Extractors and Secret Sharing Against Bounded Collusion Protocols

Eshan Chattopadhyay
Cornell University
eshanc@cornell.edu

Jesse Goodman
Cornell University
jpmgoodman@cs.cornell.edu

Vipul Goyal
CMU and NTT Research
vipul@cmu.edu

Ashutosh Kumar
University of California, Los Angeles
ashutoshk@ucla.edu

Xin Li
Johns Hopkins University
lixints@cs.jhu.edu

Raghu Meka
University of California, Los Angeles
raghum@cs.ucla.edu

David Zuckerman
University of Texas at Austin
diz@cs.utexas.edu

2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS) | 978-1-7281-9621-3/20/$31.00 ©2020 IEEE | DOI: 10.1109/FOCS46700.2020.00117

***Abstract*—In a recent work, Kumar, Meka, and Sahai (FOCS 2019) introduced the notion of *bounded collusion protocols* (BCPs). BCPs are multiparty communication protocols in which N parties, holding n bits each, attempt to compute some joint function of their inputs, $f : (\{0,1\}^n)^N \to \{0,1\}$. In each round, p parties (the *collusion bound*) work together to write a single bit on a public blackboard, and the protocol continues until every party knows the value of f.**

BCPs are a natural generalization of the well-studied *number-in-hand* (NIH) and *number-on-forehead* (NOF) models, which are just endpoints on this rich spectrum of protocols (corresponding to $p = 1$ and $p = N - 1$, respectively). In this work, we investigate BCPs more thoroughly, and answer questions about them in the context of communication complexity, randomness extractors, and secret sharing.

1. First, we provide explicit lower bounds against BCPs. Our lower bounds offer a tradeoff between collusion and complexity, and are of the form $n^{\Omega(1)}$ when $p = 0.99N$ parties collude. This bound is independent of the relationship between N, n, whereas all previous bounds became trivial when $N > 1.1 \log n$.

2. Second, we provide explicit leakage-resilient extractors against BCPs. Also known as cylinder-intersection extractors, these objects are multi-source extractors of the form $\mathsf{Ext} : (\{0,1\}^n)^N \to \{0,1\}$, whose output looks uniform even conditioned on the bits produced ("leaked") by a BCP executed over the inputs of the extractor. Our extractors work for sources with min-entropy $k \geq \mathrm{polylog}(n)$ against BCPs with collusion $p \leq N - 2$. Previously, all such extractors required min-entropy $k \geq 0.99n$ even when $p \leq O(1)$.

3. Third, we provide efficient leakage-resilient secret sharing schemes against BCPs. These cryptographic primitives are standard t-out-of-N secret sharing schemes, equipped with an additional guarantee that the secret remains hidden even if the individuals participate in a BCP using their shares. Our schemes can handle collusion up to $p \leq O(t/\log t)$, whereas the previous best scheme required $p \leq O(\log N)$.

Along the way, we also construct objects that are more general than those listed above (i.e., compilers), objects that are more specialized (and stronger) than those listed above, and resolve open questions posed by Goyal and Kumar (STOC 2018) and Kumar, Meka, and Sahai (FOCS 2019).

***Keywords*-bounded collusion protocols; multiparty communication complexity; randomness extractors; secret sharing; leakage-resilience; lower bounds**

I. Introduction

We begin by motivating our questions through the three main focus areas of this work: communication complexity, randomness extractors, and secret sharing.

Multiparty communication complexity: In a seminal work, Yao [1] initiated the study of communication complexity, where one seeks to understand how much communication is needed to compute a function f when its input is split between two parties. Since its introduction, communication complexity has blossomed into a central area of complexity theory, with connections to many other fields (see, e.g., the excellent book [2]).

Here, we focus on the multiparty setting - when the input is split amongst more than two parties. Perhaps the most natural way to define the *multiparty communication complexity* of a function f would be via *number-in-hand* (NIH) protocols ([3]–[5]). In this model, the input is split evenly across the participating parties, and each party can see only the input provided to them. The parties may communicate by writing on a public blackboard, and the protocol continues until every party knows the value of f. The NIH multiparty communication complexity of f, denoted $\mathsf{CC}^{\mathsf{NIH}}(f)$, is then defined as the number of bits that must be communicated by any such protocol (in the worst case over all inputs).

In 1983, Chandra, Furst, and Lipton [6] introduced an alternative way to define multiparty communication complexity via so-called *number-on-forehead* (NOF) protocols. Here, the input is once again split evenly across the participating parties, but each party is now able to see all inputs *except* their own (which, metaphorically, is written on their forehead). The parties still communicate via a public blackboard, and the NOF multiparty communication complexity of f, denoted $\mathsf{CC}^{\mathsf{NOF}}(f)$, is defined analogously to $\mathsf{CC}^{\mathsf{NIH}}(f)$. Because each party can see much more of the input in NOF protocols than in NIH protocols, NOF protocols are much more powerful and hence $\mathsf{CC}^{\mathsf{NOF}}(f) \leq \mathsf{CC}^{\mathsf{NIH}}(f)$.

It turns out that multiparty communication protocols offer an attractive model in which to pursue lower bounds, for two

2575-8454/20/$31.00 ©2020 IEEE
DOI 10.1109/FOCS46700.2020.00117

Authorized licensed use limited to: University of Texas at Austin. Downloaded on August 15,2021 at 16:31:35 UTC from IEEE Xplore. Restrictions apply.

reasons. First, these protocols appear to be simple enough to reason about combinatorially: if we write down a boolean function[1] $f : (\{0,1\}^n)^N \to \{-1,1\}$ in the cells of a multi-dimensional matrix M_f, one can get lower bounds on $\mathsf{CC}^{\mathsf{NIH}}(f)$ and $\mathsf{CC}^{\mathsf{NOF}}(f)$ by upper bounding the discrepancy of certain well-structured subsets of M_f. Second, these protocols appear to be rich enough to capture seemingly unrelated models of computation: NIH lower bounds imply (memory) lower bounds against streaming algorithms [7], while NOF lower bounds imply lower bounds in proof complexity [8] and circuit complexity [9]–[13] (for more connections, see the excellent survey [14]). Finally, as we will see, lower bounds against these protocols find great applicability in settings where hardness is considered "good" (like cryptography) [15].

Given these beautiful connections, it is natural to wonder whether NIH and NOF protocols can be placed into a more general framework of protocols, whose exploration could offer further insight into the above applications and lower bounds. In a recent work [15], Kumar, Meka, and Sahai introduced exactly such a family of communication protocols, called *bounded collusion protocols* (BCPs). Like the protocols we've seen before, BCPs are defined with respect to N parties, holding n bits of input each, who wish to compute a function $f : (\{0,1\}^n)^N \to \{0,1\}$ while communicating via a public blackboard. Unlike the protocols we've seen before, BCPs consider an additional parameter p (the *collusion bound*): in every round of the BCP, p parties may get together to write a bit on the blackboard, using all of the input in their possession.

BCPs define a natural spectrum of communication protocols, which is induced by the collusion bound p and gets more powerful as p increases. Furthermore, it is easy to see that NIH and NOF protocols are just endpoints on this spectrum (at $p = 1$ and $p = N - 1$, respectively). Thus, if we define $\mathsf{CC}^p(f)$ as the communication complexity of the function $f : (\{0,1\}^n)^N \to \{0,1\}$ with respect to BCPs with collusion bound p, we have:

$$\mathsf{CC}^{\mathsf{NOF}}(f) = \mathsf{CC}^{N-1}(f) \leq \mathsf{CC}^{N-2}(f) \leq \dots \leq \mathsf{CC}^2(f) \leq \mathsf{CC}^1(f) = \mathsf{CC}^{\mathsf{NIH}}(f).$$

It is relatively straightforward to come up with an explicit function $f : (\{0,1\}^n)^N \to \{0,1\}$ such that $\mathsf{CC}^{\mathsf{NIH}}(f) \geq n$ (see, for example, the books [14], [16]), whereas getting strong explicit lower bounds against NOF protocols is much more difficult: the best known results [17], [18] are of the form $\mathsf{CC}^{\mathsf{NOF}}(f) \geq \Omega(n/2^N)$. It is natural to wonder how explicit lower bounds against BCPs (and the difficulty of proving them) evolve from the NIH side of the spectrum to the NOF side of the spectrum. This leads us to our first question:

[1]Here and throughout, N and n are distinct parameters.

Question 1. *Can we provide explicit lower bounds that exhibit a collusion-complexity tradeoff against the spectrum of BCPs?*

In particular, it would be especially interesting to understand when we can obtain nontrivial explicit lower bounds in the setting $N \gg \log n$. No such bounds are known for NOF protocols, and in fact overcoming this "logarithmic barrier" is a longstanding challenge in complexity theory. Indeed, it has been shown that any significant improvements to the best NOF lower bounds would yield a breakthrough in circuit complexity, by providing new lower bounds against the circuit class ACC^0 [10], [12], [13]. Our first main result (Theorem 1) will answer the above questions, and in fact show that upon slightly reducing the collusion bound from $p = N - 1$ (NOF) to $p = 0.99N$, we can obtain explicit lower bounds of the form $\mathsf{CC}^p(f) \geq n^{\Omega(1)}$, regardless of the dependence between N, n.

Leakage-resilient extractors: The second main question that we consider will ask whether we can strengthen the above explicit worst-case lower bounds to something even stronger: *average-case* bounds (or, equivalently, *correlation bounds*). In particular, we would like to explore whether it is possible to explicitly construct a function f such that every BCP requires a large amount of communication to compute f even on just *slightly more than half* of all possible inputs.[2] We will in fact consider an even more challenging goal: obtaining such correlation bounds under non-uniform distributions on the input. In order to better understand and motivate our question, we will venture into the world of *randomness extractors*.

Extractors are fundamental objects in pseudorandomness, motivated by the observation that most applications of randomness in computer science require access to *uniform* random bits, yet the bits that we are able to harvest from nature are often biased and correlated. Extractors are algorithms that *purify* biased samples of randomness into samples that look uniform, thus offering a solution to the above problem. More formally, an extractor $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ is a single deterministic function, defined with respect to some error ϵ and family of distributions $\mathcal{X}$ over $\{0,1\}^n$. It offers the guarantee that for any source $\mathbf{X} \in \mathcal{X}$,

$$|\mathsf{Ext}(\mathbf{X}) - \mathbf{U}_m| \leq \epsilon,$$

where $\mathbf{U}_m$ is the uniform distribution over $\{0,1\}^m$, and $|\cdot|$ represents the standard statistical distance. Since their introduction, extractors have found beautiful applications in cryptography, coding theory, derandomization, and combinatorics. We refer the reader to [19], [20] for an excellent introduction to the area.

[2]Note that there is always a trivial BCP that computes f on at least half of all possible inputs: the BCP that always outputs 0, or the BCP that always outputs 1.

Authorized licensed use limited to: University of Texas at Austin. Downloaded on August 15,2021 at 16:31:35 UTC from IEEE Xplore. Restrictions apply.

One of the most well-studied models of randomness extraction is the setting where each $\mathbf{X} \in \mathcal{X}$ consists of N independent sources [21]–[25]. Here, the general extractor definition above specializes to give a so-called *extractor for independent sources* $\mathsf{Ext} : (\{0,1\}^n)^N \to \{0,1\}^m$ which guarantees that, given as input N independent sources over $\{0,1\}^n$, each with min-entropy at least k, it will output m bits that are close to uniform.[3] Recently, Kumar, Meka, and Sahai introduced a much stronger variant of such extractors, which they call *cylinder-intersection extractors* [15].

A cylinder-intersection extractor is a multi-source extractor that offers an additional guarantee that its output will look uniform *even conditioned on* the output of a BCP executed on the inputs to Ext. More formally, we define the class $\mathsf{BCP}(p,\mu)$ to consist of all functions $g : (\{0,1\}^n)^N \to \{0,1\}^\mu$ representing the bits written on the public blackboard by some BCP (its *transcript*), with collusion bound p, executed over N parties holding n bits each. Then, a cylinder-intersection extractor Ext against $\mathsf{BCP}(p,\mu)$ guarantees that, for any N independent sources $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_N)$, each with min-entropy at least k, and any BCP transcript $g \in \mathsf{BCP}(p,\mu)$,

$$|(\mathsf{Ext}(\mathbf{X}), g(\mathbf{X})) - (\mathbf{U}_m, g(\mathbf{X}))| \leq \epsilon,$$

where $\mathbf{U}_m$ is independent of $g(\mathbf{X})$. Thus, cylinder-intersection extractors are classical independent source extractors, equipped with an additional guarantee of being leakage-resilient against BCPs. As such, we will also refer to these objects as *leakage-resilient extractors (LREs) against BCPs*.

Beyond being interesting objects in their own right, there are several reasons to study such leakage-resilient extractors. First, one can observe that such extractors for min-entropy $k = n$, which output $m = 1$ bit, are actually functions with high *average-case* communication complexity (also known as *distributional communication complexity*) against BCPs. Thus, as noted in [15], asking if we can construct even stronger objects that work for $k \ll n$ is a natural way to translate our original question from communication complexity into the land of extractors. Furthermore, in very recent subsequent work, it was shown that LREs against BCPs can be used to create much improved extractors for more classical settings [26], and it is also not difficult to imagine settings in cryptography where such objects can come in handy (see the full version of this paper).

All of this motivates our second main question:

Question 2. *Can we construct leakage-resilient extractors against BCPs for min-entropy $k \ll n$?*

Our second main result(s) (Theorems 2 and 3) answer this question positively, and shows that we can construct such objects even when $k = \mathrm{polylog}(n)$. Furthermore, the tradeoff we obtain between allowed leakage μ and collusion bound p almost exactly matches the complexity-collusion tradeoff in our answer to Question 1.

Leakage-resilient secret sharing: The third main question that we consider revisits the original motivating application considered by Kumar, Meka, and Sahai [15] in their introduction of BCPs: *leakage-resilient secret sharing schemes*. Secret sharing schemes were introduced in the seminal works of Blakley [27] and Shamir [28] and have since become a classical cryptographic primitive. These schemes capture the natural setting of a central authority who wishes to share some secret (e.g., missile launch codes) among a group of N somewhat trusted individuals. Each individual is to receive a portion (or *share*) of the secret, so that any t of them may reconstruct the secret, but any fewer than t of them cannot recover any information. Formally, these schemes are known as *t-out-of-N schemes*.

