Affine-malleable Extractors, Spectrum Doubling, and Application to Privacy Amplification

Divesh Aggarwal* Kaave Hosseini† Shachar Lovett‡

November 10, 2015

Abstract

The study of seeded randomness extractors is a major line of research in theoretical computer science. The goal is to construct deterministic algorithms which can take a "weak" random source X with min-entropy k and a uniformly random seed Y of length d, and outputs a string of length close to k that is close to uniform and independent of Y. Dodis and Wichs [DW09] introduced a generalization of randomness extractors called non-malleable extractors (nmExt) where $\mathsf{nmExt}(X, Y)$ is close to uniform and independent of Y and $\mathsf{nmExt}(X, f(Y))$ for any function f with no fixed points.

We relax the notion of a non-malleable extractor and introduce what we call an affine-malleable extractor ($\mathsf{AmExt} : \mathbb{F}^n \times \mathbb{F}^d \mapsto \mathbb{F}$) where $\mathsf{AmExt}(X, Y)$ is close to uniform and independent of Y and has some limited dependence of $\mathsf{AmExt}(X, f(Y))$ - that conditioned on Y, $(\mathsf{AmExt}(X, Y), \mathsf{AmExt}(X, f(Y)))$ is close to $(U, A \cdot U + B)$ where U is uniformly distributed in $\mathbb{F}$ and $A, B \in \mathbb{F}$ are random variables independent of U.

We show under a plausible conjecture in additive combinatorics (called the Spectrum Doubling Conjecture) that the inner-product function $\langle \cdot, \cdot \rangle : \mathbb{F}^n \times \mathbb{F}^n \mapsto \mathbb{F}$ is an affine-malleable extractor. As a modest justification of the conjecture, we show that a weaker version of the conjecture is implied by the widely believed Polynomial Freiman-Ruzsa conjecture.

We also study the classical problem of privacy amplification, where two parties Alice and Bob share a weak secret X of min-entropy k, and wish to agree on secret key R of length m over a public communication channel completely controlled by a computationally unbounded attacker Eve. The main application of non-malleable extractors and its many variants has been in constructing secure privacy amplification protocols.

We show that affine-malleable extractors along with affine-evasive sets can also be used to construct efficient privacy amplification protocols. We show that our protocol, under the Spectrum Doubling Conjecture, achieves near optimal parameters and achieves additional security properties like source privacy that have been the focus of some recent results in privacy amplification.

*EPFL. Email: Divesh.Aggarwal@epfl.ch.

†UCSD. Email: kaave.hosseini@gmail.com.

‡UCSD. Email: slovett@cs.ucsd.edu.

1 Introduction

The study of randomness extractors is a major line of research in theoretical computer science. The goal is to construct deterministic algorithms which can take one or more "weak" random sources and transform them into a "strong" random source, which typically means close to uniform. The challenge is in coming with explicit constructions, as often it is easy to show that a random construction is close to optimal. These constructions come in two flavours: the first are "simple" or "natural" mathematical constructions, such as the inner product function; the others are more complex constructions, usually obtained as a combination of a number of more basic constructions. For a general background to extractors and their many applications, see e.g. [Sha02].

The focus of this paper is on the first family of "simple" mathematical constructions, and in particular the inner product function. Let $\mathbb{F} = \mathbb{F}_p$ be a finite field, and consider the inner product function over $\mathbb{F}^n$ defined as $\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i$. For a random variable X taking values in $\mathbb{F}^n$, its min-entropy is defined as $H_\infty(X) = -\log(\max \Pr[X = x])$. If $H_\infty(X) = \delta \cdot \log |\mathbb{F}|^n = \delta(n \log p)$ then we say X has min-entropy rate δ. It is well known since the work of [CG88] that the inner product function is a two-source extractor for min-entropy rate $1/2$. That is, if X, Y are two independent random variables taking values in $\mathbb{F}^n$, and $H_\infty(X), H_\infty(Y) \gg (n \log p)/2$, then $\langle X, Y\rangle$ is close to uniform in $\mathbb{F}$. This bound of $1/2$ on the min-entropy rate is tight, as can be seen be the following example: take X to be uniform over a subspace of $\mathbb{F}^n$ of dimension $n/2$, and Y to be uniform over the orthogonal subspace. The best known construction of two-source extractors, until very recently, was by Bourgain [Bou05], who showed that by pre-encoding the input as $(x, x^2) \in \mathbb{F}^{2n}$ (at least when the field has odd order) and computing the inner product over $\mathbb{F}^{2n}$, one can handle source of min-entropy rate $1/2 - \varepsilon_0$ for some absolute constant $\varepsilon_0 > 0$.

The study of two-source extractors turns out to be related to a more recent notion of extractors, called *non-malleable extractors*. These were defined by Dodis and Wichs [DW09] for the purpose of designing privacy amplification protocols. They showed that such extractors would allow for optimal protocols, and showed that a random construction achieves optimal parameters. However, they left the challenge of explicit constructions open. Subsequent works [CRS12, Li12c, DY13, DLWZ14] gave constructions for non-malleable extractors for min-entropy rate $1/2-\varepsilon_0$, based on the inner product function and the construction of two-source extractor by Bourgain. Interesting enough, Li [Li12c] showed a strong connection between non-malleable extractors and two-source extractors, thereby giving evidence that it might be very difficult to construct non-malleable extractors for min-entropy rate below $1/2 - \varepsilon_0$. In a recent breakthrough work, Chattopadhyay, Goyal and Li [CGL15] have recently obtained a non-malleable extractor with min-entropy polylog$(n \log p)$. This was closely followed by a another breakthrough result by Chattopadhyay and Zuckerman [CZ15], achieving an efficient two-source extractor where each source has polylogarithmic min-entropy. Our focus in this paper is to obtain much more modest results than these obtained by [CGL15, CZ15], however we restrict ourselves to the use of very simple function - the inner-product function.

We note here that the inner product function, if considered as a potential non-malleable extractor with a source $X \in \mathbb{F}^n$ and uniform seed in $\mathbb{F}^n$, still suffers from the barrier of orthogonal subspaces, and hence is not a non-malleable extractor for any min-entropy rate. First, let us define non-malleable extractors formally, specialized to the parameters of our function of interest, the inner product function. A function $E : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ is a non-malleable extractor for min-entropy rate k, if for any random variable X taking values in $\mathbb{F}^n$ with min-entropy k; for a seed S uniform in $\mathbb{F}^n$ and independent of X; and for any function $f : \mathbb{F}^n \to \mathbb{F}^n$, the joint distribution

$$(E(X, S), E(X, f(S))) \in \mathbb{F}^2$$

lies close to a distribution in the convex hull of $\{(U, U)\} \cup \{(U, a) : a \in \mathbb{F}\}$, where $U = U_{\mathbb{F}}$ is the

uniform distribution over $\mathbb{F}$. That is, this family of distribution contains the case where the two outputs are equal and uniform in $\mathbb{F}$ (which would be the case if say f is the identity function), or where the first term is uniform in $\mathbb{F}$ and the second is independent of it (but not necessarily uniform), or any convex combination of the above.

To see why the inner product function fails to be a non-malleable extractor, consider the following example. Let $V \subset \mathbb{F}^n$ be a subspace, let $u \in \mathbb{F}^n \setminus V$ and $v \in V^\perp$ be vectors such that $\langle u, v \rangle = 1$. If we choose X uniform in the coset $V + u$ and set $f(S) = aS + bv$ for some $a, b \in \mathbb{F}$ then approximately

$$(\langle X, S \rangle, \langle X, f(S) \rangle) \sim (U, aU + b)\,.$$

In this work, we define a weaker type of extractor, which we call *affine-malleable extractor*, motivated by the above example. A distribution over $\mathbb{F}^2$ is said to be *affine* if it lies in the convex hull of $\{(U, aU + b) : a, b \in \mathbb{F}\}$. A function $E : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ is an affine-malleable extractor for min-entropy rate k, if for any random variable X taking values in $\mathbb{F}^n$ with min-entropy k; for a seed S uniform in $\mathbb{F}^n$ and independent of X; and for any function $f : \mathbb{F}^n \to \mathbb{F}^n$, the joint distribution

$$(E(X, S), E(X, f(S))) \in \mathbb{F}^2$$

lies close to an affine distribution. We prove a number of results on affine-malleable extractors:

1. Affine-malleable extractors, combined with recent constructions of affine-evasive sets [ADL14, Agg15], allow for optimal privacy amplification protocols. Thus, for the application in privacy amplification they are as useful as non-malleable extractors.

2. The inner product function is an affine-malleable extractor for any constant min-entropy rate $\delta > 0$, assuming the dimension n is large enough. This is the main technical part of this work.

3. We give a conjecture in additive combinatorics that implies that the inner product function is an affine-malleable extractor also for much smaller min-entropy, which implies near optimal privacy amplification protocols.

The analysis of the inner product function, and in particular showing that it is an affine-malleable extractor, relies on new properties of the spectrum of large sets, which we discuss next.

