PRIVACY AMPLIFICATION AND NON-MALLEABLE EXTRACTORS
VIA CHARACTER SUMS∗

YEVGENIY DODIS†, XIN LI‡, TREVOR D. WOOLEY§, AND DAVID ZUCKERMAN ¶

Abstract. In studying how to communicate over a public channel with an active adversary, Dodis and Wichs introduced the notion of a non-malleable extractor. A non-malleable extractor dramatically strengthens the notion of a strong extractor. A strong extractor takes two inputs, a weakly-random \(x\) and a uniformly random seed \(y\), and outputs a string which appears uniform, even given \(y\). For a non-malleable extractor \(\text{nmExt}\), the output \(\text{nmExt}(x, y)\) should appear uniform given \(y\) as well as \(\text{nmExt}(x, A(y))\), where \(A\) is an arbitrary function with \(A(y) \neq y\).

We show that an extractor introduced by Chor and Goldreich is non-malleable when the entropy rate (the ratio between the entropy and the length of the weakly random string) is above half. It outputs a linear number of bits when the entropy rate is \(1/2 + \alpha\), for any \(\alpha > 0\). Previously, no explicit construction was known for any entropy rate less than 1. To achieve a polynomial running time when outputting more than one bits, we rely on a widely-believed conjecture about the distribution of prime numbers in arithmetic progressions. Our analysis involves character sum estimates, which may be of independent interest.

Using our non-malleable extractor, we obtain protocols for “privacy amplification”: key agreement between two parties who share a weakly-random secret. Our protocols work in the presence of an active adversary with unlimited computational power, and have asymptotically optimal entropy loss. When the secret has entropy rate greater than \(1/2\), the protocol follows from a result of Dodis and Wichs, and takes two (or three; for strongest security guarantees) rounds. When the secret has entropy rate \(\delta\) for any constant \(\delta > 0\), our new protocol takes a constant (polynomial in \(1/\delta\)) number of rounds. Our protocols run in polynomial time under the above well-known conjecture about primes.

Key words. randomness, extractors, non-malleable, privacy amplification, cryptography

AMS subject classifications. 68Q10, 94A60, 68W20

1. Introduction. Bennett, Brassard, and Robert [3] introduced the basic cryptographic question of privacy amplification. Suppose Alice and Bob share an \(n\)-bit secret key \(X\), which is weakly random. This could occur because the secret is a password or biometric data, neither of which is uniformly random, or because an adversary Eve managed to learn some information about a secret which previously was uniformly random. How can Alice and Bob communicate over a public channel to transform \(X\) into a nearly uniform secret key, about which Eve has negligible information? We measure the randomness in \(X\) using min-entropy.

†Department of Computer Science, New York University, New York, NY 10012, U.S.A. (dodis@cs.nyu.edu). Partially supported by NSF Grants CNS-1065288, CNS-1017471, CNS-0831299 and Google Faculty Award.
‡Department of Computer Science, Johns Hopkins University, Baltimore, MD 21218, U.S.A. (lixints@cs.jhu.edu). Partially supported by NSF Grants CCF-0634811 and CCF-0916160 and THECB ARP Grant 003658-0113-2007. Much of the work was done while the author was a graduate student at University of Texas at Austin.
§School of Mathematics, University of Bristol, University Walk, Clifton, Bristol BS8 1TW, United Kingdom. (matdw@bristol.ac.uk). Supported by a Royal Society Wolfson Research Merit Award.
¶Department of Computer Science, University of Texas at Austin, Austin, TX 78701, U.S.A. (diz@cs.utexas.edu). Partially supported by NSF Grants CCF-0634811 and CCF-0916160 and THECB ARP Grant 003658-0113-2007.
**Definition 1.1.** The min-entropy of a random variable $X$ is

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

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 assume Eve has unlimited computational power. If Eve is passive, i.e., cannot corrupt the communication between Alice and Bob, then it is not hard to use randomness extractors [23] to solve this problem. In particular, a strong extractor suffices.

**Notation.** We let $[s]$ denote the set $\{1,2,\ldots,s\}$. For $\ell$ a positive integer, $U_\ell$ 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 $U_\ell$ or $U_S$ is assumed independent of the other components. We say $W \approx_z Z$ if the random variables $W$ and $Z$ have distributions which are $\epsilon$-close in variation distance.

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

$$\langle \text{Ext}(X,Y),Y \rangle \approx_\epsilon (U_m,Y).$$

Using such an extractor, the case when Eve is passive can be solved as follows. Alice chooses a fresh random string $Y$ and sends it to Bob. They then both compute $\text{Ext}(X,Y)$. The property of the strong extractor guarantees that even given $Y$, the output is close to uniform.

The case when Eve is active, i.e., can corrupt the communication, has recently received attention. Maurer and Wolf [22] gave a one-round protocol which works when the entropy rate of the weakly-random secret $X$ is bigger than $2/3$. This was later improved by Dodis, Kanukurthi, Katz, Reyzin, and Smith [8, 18, 7] to work for entropy rate bigger than $1/2$. However in both cases the resulting nearly-uniform secret key $R$ is significantly shorter than the min-entropy of $X$. (If $R$ only has length roughly $2k - n$ if $X$ is an $(n,k)$-source.) Dodis and Wichs [12] showed that for entropy rate less than $1/2$ there exists no one-round protocol. Renner and Wolf [26] gave the first protocol for entropy rate below $1/2$ which requires $O(s)$ rounds to achieve security parameter $s$. (Recall that a protocol achieves security parameter $s$ if Eve cannot predict with advantage more than $2^{-s}$ over random. For an active adversary, we further require that Eve cannot force Alice and Bob to output different secrets and not abort with probability more than $2^{-s}$.) Kanukurthi and Reyzin [19] simplified their protocol and showed that the protocol can achieve entropy loss (the difference between the entropy of $X$ and the length of the secret key $R$) $O(s^2)$. Dodis and Wichs [12] improved the number of rounds to 2 but did not improve the entropy loss. Chandran, Kanukurthi, Ostrovsky, and Reyzin [4] improved the entropy loss to $O(s)$ but the number of rounds remained $O(s)$. The natural open question is therefore whether there is a 2-round protocol with entropy loss $O(s)$.

Dodis and Wichs showed how such a protocol could be built using non-malleable extractors, which they defined. In the following definition of (worst-case) non-malleable extractor, think of an adversary changing the value of the seed via the function $A$.

**Definition 1.3.** A function $\text{nmExt} : [N] \times [D] \rightarrow [M]$ is a $(k,\epsilon)$-non-malleable extractor if, for any source $X$ with $H_\infty(X) \geq k$ and any function $A : [D] \rightarrow [D]$ such
that \( A(y) \neq y \) for all \( y \), the following holds. When \( Y \) is chosen uniformly from \([D]\) and independent of \( X \),

\[
(nmExt(X, Y), nmExt(X, A(Y)), Y) \approx_\varepsilon (U_M, nmExt(X, A(Y)), Y).
\]

Note that this dramatically strengthens the definition of strong extractor. In a strong extractor, the output must be indistinguishable from uniform, even given the random seed. For a non-malleable extractor, a distinguisher is not only given a random seed, but also the output of the extractor with the given input and an arbitrarily correlated random seed. Note that \( nmExt(X, A(Y)) \) need not be close to uniform. The above “worst-case” definition is slightly weaker than the “average-case” definition needed by applications, but Dodis and Wichs showed that any worst-case \((k, \varepsilon)\)-non-malleable extractor is also an average-case \((k - \log(1/\varepsilon), 2\varepsilon)\)-non-malleable extractor. See Subsection 3.2.

Unfortunately, Dodis and Wichs were not able to construct such non-malleable extractors. Instead, they constructed “look-ahead extractors,” which are weaker than non-malleable extractors, but nevertheless yielded the two-round, \( O(s^2) \)-entropy loss protocol mentioned above.

Dodis and Wichs also showed the existence of non-malleable extractors. The existence of excellent standard randomness extractors can be shown by the probabilistic method in a straightforward way. For non-malleable extractors, the argument requires more work. Nevertheless, Dodis and Wichs showed that non-malleable extractors exist with \( k > 2m + 3\log(1/\varepsilon) + \log d + 9 \) and \( d > \log(n - k + 1) + 2\log(1/\varepsilon) + 7 \), for \( N = 2^n, M = 2^m \), and \( D = 2^d \).

The definition of non-malleable extractor is so strong that before our work, no explicit construction was known for any length seed achieving a one-bit output, even for min-entropy \( k = .99n \). For example, a first attempt might be \( f(x, y) = x \cdot y \), where the inner product is taken over \( GF(2) \). However, this fails, even for min-entropy \( n - 1 \). To see this, take \( X \) to be the bit 0 concatenated with \( U_{n-1} \). Let \( A(y) \) be \( y \) with the first bit flipped. Then for all \( x \) in the support of \( X \), one has \( f(x, y) = f(x, A(y)) \).

Although general Hadamard codes don’t work, we nevertheless show that a specific near-Hadamard code that comes from the Paley graph works for min-entropy \( k > n/2 \). The Paley graph function is \( nmExt(x, y) = \chi(x - y) \), where \( x \) and \( y \) are viewed as elements in a finite field \( \mathbb{F} \) of odd order \( q \) and \( \chi \) is the quadratic character \( \chi(x) = x^{(q-1)/2} \). (The output of \( \chi \) is in \( \{\pm 1\} \), which we convert to an element of \( \{0, 1\} \).) The function \( nmExt(x, y) = \chi(x + y) \) works equally well. The proof involves estimating a nontrivial character sum.

We can output \( m \) bits by computing the discrete logarithm \( \log_q (x + y) \mod M \). This extractor was originally introduced by Chor and Goldreich [5] in the context of two-source extractors. To make this efficient, we need \( M \) to divide \( q - 1 \). A widely-believed conjecture about primes in arithmetic progressions implies that such a \( q \) is not too large (see Conjecture 3.7). Our result is stated as follows.

**Theorem 1.4.** For any constants \( \alpha, \beta, \gamma > 0 \) with \( \beta + \gamma < \alpha/2 \), there is an explicit \((k = (1/2 + \alpha)n, \varepsilon)\)-non-malleable extractor \( nmExt : \{0, 1\}^n \times \{0, 1\}^d \to \{0, 1\}^m \) for \( \varepsilon = 2^{-\gamma n} \) and any \( m \leq \beta n \). It has seed length \( d = n \) and runs in polynomial time if Conjecture 3.7 holds or \( m = O(\log n) \).

As a direct corollary of Theorem 1.4 and the protocol of Dodis and Wichs, we obtain a 2-round protocol for privacy amplification with optimal entropy loss,\(^1\) when

\(^1\)For so called “pre-application robustness”; we need a third round to achieve stronger “post-
application robustness".}

2Unlike the two-round protocol, this protocol right away achieves the stronger notion of “post-application robustness".
bits. He further showed that non-malleable condensers themselves give 2-round privacy amplification protocols with optimal entropy loss. However, the constructions of non-malleable condensers in [20] also only work for min-entropy rate $1/2 + \alpha$.

Li [21], as well as Dodis and Yu [13] found a new construction of non-malleable extractors for min-entropy rate $1/2 + \alpha$. With some other ideas, Li [21] additionally gave the first construction of non-malleable extractors that break the entropy rate $1/2$ barrier. More specifically, he constructed non-malleable extractors for min-entropy rate $1/2 - \alpha$ for some universal constant $\alpha > 0$. However, the problem of constructing non-malleable extractors for smaller min-entropy rate (e.g., an arbitrary constant $\delta > 0$) remains open.

**Organization.** We begin with an overview of our privacy amplification protocol in Section 2. (Readers interested only in the non-malleable extractor can skip this section.) We discuss some preliminaries in Section 3, the non-malleable extractor in Section 4, and the character sum estimate in Section 5. Finally, we give full details of the privacy amplification protocol in Section 6. In Appendix A we give a generalization of the non-malleable extractor.

2. Overview of the Protocol for Privacy Amplification. We first describe Dodis and Wichs’ optimal two-round protocol using a non-malleable extractor. The protocol also uses a cryptographic primitive: a one-time message authentication code (MAC). Roughly speaking, a MAC uses a private uniformly random key $R$ to produce a tag $T$ for a message $m$, such that without knowing the key, the probability that an adversary can guess the correct tag $T'$ for another message $m' \neq m$ is small, even given $m$ and $T$.

Now assume that we have a non-malleable extractor $nmExt$ that works for any $(n,k)$-source $X$. Then there is a very natural two-round privacy amplification protocol. In the first round Alice chooses a fresh random string $Y$ and sends it to Bob. Bob receives a possibly modified string $Y'$. They then compute $R = nmExt(X,Y)$ and $R' = nmExt(X,Y')$ respectively. In the second round, Bob chooses a fresh random string $W'$ and sends it to Alice, together with $T' = MAC_R(W')$ by using $R'$ as the MAC key. Alice receives a possibly modified version $(W,T)$, and she checks if $T = MAC_R(W)$. If not, then Alice aborts; otherwise they compute outputs $Z = Ext(X,W)$ and $Z' = Ext(X,W')$ respectively, where $Ext$ is a seeded strong extractor. The protocol is depicted in Figure 6.1.

