Correlated-Source Extractors and Cryptography with Correlated-Random Tapes*

Vipul Goyal and Yifan Song(✉)

Carnegie Mellon University, Pittsburgh, USA
{vipul,yifans2}@cmu.edu

Abstract. In this paper, we consider the setting where a party uses correlated random tapes across multiple executions of a cryptographic algorithm. We ask if the security properties could still be preserved in such a setting. As examples, we introduce the notion of *correlated-tape zero knowledge*, and, *correlated-tape multi-party computation*, where, the zero-knowledge property, and, the ideal/real model security must still be preserved even if a party uses correlated random tapes in multiple executions.

Our constructions are based on a new type of randomness extractor which we call *correlated-source extractors*. Correlated-source extractors can be seen as a dual of non-malleable extractors, and, allow an adversary to choose several tampering functions which are applied to the randomness source. Correlated-source extractors guarantee that even given the output of the extractor on the tampered sources, the output on the original source is still uniformly random. Given (seeded) correlated-source extractors, and, *resettably-secure* computation protocols, we show how to directly get a positive result for both correlated-tape zero-knowledge and correlated-tape multi-party computation in the CRS model. This is tight considering the known impossibility results on cryptography with imperfect randomness.

Our main technical contribution is an explicit construction of a correlated-source extractor where the length of the seed is independent of the number of tamperings. Additionally, we also provide a (non-explicit) existential result for correlated source extractors with almost optimal parameters.

1 Introduction

Randomness is known to be crucial for cryptography. It is known that several basic tasks in cryptography become impossible in the absence of randomness [GO94,DOPS04]. Given this, a natural and well motivated direction is to develop an understanding of the *extent* to which randomness is necessary. Towards that end, we study the following natural question.

* Research supported in part by a grant from Northrop Grumman, a gift from DOS Networks, and, a Cylab seed funding award.

Suppose that a party uses correlated random tapes in multiple executions of a cryptographic algorithm. Can the security still be preserved? As a concrete example, suppose that the prover uses correlated random tapes in multiple executions with an adversarial verifier. Can the zero-knowledge property still be preserved? What about encrypting multiple times (under a randomized encryption scheme) using correlated random tapes? The above question can be motivated by, e.g., a scenario where a party has a defective random number generator which outputs correlated tapes under multiple invocations (even though each individual tape may have high min-entropy or even be close to uniform).

The well-known line of research on resettable security can be seen as a *special case* of our general problem. In resettable zero-knowledge [CGGM00], the prover uses the *same* random tape across multiple executions. By varying the set of parties whose random tape is fixed, one can get various variants such as resettably-sound zero-knowledge [BGGL01], simultanous resettable zero-knowledge [DGS09], and, resettably secure computation [GS09,GM11].

In this work, we initiate a systematic study of the above question. The central object of our study will be a new notion of randomness extractors which we call *correlated-source extractors*. Very informally, a (seeded) correlated source extractor $\texttt{csExt}$ on input a seed s, and a source X produces an output $\texttt{csExt}(X, s)$ which is guaranteed to be close to uniform even given $\texttt{csExt}(X_1, s), ..., \texttt{csExt}(X_t, s)$ where for all i, $X_i \neq X$ and X_i could be arbitrarily correlated with X. One could also view X_i as a result of tampering the original source X. Correlated-source extractors can be seen as a dual of non-malleable extractors [DW09], where the adversary is allowed to tamper the seed instead of the source. Non-malleable extractors have played an important role in cryptography and complexity in problems such as privacy amplification [DW09], designing two-source extractors [CZ16], and in designing non-malleable codes [CG14a,CGL16]. Correlated-source extractors are also closely related to two-source non-malleable extractors [CG14a,CGL16].

1.1 Our Results.

We introduce the notion of correlated-tape zero-knowledge. We model correlations among the different random tapes by consider an adversary which may have limited control over the random tape of the honest parties. In correlated-tape zero-knowledge, the adversary is able to specify t tampering functions $f_1, f_2, ..., f_t$ at the beginning of the protocol such that in the $i-$th execution, the prover uses $f_i(X)$ as its random tape (where X is uniformly random and can be viewed as the original random tape). Other notions like correlated-tape secure multi-party computation (MPC) and correlated-tape secure encryption schemes could also be defined analogously. We also define the the main object of our study: correlated-source extractors. Specifically,

Definition 1 (Seeded Correlated-Source Extractor). *A function* $\texttt{csExt} : \{0,1\}^* \times \{0,1\}^d \to \{0,1\}^m$ *is a seeded correlated-source extractor if the following holds: There exists a polynomial* $k(\cdot,\cdot,\cdot)$ *and a negligible function* $\epsilon(\cdot)$*, such that*

for any polynomial $t(\cdot)$, $t = t(d)$ arbitrary functions $\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_t$, whose output has the same length as the input, with no fixed points, and, a source X with min-entropy $k(t, m, d)$,

$$|\mathtt{csExt}(X, U_d) \circ \{\mathtt{csExt}(\mathcal{A}_i(X), U_d)\}_{i=1}^{t} \circ U_d - U_m \circ \{\mathtt{csExt}(\mathcal{A}_i(X), U_d)\}_{i=1}^{t} \circ U_d| < \epsilon(d)$$

where U_m and U_d are uniform strings of length m and d respectively.

Jumping ahead, in our cryptographic applications, the seed will serve as the CRS while the source X will be the local random tape generated by the party. We require the output length, and, the source length (and hence the min-entropy) to be (unbounded) polynomial in the length of the seed. Thus, fixing the CRS (i.e., the seed) doesn't necessarily fix the number of executions (represented by t). One could also define a weaker notion of correlated-source extractors where the seed fixes a bound on the number of executions. Specifically,

Definition 2 (Weak t-Correlated-Source Extractor). *A function* $\mathtt{wcsExt} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *is a weak t-correlated-source extractor for min-entropy k and error ϵ if the following holds: If X is a source in $\{0,1\}^n$ with min-entropy k, and, $\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_t$ are arbitrary functions whose output has the same length as the input, with no fixed points, then*

$$|\mathtt{wcsExt}(X, U_d) \circ \{\mathtt{wcsExt}(\mathcal{A}_i(X), U_d)\}_{i=1}^{t} \circ U_d - U_m \circ \{\mathtt{wcsExt}(\mathcal{A}_i(X), U_d)\}_{i=1}^{t} \circ U_d| < \epsilon$$

where U_m and U_d are uniform strings of length m and d respectively.

Our first main result is a construction of a correlated-source extractor:

Theorem 1. *There exists an explicit correlated-source extractor* $\mathtt{csExt}$ *with*

$$k(t, m, d) = \Theta(t^3 d + t^2 m)$$
$$\epsilon(d) = \Theta(2^{-\sqrt{d}})$$

where m is the length of the output.

Note that it is necessary for the entropy of the source to grow with the number of executions t if the tampered sources may be arbitrarily correlated with the original source. This is because the entropy of the original source may reduce given the output of the extractor on a tampered source. In section 5.4 we generalize the entropy requirements on the sources. In particular, we define what we call *closed-set correlated sources* and show correlated set extractors for such sources. For closed-set correlated source, the entropy of each individual source does not necessarily grow with the number of invocations t. Hence, this would allow us to get constructions where neither the seed length, nor the source length or its entropy grows with the number of invocations.

Going to Correlated-Tape Zero-Knowledge and Secure Computation. We note that correlated-source extractors can only allow us to handle the random tapes where *each random tape differs from every other one*. We relax this constraint by relying on techniques from resettable zero-knowledge [CGGM00] [BGGL01], and, resettably secure computation [GS09]. In resettable zero-knowledge, the prover uses the *same* random tape across multiple executions. In our setting, the random tape could either be the same or arbitrarily correlated with another random tape. Very informally, relying on resettable security would allow us to achieve security in case the random tape is the same as another one, and, relying on correlated source extractor would guarantee security in case the random tape is different from every other tape but maybe arbitrarily correlated.

This allows us to obtain positive results for correlated-tape zero-knowledge and multi-party computation in the CRS model where the only (necessary) requirement on the random tape would be sufficient min-entropy; otherwise *each random tape could be arbitrarily correlated to or even the same as other random tapes*. The seed required for the correlated source extractor would be a part of the CRS. Each party in the protocol would first apply correlated-source extractor on its (potentially tampered) random tape, and, use the resulting string as the random tape to execute a resettable secure MPC (or zero-knowledge) protocol. We note that correlated-tape zero-knowledge and similar primitives such as correlated-tape encryption are impossible to obtain in the plain model. This holds even for a single execution and follows from the known impossibility results on cryptography with imperfect and tamperable randomness [DOPS04] [ACM+14] (see section 4 for more details). We also give stronger impossibility results in Section 4.

Weak Correlated-Source Extractors. Note that basic positive result for *weak* correlated-source extractor follows from the construction of two-source non-malleable extractors in [CGL16]. In fact, two-source non-malleable extractors allow the adversary to also tamper the second source (the random seed) and only requires the second source (the random seed) to have enough min-entropy. However, directly using two-source non-malleable extractor or similar techniques cannot give a positive result for correlated source extractor. This is because two-source non-malleable extractor will require the seed length to be either as long as that of the source or linear in t. Note that this would give a positive result only for *bounded* correlated-tape zero-knowledge and secure computation where the number of executions must be fixed before choosing the CRS. We note that obtaining a construction where the seed length is independent of the number of tamperings has been a challenging problem in this line of research. *In particular, obtaining such an explicit construction for non-malleable extractors still remains an open problem.* An existential result has however been shown very recently [BACD+18].

Correlated-Source Extractors with Almost Optimal Parameters. Next, we turn our attention to the following natural question: what is the optimal entropy and

source length that a correlated-source extractor requires? Towards that end, we prove the following existential result:

Theorem 2 (Existence of Correlated-Source Extractor). *There exists a correlated-source extractor* `csExt` *as long as*

$$k(t, m, d) = \Theta(tm + d) \tag{1}$$
$$\epsilon(d) = \Theta(2^{-d}) \tag{2}$$

where m is the length of the output.

For an overview of our techniques, please refer to Section 2.

Related Works. Designing randomness extractors has been a rich line of works. Most relevant to our work are non-malleable extractors [DW09], and, two source non-malleable extractors [CG14a,CGL16]. After the initial constructions, a number of works have focused on improving the entropy requirements and the seed length [Li12a,Li12b,DLWZ14,Li15,CL16,Coh16b,Coh16c,Coh16a,Li16,Li17]. However, all known explicit constructions of non-malleable and two-source non-malleable extractors require the length of the seed to grow with the number of tamperings t. Two-source non-malleable extractors from [CGL16] were used crucially in a recent breakthrough on constructing two-source extractors [CZ16].

A number of works have studied simulating randomized algorithms using weak sources with small min-entropy [VV85,CG88,Zuc96,SSZ95,ACRT97]. Andreev et al. [ACRT97] gave a simulation of any BPP algorithm with an $(n, n^{O(1)})$-source. In contrast, we focus on multiple executions with correlated random tapes and have weaker entropy requirements.

A rich line of works have studied resettable secure protocols [CGGM00], [BGGL01,DGS09,GM11,BP13,CPS16,COPV13,COP$^+$14], where a party may use the same random tape in multiple executions. The class of correlations we handle is more general. Kalai et al. [KLRZ08] introduced network extractor protocols where there are a number of parties each having independent (but imperfect) random tapes. Their result required a strong variant of the Decisional Diffie-Hellman Assumption, and, a polylogarithmic number of parties.

Goldreich and Oren [GO94] showed that constructing zero-knowledge arguments where the prover is deterministic is impossible. Dodis et al. [DOPS04] showed that a number of basic cryptographic primitives like encryption and zero-knowledge are impossible with imperfect randomness. Austrin et al. [ACM$^+$14] similarly showed a number of impossibility results (including for zero-knowledge) in the setting of tampering randomness. These results focus on the plain model and in the setting of a single execution. Moving to the CRS model allows us to bypass these negative results. A line of research also explores cryptography with related keys and related inputs (see [ABP15] and the references therein), typically for a special class of tampering functions (such as affine functions).

2 Technical Overview

Explicit Construction with Fixed Seed Length. In this section, we will give a high level idea of our construction of correlated-source extractors. We use X for the original source and X_i for the tampered source. We use Y for the random seed.

Why existing techniques fail. All the construction ideas related to two-source non-malleable extractors (which imply the existence of *weak* correlated-source extractors) somehow separate the original seed into several independent and uniformly random "slices". A general framework to constructing non-malleable extractors (and two-source non-malleable extractors) followed by several works is based on alternating extraction [DW09] and generating an advice (which is unique w.h.p. across all the tampered executions). A critical step in such constructions is to view the original seed as a second source. In the beginning, a slice of the seed is used to extract from the source. Next, the result is used as a seed to extract from the original seed. Next, the result is again used as a seed to extract from the source, and so on. This technique relies on the length of the original seed to be long enough. In particular, during the analysis, each tampering would "fix" a part of the random seed. This means that, the effective entropy of the seed reduces as the number of tampered executions increase. By using alternating extraction where the seed plays the role of one of the source, it seems that the seed length must be linear in t.

