Leakage-Resilient Secret Sharing in Non-Compartmentalized Models

Fuchun Lin
Department of Electrical and Electronic Engineering, Imperial College London, UK
flin@ic.ac.uk

Mahdi Cheraghchi
EECS Department, University of Michigan, Ann Arbor, MI, USA
http://mahdi.ch
mahdich@umich.edu

Venkatesan Guruswami
Computer Science Department, Carnegie Mellon University, Pittsburgh, PA, USA
venkatg@cs.cmu.edu

Reihaneh Safavi-Naini
Department of Computer Science, University of Calgary, CA
rei@ucalgary.ca

Huaxiong Wang
Division of Mathematical Sciences, School of Physical and Mathematical Sciences,
Nanyang Technological University, Singapore, SG
hxwang@ntu.edu.sg

Abstract

Leakage-resilient secret sharing has mostly been studied in the *compartmentalized* models, where a leakage oracle can arbitrarily leak bounded number of bits from all shares, provided that the oracle only has access to a bounded number of shares when the leakage is taking place. We start a systematic study of leakage-resilient secret sharing against *global* leakage, where the leakage oracle can access the full set of shares simultaneously, but the access is restricted to a special class of leakage functions. More concretely, the adversary can corrupt several players and obtain their shares, as well as applying a leakage function from a specific class to the full share vector. We explicitly construct such leakage-resilient secret sharing with respect to affine leakage functions and low-degree multi-variate polynomial leakage functions, respectively. For affine leakage functions, we obtain schemes with threshold access structure that are leakage-resilient as long as there is a substantial difference between the total amount of information obtained by the adversary, through corrupting individual players and leaking from the full share vector, and the amount that the reconstruction algorithm requires for reconstructing the secret. Furthermore, if we assume the adversary is non-adaptive, we can even make the secret length asymptotically equal to the difference, as the share length grows. Specifically, we have a threshold scheme with parameters similar to Shamir's scheme and is leakage-resilient against affine leakage. For multi-variate polynomial leakage functions with degree bigger than one, our constructions here only yield ramp schemes that are leakage-resilient against such leakage. Finally, as a result of independent interest, we show that our approach to leakage-resilient secret sharing also yields a competitive scheme compared with the state-of-the-art construction in the compartmentalized models.

2012 ACM Subject Classification Security and privacy → Cryptography; Security and privacy → Information-theoretic techniques; Theory of computation → Expander graphs and randomness extractors

Keywords and phrases Leakage-resilient cryptography, Secret sharing scheme, Randomness extractor

Digital Object Identifier 10.4230/LIPIcs.ITC.2020.7

Related Version Full version at https://arxiv.org/abs/1902.06195.

1st Conference on Information-Theoretic Cryptography (ITC 2020).
Editors: Yael Tauman Kalai, Adam D. Smith, and Daniel Wichs; Article No. 7; pp. 7:1–7:24
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Funding *Fuchun Lin*: The research of Fuchun Lin was supported in part by EPSRC grant EP/S021043/1 and Singapore Ministry of Education under Research Grant MOE2016-T2-2-014(S) and RG133/17(S).
Venkatesan Guruswami: The research of Venkatesan Guruswami was supported in part by United States NSF grants CCF-1422045 and CCF-1563742.
Reihaneh Safavi-Naini: The research of Reihaneh Safavi-Naini was in part supported by Natural Sciences and Engineering Research Council of Canada, Discovery Grants Program.
Huaxiong Wang: The research of Huaxiong Wang was supported by Singapore Ministry of Education under Research Grant MOE2016-T2-2-014(S) and RG133/17(S).

Acknowledgements Part of this work was done when the first author was in Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University. We would like to thank Vipul Goyal for helpful discussions. We also would like to thank the anonymous reviewers for comments.

1 Introduction

Secret sharing, introduced independently by Blakley [11] and Shamir [42], is a fundamental cryptographic primitive with far-reaching applications; e.g., a major tool in secure multiparty computation (cf. [19]). The goal in secret sharing is to encode a secret s into a number of *shares* $\mathsf{c}_1, \ldots, \mathsf{c}_\mathsf{n}$ that are distributed among a set $[\mathsf{n}] = \{1, \ldots, \mathsf{n}\}$ of players such that the access to the secret through collaboration of players can be accurately controlled. An *authorized* subset of players is a set $A \subseteq [\mathsf{n}]$ such that the shares with indices in A can be pooled together to reconstruct the secret s. On the other hand, A is an *unauthorized* subset if the knowledge of the shares with indices in A reveals no information about the secret. The set of authorized and unauthorized sets define an access structure, where the most widely used is the so-called *threshold* structure. A threshold secret sharing scheme is defined with respect to a threshold t and satisfies the following property: Any set $A \subseteq [\mathsf{n}]$ with $|A| < \mathsf{t}$ is an unauthorized set and any set $A \subseteq [\mathsf{n}]$ with $|A| \geq \mathsf{t}$ is an authorized set. Threshold secret sharing with threshold t is also called t-out-of-n secret sharing. Any threshold secret sharing scheme sharing m-bit secrets necessarily requires shares of length (in bits) at least m, and Shamir's scheme attains this lower bound [44]. The *information ratio* defined as the ratio of the secret length to the maximum share length measures the storage efficiency of a secret sharing scheme.

Leakage-resilience of secret sharing considers an adversary that has certain form of extra information about the shares beyond the unauthorized sets of shares and studies how to keep the secret remain private. Dziembowski and Pietrzak [23] developed an n-out-of-n *intrusion-resilient* secret sharing scheme using methods from the *bounded retrieval model.* The shares of such secret sharing schemes are made artificially large so that protocols with bounded communication complexity can not retrieve them completely. The reconstruction procedure is interactive requiring the players to exchange r short messages, while the adversary can also attack in rounds but is restricted to at most $r - 1$ rounds. The idea is to exploit the fact that there exist functions which can be efficiently computed interactively using low communication complexity in r, but not in $r - 1$ rounds. Davì, Dziembowski and Venturi [20] constructed the first 2-out-of-2 secret sharing scheme that statistically hides the secret even after an *adaptive* adversary executes an arbitrary leakage protocol on the two shares with bounded communication. The relation of their scheme to a 2-out-of-2 intrusion-resilient secret sharing is that the reconstruction of intrusion-resilient secret sharing needs to access only small part of the share while their scheme does not have such restriction.

Study of leakage-resilience of secret sharing has recently gained general attention. Benhamouda, Degwekar, Ishai and Rabin [10] initiated a systematic investigation on the leakage-resilience of secret sharing in the *local leakage model*. The adversary in the local leakage model *non-adaptively* obtains a bounded number (less than $\mathtt{t}-1$) of full shares and leaks from all the other shares individually, each share using an *arbitrary* leakage function bounded only by its output length. The authors focus on investigating whether standard secret sharing schemes, such as additive secret sharing and Shamir's scheme are already leakage-resilient with respect to such local leakage. They showed that these $\mathtt{t}$-out-of-$\mathtt{n}$ secret sharing schemes, if the base field is a large prime, are leakage-resilient for some limited parameter settings. This is in sharp contrast to the results of Guruswami and Wootters [29] from an orthogonal study of Reed-Solomon codes as self-repairable codes, which translated into the setting of leakage-resilient secret sharing effectively shows that by leaking one bit using a *linear* function from each share, the secret of Shamir's scheme over finite field with characteristic 2 can be completely reconstructed. While starting with standard secret sharing schemes has the advantage of inheriting the nice algebraic structure and optimality (information ratio 1) from those schemes, the large prime field requirement for the choice of base field and the limited available leakage parameters are limiting the applications of such leakage-resilient secret sharing schemes. Concurrently and independently, Goyal and Kumar [27], motivated by the task of constructing *non-malleable secret sharing*, a new primitive of tamper-resilient secret sharing they put forward (see more details in *Related works*), constructed a 2-out-of-$\mathtt{n}$ leakage-resilient secret sharing scheme in the local leakage model through building up from a 2-out-of-2 leakage-resilient secret sharing scheme, which is in turn constructed from the inner product function. The authors showed that the same 2-out-of-$\mathtt{n}$ leakage-resilient secret sharing scheme is in fact leakage-resilient against a slightly stronger leakage adversary than the one in the local leakage model, as they took their non-malleable secret sharing results to general access structures and hence need the strengthened local leakage model [28]. Since then leakage-resilient secret sharing in the local leakage model is widely studied mostly with connection to non-malleable secret sharing [7, 43, 1]. Most of the known constructions of leakage-resilient secret sharing take a compiler approach with various features to transform a secret sharing scheme into a leakage-resilient one. Note that in this approach, local leakage-resilience is enabled through pumping independent randomness into each share and hence the obtained schemes do not have *full reconstruction* [38], which is a property that requires shares that are enough to reconstruct the secret also uniquely reconstruct the full share vector. The advantage of secret sharing without full reconstruction is that one can pump unlimited independent randomness into each share and allow the adversary to leak unlimited amount of information except some finite amount (hence it is possible for the scheme to tolerate asymptotic leakage rate 1) [7, 43]. On the other hand, the extra independent randomness eventually results in a bigger share size than necessary (small information ratio).

The local leakage adversary considered by Srinivasan and Vasudevan [43] is partially adaptive in the sense that for $\mathtt{t}$-out-of-$\mathtt{n}$ threshold schemes, the choice of each local leakage function can be based on the value of the $t-2$ shares. Kumar, Meka, and Sahai [31] proposed an *adaptive joint leakage model* as opposed to the non-adaptive and partially adaptive local leakage models. They defined a *bounded collusion protocol* as a communication complexity problem that is tailored for secret sharing. A $\mathtt{t}-1$-bounded collusion protocol is one that runs in multiple rounds, each round involves at most $\mathtt{t}-1$ players and the output of that round can depend on the inputs of the involved players and all previous transcript. This can be seen as a generalisation of the communication complexity problem for the 2-out-of-2 secret sharing of [20]. Faonio and Venturi[25] (under computational assumption) considered a *noisy* leakage model that, instead of bounding the output length of the leakage functions, bounds the min-entropy of the share conditioned on the output of the leakage function.

All the leakage models mentioned are more or less based on a compartmentalized assumption that assumes the adversary does not have access to the full share vector at one time, while nevertheless allows the adversary to apply arbitrary leakage functions to the set of shares that are accessed. In this work, we are mainly interested in leakage-resilience of secret sharing against *global* type of leakage, where a leakage adversary can access the full set of shares simultaneously, but is restricted in the type of access that is admissible. Bogdanov, Ishai, Viola and Williamson [12] obtained several instances of leakage-resilient secret sharing in such global leakage models (see more details in *Related works*). For example, it is shown that Shamir's scheme over finite fields of characteristic 2 is leakage-resilient against constant depth polynomial size circuits (AC^0) and sign polynomials of degree 2, unless the threshold t is at most polylogarithmic in n. But as mentioned before, these schemes are not leakage resilient against (local) linear leakage [29]). As far as we know, this is the only previous work that studies global leakage-resilience of secret sharing.

Our contributions. We start a systematic study of Leakage-Resilient Secret Sharing (LR-SS) with respect to a class $\mathcal{L}$ of global leakage functions. To stay compatible with the well established local leakage models, we define a $(\mathcal{L}, \beta, \theta, \varepsilon)$-leakage resilient t-out-of-n secret sharing scheme. The parameter β is the bound on the total amount of leakage (measured in bits) through leakage functions chosen from $\mathcal{L}$. The parameter θ is the bound on the number of corrupted players. The parameter ε is the leakage-resilience error (measured in statistical distance). We see that if we define a class $\mathcal{L}_{\mathsf{local}}$ of functions that mimic a local leakage adversary that leaks ℓ bits arbitrarily from each share, then we recover the local leakage model as a $(\mathcal{L}_{\mathsf{local}}, \ell(\mathsf{n}-\theta), \theta, \varepsilon)$-leakage resilient t-out-of-n secret sharing scheme. We focus on the case when the share is an element of a finite field $\mathbb{F}_q$ of characteristic 2, which is of greater practical importance. We also consider adaptive as well as non-adaptive adversary, which distinguishes between the cases where the adversary can or can not base the subsequent leakage on previously obtained transcript. The exposition is focused on t-out-of-n secret sharing and all schemes constructed satisfy the *full reconstruction* property, in particular, any t shares determine the value of the remaining $\mathsf{n}-\mathsf{t}$ shares. We place a special emphasize on achieving the best possible leakage tolerance as well as optimal information ratio, for some cases.