The analysis of the above protocol is also simple. If Eve does not change $Y$, then $R = R'$ and is (close to) uniform. Therefore by the property of the MAC the probability that Eve can change $W'$ without being detected is very small. On the other hand if Eve changes $Y$, then by the property of the non-malleable extractor, one finds that $R'$ is (close to) independent of $R$. Thus in this case, again the probability that Eve can change $W'$ without being detected is very small. In fact, in this case Eve cannot even guess the correct MAC for $W'$ with a significant probability.

The above protocol is nice, except that we only have non-malleable extractors for entropy rate $> 1/2$. As a direct corollary this gives our 2-round privacy amplification protocol for entropy rate $> 1/2$. To get a protocol for arbitrary positive entropy rate, we have to do more work.

We start by converting the shared weak random source $X$ into a somewhere high min-entropy rate source. The conversion uses recent condensers built from sum-product theorems. Specifically, any $n$-bit weak random source with linear min-entropy can be converted into a matrix with a constant number of rows, such that at least
one row has entropy rate 0.9. Moreover each row still has $\Theta(n)$ bits. Note that since Alice and Bob apply the same function to the shared weak random source, they now also share the same rows.

Now it is natural to try the two-round protocol for each row and hope that it works on the row with high min-entropy rate. More specifically, for each row $i$ we have a two round protocol that produces $R_i, R_i'$ in the first round and $Z_i, Z_i'$ in the second round. Now let $g$ be the first row that has min-entropy rate 0.9. We hope that $Z_g = Z'_g$ with high probability, and further that $Z_g, Z_g'$ are close to uniform and private. This is indeed the case if we run the two round protocol for each row sequentially (namely we run it for the first row, and then the second row, the third row, and so on), and can be argued as follows.

Assume the security parameter we need to achieve is $s$, so each of $R_i, R_i'$ has $O(\log n + s)$ bits by the property of the MAC. As long as $s$ is not too large, we can fix all these random variables up to row $g − 1$, and argue that row $g$ still has min-entropy rate $> 1/2$ (since each row has $\Theta(n)$ bits we can actually achieve a security parameter up to $s = \Omega(n)$). Note that we have essentially fixed all the information about $X$ that can be leaked to Eve. Therefore now for row $g$ the protocol succeeds and thus $Z_g = Z'_g$ with high probability, and $Z_g, Z_g'$ are close to uniform and private.

However, we don’t know which row is the good row. We now modify the above protocol to ensure that, once we reach the first good row $g$, for all subsequent rows $i$, with $i > g$, we will have that $Z_i = Z'_i$ with high probability, and further $Z_i, Z'_i$ are close to uniform and private. If this is true then we can just use the output for the last row as the final output.

To achieve this, the crucial observation is that once we reach a row $i − 1$ such that $Z_{i−1} = Z'_{i−1}$, and $Z_{i−1}, Z'_{i−1}$ are close to uniform and private, then $Z'_{i−1}$ can be used as a MAC key to authenticate $W'_i$ for the next row. Now if $W'_i = W_i$ for row $i$, then $Z_i = Z'_i$ and $Z_i, Z'_i$ will also be close to uniform and private. Therefore, we modify the two-round protocol so that in the second round for row $i$, not only do we use $T'_i = \text{MAC}_{R'_i}(W'_i)$ to authenticate $W'_i$, but also we use $L'_i = \text{MAC}_{Z'_{i−1}}(W'_i)$ to authenticate $W'_i$.

This would have worked given that $Z_{i−1} = Z'_{i−1}$, and $Z_{i−1}, Z'_{i−1}$ are close to uniform and private, except for another complication. The problem is that now $T'_i = \text{MAC}_{R'_i}(W'_i)$ could leak information about $Z_{i−1}$ to Eve, so $Z_{i−1}$ is no longer private. Fortunately, there are known constructions of MACs that work even when the key is not uniform, but instead only has large enough average conditional min-entropy in the adversary’s view. Specifically, Theorem 6.9 indicates that the security parameter of this MAC is roughly the average conditional min-entropy of the key minus half the key length, and the key length is roughly twice as long as the length of the tag. Therefore, we can choose a small tag length for $T'_i = \text{MAC}_{R'_i}(W'_i)$, and a large tag length for $L'_i = \text{MAC}_{Z'_{i−1}}(W'_i)$. For example, if the tag length for $T'_i$ is 2s, and the tag length for $T'_2$ is 4s, then the key length for $L'_i$ is 8s. Thus the average min-entropy of $Z_{i−1}$ conditioned on $T'_i$ is $8s − 2s = 6s$, and we can still achieve a security parameter of $6s − 4s = 2s$.

Finally, the discussion so far implicitly assumed that Eve follows a natural “synchronous” scheduling, where she never tries to get one party out-of-sync with another party. To solve this problem, after each Phase $i$ Bob performs a “liveness” test, where Alice has to respond to a fresh extractor challenge from Bob to convince Bob that

---

3 In fact, the result is (close to) a convex combination of such matrices. For simplicity, however, we can assume that the result is just one such matrix, since it does not affect the analysis.
Alice is still “present” in this round. This ensures that if Bob completes the protocol, Alice was “in-sync” with Bob throughout. However, Eve might be able to make Alice be out-of-sync with Bob, causing Alice to output a non-random key (and Bob reject). To solve this last problem, we add one more round at the end which ensures that Alice always outputs a random key (and Bob either outputs the same key or rejects).

With this modification, the complete protocol is depicted in Figure 6.2. Essentially, for the first good row, the property of the non-malleable extractor guarantees that Eve cannot change \( W' \) with significant probability. For all subsequent rows, by using the output \( Z'_{i-1} \) from the previous row as the MAC key, the property of the MAC guarantees that Eve cannot change \( W'_i \) with significant probability. Therefore, the output for the last row can be used to authenticate the last seed of the extractor chosen by Alice (for the reason mentioned above) to produce the final output.

Finally, we note that our final protocol has \( O(1) \) rounds and achieves asymptotically optimal entropy loss is \( O(s + \log n) \), for security parameter \( s \).

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

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

We say \( W \) is \( \varepsilon \)-close to \( Z \), denoted \( W \approx_{\varepsilon} Z \), if \( \Delta(W, Z) \leq \varepsilon \). 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 \( \ell_1 \) norm of the difference of the distributions specified by the random variables \( W \) and \( Z \), which equals \( 2\Delta(W, Z) \).

3.2. 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 [11].

Definition 3.2. The average conditional min-entropy is defined as

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

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 3.3 ([11]). For any \( s > 0 \), \( \Pr_{w \leftarrow W}[H_\infty(X|W = w) \geq \bar{H}_\infty(X|W) - s] \geq 1 - 2^{-s} \).

Lemma 3.4 ([11]). If a random variable \( B \) has at most \( 2^t \) possible values, then \( \bar{H}_\infty(A|B) \geq H_\infty(A) - t \).
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 3.5.** A \((k, \varepsilon)\)-average-case non-malleable extractor is a \((k, \varepsilon)\)-worst-case non-malleable extractor. For any \(s > 0\), a \((k, \varepsilon)\)-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.

### 3.3. Primes in arithmetic progressions

To output more than \(\log n\) bits, we will rely on a well-known conjecture about primes in arithmetic progressions. We begin with a definition.

**Definition 3.6.** Let \(p(r,a)\) denote the least prime in the arithmetic progression \(a \mod r\).

We can now state a special case of a well-known conjecture. We begin with a definition.

**Conjecture 3.7.** There exists a constant \(c > 0\), such that for \(r\) a power of 2, and \(a = 1\), one has \(p(r,a) = O(r \log^c r)\).

We don’t really need \(r\) to be a power of 2; it would suffice if the conjecture held for integers \(r_n\), where \(r_n\) is a smooth integer of about \(n\) bits computable in time polynomial in \(n\). This conjecture is widely believed for \(c = 2\), all \(r\), and all \(a\) relatively prime to \(r\). For more on this conjecture, see, for example, the discussion following equation (1) of [16]. The best unconditional conclusion is substantially weaker. Thus, one has \(p(r,a) = O(r^{5/2})\) (see [30, 17]).

### 3.4. Fourier analysis

The following definitions from Fourier analysis are standard (see e.g., [28]), although we normalize differently than in many computer science papers, such as [24]. 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 \(S\), which we also view as a function from \(S\) to \(\mathbb{R}\). Note that \(E_D[|f(D)|] = \langle f,D \rangle\). Now suppose we have functions \(h : S \to T\) and \(g : T \to \mathbb{C}\). Then

\[
\langle g \circ h,D \rangle = E_D[g(h(D))] = \langle g,h(D) \rangle.
\]

Let \(G\) be a finite abelian group, and let \(\phi\) a character of \(G\), i.e., 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 \(\hat{f}(\phi) = \langle f,\phi \rangle\). We let \(\hat{f}\) denote the vector with entries \(\hat{f}(\phi)\) for all \(\phi\). Note that for a distribution \(D\), one has \(\hat{D}(\phi) = E_D[\phi(D)]\).

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

\[
\|\hat{f}\|_{\ell^2}^2 = \langle \hat{f},\hat{f} \rangle = |G|\langle f,f \rangle = |G|\|f\|_{\ell^2}^2.
\]

Hence by Cauchy-Schwarz,

\[
\|f\|_{\ell^1} \leq \sqrt{|G|}\|f\|_{\ell^2} = \|\hat{f}\|_{\ell^2} \leq \sqrt{|G|}\|\hat{f}\|_{\ell^\infty}.
\]

(3.1)

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 \( \hat{f}(\phi, \phi') \) by \( \hat{f}(\phi, \phi') \). Note that

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

### 3.5. A non-uniform XOR lemma.

We’ll need the following extension of Vazirani’s XOR lemma. We can’t use traditional versions of the XOR lemma, because our output may not be uniform. Our statement and proof parallels Rao [24].

**Lemma 3.8.** Let \((W, W')\) be a random variable on \(G \times G\) for a finite abelian group \(G\), and suppose that for all characters \(\phi, \phi'\) on \(G\) with \(\phi\) nontrivial, one has

\[
|E_{(W, W')}[\phi(W) \phi'(W')]| \leq \alpha.
\]

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

**Proof.** As implied in the lemma statement, the value of \(f(a, b)\) is the probability assigned to \((a, b)\) by the distribution of \((W, W')\) minus that assigned by \((U, W')\).

First observe that

\[
\hat{f}(\phi, \phi') = \langle f, (\phi, \phi') \rangle = E_{(W, W')}[\phi(W) \phi'(W') - E_{(W, W')}[\phi(U) \phi'(W')].
\]

Since \(U\) and \(W'\) are independent, this last term equals

\[
E_{(U, W')}[\phi(U)]E_{(W, W')}[\phi'(W')] = E_U[\phi(U)]E_{W'}[\phi'(W')] = 0,
\]

since \(\phi\) is nontrivial. Therefore, by hypothesis, when \(\phi\) is nontrivial, one finds that \(|\hat{f}(\phi, \phi')| \leq \alpha\).

When \(\phi\) is trivial, we get

\[
\hat{f}(\phi, \phi') = E_{(W, W')}[\phi'(W')] = E_{(U, W')}[\phi'(W')] = 0.
\]

Hence \(\|f\|_{\ell\infty} \leq \sqrt{|G|G} ||\hat{f}||_{\ell\infty} \leq |G| \alpha\). 

### 4. The Non-Malleable Extractor.

Our basic extractor was introduced by Chor and Goldreich [5]. They showed that it was a two-source extractor for entropy rates bigger than \(1/2\). Dodis and Oliveira [10] showed that it was strong. Neither result implies anything about non-malleability.

To output \(m\) bits, we set \(M = 2^m\) and choose a prime \(q > M\). In our basic extractor, we require that \(M(\log q - 1)\). Later, we remove this assumption. Fix a generator \(g\) of \(F_q^\times\). We define \(\text{nmExt} : F_q^2 \rightarrow \mathbb{Z}_M\) by \(\text{nmExt}(x, y) = h(\log g(x + y))\).

Here \(\log z\) is the discrete logarithm of \(z\) with respect to \(g\), and \(h : \mathbb{Z}_{q-1} \rightarrow \mathbb{Z}_M\) is given by \(h(x) = x \mod M\).

In the special case \(m = 1\), we only require that \(q\) is odd. In this case, \(\text{nmExt}(x, y)\) corresponds to the quadratic character of \(x + y\), converted to \(\{0, 1\}\) output. This is efficient to compute. Since there is no known efficient deterministic algorithm to find an \(n\)-bit prime, we may take \(q\) to be a prime power, for example \(q = 3^t\), with \(3^{t-1} < 2^n < 3^t\).
For general $M$, we use the Pohlig-Hellman algorithm to compute the discrete log mod $M$. This runs in polynomial time in the largest prime factor of $M$. Since in our case $M = 2^m$, this is polynomial time.

