On Secret Sharing, Randomness, and Random-less Reductions for Secret Sharing

Divesh Aggarwal[*] Eldon Chung[†] Maciej Obremski[‡] João Ribeiro[§]

Abstract

Secret-sharing is one of the most basic and oldest primitives in cryptography, introduced by Shamir and Blakely in the 70s. It allows to strike a meaningful balance between availability and confidentiality of secret information. It has a host of applications most notably in threshold cryptography and multi-party computation. All known constructions of secret sharing (with the exception of those with a pathological choice of parameters) require access to uniform randomness. In practice, it is extremely challenging to generate a source of uniform randomness. This has led to a large body of research devoted to designing randomized algorithms and cryptographic primitives from imperfect sources of randomness.

Motivated by this, 15 years ago, Bosley and Dodis asked whether it is even possible to build 2-out-of-2 secret sharing without access to uniform randomness. In this work, we make progress towards resolving this question.

We answer this question for secret sharing schemes with important additional properties, i.e., either leakage-resilience or non-malleability. We prove that, unfortunately, for not too small secrets, it is impossible to construct any of 2-out-of-2 leakage-resilient secret sharing or 2-out-of-2 non-malleable secret sharing without access to uniform randomness.

Given that the problem whether 2-out-of-2 secret sharing requires uniform randomness has been open for a long time, it is reasonable to consider intermediate problems towards resolving the open question. In a spirit similar to NP-completeness, we study how the existence of a t-out-of-n secret sharing without access to uniform randomness is related to the existence of a t'-out-of-n' secret sharing without access to uniform randomness for a different choice of the parameters t, n, t', n'.

1 Introduction

Secret Sharing and its Variants. Secret-sharing is one of the most basic and oldest primitives in cryptography, introduced by Shamir and Blakely in the 70s. It allows to strike a meaningful balance between availability and confidentiality of secret information. It has a host of applications most notably in threshold cryptography and multi-party computation (see, for example, [CDN15]). In a *secret sharing scheme*, a dealer who holds a secret s chosen from a domain $\mathcal{M}$ can compute a set of *shares* by evaluating a randomized function on s which we write as $\mathbf{Share}(s) = (s_1, \ldots, s_n)$. A t-out-of-n secret sharing scheme ensures that any t (out of n) shares are sufficient to recover the secret s, but any $t-1$ shares do not reveal any information about the secret s.

After its introduction, several variants of secret sharing have been suggested that address the problem of authenticity of the secret: we want to guarantee that we reconstruct the original value,

[*] National University of Singapore. dcsdiva@nus.edu.sg

[†] National University of Singapore. eldon.chung@u.nus.edu

[‡] National University of Singapore. obremski.math@gmail.com

[§] Imperial College London. j.lourenco-ribeiro17@imperial.ac.uk

even if not all players are honest. One such variant is *robust* secret-sharing where the dealer is honest, but some unqualified set of share holders are malicious and may return incorrect shares.

In all these older schemes, the adversary is of the classic type that completely corrupts a certain subset of the players in the protocol, either to steal information or to corrupt data. Whereas the players who are not corrupt are "completely honest". In many scenarios, however, this may not be the most realistic model of attacks. Instead, it may make more sense to assume that the adversary will try to attack all share holders, and will have some partial success in all or most of the cases.

For the case of attacks against confidentiality, we can model this as *leakage-resilient* secret-sharing, where the adversary is allowed to specify a leakage function Leak and will be told the value $\mathsf{Leak}(s_1, ..., s_n)$. Then, under certain restrictions on Leak, we want that the adversary learns essentially nothing about s. Typically, so called *local leakage* is considered, where $\mathsf{Leak}(s_1, ..., s_n) = (\mathsf{Leak}_1(s_1), ..., \mathsf{Leak}_n(s_n))$ for local leakage functions Leak_i with bounded output size. This makes sense in a scenario where shares are stored in physically separated locations. It is known that some secret-sharing schemes are naturally leakage-resilient against local leakage whereas others are not [BDIR19]. Boyle, Goldwasser, and Kalai [BGK14] showed how to construct (locally) leakage-resilient verifiable secret sharing for threshold access structures.

The case of attacks that try to corrupt the secret can be modeled as *non-malleable* secret sharing where the adversary specifies tampering functions $f_1, f_2, \ldots, f_n$ which acts on the shares, and then the reconstruction algorithm is applied to the tampered shares, $f_1(s_1), \ldots, f_n(s_n)$. The requirement, simplistically speaking, is that either the original secret is reconstructed or it is destroyed, i.e., the reconstruction result is unrelated to the original secret.

Replacing uniformly random sources with weak sources. Randomization has been proven to play a very important role in many areas of computer science. It is often the case that randomized algorithms and protocols are much more efficient than their deterministic counterparts. Moreover, randomness is extremely important in cryptography: secret keys must be random, and many cryptographic tasks require fresh randomness for every use. Randomized algorithms are designed under the assumption that the computer has access to perfectly random and independent bits (i.e., a sequence of unbiased coin tosses). In practice this sequence is generated from some physical source of randomness like electromagnetic noise, measuring user dependent behavior (e.g., the time taken by the key strokes of users), or biometrics. While these sources have some inherent randomness in the sense that they have entropy, a sample from such sources does not contain independent random bits. Additionally, in some situations the original source might be close to being truly random, but an attacker might be able to learn some partial information about the source through some side-channel attacks thereby effectively making the source imperfect from the perspective of achieving security against this attacker. The difficulty in working with such imperfect randomness sources not only arises from the fact that they are not uniformly random but also because the exact distribution of these sources is unknown. One can at best assume that they satisfy some minimal randomness property, for example that none of the outcomes is highly likely.

The best one can hope for is to deterministically extract a nearly perfect random string for direct usage in the desired application. Many examples of deterministically extractable sources are known in the literature like von-Neuman sources [vN51], bit-fixing sources [CGH$^+$85, KZ07, GRS06], affine-sources [Bou07, GS08, BKS$^+$10, Rao09, DG10, Yeh11, Li11, Gab11], or other efficiently generated or efficiently recognizable sources [Eli72, Blu86, Vaz87, BBR88, LLS89, AL93, TV00, CDH$^+$00, DSS01, DGW09, Dvi12, Sha11, KvMS09, KRVZ11, BGLZ15, CL16, BEG17]. However, the most natural min-entropy source, i.e. the sources where the only guarantee is a lower bound on the min-entropy of the source[1] are known to be non-extractable [SV86, CG88, Zuc96] even if they have high min-entropy. This immediately leads to the question of whether perfect randomness is essential for applications in the sense that the source of randomness must be extractable. It has been shown [VV85, SV86, CG88, DOPS04] that for simulating probabilistic algorithms and for interactive protocols, min-entropy sources are sufficient.

This brings us to our main question.

Question 1. *Does secret sharing, or any of its variants such as leakage-resilient or non-malleable secret sharing* require *access to extractable randomness?*

This question was first asked by Bosley and Dodis [BD07] (for 2-out-of-2 secret sharing) and, to this day, it remains open.

Prior work has considered this question for other cryptographic primitives. In particular, early work by McInnes and Pinkas [MP91] showed that min-entropy sources and Santha-Vazirani sources are insufficient for information-theoretic private-key encryption of even 1-bit messages. This negative result was later extended to computationally secure private-key encryption by Dodis, Ong, Prabhakaran, and Sahai [DOPS04], and was complemented by Dodis and Spencer [DS02], who showed that, in fact, non-extractable randomness *is* sufficient for information-theoretic private-key encryption of 1-bit messages. Later, the picture was completed by groundbreaking work of Bosley and Dodis [BD07]: Encryption schemes using d bits of randomness and encrypting messages of size $b > \log d$ require extractable randomness, while those encrypting messages of size $b < \log d - \log\log d - 1$ do not.

Information-theoretic private-key encryption can be seen as a special case of 2-out-of-2 secret sharing, with one share being the private key and the other share being the ciphertext (this was already noted in [DPP06, BD07]).

Random-less Reductions for Secret Sharing. Given that the problem of whether 2-out-of-2 secret sharing requires extractable randomness has been open for 15 years, it is reasonable to consider intermediate problems towards resolving the open question. In a spirit similar to Computational Complexity, we consider how the question whether t out of n secret sharing requires extractable randomness is related to the same question for a different choice of the parameters t, n i.e.,

Question 2. *Given t, n, t', n', does the fact that t-out-of-n secret sharing require extractable randomness imply that t'-out-of-n' secret sharing require extractable randomness?*

A natural approach towards resolving this question is to try to construct a t-out-of-n secret sharing scheme from a t'-out-of-n' secret sharing scheme in a black-box manner without any additional randomness. Intuitively, since we don't have access to any additional randomness, it seems that the most obvious strategy to achieve such reductions is to choose n subsets of the set of n' shares in such a way that any t out of these n subsets contain at least t' out of the original n' shares and any $t-1$ subsets contain at most $t'-1$ of the original n' shares. In particular, there is a trivial reduction when $t = n = 2$ that chooses the first subset to contain the first of the n' shares, and the second subset to contain any $t'-1$ of the remaining shares. This shows the completeness of the extractability of 2-out-of-2 secret sharing with respect to these reductions. Such reductions can be formalized via distribution designs [SW18].

[1] A source is said to have min-entropy k if the probability that it takes any fixed value is upper bounded by 2^{-k}.

1.1 Our results

In this work, we make progress on both Question 1 and Question 2. Before we proceed to discuss our results, we formalize the notions of an extractable class of randomness sources and threshold secret sharing.

Definition 1 (Extractable class of sources)**.** *We say a class of randomness sources $\mathcal{Y}$ over $\{0,1\}^d$ is* (δ, m)-extractable *if there exists a deterministic function* $\mathsf{Ext} : \{0,1\}^d \to \{0,1\}^m$ *such that*[2] $\mathsf{Ext}(Y) \approx_\delta U_m$ *for every $Y \in \mathcal{Y}$, where U_m denotes the uniform distribution over $\{0,1\}^m$.*

Note that we may consider the support of all sources in $\mathcal{Y}$ to be contained in some set $\{0,1\}^d$ without loss of generality. Since we will be interested in studying the quality of randomness used by secret sharing schemes, we make the class of randomness sources allowed for a secret sharing scheme explicit in the definition of t-out-of-n threshold secret sharing below.

Definition 2 (Threshold secret sharing scheme)**.** *We say a tuple* $(\mathbf{Share}, \mathbf{Rec}, \mathcal{Y})$ *with* $\mathbf{Share} : \{0,1\}^b \times \{0,1\}^d \to \left(\{0,1\}^\ell\right)^n$ *and* $\mathbf{Rec} : \{0,1\}^* \to \{0,1\}^b$ *deterministic algorithms and $\mathcal{Y}$ a class of randomness sources over $\{0,1\}^d$ is a* (t, n, ε)-secret sharing scheme (for b-bit messages using d bits of randomness) *if for every randomness source $Y \in \mathcal{Y}$ the following hold:*

1. *If $\mathcal{T} \subseteq [n]$ satisfies $|\mathcal{T}| \geq t$ (i.e., $\mathcal{T}$ is* authorized*), then* $\Pr_Y[\mathbf{Rec}(\mathbf{Share}(x, Y)_\mathcal{T}) = x] = 1$ *for every $x \in \{0,1\}^b$;*

2. *If $\mathcal{T} \subseteq [n]$ satisfies $|\mathcal{T}| < t$ (i.e., $\mathcal{T}$ is* unauthorized*), then for any $x, x' \in \{0,1\}^b$ we have*

$$\mathbf{Share}(x, Y)_\mathcal{T} \approx_\varepsilon \mathbf{Share}(x', Y)_\mathcal{T},$$

where $\mathbf{Share}(x, Y)_\mathcal{T}$ denotes the shares of parties $i \in \mathcal{T}$.

1.1.1 Leakage-resilient 2-out-of-2 secret sharing requires extractable randomness

As our first contribution, we settle Question 1 for the important sub-class of *leakage-resilient* 2-out-of-2 secret sharing. Intuitively, we consider 2-out-of-2 secret sharing schemes with the additional property that the adversary learns almost nothing about the message when they obtain bounded information from each share. More formally, we have the following definition.

