Better short-seed extractors against quantum knowledge

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 $\beta < \frac{1}{2}$. Previous constructions required poly-logarithmic seed length to output such a fraction of the entropy and, in addition, required super-logarithmic seed length for small values of k. 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 extractor against quantum knowledge, in the high guessing 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 guessing entropy of the source conditioned on the quantum adversary's state is at least $(1-\beta)n$, for any $\beta < \frac{1}{2}$. Previous constructions required poly-logarithmic seed length to output $\Omega(n)$ bits from such sources.

1 Introduction

In the *privacy amplification* problem Alice and Bob share information that is only partially secret towards 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, in which case it can be solved almost optimally using extractors. 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 [14]. This situation naturally occurs in analyzing the security of some quantum key-distribution protocols [4] and in bounded-storage cryptography [16, 15].

In the quantum setting, the eavesdropper is entangled with the classical string Alice and Bob share, but does not have full knowledge about it. There are several ways to formalize this lack of knowledge. Two popular ways are:

Limited storage: The adversary may store a limited number of qubits.

Limited knowledge: The adversary, who holds the state $\rho(x)$, cannot use it to guess (with too-high probability) the string x that Alice and Bob share. This translates into requiring that the *guessing entropy* of the shared string, conditioned on ρ, is not too small. We give the formal definition for this in Section 2.

The limited knowledge model is more general — any adversary with limited storage is also an adversary with limited knowledge.

[*] 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 BSF grant 2004390.

no. of truly random bits	no. of output bits	Against classical storage	Against quantum storage
$O(n)$	$m = n - b - O(1)$	Pair-wise independence, [13]	✓[14]
$O(b + \log n)$	$m = n - b - O(1)$	Fourier analysis, collision [7]	✓[9]
$\Theta(m)$	$m \leq n - b - O(1)$	Almost pair-wise ind., [21, 11]	✓, [23]
$O(\frac{\log^2 n}{\log(n-b)})$	$(n-b)^{1-\zeta}$	Designs, [24]	✓, [6]
$O(\log n)$	$m = \Omega(n - b)$	[17, 12, 8]	✓, This paper, provided $b < \frac{1}{2}n$
$\log n + O(1)$	$m = n - b - O(1)$	Lower bound [18, 19]	✓

Table 1: Milestones in building explicit strong (n, k, b, ϵ) extractors against quantum storage, in the classical and quantum settings. To simplify the parameters, the error ϵ is a constant and $k = n$.

It turns out that the 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.

A function $E : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$ 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 the precise details.) Since we have two ways to limit the correlation between ρ and X, we are interested in two families Ω.

- When Ω contains all sources X over $\{0,1\}^n$ with min-entropy k, and all quantum systems ρ over b qubits, we call E an (n, k, b, ϵ) *strong extractor against quantum storage*.
- When Ω contains all sources X over $\{0,1\}^n$ and all quantum systems ρ such that the guessing entropy of X given ρ is at least k, we call E an (n, k, ϵ) *strong extractor against quantum knowledge*.

Extractors against classical knowledge are defined similarly, and the only difference is that ρ is required to be classical. Not every extractor against classical knowledge is also an extractor against quantum knowledge, as shown by Gavinsky et al. [10]. On the positive side, several well-known classical extractors work well against quantum storage and sometimes even against quantum knowledge. Table 1 lists some of these constructions.

In this work we show a generic reduction from the problem of constructing an (n, k, b, ϵ) strong extractor against quantum storage to the problem of constructing a $((1+\alpha)k, k, b, \epsilon)$ strong extractor against quantum storage. In other words, in general, 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 distribution X, and second, it may hold b quantum bits that contain some information about the source. The reduction shows that, without loss of generality, we may assume the classical input distribution is uniform, and so the adversary only has b qubits of information about the source. The reduction uses a purely classical object called a *strong lossless condenser*. This reduction holds for any setting of the parameters.

We then augment this with a simple construction that shows how to obtain a $((1+\alpha)k, k, b, \epsilon)$ strong extractor against quantum storage, provided that $b = \beta k$ for some $\beta < \frac{1}{2}$. Together, these two reductions give:

Theorem 1.1. *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 $\beta < \frac{1}{2}$.

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

Theorem 1.2. *For any $\beta < \frac{1}{2}$ and $\epsilon \geq 2^{-n^{\beta}}$, there exists an explicit $(n, (1-\beta)n, \epsilon)$ strong extractor against quantum knowledge, $E : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$, 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 guessing entropy and extractors against quantum adversaries. In Section 3 we give the reduction that shows it is sufficient to construct extractors for sources with nearly full min-entropy. In Section 4 we describe the construction of extractors against quantum knowledge when the guessing entropy is more than half, and give the proof of Theorem 1.2. The proof of Theorem 1.1 is given in Section 5. We discuss a slight generalization of our results in Section 6.