1.1 Spectrum of large sets

The analysis of the inner product function relies on analyzing the spectrum of subsets of $\mathbb{F}^n$. For a set $A \subset \mathbb{F}^n$, its Fourier coefficients are defined as

$$\widehat{1_A}(\gamma) = \sum_{a \in A} \omega^{\langle \gamma, x \rangle}, \qquad \gamma \in \mathbb{F}^n,$$

where $\omega = \exp(2\pi i/p)$ is a primite p-th root of unity. The *spectrum* of A is the set of large Fourier coefficients,

$$\mathrm{Spec}_\varepsilon(A) = \left\{ \gamma \in \mathbb{F}^n : |\widehat{1_A}(\gamma)| \geqslant \varepsilon |A| \right\}.$$

The spectrum of a set is extremely useful when studying arithmetic properties. Several properties of it are well known and quite useful. Parseval's identity allows to bound the size of the spectrum if A is not too small,

$$|\mathrm{Spec}_\varepsilon(A)| \leqslant O\left(\frac{\mathbb{F}^n}{\varepsilon^2 |A|} \right).$$

Chang's theorem [Cha02] allows to bound the dimension of the large spectrum (a notion which we do not formally define here) by

$$\dim \mathrm{Spec}_\varepsilon(A) \leqslant O\left(\log(\mathbb{F}^n/|A|) \cdot 1/\varepsilon^2\right).$$

An application of the Cauchy-Schwartz inequality shows that the spectrum has bounded statistical doubling. That is, the sum of two random elements in the spectrum typically lies in the spectrum with a smaller threshold,

$$\Pr_{\gamma_1,\gamma_2 \in \mathrm{Spec}_\varepsilon(A)}[\gamma_1 - \gamma_2 \in \mathrm{Spec}_{\varepsilon^2/2}(A)] \geqslant \varepsilon^2/2.$$

For a proof and some extensions, see eg the work of Shkredov [Shk08]. For our purposes, it turns out that the important property is to control the growth of the spectrum in addition. In the following, for $S \subset \mathbb{F}^n$ let $S+S = \{s_1+s_2 : s_1, s_2 \in S\}$ and $\mathbb{F}\cdot S = \{cs : c \in \mathbb{F}, s \in S\}$. The following definition will be crucial for us.

Definition 1. A set $A \subset \mathbb{F}^n$ has (ε,η)-controlled spectrum if

$$|\mathbb{F}\cdot \mathrm{Spec}_\varepsilon(A) + \mathbb{F}\cdot \mathrm{Spec}_\varepsilon(A)| \leqslant \eta|\mathbb{F}^n|.$$

To see the relation to affine-malleable extractors, let X be uniform over a set $A \subset \mathbb{F}^n$, let S be uniform in $\mathbb{F}^n$, and let $f : \mathbb{F}^n \to \mathbb{F}^n$ be a function. If $(\langle X, S\rangle, \langle X, f(S)\rangle)$ is not close to affine, then this remains true if we fix S. That is, for a noticeable fraction of values $s \in \mathbb{F}^n$, the joint distribution

$$(\langle X, s\rangle, \langle X, f(s)\rangle)$$

is also not close to affine. Hence, it must have a noticeable Fourier coefficient, which is equivalent to

$$as + bf(s) \in \mathrm{Spec}_\varepsilon(A)$$

for some $a, b \in \mathbb{F}$ not both zero and $\varepsilon \leqslant p^{-O(1)}$ which depends on the error parameters. In fact, the set of Fourier coefficients cannot be one dimensional. If we consider

$$\Lambda_s := \{(a,b) \in \mathbb{F}^2 : as + bf(s) \in \mathrm{Spec}_\varepsilon(A)\}$$

then Λ_s contains $(a_1,b_1),(a_2,b_2)$ which are not scalar multiples of each other. So, if

$$a_1 s + b_1 f(s) = \gamma_1$$
$$a_2 s + b_2 f(s) = \gamma_2$$

where $\gamma_1, \gamma_2 \in \mathrm{Spec}_\varepsilon(A)$, then

$$s = c_1\gamma_1 + c_2\gamma_2$$

for some $c_1, c_2 \in S$. As this is true for many choices of $s \in \mathbb{F}^n$ (say, more than an η fraction), we get that

$$|\mathbb{F}\cdot \mathrm{Spec}_\varepsilon(A) + \mathbb{F}\cdot \mathrm{Spec}_\varepsilon(A)| > \eta|\mathbb{F}^n|$$

and A does not have (ε,η)-controlled spectrum.

So, the question arises: can we hope that any large enough set has (ε,η)-controlled spectrum. Applying Parseval's identity we get the bound

$$|\mathbb{F}\cdot \mathrm{Spec}_\varepsilon(A) + \mathbb{F}\cdot \mathrm{Spec}_\varepsilon(A)| \leqslant |\mathbb{F}|^2|\mathrm{Spec}_\varepsilon(A)|^2 \leqslant p^2\left(\frac{p^{n+1}}{\varepsilon^2|A|}\right)^2.$$

The following is immediate from this.

Theorem 1.1. *Let p be a prime, $n \in \mathbb{N}$, and $0 < \varepsilon < 1$. For any $A \subseteq \mathbb{F}^n$, such that $|A| \geqslant \frac{1}{\varepsilon^3} \cdot p^{n/2+3}$,*

$$|\mathbb{F} \cdot \mathit{Spec}_\varepsilon(A) + \mathbb{F} \cdot \mathit{Spec}_\varepsilon(A)| \leqslant \varepsilon \cdot p^n \ .$$

Again, this threshold of min-entropy rate $1/2$ turns out to be tight, by a variant of the orthogonal subspaces construction. If we take $A = (\mathbb{F}^{n/2} \times \{0\}^{n/2}) \cup (\{0\}^{n/2} \times \mathbb{F}^{n/2})$ then it is easy to verify that $\mathrm{Spec}_{1/2}(A) = A$ and that $\mathrm{Spec}_{1/2}(A) + \mathrm{Spec}_{1/2}(A) = \mathbb{F}^n$.

So, are we again stuck at the min-entropy rate $1/2$ barrier? The answer is no. In the analysis of affine-malleable extractors above, if we can partition A as $A = \cup A_i$, where each A_i has (ε, η)-controlled spectrum, then the analysis would still work. That is, if we define X_i to be uniform on A_i then

$$(\langle X_i, S\rangle, \langle X_i, f(S)\rangle)$$

are each close to affine, and hence also $(\langle X, S\rangle, \langle X, f(S)\rangle)$ being a convex combination of these, is also close to affine. The same would hold if the partition $\cup A_i$ covered most of A. For example, in the example mentioned above of $A = (\mathbb{F}^{n/2} \times \{0\}^{n/2}) \cup (\{0\}^{n/2} \times \mathbb{F}^{n/2})$ we would take $A_1 = \mathbb{F}^{n/2} \times \{0\}^{n/2}$ and $A_2 = \{0\}^{n/2} \times \mathbb{F}^{n/2}$, both of which have a controlled spectrum.

So, the real question is the following: is it true that any large enough set A contains a subset which has (ε, η)-controlled spectrum. If this is true, then the analysis would still work. We could not prove it for smaller sets A. Two of the authors of this paper [HL15] showed a bound on $|\mathrm{Spec}_\varepsilon(A) + \mathrm{Spec}_\varepsilon(A)|$ for $|A| \ll p^{n/2}$. It does not seem easy, however, to extend their results to bound $|\mathbb{F} \cdot \mathrm{Spec}_\varepsilon(A) + \mathbb{F} \cdot \mathrm{Spec}_\varepsilon(A)|$. We show the following assuming the widely believed polynomial Freiman-Ruzsa conjecture. We refer the reader to [BDL14] and the references therein for a discussion of this conjecture.

Theorem 1.2. *Assuming the PFR conjecture, we have that there is a universal constant C such that for any $A \subseteq \mathbb{F}^n$, such that $|A| \geqslant \frac{1}{\varepsilon} \cdot p^{n\sqrt{C \log p / \log n}}$, then there exists a subset $B \subseteq A$ such that*

$$|\mathbb{F} \cdot \mathit{Spec}_\varepsilon(B) + \mathbb{F} \cdot \mathit{Spec}_\varepsilon(B)| \leqslant \varepsilon \cdot p^n \ . \tag{1.1}$$

We conjecture that PFR conjecture might not precisely capture the difficulty of bounded the doubling of the spectrum, and that Theorem 1.2 is far from tight. We believe that the above holds for $|A|$ sufficiently bigger than $1/\varepsilon$, and when ε is sufficiently small. The reason we cannot conjecture the above for $|A| \geqslant (1/\varepsilon)^{O(1)}$ is because of the following example. If B is any subset of the hamming ball $A = \{x \in \{0,1\}^n : \sum x_i \leqslant \log(1/\varepsilon)\}$ then $\mathrm{Spec}_\varepsilon(B)$ contains most vectors in $\mathbb{F}^n$ with $n/2$ zeros, and hence $|\mathbb{F} \cdot \mathrm{Spec}_\varepsilon(B) + \mathbb{F} \cdot \mathrm{Spec}_\varepsilon(B)| \approx |\mathbb{F}^n|$. We conjecture that this example gives the worst possible bound on the size of $|A|$.

Conjecture 1.3 (Spectrum Doubling Conjecture)**.** *For any $A \subseteq \mathbb{F}^n$ and any $\varepsilon \in (0, 1/p)$ such that $|A| \geqslant n^{O(\log(1/\varepsilon))} = (1/\varepsilon)^{O(\log(n))}$, there is a subset B of A such that*

$$|\mathbb{F} \cdot \mathit{Spec}_\varepsilon(B) + \mathbb{F} \cdot \mathit{Spec}_\varepsilon(B)| \leqslant \varepsilon \cdot p^n \ .$$

1.2 Privacy Amplification

One of the classical problems in information-theoretic cryptography is that of privacy amplification (PA) [BBR88, Mau92, BBCM95]. In this, two parties, Alice and Bob, share a weak secret X (of length N bits and min-entropy $k < N$) and wish to securely (with security bounded by ε) agree on an ε-close to uniform secret key R of length m bits via a communication channel fully controlled by the adversary, Eve. This problem arises in a number of applications, such as biometric authentication, leakage-resilient cryptography, and quantum cryptography.

