Secret Sharing with Binary Shares

Fuchun Lin[*] Mahdi Cheraghchi[†] Venkatesan Guruswami[‡]
Reihaneh Safavi-Naini[§] Huaxiong Wang[*]

August 9, 2018

Abstract

Secret sharing is a fundamental cryptographic primitive. One of the main goals of secret sharing is to share a long secret using small shares. In this paper we consider a family of statistical secret sharing schemes indexed by N, the number of players. The family is associated with a pair of relative thresholds τ and κ, that for a given N, specify a secret sharing scheme with privacy and reconstruction thresholds, $N\tau$ and $N\kappa$, respectively. These are non-perfect schemes with gap $N(\kappa - \tau)$ and statistical schemes with errors $\varepsilon(N)$ and $\delta(N)$ for privacy and reconstruction, respectively. We give two constructions of secret sharing families as defined above, with security against (i) an adaptive, and (ii) a non-adaptive adversary, respectively. Both constructions are modular and use two components, an invertible extractor and a stochastic code, and surprisingly in both cases, for any $\kappa > \tau$, give explicit families for sharing a secret that is a constant fraction (in bits) of N, using binary shares. We show that the construction for non-adaptive adversary is optimal in the sense that it asymptotically achieves the upper bound $N(\kappa-\tau)$ on the secret length. We relate our results to known works and discuss open questions.

1 Introduction

Secret sharing was introduced independently by Blakely [3] and Shamir [33], and is a fundamental cryptographic primitive with numerous important applications including Multiparty Computation (MPC) and Threshold Cryptography [40]. In a *threshold secret sharing scheme* a dealer shares a secret $\mathbf{s}$ among a set of N players such that: (i) a set of up to t players learn no information about the secret, and (ii) a set of $t+1$ players can (efficiently) recover the secret. The requirements for privacy is that no information be leaked to unauthorised sets and, secret recovery for authorised set is with probability 1. In *statistical secret sharing schemes* these requirements are relaxed: correctness holds with high probability and for any two secrets, the statistical distance of the shares of an unauthorised set, is small. One of the key questions in secret sharing is the required share length for a given secret length. Threshold secret sharing schemes require that the share length (in bits) is at least equal to the secret length. Statistical threshold secret sharing schemes can only marginally relax this bound. Blakley and Meadows [17] introduced ramp secret sharing schemes with the goal of improving share efficiency by relaxing threshold property. Ramp schemes are *non-perfect secret*

[*] Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, SG

[†] Department of Computing, Imperial College London, UK

[‡] Computer Science Department, Carnegie Mellon University, USA

[§] Department of Computer Science, University of Calgary, CA

sharing schemes and allow some unauthorised sets that cannot recover the secret, learn partial information about it. The access structure of a ramp scheme is described by two thresholds t and k: a set of at most t players (called *forbidden sets*) learn nothing, a set of k players can reconstruct the secret and more than t but less than k players learn partial information. It was shown that the ramp schemes constructed in [17] has the share length equal to a fraction of the secret length, that is $\frac{1}{k-t}$ of the secret length. That was later proved to be the shortest possible share length for the specified thresholds, and so optimal [36]. Families of linear ramp schemes satisfy an additional homomorphic property and have found wide applications in MPC [10, 9, 6, 7, 5, 20] and distributing cryptographic functions [16, 28].

Our contributions:

Our goal is to construct Secret Sharing Schemes (SSS) with ramp access structure that is defined by a privacy threshold and a reconstruction threshold, and with *the shortest share length.* We thus consider *binary shares.* We use the relaxed notion of ε statistical privacy for forbidden sets, and δ correctness for authorised sets, and define a family of SSS indexed by N, the number of players, as follows. For two (constant) relative thresholds τ and κ, where $\tau, \kappa \in [0, 1]$ and $\kappa > \tau$, a binary $(\varepsilon(N), \delta(N))$-SSS for N players, is a ramp SSS where the leakage measured in statistical distance of secret to a set of corrupted players of size at most $t(N) = N\tau$ is bounded by $\varepsilon(N)$, and a set of $k(N) = N\kappa$ players can find the secret, with probability at least $1 - \delta(N)$. It is assumed that $\varepsilon(N), \delta(N)$ are negligible in N. When clear from context, we omit N and write ε, δ, t and k. We consider two types of adversaries: non-adaptive adversaries who choose the set of corrupted players in one step, and adaptive adversaries who corrupt the players (up to t) adaptively, at each step taking into account all their current information (see Definition 5). It is easy to see that perfect privacy implies independence of the shares of a forbidden set from the secret, and so equivalence of adaptive and non-adaptive adversary. This argument however does not hold for statistical privacy and as seen in our constructions, our non-adaptive secure construction becomes insecure against adaptive adversaries (see Remark 2).

We give two computationally efficient constructions that, in the case of non-adaptive adversary, asymptotically achieves the best achievable secret length. In adaptive case, although the upper bound from non-adaptive case applies, achieving the bound remains an open question. The constructions are both modular with two building blocks: Stochastic Affine-Erasure Correcting Code (SA-ECC) and randomness extractors (see Section 2 for basic definitions).

Adaptive adversary:

An affine extractor AExt is called invertible if there is an efficient algorithm AExt^{-1} that samples a random pre-image of any given extractor output (see Definition 7). An affine extractor $\mathsf{AExt} : \{0, 1\}^n \to \{0, 1\}^\ell$ is $(n - t, \frac{\varepsilon}{2})$-almost perfect if the probability that each output occurs, is bounded within $\left(2^{-\ell}(1 - \frac{\varepsilon}{2}), 2^{-\ell}(1 + \frac{\varepsilon}{2})\right)$ when its source has at least $n - t$ bits of entropy (see Definition 8). There are affine extractors (see Lemma 4) that can be transformed into an invertible, almost perfect affine extractor that extracts a constant fraction of uniform bits given a constant fraction of source entropy (see comment on this in Appendix B).

A stochastic code has a randomised encoder and a deterministic decoder. It is called a stochastic affine code if for any value of the encoder randomness, the encoder is an affine function. We construct a Stochastic Affine-Erasure Correcting Code (SA-ECC) by adapting a construction in [19], which uses an erasure correcting code and a list of pseudo-random objects (generators, permutations and samplers). In particular, we show there is an explicit SA-ECC that corrects p fraction of (oblivious) adversarial erasures and achieves the rate $1 - p$ (see Theorem 5).

Theorem 3. **(informal)** Let $\mathsf{AExt} : \{0,1\}^n \to \{0,1\}^\ell$ be an invertible $(n-t, \frac{\varepsilon}{2})$-almost perfect affine extractor. Let $(\mathsf{SA\text{-}ECCenc}, \mathsf{SA\text{-}ECCdec})$ be a stochastic affine-erasure correcting code that tolerates $N-k$ erasures and decodes with success probability at least $1-\delta$. Then the $(\mathsf{Share}, \mathsf{Recst})$ defined as follows is an adaptive (ε, δ)-SSS with threshold pair (t, k).

$$\begin{cases} \mathsf{Share}(\mathbf{s}) & = \mathsf{SA\text{-}ECCenc}(\mathsf{AExt}^{-1}(\mathbf{s})); \\ \mathsf{Recst}(\tilde{\mathsf{v}}) & = \mathsf{AExt}(\mathsf{SA\text{-}ECCdec}(\tilde{\mathsf{v}})). \end{cases}$$

The (k, δ)-reconstruction property follows directly from the erasure correction guarantee of the $\mathsf{SA\text{-}ECC}$. The adaptive (t, ε)-privacy property relies on the security of the AExt: uniform output of AExt means perfect privacy, which is explained as follows. To quickly see the intuition of the construction, one can tactically assume that the secret $\mathbf{S}$ is uniformly distributed on $\{0,1\}^\ell$ (our formal proof does not use uniform secret assumption). According to the definition of an inverter, $\mathsf{AExt}^{-1}(\mathbf{S})$ is uniformly distributed on $\{0,1\}^n$. We do not need the randomness of the $\mathsf{SA\text{-}ECC}$ for providing (t, ε)-privacy. We consider any value of the encoding randomness of $\mathsf{SA\text{-}ECC}$ and only use the fact that its encoder is an affine function. After revealing the answers to up to t query bits, the conditional distribution of the $\mathsf{AExt}^{-1}(\mathbf{S})$ on any value of the query answers has at least $n-t$ bits entropy. The affine property of $\mathsf{SA\text{-}ECC}$ here plays a crucial role for guaranteeing an affine structure in this conditional distribution. This is because each codeword bit of the $\mathsf{SA\text{-}ECC}$ gives a linear equation about its message $\mathsf{AExt}^{-1}(\mathbf{S})$ (as n unknown bits). Knowing t codeword bits amount to imposing t linear equations on the n unknowns and resulting in an affine distribution. Now if the output of AExt with respect to this conditional distribution is perfectly uniform, the distribution of the answer to up to t bits query is independent of the secret (since before and after the query, the secret has the same uniform distribution). On the other hand, when the output of AExt is not perfectly uniform but with a non-zero error, one has to carefully bound this error. This is handled by requiring the almost perfect property from AExt.

With the aforementioned instantiations, the $\mathsf{SA\text{-}ECC}$ is optimal: $\frac{n}{N} = \kappa - o(1)$ (Theorem 5). The entropy requirement of AExt is $n - t = N(\kappa - \tau) + o(N)$, which is a constant fraction of N (also n), for any $\kappa > \tau$. With a constant fraction of entropy, $\ell = \Omega(n)$ according to Lemma 4. The construction in Theorem 3 shows the following.

Corollary 6. **(informal) There is an explicit family of adaptive (ε, δ)-SSS with relative threshold pair (τ, κ) that shares $\Omega(N)$ bits secret among N users for any (τ, κ) pair, $0 \le \tau < \kappa \le 1$.**

Non-adaptive adversary:

A strong seeded extractor $\mathsf{Ext} : \{0,1\}^d \times \{0,1\}^n \to \{0,1\}^\ell$ for an n-bit source and using d bits of randomness, is a family of functions $\mathsf{Ext}(\mathsf{z}, \cdot) : \{0,1\}^n \to \{0,1\}^\ell$ each labeled by a seed $\mathsf{z} \in \{0,1\}^d$. Let ε_z denote the statistical distance from uniform distribution U_ℓ of the output of $\mathsf{Ext}(\mathsf{z}, \cdot)$ (with respect to a source). By the definition of a seeded extractor, we should have that (with respect to a source) $\mathbb{E}_\mathsf{z} \varepsilon_\mathsf{z} \le \varepsilon$. A linear strong seeded extractor means that each function $\mathsf{Ext}(\mathsf{z}, \cdot)$ is linear. If the source is an affine source X with flat distribution on $\mathsf{Supp}(\mathsf{X}) \subset \{0,1\}^n$, the output of $\mathsf{Ext}(\mathsf{z}, \cdot)$ is perfectly uniform if and only if the linear function $\mathsf{Ext}(\mathsf{z}, \cdot)$ restricted to $\mathsf{Supp}(\mathsf{X})$ is surjective, namely, $\mathsf{Ext}(\mathsf{z}, \mathsf{Supp}(\mathsf{X})) = \{0,1\}^\ell$. More importantly, once the equality does not hold, $\mathsf{Ext}(\mathsf{z}, \mathsf{Supp}(\mathsf{X}))$ is at most an $\ell - 1$ dimensional subspace of $\{0,1\}^\ell$. We then have $\mathsf{SD}(\mathsf{Ext}(\mathsf{z}, \mathsf{X}); U_\ell) \ge \frac{1}{2}$. This gives that $\varepsilon_\mathsf{z} = 0$ for at least $1 - 2\varepsilon$ fraction of the seeds. Another benefit of linearity is that linear functions are efficiently invertible. So a natural inverter for a linear seeded extractor is to sample a uniform seed Z and invert the linear function labeled by Z: $(\mathsf{Z} || \mathsf{Ext}^{-1}(\mathsf{Z}, \cdot))$, where $\mathsf{Z} \overset{\$}{\leftarrow} \{0,1\}^d$. This transformation incurs an overhead of d bits. But if d is negligible in N, it is considered for free. There are explicit linear strong seeded extractors with a negligible seed length that extract all the randomness (see Lemma 8).

Our construction for non-adaptive adversary uses the same optimal $\mathsf{SA\text{-}ECC}$ as in previous construction but replaces the affine extractor with a linear strong seeded extractor.