Overview of the construction. Our idea is to generate two (or multiple) independent sources from the original source itself. A straightforward idea is to generate (X^1, X^2) from X, such that the distribution of $\{X^1, X_1^1, ..., X_t^1\}$ is independent of $\{X^2, X_1^2, ..., X_t^2\}$ (here X_i is the $i-$th tampering source and (X_i^1, X_i^2) are generated from X_i). Then we may discard the original seed and use a two-source non-malleable extractor on X^1 and X^2. However, we don't know how to prove the joint distributions of two sets are independent. Our starting idea would be to use the given random seed in obtaining such a "decomposition" of the original source. We use one part of the seed Y_1 to generate $X^1 = \mathtt{Ext}(X, Y_1)$ and another part of the seed Y_2 to generate $X^2 = \mathtt{Ext}(X, Y_2)$. By assuming the source X has enough entropy, we can guarantee that, given $\{X^2, X_1^2, ..., X_t^2\}$, X^1 is uniformly random. Note that the joint distribution of $\{X^1, X_1^1, ..., X_t^1\}$ may be dependent on that of $\{X^2, X_1^2, ..., X_t^2\}$, while two-source non-malleable extractors require the adversary tampers both sources separately. Thus, it is not sufficient to use two-source non-malleable extractors.

We first generate an advice $\mathtt{adv}$ from the source X (and $\mathtt{adv}_i$ from X_i) such that it is unique w.h.p. across all the tampered executions. Let ℓ denote the length of the advice. Then, instead of just generating (X^1, X^2) from X, we generate 2ℓ sources $(X^1, X^2, ..., X^{2\ell})$ from X such that, for every $i \in \{1, ..., 2\ell\}$, X^i is uniformly random given $\{X^j, X_1^j, ..., X_t^j\}_{j \neq i}$.

Let $\mathtt{adv}^i$ denote the $i-$th bit of $\mathtt{adv}$. Each bit $\mathtt{adv}^i$ corresponds to a pair of sources (X^{2i-1}, X^{2i}). The extractor first uses one piece of the original seed as the seed and extracts randomness from one source of the first pair (X^1, X^2) decided by the value of $\mathtt{adv}^1$, then uses the result as the seed and extracts randomness

from one source of the second pair (X^3, X^4) and so on. Specifically, in the $i-$th iteration, we choose $X^{2i-1+\mathtt{adv}^i}$. This process can be described by a function $F = F(\mathtt{adv}^i, X^{2i-1}, X^{2i}, Y, Z^{i-1})$, where Z^{i-1} is the result in the last iteration and initially, Z^0 is one piece of the original seed.

Note that, in the case that $\mathtt{adv}$ is different from all tampered $\mathtt{adv}_1, ..., \mathtt{adv}_t$, for all $j \in \{1, ..., t\}$, there exists at least one iteration (denoted by the $i-$th iteration) such that the $i-$th bits of $\mathtt{adv}$ and $\mathtt{adv}_j$ are different. We note that $X^{2i-1+\mathtt{adv}^i}$ is in fact independent of $X_j^{2i-1+\mathtt{adv}_j^i}$. Thus, hopefully, we can break the correlation between X and X_j in this iteration, i.e., Z^i is independent of Z_j^i. We also need this independence to be preserved in all later iterations. Therefore, in the end, since $\mathtt{adv}$ is different from all tampered advice, Z^ℓ is independent of $Z_1^\ell, ..., Z_t^\ell$.

Now we are ready to state our construction overview in more detail. It can be divided into two steps.

Step 1: Generating advice $\mathtt{adv}$ and limited correlated parts $X^1, ..., X^{2\ell}$

In the beginning, we generate an advice $\mathtt{adv}$ for the source X such that, with high probability, $\mathtt{adv}$ is different from $\mathtt{adv}_1, ..., \mathtt{adv}_t$ (the advice of $X_1, X_2, ..., X_t$). This idea is not new and is widely used in the constructions of non-malleable extractors (e.g. in [Coh15,CGL16]).

Recall that ℓ is the length of $\mathtt{adv}$. We generate $X^1, X^2, ..., X^{2\ell}$ by using a fresh seed Y_1^i for each part X^i. Specifically, $X^i = \mathtt{Ext}(X, Y_1^i)$ (and $X_j^i = \mathtt{Ext}(X_j, Y_1^i)$). Note that the random seeds in all executions are the same.

These sources $X^1, X^2, ..., X^{2\ell}$ directly satisfy our requirement, i.e., for every $i \in \{1, ..., 2\ell\}$, X^i is uniformly random given $\{X^j, X_1^j, ..., X_t^j\}_{j \neq i}$. To see this, let $\mathcal{X}^i = \{X^i, X_1^i, ..., X_t^i\}$. In the case that X has enough min-entropy and $X^1, ..., X^{2\ell}$ are comparatively short, X still has enough min-entropy when fixing $\mathcal{X}^1, ..., \mathcal{X}^{i-1}, \mathcal{X}^{i+1}, ..., \mathcal{X}^{2\ell}$. Also, Y^i is a fresh piece from the seed. Thus, by the property of $\mathtt{Ext}$, X^i is uniformly random and independent of $\mathcal{X}^1, ..., \mathcal{X}^{i-1}, \mathcal{X}^{i+1}, ..., \mathcal{X}^{2\ell}$.

Step 2: Breaking correlation between sources by induction

Let $\mathtt{SAME}^i$ be the set of indices of sources whose advice is different from $\mathtt{adv}$ in at least one bit of the first $(i-1)$ bits but is the same in the $i-$th bit, and $\mathtt{DIFF}^i$ be the set of indices of sources whose advice is different from $\mathtt{adv}$ in the $i-$th bit. Then after the $(i-1)-$th iteration, Z^{i-1} should have been uniformly random and independent of $\{Z_j^{i-1}\}_{j \in \mathtt{SAME}^i}$. So for $j \in \mathtt{SAME}^i$, we want this independence to be preserved in the $i-$th iteration. And we want to further break the correlation with the $j-$th tampered result where $j \in \mathtt{DIFF}^i$ in the $i-$th iteration.

In the $(i-1)-$th iteration, we should have already achieved that

$$|Z^{i-1} \circ \{Z_j^{i-1}\}_{j \in \mathtt{SAME}^{i-1} \bigcup \mathtt{DIFF}^{i-1}} - U_z \circ \{Z_j^{i-1}\}_{j \in \mathtt{SAME}^{i-1} \bigcup \mathtt{DIFF}^{i-1}}| < \epsilon$$

where z is the length of Z^{i-1}. A critical fact is that the above inequality still holds even given $\mathcal{X}^{2i-1}$ and $\mathcal{X}^{2i}$. Because what we need to prove the above property is that X^j is uniformly random when given $\{\mathcal{X}^k\}_{k \neq j}$ for every $j \in \{1, ..., 2i-2\}$. Fixing $\mathcal{X}^{2i-1}$ and $\mathcal{X}^{2i}$ does not break this condition by the way how we generated $X^1, ..., X^{2i-2}$.

For a series of correlated sources $X, X_1, ..., X_t$, there are two ways to break the correlation. If the random seeds are independent and uniformly random for different sources, then the output is uniformly random and independent of others in the case that X has enough min-entropy. If X given $X_1, X_2, ..., X_t$ still has enough min-entropy, then even if the seeds are not independent, the output of extractor is still uniformly random and independent of others. This idea is also used in the recent construction of non-malleable extractors (e.g. in [CL16]).

We note that, for the executions whose indices $j \in \mathtt{SAME}^i$, they will use $X_j^{2i-1+\mathtt{adv}_j^i}$ where $2i-1+\mathtt{adv}_j^i = 2i-1+\mathtt{adv}^i$. It means that $\{X_j^{2i-1+\mathtt{adv}_j^i}\}_{j\in\mathtt{SAME}^i}$ may be highly correlated with $X^{2i-1+\mathtt{adv}^i}$. However, since $\mathtt{SAME}^i \subseteq \mathtt{SAME}^{i-1} \bigcup \mathtt{DIFF}^{i-1}$, Z^{i-1} is uniformly random and independent of $\{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i}$. If we use Z^{i-1} as the seed to do extraction on $X^{2i-1+\mathtt{adv}^i}$, the result is independent of those of executions whose indices $j \in \mathtt{SAME}^i$. Specifically,

Sub-Step 2.1: Breaking correlation with sources in $\mathtt{SAME}^i$

Let $\mathtt{Ext}$ be a strong seeded extractor. Compute $W^i = \mathtt{Ext}(X^{2i-1+\mathtt{adv}^i}, Z^{i-1})$ (and W_j^i for the result in the $j-$th tampered execution). Then, given $\{W_j^i\}_{j\in\mathtt{SAME}^i}$, W^i is uniformly random.

Now, we want to further break the correlation with the $j-$th tampered result where $j \in \mathtt{DIFF}^i$. Currently, W^i may be correlated with $\{W_j^i\}_{j\in\mathtt{DIFF}^i}$. However, We note that we can fix $\mathcal{X}^{2i-1+(1-\mathtt{adv}^i)}$ in Sub-Step 2.1 without breaking the property of the result. Since, for $j \in \mathtt{DIFF}^i$, W_j^i only depends on $X_j^{2i-1+\mathtt{adv}_j^i}$ and Z_j^{i-1}, and $X_j^{2i-1+\mathtt{adv}_j^i}$ has already been fixed (because $\mathtt{adv}_j^i = 1-\mathtt{adv}^i$), W^i will have enough min-entropy even fixing $\{Z_j^{i-1}\}_{j\in\mathtt{DIFF}^i}$ if we choose the length of W^i to be much longer than that of Z^i. It also means that W^i given $\{W_j^i\}_{j\in\mathtt{DIFF}^i}$ still has enough min-entropy. Therefore, we can simply use a fresh piece of the original seed as the seed to do extraction on W^i. The result will be independent of those of executions whose indices are in $\mathtt{SAME}^i \bigcup \mathtt{DIFF}^i$. Specifically,

Sub-Step 2.2: Breaking correlation with sources in $\mathtt{DIFF}^i$

We use a fresh piece Y_2^i from the original random seed Y. Compute $Z^i = \mathtt{Ext}(W^i, Y_2^i)$ as the output of the $i-$th iteration.

3 Preliminaries

We use capital letters to denote random variables. We use U_r to denote the uniform distribution over $\{0,1\}^r$. For random variable X, we use $x \sim X$ to denote that x is sampled from the distribution of X.

3.1 Statistical Distance, Convex Combination of Distributions and Probability Lemma

Definition 3 (Statistical Distance). *Let D_1 and D_2 be two distributions on a set S. The statistical distance between D_1 and D_2 is defined to be:*

$$|D_1 - D_2| = \max_{T \subseteq S} |D_1(T) - D_2(T)| = \frac{1}{2} \sum_{s \in S} |\Pr[D_1 = s] - \Pr[D_2 = s]|$$

D_1 is $\epsilon-$close to D_2 if $|D_1 - D_2| \leq \epsilon$.

Definition 4 (Convex Combination). *A distribution D on a set S is a convex combination of distributions $D_1, D_2, ..., D_\ell$ on S if there exists non-negative constants (called weights) $w_1, w_2, ..., w_\ell$ with $\sum_{i=1}^{l} w_i = 1$ such that $\Pr[D = s] = \sum_{i=1}^{\ell} w_i \Pr[D_i = s]$ for all $s \in S$. We use the notation $D = \sum_{i=1}^{\ell} w_i D_i$ to denote the fact that D is a convex combination of the distributions $D_1, ..., D_\ell$ with weights $w_1, ..., w_\ell$.*

3.2 Min-entropy and Flat Distribution

The min-entropy of a source X is defined as

$$H_\infty(X) = -\log(\max_{x \in \mathtt{support}(X)} (1/\Pr[X = x]))$$

A distribution D is a flat distribution (source) if it is uniformly random over a set S. An (n, k)-source is a distribution over $\{0,1\}^n$ with min-entropy at least k. It is well known that any (n, k)-source is a convex combination of flat sources supported on sets of size 2^k.

3.3 Seeded Extractors, Non-malleable Extractors, Two-source Non-malleable Extractors and Previous Construction

Definition 5 (Strong seeded Extractor). *A function $\mathtt{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is called a strong seeded extractor for min-entropy k and error ϵ if for any $(n, k)-$source X and an independent uniformly random string U_d, we have*

$$|\mathtt{Ext}(X, U_d) \circ U_d - U_m \circ U_d| < \epsilon,$$

where U_m is independent of U_d and m is the output length of $\mathtt{Ext}$.

The following definition of $t-$non-malleable extractors is from [CRS14], which generalizes the definition in [DW09].

Definition 6 (Non-malleable Extractor). *A function $\mathtt{snmExt} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a seeded $t-$non-malleable extractor for min-entropy k and error ϵ if the following holds: If X is a source on $\{0,1\}^n$ with min-entropy k and $\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_t$ are arbitrary (tampering) functions defined on $\{0,1\}^n \to \{0,1\}^n$ with no fixed points, then*

$$|\mathtt{snmExt}(X, U_d) \circ \{\mathtt{snmExt}(X, \mathcal{A}_i(U_d))\}_{i=1}^{t} - U_m \circ \{\mathtt{snmExt}(X, \mathcal{A}_i(U_d))\}_{i=1}^{t}| < \epsilon,$$

where U_m is independent of U_d and X.

The following definition of two-source non-malleable extractors is from [CGL16], which generalizes the definition in [CG14b].