Definition 3 (Leakage-resilient secret sharing scheme)**.** *We say a tuple* $(\mathbf{Share}, \mathbf{Rec}, \mathcal{Y})$ *with* $\mathbf{Share} : \{0,1\}^b \times \{0,1\}^d \to \left(\{0,1\}^\ell\right)^2$ *and* $\mathbf{Rec} : \{0,1\}^* \to \{0,1\}^b$ *deterministic algorithms and $\mathcal{Y}$ a class of randomness sources over $\{0,1\}^d$ is an* $(\varepsilon_1, \varepsilon_2)$-leakage-resilient secret sharing scheme (for b-bit messages using d bits of randomness) *if* $(\mathbf{Share}, \mathbf{Rec}, \mathcal{Y})$ *is a $(t = 2, n = 2, \varepsilon_1)$-secret sharing scheme and the following additional property is satisfied: For any two messages $x, x' \in \{0,1\}^b$ and randomness source $Y \in \mathcal{Y}$, let* $(\mathsf{Sh}_1, \mathsf{Sh}_2) = \mathbf{Share}(x, Y)$ *and* $(\mathsf{Sh}'_1, \mathsf{Sh}'_2) = \mathbf{Share}(x', Y)$*. Then, for any leakage functions $f, g : \{0,1\}^\ell \to \{0,1\}$ it holds that*

$$f(\mathsf{Sh}_1), g(\mathsf{Sh}_2) \approx_{\varepsilon_2} f(\mathsf{Sh}'_1), g(\mathsf{Sh}'_2).$$

Leakage-resilient secret sharing has received significant attention recently, with several constructions and leakage models being analyzed [BDIR21, ADN$^+$19, KMS19, SV19, CGG$^+$20, LCG$^+$20, MPSW20]. Comparatively, Definition 3 considers a significantly weaker notion of leakage-resilience

[2] We use the notation $X \approx_\delta Y$ to denote the fact that $\Delta(X; Y) \leq \delta$, where $\Delta(\cdot; \cdot)$ corresponds to statistical distance (see Definition 8).

than all works just mentioned. In particular, we do not require leakage-resilience to hold even when the adversary has full access to one of the shares on top of the leakage. This means that our results are widely applicable. Roughly speaking, we prove that every leakage-resilient secret sharing scheme for b-bit messages either requires a huge number of bits of randomness, or we can extract several bits of perfect randomness with low error from its underlying class of randomness sources. More formally, we prove the following.

Theorem 1. *Let* $(\mathbf{Share}, \mathbf{Rec}, \mathcal{Y})$ *be an* $(\varepsilon_1, \varepsilon_2)$*-leakage-resilient secret sharing scheme for* b*-bit messages. Then, either:*

1. *The scheme uses* $d \geq \min\left(2^{\Omega(b)}, (1/\varepsilon_2)^{\Omega(1)}\right)$ *bits of randomness, or;*
2. *The class of randomness sources* $\mathcal{Y}$ *is* (δ, m)*-extractable with* $\delta \leq \max\left(2^{-\Omega(b)}, \varepsilon_2^{\Omega(1)}\right)$ *and* $m = \Omega(\min(b, \log(1/\varepsilon_2)))$*. Moreover, if* $\mathbf{Share}$ *is an efficient algorithm, then* $\mathcal{Y}$ *is efficiently* (δ, m)*-extractable.*

An important corollary of Theorem 1 is that every efficient negligible-error leakage-resilient secret sharing scheme requires extractable randomness with negligible error.

Corollary 1. *If* $(\mathbf{Share}, \mathbf{Rec}, \mathcal{Y})$ *is an* $(\varepsilon_1, \varepsilon_2)$*-leakage-resilient secret sharing scheme for* b*-bit messages running in time* $\mathrm{poly}(b)$ *with* $\varepsilon_2 = \mathsf{negl}(b)$*,*[3] *it follows that* $\mathcal{Y}$ *is* (δ, m)*-extractable with* $\delta = \mathsf{negl}(b)$ *and* $m = \Omega(\min(b, \log(1/\varepsilon_2)))$*.*

Split-state non-malleable coding requires extractable randomness. Non-malleable coding, introduced by Dziembowski, Pietrzak, and Wichs [DPW18], is another recent notion which has attracted much attention, in particular regarding the *split-state* setting (see [AO20] and references therein). Informally, a split-state non-malleable code has the guarantee that if an adversary is allowed to split a codeword in half and tamper with each half arbitrarily but separately, then the tampered codeword either decodes to the same message, or the output of the decoder is nearly independent of the original message. More formally, we have the following definition.

Definition 4 (Split-state non-malleable code [DPW18]). *We say a tuple* $(\mathbf{Enc}, \mathbf{Dec}, \mathcal{Y})$ *with* $\mathbf{Enc} : \{0,1\}^b \times \{0,1\}^d \to (\{0,1\}^\ell)^2$ *and* $\mathbf{Dec} : (\{0,1\}^\ell)^2 \to \{0,1\}^b \cup \{\perp\}$ *deterministic algorithms and* $\mathcal{Y}$ *a class of randomness sources is a* (split-state) ε-non-malleable code *if the following holds for every randomness source* $Y \in \mathcal{Y}$*:*

1. $\Pr[\mathbf{Dec}(\mathbf{Enc}(x, Y)) = x] = 1$ *for all* $x \in \{0,1\}^b$*;*
2. *For tampering functions* $f, g : \{0,1\}^\ell \to \{0,1\}^\ell$*, denote by* $\mathsf{Tamp}_x^{f,g}$ *the tampering random experiment which computes* $(L, R) = \mathbf{Enc}(x, Y)$ *and outputs* $\mathbf{Dec}(f(L), g(R))$*. Then, for any tampering functions* f *and* g *there exists a distribution* $D^{f,g}$ *over* $\{0,1\}^b \cup \{\perp, \mathsf{same}^*\}$ *such that*
$$\mathsf{Tamp}_x^{f,g} \approx_\varepsilon \mathsf{Sim}_x^{f,g}$$
for all $x \in \{0,1\}^b$*, where* $\mathsf{Sim}_x^{f,g}$ *denotes the random experiment which samples* z *according to* $D^{f,g}$ *and outputs* z *if* $z \neq \mathsf{same}^*$ *and* x *if* $z = \mathsf{same}^*$*.*

The notion of non-malleable code in the split-state model is equivalent to the notion of a **2-out-of-2 non-malleable secret sharing scheme***.*

[3] By $\varepsilon_2 = \mathsf{negl}(b)$, we mean that $\varepsilon_2 = o(1/b^c)$ for every constant $c > 0$ as $b \to \infty$.

It is known by [AKO17, Lemmas 3 and 4] that every ε-non-malleable code $(\mathbf{Enc}, \mathbf{Dec}, \mathcal{Y})$ for b-bit messages is also a $(2\varepsilon, \varepsilon)$-leakage-resilient secret sharing scheme, provided $b \geq 3$ and $\varepsilon < 1/20$. Combining this observation with Theorem 1 yields the following corollary, which states that every split-state non-malleable code either uses a huge number of bits of randomness, or requires extractable randomness with low error and large output length.

Corollary 2. *Let $(\mathbf{Enc}, \mathbf{Dec}, \mathcal{Y})$ be an ε-non-malleable code (i.e., 2-out-of-2 ε-non-malleable secret sharing scheme) for b-bit messages with $b \geq 3$ and $\varepsilon < 1/20$. Then, either:*

1. *The scheme uses $d \geq \min\left(2^{\Omega(b)}, (1/\varepsilon)^{\Omega(1)}\right)$ bits of randomness, or;*
2. *The class of randomness sources $\mathcal{Y}$ is (δ, m)-extractable with $\delta \leq \max\left(2^{-\Omega(b)}, \varepsilon^{\Omega(1)}\right)$ and $m = \Omega(\min(b, \log(1/\varepsilon)))$. Moreover, if $\mathbf{Enc}$ is an efficient algorithm, then $\mathcal{Y}$ is efficiently (δ, m)-extractable.*

As a result, an analogous version of Corollary 1 also holds for split-state non-malleable coding. This resolves Question 1 for 2-out-of-2 non-malleable secret sharing.

1.1.2 Random-less Reductions for Secret Sharing

In this section, we mention our contribution towards resolving Question 2. We focus on the following complementary scenario: Suppose we have proved that all (t, n, ε)-secret sharing schemes for b-bit messages using d bits of randomness require a (δ, m)-extractable class of randomness sources. It is then natural to wonder whether such a result can be bootstrapped to conclude that all (t', n', ε)-secret sharing schemes for the same message length b and number of randomness bits d also require (δ, m)-extractable randomness, for different threshold t' and number of parties n'. A natural approach is to set up general *black-box reductions* between different types of secret sharing which, crucially, do not use extra randomness. In fact, if we can obtain from a (t', n', ε)-secret sharing scheme $(\mathbf{Share}', \mathbf{Rec}', \mathcal{Y})$ another (t, n, ε)-secret sharing scheme $(\mathbf{Share}, \mathbf{Rec}, \mathcal{Y})$ for b-bit messages which uses the same class of randomness sources $\mathcal{Y}$, then our initial assumption would allow us to conclude that $\mathcal{Y}$ is (δ, m)-extractable.

Remarkably, we are able to obtain the desired reductions for a broad range of parameters by exploiting a connection to the construction of combinatorial objects called *distribution designs*, a term coined by Stinson and Wei [SW18] for the old technique of devising a new secret sharing scheme by giving multiple shares of the original scheme to each party. Surprisingly, although these objects have roots going back to early work on secret sharing [BL88], they have not been the subject of a general study. In this work, we obtain general and simple constructions of, and bounds for, distribution designs, which are tight in certain parameter regimes. We give two examples of reductions we derive from these results.

Corollary 3 (Informal)**.** *If every $(t = 2, n, \varepsilon)$-secret sharing scheme for b-bit messages using d bits of randomness requires a (δ, m)-extractable class of randomness sources, then so does every (t', n', ε)-secret sharing scheme for b-bit messages using d bits of randomness whenever $n \leq \binom{n'}{t'-1}$. Moreover, this is the best distribution-design-based reduction possible with $t = 2$.*

Corollary 4 (Informal)**.** *If every (t, n, ε)-secret sharing scheme for b-bit messages using d bits of randomness requires a (δ, m)-extractable class of randomness sources, then so does every $(t' = n', n', \varepsilon)$-secret sharing scheme for b-bit messages using d bits of randomness whenever $n' \leq \binom{n}{t-1}$. Moreover, this is the best distribution-design-based reduction possible with $t' = n'$.*

1.2 Other related work

Besides the results already discussed above for private-key encryption and secret sharing, the possibility of realizing other primitives using certain classes of imperfect randomness sources has also been studied. Non-extractable randomness is known to be sufficient for message authentication [MW97, DS02], signature schemes [DOPS04, ACM$^+$14], differential privacy [DLMV12, DY15, YL18], secret-key agreement [ACM$^+$14], identification protocols [ACM$^+$14], and interactive proofs [DOPS04]. On the other hand, Santha-Vazirani sources are insufficient for bit commitment, secret sharing, zero knowledge, and two-party computation [DOPS04], and in some cases this negative result even holds for Santha-Vazirani sources with efficient tampering procedures [ACM$^+$14].

In other directions, the security loss incurred by replacing uniform randomness by imperfect randomness was studied in [BBN$^+$09, BKMR15], and the scenario where a perfect common reference string is replaced by certain types of imperfect randomness has also been considered [CPs07, AOR$^+$20].

1.3 Technical overview

1.3.1 Leakage-resilient secret sharing requires extractable randomness

We present a high-level overview of our approach towards proving Theorem 1. Recall that our goal is to show that if $(\mathbf{Share}, \mathbf{Rec}, \mathcal{Y})$ is an $(\varepsilon_1, \varepsilon_2)$-leakage-resilient secret sharing for b-bit messages using d bits of randomness, then there exists a deterministic function $\mathsf{Ext} : \{0,1\}^d \to \{0,1\}^m$ such that $\mathsf{Ext}(Y) \approx_\delta U_m$ for all sources $Y \in \mathcal{Y}$, provided that the number of randomness bits d used is not huge.

