Impossibility of Efficient Information-Theoretic Fuzzy Extraction*

Benjamin Fuller
University of Connecticut, Storrs, CT 06269, United States
benjamin.fuller@uconn.edu

March 14, 2024

Abstract

Fuzzy extractors convert noisy signals from the physical world into reliable cryptographic keys. Fuzzy min-entropy measures the limit of the length of key that a fuzzy extractor can derive from a distribution (Fuller, Reyzin, and Smith, IEEE Transactions on Information Theory 2020). In general, fuzzy min-entropy that is superlogarithmic in the security parameter is required for a noisy distribution to be suitable for key derivation.

There is a wide gap between what is possible with respect to computational and information-theoretic adversaries. Under the assumption of general-purpose obfuscation, keys can be securely derived from all distributions with superlogarithmic entropy. Against information-theoretic adversaries, however, it is impossible to build a single fuzzy extractor that works for all distributions (Fuller, Reyzin, and Smith, IEEE Transactions on Information Theory 2020).

A weaker information-theoretic goal is building a fuzzy extractor for each probability distribution. This is the approach taken by Woodage et al. (Crypto 2017). Prior approaches use the full description of the probability mass function and are inefficient. We show this is inherent: **for a quarter of distributions with fuzzy min-entropy and 2^k points there is no secure fuzzy extractor that uses less $2^{\Theta(k)}$ bits of information about the distribution.**

We show an analogous result with stronger parameters for information-theoretic secure sketches. Secure sketches are frequently used to construct fuzzy extractors.

Keywords: Fuzzy extractors, information theory, information reconciliation, secure sketches.

1 Introduction

Information reconciliation and privacy amplification are the two fundamental tasks for key derivation from noisy sources. Roughly speaking, information reconciliation takes two correlated distributions w and w' and maps them to the same value while minimizing what is leaked about that value. Privacy amplification converts the uncertainty in this mapped value to a uniform value suitable for cryptography. Applications areas include quantum key agreement, biometrics, and physically uncloneable functions [BBR88, DORS08].

We focus on non-interactive versions of these problems [DORS08] as defined by secure sketches, which perform information-reconciliation, and fuzzy extractors, which perform both information-reconciliation and privacy amplification. A Secure Sketch consists of a pair of algorithms $(\mathsf{SS}, \mathsf{Rec})$ where:

1. $\mathsf{SS}(w) = ss$ should reveal as little information as possible about w; and

*The author thanks the reviewers for their helpful feedback and Luke Demarest and Alexander Russell for their helpful discussions. B.F. is supported by National Science Foundation Grants #2232813 and #2141033 and the Office of Naval Research.

2. $\mathsf{SS}(w) = ss$ should allow one to reconstruct w from a nearby w'. That is, it should be the case that for all nearby $w', \mathsf{Rec}(w', ss) = w$. In the above, "nearby" is w' such that $\mathsf{dis}(w, w') \leq t$ for distance metric dis and distance t.

These two properties are in tension because allowing recovery of w requires information about w. The most natural (inefficient) construction is for ss to be a pairwise independent [CW77] hash h of w [ST09, FRS16, WCD+17, FRS20]. The hash h should be long enough so that $\{w|h(w) = y \wedge w' s.t. \mathsf{dis}(w, w') \leq t\} = 1$ and short enough so $\{w|h(w) = y\}$ is large. Efficient constructions are also known based on error-correcting codes. This is achieved by writing down the coset of w with respect to an error-correcting code with distance t [DORS08]. In fact, upper bounds on the unpredictability of $w|ss$ are related to the size of the best error-correcting codes [DORS08, FMR20]. Given a good information reconciliation, one can achieve privacy amplification using an average-case randomness extractor [NZ93] to convert w into a uniform value.

Fuzzy extractors perform both information reconciliation and privacy amplification. They consist of a pair $(\mathsf{Gen}, \mathsf{Rep})$. Intuitively, Gen converts a value w into a uniform value, denoted as r and Rep reproduces that value for any nearby w'. Notationally, $(r, p) \leftarrow \mathsf{Gen}(w)$ should be indistinguishable from (u, p) where u is a truly random value. On the correctness side, it should be the case that for all w' such that $\mathsf{dis}(w, w') \leq t$ then $\mathsf{Rep}(w', p) = r$. Both SS and Gen are allowed to have private internal randomness.

Since noisy sources come from the physical world, an important goal is to be able to support as many distributions W as possible. This goal is the focus of this work. Throughout the Introduction, we use the notation of fuzzy extractors and note when there are material differences for secure sketches. Fuller, Reyzin, and Smith [FRS16, FRS20] identified the notion of fuzzy min-entropy, $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$, which measures the adversary's success when given oracle access to $\mathsf{Rep}(\cdot, p)$ but is unable to learn anything from the value p. Fuzzy min-entropy quantifies the weight of the heaviest ball in the probability mass function of W. That is,

$$\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) := -\mathsf{log}\left(\max_{w'} \sum_{w} \Pr[W = w | \mathsf{dis}(w, w') \leq t]\right).$$

Ideally, one would build a single fuzzy extractor that works for the family of all distributions $\mathcal{W}^{\mathtt{all}}_{\mathtt{fuzz}} = \{W | \mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) = \omega(\mathsf{log}(\lambda))\}$ for some security parameter λ. We call such a fuzzy extractor *universal* as it simultaneously works for any secureable distribution W. If one desires computational security, a universal fuzzy extractor is achievable using general obfuscation [BBC+14, BCKP14, BCKP17] or under specific number-theoretic assumptions [GZ19].

The situation for information-theoretic security is more complicated.[1] Fuller, Reyzin, and Smith [FRS20] showed that it is impossible to build a universal fuzzy extractor with information-theoretic security. More precisely, they constructed a family of distributions $\mathcal{W}$ and showed that any fuzzy extractor $(\mathsf{Gen}, \mathsf{Rep})$ must be insecure for an average member of $\mathcal{W}$. Let z be a string that indexes the family $\mathcal{W}$. We use Z to describe a uniformly chosen index for the family $\mathcal{W}$. We use the notation $z \leftarrow Z$ to indicate this choice. We use the notation W_Z to indicate sampling W uniformly from $\mathcal{W}$ with Z being a random variable that describes the choice of $W \in \mathcal{W}$. For all z, the goal is to build a good fuzzy extractor for W_z. The impossibility result shows a family $\mathcal{W}$ where any $(\mathsf{Gen}, \mathsf{Rep})$ is insecure for an average z chosen according to Z. The model tells the adversary the outcome $Z = z$ but not the individual point $w \leftarrow W_z$ that is input to Gen.

On the positive side, multiple works [HTW14, HTW16, FRS16, WCD+17, TW17, TVW18, LA18, FP19, FRS20] presented a construction that works for each $W_z \in \mathcal{W}^{\mathtt{all}}_{\mathtt{fuzz}}$. This is called the *distribution-sensitive* setting as Gen also knows the entire probability mass function described by z, denoted as

[1] Fuzzy extractors were first designed as an information-theoretic primitive because of strong connections to randomness extraction and coding theory. An important application is in quantum key agreement which does not allow computational assumptions. Many computational constructions use an information-theoretic secure sketch [WL18, WLG19]. (Exceptions exist such as the universal constructions listed above and constructions for distributions with statistical properties beyond fuzzy min-entropy [ACEK17, ABC+18, FMR20, CFP+21].)

$\mathsf{Gen}_z, \mathsf{Rec}_z$. All constructions in this line are computationally inefficient; for an input point w they look up the probability that $\Pr[W_z = w]$ and the probability of points w' where $\mathsf{dis}(w, w') \leq t$. We show this inefficiency is unavoidable:

> **Any distribution-sensitive information-theoretic fuzzy extractor requires an exponential amount of information about the distribution W_z.**

Our results are for the Hamming metric over $\{0,1\}^n$. Below we present the two informal theorems for fuzzy extractors (see Theorem 6) and secure sketches (see Theorem 11) respectively. For a value $p \in [0,1]$ let h_2 be the binary entropy of p. Both secure sketches and fuzzy extractor are frequently parameterized by an error parameter δ which controls the maximum probability they get the wrong value. We consider $\delta = 0$ for the fuzzy extractor setting and $\delta > 0$ for the secure sketch setting. (Discussion in Section 1.2.)

Theorem 1 (Informal Theorem 6)**.** *Consider $\{0,1\}^n$ and $t < n/2$ be a distance parameter. Let $\mathcal{W}_\gamma = \{W | \mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) = \gamma\}$. Let $c > 0$ be a constant and suppose that*

$$\gamma \leq n \cdot \min\left\{(1 - h_2(t/n)) + o(1), \frac{1 - \Theta(c) - h_2(1/2 - t/n)}{3}\right\}.$$

For a quarter of $W \in \mathcal{W}_\gamma$ there is no fuzzy extractor that simultaneously has

1. *no error,*
2. *is of size at most $2^{\gamma + cn}$, and*
3. *extracts keys of length $\omega(\mathsf{log}(n))$ that are within statistical distance $1/3 - \mathtt{ngl}(n)$ to a uniform key.*

Theorem 2 (Informal Theorem 11)**.** *Consider $\{0,1\}^n$ and $t < n/2$ be a distance parameter. Let $\mathcal{W}_\gamma = \{W | \mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) = \gamma\}$. Let $\delta < 1/4$ be the error of the secure sketch, let $c > 0$ be a constant and suppose that*

$$\gamma \leq n \cdot \min\left\{(1 - h_2(t/n)) + o(1), c_\delta h_2(t/n) - \Theta(c)\right\}.$$

where $1/3 \leq c_\delta \leq 2/3$ and depends on $h_2(\delta)$. For 2^{-5} fraction of $W \in \mathcal{W}_\gamma$ there is no secure sketch of size of at most $2^{\gamma + cn}$ that retains unpredictability of $w|ss$ of at least 5.

The *size* of the fuzzy extractor (resp. secure sketch) refers to the amount of information the algorithm has about z, it is not a restriction on the running time of the algorithm, our results hold for unbounded time algorithms. The relevant parameter regimes of impossibility are shown in Figure 1. The two most important parameters are the noise rate t/n and the fuzzy entropy rate γ/n. The area under the curves represents parameters where the construction is impossible for the fraction of distributions in the informal theorems unless one has algorithms of $2^{\Theta(n)}$ size. In spirit, our result rules out constructions that do not have a full description of the probability mass function written in their description.

Our results use only first and second-moment bounds. **Our theorems are crucial for the future of information-theoretic fuzzy extractors and secure sketches. To prove security for an efficient construction one must either restrict to sources with high fuzzy min-entropy or use properties of a noisy source beyond fuzzy min-entropy.** We discuss this more in Section 1.2.