Theorem 7. (informal) Let $\mathsf{Ext} : \{0,1\}^d \times \{0,1\}^n \to \{0,1\}^\ell$ be a linear strong seeded $(n-t, \frac{\varepsilon}{2})$-extractor. Let $(\mathsf{SA\text{-}ECCenc}, \mathsf{SA\text{-}ECCdec})$ be a stochastic affine-erasure correcting code that tolerates $N-k$ erasures and decodes with success probability at least $1-\delta$. Then the following coding scheme $(\mathsf{Share}, \mathsf{Recst})$ is a non-adaptive (ε, δ)-SSS with threshold pair (t, k).

$$\begin{cases} \mathsf{Share}(\mathbf{s}) & = \mathsf{SA\text{-}ECCenc}(\mathsf{Z}||\mathsf{Ext}^{-1}(\mathsf{Z}, \mathbf{s})), \text{where } \mathsf{Z} \xleftarrow{\$} \{0,1\}^d; \\ \mathsf{Recst}(\tilde{\mathsf{v}}) & = \mathsf{Ext}(\bar{\mathsf{z}}, \bar{\mathsf{x}}), \text{where } (\bar{\mathsf{z}}||\bar{\mathsf{x}}) = \mathsf{SA\text{-}ECCdec}(\tilde{\mathsf{v}}). \end{cases}$$

Again, the (k, δ)-reconstruction property follows directly from the erasure correction guarantee of the $\mathsf{SA\text{-}ECC}$. Similar to the analysis of previous construction, the affine property of $\mathsf{SA\text{-}ECC}$ helps in guaranteeing an affine distribution. Now using the above discussion about strong seeded extractors, we have the output of $\mathsf{Ext}(\mathsf{z}, \cdot)$ is perfectly uniform for $1-\varepsilon$ fraction of the seeds, which directly implies the non-adaptive (t, ε)-privacy. Here we do not need the almost perfect property as required for the previous construction. On the other hand, we crucially rely on the fact that the reading of the adversary is not adaptive. An adaptive adversary can first choose the query bits to learn about the extractor seed Z, and then once the seed is learnt (to some extent) chooses the remaining query bits. In this case, the source (the conditional distribution induced by the choice of reading positions) is not independent of the seed and we can not use the definition of the strong seeded extractor.

With the aforementioned specific instantiation, the $\mathsf{SA\text{-}ECC}$ is optimal: $\frac{n+d}{N} = \kappa - o(1)$ (Theorem 5). The entropy requirement of Ext is $n - t = n - N\tau = N(\kappa - \tau) - d - o(N)$. The extractor Ext extracts all the randomness $\ell = n - t - o(n)$ and the seed length is negligible $d = o(n)$ (see Lemma 8). We then have $\ell = N(\kappa - \tau) - o(N)$. The construction in Theorem 7 shows the following.

Corollary 9. (informal) There is an explicit family of non-adaptive (ε, δ)-SSS that shares $N(\kappa - \tau) - o(N)$ bits secret among N users for any (τ, κ) pair, $0 \le \tau < \kappa \le 1$.

Bounds and optimality:

For each value of N, we have a $(\varepsilon(N), \delta(N))$-SSS with thresholds $N\tau$ and $N\kappa$ for privacy and reconstruction respectively. Using the bound in [36, Theorem 13], gives $1 \ge \frac{\ell(N)}{N(\kappa-\tau)}$ and so $\frac{\ell(N)}{N} \le \kappa - \tau$, where $\ell(N)$ is the length of the secret. But this bound was proved for the case $\varepsilon(N) = \delta(N) = 0$. We use a connection to the *Wyner wiretap channel* to show that asymptotically the same bound holds even when $\varepsilon(N)$ and $\delta(N)$ are not zero, but negligible. (see *Connections to wiretap channel* below for more details and Appendix A for a proof).

Our construction for non-adaptive adversary is optimal in the sense that it achieves the above asymptotic bound. Our construction for adaptive adversary however fails to achieve the above asymptotic bound. There are a few obstacles that prevent it from achieving the above asymptotic bound. There are known constructions of binary affine extractors that extract all the randomness [32]. Such affine extractors can be made invertible while only incurring negligible overhead [8]. But unfortunately, these affine extractors do not have almost perfect property. Nevertheless, our construction for adaptive adversary achieves the following "relative threshold" property. For any arbitrarily small relative gap $\xi = \kappa - \tau > 0$, there are explicit (ε, δ)-SSS with security against adaptive adversary that share $\Omega(N)$ bits of secret. Known adaptive (ε, δ)-SSS before this work do not have vanishing relative gap. Example 1 in Section 3 summarises a probabilistic construction of ramp SSS from random linear codes which shows existence of adaptive $(0, 0)$-SSS with secret length $\ell = N(1 - h_2(\tau) - h_2(1 - \kappa)) - o(N)$ provided that $1 - h_2(\tau) - h_2(1 - \kappa) > 0$. This construction is not explicit and the relative gap $\kappa - \tau > \frac{1}{2}$.

Connections to wiretap channel:

In the Wyner wiretap channel model [39, 13], there are two Discrete Memoryless Channels (DMC): one for the legitimate receiver called the *main channel*, the other for the eavesdropper called the *wiretapper channel*. The coding problem for the wiretap channel model is to provide privacy against an eavesdropper observing from the wiretapper channel and reliability for the legitimate receiver observing from the main channel (see Section 2 for a brief introduction and discussion for different privacy definitions). In Appendix A, we prove the asymptotic upper bound $\frac{\ell(N)}{N} \leq \kappa - \tau$ for the non-adaptive (ε, δ)-SSS by reducing it to a wiretap code for a wiretap channel with a pair of Binary Erasure Channels (BEC).

Later, Ozarow and Wyner proposed the wiretap channel II, where an adversary observes arbitrary t out of the total n bits of the communication [29]. Let us call the wiretapper channel of the wiretap channel II model a $\frac{n-t}{n}$-*erasure channel*. The non-adaptive (ε, δ)-SSS with relative threshold pair (τ, κ) can be interpreted as a *generalised wiretap II*, where the wiretapper channel is a $(1 - \tau)$-erasure channel and the main channel is a $(1 - \kappa)$-erasure channel. Our construction in Theorem 7 gives a capacity achieving construction in this model. The contributions are three-fold. Firstly, we find the capacity of this generalised wiretap II model by proving a lower bound that matches a trivial upper bound, which can be derived through a reduction to Wyner wiretap channel with two BEC's similar to Appendix A. Secondly, we show the lower bound remains tight even when the privacy of the wiretap II is strengthened to indistinguishability. Thirdly, the lower bound is proved through giving an explicit construction. This answers an open question posted in [1]. The authors studied a generalisation of the wiretap II model, where the adversary chooses t bits to observe and erases them. They showed that the rate $1 - \tau - h_2(\tau)$ (while the quantity is non-negative) can be achieved and left open the question of whether a higher rate is achievable. Our result shows that in their model, the rate $1 - 2\tau$ can be explicitly achieved.

Related work:

Using linear codes for constructing linear secret sharing schemes dates back to [27, 22, 25]. These constructions are for threshold SS with perfect secrecy, and use special classes of linear codes. In [11], a general construction of ramp secret sharing from a linear code and a universal hash family is given. The construction uses results from [7] that relate linear codes and linear secret sharing.The construction has perfect privacy and δ-reconstruction. The construction has similarity with our non-adaptive construction if we consider the universal hash family as an extractor. The two constructions however are completely different, not only because of the type of privacy that they offer (perfect versus statistical) but the way extractor is used and the security is proved. In [11], a hash function h is randomly chosen from the family, and the pre-image of the secret under h, is encoded using an error correcting code. The security of the construction holds with overwhelming probability (depending on the choice of h). This is very different from our construction in Theorem 7 where the seed (description of the function) is part of the input to the stochastic code that is used to recover the secret from $N\kappa$ erasures, and security is proved directly (always hold). We also remark that the results in [7] shows a construction of linear ramp schemes by randomly generating the generator matrices of linear codes. The construction will support long secret and will succeed with overwhelming probability.

The idea of using the combination of a linear error correcting code and an invertible affine extractor was first proposed in [8]. The coding problem they used this construction for is a wiretap protocol with active adversary. A wiretap protocol is the wiretap channel II with a slightly generalised security definition (see Section 2 for brief introduction to wiretap channel). Our construction for adaptive adversary uses a combination of a stochastic affine-erasure correcting code and an invertible almost perfect affine extractor. On one hand, the coding part is generalised from linear to stochastic affine. On the other hand, our affine extractor is more restricted. These two changes are crucial for achieving the arbitrarily small relative threshold gap property and adaptive privacy, respectively. The construction in [8] only achieves security for uniform message.

In [21], (ε, δ)-secret sharing with adaptive and non-adaptive adversaries, as well as robust SS where the adversary can tamper with shares is considered. The goal there is to relax privacy to achieve large secrets over small alphabets, which is similar to ours. However the paper considers high privacy threshold and perfect reconstruction from the whole share set (and not a reconstruction threshold $N\kappa$). Their main results for adaptive (passive) adversary ([21, Theorem 1.2]) and active adversary ([21, Theorem 1.6]) are for binary shares when the complete share vector is input to the reconstruction algorithm.

2 Preliminaries

A binary error correcting code (or simply a code) of length n is defined by its code book, which is a subset of $\{0,1\}^n$. The encoder of the code is a deterministic function taking a message to its corresponding codeword. The decoder is also deterministic and has certain error-correcting property.

Definition 1. The Hamming weight $wt(\mathbf{c})$ of a vector $\mathbf{c} \in \{0,1\}^n$ is the number of non-zero positions in $\mathbf{c}$. The minimum distance d_C of a code C is defined as $\min\{wt(\mathbf{c}-\mathbf{c}')|\mathbf{c},\mathbf{c}' \in C, \mathbf{c} \neq \mathbf{c}'\}$. For a subspace $C \subset \{0,1\}^n$, the minimum distance is then $\min\{wt(\mathbf{c})|\mathbf{c} \in C\backslash\{0^n\}\}$. The ratio $\frac{d_C}{n}$ is referred to as the relative distance of the code.

An $[n,k,d]$-code C is defined to be a k-dimensional subspace of $\{0,1\}^n$ with $d(C) = d$.

Definition 2. The dual code $C^\perp$ of a code C consists of all $\mathbf{c}' \in \{0,1\}^n$ such that $\langle \mathbf{c}', \mathbf{c} \rangle = 0$ for all $\mathbf{c} \in C$, where $\langle \cdot, \cdot \rangle$ denotes the standard inner product. Whenever d is used to denote the minimum distance of C, $d^\perp$ is used to denote the minimum distance of $C^\perp$.

A stochastic code has a randomised encoder and a deterministic decoder. The encoder $\mathsf{Enc} : \{0,1\}^k \times \mathcal{R} \rightarrow \{0,1\}^n$ uses local randomness $R \leftarrow \mathcal{R}$ to encode a message $\mathbf{m} \in \{0,1\}^k$. The decoder is a deterministic function $\mathsf{Dec} : \{0,1\}^n \rightarrow \{0,1\}^k \cup \{\perp\}$. The decoding probability is defined over the encoding randomness $R \leftarrow \mathcal{R}$. In our construction, we use a stochastic code that can protect against an (oblivious) adversarial erasure channel, where the adversary can choose any subset of shares to erase. The channel captures the provision of the incomplete share vector for reconstruction.

Randomness extractors. Randomness extractors extract close to uniform bits from input sequences that are not uniform but have some guaranteed entropy. The closeness to uniform of the extractor output is measured by the statistical distance (half the ℓ_1-norm). For two random variables $\mathsf{X}, \mathsf{Y} \leftarrow \Omega$, the statistical distance between X and Y (or their distributions) is defined as,

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

We say X and Y are ε-close if $\mathsf{SD}(\mathsf{X},\mathsf{Y}) \leq \varepsilon$. A *randomness source* is a random variable with lower bound on its min-entropy, which is defined by $\mathsf{H}_\infty(\mathsf{X}) = -\log \max_{\mathbf{x}}\{\Pr[\mathsf{X} = \mathbf{x}]\}$. 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 3. A polynomial time function $\mathsf{AExt} : \{0,1\}^n \rightarrow \{0,1\}^m$ is an affine (k,ε)-extractor if for any affine (n,k)-source X, we have

$$\mathsf{SD}(\mathsf{AExt}(\mathsf{X}); U_m) \leq \varepsilon.$$

More specially, a *bit-fixing* (n,k)-source is a random variable $\mathsf{X} = (\mathsf{X}_1, \cdots, \mathsf{X}_n)$, where at least k of the coordinates are uniformly and independently distributed on $\{0,1\}$ while the rest have fixed values. A *bit-fixing extractor* can be similarly defined as one extracts close to uniform bits from bit-fixing source. Since every bit-fixing source is an affine source, an affine extractor is also a bit-fixing extractor.

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 4. A polynomial time function $\mathsf{Ext} : \{0,1\}^d \times \{0,1\}^n \rightarrow \{0,1\}^m$ is a strong seeded (k,ε)-extractor if for any (n,k)-source X, we have

$$\mathsf{SD}(\mathsf{S}, \mathsf{Ext}(\mathsf{S},\mathsf{X}); \mathsf{S}, U_m) \leq \varepsilon,$$

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.

Wiretap channel. In the Wyner wiretap channel model [39, 13], there are two Discrete Memoryless Channels (DMC): one for the legitimate receiver called the *main channel*, the other for the eavesdropper called the *wiretapper channel*. The coding problem is to design a *wiretap code* that provides privacy of the message against a passive adversary that observes the communication from the wiretapper channel and reliability for the legitimate receiver whose observation is corrupted due to the main channel. The privacy of the message has been defined with respect to a uniformly distributed message $\mathsf{M} \overset{\$}{\leftarrow} \{0,1\}^m$ and requires that $\mathsf{H}(\mathsf{M}|\mathsf{W}) \geq m(1-o(1))$, where $\mathsf{H}(\mathsf{M}|\mathsf{W})$ is the conditional Shannon entropy and W denotes the view of the random codeword from the wiretapper channel. The highest information rate of wiretap codes for a pair of DMC's is called the *secrecy capacity*. Basic examples of binary DMC are the Binary Symmetric Channel (BSC) and the Binary Erasure Channel (BEC). A BEC_p is a probabilistic transformation that maps inputs 0 and 1 to a symbol ? that denotes erasure with probability p and to the inputs themselves with probability $1-p$. For a main channel - wiretapper channel pair (BEC_{p_m},BEC_{p_w}) such that $p_m < p_w$, it is known that the secrecy capacity is the difference of the respective channel capacities: $(1-p_m)-(1-p_w) = p_w - p_m$.