Kumar, Meka, and Sahai study a much stronger variant of secret sharing, known as *leakage-resilient secret sharing* (LRSS). In addition to the above thresholding guarantees, an LRSS scheme guarantees that the secret will remain statistically hidden even against much stronger adversaries. The adversaries they consider are exactly the family of bounded collusion protocols $\mathsf{BCP}(p,\mu)$ (defined above) acting over the N individuals participating in the scheme.[4] It turns out that BCPs are an especially natural class against which one may want to ensure leakage-resilience of a secret sharing scheme: not only do BCPs generalize and strengthen several previous models of leakage (from non-adaptive, disjoint settings to an adaptive, joint setting), but their definition also allows one to leverage tools from communication complexity to construct such LRSS schemes.

Indeed, in [15] the authors show that by using a function with high (distributional) communication complexity against NOF protocols, it is possible to construct LRSS schemes against NOF protocols. However, there is a catch: because the best known NOF lower bounds are quite weak (recall from earlier that they are of the form $\Omega(n/2^N)$), it turns out that in order to make the schemes leakage-resilient, each individual must be provided with a secret share of size $\gg 2^N$. Because *efficiency* in secret sharing is classified by share size growing polynomially in the number of participants, N, a new idea is needed.

By applying the above NOF lower bounds to a smaller number of parties, and by reusing shares via perfect hash families, Kumar, Meka, and Sahai are able to circumvent this issue (at least for *BCPs*) and construct LRSS schemes. However, while they successfully remove the exponential dependence of share size on N, they incur an exponential dependence of share size on p. As such, they are only able

[3]The *min-entropy* of a source $\mathbf{X}$ over $\{0,1\}^n$ is at least k if and only if $\Pr[\mathbf{X} = x] \leq 2^{-k}$, for all $x \in \mathrm{support}(\mathbf{X})$.

[4]As a sanity check, it is useful to observe that a t-out-of-N secret-sharing scheme can only be equipped with leakage-resilience against $\mathsf{BCP}(p,\mu)$ when $p < t$.

Authorized licensed use limited to: University of Texas at Austin. Downloaded on August 15,2021 at 16:31:35 UTC from IEEE Xplore. Restrictions apply.

to ensure leakage-resilience against BCPs with collusion $p = O(\log N)$. This is again an artifact of the weak $\Omega(n/2^N)$ NOF lower bounds, which they use as a black box in their construction. Indeed, as observed in [15], efficient t-out-of-N LRSS schemes for $t = p + 1$ with $p = \omega(\log N)$ would resolve longstanding bottlenecks in complexity theory. However, it is not unreasonable to think that if we have stronger lower bounds against BCPs (which do not follow from applying NOF bounds in a black box manner), we may be able to avoid this barrier for $t \gg p$. Luckily, as described in the previous two sections, we are able to obtain explicit bounds of exactly this nature. This motivates our third and final question:

Question 3. *Can we use our new explicit lower bounds against BCPs to construct efficient LRSS schemes against BCPs with collusion $p = \omega(\log N)$?*

As we will see, our third main result (Theorem 4) answers this question in the affirmative, and shows that we can construct such leakage-resilient t-out-of-N schemes for collusion $p \leq O(t/\log t)$. For the special case of N-out-of-N schemes, we can handle p up to $0.99N$.

We now proceed to formally state our main theorems.

A. Summary of Our Results

In our first main theorem, we establish explicit lower bounds against the spectrum of BCPs. These bounds exhibit a collusion-complexity tradeoff, answering Question 1.

Theorem 1. *For all $N, n \in \mathbb{N}$ and $p \leq N - 1$, there exists an explicit function $f : (\{0,1\}^n)^N \to \{0,1\}$ with*

$$\mathsf{CC}^p(f) \geq \Omega\left(n^{\frac{\log(N/p)}{\log(N/p)+1}}\right).$$

In particular, Theorem 1 provides an explicit function f such that for BCPs with $p = 0.99N$ collusion, $\mathsf{CC}^p(f) \geq n^{\Omega(1)}$. Previously, the best known result [15] followed immediately from lower bounds against NOF protocols [17], and was of the form $\mathsf{CC}^p(f) \geq \Omega(n/2^p)$.[5] Thus, all previous bounds against $p = 0.99N$ collusion became trivial when $N > 1.1 \log n$.

Next, we show that we can significantly strengthen our explicit lower bounds to *average-case* lower bounds, and in fact strengthen these even further to produce leakage-resilient extractors against BCPs for polylogarithmic entropy (such extractors, even for full entropy, give explicit average-case lower bounds against BCPs). Furthermore, we achieve a collusion-leakage tradeoff that mirrors the collusion-complexity tradeoff from the previous result. We record our second main theorem, which answers Question 2 and a question of Kumar, Meka, and Sahai [15].

[5]We note that our function in Theorem 1 can also achieve this bound, which is slightly better when, e.g., $p = O(1)$.

Theorem 2. *For all $N, n, k \in \mathbb{N}$ satisfying $N \geq 3$ and $k \geq \operatorname{polylog} n$, and any $p \leq N - 2$, there exists an explicit leakage-resilient extractor $\mathsf{Ext} : (\{0,1\}^n)^N \to \{0,1\}^m$ against $\mathsf{BCP}(p, \mu)$ for min-entropy k, with output length $m = \mu$ and error $\epsilon = 2^{-\mu}$, where*

$$\mu = k^{\Omega\left(\frac{\log(N/p)}{\log(N/p)+1}\right)}.$$

In particular, Theorem 2 provides an explicit extractor for $k \geq \operatorname{polylog}(n)$ entropy that can handle $k^{\Omega(1)}$ leakage from BCPs with $p = 0.99N$ collusion. Previously, the best known result [15] followed immediately from lower bounds against NOF protocols [17], and required min-entropy $k \geq 0.99n$ even for $p \leq O(1)$ collusion. In addition to our explicit extractor, we also provide a much more general object: an explicit *compiler* that can transform *any* function with NOF lower bounds into a leakage-resilient extractor against BCPs. As NOF lower bounds are strengthened over time, our compiler is guaranteed to produce improved extractors.

Theorem 2 can handle up to $p = N - 2$ collusion. In the extreme setting of $p = N - 1$ collusion (i.e., NOF leakage), we construct the following extractor.

Theorem 3. *For all $N, n, k \in \mathbb{N}$ satisfying $N \geq 6$ and $k \geq 0.3n$, there exists an explicit leakage-resilient extractor $\mathsf{Ext} : (\{0,1\}^n)^N \to \{0,1\}$ against $\mathsf{BCP}(N-1, \mu)$ for min-entropy k with error $\epsilon = 2^{-\mu}$, where $\mu = \Omega(n/2^N)$.*

We remark that in the case $N = 3$ (resp., $4, 5$), Theorem 3 will work for min-entropy $k \geq 0.4n$ (resp., $k \geq 0.33n$).

In our final main theorem, we combine our explicit (average-case) lower bounds with new ideas to construct much improved t-out-of-N secret sharing schemes that are leakage-resilient against BCPs, answering Question 3.

Theorem 4. *For all $N \geq t \geq 2$, there exists an efficient[6] t-out-of-N secret sharing scheme* $(\mathsf{Share}, \mathsf{Rec})$ *that is leakage-resilient against* $\mathsf{BCP}(p, \mu)$ and *complete leakage of any $t-1$ shares, provided $p \leq O(t/\log t)$.*

As a crucial step towards Theorem 4, we design efficient N-out-of-N schemes that can handle collusion up to $p = 0.99N$. Previously, all efficient secret sharing schemes could only handle $p = O(\log N)$ collusion [15]. In addition to these schemes, we also provide a much more general object: an explicit *compiler* that can transform *any* given secret sharing scheme having authorized sets of size at least t into a scheme that is leakage-resilient against BCPs with collusion $p \leq O(t/\log t)$. Thus, we can also obtain leakage-resilient schemes for access structures like monotone span programs, monotoneP, and monotoneNP, by instantiating our compiler with previous schemes ([29]–[31]). Finally, in the special case of BCPs over *disjoint* subsets of individuals, our t-out-of-N schemes can handle optimal collusion $p = t - 1$, resolving a question of Goyal and Kumar [32].

[6]A scheme is *efficient* if its sharing and reconstruction functions run in time $\operatorname{poly}(N, m, \mu, \log(1/\epsilon))$ for m bit secrets.

Authorized licensed use limited to: University of Texas at Austin. Downloaded on August 15,2021 at 16:31:35 UTC from IEEE Xplore. Restrictions apply.

B. Relevant Prior Work

Bounded collusion protocols: BCPs were introduced by Kumar, Meka, and Sahai in [15]. There, the authors provided preliminary explicit lower bounds against BCPs (as discussed in the previous section), but primarily focused on the application of BCPs to secret sharing. Given their recent introduction, no other work has been done on BCPs - we hope, however, they will become a fruitful object of study.

Extractors: Generalizing the work of Santha and Vazirani [33] and Vazirani [34], Chor and Goldreich [21] initiated the study of extractors for independent sources in 1988. Since then, a beautiful line of work [22]–[25] has provided explicit constructions of these objects with near-optimal parameters. Recently, several works have emerged that study randomness extractors for "slightly dependent" sources [35]–[37]. In the current paper, we study cylinder-intersection extractors (which we also call leakage-resilient extractors against BCPs). These may indeed be viewed as extractors for "slightly dependent" sources, but the dependence model here is very different from the aforementioned works. Cylinder-intersection extractors were introduced by Kumar, Meka, and Sahai [15].

Secret Sharing: Secret sharing schemes were introduced in the seminal works of Blakley [27] and Shamir [28]. These schemes, while originally envisioned with only the goal of t-out-of-N secrecy, have since been strengthened in various ways (see the survey [30]). In the current paper, we focus on equipping secret sharing schemes with *leakage-resilience*, which has a long history in cryptography (see, e.g., the survey of Kalai and Reyzin [38]). In the context of secret sharing, leakage-resilience has recently garnered significant interest [15], [32], [39]–[46]. We refer the reader to [15] and [45] for a more detailed overview.

C. Open Problems

Improved cylinder-intersection extractors for NOF leakage: In the current paper, we construct leakage-resilient extractors against BCPs with collusion $p \leq N-2$, which work for $k \geq \mathrm{polylog}(n)$ entropy. In the NOF case ($p = N-1$), however, our extractors require $k \geq 0.3n$ entropy. It would be very nice to reduce the entropy requirement in this more difficult case, even just to $k \geq o(n)$, as we imagine this will require significantly new techniques.

Lifting theorems for the NOF model: Two-source extractors have been useful for obtaining query-to-communication lifting theorems for the case of two parties [47]. It is an interesting research direction to use our new leakage-resilient extractors against BCPs to obtain lifting theorems for the multiparty case.

Reduce the gap between p and t for LRSS: In the current paper, we design efficient t-out-of-N secret sharing schemes for collusion up to $p \leq O(t/\log t)$. However, it may still be possible to reach $p = 0.99t$ for all $t < N$, which would match our results for the special case of $t = N$. It would be interesting to reduce this gap, perhaps by designing schemes that do not rely on ramp hash families (since this gap originates from natural barriers that come from such families).

Leakage-resilient multiparty computation (MPC): Finally, it would be interesting to try to design leakage-resilient MPC protocols as a possible application of our LRSS schemes. Prior work of Goyal et al. [48] designs such protocols only for the two party case, and leaves open the design of such protocols for a higher number of parties.

D. Organization

In Section II, we provide an overview of the techniques we use to prove our main results. In Section III, we discuss some basic preliminaries, and include formal definitions for several key notions that were discussed informally in Section I. Next, in Section IV, we construct explicit lower bounds and leakage-resilient extractors against BCPs, proving Theorems 1 and 2. Finally, in Section V, we explicitly construct a leakage-resilient extractor against NOF leakage, proving Theorem 3. The proof to Theorem 4, and the remarks sprinkled throughout Section I-A, can be found in the full version of the paper.

This paper is a merge of the works [49], [50]. Together, these should be referenced as the full version of the paper.

II. Overview of Techniques

In what follows, we provide a high level overview of the techniques we use to prove our main results (recorded in Section I-A).

A. Explicit Lower Bounds Against BCPs

We start by outlining the proof of our explicit lower bounds against BCPs (Theorem 1). We recall that a function $\mathsf{hard} : (\{0,1\}^n)^N \to \{0,1\}$ has complexity $\mathsf{CC}^p(\mathsf{hard}) > \mu$ if, for any BCP protocol $g \in \mathsf{BCP}(p, \mu)$ and uniform random variable $\mathbf{X} \sim (\{0,1\}^n)^N$, it holds that $\mathsf{hard}(\mathbf{X})$ becomes a constant with probability less than 1 upon fixing $g(\mathbf{X})$ (see Remark 1). Thus, we aim to explicitly construct such a function hard with μ matching the parameters in Theorem 1. We will show that, in fact, it suffices to simply take hard to be an explicit function that exhibits the best known lower bounds against NOF protocols.

We will take hard to be the *finite field multiplication* function $\mathsf{FFM}_N : (\{0,1\}^n)^N \to \{0,1\}$. On input $(x_1, \ldots, x_N)$, this function interprets these strings as elements of $\mathbb{F}_{2^n}$, takes their product over this field, interprets this result again as a boolean string, and outputs the first bit. Using discrepancy arguments over objects known as *Hadamard tensors*, Ford and Gál show [18] that this extremely simple function has NOF complexity $\mathsf{CC}^{\mathsf{NOF}}(\mathsf{FFM}_N) \geq \Omega(n/2^N)$.

Our first key observation is that this function enjoys the special property of *self-reducibility*: if we feed it N

Authorized licensed use limited to: University of Texas at Austin. Downloaded on August 15,2021 at 16:31:35 UTC from IEEE Xplore. Restrictions apply.

uniform independent random variables, and fix any $N-K$ of them to nonzero values, then we simply obtain an instance of $\mathsf{FFM}_K : (\{0,1\}^n)^K \to \{0,1\}$ called on independent uniform random variables. We will now describe how to use this property to lift the known NOF lower bounds against FFM_N to BCP lower bounds.[7] The overview below will actually obtain lower bounds against *non-adaptive* BCPs, as there is an easy way to transform these into lower bounds against adaptive BCPs (see full version).

