On Extracting Private Randomness Over a Public Channel

Yevgeniy Dodis* Roberto Oliveira†

April 14, 2003

Abstract

We introduce the notion of a *super-strong extractor*. Given two independent weak random sources X, Y, such extractor $\text{EXT}(\cdot,\cdot)$ has the property that $\text{EXT}(X,Y)$ is statistically random even if one is given Y. Namely, $\langle Y, \text{EXT}(X,Y)\rangle \approx \langle Y, R\rangle$. Super-strong extractors generalize the notion of strong extractors [16], which assume that Y is truly random, and extractors from two weak random sources [26, 7] which only assure that $\text{EXT}(X,Y) \approx R$. We show that super-extractors have many natural applications to design of cryptographic systems in a setting when different parties have independent weak sources of randomness, but have to communicate over an insecure channel. For example, they allow one party to "help" other party extract private randomness: the "helper" simply sends Y, and the "client" gets private randomness $\text{EXT}(X,Y)$. In particular, it allows two parties to derive a nearly random key after initial agreement on only a weak shared key, without using ideal local randomness.

We show that optimal super-strong extractors exist, which are capable of extracting all the randomness from X, as long as Y has a logarithmic amount of min-entropy. This generalizes a similar result from strong extractors, and improves upon previously known bounds [7] for a weaker problem of randomness extraction from two independent random sources. We also give explicit super-strong extractors which work provided the sum of the min-entropies of X and Y is at least their block length.

Finally, we consider the setting of our problem where the public communication channels are *not authenticated*. Using the results of [13], we show that non-trivial authentication is possible when the min-entropy rate of the shared secret key is at least a half. Combining this with our explicit super-extractor construction, we get the first privacy amplification protocol over an adversarially controlled channel, where player do not have ideal local randomness.

*Department of Computer Science, New York University, 251 Mercer Street, New York, NY 10012, USA. Email: dodis@cs.nyu.edu. Parially supported by the NSF CAREER Award.

†Department of Mathematics, New York University, 251 Mercer Street, New York, NY 10012, USA. Email: oliveira@cims.nyu.edu. The work of this author was funded by a doctoral fellowship from CNPq, Brazil.

1 INTRODUCTION

IMPERFECT RANDOMNESS. Randomization has proved to be extremely useful and fundamental in many areas of computer science. Unfortunately, in many situations one does not have ideal sources of randomness, and one has to base a given application on *imperfect sources of randomness*. Among many imperfect sources considered so far, perhaps the most general and realistic source is the *weak* source [29, 7]. The only thing guaranteed about a *weak* source is that no particular string has a very high probability of occurring. This is characterized by a parameter b (called the *min-entropy* of the source) by saying that no string (of some given length ℓ) occurs with probability more than 2^{-b} (for any distribution of the source). We will call this source (ℓ, b)-weak. Unfortunately, handling such weak sources is often necessary in many applications, as it is typically hard to assume much structure on the source beside the fact that it contains some randomness. Thus, by now a universal goal in basing some application on imperfect sources is to make it work with the weak source.

The most direct way of utilizing weak sources would be to extract nearly perfect randomness from such a source. Unfortunately, it is trivial to see [7] that that no *deterministic* function can extract even one random bit from a weak source, as long as $b < \ell$ (i.e., the source is not random to begin with). This observation leaves two possible options. First, one can try to use weak sources for a given application *without* an intermediate step of extracting randomness from it. Second, one can try designing *probabilistic* extractors, and later justify where and how one can obtain the additional randomness needed for extraction.

USING A SINGLE WEAK SOURCE. A big successful line of research [27, 25, 7, 8, 29, 3] following the first approach showed that a single weak source is sufficient to simulate any probabilistic computation of decision or optimization problems (i.e., problems with a unique "correct" output which are potentially solved more efficiently using randomization; this class is called BPP). Unfortunately, most of the methods in this area are not applicable for applications of randomness, where the randomness is needed by the application *itself*, and not mainly for the purposes of efficiency. One prime example of this is cryptography. For example, secret keys have to be random, and many cryptographic primitives (such as public-key encryption) *must* be probabilistic. Indeed, it is not good enough to produce a set of secret keys most of which are "good", if one does not know which of these keys to actually use, or to say that "at least half of the ciphertext does not reveal the message". Thus, new methods are needed to base cryptographic protocols on weak sources. So far, this question has only been studied in the setting of information-theoretic symmetric-key cryptography. In this scenario, the shared secret key between the sender and the recipient is no longer random, but comes from a weak source. As a very negative result, McInnes and Pinkas [14] proved that one cannot securely encrypt even a single bit, even when using an "almost random" $(\ell, \ell-1)$-weak source. Thus, one cannot base symmetric-key encryption on weak sources. Dodis and Spencer [9] also consider the question of message authentication and show that one cannot (non-interactively) authenticate even one bit using $(\ell, \ell/2)$-weak source (this bound is tight as Maurer and Wolf [13] showed how to authenticate up to $\ell/2$ bits when $b > \ell/2$).

Basing more advanced cryptographic primitives on a single weak random sources also promises to be challenging. For example, it is not clear how to meaningfully model access a single weak source by many parties participating in a given cryptographic protocol. Additionally, moving to the *computational* setting will likely require making very non-standard cryptographic assumptions.

USING SEVERAL WEAK SOURCES. Instead, we will assume that each party will have its own weak source, which is *independent* from all the other weak random sources. In other words, while each individual party cannot assume that his source is truly random, the parties are located "far apart" so that their imperfect sources are independent from each other. For simplicity, we will restrict the number of independent sources to two for the remainder of this paper. One of the questions we will consider if it is possible to construct cryptographic protocols, like secret-key encryption or key exchange, in this new setting. In fact, rather than construct these primitives from scratch, we will try to extract nearly ideal randomness from two weak sources, and then simply

use whatever standard methods exist for the cryptographic task at hand!

This brings us to the question of randomness extraction from several independent random sources. This question originated as early as [18], who showed that one can extract randomness from "many" so called "semi-random" sources (later called *SV-sources*), which are very special cases of the weak sources. Vazirani [26] considered two SV-sources, and showed that the inner product function indeed extracts an almost random bit in this case (he also extended this construction to extract more than one bit). Chor and Goldreich [7] were the first to consider general weak sources of equal block length; let us say that the sources X and Y are (ℓ_1, b_1)-weak and (ℓ_2, b_2)-weak, for concreteness, while here we also assume $\ell_1 = \ell_2 = \ell$. First, they showed that a random function can extract almost $(b_1 + b_2 - \ell)$ nearly random bits in this setting.[1] The also gave an explicit number-theoretic construction that can essentially match this (non-optimal) bound. Moreover, they showed that the simple inner product function is also a good bit-extractor under the same condition that $b_1 + b_2 > \ell$. Recently, Trevisan and Vadhan [24] broke this the "barrier" $b_2 + b_2 > \ell$, but only for the very "imbalanced" case when $b_1 = \varepsilon^2\ell$, $b_2 = (1 - O(\varepsilon))\ell$ (for any $\varepsilon > 0$). To summarize, while non-trivial randomness extraction is possible, the known constructions and parameters seem far from optimal. Unfortunately, improving this situation seems to be extremely challenging. Indeed, it is easy to see that the question of extracting randomness from two independent sources beyond what is currently known is even harder than a notoriously hard problem of explicitly constructing certain bipartite Ramsey graphs (see [28, 17]).

EXTRACTORS. A special case of the above question has received a huge amount of attention recently. It involved the case when one of the two sources, say X, is perfect: $b_1 = \ell_1$. In this case, one invests b_1 bits of true randomness X (called the *seed*) and hopes to extract nearly $b_1 + b_2$ random bits from X and a given (b_2, ℓ_2)-weak source Y. A deterministic function EXT achieving this task has simply been called an *extractor* [16]. A *strong extractor* additionally requires X itself to be part of the extracted randomness. In this case, X is usually excluded from the output of EXT, so that the goal becomes to extract up to b_2 random bits from Y. By now, it is well known that one can indeed achieve this goal provided $b_1 \gg \log \ell_2$. Moreover, many explicit constructions of strong extractors which come very close to this bound are known by now (see [15, 23, 11, 20, 19] and the references therein). Not surprisingly, strong extractors have found many applications (e.g., see [19, 6, 12]).