Remark 1. The privacy definition above is later strengthen to the so-called *strong secrecy* [26] $\mathsf{H}(\mathsf{M}|\mathsf{W}) = m - o(1)$. Strong secrecy is also called *mutual information security for random message* in [2], i.e. $\mathsf{I}(\mathsf{M};\mathsf{W}) = o(1)$, $\mathsf{M} \overset{\$}{\leftarrow} \{0,1\}^m$, where a strictly stronger security measure that removes the uniform message condition is defined,

$$\max_{\mathsf{M} \leftarrow \{0,1\}^m} \mathsf{I}(\mathsf{M};\mathsf{W}) = o(1),$$

and shown to be equivalent to the standard cryptography semantic security and indistinguishability security, with the latter defined by,

$$\mathsf{SD}(\mathsf{W}_{\mathsf{m}_0};\mathsf{W}_{\mathsf{m}_1}) = o(1), \tag{1}$$

where $\mathsf{W}_{\mathsf{m}_i}$ denotes the wiretapper channel view of the random codeword encoding the message m_i, $i = 0, 1$, for any two messages m_0 and m_1.

Later, Ozarow and Wyner proposed the wiretap channel II, where an adversary observes arbitrary t out of the total n bits of the communication [29]. In the original wiretap channel II model, the legitimate receiver's observation is the full n bits of codeword. There have been efforts on generalising the wiretap channel II to one with a non-perfect main channel. For example, two generalised wiretap channel II models with an adversary that can erase and respectively modify the t bits observed are studied in [1]. For the erasure case, the authors show that a coding rate of $1 - \mu - h_2(\mu)$, where $\mu = \frac{t}{n}$ is achievable, provided that the quantity is non-negative. It is also commented that the same rate can be achieved even when the observed bits and erased bits are different. But whether higher coding rate can be achieved was left as an open question. Another line of works consider the Adversarial Wiretap (AWTP) model, where the adversary can read a ρ_r fraction of the codeword components and either erase [38] or additively tamper [37] a ρ_w fraction of the codeword components. This line of works consider big alphabet and do not yield results in binary.

3 (ε, δ)-SSS

Definition 5. A $(\varepsilon(N), \delta(N))$-SSS with relative threshold pair (τ, κ) for $0 \leq \tau < \kappa \leq 1$ is a pair of polynomial-time algorithms $(\mathsf{Share}, \mathsf{Recst})$,

$$\mathsf{Share}: \{0,1\}^{\ell(N)} \times \mathcal{R} \to \{0,1\}^N,$$

where $\mathcal{R}$ denote the randomness set, and

$$\mathsf{Recst}: \widetilde{\{0,1\}}^N \to \{0,1\}^{\ell(N)} \cup \{\perp\},$$

where $\widetilde{\{0,1\}}^N$ denotes the subset of $(\{0,1\} \cup \{?\})^N$ with at least $N\kappa$ components not equal to the erasure symbol "?", that satisfy the following properties.

- (κ, δ)-reconstruction: Given $k(N) = N\kappa$ correct shares of a share vector $\mathsf{Share}(\mathbf{s})$, the reconstruct algorithm Recst reconstructs the secret $\mathbf{s}$ with probability at least $1 - \delta(N)$. When $\delta(N) = 0$, we say the SSS has perfect reconstruction.
- (τ, ε)-privacy (non-adaptive/adaptive):
  - Non-adaptive: for any $\mathbf{s}_0, \mathbf{s}_1 \in \{0,1\}^{\ell(N)}$, any $A \subset [N]$ of size $|A| \leq t(N) = N\tau$,
  $$\mathsf{SD}(\mathsf{Share}(\mathbf{s}_0)_A; \mathsf{Share}(\mathbf{s}_1)_A) \leq \varepsilon(N). \tag{2}$$
  - Adaptive: for any $\mathbf{s}_0, \mathbf{s}_1 \in \{0,1\}^{\ell(N)}$ and any adversary $\mathcal{A}$ adaptively reads up to $t(N) = N\tau$ shares,
  $$\left|\mathsf{Pr}[\mathcal{A}^{\mathsf{Share}(\mathbf{s}_0)}(\mathbf{s}_0, \mathbf{s}_1) = 1] - \mathsf{Pr}[\mathcal{A}^{\mathsf{Share}(\mathbf{s}_1)}(\mathbf{s}_0, \mathbf{s}_1) = 1]\right| \leq \varepsilon(N). \tag{3}$$

When $\varepsilon(N) = 0$, we say the SSS has perfect privacy. When $\varepsilon(N) > 0$ and (2) is satisfied, we have a non-adaptive $(\varepsilon(N), \delta(N))$-SSS. If (3) is further satisfied, we have an adaptive $(\varepsilon(N), \delta(N))$-SSS.

When clear from context, write $\varepsilon, \delta, t, k, \ell$ instead of $\varepsilon(N), \delta(N), t(N), k(N), \ell(N)$.

Definition 5 with $\varepsilon = \delta = 0$ is the so called *ramp* (also called *quasi-threshold*) SSS with t-privacy and k-reconstruction constructed in [7]. It is easy to see that perfect privacy implies independence of the shares of a forbidden set from the secret, and so equivalence of adaptive and non-adaptive adversary. So any ramp scheme is an adaptive SSS. We then have the following example of adaptive $(0, 0)$-SSS.

Example 1. (follows from Theorem 4 and Corollary 1 of [7]) By sampling a generator matrix of a linear code of length N uniformly at random, with high probability one obtains a code C' with minimum distance $d_{C'} = N(1-\kappa)$ who has a sub-code $C \subset C'$ with dual distance $d_C^{\perp} = N\tau$, as long as $h_2(\tau)+h_2(1-\kappa) < 1$, for $\tau \in [0, \frac{1}{2})$ and $\kappa \in (\frac{1}{2}, 1]$. A linear SSS can be obtained from the pair of nested codes $C \subset C'$ as follows. Let $\phi : \{0,1\}^\ell \to C'/C$ be a group homomorphism taking a secret $\mathbf{s}$ to its corresponding coset of C in C'. The sharing algorithm chooses a member in the coset $\phi(\mathbf{s})$ uniformly at random and outputs it as the share vector for $\mathbf{s}$. The reconstruction uses the decoder of C' to recover the full share vector from any $N\kappa$ shares and use the full share vector to identify the coset $\phi(\mathbf{s})$. The secret length $\ell = \log|C'/C|$ is approximately $N(1 - h_2(\tau) - h_2(1 - \kappa))$.

Perfect privacy imposes stringent conditions on (ε, δ)-SSS. This is seen by relating SSS to a closely related problem called *All-Or-Nothing Transform (AONT)*, introduced by Rivest [31] in computational setting, and later [15] extended to information theoretic setting. In a nutshell, an information theoretic AONT is an invertible randomized transformation $\mathsf{T}(\cdot)$ that takes an input $\mathbf{m}$ and outputs a pair (c_p, c_s), where c_p is public and c_s secret, and a deterministic function that recovers the message from the pair. The security guarantee of the transformation is that if t out of N bits of the secret part c_s of $\mathsf{T}(\mathbf{m})$ are known, the input remains indistinguishable. In the *secret only* case, only the secret output exists. The security of information-theoretic secret only AONT is defined in non-adaptive and adaptive models identical to the $(\frac{t}{N}, \varepsilon)$-privacy of SSS in Definition 5. And in the $\varepsilon = 0$ case, there is a stringent lower bound on the output length N in terms of the fraction $\tau = \frac{t}{N}$ of leakage and the input length, for secret only AONT.

Lemma 1 ([15])**.** If $T : \{0,1\}^\ell \to \{0,1\}^N$ is a secret only AONT with perfect privacy and a τ fraction of leakage, then

$$N\tau \leq \frac{N}{2} + \frac{N}{2(2^\ell - 1)} - 1.$$

For $\tau \geq \frac{1}{2}$. Lemma 1 says that

$$\frac{N}{2} \leq N\tau \leq \frac{N}{2} + \frac{N}{2(2^\ell - 1)} - 1 \Longrightarrow N \geq 2(2^\ell - 1).$$

An N player $(0, \delta)$-SSS with threshold pair (t, k) for $\frac{t}{N} \geq \frac{1}{2}$ is always a secret only AONT with perfect privacy for $\tau \geq \frac{1}{2}$ fraction of leakage. This means that under perfect privacy, the number of secret bits that can be shared by an N player $(0, \delta)$-SSS is less than $\log N$, even given all the N shares to reconstruct.

Coding rate of $(\varepsilon(N), \delta(N))$-SSS. In this work, we are concerned with binary SSS that shares a secret of length ℓ that is a constant fraction of N ($\ell = \Omega(N)$).

Definition 6. A coding rate $R \in [0,1]$ is achievable for the pair (τ, κ) if there exists a family of $(\varepsilon(N), \delta(N))$-SSS with relative threshold pair (τ, κ) such that $\varepsilon(N)$ and $\delta(N)$ are both negligible in N and $\frac{\ell(N)}{N} \to R$. The highest achievable coding rate for a pair (τ, κ) is called its capacity.

By relating (ε, δ)-SSS with relative threshold pair (τ, κ) to Wyner wiretap codes for a pair of binary erasure channels we obtain the following coding rate upper bound.

Lemma 2. The coding rate capacity of (ε, δ)-SSS with relative threshold pair (τ, κ) is asymptotically upper-bounded by $\kappa - \tau$.

The proof is given in Appendix A.

In the rest of the paper, we give two constant rate constructions of (ε, δ)-SSS against adaptive adversary and non-adaptive adversary, respectively. The non-adaptive adversary construction is optimal in the sense that the coding rate achieves the upper bound in Lemma 2.

4 Adaptive Adversary

We now introduce the building blocks before stating the construction theorem.

It is explicit in the definition of randomness extractors that the forward direction of extracting is efficient. In some applications, we usually need to efficiently invert the process and sample a random pre-image for a given extractor output.

Definition 7 ([8])**.** Let f be a mapping from $\{0,1\}^n$ to $\{0,1\}^m$. For $v \geq 0$, a function $Inv : \{0,1\}^m \times \{0,1\}^r \to \{0,1\}^n$ is called a v-inverter for f if the following conditions hold:

- (Inversion) Given $y \in \{0,1\}^m$ such that its pre-image $f^{-1}(y)$ is nonempty, for every $z \in \{0,1\}^r$ we have $f(Inv(y,z)) = x$.
- (Uniformity) $Inv(U_m, U_r)$ is v-close to U_n.

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

The following *almost perfect* notion was defined on *resilient functions* (bit-fixing extractors) for the construction of adaptive AONT [14] [1]. Here we generalise it to affine extractors. Almost perfect property can be trivially achieved by requiring an exponentially (in m) small error in statistical distance, using the relation between ℓ_∞-norm and ℓ_1-norm.

Definition 8. An affine extractor $\mathsf{AExt} : \{0,1\}^n \to \{0,1\}^m$ is called (k, ε)-almost perfect if for any affine (n,k)-source X,

$$\left| \mathsf{Pr}[\mathsf{AExt}(\mathsf{X}) = \mathsf{y}] - \frac{1}{2^m} \right| \leq 2^{-m} \cdot \varepsilon, \text{ for any } \mathsf{y} \in \{0,1\}^m.$$

Affine source plays an important role in both of our constructions. We define a general requirement for the stochastic code used in the outer layer that facilitates an affine structure.

[1] The relaxation from (perfect) resilient functions to ϵ-almost resilient functions with a small error ϵ was first studied in [23]

Definition 9 (Stochastic Affine codes)**.** Let $\mathsf{Enc} : \{0,1\}^m \times \mathcal{R} \to \{0,1\}^n$ be the encoder of a stochastic code. We say it is a stochastic affine code if for any $\mathsf{r} \in \mathcal{R}$, the encoding function $\mathsf{Enc}(\cdot, \mathsf{r})$ specified by r is an affine function of the message. That is we have

$$\mathsf{Enc}(\mathbf{m}, \mathsf{r}) = \mathbf{m}G_{\mathsf{r}} + \Delta_{\mathsf{r}},$$

where $G_{\mathsf{r}} \in \{0,1\}^{m\times n}$ and $\Delta_{\mathsf{r}} \in \{0,1\}^n$ are determined by the code and the randomness r.

We are ready to give our construction for adaptive adversary.