We start with the class $\mathcal{L}_{\mathsf{affine}}$ of affine leakage functions over $\mathbb{F}_2$. This class of leakage functions is interesting because of the following. In practice, the shares of a secret sharing scheme are to be sent to the players in private. Suppose we send the shares over a network and some of the packets sent by intermediate servers are leaked. If only routing is used, we have the compartmentalized model where individual packets are observed by the adversary. Suppose in a real-life application where the network is utilizing linear network coding [26], then effectively we are going to have global affine leakage. Another example is the type of leakage for Shamir's scheme over finite field of characteristic 2 implied in the result of [29]. It is in effect a subset of the linear leakage functions, which is in turn subsumed by affine leakage functions. Our results for affine type of leakage are summarized as follows.

▶ **Theorem** (informal summary of Theorem 14 and Theorem 17). *Let $\mathcal{L}_{\mathsf{affine}}$ be the class of affine leakage functions over $\mathbb{F}_2$. Let $0 < \xi \leq 1$ be any real number. There is a t-out-of-n secret sharing scheme over $\mathbb{F}_q$ that is simultaneously $(\mathcal{L}_{\mathsf{affine}}, \beta, \theta, \varepsilon)$-leakage resilient for all $0 \leq \beta < (\mathsf{t}-\xi)\log q$ and $\theta \in \{0, 1, \ldots, \mathsf{t}-1\}$ such that $\theta + \frac{\beta}{\log q} \leq \mathsf{t}-\xi$, where the share size q is a large enough power of 2 and the secret length $m = \Omega(\log q)$. In particular, if the adversary is non-adaptive, the result can be strengthened to have secret length $m = \xi \log q - o(\log q)$.*

The leakage tolerance in the above construction is the best one can hope for in the following sense. The amount of information given to the adversary is $(\mathsf{t}-\xi)\log q$ bits in total (θ shares through corrupted players and β bits through global leakage) while with $\mathsf{t}\log q$ bits of information the secret (in fact the full share vector) can be completely reconstructed. Our results show that as long as $\xi > 0$, we can construct a t-out-of-n secret sharing scheme with constant information ratio that is universally leakage-resilient against affine leakage for all combinations of β and θ as long as the total amount satisfies $\theta\log q + \beta \leq (\mathsf{t}-\xi)\log q$.

The strengthened result in the case of non-adaptive adversary is furthermore optimal in the sense that the asymptotic information ratio is exactly ξ. For example, let $\xi = 1$. We have a t-out-of-n secret sharing scheme with asymptotic information ratio 1. As a plain secret sharing scheme, this scheme is almost as good as the Shamir's secret sharing scheme, except that Shamir's scheme has perfect privacy $\varepsilon = 0$ and information ratio exactly 1. On the other hand, Shamir's scheme over finite field of characteristic 2 is not leakage-resilient against even leaking one bit from each share using a linear leakage function [29], while our scheme is universally leakage-resilient against affine leakage for all combinations of β and θ as long as the total amount of information given to the adversary satisfies $\theta\log q + \beta \leq (\mathsf{t}-1)\log q$.

We venture a little bit further and explore leakage-resilience of secret sharing against stronger algebraic type of global leakage. These leakage models can occur in more sophisticated applications of secret sharing schemes such as the secure multiparty computation. Private inputs from the participants are protected by secret sharing schemes and algebraic computations on the private inputs are done through computations on the individual shares. Leakage that occurs throughout the duration of the computation is a good example for global multi-variate polynomial leakage model.

Let $\mathcal{L}_{d\text{-poly}}$ denote the class of multi-variate polynomials with degree at most d over $\mathbb{F}_2$. The LR-SS with respect to $\mathcal{L}_{d\text{-poly}}$ we construct is a slightly weaker form of t-out-of-n secret sharing called *ramp scheme*. There is a second threshold called privacy threshold $\mathsf{t}-g$ and the gap between the two thresholds is denoted by g. When $g = 1$, we recover the threshold scheme. For ramp schemes, it is possible to share a secret longer than the share length.

▶ **Theorem** (informal summary of Theorem 20). *Let $\mathcal{L}_{d\text{-poly}}$ be the class of multi-variate polynomials over $\mathbb{F}_2$ with degree at most d. Let g be an integer satisfying $\frac{g}{\mathsf{t}} > 1 - \frac{1}{c_d}$, where $c_d = \Theta(d^2 4^d)$ is a constant determined by d. There is a t-out-of-n secret sharing scheme over $\mathbb{F}_q$ with threshold gap g over the finite field $\mathbb{F}_q$ that is simultaneously $(\mathcal{L}_{d\text{-poly}}, \beta, \theta, \varepsilon)$-leakage resilient for all $\beta = (\log q)\phi$ and $\theta \leq \mathsf{t}-g$ satisfying $0 \leq \phi \leq \mathsf{t}-g-\theta$, where the share size q is a large enough power of 2 and the secret length $m = \Omega(\log q)$.*

The threshold gap g in the above result is mostly bigger than 1. We note that smaller threshold gap can be achieved using our construction if a building block of better parameters is constructed in the future.

Finally, as a result of independent interest, we consider an instantiation of our generic construction for global leakage to obtain LR-SS in local leakage model with a partially adaptive adversary (currently strongest local leakage model) [43]. It is known that the number of full shares from corrupted players that a partially adaptive adversary has before the leakage functions are chosen is at most $\mathsf{t}-2$ [1]. Our LR-SS allow the number θ of

[1] As argued in [43], based on $\mathsf{t}-1$ shares, the partially adaptive adversary can invoke the reconstruction algorithm to define a leakage function that takes a new share (not among the $\mathsf{t}-1$ shares) as input, pooling with the $\mathsf{t}-1$ shares to reconstruct the secret, and outputs one bit of the secret.

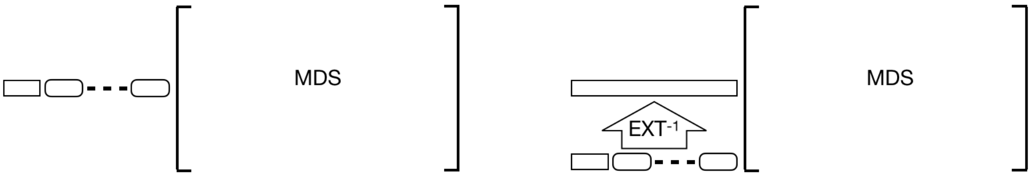

Figure 1 Shamir's scheme versus our generic construction.

corrupted players to be equal to $\mathsf{t}-1$. But due to the known impossibility result, when $\theta=\mathsf{t}-1$, the partially adaptive adversary is required to choose leakage functions base on at most $\mathsf{t}-2$ out of the $\mathsf{t}-1$ corrupted players.

▶ **Theorem** (informal summary of Theorem 24)**.** *Let $\mathcal{L}_{\mathsf{local}}$ be the class of local leakage functions. Let $\xi>0$ be a small real number. There is a t-out-of-n secret sharing scheme over $\mathbb{F}_q$ that is simultaneously $(\mathcal{L}_{\mathsf{local}},\beta,\theta,\varepsilon)$-leakage resilient against a partially adaptive adversary that corrupts θ players for $\theta\in\{0,1,\ldots,\mathsf{t}-2\}$ and, based on the shares of the θ players, chooses $\mathsf{n}-\theta$ arbitrary leakage functions each leaks $\ell=\frac{(\mathsf{t}-\theta-8\xi)\log q}{(\mathsf{n}-\theta)(1+5\xi)}$ bits for the remaining $\mathsf{n}-\theta$ shares, where the share size q is a large enough power of 2 and the secret length $m=\frac{\xi}{1+5\xi}\log q-o(\log q)$. When $\theta=\mathsf{t}-1$, the partially adaptive adversary can only choose the $\mathsf{n}-\mathsf{t}+1$ arbitrary leakage functions based on $\mathsf{t}-2$ shares*

Our scheme achieves positive information ratio, which is an absolute constant, while the scheme in [43] achieves information ratio $\Omega(1/\mathsf{n})$, which depends on n. On the other hand, when $\theta=\mathsf{t}-1$, our scheme achieves an asymptotic leakage rate of $\frac{\mathsf{t}-\theta}{\mathsf{n}-\theta}=\frac{1}{\mathsf{n}-\mathsf{t}+1}$, while the scheme in [43] achieves asymptotic leakage rate 1, which is only possible for LR-SS without full reconstruction. For LR-SS with full reconstruction, the asymptotic leakage rate of $\frac{1}{\mathsf{n}-\mathsf{t}+1}$ is close to the optimal $\frac{1}{\mathsf{n}-\mathsf{t}}$, according to the bound developed for local leakage model LR-SS that satisfy full reconstruction property [38].

Technical overview. All these results are obtained from a generic construction that can be interpreted as follows. We preprocess the secret and the randomness of the sharing algorithm of Shamir's scheme over finite field of characteristic 2 before inputting them into the sharing algorithm to enable leakage-resilience. As illustrated in **Figure 1**, Shamir's scheme can be seen as directly concatenating the secret (as a finite field element of $\mathbb{F}_{2^m}$) with $\mathsf{t}-1$ independent uniformly random finite field elements of $\mathbb{F}_{2^m}$ and multiply with a $\mathsf{t}\times\mathsf{n}$ Maximum Distance Separable (MDS) matrix over the finite field $\mathbb{F}_{2^m}$. Our construction is to input the m bits secret as binary string and the $m(\mathsf{t}-1)$ bits independent uniform randomness into the inverter of an invertible randomness extractor to obtain an $m\mathsf{t}$ bits output, which is then interpreted as a vector in $\mathbb{F}_{2^m}^{\mathsf{t}}$ and multiply with the $\mathsf{t}\times\mathsf{n}$ MDS matrix over the finite field $\mathbb{F}_{2^m}$.

A randomness extractor takes an entropy source as input and output a close to uniform distribution over a smaller space. For structured entropy source, for example, the source is a flat distribution over an affine subspace of the universal space $\{0,1\}^n$, there exists a single function that turns any such source distribution with enough entropy into a close to uniform distribution over $\{0,1\}^m$. Such extractors are called *seedless extractors.* On the other hand, for arbitrary entropy source, we need a family of functions that are labeled by *seeds* and uniformly choose one function among them to extract from such entropy source. These extractors are called *seeded extractors.* A seeded extractor is *strong* if given a source distribution, almost all functions in the family can extract close to uniform output from it. A seedless extractor is called invertible if there exists an efficient function (called *inverter*) that takes a vector in $\{0,1\}^m$ and some randomness as input and outputs a random pre-image in $\{0,1\}^n$. A seeded extractor is invertible if all the functions in the family are invertible. The inverter of a seeded extractor first samples a uniform seed and then use the inverter of the function corresponding to the seed to invert the vector in $\{0,1\}^m$.

The intuition of our generic construction illustrated in **Figure 1** is that randomness extractors can make the secret and output of the leakage functions independent, hence provide privacy and leakage-resilience. Assume that the secret is uniform. Then the random pre-image outputted by the inverter has uniform distribution. As long as this uniformly distributed pre-image conditioned on the output of the leakage functions has enough entropy (and with the right structure in the case when a seedless extractor is used), the secret remains uniform.

Our results in this work are obtained by using different randomness extractors in the generic construction.

- Affine leakage, non-adaptive adversary.
  We use a *linear* strong seeded extractor that extracts all the randomness. A seeded extractor is linear if every function in the family is linear. The linearity of the extractor function together with the fact that the source distributions induced by affine leakage functions are flat over some affine subspaces allows us to claim that almost all functions in the family output exactly uniform distribution. This observation plays an important role in achieving optimal information ratio.
- Affine leakage, adaptive adversary.
  We use an invertible affine extractor that can extract from an arbitrary fraction of entropy and output length is a constant fraction of the input length with exponentially small extractor error. In fact, we additionally apply a series of sophisticated optimization techniques to improve the parameters of the obtained scheme.
- Low-degree multi-variate polynomial leakage.
  There are a few challenges that have to be overcome when the degree of multi-variate polynomials goes beyond 1. For an affine leakage function f, the number of solutions to $f(x) = a$ for any a that admits non-empty solutions is determined by f and independent of the value a. This gives a natural bound on the min-entropy of a random x conditioned on the value of $f(x)$. When the degree is bigger than 1, we no longer have such nice structure. We identify the *average conditional min-entropy* [21] as the "right" entropy measure and use it to derive a lower bound with respect to the output length of the leakage function. Average conditional min-entropy is usually used in combination with the *average-case* strong seeded extractors. We then define a seedless analogue of average-case strong seeded extractors and verify that the explicit seedless extractor we use in our construction is in fact an *average-case seedless* extractor. We think these might have independent interest in other application scenarios of seedless extractors. There is also the challenge of explicitly constructing such extractors with parameters as good as affine extractors, which we leave it as an interesting open question.
- Local leakage, partially adaptive adversary.
  We use an average-case strong seeded extractor with exponentially small error. Such extractors necessarily requires a long seed. That is one of the reasons that we no longer have optimal information ratio (in the instantiation for affine leakage non-adaptive adversary, the seed length is negligible compared to the extractor input). There is one more modification. The random seed chosen by the inverter of the seeded extractor is no longer directly appended to the random pre-image to be encoded using the MDS code.