1.1 Proof Techniques

Our results are information-theoretic. We consider a family of distributions $\mathcal{W} = \{W_z\}$ indexed by a string z. We let $\mathcal{Z}$ denote the set of possible z and let Z denote the uniform distribution over $\mathcal{Z}$. Lastly, we use $w \leftarrow W_z$ to denote sampling a point from the distribution indexed by z. We show the impossibility of two types of fuzzy extractors:

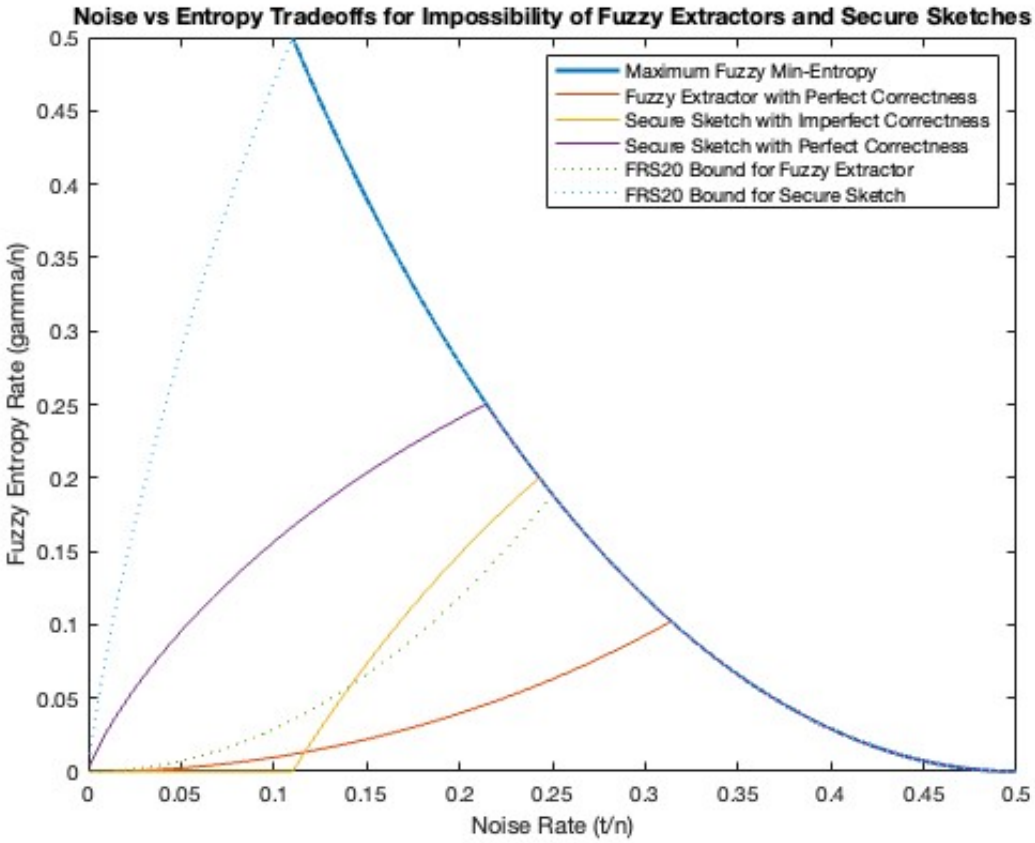

Figure 1: The region of error rate t/n (x-axis) and fuzzy entropy rate γ/n (y-axis) pairs for which the two negative results apply. The six curves are maximum fuzzy min-entropy $\gamma/n = (1 - h_2(t/n))$, Theorem 6, Theorem 11 with $\delta = .25$, Theorem 11 with $\delta = 0$, [FRS20, Theorem 5.1] and [FRS20, Theorem 7.2]. The parameter δ is how frequently the secure sketch is allowed to be incorrect. We consider fuzzy extractors with perfect correctness where $\delta = 0$.

Def. 8 (Universal) Fuzzy extractors with distributional advice. This is a triplet of algorithms $(\mathsf{advice}, \mathsf{Gen}, \mathsf{Rep})$ designed to work for all $W_z \in \mathcal{W}$ for a fixed error tolerance t. The fuzzy extractor is given information about z through a function $\mathsf{advice} = \mathsf{advice}(z)$ which is input to both Gen and Rep. The value of advice *specializes* $(\mathsf{Gen}, \mathsf{Rep})$ to the distribution described by z.
Define $w \leftarrow W_z$ and $(r, p) \leftarrow \mathsf{Gen}(w, \mathsf{advice})$, it should be true that

$$(r, p, z) \approx (u, p, z).$$

where u is uniformly and independently sampled. Since $\mathsf{advice}(\cdot)$ is a function, advice is available to the adversary.

Def. 6 Fuzzy extractors for a specific distribution $W_z \in \mathcal{W}$ that are required to have a bounded size description of $(\mathsf{Gen}, \mathsf{Rep})$.

We show impossibility of building a fuzzy extractor with distributional advice of length ℓ for $\mathcal{W}$ implies impossibility of building a space bounded fuzzy extractor for length ℓ for a constant fraction of $\mathcal{W}$ (Lemma 4). The core of our negative results is to show the impossibility of building fuzzy extractors with distributional advice.

We review Fuller, Reyzin, and Smith's [FRS20] impossibility result. Fuzzy extractor correctness says that for $(r, p) \leftarrow \mathsf{Gen}(w)$ for all w' close to w the correct key is reproduced, i.e., $\mathsf{Rep}(w', p) = r$. As such, for each value of p, one can partition the input space $\{0,1\}^n$ by what value of r the point $v \in \{0,1\}^n$ produces. Values v that could have produced r will be at least distance t from the boundary of this partition, we call the set of such v, $\mathtt{Viable}_{r,p}$. $\mathtt{Viable}_{r,p}$ can be bounded geometrically using the isoperimetric inequality [Har66]. This bound applies for any distribution over the inputs w.

Consider the following simple distinguisher for a triple r, p, z. One computes the key partition described above and the set $\mathtt{Viable}_{r,p}$. If $\mathtt{Viable}_{r,p} \cap W_z = \emptyset$ output the key is random, otherwise output key is real. The core of Fuller, Reyzin, and Smith's impossibility was to build a family $\mathcal{W}^{FRS}$ with two properties:

1. The distribution was 2-universal [CW77], so the remainder of the distribution was unknown conditioned on the input w.
2. Distributions $W_z \in \mathcal{W}^{FRS}$ shared few points and had fuzzy min-entropy.

These two properties meant that for any partition p created after seeing w for most distributions W_z where $\Pr[W_z = w] > 0$ have few parts with nonempty interiors. Thus, the above distinguisher works.

The family is as follows: let $\mathbf{C}$ be a linear error-correcting code with distance t, let $\mathbf{H}$ be its syndrome, let c be a coset. Then each $z = (\mathbf{H}, c)$ and a distribution $W_{z=(\mathbf{H},c)}$ is the uniform distribution over the set of all points $\{w \mid \mathbf{H}w = c\}$.

Moving to the distributional advice setting To set notation for the distributional advice game, we consider the following game for a tuple of algorithms $(\mathsf{advice}, \mathsf{Gen}, \mathsf{Rep})$:

1. A uniform sample $z \leftarrow Z$ which picks $W_z \in \mathcal{W}$.
2. A bounded length $\mathsf{advice} = \mathsf{advice}(z)$ is computed.
3. Sample $w \leftarrow W_z$.
4. The algorithm computes $(r, p) \leftarrow \mathsf{Gen}(w, \mathsf{advice})$.
5. The adversary is given either (r, p, z) or (u, p, z) for a uniform u.

In [FRS20], the only information that Gen has about z was the input point w. In our setting, Gen gets advice. Fuller, Reyzin, and Smith's family had a short description so advice allows Gen to align $\mathtt{Viable}$ with points in W_z. Thus, extending the result requires a long description that can't be compressed. We consider the natural candidate: the set $\mathcal{W}_\gamma$ of all distributions with fuzzy min-entropy at least γ.

We use the notation $\mathcal{W}_{n,k} = \{W | W \text{ has support size } 2^k\}$. For a positive integer γ, If one considers $k = \gamma + cn$ for some $c > 0$ there are few distributions $W_z \in \mathcal{W}_{n,k}$ where $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) < \gamma$. As long as $|\mathsf{advice}|$ is shorter than 2^k, most points in the support of W_z are unpredictable conditioned on advice.

The techniques for the secure sketch setting are similar, however, there are stronger geometric bounds on the number of viable points because secure sketches imply Shannon error correcting codes [DORS08, FMR20]. Our result considers a secure sketch that retains smooth min-entropy instead of min-entropy. This is so we can use $\mathcal{W}_{n,k}$ throughout the proof and "smooth" to a family where every distribution has fuzzy min-entropy γ. Our final result also applies to secure sketches that retain non-smooth conditional min-entropy.

Importantly, both results operate generically in the size of the maximum number of viable points for the relevant primitive. Such bounds have been well established in the literature due to their connections with coding theory. This means if one can provide a new bound on fuzzy extractor or secure sketch quality this can be directly used in our results.

1.2 Discussion

Avoiding the result Lemma 4 shows the impossibility of efficient constructions for a constant fraction of the family. This means it may be possible to secure all low-entropy distributions of practical interest.

However, new designs or analyses are required. One must use statistical properties beyond fuzzy min-entropy. Demarest, Fuller, and Russell [DFR21] provide a summary of statistical properties in addition to fuzzy min-entropy used in low-entropy computationally secure constructions, such as small, random subsets of bits having high entropy. Simhadri et al. provide a discussion on the current state of biometric cryptosystems and their limited security [SSF19]. Current information-theoretic constructions fall into three categories: 1) requiring high min-entropy, such as the coset construction from Dodis et al. [DORS08] 2) requiring bits of w to be i.i.d. [Mau93, MW96, MTV09, YD10, HMSS12] and 3) inefficient constructions that use the whole probability distribution such as Woodage et al. [WCD+17].

We provide some intuition for why high entropy distributions are easier to secure. First, from the construction perspective if the distribution has at $\mathsf{log}(|B_t|) + \omega(\mathsf{log}(\lambda))$ bits of entropy, one can write down enough bits to uniquely determine the original w from a nearby w' without removing all entropy of w (assuming a perfect error correcting code). Second, from an impossibility perspective, impossibility results (both ours and prior results) require the construction to choose `Viable` points in the construction and have some side information about the distribution to reduce the size of this set. The larger the support of the distribution the harder it is for this side information to reduce the entropy of this set. For example, Fuller, Reyzin, and Smith [FRS20] distributions, $W \in \mathcal{W}^{FRS}$, were lines that overlapped at one point, upper bounding their size.

The two natural directions stemming from this research are 1) can one use natural statistical properties to provide information-theoretic security and 2) can one *compress* inefficient information-theoretic constructions to not require the whole probability distribution of W_z.