Theorem 3. Let $\mathsf{AExt} : \{0,1\}^n \to \{0,1\}^\ell$ be an invertible $(n-t, \frac{\varepsilon}{2})$-almost perfect affine extractor and $\mathsf{AExt}^{-1} : \{0,1\}^\ell \times \mathcal{R}_1 \to \{0,1\}^n$ be its inverter that maps an $\mathbf{s} \in \{0,1\}^\ell$ to one of its pre-images chosen uniformly at random. Let $(\mathsf{SA\text{-}ECCenc}, \mathsf{SA\text{-}ECCdec})$ be a stochastic affine-erasure correcting code with encoder $\mathsf{SA\text{-}ECCenc} : \{0,1\}^n \times \mathcal{R}_2 \to \{0,1\}^N$ that tolerates $N-k$ erasures and decodes with success probability at least $1-\delta$. Then the $(\mathsf{Share}, \mathsf{Recst})$ defined as follows is an adaptive (ε, δ)-SSS with threshold pair (t, k).

$$\begin{cases} \mathsf{Share}(\mathbf{s}) & = \mathsf{SA\text{-}ECCenc}(\mathsf{AExt}^{-1}(\mathbf{s})); \\ \mathsf{Recst}(\tilde{\mathsf{v}}) & = \mathsf{AExt}(\mathsf{SA\text{-}ECCdec}(\tilde{\mathsf{v}})), \end{cases}$$

where $\tilde{\mathsf{v}}$ denotes an incomplete version of a share vector $\mathsf{v} \in \{0,1\}^N$ with some of its components replaced by erasure symbols.

Proof. The (k, δ)-reconstructability follows directly from the erasure correcting capability of the $\mathsf{SA\text{-}ECC}$. For any $\tilde{\mathsf{v}}$ with at most $N-k$ erasure symbols and the rest of its components consistent with a valid codeword $\mathsf{v} \in \{0,1\}^N$, the $\mathsf{SA\text{-}ECC}$ decoder identifies the unique codeword v with probability $1-\delta$ over the encoder randomness. The corresponding $\mathsf{SA\text{-}ECC}$ message of v is then inputted to AExt and the original secret $\mathbf{s}$ is reconstructed with the same probability.

We next prove the (t, ε)-privacy. For any $\mathsf{r} \in \mathcal{R}_2$, the affine encoder of $\mathsf{SA\text{-}ECC}$ is characterised by a matrix $G_{\mathsf{r}} \in \{0,1\}^{n\times N}$ and an offset Δ_{r}. For n unknowns $\mathbf{x} = (x_1, \cdots, x_n)$, we have

$$\mathsf{SA\text{-}ECCenc}(\mathbf{x}) = \mathbf{x}G_{\mathsf{r}} + \Delta_{\mathsf{r}} = (\mathbf{x}G_1, \cdots, \mathbf{x}G_N) + \Delta_{\mathsf{r}},$$

where $G_i = (g_{1,i}, \cdots, g_{n,i})^T$ (here the subscript "$_{\mathsf{r}}$" is omitted to avoid double subscripts) denotes the ith column of G_{r}, $i = 1, \cdots, N$. This means that knowing a component c_i of the $\mathsf{SA\text{-}ECC}$ codeword is equivalent to obtaining a linear equation $c_i \oplus \Delta_i = \mathbf{x}G_i = g_{1,i}x_1 + \cdots + g_{n,i}x_n$ about the n unknowns $x_1, \cdots, x_n$, where Δ_i (again, the subscript "$_{\mathsf{r}}$" is omitted) denotes the ith component of Δ_{r}.

Now take any distinguisher $\mathcal{A}$, any secret $\mathbf{s} \in \{0,1\}^\ell$, and any possible view $\mathsf{View}_{\mathcal{A}}^{\mathsf{Share}(\mathbf{s})}$ of $\mathcal{A}$ having oracle access to $\mathsf{Share}(\mathbf{s})$ through adaptively obtaining t components. We assume that $\mathcal{A}$ is deterministic and argue security holds for any deterministic distinguisher.

Now, we find $\Pr[\mathsf{View}_{\mathcal{A}}^{\mathsf{Share}(\mathbf{s})} = \mathsf{w}]$ for arbitrary w (assuming the probability is non-zero). Note that since $\mathsf{AExt}^{-1}(\mathbf{s})$ is uniformly selected from the set of pre-images of $\mathbf{s}$, we have

$$\begin{aligned}
\mathsf{Pr}[\mathsf{View}_{\mathcal{A}}^{\mathsf{Share}(\mathbf{s})} = \mathsf{w}] &= \mathsf{Pr}[\mathsf{View}_{\mathcal{A}}^{\mathsf{ECCenc}(\mathsf{X})} = \mathsf{w} | \mathsf{AExt}(\mathsf{X}) = \mathbf{s}] \\
&= \frac{\mathsf{Pr}[\mathsf{AExt}(\mathsf{X}) = \mathbf{s} | \mathsf{View}_{\mathcal{A}}^{\mathsf{ECCenc}(\mathsf{X})} = \mathsf{w}] \cdot \mathsf{Pr}[\mathsf{View}_{\mathcal{A}}^{\mathsf{ECCenc}(\mathsf{X})} = \mathsf{w}]}{\mathsf{Pr}[\mathsf{AExt}(\mathsf{X}) = \mathbf{s}]} \\
&\overset{(i)}{=} \frac{(1 \pm \frac{\varepsilon}{2})2^{-\ell} \cdot \mathsf{Pr}[\mathsf{View}_{\mathcal{A}}^{\mathsf{ECCenc}(\mathsf{X})} = \mathsf{w}]}{\mathsf{Pr}[\mathsf{AExt}(\mathsf{X}) = \mathbf{s}]} \\
&\overset{(ii)}{=} \frac{(1 \pm \frac{\varepsilon}{2})2^{-\ell} \cdot \frac{2^{n-\mathrm{rank}(\mathcal{A})}}{2^n}}{2^{-\ell}} \\
&= (1 \pm \frac{\varepsilon}{2}) \cdot 2^{-\mathrm{rank}(\mathcal{A})}.
\end{aligned}$$

In above, we first use the fact that $\mathsf{Pr}[\mathsf{View}_{\mathcal{A}}^{\mathsf{Share}(\mathbf{s})} = \mathsf{w}]$ can be seen as the probability of randomly selecting X from $\{0,1\}^n$, with the condition that $\mathsf{AExt}(\mathsf{X}) = \mathbf{s}$. This is true because the sets $\mathsf{AExt}^{-1}(\mathbf{s})$ for all $\mathbf{s}$, partition $\{0,1\}^n$. The shorthand "$y = 1 \pm \frac{\varepsilon}{2}$" denotes "$1 - \frac{\varepsilon}{2} \leq y \leq 1 + \frac{\varepsilon}{2}$". The $\mathrm{rank}(\mathcal{A})$ denotes the rank of $\mathcal{A}$'s choice of up to t columns of G. The equality (i) follows from the fact that AExt is an almost perfect affine extractor and the uniform X conditioned on at most t linear equations is an affine source with at least $n - t$ bits entropy. The equality (ii) follows from the assumption that X is chosen uniformly at random from $\{0,1\}^n$. The observation $\mathsf{View}_{\mathcal{A}}^{\mathsf{Share}(\mathbf{s})}$ of $\mathcal{A}$ with respect to a secret $\mathbf{s}$ has a distribution only depends on $\mathrm{rank}(\mathcal{A})$ (independent of the secret $\mathbf{s}$) and and the error for each w is bounded by $\frac{\varepsilon}{2} \cdot 2^{-\mathrm{rank}(\mathcal{A})}$. We have for any distinguisher $\mathcal{A}$,

$$\left| \mathsf{Pr}[\mathcal{A}^{\mathsf{Share}(\mathbf{s}_0)}(\mathbf{s}_0, \mathbf{s}_1) = 1] - \mathsf{Pr}[\mathcal{A}^{\mathsf{Share}(\mathbf{s}_1)}(\mathbf{s}_0, \mathbf{s}_1) = 1] \right| \leq \varepsilon. \quad \square$$

Instantiations of the construction.

There are explicit constructions of binary affine extractors that, given a constant fraction of entropy, outputs a constant fraction of random bits with exponentially small error.

Lemma 4 ([4, 24])**.** For every constant $0 < \mu \leq 1$, there is an explicit affine extractor $\mathsf{AExt} : \{0,1\}^n \to \{0,1\}^m$ for sources with min-entropy $n\mu$ with output length $m = \Omega(n)$ and error at most $2^{-\Omega(n)}$.

There are known methods for constructing an invertible affine extractor AExt' from any affine extractor AExt such that the constant fraction output size and exponentially small error properties are preserved. A simple method is to let $\mathsf{AExt}'(U_n||M) := \mathsf{AExt}(U_n) \oplus M$ (see Appendix B for a discussion).

We discuss two constructions of $\mathsf{SA\text{-}ECC}$, one non-explicit the other explicit. The non-explicit construction uses an optimal rate linear erasure list-decodable code (existence proved in [18]) and an Algebraic Manipulation Detection (AMD) code (e.g. constructed in [12]). List-decodable codes correct worst-case erasures by outputting a list of candidate codewords that include the correct codeword. The AMD pre-coding is then used to identify the correct codeword in the list, hence restoring the correct message. Since the erasure list-decodable code [18] is linear, the combination is a stochastic affine code (see Appendix C). This construction requires a linear erasure list-decodable code that corrects up to p fraction of Erasures, and achieves coding rate $1-p$. However construction of such codes is an open problem.

We next describe another rate optimal construction of SA-ECC that is also explicit and results in an SSS with vanishing relative gap, as N grows. The construction uses the same approach as the construction of stochastic code for additive oblivious [2] error, in [19, Theorem 6.1]. The main idea of the construction is to convert the adversarial erasure channel to a random erasure channel by permuting the codeword components, using a randomly selected permutation, whose description is sent to the receiver along with the codeword. The description of the permutation must be robustly (against erasure) "transmitted" to the decoder. This is done by first using a highly robust and low rate code in big alphabet, and then encoding each component of the code (alphabet symbol) using a stochastic erasure correcting code with detection property that detects a random translation. The stochastic erasure correcting code with detection property can be the above combination of a linear erasure list-decodable code and an AMD. The rate of the list-decodable code is not required to be optimal here because the *control information* is negligible compared to the so called *payload* that carries the encoded message. The control information blocks are randomly mixed into the payload blocks to form the final codeword of SA-ECC. There is one more building block that prevents (using the detection property of the stochastic erasure correcting code) the payload blocks from being mistaken for the control information blocks during the first step of decoding in which the control information must be correctly recovered. We outline this construction (see Appendix D) as part of the proof for the following theorem.

Theorem 5. For every $p \in [0, 1)$, and every $\varepsilon > 0$, there is an efficiently encodable and decodable stochastic affine code $(\mathsf{Enc}, \mathsf{Dec})$ with rate $R_{ECC} = 1-p-\varepsilon$ such that for every $\mathbf{m} \in \{0,1\}^{NR_{ECC}}$ and erasure pattern of at most p fraction, we have $\Pr[\mathsf{Dec}(\widetilde{\mathsf{Enc}(\mathbf{m})}) = \mathbf{m}] \geq 1 - \exp(-\Omega(\varepsilon^2 N/\log^2 N))$, where $\widetilde{\mathsf{Enc}(\mathbf{m})}$ denotes the partially erased random codeword and N denotes the length of the codeword.

The proof of Theorem 3 says that as long as AExt is an affine extractor with the right parameters, the SSS provides privacy. Let R_{ECC} denote the rate of the SA-ECC. Using the notations in the proof, there are $n - \tau N = n(1 - \frac{\tau}{R_{ECC}})$ bits of entropy in X, where $\tau = \frac{t}{N}$. Let us assume $\tau < R_{ECC}$. Using the AExt from Lemma 4 (more precisely, an invertible affine extractor $\mathsf{AExt}' : \{0,1\}^{n'} \to \{0,1\}^{\ell}$ constructed from AExt) with $\mu = 1 - \frac{\tau}{R_{ECC}}$, a constant fraction $\Omega(n)$ of random bits can be extracted with exponentially small error. This says that (τ, ε)-privacy is guaranteed for $\tau \in [0, R_{ECC})$. We then want to have R_{ECC} as big as possible. The stochastic affine ECC in Theorem 5 asymptotically achieves the rate $1 - (1 - \kappa) = \kappa$. We then have the following corollary.

Corollary 6. There is an explicit constant coding rate adaptive (ε, δ)-SSS with relative threshold pair (τ, κ) for any $0 \leq \tau < \kappa \leq 1$.

The construction above achieves a constant coding rate for any (τ, κ) pair satisfying $0 \leq \tau < \kappa \leq 1$. Since the binary affine extractor in Lemma 4 does not extract all the entropy from the source (meaning the output length is equal to the source entropy minus a negligible amount) and moreover the step that transforms an affine extractor into an invertible affine extractor incurs non-negligible overhead, the coding rate of the above construction is strictly smaller than the upper bound $\kappa - \tau$. We give another construction in Section 5 that uses a linear seeded extractor instead of an affine extractor, which achieves the coding rate $\kappa - \tau$. This construction, however, only achieves non-adaptive privacy.

[2] An offset of Hamming weight at most Np is chosen obliviously and added to the N-bit codeword.