Definition 7 (Two-source Non-malleable Extractor). *A function* $\mathtt{nmExt} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$ *is a two-source* $t-$*non-malleable extractor for min-entropy* k *and error* ϵ *if the following holds: If* X, Y *are independent sources on* $\{0,1\}^n$ *with min-entropy* k *and* $\mathcal{A}_1 = (f_1, g_1), \mathcal{A}_2 = (f_2, g_2), ..., \mathcal{A}_t = (f_t, g_t)$ *are arbitrary 2-split-state tampering functions where* f_i, g_i *are defined on* $\{0,1\}^n \to \{0,1\}^n$ *such that for any* i*, at least one of* f_i, g_i *has no fixed points, then*

$$|\mathtt{nmExt}(X,Y) \circ \{\mathtt{nmExt}(f_i(X), g_i(X))\}_{i=1}^t - U_m \circ \{\mathtt{nmExt}(f_i(X), g_i(X))\}_{i=1}^t| < \epsilon$$

Theorem 3 ([GUV09]). *For any constant* $\alpha > 0$*, and all integers* $n, k > 0$ *there exists a polynomial time computable strong seeded extractor* $\mathtt{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *with* $d = O(\log n + \log(1/\epsilon))$ *and* $m = (1-\alpha)k$*.*

3.4 Conditional Min-entropy

Definition 8. *The average conditional min-entropy is defined as*

$$\tilde{H}_\infty(X|W) = \log\left(\mathbf{E}_{w\sim W}[\max_x \Pr[X = x|W = w]]\right) = -\log\left(\mathbf{E}_{w\sim W}[2^{-H_\infty(X|W=w)}]\right)$$

The following result on conditional min-entropy was proved in [MW97].

Lemma 1. *Let* X, Y *be random variables such that the random variable* Y *takes at most* l *values. Then*

$$\Pr_{y\sim Y}[H_\infty(X|Y=y) \geq H_\infty(X) - \log l - \log(\frac{1}{\epsilon})] > 1 - \epsilon$$

We recall some results on conditional min-entropy from [DRS04].

Lemma 2 ([DRS04]). *If a random variable* B *can take at most* 2^ℓ *values, then* $\tilde{H}_\infty(A|BC) \geq \tilde{H}_\infty(A|C) - \ell$*.*

Lemma 3 ([DRS04]). *For any* $\delta > 0$*, if* $\mathtt{Ext}$ *is a* $(k,\epsilon)-$*extractor then it is also a* $(k + \log(1/\delta), \epsilon + \delta)$ *average case extractor.*

4 Our Model

In this section, we introduce a new model of cryptographic protocol where a party may be involved in multiple executions with correlated random tapes. We first focus on zero-knowledge and later generalize to secure multi-party computation. This captures the setting where an honest party may have a defective random number generator G which may output highly correlated strings in different executions. In the worst case, the output of G after the first execution may fully depend on the output in the first execution. Then an adversary may use the messages it received in the first execution to get information about the random tape of the honest parties in the subsequent executions.

We will formalize the above setting by considering an experiment where an adversary is given limited control of the random tape of the honest party and can interact with the honest party in multiple sessions.

Correlated-Tape Zero-Knowledge. For every $\{(x_i, w_i)\}_{i=1}^t$ where $w_i \in R_L(x_i)$, the verifier V^* will sequentially interact with the actual prover P. V^* can specify t tampering functions $f_1, f_2, ..., f_t$. To overcome known impossibility results [DOPS04] as discussed later, we also assume the existence of a CRS. In the beginning, P has a private random tape X distributed uniformly at random. In the $j-$th execution, P uses $f_j(X)$ as its random tape. We use $\tau(P, V^*, \texttt{CRS}, \{(x_i, w_i)\}_{i=1}^t)$ to denote the transcripts of t consecutive executions and the total view of V^* where in the $j-$th execution, P takes $(x_j, w_j), \texttt{CRS}$ as input and uses $f_j(X)$ as random tape, and V^* takes $x_j, \texttt{CRS}$, all previous transcripts and its previous view as input.

Definition 9. *A pair of algorithms (P, V) is a correlated-tape zero-knowledge proof system for language L, if there exist polynomials $len(\cdot), k(\cdot, \cdot)$ such that for any polynomial $t(\cdot)$, the following conditions hold:*

- Completeness: *For every security parameter $\kappa, x \in L, w \in R_L(x)$,*

$$\Pr[< P(w, X), V > (x, U_{len(\kappa)}) = 1] = 1$$

  Here w is the private input for P and X is the private random tape of P. By $< P, V > (x, U_{len(\kappa)})$, we denote the output of V when P and V interact on the common input x and a common reference string distributed uniformly random over $\{0,1\}^{len(\kappa)}$.
- Soundness: *For every algorithm A and every $x \notin L$, there exists a negligible function $\mu(\cdot)$ such that for every security parameter κ,*

$$\Pr[< A, V > (x, U_{len(\kappa)}) = 1] < \mu(\kappa)$$

- Correlated-Tape Zero Knowledge: *There exists a simulator S, such that for any $t = t(\kappa)$ functions $f_1, f_2, ..., f_t$, whose output has the same length as input, such that $H_\infty(f_i(X)) \geq k(t(\kappa), \kappa)$, any V^* and (x_i, w_i) where $w_i \in R_L(x_i)$ and $i \in [t]$ the following two distributions are computationally indistinguishable:*

$$\{\texttt{CRS} \sim U_{len(\kappa)} : \tau(P, V^*, \texttt{CRS}, \{(x_i, w_i)\}_{i=1}^t)\}$$

$$S(\{x_i\}_{i=1}^t, V^*)$$

One could consider a variant of the above definition where there is no CRS. One could also define the complementary setting where the random number generator of the verifier (rather than the prover) is defective (and one would like to ensure soundness in the correlated tape setting).

Correlated-Tape Secure Multi-Party Computation. Now let us consider the case of a multi-party computation protocol. We use P_i to denote the $i-$th party and A to denote the adversary. Suppose there are n parties in total. We use F to represent the desired functionality. Let $T \subseteq [n]$ to be the set of indices of parties that are corrupted.

Our ideal model will be the same as that in a standard definition of MPC. Specifically,

Ideal Model. There is a trusted party which computes the desired functionality based on the inputs of all parties. An execution in the ideal model proceeds as follows:

- **Inputs** All parties (including the corrupted parties) will send their inputs to the trusted party. An honest party always sends its real input. A corrupted party may send modified value depending on the strategy of the adversary. We use x_i to denote the input sent by P_i.
- **Trusted Party Computes the Result** The trusted party will use the inputs from all parties to compute the desired functionality. Let $(y'_1, y'_2, ..., y'_n) = F(x_1, x_2, ..., x_n)$
- **Trusted Party Sends out the result** For $i = 1, 2, ..., n$, the trusted party asks A whether it wants to abort. If A does not abort, then the trusted party will send y'_i to P_i. Otherwise, for all $j \geq i$, P_j will receive nothing from the trusted party.
- **Outputs** An honest party P_i always outputs the response it received from the trusted party (it will output $\perp$ if it receives nothing) together with its input x_i. The adversary A outputs an arbitrary function of its entire view so far (including the views in the previous executions).

We use $\texttt{IDEAL}_{F,A}(\{x_i^j : i \notin T\}_{j=1}^t)$ to represents the outputs of t consecutive sequential executions in the ideal world with functionality F, adversary A and input x_i^j for honest party P_i in the $j-$th execution.

Real Model. In the real model, the adversary A is allowed to specify t tampering functions for each honest party. We use f_i^j to represent the $j-$th tampering function for P_i. Then there will be a trusted party generating a uniform string as CRS. In the beginning, each party has a private random tape distributed uniformly random. Denote X_i to be the initial random tape of P_i. In the $j-$th execution in the real model, an honest party P_i uses $f_i^j(X_i)$ as its random tape. The outputs of a protocol π in the real model include the inputs and outputs of all honest parties together with the full view of the adversary so far (including the view in the previous executions). We use $\texttt{REAL}_{F,A}^{\pi}(\{x_i^j : i \notin T\}_{j=1}^t, \{X_i^j : i \notin T\}_{j=1}^t, \texttt{CRS})$ to represents the outputs of π in t consecutive sequential executions in the real world for functionality F, adversary A, common reference string CRS and input x_i^j for honest party P_i with random tape X_i^j in the $j-$th execution.

Definition 10. *A protocol π is a secure correlated-tape multi-party computation protocol for functionality F of n parties, if there exist polynomials $len(\cdot), k(\cdot,\cdot)$ such that for any polynomial $t(\cdot)$, adversary A which corrupts $(n-\ell)$ parties with the set of indices $T \subset [n]$ and security parameter κ, there exists an ideal attacker A' such that, for all inputs $\{x_i^j : i \notin T\}_{j=1}^{t(\kappa)}$ and functions $\{f_i^j : i \notin T\}_{j=1}^{t(\kappa)}$,whose output has the same length as input, such that $H_\infty(f_i^j(X_i)) \geq k(t(\kappa), \kappa)$, the following two distributions are computationally indistinguishable:*

$$\{\texttt{CRS} \sim U_{len(\kappa)} : \texttt{REAL}_{F,A}(\{x_i^j : i \notin T\}_{j=1}^{t(\kappa)}, \{f_i^j(X_i) : i \notin T\}_{j=1}^{t(\kappa)}, \texttt{CRS})\}$$

$$\texttt{IDEAL}_{F,A'}(\{x_i^j : i \notin T\}_{j=1}^{t(\kappa)})$$

Note that in the above definitions, we use the min-entropy condition to constrain the tampering functions that the adversary may choose to avoid known impossibility results on deterministic zero-knowledge [GO94].

Impossibility without CRS. We stress that a common public random string as auxiliary input is necessary for our construction. In the work [DOPS04] of Dodis, Ong, Prabhakaran and Sahai, they studied the model which uses imperfect random tape in a zero knowledge protocol without CRS. The result is negative. Note that in Definition 9 and when we set $t = 1$, one can view $f_1(X)$ as an imperfect random tape and $f_1(X)$ can be all possible flat source with min-entropy $k(t(\kappa), \kappa)$. Thus, if there is no CRS as auxiliary input, it is also impossible to construct a protocol satisfying Definition 9.

Notice that even in the CRS model, the (tampering) functions $f_1, f_2, ..., f_t$ must not *depend on* the CRS to allow for a positive result. We give a proof sketch as following. The idea is very similar to that in [DOPS04].

We focus on the case where $t(\cdot) = 1$. Without loss of generality, we can assume that the length of random tape $N \geq k(1, \kappa) + \kappa$ (It can be achieved by simply padding κ random bits and never use them). Suppose the length of the transcript is bounded by $q(\kappa)$ where q is a polynomial. Consider a distinguisher D_i which just outputs the $i-$th bit of the transcript. Now for a fixed CRS and a fixed random tape of V^*, we want to show one of the following cases happens:

- There exists two tampering functions f and f' such that $f(U_N)$ and $f'(U_N)$ are both $(N, k(1, \kappa))-$flat sources but two distributions of the transcripts can be distinguished by some D_i with noticeable probability.
- The distribution of the transcript is deterministic except for a negligible probability.

Now consider the distribution of the transcript where the prover just uses the uniform random tape X. If it is deterministic except for a negligible probability, then we are done. Otherwise, there exists some i such that the $i-$th bit in the transcript is not almost deterministic. Then we may find two sets S, S' of size $2^{k(1,\kappa)}$ such that for every $X \in S$, the $i-$th bit of the transcript is always 0, and for every $X \in S'$, the $i-$th bit of the transcript is always 1. Let f and f' be functions such that $f(U_N)$ is a flat distribution over S and $f'(U_N)$ is a flat distribution over S'. Note that D_i can distinguish these two distributions with probability 1.

Then, we fix the random tape of V^* and consider all possible CRS. We may construct $f(\cdot, \texttt{CRS})$ and $f'(\cdot, \texttt{CRS})$ such that one of the above cases happens. We say a CRS is good if the second case happens, i.e. the distribution of the transcript is deterministic except for a negligible probability. We say a CRS is bad otherwise. Note that, for a bad CRS, there exists some D_i such that it will always output 0 when using f and output 1 when using f'. It means that each bad CRS corresponds to one distinguisher. Since there are in total $q(k)$ distinguisher, then there exists D_{i^*} where over $1/q(k)$ bad CRS corresponds to

it. If all but a negligible portion of CRS is good, then the prover's behavior almost fully depends on CRS and the random tape of V^*. Otherwise, D_{i^*} may distinguish two distributions with a noticeable probability.

Therefore, except for a negligible probability, the randomness of the prover only comes from CRS and the random tape of V^*, which is impossible for a non-trivial language L.