Perfect Correctness Our result for fuzzy extractors considers perfect correctness. We do not think this is a fundamental limitation but we briefly explain the issue. As mentioned above, in the case of perfect correctness, one includes a point w in $\mathtt{Viable}_{r,p}$ if it is distance t from any point that produces a different r'. Once one allows imperfect correctness, there is no immediate test for whether a point w should be considered viable. It seems possible that one could argue for a point to be viable when most points around w produce the same key. We were not able to apply the isoperimetric inequality in this setting. If one finds a clean argument for viable points with imperfect correctness, it directly replaces Lemma 8. The rest of our argument then applies. On the other hand, for a secure sketch, one can easily bound the size of the set of points

$$\left\{ w \middle| \frac{\{w' | \mathsf{Rec}(w', p) = w \wedge \mathsf{dis}(w, w')\}}{\{w' | \mathsf{dis}(w, w')\}} \geq 1 - \delta \right\},$$

this set forms a Shannon error correcting code [FRS20, Lemma 7.3]. This is the viable set in the secure sketch case.

Differences from prior work Our fuzzy extractor result requires that $|r| = \omega(\mathsf{log}(n))$. This is in contrast to Fuller, Reyzin, and Smith [FRS20] who showed an impossibility for a key length of 3.[2] This change comes because advice can supply a lot of information about a small number of points in W_z, allowing Gen to ensure that some $\mathtt{Viable}_{r,p}$ are nonempty. Furthermore, all bounds are weaker than those of Fuller, Reyzin, and Smith. The core of the difference is that in $\mathcal{W}^{FRS}$ the adversary received entirely new information by the leftover hash lemma [HILL93, BDK+11]. In our setting, we argue about the expected number of points in the support of W_z that are included in the `Viable` region.

Our secure sketch result also considers an object that retains smooth conditional min-entropy [RW05]. Smooth conditional min-entropy is the necessary and sufficient condition for privacy amplification using a randomness extractor.

Organization The rest of this work is organized as follows, Section 2 covers preliminaries including the relevant definitions of fuzzy extractors and secure sketches. Section 3 presents the negative result for fuzzy extractors including a proof outline, and Section 4 presents the negative result for secure sketches.

[2] Our result for secure sketches requires them to retain at least 5 bits of min-entropy about the input in comparison with [FMR20] which required the sketch to maintain 3 bits of entropy.

2 Preliminaries

For distributions X, Y over the same discrete domain χ,

$$\Delta(X,Y) \stackrel{def}{=} \frac{1}{2}\sum_{x\in\chi} |\Pr[X=x] - \Pr[Y=y]| \,.$$

For a metric space $(\mathcal{M}, \mathsf{dis})$ let $B_t(x) = \{y | \mathsf{dis}(x,y) \leq t\}$. If the size of $B_t(x)$ is the same for all points x we use $|B_t|$ to denote this quantity. This is the case for the Hamming metric. All logarithms are base 2. For a set X, let U_X denote the uniform distribution over that set. For a distribution W, let $\mathtt{Supp}(W)$ denote the support of the distribution.

Indexing and sampling from a family of distributions This work considers the possibility of constructing fuzzy extractors from a finite family of distributions that we will call $\mathcal{W}$. Throughout, we will need the ability to describe a particular value in this family. We let $\mathcal{Z}$ be an index for the family $\mathcal{W}$. Each string $z \in \mathcal{Z}$ describes a distribution $W_z \in \mathcal{W}$. We use Z to describe the uniform distribution $U_{\mathcal{Z}}$, that is $W_Z \stackrel{d}{=} U_{\mathcal{W}}$. We use $w \leftarrow W_z$ to denote a sample from W_z where $w \in \{0,1\}^n$. Where appropriate we use $w \leftarrow W_Z$ to denote the two-stage process of sampling $z \leftarrow Z$ and then sampling $w \leftarrow W_z$.

2.1 Notions of Entropy

For a random variable X whose outcomes are in $\{0,1\}$, let $\Pr[X=1] = p$. The **binary entropy** of X is $h_2(X) := H(X) = -p \cdot \mathsf{log}(p) - (1-p)\cdot \mathsf{log}(1-p)$. For a discrete random variable X, **min-entropy** is $\mathrm{H}_\infty(X) := -\mathsf{log}(\max_{x_i} \Pr(X = x_i))$.

Definition 1 (Average Min Entropy)**.** *Let X be a discrete random variable and let Y be a random variable. The* average min-entropy *of $X|Y$ is*

$$\tilde{\mathrm{H}}_\infty(X|Y) := -\mathsf{log}\left(\mathop{\mathbb{E}}_{y\leftarrow Y}\left[\max_x \Pr[X=x \mid Y=y]\right]\right).$$

Definition 2 (Smooth Conditional Min Entropy)**.** *The* smooth conditional min entropy*, denoted $\tilde{\mathrm{H}}^\epsilon_\infty(X|Y)$ for two random variables X and Y is*

$$\tilde{\mathrm{H}}^\epsilon_\infty(X|Y) := \max_{(X',Y')|\Delta((X',Y'),(X,Y))\leq\epsilon} \tilde{\mathrm{H}}_\infty(X'|Y') \,.$$

The above definition combines prior definitions [RW05, DORS08, Rey11, GW11]. Renner and Wolf's definition considers the worst case Y. We focus on the average case Y. We also replace the condition when considering statistical distance similar to Gentry and Wichs [GW11].

2.2 Fuzzy Min-Entropy and Hamming Balls

Definition 3 (Fuzzy min-entropy [FRS20])**.** *For a distribution W and a distance parameter t, the fuzzy min-entropy of W, denoted $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$ is*

$$\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) := -\mathsf{log}\left(\max_{w^*}\left(\sum_w \Pr[W=w | \mathsf{dis}(w,w^*)\leq t]\right)\right).$$

Proposition 3. *For all distributions W over a metric space $(\mathcal{M}, \mathsf{dis})$, $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) \leq \mathsf{log}(|\mathcal{M}|) - \mathsf{log}(|B_t|)$.*

For $\mathcal{M} = \{0,1\}^n$ and the binary Hamming metric, Using Ash [Ash65, Lemma 4.7.2, Equation 4.7.5, p. 115] one has

$$nh_2(t/n) - 1/2\mathsf{log}(n) - 1/2 \leq \mathsf{log}(|B_t|) \leq nh_2(t/n). \tag{1}$$

and thus,

$$\mathrm{H}^{\mathbf{fuzz}}_{t,\infty}(W) \leq \mathsf{log}(|\mathcal{M}|) - \mathsf{log}(|B_t|) \leq n\left(1 - h_2\left(\frac{t}{n}\right)\right) + \frac{\mathsf{log}(n)}{2} + 1/2.$$

We now introduce the notion of β-density which measures the size of a Hamming ball in comparison to the whole metric space.

Definition 4. *Let $(\mathcal{M}, \mathsf{dis})$ be a metric space where the size of balls is center independent. The β density is*

$$\beta := \mathsf{log}\left(\frac{|\mathcal{M}| - |B_t|}{|B_t|}\right)$$

Claim 1. *For $n, t \in \mathbb{Z}^+$ with $t < n/2$ for the Hamming metric over $\{0,1\}^n$*

$$\beta \geq n\left(1 - h_2\left(\frac{t}{n}\right)\right) - 1.$$

Proof. By Equation 1 one has:

$$\beta \geq \mathsf{log}\left(2^{n\left(1-h_2\left(\frac{t}{n}\right)\right)} - 1\right) \geq \mathsf{log}\left(2^{n\left(1-h_2\left(\frac{t}{n}\right)\right)-1}\right) \geq n(1 - h_2(t/n)) - 1.$$

□

2.3 Fuzzy Extractors and Secure Sketches

Definition 5 (Secure Sketch [DORS08])**.** *For metric space $(\mathcal{M}, \mathsf{dis})$ and distribution W_z, a $(\mathcal{M}, \tilde{m}, t, \epsilon, \ell, \delta)$-*secure sketch *is a pair of algorithms $(\mathsf{SS}_z, \mathsf{Rec}_z)$ with the following properties*

1. ***Correctness*** *For all w, w' such that $\mathsf{dis}(w, w') \leq t$, then $\Pr_{ss \leftarrow \mathsf{SS}(w)}[\mathsf{Rec}_z(w', ss) = w] \geq 1 - \delta$.*
2. ***Security*** $\tilde{\mathrm{H}}^{\epsilon}_{\infty}(W_z|\mathsf{SS}_z(W_z)) \geq \tilde{m}$.
3. ***Space Bounded*** *The circuits SS_z and Rec_z require at most ℓ bits to describe. That is, $|\mathsf{SS}_z| + |\mathsf{Rec}_z| \leq \ell$.*

The use of smooth min-entropy In the above definition, the secure sketch retains smooth conditional min-entropy of W_z. Many definitions consider $\epsilon = 0$ or average min-entropy. The ϵ-smooth min-entropy can be used to bound the average min-entropy [DORS08, Appendix B].

Definition 6 (Fuzzy Extractor [DORS08])**.** *For metric space $(\mathcal{M}, \mathsf{dis})$ and probability distribution W_z, a $(\mathcal{M}, \kappa, t, \epsilon, \ell)$-*fuzzy extractor *is a pair of algorithms $(\mathsf{Gen}_z, \mathsf{Rep}_z)$ with the following properties*

1. ***Correctness*** *For all w, w' such that $\mathsf{dis}(w, w') \leq t$, then $\Pr_{r,p \leftarrow \mathsf{Gen}(w)}[\mathsf{Rep}(w', p) = r] = 1$.*
2. ***Security*** *Let $R, P \leftarrow \mathsf{Gen}_z(W_z)$ and U_κ be a uniformly distributed random variable over $\{0,1\}^\kappa$, $\Delta((R, P), (U_\kappa, P)) \leq \epsilon$.*
3. ***Space Bounded*** *The circuits Gen_z and Rep_z require ℓ bits to describe. That is, $|\mathsf{Gen}| + |\mathsf{Rec}| \leq \ell$.*

We now define fuzzy extractors and secure sketches with advice. This is an intermediate definition that will be used in proofs throughout. Let $\mathcal{W}_{n,k}$ be a family of distributions. As we show in Lemmas 4 and 5, the impossibility of building a fuzzy extractor (resp. secure sketch) with advice for the uniform distribution of $\mathcal{W}_{n,k,Z}$ from family $\mathcal{W}_{n,k}$ implies the impossibility of building a fuzzy extractor (resp. secure sketch) for a constant fraction of $W_z \in \mathcal{W}_{n,k}$.