We still need a prime or prime power $q$ such that $M | (q-1)$. Unconditionally, we get a polynomial-time algorithm to output $m = c \log n$ bits for any $c > 0$. To output more bits efficiently, we rely on a widely believed conjecture. Under Conjecture 3.7, such a prime can be found efficiently by testing $M+1, 2M+1, 3M+1, \ldots$ in succession.

Now we prove that $\text{nmExt}$ is a non-malleable extractor.

**Theorem 4.1.** The above function $\text{nmExt} : \mathbb{F}_q^2 \rightarrow \mathbb{Z}_M$ is a $(k, \varepsilon)$-non-malleable extractor for $\varepsilon = Mq^{1/4}2^{1-k/2}$.

**Proof.** The heart of our proof is a new character sum estimate, given in Theorem 5.2 (and Corollary 5.3). We now show how to deduce Theorem 4.1 from the generality that $\text{nmExt}$ is a $(\log d, \varepsilon)$-non-malleable extractor with $\varepsilon = Mq^{1/4}2^{1-k/2}$. If we use a $d$-bit seed where $2^d \leq q < 2^{d+1}$, then we can view the seed as an integer between 0 and $2^d - 1$, or simply as an element of $\mathbb{F}_q$ with min-entropy at least $(\log q) - 1$. We can handle this, and in fact much lower min-entropy in the seed, as follows. First, we recall the Definition 1.6 of a non-malleable extractor with a weakly-random seed. The following lemma shows that a non-malleable extractor with small error remains a non-malleable extractor even if the seed is weakly random.

**Lemma 4.2.** A $(k, \varepsilon)$-non-malleable extractor $\text{nmExt} : [N] \times [D] \rightarrow [M]$ is also a $(k, k', \varepsilon')$-non-malleable extractor with $\varepsilon' = (D/2^k)\varepsilon$.

**Proof.** For $y \in [D]$, let $\varepsilon_y = \Delta((\text{nmExt}(X, y), \text{nmExt}(X, A(y)), y), (U[M], \text{nmExt}(X, A(y)), y))$. Then for $Y$ chosen uniformly from $[D]$,

$$
\varepsilon \geq \Delta((\text{nmExt}(X, Y), \text{nmExt}(X, A(Y)), Y), (U[M], \text{nmExt}(X, A(Y)), Y)) = \frac{1}{D} \sum_{y \in [D]} \varepsilon_y.
$$

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

$$
\Delta((\text{nmExt}(X, Y'), \text{nmExt}(X, A(Y'))), Y'), (U[M], \text{nmExt}(X, A(Y'))), Y'))
= \sum_{y \in [D]} \Pr[Y = y] \varepsilon_y \leq 2^{-k'} \sum_{y \in [D]} \varepsilon_y \leq (D/2^k)\varepsilon.
$$

It is now simple to analyze our non-malleable extractor as a function $\text{nmExt} : \{0, 1\}^n \times \{0, 1\}^d \rightarrow \{0, 1\}^m$. Here we work over $\mathbb{F}_q$, where $q$ is the smallest prime (or prime power) congruent to 1 modulo $M = 2^m$. We let $d = \lceil \log_2 q \rceil$, which is...
n + c \log n + O(1) under Conjecture 3.7. We could even let d = n and the error would only grow by nε.

**Theorem 4.3.** Under Conjecture 3.7 with constant c, for any n, k > n/2 + \((c/2)\log n\), and m < \(k/2 - n/4 - (c/4)\log n\), the above function \(nm\text{Ext} : \{0,1\}^n \times \{0,1\}^m \rightarrow \{0,1\}^m\) is a polynomial-time computable, \((k,\varepsilon)\)-non-malleable extractor for \(\varepsilon = O(n^{c/4}2^m+n/4−k/2)\).

**Proof.** Suppose that Conjecture 3.7 holds for the constant c. Then \(q = O(n^{c/2}2^n)\), and the seed has min-entropy \(k' = d\). Applying lemma 4.2, we obtain error \(\varepsilon = \left(\frac{q}{2^d}\right)m\frac{Mq}{1/4^{1/2}} = O(n^{c/4}2^m+n/4−k/2)\).

After seeing [6], we improved our character sum to handle weak seeds, using ideas related to their work and [25]. In particular, we showed Theorem 5.4, which implies the following theorem.

**Theorem 4.4.** Under Conjecture 3.7, for \(k \geq \left(\frac{1}{2} + \alpha\right)n\) and \(k' \geq \left(\frac{7}{\alpha}\right)(m + \log \varepsilon^{-1}) + 8\log n\), the above function is a \((k,k',\varepsilon)\)-non-malleable extractor.

**Proof.** The theorem follows from Theorem 5.4 in the same way that Theorem 4.1 follows from Theorem 5.2.

### 5. A Character Sum Estimate

We now prove the necessary character sum estimate. We prove a somewhat more general statement than is needed for the one-bit extractor, as the general statement is needed to output many bits. Throughout this section, we take \(F = F_q\) to be a finite field with \(q\) elements. In addition, we suppose that \(\chi : F^\times \rightarrow C^\times\) is a nontrivial character of order \(d = q - 1\), and we extend the domain of \(\chi\) to \(F\) by taking \(\chi(0) = 0\). The following lemma is a consequence of Weil’s resolution of the Riemann Hypothesis for curves over finite fields (see [29]). In this context, we say that a polynomial \(f \in F[x]\) has \(m\) distinct roots when \(f\) has \(m\) distinct roots in the algebraic closure \(\overline{F}\) of \(F\), or equivalently that such holds in a splitting field for \(f\).

**Lemma 5.1.** Suppose that \(f \in F[x]\) is a monic polynomial having \(m\) distinct roots which is not a \(d\)th power in \(F[x]\). Then

\[
\left|\sum_{x \in F} \chi(f(x))\right| \leq (m - 1)\sqrt{q}.
\]

**Proof.** This is immediate from Theorem 2C’ of Schmidt [27] (see page 43 of the latter source).

We next consider two arbitrary characters, where the first is nontrivial; without loss of generality we may take these to be \(\chi_a(x) = (\chi(x))^a\) and \(\chi_b(x) = (\chi(x))^b\), where \(0 < a < q - 1\) and \(0 < b < q - 1\). Now we establish the main character sum estimate. Note that we need the assumption that \(a \neq 0\): if \(a = 0\) and \(b = (q - 1)/2\), we could take \(A(y) = 0\) and let \(S\) be the set of quadratic residues, and then one has no cancellation in the character sum.

#### 5.1. Character Sum for Uniform Seeds

We begin by proving the character sum corresponding to uniformly random seeds. Although this follows from the more general character sum Theorem 5.4 below, the proof is simpler and gives intuition for the general character sum. Moreover, this theorem came before [6], whereas the more general Theorem 5.4 came afterwards.
Theorem 5.2. Suppose that $\mathcal{S}$ is a non-empty subset of $\mathbb{F}$, and that $A : \mathbb{F} \rightarrow \mathbb{F}$ is any function satisfying the property that $A(y) \neq y$ for all $y \in \mathbb{F}$. Then one has

$$\sum_{y \in \mathbb{F}} \left| \sum_{s \in \mathcal{S}} \chi_a(s + y) \chi_b(s + A(y)) \right| \leq 11^{1/4} q^{5/4} |\mathcal{S}|^{1/2}.$$  

Proof. Write $\Theta = \sum_{y \in \mathbb{F}} \left| \sum_{s \in \mathcal{S}} \chi_a(s + y) \chi_b(s + A(y)) \right|$. We begin by applying Cauchy’s inequality to obtain

$$\Theta^2 \leq q \sum_{y \in \mathbb{F}} \left| \sum_{s \in \mathcal{S}} \chi_a(s + y) \chi_b(s + A(y)) \right|^2 = q \sum_{s,t \in \mathcal{S}} \psi_{s,t}(y),$$

in which we have written

$$\psi_{s,t}(y) = \chi_a(s + y) \chi_b(s + A(y)) \bar{\chi}_a(t + y) \bar{\chi}_b(t + A(y)).$$  

(5.1)

Applying Cauchy’s inequality a second time, we deduce that

$$\Theta^4 \leq q^2 |\mathcal{S}|^2 \sum_{s,t \in \mathcal{S}} \left| \psi_{s,t}(y) \right|^2.$$  

By positivity, the sum over $s$ and $t$ may be extended from $\mathcal{S}$ to the entire set $\mathbb{F}$, and thus we deduce that

$$\Theta^4 \leq q^2 |\mathcal{S}|^2 \sum_{s,t \in \mathbb{F}} \omega_{s,t}(y) \bar{\omega}_{s,t}(z).$$  

(5.2)

On recalling the definition (5.1), we may expand the right hand side of (5.2) to obtain the bound

$$\Theta^4 \leq q^2 |\mathcal{S}|^2 \sum_{y,z \in \mathbb{F}} |\nu(y, z)|^2,$$  

(5.3)

where

$$\nu(y, z) = \sum_{s \in \mathbb{F}} \chi_a(s + y) \chi_b(s + A(y)) \bar{\chi}_a(s + z) \bar{\chi}_b(s + A(z)).$$

Recall now the hypothesis that $y \neq A(y)$. It follows that, considered as an element of $\mathbb{F}[x]$, the polynomial

$$h_{y,z}(x) = (x + y)^a (x + A(y))^b (x + z)^{q-1-a}(x + A(z))^{q-1-b}$$

can be a $d$th power only when $y = z$, or when $y = A(z)$, $a = b$ and $z = A(y)$. In order to confirm this assertion, observe first that when $y \neq z$ and $y \neq A(z)$, then $h_{y,z}$ has a zero of multiplicity $a$ at $-y$. Next, when $y = A(z)$, one has $z \neq y$, and so when $a \neq b$ the polynomial $h_{y,z}$ has a zero of multiplicity $q - 1 + a - b$ at $-y$. Finally, when $y = A(z)$ and $a = b$, then provided that $z \neq A(y)$ one finds that $h_{y,z}$ has a zero of multiplicity $q - 1 - a$ at $-z$. In all of these situations it follows that $h_{y,z}$ has a zero of multiplicity not divisible by $d = q - 1$. When $y \neq z$, and $(y, z) \neq (A(z), A(y))$, therefore, the polynomial $h_{y,z}(x)$ is not a $d$th power in $\mathbb{F}[x]$, and has at most 4 distinct roots. In such a situation, it therefore follows from lemma 5.1 that

$$\nu(y, z) = \sum_{s \in \mathbb{F}} \chi(h_{y,z}(s))$$
is bounded in absolute value by \(3\sqrt{q}\). Meanwhile, irrespective of the values of \(y\) and \(z\), the expression \(\nu(y, z)\) is trivially bounded in absolute value by \(q\). Substituting these estimates into (5.3), we arrive at the upper bound

\[
\Theta^4 \leq q^2|S|^2 \sum_{y \in F} \left( |\nu(y, y)|^2 + |\nu(y, A(y))|^2 + \sum_{z \in F \setminus \{y, A(y)\}} |\nu(y, z)|^2 \right)
\]

\[
\leq q^2|S|^2 \sum_{y \in F} (q^2 + q^2 + q(3\sqrt{q})) = 11q^5|S|^2.
\]

We may thus conclude that \(\Theta \leq 11^{1/4}q^{5/4}|S|^{1/2}\). \(\Box\)

**Corollary 5.3.** Under the hypotheses of the statement of Theorem 5.4, one has

\[
\sum_{y \in T} \left| \sum_{s \in S} \chi_a(s + y)\chi_b(s + A(y)) \right| < \eta q|S|
\]

where \(\eta < 2q^{1/4}/|S|^{1/2}\).

### 5.2. Character Sum for Weak Seeds

The work in this subsection came after [6], and uses ideas related to their work and to [25].

**Theorem 5.4.** For \(0 < \alpha, \eta \leq 1/2\), suppose that \(S\) and \(T\) are non-empty subsets of \(F\) with \(|S| \geq q^{1/2+\alpha}\) and \(|T| \geq \max((1/\eta)^{7/\alpha}, (\log q)^8)\), and that \(A : F \to F\) is any function satisfying the property that \(A(y) \neq y\) for all \(y \in F\). Then for large enough \(q\), we have

\[
\sum_{y \in T} \left| \sum_{s \in S} \chi_a(s + y)\chi_b(s + A(y)) \right| < \eta |T||S|.
\]

We prove this by choosing a suitable parameter \(r\) in the following theorem.