OUR QUESTION. The general question of extracting randomness from two weak sources [18, 26, 7] concentrated on regular, non-strong extractors. Namely, one could not publish the random seed X. If the extracted randomness is to be used as the secret key of the conventional cryptographic systems, this means that one should sample X and Y from two independent weak sources, and securely "transport" X to Y. Consider, for example, the following application. Alice and Bob stay together and wish to securely communicate when Alice goes away. Then can agree on an auxiliary secret key X sampled from their common weak source. When Alice leaves far away, she gets access to an independent source Y. Assuming the parameters are right, Alice can now extract a nearly random secret key $k = \text{EXT}(X, Y)$. However, Bob only knows X, so Alice has to send Y to Bob. In cryptography, it is conventional to assume that the communication channel between Alice and Bob is public. Thus, Alice has to send Y "in the clear". With regular extractors, even from two weak sources, there is no guarantee that $\text{EXT}(X, Y)$ will look random to the eavesdropper who learns X. On the other hand, conventional strong extractors resolve this problem, but rely on a strong assumption that Alice can sample a truly random Y and send it over the channel. In the world with no "true randomness" and only weak sources, this assumption is not realizable, unless eventually two independent sources are secretly brought together.

The above example motivates our common generalization of previous work. We wish to consider *strong* extractors with *weak* seeds. We will call such extractors *super-strong*. Namely, we want to design a function EXT such that $\text{EXT}(X, Y)$ looks random even for an observer who knows Y, for any X and Y sampled from their corresponding (ℓ_1, b_1) and (ℓ_2, b_2) weak sources.

[1] A trivial strengthening of their technique can push this number to $\min(b_1, b_2)$; we will later *non-trivially* push this to $b_1 + b_2$.

OUR RESULTS. As we demonstrate, such remarkable super-strong extractors exists. In particular, we show that a random function can be used to extract essentially all the randomness from X (i.e., nearly b_1 bits), provided only that $b_2 \geq \log \ell_1$ (and also $b_1 \geq \log \ell_2$). The latter condition says that as long as the public seed Y has barely enough randomness, we can extract almost all the randomness for our target source Y. Clearly, this bound generalizes the standard setting of (strong) extractors, where one needs $\ell_2 = b_2 \geq \log \ell_1$ to extract all the randomness from X. We also remark that our analysis *non-trivially* extends the previous work. In particular, it is not just a "trivial application" of the Chernoff bound, like is the case for the existence proof for regular extractors [21, 22].[2] Proving our bound will involve a careful martingale construction, and then applying the famous Azuma's inequality to bound its deviation from the mean. Also, our bound strengthens what was known for regular (non-strong) extraction from two weak sources [7]. As mentioned, their result gave only $b_1 + b_2 - \ell$ bits. It is easy to improve it to $\min(b_1, b_2)$ bits, but getting $b_1 + b_2$ bits — which follows from our *more general* bound — does not seem possible when using standard Chernoff type bounds used by [7].

Next, we address explicit constructions of super-strong extractors. Unfortunately, the large body of work on strong extractors does not seem to be applicable to super-strong extractors. Intuitively, standard extractors use the seed to perform a random walk on some expander, or to select a hash function from a small family of functions. These arguments seem to fall apart completely once the seed comes from a weak random source. On the other hand, any explicit constructions of super-strong extractors will in particular implies extraction from two independent weak sources, for which any improvement seems very hard, as we mentioned earlier. Thus, the best we can hope for is to extend the best known constructions in this latter setting to yield super-strong extractors. And, indeed, this is exactly what we achieve. First, we show that the inner product function is a super-strong bit-extractor for the case $\ell_1 = \ell_2 = \ell$, provided $b_1 + b_2 > \ell$. This argument involves extending the combinatorial lemma of Lindsey (see Section 4). Second, we show that Vazirani's milti-bit extraction for SV-sources can be applied to weak source as well. This allows to extract $\Omega(\ell)$ bits provided $b_1 + b_2 \gg 3\ell/2$. Finally, we show that the explicit extractor of [7] based on discrete logarithms can also be also be extended to our setting, which gives a way to extract nearly $(b_1 + b_2 - \ell)/2$ random bits. Again, we remark than all these extensions actually involve non-trivial modifications to the existing arguments.

PRIVACY AMPLIFICATION. Finally, we return to applications of super-strong extractors to the setting where different parties have independent weak sources, but all the communication between them is public. The most natural such application is that of key agreement (aka *privacy amplification* [5, 4]) by public discussion: sending Y over the channel allows Alice and Bob to agree on a (nearly) random key $k = \text{EXT}(X, Y)$, provided the communication channel is *authentic*. Therefore, the remaining interesting case to consider is what happens when the channel is not only public, but *adversarially controlled* [13]. In particular, the question is whether we can build any kind of message authentication with a shared key coming from a (ℓ_1, b_1)-weak block source, and without any local randomness. Specifically, assume Alice and Bob share a key $X_1, \ldots, X_i$ (where i will be determined from their need; see below) coming from i samples from the (ℓ_1, b_1)-weak block source. When Alice gets her hands on an independent source Y, she would like to authenticate Y using $X_2 \ldots X_i$. Then, they both can agree on the key $k = \text{EXT}(X_1, Y)$, where EXT is our super-strong extractor. As we mentioned, [9] showed that non-interactive one-time authentication from Alice to Bob is impossible when $b_1 \leq \ell_1/2$. On the other hand, Maurer and Wolf [13] gave a way to non-interactively authenticate up to $\ell_1/2$ bits per each shared X_i, provided $b_1 \gg \ell_1/2$.[3] Thus, sharing $i = 1 + 2\ell_2/\ell_1$ values X_i will allows Alice to authentically transmit Y over the channel, so that both can apply a super-strong extractors to agree on a random $k = \text{EXT}(X_1, Y)$.

Combining this observation with our explicit constructions of super-strong extractor for $\ell_1 = \ell_2 = \ell$, we get that the first efficient privacy amplification *without ideal local randomness*, provided $b_1 \gg \ell/2$ and $b_2 \gg \ell - b_1$.

[2] We are not aware of any written proof for the existence of *strong* extractors, as all the references we found point to [21, 22].

[3] We remark that unlike our setting, Alice and Bob had *ideal* local randomness in the setting of [13], and used it at later stages of their application. Luckily, the authentication step was deterministic, which makes it "coincidentally applicable" to our situation.

2 Preliminaries

2.1 Basic notation

We mostly employ standard notation. The symbol $\log$ is reserved for the base 2 logarithm. For a positive integer t, U_t denotes a random variable that is uniform over $\{0,1\}^t$ and independent of all other random variables under consideration. We also write $[t] \equiv \{1, 2, \dots t\}$. For two random variables A, B taking values in the finite set $\mathcal{A}$, their *statistical distance* is $\|A - B\| \equiv \frac{1}{2}\sum_{a\in\mathcal{A}} |\Pr(A = a) - \Pr(B = a)|$, and the min-entropy of A is $H_\infty(A) \equiv \min_{a\in\mathcal{A}} -\log(\Pr(A = a))$. Finally, if C is another random variable, $C|_{A=a}$ represents the distribution of C conditioned on $A = a \in \mathcal{A}$.

2.2 Extraction vs. Super-Strong Extraction

Min-entropy quantifies the amount of hidden randomness in a source X. The objective of extractors is to purify this randomness with the aid of a (small amount of) truly random bits.