The main idea is as follows. Let $g : (\{0,1\}^n)^N \to \{0,1\}^\mu$ be the transcript (bits written on the blackboard) of an arbitrary non-adaptive BCP; that is, $g \in \mathsf{nBCP}(p,\mu)$. In order to show that $\mathsf{CC}^p(\mathsf{FFM}_N) > \mu$, we just need to show that $\mathsf{FFM}_N(\mathbf{X})$ becomes constant with probability less than 1 upon fixing $g(\mathbf{X})$, where $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_N)$ is uniform over $(\{0,1\}^n)^N$. In order to leverage the self-reducibility of FFM_N, we would like to somehow find a subset $S \subseteq [N]$ of t random variables $\{\mathbf{X}_i\}_{i\in S}$ with the guarantee that no leaked bit in g depends on all of them (in other words, during no round were *all* of the parties in S involved in a joint collusion). If we can find such a set S, then we can fix all random variables outside of that set. The value $\mathsf{FFM}_N(\mathbf{X})$ then becomes $\mathsf{FFM}_t((\mathbf{X}_i)_{i\in S})$, while $g(\mathbf{X})$ becomes $g'((\mathbf{X}_i)_{i\in S})$ for some NOF protocol g'. Thus, by applying the NOF lower bounds from [18], we immediately get bounds of the form $\mathsf{CC}^p(\mathsf{FFM}_N) \geq \Omega(n/2^t)$.

Thus, in order to get the best possible lower bounds for $\mathsf{CC}^p(\mathsf{FFM}_N)$, we would like to find such a subset S with *as small size* t as possible.[8] The first idea in this direction is to attempt to find such a subset with size $t = p+1$: this is of course always possible, since *every* such subset of this size is guaranteed to never work together in a joint collusion, simply by the fact that the collusion bound in g is p. This immediately produces bounds of the form $\mathsf{CC}^p(\mathsf{FFM}_N) \geq \Omega(n/2^{p+1}) = \Omega(n/2^p)$. While these are better than the general NOF bound $\Omega(n/2^N)$ for small p, they become trivial whenever $p > \log n$. We would like to try to handle much larger collusion.

To go beyond the logarithmic barrier, the key idea is to consider the *round bound* μ, and notice that sometimes we can actually find such a subset of size $t \ll p+1$, *if there are not too many rounds of communication*. In particular, let us fix $t \in [1, N]$ to some value that will be chosen later, and observe that there are $\binom{N}{t}$ subsets with the desired "non-joint" property at the beginning of the protocol. Furthermore, each time a bit is leaked, it can depend on at most p parties, and thus at most $\binom{p}{t}$ subsets of size t will lose the "non-joint" property that we desire. Thus, as long as $\mu\binom{p}{t} < \binom{N}{t}$, we will always be left with *at least one* subset with the desired property. By reordering this inequality and applying Stirling's formula, the condition $\mu < \Omega((N/p)^t)$ also suffices. Thus, we immediately get that for any round bound μ that obeys both $\mu < \Omega((N/p)^t)$ (so that we can find a "non-joint" subset) *and* $\mu < \Omega(n/2^t)$ (to ensure that we can apply the NOF lower bounds of the self-reduced FFM_N function), it holds that $\mathsf{CC}^p(\mathsf{FFM}_N) \geq \mu$. Thus, setting $t = \frac{\log n}{\log(N/p)+1}$ immediately gives Theorem 1, which in particular provides us with polynomial lower bounds against $p = 0.99N$ collusion, for any N, n.

[7]The use of self-reductions to prove lower bounds in communication complexity also appeared in the work of Podolskii and Sherstov [51].

[8]If we are dealing with some specialized type of BCP that guarantees the existence of such a set by definition, then of course we can stop here and get very strong complexity bounds of the form $\mathsf{CC}^p(\mathsf{FFM}_N) \geq \Omega(n/2^t)$. But this will not be the case for general BCPs with collusion bound p.

B. LREs Against BCPs for Polylogarithmic Entropy

It turns out that without too much more work, we can actually equip the function FFM_N with a few simple add-ons so that it becomes an explicit function with much stronger correlation bounds than those advertised in Theorem 1 (and outlined in the previous section). Indeed, in this section we will outline the proof of Theorem 2, which shows that we can turn FFM_N into a leakage-resilient extractor against BCPs with excellent parameters.

The first step is to augment FFM_N so that it achieves high *average-case* (or *distributional*) communication complexity against NOF protocols. Luckily, the discrepancy-based techniques used by Ford and Gál for their worst-case bounds yield average-case bounds for free. In particular, in [18] the authors show that $\mathsf{CC}^{\mathsf{NOF}}_\epsilon(\mathsf{FFM}_N) \geq \Omega(n/2^N)$ for $\epsilon = 2^{-\Omega(n/2^N)}$. Thus, their bounds show that not only do NOF protocols with $\Omega(n/2^N)$ rounds fail to compute FFM_N everywhere, but every such protocol offers no more than an exponentially small advantage over a trivial (constant) protocol. Furthermore, by Remark 2, this immediately tells us that for any $g \in \mathsf{NOF}(\mu)$, it holds that $|\mathsf{FFM}_N(\mathbf{X}) \circ g(\mathbf{X}) - \mathbf{U}_1 \circ g(\mathbf{X})| \leq \epsilon$, where $\mathbf{X}$ is uniform over $(\{0,1\}^n)^N$, $\mathbf{U}_1$ is independent from $\mathbf{X}$, and $\epsilon = 2^{-\mu}$, where $\mu = \Omega(n/2^N)$. This tells us that $\mathsf{FFM}_N : (\{0,1\}^n)^N \to \{0,1\}$ is a leakage-resilient extractor against $\mathsf{NOF}(\mu)$ for entropy $k = n$ with error ϵ, where μ, ϵ are as above.

The next step is to augment this leakage-resilient extractor so that it can *output multiple bits*, and in particular we would like $m = \mu$ bits of output. While this extension is not provided in [18], we show that it is straightforward to apply standard XOR lemmas ([52]) to character sums that appear in their paper in order to prove this result. As a result, we obtain a function $\mathsf{prodExt}_N : (\{0,1\}^n)^N \to \{0,1\}^m$ which is simply a version of FFM_N that can output multiple bits (by taking the first m bits produced from multiplying its inputs over $\mathbb{F}_{2^n}$, instead of just the first bit). It is now immediate that $\mathsf{prodExt}_N : (\{0,1\}^n)^N \to \{0,1\}^m$ is a leakage-resilient extractor against $\mathsf{NOF}(\mu)$ for entropy $k = n$ with error ϵ and output length m, where μ, ϵ, m are as above. Furthermore, at this point, we could plug $\mathsf{prodExt}$ into the analysis of Section II-A to see that $\mathsf{prodExt}$ is also an explicit leakage-resilient extractor against $\mathsf{BCP}(p,\mu)$

Authorized licensed use limited to: University of Texas at Austin. Downloaded on August 15,2021 at 16:31:35 UTC from IEEE Xplore. Restrictions apply.

for min-entropy $k = n$, with output length $m = \mu$ and error $\epsilon = 2^{-\mu}$, where $\mu = \Omega\left(n^{\frac{\log(N/p)}{\log(N/p)+1}}\right)$. This yields a "uniform" version of Theorem 2, with slightly better parameters.

We now arrive at the interesting part of the construction: dropping the entropy requirement. We start with the observation that without any further modifications, $\mathsf{prodExt}$ actually works when given sources that are missing a little min-entropy, simply by treating them as leakage. The permissible missing entropy, however, is extremely small - even for $N = O(1)$ sources, this approach could never do better than requiring $k \geq 0.99n$. Luckily, however, this "near-uniform" leakage-resilient extractor is just good enough to enable the next step, and heart, of our construction.

The main idea is to preprocess the inputs to $\mathsf{prodExt}$ using a *low-error strong two-source condenser* $\mathsf{2Cond} : (\{0,1\}^n)^2 \rightarrow \{0,1\}^\ell$ for polylogarithmic min-entropy. Given two independent sources, such an object is guaranteed to output a source with extremely high entropy, *even upon fixing one of the inputs* (with high probability). An explicit construction of such an object was only recently designed by Ben-Aroya et al. [53], but it has already found great use in extractor constructions.

We are now ready to construct our LRE that can handle polylogarithmic entropy. The idea is to simply *sacrifice one of the sources* to use as a common seed across $N-1$ condenser calls (one for every other source). Then, the outputs of these condenser calls are fed into $\mathsf{prodExt}$. More formally, we construct our *low-entropy* LRE, which we call $\mathsf{leProdExt} : (\{0,1\}^n)^N \rightarrow \{0,1\}^m$ as follows: on input $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_N)$, define $\mathbf{Y}_i := \mathsf{2Cond}(\mathbf{X}_i, \mathbf{X}_N)$, for each $i \in [N-1]$. Then, construct $\mathsf{prodExt} : (\{0,1\}^\ell)^{N-1} \rightarrow \{0,1\}^m$ as before, and simply output $\mathsf{prodExt}(\mathbf{Y}_1, \ldots, \mathbf{Y}_{N-1})$.

To prove that this leakage-resilient extractor achieves the parameters advertised in Theorem 2, all that we need to do is fix the common seed $\mathbf{X}_N$, and by a union bound, each random variable $\mathbf{Y}_i, i \in [N-1]$ obtains extremely high min-entropy. Furthermore, fixing $\mathbf{X}_N$ makes $\mathbf{Y}_i$ a deterministic function of $\mathbf{X}_i$, for each $i \in [N-1]$, and thus these random variables are all independent. This essentially completes the reduction from the low-entropy case for $\mathsf{leProdExt}$ to the near-uniform case for $\mathsf{prodExt}$, whose analysis we already know, from above. All that remains is the simple observation that leaks we must provide resilience against are still acting on $\{\mathbf{X}_i\}_{i\in[N-1]}$, whereas we need them to be acting on $\{\mathbf{Y}_i\}_{i\in[N-1]}$ (the inputs to $\mathsf{prodExt}$) in order to apply this extractor's leakage-resilience properties. We can easily take care of this, simply by fixing any additional randomness that appears in $\{\mathbf{X}_i\}_{i\in[N-1]}$ (but not in $\{\mathbf{Y}_i\}_{i\in[N-1]}$). This proves Theorem 2.

A compiler that produces LREs against BCPs: Above, we discussed how we can transform a single explicit function (FFM_N) with lower bounds against NOF protocols into a much stronger object: a leakage-resilient extractor against BCPs for polylogarithmic min-entropy. Given this construction, it is natural to wonder whether there exists a general transformation, or *compiler*, that could transform *any* explicit function with NOF lower bounds into LREs against BCPs. Such a compiler would be useful because it would be guaranteed to produce improved leakage-resilient extractors as NOF bounds are strengthened over time. In the full version of our paper, we explicitly construct exactly such an object.

The spirit of our compiler is very similar to that of our simple LRE that is discussed above. However, several complications arise from the fact that our construction above relied heavily on FFM_N in a white box manner. In particular, we relied on its *self-reducibility*, in the sense that FFM_N looks like FFM_t on any t of its inputs. If we wish to simulate this behavior in a black box manner (with an arbitrary function hard that has NOF lower bounds), one natural idea is to make a call to hard over every t-tuple of inputs, and take the bitwise XOR of the results. But even if the sources are uniform, we immediately run into a problem: if $t = \omega(1)$, then this construction is no longer efficient.

We circumvent the above issue by using a *sampler* over the sources to select t-tuples, instead of brute-forcing over all of them. But now, if we wish to drop the entropy of the sources by even just a little, this could incur too large an error in the sampler, which was expecting to receiving uniform bits. It turns out, however, that by sacrificing an additional source and using some extra two-source extractor calls, we can bypass this issue as well. As a result, we are able to successfully adapt the FFM_N-to-LRE transformation above to work with an arbitrary NOF-hard function in a black box manner, and thereby construct an explicit compiler.

C. LREs Against NOFs via Exponential Sums

Our polylogarithmic extractor described above works for all $p \leq N-2$. The case $p = N-1$ corresponding to NOF leakage seems much more difficult. In this case, relying on recent results in additive combinatorics, we are still able to construct extractors that can handle min-entropy $k \geq 0.3n$, and prove Theorem 3. Our starting point is a result of Kamp, Rao, Vadhan, and Zuckerman [54], who used exponential sum estimates of Bourgain, Glibichuk, and Konyagin [55] to construct extractors for any constant min-entropy rate, given just a constant number of independent sources.

To give our extractor construction, we begin with some notation. Let $\mathbb{F}_q$ be the prime field of cardinality q. Inspired by the extractor of Bourgain [22] and Kamp et al. [54], our N source extractor $\mathsf{Ext} : \mathbb{F}_q^N \rightarrow \{0,1\}$ is defined as

$$\mathsf{BouExt}(x_1, \ldots, x_N) = \mathsf{sign} \sin\left(\frac{2\pi \prod_{i\in[N]} x_i}{q}\right),$$

Authorized licensed use limited to: University of Texas at Austin. Downloaded on August 15,2021 at 16:31:35 UTC from IEEE Xplore. Restrictions apply.

where sign is the usual sign function defined as $\mathsf{sign}(x) = 1$ if and only if $x \geq 0$. Bourgain [22] noted that the above is an extractor for rate δ if we can obtain non-trivial upper bounds on the exponential sum

$$\left| \sum_{x_1 \in X_1} \cdots \sum_{x_N \in X_N} e_q \left(\prod_{i \in [N]} x_i \right) \right|,$$

where e_q is the exponential function, defined as $e_q(x) = \exp\left(\frac{2\pi i x}{q}\right)$, and $X_1, \ldots, X_N$ are arbitrary subsets of $\mathbb{F}_q$ of size q^δ. Bounds with optimal subset sizes have been obtained by Bourgain [56].

Intuitively, notice that the choice of x_2 is independent of the choice of x_1, and thus sums of the above form correspond to independent source extractors. To model sources correlated by NOF leakage, we need to look towards a richer class of exponential sums. For example, to model three sources against NOF leakage, we can use three indicator functions (traditionally known as cylinders) $\phi_{1,2}, \phi_{2,3}, \phi_{1,3}$, each of the form $\mathbb{F}_q^2 \to \{0,1\}$. In more detail, the cylinder $\phi_{1,2}$ decides whether or not to sum over the input x_1, x_2, modeling the correlation between the pair of sources. Correlation between every pair of sources can then be modeled as the product of these three cylinders (traditionally known as a *cylinder-intersection* [17]). Formally,

$$\phi(x_1, x_2, x_3) = \phi_{1,2}(x_1, x_2) \cdot \phi_{2,3}(x_2, x_3) \cdot \phi_{1,3}(x_1, x_3).$$

Equipping the previous exponential sum with these indicators yields a richer class of sums, which are of the form

$$\left| \sum_{x_1 \in X_1} \sum_{x_2 \in X_2} \sum_{x_3 \in X_3} \phi(x_1, x_2, x_3) \cdot e_q \left(\prod_{i \in [3]} x_i \right) \right|.$$

Fortunately, such multilinear exponential sums have been recently considered in the additive combinatorics literature, starting with the work of Petridis and Shparlinski [57], who obtained concrete bounds for the special cases of 3 and 4 sources. Very recently, Kerr and Macourt [58] generalized their result to N sources, for $N \ll \log\log q$.