In the case when the adversary is passive, i.e., only sees but does not modify the message(s) sent over the channel, this problem can be easily solved in a single round using a strong extractor Ext. This can be done by Alice sending a seed W for the extractor, and then Alice and Bob both compute the secret $\mathsf{Ext}(X, W)$ which is close to being uniformly random and independent of W by the strong extractor property.

However, it becomes significantly more involved to design an efficient protocol that is secure for the case when the adversary is active and can modify messages. This problem has been studied extensively in the recent years [MW97, RW03, DKK+12, DW09, CKOR10, DLWZ14, CRS12, Li12c, Li12a, Li12b, DY13, ADJ+14] with the goal of optimizing several natural quantities like (i) The min-entropy k (ii) The entropy loss $L = k - m$ (iii) The number of rounds of the protocol.

Prior Results. Existing one-round solutions [MW97, DKK+12] work for min-entropy $k > N/2$ and require large entropy loss $L > N - k$. It was shown by [DS02, DW09] that $k > N/2$ is necessary, and that the large entropy loss of $N - k$ is likely necessary, as well.

So, in order to achieve PA for $k < N/2$, we require at least two rounds. In a breakthrough result, Dodis and Wichs [DW09] showed non-constructively that a two-round PA protocol exists with optimal (up to constant factors) entropy loss $L = \Theta(\log(1/\varepsilon))$ for any k. This was achieved by defining and showing existence of non-malleable extractors.

Since [DW09] could not construct a non-malleable extractor, they instead defined and constructed look-ahead extractors which was done using the alternating extraction protocol of Dziembowski and Pietrzak [DP07]. Look-ahead extractor laExt is an extractor with weaker non-malleable requirements where $\mathsf{laExt}(X, Y)$ is allowed to have some limited dependence with $\mathsf{laExt}(X, \mathcal{A}(Y))$. PA protocols using look-ahead extractors were constructed only for $k > \log^2(1/\varepsilon)$. This protocol is meaningful only in the scenario when the error $\varepsilon = 2^{-O(\sqrt{N})}$ and so would not work if we require a smaller error.

Subsequent works [DLWZ14, CRS12, Li12c, DY13] gave constructions for non-malleable extractors but none of these were for $k < 0.499N$. Very recently, a non-malleable extractor has been constructed by Chattopadhyay et al [CGL15] for $k = \Omega(\log^2 N + \log^2(1/\varepsilon))$.

In another work, Li [Li12a] defined and constructed one such weaker notion called non-malleable condensers nmCond, where he weakened the extractor requirement so that $\mathsf{nmCond}(X, Y)$ is not required to be close to uniform but rather only has a better min-entropy rate guarantee as compared X. In a subsequent work [Li15], he improved the non-malleable condenser constructions thereby giving two different PA protocols using non-malleable condensers with optimal entropy loss $L = \Theta(\log(1/\varepsilon))$. However, one of these protocols works only for $k > \log^2(1/\varepsilon)$, and the other only for $k \geqslant \delta N$ for any constant δ. We note here that in order to achieve the more useful security notion of post-application authenticity [1] with the same entropy loss, we need to add one additional round as observed by Aggarwal et al [ADJ+14].

Another related goal that has been considered in the literature [DS05, BF11, ADJ+14] for PA protocols is that of *source privacy*. Intuitively, this property demands that the transcript of the protocol (even together with the derived key R!) does not reveal any "useful information" about the source X; or, equivalently (as shown by [DS05]), that the transcript does not reveal any information at all about the *distribution* of X (beyond a lower bound k on its min-entropy). For active Eve, the only works that considered this notion are [BF11, ADJ+14]. All protocols achieving source privacy considered in these works are secure only for $k > \log^2(1/\varepsilon)$, or for $k \geqslant 0.499N$.

[1] Intuitively, post-application authenticity requires that Eve cannot modify Bob's last message and cause Alice to output $R_A \neq R_B$, even if given Bob's key R_B. This is useful to protect against attacks where Bob might begin to use his key and Eve can use this information to modify the last message sent to Alice.

Result	Entropy	Rounds	Entropy Loss	Source Privacy
[DW09, ADJ+14] (non-explicit)	$k = \Omega(\log(1/\varepsilon))$	2	$\Theta(\log(1/\varepsilon))$	YES
[DKK+12]	$k > \frac{n}{2}$	1	$\frac{n}{2} + \Theta(\log(1/\varepsilon))$	YES
[DW09]	$k = \Omega(\log^2(1/\varepsilon))$	2	$\Theta(\log^2(1/\varepsilon))$	NO
[Li15, ADJ+14]	$k = \Omega\left(\min\left(\delta n, \log^2(1/\varepsilon)\right)\right)$	3	$\Theta(\log(1/\varepsilon))$	NO
[ADJ+14]	$k = \Omega\left(\min\left(\delta n, \log^2(1/\varepsilon)\right)\right)$	5	$\Theta(\log(1/\varepsilon))$	YES
This work (assuming Conjecture 1.3)	$k = \widetilde{\Omega}(\log(1/\varepsilon))$	2	$\Theta(\log(1/\varepsilon))$	YES

Table 1: Comparison of our PA protocol with previous best-known protocols.

We note that in a typical PA application we would require the error to be significantly smaller than $2^{-O(\sqrt{N}}$ (if say N is not too large), in which case all protocols that work for min-entropy above $\log^2(1/\varepsilon)$ will be useless.

Our Result. Our PA protocol using an affine-malleable extractor and an affine-evasive set can be described as follows. Like in the case of a passive Eve that we mentioned earlier, the aim of the protocol is to enable Bob to send an authenticated copy of a seed W for a strong extractor Ext. Bob chooses the seed W as a random element of an affine-evasive set $S \subset \mathbb{F}$. In the first round Alice sends a seed Y for the affine-malleable extractor amExt, which can possibly be changed by the adversary to $\mathcal{A}(Y) = Y'$. Then in the second round, Bob sends $T = W + \mathsf{amExt}(X, Y')$. As we show in Section 4, by the definition of an affine-malleable extractor and an affine-evasive set, it is hard for the adversary to change T to some $T' = W' + \mathsf{amExt}(X, Y)$, where $W' \neq W$ and W' is an element of S. Moreover, since both the messages sent in the first round and in the second are completely independent of X, we get source-privacy for free in this protocol.

We observe and give justification for the fact that the inner product function from $\mathbb{F}^n \times \mathbb{F}^n$ to $\mathbb{F}$ for a large prime p is likely to be a good affine-malleable extractor. In fact, under a plausible conjecture that we call the spectrum doubling conjecture, our 2-round PA protocol using the inner product extractor as an affine-malleable extractor achieves near-optimal min-entropy and entropy loss requirements.

Theorem 1.4. *Assuming conjecture 1.3, there exists an explicit polynomial-time, two-round (k, ε)-private, (k, m, ε)-secure privacy amplification protocol, where $k \geqslant O(\log(1/\varepsilon)\log N)$, and $k - m \leqslant (\log(1/\varepsilon))$.*

A comparison of our result (under Conjecture 1.3) and the known results is summarized in Table 1. All results achieve ε-secure post application robust PA protocols, where we assume that $\varepsilon = 2^{-n^c}$ where c is any constant in $(0, 1)$. Also δ is assumed to be a constant.

1.3 Organization of this Paper

The paper is organized as follows. In Section 2, we introduce the mathematical preliminaries and known primitives used for our construction. In Section 3, we show that the spectrum doubling bound implies that the inner-product function is an affine-malleable extractor. In Section 4, we formally introduce the problem of privacy amplification and give the construction and security proof of our new 2 round privacy amplification protocol assuming an affine-malleable extractor. In

Section 5 we prove Theorem 1.2, and then conclude with the implication on the construction of affine-malleable extractors and our privacy amplification protocol.

1.4 Conclusions and Open Questions

In this work, we initiate the study of non-malleability properties of the inner-product function, which is one of the first family of simple mathematical constructions of a randomness extractor. We show, under a plausible conjecture, that even though $\langle \cdot, \cdot \rangle : \mathbb{F}^n \times \mathbb{F}^n \mapsto \mathbb{F}$ is not a non-malleable extractor, it has strong non-malleability properties, i.e. that if X has sufficient min-entropy and Y is uniform in $\mathbb{F}^n$, then for any function $f : \mathbb{F}^n \mapsto \mathbb{F}^n$, $\langle X, f(Y) \rangle$ has distribution close to an affine function of $\langle X, Y \rangle$.

Even though we do not achieve a non-malleable extractor, we manage to construct a PA protocol (which is the main and most direct application of a non-malleable extractor) using this weaker notion of a non-malleable extractor that we call an affine-malleable extractor along with an affine-evasive set [Agg15]. We show that under Conjecture 1.3, our PA protocol achieves near optimal parameters. In addition, in our protocol, we get the source privacy property for free unlike the recent results of [BF11, ADJ+14] where the previous protocols had to be modified significantly (thereby making them more complicated) in order to achieve source privacy.

We note that our protocol is the first protocol that under a reasonable conjecture achieves privacy amplification with near optimal min-entropy. To the best of our knowledge, none of the previous constructions could even be conjectured to achieve similar parameters.

This work leaves open several questions but the most interesting question motivated by this work is to prove Conjecture 1.3. Even proving a weaker version of the conjecture like Theorem 1.2 without assuming the PFR conjecture will be very interesting. We would like to mention that we have stated Conjecture 1.3 the way it is since this is the best bound for which we do not have a counter-example. We encourage the reader to suggest a counter-example which would imply that we will need to relax the lower bound on A.

2 Preliminaries

For a set S, we let U_S denote the uniform distribution over S. For an integer $m \in \mathbb{N}$, we let U_m denote the uniform distribution over $\{0,1\}^m$, the bit-strings of length m. For a distribution or random variable X we write $x \leftarrow X$ to denote the operation of sampling a random x according to X. For a set S, we write $s \leftarrow S$ as shorthand for $s \leftarrow U_S$.

