Better short-seed quantum-proof extractors

Avraham Ben-Aroya* Amnon Ta-Shma†

Abstract

We construct a strong extractor against quantum storage that works for every min-entropy k, has logarithmic seed length, and outputs $\Omega(k)$ bits, provided that the quantum adversary has at most βk qubits of memory, for any $\beta < \frac{1}{2}$. The construction works by first condensing the source (with minimal entropy-loss) and then applying an extractor that works well against quantum adversaries when the source is close to uniform.

We also obtain an improved construction of a strong quantum-proof extractor in the high min-entropy regime. Specifically, we construct an extractor that uses a logarithmic seed length and extracts $\Omega(n)$ bits from any source over $\{0,1\}^n$, provided that the min-entropy of the source conditioned on the quantum adversary's state is at least $(1-\beta)n$, for any $\beta < \frac{1}{2}$.

1 Introduction

In the *privacy amplification* problem Alice and Bob share information that is only partially secret with respect to an eavesdropper Charlie. Their goal is to distill this information to a shorter string that is completely secret. The problem was introduced in [2, 1] for classical eavesdroppers. An interesting variant of the problem, where the eavesdropper is allowed to keep quantum information rather than just classical information, was introduced by König, Maurer and Renner [15]. This situation naturally occurs in analyzing the security of some quantum key-distribution protocols [4] and in bounded-storage cryptography [18, 16].

The shared information between Alice and Bob is modeled as a shared string $x \in \{0,1\}^n$, sampled according a distribution X. The information of the eavesdropper is modeled as a mixed state, $\rho(x)$, which might correlated with x.

The privacy amplification problem can be solved by Alice and Bob, but only by using a (hopefully short) random seed y, which can be public. Thus, Alice and Bob look for a function $E : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$ that acts on their shared input x and the public random string y, and extracts "true randomness" for any "allowed" classical distribution X and side information $\rho(X)$. More formally, E is an *ϵ-strong extractor for a family of inputs* Ω, if for any distribution X and any quantum system ρ such that $(X;\rho) \in \Omega$, the distribution $Y \circ E(X,Y) \circ \rho$ is ϵ-close to $U \circ \rho$, where U denotes the uniform distribution. (See Section 2.2 for precise details.)

Clearly, no randomness can be extracted if, for every x, it is possible to recover x from the side information $\rho(x)$. We say the *conditional min-entropy* of X with respect to $\rho(X)$ is k, if an adversary holding the state $\rho(x)$ cannot guess the string x with probability higher than 2^{-k}. Roughly speaking, if one can extract k

*The Blavatnik School of Computer Science, Tel-Aviv University, Tel-Aviv 69978, Israel. Supported by the Adams Fellowship Program of the Israel Academy of Sciences and Humanities, by the Israel Science Foundation, by the Wolfson Family Charitable Trust, and by a European Research Council (ERC) Starting Grant.

†The Blavatnik School of Computer Science, Tel-Aviv University, Tel-Aviv 69978, Israel. Supported by the FP7 FET-Open project QCS

no. of truly random bits	no. of output bits	classical	quantum-proof
$O(n)$	$m = k - O(1)$	Pair-wise independence, [14]	✓[15]
$O(n - k + \log n)$	$m = n$	Fourier analysis, collision [7]	✓[10]
$\Theta(m)$	$m = k - O(1)$	Almost pair-wise ind., [22, 12]	✓, [25]
$O(\frac{\log^2 n}{\log(k)})$	$k^{1-\zeta}$	Designs, [26]	✓, [6]
$O(\log n)$	$m = \Omega(n)$	[19, 3]	✓, This paper, provided $k > (\frac{1}{2} + \zeta)n$
$\log n + O(1)$	$m = k - O(1)$	Lower bound [19, 20]	✓

Table 1: Explicit quantum-proof (n, k, ϵ) strong extractors. To simplify parameters, the error ϵ is a constant.

almost uniform bits from a source X in spite of the side information $\rho(X)$, then the state $X \circ \rho(X)$ is close to another state with conditional min-entropy at least k.[1] Thus, in a very concrete sense, the ultimate goal is finding extractors for sources with high conditional min-entropy.[2] We say E is a *quantum-proof* (n, k, ϵ) strong extractor if it extracts randomness from every input $(X; \rho)$ with conditional min-entropy at least k.

Not every classical extractor[3] is quantum-proof, as was shown by Gavinsky et al. [11]. On the positive side, several well-known classical extractors are quantum-proof. Table 1 lists some of these constructions. We remark that the best explicit classical extractors [13, 9, 8] achieve significantly better parameters than those known to be quantum-proof.

A simpler adversarial model is the "bounded storage model" where the adversary may store a limited number of qubits. The only advantage of the bounded storage model for extractors is that it simplifies the proofs, and allows us to achieve results which currently we cannot prove in the general model. We say E is an (n, k, b, ϵ) strong extractor *against quantum storage* if it extracts randomness from every pair $(X; \rho)$ for which X has at least k min-entropy and for every x, $\rho(x)$ is a mixed state with at most b qubits.

In this paper we work with a slight generalization of the bounded storage model. We say E is a *quantum-proof* (n, f, k, ϵ) strong extractor for *flat distributions* if it extracts randomness from every input $(X; \rho)$ for which X is a flat distribution (meaning it is uniform over its support) with exactly f min-entropy and the conditional min-entropy is at least k. In Lemma 2.4 we prove the easy observation that any quantum-proof (n, f, k, ϵ) strong extractor for flat distributions is also a $(n, f, f - k, \epsilon)$ strong extractor against quantum storage.

We show a generic reduction from the problem of constructing quantum-proof (n, f, k, ϵ) strong extractors for flat distributions to the problem of constructing quantum-proof $((1 + \alpha)f, f, k, \epsilon)$ strong extractors for flat distributions, and a similar reduction for the bounded storage model. In other words, in our model the quantum adversary may have two types of information about the source: first, it may have some classical knowledge about it, reflected in the fact that the input x is taken from some classical flat distribution X, and second, it holds a quantum state that contains some information about the source. The reduction shows that without loss of generality we may assume the classical input distribution is almost uniform. The reduction uses a purely classical object called a *strong lossless condenser* and extends work done in [24] on extractors to quantum-proof extractors. This reduction holds for any setting of the parameters.