It has been shown in [17] that each possible transcript of an NOF protocol induces a cylinder-intersection on its inputs. This observation allows us to rely on the above multilinear exponential sum upper bounds to prove that BouExt is in fact an LRE against $\mathsf{BCP}(N-1, \mu)$. In the full version of this paper, we also consider more general exponential sums, which may be of independent interest.

D. Leakage-Resilient Secret Sharing Schemes

In order to construct our LRSS schemes, we start by constructing secret sharing schemes that are resilient against adaptive but disjoint leakage, and then extend it to obtain Theorem 4, which works for overlapping leakage. We begin by considering a leader-based t-out-of-N scheme, which at first sight looks artificial, but proves instrumental in both of these results. Our notion can be seen as a generalization of an idea present in the recent work of Aggarwal et al. [43], who implicitly designed leader-based 2-out-of-N schemes against non-adaptive and individual leakage, while designing general LRSS schemes in the same leakage model.

Leader-based t-out-of-N schemes: For any "leader" $\ell \in [N]$, we define and construct t-out-of-N schemes for leader ℓ, which allow the leader and any $t-1$ other parties to reconstruct the secret. More importantly, it guarantees that the transcript of any protocol over the two unauthorized subsets, namely $[N] \setminus \{\ell\}$ and $\{\ell\} \cup U$ for any $|U| = t-2$, reveals nothing about the underlying secret.

Standard SS schemes from leader-based schemes: The idea now is to share the secret Ψ using any standard t-out-of-N scheme to obtain N shares $\Psi_1, \ldots, \Psi_N \leftarrow \mathsf{Share}_t^N(\Psi)$, and make each of the N parties the leader for exactly one of these shares. That is, Ψ_i is shared using a t-out-of-N scheme for leader i. Notice that any set of less than t parties of the final scheme can only have at most $t-1$ leaders, and consequently the secret will be hidden. To prove leakage-resilience, we use a hybrid argument, which relies on the leakage-resilience that is ensured by our leader-based scheme for each choice of leader. It is not hard to generalize this result to general access structures, and we defer the details to the full version.

Leakage-resilience against BCPs: We next sketch the proof of Theorem 4. We first describe the basic construction of LRSSs [15] which we will rely on. The construction in [15] can be abstracted as follows:

1) Use a function hard for NOF protocols (in a black-box way) to get an N-out-of-N LRSS against $\mathsf{BCP}(N-1, \mu)$.
2) Use several instantiations of the above LRSS, along with *perfect hash families*, to build a t-out-of-N LRSS against $\mathsf{BCP}(p, \mu)$.

Both of these steps hit barriers at $p = \omega(\log N)$ in [15]: The first step blows up the share-length by a 2^p factor, owing to the use of NOF lower bounds; and the second step incurs another $2^{O(p)}$ factor, owing to the use of perfect hash families [59].

If we use our average-case lower bounds against BCPs, as opposed to NOF lower bounds, we can already implement step (1) above without losing a 2^p factor, as long as p is any constant fraction of N. The main hurdle is now in implementing step (2) efficiently when $p = \omega(\log N)$. But we need a new idea, as there are information theoretic lower bounds against perfect hash families [60]. We will use two special ingredients to circumvent this hurdle: *ramp hash families* and *leader-based* threshold secret sharing schemes.

Ramp hash families: Inspired by the ramp secret sharing literature [61], [62] and *covering* hash families as defined in [63], we define ramp hash families as weaker analogues of perfect hash families.

Authorized licensed use limited to: University of Texas at Austin. Downloaded on August 15,2021 at 16:31:35 UTC from IEEE Xplore. Restrictions apply.

Definition 1 (Ramp hash families). *A family of hash functions $H = \{h : [N] \to [p]\}$ is called a (p, t, N)-ramp hash function family if for all subsets $T \subseteq [N]$ of cardinality t, there exists a function h in the family such that h is surjective on T — that is, $\{h(i) : i \in T\} = [p]$.*

Perfect hash families correspond to (p, p, N)-ramp hash families and necessarily need to have size at least $2^{\Omega(p)} \log N$ [60]. But, owing to a "coupon collector" phenomenon, if $t > Cp \log p$, then there exist (p, t, N)-ramp hash families with size $\mathrm{poly}(p) \cdot \log N$. One can use the probabilistic method to argue the existence of such families.

Such a property was first studied as *covering* in the work of Alon et al. [63], who asked for the stronger requirement that a random hash function from H be surjective on any fixed set T with high probability. We will use an explicit construction of such families of size $\mathrm{poly}(p, \log N)$, due to Meka, Reingold, and Zhou [64].

Given ramp hash families as above, we can implement the second step of [15] (which is based on a classical idea of Kurosawa and Stinson [65], [66]) to get "ramp" secret sharing schemes that are leakage-resilient against collusion $p = O(t/\log t)$, but satisfy a weaker secrecy guarantee. Concretely, while any t parties can recover the secret, no p parties can reconstruct the secret. However, some set of $p+1$ shares may reveal the secret, whereas we need to ensure that no $t-1$ parties can learn anything about the secret.

Stronger leader-based t-out-of-N schemes: To fix the secrecy issue, we rely on our notion of leader-based scheme, albeit a stronger one. Apart from the reconstruction property as in the disjoint case, now we also require that the transcript of any BCP, along with all but the leader's shares, reveal nothing about the underlying secret. To achieve this, we need to strengthen our communication complexity lower bounds from Theorem 1 to also hold when one set of $N-1$ parties to collude, in addition to the usual p-party collusion. Fortunately, our techniques can easily be generalized to this setting. We can then proceed as in the disjoint case, and use these leader-based schemes to get standard t-out-of-N SS schemes, proving leakage-resilience via an appropriately modified hybrid argument.

III. Preliminaries

We provide here formal definitions of the objects discussed informally in the introduction (BCPs, communication complexity and leakage-resilient / cylinder-intersection extractors against BCPs). Additionally, we provide some basic notation and concepts regarding probability. The reader should feel free to skip this section and return should anything become unclear.

A. Basic Notation, Definitions, and Objects

Notation: Throughout, we let $\circ$ denote string concatenation. For a string $x \in \{0,1\}^n$, we let x_i denote the value at its ith coordinate, and we let x_S for some $S \subseteq [n]$ denote the concatenation of all $x_i, i \in S$, in increasing order of i. For a string $x \in (\{0,1\}^n)^N$ and $i \in [N]$, we let x_i denote its ith chunk of n consecutive bits, and for a set $S \subseteq [N]$, we define x_S as the concatenation of all chunks indexed by $i \in S$.

Probability: The *min-entropy* of a random variable $\mathbf{X}$ over $\{0,1\}^n$ is defined as $\min_{x \in \text{support}(\mathbf{X})} \log(1/\Pr[\mathbf{X} = x])$, and $\mathbf{X}$ is said to be an (n, k) *source* if it has min-entropy at least k. The *statistical distance* between distributions $\mathbf{X}$ and $\mathbf{Y}$ over $\{0,1\}^n$, is defined as $|\mathbf{X} - \mathbf{Y}| := \frac{1}{2} \sum_{v \in \{0,1\}^n} |\Pr[\mathbf{X} = v] - \Pr[\mathbf{Y} = v]|$, or equivalently, $\max_{S \subseteq \{0,1\}^n} |\Pr[\mathbf{X} \in S] - \Pr[\mathbf{Y} \in S]|$.

B. Bounded Collusion Protocols

While we believe the informal definitions of BCPs (provided in the introduction) to be much more illuminating, we provide here the formal definitions of non-adaptive and adaptive BCPs from Kumar, Meka and Sahai [15].

Definition 2. *A function $f : (\{0,1\}^n)^N \to \{0,1\}^\mu$ is in the class of* non-adaptive bounded collusion protocols $\mathsf{nBCP}_{N,n}(p, \mu)$ *if: for every $i \in [\mu]$, there exists a subset $S_i \subseteq [N]$ of size p, and a function $g_i : (\{0,1\}^n)^p \to \{0,1\}$ such that for every $x \in (\{0,1\}^n)^N$, it holds that $f(x) = (y_1, y_2, \ldots, y_\mu)$, where $y_i := g_i(x_{S_i})$, for every $i \in [\mu]$.*

Definition 3. *A function $f : (\{0,1\}^n)^N \to \{0,1\}^\mu$ is in the class of* adaptive bounded collusion protocols $\mathsf{BCP}_{N,n}(p, \mu)$ *if: for every $i \in [\mu]$, there exists a subset function $S_i : \{0,1\}^{i-1} \to \binom{[N]}{p}$, and a function $g_i : \{0,1\}^{i-1} \times (\{0,1\}^n)^p \to \{0,1\}$ such that for every $x \in (\{0,1\}^n)^N$, it holds that $f(x) = (y_1, y_2, \ldots, y_\mu)$, where $y_i := g_i(y_1, y_2, \ldots, y_{i-1}, x_{S_i(y_1, y_2, \ldots, y_{i-1})})$, for every $i \in [\mu]$.*

For ease of exposition, we will occasionally shorten the names of these classes. For example, when N, n are clear from context, we will drop the subscripts of the class names, and when p, μ are also clear from context, we simply write nBCP and BCP. Furthermore, we define number-on-forehead protocols and number-in-hand protocols in the expected way: we let $\mathsf{nNOF}(\mu) := \mathsf{nBCP}(N-1, \mu)$ and $\mathsf{NOF}(\mu) := \mathsf{BCP}(N-1, \mu)$, and let $\mathsf{nNIH}(\mu) := \mathsf{nBCP}(1, \mu)$ and $\mathsf{NIH}(\mu) := \mathsf{BCP}(1, \mu)$.

Next, we extend the definition of communication complexity against NOF and NIH protocols to BCPs in the expected way, and also provide an alternative definition that will be useful later on.

Definition 4. *The* communication complexity *of a function $f : (\{0,1\}^n)^N \to \{0,1\}$ against BCPs with collusion bound p, denoted $\mathsf{CC}^p(f)$, is defined as the minimum μ such that there exists some $g : (\{0,1\}^n)^N \to \{0,1\}^\mu \in \mathsf{BCP}(p, \mu)$ such that for every $x \in (\{0,1\}^n)^N$, it holds that $f(x) = g(x)_\mu$.*

Authorized licensed use limited to: University of Texas at Austin. Downloaded on August 15,2021 at 16:31:35 UTC from IEEE Xplore. Restrictions apply.

Remark 1. *We can equivalently define* $\mathsf{CC}^p(f)$ *as the minimum* μ *such that there exists some* $g : (\{0,1\}^n)^N \to \{0,1\}^\mu \in \mathsf{BCP}(p,\mu)$ *such that if we sample* $\mathbf{X}$ *uniformly from* $(\{0,1\}^n)^N$*, then with probability 1 over fixing* $g(\mathbf{X})$*, it holds that* $f(\mathbf{X})$ *becomes fixed.*

We now do the same for *distributional* (average-case) communication complexity.

Definition 5. *The* ϵ-distributional communication complexity *of a function* $f : (\{0,1\}^n)^N \to \{0,1\}$ *against BCPs with collusion bound p, denoted* $\mathsf{CC}^p_\epsilon(f)$*, is defined as the minimum* μ *such that there exists some* $g : (\{0,1\}^n)^N \to \{0,1\}^\mu \in \mathsf{BCP}(p,\mu)$ *such that* $f(\mathbf{X}) = g(\mathbf{X})_\mu$ *with probability at least* $(1+\epsilon)/2$ *over sampling* $\mathbf{X}$ *uniformly from* $(\{0,1\}^n)^N$.

Remark 2. *We can equivalently define* $\mathsf{CC}^p_\epsilon(f)$ *as the minimum* μ *such that there exists some* $g : (\{0,1\}^n)^N \to \{0,1\}^\mu \in \mathsf{BCP}(p,\mu)$ *such that if we sample* $\mathbf{X}$ *uniformly at random from* $(\{0,1\}^n)^N$*, then* $|f(\mathbf{X}) \circ g(\mathbf{X}) - \mathbf{U}_1 \circ g(\mathbf{X})| \geq \epsilon/2$*, where* $\mathbf{U}_1$ *is independent from* $g(\mathbf{X})$.

C. Randomness Extractors

In this paper, we study a much stronger (and more general) version of classical extractors, called *leakage-resilient extractors (LREs)*. Such extractors offer the additional guarantee that their output looks uniform *even conditioned on* leakage from certain families of functions.

Definition 6. *Let* $\mathcal{X}$ *be a family of sources over* $\{0,1\}^n$*, and* $\mathcal{F} \subseteq \{f : \{0,1\}^n \to \{0,1\}^\mu\}$ *a family of* leakage *functions. A function* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ *is a* leakage-resilient extractor (LRE) *against* $\mathcal{F}$ *for the family* $\mathcal{X}$*, with error* ϵ*, if for every* $\mathbf{X} \in \mathcal{X}, f \in \mathcal{F}$*,*

$$|\mathsf{Ext}(\mathbf{X}) \circ f(\mathbf{X}) - \mathbf{U}_m \circ f(\mathbf{X})| \leq \epsilon,$$

where $\mathbf{U}_m$ *is independent from* $f(\mathbf{X})$.

The particular type of LREs that we examine in this paper are *cylinder-intersection extractors*, which are multi-source extractors that offer leakage-resilience against BCPs:

Definition 7 ([15])**.** *A function* $\mathsf{Ext} : (\{0,1\}^n)^N \to \{0,1\}^m$ *is a* (p,N,μ)-cylinder intersection extractor *for* (n,k) *sources and error* ϵ *if* Ext *is an LRE against* $\mathsf{BCP}_{N,n}(p,\mu)$ *with error* ϵ*, for the family of sources* $\mathcal{X}$ *where each* $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_N) \in \mathcal{X}$ *consists of* N *independent* (n,k) *sources. Equivalently, we also call such objects* LREs against BCPs.

Given Definition 6, Definition 7, and Remark 2, it is easy to see that leakage-resilient extractors against BCPs are strictly stronger than average-case lower bounds against BCPs.

