{aligned}$$

By the definition of Ext_2 the second term in (A.2) can then be upper-bounded by ε_2. $\square$

Appendix B. Technical lemmas.

B.1. Min-entropy chain rules. We use the following "chain-rule-type" statements about the min-entropy. The proofs for the first two can be found in [22].

LEMMA B.1 (see [22, Lemma 3.1.10]). *For any state ρ_{ABC},*

$$H_{\min}(A|BC) \geq H_{\min}(AC|B) - H_0(C),$$

where $H_0(C) = \log \operatorname{rank} \rho_C$.

LEMMA B.2 (see [22, Lemma 3.1.9]). *For any state ρ_{ABZ} classical on Z,*

$$H_{\min}(AZ|B) \geq H_{\min}(A|B).$$

LEMMA B.3. *For any state ρ_{ABZ} classical on Z,*

$$H_{\min}(A|BZ) \geq H_{\min}(A|B) - H_0(Z),$$

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

Proof. The proof is immediate by combining Lemmas B.1 and B.2. □

B.2. Reduction step. To show that a player who can distinguish the output of Ext_C (defined in Definition 4.2) from uniform can also guess the output of the extractor C, we first show that such a player 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 [37].

LEMMA B.4. *Let ρ_{ZB} be a cq-state, where Z is a random variable on m-bit strings. If $\|\rho_{ZB} - \rho_{U_m} \otimes \rho_B\|_{\mathrm{tr}} > \varepsilon$, then there exists an $i \in [m]$ such that*

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

Using the fact that, for any *binary* random variable X and quantum system Q with $\rho_{XQ} = \sum_{i=0,1} p_i |i\rangle\langle i| \otimes \rho_Q^i$, the equality $\|\rho_{XQ} - \rho_{U_1} \otimes \rho_Q\|_{\mathrm{tr}} = \|p_0 \rho_Q^0 - p_1 \rho_Q^1\|_{\mathrm{tr}}$ holds, (B.1) can be rewritten as $\|\rho_{Z_{i[i-1]}B} - \rho_{U_1} \otimes \rho_{Z_{[i-1]}B}\|_{\mathrm{tr}} > \frac{\varepsilon}{m}$. Lemma B.4 can thus be interpreted as saying that if a player holding B 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 ith bit of Z from uniform with probability greater than $\frac{\varepsilon}{m}$.

Proof. The proof uses a hybrid argument. Let

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

Then

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

By rearranging $\|\sigma_i - \sigma_{i-1}\|_{\mathrm{tr}}$, we get the left-hand side of (B.1). ◻

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. For the reader's convenience we now restate Proposition 4.4 and give its proof.

PROPOSITION B.5 (restatement of Proposition 4.4). *Let X be a classical random variable correlated to some quantum system E; let Y be a (not necessarily uniform) seed, independent from XE; and let*

$$\|\rho_{\mathrm{Ext}_C(X,Y)E} - \rho_{U_m} \otimes \rho_Y \otimes \rho_E\|_{\mathrm{tr}} > \varepsilon, \tag{B.2}$$

where Ext_C is the extractor from Definition 4.2. Then there exists a fixed partition of the seed Y into two substrings V and W, and a classical random variable G, such that G has size $H_0(G) \le rm$, where r is one of the parameters of the weak design (Definition 4.1), $V \leftrightarrow W \leftrightarrow G$ form a Markov chain, and

$$\|\rho_{C(X,V)VWGE} - \rho_{U_1} \otimes \rho_{VWGE}\|_{\mathrm{tr}} > \frac{\varepsilon}{m}. \tag{B.3}$$

Proof. We apply Lemma B.4 to (B.2) and get that there exists an $i \in [m]$ such that

(B.4)

$$\begin{aligned}&\Bigg\| \sum_{\substack{x,y\\ C(x,y_{S_i})=0}} p_x q_y |C(x,y_{S_1})\cdots C(x,y_{S_{i-1}}),y\rangle\langle C(x,y_{S_1})\cdots C(x,y_{S_{i-1}}),y| \otimes \rho^x \\ &- \sum_{\substack{x,y\\ C(x,y_{S_i})=1}} p_x q_y |C(x,y_{S_1})\cdots C(x,y_{S_{i-1}}),y\rangle\langle C(x,y_{S_1})\cdots C(x,y_{S_{i-1}}),y| \otimes \rho^x \Bigg\|_{\mathrm{tr}} \\ &> \frac{\varepsilon}{m},\end{aligned}$$

where $\{p_x\}_{x\in\mathcal{X}}$ and $\{q_y\}_{y\in\mathcal{Y}}$ are the probability distributions of X and Y, respectively.