We then augment this with a simple construction that shows how to obtain a quantum-proof $((1 + \alpha)f, f, k = (1-\beta)f, \epsilon)$ strong extractor for flat distributions, provided that $\beta < \frac{1}{2}$. The argument here builds

[1] Such a source is said to have conditional *smooth* min-entropy k.

[2] A simple argument shows an extractor for sources with high conditional min-entropy is also an extractor for sources with high conditional smooth min-entropy.

[3] We refer to extractors that extract randomness when the side information is classical as classical extractors.

on work done in [19] on composition of extractors and extends it to quantum-proof extractors. Together, these two reductions give:

Theorem 1.1. *For any $\beta < \frac{1}{2}$ and $\epsilon \geq 2^{-k^\beta}$, there exists an explicit quantum-proof $(n, k, (1-\beta)k, \epsilon)$ strong extractor for flat sources $E : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$ with seed length $t = O(\log n + \log \epsilon^{-1})$ and output length $m = \Omega(k)$.*

Consequently,

Theorem 1.2. *For any $\beta < \frac{1}{2}$ and $\epsilon \geq 2^{-k^\beta}$, there exists an explicit $(n, k, \beta k, \epsilon)$ strong extractor against quantum storage, $E : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$, with seed length $t = O(\log n + \log \epsilon^{-1})$ and output length $m = \Omega(k)$.*

This gives the first logarithmic seed length extractor against b quantum storage that works for every min-entropy k and extracts a constant fraction of the entropy, and it is applicable whenever $b = \beta k$ for $\beta < \frac{1}{2}$.

We would like to stress that in most practical applications, and in particular in cryptographic applications such as quantum key distribution, it is generally impossible to bound the *size* of the side information. For example, in quantum key distribution where extractors are used for privacy amplification, the conditional min-entropy of the source can be estimated by measuring the noise on the channel, whereas any estimate on the adversary's memory is an unproven assumption. Thus, an extractor proven to work only against quantum storage cannot be used in quantum key distribution protocols. We nevertheless feel that proving a result in the bounded storage model may serve as a first step towards solving the general question.

In fact, the second component in the above construction also works in the general quantum-proof setting. Specifically, this gives an extractor with seed length $t = O(\log n + \log \epsilon^{-1})$ that extracts $\Omega(n)$ bits from any source with conditional min-entropy at least $(1-\beta)n$ for $\beta < \frac{1}{2}$.

Theorem 1.3. *For any $\beta < \frac{1}{2}$ and $\epsilon \geq 2^{-n^\beta}$, there exists an explicit quantum-proof $(n, (1-\beta)n, \epsilon)$ strong extractor $E : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$, with seed length $t = O(\log n + \log \epsilon^{-1})$ and output length $m = \Omega(n)$.*

The rest of the paper is organized as follows. Section 2 contains all the necessary preliminaries, including the formal definitions of min-entropy, quantum-proof extractors and extractors against quantum storage. In Section 3 we give the reduction which shows it is sufficient to construct extractors for sources with nearly full min-entropy, when working in the bounded storage or flat sources settings. In Section 4 we describe the construction of quantum-proof extractors when the conditional min-entropy is more than half, and give the proof of Theorem 1.3. The proofs of Theorems 1.1 and 1.2 are given in Section 5.

2 Preliminaries

Distributions. A distribution D on Λ is a function $D : \Lambda \to [0,1]$ such that $\sum_{a \in \Lambda} D(a) = 1$. We denote by $x \sim D$ sampling x according to the distribution D. Let U_t denote the uniform distribution over $\{0,1\}^t$. We measure the distance between two distributions with the variational distance $|D_1 - D_2|_1 = \frac{1}{2} \sum_{a \in \Lambda} |D_1(a) - D_2(a)|$. The distributions D_1 and D_2 are *ϵ-close* if $|D_1 - D_2|_1 \leq \epsilon$.

The min-entropy of D is denoted by $H_\infty(D)$ and is defined to be

$$H_\infty(D) = \min_{a : D(a) > 0} -\log(D(a)).$$

If $H_\infty(D) \geq k$ then for all a in the support of D it holds that $D(a) \leq 2^{-k}$. A distribution is *flat* if it is uniformly distributed over its support. Every distribution D with $H_\infty(D) \geq k$ can be expressed as a convex combination $\sum \alpha_i D_i$ of flat distributions $\{D_i\}$, each with min-entropy at least k. We sometimes abuse notation and identify a set X with the flat distribution that is uniform over X.

If X is a distribution over Λ_1 and $f : \Lambda_1 \to \Lambda_2$ then $f(X)$ denotes the distribution over Λ_2 obtained by sampling x from X and outputting $f(x)$. If X_1 and X_2 are *correlated* distributions we denote their joint distribution by $X_1 \circ X_2$. If X_1 and X_2 are *independent* distributions we replace $\circ$ by $\times$ and write $X_1 \times X_2$.

Mixed states. A pure state is a vector in some Hilbert space. A general quantum system is in a *mixed state* — a probability distribution over pure states. Let $\{p_i, |\phi_i\rangle\}$ denote the mixed state where the pure state $|\phi_i\rangle$ occurs with probability p_i. The behavior of the mixed state $\{p_i, |\phi_i\rangle\}$ is completely characterized by its *density matrix* $\rho = \sum_i p_i |\phi_i\rangle\langle\phi_i|$, in the sense that two mixed states with the same density matrix have the same behavior under any physical operation. Notice that a density matrix over a Hilbert space $\mathcal{H}$ belongs to $\mathrm{Hom}(\mathcal{H}, \mathcal{H})$, the set of linear transformation from $\mathcal{H}$ to $\mathcal{H}$. Density matrices are positive semi-definite operators and have trace 1.