Correlated-Source Extractors. The construction of correlated tape secure protocols is closely related to the question of designing what we call correlated-source extractors. Informally, correlated-source extractors $\texttt{csExt}$ have power to break correlations between sources with a unique random seed, i.e.,

$$|\texttt{csExt}(f_i(X), Y) \circ \{\texttt{csExt}(f_j(X), Y)\}_{j \neq i} \circ Y - U \circ \{\texttt{csExt}(f_j(X), Y)\}_{j \neq i} \circ Y| < \epsilon,$$

where U is the uniform distribution and we use Y to refer the CRS. With this object and CRS, the prover can obtain a fresh uniformly random tape in each execution. Formally, we define the notion correlated-source extractor as following:

Definition 1 (Seeded Correlated-Source Extractor). *A function* $\texttt{csExt} : \{0,1\}^* \times \{0,1\}^d \to \{0,1\}^m$ *is a seeded correlated-source extractor if the following holds: There exists a polynomial* $k(\cdot,\cdot,\cdot)$ *and a negligible function* $\epsilon(\cdot)$*, such that for any polynomial* $t(\cdot)$*,* $t = t(d)$ *arbitrary functions* $\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_t$*, whose output has the same length as the input, with no fixed points, and, a source* X *with min-entropy* $k(t, m, d)$*,*

$$|\texttt{csExt}(X, U_d) \circ \{\texttt{csExt}(\mathcal{A}_i(X), U_d)\}_{i=1}^t \circ U_d - U_m \circ \{\texttt{csExt}(\mathcal{A}_i(X), U_d)\}_{i=1}^t \circ U_d| < \epsilon(d)$$

where U_m *and* U_d *are uniform strings of length* m *and* d *respectively.*

5 Explicit Construction of Correlated-Source Extractor

In this section, we will describe our construction of correlated-source extractors. To this end, we first give an explicit construction of a weak t-correlated source extractor in section 5.1 and 5.2. Then we show that it is indeed a correlated-source extractor in section 5.3. In section 5.4, we introduce a special kind of sources which we can further lower the requirement of min-entropy.

5.1 Explicit Construction of Weak Correlated-Source Extractor

We will frequently use the following lemma in the proof.

Lemma 4. *Suppose* X, X', Y, Y' *are random variables such that* $|X \circ Y - X' \circ Y'| \leq \epsilon$*. Then, for any function* $f(x, y)$*,*

$$|f(X, Y) \circ Y - f(X', Y') \circ Y'| \leq \epsilon$$

Especially, when Y is an empty string, we have $|f(X) - f(X')| \leq \epsilon$. A formal proof can be found in the full version of this paper [GS19] in Appendix A.

Before we give our construction, we need to point out an important fact about weak correlated-source extractor:

Theorem 4. *If* `wcsExt` *is a weak t-correlated-source extractor for min entropy k and output length m, then $(t+1)m \leq k$.*

We give a formal proof in the full version [GS19] of this paper in Appendix B.

In Theorem 4, it gives us an upper bound of t, i.e. $t < n$. We will use this fact in our construction.

Theorem 5. *There exists an explicit weak $t-$correlated-source extractor* `wcsExt` *for min-entropy $k \geq O(t^3(\log^2 n + \log^2(1/\epsilon)))$, seed length $d = O(\log^2 n + \log^2(1/\epsilon))$ and output length $m = O(\log n + \log(1/\epsilon))$.*

Proof. Suppose the length of `adv` is ℓ. We separate Y into several parts. Specifically, let

$$Y = Y_{\mathtt{adv}} \circ Y_1^1 \circ Y_1^2 \circ ... \circ Y_1^{2\ell} \circ Y_2^1 \circ Y_2^2 \circ ... \circ Y_2^{\ell} \circ Y_{\mathtt{start}}$$

The first part $Y_{\mathtt{adv}}$ is used to generate `adv` for X. Then we will use $Y_1^1, Y_1^2, ..., Y_1^{2\ell}$ to generate $X^1, X^2, ..., X^{2\ell}$. $Y_2^1, Y_2^2, ..., Y_2^{\ell}$ and $Y_{\mathtt{start}}$ will be used in the construction of function F. Let $d_{\mathtt{adv}} = |Y_{\mathtt{adv}}|$, $d_1 = |Y_1^i|$, $d_2 = |Y_2^i|$ and $d_{\mathtt{start}} = |Y_{\mathtt{start}}|$.

Step 1: Construction of `adv`.

We separate X into $n/d_{\mathtt{adv}}$ parts such that each part is of length $d_{\mathtt{adv}}$. Suppose $X = X^1 \circ X^2 \circ ... \circ X^{n/d_{\mathtt{adv}}}$. Construct a polynomial in the field $GF(2^{d_{\mathtt{adv}}})$:

$$F_X(n) = \sum_{i=1}^{n/d_{\mathtt{adv}}} X^i n^{i-1}$$

Let $\mathtt{adv} = F_X(Y_{\mathtt{adv}})$ as the advice of X. Then,

$$|\mathtt{adv}| = \ell = d_{\mathtt{adv}}. \tag{3}$$

For different sources X and X', $F_X(n)$ and $F_{X'}(n)$ are different. Then, $F_X(n) - F_{X'}(n) \not\equiv 0$. It is known that, $F_X(n) - F_{X'}(n) = 0$ has at most $n/d_{\mathtt{adv}}$ roots. Since $Y_{\mathtt{adv}}$ is uniformly random and independent of sources X, X', with probability at most $n/(d_{\mathtt{adv}} 2^{d_{\mathtt{adv}}})$, $F_X(Y_{\mathtt{adv}}) = F_{X'}(Y_{\mathtt{adv}})$.

Let $\mathtt{adv}_i = F_{X_i}(Y_{\mathtt{adv}})$ be the advice of the $i-$th tampering source. By union bound, with probability at least $1-tn/(d_{\mathtt{adv}} 2^{d_{\mathtt{adv}}})$, `adv` is different from $\mathtt{adv}_1, ..., \mathtt{adv}_t$. We set

$$d_{\mathtt{adv}} = \log(tn/\epsilon_1) \tag{4}$$

Then $\epsilon_1 = tn/2^{d_{\mathtt{adv}}} > tn/(d_{\mathtt{adv}} 2^{d_{\mathtt{adv}}})$. Thus, with probability $1 - \epsilon_1$, we can successfully generate a unique advice for source X.

Let $\mathtt{ADV} = \{\mathtt{adv}, \mathtt{adv}_1, ..., \mathtt{adv}_t, Y_{\mathtt{adv}}\}$. By lemma 1, we have

$$\Pr[H_\infty(X|\mathtt{ADV}) \geq H_\infty(X) - (t+2)d_{\mathtt{adv}} - \log\frac{1}{\epsilon_2}] > 1 - \epsilon_2$$

Thus, by union bound, with probability at least $1-\epsilon_1-\epsilon_2$,

$$H_\infty(X|\mathtt{ADV}) \geq H_\infty(X) - (t+2)d_{\mathtt{adv}} - \log\frac{1}{\epsilon_2} \tag{5}$$

and $\mathtt{adv}$ is different from $\mathtt{adv}_1, ..., \mathtt{adv}_t$. We say such $\mathtt{ADV}$ is good.

Now, we fix a good $\mathtt{ADV}$. For simplicity, we omit the condition $\mathtt{ADV}$.

Step 2: Generating $X^1, X^2, ..., X^{2\ell}$.

The idea is very simple, we just apply a strong-seeded extractor with seed Y_1^i to generate X^i. Let q be the length of X^i. According to Theorem 3, there exists a strong-seeded extractor $\mathtt{Ext}_1$ for min-entropy $2q$, $d = c(\log n + \log\frac{1}{\epsilon_3})$ and ϵ_3, where c is some constant. We set

$$d_1 = c(\log n + \log\frac{1}{\epsilon_3}). \tag{6}$$

By lemma 3, $\mathtt{Ext}_1$ is also a $(2q+\log(1/\epsilon_3), 2\epsilon_3)$ average case extractor. Let

$$X^i = \mathtt{Ext}_1(X, Y_1^i)$$

Recall that $\mathcal{X}^i = \{X^i, X_1^i, ..., X_t^i\}$. For every i, when we fix $\mathcal{X}^1, ..., \mathcal{X}^{i-1}, \mathcal{X}^{i+1}, ..., \mathcal{X}^{2\ell}$ and the seeds $Y_1^1, ..., Y_1^{i-1}, Y_1^{i+1}, ..., Y_1^{2\ell}$, by lemma 1, (let $\mathtt{Set}_i = \{\mathcal{X}^j, Y_1^j\}_{j\neq i}$)

$$\tilde{H}_\infty(X|\mathtt{Set}_i) \geq H_\infty(X) - (2\ell-1)(t+1)q - (2\ell-1)d_1$$

Here we need

$$H_\infty(X) - (2\ell-1)(t+1)q - (2\ell-1)d_1 \geq 2q + \log(1/\epsilon_3). \tag{7}$$

Thus,

$$|\mathtt{Ext}_1(X, Y_1^i) \circ Y_1^i \circ \mathtt{Set}_i - U_q \circ Y_1^i \circ \mathtt{Set}_i| \leq 2\epsilon_3$$

Step 3: Construction of F.

Now we will construct a suitable function F. Recall that, we use $\mathtt{adv}^i$ for the $i-$th bit of $\mathtt{adv}$. Let $Z^0 = Y_{\mathtt{start}}$ and $Z^i = F(\mathtt{adv}^i, X^{2i-1}, X^{2i}, Y_2^i, Z^{i-1})$.

The function F include two parts. First, we will apply a strong-seeded extractor on the source $X^{2i-1+\mathtt{adv}^i}$ and seed Z_{i-1}. We use W^i to denote the output of the extractor. Then, we apply another strong-seeded extractor on the source W^i and seed Y_2^i. The output will be the final output of $F(\mathtt{adv}^i, X^{2i-1}, X^{2i}, Y_2^i, Z^{i-1})$. Let z be the length of Z^i. Then $z = |Z^0| = |Y_{\mathtt{start}}| = d_{\mathtt{start}}$. Let w be the length of W^i.

According to Theorem 3, there exists a strong-seeded extractor $\mathtt{Ext}_2$ for min-entropy $2w$, $d = c(\log n + \log\frac{1}{\epsilon_4})$ and ϵ_4, where c is some constant. Similarly, there exists a strong-seeded extractor $\mathtt{Ext}_3$ for min-entropy $2z$, $d = c(\log n + \log\frac{1}{\epsilon_5})$ and ϵ_5. (For simplicity, we use the same constant in $\mathtt{Ext}_1, \mathtt{Ext}_2, \mathtt{Ext}_3$. One can choose the largest constant.) By lemma 3, $\mathtt{Ext}_2$ is also a $(2w+\log(1/\epsilon_4), 2\epsilon_4)$ average case extractor, $\mathtt{Ext}_3$ is also a $(2z+\log(1/\epsilon_5), 2\epsilon_5)$ average case extractor. We set

$$d_{\mathtt{start}} = z = c(\log n + \log\frac{1}{\epsilon_4}) \tag{8}$$

and

$$d_2 = |Y_2^i| = c(\log n + \log \frac{1}{\epsilon_5}). \tag{9}$$

Then

$$F(\mathtt{adv}^i, X^{2i-1}, X^{2i}, Y_2^i, Z^{i-1}) = \mathtt{Ext}_3(\mathtt{Ext}_2(X^{2i-1+\mathtt{adv}^i}, Z^{i-1}), Y_2^i)$$

To show correctness, we will use induction on the length of $\mathtt{adv}$. Note that, we only consider the case that $\mathtt{adv}$ is different from $\mathtt{adv}_1, \mathtt{adv}_2, ..., \mathtt{adv}_t$. We have already fixed all advice and $Y_{\mathtt{adv}}$.

We need the following lemma:

Lemma 5. *Suppose we have the following conditions:*

- *For random variables X (of length n) and W, $|X \circ W - U_n \circ W| \leq \epsilon_1$*
- *Random variable Z of length z is correlated with X and W*
- *Y is uniformly random and independent of X, W, Z.*

Then, if $\mathtt{Ext}$ is a (k, ϵ_2) average case extractor with output length m where $k \leq n - z$ and $d \leq |Y|$, we have that

$$|\mathtt{Ext}(X, Y) \circ Y \circ W \circ Z - U_m \circ Y \circ W \circ Z| < 2\epsilon_1 + \epsilon_2$$

We give a formal proof in the full version [GS19] of this paper in Appendix C.

Before we state the main lemma, we need to define a class of sets. Let $\mathtt{DIFF}^0 = \mathtt{SAME}^0 = \emptyset$. For $i \geq 1$,

$$\mathtt{DIFF}^i = \{j | \mathtt{adv}_j^i \neq \mathtt{adv}^i\}$$
$$\mathtt{SAME}^i = (\mathtt{DIFF}^{i-1} \bigcup \mathtt{SAME}^{i-1}) / \mathtt{DIFF}^i$$

Actually, $\mathtt{DIFF}^i$ is the set of indices of the advice whose $i-$th bit is different from that of the advice of X. And $\mathtt{SAME}^i$ is the set of indices of the advice whose first $i-1$ bits are different from that of the advice of X, but the $i-$th bit is the same.

Recall that, $W^i = \mathtt{Ext}_2(X^{2i-1+\mathtt{adv}^i}, Z^{i-1})$ and $F(\mathtt{adv}^i, X^{2i-1}, X^{2i}, Y_2^i, Z^{i-1}) = \mathtt{Ext}_3(W^i, Y_2^i)$. In the $i-$th step, we want to show that Z^i is uniformly random and independent of $\{Z_j^i\}_{j \in (\mathtt{DIFF}^i \bigcup \mathtt{SAME}^i)}$. By induction hypothesis, we have that Z^{i-1} is uniformly random and independent of $\{Z_j^{i-1}\}_{j \in \mathtt{SAME}^i}$. We hope we can keep this property after computing W^i, i.e., to show that W^i is uniformly random and independent of $\{W_j^i\}_{j \in \mathtt{SAME}^i}$. Then, in the second extraction, we want to break the correlation with Z_i and $\{Z_j^i\}_{j \in \mathtt{DIFF}^i}$.