We split $y \in \{0,1\}^d$ into two strings of $t = |S_i|$ and $d-t$ 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 $g(w,x,j,\cdot) : \{0,1\}^t \to \{0,1\}$. This function depends only 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 $1 \le j \le i-1$, 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 $H_0(G) \le \sum_{j=1}^{i-1} 2^{|S_j\cap S_i|}$. We now rewrite (B.4) as

$$\begin{aligned}&\Bigg\| \sum_{\substack{x,v,w\\ C(x,v)=0}} p_x q_{v,w} |g^{w,x}(v),v,w\rangle\langle g^{w,x}(v),v,w| \otimes \rho^x \\ &\qquad - \sum_{\substack{x,v,w\\ C(x,v)=1}} p_x q_{v,w} |g^{w,x}(v),v,w\rangle\langle g^{w,x}(v),v,w| \otimes \rho^x \Bigg\|_{\mathrm{tr}} > \frac{\varepsilon}{m}.\end{aligned}$$

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

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

By rearranging this a little more, we finally get

$$\|\rho_{C(X,V)VWGE} - \rho_{U_1} \otimes \rho_{VWGE}\|_{\text{tr}} > \frac{\varepsilon}{m},$$

where G is a classical system of size $H_0(G) \leq \sum_{j=1}^{i-1} 2^{|S_j \cap S_i|}$ and $V \leftrightarrow W \leftrightarrow G$ form a Markov chain. 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 $H_0(G) \leq rm$. ∎

Appendix C. Known extractors and designs.

In this section we list the known constructions for weak designs and one-bit extractors, which we plug into Trevisan's extractor in section 5.

C.1. Weak designs.

The following weak design allows nearly all the min-entropy of the source to be extracted, but it requires a rather large seed (typically $O(\log^3 n)$ for an optimal one-bit extractor).

LEMMA C.1 (see [21, Lemma 17][16]). *For every $t, m \in \mathbb{N}$ there exists a weak $(t,1)$-design $S_1, \ldots, S_m \subset [d]$ such that $d = t \left\lceil \frac{t}{\ln 2} \right\rceil \lceil \log 4m \rceil = O(t^2 \log m)$. Moreover, such a design can be found in time* poly(m,d) *and space* poly(m).

If we wish to minimize the length of the seed, we can use the following weak design with $\log r = \Theta(t)$. We then get a seed of length $O(\log n)$ (for an optimal one-bit extractor), but extract only a sublinear amount of min-entropy from the source.

LEMMA C.2 (see [21, Lemma 15]). *For every $t, m \in \mathbb{N}$ and $r > 1$ there exists a weak (t,r)-design $S_1, \ldots, S_m \subset [d]$ such that $d = t \lceil t/\ln r \rceil = O\left(t^2/\log r\right)$. Moreover, such a design can be found in time* poly(m,d) *and space* poly(m).

The following weak design construction is much more efficient than the two previous ones and ideal for a local extractor. It uses a seed of size $O(\log^2 n)$ and can extract a constant fraction of the min-entropy (for an optimal one-bit extractor).

LEMMA C.3 (see [8, 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 $S_1, \ldots, S_m \subset [d]$, where $d = O(t^2)$. Such a design can be found in time* poly$(\log m, t)$ *and space* poly$(\log m + \log t)$.

Remark C.4. For the extractor from Lemma C.7 and an error $\varepsilon = \text{poly}(1/n)$, this design requires $m = \Omega\left((\log n)^{\log\log n}\right)$. If we are interested in a smaller m, say $m = \text{poly}(\log n)$, then we can use the weak design from Lemma C.2 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 $O(\log^2 n)$.

[16] Hartman and Raz [8] give a more efficient construction of this lemma, namely in time poly$(\log m, t)$ and space poly$(\log m + \log t)$, with the extra minor restriction that $m > t^{\log t}$.

C.2. One-bit extractors.