IV. Explicit LREs Against BCPs

In this section, we will prove Theorems 1 and 2. As discussed in the overview, the foundation of our leakage-resilient extractors is the finite field multiplication function $\mathsf{FFM}_N : (\{0,1\}^n)^N \to \{0,1\}$. This function takes as input N bitstrings of length n, treats them as elements of $\mathbb{F}_{2^n}$, takes their product over this field, interprets the result again as a bitstring over $\{0,1\}^n$, and outputs the first bit. In [18], Ford and Gál show that $\mathsf{CC}^{\mathsf{NOF}}_\epsilon(\mathsf{FFM}_N) \geq \Omega(n/2^N)$ for $\epsilon = 2^{-\Omega(n/2^N)}$, thereby proving strong *average-case* lower bounds against this very simple function.

As a first step in transforming FFM_N into a low-entropy LRE against BCPs, we slightly modify this function so that it can output many bits. In particular, for $m \leq n$, we let $\sigma_{n,m} : \mathbb{F}_{2^n} \to \mathbb{F}_{2^m}$ denote the function that interprets its input as an element of $\mathbb{F}_2^n$ and projects it onto its first m coordinates, and we define a multi-bit output version of FFM_N as follows.

Definition 8. *For any* $N, n, m \in \mathbb{N}$ *with* $n \geq m$*, the* product extractor $\mathsf{prodExt} : (\{0,1\}^n)^N \to \{0,1\}^m$ *is defined as:*

$$\mathsf{prodExt}(x_1, x_2, \ldots, x_N) := \sigma_{n,m}(x_1 \cdot x_2 \cdot \cdots \cdot x_N),$$

where the inputs are interpreted as elements of $\mathbb{F}_{2^n}$*, and the product is taken over this field.*

We remark that, using the standard encoding of $\mathbb{F}_{2^n}$, it is straightforward to perform all the above operations in $\mathrm{poly}(n,N)$ time (see, e.g., [20]). Now, because we have modified the original extractor from [18], we are not able to apply their correlation bound on FFM as a black box in order to show that $\mathsf{prodExt}$ is a basic leakage-resilient extractor. Instead, we dig into their proof and show that the main character sum in their work can be combined with standard XOR lemmas [52] to yield the following:

Lemma 1. *For all* $N, n \in \mathbb{N}$ *such that* $N \geq 2$*, the product extractor* $\mathsf{prodExt} : (\{0,1\}^n)^N \to \{0,1\}^m$ *from Definition 8 is an explicit leakage-resilient extractor against* $\mathsf{nNOF}(\mu)$ *for entropy* $k = n$ *and leakage* $\mu = \xi$*, output length* $m = \xi$*, and error* $\epsilon = 2^{-\xi}$*, where* $\xi = \Omega(n/2^N)$.

As the techniques used to prove Lemma 1 are relatively standard, we refer the reader to a full version of the paper for a proof.

A. Handling More Leakage When There Is Less Collusion

Next, we show that without any further modifications, $\mathsf{prodExt}$ can handle leakage from BCPs across a very general range of parameters, and furthermore achieve a nontrivial tradeoff between leakage (complexity) and collusion. In particular, we prove the following lemma, and note that we optimize to pick a good setting for t in Section IV-C.

Lemma 2. *For all sufficiently large* $N, n \in \mathbb{N}$ *and any* $t, p \in \mathbb{N}$ *such that* $t \leq N$ *and* $p \leq N-1$*, the product extractor* $\mathsf{prodExt} : (\{0,1\}^n)^N \to \{0,1\}^m$ *from Definition 8*

Authorized licensed use limited to: University of Texas at Austin. Downloaded on August 15,2021 at 16:31:35 UTC from IEEE Xplore. Restrictions apply.

is an explicit leakage-resilient extractor against $\mathsf{nBCP}(p,\mu)$ *for entropy* $k = n$ *and leakage* $\mu < \min\{\xi, \binom{N}{t}/\binom{p}{t}\}$, *with output length* $m = \xi$ *and error* $\epsilon = 2^{-\xi}$, *where* $\xi = \Omega(n/2^t)$.

Proof: We must show that for N independent uniform sources $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_N)$ and any $\mathsf{Leak} \in \mathsf{nBCP}(p,\mu)$ with $\mu < \xi$ and $\mu\binom{p}{t} < \binom{N}{t}$,

$$|\mathsf{prodExt}(\mathbf{X}) \circ \mathsf{Leak}(\mathbf{X}) - \mathbf{U}_m \circ \mathsf{Leak}(\mathbf{X})| \leq \epsilon,$$

where $\epsilon = 2^{-\xi}$. For brevity, we let $\mathbf{Z}_1 := \mathsf{prodExt}(\mathbf{X}) \circ \mathsf{Leak}(\mathbf{X})$ and $\mathbf{Z}_2 := \mathbf{U}_m \circ \mathsf{Leak}(\mathbf{X})$ and show $|\mathbf{Z}_1 - \mathbf{Z}_2| \leq \epsilon$. By Definition 2, there must exist some $S_1, \ldots, S_\mu \subseteq [N]$, each of size p, and some functions $g_1, \ldots, g_\mu : (\{0,1\}^n)^p \to \{0,1\}$ such that $\mathsf{Leak}(\mathbf{X}) = (g_1(\mathbf{X}_{S_1}), \ldots, g_\mu(\mathbf{X}_{S_\mu}))$. Thus:

$$\mathbf{Z}_1 := \mathsf{prodExt}(\mathbf{X}) \circ g_1(\mathbf{X}_{S_1}) \circ \cdots \circ g_\mu(\mathbf{X}_{S_\mu}),$$
$$\mathbf{Z}_2 := \mathbf{U}_m \circ g_1(\mathbf{X}_{S_1}) \circ \cdots \circ g_\mu(\mathbf{X}_{S_\mu}).$$

The goal now is to perform fixings so as to reduce the analysis to let us apply Lemma 1 over t sources. We proceed as follows. First, notice that each of the μ subsets S_i has size p, and can therefore hold at most $\binom{p}{t}$ distinct subsets of size t. Thus, since we are told $\mu\binom{p}{t} < \binom{N}{t}$, there must be some *good* $G \in \binom{[N]}{t}$ where $G \not\subseteq S_i, \forall i \in [\mu]$. Without loss of generality, we may assume $G = [t]$ (any other case uses almost exactly the same ideas that follow, but the notation gets a little cumbersome). We let $\overline{G} = [N] \setminus G$, and by definition of statistical distance, we know that $|\mathbf{Z}_1 - \mathbf{Z}_2| \leq |\mathbf{Z}_1 \circ \mathbf{X}_{\overline{G}} - \mathbf{Z}_2 \circ \mathbf{X}_{\overline{G}}|$. This means there is some fixed $X^* \in (\{0,1\}^n)^{N-t}$ such that

$$|\mathbf{Z}_1 - \mathbf{Z}_2| \leq |(\mathbf{Z}_1 \mid \mathbf{X}_{\overline{G}} = X^*) - (\mathbf{Z}_2 \mid \mathbf{X}_{\overline{G}} = X^*)|.$$

Now, to see that this quantity is bounded above by ϵ, we just have to carefully rewrite the random variables. First, we note that it is safe to assume that $X_i^* \neq \vec{0}$, for all $i \in [N-t]$: this is because we can in fact preprocess the input to redirect zeroes to ones before calling prodExt. Later on in the proof, when it is convenient, we can still assume that each $\mathbf{X}_i, i \in G$ is uniform by incurring an additive $t \cdot 2^{-n+1}$ in the error (which is always insignificant). Let us now examine the conditioned versions of $\mathbf{Z}_1$ and $\mathbf{Z}_2$.

We start by observing that $(\mathsf{prodExt}(\mathbf{X}) \mid \mathbf{X}_{\overline{G}} = X^*) = \sigma_{n,m}(\mathbf{X}_1 \cdot \cdots \cdot \mathbf{X}_t \cdot X_1^* \cdot \cdots \cdot X_{N-t}^*) = \mathsf{prodExt}(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_{t-1}, \mathbf{X}_t \cdot \pi)$, where $\pi = X_1^* \cdot \cdots \cdot X_{N-t}^*$ is some fixed nonzero value in $\mathbb{F}_{2^n}$. Furthermore, for any $i \in [\mu]$, observe that $(g_i(\mathbf{X}_{S_i}) \mid \mathbf{X}_{\overline{G}} = X^*)$ becomes the function $g_i'(\mathbf{X}_{S_i \cap G})$, which is the restriction of g_i obtained by fixing the variables $\mathbf{X}_i, i \in S_i \cap \overline{G}$ according to X^*. By definition of G, we know $G \not\subseteq S_i$, and thus if we define $S_i' := S_i \cap G$, we know $S_i' \subsetneq G$, and furthermore $(g_i(\mathbf{X}_{S_i}) \mid \mathbf{X}_{\overline{G}} = X^*) = g_i'(\mathbf{X}_{S_i'})$. Thus we have:

$$(\mathbf{Z}_1 \mid \mathbf{X}_{\overline{G}} = X^*) = \mathsf{prodExt}(\mathbf{X}_1, \ldots, \mathbf{X}_t \cdot \pi) \circ g_1'(\mathbf{X}_{S_1'}) \circ \cdots \circ g_\mu'(\mathbf{X}_{S_\mu'}), \text{ and}$$
$$(\mathbf{Z}_2 \mid \mathbf{X}_{\overline{G}} = X^*) = \mathbf{U}_m \circ g_1'(\mathbf{X}_{S_1'}) \circ \cdots \circ g_\mu'(\mathbf{X}_{S_\mu'}).$$

Suppose now that we define $\mathbf{Y}_i, i \in [t]$ as $\mathbf{Y}_i := \mathbf{X}_i$ when $i \in [t-1]$, and $\mathbf{Y}_i := \mathbf{X}_t \cdot \pi$ when $i = t$. Then, for each $i \in [\mu]$, we define g_i'' to be identical to g_i', except for the fact that if g_i' receives $\mathbf{X}_t$ as an input - say, as its jth argument - then g_i'' multiplies its jth input by *the inverse of* π (in $\mathbb{F}_{2^n}^*$) before passing all its input into g_i', and returning the result. By construction, such a function guarantees $g_i''(\mathbf{Y}_{S_i'}) = g_i'(\mathbf{X}_{S_i'})$. And thus, we see that we can write:

$$(\mathbf{Z}_1 \mid \mathbf{X}_{\overline{G}} = X^*) = \mathsf{prodExt}(\mathbf{Y}_1, \ldots, \mathbf{Y}_t) \circ g_1''(\mathbf{Y}_{S_1'}) \circ \cdots \circ g_\mu''(\mathbf{Y}_{S_\mu'}), \text{ and}$$
$$(\mathbf{Z}_2 \mid \mathbf{X}_{\overline{G}} = X^*) = \mathbf{U}_m \circ g_1''(\mathbf{Y}_{S_1'}) \circ \cdots \circ g_\mu''(\mathbf{Y}_{S_\mu'}).$$

Since $\mathbf{Y}_t$ is just a permutation of $\mathbf{X}_t$, it must have the same entropy, and furthermore note that each g_i'' acts as non-adaptive NOF leakage on $\mathbf{Y}_1, \ldots, \mathbf{Y}_t$, since $S_i' \subsetneq [t]$. Thus we can use Lemma 1 to bound the difference $|(\mathbf{Z}_1 \mid \mathbf{X}_{\overline{G}} = X^*) - (\mathbf{Z}_2 \mid \mathbf{X}_{\overline{G}} = X^*)|$ as desired, completing the proof. ■

B. Reducing the Entropy Requirement

In this section, we will show how to improve the entropy requirement of prodExt from $k = n$ all the way to $k = \operatorname{polylog} n$, ultimately proving Theorem 2. The first step we take in this direction is a modest one: we show that without any further modificiations, prodExt will still work if its inputs are missing just a little entropy. More generally, we prove the following result:

Lemma 3. *Let* $\mathsf{Ext} : (\{0,1\}^n)^N \to \{0,1\}^m$ *be a leakage-resilient extractor against* $\mathsf{nNOF}(\mu)$ *for entropy* $k = n$ *and error* ϵ. *Then for any* $0 < k \leq n$, Ext *is also a leakage-resilient extractor against* $\mathsf{nNOF}(\mu - 2)$ *for entropy* k *and error* $\epsilon \cdot 2^{N(n-k)}$.

Proof: Given N independent (n,k) sources $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_N)$ and any $\mathsf{Leak} \in \mathsf{nNOF}(\mu - 2)$, we wish to upper bound the quantity

$$|\mathsf{Ext}(\mathbf{X}) \circ \mathsf{Leak}(\mathbf{X}) - \mathbf{U}_m \circ \mathsf{Leak}(\mathbf{X})|. \tag{1}$$

We may assume that each source $\mathbf{X}_i$ is *flat*; i.e., uniform over some $T_i \subseteq \{0,1\}^n$ of size 2^k (this is a standard assumption that can be done without loss of generality; see, e.g., [20]). The main idea of this proof is to treat the missing entropy as leakage on uniform sources, by defining a function belonging to $\mathsf{nNOF}(2)$ that identifies the support of $(\mathbf{X}_1, \ldots, \mathbf{X}_N)$. In particular, we define the indicator function $\mathsf{id} : (\{0,1\}^n)^N \to \{0,1\}^2$ as the map

Authorized licensed use limited to: University of Texas at Austin. Downloaded on August 15,2021 at 16:31:35 UTC from IEEE Xplore. Restrictions apply.