Our candidate extractor Ext works as follows on input some $y \in \{0,1\}^d$:

1. Compute $(\mathsf{Sh}_1, \mathsf{Sh}_2) = \mathbf{Share}(0^b, y) \in \{0,1\}^\ell \times \{0,1\}^\ell$;
2. For appropriate leakage functions $f, g : \{0,1\}^\ell \to \{0,1\}^s$, compute $(f(\mathsf{Sh}_1), g(\mathsf{Sh}_2))$;
3. For an appropriate function $h : \{0,1\}^{2s} \to \{0,1\}^m$, output $\mathsf{Ext}(y) = h(f(\mathsf{Sh}_1), g(\mathsf{Sh}_2))$.

The proof of Theorem 1 follows from an analysis of this candidate construction, and we show the existence of appropriate functions f, g, and h via the probabilistic method. Note that the number of sources in $\mathcal{Y}$ may be extremely large. Consequently, our first step, which is similar in spirit to the first step of the related result for private-key encryption in [BD07], is to exploit the leakage-resilience of the scheme in question to show that it suffices to focus on a restricted family to prove the desired result. More precisely, it suffices to show the existence of functions f, g, and h as above satisfying

$$h(f(Z_1), g(Z_2)) \approx_{\delta'} U_m, \tag{1}$$

with δ' an appropriate error parameter, for all $(Z_1, Z_2) \in \mathcal{Z}$ defined as

$$\mathcal{Z} = \{\mathbf{Share}(U_b, y) : y \in \{0,1\}^d\},$$

which contains at most 2^d distributions. Our analysis then proceeds in three steps:

1. We show that each $(Z_1, Z_2) \in \mathcal{Z}$ is close in statistical distance to a convex combination of joint distributions $(D_{1,i}, D_{2,i})$ with the property that $\mathbf{H}_\infty(D_{1,i}) + \mathbf{H}_\infty(D_{2,i})$ is sufficiently large for all i, where $\mathbf{H}_\infty(\cdot)$ denotes the min-entropy of a distribution;

2. Exploiting the previous step, we prove that if we pick f and g uniformly at random, then with high probability over this choice it holds that the joint distribution $(f(Z_1), g(Z_2))$ is close in statistical distance to a high min-entropy distribution;

3. A well known, standard application of the probabilistic method then shows that a uniformly random function h will extract many perfectly random bits from $(f(Z_1), g(Z_2))$ with high probability over the choice of h.

While this proves that there exist functions f, g, and h such that (1) holds for a given $(Z_1, Z_2) \in \mathcal{Z}$, we need (1) to be true simultaneously for all $(Z_1, Z_2) \in \mathcal{Z}$. We resolve this by employing a union bound over the at most 2^d distributions in $\mathcal{Z}$. Therefore, if d is not extremely large, we succeed in showing the existence of appropriate functions f, g, and h, and the desired result follows. More details can be found in Section 3.

1.3.2 Random-less reductions for secret sharing

In this section, we define distribution designs and briefly discuss how they can be used to provide the desired black-box reductions between different types of threshold secret sharing, in particular Corollaries 3 and 4. Intuitively, a (t, n, t', n')-distribution design distributes shares $(\mathsf{Sh}_1, \mathsf{Sh}_2, \dots, \mathsf{Sh}_{n'})$ of some (t', n', ε)-secret sharing scheme into subsets of shares $\mathcal{S}_1, \dots, \mathcal{S}_n$, with the property that $(\mathcal{S}_1, \dots, \mathcal{S}_n)$ are now shares of a (t, n, ε)-secret sharing scheme. More formally, we have the following definition, which also appears in [SW18].

Definition 5 (Distribution design). *We say a family of sets $\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_n \subseteq [n']$ is a (t, n, t', n')-*distribution design *if for every $\mathcal{T} \subseteq [n]$ it holds that*

$$\left| \bigcup_{i \in \mathcal{T}} \mathcal{D}_i \right| \geq t'$$

if and only if $|\mathcal{T}| \geq t$.

Given a (t, n, t', n')-distribution design $\mathcal{D}_1, \dots, \mathcal{D}_n \subseteq [n']$, it is clear how to set up a black-box reduction without extra randomness from (t', n', ε)-secret sharing to (t, n, ε)-secret sharing: If $(\mathbf{Share}', \mathbf{Rec}', \mathcal{Y})$ is an arbitrary (t', n', ε)-secret sharing scheme for b-bit messages, we can obtain a (t, n, ε)-secret sharing scheme $(\mathbf{Share}, \mathbf{Rec}, \mathcal{Y})$ for b-bit messages by defining

$$\mathbf{Share}(x, y)_i = \mathbf{Share}'(x, y)_{\mathcal{D}_i}$$

for each $i \in [n]$, and

$$\mathbf{Rec}(\mathbf{Share}(x, y)_{\mathcal{T}}) = \mathbf{Rec}'\Big(\mathbf{Share}'(x, y)_{\bigcup_{i \in \mathcal{T}} \mathcal{D}_i}\Big)$$

for each $\mathcal{T} \subseteq [n]$. The following lemma is then straightforward from the definitions of threshold secret sharing and distribution designs, and this construction.

Lemma 1. *If every (t, n, ε)-secret sharing scheme for b-bit messages using d bits of randomness requires (δ, m)-extractable randomness and there exists a (t, n, t', n')-distribution design, then so does every (t', n', ε)-secret sharing scheme for b-bit messages using d bits of randomness.*

Details of our constructions of distribution designs and associated bounds can be found in Section 4. The black-box reductions then follow immediately by combining these constructions with Lemma 1.

1.4 Open questions

We obtain distribution designs for a wide variety of parameters, but for some of these constructions we could not prove optimality or find a better construction. We leave this as an open question. A naturally related question is whether there is an alternative approach to obtain a random-less reduction for secret sharing that does not use distribution designs.

Finally, we really hope this work further motivates research on the main open question whether 2-out-of-2 secret sharing (or even t-out-of-n secret sharing for any t and n) requires extractable randomness.

2 Preliminaries

2.1 Notation

Random variables are denoted by uppercase letters such as X, Y, and Z, and we write U_m for the uniform distribution over $\{0,1\}^m$. We usually denote sets by uppercase calligraphic letters like $\mathcal{S}$ and $\mathcal{T}$, and write $[n]$ for the set $\{1,2,\ldots,n\}$. Given a vector $x \in \mathcal{S}^n$ and set $\mathcal{T} \subseteq [n]$, we define $x_{\mathcal{T}} = (x_i)_{i\in\mathcal{T}}$. We denote the $\mathbb{F}_2$-inner product between vectors $x, y \in \{0,1\}^n$ by $\langle x, y\rangle$. All logarithms in this paper are taken with respect to base 2.

2.2 Probability theory

In this section, we introduce basic notions from probability theory that will be useful throughout this work.

Definition 6 (Min-entropy)**.** *The* min-entropy *of a random variable X on a set $\mathcal{X}$, denoted by $\mathbf{H}_\infty(X)$, is defined as*

$$\mathbf{H}_\infty(X) = -\log \max_{x\in\mathcal{X}} \Pr[X = x].$$

Definition 7 ((n,k)-source)**.** *We say a random variable X supported over $\{0,1\}^n$ is an (n,k)-*source *if $\mathbf{H}_\infty(X) \geq k$. When the support of the random variable is clear from context we may instead say k-source. Moreover, we say X is* flat *if it is uniformly distributed over a subset of $\{0,1\}^n$.*

Definition 8. *The* statistical distance between random variables X and Y *over a set $\mathcal{X}$, denoted by $\Delta(X,Y)$, is defined as*

$$\Delta(X,Y) = \max_{\mathcal{S}\subseteq\mathcal{X}} |\Pr[X \in \mathcal{S}] - \Pr[Y \in \mathcal{S}]| = \frac{1}{2}\sum_{x\in\mathcal{X}} |\Pr[X = x] - \Pr[Y = x]|.$$

*Moreover, we say that X and Y are ε-*close*, denoted by $X \approx_\varepsilon Y$, if $\Delta(X,Y) \leq \varepsilon$, and ε-*far *if this does not hold.*

The following lemma is a version of the well-known XOR lemma (see [Gol11] for a detailed exposition of these types of results).

Lemma 2 (XOR Lemma)**.** *If X and Y are distributions supported on $\{0,1\}^t$ such that*

$$\langle a, X\rangle \approx_\varepsilon \langle a, Y\rangle$$

for all non-zero vectors $a \in \{0,1\}^t$, then

$$X \approx_{\varepsilon'} Y$$

for $\varepsilon' = 2^{t/2}\varepsilon$.

We end this section with a standard lemma stemming from a straightforward application of the probabilistic method, which states that, with high probability, a random function extracts almost perfect randomness from a fixed source with sufficient min-entropy. By a union bound, this result also implies that a random function is a great extractor for all sufficiently small classes of flat sources (and convex combinations thereof), an observation we will exploit later on.

Lemma 3. *Fix an (n,k)-source X. Then, for every $\varepsilon > 0$ it holds that a uniformly random function $F : \{0,1\}^n \to \{0,1\}^m$ with $m \leq k - 2\log(1/\varepsilon)$ satisfies $F(X) \approx_\varepsilon U_m$ with probability at least $1 - 2e^{-\varepsilon^2 2^k}$ over the choice of F.*

Proof. See Appendix A. □

The following extension of Lemma 3, stating that a random function condenses weak sources with high probability, will also be useful.

Lemma 4. *Fix an (n,k)-source X. Then, for every $\varepsilon > 0$ it holds that a uniformly random function $F : \{0,1\}^n \to \{0,1\}^m$ satisfies $F(X) \approx_\varepsilon W$ for some W such that $\mathbf{H}_\infty(W) \geq \min(m, k - 2\log(1/\varepsilon))$ with probability at least $1 - 2e^{-\varepsilon^2 2^k}$ over the choice of F.*

Proof. For $m' = \min(m, k - 2\log(1/\varepsilon))$, let $F' : \{0,1\}^n \to \{0,1\}^{m'}$ be the restriction of F to its first m' bits. Then, Lemma 3 ensures that $F'(X) \approx_\varepsilon U_{m'}$ with probability at least $1 - 2e^{-\varepsilon^2 2^k}$ over the choice of F. Via a coupling argument, this implies that $F(X) \approx W$ for some W with $\mathbf{H}_\infty(W) \geq m'$. □

2.3 Amplifying leakage-resilience

Recall the definition of leakage-resilient secret sharing from Definition 3 already discussed in Section 1. The following lemma states that every secret sharing scheme withstanding 1 bit of leakage also withstands $t > 1$ bits of leakage from each share, at the cost of an increase in statistical error.

Lemma 5. *Let $(\mathbf{Share}, \mathbf{Rec}, \mathcal{Y})$ be an $(\varepsilon_1, \varepsilon_2)$-leakage-resilient secret sharing scheme. Then, for all secrets $x, x' \in \{0,1\}^b$, randomness source $Y \in \mathcal{Y}$, and functions $f, g : \{0,1\}^\ell \to \{0,1\}^t$ we have*

$$f(\mathsf{Sh}_1), g(\mathsf{Sh}_2) \approx_{\varepsilon'} f(\mathsf{Sh}_1'), g(\mathsf{Sh}_2')$$

with $\varepsilon' = 2^t \varepsilon_2$, where $(\mathsf{Sh}_1, \mathsf{Sh}_2) = \mathbf{Share}(x, Y)$ and $(\mathsf{Sh}_1', \mathsf{Sh}_2') = \mathbf{Share}(x', Y)$.