As a one-bit extractor, Raz, Reingold, and Vadhan [21] (as well as Trevisan [33]) used the bits of a list-decodable code. We give the parameters here as Proposition C.5 and refer to Appendix D for details on the construction and proof.

PROPOSITION C.5. *For any* $\varepsilon > 0$ *and* $n \in \mathbb{N}$ *there exists a* (k, ε)*-strong extractor with uniform seed* $\mathrm{Ext}_{n,\varepsilon} : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}$ *with* $t = O(\log(n/\varepsilon))$ *and* $k = 3\log 1/\varepsilon$.

Local extractor. Local extractors are defined as follows.

DEFINITION C.6 (ℓ-local extractor [34]). *An extractor* $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *is* ℓ-locally computable *(or* ℓ-local*) if, for every* $r \in \{0,1\}^d$*, the function* $x \mapsto \mathrm{Ext}(x, r)$ *depends on only* ℓ *bits of its input, where the bit locations are determined by* r.

Lu [15] modified Trevisan's scheme [33, 21] to use a local list-decodable code as a one-bit extractor. Vadhan [34] proposes another construction for local extractors, which is optimal up to constant factors. Both these constructions have similar parameters in the case of one-bit extractors.[17] We state the parameters of Vadhan's construction here and refer the interested reader to [15] for Lu's constructions.

LEMMA C.7 (see [34, 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 with uniform seed* $\mathrm{Ext}_{n,\varepsilon,\gamma} : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}$ *with* $t = O(\log(n/\varepsilon))$*,* $k = \gamma n$*, and* $\ell = O(\log 1/\varepsilon)$.

Weak random seed. Raz [20] shows how to transform any extractor which needs a uniform seed into one which can work with a weakly random seed.

LEMMA C.8 (see [20, Theorem 4]). *For any* (k, ε)*-strong extractor* $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$ *with uniform seed, there exists a* $(k, 2\varepsilon)$*-strong extractor* $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^{t'} \to \{0,1\}^m$ *requiring only a seed with min-entropy* $H_{\min}(Y) \geq \left(\frac{1}{2} + \beta\right) t'$*, where* $t' = 8t/\beta$.

By applying this lemma to the one-bit extractor given in Proposition C.5, we obtain the following one-bit extractor.

COROLLARY C.9. *For any* $\varepsilon > 0$ *and* $n \in \mathbb{N}$ *there exists a* (k, ε)*-strong extractor* $\mathrm{Ext}_{n,\varepsilon} : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}$ *requiring a seed with min-entropy* $\left(\frac{1}{2} + \beta\right) d$*, where* $t = O(\frac{1}{\beta}\log(n/\varepsilon))$ *and* $k = 3\log 1/\varepsilon + 3$.

Appendix D. List-decodable codes are one-bit extractors.

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 D.1 (list-decodable code [27]). *A code* $C : \{0,1\}^n \to \{0,1\}^{\bar{n}}$ *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 [33] nor [21] states it explicitly, but both papers contain an implicit proof that if $C : \{0,1\}^n \to \{0,1\}^{\bar{n}}$ is a (ε, L)-list-decodable code, then

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

is a $(\log L + \log 1/2\varepsilon, 2\varepsilon)$-strong extractor (according to Definition 3.1). We have

[17] If the extractor is used to extract m bits, then Vadhan's scheme reads fewer input bits and uses a shorter seed than does Lu's.

rewritten their proof as Theorem D.3 for completeness.[18]

There exist list-decodable codes with the following parameters.

LEMMA D.2. *For every $n \in \mathbb{N}$ and $\delta > 0$ there is a code $C_{n,\delta} : \{0,1\}^n \to \{0,1\}^{\bar{n}}$, 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.*

For example, Guruswami et al. [7] 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 [15, Corollary 1].

THEOREM D.3. *Let $C : \{0,1\}^n \to \{0,1\}^{\bar{n}}$ be an (ε, L)-list-decodable code. Then the function*

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

is a $(\log L + \log 1/2\varepsilon, 2\varepsilon)$-strong extractor.[19]

To prove this theorem we first show that a player who can distinguish the bit of $C'(X, Y)$ 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 a player who can break the extractor must have low min-entropy about X.