$(x_1, \ldots, x_N) \mapsto (y_1, y_2)$, where $y_1 = 1$ if and only if $x_i \in T_i$, for all $i \in [N-1]$, and $y_2 = 1$ if and only if $x_N \in T_N$. Furthermore, we define $\vec{1} := (1,1)$. Given these definitions, it is straightforward to verify that $\mathsf{id}(x) = \vec{1}$ if and only if $x \in T_1 \times \cdots \times T_N$, and that $\mathsf{id} \in \mathsf{nNOF}(2)$. Thus, if we define a function $\mathsf{Leak}' : (\{0,1\}^n)^N \to \{0,1\}^\mu$ that maps $x \mapsto (\mathsf{Leak}(x), \mathsf{id}(x))$, then $\mathsf{Leak}' \in \mathsf{nNOF}(\mu)$, and we may use it to upper bound Equation (1):

$$\begin{aligned}
&|\mathsf{Ext}(\mathbf{X}) \circ \mathsf{Leak}(\mathbf{X}) - \mathbf{U}_m \circ \mathsf{Leak}(\mathbf{X})| \\
&= |\mathsf{Ext}(\mathbf{U}_{Nn} \mid \mathsf{id}(\mathbf{U}_{Nn}) = \vec{1}) \circ \mathsf{Leak}(\mathbf{U}_{Nn} \mid \mathsf{id}(\mathbf{U}_{Nn}) = \vec{1}) \\
&\qquad - \mathbf{U}_m \circ \mathsf{Leak}(\mathbf{U}_{Nn} \mid \mathsf{id}(\mathbf{U}_{Nn}) = \vec{1})| \\
&\le |\mathsf{Ext}(\mathbf{U}_{Nn}) \circ \mathsf{Leak}(\mathbf{U}_{Nn}) \circ \mathsf{id}(\mathbf{U}_{Nn}) \\
&\qquad - \mathbf{U}_m \circ \mathsf{Leak}(\mathbf{U}_{Nn}) \circ \mathsf{id}(\mathbf{U}_{Nn})| / \Pr[\mathsf{id}(\mathbf{U}_{Nn}) = \vec{1}] \\
&\le \epsilon \cdot 2^{N(n-k)},
\end{aligned}$$

where the first inequality is a Markov-type inequality, and the second inequality follows from the hypothesis. ■

The key object we need in order to drop the entropy requirement of our extractor down to $k \ge \mathrm{polylog}(n)$ is a *low-error strong two-source condenser*. Ben-Aroya et al. [53] recently constructed such objects with excellent parameters; most relevant to us here will be the following specialization of one of their more general constructions:

Theorem 5 ([53])**.** *There exist universal constants $C > 0$ and $\gamma := 1/C$ such that for every $n, k, m \in \mathbb{N}$ and $\epsilon > 0$ satisfying $k \ge \log^C n$ and $k^\gamma \ge m$ and $\epsilon \ge 2^{-k^{\gamma/2}}$, there exists an explicit function $\mathsf{2Cond} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$ such that for any two independent (n,k) sources $\mathbf{X}_1, \mathbf{X}_2$, with probability $1-\epsilon$ over $x_2 \sim \mathbf{X}_2$, the output has min-entropy $H_\infty(\mathsf{2Cond}(\mathbf{X}_1, x_2)) \ge m - \sqrt{m}$.*

With this condenser in hand, we are ready to define our final, low-entropy, version of the product-extractor:

Definition 9. *For a sufficiently large constant $C \ge 1$ and any $N, n, k, m_0, m \in \mathbb{N}$ satisfying $k \ge \log^C n$ and $k^{1/C} \ge m_0 \ge m$, let $\mathsf{2Cond} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^{m_0}$ be the condenser for (n,k) sources from Theorem 5, and let $\mathsf{prodExt}(\{0,1\}^{m_0})^{N-1} \to \{0,1\}^m$ be the product extractor from Definition 8. We define the* low entropy product extractor, *$\mathsf{leProdExt} : (\{0,1\}^n)^N \to \{0,1\}^m$, as*

$$\begin{aligned}
&\mathsf{leProdExt}(x_1, x_2, \ldots, x_N) \\
&\qquad := \mathsf{prodExt}((\mathsf{2Cond}(x_i, x_N))_{i \in [N-1]}).
\end{aligned}$$

We now prove the main lemma of the section, which proves that leProdExt can in fact handle low-entropy, while simultaneously achieving a leakage-collusion tradeoff similar to Lemma 2. We remark again here that we will optimize to pick a good setting for t in Section IV-C.

Lemma 4. *There is a constant $C \ge 1$ such that for all sufficiently large $N, n \in \mathbb{N}$ and any $t, p \in \mathbb{N}$ such that $t \le N-1$ and $p \le N-2$, the low-entropy product extractor $\mathsf{leProdExt} : (\{0,1\}^n)^N \to \{0,1\}^m$ from Definition 9 is an explicit leakage-resilient extractor against $\mathsf{nBCP}(p,\mu)$ for entropy $k \ge \log^C n$ and leakage $\mu < \min\{\xi, \binom{N-1}{t}/\binom{p}{t}\}$, with output length $m = \xi$ and error $\epsilon = 2^{-\xi}$, where*

$$\xi = k^{\Omega(1)}/2^t.$$

Proof: We must show that for N independent (n,k) sources $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_N)$, and any $\mathsf{Leak} \in \mathsf{nBCP}(p,\mu)$ with $\mathsf{Leak} \in \mathsf{nBCP}(p,\mu)$ with $\mu < \xi$ and $\mu\binom{p}{t} < \binom{N-1}{t}$,

$$|\mathsf{leProdExt}(\mathbf{X}) \circ \mathsf{Leak}(\mathbf{X}) - \mathbf{U}_m \circ \mathsf{Leak}(\mathbf{X})| \le \epsilon,$$

where $\epsilon = 2^{-\xi}$. For brevity, we let $\mathbf{Z}_1 := \mathsf{leProdExt}(\mathbf{X}) \circ \mathsf{Leak}(\mathbf{X})$ and $\mathbf{Z}_2 := \mathbf{U}_m \circ \mathsf{Leak}(\mathbf{X})$ and show $|\mathbf{Z}_1 - \mathbf{Z}_2| \le \epsilon$. By Definition 2, there must exist some $S_1, \ldots, S_\mu \subseteq [N]$, each of size p, and some functions $g_1, \ldots, g_\mu : (\{0,1\}^n)^p \to \{0,1\}$ such that $\mathsf{Leak}(\mathbf{X}) = (g_1(\mathbf{X}_{S_1}), \ldots, g_\mu(\mathbf{X}_{S_\mu}))$. Thus, substituting in the definition of leProdExt, we have

$$\begin{aligned}
\mathbf{Z}_1 &:= \mathsf{prodExt}((\mathsf{2Cond}(\mathbf{X}_i, \mathbf{X}_N))_{i \in [N-1]}) \\
&\qquad \circ g_1(\mathbf{X}_{S_1}) \circ \cdots \circ g_\mu(\mathbf{X}_{S_\mu}), \text{ and} \\
\mathbf{Z}_2 &:= \mathbf{U}_m \circ g_1(\mathbf{X}_{S_1}) \circ \cdots \circ g_\mu(\mathbf{X}_{S_\mu}).
\end{aligned}$$

We remark that any parameters in the construction itself (like condenser output length, condenser error, product extractor output length, etc.) will be set at the end so that everything works out.

The goal now is to perform fixings to reduce the analysis to the analysis in Lemma 2. We proceed as follows. First, notice that each of the μ subsets S_i has size p, and can therefore hold at most $\binom{p}{t}$ distinct subsets of size t. Thus, since we are told $\mu\binom{p}{t} < \binom{N-1}{t}$, there must be some *good* $G \in \binom{[N-1]}{t}$ where $G \not\subseteq S_i, \forall i \in [\mu]$. Like in the proof to Lemma 2, we will assume, without loss of generality, that $G = [t]$.

Now, we let ϵ_1 be the error of 2Cond, meaning that with probability $1 - \epsilon_1$ over $x_N \sim \mathbf{X}_N$, $\mathsf{2Cond}(\mathbf{X}_i, x_N)$ has entropy at least $m_0 - \sqrt{m_0}$, for any single $i \in [N-1]$. Thus, by a union bound, with probability $1 - t\epsilon_1$ over $x_N \sim \mathbf{X}_N$, *every* random variable in $\{\mathsf{2Cond}(\mathbf{X}_i, x_N) : i \in G\}$ has entropy at least $m_0 - \sqrt{m_0}$. In other words, there is always some X_N^* such that the following is true:

$$\begin{aligned}
|\mathbf{Z}_1 - \mathbf{Z}_2| &\le |\mathbf{Z}_1 \circ \mathbf{X}_N - \mathbf{Z}_2 \circ \mathbf{X}_N| \\
&\le t\epsilon_1 + |(\mathbf{Z}_1 \mid \mathbf{X}_N = X_N^*) - (\mathbf{Z}_2 \mid \mathbf{X}_N = X_N^*)|,
\end{aligned}$$

where $\mathsf{2Cond}(\mathbf{X}_i, X_N^*)$ has entropy at least $m_0 - \sqrt{m_0}$, for each $i \in G$. We now define $\mathbf{Y}_i := \mathsf{2Cond}(\mathbf{X}_i, X_N^*)$ for each $i \in [N-1]$, and we notice that this collection of random variables are mutually independent (because they are single-argument deterministic functions of the independent random variables $\{\mathbf{X}_i\}_{i \in [N-1]}$). We write $\mathbf{Y} = \mathbf{Y}_1 \circ \cdots \circ \mathbf{Y}_{N-1}$. We now fix each $\mathbf{X}_i, i \notin G \cup \{N\}$ to some $X_i^* \in \{0,1\}^n$ such that the following holds:

$$\begin{aligned}
&|(\mathbf{Z}_1 \mid \mathbf{X}_N = X_N^*) - (\mathbf{Z}_2 \mid \mathbf{X}_N = X_N^*)| \\
&\quad \le |(\mathbf{Z}_1 \mid \mathbf{X}_i = X_i^*, \forall i \notin G) - (\mathbf{Z}_2 \mid \mathbf{X}_i = X_i^*, \forall i \notin G)|.
\end{aligned}$$

Authorized licensed use limited to: University of Texas at Austin. Downloaded on August 15,2021 at 16:31:35 UTC from IEEE Xplore. Restrictions apply.

Notice that as a result, each $\mathbf{Y}_i, i \notin G \cup \{N\}$, gets fixed to $Y_i^* = \mathsf{2Cond}(X_i^*, X_N^*) \in \{0,1\}^{m_0}$, while each $\mathbf{Y}_i, i \in G$ still has entropy at least $m_0 - \sqrt{m_0}$, and the collection $\{\mathbf{Y}_i : i \in G\}$ remains mutually independent. By following the exact same reasoning as in the proof to Lemma 2 about restrictions and the structure of prodExt, we know that at this point we can write

$$\begin{aligned}
&(\mathbf{Z}_1 \mid \mathbf{X}_i = X_i^*, \forall i \notin G) \\
&= \mathsf{prodExt}(\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_{t-1}, \mathbf{Y}_t \cdot \pi) \\
&\qquad \circ g_1'(\mathbf{X}_{S_1 \cap G}) \circ \cdots \circ g_\mu'(\mathbf{X}_{S_\mu \cap G}) =: \mathbf{Z}_1', \text{ and} \\
&(\mathbf{Z}_2 \mid \mathbf{X}_i = X_i^*, \forall i \notin G) \\
&= \mathbf{U}_m \circ g_1'(\mathbf{X}_{S_1 \cap G}) \circ \cdots \circ g_\mu'(\mathbf{X}_{S_\mu \cap G}) =: \mathbf{Z}_2',
\end{aligned}$$

where $\pi = Y_{t+1}^* \cdot \cdots \cdot Y_{N-1}^*$, and each g_i' is the appropriate restriction of g_i induced by the fixings of its inputs outside G. We are nearly ready to apply the leakage-resilience of prodExt against NOF protocols and complete the proof. In order to arrive at this situation, we must somehow write each leakage function g_i' as a function of random variables from $\{\mathbf{Y}_i : i \in G\}$ instead of $\{\mathbf{X}_i : i \in G\}$. It turns out this is not so difficult to do: we know that for each $i \in [N-1]$, $\mathbf{Y}_i$ is a deterministic function of $\mathbf{X}_i$. As such, for each i, we can find some randomness $\mathbf{Q}_i$ and a deterministic function sample such that $\mathbf{Q}_i$ is independent of $\mathbf{Y}_i$, and $\mathsf{sample}(\mathbf{Y}_i, \mathbf{Q}_i)$ is arbitrarily close to the distribution $\mathbf{X}_i$: for any fixed $y \sim \mathbf{Y}_i$, the function call $\mathsf{sample}(y, \mathbf{Q}_i)$ simply uses $\mathbf{Q}_i$ to sample from $(\mathbf{X}_i \mid \mathbf{Y}_i = y)$.[9]

As always, we can fix the random variables $\{\mathbf{Q}_i' : i \in G\}$ to some values $\{Q_i^*\}$ without reducing the distance between $\mathbf{Z}_1', \mathbf{Z}_2'$:

$$\begin{aligned}
|\mathbf{Z}_1' - \mathbf{Z}_1'| \leq |&(\mathbf{Z}_1' \mid \mathbf{Q}_i = Q_i^*, \forall i \in G) \\
&- (\mathbf{Z}_2' \mid \mathbf{Q}_i = Q_i^*, \forall i \in G)|.
\end{aligned}$$

Furthermore, under this conditioning, we know that each $g_i'(\mathbf{X}_{S_1 \cap G})$ obtains the form $g_i''(\mathbf{Y}_{S_1 \cap G})$ for some other deterministic function g_i'', since we saw above that $\mathbf{X}_i \approx \mathsf{sample}(\mathbf{Y}_i, \mathbf{Q}_i)$, and we have fixed all the variables $\{\mathbf{Q}_i : i \in G\}$. Thus, we may write:

$$\begin{aligned}
&(\mathbf{Z}_1' \mid \mathbf{Q}_i = Q_i^*, \forall i \in G) \\
&\quad = \mathsf{prodExt}(\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_{t-1}, \mathbf{Y}_t \cdot \pi) \\
&\qquad \circ g_1''(\mathbf{Y}_{S_1 \cap G}) \circ \cdots \circ g_\mu''(\mathbf{Y}_{S_\mu \cap G}), \text{ and} \\
&(\mathbf{Z}_2' \mid \mathbf{Q}_i = Q_i^*, \forall i \in G) \\
&\quad = \mathbf{U}_m \circ g_1''(\mathbf{Y}_{S_1 \cap G}) \circ \cdots \circ g_\mu''(\mathbf{Y}_{S_\mu \cap G})
\end{aligned}$$

Using the same trick as in the proof to Lemma 2, we may remove π from the equation above, and thus we just assume now that it is no longer present. Thus, at last, we see that it suffices to upper bound $|\mathbf{Z}_1'' - \mathbf{Z}_2''|$, where

$$\begin{aligned}
\mathbf{Z}_1'' &:= \mathsf{prodExt}(\mathbf{Y}_1, \ldots, \mathbf{Y}_t) \\
&\qquad \circ g_1''(\mathbf{Y}_{S_1 \cap G}) \circ \cdots \circ g_\mu''(\mathbf{Y}_{S_\mu \cap G}), \text{ and} \\
\mathbf{Z}_2'' &:= \mathbf{U}_m \circ g_1''(\mathbf{Y}_{S_1 \cap G}) \circ \cdots \circ g_\mu''(\mathbf{Y}_{S_\mu \cap G}).
\end{aligned}$$

These random variables have a very special structure: the collection $\{\mathbf{Y}_i : i \in [t]\}$ are independent, and each is an $(m_0, m_0 - \sqrt{m_0})$ source, due to the parameters of 2Cond. Furthermore, we know $G \cap S_i \subsetneq G$, for all $i \in [\mu]$, and thus each g_i'' is an NOF-leak on $\mathbf{Y} = \mathbf{Y}_1 \circ \cdots \circ \mathbf{Y}_t$.