Entropy and Statistical Distance. The *min-entropy* of a random variable X is defined as $\mathbf{H}_\infty(X) \stackrel{\text{def}}{=} -\log(\max_x \Pr[X = x])$. We say that X is an (n,k)-*source* if $X \in \{0,1\}^n$ and $\mathbf{H}_\infty(X) \geqslant k$. For $X \in \{0,1\}^n$, we define the *entropy rate* of X to be $\mathbf{H}_\infty(X)/n$. We also define *average (aka conditional) min-entropy* of a random variable X conditioned on another random variable Z as

$$\mathbf{H}_\infty(X|Z) \stackrel{\text{def}}{=} -\log\left(\mathbb{E}_{z \leftarrow Z}\left[\max_x \Pr[X = x|Z = z]\right]\right) = -\log\left(\mathbb{E}_{z \leftarrow Z}\left[2^{-\mathbf{H}_\infty(X|Z=z)}\right]\right)$$

where $\mathbb{E}_{z \leftarrow Z}$ denotes the expected value over $z \leftarrow Z$. We have the following lemma.

Lemma 2.1 ([DORS08])**.** *Let (X, W) be some joint distribution. Then,*

- *For any $s > 0$, $\Pr_{w \leftarrow W}[\mathbf{H}_\infty(X|W = w) \geq \mathbf{H}_\infty(X|W) - s] \geq 1 - 2^{-s}$.*
- *If Z has at most 2^ℓ possible values, then $\mathbf{H}_\infty(X|(W,Z)) \geq \mathbf{H}_\infty(X|W) - \ell$.*

The *statistical distance* between two random variables W and Z distributed over some set S is

$$\Delta(W,Z) \stackrel{def}{=} \max_{T\subseteq S} |W(T) - Z(T)| = \frac{1}{2}\sum_{s\in S} |W(s) - Z(s)|.$$

Note that $\Delta(W,Z) = \max_D(\Pr[D(W)=1] - \Pr[D(Z)=1])$, where D is a probabilistic function. We say W is ε-close to Z, denoted $W \approx_\varepsilon Z$, if $\Delta(W,Z) \leq \varepsilon$. We write $\Delta(W,Z|Y)$ as shorthand for $\Delta((W,Y),(Z,Y))$, and note that $\Delta(W,Z|Y) = \mathbb{E}_{y\leftarrow Y}\Delta(W|Y=y, Z|Y=y)$.

Lemma 2.2. *Let $\mathbb{F} = \mathbb{F}_p$ be a finite field, and let $A, B \in \mathbb{F}$ be random variables. If $\Delta(A+\alpha\cdot B, U_\mathbb{F}) \leqslant \varepsilon$ for all $\alpha \in \mathbb{F}$, then $\Delta((A,B);(U_\mathbb{F},B)) \leqslant \varepsilon\cdot\sqrt{p}$.*

The proof uses the following inequality.

$$\forall x_1,\ldots,x_k > 0,\ \sum_{i=1}^k x_i^2 \leqslant \left(\sum_{i=1}^k x_i\right)^2 \leqslant k\sum_{i=1}^k x_i^2 \tag{2.1}$$

Proof of Lemma 2.2. We first show the following claim.

Claim 2.3. *Let $C, D \in \mathcal{X}$ be a pair random variables for some set $\mathcal{X}$. Then if C', D' are identically distributed as C, D and independent from them, then*

$$\frac{4\Delta((C,D)\ ;\ (U_\mathcal{X},D)))^2}{|\mathcal{X}|} \leqslant \Pr[C=C', D=D'] - \frac{\Pr[D=D']}{|\mathcal{X}|} \leqslant 4\Delta((C,D)\ ;\ (U_\mathcal{X},D))^2\ .$$

Proof. We have

$$\sum_{x,y\in\mathcal{X}}\left(\Pr[C=x,D=y] - \frac{\Pr[D=y]}{|\mathcal{X}|}\right)^2 = \sum_{x,y\in\mathcal{X}}\Pr[C=x,D=y]^2 - \sum_{y\in\mathcal{X}}\frac{\Pr[D=y]^2}{|\mathcal{X}|}\ ,$$

and hence the result follows by Equation 2.1. □

Now, let (A',B') be independent and identically distributed as (A,B). By Claim 2.3, we have that for all $\alpha\in\mathbb{F}$,

$$\Pr[A+\alpha\cdot B = A'+\alpha\cdot B'] \leqslant \frac{1}{p} + 4\varepsilon^2\ .$$

Summing over all $\alpha\in\mathbb{F}$, we have that

$$\begin{aligned}
1+4p\varepsilon^2 &\geqslant \sum_{\alpha\in\mathbb{F}}\Pr[A+\alpha\cdot B = A'+\alpha\cdot B'] \\
&= \sum_{\alpha\in\mathbb{F}}\Pr[A+\alpha\cdot B = A'+\alpha\cdot B'|B=B']\cdot\Pr[B=B'] \\
&\quad + \sum_{\alpha\in\mathbb{F}}\Pr[A+\alpha\cdot B = A'+\alpha\cdot B'|B\neq B']\cdot\Pr[B\neq B'] \\
&= |\mathbb{F}|\cdot\Pr[A=A'|B=B']\cdot\Pr[B=B'] + \Pr[B\neq B'] \\
&= p\Pr[A=A',B=B'] + \Pr[B\neq B'] \\
&\geqslant 1+4\Delta((A,B);(U_\mathbb{F},B))^2\ ,
\end{aligned}$$

using Claim 2.3. □

2.1 Known Primitives

EXTRACTORS. An extractor [NZ96] can be used to extract uniform randomness out of a weakly-random value which is only assumed to have sufficient min-entropy. Our definition follows that of [DORS08], which is defined in terms of conditional min-entropy.

Definition 2.4 (Extractors)**.** An efficient function $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is an (average-case, strong) (k,ε)-*extractor*, if for all X, Z such that X is distributed over $\{0,1\}^n$ and $\mathbf{H}_\infty(X|Z) \geq k$, we get

$$\Delta(\ (Z, Y, \mathsf{Ext}(X;Y))\ ,\ (Z, Y, U_m)\ ) \leqslant \varepsilon$$

where $Y \equiv U_d$ denotes the coins of Ext (called the *seed*). The value $L = k - m$ is called the *entropy loss* of Ext, and the value d is called the *seed length* of Ext.

It is well known [RTS00] that the optimal entropy loss of an extractor is $2\log(1/\varepsilon) - O(1)$, which is achieved by the famous Leftover Hash Lemma [HILL99] with seed length $d = n$. To reduce the seed length to $d = O((\log(1/\varepsilon) + \log k)\log n)$, we can also use more sophisticated extractor constructions, such as those in [GUV09, DKSS09]. Altentatively, we can extract $m = (1-\delta)k$ bits using asymptotically optimal seed length $d = O(\log(1/\varepsilon) + \log n)$ [GUV09]. We will use the following.

Theorem 2.5 ([HILL99])**.** *There exists an efficient (k,ε)-extractor $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^{k-2\log(1/\varepsilon)-O(1)}$.*

AFFINE-EVASIVE SETS. Let $\mathbb{F}$ be a finite field. An affine-evasive set $S \subseteq \mathbb{F}$ is a large set which has small intersection with affine shifts of itself. For $a, b \in \mathbb{F}^2$ define $aS + b = \{as + b | s \in S\}$.

Definition 2.6. A non-empty set $S \subseteq \mathbb{F}$ is said to be *a (γ,ν)-affine-evasive subset of* $\mathbb{F}$ if $|S| = \gamma|\mathbb{F}|$, and for any $(a,b) \in \mathbb{F}^2 \setminus \{(1,0)\}$, we have

$$|S \cap (aS + b)| \leq \nu|S|\ .$$

The notion of affine evasive sets in this context initiated in [ADL13], where a somewhat lossy construction was given. This was improved in a recent work of one of the authors of this paper [Agg15].

Theorem 2.7 ([Agg15])**.** *Let $\mathbb{F}_p$ be a prime field. There exists an efficiently samplable (γ,ν)-affine-evasive subset $S \subset \mathbb{F}_p$ with $\gamma = p^{-3/4+o(1)}$ and $\nu = p^{-1/4+o(1)}$.*

3 Inner product is a candidate affine-malleable extractor

Affine-malleable Extractors. We first define the family of desired distributions.

Definition 3.1 (Affine distribution)**.** Let $\mathbb{F}$ be a finite field. A distribution D taking values in $\mathbb{F} \times \mathbb{F}$ is said to be an *affine distribution* if it is a convex combination of distributions of the form $(U, aU + b)$ where $a, b \in \mathbb{F}$ and U is the uniform distribution over $\mathbb{F}$.

An affine malleable extractor is a seeded extractor, such that two correlated applications of it on the same random source, with correlated seeds, are close to being an affine distribution.