LEMMA D.4. *Let X and Y be two independent random variables with alphabets $\{0,1\}^n$ and $[n]$, respectively. Let Y be uniformly distributed, and let 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

$$P_{X_y}(b) = \sum_{\substack{x \in \{0,1\}^n \\ x_y = b}} P_X(x).$$

The average relative Hamming distance between X and α is

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

[18] A slightly more general proof, stating that *approximate* list-decodable codes are one-bit extractors, can be found in [5, Claim 3.7].

[19] This theorem still holds in the presence of classical side information 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

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

We now wish to lower-bound the probability that the relative 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

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

Combining (D.1) and (D.2), we get

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

We are now ready to prove Theorem D.3.

Proof of Theorem D.3. 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 $\frac{1}{2}\left|C'(X,Y)\circ Y - U_1 \circ Y\right| > 2\varepsilon$, then by Lemma D.4 we know that there exists an $\alpha \in \{0,1\}^{\bar{n}}$ such that

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

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

This means that with probability at least 2ε, X takes values x such that the relative Hamming distance is $d(C(x),\alpha) \leq \frac{1}{2}-\varepsilon$. 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 (2.1), $H_{\min}(X) < \log L + \log 1/2\varepsilon$. $\square$

REFERENCES

[1] A. Ambainis, A. Nayak, A. Ta-Shma, and U. Vazirani, *Dense quantum coding and a lower bound for* 1*-way quantum automata*, in Proceedings of the 31st Symposium on Theory of Computing (STOC '99), ACM, New York, 1999, pp. 376–383.

[2] A. Ben-Aroya and A. Ta-Shma, *Better short-seed quantum-proof extractors*, Theoret. Comput. Sci., 419 (2012), pp. 17–25.

[3] C. H. Bennett, G. Brassard, C. Crépeau, and U. Maurer, *Generalized privacy amplification*, IEEE Trans. Inform. Theory, 41 (1995), pp. 1915–1923.

[4] C. H. Bennett, G. Brassard, and J.-M. Robert, *Privacy amplification by public discussion*, SIAM J. Comput., 17 (1988), pp. 210–229.

[5] A. De and T. Vidick, *Near-optimal extractors against quantum storage*, in Proceedings of the 42nd Symposium on Theory of Computing (STOC '10), ACM, New York, 2010, pp. 161–170.

[6] D. Gavinsky, J. Kempe, I. Kerenidis, R. Raz, and R. de Wolf, *Exponential separations for one-way quantum communication complexity, with applications to cryptography*, in Proceedings of the 39th Symposium on Theory of Computing (STOC '07), ACM, New York, 2007, pp. 516–525.

[7] V. Guruswami, J. Håstad, M. Sudan, and D. Zuckerman, *Combinatorial bounds for list decoding*, IEEE Trans. Inform. Theory, 48 (2002), pp. 1021–1034.

[8] T. Hartman and R. Raz, *On the distribution of the number of roots of polynomials and explicit weak designs*, Random Structures Algorithms, 23 (2003), pp. 235–263.

[9] R. Impagliazzo, R. Shaltiel, and A. Wigderson, *Extractors and pseudo-random generators with optimal seed length*, in Proceedings of the 32nd Symposium on Theory of Computing (STOC '00), ACM, New York, 2000, pp. 1–10.

[10] R. Impagliazzo and D. Zuckerman, *How to recycle random bits* in Proceedings of the 30th Symposium on Foundations of Computer Science (FOCS '89), IEEE Press, Piscataway, NJ, 1989, pp. 248–253.

[11] R. Kasher and J. Kempe, *Two-source extractors secure against quantum adversaries*, in Proceedings of the International Conference on Approximation, Randomization, and Combinatorial Optimization (APPROX-RANDOM '10), Springer, New York, 2010, pp. 656–669.

[12] R. König and R. Renner, *Sampling of min-entropy relative to quantum knowledge*, IEEE Trans. Inform. Theory, 57 (2011), pp. 4760–4787.

[13] R. König, R. Renner, and C. Schaffner, *The operational meaning of min- and max-entropy*, IEEE Trans. Inform. Theory, 55 (2009), pp. 4337–4347.

[14] R. König and B. M. Terhal, *The bounded-storage model in the presence of a quantum adversary*, IEEE Trans. Inform. Theory, 54 (2008), pp. 749–762.

[15] C.-J. Lu, *Encryption against storage-bounded adversaries from on-line strong extractors*, J. Cryptol., 17 (2004), pp. 27–42.