We use a Shamir's scheme to share this random seed into n shares. The final share of our LR-SS consists of one MDS codeword component and one share of the random seed. Through this modification, we are able to prove that the partially adaptive adversary is not able to make the leakage depend on the seed, which then remains independent of the source and leakage-resilience follows from the definition of the seeded extractor. The fact that we are using a seeded extractor with seed length bigger than the output length may seem counter-intuitive. But since the goal here is to provide leakage-resilience instead of extracting randomness, we do not have to force the extractor seed to be shorter than the extractor output.

Related works. The works most related to ours is the following. Bogdanov, Ishai, Viola and Williamson [12] initiated the study of *t-wise indistinguishability* as a relaxation of the well studied notion of *t-wise independence* and considered a pair of such distributions as a statistical secret sharing scheme sharing one bit secret. Through discovering the fact that t-wise indistinguishability implies leakage-resilience against a class of global leakage functions with *approximate degree* smaller than a quantity determined by t, they obtained the first instance of LR-SS in the global leakage model. On one hand, the leakage-resilience is implied by the privacy of the secret sharing shceme (leakage resilience for free). On the other hand, the achieved leakage-resilience is restricted to very limited leakage functions.

As distant related works, we briefly discuss a large body of works (with many overlapping references) on tamper-resilience of secret sharing. Goyal and Kumar [27] initiated the systematic study of Non-Malleable Secret Sharing (NM-SS). A secret sharing is called a NM-SS against certain type of tampering if a tampering either results in the original secret or a random secret whose distribution depends only on the particular tampering function that was applied and independent of the original secret. The recent interests in constructing LR-SS is partially due to its role as an important building block in the constructions of NM-SS. The basic tampering model for NM-SS is the *independent* tampering model where each share is arbitrarily tampered independent of each other [27, 28, 43, 7, 1, 25]. The study of NM-SS is in turn closely related to Non-Malleable Codes (NMC) proposed by Dziembowski, Pietrzak and Wichs [24]. The independent tampering model of NM-SS is corresponding to the *split-state* tampering models studied for NMC. In particular, according to [4], 2-split state NMC's are also 2-out-of-2 NM-SS's [37, 22, 3, 4, 2, 15, 34, 35, 17]. Non-compartmentalized models are well studied in the literature of NMC. Agrawal et.al [5, 6] initiated the study of non-compartmentalized tampering models for NMC. They considered non-malleability against permutation composed with *bit-wise independent tampering*, and showed that non-malleable codes in such a tampering model transform non-malleable bit-commitments into a non-malleable string-commitment. There have been other non-compartmentalized tampering families studied for non-malleable codes: *bounded fan-in* circuits [14], affine functions [16], small-depth circuits [8] and decision tree [9]. Our global affine leakage model can be seen as a leakage-resilient analogue of the non-compartmentalized affine tampering model in NMC. There is not yet a tamper-resilient analogue for our low-degree multi-variate polynomial leakage model.

Paper organisation. The rest of the paper is organised as follows. Section 2 contains the definitions of various randomness extractors that appear in this work. Section 3 starts with a general definition of LR-SS that extends the local LR-SS to include the global leakage models. It is followed by a detailed description of our generic construction. In Section 4, we study the affine leakage models and obtain two sets of results for adaptive adversary and non-adaptive adversary, respectively. Section 5 is devoted to the low-degree multi-variate polynomial leakage model. In Section 6, we extend our results to give a LR-SS in local leakage model. We conclude our results in Section 7.

2 Preliminaries

The *statistical distance* of two random variables (their corresponding distributions) is defined as follows. For $\mathsf{X}, \mathsf{Y} \leftarrow \Omega$,

$$\mathsf{SD}(\mathsf{X};\mathsf{Y}) = \frac{1}{2}\sum_{\omega\in\Omega} |\Pr(\mathsf{X}=\omega) - \Pr(\mathsf{Y}=\omega)|.$$

We say X and Y are ε-close (denoted $\mathsf{X} \overset{\varepsilon}{\sim} \mathsf{Y}$) if $\mathsf{SD}(\mathsf{X},\mathsf{Y}) \leq \varepsilon$.

We use various types of randomness extractors in our constructions. Randomness extractors extract close to uniform bits from input sequences that are not uniform but have some guaranteed entropy. See [39] and references there in for more information about randomness extractors.

A *randomness source* is a random variable with lower bound on its min-entropy, which is defined by $\mathsf{H}_\infty(\mathsf{X}) = -\log \max_{\mathsf{x}}\{\mathsf{Pr}[\mathsf{X}=\mathsf{x}]\}$. We say a random variable $\mathsf{X} \leftarrow \{0,1\}^n$ is a (n,k)*-source*, if $\mathsf{H}_\infty(\mathsf{X}) \geq k$.

For well structured sources, there exist deterministic functions that can extract close to uniform bits. An *affine* (n,k)-source is a random variable that is uniformly distributed on an affine translation of some k-dimensional sub-space of $\{0,1\}^n$. Let U_m denote the random variable uniformly distributed over $\{0,1\}^m$.

▶ **Definition 1.** *A function* $\mathsf{aExt}\colon \{0,1\}^n \to \{0,1\}^m$ *is an affine* (k,ε)*-extractor if for any affine* (n,k)*-source* X*, we have*

$$\mathsf{SD}(\mathsf{aExt}(\mathsf{X});\mathsf{U}_m) \leq \varepsilon.$$

We will use Bourgain's affine extractor (or the alternative [33] due to Li) in our constructions.

▶ **Lemma 2** ([13])**.** *For every constant* $0 < \mu \leq 1$*, there is an explicit affine extractor* $\mathsf{aExt}\colon \{0,1\}^n \to \{0,1\}^m$ *for affine* $(n, n\mu)$*-sources with output length* $m = \Omega(n)$ *and error at most* $2^{-\Omega(n)}$*.*

An algebraic set is a set of common zeros of one or more multivariate polynomials defined over a finite field. An *algebraic source* is a random variable distributed uniformly over an algebraic set. Algebraic sources are a natural generalization of affine sources that have been widely studied.

▶ **Definition 3.** *A function* $\mathsf{aExt}\colon \{0,1\}^n \to \{0,1\}^m$ *is an algebraic* (k,d,ε)*-extractor if for any degree-*d *algebraic source* $\mathsf{U}_{\mathcal{V}}$ *with algebraic set* $\mathcal{V} \subseteq \{0,1\}^n$ *and* $|\mathcal{V}| \geq 2^k$*, we have*

$$\mathsf{SD}(\mathsf{aExt}(\mathsf{U}_{\mathcal{V}});\mathsf{U}_m) \leq \varepsilon.$$

We will use Li and Zuckerman's [32] recent algebraic extractors in our constructions.

▶ **Lemma 4** ([32])**.** *For any positive integer* d*, there is an efficient* $\left((1-\frac{1}{c_d})n, d, 2^{-\Omega(\frac{n}{c_d})}\right)$*-extractor* $\mathsf{aExt}\colon \{0,1\}^n \to \{0,1\}^m$*, where* $c_d = \Theta(d^2 4^d)$ *and* $m = \Omega(\frac{n}{c_d})$*.*

More generally, *recognizable sources* are flat distributions over sets of the form $\{\mathbf{x}|f(\mathbf{x}) = \mathsf{v}\} \subseteq \{0,1\}^n$ for functions f coming from some specified class $\mathcal{C}$. The distribution $\mathsf{U}_{\{\mathbf{x}|f(\mathbf{x})=\mathsf{v}\}}$ is called the f-recognizable source. A $\mathcal{C}$-recognizable source is the set of f-recognizable sources for each $f \in \mathcal{C}$.

For general (n, k)-sources, there does not exist a deterministic function that can extract close to uniform bits from all of them simultaneously. A family of deterministic functions are needed.

▶ **Definition 5.** *A function* $\mathsf{Ext}\colon \{0,1\}^d \times \{0,1\}^n \to \{0,1\}^m$ *is a strong seeded* (k, ε)*-extractor if for any* (n, k)*-source* X*, we have*

$$\mathsf{SD}(\mathsf{S}, \mathsf{Ext}(\mathsf{S}, \mathsf{X}); \mathsf{S}, \mathsf{U}_m) \le \varepsilon,$$

where S *is chosen uniformly from* $\{0,1\}^d$*. A seeded extractor* $\mathsf{Ext}(\cdot, \cdot)$ *is called linear if for any fixed seed* $\mathsf{S} = \mathsf{s}$*, the function* $\mathsf{Ext}(\mathsf{s}, \cdot)$ *is a linear function.*

There are linear seeded extractors that extract all the randomness, for example, the Trevisan's extractor [45]. In particular, we use the following improvement of this extractor due to Raz, Reingold and Vadhan [40].

▶ **Lemma 6** ([40]). *There is an explicit linear strong* (k, ε)*-extractor* $\mathsf{Ext}\colon \{0,1\}^d \times \{0,1\}^n \to \{0,1\}^m$ *with* $d = O(\log^3(n/\varepsilon))$ *and* $m = k - O(d)$*.*

We will need another explicit strong seeded extractor from universal hash family for our constructions.

▶ **Lemma 7** ([30]). *There is an explicit linear strong* $(3m, 2^{-m})$*-extractor* $\mathsf{Ext}\colon \{0,1\}^d \times \{0,1\}^n \to \{0,1\}^m$ *with* $d = 5m$*.*

In the applications of randomness extractors, the source does not necessarily come in the form of one single random variable, but as a pair of random variables and we want to extract from one of them conditioned on the other one. In this general setting, we usually need the *average min-entropy* to measure the amount of available entropy for extraction.

▶ **Definition 8** ([21]). *The average conditional min-entropy* $\tilde{\mathsf{H}}_\infty(\mathsf{U}|\mathsf{V})$ *of two random variables* $\mathsf{U} \leftarrow \mathcal{U}$ *and* $\mathsf{V} \leftarrow \mathcal{V}$ *is defined as*

$$\tilde{\mathsf{H}}_\infty(\mathsf{U}|\mathsf{V}) = -\log\left(\sum_{\mathsf{v}\in\mathcal{V}} \mathsf{Pr}[\mathsf{V} = \mathsf{v}] \max_{\mathsf{u}\in\mathcal{U}}\{\mathsf{Pr}[\mathsf{U} = \mathsf{u}|\mathsf{V} = \mathsf{v}]\}\right).$$

The average conditional min-entropy satisfies the following property.

▶ **Lemma 9** ([21]). *Let* $\mathsf{V} \leftarrow \mathcal{V}$*. Then the average conditional min-entropy* $\tilde{\mathsf{H}}_\infty(\mathsf{U}|\mathsf{V})$ *is lower bounded as follows.*

$$\tilde{\mathsf{H}}_\infty(\mathsf{U}|\mathsf{V}) \ge \mathsf{H}_\infty(\mathsf{U}) - \log|\mathcal{V}|.$$

We need an *average-case* strong seeded extractor to extract the average conditional min-entropy from such a pair of random variables.

▶ **Definition 10** ([21]). *A function* $\mathsf{Ext}\colon \{0,1\}^d \times \{0,1\}^n \to \{0,1\}^m$ *is an average-case strong seeded* (k, ε)*-extractor if for any random variables* $\mathsf{X} \leftarrow \{0,1\}^n$ *and* $\mathsf{V} \leftarrow \mathcal{V}$ *satisfying* $\tilde{\mathsf{H}}_\infty(\mathsf{X}|\mathsf{V}) \ge k$*, we have*

$$\mathsf{SD}(\mathsf{S}, \mathsf{V}, \mathsf{Ext}(\mathsf{S}, \mathsf{X}); \mathsf{S}, \mathsf{V}, \mathsf{U}_m) \le \varepsilon,$$

where S *is chosen uniformly from* $\{0,1\}^d$*.*

Explicit constructions of randomness extractors have efficient forward direction of extraction. In some applications, we usually need to efficiently invert the process: Given an extractor output, sample a random pre-image. This is not necessarily efficient if the extractor is not a linear function, in which case we need to explicitly construct an *invertible extractor*. If the extractor is linear, sampling a random pre-image can be done in polynomial time.