The *trace distance* between density matrices ρ_1 and ρ_2 is $\|\rho_1 - \rho_2\|_{\mathrm{tr}} = \frac{1}{2}\sum_i |\lambda_i|$, where $\{\lambda_i\}$ are the eigenvalues of $\rho_1 - \rho_2$. The trace distance coincides with the variational distance when ρ_1 and ρ_2 are classical states (ρ is classical if it is diagonal in the standard basis). Similarly to probability distributions, the density matrices ρ_1 and ρ_2 are *ϵ-close* if the trace distance between them is at most ϵ.

A positive operator valued measure (POVM) is the most general formulation of a measurement in quantum computation. A POVM on a Hilbert space $\mathcal{H}$ is a collection $\{F_i\}$ of positive semi-definite operators $F_i : \mathrm{Hom}(\mathcal{H}, \mathcal{H}) \to \mathrm{Hom}(\mathcal{H}, \mathcal{H})$ that sum-up to the identity transformation, i.e., $F_i \succeq 0$ and $\sum F_i = I$. Applying a POVM $F = \{F_i\}$ on a density matrix ρ results in the distribution $F(\rho)$ that outputs i with probability $\mathrm{Tr}(F_i\rho)$.

A Boolean measurement $\{F, I - F\}$ *ϵ-distinguishes* ρ_1 and ρ_2 if $|\mathrm{Tr}(F\rho_1) - \mathrm{Tr}(F\rho_2)| \geq \epsilon$.

We shall need the following facts regarding the trace distance.

Fact 2.1. *If $\|\rho_1 - \rho_2\|_{\mathrm{tr}} = \delta$ then there exists a Boolean measurement that δ-distinguishes ρ_1 and ρ_2.*

Fact 2.2. *If ρ_1 and ρ_2 are ϵ-close then $\mathcal{E}(\rho_1)$ and $\mathcal{E}(\rho_2)$ are ϵ-close, for any physically realizable transformation $\mathcal{E}$.*

2.1 Min-entropy

To define the notion of quantum-proof extractors we first need the notion of quantum encoding of classical states.

Definition 2.1. *Let X be a distribution over some set Λ.*

- *An* encoding *of X is a collection $\rho = \{\rho(x)\}_{x\in\Lambda}$ of density matrices.*
- *An encoding ρ is a b-*storage encoding *if $\rho(x)$ is a mixed state over b qubits, for all $x \in \Lambda$.*
- *An encoding is* classical *if $\rho(x)$ is classical for all x.*

The average encoding is denoted by $\bar{\rho}_X = \mathbb{E}_{x\sim X}[\rho(x)]$.

Next we define the notion of conditional min-entropy. The conditional min-entropy of X given $\rho(X)$ measures the average success probability of predicting x given the encoding $\rho(x)$. Formally,

Definition 2.2. *The* conditional min-entropy of X given an encoding ρ *is*

$$H_\infty(X;\rho) = -\log \sup_F \mathop{\mathbb{E}}_{x\sim X}[\mathrm{Tr}(F_x\rho(x))],$$

where the supremum ranges over all POVMs $F = \{F_x\}_{x\in\Lambda}$.

We remark that there exists another definition of conditional min-entropy in the quantum setting, which is more algebraic in flavor. However, the two definitions are equivalent, as shown in [17].

Proposition 2.1 ([18, Proposition 2])**.** *If ρ is a b-storage encoding of X then $H_\infty(X;\rho) \geq H_\infty(X) - b$.*

We shall need the following standard lemmas regarding min-entropy that can be found, e.g., in [21]. The first lemma says that cutting ℓ bits from a source cannot reduce the min-entropy by more than ℓ.

Lemma 2.1. *Let $X = X_1 \circ X_2$ be a distribution over bit strings and ρ be an encoding such that $H_\infty(X;\rho) \geq k$, and suppose that X_2 is of length ℓ. Let ρ' be the encoding of X_1 defined by $\rho'(x_1) = \mathbb{E}_{x\sim(X|X_1=x_1)}[\rho(x)]$. Then, $H_\infty(X_1;\rho') \geq k - l$.*

Proof: Given any predictor P' which predicts X_1 from ρ', we can construct a predictor P for X (from ρ) as follows: P simply runs P' to obtain a prediction for the prefix x_1, and then appends it with a randomly chosen string from $\{0,1\}^\ell$. Then,

$$\begin{aligned}\Pr_{x_1\circ x_2\sim X}[P(\rho(x_1\circ x_2)) = x_1\circ x_2] &= \Pr_{x_1\circ x_2\sim X}[P'(\rho(x_1\circ x_2)) = x_1]\cdot 2^{-\ell} \\ &= \Pr_{x_1\sim X_1}[P'(\rho'(x_1)) = x_1]\cdot 2^{-\ell}.\end{aligned}$$

Thus, if $H_\infty(X_1;\rho') < k - l$ then there would have been a predictor which predicts X with probability greater than 2^{-k} and this cannot be the case since $H_\infty(X;\rho) \geq k$. ∎

The second lemma says that if a source has high min-entropy, then revealing a short prefix (with high probability) does not change much the min-entropy. The lemma is a generalization of a well known classical lemma.

Lemma 2.2. *Let $X = X_1 \circ X_2$ be a distribution and ρ be an encoding such that $H_\infty(X;\rho) \geq k$, and suppose that X_1 is of length ℓ. For a prefix x_1, let ρ_{x_1} be the encoding of X_2 defined by $\rho_{x_1}(x_2) = \rho(x_1 \circ x_2)$. Call a prefix x_1* bad *if $H_\infty(X_2 \mid X_1 = x_1;\rho_{x_1}) \leq r$ and denote by B the set of bad prefixes. Then,*

$$\Pr[X_1 \in B] \leq 2^\ell \cdot 2^r \cdot 2^{-k}.$$

Proof: Let the prefix $x_1' \in B$ be the one with the largest probability mass. Then, $\Pr[X_1 = x_1'] \geq \Pr[X_1 \in B]\cdot 2^{-\ell}$. For any $z \in B$, let A_z denote the optimal predictor that predicts X_2 from ρ_z, conditioned on $X_1 = z$. By the definition of min-entropy, for any $z \in B$,

$$\mathop{\mathbb{E}}_{x_2\sim(X_2|X_1=z)} \Pr[A_z(\rho_z(x_2)) = x_2] \geq 2^{-r}.$$

In particular this holds for $z = x_1'$.

