Non-Malleable Extractors, Two-Source Extractors and Privacy Amplification

Xin Li*
Department of Computer Science
University of Washington
Seattle, WA 98905, U.S.A.
lixints@cs.washington.edu

April 9, 2012

Abstract

Dodis and Wichs [DW09] introduced the notion of a non-malleable extractor to study the problem of privacy amplification with an active adversary. A non-malleable extractor is a much stronger version of a strong extractor. Given a weakly-random string x and a uniformly random seed y as the inputs, the non-malleable extractor nmExt has the property that $\mathsf{nmExt}(x, y)$ appears uniform even given y as well as $\mathsf{nmExt}(x, \mathcal{A}(y))$, for an arbitrary function $\mathcal{A}$ with $\mathcal{A}(y) \neq y$. Dodis and Wichs showed that such an object can be used to give optimal privacy amplification protocols with an active adversary.

Previously, there are only two known constructions of non-malleable extractors [DLWZ11, CRS12]. Both constructions only work for (n, k)-sources with $k > n/2$. Interestingly, both constructions are also two-source extractors.

In this paper, we present a strong connection between non-malleable extractors and two-source extractors. The first part of the connection shows that non-malleable extractors can be used to construct two-source extractors. If the non-malleable extractor works for small min-entropy and has a short seed length with respect to the error, then the resulted two-source extractor beats the best known construction of two-source extractors. This partially explains why previous constructions of non-malleable extractors only work for sources with entropy rate $> 1/2$, and why explicit non-malleable extractors for small min-entropy may be hard to get.

The second part of the connection shows that certain two-source extractors can be used to construct non-malleable extractors. Using this connection, we obtain the first construction of non-malleable extractors for $k < n/2$. Specifically, we give an unconditional construction for min-entropy $k = (1/2 - \delta)n$ for some constant $\delta > 0$, and a conditional (semi-explicit) construction that can potentially achieve $k = \alpha n$ for any constant $\alpha > 0$.

We also generalize non-malleable extractors to the case where there are more than one adversarial seeds, and show a similar connection between the generalized non-malleable extractors and two-source extractors.

Finally, despite the lack of explicit non-malleable extractors for arbitrarily linear entropy, we give the first 2-round privacy amplification protocol with asymptotically optimal entropy loss and communication complexity for (n, k) sources with $k = \alpha n$ for any constant $\alpha > 0$. This dramatically improves previous results and answers an open problem in [DLWZ11].

*Partially supported by NSF Grants CCF-0634811, CCF-0916160, THECB ARP Grant 003658-0113-2007, and a Simons postdoctoral fellowship.

1 Introduction

The broad area of *randomness extraction* studies the problem of converting a weakly random source into a distribution that is close to the uniform distribution in statistical distance. Over the past decades extensive research has been conducted in this area. Among which, a long line of research ([SZ99, Tre01, RRV02, LRVW03, GUV09, DW08, DKSS09] to name a few) studies the so called "seeded extractors", as defined by Nisan and Zuckerman [NZ96]. Besides its original motivation in computing with imperfect random sources, seeded extractors have found applications in coding theory, cryptography, complexity and many other areas. We refer the reader to [FS02, Vad02, **?**] for a survey on this subject. Nowadays we have nearly optimal constructions of seeded extractors [LRVW03, GUV09, DW08, DKSS09].

Another line of research focuses on the problem of extracting random bits from several independent sources [CG88, BIW04, BKS$^+$05, Raz05, Bou05, Rao06, BRSW06, Li11]. In this case, however, the best known construction is far from optimal. Specifically, the probabilistic method shows that there exists an extractor for two independent sources on n bits with each having roughly $\log n$ bits of entropy, while the best two-source extractor to date can only achieve entropy slightly below $n/2$ [Bou05]. The best known extractor for small entropy k requires $O(\log n/\log k)$ independent sources [Rao06, BRSW06]. Moreover, it seems hard to improve these results. Especially in the two-source case, after decades of efforts the entropy requirement only drops from anything above $n/2$ [CG88] to slightly below $n/2$ [Bou05].

Recently, a new kind of seeded extractors, called *non-malleable extractors* were introduced in [DW09] to give protocols for the problem of privacy amplification with an active adversary. We now give the definition of a non-malleable extractor below. As a comparison, we also give the definition of a strong seeded extractor.

Notation. We let $[s]$ denote the set $\{1, 2, \ldots, s\}$. For ℓ a positive integer, U_ℓ denotes the uniform distribution on $\{0,1\}^\ell$, and for S a set, U_S denotes the uniform distribution on S. When used as a component in a vector, each U_ℓ or U_S is assumed independent of the other components. We say $W \approx_\varepsilon Z$ if the random variables W and Z have distributions which are ε-close in variation distance.

Definition 1.1. The *min-entropy* of a random variable X is

$$H_\infty(X) = \min_{x \in \mathsf{supp}(X)} \log_2(1/\Pr[X = x]).$$

For $X \in \{0,1\}^n$, we call X an $(n, H_\infty(X))$-source, and we say X has *entropy rate* $H_\infty(X)/n$. We say X is a flat source if it is the uniform distribution over some subset $S \subset \{0,1\}^n$.

Definition 1.2. A function $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a *strong* (k, ε)*-extractor* if for every source X with min-entropy k and independent Y which is uniform on $\{0,1\}^d$,

$$(\mathsf{Ext}(X, Y), Y) \approx_\varepsilon (U_m, Y).$$

Definition 1.3. [1] A function $\mathsf{nmExt} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a (k, ε)-non-malleable extractor if, for any source X with $H_\infty(X) \geq k$ and any function $\mathcal{A} : \{0,1\}^d \to \{0,1\}^d$ such that $\mathcal{A}(y) \neq y$ for all y, the following holds. When Y is chosen uniformly from $\{0,1\}^d$ and independent of X,

$$(\mathsf{nmExt}(X, Y), \mathsf{nmExt}(X, \mathcal{A}(Y)), Y) \approx_\varepsilon (U_m, \mathsf{nmExt}(X, \mathcal{A}(Y)), Y).$$

[1] Following [DLWZ11], we define worst case non-malleable extractors, which is slightly different from the original definition of average case non-malleable extractors in [DW09]. However, the two definitions are essentially equivalent up to a small change of parameters.

As we can see from the definitions, a non-malleable extractor is a stronger version of the strong extractor, in the sense that it requires the output to be close to uniform even conditioned on both the seed Y and the output $\mathsf{nmExt}(X, \mathcal{A}(Y))$ on a different but arbitrarily correlated seed $\mathcal{A}(Y)$.

The motivation to study a non-malleable extractor, the privacy amplification problem, is a fundamental problem in symmetric cryptography that has been studied by many researchers. Bennett, Brassard, and Robert introduced this problem in [BBR88]. The basic setting is that, two parties (Alice and Bob) share an n-bit secret key X, which is weakly random. This could happen because the secret comes from a password or biometric data, which are themselves weakly random, or because an adversary Eve managed to learn some partial information about an originally uniform secret, for example via side channel attacks. We measure the entropy of X by the min-entropy defined above. The goal is to have Alice and Bob communicate over a public channel so that they can convert X into a nearly uniform secret key. Generally, we also assume that Alice and Bob have local private uniform random bits. The problem is the presence of the adversary Eve, who can see every message transmitted in the channel and may or may not change the messages. We assume that Eve has unlimited computational power.

The case where Eve is *passive*, i.e., cannot change the messages, can be solved simply by using the above mentioned strong seeded extractors. The case where Eve is *active* (i.e., can change the messages in arbitrary ways), on the other hand, is much more difficult. Historically, Maurer and Wolf [MW97] gave the first non-trivial protocol in this case. Their protocol takes one round and works when the entropy rate of the weakly-random secret X is bigger than $2/3$. Dodis, Katz, Reyzin, and Smith [DKRS06] later improved this result to give protocols that work for entropy rate bigger than $1/2$. One drawback in both cases is that the final secret key R is much shorter than the min-entropy of X. Later, Dodis and Wichs [DW09] showed that no one-round protocol exists for entropy rate less than $1/2$. The first protocol that breaks the $1/2$ entropy rate barrier is due to Renner and Wolf [RW03], where they gave a protocol that works for essentially any entropy rate. However their protocol takes $O(s)$ rounds and only achieves entropy loss $O(s^2)$, where s in the security parameter of the protocol. Kanukurthi and Reyzin [KR09] simplified their protocol, but the parameters remain essentially the same.

In [DW09], Dodis and Wichs showed that explicit non-malleable extractors can be used to give privacy amplification protocols that take an optimal 2 rounds and achieve optimal entropy loss $O(s)$. They showed that non-malleable extractors exist when $k > 2m + 3\log(1/\varepsilon) + \log d + 9$ and $d > \log(n-k+1) + 2\log(1/\varepsilon) + 7$. However, they only constructed weaker forms of non-malleable extractors and they gave a protocol that takes 2 rounds but that still has entropy loss $O(s^2)$. Chandran, Kanukurthi, Ostrovsky and Reyzin [CKOR10] improved the entropy loss to $O(s)$ but the number of rounds becomes $O(s)$ as well.

Dodis, Li, Wooley and Zuckerman [DLWZ11] constructed the first explicit non-malleable extractor. Their construction works for entropy $k > n/2$, but they use a large seed length $d = n$ and the efficiency when outputting more than $\log n$ bits relies on an unproven assumption. Cohen, Raz, and Segev [CRS12] later gave an alternative construction that also works for $k > n/2$, but uses a short seed length and does not rely on any unproven assumption. The construction in [CRS12] also allows multiple adversarial functions $\{\mathcal{A}_i\}$. By using the non-malleable extractors, these two papers thus gave 2-round privacy amplification protocols that achieve optimal entropy loss $O(s)$. However, since both constructions of non-malleable extractors are only shown to work for entropy $k > n/2$,[2] the protocols also only work for $k > n/2$. For any constant $\delta > 0$, [DLWZ11] also

[2]We remark that the 1-bit case construction in [DLWZ11] is a special case of the construction in [CRS12]. Also,

gave a protocol for $k = \delta n$ than runs in poly$(1/\delta)$ rounds and achieves optimal entropy loss $O(s)$. Recently, Li [Li12] introduced the notion of a non-malleable condenser, which is a relaxation of a non-malleable extractor. He showed that non-malleable condensers for (n, k) sources also give privacy amplification protocols that take an optimal 2 rounds and achieve optimal entropy loss $O(s)$. However, the non-malleable condensers constructed in [Li12] also only work for $k > n/2$. Thus the natural open question is whether we can construct non-malleable extractors or condensers for smaller min-entropy, and whether there are 2-round privacy amplification protocols with optimal entropy loss for smaller min-entropy.

One interesting aspect of the two known constructions of non-malleable extractors is that they are also both two-source extractors. Indeed, the construction in [DLWZ11] is in fact one of the two-source extractors introduced in [CG88], which requires the sources to have min-entropy $> n/2$, and the construction in [CRS12] is in fact the two-source extractor in [Raz05], which requires at least on of the sources to have min-entropy $> n/2$. Coincidently, when used as non-malleable extractors, both of these constructions also require the weak source to have min-entropy $> n/2$. These facts suggest that, despite the fact that these two kinds of extractors seem quite different, there may be some connections between them. However, before this work, no such connection is known.

1.1 Our results

In this paper, we present a strong connection between non-malleable extractors and two-source extractors. First, we show that non-malleable extractors can be used to construct two-source extractors. If the non-malleable extractor works for small min-entropy and has a short seed length (w.r.t. $\log(1/\epsilon)$ where ϵ is the error of the extractor), then the resulted two-source extractor beats the best known construction of two-source extractors.

Theorem 1.4. *Assume that for any $\epsilon > 0$, we have explicit constructions of (k, ϵ)-non-malleable extractors with seed length $d = 2\log(1/\epsilon) + o(n)$ and output length m. Then there exists a constant $\delta > 0$ and an explicit construction of two source extractors that take as input an $(n, (1/2 - \delta)n)$ source and an independent (n, k) source, and output m bits with error $2^{-\Omega(n)}$.*

Note that if k is small, say $k = n/3$ then this already beats the best known two-source extractors, but better results can be achieved if we have explicit constructions of generalized non-malleable extractors. We have the following definition (which already appears in [CRS12]).

Definition 1.5. A function $\mathsf{nmExt} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a (r, k, ε)-non-malleable extractor if, for any source X with $H_\infty(X) \geq k$ and any r function $\mathcal{A}_i : \{0,1\}^d \to \{0,1\}^d, i = 1, \cdots, r$ such that $\mathcal{A}_i(y) \neq y$ for all i and y, the following holds. When Y is chosen uniformly from $\{0,1\}^d$ and independent of X,

$$(\mathsf{nmExt}(X, Y), \{\mathsf{nmExt}(X, \mathcal{A}_i(Y))\}, Y) \approx_\varepsilon (U_m, \{\mathsf{nmExt}(X, \mathcal{A}_i(Y))\}, Y).$$

Here r is the number of adversarial seeds. Note that traditional non-malleable extractors are just $(1, k, \epsilon)$-non-malleable extractors according to our definition. In Appendix A we show that for any constant r, (r, k, ε)-non-malleable extractors exist with seed length $d > \frac{3}{2}\log(n-k) + 3\log(1/\epsilon) + O(1)$. Now we have the following theorem.

it is possible that the construction in [DLWZ11] can work for entropy $k \leq n/2$ (but until now nobody can prove it), but the construction in [CRS12] in general cannot work for entropy $k \leq n/2$.

Theorem 1.6. *For any constant $b > 2$ and any constant $0 < \delta < 1$, there exists a constant $C = C(\delta) = \text{poly}(1/\delta)$ such that the following holds. Assume that for any $\epsilon > 0$ there exists an explicit construction of (C, k, ϵ)-non-malleable extractors with seed length $d = b \log(1/\epsilon) + o(n)$ and output length m. Then there exists an explicit construction of two source extractors that take as input an $(n, \delta n)$ source and an independent (n, k) source, and output m bits with error $2^{-\Omega(n)}$.*

Note that if we have a (C, k, ϵ)-non-malleable extractor for $k = \delta n$ and some constant $C = C(\delta) = \text{poly}(1/\delta)$ then this will give us a two-source extractor for $(n, \delta n)$ sources. If δ is small this will be a big breakthrough for two-source extractors. This also implies that, given current techniques, the (r, k, ϵ)-non-malleable extractor in [CRS12] is probably the best that we can achieve.

Next, we show that in the opposite direction, certain two-source extractors can be used to construct non-malleable extractors. The two-source extractors we will use are those that are constructed based on the inner product function. More specifically, we will consider two-source extractors of the form $\mathsf{TExt} = \mathsf{IP}(f(X), Y)$, where IP is the inner product function over $\mathbb{F}_2$ and $f(X)$ stands for some function (encoding) of the source X. We have the following theorem.

Theorem 1.7. *Given two integers r, ℓ such that $\ell > r$. Assume that we have a two-source extractor $\mathsf{TExt} = \mathsf{IP}(f(X), W)$ such that when given an (n, k)-source X and an independent $(n_2, n_2/(r+1) - \ell)$-source W, TExt outputs 1 bit with error ϵ. Then there exists an explicit construction of (r, k, ϵ')-non-malleable extractors that output 1 bit with error $\epsilon' = O(r 2^{r-\ell} + 2^{\frac{3r}{2}} \epsilon)$.*

Using this theorem, and by combining known two-source extractors, we obtain new and improved constructions of non-malleable extractors. We give the first explicit constructions of non-malleable extractors that work for min-entropy $k < n/2$. One of them is unconditional and works for $k = (1/2 - \delta)n$ for some universal constant $\delta > 0$. The other is conditional but can potentially work for $k = \delta n$ for any constant $\delta > 0$. Specifically, we have the following theorems.

Theorem 1.8. *There exists a constant $0 < \delta < 1$ and an explicit (k, ϵ)-non-malleable extractor $\mathsf{nmExt} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$ with $k = (1/2 - \delta)n$, $m = \Omega(n)$ and $\epsilon = 2^{-\Omega(n)}$.*

Our conditional result needs to use an affine extractor and an assumption from additive combinatorics, as used in [BSZ11]. Thus we first define affine extractors and state the assumption.

Definition 1.9. An $[n, m, \rho, \epsilon]$ affine extractor is a deterministic function $f : \{0,1\}^n \to \{0,1\}^m$ such that whenever X is the uniform distribution over some affine subspace over $\mathbb{F}_2^n$ with dimension ρn, we have that for every $z \in \{0,1\}^m$,

$$|\Pr[f(X) = z] - 2^{-m}| < \epsilon.$$

Note that we bound the error by the ℓ^∞ norm instead of the traditional ℓ^1 norm, as in [BSZ11]. We will let λ denote the entropy loss rate, i.e., $\lambda = 1 - \frac{m}{\rho n}$. We note that it is straightforward to show by the probabilistic method that such extractors exist for any constant $\rho, \lambda > 0$. However the state of art constructions only achieve λ bigger than $1/2$.

[BSZ11] also introduced the Approximate Duality conjecture (ADC), which basically says that if two independent sources X, Y with linear entropy are such that $\mathsf{IP}(X, Y)$ is not close to uniform, then there exist two subsources $X' \subset X, Y' \subset Y$ with small deficiency such that $\mathsf{IP}(X', Y')$ is constant. In [BSZ11] it is shown that ADC is implied by the well-known Polynomial Freiman-Ruzsa Conjecture in additive combinatorics. For a formal definition, see Section 6.3. We now have the following theorems.

Theorem 1.10. *Assume the ADC conjecture and we have an explicit* $[n, m, \frac{2}{3}, 2^{-m}]$ *affine extractor with* $m = (1-\lambda)\frac{2}{3}n$. *Then there exists a semi-explicit* (k, ϵ)*-non-malleable extractor* $\mathsf{nmExt} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *with* $k = \frac{3\lambda}{1+2\lambda}n$, $d = \frac{3}{2+4\lambda}n$, $m = \Omega(n)$ *and* $\epsilon = 2^{-\Omega(n)}$.

Theorem 1.11. *Given a constant integer* r, *assume the ADC conjecture and we have an explicit* $[n, m, \frac{r+1}{r+2}, 2^{-m}]$ *affine extractor with* $m = (1-\lambda)\frac{r+1}{r+2}n$. *Then there exists a semi-explicit* (r, k, ϵ)*-non-malleable extractor with* $k = \frac{(r+2)\lambda}{1+(r+1)\lambda}n$, *seed length* $d = \frac{r+2}{r+1+(r+1)^2\lambda}n - 1$ *and* $\epsilon = 2^{-\Omega(n)}$.

Remark 1.12. Here we use the "*semi-explicit*" to mean that the construction may run in time 2^n. It is semi-explicit in the sense that the running time is polynomial in the length of the extractor's truth table (note that an exhaustive search takes time 2^{2^n}). If we have affine extractors with large output size such that $\lambda \to 0$, then we can essentially achieve $k = \alpha n$ for any constant $\alpha > 0$.

Finally, we give a new privacy amplification protocol for min-entropy $k = \delta n$ for any constant $\delta > 0$. Although we don't have explicit non-malleable extractors or condensers for such small k, our protocol simultaneously achieves optimal round complexity (2 rounds), asymptotically optimal entropy loss and asymptotically optimal communication complexity. This is the first optimal privacy amplification protocol for arbitrarily linear min-entropy. We have the following theorem.

Theorem 1.13. *For any constant* $0 < \delta < 1$ *there exists a constant* $0 < \beta < 1$ *such that as long as* $s \le \beta n$, *there is an efficient 2-round privacy amplification protocol for any* $(n, \delta n)$ *weak secret* X *with security parameter* s, *entropy loss* $O(s + \log n)$ *and communication complexity* $O(s + \log n)$.

Thus, in the case where $k = \delta n$, our result dramatically improves all previous results. Especially, it improves the round complexity of the protocol in [DLWZ11] from $\mathrm{poly}(1/\delta)$ to 2, and thus answers an open problem in [DLWZ11].

2 Overview of The Constructions and Techniques

In this section we give an overview of our constructions and the techniques used. In order to give a clean description, we shall be informal and imprecise sometimes.

2.1 From non-malleable extractors to two-source extractors

Given a (k, ϵ) non-malleable extractor nmExt with seed length $d = 2\log(1/\epsilon) + o(n)$, here is how we can get a two-source extractor. Assume that we have an (n, k) source X and an independent $(n, (1/2 - \delta)n)$ source Y for some constant $\delta > 0$. Our first step is to use the 1-bit condenser in [Zuc07] to convert Y into two sources $\bar{Y}_1, \bar{Y}_2$ such that each of them has $l = \Omega(n)$ bits and one of them has min-entropy at least $(1/2 + \delta)l$. Note that for an appropriately chosen δ this is indeed possible. Without loss of generality assume that $\bar{Y}_1$ has min-entropy at least $(1/2 + \delta)l$.

Our key observation here is that $\bar{Y}_2$ can now be viewed as a function of $\bar{Y}_1$. More precisely, we show that the source Y is a convex combination of sources $\{Y^i\}$ such that for each Y^i, the corresponding $\bar{Y}_1^i$ also has min-entropy at least $(1/2 + \delta)l$, and $\bar{Y}_2^i$ is a deterministic function of $\bar{Y}_1^i$. Now this looks like the setting of a non-malleable extractor, where we have one seed and another correlated seed. However, there is a small problem: $\bar{Y}_1^i$ and $\bar{Y}_2^i$ may be equal sometimes. To solve this, we let $Y_1 = \bar{Y}_1 \circ 0$ and $Y_2 = \bar{Y}_2 \circ 1$. In this way we guarantee that Y_1^i and Y_2^i are different, and Y_2^i is still a function of Y_1^i. Finally, this only increases the length of the seed by 1.

Now we are all set, and we can take the two-source extractor to be $\mathsf{TExt}(X,Y) = \mathsf{nmExt}(X,Y_1) \oplus \mathsf{nmExt}(X,Y_2)$. Note that the seed Y_1 here is not uniform. However, a simple argument shows that a non-malleable extractor with seed length d and error ϵ remains a non-malleable extractor even if the seed only has min-entropy k', with error increased to $2^{d-k'}\epsilon$. In our case, with seed length $d = l+1 = \Omega(n) = 2\log(1/\epsilon)+o(n)$ and $k' = (1/2+\delta)l$, the error is $\epsilon' = 2^{d-k'}\epsilon \approx 2^{(1/2-\delta)l}2^{-l/2} = 2^{-\Omega(n)}$. By the non-malleability of nmExt, we get that $\mathsf{TExt}(X,Y)$ is $2^{-\Omega(n)}$-close to uniform.

We note that given any strong extractor with error ϵ and seed length d, it remains an extractor even if the seed only has min-entropy k', with error increased to $2^{d-k'}\epsilon$. However, since the seed length is at least $d = \log(n-k)+2\log(1/\epsilon)-O(1)$, this will never be able to get k' below $d/2$. On the other hand, if the extractor is non-malleable as in our case, then it does allow us to break the entropy rate $1/2$ barrier, and get a two-source extractor for an $(n,(1/2-\delta)n)$ source and an (n,k) source. This shows that non-malleability is a highly non-trivial property of a seeded extractor.

Similarly, if we have (r,k,ϵ)-non-malleable extractors for larger r, then we can afford to have more correlated seeds Y_i, or equivalently, more sources in the output of the condenser. Thus we can deal with smaller entropy in Y. For example, for any constant $\delta > 0$, condensers based on sum-product theorems [BKS+05, Raz05, Zuc07] allow us to convert an $(n,\delta n)$ source into a constant D number of sources such that each of them has $l = \Omega(n)$ bits and one of them has min-entropy at least $0.9l$. If we have $(D-1,k,\epsilon)$-non-malleable extractors with suitable parameters, then we can get two-source extractors or an $(n,\delta n)$ source and an (n,k) source.

2.2 From two-source extractors to non-malleable extractors

As stated before, we focus on two-source extractors of the form $\mathsf{IP}(f(X),Y)$, where IP is the inner product function. First consider the simplest function $\mathsf{IP}(X,Y)$. Note that it is a good two-source extractor. For two independent sources on n bits, it works as long as the sum of the entropies of the two sources is greater than n. However, at first this function does not seem to be a good candidate for a non-malleable extractor. To see this, consider the inner product function over $\mathbb{F}_2$. Let X be a source that is obtained by concatenating the bit 0 with U_{n-1}, and let Y be an independent uniform seed over $\{0,1\}^n$. Now for any $y \in \{0,1\}^n$, let $\mathcal{A}(y)$ be y with the first bit flipped. Thus we see that for all x in the support of X, one has $\langle x,y\rangle = \langle x,\mathcal{A}(y)\rangle$. Therefore, the inner product function is not a non-malleable extractor even for weak sources with min-entropy $k = n-1$.

In the above example, we have that for all x in the support of X, $\mathsf{IP}(x,y) = \mathsf{IP}(x,\mathcal{A}(y))$. Or equivalently, $\mathsf{IP}(x,y+\mathcal{A}(y)) = 0$. How does this happen? Looking closely at this example, our key observation is that this is because the range of Y is *too large*. Indeed, in this example the range of Y is the entire $\{0,1\}^n$, thus for any y the adversary can choose a different $\mathcal{A}(y)$ such that $y+\mathcal{A}(y) = 10\cdots0$ so that $\forall x \in \mathsf{Supp}(X), \mathsf{IP}(x,y+\mathcal{A}(y)) = 0$.

This observation suggests that we should choose the range of Y to be a subset $S \subset \{0,1\}^n$, so that for some y's, the adversary will be unable to choose the appropriate $\mathcal{A}(y)$ from S. Equivalently, we take a shorter seed length l, choose a uniform $y \in \{0,1\}^l$ and map y to an element in $\{0,1\}^n$. This is essentially an encoding. Now let us see what properties we need the encoding to have.

We start with a construction for min-entropy $k > n/2$. Assume that we have an (n,k) source X with $k = (1/2+\delta)n$ for some constant $\delta > 0$. We take an independent and uniform $y \in \{0,1\}^l$ and encode y to $\bar{y} \in \{0,1\}^n$. For any function $\mathcal{A}$, let $\bar{y}'$ be the encoding of $\mathcal{A}(y)$. We will use an injective encoding, so that $\forall y, \bar{y}' \neq \bar{y}$. The output of the non-malleable extractor is then $\mathsf{IP}(X,\bar{Y})$.