5 Non-adaptive Adversary

The following construction achieves the upper bound in Lemma 2 and is hence capacity-achieving. The main reason is the use of a linear seeded extractor, which extracts all the randomness from the source. The linearity makes the functions invertible without incurring overhead. Moreover, the linearity of the extractor functions together with the affine structure of the $\mathsf{SA\text{-}ECC}$ has a special effect on the extractor error: either perfect extraction or big error (also observed and exploited in [11, Theorem 1] to achieve perfect privacy in their construction). This directly leads to the desired privacy error without requiring an exponentially small extractor error, which would have incurred non-negligible loss in coding rate. The drawback of using a seeded extractor is that we can only prove non-adaptive privacy for the construction (see Remark 2).

Theorem 7. Let $\mathsf{Ext}:\{0,1\}^d\times\{0,1\}^n\rightarrow\{0,1\}^\ell$ be a linear strong seeded $(n-t,\frac{\varepsilon}{4})$-extractor and $\mathsf{Ext}^{-1}(\mathsf{z},\cdot):\{0,1\}^\ell\times\mathcal{R}_1\rightarrow\{0,1\}^n$ be the inverter of the function $\mathsf{Ext}(\mathsf{z},\cdot)$ that maps an $\mathsf{s}\in\{0,1\}^\ell$ to one of its pre-images chosen uniformly at random. Let $(\mathsf{SA\text{-}ECCenc},\mathsf{SA\text{-}ECCdec})$ be a stochastic affine-erasure correcting code with encoder $\mathsf{SA\text{-}ECCenc}:\{0,1\}^{d+n}\times\mathcal{R}_2\rightarrow\{0,1\}^N$ that tolerates $N-k$ erasures and decodes with success probability at least $1-\delta$. Then the following coding scheme $(\mathsf{Share},\mathsf{Recst})$ is a non-adaptive (ε,δ)-SSS with threshold pair (t,k).

$$\begin{cases}\mathsf{Share}(\mathbf{s}) & =\mathsf{SA\text{-}ECCenc}(\mathsf{Z}||\mathsf{Ext}^{-1}(\mathsf{Z},\mathbf{s})),\text{where }\mathsf{Z}\overset{\$}{\leftarrow}\{0,1\}^d;\\ \mathsf{Recst}(\tilde{\mathsf{v}}) & =\mathsf{Ext}(\bar{\mathsf{z}},\bar{\mathsf{x}}),\text{where }(\bar{\mathsf{z}}||\bar{\mathsf{x}})=\mathsf{SA\text{-}ECCdec}(\tilde{\mathsf{v}}).\end{cases}$$

Here $\tilde{\mathsf{v}}$ denotes an incomplete version of a share vector $\mathsf{v}\in\{0,1\}^N$ with some of its components replaced by erasure symbols.

Here we give a direct proof using techniques similar to that of Theorem 3. In Appendix E, we prove a general property of linear seeded extractors, from which Theorem 7 follows as a corollary.

Proof. The algorithm $\mathsf{Share}(\cdot)$ has three parts of randomness: the uniform seed $\mathsf{Z}\overset{\$}{\leftarrow}\{0,1\}^d$, the randomness of the inverter sampled from $\mathcal{R}_1$ and the randomness of the stochastic code sampled from $\mathcal{R}_2$. In our analysis, we always consider a fixed randomness $\mathsf{r}\in\mathcal{R}_2$ of the stochastic code. We will first consider a fixed seed $\mathsf{z}\in\{0,1\}^d$ and solely use the randomness of the inverter in the argument for privacy of the scheme. We show that for most of the seed $\mathsf{z}\in\{0,1\}^d$, the scheme is perfectly secure and hence when the seed is uniformly chosen, the privacy error can be bounded by the fraction of "bad" seeds.

We next prove the (t,ε)-privacy. For any $\mathsf{r}\in\mathcal{R}_2$, the affine encoder of $\mathsf{SA\text{-}ECC}$ is characterised by a matrix $G_\mathsf{r}\in\{0,1\}^{n\times N}$ and an offset Δ_r. Here the message of the $\mathsf{SA\text{-}ECC}$ is a $(d+n)$-bit string, where the first d bits constitute the seed of Ext. Our analysis is first done for each particular seed and then average over the seed space. So we assume the first d bits of the message of $\mathsf{SA\text{-}ECC}$ are fixed values z and only consider the remaining n bits $\mathbf{x}=(x_1,\cdots,x_n)$ as unknowns. Then for any $\mathsf{r}\in\mathcal{R}_2$, we have

$$\mathsf{SA\text{-}ECCenc}(\mathsf{z}||\mathbf{x})=(\mathsf{z}||\mathbf{x})G_\mathsf{r}+\Delta_\mathsf{r}=(\mathbf{x}G'_1,\cdots,\mathbf{x}G'_N)+\hat{\Delta}(\mathsf{z})+\Delta_\mathsf{r},$$

where $G'_i=(g_{d+1,i},\cdots,g_{d+n,i})^T$, $i=1,\cdots,N$ denotes the ith column of G'_r, which is obtained from G_r by removing the top d rows, and $\hat{\Delta}(\mathsf{z})\in\{0,1\}^N$ denotes the offset obtained from multiplying z to the top d rows of G_r. This means that knowing a component c_i of the $\mathsf{SA\text{-}ECC}$ codeword is equivalent to obtaining a linear equation on the n unknowns $x_1,\cdots,x_n$:

$$c_i\oplus\Delta_i\oplus\hat{\Delta}(\mathsf{z})_i=\mathbf{x}G'_i=g_{d+1,i}x_1+\cdots+g_{d+n,i}x_n. \tag{4}$$

Now, we consider an arbitrary secret $\mathbf{s}$ and for any non-adaptive $\mathcal{A}$'s choice of $A \subset [N]$ with $|A| \leq t$, we find $\mathsf{Pr}[\mathsf{View}_A^{\mathsf{Share}(\mathbf{s})} = \mathsf{w}]$ for each $\mathsf{w} \in \{0,1\}^{|A|}$. Since we have the following decomposition with respect to the uniform seed Z,

$$\mathsf{Pr}[\mathsf{View}_A^{\mathsf{Share}(\mathbf{s})} = \mathsf{w}] = \sum_{\mathsf{z} \in \{0,1\}^d} 2^{-d} \cdot \mathsf{Pr}[\mathsf{View}_A^{\mathsf{Share}(\mathbf{s})} = \mathsf{w} | \mathsf{Z} = \mathsf{z}], \tag{5}$$

we proceed with considering a fixed seed $\mathsf{Z} = \mathsf{z}$.

Note that the inverter function $\mathsf{Ext}^{-1}(\mathsf{z}, \cdot)$ defines a partition of $\{0,1\}^n$ into 2^ℓ subsets and maps a secret to an n-tuple in the corresponding subset uniformly at random. Let X be the uniform distribution over $\{0,1\}^n$. We have $\mathsf{SA\text{-}ECCenc}(\mathsf{z}||\mathsf{Ext}^{-1}(\mathsf{z}, \mathbf{s})) = (\mathsf{SA\text{-}ECCenc}(\mathsf{z}||\mathsf{X})) \,|\, (\mathsf{Ext}(\mathsf{z}, \mathsf{X}) = \mathbf{s})$. We then use the Bayes law and compute the following.

$$\begin{aligned}
\mathsf{Pr}\left[\mathsf{View}_A^{\mathsf{Share}(\mathbf{s})} = \mathsf{w} | \mathsf{Z} = \mathsf{z}\right] &= \mathsf{Pr}[\mathsf{View}_A^{\mathsf{SA\text{-}ECCenc}(\mathsf{z}||\mathsf{X})} = \mathsf{w} | \mathsf{Ext}(\mathsf{z}, \mathsf{X}) = \mathbf{s}] \\
&= \frac{\mathsf{Pr}[\mathsf{Ext}(\mathsf{z}, \mathsf{X}) = \mathbf{s} | \mathsf{View}_A^{\mathsf{SA\text{-}ECCenc}(\mathsf{z}||\mathsf{X})} = \mathsf{w}] \cdot \mathsf{Pr}[\mathsf{View}_A^{\mathsf{SA\text{-}ECCenc}(\mathsf{z}||\mathsf{X})} = \mathsf{w}]}{\mathsf{Pr}[\mathsf{Ext}(\mathsf{z}, \mathsf{X}) = \mathbf{s}]} \\
&\overset{(i)}{=} \frac{2^{-\ell} \cdot \mathsf{Pr}[\mathsf{View}_A^{\mathsf{SA\text{-}ECCenc}(\mathsf{z}||\mathsf{X})} = \mathsf{w}]}{\mathsf{Pr}[\mathsf{Ext}(\mathsf{z}, \mathsf{X}) = \mathbf{s}]} \\
&\overset{(ii)}{=} \frac{2^{-\ell} \cdot \frac{2^{n-\mathrm{rank}(A)}}{2^n}}{2^{-\ell}} \\
&= 2^{-\mathrm{rank}(A)}.
\end{aligned}$$

The above computation requires two conditions. The equality (i) holds when the linear function $\mathsf{Ext}(\mathsf{z}, \cdot) : \{0,1\}^n \to \{0,1\}^\ell$ restricted to support of the random variable $\mathsf{X}| \left(\mathsf{View}_A^{\mathsf{SA\text{-}ECCenc}(\mathsf{z}||\mathsf{X})} = \mathsf{w}\right)$ is surjective. In this case, we say the seed $\mathsf{z} \in \{0,1\}^d$ is a *good* seed with respect to w. The equality (ii) holds when there exists an $\mathsf{x} \in \{0,1\}^n$ such that $\mathsf{SA\text{-}ECCenc}(\mathsf{z}||\mathsf{x})_A = \mathsf{w}$. The short hand $\mathrm{rank}(A)$ denotes the rank of the columns G'_i that satisfy $i \in A$ (see (4) for definition of G'_i). Consider X as unknowns for equations, the number of solutions to the linear system $\mathsf{SA\text{-}ECCenc}(\mathsf{z}||\mathsf{X})_A = \mathsf{w}$ is either 0 or equal to $2^{n-\mathrm{rank}(A)}$.

Note that whether $\mathsf{z} \in \{0,1\}^d$ is a good seed is in fact determined by z and the adversary's choice of A. Let $\mathsf{Ker}_A = \left\{\mathsf{x} \in \{0,1\}^n | \mathsf{SA\text{-}ECCenc}(\mathsf{z}||\mathsf{x})_A = 0^{|A|}\right\}$. Then the solutions to the linear system $\mathsf{SA\text{-}ECCenc}(\mathsf{z}||\mathsf{X})_A = \mathsf{w}$ can be written as a translate of the subspace Ker_A. We see that if the linear function $\mathsf{Ext}(\mathsf{z}, \cdot) : \{0,1\}^n \to \{0,1\}^\ell$ restricted to Ker_A is surjective, it is also surjective restricted to any translate of Ker_A. Let $\mathcal{G}_A \subset \{0,1\}^d$ denote the set of good seeds with respect to A. We next show that $\mathsf{Pr}[\mathsf{Z} \in \mathcal{G}_A] \geq 1 - \frac{\varepsilon}{2}$. Since there are at least t linear equations on n unknowns, Ker_A is a subspace of $\{0,1\}^n$ of dimension at least $n - t$. The flat distribution over Ker_A is an affine source of min-entropy at least $n - t$. Now applying a linear function $\mathsf{Ext}(\mathsf{z}, \cdot)$ to this affine source, the output is uniform ℓ bits if and only if $\mathsf{Ext}(\mathsf{z}, \mathsf{Ker}_A) = \{0,1\}^\ell$. On the other hand, if $\mathsf{Ext}(\mathsf{z}, \mathsf{Ker}_A) \neq \{0,1\}^\ell$, then the distribution of the output has at least $\frac{1}{2}$ statistical distance from uniform. This is because the image $\mathsf{Ext}(\mathsf{z}, \mathsf{Ker}_A)$ is a subspace of dimension at most $\ell - 1$. The security of the $(n, n-t)$-extractor Ext asserts that

$$0 \cdot \mathsf{Pr}[\mathsf{Z} \in \mathcal{G}_A] + \frac{1}{2} \cdot \mathsf{Pr}[\mathsf{Z} \notin \mathcal{G}_A] \leq \frac{1}{4}\varepsilon. \tag{6}$$

This yields the desired bound $\mathsf{Pr}[\mathsf{Z} \in \mathcal{G}_A] \geq 1 - \frac{\varepsilon}{2}$.

Finally, we have shown that the distribution $\mathsf{View}_A^{\mathsf{Share(s)}}|\,(\mathsf{Z}=\mathsf{z})$ is independent of the secret $\mathbf{s}$ when $\mathsf{z}\in\mathcal{G}_A$. According to (5), the distribution $\mathsf{View}_A^{\mathsf{Share(s)}}$ is independent of the secret $\mathbf{s}$ conditioned on the event that $\mathsf{Z}\in\mathcal{G}_A$. It then follows that for any two secrets $\mathbf{s}_0$ and $\mathbf{s}_1$, we always have $\mathsf{SD}(\mathsf{Share}(\mathbf{s}_0)_A;\mathsf{Share}(\mathbf{s}_1)_A)\leq\varepsilon$.

□

Remark 2. Note that the same argument cannot be made if the set A is chosen according to the seed z, in which case we can not define $\mathsf{W}=\mathsf{Share}(\mathbf{S})_A$ for all seeds of Ext and no longer have (6). In a real life adaptive attack, the adversary can first spend some reading budget on figuring out the value of the seed Z of Ext, and then decide the rest of the reading positions according to the seed value.