Definition 1 *Let $k \geq 0$, $\varepsilon > 0$. A (k, ε)-extractor* $\text{EXT} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *is a function such that for all n-bit random variables X with min-entropy $H_\infty(X) \geq k$ $\|\text{EXT}(X, U_d) - U_m\| \leq \varepsilon$.* EXT *is a (k, ε)-strong extractor if the function* $\text{EXT}' : (x, y) \mapsto y \circ \text{EXT}(x, y)$ *is an extractor.*

Efficient extraction from weakly random sources using small seed length is a non-trivial problem on which a lot of progress has been made recently (see references in the Introduction). However, even a minimal amount of true randomness can be very difficult to obtain in many situations. Given the impossibility of deterministic extraction [7], it is therefore natural to consider *pairs of weak sources*. We adopt a special notation for those.

Definition 2 *[7] The set $\mathsf{CG}(\ell_1, \ell_2, b_1, b_2)$ of pairs of independent (Chor-Goldreich) weak sources is the set of all pairs of independent random variables (X, Y) where X (respectively Y) is ℓ_1 (resp. ℓ_2) bits long and $H_\infty(X) \geq b_1$ (resp. $H_\infty(Y) \geq b_2$).*

We define extractors from 2 weak sources whose output remains random even if one of the input strings is revealed. This is stronger than the definition used in [7], and this extra strength is essential for cryptographic purposes.

Definition 3 *A (b_1, b_2, ε)-super-strong extractor (*SSE*) is a function* $\text{EXT} : \{0,1\}^{\ell_1} \times \{0,1\}^{\ell_2} \to \{0,1\}^m$ *such that for all pairs $(X, Y) \in \mathsf{CG}(\ell_1, \ell_2, b_1, b_2)$, we have $\|\langle Y, \text{EXT}(X, Y)\rangle - \langle Y, U_m\rangle\| \leq \varepsilon$.*

We state for later convenience the following proposition, which can be deduced from the linear programming argument in [7] (i.e. the fact that general sources of a given min-entropy are convex combinations of flat distributions with the same min-entropy).

Proposition 1 *If b_1 and b_2 are integers, then for* any *function* $\text{EXT} : \{0,1\}^{\ell_1} \times \{0,1\}^{\ell_2} \to \{0,1\}^m$ *the maximum of $\|\langle Y, \text{EXT}(X, Y)\rangle - \langle Y, U_m\rangle\|$ over all $(X, Y) \in \mathsf{CG}(\ell_1, \ell_2, b_1, b_2)$ is achieved by flat random variables, that is, by a pair (X, Y) for which X is uniform over a subset $S_X \subset \{0,1\}^{\ell_1}$ with $|S_X| = 2^{b_1}$, and Y is uniform over $S_Y \subset \{0,1\}^{\ell_2}$, $|S_Y| = 2^{b_2}$.*

3 Existence of super-strong extractors

Form now on m, $\ell_1 \geq b_1 \geq 2$ and $\ell_2 \geq b_2 \geq 2$ are positive integers and $\varepsilon > 0$ is a positive real number. The aim of this section is to prove that super-strong extractors exist for certain choices of parameters.

Theorem 1 *There exists a (b_1, b_2, ε)-*SSE $\text{EXT} : \{0,1\}^{\ell_1} \times \{0,1\}^{\ell_2} \to \{0,1\}^m$ *for any choice of parameters satisfying*

$$\begin{aligned} m &\leq b_1 - 2\log\tfrac{1}{\varepsilon} \\ b_1 &\geq \log_2(\ell_2 - b_2) + 2\log\tfrac{1}{\varepsilon} + O(1) \\ b_2 &\geq \log_2(\ell_1 - b_1) + 2\log\tfrac{1}{\varepsilon} + O(1) \end{aligned} \tag{1}$$

Our proof of Theorem 1 is non-constructive and does not provide an efficiently computable SSE. Nevertheless, Theorem 1 is important because it provides nearly tight conditions for existence of super-strong extractors, as demonstrated by Theorem 2 (proved in the Appendix A).

Theorem 2 *There exists a constant c such that if $b_1 \leq \ell_1 - c$ and $b_2 \leq \ell_2 - c$, then the first and second conditions (1) of Theorem 1 are in fact necessary for the existence of a (b_1, b_2, ε)-*SSE $\textsc{Ext} : \{0,1\}^{\ell_1} \times \{0,1\}^{\ell_2} \to \{0,1\}^m$.

We believe that some condition along the lines of the second one in Theorem 1 is also necessary for the existence of SSE's. However, our proof technique for Theorem 2, which is based on the lower bounds of Ta-Shma and Radhakrishnand [22] for (regular) extractors, does not allow us to conclude this fact (cf. Remark 1 in the Appendix A).

3.1 Reformulating Theorem 1

Theorem 1 follows from the following combinatorial theorem.

Theorem 3 *Let $0 < \delta < 1/2$ be given, and let $L_1 \geq B_1, L_2 \geq B_2$ and M be integers, all of which are at least two. Assume that*

$$\delta \geq \max\left\{ 2\sqrt{\frac{M}{B_1}}, 4\sqrt{\frac{\ln L_1 + 1 - \ln B_1}{B_2}}, 4\sqrt{\frac{\ln L_2 + 1 - \ln B_2}{B_1}} \right\} \tag{2}$$

Then there exists a matrix $H = \{H_{ij}\}_{i=1,j=1}^{L_1,L_2} \in [M]^{L_1 \times L_2}$ such that for all sets of rows $R \subset [L_1]$ of size B_1 and all sets of columns $C \subset [L_2]$ of size B_2

$$\frac{1}{2B_1B_2} \sum_{\alpha \in [M]} \sum_{i \in R} \left| \sum_{j \in C} [H_{ij} = \alpha] - \frac{B_2}{M} \right| \leq \delta \tag{3}$$

Indeed, if we set $2^{b_i} = B_i$, $2^{\ell_i} = L_i$ $(i = 1, 2)$, $2^m = M$, $\varepsilon = \delta$ and let $\textsc{Ext}(x, y) = H_{xy}$ (identifying $[M] \approx \{0,1\}^m$), we note that the conditions of Theorem 3 correspond naturally to those of Theorem 1. The result (3) of Theorem 3 corresponds to $\|\langle Y, \textsc{Ext}(X,Y)\rangle - \langle Y, U_m\rangle\| \leq \varepsilon$ where X is flat on R and Y is flat on C. By Proposition 1, this implies $\|\langle Y, \textsc{Ext}(X,Y)\rangle - \langle Y, U_m\rangle\| \leq \varepsilon$ for *all* $(X, Y) \in \mathsf{CG}(\ell_1, \ell_2, b_1, b_2)$, which is exactly the defining property of SSE's. We prove Theorem 3 below.

3.2 Proof of Theorem 3

Proof: *(of Theorem 3)* The proof is probabilistic: we choose a matrix $H \in [M]^{L_1 \times L_2}$ uniformly at random and prove that under the hypotheses this matrix has a positive probability of satisfying the desired low-discrepancy property (3). For each fixed α the indicator random variables $[H_{ij} = \alpha]$ are i.i.d Bernoulli with common mean $1/M$ and one could try to apply a Chernoff bound to bound the probability of their sum being too big for a fixed choice of column and row sets C, R[4]. This standard approach would finish with a union bound over all α, R, C. However, this does *not* work directly: we are considering *a sum of absolute values of sums* of these random variables. A two-step approach of bounding the "inside" sum first and then the "outside" sum does not seem work either, for it only provides weak bounds. We circumvent this difficulty by employing a stronger *concentration inequality* discussed below.