To show that $\mathsf{IP}(X,\bar{Y})$ is a non-malleable extractor, it suffices to show that $\mathsf{IP}(X,\bar{Y})$ is close to uniform, and that $\mathsf{IP}(X,\bar{Y}) \oplus \mathsf{IP}(X,\bar{Y}')$ is close to uniform. The first part is easy. If X has min-entropy $k > n/2$, then we can take Y to be the uniform distribution over some $l \geq n/2$ bits. Since the encoding is injective, $\bar{Y}$ will have min-entropy $l \geq n/2$. Thus $\mathsf{IP}(X, \bar{Y})$ is close to uniform. For the second part, note that $\mathsf{IP}(X, \bar{Y}) \oplus \mathsf{IP}(X, \bar{Y}') = \mathsf{IP}(X, \bar{Y} + \bar{Y}')$. Thus now we need $\bar{Y} + \bar{Y}'$ to have large min-entropy. Indeed, in the above counterexample where $l = n$, the adversary can choose $\mathcal{A}$ such that $\bar{Y} + \bar{Y}'$ is always equal to $10 \cdots 0$ and thus has entropy 0. Now when we take $l < n$ and map $\{0,1\}^l$ to $S \subset \{0,1\}^n$, we want $\bar{Y} + \bar{Y}'$ to have a large support size.

The ideal case would be that $\bar{Y} + \bar{Y}'$ also has support size $|S| = 2^l$. This can be achieved if the encoding has the following property: for every two different y_1, y_2, we have that $\bar{y}_1 + \bar{y}_1' \neq \bar{y}_2 + \bar{y}_2'$, or equivalently, $\bar{y}_1 + \bar{y}_1' + \bar{y}_2 + \bar{y}_2' \neq 0$. Indeed, if this is true then $\bar{Y} + \bar{Y}'$ also has min-entropy $l \geq n/2$, and thus $\mathsf{IP}(X, \bar{Y}) \oplus \mathsf{IP}(X, \bar{Y}')$ is close to uniform. Looking carefully at this property, we see that it can be ensured (at least almost ensured, as we will explain shortly) if we have another property: the elements in S (when viewed as vectors in $\mathbb{F}_2^n$) are 4-wise linearly independent. Indeed, assume that the elements in S are 4-wise linearly independent. Then if $\bar{y}_1 + \bar{y}_1' + \bar{y}_2 + \bar{y}_2' = 0$, the only possible situation is that $\bar{y}_1' = \bar{y}_2$ and $\bar{y}_2' = \bar{y}_1$. Thus there cannot be three different y_1, y_2, y_3 such that $\bar{y}_1 + \bar{y}_1' = \bar{y}_2 + \bar{y}_2' = \bar{y}_3 + \bar{y}_3'$. Thus the min-entropy of $\bar{Y} + \bar{Y}'$ is at least $l - 1$.

So now the question is to explicitly find a large subset $S \subset \{0,1\}^n$ such that the elements in S are 4-wise linearly independent. Note that in particular this implies that the sum of any two different pairs of elements in S cannot be the same. Thus we have $\binom{|S|}{2} \leq 2^n$. Therefore $|S|$ can be at most roughly $2^{n/2}$. On the other hand, in order to work for any min-entropy $k > n/2$, we will need $l \geq n/2$ and thus $|S| = 2^l \geq 2^{n/2}$. These are very tight upper and lower bounds. Luckily, we have explicit constructions that meet these bounds. We will think of the elements in S as columns in a parity check matrix of some binary linear code. Thus we basically need a code with block length $2^{n/2}$ and message length $2^{n/2} - n$. The 4-wise linearly independent property basically is equivalent to saying that the code has distance at least 5. This is precisely the $[2^{n/2}, 2^{n/2} - n, 5]$-BCH code. Note that although the parity check matrix has $2^{n/2}$ columns, each column is (a, a^3) for a different element $a \in \mathbb{F}^*_{2^{n/2}}$. Thus the encoding from y to $\bar{y}$ can be computed efficiently.

Once we have the encoding, we can choose $l = n/2$ and we know that $\bar{Y}$ has min-entropy l and $\bar{Y} + \bar{Y}'$ has min-entropy $l - 1$. Now it is straightforward to show that both $\mathsf{IP}(X, \bar{Y})$ and $\mathsf{IP}(X, \bar{Y} + \bar{Y}')$ are close to uniform. Thus we obtain a non-malleable extractor for entropy $k > n/2$.

Thinking about the above encoding for a moment, one realizes that the same encoding can be used in any two-source extractor of the form $\mathsf{IP}(f(X), Y)$. Specifically, assume that $\mathsf{IP}(f(X), Y)$ is a two-source extractor for an (n, k) source X and an independent $(n, n/2 - 1)$ source Y. Then by the same argument above, if we choose the seed Y to be the uniform distribution over $\{0,1\}^{n/2}$ and encode Y to $\bar{Y}$ like before, we will have that both $\bar{Y}$ and $\bar{Y} + \bar{Y}'$ have min-entropy at least $n/2 - 1$. Thus both $\mathsf{IP}(f(X), \bar{Y})$ and $\mathsf{IP}(f(X), \bar{Y}) \oplus \mathsf{IP}(f(X), \bar{Y}')$ are close to uniform. Therefore we get a non-malleable extractor for min-entropy k.

Similarly, if we have a two-source extractor $\mathsf{IP}(f(X), Y)$ for an (n, k) source X and an independent (n, k') source Y with $k' \approx n/(r+1)$, then we can use a BCH code with distance $2r + 3$ to construct a (r, k, ϵ)-non-malleable extractor. We choose the seed Y to be the uniform distribution over $\{0,1\}^{n/(r+1)}$ and encode Y to $\bar{Y}$ using the parity check matrix, i.e., $\bar{Y} = (Y, Y^3, \cdots, Y^{2r+1})$ when Y is viewed as an element in $\mathbb{F}^*_{2^{n/(r+1)}}$. Since the columns of the parity check matrix are $2(r+1)$-wise linearly independent, we can show that for any subset $S \subseteq [r]$, $\bar{Y} \oplus \bigoplus_{i \in S} \overline{\mathcal{A}_i(Y)}$ has min-entropy roughly $n/(r+1)$. Thus $\mathsf{IP}(f(X), \bar{Y}) \oplus \bigoplus_{i \in S} \mathsf{IP}(f(X), \overline{\mathcal{A}_i(Y)})$ is close to uniform. Therefore we get a (r, k, ϵ)-non-malleable extractor.

2.3 Non-malleable extractors for min-entropy $k < n/2$

We give the first construction of non-malleable extractors for min-entropy $k < n/2$ by observing that the encoding of sources in [Bou05] gives a function f such that $\mathsf{IP}(f(X), Y)$ is a two-source extractor for an $(n, (1/2-\delta)n)$ source X and an independent (n, k') source Y with $k' \approx n/2$.

Specifically, let X be a distribution over some vector space $\mathbb{F}_q^n$ and let cX be the distribution obtained by sampling $x_1, x_2, \cdots, x_c$ from c independent copies of X and computing $\sum x_i$. By Fourier analysis and the Cauchy-Schwarz inequality one can show that in order to prove $\mathsf{IP}(X, Y)$ is close to uniform, it suffices to prove that $\mathsf{IP}(cX, Y)$ is close to uniform with a smaller error, for some integer $c > 1$. In [Bou05], Bourgain showed that for a weak source X with min-entropy rate $1/2 - \delta$ for some constant $\delta > 0$, one can encode X to $\mathsf{Enc}(X)$ such that $3\mathsf{Enc}(X)$ is close to having min-entropy rate $1/2 + \delta$. Thus $\mathsf{IP}(\mathsf{Enc}(X), Y)$ is a two-source extractor that meets our needs. Therefore we obtain our non-malleable extractors for min-entropy $k = (1/2 - \delta)n$.

2.4 Non-malleable extractors for any constant min-entropy rate

In [BSZ11], Ben-Sasson and Zewi showed that affine extractors with large output size can be used to construct two source extractors for min-entropy rate $< 1/2$. Their "preimage construction" can potentially achieve any constant min-entropy rate. We observe that their encoding gives a function f such that $\mathsf{IP}(f(X), Y)$ is a two-source extractor for two independent sources with min-entropy rate δ for any constant $\delta > 0$. Specifically, they showed that if we have an affine extractor with large output size, then there is an injective mapping $F : \{0,1\}^n \to \{0,1\}^{n'}$ that maps $\{0,1\}^n$ into the preimage of a certain output of the affine extractor, such that for any weak source X with min-entropy δn, $F(\mathsf{Supp}(X))$ is not contained in any affine subspace of dimension say $(1-\delta/2)n'$. Thus when Y is a $(n', \delta n')$ source, we have that $\mathsf{IP}(F(X), Y)$ is non-constant. Next, similar as in [BSZ11], the ADC conjecture implies that in fact $\mathsf{IP}(F(X), Y)$ is close to uniform. Thus $\mathsf{IP}(F(X), Y)$ is a two-source extractor that meets our needs. Therefore we obtain a non-malleable extractor (and even a (r, k, ϵ)-non-malleable extractor) for min-entropy $k = \delta n$.

2.5 Increasing output size

We can also increase the output size to $\Omega(n)$ for all our constructions with 1 bit output. To do this, note that we encode the seed Y by using the columns of a parity check matrix of a BCH code. Equivalently, the encoding is that $\bar{Y} = (Y, Y^3)$ when we use a field $\mathbb{F}_{2^l}$ with $l = \Theta(n)$ and Y is viewed as an element in $\mathbb{F}_{2^l}^*$. Now treat $\mathbb{F}_{2^l}$ as the vector space $\mathbb{F}_2^l$ and take l elements $b_1, \cdots, b_l \in \mathbb{F}_{2^l}$ that corresponds to a basis of $\mathbb{F}_2^l$. For each b_i we define one bit $Z_i = \mathsf{IP}(f(X), b_i\bar{Y})$.

We then show that $\{Z_i\}$ satisfy the conditions of a non-uniform XOR lemma, Lemma 3.13. Specifically, let $Z_i' = \mathsf{IP}(f(X), b_i\bar{Y}')$ where $Y' = \mathcal{A}(Y)$. For any non-empty subset $S_1 \subset [l]$ and any subset $S_2 \subset [l]$, by the linearity of the inner product function, the xor of Z_i's where $i \in S_1$ and Z_j''s where $j \in S_2$ is of the form $\mathsf{IP}(f(X), t_1\bar{Y} + t_2\bar{Y}')$, with $t_1, t_2 \in \mathbb{F}_{2^l}$. Since S_1 is non-empty we have $t_1 \neq 0$. We then show that $t_1\bar{Y} + t_2\bar{Y}'$ roughly has the same min-entropy as Y (at least the min-entropy of Y minus $\log 3$). Thus $\mathsf{IP}(f(X), t_1\bar{Y} + t_2\bar{Y}')$ is close to uniform. We further show that the error is $2^{-\Omega(n)}$. Thus by Lemma 3.13 we can output $m = \Omega(n)$ bits with error $2^{-\Omega(n)}$.

2.6 Reducing seed length

In all the constructions where we encode the seed Y by a parity check matrix, the seed length is linear in the source length. However the error is also $2^{-\Omega(n)}$. If we only need to achieve a bigger error, we can reduce the seed length by using the parity check matrix of a BCH code with larger distance. Specifically, when the distance is $2t+1$ the seed length is roughly n/t. However we need to guarantee something else. For example, in the construction for min-entropy $k > n/2$, we need to show that both $\mathsf{IP}(X, \bar{Y})$ and $\mathsf{IP}(X, (\bar{Y} + \bar{Y}'))$ are still close to uniform. This can be shown as follows. Since now the columns of the parity check matrix are $2t$-wise linearly independent, both $\frac{t}{2}\bar{Y}$ and $\frac{t}{2}(\bar{Y} + \bar{Y}')$ will now have min-entropy roughly $\frac{t}{2}H_\infty(Y) = n/2$. Thus we can conclude that both $\mathsf{IP}(X, \frac{t}{2}\bar{Y})$ and $\mathsf{IP}(X, \frac{t}{2}(\bar{Y} + \bar{Y}'))$ are close to uniform, and therefore both $\mathsf{IP}(X, \bar{Y})$ and $\mathsf{IP}(X, (\bar{Y} + \bar{Y}'))$ are also close to uniform, by the Cauchy-Schwarz inequality. However the error increases according to the seed length. Calculations show that we can get seed length $d = O(\log n + \log(1/\epsilon))$.

2.7 An optimal privacy amplification protocol for $k = \delta n$

In [DLWZ11], the authors give a privacy amplification protocol for $k = \delta n$ with $C = \mathrm{poly}(1/\delta)$ rounds and entropy loss $\mathrm{poly}(1/\delta)s$, where s is the security parameter. Here we want to somehow "compress" the protocol into 2 rounds while still keeping the entropy loss to be $O(s)$. As in [DLWZ11], we first use the condenser in [BKS+05, Raz05, Zuc07] to convert the shared (n, k) source X into a somewhere rate-0.9 source $(X_1, \cdots, X_C)$ with $C = \mathrm{poly}(1/\delta)$ rows. Now the high-level idea of the protocol is as follows. In the first round, Alice samples a fresh random string Y_1 from her private random bits and sends it to Bob, where Bob receives a possibly modified version Y_1'. In the second round, Bob samples a fresh random string W' from his private random bits and tries to send it to Alice, where Alice receives a possibly modified version W. We want a protocol such that if Eve does not change Y_1, then with high probability Bob can authenticate W' to Alice and they can both output $\mathsf{Ext}(X, W')$ as the final outputs, by using a strong seeded extractor Ext. If Eve does change Y_1, then with high probability Alice should be able to detect this and reject.

The first goal is relatively easy to achieve. At the end of the first round, Alice and Bob compute $Z = \mathsf{Ext}(X, Y_1)$ and $Z' = \mathsf{Ext}(X, Y_1')$ respectively, using a strong extractor Ext. If Eve does not change Y_1 then $Z = Z'$ and is private and uniform. Thus in the second round Bob can authenticate W' to Alice by also sending a tag T' produced by a standard MAC (message authentication code) with Z as the key. We now focus on the second goal. If the extractor Ext in computing Z and Z' is non-malleable for entropy k then this can be done by using the protocol proposed by Dodis and Wichs [DW09]. However, we do not have explicit non-malleable extractors for $k = \delta n$.

Nevertheless, we will still have Alice and Bob each produce a variable V and V' respectively. We will ensure that, if Eve changes Y_1 to a different Y_1', then even given T' and V', with high probability Eve cannot come up with the correct V for Alice. If this is true then in the second round we can have Bob also send V' to Alice, where Alice receives a possibly modified version $\bar{V}$. Alice then checks both the tag T and whether $V = \bar{V}$. If either of them fails, Alice rejects. This will give us a privacy amplification protocol.

The first problem with the above strategy is that now V' may give information about Z', thus now the MAC key may not be uniform. This is easy to solve since there are constructions of MACs that work as long as the key has entropy rate $> 1/2$. Thus by limiting the size of V' to be at most half the size of Z', we can ensure that if Eve does not change Y_1, Bob can still authenticate W' to

Alice. We now explain how we produce the variables V, V'.

We actually have Alice produce C variables $V = (V_1, \cdots, V_C)$. Similarly, Bob produces $V' = (V'_1, \cdots, V'_C)$. For this, we first choose a non-malleable extractor and have Alice and Bob each apply the extractor to the somewhere rate-0.9 source $(X_1, \cdots, X_C)$, using Y_1 and Y'_1 as the seeds respectively. Let the outputs be $(\bar{X}_1, \cdots, \bar{X}_C)$ and $(\bar{X}'_1, \cdots, \bar{X}'_C)$. Note that one of the X_i's, say X_g is a rate 0.9-source. Thus we can use the non-malleable extractors in [DLWZ11, CRS12, Li12]. Now we fix Y_1, Y'_1, and we have that $\bar{X}_g$ is uniform and independent of $\bar{X}'_g$. Thus we can fix $\bar{X}'_g$ and $\bar{X}_g$ is still uniform. Next, we fix Z'. Since now Z' is a deterministic function of X, as long as the size of Z' is smaller than the size of $\bar{X}_g$, conditioned on this fixing $\bar{X}_g$ still has a lot of entropy left. We will now have Alice extract each V_i from $\bar{X}_i$, and correspondingly, Bob will extract each V'_i from $\bar{X}'_i$. Note that we can indeed ensure that the size of $\bar{X}_g$ is bigger than Z', while the size of Z' is bigger than the size of $(V'_1, \cdots, V'_C)$ just by limiting the size of each V'_i.

Ideally, we would want V, V' be such that V_g is close to uniform conditioned on $V' = (V'_1, \cdots, V'_C)$. However, we cannot achieve exactly this. Instead, what we can achieve is that V_g is close to uniform conditioned on $(V'_1, \cdots, V'_g)$. Once we have this, we can limit the size of $(V'_{g+1}, \cdots, V'_C)$ to be smaller than the size of V_g. Thus V_g still has a lot of entropy even conditioned on $V' = (V'_1, \cdots, V'_C)$. This will ensure that with high probability Eve cannot come up with the correct V_g. Since we do not know which one of $\{\bar{X}_i\}$ is $\bar{X}_g$, we will choose $(V_1, \cdots, V_C)$ such that the size of V_C is say $2s$, and for any i the size of V_i is twice the size of V_{i+1}. In this way, no matter what g is, the size of $(V'_{g+1}, \cdots, V'_C)$ is the size of V_g minus $2s$. Thus V_g still has $2s$ entropy left conditioned on V'.

Finally we explain how we can achieve the above property. We achieve this by using the "look-ahead" extractor in [DW09] based on an alternating extraction protocol. Specifically, in the first round we also have Alice sample two other random strings (Y_2, Y_3) and send them to Bob, where Bob receives (Y'_2, Y'_3). Note that after we fix (Y_1, Y'_1), (Y'_2, Y'_3) is a deterministic function of (Y_2, Y_3). Now pick a strong extractor Ext and Alice performs the following alternating extraction protocol: $S_1 = Y_3, R_1 = \mathsf{Ext}(X, S_1), S_2 = \mathsf{Ext}(Y_2, R_1), R_2 = \mathsf{Ext}(X, S_2), \cdots S_C = \mathsf{Ext}(Y_2, R_{C-1}), R_C = \mathsf{Ext}(X, S_C)$. Bob will perform the same protocol using (Y'_2, Y'_3) and produces $\{S'_i, R'_i\}$. As long as the size of (S_i, R_i) is limited, this protocol has the property that for any i, (R_i, S_i) is uniform and independent of $\{S_j, R_j, S'_j, R'_j, j < i\}$. We now modify this protocol such that whenever S_i is used to extract R_i, Alice also uses it to extract $V_i = \mathsf{Ext}(\bar{X}_i, S_i)$. Correspondingly, Bob extracts $V'_i = \mathsf{Ext}(\bar{X}'_i, S'_i)$. Now one can show that as long as the size of V_i is also limited, we have that for any i, (R_i, S_i) is uniform and independent of $\{S_j, R_j, V_j, S'_j, R'_j, V'_j, j < i\}$. Specifically, we have that S_g is uniform and independent of $\{S_j, R_j, V_j, S'_j, R'_j, V'_j, j < g\}$. Moreover, we can show that conditioned on the fixing of $\{S_j, R_j, V_j, S'_j, R'_j, V'_j, j < g\}$, both S_g and S'_g are deterministic functions of (Y_2, Y_3) and are thus independent of $\bar{X}_g$. Furthermore, $\bar{X}_g$ still has a lot of entropy left. Now since Ext is a strong extractor, we have that $V_g = \mathsf{Ext}(\bar{X}_g, S_g)$ is uniform conditioned on $\{V_j, V'_j, j < g\}$ and (S_g, S'_g). Note that we have fixed $(\bar{X}'_g, Z')$ before, while $V'_g = \mathsf{Ext}(\bar{X}'_g, S'_g)$ and T' is a function of Z'. Thus V_g is still uniform even conditioned on $(\{V'_j, j \le g\}, T')$. Thus we have achieved our goal.

One small problem with the above discussion is that in the first round Alice sends (Y_1, Y_2, Y_3) to Bob and Bob receives (Y'_1, Y'_2, Y'_3). Thus Y'_1 is a function of (Y_1, Y_2, Y_3). Therefore fixing (Y_1, Y'_1) may cause (Y_2, Y_3) to lose entropy. Thus in the alternating extraction protocol (Y_2, Y_3) may not be uniform. However, by making the size of Y_1 a constant times smaller than the size of Y_3, we can ensure that Y_3 has entropy rate $> 2/3$ conditioned on (Y_1, Y'_1). Thus we can add a step 0 in the alternating extraction protocol: $S_0 = Y_3, R_0 = \mathsf{Raz}(S_0, X), S_1 = \mathsf{Ext}(Y_2, R_0)$ and the following protocol remains the same. Here Raz is the two-source extractor in [Raz05] that works as long as one of the source has entropy rate $> 1/2$. This gives our whole privacy amplification protocol. Note that the entropy loss is $O(2^C s) = 2^{\text{poly}(1/\delta)}s$. For any constant $\delta > 0$ this is still $O(s)$. By using the improved non-malleable extractor in [Li12] that has short seed length and large output size (In fact, it suffices to use the non-malleable condensers in [Li12], instead of extractors), we can achieve randomness complexity (the number of truly random bits needed) $O(Cs) = \text{poly}(1/\delta)s$ and communication complexity $O(2^C s) = 2^{\text{poly}(1/\delta)}s$.

Organization. The rest of the paper is organized as follows. We give some preliminaries in Section 3. In Section 4 we show that non-malleable extractors can be used to construct two-source extractors. In Section 5 we show that two-source extractors based on the inner product function can be used to construct non-malleable extractors. In Section 6 we give our new and improved constructions of non-malleable extractors. In Section 7 we give our privacy amplification protocol for arbitrarily linear min-entropy. We conclude with some open problems in Section 8. The existence of generalized non-malleable extractors is proved in Appendix A and an alternative construction of non-malleable extractors for entropy $(1/2 - \delta)n$ is given in Appendix B.

3 Preliminaries

We often use capital letters for random variables and corresponding small letters for their instantiations. Let $|S|$ denote the cardinality of the set S. Let $\mathbb{Z}_r$ denote the cyclic group $\mathbb{Z}/(r\mathbb{Z})$, and let $\mathbb{F}_q$ denote the finite field of size q. All logarithms are to the base 2.

3.1 Probability distributions

Definition 3.1 (statistical distance)**.** Let W and Z be two distributions on a set S. Their *statistical distance* (variation distance) is

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

We say W is ε-close to Z, denoted $W \approx_\varepsilon Z$, if $\Delta(W, Z) \le \varepsilon$. For a distribution D on a set S and a function $h : S \to T$, let $h(D)$ denote the distribution on T induced by choosing x according to D and outputting $h(x)$. We often view a distribution as a function whose value at a sample point is the probability of that sample point. Thus $\|W - Z\|_{\ell^1}$ denotes the ℓ_1 norm of the difference of the distributions specified by the random variables W and Z, which equals $2\Delta(W, Z)$.

Definition 3.2. A function $\mathsf{TExt} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ is a *strong two source extractor* for min-entropy k_1, k_2 and error ϵ if for every independent (n_1, k_1) source X and (n_2, k_2) source Y,

$$|(\mathsf{TExt}(X, Y), X) - (U_m, X)| < \epsilon$$

and

$$|(\mathsf{TExt}(X, Y), Y) - (U_m, Y)| < \epsilon,$$

where U_m is the uniform distribution on m bits independent of (X, Y).

3.2 Somewhere Random Sources, Extractors and Condensers

Definition 3.3 (Somewhere Random sources)**.** A source $X = (X_1, \cdots, X_t)$ is $(t \times r)$ *somewhere-random* (SR-source for short) if each X_i takes values in $\{0,1\}^r$ and there is an i such that X_i is uniformly distributed.

Definition 3.4. An elementary somewhere-k-source is a vector of sources $(X_1, \cdots, X_t)$, such that some X_i is a k-source. A somewhere k-source is a convex combination of elementary somewhere-k-sources.

Definition 3.5. A function $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a $(k \to l, \epsilon)$-condenser if for every k-source X, $C(X, U_d)$ is ϵ-close to some l-source. When convenient, we call C a rate-$(k/n \to l/m, \epsilon)$-condenser.

Definition 3.6. A function $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a $(k \to l, \epsilon)$-somewhere-condenser if for every k-source X, the vector $(C(X, y)_{y \in \{0,1\}^d})$ is ϵ-close to a somewhere-l-source. When convenient, we call C a rate-$(k/n \to l/m, \epsilon)$-somewhere-condenser.

We are going to use condensers recently constructed based on the sum-product theorem. The following constructions are due to Zuckerman [Zuc07].

Theorem 3.7 ([Zuc07])**.** *There exists a constant $\alpha > 0$ such that for any constant $0 < \delta < 0.9$, there is an efficient family of rate-$(\delta \to (1+\alpha)\delta, \epsilon = 2^{-\Omega(n)})$-somewhere condensers* $\mathsf{Scond} : \{0,1\}^n \to (\{0,1\}^m)^2$ *where $m = \Omega(n)$.*

Theorem 3.8 ([BKS⁺05, Raz05, Zuc07])**.** *For any constant $\beta, \delta > 0$, there is an efficient family of rate-$(\delta \to 1 - \beta, \epsilon = 2^{-\Omega(n)})$-somewhere condensers* $\mathsf{Cond} : \{0,1\}^n \to (\{0,1\}^m)^D$ *where $D = O(1)$ and $m = \Omega(n)$.*

3.3 Average conditional min-entropy

Dodis and Wichs originally defined non-malleable extractors with respect to average conditional min-entropy, a notion defined by Dodis, Ostrovsky, Reyzin, and Smith [DORS08].

Definition 3.9. The *average conditional min-entropy* is defined as

$$\widetilde{H}_\infty(X|W) = -\log\left(\mathrm{E}_{w \leftarrow W}\left[\max_x \Pr[X = x | W = w]\right]\right) = -\log\left(\mathrm{E}_{w \leftarrow W}\left[2^{-H_\infty(X|W=w)}\right]\right).$$

Average conditional min-entropy tends to be useful for cryptographic applications. By taking W to be the empty string, we see that average conditional min-entropy is at least as strong as min-entropy. In fact, the two are essentially equivalent, up to a small loss in parameters.

We have the following lemmas.