**Theorem 5.5.** Suppose that \(S\) and \(T\) are non-empty subsets of \(F\), and that \(A : F \to F\) is any function satisfying the property that \(A(y) \neq y\) for all \(y \in F\). Then for each natural number \(r\), one has

\[
\sum_{y \in T} \left| \sum_{s \in S} \chi_a(s + y)\chi_b(s + A(y)) \right| \leq \lambda_r q^{1/(4r)}|S|^{1-1/(2r)}|T|,
\]

where

\[
\lambda_r = \left((4r - 1)^2 + (2r)^4q|T|^{-r}\right)^{1/(4r)}.
\]

**Proof.** Write

\[
\Theta = \sum_{y \in T} \left| \sum_{s \in S} \chi_a(s + y)\chi_b(s + A(y)) \right|.
\]

We begin by applying Cauchy’s inequality to obtain

\[
\Theta^2 \leq |T| \sum_{y \in T} \left| \sum_{s \in S} \chi_a(s + y)\chi_b(s + A(y)) \right|^2 = |T| \sum_{s,t \in S} \sum_{y \in T} \psi_{s,t}(y),
\]

where \(\psi_{s,t}(y) = \chi_a(s + y)\chi_b(s + A(y))\).
in which we have written
\[ \psi_{s,t}(y) = \chi_a(s+y)\chi_b(s+\mathcal{A}(y))\overline{x}_a(t+y)\overline{x}_b(t+\mathcal{A}(y)). \] (5.4)

Applying Hölder’s inequality, we deduce that
\[ \Theta^{4r} \leq |T|^{2r}|S|^{4r-2} \sum_{s,t \in S} \left| \sum_{y \in T} \psi_{s,t}(y) \right|^{2r}. \]

By positivity, the sum over \( s \) and \( t \) may be extended from \( S \) to the entire set \( \mathbb{F} \), and thus we deduce that
\[ \Theta^{4r} \leq |T|^{2r}|S|^{4r-2} \sum_{s,t \in \mathbb{F}} \sum_{y \in T^{2r}} \prod_{i=1}^{r} \psi_{s,t}(y) \overline{\psi}_{s,t}(y_{r+i}). \] (5.5)

On recalling the definition (5.4), we may expand the right hand side of (5.5) to obtain
\[ \Theta^{4r} \leq |T|^{2r}|S|^{4r-2} \sum_{y \in T^{2r}} |\nu(y)|^2, \] (5.6)

where
\[ \nu(y) = \sum_{s \in \mathbb{F}} \prod_{i=1}^{r} \chi_a(s+y_i)\chi_b(s+\mathcal{A}(y_i))\overline{x}_a(s+y_{r+i})\overline{x}_b(s+\mathcal{A}(y_{r+i})). \]

We now consider the circumstances in which, considered as an element of \( \mathbb{F}[x] \), the polynomial
\[ h_{\mathcal{Y}}(x) = \prod_{i=1}^{r} (x+y_i)^a(x+\mathcal{A}(y_i))^{q-1-a}(x+\mathcal{A}(y_{r+i}))^{q^2-b} \]

is a \( d \)th power. Consider a fixed \( 2r \)-tuple \( \mathcal{Y} \in T^{2r} \) and an index \( i \) with \( 1 \leq i \leq 2r \). If there is no index \( j \) with \( 1 \leq j \leq 2r \) and \( j \neq i \) for which
\[ y_i = y_j \quad \text{or} \quad y_i = \mathcal{A}(y_j), \] (5.7)

then in view of our hypothesis that \( y_i \neq \mathcal{A}(y_i) \), it follows that the polynomial \( h_{\mathcal{Y}}(x) \) has a zero of order precisely \( a \) at \(-y_i\) in the situation where \( 1 \leq i \leq r \), or \( q-1-a \) at \(-y_i\) in the situation where \( r+1 \leq i \leq 2r \). Write \( \mathcal{B} \) for the set of \( 2r \)-tuples \( \mathcal{Y} \in T^{2r} \) having the property that, for each index \( i \) with \( 1 \leq i \leq 2r \), there exists an index \( j \) with \( 1 \leq j \leq 2r \) and \( j \neq i \) for which (5.7) holds. It follows that if \( h_{\mathcal{Y}}(x) \) is to be a \( d \)th power in \( \mathbb{F}[x] \), then one must have \( \mathcal{Y} \in \mathcal{B} \). On the other hand, when \( \mathcal{Y} \in T^{2r} \setminus \mathcal{B} \), the polynomial \( h_{\mathcal{Y}}(x) \) is not a \( d \)th power in \( \mathbb{F}[x] \), and has at most \( 4r \) distinct roots. In such a situation, we therefore deduce from lemma 5.1 that
\[ \nu(\mathcal{Y}) = \sum_{s \in \mathbb{F}} \chi(h_{\mathcal{Y}}(s)) \]
is bounded in absolute value by \((4r-1)\sqrt{q}\). Meanwhile, irrespective of the value of \( \mathcal{Y} \), the expression \( \nu(\mathcal{Y}) \) is trivially bounded in absolute value by \( q \). Substituting these estimates into (5.6), we arrive at the upper bound
\[ \Theta^{4r} \leq |T|^{2r}|S|^{4r-2} \left( \sum_{\mathcal{Y} \in T^{2r} \setminus \mathcal{B}} |\nu(\mathcal{Y})|^2 + \sum_{\mathcal{Y} \in \mathcal{B}} |\nu(\mathcal{Y})|^2 \right) \lessgtr |T|^{2r}|S|^{4r-2} \left( (4r-1)^2 q |T|^{2r} + q^2 |\mathcal{B}| \right). \] (5.8)
It remains now only to bound \(|B|\). We establish shortly that the \(2r\)-tuples \(y\) lying in \(B\) are generated via the relations (5.7) from at most \(r\) of the coordinates \(y_i\) of \(y\). With this in mind, we begin by bounding the number of \(2r\)-tuples \(y\) generated from one such \(r\)-tuple. Suppose that there are \(l\) distinct values amongst \(y_1, \ldots, y_{2r}\), say \(y_{i_1} = v_1, \ldots, y_{i_l} = v_l\), with respective multiplicities \(a_1, \ldots, a_l\). Considering a fixed choice of \(v_1, \ldots, v_l\), the number of ways in which \(y_1, \ldots, y_{2r}\) may be selected to satisfy the multiplicity condition is \((2r)!/(a_1!a_2!\ldots a_l!)\).

Next we identify a directed graph with vertices labelled by the distinct elements \(v_1, \ldots, v_l\) of \(\mathbb{F}\) as follows. We consider the vertices \(v_1, v_2, \ldots, v_l\) in turn. At stage \(i\) we consider all elements \(v_j\) with \(A(v_j) = v_i\). If no such element \(v_j\) exists, then we add no edge. If one or more exist, on the other hand, then we select one such element \(v_j\) at random, and add a directed vertex from \(v_j\) to \(v_i\). Notice that, since \(v_1, \ldots, v_l\) are distinct, it follows that there can be at most one directed edge leaving any given vertex. Also, by construction, there is at most one directed edge arriving at any given vertex. Furthermore, in view of the criterion (5.7), any vertex \(v_k\) which possesses no edges must necessarily have multiplicity \(a_k \geq 2\). In this way, we see that the graph constructed in this manner consists of at most a union of isolated vertices, non-branching paths of the shape

\[v_{i_1} \rightarrow v_{i_2} \rightarrow \cdots \rightarrow v_{i_k},\]

and cycles of the shape

\[v_{i_1} \rightarrow v_{i_2} \rightarrow \cdots \rightarrow v_{i_k} \rightarrow v_{i_1},\]

(5.9) (5.10)

In the latter two cases, of course, one has \(k \geq 2\). For each non-branching path of type (5.9), we call the element \(v_{i_1}\) the root of the path. For each cycle of type (5.10), we call the element \(v_{i_1}\) a root of the cycle, though of course which element we label as \(v_{i_1}\) is unimportant. Notice that since \(v_{i_{m+1}} = A(v_{i_m})\) for each \(m < k\), roots uniquely determine all elements in the respective paths and cycles by repeated application of \(A\). Consequently, all of the elements \(v_1, \ldots, v_l\) are uniquely determined by the identities of the roots, and the indices defining the paths, cycles and isolated vertices of the graph.

Denote by \(z\) the number of paths and cycles in the graph, and by \(w\) the number of isolated vertices in the graph. Then on considering the multiplicities associated with the elements \(v_1, \ldots, v_l\), one finds that

\[2z + 2w \leq a_1 + \ldots + a_l = 2r.\]

The number of elements from \(T\) that can occur as roots and isolated vertices is consequently at most \(|T|^{z+w} \leq |T|^r\). We estimate the number of possible arrangements of indices defining the paths, cycles and isolated vertices as follows. Given each element \(v_i\), we can attach to it a directed path going to another element \(v_j\) in at most \(l - 1\) ways, or choose not to attach a directed path from it. Thus there are in total at most \((l - 1 + 1)^l\) possible arrangements of indices defining the paths, cycles and isolated vertices amongst the \(l\) elements \(v_1, \ldots, v_l\). Combining the estimates that we have assembled thus far, we conclude that

\[|B| \leq \sum_{1 \leq a_1, \ldots, a_l \leq 2r} \frac{(2r)!}{a_1!a_2!\ldots a_l!} l^l |T|^r \leq (1 + \ldots + 1)^{2r} l^l |T|^r \leq (2r)^{4r} |T|^r.\]
Finally, on substituting this estimate into (5.8), we obtain
\[ \Theta^{4r} \leq |T|^{4r}|S|^{4r-2}q\left((4r-1)^2 + q(2r)^4|T|^{-r}\right), \]
and the conclusion of the theorem follows on extracting \(4r\)-th roots.

A direct computation yields the following corollary.

**Corollary 5.6.** Let \(\eta\) be a positive number with \(\eta \leq 1\). Then under the hypotheses of the statement of Theorem 5.5, one has
\[
\sum_{y \in T} \sum_{s \in S} \chi_a(s + y)\chi_b(s + A(y)) < \eta|T||S|
\]
whenever \(|T| \geq (2r)^4q^{1/r}\) and \(|S| \geq 4rq^{1/2}/\eta^2r\).

**Proof.** Recall the notation of the statement of Theorem 5.5. When \(|T| > (2r)^4q^{1/r}\), we find that
\[ \chi_r^{4r} = (4r-1)^2 + (2r)^4q|T|^{-r} \leq (4r-1)^2 + 1 < 16r^2. \]
But then the upper bound (5.11) follows from Theorem 5.4 provided only that
\[ (16r^2)^{(4r)} q^{1/(4r)} |S|^{1-1/(2r)} |T| \leq \eta|T||S|, \]
as is the case whenever \(|S| \geq 4rq^{1/2}/\eta^2r\). 

**Proof.** [Proof of Theorem 5.4] We verify that the hypotheses of Theorem 5.4 imply the conditions on \(|S|\) and \(|T|\) in Corollary 5.6, for \(r = 1 + \lfloor (2\log q)/\log |T| \rfloor \geq 3\). We then have \(q^{2/r} \leq |T| \leq q^{2/(r-1)}\), and for large enough \(q\) we get \(2r \leq \log q\). Therefore \(|T| \geq (\log q)^4|T|^{1/2} \geq (2r)^4q^{1/r}\).

Moreover, \(1/\eta^2r \leq |T|^{2ar/r} \leq q^{(4a/7)(r/(r-1))} \leq q^{(4a/7)(3/2)}\), and hence for large enough \(q\) we have
\[ |S| \geq 4(\log q)q^{1/2+6a/7} \geq 4rq^{1/2}/\eta^2r, \]
as required. 

**6. Application to Privacy Amplification.** Following [19], we define a privacy amplification protocol \((P_A, P_B)\), executed by two parties Alice and Bob sharing a secret \(X \in \{0,1\}^n\), in the presence of an active, computationally unbounded adversary Eve, who might have some partial information \(E\) about \(X\) satisfying \(H_\infty(X|E) \geq k\). Informally, this means that whenever a party (Alice or Bob) does not reject, the key \(R\) output by this party is random and statistically independent of Eve’s view. Moreover, if both parties do not reject, they must output the same keys \(R_A = R_B\) with overwhelming probability.

More formally, we assume that Eve is in full control of the communication channel between Alice and Bob, and can arbitrarily insert, delete, reorder or modify messages sent by Alice and Bob to each other. In particular, Eve’s strategy \(P_E\) actually defines two correlated executions \((P_A, P_E)\) and \((P_E, P_B)\) between Alice and Eve, and Eve and Bob, called “left execution” and “right execution”, respectively. We stress that the message scheduling for both of these executions is completely under Eve’s control, and Eve might attempt to execute a run with one party for several rounds before resuming the execution with another party. However, Alice and Bob are assumed to have fresh, private and independent random tapes \(Y\) and \(W\), respectively, which are not known to Eve (who, by virtue of being unbounded, can be assumed deterministic). At the
end of the left execution \((P_A(X,Y),P_E(E))\), Alice outputs a key \(R_A \in \{0,1\}^m \cup \{\perp\}\), where \(\perp\) is a special symbol indicating rejection. Similarly, Bob outputs a key \(R_B \in \{0,1\}^m \cup \{\perp\}\) at the end of the right execution \((P_E(E),P_B(X,W))\). We let \(E'\) denote the final view of Eve, which includes \(E\) and the communication transcripts of both executions \((P_A(X,Y),P_E(E))\) and \((P_E(E),P_B(X,W))\). We can now define the security of \((P_A,P_B)\). Our definition is based on [19].