[4]Of course, for different α, β $[H_{ij} = \alpha]$ and $[H_{ij} = \beta]$ are actually dependent

THE CONCENTRATION INEQUALITY. Let $g : \prod_{i=1}^{k} \Lambda_i \to \mathbb{R}$ be a function defined on a Cartesian product (for simplicity, we assume that each Λ_i finite). We assume that g is *c-Lipschitz*; that is, if $x, y \in \prod_{i=1}^{k} \Lambda_i$ differ at at most one coordinate, $|g(x)-g(y)| \leq c$. Now let $(X_1, X_2, \ldots, X_k)$ be a k-tuple of independent random variables that takes values in $\prod_{i=1}^{k} \Lambda_i$, and define $Z \equiv g(X_1, \ldots, X_k)$. The proposed inequality is

$$\forall t \geq 0 \ \Pr(Z - \mathrm{E}(Z) > t\sqrt{k}) \leq e^{-t^2/2c^2} \tag{4}$$

This is a consequence of Azuma's Inequality (cf. [10, Chapter 2] or [2, Chapter 6]) applied to the Doob's martingale in which the values of $X_1, X_2, \ldots X_k$ are revealed one at a time. We now show how to apply this result in our present context.

CONSIDERING ONE SUBMATRIX. Let H be randomly picked from $H \in [M]^{L_1 \times L_2}$. Fix a choice of R, C as in the statement of the theorem. That is, $R \subset [L_1]$ has size B_1 and $C \subset [L_2]$ has size B_2. Consider

$$Z_{R,C} \equiv \frac{1}{2} \sum_{i \in R, \alpha \in [M]} \Big| \sum_{j \in C} [H_{ij} = \alpha] - \frac{B_2}{M} \Big| \tag{5}$$

We shall use the concentration inequality (4) to bound the probability of $Z_{R,C} > \delta B_1 B_2$. To this end, note that $Z_{R,C}$ is a function of $(H_{ij})_{(i,j) \in R \times C} \in [M]^{R \times C}$. This space is the Cartesian product of $B_1 B_2$ copies of $[M]$, the choices of H_{ij} are all independent, and $Z_{R,C}$ is also 1-Lipschitz on this space. The last assertion is proved as follows: choose $(i, j) \in R \times C$ and change the value of H_{ij} from α to β (say). This changes the value of

$$\frac{1}{2} \Big| \sum_{t \in C} [H_{it} = \alpha] - \frac{B_2}{M} \Big| + \frac{1}{2} \Big| \sum_{t \in C} [H_{it} = \beta] - \frac{B_2}{M} \Big|$$

by at most 1 while leaving all other summands in (5) unchanged; therefore, $Z_{R,C}$ changes by at most 1. This proves that $Z_{R,C}$ satisfies the assumptions of (4), and we can deduce

$$\forall t \geq 0 \ \Pr(Z_{R,C} - \mathrm{E}(Z_{R,C}) > t\sqrt{B_1 B_2}) \leq e^{-t^2/2} \tag{6}$$

We now estimate the expectation of $Z_{R,C}$. Observe that for any fixed $i \in [L_2]$ and $\alpha \in [M]$, $\sum_{j \in C} [H_{ij} = \alpha] - B_2/M$ is distributed like $\mathrm{Bin}(n, p) - np$, where $\mathrm{Bin}(n, p)$ is the Binomial distribution of n i.i.d Bernoulli summands with common mean p (in our case, $n = B_1$ and $p = 1/M$). It follows that

$$\mathrm{E}\Big(\Big| \sum_{j \in C} [H_{ij} = \alpha] - B_2/M \Big| \Big) \leq \sqrt{\mathrm{E}\Big(\Big| \sum_{j \in C} [H_{ij} = \alpha] - B_2/M \Big|^2 \Big)} < \sqrt{B_2/M} \tag{7}$$

$Z_{R,C}$ is half of the sum of MB_1 such terms, so that $\mathrm{E}(|Z_{R,C}|) < 2^{-1} B_1 \sqrt{MB_2} = 2^{-1} B_1 B_2 \sqrt{M/B_2}$ and by assumption, this last quantity is upper bounded by $\delta B_1 B_2/2$. We plug this estimate into inequality (6) and set $t = \delta\sqrt{B_1 B_2}/2$ to obtain

$$\Pr(Z_{R,C} > \delta B_1 B_2) \leq e^{-\delta^2 B_1 B_2/8} \tag{8}$$

THE UNION BOUND. Our next step is to take an union bound over all valid choices of rows R and columns C.

$$\begin{aligned} \Pr(\exists R, C \ Z_{R,C} > \delta B_1 B_2) &\leq \binom{L_1}{B_1} \binom{L_2}{B_2} e^{-\delta^2 B_1 B_2/8} \\ &< \exp\left\{ \left(\ln L_1 + 1 - \ln B_1 - \frac{\delta^2 B_2}{16} \right) B_1 + \left(\ln L_2 + 1 - \ln B_2 - \frac{\delta^2 B_1}{16} \right) B_2 \right\} \leq 1 \end{aligned} \tag{9}$$

by the assumption that $\ln L_1 + 1 - \ln B_1 \leq \delta^2 B_2/16$ and $\ln L_2 + 1 - \ln B_2 \leq \delta^2 B_1/16$. It follows that there exists a matrix H for which $Z_{R,C} \leq \delta B_1 B_2$ for all R, C. From the definition of $Z_{R,C}$ (5), we have shown the existence of a matrix H satisfying (3). This concludes the proof. □

4 Efficient constructions

This section is devoted to constructions of efficiently computable SSE. Most standard techniques applied to the construction of regular and strong extractors (e.g. using list-decodable codes, walks on expanders, mergers or even hash functions) seem to fail when the seed is not completely random. As a result, even our 1-bit SSE is only useful when the sources have the same length $\ell_1 = \ell_2 = \ell$ and their min entropies add up to more than ℓ. However, the present constructions are both useful and elegant. We propose the construction of more efficient SSE's as a challenging open problem.

4.1 Hadamard matrices and extraction of one bit

A class of 1-bit super-strong extractors which includes the inner product function is now considered, thus providing a strengthening of a result of Chor and Goldreich [7]. Identify $[L] \equiv [2^\ell] \approx \{0,1\}^\ell$ and let $H = \{H_{xy}\}_{x,y=1}^{L}$ be a $L \times L$ Hadamard matrix (i.e. a ± 1 matrix with pairwise orthogonal rows and columns). Define

$$\begin{array}{rccc} \textsc{Ext}_H: & \{0,1\}^\ell \times \{0,1\}^\ell & \to & \{0,1\} \\ & (x,y) & \longmapsto & \frac{1+H_{xy}}{2} \end{array} \tag{10}$$

We shall prove the following two results.

Theorem 4 *$\textsc{Ext}_H$ as defined above is a (b_1, b_2, ε)-SSE with $\log \frac{1}{\varepsilon} = \frac{b_1+b_2-\ell}{2} + 1$.*

Corollary 2 *The inner product function on ℓ-bit strings is a (b_1, b_2, ε)-SSE with ε as above.*

Proof: (*of Corollary 2*) Inner product is of the form $\textsc{Ext}_H$ for some Hadamard matrix H (as one can easily show). □

Proof: (*of Theorem 4*) The proof parallels that of the corresponding theorem in [7]. In particular, we also employ Lindsay's Lemma.