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.

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. 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 POVM (Positive Operator Valued Measure) 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 $\{F_i\}$ on a density matrix ρ results in answer 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 Guessing entropy

To define the notion of quantum knowledge 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 a classical mixed-state 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 guessing entropy. The guessing entropy of X given $\rho(X)$ measures the average success probability of predicting x given the encoding $\rho(x)$. Formally,

Definition 2.2. *The* guessing entropy of X given an encoding ρ *is*

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

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

Proposition 2.1 ([16, Proposition 2])**.** *If ρ is a b-storage encoding of X then $H_g(X;\rho) \geq H_\infty(X) - b$.*

We shall need the following standard lemmas regarding the guessing entropy that can be found, e.g., in [20]. The first lemma says that cutting ℓ bits from a source cannot reduce the guessing 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 $H_g(X;\rho) \geq k$, 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, $H_g(X_1;\rho') \geq k - l$.*

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

The second lemma says that if a source has high guessing entropy, then revealing a short prefix (with high probability) does not change much the guessing 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 $H_g(X;\rho) \geq k$, 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 $H_g(X_2 \mid X_1 = x_1; \rho_{x_1}) \leq r$ 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 the guessing 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 $H_g(X;\rho) \geq k$, 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 Extractors against quantum knowledge

We now define the three different classes of extractors against quantum adversaries that we deal with in this paper. We begin with the most restrictive class of extractors against quantum storage:

Definition 2.3. *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,*

$$\|U_t \circ E(X,U_t) \circ \rho(X) - U_{t+m} \circ \bar{\rho}_X\|_{\mathrm{tr}} \leq \epsilon.[1]$$

Intuitively, the above expression measures how close the mixed state $U_t \circ E(X,U_t)$ is to the completely mixed state U_{t+m}, for an adversary that holds b qubits that are correlated with X.

We define extractors again quantum knowledge:

Definition 2.4. *A function $E : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$ is an (n,k,ϵ)* strong extractor against quantum knowledge, *if for every distribution X over $\{0,1\}^n$ and every encoding ρ such that $H_g(X;\rho) \geq k$,*

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

[1] The expression $U_t \circ E(X,U_t) \circ \rho(X)$ denotes the mixed state obtained by sampling $x \sim X, y \sim U_t$ and outputting $|y, E(x,y)\rangle\langle y, E(x,y)| \otimes \rho(x)$. Similarly, $U_{t+m} \circ \bar{\rho}_X$ denotes the mixed state obtained by sampling $x \sim X, w \sim U_{t+m}$ and outputting $|w\rangle\langle w| \otimes \rho(x)$.

2.3 Lossless condensers

Definition 2.5 (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 $\log((n-k)/(n'-k)) + \log \epsilon^{-1} + O(1)$). 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 ([22, 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 $|X| \le 2^k$, 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 The bounded storage model: a reduction to full classical entropy

In this section we show that one can compose an extractor against quantum storage with any classical lossless condenser, and get an extractor that works for arbitrary min-entropy. 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 $(d + n', d + k, d + b, \epsilon_2)$ strong extractor against quantum storage.

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

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

We claim that EC is a strong extractor against quantum storage. The intuition behind this 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. More precisely, we need to take care of the condenser's seed, and this incurs a small loss in the quantum knowledge parameters.

Theorem 3.1. *Let C, E and EC be as above. Then EC is an $(n, k, b, \epsilon = \epsilon_2 + 2\epsilon_1)$ strong extractor against quantum storage.*

Proof: Assume, by contradiction, 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,

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

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 $(r, y_1) \in B$ we define the encoding

$$\rho'(r, y_1) = |y_1\rangle\langle y_1| \otimes \rho(D^{-1}(r, y_1)),$$

where $D^{-1}(r, y_1)$ is the unique element $x \in X$ such $D(x, y_1) = (r, y_1)$.

Next, we define a measurement $\left\{\overline{F}^{y_2,z}, I - \overline{F}^{y_2,z}\right\}$ that given the input $y_2 \in \{0,1\}^t, z \in \{0,1\}^m$ and $\rho'(r, y_1) = |y_1\rangle\langle y_1| \otimes \rho(x)$, 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 remark that the proof also works with lossy condensers (or extractors).

One might expect (or suspect) that if C is a $(n, k) \rightarrow_\epsilon (n', k')$ condenser, and if $H_g(X; \rho) \geq k$ then $(C(X, Y); \rho)$ is ϵ-close to having k' guessing entropy. This, however, is false, as shown by [10]. It is possible that some variant of this statement is true (with some payment in parameters).

4 The bounded knowledge model: an extractor for the high-entropy regime

In this section we describe a construction of a short-seed strong extractor against βn quantum knowledge for sources over n bits, for any $\beta < \frac{1}{2}$. In the classical setting this scenario was studied in [3], developing and improving techniques from [18] and other papers. Here we only need the techniques developed in [18].