Definition 3.2 (Affine-Malleable Extractors)**.** Let $\mathbb{F}$ be a finite field. An efficient function $\mathsf{AmExt} : \{0,1\}^n \times \{0,1\}^d \to \mathbb{F}$ is an (average-case) (k,ε)-*affine-malleable extractor*, if the following holds. Let $X \in \{0,1\}^n, Z \in \mathcal{Z}$ be jointly distributed random variables such that $\mathbf{H}_\infty(X|Z) \geq k$. Let $Y \in \{0,1\}^d$ be uniform and independent. Let $\mathcal{F} : \{0,1\}^d \times \mathcal{Z} \mapsto \{0,1\}^d$ be any adversarial function. Then, there exist a random variable $D = D(Y,Z)$ taking values in $\mathbb{F} \times \mathbb{F}$, such that the distribution $D|Y = y, Z = z$ is affine for all $y \in \{0,1\}^d, z \in \mathcal{Z}$, such that we have

$$\Delta\left((\mathsf{AmExt}(X;Y), \mathsf{AmExt}(X;\mathcal{F}(Y,Z)))\ ;\ D \mid Z, Y\right) \leqslant \varepsilon.$$

Let $\mathbb{F} = \mathbb{F}_p$ for the remainder of this section. We show that if for any large enough $\mathcal{A} \subseteq \mathbb{F}^n$, there exists $\mathcal{A}' \subseteq \mathcal{A}$ such that $\mathrm{Spec}_\varepsilon(\mathcal{A}')$ has small doubling, then the inner product over $\mathbb{F}$ is a good affine-malleable extractor. In this section, fix $\omega = e^{\frac{2\pi i}{p}}$.

Let $\phi_f(X,Y)$ be the joint distribution $\langle X,Y\rangle, \langle X, f(Y)\rangle$, and let $\mathcal{D}$ be the set of affine distributions over $\mathbb{F} \times \mathbb{F}$. We will use the following result from [ADL14]. This was shown in [ADL14] for a different definition of ϕ_f, but the proof is the same.

Lemma 3.3. *Let $\mathcal{P} \subseteq \mathbb{F}^n \times \mathbb{F}^n$. Let $\mathcal{P}_1, \ldots, \mathcal{P}_k$ be a partition of $\mathcal{P}$. Assume that for all $1 \leqslant i \leqslant k$,*

$$\Delta\left(\phi_f(X,Y)|_{(X,Y)\in\mathcal{P}_i}\ ;\ \mathcal{D} \mid Z, Y\right) \leqslant \varepsilon_i.$$

Then

$$\Delta\left(\phi(X,Y)|_{(X,Y)\in\mathcal{P}}\ ;\ \mathcal{D} \mid Z, Y\right) \leqslant \sum \varepsilon_i \frac{|\mathcal{P}_i|}{|\mathcal{P}|}.$$

We will need the following result that shows a connection between the statistical distance of the inner product from uniform, and the spectrum of a set $\mathcal{A}$.

Lemma 3.4. *Let $\mathcal{A} \subseteq \mathbb{F}^n$, and let X be uniform over $\mathcal{A}$. For all $y \in \mathbb{F}^n$, if for some $a \in \mathbb{F} \setminus \{0\}$, $a \cdot y \in \mathit{Spec}_\varepsilon(\mathcal{A})$ then $\Delta(\langle y, X\rangle, U_\mathbb{F}) \geqslant \frac{\varepsilon}{2\sqrt{p}}$. Also if $\Delta(\langle y, X\rangle, U_\mathbb{F}) \geqslant \frac{\varepsilon\sqrt{p}}{2}$, then there exists an $a \in \mathbb{F} \setminus \{0\}$ such that $a \cdot y \in \mathit{Spec}_\varepsilon(\mathcal{A})$.*

Proof. We have by the Parseval identity that

$$\sum_{c\in\mathbb{F}} \left(\Pr[\langle y, X\rangle = c] - \frac{1}{p}\right)^2 = \frac{1}{p} \sum_{a\in\mathbb{F}\setminus\{0\}} |\mathbb{E}(\omega^{a\cdot\langle y,X\rangle})|^2 .$$

Using Equation 2.1, we get that

$$\frac{1}{p} \cdot \Delta(\langle y, X\rangle, U_\mathbb{F})^2 \leqslant \frac{1}{4} \cdot \sum_{c\in\mathbb{F}} \left(\Pr[\langle y, X\rangle = c] - \frac{1}{p}\right)^2 \leqslant \Delta(\langle y, X\rangle, U_\mathbb{F})^2 .$$

Also,

$$\frac{1}{p} \max_{a\in\mathbb{F}\setminus\{0\}} \left|\mathbb{E}[\omega^{a\cdot\langle y,X\rangle}]\right|^2 \leqslant \frac{1}{p} \sum_{a\in\mathbb{F}\setminus\{0\}} \left|\mathbb{E}[\omega^{a\cdot\langle y,X\rangle}]\right|^2 \leqslant \max_{a\in\mathbb{F}\setminus\{0\}} \left|\mathbb{E}[\omega^{a\cdot\langle y,X\rangle}]\right|^2 .$$

Combining we get that

$$\sqrt{\frac{1}{4p}} \cdot \max_{a\in\mathbb{F}\setminus\{0\}} \left|\mathbb{E}[\omega^{a\cdot\langle y,X\rangle}]\right| \leqslant \Delta(\langle y, X\rangle, U_\mathbb{F}) \leqslant \sqrt{\frac{p}{4}} \cdot \max_{a\in\mathbb{F}\setminus\{0\}} \left|\mathbb{E}[\omega^{a\cdot\langle y,X\rangle}]\right| .$$

□

For $\mathcal{A} \subseteq \mathbb{F}^n$, X uniform in A, define $\mathrm{Bad}_\varepsilon(\mathcal{A}) \subset \mathbb{F}^n$ to be the set of "bad" choices for the seed as follows:

$$\mathrm{Bad}_\varepsilon(\mathcal{A}) = \{y \in \mathbb{F}^n : \exists w \in \mathbb{F}^n \text{ such that } (\langle X, y\rangle, \langle X, w\rangle) \text{ is not } \varepsilon\text{-close to an affine distribution}\}$$

We will show that if A has controlled spectrum, then $\mathrm{Bad}_\varepsilon(\mathcal{A})$ is small. Recall that for $S \subset \mathbb{F}^n$ we defined $S + S = \{s_1 + s_2 : s_1, s_2 \in S\}$ and $\mathbb{F} \cdot S = \{c \cdot s : c \in \mathbb{F}, s \in S\}$.

Lemma 3.5. *Let $\mathcal{A} \subseteq \mathbb{F}^n$. We have*

$$\mathit{Bad}_\varepsilon(\mathcal{A}) \subseteq \mathbb{F} \cdot \mathit{Spec}_{\varepsilon'}(A) + \mathbb{F} \cdot \mathit{Spec}_{\varepsilon'}(A)$$

where $\varepsilon' = 2\varepsilon/p$.

Proof. Let X be uniform in $\mathcal{A}$. Let $y, w \in \mathbb{F}^n$ be such that $(\langle X, y\rangle, \langle X, w\rangle)$ is not ε-close to any affine distribution. In particular this implies that for any $c \in \mathbb{F}$, $(\langle X, y\rangle, \langle X, w + cy\rangle)$ is not ε-close to the convex hull of distributions $\{(U, b) : b \in \mathbb{F}\}$. By Lemma 2.2, this implies that for any $c \in \mathbb{F}$, there exist $\alpha_c \in \mathbb{F}$ such that $\langle X, y + \alpha_c(w + cy)\rangle$ is not $(\varepsilon/\sqrt{p})$-close to uniform. Thus by Lemma 3.4, there exist $\beta, \in \mathbb{F} \setminus \{0\}$ such that $y + \alpha w = \beta\gamma_1$ for some $\gamma_1 \in \mathrm{Spec}_{\varepsilon'}(\mathcal{A})$. If $\alpha = 0$, then $y = \beta\gamma_1$, and the claim is satisfied. So, we assume $\alpha \neq 0$. Then, again by Lemma 3.4, $y + \alpha'(w + \frac{1}{\alpha}y) = \beta'\gamma_2$, for some $\gamma_2 \in \mathrm{Spec}_{\varepsilon'}(\mathcal{A})$. Solving for y gives that $y = \beta'\gamma_2 - \frac{\beta \cdot \alpha'}{\alpha} \cdot \gamma_1$. □

Now, we proceed to the main result of this section.

Theorem 3.6. *Fix $\varepsilon, \eta > 0$ and $1 \leqslant k \leqslant n \log p$. Assume that for any $\mathcal{A} \subseteq \mathbb{F}^n$ of size $|\mathcal{A}| \geqslant 2^k$ we have that there exists $\mathcal{B} \subseteq \mathcal{A}$ such that*

$$|\mathbb{F} \cdot \mathit{Spec}_\varepsilon(\mathcal{B}) + \mathbb{F} \cdot \mathit{Spec}_\varepsilon(\mathcal{B})| \leqslant \eta|\mathbb{F}|^n.$$

Then for any $\delta > 0$, $\langle \cdot, \cdot \rangle : \mathbb{F}^n \times \mathbb{F}^n$ is a $(k + 2\log\frac{1}{\delta}, \varepsilon + \eta + 2\delta)$-affine-malleable extractor.