Lemma 3 (Lindsay's Lemma cf. [7]) *Let $G = (G_{ij})_{i,j=1}^{T}$ S be a $T \times T$ Hadamard matrix, and R and C be subsets of $[T]$ corresponding to choices of rows and columns of G (respectively). Then $|\sum_{i\in R}\sum_{j\in C} G_{ij}| \le \sqrt{|R|\,|C|\,T}$.*

For any choice of $(q_1, \ldots, q_L) \in \{-1,+1\}^L$, the matrix $\tilde{H} = (\tilde{H}_{ij})$ whose ith row is q_i times the ith row of H is Hadamard. Hence Lindsay's Lemma applies and for all sets $R, C \subset [L]$ the sum $\sum_{i\in R}\sum_{j\in C} \tilde{H}_{ij}$, which is just $\sum_{i\in R}\left(q_i \sum_{j\in C} H_{ij}\right)$, is bounded by $\le \sqrt{|R|\,|C|\,L}$. From this fact it is easy to deduce a stronger form of Lemma 3.

$$\forall R, C \subset [N] \quad \sum_{i\in R}\left|\sum_{j\in C} H_{ij}\right| \le \sqrt{|R|\,|C|\,L} \tag{11}$$

Now let $(X,Y) \in \mathsf{CG}(\ell,\ell,b_1,b_2)$ be flat random variables and assume that X is uniform on S_X, $|S_X| = 2^{b_1}$ and Y is uniform on S_Y, $|S_Y| = 2^{b_2}$. For each $y \in S_Y$

$$\begin{aligned} \|\textsc{Ext}_H(X,y) - U_1\| &= \frac{1}{2}\left(|\Pr(\textsc{Ext}_H(X,y)=1) - \frac{1}{2}| + |\Pr(\textsc{Ext}_H(X,y)=0) - \frac{1}{2}|\right) \\ &= \frac{1}{2}|\Pr(\textsc{Ext}_H(X,y)=1) - \Pr(\textsc{Ext}_H(X,y)=0)| = \frac{1}{2}\left|\sum_{x\in S_X}\frac{H_{xy}}{|S_X|}\right| \end{aligned} \tag{12}$$

Averaging over all y and using (11) with $R = S_X$, $C = S_Y$, we obtain

$$\|\langle Y, \textsc{Ext}_H(X,Y)\rangle \;-\; \langle Y, U_1\rangle\| = \frac{1}{|S_Y|}\sum_{y\in S_Y}\|\textsc{Ext}_H(X,y) \;-\; U_1\|$$

$$= \frac{1}{2|S_Y|}\sum_{y\in S_Y}\left|\sum_{x\in S_X}\frac{H_{xy}}{|S_X|}\right| \le \frac{1}{2}\sqrt{\frac{L}{|S_X||S_Y|}} = 2^{-\frac{b_1+b_2-\ell}{2}-1} \quad (13)$$

By Proposition 1 this finishes the proof. □

4.2 Extracting many bits

We now adapt a construction from [26] based on error-correcting codes to obtain many bits from weak sources of same length and sufficiently high min-entropy. In what follows $\textsc{Ecc} : \{0,1\}^m \to \{0,1\}^\ell$ is a linear error correcting code with distance d, $\{\vec{e_i} : 1 \le i \le m\}$ is the canonical basis of $\{0,1\}^m$ as a vector space over $\mathbb{Z}_2$, and for $(x,y) = \big((x_1,\ldots,x_\ell),(y_1,\ldots,y_\ell)\big) \in \{0,1\}^\ell \times \{0,1\}^\ell$ we let $\vec{v}(x,y) \in \{0,1\}^\ell$ be the vector whose ith coordinate is $x_i y_i$. The proposed SSE is

$$\begin{array}{rccl} \textsc{Ext}: & \{0,1\}^\ell \times \{0,1\}^\ell & \to & \{0,1\}^m \\ & (x,y) & \longmapsto & \big(\textsc{Ecc}(\vec{e_1})\cdot\vec{v}(x,y)\big)\circ\cdots\circ\big(\textsc{Ecc}(\vec{e_m})\cdot\vec{v}(x,y)\big) \end{array} \quad (14)$$

Note that each bit that $\textsc{Ext}$ outputs corresponds to the inner product of matching segments of the input strings x and y. We show that

Theorem 5 *The function* $\textsc{Ext}$ *constructed above is a* (b_1,b_2,ε)-SSE *with* $\log\frac{1}{\varepsilon} = 1 + \frac{b_1+b_2+d}{2} - (\ell+m)$.

There exist efficiently encodable linear codes of codeword length ℓ, dimension $m = \delta^3\ell$ and distance $d = (\frac{1}{2}-\delta)\ell$, for all fixed $0 < \delta < \frac{1}{2}$. Plugging one such code into Theorem 5 yields an efficiently computable (b_1,b_2,ε)-SSE with $\varepsilon = \ell^{-\omega(1)}$ for all min-entropies satisfying $\frac{b_1+b_2}{2} \ge (3/4+\delta)\ell + \omega(\log\ell)$, and the number of extracted bits is $m = \delta^3 n$. It remains to prove Theorem 5, and for that we use two lemmas (the second being fairly standard).

Lemma 4 *(Parity Lemma, [26]) For any t-bit random variable T, $\|T \;-\; U_t\| \le \sum_{\vec{v}\in\{0,1\}^t\setminus\{\vec{0}\}}\big\|\big(T\cdot\vec{v}\big) \;-\; U_1\big\|$.*

Lemma 5 *If $Z = Z_1 Z_2 \ldots Z_t$ is a t-bit random variable and $W \subset [t]$, let $Z|_W$ denote the concatenation of all Z_i with $i \in W$. Then $H_\infty(Z|_W) \ge H_\infty(Z) - t + |W|$.*

Proof: *(of Theorem 5)* We need to show that for any pair $(X,Y) \in \mathsf{CG}(\ell,\ell,b_1,b_2)$

$$\|\langle Y, \textsc{Ext}(X,Y)\rangle \;-\; \langle Y, U_m\rangle\| = \sum_{y\in\{0,1\}^\ell}\Pr(Y=y)\,\|\textsc{Ext}(X,y) \;-\; U_m\| \quad (15)$$

is bounded by $\varepsilon = 2^{\ell+m-\frac{b_1+b_2+d}{2}-1}$. To this end, we apply Lemma 4 to each of the above summands.

$$\begin{aligned}\|\langle Y, \textsc{Ext}(X,Y)\rangle \;-\; \langle Y, U_m\rangle\| &\le \sum_{y\in\{0,1\}^\ell}\Pr(Y=y)\sum_{\vec{a}\in\{0,1\}^m\setminus\{\vec{0}\}}\big\|\big(\textsc{Ext}(X,y)\cdot\vec{a}\big) \;-\; U_1\big\| \\ &= \sum_{\vec{a}\in\{0,1\}^m\setminus\{\vec{0}\}}\big\|\langle Y, \big(\textsc{Ext}(X,Y)\cdot\vec{a}\big)\rangle \;-\; \langle Y, U_1\rangle\big\| \quad (16)\end{aligned}$$

It now suffices to show that for any $\vec{a} \in \{0,1\}^m\setminus\{\vec{0}\}$

$$\big\|\langle Y, \big(\textsc{Ext}(X,Y)\cdot\vec{a}\big)\rangle \;-\; \langle Y, U_1\rangle\big\| \le \frac{\varepsilon}{2^m} = 2^{\ell-\frac{b_1+b_2+d}{2}-1} \quad (17)$$

Fix some non-zero $\vec{a} = \sum_{i=1}^m a_i\vec{e_i}$ and note that by the linearity of $\textsc{Ecc}$

$$\textsc{Ext}(X,Y)\cdot\vec{a}=\sum_{i=1}^{m}a_i\big(\textsc{Ecc}(\vec{e_i})\cdot\vec{v}(X,Y)\big)=\textsc{Ecc}(\vec{a})\cdot\vec{v}(X,Y)=\sum_{i\in S}X_iY_i=\big(X|_S\big)\cdot\big(Y|_S\big) \tag{18}$$

where S is the set of all non-zero coordinates of $\textsc{Ecc}(\vec{a})$, and $X|_S$ and $Y|_S$ are defined as in the statement of Lemma 5. Applying that same lemma, we conclude that $X|_S$ ($Y|_S$) has min-entropy at least $b_1-\ell+|S|$ (respectively $b_2-\ell+|S|$). It now follows from (18), Corollary 2 and the fact that $X|_S$ and $Y|_S$ have length $|S|$ that

$$\Big\|\langle Y,\big(\textsc{Ext}(X,Y)\cdot\vec{a}\big)\rangle\ -\ \langle Y,U_1\rangle\Big\|\leq 2^{\frac{|S|-b_1-b_2+2\ell-2|S|}{2}-1}=2^{\ell-1-\frac{b_1+b_2+|S|}{2}} \tag{19}$$