Instantiations of the construction.

We will use Trevisan's extractor [35] for the Ext in Theorem 7. In particular, we use the following improvement of this extractor due to Raz, Reingold and Vadhan [30].

Lemma 8 ([30])**.** There is an explicit strong linear seeded (k,ε)-extractor $\mathsf{Ext}:\{0,1\}^d\times\{0,1\}^n\to\{0,1\}^\ell$ with $d=O(\log^3(n/\varepsilon))$ and $\ell=k-O(d)$.

We now analyse the coding rate of the (ε,δ)-SSS with $(\frac{t}{N},\frac{k}{N})$-threshold constructed in Theorem 7 with the $\mathsf{SA\text{-}ECC}$ from Theorem 5 and the Ext from Lemma 8. The secret length $\ell=n-t-O(d)$, where the seed length is $d=O(\log^3(2n/\varepsilon))$. The $\mathsf{SA\text{-}ECC}$ is from $n+d$ bits to N bits and with coding rate $R_{ECC}=\kappa-\xi$ for a small ξ determined by δ (they satisfy the relation $\delta=\exp(-\Omega(\xi^2N/\log^2N))$ according to Theorem 5). We then have $n=N(\kappa-\xi)-d$. Finally, the coding rate is

$$R=\frac{\ell}{N}=\frac{n-t-O(d)}{N}=\frac{N(\kappa-\xi)-t-O(d)}{N}=\kappa-\tau-(\xi+\frac{O(d)}{N}).$$

Since the seed length d is negligible in N for any privacy error ε, and the rate deficiency ξ of the $\mathsf{SA\text{-}ECC}$ can be chosen arbitrarily small at the expense of a bigger N, we then conclude that the rate $\kappa-\tau$ is achieved.

Corollary 9. There is an explicit construction of non-adaptive (ε,δ)-SSS with relative threshold pair (τ,κ)achieving coding rate $\kappa-\tau$.

6 Conclusion

Secret sharing is a fundamental cryptographic primitive with diverse applications. We considered (ε,δ) ramp schemes which allows bounded privacy and reconstructability errors (ε, and δ, respectively), with the goal of achieving short shares for long secrets. We gave two efficient modular constructions with security against adaptive and non-adaptive adversaries, and showed an (asymptotically) optimal instantiation of non-adaptive case. The constructions are non-linear, and although are stated in terms of binary shares, can be extended to q-ary shares also. Interesting open questions that arise from this work include, (i) tight upper bound for adaptive security, or constructions that achieve the bound for non-adaptive case, and (ii) construction of linear (ε,δ) ramp schemes with similar properties.

References

[1] Vaneet Aggarwal, Lifeng Lai, A. Robert Calderbank, and H. Vincent Poor. Wiretap channel type II with an active eavesdropper. In *IEEE International Symposium on Information Theory, ISIT 2009, June 28 - July 3, 2009, Seoul, Korea, Proceedings*, pages 1944–1948. IEEE, 2009.

[2] Mihir Bellare, Stefano Tessaro, and Alexander Vardy. Semantic security for the wiretap channel. In Reihaneh Safavi-Naini and Ran Canetti, editors, *Advances in Cryptology - CRYPTO 2012 - 32nd Annual Cryptology Conference, Santa Barbara, CA, USA, August 19-23, 2012. Proceedings*, volume 7417 of *Lecture Notes in Computer Science*, pages 294–311. Springer, 2012.

[3] G. R. BLAKLEY. Safeguarding cryptographic keys. *Proceedings of the 1979 AFIPS National Computer Conference, pp. 313–317, 1979.*

[4] J. Bourgain. On the construction of affine extractors. *Geometric and Functional Analysis, vol. 17, no. 1, p. 33-57, 2007.*

[5] Ignacio Cascudo, Hao Chen, Ronald Cramer, and Chaoping Xing. Asymptotically good ideal linear secret sharing with strong multiplication over any fixed finite field. *In Advances in Cryptology - CRYPTO 2009, 29th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 16- 20, 2009. Proceedings, pages 466–486, 2009.*

[6] H. Chen and R. Cramer. Algebraic geometric secret sharing schemes and secure multiparty computation over small fields. *In: Dwork, C. (ed.) CRYPTO 2006. LNCS, vol. 4117, pp. 516–531. Springer, Heidelberg (2006).*

[7] Hao Chen, Ronald Cramer, Shafi Goldwasser, Robbert de Haan, and Vinod Vaikuntanathan. Secure computation from random error correcting codes. *In: Naor M. (eds) Advances in Cryptology - EUROCRYPT 2007. EUROCRYPT 2007. Lecture Notes in Computer Science, vol 4515. Springer, Berlin, Heidelberg.*

[8] Mahdi Cheraghchi, Fredric Didier, and Amin Shokrollahi. Invertible extractors and wiretap protocols. *Information Theory, IEEE Transactions on 58.2 (2012): 1254-1274.*

[9] R. Cramer, I. Damgaard, and S. Dziembowski. On the complexity of verifiable secret sharing and multi-party computation. *Proceedings of STOC 2000, pp. 325– 334. ACM Press, 2000.*

[10] R. Cramer, I. Damgaard, and U. Maurer. General secure multi-party computation from any linear secret sharing scheme. *In: Preneel, B. (ed.) EUROCRYPT 2000. LNCS, vol. 1807, pp. 316–334. Springer, Heidelberg (2000).*

[11] Ronald Cramer, Ivan Bjerre Damgård, Nico Döttling, Serge Fehr, and Gabriele Spini. Linear secret sharing schemes from error correcting codes and universal hash functions. In *Advances in Cryptology-EUROCRYPT 2015*, pages 313–336. Springer, 2015.

[12] Ronald Cramer, Yevgeniy Dodis, Serge Fehr, Carles Padró, and Daniel Wichs. Detection of algebraic manipulation with applications to robust secret sharing and fuzzy extractors. In Nigel P. Smart, editor, *Advances in Cryptology - EUROCRYPT 2008, 27th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Istanbul, Turkey, April 13-17, 2008. Proceedings*, volume 4965 of *Lecture Notes in Computer Science*, pages 471–488. Springer, 2008.

[13] Imre Csiszár and Janos Körner. Broadcast channels with confidential messages. *IEEE Transactions on Information Theory*, 24(3):339–348, May 1978.

[14] Yevgeniy Dodis, Amit Sahai, and Adam Smith. On perfect and adaptive security in exposure-resilient cryptography. *In B. Pfitzmann (Ed.) EUROCRYPT 2001, LNCS 2045, pp. 301?324, 2001. Springer-Verlag Berlin Heidelberg.*

[15] Yevgeniy Dodis, Amit Sahai, and Adam Smith. On perfect and adaptive security in exposure-resilient cryptography. *B. Pfitzmann (Ed.) EUROCRYPT 2001, LNCS 2045, Springer-Verlag Berlin Heidelberg.*, pages 301–324, 2001.

[16] O. Goldreich, S. Micali, and A. Wigderson. How to play any mental game or a completeness theorem for protocols with honest majority,. *Proc. of ACM STOC '87, pp. 218–229.*

[17] Blakley G.R. and Meadows C. Security of ramp schemes. *Blakley G.R., Chaum D. (eds) Advances in Cryptology. CRYPTO 1984. Lecture Notes in Computer Science, vol . Springer, Berlin, Heidelberg*, 196, 1985.

[18] Venkatesan Guruswami. List decoding from erasures: bounds and code constructions. *IEEE Transactions on Information Theory (Volume: 49, Issue: 11, Nov. 2003).*

[19] Venkatesan Guruswami and Adam Smith. Codes for computationally simple channels: Explicit constructions with optimal rate. In *51th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2010, October 23-26, 2010, Las Vegas, Nevada, USA*, pages 723–732. IEEE Computer Society, 2010.

[20] Chen H., Cramer R., de Haan R., and Pueyo I.C. Strongly multiplicative ramp schemes from high degree rational points on curves. *In: Smart N. (eds) Advances in Cryptology – EUROCRYPT 2008. EUROCRYPT 2008. Lecture Notes in Computer Science, vol 4965. Springer, Berlin, Heidelberg.*

[21] Cheng K., Ishai Y., and Li X. Near-optimal secret sharing and error correcting codes in ac0. *In: Kalai Y., Reyzin L. (eds) Theory of Cryptography. TCC 2017. Lecture Notes in Computer Science, vol 10678. Springer, Cham.*

[22] E.D Kamin, J.W. Green, and M.E. Hellman. On secret sharing systems. *IEEE Trans. Information Th., Vol. IT-29, No 1, pp. 35-41, Jan 1983.*

[23] K. Kurosawa, T. Johansson, and D. R. Stinson. Almost k-wise independent sample spaces and their cryptologic applications. *J. Cryptology, 14 (2001), pp. 231–253.*

[24] Xin Li. A new approach to affine extractors and dispersers. *in Proceedings of the 26th IEEE Conference on Computational Complexity (CCC), 2011.*

[25] J.L Massey. Some applications of coding theory in cryptography. *Codes and Ciphers: Cryptography and Coding IV. pp. 33 47 (1995).*

[26] U. M. Maurer and S. Wolf. Information-theoretic key agreement: From weak to strong secrecy for free. *in Advances in Cryptology - Eurocrypt 2000, Lecture Notes in Computer Science. B. Preneel, 2000, p. 351.*

[27] R. J. McEliece and D. V. Sarwate. On sharing secrets and reed-solomon codes. *Comm. ACM, Vol.24, pp 583-584, Sep 1981.*

[28] Rafail Ostrovsky and Moti Yung. How to withstand mobile virus attacks. *Proc. 10th ACM Conf. on Principles of Distributed Systems, August 1991.*

[29] L. H. Ozarow and A. D. Wyner. Wire-tap channel ii. *Bell System Technical Journal, vol. 63, pp. 2135-2157, Dec. 1984.*

[30] R. Raz, O. Reingold, and S. Vadhan. Extracting all the randomness and reducing the error in trevisan's extractor. *Journal of Computer and System Sciences, vol. 65, no. 1, p. 97–128, 2002.*

[31] R. Rivest. All-or-nothing encryption and the package transform. *Proceedings of the International Workshop on Fast Software Encryption, Lecture Notes in Computer Science*, 1267:210–218, 1997.

[32] R. Shaltiel. How to get more mileage from randomness extractors. *21st Annual IEEE Conference on Computational Complexity (CCC'06), Prague, 2006, pp. 46-60. doi: 10.1109/CCC.2006.24.*

[33] A. SHAMİR. How to share a secret. *Commun. ACM, 22, pp. 612–613, 1979.*

[34] A. Smith. Scrambling adversarial errors using few random bits, optimal information reconciliation, and better private codes. *In Symposium on Discrete Algorithms, pages 395–404, 2007.*

[35] L. Trevisan. Extractors and pseudorandom generators. *Journal of the ACM, vol. 48, no. 4, p. 860-879, 2001.*

[36] Ogata W., Kurosawa K., and Tsujii S. Nonperfect secret sharing schemes. *In: Seberry J., Zheng Y. (eds) Advances in Cryptology — AUSCRYPT '92. AUSCRYPT 1992. Lecture Notes in Computer Science, . Springer, Berlin, Heidelberg*, vol 718, 1993.

[37] Pengwei Wang and Reihaneh Safavi-Naini. A model for adversarial wiretap channel. *IEEE Transactions on Information Theory (Volume: 62, Issue: 2, Feb. 2016)*, pages 970 – 983.

[38] Pengwei Wang, Reihaneh Safavi-Naini, and Fuchun Lin. Erasure adversarial wiretap channels. *Communication, Control, and Computing (Allerton), 2015 53rd Annual Allerton Conference on.*

[39] A. D. Wyner. The wire-tap channel. *The Bell System Technical Journal (Volume: 54, Issue: 8, Oct. 1975).*

[40] Desmedt Y. Threshold cryptosystems. *In: Seberry J., Zheng Y. (eds) Advances in Cryptology — AUSCRYPT '92. AUSCRYPT 1992. Lecture Notes in Computer Science, vol 718. Springer, Berlin, Heidelberg.*

Appendices

A Proof of Lemma 2

Proof. Let $(\mathsf{Share}, \mathsf{Recst})$ be a non-adaptive (ε, δ)-SSS with relative threshold pair (τ, κ). We use Share as the encoder and Recst as the decoder, and verify in the following that we obtain a Wyner wiretap code for a BEC_{p_m} main channel and a BEC_{p_w} wiretapper channel, where $p_m = 1 - \kappa - \xi$

and $p_w = 1 - \tau + \xi$, respectively, for arbitrarily small $\xi > 0$. Erasure in SSS is worst case, while it is probabilistic in the Wyner wiretap model. We however note that asymptotically, the number of random erasures approaches Np_m and Np_w, respectively, with overwhelming probability, and so a code that protects against worst case erasure can be used as a wiretap code with probabilistic erasure. In our proof we also take into account the difference in the secrecy notion in SSS and in the case of Wyner wiretap code.