Proof. Let $X \in \mathbb{F}^n, Z \in \mathcal{Z}$ be jointly distributed random variables such that $\mathbf{H}_\infty(X|Z) \geq k + 2\log 1/\delta$. Let $Y \in \mathbb{F}^n$ be uniform and independent. First, by lemma 2.1 we have that

$$\Pr_{z \leftarrow Z}[\mathbf{H}_\infty(X|Z = z) \geqslant k + \log 1/\delta] \geqslant 1 - \delta.$$

Fix z so that $\mathbf{H}_\infty(X|Z = z) \geqslant k + \log 1/\delta$. It suffices to consider the case where the random variable $X_z = (X|Z = z)$ is distributed uniformly over a set $\mathcal{A}_0 \subseteq \mathbb{F}^n$ of size $|\mathcal{A}_0| \geqslant 2^k/\delta$. Define $\mathcal{B}_0, \mathcal{A}_1, \mathcal{B}_1, \mathcal{A}_2, \ldots, \mathcal{B}_{t-1}, \mathcal{A}_t$ iteratively as follows. For all i, let $\mathcal{B}_i$ be a subset of $\mathcal{A}_i$ such that

$$|\mathbb{F} \cdot \mathrm{Spec}_\varepsilon(\mathcal{B}_i) + \mathbb{F} \cdot \mathrm{Spec}_\varepsilon(\mathcal{B}_i)| \leqslant \eta|\mathbb{F}|^n.$$

and let $\mathcal{A}_{i+1} = \mathcal{A}_i \setminus \mathcal{B}_i$. We terminate for a t such that no such $\mathcal{B}_t$ exists. By the assumption, $|\mathcal{A}_t| < 2^k \leqslant \delta|\mathcal{A}_0|$. Thus,

$$\mathcal{A}_0 = \mathcal{B}_0 \cup \cdots \cup \mathcal{B}_{t-1} \cup \mathcal{A}_t\ .$$

Let $\mathcal{C}'_j = \mathrm{Bad}_\varepsilon(\mathcal{B}_j)$ and $\mathcal{C}_j = \mathbb{F}^n \setminus \mathcal{C}'_j$. By Lemma 3.5, $|\mathcal{C}'_j| \leqslant \eta|\mathbb{F}|^n$. Consider the following partition of $\mathcal{A}_0 \times \mathbb{F}^n$:

$$\mathcal{A}_0 \times \mathbb{F}^n = \bigcup_{j=0}^{t-1} (\mathcal{B}_j \times \mathcal{C}_j) \cup \bigcup_{j=0}^{t-1} (\mathcal{B}_j \times \mathcal{C}'_j) \cup \mathcal{A}_t \times \mathbb{F}^n.$$

Let $X_{z,j} = X_z|_{X_z \in \mathcal{B}_j}$, and note that $X_{z,j}$ is uniform in $\mathcal{B}_j$. For any $y \in \mathcal{C}_j$ we know that $\langle X_{z,j}, y\rangle, \langle X_{z,j}, w\rangle$ is ε-close to affine for any $w \in \mathbb{F}^n$, and in particular for $w = \mathcal{F}(y, z)$. Hence, for any $y \in \mathcal{C}_j$ we have that $\phi_f(X_{z,j}, y)$ is ε-close to $\mathcal{D}$. To conclude, note that $\Pr[(X_z, Y) \in \cup(\mathcal{B}_j \times \mathcal{C}'_j)] \leqslant \eta$ and $\Pr[(X_z, Y) \in \mathcal{A}_t \times \mathbb{F}^n] \leqslant \delta$. By Lemma 3.3, we have that

$$\mathbb{E}_{y \in \mathbb{F}^n} \Delta((\langle X_z, y\rangle, \langle X_z, \mathcal{F}(y, z)\rangle) \mid \mathcal{D}) \leqslant \varepsilon + \eta + \delta.$$

The theorem follows by averaging over $z \leftarrow Z$. □

4 Privacy amplification protocol

We define a privacy amplification protocol (P_A, P_B), executed by two parties Alice and Bob sharing a secret $X \in \{0,1\}^N$, in the presence of an active, computationally unbounded adversary Eve, who might have some partial information E about X satisfying $\mathbf{H}_\infty(X|E) \geqslant k$. Informally, this means that whenever a party (Alice or Bob) does not reject, the key R output by this party is random and statistically independent of Eve's view. Moreover, if both parties do not reject, they must output the same keys $R_A = R_B$ with overwhelming probability.

More formally, we assume that Eve is in full control of the communication channel between Alice and Bob, and can arbitrarily insert, delete, reorder or modify messages sent by Alice and Bob to each other. In particular, Eve's strategy P_E actually defines two correlated executions (P_A, P_E) and (P_E, P_B) between Alice and Eve, and Eve and Bob, called "left execution" and "right execution", respectively. We stress that the message scheduling for both of these executions is completely under Eve's control, and Eve might attempt to execute a run with one party for several rounds before resuming the execution with another party. However, Alice and Bob are assumed to have fresh, private and independent random tapes Y and W, respectively, which are not known to Eve (who, by virtue of being unbounded, can be assumed to be deterministic). At the end of the left execution $(P_A(X,Y), P_E(E))$, Alice outputs a key $R_A \in \{0,1\}^m \cup \{\perp\}$, where $\perp$ is a special symbol indicating rejection. Similarly, Bob outputs a key $R_B \in \{0,1\}^m \cup \{\perp\}$ at the end of the right execution $(P_E(E), P_B(X,W))$. We let E' denote the final view of Eve, which includes E and the communication transcripts of both executions $(P_A(X,Y), P_E(E))$ and $(P_E(E), P_B(X,W)$. We can now define the security of (P_A, P_B). Our definition is based on [DLWZ14, ADJ+14].

Definition 4.1. An interactive protocol (P_A, P_B), executed by Alice and Bob on a communication channel fully controlled by an active adversary Eve, is a (k,ε)-*private*, (k,m,ϵ)-*privacy amplification protocol* if it satisfies the following properties whenever $\mathbf{H}_\infty(X|E) \geq k$:

1. Correctness. If Eve is passive, then $\Pr[R_A = R_B \wedge \ R_A \neq \perp \wedge R_B \neq \perp] = 1$.

2. Robustness. We start by defining the notion of *pre-application* robustness, which states that even if Eve is active, $\Pr[R_A \neq R_B \wedge \ R_A \neq \perp \wedge R_B \neq \perp] \leqslant \epsilon$.

   The stronger notion of *post-application* robustness is defined similarly, except Eve is additionally given the key R_A the moment she completed the left execution (P_A, P_E), and the key R_B the moment she completed the right execution (P_E, P_B). For example, if Eve completed the left execution before the right execution, she may try to use R_A to force Bob to output a different key $R_B \notin \{R_A, \perp\}$, and vice versa.

3. Extraction. Given a string $r \in \{0,1\}^m \cup \{\perp\}$, let $\mathsf{purify}(r)$ be $\perp$ if $r = \perp$, and otherwise replace $r \neq \perp$ by a fresh m-bit random string U_m: $\mathsf{purify}(r) \leftarrow U_m$. Letting E' denote Eve's view of the protocol, we require that

   $$\Delta((R_A, E'), (\mathsf{purify}(R_A), E')) \leq \epsilon \quad \text{and} \quad \Delta((R_B, E'), (\mathsf{purify}(R_B), E')) \leq \epsilon$$

   Namely, whenever a party does not reject, its key looks like a fresh random string to Eve.

4. Source Privacy. To define this property, we let $\mathsf{FullOutput}(X, E)$ denote the tuple (E', R_A, R_B), where Alice and Bob share a secret X and output keys R_A and R_B, respectively, and Eve starts with initial side information E and ends with final view E' at the end of the protocol.

   We require that for any two distributions (X_0, E) and (X_1, E), where $\mathbf{H}_\infty(X_0|E) \geq k$ and $\mathbf{H}_\infty(X_1|E) \geqslant k$, we have

   $$\Delta(\mathsf{FullOutput}(X_0, E), \mathsf{FullOutput}(X_1, E)) \leq \varepsilon$$

The quantity $k - m$ is called the *entropy loss* and the quantity $\log(1/\epsilon)$ is called the *security parameter* of the protocol.

Our Protocol. We obtain the following result.

Theorem 4.2. *Let $\mathbb{F}$ be a finite field. Assume that there exists*

1. *An efficiently samplable (γ, ν)-affine-evasive subset $\mathcal{M}$ of $\mathbb{F}$.*
2. *An efficient (k_1, ε')-extractor $\mathsf{Ext} : \{0,1\}^N \times \mathcal{M} \to \{0,1\}^m$.*
3. *An efficient (k_2, ε')-affine-malleable extractor $\mathsf{AmExt} : \{0,1\}^N \times \{0,1\}^d \mapsto \mathbb{F}$.*

Then there exists an explicit polynomial-time, two-round (k, ε)-private, (k, m, ε)-secure privacy amplification protocol, where

$$k \geqslant \max(k_1 + 2\log|\mathbb{F}|, k_2) \ ,$$
$$\varepsilon = 3\varepsilon' + \max(\gamma, \nu) \ .$$

Proof. Using the given building blocks, the protocol is depicted as Protocol 1.

Alice: X	Eve: E	Bob: X
Sample uniformly random $Y \in \{0,1\}^d$	$Y \longrightarrow Y'$	
		Sample uniformly random $W' \in \mathcal{M}$ $K' = \mathsf{AmExt}(X; Y')$ $T' = W' + K'$ Set final $R_B = \mathsf{Ext}(X; W')$
	$T \longleftarrow T'$	
$K = \mathsf{AmExt}(X; Y)$ $W = T - K$ **If** $W \notin \mathcal{M}$, set $R_A = \perp$ **Else** set final $R_A = \mathsf{Ext}(X; W)$		

Protocol 1: Our 2-round Privacy Amplification Protocol

Post-application Robustness. To have any chance of breaking robustness, Eve must choose T such that $W \neq W'$ because if $W = W'$ then either $R_A = R_B$ or $R_A = \perp$. Thus, we assume $W \neq W'$.

We have that $K = \mathsf{AmExt}(X, Y)$ and $K' = \mathsf{AmExt}(X, Y')$ where $Y' = Y'(Y, Z)$. By the definition AmExt,

$$(K, K'), Y, Z \approx_{\varepsilon'} D, Y, Z$$

where $D = D(Y, Z)$ has the property that $D|Y = y, Z = z$ is an affine distribution for all y, z. Next, W' is independent of D, Y, Z and $H_\infty(X|D, Y, Z) \geqslant k - 2\log|\mathbb{F}| \geqslant k_1$ since $H_\infty(X|Z) \geqslant k$, Y is independent of X, Z and D is supported on $\mathbb{F} \times \mathbb{F}$. Hence

$$(K, K'), Y, Z, W', R_B \approx_{2\varepsilon'} D, Y, Z, W', U_m$$

where U_m is the uniform distribution over $\{0,1\}^m$.