Since $|S|=(\text{weight of }\textsc{Ecc}(\vec{a}))\geq d$ by definition of S, equation (19) proves (17) and finishes the proof. □

4.3 A number-theoretic construction

A third efficient SSE construction is now presented. Its minimal min-entropy requirement is basically $b_1+b_2>\ell$, which roughly matches the Hadamard matrix construction for 1-bit extraction. However, this SSE has the drawback of requiring a pre-processing stage for efficiency to be achieved. The construction dates back to [7], in which it was shown that $\textsc{Ext}(X,Y)$ is close to random. We claim that the same is true even if Y is given to the adversary, thus establishing that this construction satisfies our definition of SSE. In what follows, $p>2$ is a prime and we take $\ell=\lfloor\log p\rfloor$ so that we can assume $\{0,1\}^\ell\subseteq\mathbb{Z}_p$. Let k be a divisor of $p-1$; our SSE will output elements of $\mathbb{Z}_k$ (the definition of a SSE easily generalizes to this case). Finally, let g be a generator of the multiplicative group $\mathbb{Z}_p^*$ and denote by $\log_g$ the base-g discrete logarithm in $\mathbb{Z}_p^*$. We define

$$\begin{array}{rccl}\textsc{Ext}: & \{0,1\}^\ell\times\{0,1\}^\ell & \to & \mathbb{Z}_k\\ & (x,y) & \longmapsto & \log_g(x-y)\mod k\end{array} \tag{20}$$

We prove in the Appendix B that approximately $m=\log k\approx\frac{b_1+b_2-\ell}{2}-\log\frac{1}{\varepsilon}$ bits can be extracted by this construction.

Theorem 6 *The function* $\textsc{Ext}$ *defined above is a* (b_1,b_2,ε)-SSE *with* $\log\frac{1}{\varepsilon}=\frac{b_1+b_2-\ell}{2}+1-\log k$.

We refer to [7] for details on the efficient implementation of $\textsc{Ext}$ and the pre-computation of p, k and g.

5 Simple authentication with weak sources

5.1 Motivation

Assume that two parties Alice and Bob share the output of a ℓ_1-bit weak random source X with min-entropy $H_\infty(X)\geq b_1$. It is then clear that a (b_1,b_2,ε)-SSE $\textsc{Ext}:\{0,1\}^{\ell_1}\times\{0,1\}^{\ell_2}\to\{0,1\}^m$ can be used by both Alice and Bob to extract an almost perfectly random secret key $S=\textsc{Ext}(X,Y)$ from the shared secret information X and a weakly random public string Y with min-entropy $H_\infty(Y)\geq b_2$. This shows that super-strong extractors trivially solve the problem of privacy amplification over passive public channels when weak random sources are used. In this Section we provide a simple protocol PA for privacy amplification over an *adversarially controlled* channel, when only weak sources of randomness are available. Following [13], we show that weak sources can be used in conjunction with the simple "$ax+b$" message authentication code (MAC) to transmit the non-secret input Y over the adversarial channel. For completeness, we prove (in the Appendix C) appropriate versions of the results of [13].

Let us specify the idealized world we shall deal with when discussing the protocol. We assume that Bob can either be close (to Alice) or far from Alice. Each one of them has a weak source (specified below). If Bob is close, they can share secret information at will, but their sources should be assumed to be arbitrarily correlated

(which is reasonable for purely physical and adversarial sources). On the other hand, if Bob is far, the sources can be assumed to be independent, but only active adversarial communication channels are available. Finally, we assume that Bob's source outputs a ℓ-bit long string Y with min-entropy $H_\infty(Y) \geq b_2$, whereas Alice's source outputs three ℓ-bit strings A, B, X, which are assumed that they form a b_1-*block source* [7]. That is, for any $a, b \in \{0,1\}^\ell$, A,$B|_{A=a}$ and $X|_{A=a,B=b}$ all have min entropy at least b_1.

Our scenario differs from that of previous work on privacy amplification (e.g. [13]). In most of those works, the secret key X satisfies a min-entropy condition (that is, the adversary does not know it completely), but Alice and Bob are capable of sampling perfectly random bits. By contrast, under our constrants, no perfect randomness is available to either Alice or Bob, and geographical distance between the sources is necesary for independence, which is a reasonable assumption for physical and adversarial sources. Whereas in [13] (for instance) it is not clear that it would not be possible for the parties to agree on a perfectly random secret key when they meet in the first place, this is impossible in our case. Therefore, the need for privacy amplification is arguably better motivated in the present work. We also note that, although our assumption on Alice's source is stronger than that on Bob's source, it is still much weaker than the capability to generate truly random bits.

5.2 The protocol

Alice and Bob's aim is to agree on a secret key S that is very close to being random from Eve's point of view. This is achieved by the protocol PA which we now describe (see also Table 1 in the Appendix D), in which we identify $\{0,1\}^\ell$ with the finite field $\mathbb{F}_{2^\ell}$ for the purpose of arithmetic operations, and $\textsc{Ext} : \{0,1\}^\ell \times \{0,1\}^\ell \to \{0,1\}^m$ is a function (we will later choose it to be a suitable SSE). Briefly, Alice and Bob share (A, B, X) when Bob is close. Then Bob moves to far, samples, Y and sends $Y, Z = AY + B$ to Alice. Eve then intercepts (Y, Z), which she substitutes for $(\tilde{Y}, \tilde{Z})$ and sends to Alice. Alice checks if $A\tilde{Y} + B = \tilde{Z}$ and, if this is satisfied, she computes $\tilde{S} = \textsc{Ext}(X, \tilde{Y})$, rejecting otherwise. In the meantime, Bob has computed $S = \textsc{Ext}(X, Y)$. As we shall see (see Theorem 7), with high probability either $S = \tilde{S}$ and Alice and Bob share a secret key, or else Alice has rejected. Note that we always assume that $\tilde{Y}$ and $\tilde{Z}$ as in Table 1 have length ℓ each by establishing that Alice rejects otherwise.

We demonstrate in the Appendix C that protocol PA is indeed secure as long as $b_1 = \frac{\ell}{2} + \omega(\log \ell)$ and a $(b_1, b_2, \ell^{-\omega(1)})$-SSE exists. For instance, the number-theoretic SSE (Theorem 6) permits agreement on a key of length $m \approx \frac{b_1 + b_2 - \ell}{2} + \omega(\log \ell)$.

Theorem 7 *If* $\textsc{Ext}$ *is a* (b_1, b_2, ε)-SSE*, the protocol* PA *has the following property. If Eve is passive, Alice never rejects,* $\tilde{S} = S$ *and* $\|\langle Y, S\rangle - \langle Y, U_m\rangle\| \leq \varepsilon$*. If Eve is active, the probability of either Alice rejecting or* $S = \tilde{S}$ *and* $\|\langle Y, S\rangle - \langle Y, U_m\rangle\| \leq \varepsilon$ *is at least* $1 - 2^{\ell - 2b_1}$.

6 Acknowledgments

We thank Yan Zhong Ding, Amit Sahai, Joel Spencer and Salil Vadhan for useful discussions.

References

[1] M. Ajtai, L. Babai, P. Hajnal, J. Komlos, P. Pudlak. Two lower bounds for branching programs. In *Proceedings of the eighteenth annual ACM symposium on Theory of computing*, 30–38, 1986.

[2] N. Alon and J. Spencer. The Probabilistic Method - 2nd ed. Wiley Interscience, New York, 2000.

[3] A. Andreev, A. Clementi, J. Rolim, L. Trevisan. Dispersers, deterministic amplification, and weak random sources. In *SIAM J. on Comput.*, 28(6):2103–2116, 1999.