Proof. Fix arbitrary secrets $x, x' \in \{0,1\}^b$ and a randomness source $Y \in \mathcal{Y}$, and define $(\mathsf{Sh}_1, \mathsf{Sh}_2) = \mathbf{Share}(x, Y)$ and $(\mathsf{Sh}_1', \mathsf{Sh}_2') = \mathbf{Share}(x', Y)$. Suppose that there exist functions $f, g : \{0,1\}^\ell \to \{0,1\}^t$ such that the distributions $(f(\mathsf{Sh}_1), g(\mathsf{Sh}_2))$ and $(f(\mathsf{Sh}_1'), g(\mathsf{Sh}_2'))$ are $(\varepsilon' = 2^t\varepsilon_2)$-far. Then, the XOR lemma implies that there is a non-zero vector $a \in \{0,1\}^{2t}$, which we may write as $a = (a^{(1)}, a^{(2)})$ for $a^{(1)}, a^{(2)} \in \{0,1\}^t$, such that the distributions

$$\langle a, (f(\mathsf{Sh}_1), g(\mathsf{Sh}_2)) \rangle = \langle a^{(1)}, f(\mathsf{Sh}_1) \rangle + \langle a^{(2)}, g(\mathsf{Sh}_2) \rangle$$

and

$$\langle a, (f(\mathsf{Sh}_1'), g(\mathsf{Sh}_2')) \rangle = \langle a^{(1)}, f(\mathsf{Sh}_1') \rangle + \langle a^{(2)}, g(\mathsf{Sh}_2') \rangle$$

are ε_2-far. Consequently, for $f', g' : \{0,1\}^\ell \to \{0,1\}$ defined as $f'(z) = \langle a^{(1)}, f(z) \rangle$ and $g'(z) = \langle a^{(2)}, g(z) \rangle$ it holds that

$$f'(\mathsf{Sh}_1), g'(\mathsf{Sh}_2) \not\approx_{\varepsilon_2} f'(\mathsf{Sh}_1'), g'(\mathsf{Sh}_2'),$$

contradicting the fact that $(\mathbf{Share}, \mathbf{Rec}, \mathcal{Y})$ is an $(\varepsilon_1, \varepsilon_2)$-leakage-resilient secret sharing scheme. □

3 Randomness extraction from leakage-resilient secret sharing schemes

In this section, we show that all 2-out-of-2 secret sharing schemes satisfying the weak leakage-resilience requirement from Definition 2 require extractable randomness with good parameters.

Theorem 2. *Given any $\gamma \in (0,1)$, there are absolute constants $c_\gamma, c'_\gamma, c''_\gamma > 0$ such that the following holds: Suppose $(\mathbf{Share}, \mathbf{Rec}, \mathcal{Y})$ is an $(\varepsilon_1, \varepsilon_2)$-leakage-resilient secret sharing scheme for b-bit messages using d bits of randomness. Then, if $b \geq c_\gamma$ and $d \leq 2^{c'_\gamma b}$ it holds that $\mathcal{Y}$ is (δ, m)-extractable with $\delta \leq 2^b \varepsilon_2 + 2^{-c''_\gamma b}$ and $m \geq (1-\gamma)b$.*

We prove Theorem 2 via a sequence of lemmas by showing the existence of an extractor $\mathsf{Ext} : \{0,1\}^d \to \{0,1\}^m$ for the class $\mathcal{Y}$ with appropriate parameters. Our construction works as follows: On input $y \in \{0,1\}^d$, the extractor Ext computes $(L_y, R_y) = \mathbf{Share}(0^b, y)$, applies special leakage functions $f, g : \{0,1\}^\ell \to \{0,1\}^b$ to be determined in order to obtain local leakage $(f(L_y), g(R_y))$, and finally outputs $\mathsf{Ext}(y) = h(f(L_y), g(R_y))$ for an appropriate function $h : \{0,1\}^{2b} \to \{0,1\}^m$. Our goal is to show that

$$\mathsf{Ext}(Y) \approx_\delta U_m \tag{2}$$

for all sources $Y \in \mathcal{Y}$. Similarly in spirit to [BD07], our first lemma shows that in order to prove (2) we can instead focus on extracting randomness from the family of distributions

$$\mathcal{Z} = \{\mathbf{Share}(U_b, y) : y \in \{0,1\}^d\}.$$

Lemma 6. *Fix functions $f, g : \{0,1\}^\ell \to \{0,1\}^b$ and $h : \{0,1\}^{2b} \to \{0,1\}^m$, and suppose that*

$$\mathsf{Ext}'(Z) = h(f(Z_1), g(Z_2)) \approx_{\delta'} U_s \tag{3}$$

for all $Z = (Z_1, Z_2) \in \mathcal{Z}$. Then, it holds that Ext *given by* $\mathsf{Ext}(y) = h(f(L_y), g(R_y))$*, where $(L_y, R_y) = \mathbf{Share}(0^b, y)$, satisfies*

$$\mathsf{Ext}(Y) \approx_\delta U_m$$

for all $Y \in \mathcal{Y}$ with $\delta = 2^b \varepsilon_2 + \delta'$.

Proof. Lemma 5 implies that

$$f(L_Y), g(R_Y) \approx_{\varepsilon'} f(L'_Y), g(R'_Y),$$

where $(L'_Y, R'_Y) = \mathbf{Share}(U_b, Y)$ holds with $\varepsilon' = 2^b \varepsilon_2$ for all $Y \in \mathcal{Y}$, and so $\mathsf{Ext}(Y) \approx_{\varepsilon'} h(f(L'_K), g(R'_K))$. Since (3) holds for all $Z \in \mathcal{Z}$ and $\mathbf{Share}(U_b, Y)$ is a convex combination of distributions in $\mathcal{Z}$, it follows that $h(f(L'_Y), g(R'_Y)) \approx_{\delta'} U_m$. The triangle inequality yields the desired result. □

Given Lemma 6, we will focus on proving (3) for appropriate functions f, g, and h and error δ' in the remainder of this section. We show the following lemma, which implies Theorem 2 together with Lemma 6.

Lemma 7. *Given any $\gamma \in (0,1)$, there are absolute constants $c_\gamma, c'_\gamma, c''_\gamma > 0$ such that if $b \geq c_\gamma$ and $d \leq 2^{c'_\gamma b}$, then there exist functions $f, g : \{0,1\}^\ell \to \{0,1\}^b$ and $h : \{0,1\}^{2b} \to \{0,1\}^m$ such that*

$$\mathsf{Ext}'(Z) = h(f(Z_1), g(Z_2)) \approx_{\delta'} U_m$$

for all $Z = (Z_1, Z_2) \in \mathcal{Z}$ with $\delta' \leq 2^{-c''_\gamma b}$ and $m \geq (1-\gamma)b$.

The roadmap for the proof ahead is that we are first going to fix a $Z \in \mathcal{Z}$, and then do the following:

1. Justify that $Z = (Z_1, Z_2)$ is statistically close to an appropriate convex combination of distributions with linear min-entropy that suit our purposes. (Lemma 8)

2. Show that if we pick f and g uniformly at random, then with high probability over this choice it holds that $(f(Z_1), g(Z_2))$ is statistically close to a distribution with decent min-entropy. (Lemma 9)

3. Note that a random function h extracts uniformly random bits from $(f(Z_1), g(Z_2))$ with high probability, provided that this distribution contains enough min-entropy. A union bound over the 2^d distributions in $\mathcal{Z}$ concludes the argument.

Lemma 8. *Fix $\beta \in (0,1)$ and an integer $r > 0$. Then, for all $(Z_1, Z_2) \in \mathcal{Z}$ it holds that (Z_1, Z_2) is $\left(r \cdot 2^{-(1-\beta-1/r)}\right)$-close to a distribution $D = \sum_{i \in \mathcal{I}} p_i \cdot (D_{1,i}, D_{2,i})$ where for each $i \in \mathcal{I} \subseteq [r]$ it holds that $D_{1,i}, D_{2,i} \in \{0,1\}^\ell$, and $\mathbf{H}_\infty(D_{1,i}) \geq \left(\beta - (\frac{i-1}{r})\right)b$ and $\mathbf{H}_\infty(D_{2,i} | D_{1,i} = \mathsf{sh}_1) \geq \left(\frac{i-1}{r}\right)b$ for every $\mathsf{sh}_1 \in \mathsf{supp}(D_{1,i})$.*

Proof. Fix some $y \in \{0,1\}^d$ and set $(Z_1, Z_2) = \mathbf{Share}(U_b, y)$. It will be helpful for us to see $\mathbf{Share}(\cdot, y)$ as a bipartite graph G with left and right vertex sets $\{0,1\}^\ell$ and an edge between sh_1 and sh_2 if $(\mathsf{sh}_1, \mathsf{sh}_2) \in \mathsf{supp}(Z_1, Z_2)$. Then, (Z_1, Z_2) is the uniform distribution on the 2^b edges of G. For every left vertex $\mathsf{sh}_1 \in \{0,1\}^\ell$, we define its neighborhood

$$\mathcal{A}(\mathsf{sh}_1) = \{\mathsf{sh}_2 : (\mathsf{sh}_1, \mathsf{sh}_2) \in \mathsf{supp}(Z_1, Z_2)\}$$

and its degree

$$\mathsf{deg}(\mathsf{sh}_1) = |\mathcal{A}(\mathsf{sh}_1)|.$$

Note that $(Z_2 | Z_1 = \mathsf{sh}_1)$ is uniformly distributed over $\mathcal{A}(\mathsf{sh}_1)$, and so $\mathbf{H}_\infty(Z_2 | Z_1 = \mathsf{sh}_1) = \log \mathsf{deg}(\mathsf{sh}_1)$.

Partition $\mathsf{supp}(Z_1)$ into sets

$$\mathcal{S}_i = \left\{\mathsf{sh}_1 : 2^{(\frac{i-1}{r})b} \leq \mathsf{deg}(\mathsf{sh}_1) < 2^{(\frac{i}{r})b}\right\}$$

for $i \in [r]$. With this definition in mind, we can express (Z_1, Z_2) as

$$\sum_{i \in [r]} \Pr[Z_1 \in \mathcal{S}_i](Z_1, Z_2 | Z_1 \in S_i),$$

where $(Z_1, Z_2 | Z_1 \in S_i)$ denotes the distribution (Z_1, Z_2) conditioned on the event that $Z_1 \in \mathcal{S}_i$. Call a non-empty set $\mathcal{S}_i$ *good* if $\sum_{\mathsf{sh}_1 \in \mathcal{S}_i} \mathrm{deg}(\mathsf{sh}_1) \geq 2^{(\beta+1/r)b}$. Otherwise the set $\mathcal{S}_i$ is *bad*. Let $\mathcal{I}$ denote the set of indices $i \in [r]$ such that $\mathcal{S}_i$ is good. We proceed to show that we can take the target distribution D in the lemma statement to be $D = \sum_{i \in \mathcal{I}} p_i \cdot (D_{1,i}, D_{2,i})$ for $p_i = \frac{\Pr[Z_1 \in \mathcal{S}_i]}{\Pr[Z_1 \text{ lands on good set}]}$ with $(D_{1,i}, D_{2,i}) = (Z_1, Z_2 | Z_1 \in \mathcal{S}_i)$ when $i \in \mathcal{I}$.

To see this, consider the case where $\mathcal{S}_i$ is good, i.e., we have $\sum_{\mathsf{sh}_1 \in \mathcal{S}_i} \mathrm{deg}(\mathsf{sh}_1) \geq 2^{(\beta+1/r)b}$. For each $\mathsf{sh}_1 \in \mathcal{S}_i$, we have

$$\begin{aligned}\Pr[Z_1 = \mathsf{sh}_1 | Z_1 \in \mathcal{S}_i] &= \frac{\mathsf{deg}(\mathsf{sh}_1)}{\sum_{s \in \mathcal{S}_i} \mathsf{deg}(s)} \\ &\leq \frac{2^{\frac{i}{r}b}}{2^{(\beta+1/r)b}}\end{aligned}$$

$$= 2^{-\left(\beta - (\frac{i-1}{r})\right)b}.$$

Furthermore, for any $\mathsf{sh}_1 \in \mathcal{S}_i$ and sh_2 we know that

$$\Pr[Z_2 = \mathsf{sh}_2 | Z_1 = \mathsf{sh}_1] \leq 2^{-\left(\frac{i-1}{r}\right)b}.$$

Combining these two observations shows that in this case we have $\mathbf{H}_\infty(Z_1 | Z_1 \in \mathcal{S}_i) \geq \left(\beta - (\frac{i-1}{r})\right)b$ and $\mathbf{H}_\infty(Z_2 | Z_1 = \mathsf{sh}_1) \geq \left(\frac{i-1}{r}\right)b$ for all valid fixings $\mathsf{sh}_1 \in \mathcal{S}_i$.