Lemma 3.10 ([DORS08])**.** *For any $s > 0$, $\Pr_{w \leftarrow W}[H_\infty(X|W = w) \geq \widetilde{H}_\infty(X|W) - s] \geq 1 - 2^{-s}$.*

Lemma 3.11 ([DORS08])**.** *If a random variable B has at most 2^ℓ possible values, then $\widetilde{H}_\infty(A|B) \geq H_\infty(A) - \ell$.*

To clarify which notion of min-entropy and non-malleable extractor we mean, we use the term *worst-case non-malleable extractor* when we refer to our Definition 1.3, which is with respect to traditional (worst-case) min-entropy, and *average-case non-malleable extractor* to refer to the original definition of Dodis and Wichs, which is with respect to average conditional min-entropy.

Corollary 3.12. *A* (k,ε)*-average-case non-malleable extractor is a* (k,ε)*-worst-case non-malleable extractor. For any* $s>0$*, a* (k,ε)*-worst-case non-malleable extractor is a* $(k+s,\varepsilon+2^{-s})$*-average-case non-malleable extractor.*

Throughout the rest of our paper, when we say non-malleable extractor, we refer to the worst-case non-malleable extractor of Definition 1.3.

3.4 Fourier analysis

We give some basic and standard facts about Fourier analysis here. We normalize as in [DLWZ11]. For functions f,g from a set S to $\mathbb{C}$, we define the inner product $\langle f,g\rangle = \sum_{x\in S} f(x)\overline{g(x)}$. Let D be a distribution on S, sometimes we will also view it as a function from S to $\mathbb{R}$. Note that $\mathrm{E}_D[f(D)] = \langle f, D\rangle$. Now suppose we have functions $h : S \to T$ and $g : T \to \mathbb{C}$. Then

$$\langle g\circ h, D\rangle = \mathrm{E}_D[g(h(D))] = \langle g, h(D)\rangle.$$

Let G be a finite abelian group, we say ϕ is a character of G if it is a homomorphism from G to $\mathbb{C}^\times$. We call the character that maps all elements to 1 the trivial character. Define the Fourier coefficient $\widehat{f}(\phi) = \langle f,\phi\rangle$, and let $\widehat{f}$ denote the vector with entries $\widehat{f}(\phi)$ for all ϕ. Note that for a distribution D, one has $\widehat{D}(\phi) = \mathrm{E}_D[\phi(D)]$.

Since the characters divided by $\sqrt{|G|}$ form an orthonormal basis, the inner product is preserved up to scale: $\langle \widehat{f},\widehat{g}\rangle = |G|\langle f,g\rangle$. As a corollary, we obtain Parseval's equality:

$$\|\widehat{f}\|_{\ell^2}^2 = \langle \widehat{f},\widehat{f}\rangle = |G|\langle f,f\rangle = |G|\|f\|_{\ell^2}^2.$$

Hence by Cauchy-Schwarz,

$$\|f\|_{\ell^1} \le \sqrt{|G|}\|f\|_{\ell^2} = \|\widehat{f}\|_{\ell^2} \le \sqrt{|G|}\|\widehat{f}\|_{\ell^\infty}. \tag{1}$$

For functions $f,g : S\to\mathbb{C}$, we define the function $(f,g) : S\times S\to\mathbb{C}$ by $(f,g)(x,y) = f(x)g(y)$. Thus, the characters of the group $G\times G$ are the functions (ϕ,ϕ'), where ϕ and ϕ' range over all characters of G. We abbreviate the Fourier coefficient $\widehat{(f,g)}((\phi,\phi'))$ by $\widehat{(f,g)}(\phi,\phi')$. Note that

$$\widehat{(f,g)}(\phi,\phi') = \sum_{(x,y)\in G\times G} f(x)g(y)\phi(x)\phi'(y) = \left(\sum_{x\in G} f(x)\phi(x)\right)\left(\sum_{y\in G} g(x)\phi'(x)\right) = \widehat{f}(\phi)\widehat{g}(\phi').$$

In this paper, in the additive group of $\mathbb{F}_p$ we use the characters $e_r(s) = e^{2\pi i rs/p}$ for $r\in\mathbb{F}_p$. It is easy to verify that $\{e_r, r\in F_p\}$ indeed are characters and these characters divided by $\sqrt{p}$ form an orthonormal basis. Note that the trivial character corresponds to the case $r=0$.

We next generalize the characters to the additive group of the field $\mathbb{F}_{p^l}$. In this case, for any $r\in\mathbb{F}_{p^l}$, we use the character $e_r(s) = e^{2\pi i(r\cdot s)/p}$, where r and s are viewed as vectors in $\mathbb{F}_p^l$ and $\cdot$ indicates the inner product function in $\mathbb{F}_p^l$. Again it is easy to verify that these indeed are characters and they form an orthonormal basis (up to a normalization factor of $p^{l/2}$).

3.5 Non-uniform XOR lemma

The following non-uniform XOR lemmas are proved in [DLWZ11].

Lemma 3.13. *Let (W, W') be a random variable on $G \times G$ for a finite abelian group G, and suppose that for all characters ψ, ψ' on G with ψ nontrivial, one has*

$$|\operatorname{E}_{(W,W')}[\psi(W)\psi'(W')]| \leq \epsilon.$$

Then the distribution of (W, W') is $\epsilon|G|$ close to (U, W'), where U is the uniform distribution on G independent of W'. Moreover, for $f : G \times G \to \mathbb{R}$ defined as the difference of distributions $(W, W') - (U, W')$, we have $\|\widehat{f}\|_{\ell^\infty} \leq \epsilon$.

Lemma 3.14. *For every cyclic group $G = \mathbb{Z}_N$ and every integer $M \leq N$, there is an efficiently computable function $\sigma : \mathbb{Z}_N \to \mathbb{Z}_M = H$ such that the following holds. Let (W, W') be a random variable on $G \times G$, and suppose that for all characters ψ, ψ' on G with ψ nontrivial, one has*

$$|\operatorname{E}_{(W,W')}[\psi(W)\psi'(W')]| \leq \epsilon.$$

Then the distribution $(\sigma(W), \sigma(W'))$ is $O(\epsilon M \log N + M/N)$-close to the distribution (U, W') where U stands for the uniform distribution over H independent of W'.

The following non-uniform XOR lemma is proved in [CRS12].

Lemma 3.15. *Let X be a random variable over $\{0,1\}^m$ and Y be a random variable over $\{0,1\}^n$. For any subset $\sigma \subseteq [m]$ and $\tau \subseteq [n]$, let $X_\sigma = \oplus_{i\in\sigma} X_i$ and $Y_\tau = \oplus_{j\in\tau} Y_j$. Assume that for any non-empty $\sigma \subseteq [m]$ and any $\tau \subseteq [n]$, we have $X_\sigma \oplus Y_\tau \approx_\epsilon U$. Then*

$$|(X, Y) - (U_m, Y)| \leq ((2^m - 1) \cdot 2^n)^{1/2}\epsilon.$$

3.6 Strong non-malleable extractor

The following theorem is proved in [Rao07].

Theorem 3.16. *[Rao07] Let $\mathsf{TExt} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ be any two source extractor for min-entropy k_1, k_2 with error ϵ. Then if X is an (n_1, k_1) source and Y is an independent (n_1, k_2') source, we have*

$$|(\mathsf{TExt}(X, Y), Y) - (U_m, Y)| \leq 2^m(2^{k_2 - k_2' + 1} + \epsilon).$$

Here we prove a similar theorem that will enable our non-malleable extractor to be "strong".

Theorem 3.17. *Let $\mathsf{TExt} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ be a two source extractor for min-entropy k_1, k_2 and $\mathcal{A}_i : \{0,1\}^{n_2} \to \{0,1\}^{n_2}, i = 1, \cdots, r$ be r deterministic functions such that for any (n_1, k_1) source X and any independent (n_2, k_2) source Y,*

$$|(\mathsf{TExt}(X, Y), \{\mathsf{TExt}(X, \mathcal{A}_i(Y))\}) - (U_m, \{\mathsf{TExt}(X, \mathcal{A}_i(Y))\})| \leq \epsilon.$$

Then for any (n_2, k_2') source Y' independent of X,

$$|(\mathsf{TExt}(X, Y'), \{\mathsf{TExt}(X, \mathcal{A}_i(Y'))\}, Y') - (U_m, \{\mathsf{TExt}(X, \mathcal{A}_i(Y'))\}, Y')| \leq 2^{(r+1)m}(2^{k_2 - k_2' + 1} + \epsilon).$$

Proof. Let $W = \mathsf{TExt}(X, Y)$ and $W'_i = \mathsf{TExt}(X, \mathcal{A}_i(Y))$ for any $i \in [r]$. Let $\bar{W}$ be the vector $(W'_1, \cdots, W'_r)$. Let $\bar{z}$ be the vector $(z'_1, \cdots, z'_r) \in (\{0,1\}^m)^r$. For any $(z, \bar{z}) \in \{0,1\}^m \times (\{0,1\}^m)^r$, define the set of bad y's for $(z, \bar{z})$ to be

$$B_{z,\bar{z}} = \{y : |\Pr[W = z, \bar{W} = \bar{z}] - 2^{-m} \Pr[\bar{W} = \bar{z}]| > \epsilon\}.$$

Then we must have

Claim 3.18. *For every $(z, \bar{z})$, $|B_{z,\bar{z}}| < 2 \cdot 2^{k_2}$.*

To see this, assume for the sake of contradiction that $|B_{z,\bar{z}}| \geq 2 \cdot 2^{k_2}$ for some $(z, \bar{z})$. Let

$$B^+_{z,\bar{z}} = \{y : \Pr[W = z, \bar{W} = \bar{z}] - 2^{-m} \Pr[\bar{W} = \bar{z}] > \epsilon\}$$

and

$$B^-_{z,\bar{z}} = \{y : \Pr[W = z, \bar{W} = \bar{z}] - 2^{-m} \Pr[\bar{W} = \bar{z}] < -\epsilon\}.$$

Then $|B_{z,\bar{z}}| = |B^+_{z,\bar{z}}| + |B^-_{z,\bar{z}}|$ and thus one of them must have size $\geq 2^{k_2}$. Without loss of generality assume that $|B^+_{z,\bar{z}}| \geq 2^{k_2}$. Then we can let Y to be the uniform distribution over $|B^+_{z,\bar{z}}|$ and Y is independent of X, but $|(W, \bar{W}) - (U_m, \bar{W})| > \epsilon$, which is a contradiction.

Let $B = \cup_{z,\bar{z}} B_{z,\bar{z}}$. We have $|B| < 2^{(r+1)m} \cdot 2 \cdot 2^{k_2} = 2^{(r+1)m+1} 2^{k_2}$. Now we can bound $|(W, \bar{W}, Y') - (U_m, \bar{W}, Y')|$ when Y' is an independent (n_2, k'_2) source, as follows.

$$\begin{aligned}
&|(W, \bar{W}, Y') - (U_m, \bar{W}, Y')| \\
\leq &\sum_{y \in \mathsf{Supp}(Y')} 2^{-k'_2} |(W, \bar{W})|_{Y'=y} - (U_m, \bar{W})|_{Y'=y}| \\
= &\sum_{y \in \mathsf{Supp}(Y') \cap B} 2^{-k'_2} |(W, \bar{W})|_{Y'=y} - (U_m, \bar{W})|_{Y'=y}| + \sum_{y \in \mathsf{Supp}(Y') \setminus B} 2^{-k'_2} |(W, \bar{W})|_{Y'=y} - (U_m, \bar{W})|_{Y'=y}| \\
< &2^{-k'_2} 2^{(r+1)m+1} 2^{k_2} + 2^{(r+1)m} \epsilon \\
= &2^{(r+1)m} (2^{k_2 - k'_2 + 1} + \epsilon).
\end{aligned}$$

■

3.7 Basic properties of the inner product function

Here we prove some basic properties of the inner product function.

Lemma 3.19. *Let $\mathbb{F}_p$ be a field and X, Y be two independent random variables over $\mathbb{F}_p^l$. Assume that X has min-entropy k_1 and Y has min-entropy k_2. Let $Z = \mathsf{IP}(X, Y) = X \cdot Y$ be the inner product function where the operation is in $\mathbb{F}_p$. For any non-trivial character e_r where $r \in \mathbb{F}_p$,*

$$|E_{X,Y}[e_r(Z)]|^2 \leq p^l 2^{-(k_1+k_2)}.$$

Proof. Note that if a weak random source W has min-entropy k, then $\|W\|_{\ell^\infty} \leq 2^{-k}$, and $\|W\|_{\ell^2}^2 = \sum_w (\Pr[W = w])^2 \leq 2^{-k} \sum_w \Pr[W = w] = 2^{-k}$.

For a fixed $Y = y$,

$$E_X[e_r(x \cdot y)] = E_X[e_{ry}(X)] = \langle e_{ry}, X \rangle = \overline{\widehat{X}(e_{ry})}.$$

Thus

$$E_{X,Y}[e_r(Z)] = E_Y[E_X[e_r(x \cdot y)]] = E_Y[\overline{\widehat{X}(e_{ry})}] = \langle Y, \widehat{X} \rangle.$$

Therefore by Cauchy-Schwartz,

$$\begin{aligned}
(E_{X,Y}[e_r(Z)])^2 &\leq \langle Y, Y \rangle \cdot \langle \widehat{X}, \widehat{X} \rangle \\
&= \|Y\|_{\ell^2}^2 \|\widehat{X}\|_{\ell^2}^2 = p^l \|Y\|_{\ell^2}^2 \|X\|_{\ell^2}^2 \\
&\leq p^l 2^{-k_1} 2^{-k_2} = p^l 2^{-(k_1+k_2)}.
\end{aligned}$$

□

Now for any weak random source W, we let $2W = W + W$ stand for the distribution that is obtained by first sampling w_1, w_2 from two independent and identical distributions according to W, and then computing $w_1 + w_2$. Similarly $W - W$ is obtained by first sampling w_1, w_2 and then computing $w_1 - w_2$. Similarly we define cW to be the distribution by sampling w_i from c independent and identical distributions according to W, and then computing the sum. We now have the following lemma.

Lemma 3.20. *Let X, Y be two independent random variables over $\mathbb{F}_p^l$. For any two integers c_1, c_2, let $X_{c_1} = 2^{c_1}X - 2^{c_1}X$ and $Y_{c_2} = 2^{c_2}Y - 2^{c_2}Y$. Then for any non-trivial character ψ,*

$$|E_{X,Y}[\psi(X \cdot Y)]| \leq |E_{X_{c_1}, Y_{c_2}}[\psi(X_{c_1} \cdot Y_{c_2})]|^{1/2^{c_1+c_2+2}}.$$

Proof. First note

$$|E_{X,Y}[\psi(X \cdot Y)]| = |E_Y[E_X[\psi(X \cdot Y)]]| \leq E_Y |E_X[\psi(X \cdot Y)]|.$$

Note that $\psi(s) = e^{2\pi i r s/p}$ for some $r \in \mathbb{F}_p$. Thus by Jensen's inequality,

$$\begin{aligned}
(E_{X,Y}[\psi(X \cdot Y)])^2 &\leq E_Y |E_X[\psi(X \cdot Y)]|^2 = E_Y[E_X[\psi(X \cdot Y)]\overline{E_X[\psi(X \cdot Y)]}] \\
&= |E_Y \sum_{x_1, x_2} X(x_1)X(x_2)\psi((x_1 - x_2) \cdot Y)| \\
&= |E_Y E_{X-X}[\psi((X - X) \cdot Y)]| \\
&= |E_{X_1, Y}[\psi(X_1 \cdot Y)]|
\end{aligned}$$

where $X_1 = X - X$.

Apply the above procedure again, we get that

$$(E_{X,Y}[\psi(X \cdot Y)])^4 \leq |E_{X_1,Y}[\psi(X_1 \cdot Y)]|^2 \leq |E_{X_2,Y}[\psi(X_2 \cdot Y)]|,$$

where $X_2 = X_1 - X_1 = 2X - 2X$.

Repeat the procedure for c_1 times, we get that

$$(E_{X,Y}[\psi(X \cdot Y)])^{2^{c_1+1}} \leq |E_{X_{c_1},Y}[\psi(X_{c_1} \cdot Y)]|,$$

where $X_{c_1} = 2^{c_1}X - 2^{c_1}X$.

similarly, we can apply the argument to Y for another c_2 times, and we get

$$(E_{X,Y}[\psi(X \cdot Y)])^{2^{c_1+c_2+2}} \leq |E_{X_{c_1},Y_{c_2}}[\psi(X_{c_1} \cdot Y_{c_2})]|,$$

where $X_{c_1} = 2^{c_1}X - 2^{c_1}X$ and $Y_{c_2} = 2^{c_2}Y - 2^{c_2}Y$. Thus the lemma is proved. $\square$

3.8 Incidence theorems

We need the following theorems about point line incidences. For a field $\mathbb{F}$, we call a subset $\ell \subset F \times F$ a line if there exist $a, b \in \mathbb{F}$ such that $\ell = \{(x, ax + b)\}$ for all $x \in \mathbb{F}$. Let $P \subset F \times F$ be a set of points and L be a set of lines, we say that a point (x, y) has an incidence with a line ℓ if $(x, y) \in \ell$. The following theorem provides a bound on the number of incidences that can be generated from K points and K lines.

Theorem 3.21. *[BKT04, Kon03] There exist universal constants $\alpha > 0, 0.1 > \beta > 0$ such that for any field $\mathbb{F}_q$ where q is either prime or 2^p for p prime, if L, P are sets of K lines and K points respectively, with $K \leq q^{2-\beta}$, the number of incidences $I(P, L) \leq O(K^{3/2-\alpha})$.*

3.9 BCH codes

In this paper we will only focus on BCH codes over $\mathbb{F}_2$. Given two parameters $m, t \in \mathbb{N}$, a BCH code is a linear code with block length $n = 2^m - 1$, message length roughly $n - mt$ and distance $d \geq 2t + 1$. Specifically, we have the following theorem.

Theorem 3.22. *For all integers m and t there exists an explicit $[n, n - mt, 2t + 1]$-BCH code[3], with $n = 2^m - 1$.*

Since a BCH code is a linear code, we can take its parity check matrix. Note that this is a $mt \times n$ matrix. Let α be a primitive element in $\mathbb{F}^*_{2^m}$, the i'th column of the parity check matrix is of the form $(\alpha^i, (\alpha^i)^3, (\alpha^i)^5, \cdots, (\alpha^i)^{2t-1})$, for $i = 0, 1, \cdots, n - 1$. Since α is a generator in $\mathbb{F}^*_{2^m}$, equivalently, for $y \in \mathbb{F}^*_{2^m}$ we can think of the y'th column to be $(y, y^3, \cdots, y^{2t-1})$.

4 From Non-Malleable Extractors to Two-Source Extractors

In this section we show that non-malleable extractors can be used to construct two-source extractors. First we have the following lemmas.

Lemma 4.1. *Let X be a probability distribution on $\{0,1\}^{n_1}$ and $Y, Y_1, \cdots, Y_m$ be probability distributions on $\{0,1\}^{n_2}$. Assume that there exists a function $f : \{0,1\}^{n_1} \to \{0,1\}^{n_2}$ and positive*

[3] In fact, the message length may not be exactly $n - mt$, but for simplicity we will assume that it is exactly $n - mt$. The small error does not affect our analysis. Also, for small t the message length is exactly $n - mt$.

numbers $v_1, \cdots v_m$ with $\sum_i v_i = 1$ such that $Y = f(X)$ and $Y = \sum_i v_i Y_i$. Then there exist probability distributions $X_1, \cdots, X_m \in \{0,1\}^{n_1}$ such that

$$X = \sum_i v_i X_i \text{ and } Y_i = f(X_i).$$

Proof. We define the distributions $\{X_i\}$ as follows. Let S be the support of Y. For any $y \in S$, let $S_y = \{x \in \mathsf{Supp}(X) : f(x) = y\}$, i.e., S_y is the set of preimages of y. Let $p(y) = \Pr[Y = y]$ and $\forall i, p_i(y) = \Pr[Y_i = y]$. Thus we have

$$p(y) = \sum_i v_i p_i(y)$$

and

$$p(y) = \sum_{x \in S_y} \Pr[X = x].$$

Now for any $x \in \mathsf{Supp}(X)$, let $y = f(x)$. Let $\Pr[X_i = x] = q_i(x) = \frac{p_i(y)}{p(y)} \Pr[X = x]$. First note that

$$\sum_{x \in S_y} q_i(x) = \sum_{x \in S_y} \frac{p_i(y)}{p(y)} \Pr[X = x] = \frac{p_i(y)}{p(y)} \sum_{x \in S_y} \Pr[X = x] = \frac{p_i(y)}{p(y)} p(y) = p_i(y).$$

Thus we have that $Y_i = f(X_i)$. Next note that

$$\sum_x q_i(x) = \sum_y \sum_{x \in S_y} q_i(x) = \sum_y p_i(y) = 1.$$

Therefore for any i, X_i is indeed a probability distribution. Finally, note that

$$\sum_i v_i q_i(x) = \sum_i \frac{v_i p_i(y)}{p(y)} \Pr[X = x] = \Pr[X = x].$$

Thus we have that $X = \sum_i v_i X_i$. □

Lemma 4.2. *Let X be a probability distribution on $\{0,1\}^{n_1}$. Assume that there exists a function $f : \{0,1\}^{n_1} \to \{0,1\}^{n_2}$ such that $Y = f(X)$ is an (n_2, k)-source. Then there exits an integer m and (flat) (n_1, k)-sources $X_1, \cdots, X_m$, bijections $f_1, \cdots, f_m : \{0,1\}^{n_1} \to \{0,1\}^{n_2}$, positive numbers $v_1, \cdots v_m$ with $\sum_i v_i = 1$ such that*

$$X = \sum_i v_i X_i \text{ and } f(X_i) = f_i(X_i).$$

Proof. First, note that every (n_2, k)-source is a convex combination of flat (n_2, k)-sources. By Lemma 4.1, if $Y = f(X)$ and $Y = \sum_i v_i Y_i$, then there exist probability distributions $X_1, \cdots, X_m$ on $\{0,1\}^{n_1}$ such that $X = \sum_i v_i X_i$ and $Y_i = f(X_i)$. Thus it suffices to consider only the case where Y is a flat (n_2, k)-source.

Now let Y be a flat (n_2, k)-source. For any $y \in \mathsf{Supp}(Y)$, let $S_y = \{x \in \mathsf{Supp}(X) : f(x) = y\}$. For any $x \in \mathsf{Supp}(X)$, let $p(x) = \Pr[X = x]$ and for any $y \in \mathsf{Supp}(Y)$, let $p(y) = \Pr[Y = y]$. Thus we have $\sum_{x \in S_y} p(x) = p(y) = 2^{-k}$. We now decompose X into a convex combination and produce the bijections $f_1, \cdots, f_m$ as follows.

Let $i = 1$. While $\cup_y S_y$ is not empty, do the following.

1. Pick x from $\cup_y S_y$ such that $p(x)$ is the minimum. Assume that $x \in S_y$. Now for all $y' \in \mathsf{Supp}(Y), y' \neq y$, pick an arbitrary $x' \in S_{y'}$. Thus for any x', $p(x') \geq p(x)$. We now let the source X_i be the uniform distribution over the set of the chosen x's, and we let the bijection f_i be such that $f_i(x) = y$ and $f_i(x') = y'$ for all $y' \neq y$. This clearly satisfies the property that $f(X_i) = f_i(X_i)$. Next, we let v_i be the total probability mass of x's, i.e., $v_i = 2^k p(x)$.

2. We now want to subtract the probability mass from both X and Y. Thus for any x', we let $p(x') = p(x') - p(x)$ and if $p(x') = 0$, remove x' from $S_{y'}$. Specifically, we remove x from S_y. Similarly, for any $y \in \mathsf{Supp}(Y)$, let $p(y) = p(y) - p(x)$. Note after this we still have that for any $y \in \mathsf{Supp}(Y)$, $\sum_{x \in S_y} p(x) = p(y)$.

3. Finally, let $i = i + 1$.

Note that in the above algorithm, in each iteration at least one element will be removed from $\cup_y S_y$. Thus the algorithm will terminate after finite steps, and we obtain $X_1, \cdots, X_m$, $f_1, \cdots, f_m$ and $v_1, \cdots v_m$. Note that in each iteration the $p(y)$'s are always the same. Thus in each step we can always obtain a flat (n_1, k)-source X_i and a bijection f_i such that $f_i(X_i) = f(X_i)$. Note that the algorithm terminates only when $\cup_y S_y$ is empty. Since we always have that for any $y \in \mathsf{Supp}(Y)$, $\sum_{x \in S_y} p(x) = p(y)$, when the algorithm terminates we must have that for any $y \in \mathsf{Supp}(Y)$, $p(y) = 0$. Thus we have decomposed X into a convex combination of flat sources $X_1, \cdots, X_m$. In other words, $\sum_i v_i = 1$. $\square$

Lemma 4.3. *Let X be a probability distribution over $\{0,1\}^{n_1}$, Y be a probability distribution over $\{0,1\}^{n_2}$ and $f : \{0,1\}^{n_1} \to \{0,1\}^{n_2}$ be any deterministic function. Assume that $|f(X) - Y| \leq \epsilon$ for some $0 < \epsilon < 1$, then there exists a probability distribution X' over $\{0,1\}^{n_1}$ such that*

$$|X' - X| \leq \epsilon \text{ and } Y = f(X').$$

Proof. For any $y \in \mathsf{Supp}(Y)$, let $p(y) = \Pr[f(X) = y]$ and $q(y) = \Pr[Y = y]$. Let $S_y = \{x \in \mathsf{Supp}(X) : f(x) = y\}$. Thus we have that $p(y) = \sum_{x \in S_y} \Pr[X = x]$. Let $W = \{y \in \mathsf{Supp}(Y) : p(y) > q(y)\}$ and $V = \{y \in \mathsf{Supp}(Y) : p(y) < q(y)\}$. Thus we have that $\sum_{y \in W} |p(y) - q(y)| = \sum_{y \in V} |p(y) - q(y)| = \epsilon$.

We now gradually change the probability distribution X into X', as follows. First let X' be the same probability distribution as X, then, while W is not empty or V is not empty, do the following.