Thus, to conclude, we combine Lemma 1 with Lemma 3 to see that $\mathsf{prodExt} : (\{0,1\}^{m_0})^t \to \{0,1\}^m$ is a leakage-resilient extractor against $\mathsf{nNOF}(\mu_2 - 2)$ for entropy $m_0 - \sqrt{m_0}$, error $\epsilon_3 = \epsilon_2 \cdot 2^{t\sqrt{m_0}}$, and output $m = \gamma m_0 / 2^t$, where $\epsilon_2 = 2^{-m}$ and $\mu_2 = m$, for some small universal constant $\gamma > 0$. Furthermore, we know the sources $\mathbf{Y}_i, i \in [t]$ end up with the promised entropy $m_0 - \sqrt{m_0}$ as long as $k \geq \log^C n$ and $k^{1/C} \geq m_0$, with error $\epsilon_1 = 2^{-k^{1/(2C)}}$, for some universal constant $C \geq 1$. Thus, the condenser and extractor will both work as long as this entropy is guaranteed, and as long as $\mu < \mu_2 - 2$ (because this means the concatenation of leaks $g_1'', \ldots, g_\mu''$ is in $\mathsf{nNOF}(\mu_2 - 2)$). Furthermore, its error will be

$$\epsilon = |\mathbf{Z}_1 - \mathbf{Z}_2| \leq t\epsilon_1 + |\mathbf{Z}_1'' - \mathbf{Z}_2''| = t\epsilon_1 + \epsilon_3.$$

Finally, recall that we required $\mu\binom{p}{t} < \binom{N-1}{t}$ at the beginning to ensure we could find a good set G. Thus, there exists some small constant $c > 0$ and function $\xi(k,t) := k^c/2^t$ such that as long as $k \geq \log^C n$, and $\mu < \min\{\xi, \binom{N-1}{t}/\binom{p}{t}\}$, and $m \leq \xi$, it holds that $|\mathbf{Z}_1 - \mathbf{Z}_2| = \epsilon \leq 2^{-\xi}$, which completes the proof. ■

C. Adding Adaptivity and Wrapping Up

Finally, all that remains is to add adaptivity to Lemma 2 and Lemma 4, and set t, in order to obtain our main theorems about the product extractor (Theorems 1 and 2). Our main tool in this direction is the following general adaptivity lemma, which shows that any LRE against non-adaptive BCPs actually works against adaptive BCPs as well - at the expense of some loss in the error.

Lemma 5. *Let* Ext $: (\{0,1\}^n)^N \to \{0,1\}^m$ *be a leakage-resilient extractor against* $\mathsf{nBCP}(p, \mu)$ *for entropy* k*, with error* ϵ*. Then* Ext *is also a leakage-resilient extractor against* $\mathsf{BCP}(p, \mu)$ *for entropy* k *with error at most* $(2^\mu + 1)\sqrt{\epsilon}$.

The proof of this lemma works, in some sense, by approximating adaptive BCPs by non-adaptive BCPs. It can be found in the full version of the paper.

Finally, we note that by applying Lemma 5 to Lemma 2 and setting $t = \frac{\log(n)}{\log(N/p)+1}$, we immediately obtain Theorem 1. By applying Lemma 5 to Lemma 4 and setting $t = \frac{\log(k^{\Omega(1)})}{\log((N-1)/p)+1}$, we immediately obtain Theorem 2.

[9] We ignore the error from this approximation, as it can be made arbitrarily small and thereby absorbed by any other error appearing in this proof.

Authorized licensed use limited to: University of Texas at Austin. Downloaded on August 15,2021 at 16:31:35 UTC from IEEE Xplore. Restrictions apply.

V. An Explicit LRE Against NOFs

In this section, we will prove Theorem 3, and show that the function $\mathsf{BouExt} : \mathbb{F}_q^N \to \{0,1\}$, defined by

$$\mathsf{BouExt}(x_1, \ldots, x_N) = \mathsf{sign} \sin\left(\frac{2\pi \prod_{i\in[N]} x_i}{q}\right),$$

is an explicit LRE against number on forehead protocols.[10]

We follow the approach of the seminal work of Babai, Nisan, and Szegedy [17], who proved lower bounds for number-on-forehead (NOF) protocols. We begin by recalling the relevant definitions and lemma from [17].

Definition 10 (**τ-Components of a protocol** Π)**.** *Let Π be a multiparty protocol on N parties and τ be any transcript. The τ-*component *$X_{\Pi,\tau}$ is defined as the set of N-tuples $x \in \mathbb{F}^N$ such that on input x the protocol Π results in exactly τ being written on the board.*

Definition 11 (**Cylinders and cylinder-intersections**)**.** *A subset Y of N-tuples in called a* cylinder *in dimension $i \in [N]$ if membership in Y does not depend on coordinate i. A subset Y of N-tuples in called a* cylinder-intersection *if Y is an intersection of cylinders.*

Definition 12 (**Discrepancy**)**.** *Let $f : \mathbb{F}^N \to \{0,1\}$ be a boolean function. The discrepancy of f, denoted $\Gamma(f)$, is defined as*

$$\max_Y \left|\Pr[f(x) = 1 \ \& \ x \in Y] - \Pr[f(x) = 0 \ \& \ x \in Y]\right|,$$

where Y ranges over cylinder-intersections and x is chosen uniformly over $\mathbb{F}^N$.

Lemma 6 ([17])**.** *For any N-party NOF protocol Π and transcript τ, the τ-component $X_{\Pi,\tau}$ is a cylinder-intersection.*

Our extractor will use these notions in the context of *exponential sums*, which we recall next.

A. Exponential Sums

Let $\mathbb{F}_q$ be a prime field of cardinality q. Let $e_q(x) = \exp(2\pi i x/q)$. Building upon the exponential sums of [55], [56], Petridis and Shparlinski [57] defined the following class of exponential sums.[11]

Definition 13 (**Multilinear exponential sums** [57])**.** *For any K and N, the multilinear exponential sum SUM_N is defined as*

$$\mathsf{SUM}_N(K) = \max_{X,\eta,\phi} \left| \sum_{x_1\in X_1} \cdots \sum_{x_N \in X_N} \phi(x) e_q\left(\eta \prod_{i\in[N]} x_i\right)\right|,$$

where the maximum is over all subsets $X = X_1 \times \cdots \times X_N \subseteq \mathbb{F}_q^N$ with $|X_i| \le K$ for all $i \in [N]$, all $\eta \in \mathbb{F}_q^*$, and all cylinder-intersections $\phi : \mathbb{F}_q^N \to \{0,1\}$. Here and throughout, $\phi(x) = \phi_1(x) \cdot \cdots \cdot \phi_N(x)$, where ϕ_i does not depend on x_i for each $i \in [N]$.

Extending the results of Petridis and Shparlinski [57], Kerr and Macourt [58] obtained bounds on SUM_N for any $N \ll \log\log q$. We state the most suitable bounds of Petridis and Shparlinksi [57] and Kerr and Macourt [58] that we will use[12] for small constants (3,4,6). In our notation, their results translate to:

Theorem 6. *For some constant $C > 0$, we have:*

- *[57]:* $\mathsf{SUM}_3(K) \le Cq^{1/8}K^{43/16}$.
- *[57]:* $\mathsf{SUM}_4(K) \le Cq^{1/16}K^{61/16}$.
- *[58]:* $\mathsf{SUM}_6(K) \le Cq^{1/64}K^{3045/512+o(1)}$.

Furthermore, in the full version, we show that for all $N \ge 6$,

$$\mathsf{SUM}_N(q^{0.3}) \le \frac{q^{0.3N}}{q^{\Omega(1)/2^N}}.$$

Finally, we recall an observation from Bourgain [22, Remark 3.3] (see also [67]), which shows that it suffices to obtain upper bounds on multilinear exponential sums in order to obtain upper bounds on sums that involve BouExt.

Lemma 7 ([22])**.** *For any $X = X_1 \times \cdots \times X_N \subseteq \mathbb{F}_q^N$ and $\phi : \mathbb{F}_q^N \to \{0,1\}$, we have*

$$\left|\sum_{x_1\in X_1} \cdots \sum_{x_N\in X_N} \phi(x)(-1)^{\mathsf{BouExt}(x)}\right| \le$$

$$(C\log q) \max_{\eta\in\mathbb{F}_q^*} \left|\sum_{x_1\in X_1} \cdots \sum_{x_N\in X_N} \phi(x) e_q\left(\eta \prod_{i\in[N]} x_i\right)\right|,$$

where C is a universal constant and $x = (x_1, \ldots, x_N)$.

B. The Extractor

We are now ready to prove Theorem 3. A technicality is that while Theorem 3 was stated with inputs to each party being elements of $\{0,1\}^n$, we will on the other hand work with inputs to each party being elements of $\mathbb{F}_q$ for prime $q \approx 2^n$. We assume that we have access to such a prime.[13]

Theorem 7. *Fix a sufficiently large prime q. Let $n = \log q$, and $\mathsf{BouExt} : \mathbb{F}_q^N \to \{0,1\}$ be defined as above. Then for all $N \ge 6$, BouExt is an $(N-1, N, \mu)$-cylinder intersection extractor with error ϵ for all $(n, 0.3n)$-sources, where $\epsilon = 2^{-\mu}$ and $\mu = \Omega(n/2^N)$.*

Proof: We begin with the observation of Chor and Goldreich [21] that any source $\mathbf{X}_i$ distributed on $\mathbb{F}_q$ with

[10] We can use standard XOR lemmas to output multiple bits [52].

[11] Petridis and Shparlinski [57] use the term *weight functions* to denote cylinder-intersections. Their functions are more general, and can output any complex number with absolute value at most 1.

[12] This will help us achieve the best min-entropy rate later.

[13] We could potentially avoid this technicality by assuming Cramer's conjecture on primes, or by using part of the input to generate the prime at random (we only need average-case lower bounds). We do not delve into this issue here.

Authorized licensed use limited to: University of Texas at Austin. Downloaded on August 15,2021 at 16:31:35 UTC from IEEE Xplore. Restrictions apply.

min-entropy rate $\delta := 0.3$ is a convex combination of uniform sources on q^δ sized subsets $X_i \subseteq \mathbb{F}_q$. Therefore, we only need to focus on q^δ sized subsets. Fix any $X = X_1 \times \cdots \times X_N \subseteq \mathbb{F}_q^N$ such that $|X_i| = q^\delta$ for each $i \in [N]$.

Fix any number-on-forehead protocol Π with at most μ bits of communication. Let Γ be the set of transcripts that can be produced by executing Π on some $x = (x_1, \ldots, x_N) \in X$. Recall the notion of a τ-component from Definition 10: $X_{\Pi,\tau}$ denotes the set of $x \in X$ that result in transcript τ when protocol Π is executed on x. By Lemma 6, for each transcript τ, the τ-component $X_{\Pi,\tau}$ is a cylinder-intersection. Denote it by ϕ_τ.

To show that BouExt is a $(N-1, N, \mu)$-cylinder intersection extractor with error ϵ, it suffices to upper bound

$$|(\mathsf{BouExt}(\mathbf{X}), \Pi(\mathbf{X})) - (\mathbf{U}_1, \Pi(\mathbf{X}))|$$

where $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_N)$, and each $\mathbf{X}_i$ is uniformly distributed over the subset X_i of size q^δ. By definition of statistical distance, this is equal to

$$= \frac{1}{2} \sum_{b \in \{0,1\}} \sum_{\tau \in \Gamma} \Big| \Pr[\Pi(\mathbf{X}) = \tau \text{ and } \mathsf{BouExt}(\mathbf{X}) = b]$$

$$- \Pr[\Pi(\mathbf{X}) = \tau \text{ and } \mathbf{U}_1 = b] \Big|$$

$$= \frac{1}{2} \sum_{\tau \in \Gamma} \Pr[\Pi(\mathbf{X}) = \tau] \Big| \Pr[\mathsf{BouExt}(\mathbf{X}) = 1 \mid \Pi(\mathbf{X}) = \tau]$$

$$- \Pr[\mathsf{BouExt}(\mathbf{X}) = 0 \mid \Pi(\mathbf{X}) = \tau] \Big|$$

Next, note that the condition $\Pi(\mathbf{X}) = \tau$ is equivalent to $\mathbf{X}$ being in the τ-component, which in turn, is equivalent to $\mathbf{X}$ being in the corresponding cylinder-intersection ϕ_τ. Substituting this in, we get

$$= \frac{1}{2} \sum_{\tau \in \Gamma} \Pr[\Pi(\mathbf{X}) = \tau] \Big| \Pr[\mathsf{BouExt}(\mathbf{X}) = 1 \mid \phi_\tau(\mathbf{X}) = 1]$$

$$- \Pr[\mathsf{BouExt}(\mathbf{X}) = 0 \mid \phi_\tau(\mathbf{X}) = 1] \Big|$$

Computing this conditional probability, we obtain

$$= \frac{1}{2} \sum_{\tau \in \Gamma} \Pr[\Pi(\mathbf{X}) = \tau] \frac{1}{|X_{\Pi,\tau}|}$$

$$\cdot \left| \sum_{x_1 \in X_1} \cdots \sum_{x_N \in X_N} \phi_\tau(x) (-1)^{\mathsf{BouExt}(x)} \right|$$

where $x = (x_1, \ldots, x_N)$. Moreover, since $\mathbf{X}$ is uniform over X, we know $\Pr[\Pi(\mathbf{X}) = \tau]$ is equal to $|X_{\Pi,\tau}|/|X|$. Plugging this in, the above simplifies to

$$= \frac{1}{2} \sum_{\tau \in \Gamma} \frac{1}{|X|} \left| \sum_{x_1 \in X_1} \cdots \sum_{x_N \in X_N} \phi_\tau(x) (-1)^{\mathsf{BouExt}(x)} \right|$$

We can then use the connection to exponential sums (Lemma 7) to upper bound the above quantity by

$$\leq \frac{1}{2} \sum_{\tau \in \Gamma} \frac{(C \log q) \mathsf{SUM}_N(q^\delta)}{|X|}$$

Now, as there can be at most 2^μ transcripts in Γ, we obtain

$$\leq \frac{(C \log q) \cdot 2^\mu \cdot \mathsf{SUM}_N(q^\delta)}{|X|}.$$

Finally, since $|X| = q^{\delta N}$, we can use Theorem 6 to see that, for $N \geq 6$, we have

$$|(\mathsf{BouExt}(\mathbf{X}), \Pi(\mathbf{X})) - (\mathbf{U}_1, \Pi(\mathbf{X}))| \leq C \cdot 2^\mu \cdot q^{-\Omega(1)/2^N}.$$

Substituting in a sufficiently small $\mu = q^{\Omega(1)/2^N}$ proves our theorem. ■

VI. Acknowledgements

Eshan Chattopadhyay and Jesse Goodman are supported by NSF grant CCF-1849899. Vipul Goyal is supported in part by NSF grant 1916939, a gift from Ripple, a JP Morgan Faculty Fellowship, and a Cylab seed funding award. Ashutosh Kumar is supported by NSF grants CCF-1553605, 1619348, and 2007682; DARPA under Cooperative Agreement No: HR0011-20-2-0025; and US-Israel BSF grant 2012366. Xin Li is supported by NSF Award CCF-1617713 and NSF CAREER Award CCF-1845349. Raghu Meka is supported by NSF grants CCF-1553605 and 2007682. David Zuckerman is supported by NSF Grant CCF-1705028 and a Simons Investigator Award (#409864).