To conclude the proof, consider D as above, which we have shown satisfies the properties described in the lemma statement. Noting that D corresponds exactly to (Z_1, Z_2) conditioned on Z_1 landing on a good set, we have

$$\Delta((Z_1, Z_2); D) \leq \Pr[Z_1 \text{ lands in a bad set}].$$

It remains to bound this probability on the right-hand side. Assuming the set $\mathcal{S}_i$ is bad, it holds that $\sum_{\mathsf{sh}_1 \in \mathcal{S}_i} \deg(\mathsf{sh}_1) < 2^{(\beta+1/r)b}$. Therefore, since (Z_1, Z_2) takes on any edge with probability 2^{-b}, it holds that Z_1 lands in $\mathcal{S}_i$ with probability at most $2^{-b} \cdot 2^{(\beta+1/r)b} = 2^{-(1-\beta-1/r)b}$. There are at most r bad sets, so by a union bound we have $\Pr[Z_1 \text{ lands in a bad set}] \leq r \cdot 2^{-(1-\beta-1/r)b}$. □

Lemma 9. *Fix $\alpha, \beta \in (0,1)$ and an integer r. Then, with probability at least $1 - 3r \cdot e^{b - \alpha^2 2^{\min(b/r, (\beta-1/r)b)}}$ over the choice of uniformly random functions $f, g : \{0,1\}^\ell \to \{0,1\}^b$ it holds that $(f(Z_1), g(Z_2))$ is $\left(2\alpha + r \cdot 2^{-(1-\beta-1/r)b}\right)$-close to a $(2b, (\beta - 1/r)b - 4\log(1/\alpha))$-source.*

Proof. Suppose we pick functions $f, g : \{0,1\}^\ell \to \{0,1\}^b$ uniformly at random. We begin by expressing $(f(Z_1), g(Z_2))$ as

$$\sum_{i \in [r]} \Pr[Z_1 \in \mathcal{S}_i](f(Z_1), g(Z_2) | Z_1 \in S_i),$$

which by Lemma 8 is $\left(r \cdot 2^{-(1-\beta-1/r)b}\right)$-close to

$$\sum_{i \in \mathcal{I}} \Pr[Z_1 \in \mathcal{S}_i](f(D_{1,i}), g(D_{2,i})).$$

We proceed by cases:

- $\frac{i-1}{r} \geq \beta - 1/r$: We know from Lemma 8 that $\mathbf{H}_\infty(D_{2,i} | D_{1,i} = \mathsf{sh}_1) \geq (\beta - 1/r)b$ for all $\mathsf{sh}_1 \in \mathsf{supp}(D_{1,i})$. By Lemma 4, we have

  $$(g(D_{2,i}) | D_{1,i} = \mathsf{sh}_1) \approx_\alpha V$$

  for some V with $\mathbf{H}_\infty(V) \geq (\beta - 1/r)b - 2\log(1/\alpha)$ with probability at least $1 - 2e^{-\alpha^2 2^{(\beta-1/r)b}}$ over the choice of g. Since this holds for any valid fixing $D_{1,i} = \mathsf{sh}_1$, we conclude via a union bound over the at most 2^b possible fixings that

  $$f(D_{1,i}), g(D_{2,i}) \approx_\alpha W_i$$

  for some W_i with $\mathbf{H}_\infty(W_i) \geq (\beta - 1/r)b - 2\log(1/\alpha)$ with probability at least $1 - 2e^{b - \alpha^2 2^{(\beta-1/r)b}}$ over the choice of f and g.

- $1/r \leq \frac{i-1}{r} < \beta - 1/r$: We know from Lemma 8 that $\mathbf{H}_\infty(D_{1,i}) \geq \left(\beta - \frac{i-1}{r}\right)b$ and $\mathbf{H}_\infty(D_{2,i}|D_{1,i} = \mathsf{sh}_1) \geq \left(\frac{i-1}{r}\right)b$ for all $\mathsf{sh}_1 \in \mathsf{supp}(D_{1,i})$. First, by Lemma 4 we conclude that with probability at least
$$1 - 2e^{-\alpha^2 2^{\left(\beta - \frac{i-1}{r}\right)b}} \geq 1 - 2e^{-\alpha^2 2^{b/r}}$$
over the choice of f it holds that
$$f(D_{1,i}) \approx_\alpha V_1 \tag{4}$$
for some V_1 with $\mathbf{H}_\infty(V_1) \geq (\beta - \frac{i-1}{r})b - 2\log(1/\alpha)$. Analogously, for every $\mathsf{sh}_1 \in \mathsf{supp}(D_{1,i})$, we can again invoke Lemma 4 to see that with probability at least
$$1 - 2e^{-\alpha^2 2^{\left(\frac{i-1}{r}\right)b}} \geq 1 - 2e^{-\alpha^2 2^{b/r}}$$
over the choice of g, for any $\mathsf{sh}_1 \in \mathsf{supp}(D_{1,i})$ it holds that
$$(g(D_{2,i})|D_{1,i} = \mathsf{sh}_1) \approx_\alpha V_{2,\mathsf{sh}_1} \tag{5}$$
for some V_{2,sh_1} with $\mathbf{H}_\infty(V_{2,\mathsf{sh}_1}) \geq \left(\frac{i-1}{r}\right)b - 2\log(1/\alpha)$. By a union bound over the at most 2^b possible fixings sh_1, we conclude that (5) holds simultaneously for all $\mathsf{sh}_1 \in \mathsf{supp}(D_{1,i})$ with probability at least $1 - 2e^{b-\alpha^2 2^{b/r}}$ over the choice of g. An additional union bound shows that this holds simultaneously along (4) with probability at least $1 - 3e^{b-\alpha^2 2^{b/r}}$ over the choice of f and g, which implies that
$$f(D_{1,i}), g(D_{2,i}) \approx_{2\alpha} W_i$$
for some W_i with
$$\mathbf{H}_\infty(W_i) \geq \left(\beta - \frac{i-1}{r}\right)b - 2\log(1/\alpha) + \left(\frac{i-1}{r}\right)b - 2\log(1/\alpha) = \beta b - 4\log(1/\alpha).$$

- $i = 1$: In this case, by Lemma 8 we know that $\mathbf{H}_\infty(D_{1,i}) \geq \beta b$. Therefore, Lemma 4 implies that $f(D_{1,i}) \approx_\alpha V_1$ for some V_1 such that $\mathbf{H}_\infty(V_1) \geq \beta b - 2\log(1/\alpha)$ with probability at least $1 - 2e^{-\alpha^2 2^{\beta b}} \geq 1 - 2e^{-\alpha^2 2^{b/r}}$. This implies that $f(D_{1,i}), g(D_{2,i}) \approx_\alpha W_i$ for some W_i with $\mathbf{H}_\infty(W_i) \geq \beta b - 2\log(1/\alpha)$.

Finally, a union bound over the at most r indices $i \in \mathcal{I}$ yields the desired statement. □

We are now ready to prove Lemma 7 with the help of Lemma 9.

Proof of Lemma 7. Fix some $\gamma \in (0,1)$. Then, we set $\beta = 1-\gamma/2 > 1-\gamma$, $\alpha = 2^{-cb}$ for a sufficiently small constant $c > 0$, and $r > 0$ a sufficiently large integer so that
$$1 - \gamma \leq \beta - 1/r - 6c \tag{6}$$
and
$$1/r + 6c \leq \frac{\min(\beta, 1-\beta)}{100}. \tag{7}$$
According to Lemma 9, we know that for any given $Z = (Z_1, Z_2) \in \mathcal{Z}$ it holds that $(f(Z_1), g(Z_2))$ is $(2\alpha + r \cdot 2^{-(1-\beta-1/r)b})$-close to some $(2b, (\beta - 1/r)b - 4\log(1/\alpha))$-source W with probability at least $1 - 3r \cdot e^{b-\alpha^2 2^{\min(b/r, (\beta-1/r)b)}}$ over the choice of f and g.

Let $m = (1-\gamma)b$ and pick a uniformly random function $h : \{0,1\}^{2b} \to \{0,1\}^m$. Then, since $m \leq \mathbf{H}_\infty(W) - 2\log(1/\alpha)$ by (6), Lemma 3 implies that $h(W) \approx_\alpha U_m$, and hence
$$h(f(Z_1), g(Z_2)) \approx_{3\alpha + r\cdot 2^{-(1-\beta-1/r)b}} U_m, \tag{8}$$

with probability at least

$$1 - 2e^{-\alpha^2 2^{(\beta-1/r)b-4\log(1/\alpha)}} - 3r \cdot e^{b-\alpha^2 2^{\min(b/r,(\beta-1/r)b)}} \geq 1 - 5r \cdot e^{b-\alpha^2 2^{\min(b/r,(\beta-1/r)b)-4\log(1/\alpha)}}$$

over the choice of f, g, and h, via a union bound.

Now, observe that from (7), if $b \geq c_\gamma$ for a sufficiently large constant $c_\gamma > 0$, it follows that

$$5r \cdot e^{b-\alpha^2 2^{\min(b/r,(\beta-1/r)b)-4\log(1/\alpha)}} \leq 2^{-2^{2c'_\gamma b}}$$

for some constant $c'_\gamma > 0$. Moreover, under (7) we also have that

$$\delta' := 3\alpha + r \cdot 2^{-(1-\beta-1/r)b} \leq 2^{-c''_\gamma b}$$

for some constant $c''_\gamma > 0$. Finally, a union bound over the 2^d distributions in $\mathcal{Z}$ shows that (8) holds simultaneously for all $Z \in \mathcal{Z}$ with probability at least $1 - 2^{d-2^{2c'_\gamma b}}$. Consequently, if $d \leq 2^{c'_\gamma b}$ it follows that there exist functions f, g, and h such that (8) holds for all $Z \in \mathcal{Z}$ with the appropriate error δ' and output length m. □

3.1 The main result

In this section, we use Theorem 2 to obtain the main result of this section, along with an important corollary.

Theorem (Theorem 1, restated)**.** *Let* $(\mathbf{Share}, \mathbf{Rec}, \mathcal{Y})$ *be an* $(\varepsilon_1, \varepsilon_2)$*-leakage-resilient secret sharing scheme for* b*-bit messages. Then, either:*

- *The scheme uses* $d \geq \min\left(2^{\Omega(b)}, (1/\varepsilon_2)^{\Omega(1)}\right)$ *bits of randomness, or;*
- *The class of randomness sources* $\mathcal{Y}$ *is* (δ, m)*-extractable with* $\delta \leq \max\left(2^{-\Omega(b)}, \varepsilon_2^{\Omega(1)}\right)$ *and* $m = \Omega(\min(b, \log(1/\varepsilon_2)))$*. Moreover, if* **Share** *is an efficient algorithm, then* $\mathcal{Y}$ *is efficiently* (δ, m)*-extractable.*

Proof. Given the scheme $(\mathbf{Share}, \mathbf{Rec}, \mathcal{Y})$ from the theorem statement, let $b' = \min\left(b, \left\lceil \frac{\log(1/\varepsilon_2)}{100} \right\rceil\right)$ and consider the modified scheme $(\mathbf{Share}', \mathbf{Rec}', \mathcal{Y})$ for b'-bit messages obtained by appending $0^{b-b'}$ to every b'-bit message and running the original scheme $(\mathbf{Share}, \mathbf{Rec}, \mathcal{Y})$. Applying Theorem 2 to $(\mathbf{Share}', \mathbf{Rec}', \mathcal{Y})$ we conclude that either $\mathbf{Share}'$, and hence **Share**, uses at least $2^{\Omega(b')} = \min\left(2^{\Omega(b)}, (1/\varepsilon_2)^{\Omega(1)}\right)$ bits of randomness, or $\mathcal{Y}$ is (δ, m)-extractable with $\delta \leq 2^{-\Omega(b')} = \max\left(2^{-\Omega(b)}, \varepsilon_2^{\Omega(1)}\right)$ and $m = \Omega(b') = \Omega(\min(b, \log(1/\varepsilon_2)))$. □

3.2 Efficient leakage-resilient secret sharing requires efficiently extractable randomness

In this section, we show that every low-error leakage-resilient secret sharing scheme $(\mathbf{Share}, \mathbf{Rec}, \mathcal{Y})$ for b-bit messages where **Share** runs in time $\text{poly}(b)$ admits an efficient low-error extractor for $\mathcal{Y}$ running in time $\text{poly}(b)$. Similarly to [BD07, Section 3.1], this is done by replacing the uniformly random functions f, g, and h in the proof of Theorem 2 by *t-wise independent functions*, for an appropriate parameter t.