**Definition 6.1.** 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 \((k,m,\varepsilon)\)-privacy amplification protocol if it satisfies the following properties whenever \(H_\infty(X|E) \geq k\):

1. **Correctness.** If Eve is passive, then \(\Pr[R_A = R_B \land R_A \neq \perp \land R_B \neq \perp] = 1\).
2. **Robustness.** We start by defining the notion of pre-application robustness, which states that even if Eve is active, \(\Pr[R_A \neq R_B \land R_A \neq \perp \land R_B \neq \perp] \leq \varepsilon\). 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 \(\text{purify}(r)\) be \(\perp\) if \(r = \perp\), and otherwise replace \(r \neq \perp\) by a fresh \(m\)-bit random string \(U_m\); \(\text{purify}(r) \leftarrow U_m\). Letting \(E'\) denote Eve’s view of the protocol, we require that

\[\Delta((R_A,E'),(\text{purify}(R_A),E')) \leq \varepsilon \quad \text{and} \quad \Delta((R_B,E'),(\text{purify}(R_B),E')) \leq \varepsilon\]

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 \(\log(1/\varepsilon)\) is called the security parameter of the protocol.

**6.1. Case of \(k > n/2\).** Given a security parameter \(s\), Dodis and Wichs showed that a non-malleable extractor, which extracts at least \(2\log n + 2s + 4\) number of bits with error \(\varepsilon = 2^{-s-2}\), yields a two-round protocol for privacy amplification with optimal entropy loss, which also uses any (regular) extractor \(\text{Ext}\) with optimal entropy loss and any asymptotically good one-time message-authentication code \(\text{MAC}\) (see Definition 6.8), is depicted in Figure 6.1.

Using the bound from Theorem 4.3 and setting \(\varepsilon = 2^{-s}\) and \(m = s\), we get the following theorem.

**Theorem 6.2.** Under Conjecture 3.7 with constant \(c\), for any \(s > 0\) there is a polynomial time computable \((k,\varepsilon)\)-non-malleable extractor with \(m = s\) and \(\varepsilon = 2^{-s}\), as long as \(k \geq n/2 + (c/2)\log n + 4s + O(1)\).

Using this theorem, we obtain the following.

**Theorem 6.3.** Under Conjecture 3.7 with constant \(c\), there is a polynomial-time, two-round protocol for privacy amplification with pre-application robustness, with security parameter \(s\) and entropy loss \(O(\log n + s)\), when the min-entropy \(k\) of the \(n\)-bit secret satisfies \(k \geq n/2 + (c/2 + 8)\log n + 8s + O(1)\).

**Post-application Robustness.** Unfortunately, the 2-round protocol above does not achieve very good parameters for the stronger notion of post-application robustness. Intuitively, this is because the attacker Eve will get the derived key \(R_B\) before she needs to modify \((W',T')\) into \((W,T)\). This will reduce the entropy of \(X\) by \(m = |R_B|\) bits, and will require that the remaining entropy \(k - m \geq n/2 + (c/2)\log n + 4s + O(1)\).
in order to use the results of Theorem 6.2. In other words, we can only extract $m \approx k - n/2$ bits with post-application robustness using this technique, which is not better than what can be achieved by a 1-round protocol of Dodis, Kanukurthi, Katz, Reyzin, and Smith [7]. Fortunately, as communicated to us by [1], we can easily fix the situation at the cost of adding a third round (so that Eve no longer gets $R_B$ at the end of the second round), as follows. We use the 2-round protocol above to agree on a short (namely, $O(s + \log n)$-bit) key $K$ for a one-time MAC. Then, in the third round, Alice picks a fresh seed $S$ for a regular extractor $\text{Ext}$, and sends $S, \text{MAC}_K(S)$ to Bob. Both parties finally output a key $R_A = R_B = \text{Ext}(X; S)$ of length $k - O(s + \log n)$. The analysis of post-application analysis follows rather immediately, since the MAC key $K$ is uniform even conditioned on the derived key $R_A$ (which Alice would give to Eve in the post-application robustness game), and the non-malleable extractor was used before Eve learned $R_A$. We get:

**Theorem 6.4.** Under Conjecture 3.7 with constant $c$, there is a polynomial-time, three-round protocol for privacy amplification with post-application robustness, with security parameter $s$ and entropy loss $O(\log n + s)$, when the min-entropy $k$ of the $n$-bit secret satisfies $k \geq n/2 + (c/2 + 8) \log n + 8s + O(1)$.

*Using Weak Local Randomness.* We notice that we can use Theorem 4.4 to argue that Alice does not need perfect local randomness $Y$ to run the protocol in Figure 6.1. Indeed, since the output of the non-malleable extractor is only $O(s)$-bit long, we only need the min-entropy of $Y$ to be $O(s)$. Similarly, Bob could use a two-source extractor $\text{Ext}$ with a weak seed $W$ constructed by Raz [25]. Assuming the entropy rate of $X$ is above $1/2 + \alpha$ for some $\alpha > 0$, this extractor extracts $\Omega(n)$ bits from $X$, and only needed the min-entropy of $W$ to be $O(s)$ as well. To summarize, Alice and Bob can each use local sources of randomness of min-entropy only $O(s)$, and still extract $\Omega(n)$ secret bits from $X$.

**6.2. Case of $k = \delta n$.** Here we give our protocol for arbitrary positive entropy rate. We first give some preliminaries.

**6.2.1. Prerequisites from previous work.**

**Definition 6.5.** An elementary somewhere-$k$-source is a vector of sources $(X_1, \cdots, X_C)$, such that some $X_i$ is a $k$-source. A somewhere $k$-source is a convex combination of elementary somewhere-$k$-sources.
DEFINITION 6.6. A function $\text{Cond} : \{0,1\}^n \rightarrow (\{0,1\}^{n'})^C$ is a $(k \rightarrow k', \varepsilon)$-somewhere-condenser if for every $k$-source $X$, the vector $(X_1, \ldots, X_C) = \text{Cond}(X)$ is $\varepsilon$-close to a somewhere-$k'$-source. When convenient, we call $\text{Cond}$ a rate-$(k/n \rightarrow k'/n', \varepsilon)$-somewhere-condenser.

We are going to use condensers recently constructed based on the sum-product theorem. Specifically, we have the following theorem.

THEOREM 6.7 ([2, 25, 31]). For any $\delta > 0$ and constant $\beta > 0$, there is an efficient family of rate-$(\delta \rightarrow 1 - \beta, \varepsilon = 2^{-\Omega(\delta n)})$-somewhere condensers $\text{Cond} : \{0,1\}^n \rightarrow (\{0,1\}^{n'})^C$, where $C = \text{poly}(1/\delta)$ and $n' = \text{poly}(\delta)n$.

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

DEFINITION 6.8. A function family $\{\text{MAC}_R : \{0,1\}^d \rightarrow \{0,1\}^v\}$ is a $\varepsilon$-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) $A : \{0,1\}^n \rightarrow \{0,1\}^d \times \{0,1\}^v$,

\[
\Pr_{R}[\text{MAC}_R(W') = T' \land W' \neq w \mid (W', T') = A(\text{MAC}_R(w))] \leq \varepsilon,
\]

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

THEOREM 6.9 ([19]). For any message length $d$ and tag length $v$, there exists an efficient family of $(\frac{d}{2} \cdot 2^{-v})$-secure MACs with key length $\ell = 2v$. In particular, this MAC is $\varepsilon$-secure when $v = \log d + \log(1/\varepsilon)$.

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 $A_2$,

\[
\Pr_{(R, E)}[\text{MAC}_R(W') = T' \land W' \neq W \mid W = A_1(E), (W', T') = A_2(\text{MAC}_R(W), E)] \leq \frac{d}{v} \cdot 2^{v-H_\infty(R|E)}.
\]

(In the special case when $R = U_2^n$, and independent of $E$, we get the original bound.)

Finally, we will also need to use any strong seeded $(k, \varepsilon)$-extractor with optimal entropy loss $O(\log(1/\varepsilon))$. A simple extractor that achieves this is the one from the leftover hash lemma, which uses a linear-length seed. We can also use more sophisticated constructions such as those in [15, 14], and the non-malleable extractor with short seed length [6] to reduce the communication complexity of the protocol.

6.2.2. The protocol. Now we give our privacy amplification protocol for the setting when $H_\infty(X|E) = k \geq \delta n$. We assume that the error $\varepsilon$ we seek satisfies $2^{-\Omega(\delta n)} < \varepsilon < 1/n$. In the description below, it will be convenient to introduce an “auxiliary” security parameter $s$. Eventually, we will set $s = \log(C/\varepsilon) + O(1) = \log(1/\varepsilon) + O(1)$, so that $O(C)/2^s < \varepsilon$, 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 $\text{Cond} : \{0,1\}^n \rightarrow (\{0,1\}^{n'})^C$ be a rate-$(\delta \rightarrow 0.9, 2^{-s})$-somewhere-condenser. Specifically, we will use the one from Theorem 6.7, where $C = \text{poly}(1/\delta)$ and $n' = \text{poly}(\delta)n = O(n)$ and $2^{-s} > 2^{-\Omega(\delta n)}$.

- Let $\text{nmExt} : \{0,1\}^n \times \{0,1\}^d \rightarrow \{0,1\}^{m'}$ be a $(0.9n', 2^{-s})$-non-malleable extractor. Specifically, we will use the one from Theorem 6.2 (which is legal since $0.9n' \gg n'/2 + O(\log n') + 8s + O(1)$) and set the output length $m' = 4s$ (see the description of $\text{MAC}$ below for more on $m'$.)
• Let \( \text{Ext} : \{0, 1\}^n \times \{0, 1\}^d \rightarrow \{0, 1\}^m \) be a \((k', 2^{-s})\)-extractor with optimal entropy loss \( k' - m = O(s) \). Specifically, we will set \( k' = k - (7C + 11)s = k - O(s) \), which means that \( m = k - O(s) \) as well. We will use the notation \( \text{Ext}_{a \ldots b}(X; W) \), where \( 1 \leq a \leq b \leq m \), to denote the sub-string of extracted bits from bit position \( a \) to bit position \( b \). We assume the seed length \( d \leq n \) (e.g., by using a universal hash function, but more seed-efficient extractors will work too, reducing the communication complexity).

• Let \( \text{MAC} \) be the one-time, \( 2^{-s}\)-secure MAC for \( d \)-bit messages, whose key length \( \ell' = m' \) (the output length of \( \text{nmExt} \)). Using the construction from Theorem 6.9, we set the tag length \( \nu' = s + \log d \leq 2s \) (since \( d \leq n \leq 1/\varepsilon \leq 2^s \)), which means that the key length \( \ell = m' = 2\nu' \leq 4s \).

• Let \( \text{lrMAC} \) be the another one-time ("leakage-resilient") MAC for \( d \)-bit messages, but with tag length \( \nu = 2\nu' \leq 4s \) and key length \( \ell = 2\nu \leq 8s \). We will later use the second part of Theorem 6.9 to argue good security of this MAC even when \( \nu' \) bits of partial information about its key is leaked to the attacker. To not confuse the two MACs, we will use \( Z \) (instead of \( R \)) to denote the key of \( \text{lrMAC} \) and \( L \) (instead of \( T \)) to denote the tag of \( \text{lrMAC} \).

Using the above building blocks, the protocol is given in Figure 6.2. 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 \( W'_1 \), but Alice sees potentially different \( W_1 \), etc. Also, for any random variable \( G \) used in describing our protocol, we use the notation \( G = \perp \) to indicate that \( G \) was never assigned any value, because the party who was supposed to assign \( G \) rejected earlier. The case of final keys \( R_A \) and \( R_B \) becomes a special case of this convention.

Our protocol proceeds in \( C + 1 \) Phases. During the first \( C \) Phases, we run \( C \) sequential copies of the two-round protocol for the entropy-rate greater than \( 1/2 \) case (see Figure 6.1), but use the derived secret \( X_i \) (output by the somewhere-condenser) instead of \( X \) during the \( i \)-th run. Intuitively, since one of the values \( X_i \) is expected to have entropy rate above \( 1/2 \), we hope that the key \( Z_i \) extracted in this Phase is secret and uniform. However, there are several complications we must resolve to complete this template into a secure protocol.

The first complication is that Eve might not choose to execute its run with Alice in a “synchronous” manner with its execution with Bob. We prevent such behavior of Eve by introducing “liveness tests”, where after each Phase Alice has to prove that she participated during that Phase. Such tests were implicit in the original paper of Renner and Wolf [26], and made explicit by Khanakurthi and Reyzin [19]. Each liveness test (except for the last one in Phase \( C + 1 \), to be discussed) consists of Bob sending Alice a seed \( W'_i \) for the extractor \( \text{Ext} \) (which is anyway sent during the \( i \)-th Phase), and Alice responding with the first \( s \) bits of the extracted output. Intuitively, although Eve may choose to maul the extracted seed (which might be possible for all Phases, where the entropy rate of \( X_i \) is below \( 1/2 \)), Eve cannot predict the correct output without asking Alice something. And since Bob does uses a new liveness test between every two Phases, this effectively forces Eve to follow a natural “synchronous” interleaving between the left and the right executions.