Definition 7 (Secure Sketch with distributional advice)**.** *Let $\mathcal{W}$ be a family of distributions indexed by z and let $\mathcal{Z}$ denote the set of such z. Let Z be a random variable describing the uniform selection of a $W_z \in \mathcal{W}$. For metric space $(\{0,1\}^n, \mathsf{dis})$, a $(\{0,1\}^n, \mathcal{W}, \tilde{m}, t, \epsilon, \ell, \delta)$*-secure sketch with distributional advice *is a triplet of algorithms $(\mathsf{Gen}, \mathsf{Rep}, \mathsf{Advice})$ with the following properties:*

1. ***Correctness*** *For all w, w' such that $\mathsf{dis}(w, w') \leq t$, let $\Pr_{ss \leftarrow \mathsf{SS}(w)}[\mathsf{Rec}(w', ss) = w] \geq 1 - \delta$.*
2. ***Security*** *Let Advice be a function with output in $\{0,1\}^\ell$. For all distributions $W_z \in \mathcal{W}$, define $\mathsf{advice}_z := \mathsf{Advice}(z)$ and let $SS \leftarrow \mathsf{SS}(W_z, \mathsf{advice}_z)$. Then, $\mathbb{E}_{z \leftarrow Z}[\tilde{\mathrm{H}}^{\epsilon}_{\infty}(W_z | SS, Z = z)] \geq \tilde{m}$.*

Definition 8 (Fuzzy Extractor with distributional advice)**.** *Let $\mathcal{W}$ be a family of distributions indexed by z. Let Z be a random variable describing the uniform selection of a $W_z \in \mathcal{W}$. For metric space $(\{0,1\}^n, \mathsf{dis})$, a $(\{0,1\}^n, \mathcal{W}, \kappa, t, \epsilon, \ell)$*-fuzzy extractor with distributional advice *is a triplet of algorithms $(\mathsf{Gen}, \mathsf{Rep}, \mathsf{Advice})$ with the following properties:*

1. ***Correctness*** *For all w, w' such that $\mathsf{dis}(w, w') \leq t$, $\Pr_{(r,p) \leftarrow \mathsf{Gen}(w)}[\mathsf{Rep}(w', p) = r] = 1$.*
2. ***Security*** *Let Advice be a function with output in $\{0,1\}^\ell$. For a distribution $W_z \in \mathcal{W}$, define $\mathsf{advice}_z := \mathsf{Advice}(z)$, let $(R_z, P_z) \leftarrow \mathsf{Gen}(W, \mathsf{advice}_z)$ and U_κ be a uniformly distributed random variable over $\{0,1\}^\kappa$ it holds that*

$$\Delta((R_Z, P_Z, Z), (U_\kappa, P_Z, Z)) = \mathop{\mathbb{E}}_{z \leftarrow Z} \Delta((R_z, P_z, z), (U_\kappa, P_z, z) \leq \epsilon.$$

Lemma 4. *Let $\mathcal{W}$ be a distribution family indexed by set $\mathcal{Z}$. Suppose that no $(\mathcal{M}, \mathcal{W}, \kappa, t, \epsilon, \ell)$-fuzzy extractor with distributional advice exists. For all families $\mathcal{W}' \subseteq \mathcal{W}$ indexed by $\mathcal{Z}' \subseteq \mathcal{Z}$ where $|\mathcal{Z}'|/|\mathcal{Z}| \leq 1 - \zeta$. There is some $z \in \mathcal{Z}'$ such that there is no $(\{0,1\}^n, \kappa, t, (\epsilon - \zeta))/(1 - \zeta), \ell)$ fuzzy extractor $(\mathsf{Gen}_z, \mathsf{Rep}_z)$.*

Proof of Lemma 4. We proceed by contrapositive. Let $\mathcal{W}'$ be some subset of $\mathcal{W}$ with relative size at least $1 - \zeta$ where for every $W_z \in \mathcal{W}'$ there exists an $(\{0,1\}^n, \kappa, t, (\epsilon - \zeta)/(1 - \zeta), \ell)$-fuzzy extractor. We denote these algorithms by $(\mathsf{Gen}_z, \mathsf{Rep}_z)$ respectively. We now describe how to build the fuzzy extractor $(\mathsf{Gen}, \mathsf{Rep}, \mathsf{advice})$ with distributional advice. Let

$$\mathsf{advice}(z) = \begin{cases} (\mathsf{Gen}_z, \mathsf{Rep}_z) & z \in \mathcal{Z}' \\ \perp & \text{otherwise.} \end{cases}$$

In both cases, $\mathsf{advice}(z)$ has length at most ℓ. Then define $\mathsf{Gen}(x, C)$ as follows: if $C = \perp$ sample a random key r output (r, r), otherwise interpret C as two circuits $\mathsf{Gen}', \mathsf{Rep}'$ and output $\mathsf{Gen}'(x)$. Define $\mathsf{Rep}(x, p, C)$ interpret C if $C = \perp$ output p, otherwise parse C as two circuits $\mathsf{Gen}', \mathsf{Rep}'$ and output $\mathsf{Rep}'(x', p)$. Then

$$\begin{aligned} \Delta((R_Z, P_Z, Z), (U_\kappa, P_Z, Z)) &= \Delta((R_Z, P_Z, Z), (U_\kappa, P_Z, Z) | Z \in \mathcal{Z}') \Pr[Z \in \mathcal{Z}'] \\ &\quad + \Delta((R_Z, P_Z, Z), (U_\kappa, P_Z, Z) | Z \notin \mathcal{Z}') \Pr[Z \notin \mathcal{Z}'] \\ &\leq \frac{\epsilon - \zeta}{1 - \zeta} * (1 - \zeta) + 1 * \zeta = \epsilon. \end{aligned}$$

Recall that Z denotes the uniform random variable over the set $\mathcal{Z}$. $(\mathsf{Gen}, \mathsf{Rep}, \mathsf{advice})$ is a $(\mathcal{M}, \mathcal{W}, \kappa, t, \epsilon, \ell)$ fuzzy extractor with distributional advice. □

Interpretation In the setting when $\epsilon = \Theta(1)$ then setting $\zeta = \epsilon/2$ implies that for all subsets $\mathcal{W}' \subseteq \mathcal{W}$ where $\Pr[Z \in Z'] \geq 1 - \epsilon/2 = 1 - \Theta(1)$ there is no $(\{0,1\}^n, \kappa, t, \epsilon/(2 - \epsilon), \ell)$-fuzzy extractor for some element of $\mathcal{W}'$. This shows that at least $\epsilon/2 = \Theta(1)$ fraction of elements in $\mathcal{W}$ do not have $(\{0,1\}^n, \kappa, t, \epsilon/(2 - \epsilon), \ell)$-fuzzy extractors.

Lemma 5. *Let $\mathcal{W}$ be a distribution family indexed by $\mathcal{Z}$ and suppose that no $(\{0,1\}^n, \mathcal{W}, \tilde{m}, t, \epsilon, \delta, \ell)$-secure sketch with distributional advice exists. For all families $\mathcal{W}' \subseteq \mathcal{W}$ indexed by $\mathcal{Z}' \subseteq \mathcal{Z}$ where $|\mathcal{Z}'|/|\mathcal{Z}| \geq 1-2^{-\zeta}$ there is some $z' \in \mathcal{Z}'$ for which no $(\{0,1\}^n, \tilde{m}+1, t, \epsilon, \delta, \ell)$-secure sketch $(\mathsf{SS}_{z'}, \mathsf{Rec}_{z'})$ exists if $\zeta \geq \tilde{m}+1$.*

Proof. The proof of Lemma 5 follows the structure of the proof of Lemma 4. That is,

$$\mathsf{advice}(z) = \begin{cases} (\mathsf{Gen}_z, \mathsf{Rep}_z) & z \in \mathcal{Z}' \\ \perp & \text{otherwise.} \end{cases}$$

$$\mathsf{SS}(w, C) = \begin{cases} (w, w) & C = \perp \\ \mathsf{SS}'(w) & C = \mathsf{SS}', \mathsf{Rec}'. \end{cases}$$

$$\mathsf{Rec}(w', p, C) = \begin{cases} p & C = \perp \\ \mathsf{Rec}'(w', p) & C = \mathsf{SS}', \mathsf{Rec}'. \end{cases}$$

Then consider the following equation for computing the remaining smooth conditional min-entropy.

$$\begin{aligned} \mathop{\mathbb{E}}_{z \leftarrow Z}[\tilde{\mathrm{H}}^{\epsilon}_{\infty}(W_z | \mathsf{SS}(W_z), z)] &= -\mathsf{log}\Big(\Pr[Z \in \mathcal{Z}'] \mathop{\mathbb{E}}_{Z} 2^{-\tilde{\mathrm{H}}^{\epsilon}_{\infty}(W_Z | ss, Z \in \mathcal{Z}')} + \Pr[Z \notin \mathcal{Z}'] \mathop{\mathbb{E}}_{Z}\Big[2^{-\tilde{\mathrm{H}}^{\epsilon}_{\infty}(W_Z | ss, Z \notin \mathcal{Z}'])}\Big]\Big) \\ &\leq -\mathsf{log}\big(1 \cdot 2^{-\tilde{m}} + 2^{-\zeta} \cdot 1\big) \leq \min\{\tilde{m}+1, \zeta\} - 1 \leq \tilde{m}. \end{aligned}$$

□

Interpretation Setting $\zeta = \max\{\tilde{m}, 1\}$ shows that at least $2^{-\tilde{m}}$ of the distributions have no secure sketch. Later in this work, we consider $\tilde{m} = \Theta(1)$ which suffices to that show that a constant fraction of distributions have no secure sketch.

3 Fuzzy Extractors

Before introducing our main theorem we introduce two families of distributions that are used for our negative results. Consider the following index sets:

$$\begin{aligned} \mathcal{Z}_{n,k} &= \left\{ z \subseteq \{0,1\}^n \middle| |z| = 2^k \right\}, \\ \mathcal{Z}_{n,k,t,\gamma} &= \left\{ z \subseteq \{0,1\}^n \middle| |z| = 2^k, \mathrm{H}^{\mathbf{fuzz}}_{t,\infty}(U_z) \geq \gamma \right\}. \end{aligned}$$

That is, $\mathcal{Z}_{n,k,t,\gamma}$ and $\mathcal{Z}_{n,k}$ are sets of sets. Throughout, we use $\alpha := \mathsf{log}\left(\binom{2^n}{2^k}\right) = |\mathcal{Z}_{n,k}|$. In either case, one can specify the particular choice of z by listing the 2^k points. We use the notation

$$\begin{aligned} \mathcal{W}_{n,k} &= \{W_z | z \in \mathcal{Z}_{n,k} \wedge \forall w \in W_z, \Pr[W_z = w] = 1/2^k\}, \\ \mathcal{W}_{n,k,t,\gamma} &= \{W_z | z \in \mathcal{Z}_{n,k,t,\gamma} \wedge \forall w \in W_z, \Pr[W_z = w] = 1/2^k\}. \end{aligned}$$

$\mathcal{W}_{n,k}$ is the family of uniform distributions W_z over a set $z, |z| = 2^k$. $\mathcal{W}_{n,k,t,\gamma}$ adds the requirement that $\mathrm{H}^{\mathbf{fuzz}}_{t,\infty}(W_z) \geq \gamma$. We use $Z_{n,k}$ to denote the uniform distribution over $\mathcal{Z}_{n,k}$ and $\mathcal{W}_{n,k,Z}$ to denote the uniform choice of some W_z where $z \leftarrow Z_{n,k}$, similarly, we use $\mathcal{Z}_{n,k,t,\gamma}$ to denote the uniform distribution over $Z_{n,k,t,\gamma}$ and $\mathcal{W}_{n,k,t,\gamma,Z}$ to denote the uniform choice of some W_z where $z \leftarrow \mathcal{W}_{n,k,t,\gamma,Z}$. For $z = (z_1, ..., z_{2^k})$ let $w_{z_1}, ..., w_{z_{2^k}}$ denote the support of W_z. We also summarize notation in Table 1.