[16] U. M. Maurer, *Conditionally-perfect secrecy and a provably-secure randomized cipher*, Journal of Cryptology, 5 (1992), pp. 53–66.

[17] N. Nisan and A. Wigderson, *Hardness vs. randomness*, J. Comput. System Sci., 49 (1994), pp. 149–167.

[18] N. Nisan and D. Zuckerman, *Randomness is linear in space*, J. Comput. System Sci., 52 (1996), pp. 43–52.

[19] J. Radhakrishnan and A. Ta-Shma, *Bounds for dispersers, extractors, and depth-two superconcentrators*, SIAM J. Discrete Math., 13 (2000), pp. 2–24.

[20] R. Raz, *Extractors with weak random seeds*, in Proceedings of the 37th Symposium on Theory of Computing (STOC '05), ACM, New York, 2005, pp. 11–20.

[21] R. Raz, O. Reingold, and S. Vadhan, *Extracting all the randomness and reducing the error in Trevisan's extractors*, J. Comput. System Sci., 65 (2002), pp. 97–128.

[22] R. Renner, *Security of Quantum Key Distribution*, Ph.D. thesis, Department of Computer Science, Swiss Federal Institute of Technology Zurich, Zurich, 2005.

[23] R. Shaltiel, *Recent developments in explicit constructions of extractors*, Bull. European Assoc. Theoret. Comput. Sci., 77 (2002), pp. 67–95.

[24] R. Shaltiel and C. Umans, *Simple extractors for all min-entropies and a new pseudorandom generator*, J. ACM, 52 (2005), pp. 172–216.

[25] R. Solcà, *Efficient Simulation of Random Quantum States and Operators*, Master's thesis, Department of Physics, Swiss Federal Institute of Technology, Zurich, 2010.

[26] A. Srinivasan and D. Zuckerman, *Computing with very weak random sources*, SIAM J. Comput., 28 (1999), pp. 1433–1459.

[27] M. Sudan, *List decoding: Algorithms and applications*, SIGACT News, 31 (2000), pp. 16–27.

[28] A. Ta-Shma, *Short seed extractors against quantum storage*, in Proceedings of the 41st Symposium on Theory of Computing (STOC '09), ACM, New York, 2009, pp. 401–408.

[29] A. Ta-Shma, C. Umans, and D. Zuckerman, *Loss-less condensers, unbalanced expanders, and extractors*, in Proceedings of the 33rd Symposium on Theory of Computing (STOC '01), ACM, New York, 2001, pp. 143–152.

[30] A. Ta-Shma, D. Zuckerman, and S. Safra, *Extractors from Reed-Muller codes*, J. Comput. System Sci., 72 (2006), pp. 786–812.

[31] M. Tomamichel, R. Colbeck, and R. Renner, *Duality between smooth min- and max-entropies*, IEEE Trans. Inform. Theory, 56 (2010), pp. 4674–4681.

[32] M. Tomamichel, C. Schaffner, A. Smith, and R. Renner, *Leftover hashing against quantum side information*, in Proceedings of 2010 International Symposium on Information Theory (ISIT), IEEE Press, Piscataway, NJ, 2010, pp. 2703–2707.

[33] L. Trevisan, *Extractors and pseudorandom generators*, J. ACM, 48 (2001), pp. 860–879.

[34] S. P. Vadhan, *Constructing locally computable extractors and cryptosystems in the bounded-*

storage model, J. Cryptol., 17 (2004), pp. 43–77.

[35] A. Wigderson and D. Zuckerman, *Expanders that beat the eigenvalue bound: Explicit construction and applications*, Combinatorica, 19 (1999), pp. 125–138.

[36] F. Xu, B. Qi, X. Ma, H. Xu, H. Zheng, and H.-K. Lo, *An ultrafast quantum random number generator based on quantum phase fluctuations*, preprint available at http://arxiv.org/abs/1109.0643 (2011).

[37] A. C.-C. Yao, *Theory and applications of trapdoor functions (extended abstract)*, in Proceedings of the 23rd Symposium on Foundations of Computer Science (FOCS '82), IEEE Press, Piscataway, NJ, 1982, pp. 80–91.

[38] D. Zuckerman, *General weak random sources*, in Proceedings of the 31st Symposium on Foundations of Computer Science (FOCS '90), IEEE Press, 1990, Piscataway, NJ, pp. 534–543.