Now, define a predictor P for X from ρ by

$$P(\rho(x)) = x_1' \circ A_{x_1'}(\rho(x)),$$

that is, P simply "guesses" that the prefix is x'_1 and then applies the optimal predictor $A_{x'_1}$. The average success probability of P is

$$\begin{aligned}
\mathop{\mathbb{E}}_{x\sim X}\left[\Pr[P(\rho(x))=x]\right] &= \mathop{\mathbb{E}}_{x_1\sim X_1}\left[\mathop{\mathbb{E}}_{x_2\sim(X_2|X_1=x_1)}\left[\delta_{x_1,x'_1}\cdot\Pr[A_{x'_1}(\rho_{x'_1}(x_2))=x_2]\right]\right] \\
&= \Pr[X_1=x'_1]\cdot\mathop{\mathbb{E}}_{x_2\sim(X_2|X_1=x'_1)}\left[\Pr[A_{x'_1}(\rho_{x'_1}(x_2))=x_2]\right] \\
&\geq \Pr[X_1\in B]\cdot 2^{-\ell}\cdot 2^{-r}
\end{aligned}$$

On the other hand, since $H_\infty(X;\rho)\geq k$, the average success probability of P is at most 2^{-k}. Altogether, $\Pr[X_1\in B]\leq 2^\ell\cdot 2^r\cdot 2^{-k}$. ∎

2.2 Quantum-proof extractors

We now define the three different classes of extractors against quantum adversaries that we deal with in this paper. We begin with the most general (and natural) definition:

Definition 2.3. *A function $E:\{0,1\}^n\times\{0,1\}^t\to\{0,1\}^m$ is a* quantum-proof (n,k,ϵ) strong extractor *if for every distribution X over $\{0,1\}^n$ and every encoding ρ such that $H_\infty(X;\rho)\geq k$,*

$$\|U_t\circ E(X,U_t)\circ\rho(X)-U_{t+m}\times\bar{\rho}_X\|_{\mathrm{tr}}\leq\epsilon.$$

We use $\circ$ to denote correlated values. Thus, $U_t\circ E(X,U_t)\circ\rho(X)$ denotes the mixed state obtained by sampling $x\sim X, y\sim U_t$ and outputting $|y,E(x,y)\rangle\langle y,E(x,y)|\otimes\rho(x)$. Notice that all 3 registers are correlated. When a register is independent of the others we use $\times$ instead of $\circ$. Thus, $U_{t+m}\times\bar{\rho}_X$ denotes the mixed state obtained by sampling $x\sim X, w\sim U_{t+m}$ and outputting $|w\rangle\langle w|\otimes\rho(x)$.

Next we define quantum-proof extractors for *flat distributions*:

Definition 2.4. *A function $E:\{0,1\}^n\times\{0,1\}^t\to\{0,1\}^m$ is a* quantum-proof (n,f,k,ϵ) strong extractor for flat distributions *if for every* flat *distribution X over $\{0,1\}^n$ with exactly f min-entropy and every encoding ρ of X with $H_\infty(X;\rho)\geq k$,*

$$\|U_t\circ E(X,U_t)\circ\rho(X)-U_{t+m}\times\bar{\rho}_X\|_{\mathrm{tr}}\leq\epsilon.$$

We remark that in the classical setting every extractor for flat distributions is also an extractor for general distributions, since every distribution with min-entropy k can be expressed as a convex combination of flat distributions over 2^k elements.

Finally we define extractors against quantum storage:

Definition 2.5. *A function $E:\{0,1\}^n\times\{0,1\}^t\to\{0,1\}^m$ is an (n,k,b,ϵ)* strong extractor against quantum storage *if for every distribution X over $\{0,1\}^n$ with $H_\infty(X)\geq k$ and every b-storage encoding ρ of X,*

$$\|U_t\circ E(X,U_t)\circ\rho(X)-U_{t+m}\times\bar{\rho}_X\|_{\mathrm{tr}}\leq\epsilon.$$

The next lemma shows it sufficient to consider only flat distributions when arguing about the correctness of extractors against quantum storage.

Lemma 2.3. *If E is* not *an (n, k, b, ϵ) strong extractor against quantum storage then there exists a set X of cardinality 2^k and a b-storage encoding ρ such that E fails on $(X; \rho)$, that is,*

$$\|U_t \circ E(X, U_t) \circ \rho(X) - U_{t+m} \times \bar{\rho}_X\|_{\mathrm{tr}} > \epsilon.$$

Proof: We prove the contrapositive, i.e., we assume that E works for flat distributions of min-entropy exactly k and prove that it also works for general distributions with at least k min-entropy.

Suppose X is a distribution with $H_\infty(X) \geq k$. Then X can expressed as a convex combination of flat distributions X_i each with $H_\infty(X_i) = k$. If ρ is a b-storage encoding of X then it is also a b-storage encoding of each of these flat distributions X_i. Thus, by assumption,

$$\|U_t \circ E(X_i, U_t) \circ \rho(X_i) - U_{t+m} \times \bar{\rho}_{X_i}\|_{\mathrm{tr}} \leq \epsilon.$$

Now by convexity,

$$\|U_t \circ E(X, U_t) \circ \rho(X) - U_{t+m} \times \bar{\rho}_X\|_{\mathrm{tr}} \leq \epsilon,$$

as desired. ■

Combining this with Proposition 2.1 we get:

Lemma 2.4. *Every quantum-proof (n, f, k, ϵ) strong extractor for flat distributions, is an $(n, f, f - k, \epsilon)$ strong extractor against quantum storage.*

2.3 Lossless condensers

Definition 2.6 (strong condenser)**.** *A mapping $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^{n'}$ is an $(n, k_1) \to_\epsilon (n', k_2)$* strong condenser *if for every distribution X with k_1 min-entropy, $U_d \circ C(X, U_d)$ is ϵ-close to a distribution with $d + k_2$ min-entropy.*

One typically wants to maximize k_2 and bring it close to k_1 while minimizing n' (it can be as small as $k_1 + O(\log \epsilon^{-1})$) and d (it can be as small as $\log((n-k)/(n'-k)) + \log \epsilon^{-1} + O(1)$). For a discussion of the parameters, see [3, Appendix B]. We call the condenser *lossless* if $k_2 = k_1$.