1. Pick $y \in W \cup V$ such that $|p(y) - q(y)| = min\{|p(y') - q(y')|, y' \in W \cup V\}$.

2. If $y \in W$, we decrease $p(y)$ to $q(y)$. Specifically, let $\delta = \tau = p(y) - q(y)$. We pick the elements $x \in S_y$ one by one in an arbitrary order and while $\tau > 0$, do the following. Let $\tau' = min(\Pr[X' = x], \tau)$, $\Pr[X' = x] = \Pr[X' = x] - \tau'$ and $\tau = \tau - \tau'$. Note that since $p(y) = \delta + q(y) \geq \delta$, this process will indeed end when $\tau = 0$ and now $p(y) = q(y)$. Now to ensure that X' is still a probability distribution, we pick any $\bar{y} \in V$ and increase $p(\bar{y})$ to $p(\bar{y}) + \delta$. To do this, simply pick any $x \in S_{\bar{y}}$ and let $\Pr[X' = x] = \Pr[X' = x] + \delta$. Note that after this change we still have that $p(\bar{y}) \leq q(\bar{y})$. Finally, remove y from W and if $p(\bar{y}) = q(\bar{y})$, remove $\bar{y}$ from V.

3. If $y \in V$, we increase $p(y)$ to $q(y)$. Specifically, let $\delta = q(y) - p(y)$. Pick any $x \in S_y$ and let $\Pr[X' = x] = \Pr[X' = x] + \delta$. Now to ensure that X' is still a probability distribution, we pick any $\bar{y} \in W$ and decrease $p(\bar{y})$ to $p(\bar{y}) - \delta$. To do this, let $\tau = \delta$. We pick the elements $x \in S_{\bar{y}}$ one by one in an arbitrary order and while $\tau > 0$, do the following. Let $\tau' = min(\Pr[X' = x], \tau)$, $\Pr[X' = x] = \Pr[X' = x] - \tau'$ and $\tau = \tau - \tau'$. Note that since $p_{\bar{y}} \geq \delta + q_{\bar{y}}$, this process will indeed end when $\tau = 0$ and we still have $p(\bar{y}) \geq q(\bar{y})$. Finally, remove y from V and if $p(\bar{y}) = q(\bar{y})$, remove $\bar{y}$ from W.

Note that in each iteration, at least one element will be removed from $W \cup V$. Thus the iteration will end after finite steps. When it ends, we have that $\forall y, p(y) = q(y)$. Thus $f(X') = Y$. Also, it is clear from the algorithm that $|X' - X| = \sum_{y \in W} |p(y) - q(y)| \leq \epsilon$. $\square$

We also need the following definition and theorem about non-malleable extractors with weak random seeds.

Definition 4.4. [DLWZ11] A function $\mathsf{nmExt} : [N] \times [D] \to [M]$ is a (k, k', ε)-non-malleable extractor if, for any source X with $H_\infty(X) \geq k$, any seed Y with $H_\infty(Y) \geq k'$, and any function $\mathcal{A} : [D] \to [D]$ such that $\mathcal{A}(y) \neq y$ for all y, the following holds:

$$(\mathsf{nmExt}(X, Y), \mathsf{nmExt}(X, \mathcal{A}(Y)), Y) \approx_\varepsilon (U_{[M]}, \mathsf{nmExt}(X, \mathcal{A}(Y)), Y).$$

A non-malleable extractor with small error will remain to be non-malleable even if the seed is somewhat weak random.

Lemma 4.5. *[DLWZ11] A (k, ϵ)-non-malleable extractor $\mathsf{nmExt} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is also a (k, k', ϵ')-non-malleable extractor with $\epsilon' = 2^{d-k'}\epsilon$.*

Now we can show how non-malleable extractors can be used to construct two-source extractors. In [DW09], Dodis and Wichs showed that non-malleable extractors for (n, k)-sources exist when when $k > 2m + 3\log(1/\varepsilon) + \log d + 9$ and $d > \log(n - k + 1) + 2\log(1/\varepsilon) + 7$. We now show the following theorem.

Theorem 4.6. *Assume that for any $\epsilon > 0$, we have explicit constructions of (k, ϵ)-non-malleable extractors nmExt with seed length $d = 2\log(1/\epsilon) + o(n)$ and output length m. Then there exists a constant $\delta > 0$ and an explicit construction of two source extractors that take as input an $(n, (1/2 - \delta)n)$ source and an independent (n, k) source, and output m bits with error $2^{-\Omega(n)}$.*

Proof. Let Y be an $(n, (1/2 - \delta)n)$ source and X be an independent (n, k) source. We construct the two-source extractor TExt as follows. First we use the somewhere-condenser in Theorem 3.7 to convert Y into a source $\bar{Y} = \mathsf{Scond}(Y)$with two rows $(\bar{Y}_1, \bar{Y}_2)$. By Theorem 3.7, $\bar{Y}$ is $2^{-\Omega(n)}$-close to a somewhere rate-$(1+\alpha)(1/2-\delta)$-source. We choose $\delta > 0$ such that $(1+\alpha)(1/2-\delta) = (1/2+\delta)$. Thus now $\bar{Y}$ is $2^{-\Omega(n)}$-close to a somewhere rate-$(1/2+\delta)$-source. Note that each row of $\bar{Y}$ has $l = \Omega(n)$ bits.

Now let $Y_1 = (\bar{Y}_1 \circ 0)$ and $Y_2 = (\bar{Y}_2 \circ 1)$. The two-source extractor is defined as

$$\mathsf{TExt}(X, Y) = \mathsf{nmExt}(X, Y_1) \oplus \mathsf{nmExt}(X, Y_2).$$

We now show that this is indeed a two-source extractor for X and Y. First note that $\mathsf{Scond}(Y)$ is $2^{-\Omega(n)}$-close to a somewhere rate-$(1/2+\delta)$-source. Lemma 4.3 implies that there exists another source Y' such that $|Y - Y'| \le 2^{-\Omega(n)}$ and $\mathsf{Scond}(Y')$ is a somewhere rate-$(1/2+\delta)$-source. In the following analysis we will treat Y as Y', and this will add at most $2^{-\Omega(n)}$ to the error.

Thus we now have that $\bar{Y} = \mathsf{Scond}(Y)$ is a somewhere rate-$(1/2+\delta)$-source, and we want to show that $\mathsf{TExt}(X, Y)$ is close to uniform. Note that a somewhere rate-$(1/2+\delta)$-source is a convex combination of elementary somewhere rate-$(1/2+\delta)$-sources. Lemma 4.1 now implies that there exist sources $Y^1, \cdots, Y^t$ such that Y is a convex combination of $Y^1, \cdots, Y^t$ and for each Y^i, $\mathsf{Scond}(Y^i)$ is an elementary somewhere rate-$(1/2+\delta)$-source. Thus we only need to show that for each Y^i, $\mathsf{TExt}(X, Y^i)$ is close to uniform. Equivalently and for simplicity, we can assume without loss of generality that $\mathsf{Scond}(Y)$ is an elementary somewhere rate-$(1/2+\delta)$-source.

Now, again without loss of generality we assume that $\bar{Y}_1$ is a $(l, (1/2+\delta)l)$-source. Consider the function $f(Y) = \bar{Y}_1$. Lemma 4.2 implies that there exist flat $(n, (1/2+\delta)l)$ sources $Y^1, \cdots, Y^t$ and bijections $f_1, \cdots, f_t$ such that Y is a convex combination of $Y^1, \cdots, Y^t$, and for each $i \in [t]$, $f(Y^i) = f_i(Y^i)$. Thus we only need to show that for each Y^i, $\mathsf{TExt}(X, Y^i)$ is close to uniform. Consider such a Y^i. Since $f(Y^i) = f_i(Y^i)$ and f_i is a bijection, $\bar{Y}_1^i$ is a $(l, (1/2+\delta)l)$-source. Moreover, we can take g_i to be the inverse function of f_i, and now $Y^i = g_i(\bar{Y}_1^i)$. Note that $\bar{Y}_2^i$ is a deterministic function of Y^i, thus we have that now $\bar{Y}_2^i$ is a deterministic function of $\bar{Y}_1^i$. Finally, note that $Y_1^i \leftrightarrow \bar{Y}_1^i$ and $Y_2^i \leftrightarrow \bar{Y}_2^i$ are both bijections, we thus have the following claim.

Claim 4.7. *Y_1^i is a $(l+1, (1/2+\delta)l)$ source, $Y_2^i = h_i(Y_1^i)$ where h_i is a deterministic function, and $\forall y \in \mathsf{Supp}(Y_1^i), h_i(y) \neq y$.*

In other words, Y_2^i can be viewed exactly as the seed modified by an adversary in a non-malleable extractor. Now since we are using a non-malleable extractor with seed length $d = l+1 = 2\log(1/\epsilon) + o(n) = \Omega(n)$ and the seed has min-entropy $(1/2+\delta)l$, by Lemma 4.5 the error of the non-malleable extractor is

$$2^{d-k'}\epsilon = 2^{l+1-(1/2+\delta)l} 2^{(o(n)-l-1)/2} = 2^{-(\delta l - o(n))} = 2^{-\Omega(n)}.$$

Therefore, we have that $|(\mathsf{nmExt}(X, Y_1^i), \mathsf{nmExt}(X, Y_2^i)) - (U_m, \mathsf{nmExt}(X, Y_2^i))| \le 2^{-\Omega(n)}$. Thus $\mathsf{TExt}(X, Y^i) = \mathsf{nmExt}(X, Y_1^i) \oplus \mathsf{nmExt}(X, Y_2^i)$ is $2^{-\Omega(n)}$-close to uniform. So $\mathsf{TExt}(X, Y)$ is also $2^{-\Omega(n)}$-close to uniform. ■

We can generalize this theorem to work for sources with smaller min-entropy. For this we need (r, k, ϵ)-non-malleable extractors with $r > 1$. We will prove the following theorem in Appendix A.

Theorem 4.8. *For any constant $r \ge 1$, there exists a (r, k, ε)-non-malleable extractor as long as*

$$d > \frac{3}{2}\log(n-k) + 3\log(1/\epsilon) + O(1)$$

$$k > (r+1)m + \frac{d}{3} + 2\log(1/\epsilon) + \log(d) + O(1).$$

We can also define (r, k, ϵ)-non-malleable extractors with weak seed.

Definition 4.9. A function $\{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a (r, k, k', ε)-non-malleable extractor if, for any source X with $H_\infty(X) \ge k$, any seed Y with $H_\infty(Y) \ge k'$, and any r function $\mathcal{A}_i : \{0,1\}^d \to \{0,1\}^d, i = 1, \cdots, r$ such that $\mathcal{A}_i(y) \neq y$ for all i and y, the following holds:

$$(\mathsf{nmExt}(X, Y), \{\mathsf{nmExt}(X, \mathcal{A}_i(Y))\}, Y) \approx_\varepsilon (U_m, \{\mathsf{nmExt}(X, \mathcal{A}_i(Y))\}, Y).$$

Similarly we have the following lemma.

Lemma 4.10. *A (r, k, ε)-non-malleable extractor* $\mathsf{nmExt} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *is also a (r, k, k', ε')-non-malleable extractor with $\varepsilon' = 2^{d-k'}\varepsilon$.*

Proof. For $y \in \{0,1\}^d$, let $\varepsilon_y = \Delta((\mathsf{nmExt}(X, y), \{\mathsf{nmExt}(X, \mathcal{A}_i(y))\}, y), (U_m, \{\mathsf{nmExt}(X, \mathcal{A}_i(y))\}, y))$. Then for Y chosen uniformly from $\{0,1\}^d$,

$$\varepsilon \geq \Delta((\mathsf{nmExt}(X, Y), \{\mathsf{nmExt}(X, \mathcal{A}_i(Y))\}, Y), (U_m, \{\mathsf{nmExt}(X, \mathcal{A}_i(Y))\}, Y)) = \frac{1}{2^d} \sum_{y \in \{0,1\}^d} \varepsilon_y.$$

Thus, for Y' with $H_\infty(Y') \geq k'$, we get

$$\begin{aligned}\Delta((\mathsf{nmExt}(X, Y'), \{\mathsf{nmExt}(X, \mathcal{A}_i(Y'))\}, Y'), (U_m, \{\mathsf{nmExt}(X, \mathcal{A}_i(Y'))\}, Y')) \\ = \sum_{y \in \{0,1\}^d} \Pr[Y = y]\varepsilon_y \leq 2^{-k'} \sum_{y \in \{0,1\}^d} \varepsilon_y \leq 2^{d-k'}\varepsilon.\end{aligned}$$

□

Now we have the following theorem.

Theorem 4.11. *For any constant $b > 2$ and any constant $0 < \delta < 1$, there exists a constant $C = C(\delta) = \mathrm{poly}(1/\delta)$ such that the following holds. Assume that for any $\epsilon > 0$ there exists an explicit construction of (C, k, ϵ)-non-malleable extractors* nmExt *with seed length $d = b\log(1/\epsilon) + o(n)$ and output length m. Then there exists an explicit construction of two source extractors that take as input an $(n, \delta n)$ source and an independent (n, k) source, and output m bits with error $2^{-\Omega(n)}$.*

Proof. Let Y be an $(n, \delta n)$ source and X be an independent (n, k) source. We construct the two-source extractor TExt as follows. First we use the somewhere-condenser in Theorem 3.8 to convert Y into a source $\bar{Y} = \mathsf{Cond}(Y)$ with $D = \mathrm{poly}(1/\delta)$ rows $(\bar{Y}_1, \cdots, \bar{Y}_D)$ such that $\bar{Y}$ is $2^{-\Omega(n)}$-close to a somewhere rate-$\frac{b}{b+1}$-source. Note that each row of $\bar{Y}$ has $l = \Omega(n)$ bits.

Now for each row j we let Y_j be $\bar{Y}_j$ concatenated with the binary expression of $j - 1$. The two-source extractor is defined as

$$\mathsf{TExt}(X, Y) = \bigoplus_j \mathsf{nmExt}(X, Y_j).$$

Let $C = D - 1$. Again, we want to show that $\mathsf{TExt}(X, Y)$ is close to uniform. The proof is similar to the proof in Theorem 4.6. Specifically, we can assume that Y is such that $\mathsf{Cond}(Y)$ is indeed a somewhere rate-$\frac{b}{b+1}$-source. This only adds $2^{-\Omega(n)}$ to the error. Next, we can assume that $\mathsf{Cond}(Y)$ is an an elementary somewhere rate-$\frac{b}{b+1}$-source.

Now, without loss of generality we assume that $\bar{Y}_1$ is a $(l, \frac{b}{b+1}l)$-source. Consider the function $f(Y) = \bar{Y}_1$. Lemma 4.2 implies that there exist flat $(n, \frac{b}{b+1}l)$ sources $Y^1, \cdots, Y^t$ and bijections $f_1, \cdots, f_t$ such that Y is a convex combination of $Y^1, \cdots, Y^t$, and for each $i \in [t]$, $f(Y^i) = f_i(Y^i)$. Thus we only need to show that for each Y^i, $\mathsf{TExt}(X, Y^i)$ is close to uniform. Consider such a Y^i. Since $f(Y^i) = f_i(Y^i)$ and f_i is a bijection, $\bar{Y}_1^i$ is a $(l, \frac{b}{b+1}l)$-source. Moreover, we can take g_i to be the inverse function of f_i, and now $Y^i = g_i(\bar{Y}_1^i)$. Note that for any $j, j \neq 1$, $\bar{Y}_j^i$ is a deterministic function of Y^i, thus we have that now $\bar{Y}_j^i$ is a deterministic function of $\bar{Y}_1^i$. Finally, note that for any j, $Y_j^i \leftrightarrow \bar{Y}_j^i$ is a bijection, we thus have the following claim.

Claim 4.12. *Y_1^i is a $(l+O(1), \frac{b}{b+1}l)$ source, for any $j, j \neq 1$, $Y_j^i = h_{ij}(Y_1^i)$ where h_{ij} is a deterministic function, and $\forall j, j \neq 1, \forall y, h_{ij}(y) \neq y$.*

Note that now we have C modified seeds for the non-malleable extractor. Since we are using a non-malleable extractor with seed length $d = l + O(1) = b\log(1/\epsilon) + o(n) = \Omega(n)$ and the seed has min-entropy $\frac{b}{b+1}l$, by Lemma 4.10 the error of the non-malleable extractor is

$$2^{d-k'}\epsilon = 2^{l+O(1)-\frac{b}{b+1}l}2^{(o(n)-l-O(1))/b} \leq 2^{-(\frac{1}{b(b+1)}l-o(n))} = 2^{-\Omega(n)}.$$

Therefore, we have that $|(\mathsf{nmExt}(X, Y_1^i), \{\mathsf{nmExt}(X, Y_j^i)\}) - (U_m, \{\mathsf{nmExt}(X, Y_j^i)\})| \leq 2^{-\Omega(n)}$. Thus $\mathsf{TExt}(X, Y^i) = \bigoplus_j \mathsf{nmExt}(X, Y_j^i)$ is $2^{-\Omega(n)}$-close to uniform. So $\mathsf{TExt}(X, Y)$ is also $2^{-\Omega(n)}$-close to uniform. ■

5 From Two-Source Extractors to Non-Malleable Extractors

In this section we show how to use a certain kind of two-source extractors to construct non-malleable extractors.

We first define the following encoding of a string $y \in \{0,1\}^d$.

Definition 5.1. Given an integer s, we choose a BCH code with $t = 2$ and $m = s+1$, thus the block length is $n = 2^{s+1} - 1$ and the parity check matrix is a $mt \times n$ matrix. For any $y \in \{0,1\}^s$, let S_Y stand for the integer whose binary expression is y. We encode y to $\bar{y}$ such that $\bar{y}$ is the S_y'th column in the parity check matrix (i.e., $\mathsf{Enc}(y) = \bar{y} = (y, y^3)$ when y is viewed as an element in $\mathbb{F}^*_{2^{s+1}}$).

We have the following theorem.

Theorem 5.2. *Assume that we have a two-source extractor $\mathsf{TExt} = \mathsf{IP}(f(X), W)$ such that when given an (n_1, k)-source X and an independent $(n_2, n_2/2 - \ell)$-source W, TExt outputs 1 bit with error ϵ. Let $n_2' = \lfloor \frac{n_2}{2} \rfloor - 1$ and let Y be the uniform distribution over $\{0,1\}^{n_2'}$. Define a seeded extractor*

$$\mathsf{nmExt}(X, Y) = \mathsf{IP}(f(X), \mathsf{Enc}(Y)).$$

Then nmExt is a (k, ϵ')-non-malleable extractor with error $\epsilon' = O(2^{-\ell} + \epsilon)$.

Proof. First let Y' be a source over $\{0,1\}^{n_2'}$ with min-entropy $n_2/2 - \ell + 1$. Let $\mathcal{A} : \{0,1\}^{n_2'} \to \{0,1\}^{n_2'}$ be any deterministic function such that $\forall y, \mathcal{A}(y) \neq y$.

Note that the BCH code has distance $2t + 1 = 5 > 4$, thus any 4 columns in the parity check matrix must be linearly independent. This in particular implies that every two different columns must be different. Thus $\mathsf{Enc}(Y') = \bar{Y'}$ has min-entropy $n_2/2 - \ell + 1$. Therefore by the assumption we have that

$$\mathsf{nmExt}(X, Y') \approx_\epsilon U.$$

Next, note that

$$\begin{aligned}\mathsf{nmExt}(X, Y') \oplus \mathsf{nmExt}(X, \mathcal{A}(Y')) &= \mathsf{IP}(f(X), \mathsf{Enc}(Y')) \oplus \mathsf{IP}(f(X), \mathsf{Enc}(\mathcal{A}(Y'))) \\ &= \mathsf{IP}(f(x), \overline{Y'} + \overline{\mathcal{A}(Y')}).\end{aligned}$$

For two different y_1, y_2, if $\overline{y_1} + \overline{\mathcal{A}(y_1)} = \overline{y_2} + \overline{\mathcal{A}(y_2)}$, then $\overline{y_1}, \overline{\mathcal{A}(y_1)}, \overline{y_2}, \overline{\mathcal{A}(y_2)}$ are linearly dependent. Note that $\overline{\mathcal{A}(y_1)} = \mathsf{Enc}(\mathcal{A}(y_1))$ and $\overline{\mathcal{A}(y_2)} = \mathsf{Enc}(\mathcal{A}(y_2))$ are also some columns of the parity check matrix. Since $\mathcal{A}(y_1) \neq y_1$ and $\mathcal{A}(y_2) \neq y_2$, we have that $\overline{\mathcal{A}(y_1)} \neq \overline{y_1}$ and $\overline{\mathcal{A}(y_2)} \neq \overline{y_2}$. Thus we must have $\overline{\mathcal{A}(y_1)} = \overline{y_2}$ and $\overline{\mathcal{A}(y_2)} = \overline{y_1}$.

Therefore, the min-entropy of $\overline{Y'} + \overline{\mathcal{A}(Y')}$ is at least $n_2/2 - \ell$ since the probability of getting any particular element in the support is at most $2 \cdot 2^{-(n_2/2-\ell+1)} = 2^{-(n_2/2-\ell)}$. Thus by the assumption we have

$$\mathsf{nmExt}(X, Y') \oplus \mathsf{nmExt}(X, \mathcal{A}(Y')) \approx_\epsilon U.$$

Thus by the non-uniform XOR lemma, Lemma 3.13, we have

$$|(\mathsf{nmExt}(X, Y'), \mathsf{nmExt}(X, \mathcal{A}(Y'))) - (U, \mathsf{nmExt}(X, \mathcal{A}(Y')))| \leq 2\epsilon.$$

Now note that Y has min-entropy $n_2' = \lfloor \frac{n_2}{2} \rfloor - 1$, thus by Theorem 3.17,

$$|(\mathsf{nmExt}(X, Y), \mathsf{nmExt}(X, \mathcal{A}(Y)), Y) - (U, \mathsf{nmExt}(X, \mathcal{A}(Y)), Y)| \leq 2^2(2^{-(\ell-4)} + 2\epsilon) = O(2^{-\ell} + \epsilon).$$

■

Similarly, we can generalize the above theorem to the case of (r, k, ε)-non-malleable extractors. We have the following definition and theorem.

Definition 5.3. Given two integers r, s, we choose a BCH code with $t = r + 1$ and $m = s + 1$, thus the block length is $n = 2^{s+1} - 1$ and the parity check matrix is a $mt \times n$ matrix. For any $y \in \{0,1\}^s$, let S_Y stand for the integer whose binary expression is y. We encode y to $\bar{y}$ such that $\bar{y}$ is the S_y'th column in the parity check matrix (i.e., $\mathsf{Enc}_r(y) = \bar{y} = (y, y^3, \cdots, y^{2r+1})$ when y is viewed as an element in $\mathbb{F}^*_{2^{s+1}}$).

Theorem 5.4. *Given two integers r, ℓ such that $\ell > r$. Assume that we have a two-source extractor $\mathsf{TExt} = \mathsf{IP}(f(X), W)$ such that when given an (n_1, k)-source X and an independent $(n_2, n_2/(r+1) - \ell)$-source W, TExt outputs 1 bit with error ϵ. Let $n_2' = \lfloor \frac{n_2}{r+1} \rfloor - 1$ and let Y be the uniform distribution over $\{0,1\}^{n_2'}$. Define a seeded extractor*

$$\mathsf{nmExt}(X, Y) = \mathsf{IP}(f(X), \mathsf{Enc}_r(Y)).$$

Then nmExt is a (r, k, ϵ')-non-malleable extractor with error $\epsilon' = O(r2^{r-\ell} + 2^{\frac{3r}{2}}\epsilon)$.

Proof. First let Y' be a source over $\{0,1\}^{n_2'}$ with min-entropy $n_2/(r+1) - \ell + \log(r+1)$. Let $\mathcal{A}_i : \{0,1\}^{n_2'} \to \{0,1\}^{n_2'}, i = 1, \cdots, r$ be r deterministic function such that for any i, $\forall y, \mathcal{A}_i(y) \neq y$.

Note that the BCH code has distance $2t + 1 = 2r + 3 > 2r + 2$, thus any $2r + 2$ columns in the parity check matrix must be linearly independent. This in particular implies that every two different columns must be different. Thus $\mathsf{Enc}(Y') = \bar{Y}'$ has min-entropy $n_2/(r+1) - \ell + \log(r+1)$. Therefore by the assumption we have that

$$\mathsf{nmExt}(X, Y') \approx_\epsilon U.$$

Next, choose any non-empty subset $S \subseteq [r]$, note that

$$\mathsf{nmExt}(X,Y') \oplus \bigoplus_{i\in S} \mathsf{nmExt}(X,\mathcal{A}_i(Y')) = \mathsf{IP}(f(X),\mathsf{Enc}(Y')) \oplus \bigoplus_{i\in S} \mathsf{IP}(f(X),\mathsf{Enc}(\mathcal{A}_i(Y')))$$
$$= \mathsf{IP}(f(x), \overline{Y'} + \sum_{i\in S} \overline{\mathcal{A}_i(Y')}).$$

For two different y_1, y_2, if $\overline{y_1}+\sum_{i\in S}\overline{\mathcal{A}_i(y_1)} = \overline{y_2}+\sum_{i\in S}\overline{\mathcal{A}_i(y_2)}$, then $\overline{y_1}, \{\overline{\mathcal{A}_i(y_1)}, i\in S\}, \overline{y_2}, \{\overline{\mathcal{A}_i(y_2)}, i\in S\}$ are linearly dependent. Without loss of generality we assume that all $\mathcal{A}_i(y_1)$ are different and all $\mathcal{A}_i(y_2)$ are different, since if not then some of them sum to 0 and this only decreases the number of items. Note that $\overline{\mathcal{A}_i(y_1)} = \mathsf{Enc}_r(\mathcal{A}_i(y_1))$ and $\overline{\mathcal{A}_i(y_2)} = \mathsf{Enc}_r(\mathcal{A}_i(y_2))$ are also some columns of the parity check matrix. Since the columns are $2r+2$-wise linearly independent, and the total number of items here is at most $2r+2$, we must have that the items in $\{\overline{y_1}, \{\overline{\mathcal{A}_i(y_1)}, i\in S\}\}$ and the items in $\{\overline{y_2}, \{\overline{\mathcal{A}_i(y_2)}, i\in S\}\}$ form a perfect matching, where an edge in the matching is of the form $\overline{y_1} = \overline{\mathcal{A}_j(y_2)}$, $\overline{\mathcal{A}_i(y_1)} = \overline{y_2}$ or $\overline{\mathcal{A}_i(y_1)} = \overline{\mathcal{A}_j(y_2)}$.