Theorem 6. *Let $\gamma, n, \kappa, t, \ell, \nu, \gamma \in \mathbb{Z}^+$ be parameters where $t < n/2$. Denote $\alpha := \mathsf{log}\left(\binom{2^n}{2^k}\right)$, $\mu := n \cdot h_2\left(\frac{1}{2} - \frac{t}{n}\right)$. For a $1/4$ of the values $z \in \mathcal{Z}_{n,k,t,\gamma}$ there is no $(\{0,1\}^n, W_z, \kappa, t, \epsilon, \ell)$-fuzzy extractor for*

Notation	Meaning
k	log size support of input distribution
ℓ	Length of advice and circuit size of fuzzy extractor
$\tilde{m}$	Residual min-entropy of fuzzy extractor conditioned on helper
n	Dimension of input points
t	Distance for correction
B_t	Hamming Ball of radius t
U	Uniform distribution
ϵ	Statistical Distance Parameter for fuzzy extractor, Smoothness parameter for min-entropy of secure sketch
$\alpha = \mathsf{log}\left(\binom{2^n}{2^k}\right)$	Log size of number of distributions in $\mathcal{W}_{n,k}$.
β	Ratio of Metric Space Size to Size of Hamming Ball (Def 4)
γ	lower bound on fuzzy min-entropy of distributions
κ	key length of fuzzy extractor
ν	Number of points that adversary "describes" in advice
$\mu = n \cdot h_2\left(\frac{1}{2} - \frac{t}{n}\right)$	Bound on log of maximum number of viable points (Lem 8)
$\mathcal{W}_{n,k}$	Set of all distributions with k points
$\mathcal{Z}_{n,k}$	Set of indices for $\mathcal{W}_{n,k}$
$Z_{n,k}$	Uniform choice of $z \leftarrow \mathcal{Z}_{n,k}$
$\mathcal{W}_{n,k,Z}$	Uniform choice of W_z from $\mathcal{W}_{n,k}$
$\mathcal{W}_{n,k,t,\gamma}$	Restriction of $\mathcal{W}_{n,k}$ to distributions W_z with $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W_z) \geq \gamma$
$\mathcal{Z}_{n,k,t,\gamma}$	Set of indices for $\mathcal{W}_{n,k,t,\gamma}$
$Z_{n,k,t,\gamma}$	Uniform choice of z from $\mathcal{Z}_{n,k,t,\gamma}$
$\mathcal{W}_{n,k,t,\gamma,Z}$	Uniform selection of W_z from $\mathcal{W}_{n,k,t,\gamma}$

Table 1: Summary of notation.

$\epsilon < 1/3 - (\epsilon_1 + \epsilon_2 + \epsilon_3)/3$. *For*

$$\begin{aligned}
\mathsf{log}(\epsilon_1) &:= -\left(\kappa + \frac{\alpha - \ell}{2^k} - \mu - 2k + \mathsf{log}(\nu)\right), \\
\epsilon_2 &:= \frac{\nu + 1}{2^{\kappa - 1}}, \\
\epsilon_3 &:= \left(e 2^{\gamma - \beta}\right)^{2^{k-\gamma}} 2^{n-1}.
\end{aligned}$$

Lemma 4 states that to show the hardness of building an efficient fuzzy extractor (Definition 6) for a constant fraction of $W_z \in \mathcal{W}_{n,k,t,\gamma}$ it suffices to show the hardness of building a fuzzy extractor with distributional advice (Definition 8) for the family $\mathcal{W}_{n,k,t,\gamma}$. The proof of Theorem 6 focuses on the distributional advice setting using the following structure:

1. We show that few elements in $\mathcal{Z}_{n,k}$ are not in $\mathcal{Z}_{n,k,t,\gamma}$. Lemma 7 shows that the statistical distance between $\mathcal{W}_{n,k,Z}$ and $\mathcal{W}_{n,k,t,\gamma,Z}$ is small. This shows that the hardness of building a fuzzy extractor with distributional advice for the family $\mathcal{W}_{n,k}$ implies hardness for $\mathcal{W}_{n,k,t,\gamma}$. For the remainder of the proof, we consider $\mathcal{W}_{n,k}$.
2. Lemma 9 shows one cannot build a fuzzy extractor with distribution advice for $\mathcal{W}_{n,k}$. Proving this requires several steps

(a) Lemma 8 which bounds the number of "viable" points for most public values p. This lemma bounds the total number of points and holds even if $\mathsf{Gen}, \mathsf{Rep}$ have access to an arbitrary advice string. We switch to considering a fixed value of advice. At the end of the proof we average across the distribution of advice.

(b) Claim 3 shows the majority of the support of $Z_{n,k}|\mathsf{Advice} = \mathsf{advice}$ is difficult to predict and thus unlikely to be included in the set of viable points. We call such points **Hard Points**. There are some points that the adversary has a large amount of information on that we call **Free Points**.

(c) Corollary 10 puts together the above two steps to show that the adversary includes a small number of points from the particular distribution W_z in the viable set.

Since the construction cannot align the viable points with the distribution there exists a distinguisher that can distinguish a uniform triple from a key triple.

The rest of this section is organized as follows:

Section 3.1 We present and prove Lemmas 7 and 9,

Section 3.2 We present the proof of Theorem 6 combining Lemmas 7 and 9, and

Section 3.3 We present our preferred setting of parameters.

3.1 Main Technical Lemmas and Proofs

Lemma 7. *Fix $n, t, k, \gamma \in \mathbb{Z}^+$ where $t < n/2$, then $\Delta(\mathcal{W}_{n,k,Z}, \mathcal{W}_{n,k,t,\gamma,Z}) \leq \left(e2^{\gamma-\beta}\right)^{2^{k-\gamma}} 2^n$.*

Proof of Lemma 7. We first show the fraction of items in $\mathcal{Z}_{n,k}$ that are not in $\mathcal{Z}_{n,k,t,\gamma}$ is at most $\left(e2^{\gamma-\beta}\right)^{2^{k-\gamma}} 2^n$. A string $z \in \mathcal{Z}_{n,k} \setminus \mathcal{Z}_{n,k,t,\gamma}$ if and only if there exists some point y such that there are at least $2^{-\gamma}2^k = 2^{k-\gamma}$ points within distance t of y. Fix some arbitrary y^*. The point y^* defines a set A_{y^*} of all points within distance t and note that $|A_{y^*}| = |B_t|$. We now show that the probability that a value in $z \leftarrow Z_{n,k}$ has a large intersection with A_{y^*} is small.

Claim 2. *Let $n, k, a \in \mathbb{Z}^+$ where $\mathsf{log}(a) < k$. Let $a^* = a/(1 - a2^{-k})$. Let $Z_{n,k}$ be the uniform distribution over $\mathcal{Z}_{n,k}$. Let A be a fixed subset of size $a \cdot 2^{n-k}$. Then $\mathbb{E}[|Z_{n,k} \cap A|] = a$ and any $\zeta > 0$,*

$$\Pr[|Z_{n,k} \cap A| \geq a^*(1+\zeta)] \leq \left[\frac{e^\zeta}{(1+\zeta)^{1+\zeta}}\right]^{a^*}.$$

Proof. For the purposes of bookkeeping, arrange the elements of A in an arbitrary order and note that $|A| = a2^{n-k} < 2^k 2^{n-k} = 2^n$ so $A \subset \{0,1\}^n$, and let

$$X_1, \ldots, X_{a2^{n-k}}$$

be indicator random variables so that $X_i = 1$ if and only if the ith element of A lies in $Z_{n,k}$. Note that for any individual i, $\Pr[X_i = 1] = a2^{n-k}/2^n = a2^{-k}$ and thus $\mathbb{E}[|Z_{n,k} \cap A|] = \sum_i \mathbb{E}[X_i] = 2^k \mathbb{E}[X_i] = 2^k(a2^{-k}) = a$ by linearity of expectation. Observe that under any conditioning on the variables $X_1, \ldots, X_t$,

$$\Pr[X_{t+1} = 1] \leq \frac{2^k}{2^n - a2^{n-k}} = \frac{2^k}{2^n(1 - a2^{-k})}.$$

Let Y_i be a sequence of i.i.d. random variables (with the same index set) for which

$$\Pr[Y_i = 1] = \frac{2^k}{2^n(1 - a2^{-k})}.$$

It follows that the random variable $\sum_i X_i$ is stochastically dominated by the random variable $\sum_i Y_i$. Observe that $\mathbb{E}[\sum Y_i] = a^*$. Applying a standard Chernoff upper tail bound to the Y_i then yields the result. This completes the proof of Claim 2. □

We now continue using the notation of Claim 2, let

$$a^* = \frac{2^k |B_t|}{2^n - |B_t|} = 2^{k-\beta}.$$

Fix the value of ζ such that

$$1 + \zeta = \frac{2^{-\gamma}(2^n - |B_t|)}{|B_t|} \geq 2^{\beta-\gamma}.$$

then the probability $Z_{n,k}$ intersects with A_{y^*} in at least $2^{k-\gamma}$ places is at most

$$\begin{aligned}
\Pr[|Z_{n,k} \cap A_{y^*}| \geq a^*(1+\zeta)] \leq \\
\Pr[|Z_{n,k} \cap A_{y^*}| \geq 2^{k-\gamma}] &\leq \left(\left(e 2^{\gamma-\beta}\right)^{2^{\beta-\gamma}} \right)^{2^{k-\beta}} \\
&= \left(e 2^{\gamma-\beta}\right)^{2^{k-\gamma}}
\end{aligned}$$

Now we consider a union bound across all y^*. That is

$$\Pr_{z \leftarrow Z_{n,k}} [\mathrm{H}^{\mathbf{fuzz}}_{t,\infty}(W_z) \geq \gamma] = \left(e 2^{\gamma-\beta}\right)^{2^{k-\gamma}} 2^n.$$

The difference between $\mathcal{W}_{n,k,Z}$ and $\mathcal{W}_{n,k,t,\gamma,Z}$ is exactly the set of $z \in \mathcal{Z}_{n,k}$ that are not present in $\mathcal{Z}_{n,k,t,\gamma}$. This probability mass is uniformly distributed in both cases, thus

$$\Delta(\mathcal{W}_{n,k,Z}, \mathcal{W}_{n,k,t,\gamma,Z}) \leq \left(e 2^{\gamma-\beta}\right)^{2^{k-\gamma}} 2^n.$$

completing the proof of Lemma 7. □

Fuller, Reyzin, and Smith [FRS16, FRS20] bound the number of points that could have produced the output of a fuzzy extractor. Such points are called *viable*. We present a stronger version of their lemma that is contained in their proof. The difference is we bound the union of viable points across different values of key while they only bound the size of a single key corresponding to the true point. Their argument is purely geometric, so it also applies to fuzzy extractors with distributional advice. We state the stronger version of [FRS20, Lemma 5.2].