▶ **Definition 11** ([18]). *Let f be a mapping from $\{0,1\}^n$ to $\{0,1\}^m$. For $\mu \geq 0$, a function* $\mathsf{Inv}\colon \{0,1\}^m \times \{0,1\}^r \to \{0,1\}^n$ *is called a μ-inverter for f if the following conditions hold:*

- *(Inversion) Given* $\mathsf{y} \in \{0,1\}^m$ *such that its pre-image* $f^{-1}(\mathsf{y})$ *is nonempty, for every* $\mathsf{r} \in \{0,1\}^r$ *we have* $f(\mathsf{Inv}(\mathsf{y},\mathsf{r})) = \mathsf{y}$.
- *(Uniformity)* $\mathsf{Inv}(\mathsf{U}_m, \mathsf{U}_r)$ *is μ-close to* U_n.

A μ-inverter is called efficient if there is a randomized algorithm that runs in worst-case polynomial time and, given $\mathsf{y} \in \{0,1\}^m$ *and* r *as a random seed, computes* $\mathsf{Inv}(\mathsf{y},\mathsf{r})$. *We call a mapping μ-invertible if it has an efficient μ-inverter, and drop the prefix μ from the notation when it is zero. We abuse the notation and denote the inverter of f by f^{-1}.*

Finally, we need the following simple lemma whose proof can be found in the full version.

▶ **Lemma 12.** *Let* V, V' *be two random variables distributed over the set $\mathcal{V}$ and* W, W' *over $\mathcal{W}$ satisfying* $\mathsf{SD}(\mathsf{V},\mathsf{W};\mathsf{V}',\mathsf{W}') \leq \varepsilon$. *Let $\mathcal{E} \subset \mathcal{W}$ be an event. Then we have the following.*

$$\mathsf{SD}(\mathsf{V}|\mathsf{W} \in \mathcal{E}; \mathsf{V}'|\mathsf{W}' \in \mathcal{E}) \leq \frac{2\varepsilon}{\mathsf{Pr}[\mathsf{W}' \in \mathcal{E}]}.$$

3 Leakage Resilient Secret Sharing

In this section, we define a general leakage model for secret sharing, which can be viewed as an extension of the local leakage model proposed in [10, 27] to include non-compartmentalised leakage models.

▶ **Definition 13.** *Let* t *and* n *satisfying* $2 \leq \mathsf{t} \leq \mathsf{n}$ *be two integers. Let $\mathbb{F}_q$ be the finite field of q elements for q a power of 2. Let $\mathcal{L}$ be a set of Boolean functions of* $\mathsf{n}\log q$ *bits input. Let β be an integer denoting the bound on the total number of bits leaked. Let* $\theta < \mathsf{t}$ *be an integer and ε be a small positive real. A $(\mathcal{L}, \beta, \theta, \varepsilon)$-leakage resilient* t*-out-of-*n *secret sharing scheme over the finite field $\mathbb{F}_q$ is defined by a pair of polynomial-time algorithms* $(\mathsf{Share}, \mathsf{Rec})$, *where* Share *is a randomized mapping of an input* $\mathsf{s} \in \{0,1\}^m$, *for $m \leq \log q$, to a share vector* $\mathbf{Sh} = (\mathsf{Sh}_1, \ldots, \mathsf{Sh}_\mathsf{n})$ *and the reconstruction algorithm* Rec *is a deterministic function mapping a set* $A \subseteq [\mathsf{n}]$ *and the corresponding shares* $\mathbf{Sh}_A = (\mathsf{Sh}_i)_{i \in A}$ *to a secret in* $\{0,1\}^m$, *such that the following properties hold:*

- *Correctness:* $\mathsf{Rec}(A, \mathbf{Sh}_A)$ *outputs the secret* s *for all sets* $A \subseteq [\mathsf{n}]$ *where* $|A| \geq \mathsf{t}$.
- *Privacy and leakage resiliency:*
    - *Non-adaptive adversary: for any pair* $\mathsf{s}^0, \mathsf{s}^1 \in \{0,1\}^m$ *of secrets with share vectors* $\mathbf{Sh}^0$ *and* $\mathbf{Sh}^1$, *any* $A \subseteq [\mathsf{n}]$ *of size* $|A| \leq \theta$, *any β functions* $f_i \in \mathcal{L}$, $i = 1, \ldots, \beta$,

      $$\mathsf{SD}\left(\mathsf{Leak}_{A,\beta}(\mathbf{Sh}^0), \mathbf{Sh}^0_A; \mathsf{Leak}_{A,\beta}(\mathbf{Sh}^1), \mathbf{Sh}^1_A\right) \leq \varepsilon, \qquad (1)$$

      where $\mathsf{Leak}_{A,\beta}(\mathbf{Sh}^b)$ *denotes the output of the Boolean functions* $f_1, \ldots, f_\beta$ *on input* $\mathbf{Sh}^b$ *for* $b \in \{0,1\}$. *In the case when $\mathcal{L}$ is defined such that the input of these Boolean functions* $f_1, \ldots, f_\beta$ *are restricted to one share and otherwise not restricted, we recover the local leakage model.*

- *Adaptive adversary: for any pair* $\mathsf{s}^0, \mathsf{s}^1 \in \{0,1\}^m$ *of secrets with share vectors* $\mathbf{Sh}^0$ *and* $\mathbf{Sh}^1$*,*

$$\mathsf{SD}\left(\mathsf{LEAK}(\mathbf{Sh}^0, \mathcal{L}, \beta, \theta); \mathsf{LEAK}(\mathbf{Sh}^1, \mathcal{L}, \beta, \theta)\right) \leq \varepsilon, \tag{2}$$

where $\mathsf{LEAK}(\mathbf{Sh}^b, \mathcal{L}, \beta, \theta)$ *denotes the transcript of the following interactive protocol between the adversary* $\mathcal{A}$ *and an oracle* $\mathcal{O}$ *holding* $\mathbf{Sh}^b$*. For* $i = 1, \ldots, \beta$ *and* $j = 1, \ldots, \theta$*,* $\mathcal{A}$ *chooses* $f_i \in \mathcal{L}$ *and* $\mathsf{n}_j \in [\mathsf{n}]$ *based on all previous communication, and* $\mathcal{O}$ *answers with* $f_i(\mathbf{Sh}^b)$ *and* $\mathbf{Sh}^b_{\mathsf{n}_j}$*, respectively.*

When $\beta = 0$ *and* $\theta = \mathsf{t} - 1$ *is achieved, we recover the statistical privacy of a threshold scheme. When* $\beta = 0$ *and only* $\theta < \mathsf{t} - 1$ *is achieved, the scheme is a ramp scheme with statistical privacy for privacy threshold* θ*.*

Generic Construction

$$\begin{cases} \mathsf{Share}(\cdot) & = \mathsf{ECCenc}(\mathsf{EXT}^{-1}(\cdot)); \\ \mathsf{Rec}(\cdot) & = \mathsf{EXT}(\mathsf{ECCdec}(\cdot)). \end{cases}$$

The ECC is an erasure correcting code with encoder/decoder pair $(\mathsf{ECCenc}, \mathsf{ECCdec})$ and EXT is an invertible randomness extractor with inverter EXT^{-1}.

In most of the instantiations, EXT is a seedless extractor. All the efficient extractors mentioned in the preliminary section (if not already invertible) can be transformed into one that is invertible, at the cost of increasing the input length, using the following method. Let $\mathsf{EXT}_0 \colon \{0,1\}^{n_0} \to \{0,1\}^m$ be an (n_0, k)-extractor with error ε. Let $n = n_0 + m$. Then $\mathsf{EXT} : \{0,1\}^{n_0+m} \to \{0,1\}^m$ defined as follows is a ε-invertible $(n, k+m)$-extractor with error ε.

$$\mathsf{EXT}(\mathsf{x}||\mathsf{y}) = \mathsf{EXT}_0(\mathsf{x}) + \mathsf{y},$$

where "||" denotes concatenation. The inverter of EXT is

$$\mathsf{EXT}^{-1}(\mathsf{s}) = (\mathsf{U}_{n_0}||\mathsf{EXT}_0(\mathsf{U}_{n_0}) + \mathsf{s}),$$

where the two copies of U_{n_0} denote the same random variable. In the case when EXT is a seeded extractor, its inverter EXT^{-1} uniformly chooses a seed and invert according to the function labeled by the seed. More concretely, let $\mathsf{Ext} \colon \{0,1\}^d \times \{0,1\}^n \to \{0,1\}^m$ be an invertible seeded extractor. We use the short hand $\mathsf{Ext}_\mathsf{z}(\cdot) = \mathsf{Ext}(\mathsf{z}, \cdot)$ and $\mathsf{EXT}(\cdot) = \mathsf{Ext}_{\mathsf{U}_d}(\cdot)$. Then the inverter $\mathsf{EXT}^{-1} \colon \{0,1\}^m \to \{0,1\}^n$ is defined as follows.

$$\mathsf{EXT}^{-1}(\mathsf{s}) = \mathsf{EXT}^{-1}_{\mathsf{U}_d}(\mathsf{s}),\ \mathsf{U}_d \overset{\$}{\leftarrow} \{0,1\}^d.$$

Though not explicitly reflected in the notations above, the uniform seed chosen in the process of inverting is also recorded so that the reconstruction algorithm can correctly recover the secret. For example, in the construction for global affine leakage (Theorem 14), the seed is directly appended to the n-bit pre-image.

In all the instantiations, ECC is a linear Maximum Distance Separable (MDS) code over a large enough finite field $\mathbb{F}_q$ of characteristic 2. For example, we require the share size q satisfy $q > \mathsf{n}$ if we use a $[\mathsf{n}, \mathsf{t}, \mathsf{n} - \mathsf{t} + 1]_q$ Reed-Solomon code and the input length $n = \mathsf{t} \log q$ of the invertible extractor EXT should be big enough to achieve the error ε.

4 Affine Leakage Models

In this section, we study LR-SS with respect to the class $\mathcal{L}_{\mathsf{affine}}$ of affine functions over $\mathbb{F}_2$. We have two instantiations of the generic construction for non-adaptive and adaptive adversary, respectively.

4.1 Affine Leakage Non-Adaptive Adversary

We begin with an instantiation that gives optimal parameters in terms of both leakage tolerance and information ratio, though we can only prove security against a non-adaptive adversary.

▶ **Theorem 14.** *Let $\mathcal{L}_{\mathsf{affine}}$ be the class of affine functions over $\mathbb{F}_2$. Let $0 < \xi \leq 1$ be any real number. There is a family of t-out-of-n secret sharing schemes over $\mathbb{F}_q$ (labeled by $\log q$) that is simultaneously $(\mathcal{L}_{\mathsf{affine}}, \beta, \theta, \varepsilon)$-leakage resilient against a non-adaptive adversary for $0 \leq \beta \leq (\mathsf{t} - \xi)\log q$ and $\theta \in \{0, 1, \ldots, \mathsf{t} - 1\}$ such that $\theta + \frac{\beta}{\log q} \leq \mathsf{t} - \xi$, where the share size q (determined by ε and satisfies $q > \mathsf{n}$) is a large enough power of 2 and the secret length $m = \xi \log q - o(\log q)$.*

Proof. We instantiate the generic construction using a linear strong seeded extractor $\mathsf{Ext}\colon \{0,1\}^d \times \{0,1\}^n \to \{0,1\}^m$ and a linear MDS code.

- $\mathsf{EXT} : \{0,1\}^n \to \{0,1\}^m$ is instantiated as follows.

  $$\mathsf{EXT}(\mathsf{x}) = \mathsf{Ext}(\mathsf{U}_d, \mathsf{x}),$$

  where Ext is the linear strong seeded extractor from Lemma 6 with parameters (k_E, ε_E). The uniform seed sampled during inverting is directly appended to the n-bit pre-image $\mathsf{EXT}^{-1}(\mathsf{s})$.