From now on, fix $Y = y, Z = z$. We have that $D|Y = y, Z = z$ is a convex combination of $(U, aU + b)$ for $a, b \in \mathbb{F}$, and $U = U_{\mathbb{F}}$. Furthermore, we have $T = f(T')$ where $f : \mathbb{F}_p \to \mathbb{F}_p$ is some

function (which may depend on y, z). Thus, it is sufficient to show that for all $a, b \in \mathbb{F}$, and all functions $f : \mathbb{F}_p \to \mathbb{F}_p$,

$$\Pr\left(f(W' + aU + b) - U \in \mathcal{M} \quad \wedge \quad f(W' + aU + b) - U \neq W'\right) \leqslant \max(\gamma, \nu) \,, \tag{4.1}$$

where U is uniform in $\mathbb{F}$. This is because Equation 4.1 implies that

$$\Pr(R_A \neq \perp \ \wedge \ W' \neq W) \leqslant \max(\gamma, \nu) + 2\varepsilon' \,,$$

which implies post-application robustness. To prove Equation 4.1, we consider two cases.

CASE 1: $a = 0$. In this case, $f(W' + aU + b) - U$ is uniform in $\mathbb{F}$. From Definition 2.6

$$\Pr\left(f(W' + aU + b) - U \in \mathcal{M}\right) = \gamma \,.$$

CASE 2: $a \neq 0$. Let $C = W' + aU + b$. Note that C is independent of W' and uniform in $\mathbb{F}$. Then

$$f(W' + aU + b) - U = f(C) - \frac{C - W' - b}{a} \,.$$

Fix $C = c$, and let $a' = \frac{1}{a}$, $b' = f(c) - \frac{c-b}{a}$. Thus it is sufficient to show the following:

$$\Pr\left(a'W' + b' \in \mathcal{M} \quad \wedge \quad a'W' + b' \neq W'\right) \leqslant \nu \,.$$

Note that if $(a', b') = (1, 0)$, then $a'W' + b' = W'$. So, without loss of generality we assume that $(a', b') \neq (1, 0)$. Then, Definition 2.6 implies the result.

Extraction. Note that Y is independent of X. Thus, using Lemma 2.1,

$$\mathbf{H}_\infty(X | Z, Y, K') \geqslant k - \log|\mathbb{F}| \,.$$

Thus, using the definition of Ext, we have that R_B is ε'-close to uniform given the transcript of Eve, $E' = Z, Y, W' + K'$. Also, with probability $1 - 2\varepsilon' - \max(\gamma, \nu)$, $R_A = R_B$ or $R_A = \perp$. Thus, conditioned on the event that $R_A = R_B$ or $R_A = \perp$, R_A is ε' close to $\mathsf{purify}(R_A)$ given E'. Thus,

$$\Delta((R_A, E'), (\mathsf{purify}(R_A), E')) \leq 3\varepsilon' + \max(\gamma, \nu) \,.$$

Source Privacy. The source privacy follows easily from the following observations: R_B is ε'-close to uniform given the final transcript of Eve $E' = Z, Y, T'$. Also, Y, T' are uniform and independent of X, Z. Finally note that by the robustness property, $R_B = R_A$ if $T' = T$ and $R_A = \perp$, otherwise except with probability at most $2\varepsilon' + \max(\gamma, \nu)$. Thus, the knowledge of R_B doesn't provide any additional information than what can be concluded from (R_A, E'), except with probability $2\varepsilon' + \max(\gamma, \nu)$. Thus

$$\Delta(\mathsf{FullOutput}(X_0, E), \mathsf{FullOutput}(X_1, E)) \leq 2\varepsilon' + \max(\gamma, \nu) + \varepsilon' = 3\varepsilon' + \max(\gamma, \nu) \,.$$

□

5 Spectrum doubling

5.1 Weak Spectrum Doubling from PFR Conjecture

In this section, we prove Theorem 1.2. We will need the following approximate duality conjecture that was proved in [BDL14] assuming the polynomial Freiman-Ruzsa conjecture.

Theorem 5.1. *[BDL14, Theorem 6.3] Let V be a subspace of $\mathbb{F}^n$ of dimension r. Let $A \subseteq V$, and let $\varepsilon \geqslant p^{-\sqrt{r}}$. Then, assuming the polynomial Freiman Ruzsa conjecture, there exist a universal constant c, and $A' \subseteq A$ and $B' \subseteq V$ such that $|A'| \geqslant p^{-cr \log p/\log r}|A|$, $|B'| \geqslant p^{-cr \log p/\log r}|Spec_\varepsilon(A) \cap V|$, and $\langle a, b\rangle = 0$ for all $a \in A'$, and $b \in B'$.*

Note that in Theorem 6.3 from [BDL14], the authors only considered the case when p is small, and hence did not explicitly state the dependence of $|A'|$ and $|B'|$ on p.

Proof of Theorem 1.2. Let $\delta = \sqrt{\varepsilon}/p$, $\alpha = \sqrt{c \log p/\log n}$, where c is the constant from Theorem 5.1. We can assume that $\alpha < 1/100$, since otherwise, the result holds trivially by choosing C sufficiently larger than c. We will now show the result for $C = c$. If $|\mathrm{Spec}_\varepsilon(A)| \leqslant \delta \cdot p^{n/2}$, then clearly Equation 1.1 holds for $B = A$. So, without loss of generality, we assume that $|\mathrm{Spec}_\varepsilon(A)| > \delta \cdot p^{n/2}$. By Theorem 5.1, there exist $A_1 \subseteq A$, $B_1 \subseteq \mathbb{F}^n$ such that $|A_1| \geqslant \frac{1}{\varepsilon} \cdot p^{(\alpha-\alpha^2)n}$, and $|B_1| \geqslant \delta \cdot p^{n(1/2-\alpha^2)} \geqslant \delta \cdot p^{n/3}$, and every element of A_1 is orthogonal to every element of B_1.

Thus, $A_1 \subset V_1$ for a subspace V_1 of size at most $\frac{1}{\delta} \cdot p^{2n/3} < \frac{1}{\delta^2} \cdot p^{2n/3}$. Now, repeating the argument, we have that if $|\mathrm{Spec}_\varepsilon(A_1) \cap V_1| \leqslant \delta \cdot \sqrt{|V_1|}$, then Equation 1.1 holds for $B = A_1$. So, without loss of generality, we assume that $|\mathrm{Spec}_\varepsilon(A_1) \cap V_1| > \delta \cdot \sqrt{|V_1|}$. By Theorem 5.1, there exist $A_2 \subseteq A_1$, $B_2 \subseteq V_1$ such that $|A_2| \geqslant \frac{1}{\varepsilon} \cdot p^{(\alpha-2\alpha^2)n}$, and $|B_2| \geqslant \delta \cdot \sqrt{|V_1|} \cdot p^{-\alpha^2 n}$, and every element of A_1 is orthogonal to every element of B_1.

Thus, $A_2 \subset V_2$ for a subspace V_2 of size at most

$$\frac{1}{\delta} \cdot \sqrt{|V_1|} \cdot p^{\alpha^2 n} \leqslant \frac{1}{\delta^2} \cdot p^{4n/9} ,$$

as long as $\frac{2}{3} \cdot \frac{1}{2} + \alpha^2 < (\frac{2}{3})^2$.

Continuing in this manner, we get $A_2 \supseteq \ldots \supseteq A_i$ such that either one of them already satisfies Equation 1.1, or $|A_i| \geqslant \frac{1}{\varepsilon} \cdot p^{n(\alpha-i\alpha^2)}$, and it is contained in a subspace V_i of $\mathbb{F}^n$ of size at most

$$\frac{1}{\delta^2} \cdot p^{(2/3)^i n} ,$$

as long as $(\frac{2}{3})^{i-1} \cdot \frac{1}{2} + \alpha^2 < (\frac{2}{3})^i$. This condition holds as long as $i < i_{\max} = \lfloor 3/2 \log_{2/3} \alpha \rfloor$.

Now, we show that the above process terminates before $i < i_{\max}$, i.e., we find some $j \in [i_{\max}]$ such that $B = A_j$ satisfies Equation 1.1. Note that the above algorithm for generating A_i, V_i terminates before the smallest j such that

$$\frac{1}{\varepsilon} \cdot p^{n(\alpha-j\alpha^2)} > \frac{1}{\delta^2} \cdot p^{(2/3)^j n} ,$$

which clearly holds for $j = i_{\max}$ since for $j = i_{\max}$, the right hand side is less than $\frac{p^2}{\varepsilon} \cdot p^{n \cdot \alpha^{3/2}}$, and the left hand side is greater than $p^{n\alpha/2} \cdot \frac{1}{\varepsilon}$. □

5.2 Applications

Affine-malleable Extractors. We first mention the results of applying spectrum doubling to obtain affine-malleable extractors. By combining Theorem 3.6 and Theorem 1.1, we get the following.

Theorem 5.2. *For any $\varepsilon > 0$, and $k \geqslant \frac{n \log p}{2} + O(\log p + \log(1/\varepsilon))$, $\langle \cdot, \cdot \rangle : \mathbb{F}^n \times \mathbb{F}^n$ is a (k, ε)-affine-malleable extractor.*

Also, for going below entropy rate 1/2, by combining Theorem 3.6 with Theorem 1.2, we get the following under the PFR conjecture.

Theorem 5.3. *Assuming the PFR conjecture, we have that for any $\varepsilon > 0$, and*

$$k \geqslant O(n \log p \sqrt{\log p / \log n} + \log(1/\varepsilon)) \ ,$$

$\langle \cdot, \cdot \rangle : \mathbb{F}^n \times \mathbb{F}^n$ is a (k, ε)-affine-malleable extractor.