Now we claim that for any y, there are at most r different y_j's such that $\overline{y} + \sum_{i\in S}\overline{\mathcal{A}_i(y)} = \overline{y_j}+\sum_{i\in S}\overline{\mathcal{A}_i(y_j)}$. To see this, assume for the sake of contradiction that there are $r+1$ such different y_j's. Then by the above discussion we see that for each j, there exists an $i\in S$ such that $\overline{y_j} = \overline{\mathcal{A}_i(y)}$. Since $|S|\le r$ we must have two different y_j and y_l and an $i\in S$ such that $\overline{y_j} = \overline{\mathcal{A}_i(y)} = \overline{y_l}$. Note that Enc_r is injective, thus we have $y_j = y_l$, a contradiction.

Therefore, the min-entropy of $\overline{Y'}+\sum_{i\in S}\overline{\mathcal{A}_i(Y')}$ is at least $n_2/(r+1)-\ell+\log(r+1)-\log(r+1) = n_2/(r+1)-\ell$. Therefore by the assumption we have that

$$\mathsf{nmExt}(X,Y') \oplus \bigoplus_{i\in S} \mathsf{nmExt}(X,\mathcal{A}_i(Y')) \approx_\epsilon U.$$

Thus by the non-uniform XOR lemma, Lemma 3.15, we have

$$|(\mathsf{nmExt}(X,Y'), \{\mathsf{nmExt}(X,\mathcal{A}_i(Y'))\}) - (U, \{\mathsf{nmExt}(X,\mathcal{A}_i(Y'))\})| \le 2^{\frac{r}{2}}\epsilon.$$

Now note that Y has min-entropy $n_2' = \lfloor \frac{n_2}{r+1} \rfloor - 1$, thus by Theorem 3.17,

$$|(\mathsf{nmExt}(X,Y), \{\mathsf{nmExt}(X,\mathcal{A}_i(Y))\}, Y) - (U, \mathsf{nmExt}(X, \{\mathsf{nmExt}(X,\mathcal{A}_i(Y))\}, Y)|$$
$$\le 2^{r+1}(2^{-(\ell-\log(r+1)-3)} + 2^{\frac{r}{2}}\epsilon) = O(r2^{r-\ell} + 2^{\frac{3r}{2}}\epsilon).$$

■

6 Improved Constructions of Non-Malleable Extractors

Now we can use the above theorems to construct new non-malleable extractors. As a warm up, we first give a new construction of non-malleable extractors for min-entropy rate $> 1/2$.

6.1 A Non-Malleable Extractor for Entropy Rate $> 1/2$

For this purpose, simply notice that the inner product function itself is a two-source extractor for an $(n, (1/2+\delta)n)$ source and another independent source on n bits with min-entropy slightly below $n/2$. Specifically, we have

Theorem 6.1. *[CG88, Vaz85] For every constant $\delta > 0$, if X is an (n, k_1) source, Y is an independent (n, k_2) source and $k_1 + k_2 \geq (1 + \delta)n$, then*

$$\mathsf{IP}(X, Y) \approx_\epsilon U$$

with $\epsilon = 2^{-\Omega(\delta n)}$.

Thus we can use the following construction. Given an (n, k)-source X with $k = (1/2+\delta)n$, take an independent uniform seed $Y \in \{0,1\}^{n/2-1}$ and encode Y to $\bar{Y}$ such that $\bar{Y} = \mathsf{Enc}(Y) = (Y, Y^3)$ when Y is viewed as an element in $\mathbb{F}^*_{2^{n/2}}$. Our non-malleable extractor is now defined as

$$\mathsf{nmExt}(X, Y) = \mathsf{IP}(X, \mathsf{Enc}(Y)) = \mathsf{IP}(X, \bar{Y}),$$

where IP is the inner product function over $\mathbb{F}_2$.

Theorem 6.2. *For any constant $\delta > 0$, the function nmExt defined as above is a $((1/2+\delta)n, 2^{-\Omega(\delta n)})$ non-malleable extractor.*

Proof. By Theorem 6.1, $\mathsf{IP}(X, Y)$ is a two-source extractor for an $(n, (1/2 + \delta)n)$ source and an independent $(n, (1/2 - \delta/2)n)$ source with error $2^{-\Omega(\delta n)}$. Thus by Theorem 5.2, nmExt is a $((1/2 + \delta)n, \epsilon)$ non-malleable extractor with $\epsilon = O(2^{-\delta n/2} + 2^{-\Omega(\delta n)}) = 2^{-\Omega(\delta n)}$. ∎

6.2 Non-Malleable Extractors for Entropy Rate $< 1/2$

In this section we give one of our main constructions, namely a non-malleable extractor for weak sources with min-entropy rate $1/2 - \delta$ for some universal constant $\delta > 0$. We have the following construction.

Given an (n, k)-source X with $k = (1/2 - \delta)n$, we first pick a prime p that is close to n. By Bertrand's postulate and [BHP01], there exists $n_0 \in \mathbb{N}$ such that for every $n \geq n_0$, there exists a prime between n and $n + O(n^{0.525})$. We will pick a prime p in this range. Note that the prime can be found in polynomial time in n. Take the field $\mathbb{F}_q$ where $q = 2^p$ and let g be a generator in $\mathbb{F}^*_q$. The construction is as follows.

- Treat X as an element in $\mathbb{F}^*_q$ and encode X such that $\mathsf{Enc}(X) = (X, g^X)$.
- Take an independent and uniform seed $Y \in \{0,1\}^{p-1}$ and encode Y to $\bar{Y}$ such that $\bar{Y} = (Y, Y^3)$ when Y is viewed as an element in $\mathbb{F}^*_{2^p}$.
- Output $\mathsf{nmExt}(X, Y) = \mathsf{IP}(\mathsf{Enc}(X), \bar{Y})$ where IP is the inner product function over $\mathbb{F}_2$.

To prove our construction is a non-malleable extractor, we are going to use Theorem 5.2. To this end, we first prove the following lemma.

Lemma 6.3. *There exists a constant $\delta > 0$ such that for any (n, k)-source X with $k = (1/2 - \delta)n$, and any independent $(2p, k_2)$ source Y with $k_2 \geq (1 - \delta)p$,*

$$|\mathsf{IP}(\mathsf{Enc}(X), Y) - U| \leq \epsilon,$$

where $\epsilon = 2^{-\Omega(n)}$.

Proof. We think of X as a distribution in $\mathbb{F}_q^*$ that has min-entropy k. This increases the error by at most 2^{-k} (for the element 0). By the XOR lemma, we only need to show that for the only non-trivial character ψ (since we only output 1 bit),

$$|E_{X,Y}[\psi(\mathsf{IP}(\mathsf{Enc}(X), Y))]| \leq 2^{-\Omega(n)}.$$

Let $X' = 4\mathsf{Enc}(X) - 4\mathsf{Enc}(X)$, by Lemma 3.20 we have

$$|E_{X,Y}[\psi(\mathsf{IP}(\mathsf{Enc}(X), Y))]| \leq |E_{X',Y}[\psi(X' \cdot Y)]|^{\frac{1}{8}}.$$

We next bound $|E_{X',Y}[\psi(X' \cdot Y)]|$. First we show that X' is close to a source with min-entropy rate $> 1/2$. We have the following claim.

Claim 6.4. *There is a universal constant $\delta > 0$ such that if X is any weak source with min-entropy $(1/2 - \delta)n$, $3\mathsf{Enc}(X)$ is $2^{-\Omega(n)}$-close to a source with min-entropy $(1/2 + \delta)(2p)$.*

Proof of the claim. Note that $k = (1/2 - \delta)n$ and p is between n and $n(1 + \frac{1}{2\ln^2 n})$. Thus for sufficiently large n we have that $k \geq (1/2 - 1.01\delta)p$. Note that we choose the field $\mathbb{F}_q$ where $q = 2^p$. Thus the sum of $\mathsf{Enc}(X) + \mathsf{Enc}(X)$ when viewing $\mathsf{Enc}(X)$ as a vector in $\mathbb{F}_2^{2p}$ is the same as when viewing $\mathsf{Enc}(X)$ as a vector in $\mathbb{F}_q^2$. In the following we will view $\mathsf{Enc}(X)$ as a vector in $\mathbb{F}_q^2$. We show that $3\mathsf{Enc}(X)$ has a larger min-entropy rate.

First consider the distribution $2\mathsf{Enc}(X)$. Note that the distribution is of the form $(X + X, g^X + g^X)$. Let $\bar{X} = g^X$ and note that g^x is a bijection in F_q^*. Thus $\bar{X}$ has the same min-entropy as X. Now the support of $2\mathsf{Enc}(X)$ is of the form $(\log_g(\bar{x}_1\bar{x}_2), \bar{x}_1 + \bar{x}_2)$. For any (b, a) in this support, we have that $\bar{x}_1\bar{x}_2 = g^b$ and $\bar{x}_1 + \bar{x}_2 = a$. Thus there are at most 2 different pairs of $(\bar{x}_1, \bar{x}_2)$ that satisfy both equations. Therefore the min-entropy of $2\mathsf{Enc}(X)$ is at least $2H_\infty(X) - 1$. We can also assume that $a \neq 0$ since this only increases the error by at most $2^{-H_\infty(X)}$. Now let $k = H_\infty(X) - 1$, we have that $\mathsf{Enc}(X)$ has min-entropy at least k and $2\mathsf{Enc}(X)$ has min-entropy at least $2k$.

Now consider $3\mathsf{Enc}(X)$. Every element in the support of $3\mathsf{Enc}(X)$ has the form $(\log_g(\bar{x}_1\bar{x}_2\bar{x}_3), \bar{x}_1 + \bar{x}_2 + \bar{x}_3)$, which determines the point $(\bar{x}_1\bar{x}_2\bar{x}_3, \bar{x}_1 + \bar{x}_2 + \bar{x}_3)$. Let $a = \bar{x}_1 + \bar{x}_2$ and $b = \bar{x}_1\bar{x}_2$, this point is

$$(b\bar{x}_3, a + \bar{x}_3).$$

Let $\tilde{x}_3 = a + \bar{x}_3$, then

$$(a + \bar{x}_3, b\bar{x}_3) = (\tilde{x}_3, b\tilde{x}_3 - ab).$$

For a fixed $(a = \bar{x}_1 + \bar{x}_2, b = \bar{x}_1\bar{x}_2)$ define the line

$$\ell_{a,b} = \{(x, bx - ab) | x \in \mathbb{F}_q\}.$$

Thus we have a set of lines $L = \{\ell_{a,b}\}$. Note that $a \neq 0$ and $b \neq 0$. Thus for different (a, b), the line $\ell_{a,b}$ is also different. Note that x_3 is sampled from X_3, which has min-entropy k and (a, b) is sampled from $\mathsf{Enc}(X_1) + \mathsf{Enc}(X_2)$, which has min-entropy $2k$. Further note that these two distributions are independent. Since every weak source with min-entropy k is a convex combination of flat k sources, without loss of generality we can assume that X_3 and $\mathsf{Enc}(X_1) + \mathsf{Enc}(X_2)$ are both flat sources. Thus L has size 2^{2k}.

Now let α, β be the two constants in Theorem 3.21. Assume that $3\mathsf{Enc}(X)$ is ϵ-far from any source with min-entropy $(1+\alpha/2)2k$. Since $3\mathsf{Enc}(X)$ determines the distribution $(A+\bar{X}_3, B\bar{X}_3)$, this distribution is also ϵ-far from any source with min-entropy $(1+\alpha/2)2k$. Thus there must exist some set M of size at most $2^{(1+\alpha/2)2k}$ such that

$$\Pr_{(a,b)\leftarrow 2\mathsf{Enc}(X), x_3 \leftarrow X}[(a+\bar{x}_3, b\bar{x}_3) \in M] \geq \epsilon.$$

Note that whenever $(a+\bar{x}_3, b\bar{x}_3) \in M$, this point has an incidence with the line $\ell_{a,b}$. Further note that whenever (a,b) is different or x_3 is different, the incidence is also different. Thus by the above inequality the number of incidences between the set of points M and the set of lines L is at least

$$\Pr_{(a,b)\leftarrow 2\mathsf{Enc}(X), x_3 \leftarrow X}[(a+\bar{x}_3, b\bar{x}_3) \in M]2^k 2^{2k} \geq \epsilon 2^{3k}.$$

On the other hand, since L has size 2^{2k} and M has size $2^{(1+\alpha/2)2k} \leq 2^{(1+\alpha/2)2(1/2-\delta)p} < 2^{(1+\alpha/2)p} \leq q^{2-\beta}$, by Theorem 3.21, the number of incidences between M and L is at most $O(2^{(3/2-\alpha)(2+\alpha)k}) < 2^{3k(1-\alpha/6)} = 2^{-\alpha k/2}2^{3k}$.

Thus we must have $\epsilon < 2^{-\alpha k/2}$.

Thus we have shown that $3\mathsf{Enc}(X)$ is $2^{-\alpha k/2}$-close to having min-entropy $(1+\alpha/2)2k$. By choosing δ appropriately, we get that $3\mathsf{Enc}(X)$ is $2^{-\Omega(n)}$-close to having min-entropy $(1/2+\delta)2p$. □

Now note that Y is a weak source over $\{0,1\}^{2p}$ with min-entropy $k_2 \geq (1-\delta)p$. Also note that the min-entropy of X' is at least the min-entropy of $3\mathsf{Enc}(X)$. Thus by Lemma 3.19 we have that

$$|E_{X',Y}[\psi(X' \cdot Y)]| \leq 2^{2p}2^{-(1/2+\delta)2p}2^{-(1-\delta)p} + 2^{-\Omega(n)} = 2^{-\Omega(n)}.$$

Therefore

$$|E_{X,Y}[\psi(\mathsf{IP}(\mathsf{Enc}(X), Y))]| \leq 2^{-\Omega(n)}.$$

□

Now we can prove our construction is a non-malleable extractor.

Theorem 6.5. *For any (n,k)-source X with $k = (1/2-\delta)n$, the function* nmExt *defined above is a (k,ϵ)-non-malleable extractor with $\epsilon = 2^{-\Omega(n)}$.*

Proof. By Lemma 6.3, $\mathsf{IP}(\mathsf{Enc}(X), Y)$ is a two-source extractor for an (n,k) source and an independent $(p,(1-\delta)p)$ source with error $2^{-\Omega(n)}$. Therefore by Theorem 5.2, nmExt is a (k,ϵ)-non-malleable extractor with error $\epsilon = O(2^{-\delta p} + 2^{-\Omega(n)}) = 2^{-\Omega(n)}$. ■

6.3 Achieving Even Smaller Min-Entropy

In this section we show that we can construct non-malleable extractors for even smaller min-entropy rate (potentially any constant arbitrarily close to 0), if we assume that we have affine extractors with large enough output size, and the Approximate Duality Conjecture (or the Polynomial Freiman-Ruzsa Conjecture) as in [BSZ11].

Recall the definition of an affine extractor.

Definition 6.6. An $[n, m, \rho, \epsilon]$ affine extractor is a deterministic function $f : \{0,1\}^n \to \{0,1\}^m$ such that whenever X is the uniform distribution over some affine subspace over $\mathbb{F}_2^n$ with dimension ρn, we have that for every $z \in \{0,1\}^m$,

$$|\Pr[f(X) = z] - 2^{-m}| < \epsilon.$$

Now we define the duality measure of two sets as in [BSZ11].

Definition 6.7. [BSZ11] Given two sets $A, B \subseteq \mathbb{F}_2^n$, their duality measure is defined as

$$\mu^{\perp}(A, B) = \left| E_{a \in A, b \in B}[(-1)^{\langle a,b \rangle}] \right|.$$

The following conjecture is introduced in [BSZ11] and is shown in that paper to be implied by the well-known Polynomial Freiman-Ruzsa Conjecture in additive combinatorics.

Conjecture 6.8. *(Approximate Duality (ADC)) [BSZ11] For every pair of constants $\alpha, \delta > 0$ there exist a constant $\zeta > 0$ and an integer r, both depending on α and δ such that the following holds for sufficiently large n. If $A, B \subseteq \mathbb{F}_2^n$ satisfy $|A|, |B| \geq 2^{\alpha n}$ and $\mu^{\perp}(A, B) \geq 2^{-\zeta n}$, then there exists a pair of subsets*

$$A' \subseteq A, A' \geq \frac{|A|}{2^{\delta n+1}} \text{ and } B' \subseteq B, |B'| \geq \left(\frac{\mu^{\perp}(A,B)}{2}\right)^r \cdot \frac{|B|}{2^{\delta n}}$$

such that $\mu^{\perp}(A', B') = 1$.

We now have the following construction.

Construction 6.9. Given any (n, k) source X and a constant $0 < \lambda < 1$, let $f : \{0,1\}^{n'} \to \{0,1\}^{m'}$ be an $[n', m' = (1-\lambda)\frac{2}{3}n', \frac{2}{3}, 2^{-m'}]$ affine extractor such that $n = n' - m'$. For any $z \in \{0,1\}^{m'}$, let $f^{-1}(z) = \{x : f(x) = z\}$. Then there exists $z \in \{0,1\}^{m'}$ such that $|f^{-1}(z)| \geq 2^n$. Let $F : \{0,1\}^n \to f^{-1}(z)$ be (any) injective map. Now take an independent uniform seed $Y \in \{0,1\}^{n'/2-1}$ and encode Y to $\bar{Y}$ such that $\bar{Y} = \mathsf{Enc}(Y) = (Y, Y^3)$ when Y is viewed as an element in $\mathbb{F}^*_{2^{n'/2}}$. Our non-malleable extractor is now defined as

$$\mathsf{nmExt}(X, Y) = \mathsf{IP}(F(X), \mathsf{Enc}(Y)) = \mathsf{IP}(F(X), \bar{Y}),$$

where IP is the inner product function taken over $\mathbb{F}_2$.

Remark 6.10. Note that here the function F may not be efficiently computable (in time $\mathrm{poly}(n)$). However, the time to compute F is polynomial in the length of the truth table of our final extractor.

Again, we will show that our construction is a non-malleable extractor by using Theorem 5.2. To this end, we first show the lemma.

Lemma 6.11. *For any (n, k) source X with $k = \frac{2.5\lambda}{1+2\lambda}n$ and any independent $(n', \frac{n'}{3}+1)$ source Y, $\mathsf{IP}(F(X), Y)$ is non-constant.*

Proof. As usual we can assume without loss of generality that X and Y are flat sources. If $\mathsf{IP}(F(X), Y)$ is a constant, then $\mathsf{Supp}(F(X))$ and $\mathsf{Supp}(Y)$ must be contained in two affine subspaces with dimension d_1, d_2 such that $d_1 + d_2 \leq n'$. Note that $d_2 > \frac{n'}{3}$ since Y has min-entropy $\frac{n'}{3} + 1$. We next show that $d_1 > \frac{2}{3}n'$ and thus reach a contradiction.

To see this, let $S = \mathsf{Supp}(F(X))$. It suffices to show that S is not contained in any affine subspace of dimension $\frac{2}{3}n'$. Let A be such an affine subspace. We have

$$|A \cap S| \leq |A \cap f^{-1}(z)| < 2 \cdot 2^{-m'} 2^{\frac{2}{3}n'} = 2^{\frac{2\lambda}{3}n'+1},$$

where the last inequality follows from the fact that f is an affine extractor. Now note that $|S| = 2^{\frac{2.5\lambda}{1+2\lambda}n} = 2^{\frac{2.5\lambda}{3}n'}$. Thus we have that $|A \cap S| < |S|$ and therefore S cannot be contained in A. $\square$

Now we have the following lemma.

Lemma 6.12. *There exists a constant $\zeta = \zeta(\lambda)$ such that for any (n, k) source X with $k = \frac{3\lambda}{1+2\lambda}n$ and any independent $(n', \frac{7n'}{15})$ source Y, $\mathsf{IP}(F(X), Y)$ is $2^{-\zeta n}$-close to uniform.*

Proof. Let $v = \frac{n}{n'} = \frac{1+2\lambda}{3}$, $\alpha = min\{\lambda, \frac{1}{3}\}$ and $\delta = \frac{\lambda}{8}$. Let ζ' and r be the constant and the integer guaranteed by conjecture 6.8 for α and δ. Let $\zeta = min\{\frac{\zeta'}{v}, \frac{1-\lambda}{8r}\}$. We will prove the lemma by way of contradiction.

Let X and Y be two independent sources as in the statement of the lemma. Again we assume without loss of generality that both X and Y are flat sources. Let $A = \mathsf{Supp}(X)$ and $\bar{A} = \{F(a)|a \in A\} \subseteq \mathbb{F}_2^{n'}$. Let $\bar{B} = \mathsf{Supp}(Y) \subseteq \mathbb{F}_2^{n'}$. Note that F is an injective function. Thus $|\bar{A}| = 2^{\frac{3\lambda}{1+2\lambda}n} = 2^{\lambda n'} \geq 2^{\alpha n'}$ and $|\bar{B}| = 2^{\frac{7n'}{15}} > 2^{\frac{n'}{3}} \geq 2^{\alpha n'}$.

Assume for the sake of contradiction that the error of $\mathsf{IP}(F(X), Y)$, which is equal to $\frac{1}{2}\mu^{\perp}(\bar{A}, \bar{B})$, is greater than $2^{-\zeta n} \geq 2^{-\zeta' n'}$. Then by the ADC conjecture (conjecture 6.8) there exist $A' \subseteq \bar{A}$ and $B' \subseteq \bar{B}$ such that

$$|A'| \geq \frac{|\bar{A}|}{2^{\delta n'+1}} \geq 2^{\frac{5\lambda}{6}n'} = 2^{\frac{2.5\lambda}{1+2\lambda}n} \text{ and } |B'| \geq \frac{|\bar{B}|}{2^{\delta n' + r\zeta n}} \geq \frac{2^{\frac{7n'}{15}}}{2^{\frac{n'}{8}}} > 2^{\frac{n'}{3}+1},$$

and $\mu^{\perp}(A', B') = 1$.

Let A'' be the preimages of A' under F. Since F is injective, we must have $|A''| \geq 2^{\frac{2.5\lambda}{1+2\lambda}n}$. Thus if we let X' and Y' be the uniform distribution over A'' and B' respectively, we get two independent sources that satisfy the conditions in Lemma 6.11. However $\mathsf{IP}(F(X'), Y')$ is a constant, which contradicts Lemma 6.11. Thus we must have that $\mathsf{IP}(F(X), Y)$ is $2^{-\zeta n}$-close to uniform. $\square$

Now we can prove the following theorem.

Theorem 6.13. nmExt *is a (k, ϵ)-non-malleable extractor with $k = \frac{3\lambda}{1+2\lambda}n$, seed length $d = \frac{3}{2+4\lambda}n - 1$ and $\epsilon = 2^{-\Omega(n)}$.*

Proof. The seed length is clearly $d = n'/2 - 1 = \frac{3}{2+4\lambda}n - 1$. Let $\zeta = \zeta(\lambda)$ be as in Lemma 6.12. By Lemma 6.12, $\mathsf{IP}(F(X), Y)$ is a two-source extractor for an $(n, k = \frac{3\lambda}{1+2\lambda}n)$ source and an independent $(n', \frac{7n'}{15})$ source with error $2^{-\zeta n}$. Thus by Theorem 5.2, nmExt is a (k, ϵ)-non-malleable extractor with error $\epsilon = O(2^{-n'/30} + 2^{-\zeta n}) = 2^{-\Omega(n)}$. ■

Similarly, we can generalize our construction to (r, k, ϵ)-non-malleable extractors. We have the following construction and theorem.

Construction 6.14. Given a constant integer r, any (n,k) source X and a constant $0<\lambda<1$, let $f:\{0,1\}^{n'}\to\{0,1\}^{m'}$ be an $[n', m'=(1-\lambda)\frac{r+1}{r+2}n', \frac{r+1}{r+2}, 2^{-m'}]$ affine extractor such that $n=n'-m'$. For any $z\in\{0,1\}^{m'}$, let $f^{-1}(z)=\{x: f(x)=z\}$. Then there exists $z\in\{0,1\}^{m'}$ such that $|f^{-1}(z)|\geq 2^n$. Let $F:\{0,1\}^n\to f^{-1}(z)$ be (any) injective map. Now take an independent uniform seed $Y\in\{0,1\}^{n'/(r+1)-1}$ and encode Y to $\bar{Y}$ such that $\bar{Y}=\mathsf{Enc}_r(Y)=(Y,Y^3,\cdots,Y^{2r+1})$ when Y is viewed as an element in $\mathbb{F}^*_{2^{n'/(r+1)}}$. Define a seeded extractor

$$\mathsf{nmExt}(X,Y)=\mathsf{IP}(F(X),\mathsf{Enc}_r(Y))=\mathsf{IP}(F(X),\bar{Y}),$$

where IP is the inner product function taken over $\mathbb{F}_2$.

Theorem 6.15. nmExt *is a* (r,k,ϵ)*-non-malleable extractor with* $k=\frac{(r+2)\lambda}{1+(r+1)\lambda}n$*, seed length* $d=\frac{r+2}{r+1+(r+1)^2\lambda}n-1$ *and* $\epsilon=2^{-\Omega(n)}$.

By using Theorem 5.4 instead of Theorem 5.2, the proof of this theorem is very similar to the proof of Theorem 6.13. We thus omit the proof here.

6.4 Increasing the Output Size and Reducing the Seed Length

In this section we show that we can increase the output size and reduce the seed length for the constructions in Subsection 6.1, Subsection 6.2 and Subsection 6.3. All these constructions share the same pattern: the seed Y is encoded using the parity check matrix of a BCH code, and then the output is the inner product function of the encoded source and the encoded seed over $\mathbb{F}_2$.

We only discuss the construction in Subsection 6.2, but the method can be applied to all the other constructions in the same way. We start by showing how to increase the output size to $m=\Omega(n)$.

6.4.1 Increasing the output size

Recall that in the construction we used a field $\mathbb{F}_{2^p}$ for a prime p. Given the finite filed $\mathbb{F}_{2^p}$, the elements of this field form a vector space of dimension p over $\mathbb{F}_2$. Let $b_1,\cdots,b_p\in\mathbb{F}_{2^p}$ be a basis for this vector space. Now recall that in the construction we encode the seed Y to $\bar{Y}=(Y,Y^3)$, when viewing Y as an element in $\mathbb{F}^*_{2^p}$. Now for each b_i, let $\bar{Y^i}=(b_iY,b_iY^3)$ and define one bit $Z_i=\mathsf{IP}(\mathsf{Enc}(X),\bar{Y^i})$. We now show that $\{Z_i\}$ satisfy the conditions of a non-uniform XOR lemma.