Intuitively, the extractor E that we construct works as follows. First, it divides the source to two parts of equal length. Since the guessing entropy is bigger than $n/2$, for almost any fixing of the first part of the source, the distribution on the second part has $\Omega(n)$ guessing 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 from 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 an $(\frac{n}{2}, \frac{n}{2} - b, \epsilon_1)$ extractor against quantum knowledge, and,
- $E_2 : \{0,1\}^{n/2} \times \{0,1\}^{d_2} \to \{0,1\}^{d_1}$ is an $(\frac{n}{2}, k, \epsilon_2)$ strong extractor against quantum knowledge.

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 and let $k = \frac{n}{2} - b - \log \epsilon^{-1}$. Then E is an $(n, n-b, \epsilon+\epsilon_1+\epsilon_2)$ strong extractor against quantum knowledge.*

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 $H_g(X;\rho) \geq n-b$. 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 $H_g(X_2 \mid X_1 = x_1; \rho_{x_1}) \leq k$. 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, $H_g(X_2 \mid X_1 = x_1; \rho_{x_1}) > k$, that is, the extractor E_2 is applied on a distribution with k guessing 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, $H_g(X;\rho) \geq n-b$, 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 $H_g(X_1;\rho') \geq n - b - n/2 = n/2 - b$. 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

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

■

4.1 Plugging in explicit constructions

We use Trevisan's extractor, which was already shown to be secure against quantum knowledge 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 an $(\frac{n}{2}, \frac{n}{2} - b, \epsilon_1)$ strong extractor against quantum knowledge.*

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 an $(\frac{n}{2}, k, \epsilon_2)$ strong extractor against quantum knowledge, for $k > n^{\gamma_2}$.*

Plugging these two constructions into Theorem 4.1 gives Theorem 1.2 which we now restate.

Theorem 1.2. *For any $\beta < \frac{1}{2}, \gamma > 0$ and $\epsilon \geq 2^{-n^{(1-\gamma)/2}}$, there exists an explicit $(n, (1-\beta)n, \epsilon)$ strong extractor against quantum knowledge, $E : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$, 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 [12].

Theorem 5.1 ([12]). *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 1.2 into Theorem 3.1 gives Theorem 1.1, which we now restate.

Theorem 1.1. *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 1.2, when the source length is set to be $2(1-\beta)(1-\zeta)k$, is a $\big(2(1-\beta)(1-\zeta)k, (1-\beta)k, \epsilon\big)$ strong extractor against quantum knowledge. 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 $\alpha > 0$. ■

6 A slight generalization

We remark that Theorem 1.1 also applies to an intermediate model that is stronger than the bounded storage model, yet weaker than the bounded knowledge model.

Definition 6.1. *A function $E : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$ is an (n,k,b,ϵ)* strong extractor against quantum knowledge for flat sources, *if for every* flat *distribution X over $\{0,1\}^n$ with $H_\infty(X) \geq k$ and every encoding ρ of X with $H_g(X;\rho) \geq k-b$,*

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

Observe that every (n,k,b,ϵ) strong extractor against quantum knowledge for flat sources, is also an (n,k,b,ϵ) strong extractor against quantum storage. This follows from the fact that if $H_\infty(X) \geq k$ then X can expressed as a convex combination of flat distributions, each with $H_\infty(X) \geq k$. (If ρ is a b-storage encoding of X then it is also a b-storage encoding of each of these flat distributions.)

Theorem 3.1 still works in the new model and with the same parameters, i.e., if E is an extractor against quantum knowledge for flat sources, then the EC (from Section 3) is an extractor against quantum knowledge for flat sources. The proof of Theorem 3.1 carries over to the new model, except that now the assumption on ρ is $H_g(X;\rho) \geq k-b$ and one needs to prove that $H_g(B;\rho') \geq k-b$ (using the notation from the proof of Theorem 3.1). This is shown in the following claim:

Claim 6.1. *$H_g(B;\rho') \geq k-b$.*

Proof: Assume, for contradiction, that $H_g(B;\rho') < k-b$. Then, there exists a predictor W' such that

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

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 $D^{-1}(\widetilde{b})$.

The success probability of the predictor W is

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

This contradicts the fact that $H_g(X;\rho) \geq k-b$. ■

Since Theorem 3.1 works in this more general model, and since the extractor from Theorem 1.2 already works in the more general setting of the bounded knowledge model, we get:

Theorem 6.1. *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 knowledge 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.

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 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

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

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

[16] 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.

[17] C. Lu, O. Reingold, S. Vadhan, and A. Wigderson. Extractors: Optimal up to constant factors. In *Proc. 35th ACM Symp. on Theory of Computing (STOC)*, pages 602–611, 2003.

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

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

[20] 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.

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

[22] 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.

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

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