Lemma 8. *For $n \in \mathbb{Z}^+$, suppose $\mathcal{M}$ is $\{0,1\}^n$ with the Hamming Metric and $\kappa \geq 2$, $0 \leq t \leq n/2$, $\epsilon > 0, \ell \in \mathbb{Z}^+$. Suppose $(\mathsf{Gen}, \mathsf{Rep})$ is a $(\mathcal{M}, \mathcal{W}, \kappa, t, \ell, \epsilon)$-fuzzy extractor with distributional advice for some distribution family $\mathcal{W}$ over $\mathcal{M}$. For any fixed p, for any value $\mathsf{advice} \in \{0,1\}^\ell$, there is a set $\mathsf{GoodKey}_p \subseteq \{0,1\}^\kappa$ of size at least $2^{\kappa-1}$ such that,*

$$\mu := \sum_{\mathsf{key} \in \mathsf{GoodKey}_p} (\log(|\{v \in \mathcal{M} | (\mathsf{key}, p) \in \mathsf{supp}(\mathsf{Gen}(v, \mathsf{advice}))\}|)) \leq n \cdot h_2\left(\frac{1}{2} - \frac{t}{n}\right).$$

We now present our main technical lemma and its proof.

Lemma 9. *Suppose $\mathcal{M}$ is $\{0,1\}^n$ with the Hamming Metric, $\kappa \geq 2$, $0 \leq t \leq n/2$, $\epsilon > 0, \ell \in \mathbb{Z}^+$. For the family $\mathcal{W}_{n,k}$ there is no $(\mathsf{Gen}, \mathsf{Rep})$ that is a $(\mathcal{M}, \mathcal{W}_{n,k}, \kappa, t, \ell, \epsilon)$-fuzzy extractor with distributional advice for*

$$\epsilon < 1/2 - (\epsilon_1 + \epsilon_2)$$

where

$$\epsilon_1 = 2^{-\kappa-\frac{\alpha-\ell}{2^k}+1+\mu+2k-\log(\nu)},$$
$$\epsilon_2 = \frac{\nu+1}{2^\kappa},$$
$$\mu \leq n \cdot h_2\left(\frac{1}{2}-\frac{t}{n}\right),$$
$$\alpha = \log\left(\binom{2^n}{2^k}\right).$$

Furthermore, there exists an algorithm $\mathcal{D}$ that always outputs 1 when given samples of the form r, p, z that are correctly generated by the fuzzy extractor.

Proof of Lemma 9. Before proceeding we introduce some additional notation. Let $(\mathsf{Gen}, \mathsf{Rep}, \mathsf{Advice})$ be a fuzzy extractor with distributional advice. Let a be some string. For a tuple (v, p, r, a) define

$$\texttt{Viable}(v, p, r, a) = \begin{cases} 1 & \Pr[\mathsf{Gen}(v, a) = (r, p)] > 0 \\ 0 & \text{otherwise} \end{cases}.$$

Recall that $|\mathcal{Z}_{n,k}| = 2^\alpha = \binom{2^n}{2^k}$. Let A be a random variable denoting $\mathsf{Advice}(Z_{n,k})$. Thus,

$$\mathrm{H}_\infty(Z_{n,k}) = \alpha,$$
$$\tilde{\mathrm{H}}_\infty(Z_{n,k}|A) \geq \alpha - \ell.$$

Define

$$2^{-\alpha_{a,w}} := \Pr_{z \leftarrow Z_{n,k}|A=a}[w \in \mathtt{Supp}(W_z)],$$
$$2^{-\alpha_a} := 2^{-\tilde{\mathrm{H}}_\infty(Z_{n,k}|A=a)} = \max_{W \subseteq \{0,1\}^n || W| = 2^k} \left(\prod_{w \in W} 2^{-\alpha_{a,w}}\right).$$

Note that

$$-\log\left(\mathop{\mathbb{E}}_{a \leftarrow A}\left(2^{-\alpha_a}\right)\right) = \tilde{\mathrm{H}}_\infty(Z_{n,k}|A) \leq \alpha - \ell.$$

Define the notation $\texttt{HViable}(w, p, \mathsf{key}, a, \mathcal{E})$:

1. 0 if $w \in \mathcal{E}$ or $\forall z \in \mathtt{Supp}(Z_{n,k}|A = a), w \notin \mathtt{Supp}(W_z)$,
2. $\Pr_{z \leftarrow Z_{n,k}|A=a}\left[\texttt{Viable}(w, p, \mathsf{key}, a) = 1 \wedge w \in \mathtt{Supp}(W_z)\right]$ otherwise.

Claim 3 bounds how much information a fixed a contains about the points in the distribution (we consider the expectation across $a \in A$ after Corollary 10).

Claim 3. *Let μ and $\mathsf{GoodKey}$ be defined as in Lemma 8. Let A be a distribution and let a be a fixed value such that $\mathrm{H}_\infty(\mathcal{W}_{n,k}|A = a) = \alpha_a$. Fix some value w and some value p. Define $\mathsf{GoodKey}_p$ as in Lemma 8. Each value $w^* \in \{0,1\}^n$ defines a set $\mathcal{E}_{a,w^*}$ where $|\mathcal{E}_{a,w^*}| \leq \nu + 1$ such that*

$$\log\left(\sum_{\mathsf{key} \in \mathsf{GoodKey}_p} \texttt{HViable}(w, p, \mathsf{key}, a, \mathcal{E}_{a,w^*})\right) \leq -\frac{\alpha_a}{2^k} + 1 + \mu + k - \log(\nu).$$

Proof of Claim 3. Let A be an random variable over $\{0,1\}^\ell$. Fix some value a. Let $w_{1,a}, ..., w_{2^k,a}$ denote an arbitrary subset of $\{0,1\}^n$ of size 2^k. Then

$$\sum_{i=1}^{2^k} \alpha_{a,w_{i,a}} \leq \alpha_a.$$

Then one has that

$$\mathop{\mathbb{E}}_{i \leftarrow U_{\{0,1\}^k}} \left(\alpha_{a,w_{i,a}}\right) \leq \frac{\alpha_a}{2^k}.$$

We need an elementary lemma which states that not too many of α_{a,w_i} are much larger than $\alpha_a/2^k$:

Claim 4. *Define* $\mathcal{E}_a \subset \{w_{1,a}, ..., w_{2^k,a}\}$ *as*

$$\mathcal{E}_a = \left\{ w \middle| \alpha_{a,w} \geq \frac{\alpha_a}{2^k} - (k - \mathsf{log}(\nu)) \right\}.$$

Then

$$|\mathcal{E}_a| \leq \nu.$$

Proof of Claim 4. By Markov's inequality

$$\Pr_{i \leftarrow U_{\{0,1\}^k}} \left[\Pr_{z \leftarrow Z_{n,k}|A=a} [w_{i,a} \in \mathtt{Supp}(W_z)] \geq \beta \mathop{\mathbb{E}}_i \Pr_{z \leftarrow Z_{n,k}|A=a} [w_{i,a} \in \mathtt{Supp}(W_z)] \right] =$$

$$\Pr_{i \leftarrow U_{\{0,1\}^k}} \left[\Pr_{z \leftarrow Z_{n,k}|A=a} [w_{i,a} \in \mathtt{Supp}(W_z)] \geq \beta 2^{-\alpha_{a,w}} \right] \leq 1/\beta.$$

Setting $\beta = 2^k/\nu$ implies the statement of the Claim. □

We now continue with the proof of Claim 3. By Claim 4 it is true that there exists a set $\mathcal{E}_a \subseteq \{0,1\}^n$ of size at most ν where such that for all $w \notin \mathcal{E}_a$ is true that $\alpha_{a,w} < \alpha_a/2^k + (k - \mathsf{log}(\nu))$. Let w^* denote the point that will be given to Gen that is $(\mathsf{key}, p) \leftarrow \mathsf{Gen}(w^*, a)$. We define the set $\mathcal{E}_{a,w^*} = \mathcal{E}_a \cup \{w^*\}$. Then,

$$\Pr_{z \leftarrow Z_{n,k}|A=a \wedge \mathsf{key} \in \mathsf{GoodKey}, p} [w \in \mathtt{Supp}(W_z) | w \notin \mathcal{E}_{a,w^*}] \geq$$

$$\Pr_{z \leftarrow Z_{n,k}|A=a \wedge w^* \in \mathtt{Supp}(W_z)} [w \in \mathtt{Supp}(W_z) | w \notin \mathcal{E}_{a,w^*}] \geq 2^{-\left(\frac{\alpha_a}{2^k} - k + \mathsf{log}(\nu) - 1\right)}.$$

This is because conditioning on a single bit that $w^* \in \mathtt{Supp}(W_z)$ increases the predictability of an random variable by at most a factor of 2. We now proceed to bounding $\mathtt{HViable}(w, p, \mathsf{key}, a, \mathcal{E}_{a,w^*})$. By Lemma 8 we know that there are at most 2^μ points in $\mathtt{Viable}(w, p, \mathsf{key}, a)$. Thus, by union bound over the set of viable points,

$$\mathtt{HViable}(w, p, \mathsf{key}, a, \mathcal{E}_{a,w^*}) \leq 2^{-\frac{\alpha_a}{2^k} + 1 + \mu + k - \mathsf{log}(\nu)}.$$

This completes the proof of Claim 3. □

Corollary 10. *Let* μ *and* $\mathsf{GoodKey}$ *be defined as in Lemma 8. Fix an arbitrary point* $w^* \in \{0,1\}^n$ *and some* a *and define* $\mathcal{E}_{a,w^*}$ *as in Claim 3. By Claim 3 on average across* $z \leftarrow Z_{n,k}|A = a$ *(by union bound across the points in* W_z*) one has that:*

$$\mathop{\mathbb{E}}_{z \leftarrow Z_{n,k}|A=a} \left| \left\{ w \middle| \begin{array}{r} w \in \mathtt{Supp}(W_z) \\ w \notin \mathcal{E}_{a,w^*} \\ \exists \mathsf{key} \in \mathsf{GoodKey}_p, \mathtt{Viable}(w, p, \mathsf{key}, a) \end{array} \right\} \right| \leq 2^{-\frac{\alpha_a}{2^k} + 1 + \mu + 2k - \mathsf{log}(\nu)}.$$

Then on average across $a \in A$ (using the fact that $\mathbb{E}_{a \leftarrow A}(2^{-\alpha_a}) = 2^{-(\alpha-\ell)}$) one has

$$\mathop{\mathbb{E}}_{a \leftarrow A}\left(\mathop{\mathbb{E}}_{z \leftarrow Z_{n,k}|A=a}\left|\left\{w \;\middle|\; \begin{array}{r} w \in \mathtt{Supp}(W_z) \\ w \notin \mathcal{E}_{a,w^*} \\ \exists \mathsf{key} \in \mathsf{GoodKey}_p, \mathtt{Viable}(w,p,\mathsf{key},a) \end{array}\right\}\right|\right) \leq 2^{-\frac{\alpha-\ell}{2^k}+1+\mu+2k-\log(\nu)}.$$

And finally,

$$\mathop{\mathbb{E}}_{a \leftarrow A}\left(\mathop{\mathbb{E}}_{z \leftarrow Z_{n,k}|A=a}\left|\left\{w \;\middle|\; \begin{array}{r} w \in \mathtt{Supp}(W_z) \\ \exists \mathsf{key} \in \mathsf{GoodKey}_p, \mathtt{Viable}(w,p,\mathsf{key},a) \end{array}\right\}\right|\right) \leq 2^{-\frac{\alpha-\ell}{2^k}+1+\mu+2k-\log(\nu)} + \nu + 1.$$

With Corollary 10 in hand we are ready to prove Lemma 9. Consider the following distinguisher $\mathcal{D}$ for triples of the form r, p, z:

1. If $r \notin \mathsf{GoodKey}_p$ output 1,
2. If $\sum_{w \in \mathtt{Supp}(W_z)} \mathtt{Viable}(w, r, p, \mathsf{advice}(z)) = 0$ output 0,
3. Else output 1.