  According to Lemma 6, there is an explicit linear strong (k_E, ε_E)-extractor $\mathsf{Ext}\colon \{0,1\}^d \times \{0,1\}^n \to \{0,1\}^m$ with $d = O(\log^3(n/\varepsilon_E))$ and $m = k_E - O(d)$.
- ECC is a linear MDS code with parameter $[\mathsf{n}, \mathsf{t}, \mathsf{n} - \mathsf{t} + 1]$ over $\mathbb{F}_q$, where $\log q = \frac{d+n}{\mathsf{t}}$. The output of $\mathsf{ECCenc}\colon \{0,1\}^{d+n} \to \mathbb{F}_q^{\mathsf{n}}$ is the share vector.

Reconstruction from any t shares follows from the functionality of ECC and the invertibility guarantee of the EXT, which insures that any correctly recovered pre-image is mapped back to the original secret.

We next prove privacy and leakage resiliency, which will follow naturally from Lemma 15. We first recall this general property of a linear strong extractor, which is proved in [36].

▶ **Lemma 15** ([36]). *Let $\mathsf{Ext}\colon \{0,1\}^d \times \{0,1\}^n \to \{0,1\}^m$ be a linear strong (k, ε)-extractor, $f\colon \{0,1\}^{d+n} \to \{0,1\}^a$ be an affine function with output length $a \leq n - k$. For any $\mathsf{m}, \mathsf{m}' \in \{0,1\}^m$, let $(\mathsf{Z}, \mathsf{X}) = (\mathsf{U}_d, \mathsf{U}_n)|(\mathsf{Ext}(\mathsf{U}_d, \mathsf{U}_n) = \mathsf{m})$ and $(\mathsf{Z}', \mathsf{X}') = (\mathsf{U}_d, \mathsf{U}_n)|(\mathsf{Ext}(\mathsf{U}_d, \mathsf{U}_n) = \mathsf{m}')$. We have*

$$\mathsf{SD}(f(\mathsf{Z}, \mathsf{X}); f(\mathsf{Z}', \mathsf{X}')) \leq 8\varepsilon. \tag{3}$$

The inverter EXT^{-1} takes a secret, which is a particular extractor output $\mathsf{s} \in \{0,1\}^m$, and uniformly samples a seed $\mathsf{z} \in \{0,1\}^d$ of Ext before uniformly finds an $\mathsf{x} \in \{0,1\}^n$ such that $\mathsf{Ext}(\mathsf{z}, \mathsf{x}) = \mathsf{s}$. This process of obtaining (z, x) is the same as sampling uniformly and independently $(\mathsf{U}_d, \mathsf{U}_n) \overset{\$}{\leftarrow} \{0,1\}^{d+n}$ and then restricting to $\mathsf{Ext}(\mathsf{U}_d, \mathsf{U}_n) = \mathsf{s}$. We define the random variable pair

$$(\mathsf{Z}, \mathsf{X}) := (\mathsf{U}_d, \mathsf{U}_n)|(\mathsf{Ext}(\mathsf{U}_d, \mathsf{U}_n) = \mathsf{s}) \tag{4}$$

and refer to it as the pre-image of s.

Let $\Pi_A : \mathbb{F}_q^{\mathsf{n}} \to \mathbb{F}_q^{\theta}$ be the projection function that maps a share vector to the θ shares with index set $A \subseteq [\mathsf{n}]$ and $|A| = \theta$ chosen by the non-adaptive adversary. Observe that the combination $(\Pi_A \circ \mathsf{ECCenc}) : \{0,1\}^{d+n} \to \{0,1\}^{(\log q)\theta}$ is an affine function. Moreover, for any affine leakage function $l : \{0,1\}^{(\log q)\mathsf{n}} \to \{0,1\}^{\beta}$, the composition $(l \circ \mathsf{ECCenc}) : \{0,1\}^{d+n} \to \{0,1\}^{\beta}$ is also an affine function. So the view of the adversary is simply the output of the affine function $f = (\Pi_A \circ \mathsf{ECCenc} || l \circ \mathsf{ECCenc})$, where "$||$" denotes concatenation, applied to the random variable tuple (Z, X) defined in (4).

We want to prove that the statistical distance of the views of the adversary for a pair of secrets s and s' can be made arbitrarily small. The views of the adversary are the outputs of the affine function f with inputs $(\mathsf{Z}^0, \mathsf{X}^0)$ and $(\mathsf{Z}^1, \mathsf{X}^1)$ for the secret s^0 and s^1, respectively. Let $k_E = n - (\log q)\theta - \beta$. According to Lemma 15, we then have that the privacy and leakage resiliency error is $8\varepsilon_E$.

Finally, since $\theta + \frac{\beta}{\log q} < \mathsf{t} - \xi$, we have $\mathsf{t} - \theta - \frac{\beta}{\log q} > \xi$. Let $d+n$ be a multiple of t. We then have a family of schemes labeled by $\log q$. The privacy and leakage resiliency error $8\varepsilon_E$ is negligible in n, and hence is obviously negligible in $\log q$. The secret length can be chosen as follows.

$$m = n - (\log q)\theta - \beta - O(d) = (\log q)(\mathsf{t} - \theta - \frac{\beta}{\log q}) - o(\log q) = \xi \log q - o(\log q),$$

where the seed length is $d = O(\log^3(2n/\varepsilon))$ and $d + n = (\log q)\mathsf{t}$. ◀

▶ Remark 16. Let $\xi = 1$ and consider $\theta = \mathsf{t} - 1$ and $\beta = 0$. In this case, we recover a t-out-of-n secret sharing scheme with information ratio $\frac{m}{\log q} \approx 1$, for large enough q. This is almost as good as the Shamir's secret sharing scheme, except that Shamir's scheme has perfect privacy and information ratio exactly 1 while we only have statistical privacy with a negligible privacy error and information ratio close to 1. On the other hand, as demonstrated in [29], Shamir's scheme over finite field of characteristic 2 is completely vulnerable in the face of a non-standard leakage adversary, in particular, even leaking one bit from each share allows reconstruction of the complete secret. The t-out-of-n secret sharing scheme in Theorem 14, when we set $m = \log q - o(\log q)$, is leakage-resilient against a non-adaptive adversary who obtains any $\theta \le \mathsf{t} - 1$ shares and leak up to $\beta = (\log q)\phi$ bits for any $0 \le \phi \le \mathsf{t} - 1 - \theta$ through applying affine leakage functions.

4.2 Affine Leakage Adaptive Adversary

The instantiation for adaptive adversary does not have optimal information ratio. We manage to maintain optimal leakage tolerance.

▶ **Theorem 17.** *Let $\mathcal{L}_{\mathsf{affine}}$ be the class of affine functions over $\mathbb{F}_2$. Let $0 < \xi \le 1$ be any real number. There is a family of t-out-of-n secret sharing schemes over $\mathbb{F}_q$ (labeled by $\log q$) that is simultaneously $(\mathcal{L}_{\mathsf{affine}}, \beta, \theta, \varepsilon)$-leakage resilient against an adaptive adversary for $0 \le \beta < (\mathsf{t} - \xi)\log q$ and $\theta \in \{0, 1, \ldots, \mathsf{t} - 1\}$ such that $\theta + \frac{\beta}{\log q} \le \mathsf{t} - \xi$, where the share size q (determined by ε and satisfies $q > \mathsf{n}$) is a large enough power of 2 and the secret length $m = \Omega(\log q)$.*

We use an invertible seedless extractors that can extract from affine recognizable sources with any constant fraction of entropy with output length also a constant fraction of the input and with exponentially small error. There are known constructions [13, 33] of affine extractors that can extract from any constant fraction of entropy and output a constant fraction of random bits with exponentially small error. We can directly instantiate our generic construction with these extractors by transforming them into invertible extractors as described in the generic construction.

Here we exploit a classical approach to building affine extractors through composing with a seeded extractor and use the trick in [18] to make it invertible at a lower cost than the above method. But the obtained affine extractor does not have exponentially small error and we can not directly use it in the generic construction intuition due to the exponential grow of error when transforming from extractor-based security to secret sharing security. We then use a more delicate analysis of the error terms that circumvents the problem.

We first recall the classical framework of constructing seedless extractors from composing with seeded extractors. Seeded extractors are known to explicitly extract all the entropy and are not restricted by source structures. Moreover, there are known constructions of *linear* seeded extractors perform almost as well as the best seeded extractors. The elegant idea of this framework is to use a seedless extractor to extract a short output from the structured source, which then serves as the seed for a seeded extractor to extract all the entropy from the same source. For this idea to work, the dependence of the extracted seed on the source has to be carefully analyzed (and removed).

▶ **Lemma 18** ([41]). *Let $\mathcal{D}$ be a class of distributions over $\{0,1\}^n$. Let $\mathsf{E}:\{0,1\}^n \to \{0,1\}^d$ be a seedless extractor for $\mathcal{D}$ with error ϵ. Let $\mathsf{F}:\{0,1\}^d \times \{0,1\}^n \to \{0,1\}^m$. Let X be a distribution in $\mathcal{D}$ and assume that for every $\mathsf{z} \in \{0,1\}^d$ and $\mathsf{y} \in \{0,1\}^m$, the distribution* $(\mathsf{X}|\mathsf{F}(\mathsf{z},\mathsf{X})=\mathsf{y})$ *belongs to $\mathcal{D}$. Then*

$$\mathsf{SD}(\mathsf{E}(\mathsf{X}),\mathsf{F}(\mathsf{E}(\mathsf{X}),\mathsf{X});\mathsf{U}_d,\mathsf{F}(\mathsf{U}_d,\mathsf{X})) \le 2^{d+3}\epsilon.$$

An example of such a class of distributions is the affine source, in which case we can use an *affine* extractor $\mathsf{F}=\mathsf{aExt}$ and a *linear* seeded extractor $\mathsf{E}=\mathsf{Ext}$. An affine source X conditioned on $\mathsf{Ext}(\mathsf{z},\mathsf{X})=\mathsf{y}$, which amounts to a set of linear equations, is still an affine source for aExt. With appropriate choice of parameters, we obtain a better (than aExt) affine extractor $\mathsf{aExt}'(\mathsf{X}):=\mathsf{Ext}(\mathsf{aExt}(\mathsf{X}),\mathsf{X})$. With an increase of d bits (instead of m bits as described below generic construction) in the input, we have the following invertible affine extractor.

$$\mathsf{EXT}(\mathsf{Sd}||\mathsf{X}):=\mathsf{Ext}(\mathsf{aExt}(\mathsf{X})+\mathsf{Sd},\mathsf{X}),$$

whose inverter is $\mathsf{EXT}^{-1}(\mathsf{s}):=\left(\mathsf{aExt}(\mathsf{Ext}_{\mathsf{Z}}^{-1}(\mathsf{s}))+\mathsf{Z}||\mathsf{Ext}_{\mathsf{Z}}^{-1}(\mathsf{s})\right)$, where $\mathsf{Z}\xleftarrow{\$}\{0,1\}^d$.

Proof of Theorem 17. We instantiate the generic construction using an affine extractor $\mathsf{aExt}\colon\{0,1\}^n\to\{0,1\}^d$, a linear strong seeded extractor $\mathsf{Ext}\colon\{0,1\}^d\times\{0,1\}^n\to\{0,1\}^m$ and a linear MDS code.

- $\mathsf{EXT}:\{0,1\}^{d+n}\to\{0,1\}^m$ is instantiated as follows.

  $$\mathsf{EXT}(\mathsf{z}||\mathsf{x})=\mathsf{Ext}(\mathsf{aExt}(\mathsf{x})+\mathsf{z},\mathsf{x}),$$

  where Ext is the linear strong seeded extractor from Lemma 6 with parameters (k_E,ε_E) and $\varepsilon_E<\frac{1}{8}$; aExt is the seedless extractor from Lemma 2 with parameters (k_A,ε_A). The inverter

  $$\mathsf{EXT}^{-1}(\mathsf{s}):=\left(\mathsf{aExt}(\mathsf{Ext}_{\mathsf{Z}}^{-1}(\mathsf{s}))+\mathsf{Z}||\mathsf{Ext}_{\mathsf{Z}}^{-1}(\mathsf{s})\right),$$

  where $\mathsf{Z}\xleftarrow{\$}\{0,1\}^d$
  According to Lemma 6, there is an explicit linear strong (k_E,ε_E)-extractor $\mathsf{Ext}\colon\{0,1\}^d\times\{0,1\}^n\to\{0,1\}^m$ with $d=O(\log^3(n/\varepsilon_E))$ and $m=k_E-O(d)$.
  According to Lemma 2, for every constant $0<\mu\le 1$, there is an explicit affine extractor $\mathsf{aExt}\colon\{0,1\}^n\to\{0,1\}^d$ for affine (n,k_A)-sources, where $k_A=n\mu$, with output length $d=\Omega(n)$ and error ε_A at most $2^{-\Omega(n)}$.