The property of lossless condensers that we shall use is the following.

Fact 2.3 ([23, Lemma 2.2.1])**.** *Let $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^{n'}$ be an $(n, k) \to_\epsilon (n', k)$ lossless condenser. Consider the mapping*

$$C' : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^{n'} \times \{0,1\}^d$$

$$C'(x, y) = C(x, y) \circ y.$$

Then, for every set $X \subseteq \{0,1\}^n$ of size $|X| \leq 2^k$, there exists a mapping $C'' : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^{n'} \times \{0,1\}^d$ that is injective on $X \times \{0,1\}^d$ and agrees with C' on at least $1 - \epsilon$ fraction of the set $X \times \{0,1\}^d$.

3 A reduction to full classical entropy

A popular approach for constructing explicit extractors in the classical setting is as follows:

- Construct an explicit extractor for the *high* min-entropy regime, i.e. for sources X distributed over $\{0,1\}^n$ that have k min-entropy for some large k close to n, and,
- Show a reduction from the general case to the high min-entropy case.

In the classical setting this is often achieved by composing an extractor for the high min-entropy regime with a classical lossless condenser. Specifically, assume:

- $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^{n'}$ is an $(n,k) \to_{\epsilon_1} (n',k)$ strong lossless condenser, and,
- $E : \{0,1\}^{d+n'} \times \{0,1\}^t \to \{0,1\}^m$ is a $(d+n', d+k, \epsilon_2)$ strong extractor.

Define $EC : \{0,1\}^n \times (\{0,1\}^d \times \{0,1\}^t) \to \{0,1\}^m$ by

$$EC(x,(y_1,y_2)) \quad = \quad E((C(x,y_1),y_1),y_2).$$

In the classical setting, [24, Section 5] prove that EC is a strong $(n,k,\epsilon_1+\epsilon_2)$ extractor. In this section we try to generalize this result to the quantum setting. We prove:

Theorem 3.1. *Let C and EC be as above.*

- *If E is a quantum-proof $(d+n', d+k, k_2, \epsilon_2)$ strong extractor for flat distributions, then EC is a $(n,k,k_2,\epsilon=\epsilon_2+2\epsilon_1)$ strong extractor for flat distributions.*
- *If E is a $(d+n', d+k, d+b, \epsilon_2)$ strong extractor against quantum storage, then EC is an $(n,k,b,\epsilon=\epsilon_2+2\epsilon_1)$ strong extractor against quantum storage.*

The intuition behind the theorem is the following. When the condenser C is applied on a flat source, it is essentially a one-to-one mapping between the source X and its image $C(X)$. Therefore, roughly speaking, any quantum information about x can be translated to quantum information about $C(x)$ and vice-versa. To make this precise we need to take care of the condenser's seed, and this incurs a small loss in the parameters.

We first prove the second item.

Proof (second item): Assume, by contradiction that EC is not an $(n,k,b,\epsilon=\epsilon_2+2\epsilon_1)$ strong extractor against quantum storage. Then, by Lemma 2.3, there exists a subset $X \subseteq \{0,1\}^n$ of cardinality 2^k and a b-storage encoding ρ of X such that, given this encoding, the output of the extractor EC is not ϵ-close to uniform. That is,

$$\|U_{t+d} \circ EC(X, U_{t+d}) \circ \rho(X) - U_{t+d+m} \times \bar{\rho}_X\|_{\mathrm{tr}} \quad > \quad \epsilon.$$

In particular, by Fact 2.1, there exists some Boolean measurement that ϵ-distinguishes the two distributions. Since the first two components are classical, we can represent this measurement as follows. For every $y \in \{0,1\}^{t+d}$ and $z \in \{0,1\}^m$ there exists a Boolean measurement $\{F^{y,z}, I - F^{y,z}\}$ on the quantum component such that

$$\left| \mathop{\mathbb{E}}_{x\sim X,\ y\sim U} \left[\mathrm{Tr}\left(F^{y,EC(x,y)}\rho(x)\right)\right] - \mathop{\mathbb{E}}_{y,z\sim U} \left[\mathrm{Tr}\left(F^{y,z}\bar{\rho}_X\right)\right] \right| \quad > \quad \epsilon.$$

We now show how this can be used to break the extractor E. Consider the set $A = X \times \{0,1\}^d$. By Fact 2.3, there exists a mapping D that is injective on A and agrees with the condenser on at least $1-\epsilon_1$ fraction of A. Denoting $B = D(A)$, it is clear that $H_\infty(B) \geq d+k$.

For $(\tilde{x},\tilde{y}) \in B$ we define the encoding

$$\rho'(\tilde{x},\tilde{y}) = |y_1\rangle\langle y_1| \otimes \rho(D^{\leftarrow}(\tilde{x},\tilde{y})),$$

where $(x,y_1) = D^{-1}(\tilde{x},\tilde{y}) \in A$ is the unique element such that $D(x,y_1) = (\tilde{x},\tilde{y})$, and $D^{\leftarrow}(\tilde{x},\tilde{y}) = x$.

Next, we define a measurement $\{\overline{F}^{y_2,z}, I - \overline{F}^{y_2,z}\}$ that given the input $y_2 \in \{0,1\}^t, z \in \{0,1\}^m$ and $\rho'(\tilde{x},\tilde{y}) = |y_1\rangle\langle y_1| \otimes \rho(x)$, sets $y = (y_1,y_2)$ and applies the measurement $\{F^{y,z}, I - F^{y,z}\}$ on the quantum register $\rho(x)$.