We say that a family of functions $\mathcal{F}_t$ from $\{0,1\}^p$ to $\{0,1\}^q$ is *t-wise independent* if for F sampled uniformly at random from $\mathcal{F}_t$ it holds that the random variables $F(x_1), F(x_2), \ldots, F(x_t)$

are independent and uniformly distributed over $\{0,1\}^q$ for any distinct $x_1,\dots,x_t \in \{0,1\}^p$. There exist t-wise independent families of functions $\mathcal{F}_t$ such that every $f \in \mathcal{F}_t$ can be computed in time $\mathrm{poly}(b)$ and can be described by $\mathrm{poly}(b)$ bits whenever p, q, and t are $\mathrm{poly}(b)$ [Dod00, TV00, BD07]. Therefore, since **Share** runs in time $\mathrm{poly}(b)$, it suffices to show the existence of functions f, g, and h belonging to appropriate $\mathrm{poly}(b)$-wise independent families of functions such that $\mathsf{Ext}(Y) = h(f(\mathsf{Sh}_1), g(\mathsf{Sh}_2))$ is statistically close to uniform, where $(\mathsf{Sh}_1, \mathsf{Sh}_2) = \mathbf{Share}(0^b, Y)$, for every source $Y \in \mathcal{Y}$. We accomplish this with the help of some auxiliary lemmas. The first lemma states a standard concentration bound for the sum of t-wise independent random variables.

Lemma 10 ([Dod00, Theorem 5])**.** *Fix an even integer $t \geq 2$ and suppose that $X_1,\dots,X_N$ are t-wise independent random variables in $[0,1]$. Let $X = \sum_{i=1}^N X_i$ and $\mu = \mathbb{E}[X]$. Then, it holds that*

$$\Pr[|X-\mu| \geq \varepsilon \cdot \mu] \leq 3\left(\frac{t}{\varepsilon^2 \mu}\right)^{t/2}$$

for every $\varepsilon < 1$.

We can use Lemma 10 to derive analogues of Lemmas 3 and 4 for t-wise independent functions.

Lemma 11. *Suppose $f : \{0,1\}^p \to \{0,1\}^q$ is sampled uniformly at random from a $2t$-wise independent family of functions with $q \leq k - \log t - 2\log(1/\varepsilon) - 5$ and $t \geq q$, and let Y be a (p,k)-source. Then, it follows that*

$$f(Y) \approx_\varepsilon U_q$$

with probability at least $1 - 2^{-t}$ over the choice of f.

Proof. Fix a (p,k)-source Y and suppose $f : \{0,1\}^p \to \{0,1\}^q$ is sampled from a family of $2t$-wise independent functions. Note that

$$\Delta(f(Y); U_q) = \frac{1}{2} \sum_{z \in \{0,1\}^q} |\Pr[f(Y) = z] - 2^{-q}|.$$

For each $y \in \{0,1\}^p$ and $z \in \{0,1\}^q$, consider the random variable $W_{y,z} = \Pr[Y = y] \cdot \mathbf{1}_{\{f(y)=z\}}$. Then, we may write

$$\Delta(f(Y); U_q) = \frac{1}{2} \sum_{z \in \{0,1\}^q} \left| \sum_{y \in \{0,1\}^p} W_{y,z} - 2^{-q} \right|.$$

Note that the $W_{y,z}$'s are $2t$-wise independent, $\mathbb{E}[\sum_{y \in \{0,1\}^n} W_{y,z}] = 2^{-q}$, and that $2^k \cdot W_{y,z} \in [0,1]$. Therefore, an application of Lemma 10 with the random variables $(2^k \cdot W_{y,z})_{y \in \{0,1\}^p, z \in \{0,1\}^q}$ shows that

$$\Pr\left[\left| \sum_{y \in \{0,1\}^p} W_{y,z} - 2^{-q} \right| > 2\varepsilon \cdot 2^{-q} \right] \leq 3\left(\frac{t \cdot 2^q}{2\varepsilon^2 2^k}\right)^t.$$

Therefore, a union bound over all $z \in \{0,1\}^q$ shows that $f(Y) \approx_\varepsilon U_q$ fails to hold with probability at most $3 \cdot 2^q \cdot 2^{-t}\left(\frac{t \cdot 2^q}{\varepsilon^2 \cdot 2^k}\right)^t \leq 2^{-t}$ over the choice of f, where the inequality follows by the upper bound on q. □

The proof of the following lemma is analogous to the proof of Lemma 4, but using Lemma 11 instead of Lemma 3.

Lemma 12. *Suppose $f : \{0,1\}^p \to \{0,1\}^q$ is sampled uniformly at random from a $2t$-wise independent family of functions with $t \geq q$, and let Y be a (p,k)-source. Then, it follows that $f(Y) \approx_\varepsilon W$ for some W such that $\mathbf{H}_\infty(W) \geq \min(q, k - \log t - 2\log(1/\varepsilon) - 5)$ with probability at least $1 - 2^{-t}$ over the choice of f.*

Following the reasoning used in the proof of Theorem 2 but sampling $f, g : \{0,1\}^\ell \to \{0,1\}^b$ and $h : \{0,1\}^b \to \{0,1\}^m$ from $2t$-wise independent families of functions with $t = 100\max(b,d) = \mathrm{poly}(b)$, and using Lemmas 11 and 12 in place of Lemmas 3 and 4, respectively, yields the following analogous result. Informally, it states that efficient low-error leakage-resilient secret sharing schemes require efficiently extractable randomness.

Theorem 3. *There exist absolute constants $c, c' > 0$ such that the following holds for b large enough: Suppose $(\mathbf{Share}, \mathbf{Rec}, \mathcal{Y})$ is a $(\varepsilon_1, \varepsilon_2)$-leakage-resilient secret sharing for b-bit messages using d bits of randomness such that* **Share** *runs in time* $\mathrm{poly}(b)$*. Then, there exists a deterministic extractor $\mathsf{Ext} : \{0,1\}^d \to \{0,1\}^m$ computable by a family of* $\mathrm{poly}(b)$*-sized circuits with output length $m \geq c \cdot b$ such that*

$$\mathsf{Ext}(Y) \approx_\delta U_m$$

with $\delta = 2^b \varepsilon_2 + 2^{-c' \cdot b}$ for every $Y \in \mathcal{Y}$.

4 Random-less reductions for secret sharing

In this section, we study black-box deterministic reductions between different types of threshold secret sharing. Such reductions from (t', n', ε)-secret sharing schemes to (t, n, ε)-secret sharing schemes (for the same message length b and number of randomness bits d) would allow us to conclude that if all these (t, n, ε)-secret sharing schemes require a (δ, m)-extractable class of randomness sources, then so do all (t', n', ε)-secret sharing schemes. We provide reductions which work over a large range of parameters and prove complementary results showcasing the limits of such reductions. As already discussed in Section 1, our starting point for devising black-box reductions is the notion of a *distribution design* as formalized by Stinson and Wei [SW18] (with roots going back to early work on secret sharing [BL88]), which we defined in Definition 5. As stated in Lemma 1, the existence of a (t, n, t', n')-distribution design yields the desired reduction from (t', n', ε)-secret sharing to (t, n, ε)-secret sharing. Therefore, we focus directly on the study of distribution designs in this section.

We begin with a naive construction.

Theorem 4. *There exists a (t, n, t', n')-distribution design whenever $t' \geq t$ and $n' \geq n + (t' - t)$. In particular, if every (t, n, ε)-secret sharing scheme for b-bit messages and using d bits of randomness requires a (δ, m)-extractable class of randomness sources, then so does every (t', n', ε)-secret sharing scheme for b-bit messages using d bits of randomness whenever $t' \geq t$ and $n' \geq n + (t' - t)$.*

Proof. Consider the (t, n, t', n')-distribution design $\mathcal{D}_1, \dots, \mathcal{D}_n$ obtained by setting $\mathcal{D}_i = \{i\} \cup \{n' - (t' - t) + 1, n' - (t' - t) + 2, \dots, n'\}$, which is valid exactly when the conditions of the theorem are satisfied. □

The following result shows the limits of distribution designs, and will be used to show the optimality of our constructions when $t = 2$ or $t' = n'$.

Theorem 5. *A (t, n, t', n')-distribution design exists only if $\binom{n'}{t'-1} \geq \binom{n}{t-1}$ and $t' \geq t$.*

Proof. Consider an arbitrary (t, n, t', n')-distribution design $\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_n$. First, note that it must be the case that all the $\mathcal{D}_i$'s are non-empty. This implies that we must have $t' \geq t$. Second, to see that $\binom{n'}{t'-1} \geq \binom{n}{t-1}$, consider all $\binom{n}{t-1}$ distinct subsets $\mathcal{T} \subseteq [n]$ of size $t-1$, and denote $\mathcal{D}_\mathcal{T} = \bigcup_{i \in \mathcal{T}} \mathcal{D}_i$. By the definition of distribution design, it must hold that

$$|\mathcal{D}_\mathcal{T}| \leq t' - 1.$$

Consider now modified sets $\widehat{\mathcal{D}_\mathcal{T}}$ obtained by adding arbitrary elements to $\mathcal{D}_\mathcal{T}$ so that $|\widehat{\mathcal{D}_\mathcal{T}}| = t' - 1$. Then, from the definition of distribution design, for any two distinct subsets $\mathcal{T}, \mathcal{T}' \subseteq [n]$ of size $t-1$ it must be the case that

$$\left|\widehat{\mathcal{D}_\mathcal{T}} \cup \widehat{\mathcal{D}_{\mathcal{T}'}}\right| \geq t'.$$

This implies that $\widehat{\mathcal{D}_\mathcal{T}} \neq \widehat{\mathcal{D}_{\mathcal{T}'}}$ for all distinct subsets $\mathcal{T}, \mathcal{T}' \subseteq [n]$ of size $t-1$, which can only hold if $\binom{n'}{t'-1} \geq \binom{n}{t-1}$. □

We now show that Theorem 5 is tight for a broad range of parameters. In particular, when $t = 2$ or $t' = n'$ we are able to characterize exactly under which parameters a (t, n, t', n')-distribution design exists.