- ECC is a linear MDS code with parameter $[\mathsf{n}, \mathsf{t}, \mathsf{n} - \mathsf{t} + 1]$ over $\mathbb{F}_q$, where $\log q = \frac{d+n}{\mathsf{t}}$. The output of $\mathsf{ECCenc} \colon \{0,1\}^{d+n} \to \mathbb{F}_q^{\mathsf{n}}$ is the share vector.

Reconstruction from any t shares follows from the functionality of ECC and the invertibility guarantee of the EXT, which insures that any correctly recovered pre-image is mapped back to the original secret.

We next prove privacy and leakage resiliency. Consider a uniform secret U_m. By the uniformity guarantee of the inverter, we have $\mathsf{Share}(\mathsf{U}_m) = \mathsf{ECCenc}(\mathsf{Sd}||\mathsf{U}_n)$. Our analysis is done for any fixed $\mathsf{Sd} = \mathsf{sd}$. This captures a stronger adversary who on top of adaptively reading t shares, also has access to Sd through an oracle. It is easy to see that the fixing of $\mathsf{Sd} = \mathsf{sd}$ does not alter the distribution of the source U_n, which remains uniform over $\{0,1\}^n$. Let $\mathsf{V} \colon = \mathsf{LEAK}(\mathsf{ECCenc}(\mathsf{sd}||\mathsf{U}_n), \mathcal{L}_{\mathsf{affine}}, \beta, \theta)$ denote the view of the adversary on the encoding of a uniform source for the fixed $\mathsf{Sd} = \mathsf{sd}$. Let $\mathsf{Z} \colon = \mathsf{aExt}(\mathsf{U}_n) + \mathsf{sd}$ denote the seed of the strong linear extractor Ext. Finally, let $\mathsf{S} \colon = \mathsf{Ext}(\mathsf{Z}, \mathsf{U}_n)$. We study the random variable tuple $(\mathsf{V}, \mathsf{Z}, \mathsf{S})$ to complete the proof.

The pair $(\mathsf{Z}, \mathsf{S})|\mathsf{V} = \mathsf{v}$ for any fixed $\mathsf{V} = \mathsf{v}$ is by definition $(\mathsf{aExt}(\mathsf{U}_n) + \mathsf{sd}, \mathsf{Ext}(\mathsf{aExt}(\mathsf{U}_n) + \mathsf{sd}, \mathsf{U}_n))|\mathsf{V} = \mathsf{v}$. The distribution $(\mathsf{U}_n|\mathsf{V} = \mathsf{v})$ is an affine source with at least $n - (\log q)\theta - \beta$ entropy. Let $k_A = n - \theta \log q - \beta - m$. According to Lemma 18, we have

$$(\mathsf{Z}, \mathsf{S})|\mathsf{V} = \mathsf{v} \overset{2^{d+3}\varepsilon_A}{\sim} (\mathsf{U}_d, \mathsf{Ext}(\mathsf{U}_d, \mathsf{U}_n))|\mathsf{V} = \mathsf{v}.$$

Our concern is the relation between S and V, and therefore would like to further condition on values of Z. Let $k_E = n - (\log q)\theta - \beta - d$ and consider the linear strong extractor Ext. In this step, we crucially use the linearity of Ext and the underlying linear space structure of the affine source $\mathsf{U}_n|\mathsf{V} = \mathsf{v}$ to claim that there is a subset $\mathcal{G} \subset \{0,1\}^d$ of good seeds such that $\mathsf{Pr}[\mathsf{U}_d \in \mathcal{G}] \geq 1 - 4\varepsilon_E$ and for any $\mathsf{z} \in \mathcal{G}$, the distribution of $\mathsf{Ext}(\mathsf{z}, \mathsf{U}_n)|\mathsf{V} = \mathsf{v}$ is exactly uniform. This is true because $\mathsf{Ext}(\mathsf{z}, \mathsf{U}_n)|\mathsf{V} = \mathsf{v}$ is an affine source. If its entropy is m, then it is exactly uniform. If its entropy is less than m, its statistical distance $\varepsilon_E^{\mathsf{z}}$ from uniform is at least $\frac{1}{2}$. Using an averaging argument we have that at least $1 - 4\varepsilon_E$ fraction of the seeds should satisfy $\varepsilon_E^{\mathsf{z}} < \frac{1}{4}$, and hence $\varepsilon_E^{\mathsf{z}} = 0$. We then use Lemma 12 with respect to the event $\mathsf{Z} \in \mathcal{G}$ to claim that

$$(\mathsf{S}|(\mathsf{V} = \mathsf{v}, \mathsf{Z} \in \mathcal{G})) \overset{\frac{2^{d+4}\varepsilon_A}{1-4\varepsilon_E}}{\sim} (\mathsf{Ext}(\mathsf{U}_d, \mathsf{X})|(\mathsf{V} = \mathsf{v}, \mathsf{U}_d \in \mathcal{G})),$$

where the right hand side is exactly U_m. Note that the subset $\mathcal{G}$ is determined by the choice of the θ shares and by the leakage adversary, hence remains the same for any value of $\mathsf{V} = \mathsf{v}$. We then have

$$((\mathsf{V}, \mathsf{S})|\mathsf{Z} \in \mathcal{G}) \overset{\frac{2^{d+4}\varepsilon_A}{1-4\varepsilon_E}}{\sim} (\mathsf{V}, \mathsf{U}_m).$$

Another application of Lemma 12 with respect to the event $\mathsf{S} = \mathsf{s}$ gives

$$(\mathsf{V}|(\mathsf{Z} \in \mathcal{G}, \mathsf{S} = \mathsf{s})) \overset{\frac{2^{(m+1)+(d+4)}\varepsilon_A}{1-4\varepsilon_E}}{\sim} \mathsf{V}.$$

We finally bound the privacy and leakage resiliency error as follows.

$$\begin{aligned}
&\mathsf{SD}((\mathsf{V}|\mathsf{S} = \mathsf{s}_0); (\mathsf{V}|\mathsf{S} = \mathsf{s}_1)) \\
&\leq 2\mathsf{SD}((\mathsf{V}|\mathsf{S} = \mathsf{s}); \mathsf{V}) \\
&= 2\mathsf{Pr}[\mathsf{Z} \in \mathcal{G}] \cdot \mathsf{SD}((\mathsf{V}|(\mathsf{Z} \in \mathcal{G}, \mathsf{S} = \mathsf{s})); \mathsf{V}) + 2\mathsf{Pr}[\mathsf{Z} \notin \mathcal{G}] \cdot \mathsf{SD}((\mathsf{V}|(\mathsf{Z} \notin \mathcal{G}, \mathsf{S} = \mathsf{s})); \mathsf{V}) \\
&\leq 2\left(1 \cdot \frac{2^{(m+1)+(d+4)}\varepsilon_A}{1-4\varepsilon_E} + (4\varepsilon_E + \varepsilon_A) \cdot 1\right) \\
&< 2^{(m+1)+(d+4)+2}\varepsilon_A + 8\varepsilon_E.
\end{aligned}$$

Note that in the error bound above, the exponential term $2^{(m+1)+(d+4)+2}$ only appears as the multiplier of ε_A, the error of aExt. In order to cancel out the exponential multiplier $2^{(m+1)+(d+4)+2}$, we require aExt to have an exponentially small error $\varepsilon_A = \frac{\varepsilon/2}{2^{(m+1)+(d+4)+2}}$, which can be trivially done by setting $k_A = \Omega(n)$. If we want to have secret length $m = \Omega(\log q) = \Omega(\frac{d+n}{\mathsf{t}})$, we also need to set $k_E = m + O(d) = \Omega(n)$. Both are achieved by requiring

$$n - (\log q)\theta - \beta - d = \Omega(n),$$

which in turn, given that $d = O(\log^3(n/\varepsilon_E))$, can be achieved by requiring

$$n(1 - \frac{\theta}{\mathsf{t}} - \frac{\beta}{n}) = \Omega(n) \text{ or } 1 - \frac{\theta}{\mathsf{t}} - \frac{\beta}{t \log q} > 0.$$

This shows that for any $\theta + \frac{\beta}{\log q} < \mathsf{t}$, privacy and leakage resiliency error ε can be achieved through a large enough q and the secret length is $m = \Omega(\log q)$. This concludes the proof. ◀

▶ Remark 19. Directly using the affine extractor aExt and transform it using the method in the generic construction will result in an invertible affine extractor EXT: $\{0,1\}^{m+n} \to \{0,1\}^m$ with exponentially small error. Instantiating the generic construction using this extractor also gives a secret sharing scheme that is simultaneously $(\mathcal{L}_{\mathsf{affine}}, \beta, \theta, \varepsilon)$-leakage resilient for any $\theta + \frac{\beta}{\log q} < \mathsf{t}$. But the information ratio of this instantiation is $\frac{m}{(m+n)/\mathsf{t}}$, while in Theorem 17, the information ratio is $\frac{m}{(d+n)/\mathsf{t}}$. Recall that $d = O(\log^3(n/\varepsilon_E))$ and $m = \Omega(n)$. We then have $\frac{m}{(d+n)/\mathsf{t}} \approx \frac{m}{n/\mathsf{t}} > \frac{m}{(m+n)/\mathsf{t}}$.

5 Low-degree Multi-variate Polynomial Leakage

We next consider the class $\mathcal{L}_{d\text{-}\mathsf{poly}}$ of global leakage functions that are multi-variate polynomials in binary variables $x_1, \ldots, x_{\mathsf{n}\log q}$ with degree at most d, as natural extension of the affine (degree 1) leakage functions $\mathcal{L}_{\mathsf{affine}}$. From now on, we assume the degree of the algebraic leakage functions are bigger than 1 (for $d = 1$, our constructions in previous section give better parameters).

We have seen in Theorem 14 that as long as the adversary is non-adaptive and restricted to affine leakage, we can instantiate the generic construction with *any* seeded extractor to obtain a LR-SS against global affine leakage. Unfortunately, we can already give an example of seeded extractor that is not sufficient for providing privacy and leakage resiliency against a non-adaptive adversary who globally leaks through degree 2 multi-variate polynomials. It is well known that the inner product function $\{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ gives a good seeded extractor. This function can be described by a degree 2 polynomial in $2n$ variables. Assuming our generic construction is instantiated using the inner product seeded extractor and a non-adaptive adversary chooses exactly the corresponding degree 2 multi-variate polynomial to leak, the single bit leakage is the secret itself and the scheme is not leakage-resilient. Note that this example does not rule out the possibility of obtaining non-adaptive LR-SS against algebraic leakage of degree $d > 1$ using a specially chosen seeded extractor, for example, not computable by degree d polynomials. In the standard applications of seeded extractors, we only require the extractor function to be efficient. Here we need extractor functions to at least have degree more than d. For simplicity, in this work, we only use seedless extractors (for algebraically recognizable sources), which by definition already takes the structure of the leakage functions into account. Instantiating with a seedless extractor has an advantage of providing security against both a non-adaptive adversary and an adaptive adversary.

▶ **Theorem 20.** *Let $\mathcal{L}_{d\text{-poly}}$ be the class of multi-variate polynomials over $\mathbb{F}_2$ with degree at most d. Let g be an integer satisfying $\frac{g}{\mathsf{t}} > 1 - \frac{1}{c_d}$, where $c_d = \Theta(d^2 4^d)$ is a constant determined by d. There is a family of* t*-out-of-*n *secret sharing scheme with threshold gap g over the finite field $\mathbb{F}_q$ (labeled by $\log q$) that is simultaneously $(\mathcal{L}_{d\text{-poly}}, \beta, \theta, \varepsilon)$-leakage resilient for all $\beta = (\log q)\phi$ and $\theta \leq \mathsf{t} - g$ satisfying $0 \leq \phi \leq \mathsf{t} - g - \theta$. The share size q (determined by ε and satisfying $q > \mathsf{n}$) is a large enough power of 2 and the secret length $m = \Omega(\log q)$.*

Before we prove Theorem 20 through instantiating our generic construction, we need to find a suitable measure on the amount of remaining entropy conditioned on the leaked information. The average min-entropy is usually used in combination with the so-called *average-case* extractors, which are strong seeded extractors that have the extractor guarantee averaging over an extra random variable that is related to the source. As far as we know, average-case extractors have not been defined in the seedless case. We then include a brief discussion here that might be of independent interest.