First note that by perfect correctness it is always the case that when given key, p, z that $\mathcal{D}$ outputs 1. We proceed to bound the probability that $\mathcal{D}$ outputs 1 when given U_κ, p, z. Note that the probability that $\Pr[U_\kappa \in \mathtt{GoodKey}_p] \geq 1/2$ by the definition of $\mathtt{GoodKey}_p$.

We bound the number of parts with at least one point in viable. We begin by assuming that all points in viable are in different values r so the bound on the size of

$$\left\{w \;\middle|\; \begin{array}{r} w \in \mathtt{Supp}(W_z) \\ \exists \mathsf{key} \in \mathsf{GoodKey}_p, \mathtt{Viable}(w,p,\mathsf{key},a) \end{array}\right\}$$

gives a bound on the number of keys for which $\mathcal{D}$ could output 1. By Corollary 10

$$\mathop{\mathbb{E}}_{a \leftarrow A}\left(\mathop{\mathbb{E}}_{z \leftarrow Z_{n,k}|A=a}\left|\left\{w \;\middle|\; \begin{array}{r} w \in \mathtt{Supp}(W_z) \\ \exists \mathsf{key} \in \mathsf{GoodKey}_p, \mathtt{Viable}(w,p,\mathsf{key},a) \end{array}\right\}\right|\right) \leq 2^{-\frac{\alpha-\ell}{2^k}+1+\mu+2k-\log(\nu)} + \nu + 1.$$

Thus, the fraction of non-empty parts in $\mathsf{Goodkey}_p$ on average is at most

$$2^{-\frac{\alpha-\ell}{2^k}+1+\mu+2k-\log(\nu)} + \nu + 1.$$

Thus, the probability that $\mathcal{D}$ outputs 0 when given U_κ, p, z is at least $1/2 - (\epsilon_1 + \epsilon_2)$ where

$$\epsilon_1 := 2^{-\kappa-\frac{\alpha-\ell}{2^k}+2+\mu+2k-\log(\nu)},$$
$$\epsilon_2 := \frac{\nu+1}{2^{\kappa-1}}.$$

This completes the Proof of Lemma 9. □

3.2 Proof of Theorem 6

Proof of Theorem 6. Define the pair of random variables $(R_{Z_{n,k}}, P_{Z_{n,k}}) \leftarrow \mathsf{Gen}(\mathcal{W}_{n,k,Z}, \mathsf{advice}(Z_{n,k}))$. Restating Lemma 9 one has that

$$\Delta((R_{Z_{n,k}}, P_{Z_{n,k}}, Z_{n,k}, (U_\kappa, P_{Z_{n,k}}, Z_{n,k})) \geq 1/2 - (\epsilon_1 + \epsilon_2).$$

Let $\mathcal{D}$ be one distinguisher that always outputs 1 on any value key, p, z for any distribution W_z, then

$$\Pr[\mathcal{D}((R_{Z_{n,k}}, P_{Z_{n,k}}, Z_{n,k})) = 1] = 1$$
$$\Pr[\mathcal{D}(U_\kappa, P_{Z_{n,k}}, Z_{n,k}) = 1] \leq 1/2 + (\epsilon_1 + \epsilon_2).$$

Recall that $\Delta(\mathcal{W}_{n,k,Z}, \mathcal{W}_{n,k,t,\gamma,Z}) \leq \epsilon_3$ by Lemma 7 . Define the pair of random variables

$$(R_{Z_{n,k,t,\gamma}}, P_{Z_{n,k,t,\gamma}}) \leftarrow \mathsf{Gen}(\mathcal{W}_{n,k,t,\gamma,Z}, \mathsf{advice}(Z_{n,k,t,\gamma})).$$

by the information processing lemma it is thus true that

$$\begin{aligned}\Pr[\mathcal{D}(R_{Z_{n,k,t,\gamma}}, P_{Z_{n,k,t,\gamma}}, Z_{n,k,t,\gamma}) = 1] = 1,\\ \Pr[\mathcal{D}(U_n, P_{Z_{n,k,t,\gamma}}, Z_{n,k,t,\gamma}) = 1] \leq 1/2 + (\epsilon_1 + \epsilon_2 + \epsilon_3).\end{aligned}$$

Finally, the theorem follows by application of Lemma 4 with the setting of $\zeta = 1/4$. □

3.3 Analysis of parameters

We separately consider ϵ_1, ϵ_2 and ϵ_3. We refer to these three terms as hard points, free points, and distributional closeness respectively. This is because ϵ_1 describes how much information the advice has about hard points in W_z, ϵ_2 considers a small number of points that are more thoroughly described by advice, and ϵ_3 controls the statistical distance between $\mathcal{W}_{n,k}$ and $\mathcal{W}_{n,k,t,\gamma}$. We consider parameters in order of simplicity.

3.3.1 Free Points - ϵ_2

For ϵ_2 to be negligible it suffices that $\nu/2^{\kappa} = \mathtt{ngl}(\lambda)$. Meaningful security requires $\kappa = \omega(\mathsf{log}(n))$. We set

Condition 1 $\nu = 2^{c_\kappa \kappa}$ for some constant $0 < c_\kappa < 1$ yielding

$$\epsilon_2 = \frac{\nu + 1}{2^{\kappa - 1}} = 2^{(c_\kappa - 1)\kappa - \kappa + 1} = \mathtt{ngl}(n).$$

3.3.2 Distributional Closeness - ϵ_3

Recall that $\epsilon_3 := \left(e2^{\gamma-\beta}\right)^{2^{k-\gamma}} 2^{n+1}$. Consider the following settings:

Condition 2 That $\gamma \leq \beta - \mathsf{log}(2e)$, which implies $2^{\gamma-\beta} \leq \frac{1}{2e}$, and

Condition 3 For constant $0 < c_{|k|} < 1$, we set $k = \gamma + c_{|k|} n$ which implies $2^{k-\gamma} \geq n + 1 + \omega(\mathsf{log}(n))$.

Together, these settings imply that

$$\epsilon_3 = \left(e2^{\gamma-\beta}\right)^{2^{k-\gamma}} 2^{n+1} \leq \left(\frac{1}{2}\right)^{n+\omega(\mathsf{log}(n))+1} 2^{n+1} = 2^{-\omega(\mathsf{log}(n))} = \mathtt{ngl}(n).$$

Discussion By Proposition 3 for any W in $\{0,1\}^n$ it is true that

$$\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) \leq n\left(1 - h_2\left(\frac{t}{n}\right)\right) + \frac{\mathsf{log}(n)}{2} + 1/2.$$

Thus, the additional constraint that

$$\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) := \gamma \leq \beta - \mathsf{log}(2e) \leq n\left(1 - h_2\left(\frac{t}{n}\right)\right) + \frac{\mathsf{log}(n)}{2} + \frac{1}{2} - \mathsf{log}(2e)\,.$$

imposes an additive $\mathsf{log}(2e)$ impact on the maximal fuzzy min-entropy that can be supported.

3.3.3 Hard Points - ϵ_1

We now turn to our analysis of ϵ_1. Recall $\mathsf{log}(\epsilon_1) := -\left(\kappa + \frac{\alpha-\ell}{2^k} - \mu - 2k + \mathsf{log}(\nu)\right)$ and that

$$\begin{aligned}(n/k)^k \leq \binom{n}{k} < ((ne)/k)^k \\ \mu \leq nh_2(1/2 - t/n), \\ \alpha = \mathsf{log}\left(\binom{2^n}{2^k}\right) \geq \mathsf{log}\left(2^{n2^k}/2^{k2^k}\right) = (n-k)2^k,\end{aligned}$$

Recall that $\nu = 2^{c_\kappa \kappa}$. This implies that

$$\begin{aligned}-\mathsf{log}(\epsilon_1) &= \kappa + \frac{\alpha-\ell}{2^k} - \mu - 2k + \mathsf{log}(\nu) \\ &\geq \kappa + \frac{\alpha-\ell}{2^k} - nh_2(1/2 - t/n) - 2k + \mathsf{log}(\nu)\,, \\ &\geq \kappa + \frac{(n-k)2^k - \ell}{2^k} - nh_2(1/2 - t/n) - 2k + c_\kappa \kappa, \\ &\geq (1+c_\kappa)\kappa + \frac{(n-k)2^k - 2k2^k - \ell}{2^k} - nh_2(1/2 - t/n) \\ &> (1+c_\kappa)\kappa + \frac{(n-3k)2^k - \ell}{2^k} - nh_2(1/2 - t/n).\end{aligned}$$

We now focus on parameters when $\psi := \frac{(n-3k)2^k-\ell}{2^k} - nh_2(1/2 - t/n) \geq 0$.