Theorem 6. *There exists a $(t = 2, n, t', n')$-distribution design if and only if $n \leq \binom{n'}{t'-1}$. In particular, if every $(t = 2, n, \varepsilon)$-secret sharing scheme for b-bit messages using d bits of randomness requires (δ, m)-extractable randomness, then so does every (t', n', ε)-secret sharing scheme for b-bit messages using d bits of randomness whenever $n \leq \binom{n'}{t'-1}$.*

Proof. Note that the condition $n \leq \binom{n'}{t'-1}$ implies that we can take $\mathcal{D}_1, \ldots, \mathcal{D}_n$ to be distinct subsets of $[n']$ of size $t'-1$, and so $|\mathcal{D}_i \cup \mathcal{D}_j| \geq t'$ for any distinct indices i and j. The reverse implication follows from Theorem 5. □

Theorem 7. *There exists a $(t, n, t' = n', n')$-distribution design if and only if $n' \geq \binom{n}{t-1}$. In particular, if every (t, n, ε)-secret sharing scheme for b-bit messages using d bits of randomness requires (δ, m)-extractable randomness, then so does every (n', n', ε)-secret sharing scheme for b-bit messages using d bits of randomness whenever $n' \geq \binom{n}{t-1}$.*

Proof. We show that a (t, n, n', n')-distribution design exists whenever $n' = \binom{n}{t-1}$, which implies the desired result. Let $\mathcal{P}$ denote the family of all subsets of $[n]$ of size $t-1$, and set $n' = |\mathcal{P}| = \binom{n}{t-1}$ (we may use any correspondence between elements of $\mathcal{P}$ and integers in $[n']$). Then, we define the set $\mathcal{D}_i \subseteq \mathcal{P}$ for $i \in [n]$ to contain all elements of $\mathcal{P}$ except the subsets of $[n]$ which contain i. We argue that $\mathcal{D}_1, \ldots, \mathcal{D}_n$ is a distribution design with the desired parameters. First, observe that for any distinct indices $i_1, i_2, \ldots, i_{t-1} \in [n]$ it holds that

$$\bigcup_{j=1}^{t-1} \mathcal{D}_{i_j} = \mathcal{P} \setminus \{\{i_1, i_2, \ldots, i_{t-1}\}\}.$$

On the other hand, since $\{i_1, \ldots, i_{t-1}\} \in \mathcal{D}_{i_t}$ for any index $i_t \neq i_1, \ldots, i_{t-1}$, it follows that $\bigcup_{j=1}^{t} \mathcal{D}_{i_j} = \mathcal{P}$, as desired.

The reverse implication follows from Theorem 5. □

4.1 Distribution designs from partial Steiner systems

In this section, we show that every partial Steiner system is also a distribution design which beats the naive construction from Theorem 4 for certain parameter regimes. Such set systems have been previously used in seminal constructions of pseudorandom generators and extractors [NW94, Tre01], and are also called combinatorial designs.

Definition 9 (Partial Steiner system)**.** *We say a family of sets $\mathcal{D}_1, \dots, \mathcal{D}_n \subseteq [n']$ is an (n, n', ℓ, a)-* partial Steiner system *if it holds that $|\mathcal{D}_i| = \ell$ for every $i \in [n]$ and $|\mathcal{D}_i \cap \mathcal{D}_j| \leq a$ for all distinct $i, j \in [n]$.*

The conditions required for the existence of a partial Steiner system are well-understood, as showcased in the following result from [EFF85, NW94, Tre01], which is nearly optimal [Röd85, RRV02].

Lemma 13 ([EFF85, NW94, Tre01])**.** *Fix positive integers n, ℓ, and $a \leq \ell$. Then, there exists an (n, n', ℓ, a)-partial Steiner system for every integer $n' \geq e \cdot n^{1/a} \cdot \frac{\ell^2}{a}$.*

Noting that every partial Steiner system with appropriate parameters is also a distribution design, we obtain the following theorem.

Theorem 8. *Fix an integer $a \geq 1$. Then, there exists a (t, n, t', n')-distribution design whenever $t' \geq t^2 + \frac{at(t-1)^2}{2}$ and $n' \geq \frac{en^{1/a}}{a} \cdot \left(1 + \frac{t'}{t} + \frac{a(t-1)^2}{2}\right)$.*

Proof. Fix an integer $a \geq 1$ and an (n, n', ℓ, a)-partial Steiner system $\mathcal{D}_1, \dots, \mathcal{D}_n \subseteq [n']$ with $\ell = \left\lceil \frac{t'}{t} + \frac{a(t-1)}{2} \right\rceil$. By Lemma 13 and the choice of ℓ, such a partial Steiner system is guaranteed to exist whenever n' satisfies the condition in the theorem statement. We proceed to argue that this partial Steiner system is also a (t, n, t', n')-distribution design. First, fix an arbitrary set $\mathcal{T} \subseteq [n]$ of size $t - 1$. Then, we have

$$|\mathcal{D}_{\mathcal{T}}| \leq \ell(t-1) \leq t' - 1,$$

where the rightmost inequality holds by our choice of ℓ and the condition on t' and t in the theorem statement. Second, fix an arbitrary set $\mathcal{T} \subseteq [n]$ of size t. Then, it holds that

$$\begin{aligned} |\mathcal{D}_{\mathcal{T}}| &\geq \ell + (\ell - a) + (\ell - 2a) + \cdots + (\ell - a(t-1)) \\ &= \ell \cdot t - \frac{at(t-1)}{2} \\ &\geq t', \end{aligned}$$

where the last equality follows again from our choice of ℓ and the condition on t' and t in the theorem statement. □

When n is sufficiently larger than t and t' and t' is sufficiently larger than t, the parameters in Theorem 8 cannot be attained by the naive construction from Theorem 4. For example, if $t' \geq 3t^3$ then we can choose $a = \left\lceil \frac{t'}{t(t-1)^2} \right\rceil$, in which case we have

$$t^2 + \frac{at(t-1)^2}{2} \leq t^2 + \frac{t' + t(t-1)^2}{2} \leq t'. \tag{9}$$

Moreover, we have

$$\frac{en^{1/a}}{a}\left(1 + \frac{t'}{t} + \frac{a(t-1)^2}{2}\right) = \frac{en^{1/a}}{a} + \frac{en^{1/a}}{a} \cdot \frac{t'}{t} + \frac{en^{1/a}(t-1)^2}{2}$$

$$\le \frac{et(t-1)^2 n^{1/3}}{t'} + 2e(t-1)^2 n^{1/3}$$
$$\le 3e(t-1)^2 n^{1/3}. \tag{10}$$

Combining (9) and (10) with Theorem 8, we obtain the following result improving on the naive Theorem 4.

Corollary 5. *Suppose $t' \geq 3t^3$ and $t \leq n^{1/c}$ for some constant $c > 1$. Then, there exists a (t, n, t', n')-distribution design for any $n' \geq \max(t', 3en^{1/3+2c})$. In particular, we may choose n' significantly smaller than n.*

References

[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.

[ADN$^+$19] Divesh Aggarwal, Ivan Damgård, Jesper Buus Nielsen, Maciej Obremski, Erick Purwanto, João Ribeiro, and Mark Simkin. Stronger leakage-resilient and non-malleable secret sharing schemes for general access structures. In Alexandra Boldyreva and Daniele Micciancio, editors, *Advances in Cryptology – CRYPTO 2019*, pages 510–539, Cham, 2019. Springer International Publishing.

[AKO17] Divesh Aggarwal, Tomasz Kazana, and Maciej Obremski. Inception makes non-malleable codes stronger. In Yael Kalai and Leonid Reyzin, editors, *Theory of Cryptography*, pages 319–343, Cham, 2017. Springer International Publishing.

[AL93] Miklos Ajtai and Nathan Linial. The influence of large coalitions. *Combinatorica*, 13(2):129–145, 1993.

[AO20] Divesh Aggarwal and Maciej Obremski. A constant rate non-malleable code in the split-state model. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1285–1294, 2020.

[AOR$^+$20] Divesh Aggarwal, Maciej Obremski, João Ribeiro, Luisa Siniscalchi, and Ivan Visconti. How to extract useful randomness from unreliable sources. In Anne Canteaut and Yuval Ishai, editors, *Advances in Cryptology – EUROCRYPT 2020*, pages 343–372, Cham, 2020. Springer International Publishing.

[BBN$^+$09] Mihir Bellare, Zvika Brakerski, Moni Naor, Thomas Ristenpart, Gil Segev, Hovav Shacham, and Scott Yilek. Hedged public-key encryption: How to protect against bad randomness. In Mitsuru Matsui, editor, *Advances in Cryptology – ASIACRYPT 2009*, pages 232–249, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.

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

[BD07] Carl Bosley and Yevgeniy Dodis. Does privacy require true randomness? In Salil P. Vadhan, editor, *Theory of Cryptography*, pages 1–20, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg.

[BDIR19] Fabrice Benhamouda, Akshay Degwekar, Yuval Ishai, and Tal Rabin. On the local leakage resilience of linear secret sharing schemes. Cryptology ePrint Archive, Report 2019/653, 2019. `https://eprint.iacr.org/2019/653`.

[BDIR21] Fabrice Benhamouda, Akshay Degwekar, Yuval Ishai, and Tal Rabin. On the local leakage resilience of linear secret sharing schemes. *J. Cryptol.*, 34(2):10, 2021.

[BEG17] Salman Beigi, Omid Etesami, and Amin Gohari. Deterministic randomness extraction from generalized and distributed santha–vazirani sources. *SIAM Journal on Computing*, 46(1):1–36, 2017.

[BGK14] Elette Boyle, Shafi Goldwasser, and Yael Tauman Kalai. Leakage-resilient coin tossing. *Distributed Comput.*, 27(3):147–164, 2014.

[BGLZ15] Abhishek Bhowmick, Ariel Gabizon, Thái Hoàng Lê, and David Zuckerman. Deterministic extractors for additive sources: Extended abstract. In Tim Roughgarden, editor, *Proceedings of the 2015 Conference on Innovations in Theoretical Computer Science, ITCS 2015, Rehovot, Israel, January 11-13, 2015*, pages 277–286. ACM, 2015.

[BKMR15] Michael Backes, Aniket Kate, Sebastian Meiser, and Tim Ruffing. Secrecy without perfect randomness: Cryptography with (bounded) weak sources. In Tal Malkin, Vladimir Kolesnikov, Allison Bishop Lewko, and Michalis Polychronakis, editors, *Applied Cryptography and Network Security*, pages 675–695, Cham, 2015. Springer International Publishing.

[BKS$^+$10] Boaz Barak, Guy Kindler, Ronen Shaltiel, Benny Sudakov, and Avi Wigderson. Simulating independence: New constructions of condensers, ramsey graphs, dispersers, and extractors. *J. ACM*, 57(4):20:1–20:52, 2010.

[BL88] Josh Benaloh and Jerry Leichter. Generalized secret sharing and monotone functions. In Shafi Goldwasser, editor, *Advances in Cryptology - CRYPTO '88, 8th Annual International Cryptology Conference, Santa Barbara, California, USA, August 21-25, 1988, Proceedings*, volume 403 of *Lecture Notes in Computer Science*, pages 27–35. Springer, 1988.

[Blu86] Manuel Blum. Independent unbiased coin flips from a correlated biased source—a finite state Markov chain. *Combinatorica*, 6(2):97–108, 1986.

[Bou07] Jean Bourgain. On the construction of affine extractors. *GAFA Geometric And Functional Analysis*, 17(1):33–57, 2007.

[CDH$^+$00] Ran Canetti, Yevgeniy Dodis, Shai Halevi, Eyal Kushilevitz, and Amit Sahai. Exposure-resilient functions and all-or-nothing transforms. In Bart Preneel, editor, *Advances in Cryptology - EUROCRYPT 2000, International Conference on the Theory and Application of Cryptographic Techniques, Bruges, Belgium, May 14-18, 2000, Proceeding*, volume 1807 of *Lecture Notes in Computer Science*, pages 453–469. Springer, 2000.

[CDN15] Ronald Cramer, Ivan Bjerre Damgård, and Jesper Buus Nielsen. *Secure Multiparty Computation and Secret Sharing.* Cambridge University Press, 2015.

[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.

[CGG$^+$20] Eshan Chattopadhyay, Jesse Goodman, Vipul Goyal, Ashutosh Kumar, Xin Li, Raghu Meka, and David Zuckerman. Extractors and secret sharing against bounded collusion protocols. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1226–1242, 2020.

[CGH$^+$85] Benny Chor, Oded Goldreich, Johan Håstad, Jeol Friedman, Steven Rudich, and Roman Smolensky. The bit extraction problem or t-resilient functions. In *Proceedings of the 26th IEEE Symposium on Foundation of Computer Science*, pages 396–407, 1985.

[CL16] Eshan Chattopadhyay and Xin Li. Extractors for sumset sources. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 299–311. ACM, 2016.