References

[1] A. C.-C. Yao, "Some complexity questions related to distributive computing (preliminary report)," in *STOC*. ACM, 1979, pp. 209–213.

[2] E. Kushilevitz and N. Nisan, *Communication Complexity*. Cambridge University Press, 2006.

[3] J. M. Phillips, E. Verbin, and Q. Zhang, "Lower bounds for number-in-hand multiparty communication complexity, made easy," in *SODA*. SIAM, 2012, pp. 486–501.

[4] M. Braverman, F. Ellen, R. Oshman, T. Pitassi, and V. Vaikuntanathan, "A tight bound for set disjointness in the message-passing model," in *FOCS*. IEEE, 2013, pp. 668–677.

[5] M. Braverman and R. Oshman, "On information complexity in the broadcast model," in *Principles of Distributed Computing*. ACM, 2015, pp. 355–364.

[6] A. K. Chandra, M. L. Furst, and R. J. Lipton, "Multi-party protocols," in *STOC*. ACM, 1983, pp. 94–99.

[7] N. Alon, Y. Matias, and M. Szegedy, "The space complexity of approximating the frequency moments," *Journal of Computer and System Sciences*, vol. 58, no. 1, pp. 137–147, 1999.

Authorized licensed use limited to: University of Texas at Austin. Downloaded on August 15,2021 at 16:31:35 UTC from IEEE Xplore. Restrictions apply.

[8] P. Beame, T. Pitassi, and N. Segerlind, "Lower bounds for Lovász–Schrijver systems and beyond follow from multiparty communication complexity," *SIAM Journal on Computing*, vol. 37, no. 3, pp. 845–869, 2007.

[9] E. Allender, "A note on the power of threshold circuits," in *FOCS*. IEEE Computer Society, 1989, pp. 580–584.

[10] A.-C. Yao, "On ACC and threshold circuits," in *FOCS*. IEEE, 1990, pp. 619–627.

[11] J. Håstad and M. Goldmann, "On the power of small-depth threshold circuits," *Computational Complexity*, vol. 1, no. 2, pp. 113–129, 1991.

[12] A. Razborov and A. Wigderson, "$n \log n$ lower bounds on the size of depth-3 threshold cicuits with AND gates at the bottom," *Information Processing Letters*, vol. 45, no. 6, pp. 303–307, 1993.

[13] R. Beigel and J. Tarui, "On ACC," *Computational Complexity*, vol. 4, no. 4, pp. 350–366, 1994.

[14] T. Lee and A. Shraibman, *Lower bounds in communication complexity*. Now Publishers Inc, 2009.

[15] A. Kumar, R. Meka, and A. Sahai, "Leakage-resilient secret sharing against colluding parties," in *FOCS*. IEEE, 2019, pp. 636–660.

[16] S. Arora and B. Barak, *Computational complexity: a modern approach*. Cambridge University Press, 2009.

[17] L. Babai, N. Nisan, and M. Szegedy, "Multiparty protocols, pseudorandom generators for logspace, and time-space tradeoffs," *Journal of Computer and System Sciences*, vol. 45, no. 2, pp. 204–232, 1992.

[18] J. Ford and A. Gál, "Hadamard tensors and lower bounds on multiparty communication complexity," *Computational Complexity*, vol. 22, no. 3, pp. 595–622, 2013.

[19] R. Shaltiel, "An introduction to randomness extractors," in *International colloquium on automata, languages, and programming*, 2011, pp. 21–41.

[20] S. P. Vadhan, "Pseudorandomness," *Foundations and Trends® in Theoretical Computer Science*, vol. 7, no. 1–3, pp. 1–336, 2012.

[21] B. Chor and O. Goldreich, "Unbiased bits from sources of weak randomness and probabilistic communication complexity," *SIAM J. Comput.*, vol. 17, no. 2, pp. 230–261, 1988.

[22] J. Bourgain, "More on the sum-product phenomenon in prime fields and its applications," *International Journal of Number Theory*, vol. 01, no. 01, pp. 1–32, 2005.

[23] B. Barak, R. Impagliazzo, and A. Wigderson, "Extracting randomness using few independent sources," *SIAM J. Comput.*, vol. 36, no. 4, pp. 1095–1118, Dec. 2006.

[24] X. Li, "Three-source extractors for polylogarithmic min-entropy," in *2015 FOCS*. IEEE, 2015, pp. 863–882.

[25] E. Chattopadhyay and D. Zuckerman, "Explicit two-source extractors and resilient functions," *Annals of Mathematics*, vol. 189, no. 3, pp. 653–705, 2019.

[26] E. Chattopadhyay and J. Goodman, "Explicit designs and extractors," *arXiv preprint arXiv:2007.07772*, 2020.

[27] G. R. Blakley, "Safeguarding cryptographic keys," in *AFIPS National Computer Conference (NCC '79)*. Los Alamitos, CA, USA: IEEE Computer Society, 1979, pp. 313–317.

[28] A. Shamir, "How to share a secret," *Communications of the ACM*, vol. 22, no. 11, pp. 612–613, 1979.

[29] M. Karchmer and A. Wigderson, "On span programs," in *Structure in Complexity Theory Conference, 1993., Proceedings of the Eighth Annual*. IEEE, 1993, pp. 102–111.

[30] A. Beimel, "Secret-sharing schemes: a survey," in *International Conference on Coding and Cryptology*. Springer Berlin Heidelberg, 2011, pp. 11–46.

[31] I. Komargodski, M. Naor, and E. Yogev, "Secret-sharing for NP," in *International Conference on the Theory and Application of Cryptology and Information Security*, 2014, pp. 254–273.

[32] V. Goyal and A. Kumar, "Non-malleable secret sharing," in *STOC*. ACM, 2018, pp. 685–698.

[33] M. Santha and U. V. Vazirani, "Generating quasi-random sequences from semi-random sources," *Journal of Computer and System Sciences*, vol. 33, pp. 75–87, 1986.

[34] U. V. Vazirani, "Strong communication complexity or generating quasi-random sequences from two communicating semi-random sources," *Combinatorica*, vol. 7, pp. 375–392, 1987.

[35] D. Aggarwal, M. Obremski, J. Ribeiro, L. Siniscalchi, and I. Visconti, "How to extract useful randomness from unreliable sources," in *Eurocrypt*, 2020, pp. 343–372.

[36] E. Chattopadhyay, J. Goodman, V. Goyal, and X. Li, "Extractors for adversarial sources via extremal hypergraphs," in *STOC*, 2020, pp. 1184–1197.

[37] M. Ball, O. Goldreich, and T. Malkin, "Randomness extraction from somewhat dependent sources," 2020.

[38] Y. T. Kalai and L. Reyzin, "A survey of leakage-resilient cryptography," *IACR Cryptology ePrint Archive*, vol. 2019, p. 302, 2019. [Online]. Available: https://eprint.iacr.org/2019/302

[39] S. Dziembowski and K. Pietrzak, "Intrusion-resilient secret sharing," in *FOCS*. IEEE, 2007, pp. 227–237.

[40] F. Davì, S. Dziembowski, and D. Venturi, "Leakage-resilient storage," in *International Conference on Security and Cryptography for Networks*, 2010, pp. 121–137.

[41] V. Goyal and A. Kumar, "Non-malleable secret sharing for general access structures," in *CRYPTO*, 2018, pp. 501–530.

Authorized licensed use limited to: University of Texas at Austin. Downloaded on August 15,2021 at 16:31:35 UTC from IEEE Xplore. Restrictions apply.

[42] F. Benhamouda, A. Degwekar, Y. Ishai, and T. Rabin, "On the local leakage resilience of linear secret sharing schemes," in *CRYPTO*, 2018, pp. 531–561.

[43] D. Aggarwal, I. Damgard, J. B. Nielsen, M. Obremski, E. Purwanto, J. Ribeiro, and M. Simkin, "Stronger leakage-resilient and non-malleable secret sharing schemes for general access structures," in *CRYPTO*, 2019, pp. 510–539.

[44] A. Srinivasan and P. N. Vasudevan, "Leakage resilient secret sharing and applications," in *CRYPTO*, 2019, pp. 480–509.

[45] F. Lin, M. Cheraghchi, V. Guruswami, R. Safavi-Naini, and H. Wang, "Leakage-resilient non-malleable secret sharing in non-compartmentalized models," *CoRR*, vol. abs/1902.06195, 2019.

[46] J. B. Nielsen and M. Simkin, "Lower bounds for leakage-resilient secret sharing," in *Eurocrypt*, 2020, pp. 556–577.

[47] M. Göös, S. Lovett, R. Meka, T. Watson, and D. Zuckerman, "Rectangles are nonnegative juntas," *SIAM Journal on Computing*, vol. 45, no. 5, pp. 1835–1869, 2016.

[48] V. Goyal, Y. Ishai, H. K. Maji, A. Sahai, and A. A. Sherstov, "Bounded-communication leakage resilience via parity-resilient circuits," in *FOCS*. IEEE, 2016, pp. 1–10.

[49] E. Chattopadhyay, J. Goodman, V. Goyal, and X. Li, "Leakage-resilient extractors and secret-sharing against bounded collusion protocols." in *Electronic Colloquium on Computational Complexity (ECCC)*, vol. 27, 2020, p. 60.

[50] A. Kumar, R. Meka, and D. Zuckerman, "Bounded collusion protocols, cylinder-intersection extractors and leakage-resilient secret sharing." in *Electronic Colloquium on Computational Complexity (ECCC)*, vol. 27, 2020, p. 55.

[51] V. V. Podolskii and A. A. Sherstov, "Inner product and set disjointness: Beyond logarithmically many parties," *arXiv preprint arXiv:1711.10661*, 2017.

[52] A. Rao, "An exposition of Bourgain's 2-source extractor," *Electronic Colloquium on Computational Complexity (ECCC)*, vol. 14, no. 034, 2007.

[53] A. Ben-Aroya, G. Cohen, D. Doron, and A. Ta-Shma, "Two-source condensers with low error and small entropy gap via entropy-resilient functions," in *APPROX/RANDOM 2019*, 2019.

[54] J. Kamp, A. Rao, S. Vadhan, and D. Zuckerman, "Deterministic extractors for small-space sources," in *STOC*. ACM, 2006, pp. 691–700.

[55] J. Bourgain, A. A. Glibichuk, and S. V. Konyagin, "Estimates for the number of sums and products and for exponential sums in fields of prime order," *Journal of the London Mathematical Society*, vol. 73, pp. 380–398, 4 2006.

[56] J. Bourgain, "Multilinear exponential sums in prime fields under optimal entropy condition on the sources," *Geometric and Functional Analysis*, vol. 18, no. 5, pp. 1477–1502, 2009.

[57] G. Petridis and I. E. Shparlinski, "Bounds of trilinear and quadrilinear exponential sums," *Journal d'Analyse Mathématique*, vol. 138, no. 2, pp. 613–641, 2019.

[58] B. Kerr and S. Macourt, "Multilinear exponential sums with a general class of weights," *arXiv preprint arXiv:1901.00975*, 2019.

[59] M. L. Fredman, J. Komlós, and E. Szemerédi, "Storing a sparse table with $O(1)$ worst case access time," *Journal of the ACM (JACM)*, vol. 31, no. 3, pp. 538–544, 1984.

[60] M. L. Fredman and J. Komlós, "On the size of separating systems and families of perfect hash functions," *SIAM Journal on Algebraic Discrete Methods*, vol. 5, no. 1, pp. 61–68, 1984.

[61] G. R. Blakley and C. Meadows, "Security of ramp schemes," in *Crypto*, 1984, pp. 242–268.

[62] K. Kurosawa, K. Okada, K. Sakano, W. Ogata, and S. Tsujii, "Nonperfect secret sharing schemes and matroids," in *Eurocrypt*, 1993, pp. 126–141.

[63] N. Alon, M. Dietzfelbinger, P. B. Miltersen, E. Petrank, and G. Tardos, "Linear hash functions," *Journal of the ACM (JACM)*, vol. 46, no. 5, pp. 667–683, 1999.

[64] R. Meka, O. Reingold, and Y. Zhou, "Deterministic coupon collection and better strong dispersers," in *APPROX/RANDOM 2014*, 2014.

[65] S. Blackburn, "Combinatorics and threshold cryptography," *Research Notes in Mathematics*, vol. 403, pp. 44–70, 1999.

[66] Y. Desmedt, "Some recent research aspects of threshold cryptography," in *Information Security*, E. Okamoto, G. Davida, and M. Mambo, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 1998, pp. 158–173.

[67] N. Hegyvári and F. Hennecart, "Explicit constructions of extractors and expanders," *Acta Arithmetica*, vol. 140, no. 3, pp. 233–249, 2009.

Authorized licensed use limited to: University of Texas at Austin. Downloaded on August 15,2021 at 16:31:35 UTC from IEEE Xplore. Restrictions apply.