Lemma 6.16. *Given any* (n,k)*-source* X *with* $k=(1/2-\delta)n$ *and an independent seed* $Y\in\{0,1\}^{p-1}$ *with min-entropy* $(1-\delta)p+2$*, let* $\mathcal{A}:\{0,1\}^{p-1}\to\{0,1\}^{p-1}$ *be any deterministic function such that* $\forall y,\mathcal{A}(y)\neq y$*. For any* i*, let* $Z'_i=\mathsf{IP}(\mathsf{Enc}(X),\bar{Y^{i'}})$*, where* $\bar{Y^{i'}}=(b_iY',b_iY'^3)$ *and* $Y'=\mathcal{A}(Y)$*. Then for any non-empty subset* $S_1\subseteq[p]$ *and any subset* $S_2\subseteq[p]$*, we have that*

$$|\bigoplus_{i\in S_1}Z_i\oplus\bigoplus_{j\in S_2}Z'_j-U|\leq 2^{-\Omega(n)}.$$

Proof. Note that

$$\bigoplus_{i\in S_1}Z_i=\mathsf{IP}(\mathsf{Enc}(X),\sum_{i\in S_1}\bar{Y^i})=\mathsf{IP}(\mathsf{Enc}(X),t_1(Y,Y^3)),$$

where $t_1=\sum_{i\in S_1}b_i\in\mathbb{F}_{2^p}$, and

$$\bigoplus_{j \in S_2} Z'_j = \mathsf{IP}(\mathsf{Enc}(X), \sum_{j \in S_2} \bar{Y^j}') = \mathsf{IP}(\mathsf{Enc}(X), t_2(Y', Y'^3)),$$

where $t_2 = \sum_{j \in S_2} b_j \in \mathbb{F}_{2^p}$.

Thus

$$\bigoplus_{i \in S_1} Z_i \oplus \bigoplus_{j \in S_2} Z'_j = \mathsf{IP}(\mathsf{Enc}(X), \tilde{Y}),$$

where $\tilde{Y} = t_1(Y, Y^3) + t_2(Y', Y'^3)$.

We now bound the min-entropy of $\tilde{Y}$ and have the following claim.

Claim 6.17. *$H_\infty(\tilde{Y}) > (1 - \delta)p$.*

Proof. We have two cases.

Case 1: $S_2 = \phi$. In this case $\tilde{Y} = t_1(Y, Y^3)$. Since $S_1 \neq \phi$, we have $t_1 \neq 0$. Thus $\tilde{Y}$ has the same min-entropy as Y, which is $(1 - \delta)p + 2 > (1 - \delta)p$.

Case 2: $S_2 \neq \phi$. In this case we have $t_1 \neq 0$ and $t_2 \neq 0$. We need to bound the min-entropy of $\tilde{Y} = t_1(Y, Y^3) + t_2(Y', Y'^3)$. Again, if for every two different y_1, y_2, we have $t_1(y_1, y_1^3) + t_2(y'_1, y'^3_1) \neq t_1(y_2, y_2^3) + t_2(y'_2, y'^3_2)$, then $\tilde{Y}$ will have the same min-entropy of Y. We now show that any element in $\mathsf{Supp}(\tilde{Y})$ can come from at most 3 different elements in $\mathsf{Supp}(Y)$.

To show this, assume for the sake of contradiction that there are 4 different y_1, y_2, y_3, y_4 such that $t_1(y_i, y_i^3) + t_2(y'_i, y'^3_i)$ are the same for $i = 1, 2, 3, 4$. First consider y_1, y_2, we have $t_1(y_1, y_1^3) + t_2(y'_1, y'^3_1) = t_1(y_2, y_2^3) + t_2(y'_2, y'^3_2)$. Since $t_1 \neq 0$, let $r = t_2/t_1 \in \mathbb{F}_{2^p}$. Thus $r \neq 0$ and we have $(y_1, y_1^3) + r(y'_1, y'^3_1) = (y_2, y_2^3) + r(y'_2, y'^3_2)$. We first consider the case where $r = 1$. In this case, the vectors (y_1, y_1^3), (y'_1, y'^3_1), (y_2, y_2^3) and (y'_2, y'^3_2) are linearly dependent over $\mathbb{F}_2$. However we know that the columns of the parity check matrix of the BCH code are 4-wise linearly independent. Thus we must have $y'_1 = y_2$ and $y'_2 = y_1$. Thus in this case the element in $\mathsf{Supp}(\tilde{Y})$ comes from at most 2 different elements in $\mathsf{Supp}(Y)$. Now if $r \neq 1$, we have

$$y_1 + ry'_1 = y_2 + ry'_2$$

and

$$(y_1)^3 + r(y'_1)^3 = (y_2)^3 + r(y'_2)^3.$$

Hence we get

$$y_1 - y_2 = r(y'_2 - y'_1)$$

and

$$(y_1^2 + y_1 y_2 + y_2^2)(y_1 - y_2) = r(y'^2_2 + y'_1 y'_2 + y'^2_1)(y'_2 - y'_1).$$

Since $y_1 \neq y_2$ and $r \neq 0$, we must have that $y'_1 \neq y'_2$. Thus we get

$$y_1^2 + y_1 y_2 + y_2^2 = y'^2_2 + y'_1 y'_2 + y'^2_1.$$

Similarly we can get

$$y_1^2 + y_1 y_3 + y_3^2 = y_3'^2 + y_1' y_3' + y_1'^2.$$

Thus

$$(y_1 + y_2 + y_3)(y_2 - y_3) = (y_1' + y_2' + y_3')(y_2' - y_3').$$

Also, from $y_2 + ry_2' = y_3 + ry_3'$ we get

$$y_2 - y_3 = r(y_3' - y_2').$$

Since $y_2 \neq y_3$, we have

$$y_1' + y_2' + y_3' = -r(y_1 + y_2 + y_3).$$

Similarly we can get

$$y_1' + y_2' + y_4' = -r(y_1 + y_2 + y_4).$$

Therefore

$$y_4' - y_3' = r(y_3 - y_4).$$

On the other hand, from $y_3 + ry_3' = y_4 + ry_4'$ we get

$$y_4' - y_3' = 1/r(y_3 - y_4).$$

Thus

$$(r^2 - 1)(y_3 - y_4) = 0.$$

Since $r \neq 1$, $r^2 - 1 \neq 0$. Thus we have $y_3 = y_4$, a contradiction.
Therefore, the min-entropy of $\tilde{Y}$ is at least $H_\infty(Y) - \log 3 = (1-\delta)p + 2 - \log 3 > (1-\delta)p$. □

Now, by Lemma 6.3, the lemma follows. □

Now we have the following theorem.

Theorem 6.18. *There exists a constant $0 < \delta < 1$ such that for any $n \in \mathbb{N}$, $k = (1/2 - \delta)n$, there exists an explicit (k, ϵ)-non-malleable extractor* $\mathsf{nmExt} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$ *with $m = \Omega(n)$ and $\epsilon = 2^{-\Omega(n)}$.*

Proof. By Lemma 6.16 and Lemma 3.13, we can choose $m = \Omega(n)$ bits from $\{Z_i\}$ such that when we have nmExt output $Z_1 \circ \cdots \circ Z_m$, we get

$$|(\mathsf{nmExt}(X,Y), \mathsf{nmExt}(X, \mathcal{A}(Y))) - (U_m, \mathsf{nmExt}(X, \mathcal{A}(Y)))| \le 2^{-\Omega(n)}.$$

Note that in Lemma 6.16 the seed Y only has min-entropy $(1-\delta)p+2$. Thus if we use a uniform seed $Y \in \{0,1\}^{p-1}$, by Theorem 3.17 we have that

$$|(\mathsf{nmExt}(X,Y), \mathsf{nmExt}(X, \mathcal{A}(Y)), Y) - (U_m, \mathsf{nmExt}(X, \mathcal{A}(Y)), Y)| \le \epsilon,$$

where $\epsilon = 2^{2m}(2^{-\delta p + 4} + 2^{-\Omega(n)}) = 2^{-\Omega(n)}$ when $m = \Omega(n)$ and is small enough. ■

6.4.2 Reducing the seed length

In the constructions mentioned above, we use a BCH code with distance 5. Thus the columns of the parity check matrix are 4-wise linearly independent. To reduce the seed length, we are going to use a BCH code with larger distance. Specifically, we will choose a $[2^\ell - 1, 2^\ell - 1 - 2t\ell, 4t+1]$-BCH code with $\ell = p/t$ for some parameter t to be chosen later. Note that the parity check matrix is a $2p \times (2^\ell - 1)$ matrix[4]. Thus the columns of the matrix are $D = 4t$-wise linearly independent. The detailed construction is as follows.

- Given an (n,k)-source X with $k = (1/2 - \delta)n$, pick a prime p such that $n \le p \le n(1 + \frac{1}{2\ln^2 n})$.
- Let $q = 2^p$ and g be a generator in $\mathbb{F}_q^*$. Treat X as an element in $\mathbb{F}_q^*$ and encode X such that $\mathsf{Enc}(X) = (X, g^X)$.
- Let $\ell = p/t$. Take the parity check matrix of a $[2^\ell - 1, 2^\ell - 1 - 2t\ell, 4t+1]$-BCH code. Note that it is a $2p \times (2^\ell - 1)$ matrix. Take an independent and uniform seed $Y \in \{0,1\}^{\ell-1}$ and let S_Y stand for the integer whose binary expression is Y. We encode Y to $\bar{Y}$ such that $\bar{Y}$ is the S_Y'th column in the parity check matrix.
- Output $\mathsf{nmExt}(X,Y) = \mathsf{IP}(\mathsf{Enc}(X), \bar{Y})$ where IP is the inner product function taken over $\mathbb{F}_2$.

As in Subsection 6.2, we have Claim 6.4. We now want to argue about the min-entropy of $t\bar{Y}$ and $t(\bar{Y} + \bar{\mathcal{A}(Y)})$.

Lemma 6.19. *Assume Y has min-entropy k_2, then $t\bar{Y}$ is $t^2 2^{-(k_2+1)}$-close to having min-entropy $t(k_2 - \log t)$, and $t(\bar{Y} + \bar{Y'})$ is $t^2 2^{-(k_2+2)} + t(t2^{-\frac{2}{3}k_2})^{\log t}$-close to having min-entropy $t((1 - \frac{\log t}{3t})k_2 - 3\log t)$.*

Proof. Without loss of generality assume that Y is a flat source. Let $K = 2^{k_2}$. First consider $t\bar{Y}$. Note that $\bar{Y}$ has the same min-entropy as Y and is also a flat source, since every two columns of the parity check matrix are different. The support of $t\bar{Y}$ has the form $\bar{y}_1 + \cdots + \bar{y}_t$. Consider the case where all $\bar{y}_i$'s are different. This takes up a probability mass of

$$\frac{K!}{(K-t)!} \cdot K^{-t} = 1 \cdot (1 - \frac{1}{K}) \cdots (1 - \frac{t-1}{K}) > 1 - \sum_{i=1}^{t-1} \frac{i}{K} > 1 - \frac{t^2}{2K}.$$

Since the columns of the parity check matrix are $4t$-wise linearly independent. For every two different sets $\{\bar{y}_i\}$'s, their sum cannot be the same. Therefore, the probability mass of getting a particular value is at most $t!K^{-t} \le 2^{-t(k_2 - \log t)}$. Thus $t\bar{Y}$ is $t^2 2^{-(k_2+1)}$-close to having min-entropy $t(k_2 - \log t)$.

Next consider $t(\bar{Y} + \bar{\mathcal{A}(Y)})$. Let $\mathcal{A}(Y) = Y'$ and $Y'' = \bar{Y} + \bar{Y'}$. Note that for every $s \in \mathsf{Supp}(Y'')$, $s \neq 0$ since $\forall y, \mathcal{A}(y) \neq y$. Also note that Y'' has min-entropy at least $k_2 - 1$ since if $\bar{y}_1 + \bar{y'}_1 = \bar{y}_2 + \bar{y'}_2$ for $y_1 \neq y_2$, then we must have $\bar{y'}_1 = \bar{y}_2$ and $\bar{y'}_2 = \bar{y}_1$. Without loss of generality assume that Y'' is a flat source with min-entropy $k_2 - 1$. Let $K_2 = 2^{k_2 - 1}$. Note that now in the support of Y'' there are no two different y_1, y_2 such that $\bar{y}_1 + \bar{y'}_1 = \bar{y}_2 + \bar{y'}_2$ (since this will be absorbed into the same element).

[4]Actually p is not divisible by t, thus $\ell t < p$. However for simplicity we will assume that the matrix has $2p$ rows. For example we can add 0's in the end, the small error does not affect our analysis.

We now consider tY''. An element in its support has the form $\sum_{i=1}^{t}(\bar{y}_i+\bar{y'}_i)$. We first get rid of those elements in $\mathsf{Supp}(tY'')$ such that some of the $\{\bar{y}_i+\bar{y'}_i\}$'s are the same. By the same argument as above this takes up a probability mass of at most $\frac{t^2}{2K_2}$. Now, for a particular set $\{\bar{y}_i+\bar{y'}_i\}_{i\in[t]}$, we consider how many different sets can have the same sum.

Since the columns of the parity check matrix are $4t$-wise linearly independent, if the sum of two different sets $\{\bar{y}_i+\bar{y'}_i\}_{i\in[t]}$ are the same, then except those $\bar{y}_i+\bar{y'}_i$'s that are common in both sets, the rest of $\bar{y}_i+\bar{y'}_i$'s must form cycles. By cycle we mean a set of l elements such that $\bar{y'}_1=\bar{y}_2, \bar{y'}_2=\bar{y}_3, \cdots, \bar{y'}_l=\bar{y}_1$ so that the sum is 0. Note that $l\geq 3$ since the support of Y'' has no 2-cycles. Let S_1, S_2 be the two sets $\{\bar{y}_i+\bar{y'}_i\}_{i\in[t]}$. Now, the elements in a cycle can come from both sets or just from one set. If the elements from a cycle comes only from S_2, then this cycle can be replaced by any other cycle with the same length, and the sum of S_1 and S_2 are still the same. On the other hand, if the elements of a cycle comes from both S_1 and S_2, then the elements in this cycle are completely determined by S_1 since cycles are disjoint. Therefore, let r be the number of common elements in S_1, S_2, and let l be the total length of cycles whose elements only come from the rest elements of S_2, and note that cycles have length at least 3, we have that if $l\geq \log t$, then the total probability mass of these elements in $\mathsf{Supp}(tY'')$ is at most

$$\sum_{\log t\leq l\leq t}\binom{t}{l}\binom{\frac{K_2}{3}}{\frac{l}{3}}l!(K_2)^{-l}\leq \sum_{\log t\leq l\leq t}t^l\left(\frac{K_2}{3}\right)^{\frac{l}{3}}(K_2)^{-l}<\sum_{\log t\leq l\leq t}(t(K_2)^{-\frac{2}{3}})^l<t(t2^{-\frac{2}{3}k_2})^{\log t}.$$

On the other hand, if $l<\log t$, then the probability that tY'' gets a particular value is at most

$$\sum_{0\leq l\leq \log t}\sum_{0\leq r\leq t-l}\binom{t}{l}\binom{t-l}{r}\binom{\frac{K_2}{3}}{\frac{l}{3}}t!(K_2)^{-t}<t\log t\cdot t^{2t}\left(\frac{K_2}{3}\right)^{\frac{\log t}{3}}(K_2)^{-t}<t\log t(t^2(K_2)^{-(1-\frac{\log t}{3t})})^t.$$

Thus the min-entropy is at least $t((1-\frac{\log t}{3t})k_2-3\log t)$. $\square$

Now for an (n,k)-source X with $k=(1/2-\delta)n$, we know that $\mathsf{3Enc}(X)$ is $2^{-\Omega(n)}$-close to having min-entropy $(1/2+\delta)(2p)$. Assume that we want our non-malleable extractor to have error $\epsilon\leq 1/n$. We'll choose a parameter $t<n/C\log n$ for a sufficiently large constant $C>1$. When Y is uniform over $\ell=p/t$ bits, $t\bar{Y}$ is close to having min-entropy $t(k_2-\log t)>(1-1/C)p>(1/2-\delta/2)(2p)$, and $t(\bar{Y}+\bar{Y}')$ is close to having min-entropy $t((1-\frac{\log t}{3t})k_2-3\log t)>(1-1/C)p>(1/2-\delta/2)(2p)$. When $t\bar{Y}$ and $t(\bar{Y}+\bar{Y}')$ indeed have this min-entropy, by Lemma 3.19 we have that both $\mathsf{IP}(\mathsf{3Enc}(X),t\bar{Y})$ and $\mathsf{IP}(\mathsf{3Enc}(X),t(\bar{Y}+\bar{Y}'))$ are $2^{-\Omega(n)}$-close to uniform. Thus we can take $t=\Omega(n/(\log(1/\epsilon)))$ and by Lemma 3.20 and Theorem 3.17 we have that the error of the non-malleable extractor is at most ϵ, and the seed length is roughly $p/t=O(n/t)=O(\log(1/\epsilon))$. Thus we have the following theorem.

Theorem 6.20. *There exists a universal constant $\delta>0$ such that for every $n\in\mathbb{N}$ and ϵ such that $2^{-\Omega(n)}\leq\epsilon\leq 1/\mathrm{poly}(n)$, there exists an explicit (k,ϵ) non-malleable extractor $\mathsf{nmExt}:\{0,1\}^n\times\{0,1\}^d\to\{0,1\}$ for $k=(1/2-\delta)n$ and seed length $d=O(\log n+\log(1/\epsilon))$.*

7 An Optimal Privacy Amplification Protocol for Arbitrarily Linear Entropy

In this section we present our privacy amplification protocol for (n,k)-sources X with $k = \delta n$ for any constant $\delta > 0$. Following [KR09] and [DLWZ11], we define a privacy amplification protocol (P_A, P_B). The protocol is executed by two parties Alice and Bob, who share a secret $X \in \{0,1\}^n$. An active, computationally unbounded adversary Eve might have some partial information E about X satisfying $\widetilde{H}_\infty(X|E) \geqslant k$. Since Eve is unbounded, we can assume without loss of generality that she is deterministic. Informally, we want the protocol to be such that whenever a party (Alice or Bob) does not reject, the key R output by this party is random and 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 has full control of the communication channel between the two parties. This means that Eve can arbitrarily insert, delete, reorder or modify messages sent by Alice and Bob to each other. In particular, Eve's strategy P_E 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. Alice and Bob are assumed to have fresh, private and independent random bits Y and W, respectively. Y and W are not known to Eve. In the protocol we use $\perp$ as a special symbol to indicate rejection. At the end of the left execution $(P_A(X,Y), P_E(E))$, Alice outputs a key $R_A \in \{0,1\}^m \cup \{\perp\}$. 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).

Definition 7.1. 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, m, ϵ)-*privacy amplification protocol* if it satisfies the following properties whenever $\widetilde{H}_\infty(X|E) \geq k$:

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.

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

7.1 Prerequisites from previous work

One-time message authentication codes (MACs) use a shared random key to authenticate a message in the information-theoretic setting.

Definition 7.2. A function family $\{\mathsf{MAC}_R : \{0,1\}^d \to \{0,1\}^v\}$ is a ϵ-secure one-time MAC for messages of length d with tags of length v if for any $w \in \{0,1\}^d$ and any function (adversary) $A : \{0,1\}^v \to \{0,1\}^d \times \{0,1\}^v$,

$$\Pr_R[\mathsf{MAC}_R(W') = T' \wedge W' \neq w \mid (W', T') = A(\mathsf{MAC}_R(w))] \leq \epsilon,$$

where R is the uniform distribution over the key space $\{0,1\}^\ell$.

Theorem 7.3 ([KR09]). *For any message length d and tag length v, there exists an efficient family of $(\lceil \frac{d}{v} \rceil 2^{-v})$-secure MACs with key length $\ell = 2v$. In particular, this MAC is ε-secure when $v = \log d + \log(1/\epsilon)$.*
More generally, this MAC also enjoys the following security guarantee, even if Eve has partial information E about its key R. Let (R, E) be any joint distribution. Then, for all attackers A_1 and A_2,

$$\begin{aligned}\Pr_{(R,E)}[&\mathsf{MAC}_R(W') = T' \wedge W' \neq W \mid W = A_1(E),\\ &(W', T') = A_2(\mathsf{MAC}_R(W), E)] \leq \left\lceil \frac{d}{v} \right\rceil 2^{v - \widetilde{H}_\infty(R|E)}.\end{aligned}$$

(In the special case when $R \equiv U_{2v}$ and independent of E, we get the original bound.)

Remark 7.4. Note that the above theorem indicates that the MAC works even if the key R has average conditional min-entropy rate $> 1/2$.

Sometimes it is convenient to talk about average case seeded extractors, where the source X has average conditional min-entropy $\widetilde{H}_\infty(X|Z) \geq k$ and the output of the extractor should be uniform given Z as well. The following lemma is proved in [DORS08].

Lemma 7.5. *[DORS08] For any $\delta > 0$, if Ext is a (k, ϵ) extractor then it is also a $(k+\log(1/\delta), \epsilon+\delta)$ average case extractor.*

For a strong seeded extractor with optimal parameters, we use the following extractor constructed in [GUV09].

Theorem 7.6 ([GUV09]). *For every constant $\alpha > 0$, and all positive integers n, k and any $\epsilon > 0$, there is an explicit construction of a strong (k, ϵ)-extractor $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with $d = O(\log n + \log(1/\epsilon))$ and $m \geq (1 - \alpha)k$. It is also a strong (k, ϵ) average case extractor with $m \geq (1 - \alpha)k - O(\log n + \log(1/\epsilon))$.*

We need the following construction of strong two-source extractors in [Raz05].

Theorem 7.7 ([Raz05]). *For any n_1, n_2, k_1, k_2, m and any $0 < \delta < 1/2$ with*

- $n_1 \geq 6 \log n_1 + 2 \log n_2$

- $k_1 \geq (0.5 + \delta)n_1 + 3 \log n_1 + \log n_2$
- $k_2 \geq 5 \log(n_1 - k_1)$
- $m \leq \delta \min[n_1/8, k_2/40] - 1$

There is a polynomial time computable strong 2-source extractor $\mathsf{Raz} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ *for min-entropy* k_1, k_2 *with error* $2^{-1.5m}$.

Theorem 7.8. *[DLWZ11, CRS12, Li12] For every constant* $\delta > 0$*, there exists a constant* $\beta > 0$ *such that for every* $n, k \in \mathbb{N}$ *with* $k \geq (1/2 + \delta)n$ *and* $\epsilon > 2^{-\beta n}$ *there exists an explicit* (k, ϵ) *non-malleable extractor with seed length* $d = O(\log n + \log \epsilon^{-1})$ *and output length* $m = \Omega(n)$.

The following standard lemma about conditional min-entropy is implicit in [NZ96] and explicit in [MW97].

Lemma 7.9 ([MW97])**.** *Let* X *and* Y *be random variables and let* $\mathcal{Y}$ *denote the range of* Y*. Then for all* $\epsilon > 0$*, one has*

$$\Pr_Y \left[H_\infty(X|Y = y) \geq H_\infty(X) - \log |\mathcal{Y}| - \log \left(\frac{1}{\epsilon} \right) \right] \geq 1 - \epsilon.$$

7.2 The privacy amplification protocol

We first define the following alternating extraction protocol.

Quentin: Q, S_0		Wendy: $X, \bar{X} = (X_1, \cdots, X_t)$
S_0	$\xrightarrow{S_0}$	
	$\xleftarrow{R_0}$	$R_0 = \mathsf{Raz}(S_0, X)$
$S_1 = \mathsf{Ext}_q(Q, R_0)$	$\xrightarrow{S_1}$	
	$\xleftarrow{R_1}$	$R_1 = \mathsf{Ext}_w(X, S_1)$, $V_1 = \mathsf{Ext}_v(X_1, S_1)$
$S_2 = \mathsf{Ext}_q(Q, R_1)$	$\xrightarrow{S_2}$	
	$\xleftarrow{R_2}$	$R_2 = \mathsf{Ext}_w(X, S_2)$, $V_2 = \mathsf{Ext}_v(X_2, S_2)$
	$\cdots$	
$S_t = \mathsf{Ext}_q(Q, R_{t-1})$	$\xrightarrow{S_t}$	
		$R_t = \mathsf{Ext}_w(X, S_t)$, $V_t = \mathsf{Ext}_v(X_t, S_t)$

Figure 1: Alternating Extraction.

Alternating Extraction. Assume that we have two parties, Quentin and Wendy. Quentin has a source Q, Wendy has a source X and a source $\bar{X} = (X_1 \circ \cdots \circ X_t)$ with t rows. Also assume that Quentin has a weak source S_0 with entropy rate $> 1/2$ (which may be correlated with Q). Suppose that (Q, S_0) is kept secret from Wendy and $(X, \bar{X})$ is kept secret from Quentin. Let Ext_q,

Ext_w, Ext_v be strong seeded extractors with optimal parameters, such as that in Theorem 7.6. Let Raz be the strong two-source extractor in Theorem 7.7. Let s, d be two integer parameters for the protocol. The *alternating extraction protocol* is an interactive process between Quentin and Wendy that runs in $t+1$ steps.

In the 0'th step, Quentin sends S_0 to Wendy, Wendy computes $R_0 = \mathsf{Raz}(S_0, X)$ and replies R_0 to Quentin, Quentin then computes $S_1 = \mathsf{Ext}_q(Q, R_0)$. In this step R_0, S_1 each outputs d bits. In the first step, Quentin sends S_1 to Wendy, Wendy computes $V_1 = \mathsf{Ext}_v(X_1, S_1)$ and $R_1 = \mathsf{Ext}_w(X, S_1)$. She sends R_1 to Quentin and Quentin computes $S_2 = \mathsf{Ext}_q(Q, R_1)$. In this step V_1 outputs $2^{t-1}s$ bits, and R_1, S_2 each outputs d bits. In each subsequent step i, Quentin sends S_i to Wendy, Wendy computes $V_i = \mathsf{Ext}_v(X_i, S_i)$ and $R_i = \mathsf{Ext}_w(X, S_i)$. She replies R_i to Quentin and Quentin computes $S_{i+1} = \mathsf{Ext}_q(Q, R_i)$. In step i, V_i outputs $2^{t-i}s$ bits, and R_i, S_{i+1} each outputs d bits. Therefore, this process produces the following sequence:

$$\begin{aligned} S_0, R_0 = \mathsf{Raz}(S_0, X), & S_1 = \mathsf{Ext}_q(Q, R_0), V_1 = \mathsf{Ext}_v(X_1, S_1), R_1 = \mathsf{Ext}_w(X, S_1), \cdots, \\ & S_t = \mathsf{Ext}_q(Q, R_{t-1}), V_t = \mathsf{Ext}_v(X_t, S_t), R_t = \mathsf{Ext}_w(X, S_t). \end{aligned}$$

Look-Ahead Extractor. Now we can define our look-ahead extractor. Let $Y = (Q, S_0)$ be a seed, the look-ahead extractor is defined as

$$\mathsf{laExt}((X, \bar{X}), Y) = \mathsf{laExt}((X, \bar{X}), (Q, S_0)) \stackrel{def}{=} V_1, \cdots, V_t.$$

Note that the look-ahead extractor can be computed by each party (Alice or Bob) alone in our final protocol. Now we give our protocol for privacy amplification.