[CPs07] Ran Canetti, Rafael Pass, and abhi shelat. Cryptography from sunspots: How to use an imperfect reference string. In *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07)*, pages 249–259, 2007.

[DG10] Matt DeVos and Ariel Gabizon. Simple affine extractors using dimension expansion. In *Proceedings of the 25th Annual IEEE Conference on Computational Complexity, CCC 2010, Cambridge, Massachusetts, USA, June 9-12, 2010*, pages 50–57. IEEE Computer Society, 2010.

[DGW09] Zeev Dvir, Ariel Gabizon, and Avi Wigderson. Extractors and rank extractors for polynomial sources. *Comput. Complex.*, 18(1):1–58, 2009.

[DLMV12] Yevgeniy Dodis, Adriana López-Alt, Ilya Mironov, and Salil Vadhan. Differential privacy with imperfect randomness. In Reihaneh Safavi-Naini and Ran Canetti, editors, *Advances in Cryptology – CRYPTO 2012*, pages 497–516, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[Dod00] Yevgeniy Dodis. *Exposure-resilient cryptography.* PhD thesis, Massachusetts Institute of Technology, 2000.

[DOPS04] Yevgeniy Dodis, Shien Jin Ong, Manoj Prabhakaran, and Amit Sahai. On the (im)possibility of cryptography with imperfect randomness. In *45th Annual IEEE Symposium on Foundations of Computer Science*, pages 196–205, 2004.

[DPP06] Yevgeniy Dodis, Krzysztof Pietrzak, and Bartosz Przydatek. Separating sources for encryption and secret sharing. In Shai Halevi and Tal Rabin, editors, *Theory of Cryptography*, pages 601–616, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[DPW18] Stefan Dziembowski, Krzysztof Pietrzak, and Daniel Wichs. Non-malleable codes. *J. ACM*, 65(4), April 2018.

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

[DSS01] Yevgeniy Dodis, Amit Sahai, and Adam Smith. On perfect and adaptive security in exposure-resilient cryptography. In Birgit Pfitzmann, editor, *Advances in Cryptology—EUROCRYPT 2001*, volume 2045 of *LNCS*, pages 301–324. Springer-Verlag, 2001.

[Dvi12] Zeev Dvir. Extractors for varieties. *Comput. Complex.*, 21(4):515–572, 2012.

[DY15] Yevgeniy Dodis and Yanqing Yao. Privacy with imperfect randomness. In Rosario Gennaro and Matthew Robshaw, editors, *Advances in Cryptology – CRYPTO 2015*, pages 463–482, Berlin, Heidelberg, 2015. Springer Berlin Heidelberg.

[EFF85] Paul Erdös, Peter Frankl, and Zoltán Füredi. Families of finite sets in which no set is covered by the union of r others. *Israel Journal of Mathematics*, 51(1-2):79–89, 1985.

[Eli72] Peter Elias. The efficient construction of an unbiased random sequence. *Ann. Math. Stat.*, 43(3):865–870, 1972.

[Gab11] Ariel Gabizon. *Deterministic Extractors for Affine Sources over Large Fields*, pages 33–53. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

[Gol11] Oded Goldreich. Three XOR-lemmas — An exposition. In Oded Goldreich, editor, *Studies in Complexity and Cryptography. Miscellanea on the Interplay between Randomness and Computation*, pages 248–272. Springer Berlin Heidelberg, 2011.

[GRS06] Ariel Gabizon, Ran Raz, and Ronen Shaltiel. Deterministic extractors for bit-fixing sources by obtaining an independent seed. *SIAM J. Comput.*, 36(4):1072–1094, 2006.

[GS08] Jens Groth and Amit Sahai. Efficient non-interactive proof systems for bilinear groups. In Nigel P. Smart, editor, *Advances in Cryptology - EUROCRYPT 2008*, volume 4965 of *LNCS*, pages 415–432. Springer-Verlag, 2008.

[KMS19] Ashutosh Kumar, Raghu Meka, and Amit Sahai. Leakage-resilient secret sharing against colluding parties. In *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 636–660, 2019.

[KRVZ11] Jesse Kamp, Anup Rao, Salil P. Vadhan, and David Zuckerman. Deterministic extractors for small-space sources. *J. Comput. Syst. Sci.*, 77(1):191–220, 2011.

[KvMS09] Jeff Kinne, Dieter van Melkebeek, and Ronen Shaltiel. Pseudorandom generators and typically-correct derandomization. In Irit Dinur, Klaus Jansen, Joseph Naor, and José D. P. Rolim, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, 12th International Workshop, APPROX 2009, and 13th International Workshop, RANDOM 2009, Berkeley, CA, USA, August 21-23, 2009. Proceedings*, volume 5687 of *Lecture Notes in Computer Science*, pages 574–587. Springer, 2009.

[KZ07] Jesse Kamp and David Zuckerman. Deterministic extractors for bit-fixing sources and exposure-resilient cryptography. *SIAM J. Comput.*, 36(5):1231–1247, 2007.

[LCG$^+$20] Fuchun Lin, Mahdi Cheraghchi, Venkatesan Guruswami, Reihaneh Safavi-Naini, and Huaxiong Wang. Leakage-resilient secret sharing in non-compartmentalized models. In Yael Tauman Kalai, Adam D. Smith, and Daniel Wichs, editors, *1st Conference on Information-Theoretic Cryptography (ITC 2020)*, volume 163 of *Leibniz International*

Proceedings in Informatics (LIPIcs), pages 7:1–7:24, Dagstuhl, Germany, 2020. Schloss Dagstuhl–Leibniz-Zentrum für Informatik.

[Li11] Xin Li. A new approach to affine extractors and dispersers. In *Proceedings of the 26th Annual IEEE Conference on Computational Complexity, CCC 2011, San Jose, California, USA, June 8-10, 2011*, pages 137–147. IEEE Computer Society, 2011.

[LLS89] David Lichtenstein, Nathan Linial, and Michael Saks. Some extremal problems arising from discrete control processes. *Combinatorica*, 9(3):269–287, 1989.

[MP91] James L. McInnes and Benny Pinkas. On the impossibility of private key cryptography with weakly random keys. In Alfred J. Menezes and Scott A. Vanstone, editors, *Advances in Cryptology-CRYPTO' 90*, pages 421–435, Berlin, Heidelberg, 1991. Springer Berlin Heidelberg.

[MPSW20] Hemanta Maji, Anat Paskin-Cherniavsky, Tom Suad, and Mingyuan Wang. On leakage-resilient secret sharing. Cryptology ePrint Archive, Report 2020/1517, 2020. `https://eprint.iacr.org/2020/1517`.

[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.

[NW94] Noam Nisan and Avi Wigderson. Hardness vs randomness. *Journal of Computer and System Sciences*, 49(2):149–167, 1994.

[Rao09] Anup Rao. Extractors for low-weight affine sources. In *Proceedings of the 24th Annual IEEE Conference on Computational Complexity, CCC 2009, Paris, France, 15-18 July 2009*, pages 95–101. IEEE Computer Society, 2009.

[Röd85] Vojtěch Rödl. On a packing and covering problem. *European Journal of Combinatorics*, 6(1):69–78, 1985.

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

[Sha11] Ronen Shaltiel. Weak derandomization of weak algorithms: Explicit versions of yao's lemma. *Comput. Complex.*, 20(1):87–143, 2011.

[SV86] Miklos Santha and Umesh V. Vazirani. Generating quasi-random sequences from semi-random sources. *Journal of Computer and System Sciences*, 33(1):75–87, 1986.

[SV19] Akshayaram Srinivasan and Prashant Nalini Vasudevan. Leakage resilient secret sharing and applications. In Alexandra Boldyreva and Daniele Micciancio, editors, *Advances in Cryptology – CRYPTO 2019*, pages 480–509, Cham, 2019. Springer International Publishing.

[SW18] Douglas R. Stinson and Ruizhong Wei. Combinatorial repairability for threshold schemes. *Des. Codes Cryptogr.*, 86(1):195–210, 2018.

[Tre01] Luca Trevisan. Extractors and pseudorandom generators. *J. ACM*, 48(4):860–879, July 2001.

[TV00] Luca Trevisan and Salil Vadhan. Extracting randomness from samplable distributions. In *41st Annual Symposium on Foundations of Computer Science*, pages 32–42, Redondo Beach, California, November 2000. IEEE.

[Vaz87] U. Vazirani. Efficiency considerations in using semi-random sources. In *Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing*, STOC '87, page 160–168, New York, NY, USA, 1987. Association for Computing Machinery.

[vN51] John von Neumann. Various techniques used in connection with random digits. *Monte Carlo Method, U.S. National Bureau of Stadndarts, Applied Mathematics Series*, 12:36–38, 1951.

[VV85] Umesh V. Vazirani and Vijay V. Vazirani. Random polynomial time is equal to semi-random polynomial time. In *Proceedings of the Twenty-Sixth Annual ACM Symposium on the Theory of Computing*, pages 417–428, Montréal, Québec, Canada, 23–25 May 1985.

[Yeh11] Amir Yehudayoff. Affine extractors over prime fields. *Comb.*, 31(2):245–256, 2011.

[YL18] Yanqing Yao and Zhoujun Li. Differential privacy with bias-control limited sources. *IEEE Transactions on Information Forensics and Security*, 13(5):1230–1241, 2018.

[Zuc96] David Zuckerman. Simulating bpp using a general weak random source. *Algorithmica*, 16(4/5):367–391, 1996.

A Proof of Lemma 3

Fix an (n,k)-source X and pick a function $F:\{0,1\}^n\to\{0,1\}^m$ with $m\leq k-2\log(1/\varepsilon)$ uniformly at random. It suffices to bound the probability that

$$|\Pr[F(X)\in\mathcal{T}]-\mu(\mathcal{T})|\leq\varepsilon$$

holds for every set $\mathcal{T}\subseteq\{0,1\}^m$, where $\mu(\mathcal{T})=|\mathcal{T}|/2^m$ denotes the density of $\mathcal{T}$. Fix such a set $\mathcal{T}$, and let $Z_x=\Pr[X=x]\cdot\mathbf{1}_{F(x)\in\mathcal{T}}$. Then, we have $\Pr[F(X)\in\mathcal{T}]=\sum_{x\in\{0,1\}^n}Z_x$ and $\mathbb{E}\left[\sum_{x\in\{0,1\}^n}Z_x\right]=\mu(\mathcal{T})$. As a result, since $Z_x\in[0,\Pr[X=x]]$ for all $x\in\{0,1\}^n$, Hoeffding's inequality[4] implies that

$$\begin{aligned}\Pr\left[\left|\sum_{x\in\{0,1\}^n}Z_x-\mu(\mathcal{T})\right|>\varepsilon\right]&\leq 2\cdot\exp\left(-\frac{2\varepsilon^2}{\sum_{x\in\{0,1\}^n}\Pr[X=x]^2}\right)\\&\leq 2\cdot e^{-2\varepsilon^2 2^k}.\end{aligned}$$

The last inequality follows from the fact that $\sum_{x\in\{0,1\}^n}\Pr[X=x]^2\leq\max_{x\in\{0,1\}^n}\Pr[X=x]\leq 2^{-k}$, since X is an (n,k)-source. Finally, a union bound over all 2^{2^m} sets $\mathcal{T}\subseteq\{0,1\}^m$ shows that the event in question holds with probability at least

$$1-2\cdot 2^{2^m}\cdot e^{-2\varepsilon^2 2^k}\geq 1-2e^{-\varepsilon^2 2^k}$$

over the choice of F, given the upper bound on m.

[4] The version of Hoeffding's inequality we use here states that if $X_1,\dots,X_N$ are independent random variables and $X_i\in[m_i,M_i]$ for each i, then $\Pr\left[\left|\sum_{i=1}^N X_i-\mu\right|>\varepsilon\right]\leq 2\cdot\exp\left(-\frac{2\varepsilon^2}{\sum_{i=1}^N(M_i-m_i)^2}\right)$, where $\mu=\mathbb{E}\left[\sum_{i=1}^N X_i\right]$.