The N-bit output $\mathsf{Y} = Y_1, \cdots, Y_N$ of a BEC_p has a distribution where each bit is identically independently erased with probability p. By the Chernoff-Hoeffding bounds, the fraction η of erasures satisfies the following. For arbitrarily small $\xi > 0$,

$$\begin{cases} \mathsf{Pr}[\eta \geq p + \xi] & \leq \left(\left(\frac{p}{p+\xi}\right)^{p+\xi} \left(\frac{1-p}{1-p-\xi}\right)^{1-p-\xi} \right)^N ; \\ \mathsf{Pr}[\eta \leq p - \xi] & \leq \left(\left(\frac{p}{p-\xi}\right)^{p-\xi} \left(\frac{1-p}{1-p+\xi}\right)^{1-p+\xi} \right)^N . \end{cases}$$

Applying the two inequalities to BEC_{p_m} and BEC_{p_w}, respectively, we obtain the following conclusions. The probability that the main channel BEC_{p_m} has at most $p_m + \xi = 1 - \kappa$ fraction of erasures and the probability that the wiretapper channel BEC_{p_w} has at least $p_w - \xi = 1 - \tau$ fraction of erasures are both at most $\exp(-\Omega(N))$ for arbitrarily small $\xi > 0$.

We are ready to prove the Wyner wiretap reliability and secrecy properties as defined in [39, 13].

- We show correct decoding with probability $1 - o(1)$. When the erasures are below $p_m + \xi = 1 - \kappa$ fraction, it follows directly from the (κ, δ)-reconstructability of SSS that the decoding error is bounded from above by δ, which is arbitrarily small for big enough N, where the probability is over the randomness of the encoder. When the erasures are not below $p_m + \xi = 1 - \kappa$ fraction, we do not have correct decoding guarantee. But as argued above, this only occurs with a negligible probability over the randomness of the BEC_{p_m}. Averaging over the channel randomness of the BEC_{p_m}, we have correct decoding with probability $1 - o(1)$.

- We show random message equivocation secrecy $\mathsf{H}(\mathbf{S}|\mathsf{W}) \geq \ell(1 - o(1))$, where $\mathbf{S}$ is a uniform message and $\mathsf{W} = \mathsf{BEC}_{p_w}(\mathsf{Share}(\mathbf{S}))$ is the view from the wiretapper channel. We in fact first prove the wiretap indistinguishability security as defined in (1) and then deduce that it implies Wyner wiretap secrecy as defined in [39, 13]. For each of the erasure patterns (say $A \subset [N]$ are not erased) of the wiretapper channel BEC_{p_w} that exceeds $p_w - \xi = 1 - \tau$ fraction (equivalently, $|A| \leq N\tau$), the (τ, ε)-privacy gives that for any two messages, the corresponding wiretapper channel views $\mathsf{W}|(\mathbf{S} = \mathbf{s}_0, A \text{ not erased})$ and $\mathsf{W}|(\mathbf{S} = \mathbf{s}_1, A \text{ not erased})$ are indistinguishable with error ε, which is arbitrarily small for big enough N. The distribution $(\mathsf{W}|\mathbf{S} = \mathbf{s}_0)$ and $(\mathsf{W}|\mathbf{S} = \mathbf{s}_1)$ are convex combinations of $\mathsf{W}|(\mathbf{S} = \mathbf{s}_0, A \text{ not erased})$ and $\mathsf{W}|(\mathbf{S} = \mathbf{s}_1, A \text{ not erased})$, respectively, for all the error patterns A of the wiretapper channel BEC_{p_w}. As argued before, the probability that the erasures does not exceed $p_w - \xi = 1 - \tau$ fraction is negligible. We average over the channel randomness of the wiretapper channel BEC_{p_w} and claim that the statistical distance of $(\mathsf{W}|\mathbf{S} = \mathbf{s}_0)$ and $(\mathsf{W}|\mathbf{S} = \mathbf{s}_1)$ is arbitrarily small for big enough N. According to Remark 1, this is strictly stronger than $\mathsf{H}(\mathbf{S}|\mathsf{W}) \geq \ell(1 - o(1))$, where $\mathbf{S}$ is a uniform message. The deduction takes a few steps. The wiretap indistinguishability security as defined in (1) is equivalent to wiretap mutual information security. The wiretap mutual information security is stronger than its random message analogue, which in turn is stronger than the Wyner wiretap secrecy as defined in [39, 13].

Finally we use the coding rate upper bound of the Wyner wiretap code to bound the coding rate of (ε, δ)-SSS. We have shown that a (ε, δ)-SSS with relative threshold pair (τ, κ) is a wiretap code for the pair (BEC_{p_m},BEC_{p_w}). According to [39, 13], the achievable coding rate for Wyner wiretap code with a BEC_{p_m} main channel and a BEC_{p_w} wiretapper channel is $(1-p_m)-(1-p_w) = p_w - p_m = \kappa - \tau + 2\xi$. Since this holds for any constant $\xi > 0$, we obtain an upper bound of $\kappa - \tau$ for (ε, δ)-SSS with relative threshold pair (τ, κ). □

B One-Time-Pad Trick of Inverting Extractors

There is a well known way to transform an efficient function into one that is also efficiently invertible through a "One-Time-Pad" trick. We give a proof for the special case of affine extractors, for completeness.

Lemma 10. Let $\mathsf{AExt} : \{0,1\}^n \to \{0,1\}^m$ be an affine (n, k)-extractor with error ε. Then $\mathsf{AExt}' : \{0,1\}^{n+m} \to \{0,1\}^m$ defined as follows is a ε-invertible affine $(n + m, k + m)$-extractor with error ε.

$$\mathsf{AExt}'(\mathsf{z}) = \mathsf{AExt}'(\mathsf{x}||\mathsf{y}) = \mathsf{AExt}(\mathsf{x}) + \mathsf{y},$$

where the input $\mathsf{z} \in \{0,1\}^{n+m}$ is separated into two parts: $\mathsf{x} \in \{0,1\}^n$ and $\mathsf{y} \in \{0,1\}^m$.

Proof. Let Z be a random variable with flat distribution supported on an affine subspace of $\{0,1\}^{n+m}$ of dimension at least $k + m$. Separate Z into two parts $\mathsf{Z} = (\mathsf{X}||\mathsf{Y})$, where $\mathsf{X} \in \{0,1\}^n$ and $\mathsf{Y} \in \{0,1\}^m$. Then conditioned on any $\mathsf{Y} = \mathsf{y}$, X has a distribution supported on an affine subspace of $\{0,1\}^n$ of dimension at least k. This asserts that conditioned on any $\mathsf{Y} = \mathsf{y}$,

$$\mathsf{SD}(\mathsf{AExt}(\mathsf{X}) + \mathbf{y}; U_{\{0,1\}^m}) \leq \varepsilon.$$

Averaging over the distribution of Y concludes the extractor proof.

We next show an efficient inverter AExt'^{-1} for AExt'. For any $\mathsf{s} \in \{0,1\}^m$, define

$$\mathsf{AExt}'^{-1}(\mathsf{s}) = (U_n||\mathsf{AExt}(U_n) + \mathsf{s}).$$

The randomised function AExt'^{-1} is efficient and $\mathsf{AExt}'^{-1}(U_m) \overset{\varepsilon}{\sim} U_{n+m}$. □

This transformation only gives a v-invertible affine extractor with $v = \varepsilon$. But it is easy to see that Theorem 3 can be restated with respect to a v-invertible affine extractor. On the other hand, in the full version of this paper, we explicitly construct a v-invertible affine extractor with $v = 0$ that can be directly used in Theorem 3.

C Non-explicit Stochastic Affine ECC

The following construction using a linear optimal erasure list-decodable code and an AMD code gives an SA-ECC that can be used in our constructions of (ε, δ)-SSS. The construction is simply a concatenation of the systematic AMD in [12] and the erasure list-decodable code whose existence is given below. The SA-ECC obtained further satisfies a special detection property that is needed in Theorem 5.

Lemma 11 ([18])**.** For every integer $L \geq 1$ and every $p \in [0,1]$, the achievable rate of binary linear list-decodable codes that corrects a p fraction of erasures by outputting a list of L messages is at least

$$1 - \frac{p}{r}\log(2^r - 1) - \frac{h_2(p)}{r}, \qquad (7)$$

where $r = \lceil \log(L + 1) \rceil$.

The bound in (7) shows that if we allow a big enough list size, the rate R_{ECC} can be made arbitrarily close to $1-p$. Note that Lemma 11 only states the existence of high rate linear codes with efficient algorithm that outputs a list of L messages. It is still an open problem to explicitly construct codes with rate matching this bound.

D Proof Sketch of Theorem 5

Proof. We deter the details of the construction to the full version of this paper. For now we refer to [19, Theorem 6.1] and point out the adaptations needed. There are six building blocks involved in the construction: SC, RS, Samp, KNR, POLY_t and REC. We replace the first and last building blocks.

The first building block is a *Stochastic Code* (SC). We need two properties from this building block: detect (output $\perp$) when the codeword is masked by a random offset and correct from erasures of no more than $1-\kappa+\varepsilon$ fraction. While the former property is always satisfied by the original SC used in [19], the latter property might not hold. When $1-\kappa$ is small, we can let the decoder of the SC used in [19] set the erased bits to 0 and decode from this "error". But when $1-\kappa>\frac{1}{2}$, this trick no longer works. We use the construction from Appendix C for SC in our construction of $\mathsf{SA\text{-}ECC}$.

The last building block is a *Random Error Code* (REC). We also need two properties from this building block: correct from random erasures of $1-\kappa$ fraction and the encoder is a linear function. We need the latter property for affine property of the $\mathsf{SA\text{-}ECC}$ constructed. Explicit linear codes at rate $1-p$ that correct p fraction of random erasures are known. We can use any explicit construction of capacity achieving codes for $\mathrm{BEC}_{1-\kappa}$ for REC and use a similar argument of [34].

We now refer to the **Algorithm 1.** in the proof of [19, Theorem 6.1] and show that, with the SC and REC replaced accordingly, we do have a $\mathsf{SA\text{-}ECC}$. The error correction capability and optimal rate follow similarly as in the proof of [19, Theorem 6.1]. We next show affine property. **Phase 1** and **Phase 2** are about the *control information*, which are part of the encoding randomness r of the $\mathsf{SA\text{-}ECC}$ to be fixed to constant value in the analysis of affine property. During **Phase 3**, the message $\mathbf{m}$ is linearly encoded (our REC is linear) and then permuted, followed by adding a translation term Δ_{r}. Since permutation is a linear transformation, we combine the two linear transformations and write $\mathbf{m}G_{\mathsf{r}}+\Delta_{\mathsf{r}}$, where G_{r} is a binary matrix. Finally, during **Phase 4**, some blocks that contain the control information are inserted into $\mathbf{m}G_{\mathsf{r}}+\Delta_{\mathsf{r}}$. We add dummy zero columns into G_{r} and zero blocks into Δ_{r} to the corresponding positions where the control information blocks are inserted. Let $\mathbf{m}\hat{G}_{\mathsf{r}}+\hat{\Delta}_{\mathsf{r}}$ be the vector after padding dummy zeros. Let $\hat{\Delta}'_{\mathsf{r}}$ be the vector obtained from padding dummy zero blocks, complementary to the padding above, to the control information blocks. We then write the final codeword of the $\mathsf{SA\text{-}ECC}$ in the form $\mathbf{m}\hat{G}_{\mathsf{r}}+(\hat{\Delta}_{\mathsf{r}}+\hat{\Delta}'_{\mathsf{r}})$, which is indeed an affine function of the message $\mathbf{m}$.

□

E Alternative proof of Theorem 7

In this alternative proof, we first prove a general property of a strong linear (k,ε)-extractor $\mathsf{Ext}\colon\{0,1\}^d\times\{0,1\}^n\rightarrow\{0,1\}^m$. Roughly speaking, we prove that the pre-images (the pre-image of $\mathbf{m}\in\{0,1\}^m$ is a random variable tuple (Z,X) that satisfies the condition $\mathsf{Ext}(\mathsf{Z},\mathsf{X})=\mathbf{m}$) of any two extractor outputs can not be distinguished by any affine function $f_A\colon\{0,1\}^{d+n}\rightarrow\{0,1\}^t$ with $t\leq n-k$. The privacy of the SSS in Theorem 7 then follows trivially as a natural consequence of this property.

E.1 Abstraction of privacy proof as showing an extractor property

Let $\Pi_A : \{0,1\}^N \rightarrow \{0,1\}^t$ be the projection function that maps a share vector to the t shares with index set $A \subset [N]$ chosen by the non-adaptive adversary.

The first step of the abstraction is we interpret the combination of the projection function Π and the $\mathsf{SA\text{-}ECCenc} : \{0,1\}^{n+d} \rightarrow \{0,1\}^N$ (for any fixed randomness $\mathbf{r}$) as the affine function $f_A \colon \{0,1\}^{d+n} \rightarrow \{0,1\}^{\rho n}$ mentioned above. So the view of the adversary is simply the output of the affine function $f_A = \Pi \circ \mathsf{SA\text{-}ECCenc}$ applied to a random variable tuple (Z, X) to be defined below.