Now,

$$\left| \mathop{\mathbb{E}}_{b\sim B,\ y_2\sim U_t} \left[\mathrm{Tr}\left(\overline{F}^{y_2,E(b,y_2)}\rho'(b)\right)\right] - \mathop{\mathbb{E}}_{x\sim X,\ y\sim U_{d+t}} \left[\mathrm{Tr}\left(F^{y,EC(x,y)}\rho(x)\right)\right] \right| \quad \leq \quad \epsilon_1,$$

since the flat distribution over B is ϵ_1-close to the distribution obtained by sampling $x \in X$, $y_1 \in U_d$ and outputting $(C(x,y_1),y_1)$. For the same reason, averaging over B for $\overline{F}$ is almost as averaging over X for F. Namely,

$$\left| \mathop{\mathbb{E}}_{y_2,z\sim U} \left[\mathrm{Tr}\left(\overline{F}^{y_2,z}\bar{\rho'}_B\right)\right] - \mathop{\mathbb{E}}_{y,z\sim U} \left[\mathrm{Tr}\left(F^{y,z}\bar{\rho}_X\right)\right] \right| \quad \leq \quad \epsilon_1.$$

It follows that

$$\left| \mathop{\mathbb{E}}_{b\sim B,\ y_2\sim U} \left[\mathrm{Tr}\left(\overline{F}^{y_2,E(b,y_2)}\rho'(b)\right)\right] - \mathop{\mathbb{E}}_{y_2,z\sim U} \left[\mathrm{Tr}\left(\overline{F}^{y_2,z}\bar{\rho'}_B\right)\right] \right| \quad \geq$$

$$\left| \mathop{\mathbb{E}}_{x\sim X,\ y\sim U} \left[\mathrm{Tr}\left(F^{y,EC(x,y)}\rho(x)\right)\right] - \mathop{\mathbb{E}}_{y,z\sim U} \left[\mathrm{Tr}\left(F^{y,z}\bar{\rho}_X\right)\right] \right| - 2\epsilon_1 \quad > \quad \epsilon - 2\epsilon_1 = \epsilon_2.$$

Clearly ρ' is a $(d+b)$-storage encoding of B. This contradicts the fact that E is a strong extractor against $d+b$ quantum storage. ■

We now prove the first item.

Proof (first item): Assume, for contradiction, that EC is not a quantum-proof (n,k,k_2,ϵ) strong extractor for flat distributions. Then there exists a subset $X \subseteq \{0,1\}^n$ of cardinality exactly 2^k and an encoding ρ of X such that the conditional min-entropy is at least k_2 but given this encoding the output of the extractor EC is not ϵ-close to uniform. The proof proceeds as before, defining the Boolean measurement F, the sets A and B, the encoding ρ' and the measurement $\overline{F}$. If we can show that $H_\infty(B;\rho') \geq k_2$ then we break the extractor E and reach a contradiction. Indeed:

Claim 3.1. $H_\infty(B;\rho') \geq k_2$.

Proof: Assume, for contradiction, that $H_\infty(B;\rho') < k_2$. Then, there exists a predictor W' such that

$$\Pr_{b\sim B}[W'(\rho'(b)) = b] > 2^{-k_2}.$$

Define a new predictor, W, that given $\rho(x)$ works as follows. First W chooses $y \sim U_d$ and runs W' on $|y\rangle\langle y| \otimes \rho(x)$ to get some answer $\widetilde{b}$. It then outputs $D^{\leftarrow}(\widetilde{b})$.

The success probability of the predictor W is

$$\begin{aligned}\Pr_{x\sim X}[W(\rho(x))=x] &= \Pr_{x\sim X, y\in\{0,1\}^d}[D^{\leftarrow}(W'(|y\rangle\langle y|\otimes\rho(x)))=x] \\ &\geq \Pr_{x\sim X, y\in\{0,1\}^d}[W'(|y\rangle\langle y|\otimes\rho(x))=D(x,y)] \\ &= \Pr_{b\sim B}[W'(\rho'(b))=b] > 2^{-k_2}.\end{aligned}$$

This contradicts the fact that $H_\infty(X;\rho)\geq k_2$. ■

■

We remark that we do not know how to extend the proof to work with lossy condensers.

4 An explicit quantum-proof extractor for the high-entropy regime

In this section we describe a construction of a short-seed quantum-proof (n,k,ϵ) strong extractor that works whenever $k \gg n/2$. In the classical setting this scenario was studied in [3], developing and improving techniques from [19] and other papers. Here we only need the techniques developed in [19].

Intuitively, the extractor E that we construct works as follows. First, it divides the source to two parts of equal length. Since the min-entropy is larger than $n/2$, for almost any fixing of the first part of the source, the distribution on the second part has $\Omega(n)$ min-entropy. Hence, applying an extractor E_2 on the second part results in output bits that are close to uniform. Since this is true for almost every fixing of the first part, these output bits are essentially independent of the first part of the source. Therefore, these output bits can serve as a seed for another extractor, E_1, that is applied on the first part of the source.

Formally, assume:

- $E_1:\{0,1\}^{n/2}\times\{0,1\}^{d_1}\to\{0,1\}^{m_1}$ is a quantum-proof $(\frac{n}{2},\frac{n}{2}-b,\epsilon_1)$ strong extractor, and,
- $E_2:\{0,1\}^{n/2}\times\{0,1\}^{d_2}\to\{0,1\}^{d_1}$ is a quantum-proof $(\frac{n}{2},k,\epsilon_2)$ strong extractor.

Define $E:\{0,1\}^n\times\{0,1\}^{d_2}\to\{0,1\}^{m_1}$ by

$$E(x,y)=E_1(x_1,E_2(x_2,y)),$$

where $x=x_1\circ x_2$ and $x_1,x_2\in\{0,1\}^{n/2}$.