[4] C. H. Bennett, G. Brassard, C. Crépeau, U. Maurer. Generalized Privacy Amplification. In *IEEE Transaction on Information Theory*, vol. 41, no. 6, pp. 1915–1923, 1995

[5] C. H. Bennett, G. Brassard, and J.-M. Robert. How to reduce your enemy's information. In *Proc. of CRYPTO '85*, Lecture Notes in Computer Science, vol. 218, pp. 468–476, Springer-Verlag, 1986.

[6] R. Canetti, Y. Dodis, S. Halevi, E. Kushilevitz and A. Sahai. Exposure-Resilient Functions and All-Or-Nothing-Transforms. In *Proc. of EuroCrypt*, pp. 453–469, 2000.

[7] B. Chor, O. Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity. *SIAM J. Comput.*, 17(2):230–261, 1988.

[8] A. Cohen, A. Wigderson. Dispersers, deterministic amplification, and weak random sources. In *Proc. of FOCS*, pp. 14–19, 1989.

[9] Y. Dodis, J. Spencer. On the (Non-)Universality of the One-Time Pad. In *Proc. of FOCS*, 2002.

[10] S. Janson, T. Luksak and A. Ruciński. Random Graphs. Wiley Interscience, New York, 2000.

[11] C. Lu, O. Reingold, S. Vadhan and A. Wigderson. Extractors: Optimal Up to Constant Factors. In *Proc. of STOC*, 2003.

[12] U. Maurer and S. Wolf. Strengthening security of information-theoretic secret-key agreement. In *Proceedings of EUROCRYPT 2000*, Lecture Notes in Computer Science, Springer-Verlag, 2000.

[13] U. Maurer and S. Wolf. Privacy Amplification Secure Against Active Adversaries. In *Proc. of CRYPTO*, Lecture Notes in Computer Science, Springer-Verlag, vol. 1294, pp. 307–321, 1997.

[14] J. McInnes, B. Pinkas. On the Impossibility of Private Key Cryptography with Weakly Random Keys. In *Proc. of CRYPTO*, pp. 421–435, 1990.

[15] N. Nisan, A. Ta-Shma. Extracting Randomness: a survey and new constructions. In *JCSS*, 58(1):148–173, 1999.

[16] N. Nisan, D. Zuckerman. Randomness is Linear in Space. In *JCSS*, 52(1):43–52, 1996.

[17] L. Rónyai, L. Babai, M. Ganapathy On the number of zero-patterns in a sequence of polynomials *Journal of the AMS*, 2002.

[18] M. Sántha, U. Vazirani. Generating Quasi-Random Sequences from Semi-Random Sources. *Journal of Computer and System Sciences*, 33(1):75–87, 1986.

[19] R. Shaltiel. Recent developments in Explicit Constructions of Extractors. *Bulletin of the EATCS*, 77:67–95, 2002.

[20] R. Shaltiel and C. Umans. Simple extractors for all min-entropies and a new pseudo-random generator. In *Proceedings of FOCS 2001*, pp.648-657, IEEE Computer Society, 2001.

[21] M. Sipser. Expanders, Randomness or Time versus Space. In *Journal of Computer and Systems Sciences* 36, pp. 379-383, 1988.

[22] A. Ta-Shma and J. Radhakrishnand. Bounds for Dispersers, Extractors, and Depth-Two Superconcentrators. In *SIAM Journal on Discrete Mathematics*, 13(1):2–24, 2000.

[23] L. Trevisan. Construction of Extractors Using PseudoRandom Generators. In Proc. of STOC, pp. 141–148, 1999.

[24] L. Trevisan, S. Vadhan. Extracting Randomness from Samplable Distributions. In *Proc. of FOCS*, 2000.

[25] U. Vazirani. Randomness, Adversaries and Computation. *PhD Thesis*, University of California, Berkeley, 1986.

[26] U. Vazirani. Strong Communication Complexity or Generating Quasi-Random Sequences from Two Communicating Semi-Random Sources. *Combinatorica*, 7(4):375–392, 1987.

[27] U. Vazirani, V. Vazirani. Random polynomial time is equal to slightly-random polynomial time. In *Proc. of* 26*th FOCS*, pp. 417–428, 1985.

[28] A. Wigderson. Open problems. Notes from *DIMACS Workshop on Pseudorandomness and Explicit Combinatorial Constructions*, 1999

[29] D. Zuckerman. Simulating BPP Using a General Weak Random Source. *Algorithmica*, 16(4/5):367-391, 1996.

A Lower bounds - proof of Theorem 2

In this subsection we prove Theorem 2.

Proof: *(of Theorem 2)* We use the following lower bounds of Ta-Shma and Radhakrishnan [22] for (regular) extractors.

Theorem 8 *[22] There exists a constant c such that the following holds. Let* EXT' $: \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *be a regular (k,ε)-extractor with $d \leq m-2$ and $k \leq n-c$. Then $d \geq \log(n-k) + 2\log\frac{1}{\varepsilon} - O(1)$ and $d+k-m \geq 2\log\frac{1}{\varepsilon} - O(1)$.*

Let Y be uniform on $\{0,1\}^{b_2} \circ 0^{\ell_2 - b_2}$ and note that it has min-entropy b_2. Now set

$$\begin{array}{rccc} \text{EXT}_1: & \{0,1\}^{\ell_1} \times \{0,1\}^{b_2} & \to & \{0,1\}^{m+b_2} \\ & (x,y) & \longmapsto & y \circ \text{EXT}(x, y \circ 0^{\ell_2 - b_2}) \end{array} \tag{21}$$

For all ℓ_1-bit long random variables X with min-entropy $H_\infty(X) \geq b_1$

$$\|\text{EXT}_1(X, U_{b_2}) \ - \ U_{m+b_2}\| = \|\langle Y, \text{EXT}(X,Y)\rangle \ - \ \langle Y, U_m\rangle\| \leq \varepsilon \tag{22}$$

since $(X,Y) \in \mathsf{CG}(\ell_1, \ell_2, b_1, b_2)$ and EXT is a SSE with the adequate parameters. It follows that EXT_1 is a (b_1, ε)- extractor. By Theorem 8 we conclude

$$b_2 \leq \log(\ell_1 - b_1) + 2\log\frac{1}{\varepsilon} - O(1) \textbf{ and } b_2 + b_1 - m - b_2 = b_1 - m \geq 2\log\frac{1}{\varepsilon} \tag{23}$$

□

Remark 1 *In an attempt to prove that $b_1 \leq \log(\ell_1 - b_1) + 2\log\frac{1}{\varepsilon} - O(1)$ is necessary, one might be tempted to reverse the process in the above proof and built a regular extractor* EXT_2 *out of* EXT *for which the random variable X that is uniform on $\{0,1\}^{b_1} \circ 0^{\ell_1 - b_1}$ is the seed. The reason why this does not work is that the output length m in this case is smaller than the effective seed length b_1, and Theorem 8 does not apply to this case.*

B Proof of Theorem 6

Proof: *(of Theorem 6)* We first *claim* that the following inequality holds for all subsets $A, B, C \subseteq \mathbb{Z}_p$: setting $\Phi_C \equiv \max_{1 \leq j \leq p-1} |\sum_{c \in C} e^{\frac{2\pi i c j}{p}}|$,

$$\sum_{a \in A} \left| \#\{b \in B : a - b \in C\} - \frac{|B|\,|C|}{p} \right| \leq \Phi_C \sqrt{|A|\,|B|} \tag{24}$$

Inequality (24) is proven subsequently. Assuming it for the moment, choose $\alpha \in \mathbb{Z}_k$ and set

$$\mathcal{C} \equiv \{c \in C \mid \log_g(c) = \alpha \mod k\}$$

Following [7, Section 3.2], we note that $|C| = p/k$ and $\Phi_C < \sqrt{p}$. Hence for all $A, B \subseteq C$

$$\sum_{a \in A} \left| \#\{b \in B : \log_g(a-b) = \alpha\} - \frac{|B|}{k} \right| \leq \sqrt{p|A|\,|B|} \tag{25}$$