We have the following main lemma:

Lemma 6. *Suppose $\mathtt{DIFF}^i$ and $\mathtt{SAME}^i$ are the same as above. Let $\eta_i = \frac{4^i - 1}{3}(8\epsilon_3 + 4\epsilon_4 + 2\epsilon_5)$. Then we have*

$$|Z^0 \circ \{\mathcal{X}_s\}_{s=1}^{2\ell} \circ \{Y_1^s\}_{s=1}^{2\ell} - U_z \circ \{\mathcal{X}_s\}_{s=1}^{2\ell} \circ \{Y_1^s\}_{s=1}^{2\ell}| = 0$$

and for every $1 \le i \le \ell$*, we have*

$$|Z^i \circ \{Z^i_j\}_{j\in(\mathtt{DIFF}^i \bigcup \mathtt{SAME}^i)} \circ \{\mathcal{X}_s\}_{s=2i+1}^{2\ell} \circ \{Y_1^s\}_{s=1}^{2\ell} \circ Z^0 \circ \{Y_2^s\}_{s=1}^{i}$$
$$-U_z \circ \{Z^i_j\}_{j\in(\mathtt{DIFF}^i \bigcup \mathtt{SAME}^i)} \circ \{\mathcal{X}_s\}_{s=2i+1}^{2\ell} \circ \{Y_1^s\}_{s=1}^{2\ell} \circ Z^0 \circ \{Y_2^s\}_{s=1}^{i}| \le \eta_i$$

Proof. We prove the lemma by induction. When $i = 0$, $\mathtt{DIFF}^0 \bigcup \mathtt{SAME}^0 = \emptyset$. We want to show that

$$|Z^0 \circ \{\mathcal{X}_s\}_{s=1}^{2\ell} \circ \{Y_1^s\}_{s=1}^{2\ell} - U_z \circ \{\mathcal{X}_s\}_{s=1}^{2\ell} \circ \{Y_1^s\}_{s=1}^{2\ell}| = 0$$

Note that $Z^0 = Y_{\mathtt{start}}$ and $Y_{\mathtt{start}}$ is uniformly random and independent of X and $\{Y_1^s\}_{s=1}^{2\ell}$. Thus, given $\{\mathcal{X}_s\}_{s=1}^{2\ell}$ and $\{Y_1^s\}_{s=1}^{2\ell}$, Z^0 is uniformly random. The statement holds.

For $i = 1$, we want to show that

$$|Z^1 \circ \{Z^1_j\}_{j\in(\mathtt{DIFF}^1 \bigcup \mathtt{SAME}^1)} \circ \{\mathcal{X}_s\}_{s=3}^{2\ell} \circ \{Y_1^s\}_{s=1}^{2\ell} \circ Z^0 \circ Y_2^1$$
$$-U_z \circ \{Z^1_j\}_{j\in(\mathtt{DIFF}^1 \bigcup \mathtt{SAME}^1)} \circ \{\mathcal{X}_s\}_{s=3}^{2\ell} \circ \{Y_1^s\}_{s=1}^{2\ell} \circ Z^0 \circ Y_2^1| \le \eta_1$$

Without loss of generality, assume $\mathtt{adv}^1 = 0$. Then, for $j \in \mathtt{DIFF}^1$, $\mathtt{adv}^1_j = 1$ and $\mathtt{SAME}^1 = \emptyset$.

In step 2, we have

$$|X^1 \circ \mathcal{X}_2 \circ \{\mathcal{X}_s\}_{s=3}^{2\ell} \circ \{Y_1^s\}_{s=1}^{2\ell} - U_q \circ \mathcal{X}_2 \circ \{\mathcal{X}_s\}_{s=3}^{2\ell} \circ \{Y_1^s\}_{s=1}^{2\ell}| \le 2\epsilon_3$$

Note that Z^0 is uniformly random and independent of $\mathcal{X}_2 \circ \{\mathcal{X}_s\}_{s=3}^{2\ell} \circ \{Y_1^s\}_{s=1}^{2\ell}$. By lemma 5, (here X is X^1, W is $\mathcal{X}_2 \circ \{\mathcal{X}_s\}_{s=3}^{2\ell} \circ \{Y_1^s\}_{s=1}^{2\ell}$, Z is empty, Y is Z^0 and $\mathtt{Ext}_2$ is a $(2w + \log(1/\epsilon_4), 2\epsilon_4)$ average case extractor),

$$|W^1 \circ Z^0 \circ \mathcal{X}_2 \circ \{\mathcal{X}_s\}_{s=3}^{2\ell} \circ \{Y_1^s\}_{s=1}^{2\ell} - U_w \circ Z^0 \circ \mathcal{X}_2 \circ \{\mathcal{X}_s\}_{s=3}^{2\ell} \circ \{Y_1^s\}_{s=1}^{2\ell}| < 4\epsilon_3 + 2\epsilon_4$$

Here, we require that

$$|X^1| = q \ge 2w + \log(1/\epsilon_4). \tag{10}$$

Since for every $j \in \mathtt{DIFF}^1$, $W^1_j = \mathtt{Ext}_2(X^2_j, Z^0)$ is a deterministic function of Z^0 and $\mathcal{X}_2$. Thus

$$|W^1 \circ Z^0 \circ \{W^1_j\}_{j\in\mathtt{DIFF}^1} \circ \{\mathcal{X}_s\}_{s=3}^{2\ell} \circ \{Y_1^s\}_{s=1}^{2\ell}$$
$$-U_w \circ Z^0 \circ \{W^1_j\}_{j\in\mathtt{DIFF}^1} \circ \{\mathcal{X}_s\}_{s=3}^{2\ell} \circ \{Y_1^s\}_{s=1}^{2\ell}| < 4\epsilon_3 + 2\epsilon_4$$

Note that Y_2^1 is uniformly random and independent of W^1 and $Z^0 \circ \{W^1_j\}_{j\in\mathtt{DIFF}^1} \circ \{\mathcal{X}_s\}_{s=3}^{2\ell} \circ \{Y_1^s\}_{s=1}^{2\ell}$. By lemma 5, (here X is W^1, W is $Z^0 \circ \{W^1_j\}_{j\in\mathtt{DIFF}^1} \circ \{\mathcal{X}_s\}_{s=3}^{2\ell} \circ \{Y_1^s\}_{s=1}^{2\ell}$, Z is empty, Y is Y_2^1 and $\mathtt{Ext}_3$ is a $(2z + \log(1/\epsilon_5), 2\epsilon_5)$ average case extractor),

$$|Z^1 \circ Y_2^1 \circ Z^0 \circ \{W^1_j\}_{j\in\mathtt{DIFF}^1} \circ \{\mathcal{X}_s\}_{s=3}^{2\ell} \circ \{Y_1^s\}_{s=1}^{2\ell}$$
$$-U_z \circ Y_2^1 \circ Z^0 \circ \{W^1_j\}_{j\in\mathtt{DIFF}^1} \circ \{\mathcal{X}_s\}_{s=3}^{2\ell} \circ \{Y_1^s\}_{s=1}^{2\ell}| < 8\epsilon_3 + 4\epsilon_4 + 2\epsilon_5$$

Here, we require that

$$|W^1| = w \geq 2z + \log(1/\epsilon_5). \tag{11}$$

Since for every $j \in \mathtt{DIFF}^1$, $Z_j^1 = \mathtt{Ext}_3(W_j^1, Y_2^1)$ is a deterministic function of Y_2^1 and $\{W_j^1\}_{j\in\mathtt{DIFF}^1}$. Thus

$$\begin{aligned}&|Z^1 \circ Y_2^1 \circ Z^0 \circ \{Z_j^1\}_{j\in\mathtt{DIFF}^1} \circ \{\mathcal{X}_s\}_{s=3}^{2\ell} \circ \{Y_1^s\}_{s=1}^{2\ell}\\ &-U_z \circ Y_2^1 \circ Z^0 \circ \{Z_j^1\}_{j\in\mathtt{DIFF}^1} \circ \{\mathcal{X}_s\}_{s=3}^{2\ell} \circ \{Y_1^s\}_{s=1}^{2\ell}| < 8\epsilon_3 + 4\epsilon_4 + 2\epsilon_5\end{aligned}$$

Note that $\mathtt{SAME}^1 = \emptyset$. It is exactly what we want to prove in the case $i = 1$.

Now suppose the lemma is correct for $i-1$, consider the case for i. According to induction hypothesis, we have that

$$\begin{aligned}&|Z^{i-1} \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ \{\mathcal{X}_s\}_{s=2i-1}^{2\ell} \circ \{Y_1^s\}_{s=1}^{2\ell} \circ Z^0 \circ \{Y_2^s\}_{s=1}^{i-1}\\ &-U_z \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ \{\mathcal{X}_s\}_{s=2i-1}^{2\ell} \circ \{Y_1^s\}_{s=1}^{2\ell} \circ Z^0 \circ \{Y_2^s\}_{s=1}^{i-1}| \leq \eta_{i-1}\end{aligned}$$

For simplicity, we define

$$\mathfrak{X}_i = \{\mathcal{X}_s\}_{s=2i+1}^{2\ell},\ T = \{Y_1^s\}_{s=1}^{2\ell} \circ Z^0,\ \mathcal{Y}_i = \{Y_2^s\}_{s=1}^{i}$$

Thus, we may rewrite the induction hypothesis for the case $i-1$ by

$$|Z^{i-1} \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ \mathfrak{X}_{i-1} \circ T \circ \mathcal{Y}_{i-1} - U_z \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ \mathfrak{X}_{i-1} \circ T \circ \mathcal{Y}_{i-1}| \leq \eta_{i-1}$$

Let Z' be a uniformly random string over $\{0,1\}^z$ and independent of $X, X_1, ..., X_t$ and Y. Then, we may use Z' instead of U_z, i.e.,

$$|Z^{i-1} \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ \mathfrak{X}_{i-1} \circ T \circ \mathcal{Y}_{i-1} - Z' \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ \mathfrak{X}_{i-1} \circ T \circ \mathcal{Y}_{i-1}| \leq \eta_{i-1}$$

Without loss of generality, assume $\mathtt{adv}^i = 0$. Then, $\mathtt{adv}_j^i = 1$ for $j \in \mathtt{DIFF}^i$ and $\mathtt{adv}_j^i = 0$ for $j \in \mathtt{SAME}^i$. We have

$$W^i = \mathtt{Ext}_2(X^{2i-1}, Z^{i-1})$$

Note that, in step 2, we have

$$|X^{2i-1} \circ \mathcal{X}_{2i} \circ \mathfrak{X}_i \circ \{Y_1^s\}_{s=1}^{2\ell} - U_q \circ \mathcal{X}_{2i} \circ \mathfrak{X}_i \circ \{Y_1^s\}_{s=1}^{2\ell}| \leq 2\epsilon_3$$

Also note that $\mathcal{Y}_{i-1}$ is independent of $X^{2i-1} \circ \mathcal{X}_{2i} \circ \mathfrak{X}_i \circ \{Y_1^s\}_{s=1}^{2\ell}$. Therefore,

$$|X^{2i-1} \circ \mathcal{X}_{2i} \circ \mathfrak{X}_i \circ \{Y_1^s\}_{s=1}^{2\ell} \circ \mathcal{Y}_{i-1} - U_q \circ \mathcal{X}_{2i} \circ \mathfrak{X}_i \circ \{Y_1^s\}_{s=1}^{2\ell} \circ \mathcal{Y}_{i-1}| \leq 2\epsilon_3$$

Since Z' is independent of X and Y, it is independent of $X^{2i-1} \circ \mathcal{X}_{2i} \circ \mathfrak{X}_i \circ \{Y_1^s\}_{s=1}^{2\ell} \circ \mathcal{Y}_{i-1}$ and $\{W_j^i\}_{j\in\mathtt{SAME}^i} \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ Z^0$. By lemma 5, (here X is X^{2i-1}, W is $\mathcal{X}_{2i} \circ \mathfrak{X}_i \circ \{Y_1^s\}_{s=1}^{2\ell} \circ \mathcal{Y}_{i-1}$, Z is $\{W_j^i\}_{j\in\mathtt{SAME}^i} \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ Z^0$, Y is Z' and $\mathtt{Ext}_2$ is a $(2w + \log(1/\epsilon_4), 2\epsilon_4)$ average case extractor),

$$|\mathtt{Ext}_2(X^{2i-1}, Z') \circ Z' \circ \mathcal{X}_{2i} \circ \mathfrak{X}_i \circ T \circ \{W_j^i\}_{j\in\mathtt{SAME}^i} \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ \mathcal{Y}_{i-1}$$

$$-U_w \circ Z' \circ \mathcal{X}_{2i} \circ \mathfrak{X}_i \circ T \circ \{W_j^i\}_{j\in\mathtt{SAME}^i} \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ \mathcal{Y}_{i-1}| < 4\epsilon_3 + 2\epsilon_4$$

Here, we require that

$$\begin{aligned} &|X^{2i-1}| - |\{W_j^i\}_{j\in\mathtt{SAME}^i} \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ Z^0| \\ &\geq q - (tw + tz + z) \geq 2w + \log(1/\epsilon_4) \end{aligned} \tag{12}$$

Recall that,

$$|Z^{i-1} \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ \mathfrak{X}_{i-1} \circ T \circ \mathcal{Y}_{i-1} - Z' \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ \circ \mathfrak{X}_{i-1} \circ T \circ \mathcal{Y}_{i-1}| \leq \eta_{i-1}$$

Notice that $\{W_j^i\}_{j\in\mathtt{SAME}^i}$ is a deterministic function of $\{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i}$ and $\mathcal{X}_{2i-1}$. Also, $W^i = \mathtt{Ext}_2(X^{2i-1}, Z^{i-1})$.