7.2.1 The protocol

Now we give our privacy amplification protocol for the setting when $\widetilde{H}_\infty(X|E) = k \geqslant \delta n$. We assume that the error ϵ we seek satisfies $2^{-\Omega(\delta n)} < \epsilon < 1/n$. In the description below, it will be convenient to introduce an "auxiliary" security parameter s. Eventually, we will set $s = \log(C/\epsilon) + O(1) = \log(1/\epsilon) + O(1)$, so that $O(C)/2^s < \epsilon$, for a sufficiently large $O(C)$ constant related to the number of "bad" events we will need to account for. We will need the following building blocks:

- Let $\mathsf{Cond} : \{0,1\}^n \to (\{0,1\}^{n'})^C$ be a rate-$(\delta \to 0.9, 2^{-s})$-somewhere-condenser. Specifically, we will use the one from Theorem 3.8, where $C = \mathrm{poly}(1/\delta) = O(1)$, $n' = \mathrm{poly}(\delta)n = \Omega(n)$ and $2^{-s} \gg 2^{-\Omega(\delta n)}$.
- Let $\mathsf{nmExt} : \{0,1\}^{n'} \times \{0,1\}^{d'} \to \{0,1\}^{m'}$ be a $(0.8n', 2^{-s})$-non-malleable extractor. Specifically, we will use the one from Theorem 7.8 and set the output length $m' = 6 \cdot 2^C s$.
- Let $\mathsf{Ext}, \mathsf{Ext}_q, \mathsf{Ext}_w, \mathsf{Ext}_v$ be seeded extractors with error 2^{-s}, seed length $d = O(\log n + s)$ and optimal entropy loss $O(s)$ as in Theorem 7.6. $\mathsf{Ext}_q, \mathsf{Ext}_w, \mathsf{Ext}_v$ will be used in laExt.
- Let Raz be the strong two-source extractor in Theorem 7.7. This will be used in laExt.
- Let lrMAC be a one-time ("leakage-resilient") MAC for d-bit messages, with key length $2^C(6s)$ and tag length $2^C(3s)$. We will later use the second part of Theorem 7.3 to argue good security of this MAC even when some bits of partial information about its key is leaked to the attacker.

Using the above building blocks, the protocol is given in Figure 2. To emphasize the presence of Eve, we will use 'prime' to denote all the protocol values seen or generated by Bob; e.g., Bob picks W', but Alice sees potentially different W, etc.

Alice: X	Eve: E	Bob: X
$(X_1, \ldots X_C) = \mathsf{Cond}(X)$. Sample random $Y = (Y_1, Y_2, Y_3)$ such that $\lvert Y_1\rvert = max\{d, d'\}, \lvert Y_3\rvert = 30max\{d, d'\} + 3s$, $\lvert Y_2\rvert = 4Cd + 31max\{d, d'\} + 4s$		$(X_1, \ldots X_C) = \mathsf{Cond}(X)$.
	$(Y_1, Y_2, Y_3) \to (Y_1', Y_2', Y_3')$	
		Sample random W' with d bits. $Z' = \mathsf{Ext}(X; Y_1')$ with $2^C(6s)$ bits. $\bar{X}' = (\bar{X}_1', \ldots, \bar{X}_C')$, where $\bar{X}_i' = \mathsf{nmExt}(X_i, Y_1')$. $V' = (V_1', \ldots, V_C') = \mathsf{laExt}((X, \bar{X}'), (Y_2', Y_3'))$ with parameters $(2s, d)$. $T' = \mathsf{lrMAC}_{Z'}(W')$. Set final $R_B = \mathsf{Ext}(X; W')$.
	$(W, T, \bar{V}) \leftarrow (W', T', V')$	
$Z = \mathsf{Ext}(X; Y_1)$ with $2^C(6s)$ bits. $\bar{X} = (\bar{X}_1, \ldots, \bar{X}_C)$, where $\bar{X}_i = \mathsf{nmExt}(X_i, Y_1)$. $V = (V_1, \ldots, V_C) = \mathsf{laExt}((X, \bar{X}), (Y_2, Y_3))$ with parameters $(2s, d)$. **If** $T \neq \mathsf{lrMAC}_Z(W)$ or $V \neq \bar{V}$ *reject*. Set final $R_A = \mathsf{Ext}(X; W)$.		

Figure 2: 2-round Privacy Amplification Protocol for $\widetilde{H}_\infty(X|E) \geq \delta n$.

Theorem 7.10. *For any constant $\delta > 0$, the above protocol is a privacy amplification protocol with security parameter $\log(1/\epsilon)$, entropy loss $2^{\mathrm{poly}(1/\delta)} \log(1/\epsilon)$, randomness complexity $\mathrm{poly}(1/\delta)\log(1/\epsilon)$ and communication complexity $2^{\mathrm{poly}(1/\delta)} \log(1/\epsilon)$.*

Proof. The proof can be divided into two cases: whether the adversary changes Y_1 or not. Note that Y_1, Y_2, Y_3 and W all have size $O(s)$.

Case 1: The adversary does not change Y_1. In this case, note that $Z = Z'$ and is 2^{-s}-close to uniform in Eve's view (even conditioned on Y_1, Y_2, Y_3). Note that the size of $(V_1', \ldots, V_C')$ is at most $\sum_i 2^{C-i}(2s) < 2^C(2s)$, and the size of Z is $2^C(6s)$. Therefore, by Lemma 3.11 even if conditioned on $(V_1', \ldots, V_C')$, the average conditional min-entropy of Z is at least $2^C(6s) - 2^C(2s) = 2^C(4s)$. Therefore by theorem 7.3 the probability that Eve can change W' to a different W without causing Alice to reject is at most

$$\left\lceil \frac{O(s)}{2^C(3s)} \right\rceil 2^{2^C(3s) - 2^C(4s)} + 2^{-s} \leq O(2^{-s}).$$

When $W = W'$, by theorem 7.6 $R_A = R_B$ and is 2^{-s}-close to uniform in Eve's view.

Case 2: The adversary does change Y_1. In this case, first note that by Theorem 3.8, $\mathsf{Cond}(X) = (X_1, \dots X_C)$ is 2^{-s}-close to a somewhere rate-0.9-source with C rows, and each row has length $\Omega(n)$. In the following we will simply treat it as a somewhere rate-0.9-source, since this only adds 2^{-s} to the error. We assume that $X_g, 1 \le g \le C$ is a rate 0.9-source [5].

Now since the adversary changes Y_1 to $Y_1' \neq Y_1$, by Theorem 7.8 we have that

$$(\bar{X}_g, \bar{X}_g', Y_1) \approx_{2^{-s}} (U_{m'}, \bar{X}_g', Y_1).$$

As the first step for the following analysis, we now fix Y_1, Y_1' and $Z', \bar{X}_g'$. Note that Y_1' is a deterministic function of (Y_1, Y_2, Y_3), and after fixing Y_1', $(Z', \bar{X}_g')$ is a deterministic function of X. Thus by Lemma 3.11 we have the following claim.

Claim 7.11. *After the fixings of $(Y_1, Y_1', Z', \bar{X}_g')$, $\bar{X}_g$ is a deterministic function of X and is 2^{-s} close to a source with average conditional min-entropy $m' - 2^C(6s)$.*

Note that by Lemma 3.11, after this fixing, the average conditional min-entropy of X is at least $k - m' - 2^C(6s)$. Now we analyze the sequences $(V_1, \dots, V_C)$ produced by laExt, or equivalently, the alternating extraction process. Note here $(Q, S_0) = (Y_2, Y_3)$ and $(Q', S_0') = (Y_2', Y_3')$. First we have the following claim.

Lemma 7.12. *In step 0, we have*

$$(R_0, S_0, S_0') \approx_{2^{-s}} (U_d, S_0, S_0')$$

and

$$(S_1, R_0, S_0, R_0', S_0') \approx_{5 \cdot 2^{-s}} (U_d, R_0, S_0, R_0', S_0').$$

Moreover, conditioned on (S_0, S_0'), (R_0, R_0') are both deterministic functions of X; conditioned on (R_0, S_0, R_0', S_0'), (S_1, S_1') are both deterministic functions of Q.

Proof of the claim. Note that previously we have fixed Y_1, Y_1'. Since Y_1 is independent of Y_3 and Y_1' is a deterministic function of Y, by Lemma 7.9 we have that $S_0 = Y_3$ is 2^{-s}-close to a source with min-entropy $29max\{d, d'\} + 2s$. Note that Y_3 and X are still independent. Thus by Theorem 7.7 we have that

$$(R_0, S_0) \approx_{2^{-s}} (U_d, S_0).$$

Since conditioned on S_0, R_0 is a deterministic function of X, which is independent of Y, we also have that

$$(R_0, S_0, S_0') \approx_{2^{-s}} (U_d, S_0, S_0').$$

Now we fix (S_0, S_0') and (R_0, R_0') are both deterministic functions of X. Note that $S_0 = Y_3$ is independent of Y_2 and $S_0' = Y_3'$ is a deterministic function of Y. Thus by Lemma 7.9 we have that conditioned on these fixings $Q = Y_2$ is 2^{-s}-close to a source with entropy $4Cd$. Since R_0, R_0' are both deterministic functions of X, they are independent of Q. Therefore by Theorem 7.6 we have

[5] In general a somewhere rate-0.9-source is a convex combination of elementary somewhere rate-0.9-sources, but without loss of generality we can assume it is an elementary somewhere rate-0.9-source.

$$(S_1, R_0, R_0') \approx_{2^{-s}} (U_d, R_0, R_0').$$

Thus altogether we have that

$$(S_1, R_0, S_0, R_0', S_0') \approx_{5 \cdot 2^{-s}} (U_d, R_0, S_0, R_0', S_0')$$

Moreover, conditioned on (R_0, S_0, R_0', S_0'), (S_1, S_1') are both deterministic functions of Q. ■

Now we fix (R_0, S_0, R_0', S_0'). Note that after this fixing, S_1, S_1' are both functions of $Q = Y_2$. Note that Q has min-entropy at least $4Cd$.

For $i = 0, \cdots, C$, let $View_i = (S_0, \cdots, S_i, R_0, \cdots, R_i, V_1, \cdots, V_i)$. Similarly define $View_i'$ to be the corresponding variables produced by Bob. We now have the following lemma.

Lemma 7.13. *For any i, we have that*

$$(R_i, View_{i-1}, View_{i-1}', S_i, S_i') \approx_{(2i+4)2^{-s}} (U_d, View_{i-1}, View_{i-1}', S_i, S_i')$$

and

$$(S_{i+1}, View_i, View_i') \approx_{(2i+5)2^{-s}} (U_d, View_i, View_i').$$

Moreover, conditioned on $(View_{i-1}, View_{i-1}', S_i, S_i')$, (R_i, R_i', V_i, V_i') are all deterministic functions of X; conditioned on $(View_i, View_i')$, (S_{i+1}, S_{i+1}') are both deterministic functions of Q.

Proof. We prove the lemma by induction on i. When $i = 0$, the statements are already proved in Lemma 7.12. Now we assume that the statements hold for $i = j$ and we prove them for $i = j + 1$.

We first fix $(View_j, View_j')$. Since now (S_{j+1}, S_{j+1}') are both deterministic functions of Q, they are independent of X. Moreover S_j is $(2j + 5)2^{-s}$-close to uniform. Note that the average conditional min-entropy of X is at least $k - m' - 2^C(6s) - 2^C(4s) - 2Cd = k - m' - 2^C(10s) - 2Cd$. Therefore by Theorem 7.6 we have that

$$(R_{j+1}, View_j, View_j', S_{j+1}, S_{j+1}') \approx_{(2j+6)2^{-s}} (U_d, View_j, View_j', S_{j+1}, S_{j+1}').$$

Moreover, conditioned on $(View_j, View_j', S_{j+1}, S_{j+1}')$, $(R_{j+1}, R_{j+1}', V_{j+1}, V_{j+1}')$ are all deterministic functions of X.

Next, since conditioned on $(View_j, View_j', S_{j+1}, S_{j+1}')$, (R_{j+1}, R_{j+1}') are both deterministic functions of X, they are independent of Q. Moreover R_{j+1} is $(2j + 6)2^{-s}$-close to uniform. Note that the average conditional min-entropy of Q is at least $4Cd - 2Cd = 2Cd$. Therefore by Theorem 7.6 we have that

$$\begin{aligned}&(S_{j+2}, View_j, View_j', S_{j+1}, S_{j+1}', R_{j+1}, R_{j+1}', V_{j+1}, V_{j+1}')\\ \approx_{(2j+7)2^{-s}}&(U_d, View_j, View_j', S_{j+1}, S_{j+1}', R_{j+1}, R_{j+1}', V_{j+1}, V_{j+1}').\end{aligned}$$

Namely,

$$(S_{j+2}, View_{j+1}, View_{j+1}') \approx_{(2(j+1)+5)2^{-s}} (U_d, View_{j+1}, View_{j+1}').$$

Moreover, conditioned on $(View_{j+1}, View_{j+1}')$, (S_{j+2}, S_{j+2}') are deterministic functions of Q. ■

Now we consider step g in the alternating extraction protocol. By Claim 7.11, conditioned on $(Y_1, Y_1', Z', \bar{X}_g')$, $\bar{X}_g$ is a deterministic function of X and is 2^{-s} close to a source with average conditional min-entropy $m' - 2^C(6s)$. Since $m' = \Omega(\delta n)$, when s is significantly smaller than δn (but we can still achieve up to $s = \Omega(\delta n)$), we have that $m' \geq 2^C(12s) + 2Cd$ and thus $m' - 2^C(6s) \geq 2^C(6s) + 2Cd$.

We now have the following lemma.

Lemma 7.14. *Conditioned on the fixing of $(View_{g-1}, View_{g-1}')$, $\bar{X}_g$ is a deterministic function of X, and the average conditional min-entropy of $\bar{X}_g$ is at least $2^C(2s)$.*

Proof. We first use induction to prove that for any i, conditioned on the fixing of $(View_i, View_i')$, $\bar{X}_g$ is a deterministic function of X. When $i = 0$, we first fix (S_0, S_0'), which is independent of $\bar{X}_g$. After this fixing (R_0, R_0') is a deterministic function of X. Thus we can now fix (R_0, R_0') and conditioned on this fixing, $\bar{X}_g$ is still a deterministic function of X. Thus the statement holds for $i = 0$.

Now assume that the statement holds for $i = j$, we show that it also holds for $i = j + 1$. Specifically, when $(View_j, View_j')$ is fixed, (S_{j+1}, S_{j+1}') is a deterministic function of Q, which is independent of X. Thus we can fix (S_{j+1}, S_{j+1}'). Now after this fixing, $(R_{j+1}, R_{j+1}', V_{j+1}, V_{j+1}')$ is a deterministic function of X. Thus we can now fix $(R_{j+1}, R_{j+1}', V_{j+1}, V_{j+1}')$ and conditioned on this fixing, $\bar{X}_g$ is still a deterministic function of X. Thus the statement holds for $i = j + 1$. Therefore, for any i, conditioned on the fixing of $(View_i, View_i')$, $\bar{X}_g$ is a deterministic function of X.

Finally, note that only the fixings of (R_j, R_j', V_j, V_j') can cause $\bar{X}_g$ to lose entropy. Since the total size of $\{(R_j, R_j', V_j, V_j')\}$ is at most $\sum_{i=1}^{C} 2^{C-i}(4s) + 2Cd < 2^C(4s) + 2Cd$, by Lemma 3.11 the average conditional min-entropy of $\bar{X}_g$ is at least $2^C(2s)$. ∎

Now by Lemma 7.13, conditioned on the fixing of $(View_{g-1}, View_{g-1}')$, $S_g \approx_{(2g+3)2^{-s}} U_d$ and S_g, S_g' are both deterministic functions of Q, which is independent of X. Thus S_g and $\bar{X}_g$ are independent. Therefore by Theorem 7.6 we have

$$(V_g, S_g, S_g') \approx_{2^{-s}} (U_{2^{C-g}(2s)}, S_g, S_g').$$

Adding back all the errors, and note that we have fixed $(Y_1, Y_1', Z', \bar{X}_g')$ and $(View_{g-1}, View_{g-1}')$, we have that

$$(V_g, S_g, S_g', View_{g-1}, View_{g-1}', Z', \bar{X}_g') \approx_{O(C2^{-s})} (U_{2^{C-g}(2s)}, S_g, S_g', View_{g-1}, View_{g-1}', Z', \bar{X}_g').$$

In particular, note that $V_g' = \mathsf{Ext}_v(\bar{X}_g', S_g')$ and for a fixed message w', $T' = \mathsf{IrMAC}_{Z'}(w')$ is a function of Z'. Thus we have that

$$(V_g, View_{g-1}', T', V_g') \approx_{O(C2^{-s})} (U_{2^{C-g}(2s)}, View_{g-1}', T', V_g').$$

This implies that

$$(V_g, T', V_1', \cdots, V_g') \approx_{O(C2^{-s})} (U_{2^{C-g}(2s)}, T', V_1', \cdots, V_g').$$

Now note the size of $(V_{g+1}', \cdots, V_C')$ is at most $\sum_{i=g+1}^{C} 2^{C-i}(2s) = 2^{C-g}(2s) - 2s$, and that V_g has size $2^{C-g}(2s)$. Therefore, if V_g is uniform conditioned on $(T', V_1', \cdots, V_g')$, then by Lemma 7.9

we have that with probability $1 - 2^{-s}$ over the fixings of $(T', V'_1, \cdots, V'_C)$, V_g is a source with min-entropy s. Thus the probability that Eve can come up with the correct V_g is at most $2 \cdot 2^{-s}$. Adding back the error, we have that in the case that Eve changes Y_1, the probability that Alice does not reject is at most $O(C2^{-s})$. For an appropriately chosen $s = \log(1/\epsilon) + O(1)$ this is at most ϵ.

Finally, note that the entropy loss of the protocol is $O(2^C s) = 2^{\mathrm{poly}(1/\delta)} \log(1/\epsilon) = O(\log(1/\epsilon))$, the randomness complexity is $O(Cd + s) = O(Cs) = \mathrm{poly}(1/\delta) \log(1/\epsilon) = O(\log(1/\epsilon))$ and the communication complexity is $O(2^C s) = 2^{\mathrm{poly}(1/\delta)} \log(1/\epsilon) = O(\log(1/\epsilon))$. ■

8 Conclusions and Open Problems

In this paper we present a strong connection between non-malleable extractors and two-source extractors. First, we show that non-malleable extractors can be used to construct two-source extractors. If the non-malleable extractor works for small min-entropy and has a short seed length with respect to the error, then the resulted two-source extractor beats the best known construction of two-source extractors. Second, we show that two-source extractors of the form $\mathsf{IP}(f(X), Y)$ can be used to construct non-malleable extractors. Using this connection, we give the first explicit constructions of non-malleable extractors for min-entropy $k < n/2$.

The most important message from this part is perhaps that, non-malleable extractors and two-source extractors, although seemingly different, are closely related. Thus, future research should probably consider these two kinds of extractors together, as improvements in one kind may lead to improvements in the other. Our connection also suggests that it may be hard to construct non-malleable extractors for small entropy. However, strictly speaking, our result only shows that it may be hard to construct non-malleable extractors for small entropy with short seed length with respect to the error. It is totally possible that we can get explicit non-malleable extractors for small entropy with large seed length. Moreover, the weaker notion of non-malleable condensers introduced in [Li12] is a hopeful alternative.

We also give the first privacy amplification protocol for $k = \delta n$ that simultaneously achieves optimal round complexity (2 rounds), asymptotically optimal entropy loss and communication complexity. However, our entropy loss is $2^{\mathrm{poly}(1/\delta)} s$, which has a large hidden constant for small δ. As a comparison, the protocol in [DLWZ11] runs in $\mathrm{poly}(1/\delta)$ rounds but only has entropy loss $\mathrm{poly}(1/\delta)s$. Thus for practical purposes it is interesting to see if we can reduce the hidden constant. In particular, it remains an interesting open problem to construct non-malleable extractors or non-malleable condensers for arbitrarily linear min-entropy.

Acknowledgments

We are grateful to David Zuckerman for many valuable discussions, especially for his suggestion to use the BCH code.

References

[BBR88] C.H. Bennett, G. Brassard, and J.-M. Robert. Privacy amplification by public discussion. *SIAM Journal on Computing*, 17:210–229, 1988.

[BHP01] R.C. Baker, G. Harman, and J. Pintz. The difference between consecutive primes. In *Proceedings of the London Mathematical Society*, 2001.

[BIW04] Boaz Barak, R. Impagliazzo, and Avi Wigderson. Extracting randomness using few independent sources. In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 384–393, 2004.

[BKS$^+$05] Boaz Barak, Guy Kindler, Ronen Shaltiel, Benny Sudakov, and Avi Wigderson. Simulating independence: New constructions of condensers, Ramsey graphs, dispersers, and extractors. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing*, pages 1–10, 2005.

[BKT04] Jean Bourgain, Nets Katz, and Terence Tao. A sum-product estimate in finite fields, and applications. *Geometric and Functional Analysis*, 14:27–57, 2004.

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

[BRSW06] Boaz Barak, Anup Rao, Ronen Shaltiel, and Avi Wigderson. 2 source dispersers for $n^{o(1)}$ entropy and Ramsey graphs beating the Frankl-Wilson construction. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, 2006.

[BSZ11] Eli Ben-Sasson and Noga Zewi. From affine to two-source extractors via approximate duality. In *Proceedings of the 43rd Annual ACM Symposium on Theory of Computing*, 2011.

[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.

[CKOR10] N. Chandran, B. Kanukurthi, R. Ostrovsky, and L. Reyzin. Privacy amplification with asymptotically optimal entropy loss. In *Proceedings of the 42nd Annual ACM Symposium on Theory of Computing*, pages 785–794, 2010.

[CRS12] Gil Cohen, Ran Raz, and Gil Segev. Non-malleable extractors with short seeds and applications to privacy amplification. In *Proceedings of the 27th Annual IEEE Conference on Computational Complexity*, 2012.

[DKRS06] Y. Dodis, J. Katz, L. Reyzin, and A. Smith. Robust fuzzy extractors and authenticated key agreement from close secrets. In *CRYPTO*, pages 232–250, 2006.

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

[DLWZ11] Yevgeniy Dodis, Xin Li, Trevor D. Wooley, and David Zuckerman. Privacy amplification and non-malleable extractors via character sums. In *Proceedings of the 52nd Annual IEEE Symposium on Foundations of Computer Science*, 2011.

[DORS08] Y. Dodis, R. Ostrovsky, L. Reyzin, and A. Smith. Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. *SIAM Journal on Computing*, 38:97–139, 2008.

[DW08] Zeev Dvir and Avi Wigderson. Kakeya sets, new mergers and old extractors. In *Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science*, 2008.

[DW09] Yevgeniy Dodis and Daniel Wichs. Non-malleable extractors and symmetric key cryptography from weak secrets. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing*, page 601610, 2009.

[FS02] Lance Fortnow and Ronen Shaltiel. Recent developments in explicit constructions of extractors, 2002.

[GUV09] Venkatesan Guruswami, Christopher Umans, and Salil Vadhan. Unbalanced expanders and randomness extractors from Parvaresh-Vardy codes. *Journal of the ACM*, 56(4), 2009.

[Kon03] S. Konyagin. A sum-product estimate in fields of prime order. Technical report, Arxiv, 2003. `http://arxiv.org/abs/math.NT/0304217`.

[KR09] B. Kanukurthi and L. Reyzin. Key agreement from close secrets over unsecured channels. In *EUROCRYPT*, pages 206–223, 2009.

[Li11] Xin Li. Improved constructions of three source extractors. In *Proceedings of the 26th Annual IEEE Conference on Computational Complexity*, 2011.

[Li12] Xin Li. Design extractors, non-malleable condensers and privacy amplification. In *Proceedings of the 44th Annual ACM Symposium on Theory of Computing*, 2012.

[LRVW03] C. J. Lu, Omer Reingold, Salil Vadhan, and Avi Wigderson. Extractors: Optimal up to constant factors. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing*, pages 602–611, 2003.

[MW97] Ueli M. Maurer and Stefan Wolf. Privacy amplification secure against active adversaries. In *CRYPTO '97*, 1997.

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

[Rao06] Anup Rao. Extractors for a constant number of polynomially small min-entropy independent sources. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, 2006.

[Rao07] Anup Rao. An exposition of Bourgain's 2-source extractor. Technical Report TR07-34, ECCC: Electronic Colloquium on Computational Complexity, 2007. `http://eccc.hpi-web.de/eccc-reports/2007/TR07-034/index.html`.

[Raz05] Ran Raz. Extractors with weak random seeds. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing*, pages 11–20, 2005.

[RRV02] Ran Raz, Omer Reingold, and Salil Vadhan. Extracting all the randomness and reducing the error in trevisan's extractors. *JCSS*, 65(1):97–128, 2002.

[RW03] R. Renner and S. Wolf. Unconditional authenticity and privacy from an arbitrarily weak secret. In *CRYPTO*, pages 78–95, 2003.

[SZ99] Aravind Srinivasan and David Zuckerman. Computing with very weak random sources. *SIAM Journal on Computing*, 28:1433–1459, 1999.

[Tre01] Luca Trevisan. Extractors and pseudorandom generators. *Journal of the ACM*, pages 860–879, 2001.

[Vad02] Salil Vadhan. Randomness extractors and their many guises: Invited tutorial. In *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science*, 2002.