We deduce from (25) that for any choice of flat random variables $(X, Y) \in \mathsf{CG}(\ell_1, \ell_2, b_1, b_2)$with respective supports S_X, S_Y of sizes 2^{b_1}, 2^{b_2}

$$\begin{aligned} &\|\langle Y, \textsc{Ext}(X,Y)\rangle \ - \ \langle Y, U\rangle\| \\ &\qquad = \frac{1}{2} \sum_{\alpha \in \mathbb{Z}_k} \sum_{y \in S_Y} \left| \frac{\#\{x \in S_X : \log_g(x-y) = \alpha\}}{2^{b_1+b_2}} - \frac{1}{2^{b_1} k} \right| \leq \frac{k}{2} \sqrt{\frac{p}{2^{b_1+b_2}}} = \varepsilon \end{aligned} \tag{26}$$

and this implies the theorem by Proposition 1. □

Proof: *(of inequality (24))* This proof uses the method of trigonometric sums (a.k.a. Fourier Analysis on $\mathbb{Z}_p$) and follows closely that of Lemma 6 in [1]. For simplicity, we prove the equivalent inequality

$$\sum_{a \in A} \left| \#\{b \in B : \exists c \in C \ a + b = c\} - \frac{|B|\,|C|}{p} \right| \leq \Phi_C \sqrt{|A|\,|B|} \tag{27}$$

Let $\omega \equiv e^{\frac{2\pi i}{p}}$ and define

$$\psi_T(j) \equiv \sum_{t \in T} \omega^{tj} \ (j \in \mathbb{Z}_p,\, T \subseteq \mathbb{Z}_p)$$

For any $a \in A$, the number of $b \in B$ satisfying $a + b = c$ for some $c \in C$ is precisely

$$\frac{1}{p} \sum_{j=0}^{p-1} \omega^{ja} \psi_B(j) \psi_C(-j) = \frac{|B|\,|C|}{p} + \frac{1}{p} \sum_{j=1}^{p-1} \omega^{ja} \psi_B(j) \psi_C(-j) \tag{28}$$

as a simple calculation shows. Hence for any choice of $q_a \in \pm 1$, $a \in A$

$$q_a \left(\#\{b \in B : \exists c \in C \ a + b = c\} - \frac{|B|\,|C|}{p} \right) = \frac{q_a}{p} \sum_{j=1}^{p-1} \omega^{ja} \psi_B(j) \psi_C(-j)$$

We can now sum over $a \in A$; letting $\tilde{\psi}_A(j) \equiv \sum_{a \in A} q_a \omega^{aj}$

$$\sum_{a \in A} q_a \left(\#\{b \in B : \exists c \in C \ a + b = c\} - \frac{|B|\,|C|}{p} \right) = \frac{1}{p} \sum_{j=1}^{p-1} \tilde{\psi}_A(j) \psi_B(j) \psi_C(-j) \tag{29}$$

By an appropriate choice of the q_a's, it is possible to conclude that in fact

$$\sum_{a\in A}\left|\#\{b\in B:\exists c\in C\ a+b=c\}-\frac{|B|\,|C|}{p}\right|=\frac{1}{p}\sum_{j=1}^{p-1}\tilde{\psi}_A(j)\psi_B(j)\psi_C(-j) \tag{30}$$

Applying Cauchy Schwartz to the RHS and noting that

$$\sum_{j=1}^{p-1}|\tilde{\psi}_A(j)|^2\le p|A| \textbf{ and } \sum_{j=1}^{p-1}|\psi_B(j)\psi_C(-j)|^2\le\Phi_C^2 p|B|$$

where $\Phi_C\equiv\sup_{1\le j\le p-1}|\psi_C(j)|$, we can finally bound

$$\sum_{a\in A}\left|\#\{b\in B:\exists c\in C\ a+b=c\}-\frac{|B|\,|C|}{p}\right|\le\Phi_C\sqrt{|A||B|} \tag{31}$$

This finishes the proof. □

C Proof of Theorem 7

We first prove a lemma on the "$ax+b$" MAC that is similar to Theorem 6 in [13].

Lemma 6 *Let $g:\mathbb{F}_{2^\ell}\times\mathbb{F}_{2^\ell}\to\mathbb{F}_{2^\ell}\times\mathbb{F}_{2^\ell}$ be a function, $w\in\mathbb{F}_{2^\ell}$, and $\langle C,D\rangle$ be two ℓ-bit long strings with joint min-entropy $H_\infty(\langle C,D\rangle)\ge 2b_1$. Define $(T,V)\equiv g(w,Cw+D)$. Then*

$$\Pr\big((T,V)\ne(w,Cw+D)\textit{ and }CT+D=V\big)\le 2^{\ell-2b_1}$$

Proof: *(of Lemma 6)* Fix $C=c$, $D=d$ and let $s\equiv cw+d$. The pair (T,V) is then completely determined by the value of s. Now assume that (c,d) is *bad*, that is, $(T,V)=(t(s),v(s))\ne(w,s)$ but $ct(s)+d=v(s)$. The pair (c,d) must then satisfy the following system of equations

$$\begin{pmatrix}1 & w\\ 1 & t(s)\end{pmatrix}\begin{pmatrix}c\\ d\end{pmatrix}=\begin{pmatrix}s\\ v(s)\end{pmatrix} \tag{32}$$

for $s\in\mathbb{F}_{2^\ell}$. It cannot be true that $t(s)=w$, for that would imply that $v(s)=ct(s)+d=cw+d$ and $(t(s),v(s))=(w,cw+d)$, so it must hold that $w\ne t(s)$. This implies that the matrix in (32) is non-singular and that the corresponding system of equations has exactly one solution. We conclude that for each possible value of $s=cw+d$ there can be at most one *bad* pair (c,d), and that this pair is completely determined by (32). Since there are 2^ℓ possible values for s and each pair (c,d) has probability $\le 2^{-2b_1}$, the probability of the sampled value of (C,D) being bad is $\le 2^{\ell-2b_1}$. This is precisely the desired result. □

Proof: *(of Theorem 7)* It suffices to treat the case of a deterministic active adversary. That is, Eve's strategy for producing $\langle\tilde{Y},\tilde{Z}\rangle$ is to use a deterministic function of Y and $Z\equiv AY+B$. Lemma 6 implies that for any value y of Y the probability that $\tilde{Z}=A\tilde{Y}+B$ and $\tilde{Y}\ne y$ is at most $2^{\ell-2b_1}$. Assuming that this event does not happen, Alice does not reject and $S=\tilde{S}$. Moreover,

$$\|\langle Y,AY+B,S\rangle\ -\ \langle Y,AY+B,U_m\rangle\|\le\max_{a,b\in\{0,1\}^{\ell_1}}\|\langle Y,\textsc{Ext}(X|_{A=a,B=b}Y)\rangle\ -\ \langle Y,U_m\rangle\|\le\varepsilon \tag{33}$$

by the SSE property and the block source condition on A,B,X. □

D The Protocol PA

PART 1 - Bob is close		
Alice	**(secret info)**	**Bob**
samples (A, B, X)	$\xrightarrow{(A,B,X)}$	stores (A, B, X)

PART 2 - Bob is far		
Alice	**Eve (channel)**	**Bob**
		samples Y $Z \equiv AY + B$
receives $(\tilde{Y}, \tilde{Z})$	$\xleftarrow{(Y,Z)\rightarrow(\tilde{Y},\tilde{Z})}$	sends (Y, Z)
if $A\tilde{Y} + B \neq \tilde{Z}$ reject else $\tilde{S} \equiv \text{EXT}(X, Y)$ accept		$S \equiv \text{EXT}(X, Y)$ accept

Table 1: Description of protocol PA for privacy amplification. All random variables X, Y, A, B, $\tilde{Y}$ and $\tilde{Z}$ take values in $\{0,1\}^\ell \approx \mathbb{F}_{2^\ell}$.