Thus, by reordering the composition parts, we have

$$\begin{aligned} &|W^i \circ Z^{i-1} \circ \mathcal{X}_{2i} \circ \mathfrak{X}_i \circ T \circ \{W_j^i\}_{j\in\mathtt{SAME}^i} \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ \mathcal{Y}_{i-1} \\ &-\mathtt{Ext}_2(X^{2i-1}, Z') \circ Z' \circ \mathcal{X}_{2i} \circ \mathfrak{X}_i \circ T \circ \{W_j^i\}_{j\in\mathtt{SAME}^i} \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ \mathcal{Y}_{i-1}| \\ \leq\ &\eta_{i-1} \end{aligned}$$

Still, since $\{W_j^i\}_{j\in\mathtt{SAME}^i}$ is a deterministic function of $\{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i}$ and $\mathcal{X}_{2i-1}$, we have

$$\begin{aligned} &|Z' \circ \mathcal{X}_{2i} \circ \mathfrak{X}_i \circ T \circ \{W_j^i\}_{j\in\mathtt{SAME}^i} \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ \mathcal{Y}_{i-1} \\ &-Z^{i-1} \circ \mathcal{X}_{2i} \circ \mathfrak{X}_i \circ T \circ \{W_j^i\}_{j\in\mathtt{SAME}^i} \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ \mathcal{Y}_{i-1}| \\ \leq\ &|Z' \circ \mathfrak{X}_{i-1} \circ T \circ \{W_j^i\}_{j\in\mathtt{SAME}^i} \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ \mathcal{Y}_{i-1} \\ &-Z^{i-1} \circ \mathfrak{X}_{i-1} \circ T \circ \{W_j^i\}_{j\in\mathtt{SAME}^i} \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ \mathcal{Y}_{i-1}| \\ =\ &|Z' \circ \mathfrak{X}_{i-1} \circ T \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ \mathcal{Y}_{i-1} \\ &-Z^{i-1} \circ \mathfrak{X}_{i-1} \circ T \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ \mathcal{Y}_{i-1}| \\ <\ &\eta_{i-1} \end{aligned}$$

In total, we have

$$\begin{aligned} &|W^i \circ Z^{i-1} \circ \mathcal{X}_{2i} \circ \mathfrak{X}_i \circ T \circ \{W_j^i\}_{j\in\mathtt{SAME}^i} \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ \mathcal{Y}_{i-1} \\ &-U_w \circ Z^{i-1} \circ \mathcal{X}_{2i} \circ \mathfrak{X}_i \circ T \circ \{W_j^i\}_{j\in\mathtt{SAME}^i} \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ \mathcal{Y}_{i-1}| \\ \leq\ &|W^i \circ Z^{i-1} \circ \mathcal{X}_{2i} \circ \mathfrak{X}_i \circ T \circ \{W_j^i\}_{j\in\mathtt{SAME}^i} \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ \mathcal{Y}_{i-1} \\ &-\mathtt{Ext}_2(X^{2i-1}, Z') \circ Z' \circ \mathcal{X}_{2i} \circ \mathfrak{X}_i \circ T \circ \{W_j^i\}_{j\in\mathtt{SAME}^i} \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ \mathcal{Y}_{i-1}| \\ &+|\mathtt{Ext}_2(X^{2i-1}, Z') \circ Z' \circ \mathcal{X}_{2i} \circ \mathfrak{X}_i \circ T \circ \{W_j^i\}_{j\in\mathtt{SAME}^i} \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ \mathcal{Y}_{i-1} \\ &-U_w \circ Z' \circ \mathcal{X}_{2i} \circ \mathfrak{X}_i \circ T \circ \{W_j^i\}_{j\in\mathtt{SAME}^i} \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ \mathcal{Y}_{i-1}| \\ &+|U_w \circ Z' \circ \mathcal{X}_{2i} \circ \mathfrak{X}_i \circ T \circ \{W_j^i\}_{j\in\mathtt{SAME}^i} \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ \mathcal{Y}_{i-1} \\ &-U_w \circ Z^{i-1} \circ \mathcal{X}_{2i} \circ \mathfrak{X}_i \circ T \circ \{W_j^i\}_{j\in\mathtt{SAME}^i} \circ \{Z_j^{i-1}\}_{j\in\mathtt{SAME}^i} \circ \mathcal{Y}_{i-1}| \\ <\ &\eta_{i-1} + 4\epsilon_3 + 2\epsilon_4 \end{aligned}$$

$$
\begin{aligned}
&+|Z' \circ \mathcal{X}_{2i} \circ \mathfrak{X}_i \circ T \circ \{W_j^i\}_{j\in \mathtt{SAME}^i} \circ \{Z_j^{i-1}\}_{j\in \mathtt{SAME}^i} \circ \mathcal{Y}_{i-1} \\
&-Z^{i-1} \circ \mathcal{X}_{2i} \circ \mathfrak{X}_i \circ T \circ \{W_j^i\}_{j\in \mathtt{SAME}^i} \circ \{Z_j^{i-1}\}_{j\in \mathtt{SAME}^i} \circ \mathcal{Y}_{i-1}| \\
&\leq 2\eta_{i-1} + 4\epsilon_3 + 2\epsilon_4
\end{aligned}
$$

Note that Y_2^i is uniformly random and independent of W^i and $Z^{i-1} \circ \mathcal{X}_{2i} \circ \mathfrak{X}_i \circ T \circ \{W_j^i\}_{j\in \mathtt{SAME}^i} \circ \{Z_j^{i-1}\}_{j\in \mathtt{SAME}^i \bigcup \mathtt{DIFF}^i} \circ \mathcal{Y}_{i-1}$. By lemma 5, (here X is W^i, W is $Z^{i-1} \circ \mathcal{X}_{2i} \circ \mathfrak{X}_i \circ T \circ \{W_j^i\}_{j\in \mathtt{SAME}^i} \circ \{Z_j^{i-1}\}_{j\in \mathtt{SAME}^i} \circ \mathcal{Y}_{i-1}$, Z is $\{Z_j^{i-1}\}_{j\in \mathtt{DIFF}^i}$, Y is Y_2^i and $\mathtt{Ext}_3$ is a $(2z + \log(1/\epsilon_5), 2\epsilon_5)$), we have

$$
\begin{aligned}
&|Z^i \circ Y_2^i \circ Z^{i-1} \circ \mathcal{X}_{2i} \circ \mathfrak{X}_i \circ T \circ \{W_j^i\}_{j\in \mathtt{SAME}^i} \circ \{Z_j^{i-1}\}_{j\in \mathtt{SAME}^i \bigcup \mathtt{DIFF}^i} \circ \mathcal{Y}_{i-1} \\
&-U_z \circ Y_2^i \circ Z^{i-1} \circ \mathcal{X}_{2i} \circ \mathfrak{X}_i \circ T \circ \{W_j^i\}_{j\in \mathtt{SAME}^i} \circ \{Z_j^{i-1}\}_{j\in \mathtt{SAME}^i \bigcup \mathtt{DIFF}^i} \circ \mathcal{Y}_{i-1}| \\
&< 4\eta_{i-1} + 8\epsilon_3 + 4\epsilon_4 + 2\epsilon_5 = \eta_i
\end{aligned}
$$

Here, we need

$$
|W^i| - |\{Z_j^{i-1}\}_{j\in \mathtt{DIFF}^i}| \geq w - tz \geq 2z + \log(1/\epsilon_5). \tag{13}
$$

Note that $\{Z_j^i\}_{j\in \mathtt{SAME}^i}$ is a deterministic function of Y_2^i and $\{W_j^i\}_{j\in \mathtt{SAME}^i}$. And $\{Z_j^i\}_{j\in \mathtt{DIFF}^i}$ is a deterministic function of Y_2^i, $\{Z_j^i\}_{j\in \mathtt{DIFF}^i}$ and $\mathcal{X}_{2i}$. Thus, we have (We will discard $\{W_j^i\}_{j\in \mathtt{SAME}^i}$ and Z^{i-1})

$$
\begin{aligned}
&|Z^i \circ Y_2^i \circ \{Z_j^i\}_{j\in \mathtt{SAME}^i \bigcup \mathtt{DIFF}^i} \circ \mathfrak{X}_i \circ T \circ \mathcal{Y}_{i-1} \\
&-U_z \circ Y_2^i \circ \{Z_j^i\}_{j\in \mathtt{SAME}^i \bigcup \mathtt{DIFF}^i} \circ \mathfrak{X}_i \circ T \circ \mathcal{Y}_{i-1}| < \eta_i
\end{aligned}
$$

It is exactly what we want. Thus the statement is true for the case i. □

By assumption, all advice are different. Then $\mathtt{DIFF}^\ell \bigcup \mathtt{SAME}^\ell$ include all indices of advice. Therefore,

$$
\begin{aligned}
&|Z^\ell \circ \{Z_j^\ell\}_{j=1}^t \circ \{Y_1^s\}_{s=1}^{2\ell} \circ \{Y_2^s\}_{s=1}^{\ell} \circ Z^0 - U_z \circ \{Z_j^\ell\}_{j=1}^t \circ \{Y_1^s\}_{s=1}^{2\ell} \circ \{Y_2^s\}_{s=1}^{\ell} \circ Z^0| \\
&< \frac{4^\ell - 1}{3} (8\epsilon_3 + 4\epsilon_4 + 2\epsilon_5)
\end{aligned}
$$

So far, we are in the condition that $\mathtt{ADV}$ is good. Make everything together,

$$
\begin{aligned}
&|Z^\ell \circ \{Z_j^\ell\}_{j=1}^t \circ \{Y_1^s\}_{s=1}^{2\ell} \circ \{Y_2^s\}_{s=1}^{\ell} \circ Z^0 \circ \mathtt{ADV} \\
&-U_z \circ \{Z_j^\ell\}_{j=1}^t \circ \{Y_1^s\}_{s=1}^{2\ell} \circ \{Y_2^s\}_{s=1}^{\ell} \circ Z^0 \circ \mathtt{ADV}| \\
&< \Pr[\mathtt{ADV} \text{ is bad}] + \Pr[\mathtt{ADV} \text{ is good}] \cdot \frac{4^\ell - 1}{3} (8\epsilon_3 + 4\epsilon_4 + 2\epsilon_5) \\
&< \epsilon_1 + \epsilon_2 + \frac{4^\ell - 1}{3} (8\epsilon_3 + 4\epsilon_4 + 2\epsilon_5)
\end{aligned}
$$

The total error is $\epsilon_1 + \epsilon_2 + \frac{4^\ell - 1}{3} (8\epsilon_3 + 4\epsilon_4 + 2\epsilon_5)$. Let

$$
\epsilon_1 = \epsilon_2 = \frac{\epsilon}{e^2} = O(\epsilon).
$$

By (3), (4) and the fact that $t < n$, we have

$$\ell = d_{\mathtt{adv}} = \log(tn/\epsilon_1) = O(\log n + \log(1/\epsilon))$$

Let

$$\epsilon_3 = \epsilon_4 = \epsilon_5 = \frac{\epsilon}{14e^{2\ell}} = O(\frac{\epsilon^3}{t^2 n^2})$$

Then we have

$$\epsilon_1 + \epsilon_2 + \frac{4^\ell - 1}{3}(8\epsilon_3 + 4\epsilon_4 + 2\epsilon_5) < \epsilon$$

By (6), (8), (9) and the fact that $t < n$, we have

$$d_1 = d_2 = d_{\mathtt{start}} = O(\log n + \log(1/\epsilon))$$

Therefore, the length of the seed

$$|Y| = d_{\mathtt{adv}} + 2\ell d_1 + \ell d_2 + d_{\mathtt{start}} = O(\ell(\log n + \log(1/\epsilon))) = O(\log^2 n + \log^2(1/\epsilon))$$

All requirements for the min-entropy of $X|\mathtt{ADV}$ are (7), (10), (11), (12), (13), i.e.

$$\begin{aligned} H_\infty(X|\mathtt{ADV}) - (2\ell - 1)(t+1)q - (2\ell - 1)d_1 &\geq 2q + \log(1/\epsilon_3) \\ q &\geq 2w + \log(1/\epsilon_4) \\ w &\geq 2z + \log(1/\epsilon_5) \\ q - (tw + tz + z) &\geq 2w + \log(1/\epsilon_4) \\ w - tz &\geq 2z + \log(1/\epsilon_5) \end{aligned}$$

Therefore, we have

$$\begin{aligned} z &= O(\log n + \log(1/\epsilon)) \\ w &= O(t(\log n + \log(1/\epsilon))) \\ q &= O(t^2(\log n + \log(1/\epsilon))) \\ H_\infty(X|\mathtt{ADV}) &\geq O(t^3(\log^2 n + \log^2(1/\epsilon))) \end{aligned}$$

Note that in (5), for a good $\mathtt{ADV}$, we have

$$H_\infty(X|\mathtt{ADV}) \geq H_\infty(X) - (t+2)d_{\mathtt{adv}} - \log(1/\epsilon_2)$$

Thus, we set

$$\begin{aligned} H_\infty(X) &= O(t^3(\log^2 n + \log^2(1/\epsilon))) + (t+2)d_{\mathtt{adv}} + \log(1/\epsilon_2) \\ &= O(t^3(\log^2 n + \log^2(1/\epsilon))) \end{aligned}$$

Note that the final output is Z^ℓ. Then, the output length $m = z = O(\log n + \log(1/\epsilon))$.