▶ **Definition 21.** *A (k, ε)-seedless extractor $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ for $\mathcal{C}$-recognizable sources is called average-case, if for all pair of random variables (X, V), where X is a $\mathcal{C}$-recognizable source and V is a random variable arbitrarily related to X satisfying that $\mathsf{X}|\mathsf{V} = \mathsf{v}$ is a $\mathcal{C}$-recognizable source for any $\mathsf{V} = \mathsf{v}$ and $\tilde{\mathsf{H}}_\infty(\mathsf{X}|\mathsf{V}) \geq k$, we have*

$$\mathsf{SD}(\mathsf{V}, \mathsf{Ext}(\mathsf{X}); \mathsf{V}, \mathsf{U}_m) \leq \varepsilon.$$

While any strong seeded extractor can be trivially converted into an average-case extractor at the cost of an increase in the extractor error and a proportional strengthening on the min-entropy requirement, there are extractor constructions that directly provide an average-case extractor [21]. This is also the case for seedless extractors.

▶ **Lemma 22** (modified from [21])**.** *For any $\delta > 0$, if* Ext *is a $(k - \log 1/\delta, \varepsilon)$-seedless extractor for $\mathcal{C}$-recognizable sources, then* Ext *is also an average-case $(k, \varepsilon + \delta)$-seedless extractor for $\mathcal{C}$-recognizable sources.*

Sometimes the strengthening on entropy requirement can be quite costly ($\log 1/\delta$ can be a large value if the extractor error $\varepsilon + \delta$ has to be exponentially small). Luckily, by examining the construction of the extractor in Lemma 4, we assert that it is already an average-case seedless extractor.

▶ **Lemma 23.** *For any integer $d > 0$, there is an efficient average-case $\left((1 - \frac{1}{c_d})n, d, 2^{-\Omega(\frac{n}{c_d})}\right)$-extractor $\mathsf{aExt} \colon \{0,1\}^n \to \{0,1\}^m$, where $c_d = \Theta(d^2 4^d)$ and $m = \Omega(\frac{n}{c_d})$.*

We now instantiate our generic construction with the average case seedless extractor for low-degree multi-variate polynomials in Lemma 23 to give a proof for Theorem 20.

Proof for Theorem 20. We instantiate the generic construction with the average case seedless extractor for sources recognizable by low-degree multi-variate polynomial functions in Lemma 23.

- $\mathsf{EXT} : \{0,1\}^n \times \{0,1\}^m \to \{0,1\}^m$ is instantiated as follows.

  $$\mathsf{EXT}(\mathsf{x}||\mathsf{y}) = \mathsf{aExt}(\mathsf{x}) + \mathsf{y},$$

  where $\mathsf{aExt} \colon \{0,1\}^n \to \{0,1\}^m$ is a seedless extractor for degree d multi-variate polynomials that can extract from $\frac{g}{\mathsf{t}} - \xi$ fraction of entropy, for an integer $1 \leq g \leq \mathsf{t} - 1$ and a small real $\xi > 0$, with extractor error ε_A. The inverter of EXT is

  $$\mathsf{EXT}^{-1}(\mathsf{s}) = (\mathsf{U}_n||\mathsf{aExt}(\mathsf{U}_n) + \mathsf{s}).$$

- ECC is a linear MDS code with parameter $[\mathsf{n}, \mathsf{t}, \mathsf{n}-\mathsf{t}+1]$ over $\mathbb{F}_q$, where $\log q = \frac{m+n}{\mathsf{t}}$. The output of $\mathsf{ECCenc}\colon \{0,1\}^{m+n} \to \mathbb{F}_q^{\mathsf{n}}$ is the share vector.

The composition of ECCenc and the leakage function $f \in \mathcal{L}_{d\text{-}\mathsf{poly}}$ remains a multi-variate polynomials over $\mathbb{F}_2$ with degree at most d. Since our analysis will be focusing on the extractor aExt, we derive the entropy bound on the first n bits of $\mathsf{EXT}^{-1}(\mathsf{U}_m)$. According to Lemma 9, the average (conditional) min-entropy of the random variable U_n conditioned on the view $\mathsf{V} = \mathsf{LEAK}(\mathsf{EXT}^{-1}(\mathsf{U}_m), \mathcal{L}_{d\text{-}\mathsf{poly}}, \beta, \theta)$ of an adaptive adversary corresponding to a leakage strategy $(\mathcal{L}_{d\text{-}\mathsf{poly}}, \beta, \theta)$ is bounded as follows.

$$\tilde{\mathsf{H}}_\infty(\mathsf{U}_n|\mathsf{V}) \geq n - (\log q)\theta - \beta.$$

Under the condition that $\theta + \frac{\beta}{\log q} \leq \mathsf{t} - g$, we have

$$\begin{aligned}\tilde{\mathsf{H}}_\infty(\mathsf{U}_n|\mathsf{V}) &\geq n - (\log q)(\mathsf{t}-g)\\ &= n - \tfrac{(m+n)(\mathsf{t}-g)}{\mathsf{t}}\\ &= \tfrac{n}{\mathsf{t}} - \tfrac{m(\mathsf{t}-g)}{\mathsf{t}},\end{aligned}$$

which is bigger than a $\frac{1}{\mathsf{t}} - \xi$ fraction of n for large enough n.

It then follows from the average case extractor property of aExt that

$$(\mathsf{V}, \mathsf{aExt}(\mathsf{U}_n)) \overset{\varepsilon_A}{\sim} (\mathsf{V}, \mathsf{U}_m).$$

From an application of Lemma 12, we obtain for a secret s,

$$(\mathsf{V}|\mathsf{aExt}(\mathsf{U}_n) = \mathsf{s}) \overset{2^m\varepsilon_A}{\sim} \mathsf{V}.$$

We then conclude that the scheme is a $(\mathcal{L}_{d\text{-}\mathsf{poly}}, \beta, \theta, 2^{m+2}\varepsilon_A)$-LR-SS.

According to Lemma 23, for any positive integer d, there is an efficient average case extractor $\mathsf{aExt}\colon \{0,1\}^n \to \{0,1\}^m$ for algebraic sources with parameters $\left((1-\frac{1}{c_d})n, d, 2^{-\Omega(\frac{n}{c_d})}\right)$, where $c_d = \Theta(d^2 4^d)$ and $m = \Omega(\frac{n}{c_d})$. We then require the threshold gap g to satisfy $\frac{g}{\mathsf{t}} > 1 - \frac{1}{c_d}$. In particular, once $c_d > 2$, we must have $g > \mathsf{t} - \frac{\mathsf{t}}{c_d} > \frac{\mathsf{t}}{2} \geq 1$ and do not have a threshold scheme. On the other hand, if we are satisfied with obtaining a leakage resilient ramp scheme, then privacy and leakage resiliency are guaranteed for all θ and β satisfying $\theta + \frac{\beta}{\log q} \leq \mathsf{t} - g < \frac{t}{c_d}$. The scheme shares a secret of $m = \Omega(\frac{n}{c_d})$ bits and each share contains $\log q = \frac{m+n}{\mathsf{t}}$ bits. So the information ratio is positive (a constant determined by c_d and t). Moreover, the leakage-resilience error $\varepsilon = 2^{m+2}\varepsilon_A$ can be made arbitrarily small. ◀

6 Local Leakage with Full Reconstruction

Srinivasan and Vasudevan considered a partial adaptive local leakage model that they called *strong local leakage* model and constructed such schemes for applications in leakage resilient secure multiparty computation. This model is the strongest local leakage model. We show an instantiation of our generic construction that yields competitive schemes in this model to illustrate the applicability of our generic construction to local leakage models.

▶ **Theorem 24.** *Let $\mathcal{L}_{\mathsf{local}}$ be the class of local functions. There is a family of t-out-of-n secret sharing schemes over $\mathbb{F}_q$ (labeled by $\log q$) that is simultaneously $(\mathcal{L}_{\mathsf{local}}, \ell(\mathsf{n}-\theta), \theta, \varepsilon)$-leakage resilient against a partially adaptive adversary that fully corrupts θ players for $\theta \in \{0, 1, \ldots, \mathsf{t}-2\}$ and, based on the shares of the θ players, chooses $\mathsf{n}-\theta$ arbitrary leakage* *functions each leaks* $\ell = \frac{(\mathsf{t}-\theta-8\xi)\log q}{(\mathsf{n}-\theta)(1+5\xi)}$ *bits for the remaining* $\mathsf{n}-\theta$ *shares, where* $\xi > 0$ *is a small real number, the share size* q *(determined by* ε *and* $q > \mathsf{n}$*) is a large enough power of* 2 *and the secret length* $m = \frac{\xi}{1+5\xi}\log q - o(\log q)$*. When* $\theta = \mathsf{t}-1$*, the partially adaptive adversary can only choose the* $\mathsf{n}-\mathsf{t}+1$ *arbitrary leakage functions based on* $\mathsf{t}-2$ *shares*

Proof. We instantiate the generic construction using a linear strong seeded extractor $\mathsf{Ext}\colon \{0,1\}^d \times \{0,1\}^n \to \{0,1\}^m$ and a linear MDS code. We also use a t-out-of-n Shamir's secret sharing to protect the seed such that the leakage at an individual share is independent of the seed.

- $\mathsf{EXT} : \{0,1\}^n \to \{0,1\}^m$ is instantiated as follows.

  $$\mathsf{EXT(x)} = \mathsf{Ext}(\mathsf{U}_d, \mathsf{x}),$$

  where Ext is the linear strong seeded extractor from 7 with parameters entropy requirement $3m$, extractor error $2^{-m}\varepsilon$ and seed length $d = 5m$. The uniform seed sampled during inverting is shared using t-out-of-n Shamir's secret sharing into $(\mathsf{SdSh}_1, \ldots, \mathsf{SdSh}_\mathsf{n}) \in \mathbb{F}^{\mathsf{n}}_{2^{5m}}$.
- ECC is a linear MDS code with parameter $[\mathsf{n}, \mathsf{t}, \mathsf{n}-\mathsf{t}+1]$ over $\mathbb{F}_{q_0}$, where $\log q_0 = \frac{d+n}{\mathsf{t}}$. The output of $\mathsf{ECCenc}\colon \{0,1\}^{d+n} \to \mathbb{F}^{\mathsf{n}}_{q_0}$ is $(\mathsf{SorSh}_1, \ldots, \mathsf{SorSh}_\mathsf{n})$. The final share vector is $((\mathsf{SdSh}_1||\mathsf{SorSh}_1), \ldots, (\mathsf{SdSh}_\mathsf{n}||\mathsf{SorSh}_\mathsf{n})) \in \mathbb{F}^{\mathsf{n}}_q$, where $q = 2^{5m}q_0$.

The leakage is independent of the seed. The seed is referring to the $5m$-bit uniform and independent seed sampled when inverting the extractor EXT. The seed is shared using the t-out-of-n Shamir's secret sharing and the shares are appended to the payload shares. The partially adaptive adversary corrupts arbitrary choice of θ players and obtains θ full shares. Since $\theta \leq \mathsf{t}-2$, the adversary not only obtains no information about the seed from the θ shares of the Shamir's scheme, but also can not obtain any information about the seed even if one more share is given. It then follows that the choice of individual leakage functions for the remaining $\mathsf{n}-\theta$ shares are independent of the seed. Note that the output of these individual leakage functions can depend on the seed, since Shamir's scheme over finite field of characteristic 2 is known to be not leakage resilient to (even non-adaptive) local leakage. We only need the fact that the choice (made after obtaining θ shares and before receiving information about the rest of the $\mathsf{n}-\theta$ shares) of these functions is independent of the seed for our security proof. This is because we use a strong seeded extractor to provide privacy and leakage resiliency for the payload scheme and by definition the security of a strong seeded extractor holds even if the seed is revealed, as long as the source is independent of the seed. Here the source is the uniform n-bit string conditioned on the information about it contained in the corrupted θ shares and the outputs of the individual leakage functions.