[Vaz85] Umesh Vazirani. Towards a strong communication complexity theory or generating quasi-random sequences from two communicating slightly-random sources (extended abstract). In *Proceedings of the 17th Annual ACM Symposium on Theory of Computing*, 1985.

[Zuc07] David Zuckerman. Linear degree extractors and the inapproximability of max clique and chromatic number. In *Theory of Computing*, pages 103–128, 2007.

A The Existence of Generalized Non-Malleable Extractors

In this section we prove the existence of non-malleable extractors with more than one adversarial seeds. First, we have the following definition.

Definition A.1. A function $\mathsf{nmExt} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a (r, k, ε)-non-malleable extractor if, for any source X with $H_\infty(X) \geq k$ and any r function $\mathcal{A}_i : \{0,1\}^d \to \{0,1\}^d, i = 1, \cdots, r$ such that $\mathcal{A}_i(y) \neq y$ for all i and y, the following holds. When Y is chosen uniformly from $\{0,1\}^d$ and independent of X,

$$(\mathsf{nmExt}(X, Y), \{\mathsf{nmExt}(X, \mathcal{A}_i(Y))\}, Y) \approx_\varepsilon (U_m, \{\mathsf{nmExt}(X, \mathcal{A}_i(Y))\}, Y).$$

We will prove the following theorem.

Theorem A.2. *For any constant $r \geq 1$, there exists a (r, k, ε)-non-malleable extractor as long as*

$$d > \frac{3}{2}\log(n-k) + 3\log(1/\epsilon) + O(1)$$

$$k > (r+1)m + \frac{d}{3} + 2\log(1/\epsilon) + \log(d) + O(1)$$

We prove the theorem by using the probabilistic method, similar to the existence proof in [DW09]. A function $f : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a (r, k, ε)-non-malleable extractor if for all (n, k) sources X, all adversarial functions $\{\mathcal{A}_i\}$ and all distinguishers $\mathcal{D}$, we have

$$|\Pr[\mathcal{D}(f(X,Y),\{f(X,\mathcal{A}_i(Y))\},Y)=1]-\Pr[\mathcal{D}(U_m,\{f(X,\mathcal{A}_i(Y))\},Y)=1]|\leq\epsilon.$$

As usual, it suffices to consider flat sources X. For the purpose of a union bound, we fix some $\mathcal{D},\{\mathcal{A}_i\}$ and a source X which is uniformly distributed on some subset $\mathsf{Supp}(X)\subseteq\{0,1\}^n$ with $|\mathsf{Supp}(X)|=2^k$. We use $\mathbf{F}$ to denote the uniform distribution over the space of all functions $f:\{0,1\}^n\times\{0,1\}^d\to\{0,1\}^m$.

For each $u\in\{0,1\}^{rm}$ and $y\in\{0,1\}^d$, define

$$\mathsf{Count}(u,y)=|\{u_2\in\{0,1\}^m:\mathcal{D}(u_2,u,y)=1\}|.$$

For each $x\in\mathsf{Supp}(X)$, $y\in\{0,1\}^d$, define the following random variables (where the randomness comes from $\mathbf{F}$).

$$\mathbf{L}(x,y)=\mathcal{D}(\mathbf{F}(x,y),\{\mathbf{F}(x,\mathcal{A}_i(y))\},y)$$

$$\mathbf{R}(x,y)=\frac{\mathsf{Count}(\{\mathbf{F}(x,\mathcal{A}_i(y))\},y)}{2^m}$$

$$\mathbf{Q}(x,y)=\mathbf{L}(x,y)-\mathbf{R}(x,y)$$

and let

$$\overline{\mathbf{Q}}=\frac{\sum_{x,y}\mathbf{Q}(x,y)}{2^{k+d}}.$$

Thus, $\overline{\mathbf{Q}}$ is essentially the quantity

$$p(f)=\Pr_{f\leftarrow\mathbf{F}}[\mathcal{D}(f(X,Y),\{f(X,\mathcal{A}_i(Y))\},Y)=1]-\Pr_{f\leftarrow\mathbf{F}}[\mathcal{D}(U_m,\{f(X,\mathcal{A}_i(Y))\},Y)=1].$$

Therefore, we want to upper bound

$$\Pr[|\overline{\mathbf{Q}}|>\epsilon]=\Pr_{f\leftarrow\mathbf{F}}[|p(f)|>\epsilon].$$

Again, it is easy to notice that for any (x,y), $\mathbb{E}(\mathbf{L}(x,y))=\mathbb{E}(\mathbf{R}(x,y))$ and thus $\mathbb{E}(\mathbf{Q}(x,y))=0$ and $\mathbb{E}(\overline{\mathbf{Q}})=0$. However, the variables $\mathbf{Q}(x,y)$ are not necessarily independent of each other (in particular, the adversarial seeds can form cycles), so we cannot use a simple Chernoff bound here. Now we represent the functions $\{\mathcal{A}_i\}$ as a directed graph $G=(V,E)$ where the vertex set is $V=\{0,1\}^d$ and there is an edge from y to y' iff $\exists i,\mathcal{A}_i(y)=y'$. Note that each $\mathcal{A}_i$ is a function, thus the out-degree of every vertex is exactly r. The following lemma is proved in [CRS12].

Lemma A.3. *Let $G=(V,E)$ be a directed graph without self-loops. Assume that the out-degree of each vertex is at most r, where parallel edges are allowed. Then, there exists a subset of the vertices $V'\subseteq V$, such that the induced graph $H=(V',E')$ of G is acyclic, and $|V'|\geq|V|/(r+1)$.*

Let $s=1/(r+1)$. We now use Lemma A.3 to decompose G into $t+1$ subgraphs H_j as follows. In each step $j,1\leq j\leq t$, we use the lemma to pick a s fraction of vertices from the remaining vertices to form a subset V_j and an induced graph H_j, and delete these vertices from G. After t steps the remaining graph is H_{t+1}. Thus we have that $|V_j|=s(1-s)^{j-1}|V|$ for $j\leq t$ and $|V_{t+1}|=(1-s)^t|V|$.

The following lemma is proved in [DW09].

Lemma A.4. *For $V' \subseteq V$, let H be the restriction of G to the vertices V' and assume that the graph H is acyclic. Then the set $\{\boldsymbol{Q}(x,y)_{x\in \mathsf{Supp}(X), y\in V'}\}$ of random variables can be enumerated by $\boldsymbol{Q}_1, \cdots, \boldsymbol{Q}_\ell$ for $l = |V'|2^k$ such that $\mathbb{E}[\boldsymbol{Q}_i | \boldsymbol{Q}_1, \cdots, \boldsymbol{Q}_{i-1}] = 0$ for all $1 \le i \le l$.*

We can now prove the theorem.

Proof of Theorem A.2. By Lemma A.3 and Lemma A.4, we can partition $\{\mathbf{Q}(x,y)\}$ into t (enumerated) sets $\{\mathbf{Q}_1^j, \cdots, \mathbf{Q}_{l_j}^j\}$ where $l_j = |V_j|2^k$ for $j = 1, \cdots, t$ and a remaining set $\{\mathbf{Q}(x,y)_{x\in \mathsf{Supp}(X), y\in V_{t+1}}\}$. For each $j \le t$, $1 \le i \le l_j$, we have that $\mathbb{E}[\mathbf{Q}_i^j | \mathbf{Q}_1^j, \cdots, \mathbf{Q}_{i-1}^j] = 0$. Now for each $j \le t$, $1 \le i \le l_j$, define $S_i^j = \sum_{\ell=1}^{i} \mathbf{Q}_i^j$. Then for any $j \le t$, $S_1^j, \cdots, S_{l_j}^j$ is a martingale.

We first show that if for any $j \le t$, $|S_{l_j}^j| \le \frac{\epsilon}{2} l_j$ and $|V_{t+1}| \le \frac{\epsilon}{4} 2^d$, then $|\overline{\mathbf{Q}}| \le \epsilon$. Indeed, note that for any (x,y), $\mathbf{Q}(x,y) \le 2$. Thus in this case we have that

$$
\begin{aligned}
|\overline{\mathbf{Q}}| = \left| \frac{\sum_{j=1}^t S_{l_j}^j + \sum_{x\in \mathsf{Supp}(X), y\in V_{t+1}} \mathbf{Q}_{x,y}}{2^{k+d}} \right| &\le \frac{1}{2^{k+d}} \left(\sum_{j=1}^t \left| S_{l_j}^j \right| + \left| \sum_{x\in \mathsf{Supp}(X), y\in V_{t+1}} \mathbf{Q}_{x,y} \right| \right) \\
&\le \frac{1}{2^{k+d}} \left(\sum_{j=1}^t \frac{\epsilon}{2} l_j + 2\cdot \frac{\epsilon}{4} 2^{d+k} \right) = \frac{1}{2^{k+d}} \left(\frac{\epsilon}{2} \left(\sum_{j=1}^t l_j \right) + \frac{\epsilon}{2} 2^{d+k} \right) \\
&\le \frac{1}{2^{k+d}} \left(\frac{\epsilon}{2} 2^{d+k} + \frac{\epsilon}{2} 2^{d+k} \right) = \epsilon.
\end{aligned}
$$

Next, by Azuma's inequality we have that for any $j \le t$,

$$
\Pr[|S_{l_j}^j| > \frac{\epsilon}{2} l_j] \le e^{-\frac{1}{32} l_j \epsilon^2} \le e^{-\frac{1}{32} s(1-s)^{t-1} 2^{d+k} \epsilon^2}
$$

since $l_j = |V_j|2^k \ge s(1-s)^{t-1}2^{d+k}$.

Therefore, by the union bound,

$$
\Pr[\exists j \le t, |S_{l_j}^j| > \frac{\epsilon}{2} l_j] \le t e^{-\frac{1}{32} s(1-s)^{t-1} 2^{d+k} \epsilon^2}.
$$

Now we will apply the union bound to all possible $X, \{\mathcal{A}_i\}, \mathcal{D}$. Let $N = 2^n, K = 2^k, D = 2^d, M = 2^m$. Then there are $\binom{N}{K}$ possible sources X, there are D^{rD} possible adversaries $\{\mathcal{A}_i\}$ and there are $2^{DM^{r+1}}$ possible distinguishers $\mathcal{D} : \{0,1\}^m \times (\{0,1\}^m)^r \times \{0,1\}^d \to \{0,1\}$. Thus to ensure that there exists a function that is a (r,k,ϵ)-non-malleable extractor, all we need is to satisfy the following inequalities:

$$
\binom{N}{K} D^{rD} 2^{DM^{r+1}} t e^{-\frac{1}{32} s(1-s)^{t-1} 2^{d+k} \epsilon^2} < 1
$$

and

$$
|V_{t+1}| = (1-s)^t 2^d \le \frac{\epsilon}{4} 2^d.
$$

Choose t such that $(1-s)^t = 2^{-\frac{d}{3}}$. Now it is easy to check that both of these conditions are satisfied when the statements in the theorem hold. ■

B Another Construction of Non-Malleable Extractors for Entropy Rate $1/2 - \delta$

Here we give another construction of non-malleable extractors for (n, k) sources with $k = (1/2-\delta)n$ for some constant $\delta > 0$.

Given an (n, k)-source X with $k = (1/2 - \delta)n$, we first pick a prime p such that $2^n < p < 2^{n+1}$. By Bertrand's postulate, there is always such a prime. Now treat X as an element in the field $\mathbb{F}_p$. Next we take an independent and uniform seed $Y \in \{0, 1\}^n$ and again treat Y as an element in $\mathbb{F}_p$. Encode X, Y such that $\mathsf{Enc}(X) = (X, X^2)$ and $\mathsf{Enc}(Y) = (Y, Y^2)$. The operations are in $\mathbb{F}_p$. Our non-malleable extractor is defined as

$$\mathsf{nmExt}(X, Y) = \mathsf{IP}(\mathsf{Enc}(X), \mathsf{Enc}(Y)) \bmod M$$

for some integer $M = 2^m$ that we will choose later. Note that $\mathsf{Enc}(X)$ and $\mathsf{Enc}(Y)$ are vectors in $\mathbb{F}_p^2$ and IP is the inner product function taken over $\mathbb{F}_p$.

Again, we show that for any weak source X with min-entropy $(1/2 - \delta)n$, $3\mathsf{Enc}(X)$ is close to a weak source that has min-entropy $(1/2 + \delta)\log(p^2)$.

Lemma B.1. *Let $\mathbb{F} = \mathbb{F}_p$ for p prime and X be a random variable over $\mathbb{F}$. There is a universal constant $\delta > 0$ such that if X is any weak source with min-entropy $(1/2 - \delta)n$, $3\mathsf{Enc}(X)$ is $p^{-\Omega(1)}$-close to a source with min-entropy $(1/2 + \delta)\log(p^2)$.*

Proof. Note that X has min-entropy $(1/2-\delta)n > (1/2-\delta)\log p - 1$. First consider the distribution $2\mathsf{Enc}(X)$. Note that the distribution is of the form $(X + X, X^2 + X^2)$. For any (a, b) in the support of $2\mathsf{Enc}(X)$, we have that $a = x_1 + x_2$ and $b = x_1^2 + x_2^2$. Thus there are at most 2 different pairs of (x_1, x_2) that satisfy both equations. Therefore the min-entropy of $2\mathsf{Enc}(X)$ is at least $2H_\infty(X) - 1$. Now let $k = H_\infty(X) - 1$, we have that $\mathsf{Enc}(X)$ has min-entropy at least k and $2\mathsf{Enc}(X)$ has min-entropy at least $2k$. We now have the following claim.

Claim B.2. *Let α, β be the two constants in Theorem 3.21. Then $3\mathsf{Enc}(X)$ is $2^{-\Omega(k)}$-close to a source with min-entropy $(1 + \alpha/2)2k$.*

Proof of the claim. Note that an element in the support of $3\mathsf{Enc}(X)$ has the form $(x_1+x_2+x_3, x_1^2+x_2^2+x_3^2)$. This determines the point

$$\begin{aligned}&(x_1 + x_2 + x_3, (x_1 + x_2 + x_3)^2 - (x_1^2 + x_2^2 + x_3^2))\\ =&((x_1 + x_2) + x_3, 2(x_1 + x_2)x_3 + (x_1 + x_2)^2 - (x_1^2 + x_2^2))\end{aligned}$$

Let $a = x_1 + x_2$ and $b = x_1^2 + x_2^2$, this point is

$$(a + x_3, 2ax_3 + a^2 - b).$$

Let $\bar{x}_3 = a + x_3$, then

$$(a + x_3, 2ax_3 + a^2 - b) = (\bar{x}_3, 2a\bar{x}_3 - a^2 - b).$$

For a fixed $(a = x_1 + x_2, b = x_1^2 + x_2^2)$ define the line

$$\ell_{a,b} = \{(x, 2ax - a^2 - b) | x \in \mathbb{F}\}.$$

Note that for different (a, b), the line $\ell_{a,b}$ is also different. Thus we have a set of lines $L = \{\ell_{a,b}\}$. Note that x_3 is sampled from X_3, which has min-entropy k and (a, b) is sampled from $\mathsf{Enc}(X_1) + \mathsf{Enc}(X_2)$, which has min-entropy $2k$. Further note that these two distributions are independent. Since every weak source with min-entropy k is a convex combination of flat k sources, without loss of generality we can assume that X_3 and $\mathsf{Enc}(X_1) + \mathsf{Enc}(X_2)$ are both flat sources. Thus L has size 2^{2k}.

Now assume that $3\mathsf{Enc}(X)$ is ϵ-far from any source with min-entropy $(1+\alpha/2)2k$. Since $3\mathsf{Enc}(X)$ determines the distribution $(A+X_3, 2AX_3+A^2-B)$, this distribution is also ϵ-far from any source with min-entropy $(1+\alpha/2)2k$. Thus there must exist some set M of size at most $2^{(1+\alpha/2)2k}$ such that

$$\Pr_{(a,b)\leftarrow 2\mathsf{Enc}(X), X_3 \leftarrow X}[(a + x_3, 2ax_3 + a^2 - b) \in M] \geq \epsilon.$$

Note that whenever $(a + x_3, 2ax_3 + a^2 - b) \in M$, this point has an incidence with the line $\ell_{a,b}$. Further note that whenever (a, b) is different or x_3 is different, the incidence is also different. Thus by the above inequality the number of incidences between the set of points M and the set of lines L is at least

$$\Pr_{(a,b)\leftarrow 2\mathsf{Enc}(X), X_3 \leftarrow X}[(a + x_3, 2ax_3 + a^2 - b) \in M]2^k 2^{2k} \geq \epsilon 2^{3k}.$$

On the other hand, since L has size 2^{2k} and M has size $2^{(1+\alpha/2)2k} \leq 2^{(1+\alpha/2)2(1/2-\delta)\log p} < 2^{(1+\alpha/2)\log p} \leq p^{2-\beta}$, by Theorem 3.21, the number of incidences between M and L is at most $O(2^{(3/2-\alpha)(2+\alpha)k}) < 2^{3k(1-\alpha/6)} = 2^{-\alpha k/2}2^{3k}$.

Thus we must have $\epsilon < 2^{-\alpha k/2}$. ∎

By choosing δ appropriately and noting that $k \geq (1/2 - \delta)\log p - 2$, the lemma is proved. ∎

Now we can use the non-uniform XOR lemma to argue that our extractor is non-malleable. Specifically, we have the following lemma.

Lemma B.3. *Let δ be the constant in Lemma B.1. Given any (n, k)-source X with $k = (1/2-\delta)n$, and Y an independent source over $\{0, 1\}^n$ with min-entropy $(1 - \delta)n$, let $W = \mathsf{IP}(\mathsf{Enc}(X), \mathsf{Enc}(Y))$ and $W' = \mathsf{IP}(\mathsf{Enc}(X), \mathsf{Enc}(Y'))$ where $Y' = \mathcal{A}(Y)$ and $\forall y \in \{0, 1\}^n, \mathcal{A}(y) \neq y$. For any two characters $\psi(s) = e^{2\pi its/p}$ and $\psi'(s) = e^{2\pi it's/p}$ where $t, t' \in \mathbb{F}_p$ and $t \neq 0$,*

$$|E_{W,W'}[\psi(W)\psi'(W')]| \leq 2^{-\Omega(n)}.$$

Proof. Note that W, W' are deterministic functions of X, Y. Thus

$$E_{W,W'}[\psi(W)\psi'(W')] = E_{X,Y}[\psi(W)\psi'(W')].$$

Depending on whether ψ' is trivial, we have two cases.

Case 1: $t' = 0$. This corresponds to the case where ψ' is the trivial character. In this case $\psi'(W')$ is always 1. Thus

$$E_{W,W'}[\psi(W)\psi'(W')] = E_{X,Y}[\psi(W)] = E_{X,Y}[\psi(\mathsf{Enc}(X) \cdot \mathsf{Enc}(Y))].$$

Note that $\mathsf{Enc}(Y)$ has the same min-entropy as Y, which is $(1-\delta)n$. Now consider $\mathsf{Enc}(X)$. Since X has min-entropy $(1/2-\delta)n$, by Lemma B.1 $3\mathsf{Enc}(X)$ is $p^{-\Omega(1)}$-close to having min-entropy $(1/2+\delta)\log(p^2)$. Now note that the min-entropy of $4\mathsf{Enc}(X)-4\mathsf{Enc}(X)$ is at least the min-entropy of $4\mathsf{Enc}(X)$, and which in turn is at least the min-entropy of $3\mathsf{Enc}(X)$. Thus $4\mathsf{Enc}(X)-4\mathsf{Enc}(X)$ is $p^{-\Omega(1)}$-close to having min-entropy $(1/2+\delta)\log(p^2)$. Since $(1/2+\delta)\log(p^2)+(1-\delta)n > (1+2\delta)\log p + (1-\delta)(\log p - 1) > (2+\delta)\log p - 1$, by Lemma 3.20 we have

$$|E_{W,W'}[\psi(W)\psi'(W')]| = |E_{X,Y}[\psi(\mathsf{Enc}(X) \cdot \mathsf{Enc}(Y))]| \leq (p^2 2^{1-(2+\delta)\log p})^{1/16} + p^{-\Omega(1)} = 2^{-\Omega(n)}.$$

Case 2: $t' \neq 0$. This corresponds to the case where ψ' is non-trivial. In this case, note that

$$\psi(W)\psi'(W') = e^{2\pi i t(\mathsf{Enc}(X)\cdot\mathsf{Enc}(Y))} e^{2\pi i t'(\mathsf{Enc}(X)\cdot\mathsf{Enc}(Y'))} = e^{2\pi i t(\mathsf{Enc}(X)\cdot(\mathsf{Enc}(Y)+r\mathsf{Enc}(Y')))},$$

where $r = t'/t \in \mathbb{F}_p$ and $r \neq 0$ since $t \neq 0$ and $t' \neq 0$.

Let $\widetilde{\mathsf{Enc}(Y)} = \mathsf{Enc}(Y) + r\mathsf{Enc}(Y')$, then

$$E_{W,W'}[\psi(W)\psi'(W')] = E_{X,Y}[\psi(W)\psi'(W')] = E_{X,Y}[\psi(\mathsf{Enc}(X) \cdot \widetilde{\mathsf{Enc}(Y)})].$$

Now again by the same argument as above we have that $4\mathsf{Enc}(X)-4\mathsf{Enc}(X)$ is $p^{-\Omega(1)}$-close to having min-entropy $(1/2+\delta)\log(p^2)$. Now we only need to bound the min-entropy of $\widetilde{\mathsf{Enc}(Y)}$.

If for every two different y_1, y_2, we have that $\mathsf{Enc}(y_1) + r\mathsf{Enc}(y_1') \neq \mathsf{Enc}(y_2) + r\mathsf{Enc}(y_2')$, then obviously $\widetilde{\mathsf{Enc}(Y)}$ will have the same min-entropy as Y. Now assume that for some two different y_1, y_2, we have $\mathsf{Enc}(y_1) + r\mathsf{Enc}(y_1') = \mathsf{Enc}(y_2) + r\mathsf{Enc}(y_2')$.

This gives us

$$y_1 + ry_1' = y_2 + ry_2'$$

and

$$(y_1)^2 + r(y_1')^2 = (y_2)^2 + r(y_2')^2.$$

Hence we get

$$y_1 - y_2 = r(y_2' - y_1')$$

and

$$(y_1+y_2)(y_1-y_2) = r(y_2'+y_1')(y_2'-y_1').$$

Since $y_1 \neq y_2$ and $r \neq 0$, we must have that $y_1' \neq y_2'$. Thus we get

$$y_1 + y_2 = y_2' + y_1'.$$

Therefore we can completely solve the equations and get

$$y_1' = ((r+1)y_2 + (r-1)y_1)/2r, \;\; y_2' = ((r+1)y_1 + (r-1)y_2)/2r.$$

Thus any element in $\mathsf{Supp}(\widetilde{\mathsf{Enc}(Y)})$ can come from at most 2 elements in $\mathsf{Supp}(Y)$. To see this, assume for the sake of contradiction that we have $\mathsf{Enc}(y_1) + r\mathsf{Enc}(y_1') = \mathsf{Enc}(y_2) + r\mathsf{Enc}(y_2') = \mathsf{Enc}(y_3) + r\mathsf{Enc}(y_3')$ for three different y_1, y_2, y_3. Thus by above we have

$$y_1' = ((r+1)y_2 + (r-1)y_1)/2r$$

and

$$y_1' = ((r+1)y_3 + (r-1)y_1)/2r.$$

Note that $r \neq -1$ since otherwise this would imply that $y_1' = y_1$ which contradicts the assumption that $\forall y, \mathcal{A}(y) \neq y$. Thus we get $y_2 = y_3$, another contradiction.

Therefore the min-entropy of $\widetilde{\mathsf{Enc}(Y)}$ is at least $H_\infty(Y) - 1 = (1-\delta)n - 1$. Now since $(1/2 + \delta)\log(p^2) + (1-\delta)n - 1 > (1+2\delta)\log p + (1-\delta)(\log p - 1) - 1 > (2+\delta)\log p - 2$, by Lemma 3.20 we have

$$|E_{W,W'}[\psi(W)\psi'(W')]| = |E_{X,Y}[\psi(\mathsf{Enc}(X) \cdot \widetilde{\mathsf{Enc}(Y)})]| \leq (p^2 2^{2-(2+\delta)\log p})^{1/16} + p^{-\Omega(1)} = 2^{-\Omega(n)}.$$

□

Now we can prove the following theorem.

Theorem B.4. *Let δ be the constant from Lemma B.1. Given any (n,k) source X with $k = (1/2 - \delta)n$ and an independent uniform seed $Y \in \{0,1\}^n$, as well as any deterministic function $\mathcal{A} : \{0,1\}^n \to \{0,1\}^n$ such that $\forall y, \mathcal{A}(y) \neq y$,*

$$|(\mathsf{nmExt}(X,Y), \mathsf{nmExt}(X, \mathcal{A}(Y)), Y) - (U_m, \mathsf{nmExt}(X, \mathcal{A}(Y)), Y)| \leq \epsilon,$$

where $\epsilon = 2^{-\Omega(n)}$ and output size $m = \Omega(n)$.

Proof. Let $Z = \mathsf{nmExt}(X,Y)$ and $Z' = \mathsf{nmExt}(X, \mathcal{A}(Y))$. By Lemma B.3 and Lemma 3.14, we can choose an $m = \Omega(n)$ and $M = 2^m$ such that when $\mathsf{nmExt}(X,Y) = \mathsf{IP}(\mathsf{Enc}(X), \mathsf{Enc}(Y)) \bmod M$ and Y is an $(n, (1-\delta)n)$ source independent of X, we have

$$|(Z, Z') - (U_m, Z')| \leq \epsilon',$$

where $\epsilon' = O(n2^m 2^{-\Omega(n)} + 2^{m-n}) = 2^{-\Omega(n)}$.

Therefore when Y is an independent uniform distribution over $\{0,1\}^n$, by Theorem 3.17 we have

$$|(Z, Z', Y) - (U_m, Z', Y)| \leq \epsilon,$$

where $\epsilon = 2^{2m}(2^{1-\delta n} + \epsilon')$.

Note that $\epsilon' = O(n2^m 2^{-\Omega(n)} + 2^{m-n})$. Thus we can take $m = \Omega(n)$ and $\epsilon = 2^{2m}(2^{1-\delta n} + \epsilon') = 2^{-\Omega(n)}$. Thus the theorem is proved. ■