Thus, there exists an explicit construction of $\mathtt{wcsExt}$ for

$$\begin{aligned} H_\infty(X) &\geq O(t^3(\log^2 n + \log^2(1/\epsilon))) \\ |Y| &= O(\log^2 n + \log^2(1/\epsilon)) \\ m &= O(\log n + \log(1/\epsilon)) \end{aligned}$$

□

Also, we may generalize our result to an average case weak t-correlated-source extractor.

Definition 11 (Average Case weak t-Correlated-Source Extractor). *A function* $\texttt{wcsExt} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *is an average case weak* t*-correlated-source extractor for average conditional min-entropy* k *and error* ϵ *if the following holds: If* X *is a source in* $\{0,1\}^n$*,* W *is some random variable such that* $\tilde{H}_\infty(X|W) \geq k$*,* $\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_t$ *are arbitrary tampering functions defined on* $\{0,1\}^n \to \{0,1\}^n$ *with no fixed points, then*

$$|\texttt{wcsExt}(X, U_d) \circ \{\texttt{wcsExt}(\mathcal{A}_i(X), U_d)\}_{i=1}^t \circ U_d \circ W$$
$$-U_m \circ \{\texttt{wcsExt}(\mathcal{A}_i(X), U_d)\}_{i=1}^t \circ U_d \circ W| < \epsilon$$

where U_m *is independent of* U_d *and* X*.*

We have the following lemma.

Lemma 7. *For any* δ*, if* $\texttt{wcsExt}$ *is a weak* $t-$*correlated-source extractor for min-entropy* k *and error* ϵ*, then it is also an average case* t*-correlated source extractor for average conditional min-entropy* $k + \log 1/\delta$ *and error* $\epsilon + \delta$*.*

The proof can be easily generalized from Lemma 2.3 in [DRS04].

Therefore, combining Theorem 5 and Lemma 7 by setting $\delta = \epsilon$, we have

Theorem 6. *There exists an explicit average case weak* t*-correlated source extractor* $\texttt{wcsExt}$ *for average conditional min-entropy* $k \geq O(t^3(\log^2 n + \log^2(1/\epsilon)))$*, seed length* $d = O(\log^2 n + \log^2(1/\epsilon))$ *and output length* $m = O(\log n + \log(1/\epsilon))$*.*

5.2 Boosting the Output Length

In the above construction, a major limitation is that the output length is only $O(\log n + \log 1/\epsilon)$. To boost the output length, we separate X into $2\ell + 1$ parts instead of 2ℓ parts in Theorem 5. We may set the length of the last part to be long enough. It can be viewed as the case that we append 0 to all advice. Then the length of the advice becomes $\ell + 1$. Since the last bit is 0, we will never choose $X^{2\ell+2}$. Thus, we only need one more part.

In Lemma 6, we have shown that, for every i, W^i given $\{W_j^i\}_{j \in \texttt{SAME}^i}$ is uniformly random. When $i = \ell + 1$, $\texttt{SAME}^{\ell+1} = [t]$. Thus, $W^{\ell+1}$ given $\{W_j^i\}_{j \in [t]}$ is uniformly random. $W^{\ell+1}$ will be the final output of our extractor.

Denote the length of $W^{\ell+1}$ to be m. We need the length of $X^{2\ell+1}$ to be $O(tm)$. Then, the min-entropy requirement for the original source becomes $O(t^3(\log^2 n + \log^2(1/\epsilon)) + t^2 m)$.

We have the following theorem.

Theorem 7. *There exists an explicit weak* $t-$*correlated-source extractor* $\texttt{wcsExt}$ *where* $k \geq O(t^3(\log^2 n + \log^2(1/\epsilon)) + t^2 m)$ *and* $d = O(\log^2 n + \log^2(1/\epsilon))$*, where* m *is the output length.*

A formal proof can be found in the full version [GS19] of this paper in Appendix D.

5.3 Explicit Construction of Correlated-Source Extractor

We show that, our explicit construction in Theorem 7 is indeed a correlated-source extractor.

To see this, we set

$$\epsilon(d) = \Theta(2^{-\sqrt{d}})$$

and

$$k(t, m, d) = \Theta(t^3 d + t^2 m)$$

Clearly, $\epsilon(\cdot)$ is a negligible function and $k(\cdot, \cdot, \cdot)$ is a polynomial. Then, we only need to show that $d \geq O(\log^2 n + \log^2(1/\epsilon))$ and $k(t, m) \geq O(t^3(\log^2 n + \log^2(1/\epsilon)) + t^2 m)$. Note that the source length is bounded by a polynomial of d, and $\log^2(1/\epsilon) = \Theta(d)$. Therefore, $d \geq O(\log^2 n + \log^2(1/\epsilon))$. Further, we have $k(t, m, d) = \Theta(t^3 d + t^2 m) \geq O(t^3(\log^2 n + \log^2(1/\epsilon)) + t^2 m)$.

Thus, we have the following theorem.

Theorem 1. *There exists an explicit correlated-source extractor* `csExt` *with*

$$k(t, m, d) = \Theta(t^3 d + t^2 m)$$
$$\epsilon(d) = \Theta(2^{-\sqrt{d}})$$

where m is the length of the output.

We can also define what we call an average case correlated-source extractor in a similar way.

Definition 12 (Average Case Correlated-Source Extractor). *A function* $\texttt{csExt} : \{0,1\}^* \times \{0,1\}^d \to \{0,1\}^m$ *is an average case correlated-source extractor if the following holds: There exists a polynomial $k(\cdot, \cdot, \cdot)$ and a negligible function $\epsilon(\cdot)$, such that for any polynomial $t(\cdot)$, $t = t(d)$ arbitrary functions $\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_t$, whose output has the same length as the input, with no fixed points, a source X and a random variable W such that $\tilde{H}_\infty(X|W) \geq k(t, m, d)$,*

$$|\texttt{csExt}(X, U_d) \circ \{\texttt{csExt}(\mathcal{A}_i(X), U_d)\}_{i=1}^t \circ U_d \circ W$$
$$-U_m \circ \{\texttt{csExt}(\mathcal{A}_i(X), U_d)\}_{i=1}^t \circ U_d \circ W| < \epsilon(d)$$

where U_m is independent of U_d and X.

If we are using an average case weak t-correlated source extractor, then we will get an average case correlated source extractor.

Theorem 8. *There exists an explicit average case correlated-source extractor* `csExt` *with*

$$k(t, m, d) = \Theta(t^3 d + t^2 m)$$
$$\epsilon(d) = \Theta(2^{-\sqrt{d}})$$

where m is the length of the output.

5.4 Generalizing the Entropy Requirements

In our construction, the min-entropy requirement on the source (denoted by k) grows with t. This is inherent since the total entropy of all the sources together may only be k (since each source may have zero min-entropy given any other source) which must be at least $t \cdot m$ where m is the size of the output of the extractor. A natural question is: could we place a stronger independence condition on the different sources which allows us to obtain a construction requiring the sources to have lower min-entropy? We outline such an extension in this section.

Definition 13 (Closed-Set Correlated Sources). *We say a sequence of sources $X_1, X_2, ..., X_\ell$ is a $(t,k)-$closed-set correlated sources if for every X_i,*

- *There exists a set of sources S_i such that $X_i \in S_i$ and $|S_i| \leq t$*
- *When given all sources outside S_i, X_i still has enough min-entropy, i.e.,*

$$\tilde{H}_\infty(X_i | \{X_j\}_{j=1}^{\ell} / S_i) \geq k$$

For a $(t,k)-$closed-set correlated sources, we can use an average case correlated source extractor on the set S_i, viewing X_i as the original source and $X_j \in S_i$ as the tampering source. Thus, we have the following corollary.

Corollary 1. *Let* `csExt` *be an average case correlated-source extractor constructed in Theorem 6. Let Y be a random seed of length specified in Theorem 8. For a closed-set correlated sources $X_1, X_2, ..., X_\ell$,*

$$|\texttt{csExt}(X_i, Y) \circ \{\texttt{csExt}(X_j, Y)\}_{j \neq i} - U_m \circ \{\texttt{csExt}(X_j, Y)\}_{j \neq i}| < \epsilon$$

6 Constructing Secure Correlated-Tape Multi-Party Computation Protocol

We use correlated-source extractor and resettable multi-party computation protocol based on [GS09] as building blocks. Suppose `csExt` is a correlated-source extractor, π' is a resettably secure multi-party computation protocol for ideal functionality F (in the standard setting). We construct a correlated-tape secure MPC π as follows:

In the protocol π, each party will first run `csExt` with its secret random tape and CRS. Then use the output of `csExt` as the new random tape and follow the steps in π'. We have the following theorem.

Theorem 9. *Let π, π' be defined as above. For every security parameter κ, suppose $q(\kappa)$ is the length of the random tape that π' needs. Let* `csExt` *be a correlated-source extractor in Theorem 1 with $d = \kappa, m = q(\kappa)$ and polynomials $k'(\cdot,\cdot,\cdot), \epsilon'(\cdot)$. Let $len(\kappa) = \kappa$ and $k(t,\kappa) = k'(t, q(\kappa), \kappa) + t\kappa$. Then π is a correlated-tape multi-party computation protocol.*

Proof. Let T be the set of corrupted parties which controlled by the adversary. We define a pattern $\boldsymbol{S} = (s_1, s_2, ..., s_t)$ where $s_j \in [t]$. If for two patterns $\boldsymbol{S}, \boldsymbol{S'}$, there exists a permutation $p : [t] \to [t]$ such that $s_j = p(s'_j)$ for every $j \in [t]$, we view them as the same pattern. We say an input X_i is consistent with $\boldsymbol{S}$ respect to $\{f_i^j\}_{j=1}^t$, if for every $j_1, j_2 \in [t]$, $f_i^{j_1}(X_i) = f_i^{j_2}(X_i)$ if and only if $s_{j_1} = s_{j_2}$. Let

$$\texttt{Pattern}[\boldsymbol{S}, i] = \{X_i |\ X_i \text{ is consistent with } \boldsymbol{S} \text{ respect to } \{f_i^j\}_{j=1}^t\}$$

Note that there are at most $t^t = 2^{t \log t} = 2^{o(t\kappa)}$ patterns in total. Let $\texttt{ratio}[\boldsymbol{S}, i] \in [0, 1]$ be the ratio of X_i which is consistent with $\boldsymbol{S}$ respect to $\{f_i^j\}_{j=1}^t$. Indeed $\{\texttt{Pattern}[\boldsymbol{S}, i]\}_{\boldsymbol{S}}$ is a partition of all X_i and thus

$$\sum_{\boldsymbol{S}} \texttt{ratio}[\boldsymbol{S}, i] = 1$$

After sampling X_i for P_i, we will reveal the pattern information to the adversary. Let

$$\texttt{BAD}_i = \{\boldsymbol{S} : \texttt{ratio}[\boldsymbol{S}, i] \leq \frac{1}{2^{t\kappa}}\}$$

Then, we show that, for every $\{\boldsymbol{S_i} : \boldsymbol{S_i} \notin \texttt{BAD}_i\}_{i \notin T}$, there exists an adversary A' in π' such that the following two distributions are computationally indistinguishable:

$$\{\texttt{CRS} \sim U_{len(\kappa)} : \texttt{REAL}_{F,A'}(\{x_i^j : i \notin T\}_{j=1}^t, \{X_i^j : i \notin T\}_{j=1}^t, \texttt{CRS})\}$$

$$\{\texttt{CRS} \sim U_{len(\kappa)} : \texttt{REAL}_{F,A}(\{x_i^j : i \notin T\}_{j=1}^t, \{f_i^j(X_i) : i \notin T\}_{j=1}^t, \texttt{CRS})\}$$

where $\{X_i^j : i \notin T\}_{j=1}^t$ is sampled based on the strategy of A' we will mention later.

We design A' to follow the strategy: After receiving $\{\boldsymbol{S_i} : \boldsymbol{S_i} \notin \texttt{BAD}_i\}_{i \notin T}$, for party P_i, in the $j-$th round, if there exists $j^* \in [j-1]$ such that $(\boldsymbol{S_i})_j = (\boldsymbol{S_i})_{j^*}$, then let P_i use the same random tape as the j^*-th round, otherwise let P_i use a fresh random tape.