Theorem 4.1. *Let E_1,E_2 and E be as above with $k=\frac{n}{2}-b-\log\epsilon^{-1}$. Then E is a quantum-proof $(n,n-b,\epsilon+\epsilon_1+\epsilon_2)$ strong extractor.*

Proof: Let $X=X_1\circ X_2$ be a distribution on $\{0,1\}^n=\{0,1\}^{n/2}\times\{0,1\}^{n/2}$ and ρ be an encoding such that $H_\infty(X;\rho)\geq n-b$. For a prefix $x_1\in\{0,1\}^{n/2}$, let ρ_{x_1} be the encoding of X_2 defined by $\rho_{x_1}(x_2)=\rho(x_1\circ x_2)$. A prefix x_1 is said to be *bad* if $H_\infty(X_2\mid X_1=x_1;\rho_{x_1})\leq k$. By Lemma 2.2, the probability x_1 (sampled from X_1) is bad is at most

$$\frac{2^{n/2}\cdot 2^k}{2^{n-b}}=\frac{2^{n/2}\cdot 2^{n/2-b-\log\epsilon^{-1}}}{2^{n-b}}=\epsilon.$$

Whenever x_1 is not bad, $H_\infty(X_2 \mid X_1 = x_1; \rho_{x_1}) > k$, that is, the extractor E_2 is applied on a distribution with k min-entropy. Therefore, by the assumption on E_2, its output is ϵ_2-close to uniform. That is, for every good x_1,

$$\|U_{d_2} \circ x_1 \circ E_2(X_2, U_{d_2}) \circ \rho_{x_1}(X_2) - U_{d_2} \circ x_1 \circ U_{d_1} \circ \rho_{x_1}(X_2)\|_{\mathrm{tr}} \leq \epsilon_2.$$

Hence, the distribution $U_{d_2} \circ X_1 \circ E_2(X_2, U_{d_2}) \circ \rho(X)$ is $(\epsilon + \epsilon_2)$-close to $U_{d_2} \circ X_1 \circ U_{d_1} \circ \rho(X)$. In particular,

$$\begin{aligned} & \|U_{d_2} \circ E(X, U_{d_2}) \circ \rho(X) - U_{d_2+d_1} \circ \bar{\rho}_X\|_{\mathrm{tr}} \\ = \; & \|U_{d_2} \circ E_1(X_1, E_2(X_2, U_{d_2})) \circ \rho(X) - U_{d_2+d_1} \circ \bar{\rho}_X\|_{\mathrm{tr}} \\ \leq \; & \epsilon + \epsilon_2 + \|U_{d_2} \circ E_1(X_1, U_{d_1}) \circ \rho(X) - U_{d_2+d_1} \circ \bar{\rho}_X\|_{\mathrm{tr}}, \end{aligned}$$

where the last inequality follows from Fact 2.2.

Since, $H_\infty(X; \rho) \geq n - b$, by Lemma 2.1, if we define an encoding ρ' of X_1 by $\rho'(x_1) = \mathbb{E}_{x \sim (X|X_1=x_1)}[\rho(x)]$, then $H_\infty(X_1; \rho') \geq n - b - n/2 = n/2 - b$. Therefore, by the assumption on E_1 we get

$$\|E_1(X_1, U_{d_1}) \circ \rho(X) - U_{m_1} \otimes \bar{\rho}_X\|_{\mathrm{tr}} \leq \epsilon_1,$$

and thus

$$\|U_{d_2} \circ E(X, U_{d_2}) \circ \rho(X) - U_{d_2+d_1} \otimes \bar{\rho}_X\|_{\mathrm{tr}} \quad \leq \quad \epsilon + \epsilon_1 + \epsilon_2.$$

■

4.1 Plugging in explicit constructions

We use Trevisan's extractor, which was already shown to be quantum-proof in [6, 5]. Specifically, we use the following two instantiations of this extractor:

Theorem 4.2 ([5])**.** *For every constant $\delta > 0$, there exists $E_1 : \{0,1\}^{\frac{n}{2}} \times \{0,1\}^{O(\log^2(n/\epsilon_1))} \to \{0,1\}^{(1-\delta)(\frac{n}{2}-b)}$ which is a quantum-proof $(\frac{n}{2}, \frac{n}{2} - b, \epsilon_1)$ strong extractor.*

Theorem 4.3 ([5])**.** *For every constants $\gamma_1, \gamma_2 > 0$, there exists $E_2 : \{0,1\}^{\frac{n}{2}} \times \{0,1\}^{O(\log(n/\epsilon_2))} \to \{0,1\}^{k^{1-\gamma_1}}$ which is a quantum-proof $(\frac{n}{2}, k, \epsilon_2)$ strong extractor, for $k > n^{\gamma_2}$.*

Plugging these two constructions into Theorem 4.1 gives Theorem 1.3 which we now restate.

Theorem 1.3. *For any $\beta < \frac{1}{2}, \gamma > 0$ and $\epsilon \geq 2^{-n^{(1-\gamma)/2}}$, there exists an explicit quantum-proof $(n, (1-\beta)n, \epsilon)$ strong extractor $E : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$, with seed length $t = O(\log n + \log \epsilon^{-1})$ and output length $m = \Omega(n)$.*

Proof: We set $\epsilon_1 = \epsilon_2 = \epsilon$, $b = \beta n$, $k = \frac{n}{2} - \beta n - \log \epsilon^{-1}$, $\gamma_2 = \delta = \frac{1}{2}$ and $\gamma_1 < \gamma$. In order to apply Theorem 4.1 we need to verify that the output length of E_2 is not shorter than the seed length of E_1. This is indeed the case since

$$k^{1-\gamma_1} \geq (\frac{n}{2} - \beta n - n^{\frac{1-\gamma}{2}})^{1-\gamma_1} \geq n^{1-\gamma} \geq O(\log^2(\frac{n}{\epsilon})).$$

The output length of E is $\frac{1}{2}(\frac{1}{2} - \beta)n = \Omega(n)$. ■

5 The final extractor for the bounded storage model

We need the classical lossless condenser of [13].

Theorem 5.1 ([13]). *For every $\alpha > 0$ there exists an $(n, k) \to_\epsilon ((1+\alpha)k, k)$ strong lossless condenser C with seed length $O(\log n + \log \epsilon^{-1})$.*

Plugging the condenser C and the extractor E of Theorem 1.3 into Theorem 3.1 gives Theorem 1.2, which we now restate.