The second complication comes from the fact that after a “good” (rate above \( 1/2 \)) Phase \( i \) is completed, the remaining phases might use low-rate sources \( X_{i+1}, \ldots, X_C \). Hence, one needs a mechanism to make sure that once a good key is extracted in some \textit{a-priori unknown} phase, good keys will be extracted in future phases as well, even if the remaining derived sources \( X_i \) have low entropy-rate. This is done by using a
second message authentication code lrMAC, keyed by a value $Z'_{i-1}$ extracted by Bob in the previous Phase $(i-1)$, to authenticated the seed $W'_i$ sent in Phase $i$. The only subtlety is that Bob still sends the original MAC of $W'_i$, and this MAC might be correlated with the previous extracted key $Z_{i-1}$ (especially if the Phase $i$ uses “bad-rate” $X_i$). Luckily, by using the “leakage-resilient” property of our second MAC (stated in Theorem 6.9), and setting the parameters accordingly, we can ensure that $Z'_{i-1}$ has enough entropy to withstand the “leakage” of the original MAC of $W'_i$.

The template above already ensures the robustness of the protocol, if we were to extract the key $Z_C$ (or $Z'_C$ for Bob) derived at the end of Phase $C$. Unfortunately, it does not necessarily ensure that Alice outputs a random key (i.e., it does not guarantee the extraction property for Alice). Specifically, by making Alice’s execution run faster
than Bob’s execution, it might be possible for Eve to make Alice successfully accept a non-random seed $W_C$, resulting in non-random key $Z_C$. Intuitively, since all the $X_i$’s except for one might have low entropy rate, our only hope to argue security should come from the non-malleability on $\text{nmmExt}$ in the “good” Phase $i$. However, since Bob is behind (say, at Phase $j < i$) Alice during the good Phase $i$, Bob will use a wrong source $X_j$ for the non-malleable extractor, and we cannot use the non-malleability of $\text{nmmExt}$ to argue that Eve cannot fool Alice into accepting a wrong seed $W_i$ (and, then, wrong $W_{i+1}, \ldots, W_C$). Of course, in this case we know Bob will eventually reject, since Eve won’t be able to answer the remaining liveness tests. However, Alice’s key $Z_C$ is still non-random, violating extraction.

This is the reason for introducing the last Phase $C + 1$. During this phase Alice (rather than Bob) picks the last seed $W_{C+1}$ and uses it to extract her the final key $R_A$. Therefore, $R_A$ is now guaranteed to be random. However, now we need to show how to preserve robustness and Bob’s extraction. This is done by Alice sending the MAC of $W_{C+1}$ using the key $Z_C$ she extracted during the previous round. (We call this MAC $S_C$ rather than $T_{C+1}$, since it also serves as a liveness test for Alice during Phase $(C+1)$.) From the previous discussion, we know that, with high probability, (a) either $Z_C$ is non-random from Eve’s perspective, but then Bob will almost certainly reject (ensuring robustness and preserving Bob’s extraction); or (b) $Z_C = Z_C'$ is random and secret from Eve, in which case the standard MAC security suffices to ensure both robustness and Bob’s extraction.

We detail the formal proof following the above intuition in the next section, which also establishes the desired parameters (including post-application robustness) promised by Theorem 1.5.

### 6.2.3. Security Proof of Our Protocol (Proof of Theorem 1.5).

We start by noticing that our protocol takes $2C + 1 = \text{poly}(1/\delta) = O(1)$ rounds and achieves entropy loss $k - m = O(Cs) = O(\text{poly}(1/\delta) \log(1/\varepsilon))$, as claimed. Also, the protocol obviously satisfies the correctness requirement.

We will also assume that the side information $E$ is empty (or fixed to a constant), since by Lemma 3.3, with probability $1 - 2^{-s}$, $H_\infty(X|E = e) \geq \delta n - s$, which will not affect any of our bounds. Before proving robustness and extraction properties of our protocol, we start with the following simple observation.

**Lemma 6.10.** Let $E'$ be Eve’s view at the end of her attack (without the keys $R_A$ and $R_B$ used in the post-application robustness experiment). Then, for any deterministic functions $f$ and $g$, we have

$$H_\infty(f(X) \mid g(E')) \geq H_\infty(f(X)) - (7C - 3)s$$

In particular, recalling that $k' = H_\infty(X) - (7C + 11)s$, we have $H_\infty(X|g(E')) \geq k' + 14s$.

**Proof.** Clearly, if it sufficient to prove the claim for $g$ being identity, as it gives the predictor the most information to guess $f(X)$. Also notice that, at best, if neither party rejects, Eve’s view $E' = (\tilde{Y}, \tilde{S}, \tilde{W}, \tilde{T}, \tilde{E}, W_{C+1})$, where $\tilde{Y} = \{Y_1, \ldots, Y_C\}$, $\tilde{S} = \{S_1, \ldots, S_C\}$, $\tilde{W} = \{W_1, \ldots, W_C\}$, $\tilde{T} = \{T_1, \ldots, T_C\}$ and $\tilde{E} = \{E_2, \ldots, E_C\}$. Since $\tilde{Y}$, $\tilde{W}$ and $W_{C+1}$ are independent of $X$ (and, thus, $f(X)$), using Lemma 3.4...
and recalling $|S_i| = s$ for $i < C$, $|S_C| = |T'_i| = v' \leq 2s$, $|L'_i| = v \leq 4s$, we have

$$
\bar{H}_\infty(f(X)|E') \geq \bar{H}_\infty(f(X)|Y, \bar{W}', W_{C+1})) - |\bar{S}| - |\bar{T}'| - |\bar{L}'|
= H_\infty(f(X)) - (C - 1)s - v' - C'v - (C - 1)v
\geq H_\infty(f(X)) - (C - 1)s - 2(C + 1)s - (C - 1)4s
= H_\infty(f(X)) - (7C - 3)s
$$

Lemma 6.11.

Next, we will argue the extraction property for Alice.

**Lemma 6.11.**

$$
\Delta((R_A, E'), (\text{purify}(R_A), E')) \leq 2^{-s+1}
$$

**Proof.** Since purify($R_A$) = $R_A$ when Alice rejects (i.e., $R_A = \perp$), it is sufficient to show that Alice’s key is close to uniform conditioned on Alice not rejecting, i.e.

$$
\Delta((\text{Ext}(X; W_{C+1}), E'), (U_m, E')) \leq 2^{-s+1}
$$

(6.1)

By Lemma 6.10, $\bar{H}_\infty(X|E') \geq k' + 14s$. Using Lemma 3.3, we get that $\Pr_{e' \leftarrow E'}[H_\infty(X|E' = e') \geq k'] \leq 1 - 2^{-s}$. Since Ext is $(k', 2^{-s})$-extractor, Equation (6.1) immediately follows the triangle inequality and the security of the extractor, by conditioning on whether or not $H_\infty(X|E' = e') \geq k'$. 

Next, we notice that in order to violate either robustness of Bob’s extraction, Eve must make Bob accept (i.e., $R_B = \perp$). Therefore, we start by examining how Eve might cause Bob to accept. Notice, since Alice sends $C + 1$ messages, including the first and the last message, Eve can make $C + 1$ calls to Alice, which we call Alice$_1, \ldots, Alice_{C+1}$, where, for each call Alice$_i$, $1 \leq i \leq C + 1$, Eve gets back the message sent by Alice during Phase $i$. Additionally, Alice also computes her key $R_A$ in response to Alice$_{C+1}$ (and gives $R_A$ to Eve, in addition to $S_C$ and $W_{C+1}$, for post-application robustness). Similarly, Eve can also make $C + 1$ calls to Bob, denoted Bob$_1, \ldots, Bob_{C+1}$, where each call Bob$_i$ expects as input the message that Alice supposedly sent to Bob in Phase $i$. When $i \leq C$, Bob responds to such a message with his own message in Phase $i$. When $i = C + 1$, Bob computes his key $R_B$ (and gives $R_B$ to Eve for post-application robustness). Clearly, the $(C + 1)$ calls to Alice must be made in order, and the same the $(C + 1)$ calls to Bob. However, a malicious Eve might attempt to interleave the calls in some adversarial manner to make Bob accept. We say that Eve is **synchronous** if he makes his oracle calls in the ("synchronous") order Alice$_1, Bob_1, Alice_2, Bob_2, \ldots, Alice_{C+1}, Bob_{C+1}$. We notice that, without loss of generality, Eve always starts by making the Alice$_1$ call, since this call has no inputs Eve needs to provide. Namely, Eve must as well find out the values $Y_1$ first, and, if she wants, delay using this value until later. With this convention in mind, we show that Eve must be synchronous in order to make Bob accept.

**Lemma 6.12.**

$$
\Pr[R_B \neq \perp \land \text{Eve is not synchronous}] \leq \frac{3C}{2^s}
$$

(6.2)

**Proof.** As we said, we assume Eve always makes the call Alice$_1$ first. After that, Eve makes $C + 1$ calls to Bob and $C$ calls to Alice in some order. We claim that for every $1 \leq i \leq C$, Eve must make at least one call to some Alice$_j$ in between two
successive calls $Bob_i$ and $Bob_{i+1}$. If we show this (with total failure probability from Equation (6.2)), Eve must be synchronous, since the synchronous scheduling is the only scheduling that starts with $Alice_1$ and has a fresh call to Alice between $Bob_1$ and $Bob_2$. $Bob_2$ and $Bob_3$, \ldots, $Bob_C$ and $Bob_{C+1}$.

Given $1 \leq i < C$, let $F_i$ denote the event that Eve’s scheduling of calls made two successive calls $Bob_i$ and $Bob_{i+1}$ without a fresh call to some $Alice_j$, and Bob does not reject after the call $Bob_{i+1}$. We claim that $\Pr[F_i] \leq 3/2^s$. The bound from Equation (6.2) then follows by simply taking the union bound over all $i$. We consider two cases:

**Case 1**: $1 \leq i < C$. In this case, after the call $Bob_i(\cdot, \cdot)$ is made, Bob picks a fresh seed $W'_i$, and returns it as part of the output. By assumption, Eve immediately makes a call $Bob_{i+1}(S'_i, \cdot)$, without any intermediate calls to Alice, and Bob rejects if $S'_i \neq \text{Ext}_1(\ldots, X; W'_i)$. Thus, to establish our claim it is sufficient to show that $\Pr[S'_i \neq \text{Ext}_1(\ldots, X; W'_i)] \leq 3/2^s$. Intuitively, the bound on $\Pr[F_i]$ now follows from the fact that $\text{Ext}$ is a good (strong) $(k', 2^{-s})$-extractor, since, conditioned on Eve’s information so far, the $s$-bit value $\text{Ext}_1(\ldots, X; W'_i)$ is $2^{-s}$-close to random, and, hence, cannot be predicted with probability better than $2^{-s} + 2^{-s}$ (the third $2^{-s}$ is due to Lemma 3.3, since our extractor is assumed to be worst case, and is not needed for universal hash function extractors [11]).

A bit more formally, let $E_i$ be Eve’s view before the call to $Bob_i$ is made, and $E'_i = (E_i, W'_i, T'_i, L'_i)$ be Eve’s view after the call to $Bob_i$ is made. We notice that $E'_i$ is a deterministic function of $E^*_i = (E_i, Z'_{i-1}, R'_i)$ and $W'_i$, since $L'_i = \text{lrMAC}_{Z'_{i-1}}(W'_i)$ and $T'_i = \text{MAC}_{R'_i}(W'_i)$. Moreover, $W'_i$ is freshly chosen even conditioned on $E^*_i$. Thus, $\Pr[F_i] \leq \Pr[Eve(E^*_i, W'_i) = \text{Ext}_1(\ldots, X; W'_i)]$, where $W'_i$ is independent of $(X, E^*_i)$. We also note that $\tilde{H}_\infty(X|E_i) > k' + 14s$, by Lemma 6.10, since $E_i$ is a function of $E'$. Thus, $H_\infty(X|E^*_i) \geq H_\infty(X|E_i) - |Z'_{i-1}| - |R'_i| \geq k' + 14s - 4s - 8s = k' + 2s$. Using Lemma 3.3, $\Pr[e_i | H_\infty(X|E^*_i = e^*_i) \geq k' + 2s$, and the rest follows from the fact that in this case $(W'_i, \text{Ext}_1(\ldots, X; W'_i))$ is $2^{-s}$-close to $(W'_i, U_s)$, as mentioned earlier.