We can get the following further strengthening to obtain near optimal parameters under Conjecture 1.3.

Theorem 5.4. *Assuming Conjecture 1.3, we have that for any $\varepsilon > 0$, and $k \geqslant O(\log(n \log p) \cdot \log(1/\varepsilon))$, $\langle \cdot, \cdot \rangle : \mathbb{F}^n \times \mathbb{F}^n$ is a (k, ε)-affine-malleable extractor.*

Privacy Amplification. By instantiating Theorem 4.2 (where $N = n \log p$) with the seeded extractor from Theorem 2.5, the affine evasive sets from Theorem 2.7 and the affine-malleable extractor from Theorem 5.2 with $\varepsilon = 1/p$, we obtain the following.

Theorem 5.5. *For any $N \in \mathbb{N}$, $\varepsilon > 0$, and for $k \geqslant N/2 + O(\log(1/\varepsilon))$ there exists an explicit polynomial-time, two-round (k, ε)-private, $(k, k - O(\log(1/\varepsilon)), \varepsilon)$-secure privacy amplification protocol.*

If instead, we instantiate with the affine-malleable extractor from Theorem 5.4, then we get Theorem 1.4.

References

[ADJ+14] Divesh Aggarwal, Yevgeniy Dodis, Zahra Jafargholi, Eric Miles, and Leonid Reyzin. Amplifying privacy in privacy amplification. In *Advances in Cryptology - CRYPTO 2014 - 34th Annual Cryptology Conference, Santa Barbara, CA, USA, August 17-21, 2014, Proceedings, Part II*, pages 183–198, 2014.

[ADL13] Divesh Aggarwal, Yevgeniy Dodis, and Shachar Lovett. Non-malleable codes from additive combinatorics. In *Electronic Colloquium on Computational Complexity (ECCC)*, volume 20, page 81, 2013.

[ADL14] Divesh Aggarwal, Yevgeniy Dodis, and Shachar Lovett. Non-malleable codes from additive combinatorics. In *STOC*. ACM, 2014.

[Agg15] Divesh Aggarwal. Affine-evasive sets modulo a prime. *Information Processing Letters*, 115(2):382–385, 2015.

[BBCM95] Charles H. Bennett, Gilles Brassard, Claude Crépeau, and Ueli M. Maurer. Generalized privacy amplification. *IEEE Transactions on Information Theory*, 41(6):1915–1923, 1995.

[BBR88] Charles H. Bennett, Gilles Brassard, and Jean-Marc Robert. Privacy amplification by public discussion. *SIAM Journal on Computing*, 17(2):210–229, 1988.

[BDL14] Abhishek Bhowmick, Zeev Dvir, and Shachar Lovett. New bounds for matching vector families. *SIAM J. Comput.*, 43(5):1654–1683, 2014.

[BF11] Niek J. Bouman and Serge Fehr. Secure authentication from a weak key, without leaking information. In *Advances in Cryptology - EUROCRYPT 2011 - 30th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Tallinn, Estonia, May 15-19, 2011. Proceedings*, pages 246–265, 2011.

[Bou05] Jean Bourgain. More on the sum-product phenomenon in prime fields and its applications. *International Journal of Number Theory*, 1(01):1–32, 2005.

[CG88] Benny Chor and Oded Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity. *SIAM Journal on Computing*, 17(2):230–261, 1988.

[CGL15] Eshan Chattopadhyay, Vipul Goyal, and Xin Li. Non-malleable extractors and codes, with their many tampered extensions. *arXiv preprint arXiv:1505.00107*, 2015.

[Cha02] Mei-Chu Chang. A polynomial bound in freiman's theorem. *Duke Mathematical Journal*, 113(3):399–420, 2002.

[CKOR10] Nishanth Chandran, Bhavana Kanukurthi, Rafail Ostrovsky, and Leonid Reyzin. Privacy amplification with asymptotically optimal entropy loss. In *Proceedings of the 42nd ACM Symposium on Theory of Computing, STOC 2010, Cambridge, Massachusetts, USA, 5-8 June 2010*, pages 785–794, 2010.

[CRS12] Gil Cohen, Ran Raz, and Gil Segev. Non-malleable extractors with short seeds and applications to privacy amplification. In *Computational Complexity (CCC), 2012 IEEE 27th Annual Conference on*, pages 298–308. IEEE, 2012.

[CZ15] Eshan Chattopadhyay and David Zuckerman. Explicit two-source extractors and resilient functions. *Electronic Colloquium on Computational Complexity (ECCC)*, TR15-119, 2015.

[DKK+12] Yevgeniy Dodis, Bhavana Kanukurthi, Jonathan Katz, Leonid Reyzin, and Adam Smith. Robust fuzzy extractors and authenticated key agreement from close secrets. *IEEE Transactions on Information Theory*, 58(9):6207–6222, 2012.

[DKSS09] Zeev Dvir, Swastik Kopparty, Shubhangi Saraf, and Madhu Sudan. Extensions to the method of multiplicities, with applications to kakeya sets and mergers. In *Foundations of Computer Science, 2009. FOCS'09. 50th Annual IEEE Symposium on*, pages 181–190. IEEE, 2009.

[DLWZ14] Yevgeniy Dodis, Xin Li, Trevor D. Wooley, and David Zuckerman. Privacy amplification and nonmalleable extractors via character sums. *SIAM J. Comput.*, 43(2):800–830, 2014.

[DORS08] Yevgeniy Dodis, Rafail Ostrovsky, Leonid Reyzin, and Adam Smith. Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. *SIAM Journal on Computing*, 38(1):97–139, 2008.

[DP07] Yevgeniy Dodis and Prashant Puniya. Feistel networks made public, and applications. In Moni Naor, editor, *Advances in Cryptology - EUROCRYPT 2007*, volume 4515 of *Lecture Notes in Computer Science*, pages 534–554. Springer-Verlag, 2007.

[DS02] Yevgeniy Dodis and Joel Spencer. On the (non-)universality of the one-time pad. In *43rd Annual Symposium on Foundations of Computer Science*, pages 376–385. IEEE, 2002.

[DS05] Yevgeniy Dodis and Adam Smith. Entropic security and the encryption of high entropy messages. In *Theory of Cryptography, Second Theory of Cryptography Conference, TCC 2005, Cambridge, MA, USA, February 10-12, 2005, Proceedings*, pages 556–577, 2005.

[DW09] Yevgeniy Dodis and Daniel Wichs. Non-malleable extractors and symmetric key cryptography from weak secrets. In Michael Mitzenmacher, editor, *Proceedings of the 41st Annual ACM Symposium on Theory of Computing*, pages 601–610, Bethesda, MD, USA, 2009. ACM.

[DY13] Yevgeniy Dodis and Yu Yu. Overcoming weak expectations. In *TCC*, pages 1–22, 2013.

[GUV09] Venkatesan Guruswami, Christopher Umans, and Salil P. Vadhan. Unbalanced expanders and randomness extractors from parvaresh–vardy codes. *J. ACM*, 56(4), 2009.

[HILL99] J. Håstad, R. Impagliazzo, L.A. Levin, and M. Luby. Construction of pseudorandom generator from any one-way function. *SIAM Journal on Computing*, 28(4):1364–1396, 1999.

[HL15] Kaave Hosseini and Shachar Lovett. On the structure of spectrum of small sets. arXiv, math/1504.01059, 2015. http://arxiv.org/abs/1504.01059.

[Li12a] Xin Li. Design extractors, non-malleable condensers and privacy amplification. In *Proceedings of the 44th Symposium on Theory of Computing Conference, STOC 2012, New York, NY, USA, May 19 - 22, 2012*, pages 837–854, 2012.

[Li12b] Xin Li. Non-malleable condensers for arbitrary min-entropy, and almost optimal protocols for privacy amplification. *CoRR*, abs/1211.0651, 2012.

[Li12c] Xin Li. Non-malleable extractors, two-source extractors and privacy amplification. In *FOCS*, pages 688–697, 2012.

[Li15] Xin Li. Non-malleable condensers for arbitrary min-entropy, and almost optimal protocols for privacy amplification. In *Theory of Cryptography - 12th Theory of Cryptography Conference, TCC 2015, Warsaw, Poland, March 23-25, 2015, Proceedings, Part I*, pages 502–531, 2015.

[Mau92] Ueli Maurer. Conditionally-perfect secrecy and a provably-secure randomized cipher. *Journal of Cryptology*, 5(1):53–66, 1992.

[MW97] Ueli Maurer and Stefan Wolf. Privacy amplification secure against active adversaries. In Burton S. Kaliski, Jr., editor, *Advances in Cryptology—CRYPTO '97*, volume 1294 of *LNCS*, pages 307–321. Springer-Verlag, 1997.

[NZ96] Noam Nisan and David Zuckerman. Randomness is linear in space. *Journal of Computer and System Sciences*, 52(1):43–53, 1996.

[RTS00] Jaikumar Radhakrishnan and Amnon Ta-Shma. Bounds for dispersers, extractors, and depth-two superconcentrators. *SIAM Journal on Computing*, 13(1):2–24, 2000.

[RW03] Renato Renner and Stefan Wolf. Unconditional authenticity and privacy from an arbitrarily weak secret. In Dan Boneh, editor, *Advances in Cryptology—CRYPTO 2003*, volume 2729 of *LNCS*, pages 78–95. Springer-Verlag, 2003.

[Sha02] Ronen Shaltiel. Recent developments in explicit constructions of extractors. *Bulletin of the EATCS*, 77:67–95, 2002.

[Shk08] Il'ya Dmitrievich Shkredov. On sets of large trigonometric sums. *Izvestiya: Mathematics*, 72(1):149, 2008.