Since the random tapes of each parties are independent, we only need to show that, for $\boldsymbol{S_i}$, we have:

$$\{X_i \sim \texttt{Pattern}[\boldsymbol{S_i}, i], \texttt{CRS} \sim U_{len(\kappa)} : \{\texttt{csExt}(f_i^j(X_i), \texttt{CRS})\}_{j=1}^t\} =_c \{X_i^j\}_{j=1}^t$$

Let $\texttt{Index}(\boldsymbol{S}) = \{j : \forall j' \in [j-1], s_j \neq s_{j'}\}$. Then it is sufficient to show that

$$|\{\texttt{csExt}(f_i^j(X_i), \texttt{CRS})\}_{j \in \texttt{Index}(\boldsymbol{S_i})} - U_{q(\kappa)|\texttt{Index}(\boldsymbol{S})|}| \leq \mu(\kappa)$$

where $\mu(\cdot)$ is a negligible function. Note that, for every possible output y of f_i^j,

$$\begin{aligned}\Pr[f_i^j(X_i) = y] &\geq \Pr[f_i^j(X_i) = y \text{ and } X_i \in \texttt{Pattern}[\boldsymbol{S_i}, i]] \\ &= \Pr[f_i^j(X_i) = y |\ X_i \in \texttt{Pattern}[\boldsymbol{S_i}, i]] \Pr[X_i \in \texttt{Pattern}[\boldsymbol{S_i}, i]]\end{aligned}$$

$$\geq \frac{1}{2^{t\kappa}} \Pr[f_i^j(X_i) = y |\ X_i \in \mathtt{Pattern}[\boldsymbol{S_i}, i]]$$

By condition, $H_\infty(f_i^j(X_i)) \geq k(t,\kappa)$. Thus $\Pr[f_i^j(X_i) = y] \leq \frac{1}{2^{k(t,\kappa)}}$. We have

$$\Pr[f_i^j(X_i) = y |\ X_i \in \mathtt{Pattern}[\boldsymbol{S_i}, i]] \leq \frac{1}{2^{k(t,\kappa)-t\kappa}} = \frac{1}{2^{k'(t,q(\kappa),\kappa)}}$$

Therefore, given $X_i \in \mathtt{Pattern}[\boldsymbol{S_i}, i]$, $f_i^j(X_i)$ still has enough min-entropy to use correlated-source extractor $\mathtt{csExt}$. For every $j \in \mathtt{Index}(\boldsymbol{S_i})$,

$$\begin{aligned}&|\mathtt{csExt}(f_i^j(X_i), \mathtt{CRS}) \circ \{\mathtt{csExt}(f_i^{j'}(X_i), \mathtt{CRS})\}_{j' \neq j, j' \in \mathtt{Index}(\boldsymbol{S_i})} \\ &- U_{q(\kappa)} \circ \{\mathtt{csExt}(f_i^{j'}(X_i), \mathtt{CRS})\}_{j' \neq j, j' \in \mathtt{Index}(\boldsymbol{S_i})}| \leq \epsilon'(\kappa)\end{aligned}$$

By union bound,

$$|\{\mathtt{csExt}(f_i^j(X_i), \mathtt{CRS})\}_{j \in \mathtt{Index}(\boldsymbol{S_i})} - U_{q(\kappa)|\mathtt{Index}(\boldsymbol{S})|}| \leq |\mathtt{Index}(\boldsymbol{S_i})|\epsilon'(\kappa) \leq t\epsilon'(\kappa)$$

Therefore, if $\boldsymbol{S_i} \notin \mathtt{BAD}_i$, the error is bounded by some negligible probability and further, A' satisfies our requirement.

Note that

$$\Pr[X_i \in \mathtt{Pattern}[\boldsymbol{S}, i] \text{ where } \boldsymbol{S} \in \mathtt{BAD}_i] \leq \frac{2^{t\log(t)}}{2^{t\kappa}} = \frac{1}{2^{O(t\kappa)}}$$

Thus, the distinguishable advantage is bounded by the sum of the probability that some $\boldsymbol{S_i} \in \mathtt{BAD}_i$ and the probability that one can distinguish the two distributions generated by A' and A given all $\boldsymbol{S_i} \notin \mathtt{BAD}_i$, which is still a negligible probability over security parameter κ. □

Correlated-Source Extractors with Almost Optimal Parameters. We give a non-explicit construction of correlated-source extractors with almost optimal parameters. For lack of space, this result can be found in the full version [GS19] of this paper in Appendix E.

References

ABP15. Michel Abdalla, Fabrice Benhamouda, and Alain Passelègue. An algebraic framework for pseudorandom functions and applications to related-key security. In *Advances in Cryptology - CRYPTO 2015 - 35th Annual Cryptology Conference, Santa Barbara, CA, USA, August 16-20, 2015, Proceedings, Part I*, pages 388–409, 2015.

ACM+14. Per Austrin, Kai-Min Chung, Mohammad Mahmoody, Rafael Pass, and Karn Seth. On the impossibility of cryptography with tamperable randomness. In Juan A. Garay and Rosario Gennaro, editors, *Advances in Cryptology – CRYPTO 2014*, pages 462–479, Berlin, Heidelberg, 2014. Springer Berlin Heidelberg.

ACRT97. A. E. Andreev, A. E. F. Clementi, J. D. P. Rolim, and L. Trevisan. Weak random sources, hitting sets, and bpp simulations. In *Proceedings 38th Annual Symposium on Foundations of Computer Science*, pages 264–272, Oct 1997.

BACD+18. Avraham Ben-Aroya, Eshan Chattopadhyay, Dean Doron, Xin Li, and Amnon Ta-Shma. A new approach for constructing low-error, two-source extractors. In *Proceedings of the 33rd Computational Complexity Conference*, CCC '18, pages 3:1–3:19, Germany, 2018. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

BGGL01. B. Barak, O. Goldreich, S. Goldwasser, and Y. Lindell. Resettably-sound zero-knowledge and its applications. In *Proceedings 2001 IEEE International Conference on Cluster Computing*, pages 116–125, Oct 2001.

BP13. Nir Bitansky and Omer Paneth. On the impossibility of approximate obfuscation and applications to resettable cryptography. In *Proceedings of the Forty-fifth Annual ACM Symposium on Theory of Computing*, STOC '13, pages 241–250, New York, NY, USA, 2013. ACM.

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

CG14a. Mahdi Cheraghchi and Venkatesan Guruswami. Non-malleable coding against bit-wise and split-state tampering. In Yehuda Lindell, editor, *Theory of Cryptography*, pages 440–464, Berlin, Heidelberg, 2014. Springer Berlin Heidelberg.

CG14b. Mahdi Cheraghchi and Venkatesan Guruswami. Non-malleable coding against bit-wise and split-state tampering. In Yehuda Lindell, editor, *Theory of Cryptography*, pages 440–464, Berlin, Heidelberg, 2014. Springer Berlin Heidelberg.

CGGM00. Ran Canetti, Oded Goldreich, Shafi Goldwasser, and Silvio Micali. Resettable zero-knowledge (extended abstract). In *Proceedings of the Thirty-second Annual ACM Symposium on Theory of Computing*, STOC '00, pages 235–244, New York, NY, USA, 2000. ACM.

CGL16. Eshan Chattopadhyay, Vipul Goyal, and Xin Li. Non-malleable extractors and codes, with their many tampered extensions. In *Proceedings of the Forty-eighth Annual ACM Symposium on Theory of Computing*, STOC '16, pages 285–298, New York, NY, USA, 2016. ACM.

CL16. E. Chattopadhyay and X. Li. Explicit non-malleable extractors, multi-source extractors, and almost optimal privacy amplification protocols. In *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 158–167, Oct 2016.

Coh15. G. Cohen. Local correlation breakers and applications to three-source extractors and mergers. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, pages 845–862, Oct 2015.

Coh16a. G. Cohen. Making the most of advice: New correlation breakers and their applications. In *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 188–196, Oct 2016.

Coh16b. Gil Cohen. Non-Malleable Extractors - New Tools and Improved Constructions. In Ran Raz, editor, *31st Conference on Computational Complexity (CCC 2016)*, volume 50 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 8:1–8:29, Dagstuhl, Germany, 2016. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

Coh16c. Gil Cohen. Non-malleable extractors with logarithmic seeds. *Electronic Colloquium on Computational Complexity (ECCC)*, 23:30, 2016.

COP+14. Kai-Min Chung, Rafail Ostrovsky, Rafael Pass, Muthuramakrishnan Venkitasubramaniam, and Ivan Visconti. 4-round resettably-sound zero knowledge. In Yehuda Lindell, editor, *Theory of Cryptography*, pages 192–216, Berlin, Heidelberg, 2014. Springer Berlin Heidelberg.

COPV13. K. M. Chung, R. Ostrovsky, R. Pass, and I. Visconti. Simultaneous resettability from one-way functions. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, pages 60–69, Oct 2013.

CPS16. Kai-Min Chung, Rafael Pass, and Karn Seth. Non-black-box simulation from one-way functions and applications to resettable security. *SIAM Journal on Computing*, 45(2):415–458, 2016.

CRS14. Gil Cohen, Ran Raz, and Gil Segev. Nonmalleable extractors with short seeds and applications to privacy amplification. *SIAM Journal on Computing*, 43(2):450–476, 2014.

CZ16. Eshan Chattopadhyay and David Zuckerman. Explicit two-source extractors and resilient functions. In *Proceedings of the Forty-eighth Annual ACM Symposium on Theory of Computing*, STOC '16, pages 670–683, New York, NY, USA, 2016. ACM.

DGS09. Y. Deng, V. Goyal, and A. Sahai. Resolving the simultaneous resettability conjecture and a new non-black-box simulation strategy. In *2009 50th Annual IEEE Symposium on Foundations of Computer Science*, pages 251–260, Oct 2009.

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

DOPS04. Y. Dodis, Shien Jin Ong, M. Prabhakaran, and A. Sahai. On the (im)possibility of cryptography with imperfect randomness. *Foundations of Computer Science Annual Symposium on*, pages 196–205, 2004.

DRS04. Yevgeniy Dodis, Leonid Reyzin, and Adam Smith. Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. In Christian Cachin and Jan L. Camenisch, editors, *Advances in Cryptology - EUROCRYPT 2004*, pages 523–540, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.

DW09. Yevgeniy Dodis and Daniel Wichs. Non-malleable extractors and symmetric key cryptography from weak secrets. In *Proceedings of the Forty-first Annual ACM Symposium on Theory of Computing*, STOC '09, pages 601–610, New York, NY, USA, 2009. ACM.

GM11. V. Goyal and H. K. Maji. Stateless cryptographic protocols. In *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, pages 678–687, Oct 2011.

GO94. Oded Goldreich and Yair Oren. Definitions and properties of zero-knowledge proof systems. *Journal of Cryptology*, 7(1):1–32, Dec 1994.

GS09. Vipul Goyal and Amit Sahai. Resettably secure computation. In Antoine Joux, editor, *Advances in Cryptology - EUROCRYPT 2009*, pages 54–71, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.

GS19. Vipul Goyal and Yifan Song. Correlated-source extractors and cryptography with correlated-random tapes. Cryptology ePrint Archive, 2019.

GUV09. Venkatesan Guruswami, Christopher Umans, and Salil Vadhan. Unbalanced expanders and randomness extractors from parvaresh–vardy codes. *J. ACM*, 56(4):20:1–20:34, July 2009.

KLRZ08. Y. T. Kalai, X. Li, A. Rao, and D. Zuckerman. Network extractor protocols. In *2008 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 654–663, Oct 2008.

Li12a. X. Li. Non-malleable extractors, two-source extractors and privacy amplification. In *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, pages 688–697, Oct 2012.

Li12b. Xin Li. Design extractors, non-malleable condensers and privacy amplification. In *Proceedings of the Forty-fourth Annual ACM Symposium on Theory of Computing*, STOC '12, pages 837–854, New York, NY, USA, 2012. ACM.

Li15. Xin Li. Non-malleable condensers for arbitrary min-entropy, and almost optimal protocols for privacy amplification. In Yevgeniy Dodis and Jesper Buus Nielsen, editors, *Theory of Cryptography*, pages 502–531, Berlin, Heidelberg, 2015. Springer Berlin Heidelberg.

Li16. X. Li. Improved two-source extractors, and affine extractors for polylogarithmic entropy. In *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 168–177, Oct 2016.

Li17. Xin Li. Improved non-malleable extractors, non-malleable codes and independent source extractors. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2017, pages 1144–1156, New York, NY, USA, 2017. ACM.

MW97. Ueli Maurer and Stefan Wolf. Privacy amplification secure against active adversaries. In Burton S. Kaliski, editor, *Advances in Cryptology — CRYPTO '97*, pages 307–321, Berlin, Heidelberg, 1997. Springer Berlin Heidelberg.

SSZ95. Michael Saks, Aravind Srinivasan, and Shiyu Zhou. Explicit dispersers with polylog degree. In *Proceedings of the Twenty-seventh Annual ACM Symposium on Theory of Computing*, STOC '95, pages 479–488, New York, NY, USA, 1995. ACM.

VV85. U. V. Vazirani and V. V. Vazirani. Random polynomial time is equal to slightly-random polynomial time. In *26th Annual Symposium on Foundations of Computer Science (sfcs 1985)*, pages 417–428, Oct 1985.

Zuc96. D. Zuckerman. Simulating bpp using a general weak random source. *Algorithmica*, 16(4):367–391, Oct 1996.