Theorem 1.2. *For any $\beta < \frac{1}{2}$ and $\epsilon \geq 2^{-k^\beta}$, there exists an explicit $(n, k, \beta k, \epsilon)$ strong extractor against quantum storage, $E : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$, with seed length $t = O(\log n + \log \epsilon^{-1})$ and output length $m = \Omega(k)$.*

Proof: Let $\zeta > 0$ be a constant to be fixed later. The extractor E from Theorem 1.3, when the source length is set to be $2(1-\beta)(1-\zeta)k$, is a quantum-proof $\big(2(1-\beta)(1-\zeta)k, (1-\beta)k, \epsilon\big)$ strong extractor. In particular, it is a $\big(2(1-\beta)(1-\zeta)k, k, \beta k, \epsilon\big)$ strong extractor against quantum storage. Its output length is $\Omega(k)$. The theorem follows by applying Theorem 3.1, using the condenser of Theorem 5.1 with $\alpha = 2(1-\beta)(1-\zeta)-1$. Since $\beta < \frac{1}{2}$ there is a way to fix ζ such that $\alpha > 0$. ■

Since Theorem 3.1 works in the more general model of flat distributions, and since the extractor from Theorem 1.3 already works in the most general setting, we get Theorem 1.1:

Theorem 1.1. *For any $\beta < \frac{1}{2}$ and $\epsilon \geq 2^{-k^\beta}$, there exists an explicit quantum-proof $(n, k, (1-\beta)k, \epsilon)$ strong extractor for flat distributions, $E : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$, with seed length $t = O(\log n + \log \epsilon^{-1})$ and output length $m = \Omega(k)$.*

Acknowledgements. We thank Roy Kasher for pointing out an error in an earlier version of the paper. We thank Christopher Portmann for helpful comments. We thank the anonymous referees for many helpful suggestions that helped improve the paper.

References

[1] C.H. Bennett, G. Brassard, C. Crepeau, and U. Maurer. Generalized privacy amplification. *IEEE Transactions on Information Theory*, 41(6, Part 2):1915–1923, 1995.

[2] C.H. Bennett, G. Brassard, and J.M. Robert. Privacy amplification by public discussion. *SIAM Journal on Computing*, 17(2):210–229, 1988.

[3] M. Capalbo, O. Reingold, S. Vadhan, and A. Wigderson. Randomness conductors and constant-degree expansion beyond the degree/2 barrier. In *Proc. 34th ACM Symp. on Theory of Computing (STOC)*, pages 659–668, 2002.

[4] M. Christandl, R. Renner, and A. Ekert. A Generic Security Proof for Quantum Key Distribution, 2004. arXiv:quant-ph/0402131.

[5] A. De, C. Portmann, T. Vidick, and R. Renner. Trevisan's extractor in the presence of quantum side information, 2009. arXiv:0912.5514.

[6] A. De and T. Vidick. Near-optimal extractors against quantum storage. In *Proc. 42nd ACM Symp. on Theory of Computing (STOC)*, 2010.

[7] Y. Dodis and A. Smith. Correcting errors without leaking partial information. In *Proc. 37th ACM Symp. on Theory of Computing (STOC)*, pages 654–663, 2005.

[8] Z. Dvir, S. Kopparty, S. Saraf, and M. Sudan. Extensions to the method of multiplicities, with applications to kakeya sets and mergers. In *Proc. 50th IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 181–190. IEEE, 2009.

[9] Z. Dvir and A. Wigderson. Kakeya sets, new mergers and old extractors. In *Proc. 49th IEEE Symp. on Foundations of Computer Science (FOCS)*, pages 625–633, 2008.

[10] S. Fehr and C. Schaffner. Randomness extraction via δ-biased masking in the presence of a quantum attacker. In *Proc. Fifth Theory of Cryptography Conference (TCC)*, pages 465–481, 2008.

[11] D. Gavinsky, J. Kempe, I. Kerenidis, R. Raz, and R. de Wolf. Exponential separations for one-way quantum communication complexity, with applications to cryptography. *SIAM Journal on Computing*, 38(5):1695–1708, 2008.

[12] O. Goldreich and A. Wigderson. Tiny families of functions with random properties: a quality-size trade-off for hashing. *Random Structures & Algorithms*, 11(4):315–343, 1997.

[13] V. Guruswami, C. Umans, and S. Vadhan. Unbalanced expanders and randomness extractors from Parvaresh-Vardy codes. *Journal of the ACM*, 56(4):1–34, 2009.

[14] R. Impagliazzo, L. Levin, and M. Luby. Pseudo-random generation from one-way functions. In *Proc. 21st ACM Symp. on Theory of Computing (STOC)*, pages 12–24, 1989.

[15] R. König, U. Maurer, and R. Renner. On the power of quantum memory. *IEEE Transactions on Information Theory*, 51(7):2391–2401, 2005.

[16] R. König and R. Renner. Sampling of min-entropy relative to quantum knowledge, 2007. arXiv:0712.4291.

[17] R. Konig, R. Renner, and C. Schaffner. The operational meaning of min-and max-entropy. *IEEE Transactions on Information theory*, 55(9):4337–4347, 2009.

[18] R. König and B. Terhal. The bounded-storage model in the presence of a quantum adversary. *IEEE Transactions on Information Theory*, 54(2):749–762, 2008.

[19] N. Nisan and D. Zuckerman. Randomness is linear in space. *Journal of Computer and System Sciences*, 52(1):43–52, 1996.

[20] J. Radhakrishnan and A. Ta-Shma. Bounds for dispersers, extractors, and depth-two superconcentrators. *SIAM Journal on Discrete Mathematics*, 13(1):2–24, 2000.

[21] R. Renner. *Security of Quantum Key Distribution*. PhD thesis, Swiss Federal Institute of Technology (ETH) Zurich, September 2005. available at http://arxiv.org/abs/quant-ph/0512258.

[22] A. Srinivasan and D. Zuckerman. Computing with very weak random sources. *SIAM Journal on Computing*, 28(4):1433–1459, 1999.

[23] A. Ta-Shma, C. Umans, and D. Zuckerman. Loss-less condensers, unbalanced expanders, and extractors. In *Proc. 33th ACM Symp. on Theory of Computing (STOC)*, 2001.

[24] A. Ta-Shma, C. Umans, and D. Zuckerman. Lossless condensers, unbalanced expanders, and extractors. *Combinatorica*, 27(2):213–240, 2007.

[25] M. Tomamichel, C. Schaffner, A. Smith, and R. Renner. Leftover hashing against quantum side information, 2010. arXiv:1002.2436.

[26] L. Trevisan. Extractors and pseudorandom generators. *Journal of the ACM*, 48(4):860–879, 2001.