**Case 2**: $i = C$. In this case, after the call $Bob_C(\cdot, \cdot)$ is made, Bob picks a fresh seed $W'_C$, and returns it as part of the output. By assumption, Eve immediately makes a call $Bob_{C+1}(S'_C, W'_{C+1})$, without any intermediate calls to Alice, and Bob rejects if $S'_C \neq \text{MAC}_{Z'_C}(W'_{C+1})$, where $Z'_C = \text{Ext}_{1 \ldots m'}(X; W'_C)$. Thus, to establish our claim it is sufficient to show that $\Pr[S'_C \neq \text{MAC}_{Z'_C}(W'_{C+1})] \leq 3/2^s$. Completely similar to the previous case, we can argue that the value $Z'_C$ used by Bob is $2^{-s}$-close to $U_{m'}$ conditioned on Eve’s view so far. Moreover, the $2^{-s}$-security of MAC ensures that, when the key $Z'_C$ is truly uniform, Eve cannot successfully forge a valid tag $\text{MAC}_{Z'_C}(W'_{C+1})$ of any (even adversarially chosen) message $W'_{C+1}$ with probability greater than $2^{-s}$, completing the proof of this case as well. ☐

Therefore, from now on we assume that Eve is indeed synchronous. Moreover, since Eve must make Bob accept, we assume Eve finishes the both left and right execution (with the last call to $Bob_{C+1}$, hoping that Bob will accept). Also, by Theorem 6.7, we have that $(X_1, \ldots, X_C)$ is $2^{-\Omega(2^m)}$-close to a somewhere rate-0.9 source. Thus, we will ignore the error and think of $(X_1, \ldots, X_C)$ as indeed being a somewhere rate-0.9 source, as it only adds $2^{-\Omega(2^m)} \ll 2^{-s}$ to the total probability of error. Also, it is sufficient to show robustness and extraction for Bob properties assuming that $(X_1, \ldots, X_C)$ is an elementary somewhere rate-0.9 source, since $(X_1, \ldots, X_C)$ is a convex combination of elementary somewhere rate-0.9 sources. Hence, from now one we assume that some “good” index $1 \leq g \leq C$ satisfies $H_\infty(X_g) \geq 0.9n'$. We stress that this index $g$ is not known to Alice and Bob, but could be known to Eve. We start
by showing that, with high probability, Eve must forward a correct seed \( W_g \) to Alice in the “good” Phase \( g \).

**Lemma 6.13.** Assuming Eve is synchronous,

\[
\Pr[R_B \neq \perp \land W_g \neq W'_g] \leq \frac{3}{2^w}
\]  

(6.3)

**Proof.** Let \( E'_{g-1} \) be Eve’s view before the call to Alice. Note that \( X_g \) is a deterministic function of \( X \), and \( \left( E'_{g-1}, S_{g-1}, L'_g \right) \) is a deterministic function of Eve’s transcript \( E' \). Thus, by Lemma 6.10,

\[
\tilde{H}_\infty(X_g | (E'_{g-1}, S_{g-1}, L'_g)) \geq H_\infty(X_g) - (7C - 3) s
\]

\[
\geq 0.9n' - (7C - 3) s
\]

\[
= (n'/2 + O(\log n') + 8s + O(1)) + s - (0.4n' - O(Cs + \log n))
\]

\[
\geq (n'/2 + O(\log n') + 8s + O(1)) + s
\]

where the last inequality follows since \( n' = \text{poly}(1/\delta) n \gg O(Cs + \log n) \). By Lemma 3.3, with probability \( 1 - 2^{-s} \) over the fixings of \( E'_{g-1}, S_{g-1}, L'_g \), the min-entropy of \( X_g \) conditioned on these fixings is at least \( n'/2 + O(\log n') + 8s + O(1) \).

Notice also that the seed \( Y_g \) is independent of \( E'_{g-1}, S_{g-1}, L'_g \). Moreover, for the argument in this lemma, we will “prematurely” give Eve the value \( L'_g \) already after the call to Alice (instead of waiting to get it from the call to Bob). Let us now summarize the resulting task of Eve in order to make \( W_g \neq W'_g \), and argue that Eve is unlikely to succeed.

After the call to Alice, with high probability the min-entropy of \( X_g \) conditioned on Eve’s view is greater than \( n'/2 + O(\log n') + 8s + O(1) \), so that we can apply the non-malleability guarantees of \( \text{nmExt} \) given by Theorem 6.2. Alice then picks a random seed \( Y_g \) for \( \text{nmExt} \) and gives it to Eve. (Synchronous) Eve then forwards some related seed \( Y'_g \) to Bob (and another value \( S'_{g-1} \) that we ignore here), and learns some message \( W'_g \) and the tag \( T'_g \) of \( W'_g \), under key \( R'_g = \text{nmExt}(X_g; Y'_g) \) (recall, we assume Eve already knows \( L'_g \) from before). To win the game, Eve must produce a value \( W_g \neq W'_g \) and a valid tag \( T'_g \) of \( W_g \) under the original key \( R_g = \text{nmExt}(X_g; Y_g) \).

We consider two cases. First, if Eve sets \( Y'_g = Y_g \), then \( R_g = R'_g \) is \( 2^{-s} \)-close to uniform by Theorem 6.2. Now, if \( R_g \) was truly uniform, by the one-time unforgeability of MAC, the probability that Eve can produce a valid tag \( T_g \) of a new message \( W_g \neq W'_g \) is at most \( 2^{-s} \). Hence, Eve cannot succeed with probability more than \( 2^{-s+1} \) even with \( R_g \) which is \( 2^{-s} \)-close to uniform, implying the bound stated in the lemma (since we also lost \( 2^{-s} \) by using Lemma 3.3 at the beginning).

On the other hand, if Eve makes \( Y'_g \neq Y_g \), Theorem 6.2 implies that \( \Delta((R_g, R'_g), (U_m, R'_g)) \leq 2^{-s} \). Thus, the tag \( T'_g \) under \( R'_g \) is almost completely useless in predicting the tag of \( W_g \) under (nearly random) \( R_g \). Therefore, by \( 2^{-s} \) security of MAC, once again the probability that Eve can successfully change \( W_g \) without being detected is at most \( 2^{-s+1} \) (giving again the final bound \( 3/2^w \)).

Now, we want to show that, once Eve forwards correct \( W_g = W'_g \) to Alice in Phase \( g \), Eve must forward correct seeds \( W_i = W'_i \) in all future phases \( i = g + 1, \ldots, C \). We start by the following observation, which states that the derived keys \( Z'_{i-1} \) used by Bob in InMAC look random to Eve **whenever Eve forwards a correct key** \( W_{i-1} = W'_{i-1} \) to Alice.

**Lemma 6.14.** Assume Eve is synchronous, \( 2 \leq i \leq C \), and Eve forwards a correct value \( W_{i-1} = W'_{i-1} \) to Alice during her call to Alice. Also, let \( E_i \) be Eve’s view after
the call to Alice(\(W_{i-1}, \cdots\)). Then

\[
\Delta((Z'_{i-1}, E_i), (U_\ell, E_i)) \leq \frac{3}{2^s} \tag{6.4}
\]

Proof. Notice that \(E_i = (E_{i-1}, W'_{i-1}, T'_{i-1}, L'_{i-1}, S_{i-1}, Y_i)\), where \(E_{i-1}\) is Eve's view after the call to Alice\(_{i-1}\). For convenience, we replace the two tags \(T'_{i-1}, L'_{i-1}\) of \(W'_{i-1}\) by the corresponding MAC keys \(R'_{i-1}, Z'_{i-2}\), respectively, since this gives Eve only more information. Also, since \(W_{i-1} = W'_{i-1}\), we know that the value \(S_{i-1} = \text{Ext}_{1..s}(X; W_{i-1}) = \text{Ext}_{1..s}(X; W'_{i-1})\). Recalling that \(Z'_{i-1} = \text{Ext}_{s+1..s+\ell}(X; W'_{i-1})\), and denoting “side information” by \(E^*_i = (E_{i-1}, R'_{i-1}, Z'_{i-2}, Y_i)\), it is enough to argue

\[
\Delta((E^*_i, W'_{i-1}, \text{Ext}_{1..s}(X; W'_{i-1})), \text{Ext}_{s+1..s+\ell}(X; W'_{i-1})), (E^*_i, W'_{i-1}, \text{Ext}_{1..s}(X; W'_{i-1}), U_\ell)) \leq \frac{3}{2^s} \tag{6.5}
\]

where we notice that \(E^*_i\) is independent of the choice of random \(W'_{i-1}\). In turn, Equation (6.5) follows from the fact that Ext is \((k', 2^{-s})\)-extractor provided we can show that \(\tilde{H}_\infty(X|E^*_i) \geq k + s\). Indeed, the first error term \(2^{-s}\) comes from Lemma 3.3 to argue that \(\Pr[\tilde{H}_\infty(X|E^*_i) = e^*_i] \geq k \geq 1 - 2^{-s}\), and the other two error terms follow from the triangle inequality and the security of the extractor (first time applies on the first \(s\) extracted bits, and then on all \(s + \ell\) extracted bits).

So we show that \(\tilde{H}_\infty(X|E^*_i) \geq k + s\).

\[
\tilde{H}_\infty(X|E^*_i) = \tilde{H}_\infty(X|E_{i-1}, R'_{i-1}, Z'_{i-2}, Y_i)
\geq \tilde{H}_\infty(X|E_{i-1}, Y_i) - |R'_{i-1}| - |Z'_{i-2}|
= \tilde{H}_\infty(X|E_{i-1}) - m' - \ell
\geq k' + 14s - 4s - 8s
= k' + 2s
\]

where the first inequality used Lemma 3.4, the second equality used the fact that \(Y_i\) is independent of \((X, E_{i-1})\), and the second inequality used Lemma 6.10, since \(E_{i-1}\) is deterministic function of \(E'\). \(\square\)

Next, we use Lemma 6.13 and Lemma 6.14 to show that, with high probability, Alice and Bob must agree on the same key \(Z_C = Z'_C\) when they reach the last Phase \((C + 1)\).

**Lemma 6.15.** Assuming Eve is synchronous,

\[
\Pr[R_B \neq \perp \land Z_C \neq Z'_C] \leq \frac{4C}{2^s} \tag{6.6}
\]

Proof. Since \(Z_C = \text{Ext}_{1..m'}(X; W_C)\) and \(Z'_C = \text{Ext}_{1..m'}(X; W'_C)\), we get

\[
\Pr[R_B \neq \perp \land Z_C \neq Z'_C] \leq \Pr[R_B \neq \perp \land W_C \neq W'_C]
\leq \Pr[R_B \neq \perp \land W_g \neq W'_g]
+ \sum_{i=g+1}^C \Pr[R_B \neq \perp \land W_{i-1} = W'_{i-1} \land W_i \neq W'_i]
\leq \frac{3}{2^s} + (C - 1) \cdot \max_{i > g} \Pr[R_B \neq \perp \land W_{i-1} = W'_{i-1} \land W_i \neq W'_i]
\]

where the second inequality states that in order for \(W_C \neq W'_C\), either we must already have \(W_g \neq W'_g\) (which, by Lemma 6.13, happens with probability at most \(3/2^s\)), or
there must be some initial Phase $i > g$ where $W_{i-1} = W'_{i-1}$ still, but $W_i \neq W'_i$. Thus, to establish Equation (6.6), it suffices to show that, for any Phase $g < i \leq C$,

$$Pr[R_B \neq \bot \land W_{i-1} = W'_{i-1} \land W_i \neq W'_i] \leq \frac{4}{2^s} \quad (6.7)$$

Intuitively, this property follows from the unforgeability of lrMAC, since Eve must be able to forge a valid tag $L_i$ of $W_i \neq W'_i$, given a valid tag of $W'_i$ (under the same $Z_{i-1} = Z'_{i-1}$ since $W_{i-1} = W'_{i-1}$). The subtlety comes from the fact that Eve also learns the $v'$-bit value $T_i' = \text{MAC}_{R'_i}(W'_i)$, which could conceivably be correlated with the key $Z'_{i-1}$ for lrMAC. Luckily, since the tag length $v$ of lrMAC is twice as large as $v'$, Theorem 6.9 states that lrMAC is still unforgeable despite this potential “key leakage”.

More formally, if Eve forwards a correct value $W_{i-1} = W'_{i-1}$, both Alice and Bob use the same key $Z'_{i-1} = Z_{i-1} = \text{Ext}_{s+1,s+v}(X;W'_{i-1})$ to lrMAC during Phase $i$. Moreover, by Lemma 6.14, we know that this key $Z_{i-1}$ looks random to Eve right before the call to $Bob_i$: $\Delta((Z'_{i-1}, E_i), (U_i, E_i)) \leq \frac{3}{2^s}$, where $E_i$ is Eve’s view after the call to $Alice_i(W_{i-1}, \ldots)$. After the call to $Bob_i$, Eve learns the tag $L_i'$ of $W'_i$, and also a $v'$-bit value $T_i'$, which, for all we know, might be correlated with the key $Z'_{i-1}$. Therefore, to argue the bound in Equation (6.7), it is sufficient to argue that Eve can succeed with probability at most $2^{-s}$ in the following “truncated” experiment. After the call to $Alice_i$, the actual key $Z'_{i-1}$ is replaced by uniform $Z'_{i-1} \leftarrow U_i$. Then a random message $W'_i$ is chosen, its tag $L'_i$ is given to Eve, and Eve is also obtained to allow arbitrary $v'$ bits of information about $Z'_{i-1}$. Eve succeeds if she can produce a valid tag $L_i$ (under $Z'_{i-1}$) of a different message $W_i \neq W'_i$. This is precisely the precondition of the second part of Theorem 6.9, where $H_{\infty}(Z_{i-1}|E) \geq \ell - v' = 2v - v'/2 = 3v/2$. Hence, Eve’s probability of success is at most $d2^{v-3v/2} = d2^{-v/2} = d2^{-v'} \leq 2^{-s}$. □

We need one more observation before we can finally argue Bob’s extraction and robustness. Namely, at the end of Phase $C$, (synchronous) Eve has almost no information about the authentication key $Z'_C$ used by the Bob (and Alice, by Lemma 6.15) in the final Phase $C+1$.