Achievable parameters.

$$\mathsf{U}_n = \mathsf{EXT}^{-1}(\mathsf{U}_m).$$

Let V be the adversary's view. We have

$$\tilde{\mathsf{H}}_\infty(\mathsf{U}_n|\mathsf{V}) \geq n - \theta\log q_0 - (\mathsf{n}-\theta)\ell.$$

The total entropy in the uniform n-bit string U_n is

$$n = \mathsf{t}\log q_0 - d = \mathsf{t}\log q_0 - 5m.$$

The amount of information about U_n contained in the θ shares is $\theta \log q_0$. The amount of information about U_n contained in the outputs of the leakage functions is

$$(\mathsf{n}-\theta)\ell = (\mathsf{n}-\theta)\frac{(\mathsf{t}-\theta-8\xi)\log q}{(\mathsf{n}-\theta)(1+5\xi)} = (\mathsf{t}-\theta-8\xi)\log q_0.$$

The remaining *average conditional min-entropy* is

$$\mathsf{t}\log q_0 - 5m - \theta\log q_0 - (\mathsf{t}-\theta-8\xi)\log q_0 = 8\xi\log q_0 - 5m,$$

where $m = \xi\log q_0 - o(\log q_0)$. This asserts that the remaining entropy is at least $3m$, sufficient for the average case extractor Ext. ◀

▶ Remark 25. When we set $\theta = \mathsf{t}-1$, our scheme is comparable with the scheme constructed in [43]. Both schemes allow the partial adaptive adversary to choose the individual leakage functions according to $\mathsf{t}-2$ shares and fully leak $\mathsf{t}-1$ shares. Our scheme achieves positive information ratio, which is an absolute constant, while the scheme in [43] achieves information ratio $\Omega(1/\mathsf{n})$. On the other hand, our scheme achieves an asymptotic leakage rate of $\frac{\mathsf{t}-\theta}{\mathsf{n}-\theta} = \frac{1}{\mathsf{n}-\mathsf{t}+1}$, while the scheme in [43] achieves asymptotic leakage rate 1. Note that our scheme belong to the sub-class of t-out-of-n secret sharing schemes that have *full reconstruction* property [38], since any t shares uniquely determine the rest of the shares. It is shown in [38], for such schemes, it is required that

$$\log q \geq \frac{\ell(\mathsf{n}-\mathsf{t})}{\mathsf{t}-\theta}, \text{ for any } \ell \geq 1.$$

The above inequality implies an upper bound on the asymptotic leakage rate.

$$\frac{\ell}{\log q} \leq \frac{\mathsf{t}-\theta}{\mathsf{n}-\mathsf{t}}.$$

When $\theta = \mathsf{t}-1$, we have

$$\frac{\ell}{\log q} \leq \frac{1}{\mathsf{n}-\mathsf{t}},$$

which holds even for non-adaptive adversary. This indicates that our scheme is already close to optimal with respect to leakage rate.

7 Conclusion

We started a systematic study of leakage-resilient secret sharing against *global* leakage, where the leakage oracle can access the full set of shares simultaneously, but the access is restricted to a special class of leakage functions. We studied such leakage-resilient secret sharing with respect to affine leakage functions and low-degree multi-variate polynomial leakage functions. We explicitly constructed threshold schemes with best leakage tolerance against affine functions. If the adversary is non-adaptive, our scheme is optimal both in terms of leakage tolerance and information ratio. For multi-variate polynomial leakage functions with degree bigger than one, our construction only yielded ramp schemes. As a result of independent interest, we showed that our approach to leakage-resilient secret sharing also yielded a competitive scheme compared with the state-of-the-art construction in the compartmentalized models.

References

1 Divesh Aggarwal, Ivan Damgård, Jesper Buus Nielsen, Maciej Obremski, Erick Purwanto, Joao Ribeiro, and Mark Simkin. Stronger leakage-resilient and non-malleable secret-sharing schemes for general access structures. *IACR Cryptology ePrint Archive*, page htps://eprint.iacr.org/2018/1147, 2018.

2 Divesh Aggarwal, Yevgeniy Dodis, Tomasz Kazana, and Maciej Obremski. Non-malleable reductions and applications. In *ACM SIGACT Symposium on Theory of Computing, STOC 2015*, pages 459–468, 2015.

3 Divesh Aggarwal, Yevgeniy Dodis, and Shachar Lovett. Non-malleable codes from additive combinatorics. *SIAM J. Comput.*, 47(2):524–546, 2018.

4 Divesh Aggarwal, Stefan Dziembowski, Tomasz Kazana, and Maciej Obremski. Leakage-resilient non-malleable codes. In *Theory of Cryptography Conference, TCC 2015*, pages 398–426, 2015.

5 Shashank Agrawal, Divya Gupta, Hemanta K. Maji, Omkant Pandey, and Manoj Prabhakaran. Explicit non-malleable codes against bit-wise tampering and permutations. In *Advances in Cryptology - CRYPTO 2015*, pages 538–557, 2015.

6 Shashank Agrawal, Divya Gupta, Hemanta K. Maji, Omkant Pandey, and Manoj Prabhakaran. A rate-optimizing compiler for non-malleable codes against bit-wise tampering and permutations. In *Theory of Cryptography Conference, TCC 2015*, pages 375–397, 2015.

7 Saikrishna Badrinarayanan and Akshayaram Srinivasan. Revisiting non-malleable secret sharing. *IACR Cryptology ePrint Archive*, page https://eprint.iacr.org/2018/1144, 2018.

8 Marshall Ball, Dana Dachman-Soled, Siyao Guo, Tal Malkin, and Li-Yang Tan. Non-malleable codes for small-depth circuits. In *IEEE Annual Symposium on Foundations of Computer Science, FOCS*, pages 826–837, 2018.

9 Marshall Ball, Siyao Guo, and Daniel Wichs. Non-malleable codes for decision trees. Cryptology ePrint Archive, Report 2019/379, 2019.

10 Fabrice Benhamouda, Akshay Degwekar, Yuval Ishai, and Tal Rabin. On the local leakage resilience of linear secret sharing schemes. In *Advances in Cryptology - CRYPTO 2018*, pages 531–561, 2018.

11 George R. Blakley. Safeguarding cryptographic keys. In *Proceedings of the 1979 AFIPS National Computer Conference*, pages 313–317, 1979.

12 Andrej Bogdanov, Yuval Ishai, Emanuele Viola, and Christopher Williamson. Bounded indistinguishability and the complexity of recovering secrets. *Advances in Cryptology -CRYPTO*, pages 593–618, 2016.

13 Jean Bourgain. On the construction of affine extractors. *Geometric and Functional Analysis*, 17(1):33–57, 2007.

14 Nishanth Chandran, Bhavana Kanukurthi, and Srinivasan Raghuraman. Information-theoretic local non-malleable codes and their applications. In *Theory of Cryptography - TCC*, pages 367–392, 2016.

15 Eshan Chattopadhyay, Vipul Goyal, and Xin Li. Non-malleable extractors and codes, with their many tampered extensions. In *ACM SIGACT Symposium on Theory of Computing, STOC 2016*, pages 285–298, 2016.

16 Eshan Chattopadhyay and Xin Li. Non-malleable codes and extractors for small-depth circuits, and affine functions. In *ACM SIGACT Symposium on Theory of Computing, STOC 2017*, pages 1171–1184, 2017.

17 Eshan Chattopadhyay and Xin Li. Non-malleable extractors and codes in the interleaved split-state model and more. *CoRR*, page http://arxiv.org/abs/1804.05228, 2018.

18 Mahdi Cheraghchi, Frédéric Didier, and Amin Shokrollahi. Invertible extractors and wiretap protocols. *IEEE Trans. Information Theory*, 58(2):1254–1274, 2012. doi:10.1109/TIT.2011.2170660.

19 Ronald Cramer, Ivan Damgård, and Jesper Buus Nielsen. *Secure Multiparty Computation and Secret Sharing*. Cambridge University Press, 2015. URL: `http://www.cambridge.org/de/academic/subjects/computer-science/cryptography-cryptology-and-coding/secure-multiparty-computation-and-secret-sharing?format=HB&isbn=9781107043053`.

20 Francesco Davì, Stefan Dziembowski, and Daniele Venturi. Leakage-resilient storage. In *Security and Cryptography for Networks SCN*, pages 121–137, 2010.

21 Yevgeniy Dodis, Rafail Ostrovsky, Leonid Reyzin, and Adam Smith. Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. *SIAM journal on computing*, 38(1):97–139, 2008.

22 Stefan Dziembowski, Tomasz Kazana, and Maciej Obremski. Non-malleable codes from two-source extractors. In *Advances in Cryptology - CRYPTO 2013*, pages 239–257, 2013.

23 Stefan Dziembowski and Krzysztof Pietrzak. Intrusion-resilient secret sharing. In *Foundations of Computer Science FOCS 2007*, pages 227–237, 2007.

24 Stefan Dziembowski, Krzysztof Pietrzak, and Daniel Wichs. Non-malleable codes. *J. ACM*, 65(4):20:1–20:32, 2018.

25 Antonio Faonio and Daniele Venturi. Non-malleable secret sharing in the computational setting: Adaptive tampering, noisy-leakage resilience, and improved rate. *IACR Cryptology ePrint Archive*, page https://eprint.iacr.org/2019/105, 2019.

26 Christina Fragouli and Emina Soljanin. *Network Coding Fundamentals*, volume 2. Foundations and Trends in Networking, 2007.

27 Vipul Goyal and Ashutosh Kumar. Non-malleable secret sharing. In *ACM SIGACT Symposium on Theory of Computing, STOC 2018*, pages 685–698, 2018.

28 Vipul Goyal and Ashutosh Kumar. Non-malleable secret sharing for general access structures. In *Advances in Cryptology - CRYPTO 2018*, pages 501–530, 2018.

29 Venkatesan Guruswami and Mary Wootters. Repairing reed-solomon codes. *IEEE Trans. Information Theory*, 63(9):5684–5698, 2017.

30 Masahito Hayashi and Toyohiro Tsurumaru. More efficient privacy amplification with less random seeds via dual universal hash function. *IEEE Transactions on Information Theory, Vol 62, Issue 4, 2213 - 2232.*, 2016.

31 Ashutosh Kumar, Raghu Meka, and Amit Sahai. Leakage-resilient secret sharing against colluding parties. *Foundations of Computer Science, FOCS 2019*, 2019.

32 Fu Li and David Zuckerman. Improved extractors for recognizable and algebraic sources. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM*, pages 72:1–72:22, 2019.

33 Xin Li. A new approach to affine extractors and dispersers. *IEEE Conference on Computational Complexity, CCC 2011*, pages 137–147, 2011.

34 Xin Li. Improved non-malleable extractors, non-malleable codes and independent source extractors. In *ACM SIGACT Symposium on Theory of Computing, STOC 2017*, pages 1144–1156, 2017.

35 Xin Li. Pseudorandom correlation breakers, independence preserving mergers and their applications. *Electronic Colloquium on Computational Complexity (ECCC)*, 25:28, 2018.

36 Fuchun Lin, Mahdi Cheraghchi, Venkatesan Guruswami, Reihaneh Safavi-Naini, and Huaxiong Wang. Secret sharing with binary shares. In *Innovations in Theoretical Computer Science Conference, ITCS 2019*, pages 53:1–53:20, 2019.

37 Feng-Hao Liu and Anna Lysyanskaya. Tamper and leakage resilience in the split-state model. In *Advances in Cryptology - CRYPTO*, pages 517–532, 2012.

38 Jesper Buus Nielsen and Mark Simkin. Lower bounds for leakage-resilient secret sharing. *IACR Cryptology ePrint Archive*, page https://eprint.iacr.org/2019/181, 2019.

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

40 Ran Raz, Omer Reingold, and Salil P. Vadhan. Extracting all the randomness and reducing the error in Trevisan's extractors. *J. Comput. Syst. Sci.*, 65(1):97–128, 2002.

41 Ronen Shaltiel. How to get more mileage from randomness extractors. In *IEEE Conference on Computational Complexity (CCC) 2006*, pages 46–60, 2006.

42 Adi Shamir. How to share a secret. *Commun. ACM*, 22(11):612–613, 1979.

43 Akshayaram Srinivasan and Prashant Nalini Vasudevan. Leakage resilient secret sharing and applications. *Advances in Cryptology - CRYPTO 2019*, pages 480–509, 2019.

44 Douglas R. Stinson. An explication of secret sharing schemes. *Des. Codes Cryptography*, 2(4):357–390, 1992.

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