The second step of the abstraction is concerning how we obtain the random variable tuple (Z, X). The sharing algorithm of the SSS (before applying the stochastic affine code) takes a secret, which is a particular extractor output $\mathbf{m} \in \{0,1\}^m$, 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 $\mathsf{Ext}(\mathsf{z}, \mathsf{x}) = \mathbf{m}$. This process of obtaining (z, x) is the same as sampling $(U_d, U_n) \overset{\$}{\leftarrow} \{0,1\}^{n+d}$ and then restrict to $\mathsf{Ext}(U_d, U_n) = \mathbf{m}$. In the rest of the proof, we define the random variable tuple

$$(\mathsf{Z}, \mathsf{X}) := (U_d, U_n) | \left(\mathsf{Ext}(U_d, U_n) = \mathbf{m}\right) \tag{8}$$

and refer to it as the pre-image of $\mathbf{m}$.

We can now formulate the privacy of the SSS in this context. We want to prove that the statistical distance of the views of the adversary for a pair of secrets can be made arbitrarily small. The views of the adversary are the outputs of the affine function f_A with inputs (Z, X) and $(\mathsf{Z}', \mathsf{X}')$ for the secret $\mathbf{m}$ and $\mathbf{m}'$, respectively. It is sufficient to show that no affine function f_A with t bits output can distinguish the pre-images (Z, X) and $(\mathsf{Z}', \mathsf{X}')$.

E.2 Proof of the property

We next prove a property of strong linear extractors. For the property to hold, we in fact only need the extractor to be able to extract from affine sources. But since seeded extractors for general sources with good parameters are not more difficult to construct than that for affine sources, we state the property with a condition stronger than necessary.

Theorem 12. Let $\mathsf{Ext} \colon \{0,1\}^d \times \{0,1\}^n \rightarrow \{0,1\}^m$ be a strong linear (k, ε)-extractor. Let $f_A \colon \{0,1\}^{d+n} \rightarrow \{0,1\}^t$ be any affine function with output length $t \leq n - k$. Let $(\mathsf{Z}, \mathsf{X}) = ((U_d, U_n) | \mathsf{Ext}(U_d, U_n) = \mathbf{m})$ and $(\mathsf{Z}', \mathsf{X}') = ((U_d, U_n) | \mathsf{Ext}(U_d, U_n) = \mathbf{m}')$ for any $\mathbf{m}, \mathbf{m}' \in \{0,1\}^m$. We have

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

Proof. Without loss of generality, we assume that the linear function $\mathsf{Ext}(\cdot, \mathsf{z})$, for every seed z, has the entire $\{0,1\}^m$ as its image [3]. Without loss of generality, it suffices to assume that f_A is of the form $f_A(\mathsf{Z}, \mathsf{X}) = (\mathsf{Z}, W(\mathsf{X}))$ for some affine function $W \colon \{0,1\}^n \rightarrow \{0,1\}^t$ (this is because for any arbitrary f_A, the information contained in $f_A(\mathsf{Z}, \mathsf{X})$ can be obtained from $(\mathsf{Z}, W(\mathsf{X}))$ for a suitable choice of W).

[3] If this condition is not satisfied for some choice z of the seed, there must be linear dependencies between the k output bits of $\mathsf{Ext}(\cdot, \mathsf{z})$. Therefore, for this choice Ext can never be an extractor and arbitrarily changing $\mathsf{Ext}(\cdot, \mathsf{z})$ to be an arbitrary full rank linear function will not change the overall performance of the extractor.

Let $\mathcal{D}$ be the uniform distribution on the image of W. For the above pairwise guarantee to hold, it suffices to show that for every fixed choice of $\mathbf{m} \in \{0,1\}^m$, the distribution of $f_A(\mathsf{Z}, \mathsf{X})$ is (4ε)-close to $U_d \times \mathcal{D}$, where U_d is the uniform distribution on $\{0,1\}^d$.

Let $\mathsf{K} \leftarrow \{0,1\}^n$ be a random variable uniformly distributed over the kernel of the linear transformation defined by W, and note that it has entropy at least $n - t \geq k$. The extractor Ext thus guarantees that $\mathsf{Ext}(\mathsf{K}, \mathsf{Z})$, for a uniform and independent seed Z, is ε-close to uniform. By averaging, it follows that for at least $1 - 4\epsilon$ fraction of the choices of the seed $\mathsf{z} \in \{0,1\}^d$, the distribution of $\mathsf{Ext}(\mathsf{K}, \mathsf{z})$ is $(1/4)$-close to uniform. We now use the following claim:

Claim 1. *Let U be uniformly distributed on $\{0,1\}^m$ and U' be any affine source that is not uniform on $\{0,1\}^m$. Then, the statistical distance between U and U' is at least $1/2$.* □

Since Ext is a linear function for every seed, the distribution of $\mathsf{Ext}(\mathsf{z}, \mathsf{K})$ for any seed z is an affine source. Therefore, the above claim allows us to conclude that for at least $1 - 4\epsilon$ fraction of the choices of s, the distribution of $\mathsf{Ext}(\mathsf{z}, \mathsf{K})$ is *exactly* uniform. Let $\mathcal{G} \subseteq \{0,1\}^d$ be the set of such choices of the seed. Observe that if $\mathsf{Ext}(\mathsf{z}, \mathsf{K})$ is uniform for some seed z, then for any affine translation of K, namely, $\mathsf{K} + v$ for any $v \in \{0,1\}^n$, we have that $\mathsf{Ext}(\mathsf{z}, \mathsf{K} + v)$ is uniform as well. This is due to the linearity of the extractor.

According to (8), the distribution $(\mathsf{Z}, W(\mathsf{X}))$ can be obtained as $(U_d, W(U_n))|(\mathsf{Ext}(U_d, U_n) = \mathbf{m})$. We take two steps to get there. Step one, we find out the distribution $(U_d, W(U_n))|(\mathsf{Ext}(U_d, U_n) = \mathbf{m}, U_d = \mathsf{z})$ for a particular seed z (Proposition 2). Step two, the distribution we finally want is the convex combination of the distributions obtained in Step one (Proposition 3).

Consider a uniformly distributed random variable $\mathsf{M} \overset{\$}{\leftarrow} \{0,1\}^m$, and an independent and uniform $\mathsf{Z} \overset{\$}{\leftarrow} \{0,1\}^d$. Let (Z, Y) be the pre-image of the random variable M and define the shorthand $\mathsf{W} := W(\mathsf{Y})$. The rest of the proof is focus on the three random variables W, M and Z.

Proposition 2. *Let $\mathsf{z} \in \mathcal{G}$ and consider any $\mathbf{m} \in \{0,1\}^m$. Then, the conditional distribution of $\mathsf{W}|(\mathsf{Z} = \mathsf{z}, \mathsf{M} = \mathbf{m})$ is exactly $\mathcal{D}$.* □

Note that the distribution of (Z, Y) is uniform on $\{0,1\}^{d+n}$. Therefore, the distribution of (Z, W) is exactly $U_d \times \mathcal{D}$. In particular, for any $\mathsf{z} \in \{0,1\}^d$, the conditional distribution $\mathsf{W}|(\mathsf{Z} = \mathsf{z})$ is exactly $\mathcal{D}$.

Fix any $\mathsf{z} \in \mathcal{G}$ and let $w \in \{0,1\}^t$ be any element in the image of W. The conditional distribution $\mathsf{Y}|(\mathsf{Z} = \mathsf{z})$ is uniform over $\{0,1\}^n$ and the conditional distribution $\mathsf{Y}|(\mathsf{Z} = \mathsf{z}, \mathsf{W} = w)$ is uniform over a translation of K. By the above argument, recalling $\mathsf{M} = \mathsf{Ext}(\mathsf{Z}, \mathsf{Y})$, we therefore know that the conditional distribution of $\mathsf{M}|(\mathsf{Z} = \mathsf{z}, \mathsf{W} = w)$ is exactly uniform over $\{0,1\}^m$. Since the conditional distribution of $\mathsf{W}|(\mathsf{Z} = \mathsf{z})$ is $\mathcal{D}$, this means that the conditional distribution of $(\mathsf{M}, \mathsf{W})|(\mathsf{Z} = \mathsf{z})$ is exactly $U_m \times \mathcal{D}$. We have therefore proved Proposition 2.

Proposition 3. *For any $\mathsf{m} \in \{0,1\}^m$, the conditional distribution of $(\mathsf{Z}, \mathsf{W})|(\mathsf{M} = \mathbf{m})$ is (4ϵ)-close to $U_d \times \mathcal{D}$.*

It suffices to note that the distribution of $(\mathsf{Z}, \mathbf{W})|(\mathsf{M} = \mathbf{m})$ is a convex combination of the distributions $(\mathsf{Z}, \mathbf{W})|(\mathsf{M} = \mathbf{m}, \mathsf{Z} = \mathsf{z})$ and then use the result of Proposition 2 along with the fact that $\Pr[\mathsf{Z} \in \mathcal{G}] \leq 4\varepsilon$. A detailed derivation follows.

Recall that for any $\mathsf{z} \in \{0,1\}^d$, the conditional distribution of $\mathsf{W}|(\mathsf{Z} = \mathsf{z})$ is exactly $\mathcal{D}$ (since $\mathsf{Y}|(\mathsf{Z} = \mathsf{z})$ is uniform over $\{0,1\}^n$). Consider any event $\mathcal{E} \subseteq \{0,1\}^{d+t}$ and let $p := \Pr[(\mathsf{Z}, \mathbf{W}) \in \mathcal{E}]$.

Since Z and W are independent, we have that

$$p = 2^{-d} \sum_{(\mathsf{z},w)\in\mathcal{E}} \mathcal{D}(w),$$

where $\mathcal{D}(w)$ denotes the probability assigned to the outcome w by $\mathcal{D}$. On the other hand, we shall write down the same probability in the conditional probability space $\mathsf{M} = \mathbf{m}$ and show that it is different from p by at most 4ϵ, concluding the claim on the statistical distance. We have

$$\begin{aligned}
\Pr[(\mathsf{Z}, \mathbf{W}) \in \mathcal{E} | \mathsf{M} = \mathbf{m}] &= \sum_{(\mathsf{z},w)\in\mathcal{E}} \Pr[\mathsf{Z} = \mathsf{z}, W = w | \mathsf{M} = \mathbf{m}] \\
&= \sum_{(\mathsf{z},w)\in\mathcal{E}, \mathsf{z}\in\mathcal{G}} \Pr[\mathsf{Z} = \mathsf{z}, \mathsf{W} = w | \mathsf{M} = \mathbf{m}] + \sum_{(\mathsf{z},w)\in\mathcal{E}, \mathsf{z}\notin\mathcal{G}} \Pr[\mathsf{Z} = \mathsf{z}, \mathsf{W} = w | \mathsf{M} = \mathbf{m}].
\end{aligned}$$

Note that

$$\eta := \sum_{(\mathsf{z},w)\in\mathcal{E}, \mathsf{z}\notin\mathcal{G}} \Pr[\mathsf{Z} = \mathsf{z}, \mathsf{W} = w | \mathsf{M} = \mathbf{m}] \le \Pr[\mathsf{Z} \notin \mathcal{G} | \mathsf{M} = \mathbf{m}] \le 4\epsilon,$$

since M and Z are independent. Therefore,

$$\begin{aligned}
\Pr[(\mathsf{Z}, \mathbf{W}) \in \mathcal{E} | \mathsf{M} = \mathbf{m}] &= \sum_{(\mathsf{z},w)\in\mathcal{E}, \mathsf{z}\in\mathcal{G}} \Pr[\mathsf{Z} = \mathsf{z}, \mathsf{W} = w | \mathsf{M} = \mathbf{m}] + \eta \\
&= 2^{-d} \sum_{(\mathsf{z},w)\in\mathcal{E}, \mathsf{z}\in\mathcal{G}} \Pr[W = w | \mathsf{M} = \mathbf{m}, \mathsf{Z} = \mathsf{z}] + \eta \tag{9} \\
&= 2^{-d} \sum_{(\mathsf{z},w)\in\mathcal{E}, \mathsf{z}\in\mathcal{G}} \mathcal{D}(w) + \eta \tag{10} \\
&= 2^{-d} \Big(\sum_{(\mathsf{z},w)\in\mathcal{E}} \mathcal{D}(w) - \sum_{(\mathsf{z},w)\in\mathcal{E}, \mathsf{z}\notin\mathcal{G}} \mathcal{D}(w) \Big) + \eta
\end{aligned}$$

where (9) uses the independence of X and S and (10) follows from Proposition 2. Observe that

$$\eta' := 2^{-d} \sum_{(\mathsf{z},w)\in\mathcal{E}, \mathsf{z}\notin\mathcal{G}} \mathcal{D}(w) = 2^{-d} \sum_{\mathsf{z}\notin\mathcal{G}} \sum_{\substack{w: \\ (\mathsf{z},w)\in\mathcal{E}}} \mathcal{D}(w) \le 2^{-d}(2^d - |\mathcal{G}|) \le 4\epsilon.$$

Therefore,

$$\Pr[(\mathsf{Z}, \mathbf{W}) \in \mathcal{E} | \mathsf{M} = \mathbf{m}] = p + \eta - \eta' = p \pm 4\epsilon = \Pr[(\mathsf{Z}, \mathbf{W}) \in \mathcal{E}] \pm 4\epsilon,$$

since $0 \le \eta \le 4\varepsilon$ and $0 \le \eta' \le 4\varepsilon$. The claim follows. □