**Lemma 6.16.** Assume Eve is synchronous, and let $E'_C$ be Eve’s view after the call to $Bob_C$. Then

$$\Delta((Z'_C, E'_C \mid R_B \neq \bot), (U_m, E'_C \mid R_B \neq \bot)) \leq \frac{2}{2^s} \quad (6.8)$$

Additionally, $H_{\infty}(X|(E'_C, Z'_C)) \geq k' + 10s$.

**Proof.** The proof is similar to, but simpler than, the proof of Lemma 6.14. We notice that $E'_C = (E_C, W'_C, T'_C, L'_C)$, where $E_C$ is Eve’s view after the call to $Alice_C$. For convenience, we replace the two tags $T'_C, L'_C$, of $W'_C$ by the corresponding MAC keys $R'_C, Z'_{C-1}$, respectively, since this gives Eve only more information. Recalling that $Z'_C = \text{Ext}_{1,m'}(X;W'_C)$, and denoting “side information” by $E'_C = (E_C, R'_C, Z'_{C-1})$, it is enough to argue

$$\Delta((E'_C, W'_C, \text{Ext}_{1,m'}(X;W'_C)) ; (E'_C, W'_C, U_m')) \leq \frac{2}{2^s} \quad (6.9)$$

where we notice that $E'_C$ is independent of the choice of random $W'_C$. In turn, Equation (6.9) follows from the fact that $\text{Ext}$ is $(k', 2^{-s})$-extractor provided we can show
that $\tilde{H}_\infty(X|E'_C) \geq k + s$, where the extra error term $2^{-s}$ comes from Lemma 3.3 to argue that $\Pr_{e_C}[H_\infty(X|E'_C = e'_C) \geq k] \geq 1 - 2^{-s}$.

So we show that $\tilde{H}_\infty(X|E'_C) \geq k + s$.

$$\tilde{H}_\infty(X|E'_C) = \tilde{H}_\infty(X|E'_C, R'_C, Z'_{C-1})$$
$$\geq \tilde{H}_\infty(X|E'_C) - |R'_C| - |Z'_{C-2}|$$
$$= \tilde{H}_\infty(X|E'_C) - m' - \ell$$
$$\geq k' + 14s - 4s - 8s$$
$$= k' + 2s$$

where the first inequality used Lemma 3.4, and the second inequality used Lemma 6.10, since $E'_C$ is deterministic function of $E'$.

The final claim $\tilde{H}_\infty(X|(E'_C, Z'_{C})) \geq k' + 10s$ follows from Lemma 3.4 and fact that $\tilde{H}_\infty(X|E'_C) \geq k' + 14s$ (Lemma 6.10) and $|Z'_{C}| = m' \leq 4s$. □

Lemma 6.15 and Lemma 6.16 imply that, in order for the synchronous Eve to have a non-trivial chance to make Bob accept, at the end of Phase $C$ Alice and Bob must agree on a key $Z_C = Z'_C$ which looks random to Eve. Moreover, $X$ still has a lot of entropy given $Z'_C$ and Eve’s view so far. Thus, to show both (post-application) robustness and extraction for Bob, it is sufficient to show these properties for a very simply one-round key agreement protocol, which emulates the final Phase $(C+1)$ of our protocol with Alice and Bob sharing a key $Z_C = Z'_C$ which is assumed to be random and independent from Eve’s view so far. We start with post-application robustness.

**Post-Application Robustness:** To cause Bob output a different key than Alice in Phase $(C+1)$, Eve must modify Alice seed $W_{C+1}$ to $W'_{C+1} \neq W_{C+1}$, and then forge a valid tag $S'_C$ of $W'_{C+1}$ under the shared key $Z_C = Z'_C$. For pre-application robustness, the unforgeability of MAC immediately implies that Eve’s probability of success is at most $2^{-s}$. However, in the post-application robustness experiment, Eve is additionally given Alice’s final key $R_A = \text{Ext}(X; W_{C+1})$. Luckily, since $X$ has more than $k' + s$ bits of min-entropy even conditioned of the MAC key $Z_C$, security of the extractor implies that that the joint distribution of $Z_C$ and $R_A$ looks like a pair of independent random strings. In particular, Eve still cannot change the value of the seed $W_{C+1}$ in Phase $(C+1)$, despite being additionally given Alice’s key $R_A$, since that key looks random and independent of the MAC key $Z_C = Z'_C$.

**Extraction for Bob:** We just argued (pre-application) robustness of our protocol, which — for synchronous Eve — means that if Bob does not reject, then, with high probability, he outputs the same key $R_B = \text{Ext}(X; W'_{C+1})$ as Alice’s key $R_A = \text{Ext}(X; W_{C+1})$. Thus, Bob’s extraction is implied by Alice’s extraction, which was already argued in Lemma 6.11. Alternatively, Alice’s extraction can be seen directly, as she chooses a fresh seed $W_{C+1}$ and $\tilde{H}_\infty(X|E'_C, Z_C) \geq k' + 10s$.

This concludes the proof of Theorem 1.5.

**7. Future Directions.** There are several natural open questions. First, can we give a non-malleable extractor which outputs even one bit for arbitrarily linear entropy $\delta n$? As far as we know, it is possible that our extractor works for lower min-entropy (although the Cohen-Raz-Segev extractor [6] in full generality does not). Second, can we achieve optimal round complexity (2 rounds) and entropy loss ($O(s)$) for post-application robustness even when $k > n/2$ (recall, we had to spend a third round to achieve post-application robustness)? More ambitiously, can we achieve a
fixed-constant (e.g., 2 or 3) round protocol for weak secrets with arbitrarily linear entropy $\delta n$. In principle, this problem would be solved if an efficient non-malleable extractor is constructed for entropy $\delta n$. Finally, can we generalize our techniques to sublinear entropy?

**Appendix A. Generalizing the Non-Malleable Extractor.**

We now generalize our earlier results to show that we get a non-malleable extractor even if $M$ does not divide $q - 1$. We still use the same function $\text{nmExt}(x, y) = h(\log_2(x + y))$, with $h : \mathbb{Z}_{q-1} \to \mathbb{Z}_M$ given by $h(x) = x \mod M$.

**Theorem A.1.** There exists a constant $c > 0$ such that for any $n$, $k > n/2 + \log n + c$, and $m \leq k/2 - n/4 - c$, if we let $h$ be as above for $M = 2^n$, then the following holds. The function $\text{nmExt}(x, y) = h(\log_2(x + y))$ is a $(1, \varepsilon)$-non-malleable extractor for $\varepsilon = O(n2^{m+n/4-k/2})$.

The main ingredient in our proof is Rao’s generalization of Vazirani’s XOR lemma [24].

**A.1. A generalized XOR lemma.** We now extend Rao’s generalization of Vazirani’s XOR lemma. We need to modify his lemma because our output won’t necessarily be uniform.

**Lemma A.2.** For every positive integers $M \leq N$, the function $h : \mathbb{Z}_N \to H = \mathbb{Z}_M$ defined above satisfies the following property. Let $W, W'$ be any random variables on $\mathbb{Z}_N$ such that for all characters $\phi, \phi'$ on $\mathbb{Z}_N$ with $\phi$ nontrivial, we have $|E_{(W, W')}[\phi(W)\phi'(W')]| \leq \alpha$. Then $h(W), h(W')$ is $O(\alpha M \log N + M/N)$-close to the distribution $(U, W')$, where $U$ is the uniform distribution on $H$ which is independent of $W$.

To prove Theorem A.1 assuming Lemma A.2, we set $N = q - 1$, $(W, W') = (\log_2(X + Y), \log_2(X + A(Y)))$, and we condition on $Y = y$. Note that for $\phi$ an additive character, the function $\chi(x) = \phi(\log_2(x))$ is a multiplicative character. Therefore, Theorem 5.4 shows that $((W, W'))|Y = y)$ satisfies the hypotheses of Lemma A.2 for some $\alpha_y$, where $E_{y \in \mathbb{Z}_N}[\alpha_y] \leq \alpha$ for $\alpha < q^{1/4}2^{1-k/2} < 2^{n/4 + k/2}$. Thus, by Lemma A.2, one finds that $((h(W), h(W'))|Y = y)$ is $O(\alpha_M \log N + M/N)$-close to $((U, h(W')|Y = y)$ for every $y$. Since this expression is linear in $\alpha_y$, we conclude that $(h(W), h(W'), Y)$ is $O(\alpha M \log N + M/N)$-close to $(U, h(W'), Y)$, as required.

We now turn to the proof of Lemma A.2. First note that Lemma 3.8 is a special case. To handle $h$ in the case that $M \not| (q - 1)$, note that a character on a group $G$ has one Fourier coefficient $|G|$ and the rest 0. We show that if the $\ell_1$-norm of $\phi \circ h$ is not much bigger than this, then we get the desired conclusion.

**Lemma A.3.** Let $G$ and $H$ be finite abelian groups. Let $(W, W')$ be a distribution on $G \times G$ with $|E_{(W, W')}[\psi(W, W')]| \leq \alpha$ for all nontrivial characters $\psi$ and all characters $\psi'$. Let $h : G \to H$ be a function such that for every character $\phi$ of $H$, we have that

$$\|\widehat{\phi \circ h}\|_{\ell_1} \leq b|G|.$$  

Then $\|(h(W), h(W')) - (h(U), h(W'))\|_{\ell_1} \leq b\alpha|H|$.

**Proof.** Let $g : H \times H \to \mathbb{C}$ be the difference of distributions $(h(W), h(W')) - (h(U), h(W'))$, and let $f : G \times G \to \mathbb{C}$ be the difference of distributions $(W, W') - (U, W')$. By Lemma 3.8, we have $\|f\|_{\ell_\infty} \leq \alpha$. Let $\phi$ and $\phi'$ be any characters of $H$,
with $\phi$ nontrivial. Then

$$|\hat{g}(\phi, \phi')| = |\langle (\phi, \phi'), g = (h(W), h(W')) - (h(U), h(W')) \rangle|$$

$$= |\langle (\phi, \phi') \circ h, f = (W, W') - (U, W') \rangle|$$

$$= |\langle (\phi, \phi') \circ h, \tilde{f} \rangle|/|G|^2$$

$$\leq \|\phi \circ h\|_{\ell^1} \|\tilde{f}\|_{\ell^\infty}/|G|^2$$

$$\leq \|\phi \circ h, \phi' \circ h\|_{\ell^1} \cdot \alpha/|G|^2.$$

But now

$$\|\phi \circ h, \phi' \circ h\|_{\ell^1} = |\langle \phi \circ h, \phi' \circ h \rangle| \leq \|\phi \circ h\|_{\ell^1} \|\phi' \circ h\|_{\ell^\infty} \leq (b(|G|)|G|).$$

Putting these together yields $|\hat{g}(\phi, \phi')| \leq b|\alpha|$. When $\phi$ is trivial, as in Lemma 3.8, one has $\hat{g}(\phi, \phi') = 0$. By (3.1), this implies $\|g\|_{\ell^1} \leq |H|b|\alpha|$, as required. □

We bound $b$ using the following lemma by Rao, renormalized to our setting.

**Lemma A.4.** Let $M < N$ be integers, and let $h : \mathbb{Z}_N \to \mathbb{Z}_M$ be the function $h(x) = x \mod M$. Then for every character $\phi$ of $\mathbb{Z}_M$, we have $\|\phi \circ h\|_{\ell^1} = O(N \log N)$.

Thus, we may take $b = O(\log N)$ in Lemma A.3. Finally, we use the following simple lemma from Rao.

**Lemma A.5.** Let $M < N$ be integers, and let $h : \mathbb{Z}_N \to \mathbb{Z}_M$ be the function $h(x) = x \mod M$. Then for $U$ the uniform distribution on $\mathbb{Z}_N$, we have that $h(U)$ has $2M/N$-close to the uniform distribution on $\mathbb{Z}_M$.

**Appendix B. Acknowledgments.**

We are grateful to Gil Segev for finding an error in an earlier version of this paper, and to Salil Vadhan for a helpful discussion. We would also like to thank Divesh Aggarwal for pointing to us the problem with post-application robustness, and the anonymous referees at FOCS 2011 for useful comments.

**References**


Yevgeniy Dodis and Yu Yu, "Personal communication.