Condition 4 Let $0 < c_\ell < 1$ be a parameter such that $\ell \leq 3c_\ell n 2^k$,

Condition 5 Suppose that

$$\gamma \leq \frac{n(1 - 3c_{|k|} - c_\ell - h_2(1/2 - t/n))}{3}.$$

Then it holds that

$$\begin{aligned}\psi :&= \frac{(n-3k)2^k - \ell}{2^k} - nh_2(1/2 - t/n) \\ &\geq \frac{(n-3k)2^k - c_\ell n 2^k}{2^k} - nh_2(1/2 - t/n) \\ &\geq \frac{n(1 - 3(\gamma/n + c_{|k|}) - c_\ell 2^k)}{2^k} - nh_2(1/2 - t/n) \\ &\geq n(1 - 3c_{|k|} - c_\ell) - 3\gamma - nh_2(1/2 - t/n) \geq 0\end{aligned}$$

which suffices to ensure that

$$\mathsf{log}(\epsilon_1) \leq -\left((1+c_\kappa)\kappa + \psi\right) \leq -(1+c_\kappa)\kappa = -\omega(\mathsf{log}(n)).$$

3.3.4 Overall Parameters

Combining Conditions 2 and 5 one obtains a negligible statistical distance as long as for constants $c_\kappa, c_{|k|}, c_\ell \in (0,1)$ one has:

$$\begin{aligned}
\nu &= 2^{c_\kappa \kappa},\\
k &= \gamma + c_{|k|} n,\\
\ell &\leq 3 c_\ell n 2^k,\\
0 \leq \frac{\gamma}{n} &\leq \min\left\{(1 - h_2(t/n)) + \frac{\mathsf{log}(n) + 1 - 2\mathsf{log}(2e)}{2n}, \frac{1 - 3c_{|k|} - c_\ell - h_2(1/2 - t/n)}{3}\right\}.
\end{aligned}$$

4 Secure Sketches

This section creates an upper bound on the quality of efficient secure sketches. This bound is stronger than Theorem 6 due to the stronger geometry established by the secure sketch correctness requirement.

Theorem 11. *Let $n, t, \ell, \gamma, \nu, \tilde{m} \in \mathbb{Z}^+, \epsilon, \delta \in [0,1]$ be parameters where $t < n/2$ and denote $\mu := (n(1 - h_2(t/n)) + h_2(2\delta))/(1 - 2\delta)$. For a $2^{-\tilde{m}}$ fraction of the distributions in the family $\mathcal{W}_{n,k,t,\gamma}$ (indexed by $z \in \mathcal{Z}_{n,k,t,\gamma}$) there is no $(\{0,1\}^n, W_z, \tilde{m}, t, \epsilon', \delta, \ell)$-secure sketch for*

$$\tilde{m} \geq 2 - \mathsf{log}(1 - 2\epsilon) + 2\max\left\{-\frac{\alpha - \ell}{2^k} + 1 + \mu + 2k - \mathsf{log}(\nu), \mathsf{log}(\nu + 1)\right\}.$$

where $\epsilon' = \epsilon - \left(e 2^{\gamma - \beta}\right)^{2^{k-\gamma}} 2^n$.

Proof of Theorem 11. Good secure sketches are bounded in size as they imply good Shannon error correcting codes [FRS20, Lemma 7.3]. This is true if one considers a secure sketch that retains smooth min-entropy with no loss in parameters because it only relies on the correctness of the secure sketch (not the security property).

Lemma 12. *Let $n, t, \tilde{m} \in \mathbb{Z}^+, \epsilon, \delta \in [0,1]$ be parameters where $t < n/2$ and denote $\mu := (n(1 - h_2(t/n)) + h_2(2\delta))/(1 - 2\delta)$. Let $\mathcal{W}_{n,k}$ be a family indexed by set $\mathcal{Z}_{n,k}$ and let $Z_{n,k}$ denote the uniform distribution over $Z_{n,k}$. Suppose $(\mathsf{SS}, \mathsf{Rec}, \mathsf{Advice})$ is a $(\{0,1\}^n, \mathcal{W}_{n,k}, \tilde{m}, t, \epsilon_{\mathsf{SS}}, \ell, \delta)$-secure sketch with distribution advice. For every $v \in \{0,1\}^n$ and any value $a \in \{0,1\}^\ell$ there exists a set $\mathsf{GoodSketch}_{v,a}$ where $\Pr[\mathsf{SS}(v,a) \in \mathsf{GoodSketch}_{v,a}] \geq 1/2$ and for any fixed ss,*

$$\mu := \mathsf{log}(|\{v \in \{0,1\}^n | ss \in \mathsf{GoodSketch}_{v,a}\}|) \leq \frac{n - \mathsf{log}(|B_t|) + h_2(2\delta)}{1 - 2\delta} \leq \frac{n(1 - h_2(t/n)) + h_2(2\delta)}{1 - 2\delta}.$$

Define $\alpha := \mathrm{H}_\infty(Z_{n,k})$ and note that $\tilde{\mathrm{H}}_\infty(Z_{n,k}|A) \geq \alpha - \ell$ and define $\alpha_a := \tilde{\mathrm{H}}_\infty(Z_{n,k}|A = a)$. Note that

$$\mathop{\mathbb{E}}_{a \leftarrow A}\left(2^{-\alpha_a}\right) = 2^{-\tilde{\mathrm{H}}_\infty(Z_{n,k}|A)} \geq 2^{-(\alpha - \ell)}.$$

For a triplet (v, ss, a) define $\mathtt{Viable}(v, ss, a, z) = 1$ if

1. $\Pr[\mathsf{SS}(v,a) = ss] > 0$,
2. $ss \in \mathsf{GoodSketch}_{v,a}$, and
3. $v \in \mathtt{Supp}(W_z)$.

Otherwise set $\mathtt{Viable}(v, ss, a, z) = 0$. Define $\mathtt{Viable}(v, ss, a) = \Pr_{z \leftarrow Z_{n,k}|A=a}[\mathtt{Viable}(v, ss, a, z) = 1]$. Define $\mathtt{HViable}(v, ss, a, \mathcal{E}) = \mathtt{Viable}(v, ss, a)$ if $v \notin \mathcal{E}$ and 0 otherwise. We present an analog of Claim 3 adapted to the secure sketch setting.

Claim 5. *Let μ and* $\mathsf{GoodSketch}$ *be defined as in Lemma 12. Let A be a distribution and let a be a fixed value such that $\mathrm{H}_\infty(\mathcal{W}_{n,k}|A=a)=\alpha_a$. Fix some value v. Each value $w^*\in\{0,1\}^n$ defines a set $\mathcal{E}_{a,w^*}$ where $|\mathcal{E}_{a,w^*}|\le\nu+1$ such that*

$$\mathsf{log}\left(\sum_{ss}\mathtt{HViable}(v,ss,a,\mathcal{E}_{a,w^*})\right)\le-\frac{\alpha_a}{2^k}+1+\mu+k-\mathsf{log}(\nu)\,.$$

The proof of Claim 5 follows the structure of the proof of Claim 3 and is omitted. Claim 5 suffices to bound how many points are "viable" from the output of the secure sketch.

Corollary 13. *Let μ be defined as in Lemma 12. Fix an arbitrary point $w^*\in\{0,1\}^n$, some ss and some a. Define $\mathcal{E}_{a,w^*}$ as in Claim 5. By Claim 5 on average across $z\leftarrow Z_{n,k}|A=a$ (by union bound across the points in W_z) one has that:*

$$\mathsf{log}\left(\mathop{\mathbb{E}}_{z\leftarrow Z_{n,k}|A=a}\left|\left\{w\;\middle|\;\begin{matrix}w\notin\mathcal{E}_{a,w^*}\\ \mathtt{Viable}(w,ss,a,z)\end{matrix}\right\}\right|\right)\le-\frac{\alpha_a}{2^k}+1+\mu+2k-\mathsf{log}(\nu)\,.$$

Then on average across $a\in A$ (using the fact that $\mathbb{E}_{a\leftarrow A}\left(2^{-\alpha_a}\right)=2^{-(\alpha-\ell)}$) one has

$$\mathsf{log}\left(\mathop{\mathbb{E}}_{a\leftarrow A}\left(\mathop{\mathbb{E}}_{z\leftarrow Z_{n,k}|A=a}\left|\left\{w\;\middle|\;\begin{matrix}w\notin\mathcal{E}_{a,w^*}\\ \mathtt{Viable}(w,ss,a,z)\end{matrix}\right\}\right|\right)\right)\le-\frac{\alpha-\ell}{2^k}+1+\mu+2k-\mathsf{log}(\nu)\,.$$

And finally,

$$\mathop{\mathbb{E}}_{a\leftarrow A}\left(\mathop{\mathbb{E}}_{z\leftarrow Z_{n,k}|A=a}|\{w|\mathtt{Viable}(w,ss,a,z)\}|\right)\le2^{-\frac{\alpha-\ell}{2^k}+1+\mu+2k-\mathsf{log}(\nu)}+\nu+1.$$

Lemma 14. *Let all parameters be as in Corollary 13 with $\nu\in\mathbb{Z}^+$. For the family $\mathcal{W}_{n,k,Z}$ there is no $(\{0,1\}^n,\mathcal{W}_{n,k,Z},\tilde{m},t,\epsilon,\ell,\delta)$-secure sketch with distributional advice if*

$$\tilde{m}>-\mathsf{log}(1-2\epsilon)+1+2\max\left\{-\frac{\alpha-|\mathsf{advice}|}{2^k}+1+\mu+2k-\mathsf{log}(\nu)\,,\mathsf{log}(\nu+1)\right\}.$$

Furthermore, there exists an algorithm $\mathcal{D}$ that always outputs 1 when given samples of the form w,ss,z that are correctly generated by the secure sketch.

Proof of Lemma 14. First recall for every $v\in\{0,1\}^n$ and advice string a there exists a set $\mathsf{GoodSketch}_{v,a}$ where $\Pr[\mathsf{SS}(v,a)\in\mathsf{GoodSketch}_{v,a}]\ge1/2$. We first need an elementary claim which says that just predicting w when the sketch is in good sketch implies a predictor for the full setting with only a single bit of loss.

Claim 6. *Let (X,Y) be a pair of random variables and, $S(X,Y)$ be a set, let f be a randomized function taking inputs on the domain of (X,Y). Then*

$$\tilde{\mathrm{H}}^\epsilon_\infty(X|Y,f(X,Y)\in S(X,Y))\ge\tilde{\mathrm{H}}^\epsilon_\infty(X|Y)+\mathsf{log}(\Pr[f(X,Y)\in S(X,Y)])\,.$$

Proof of Claim 6. Let X',Y' be a distribution such that $\Delta((X,Y),(X',Y'))\le\epsilon$. By [FRS20, Lemma 7.8] for any event η

$$\tilde{\mathrm{H}}_\infty(X'|Y',\eta)\ge\tilde{\mathrm{H}}_\infty(X'|Y')+\mathsf{log}(\Pr[\eta])\,.$$

Let η denote the event that $f(X,Y)\in S(X,Y)$. The proof completes by noting that $\Delta((X,Y),(X',Y'))\le\epsilon$ implies that

$$\Delta((X,Y,f(X,Y)\stackrel{?}{\in}S(X,Y)),(X',Y',f(X',Y')\stackrel{?}{\in}S(X',Y')))\le\epsilon$$

by the information processing lemma. This is turn implies that

$$\tilde{\mathrm{H}}^{\epsilon}_{\infty}(X|Y, f(X,Y) \in S(X,Y)) \geq \tilde{\mathrm{H}}^{\epsilon}_{\infty}(X|Y) + \Pr[f(X,Y) \in S(X,Y)].$$

This completes the proof of Claim 6. □

We now proceed with the proof of Lemma 14. Define $\mathcal{D}(v, ss, z) = 1$ if $ss \notin \mathsf{GoodSketch}_v$ or $\mathtt{Viable}(v, ss, \mathsf{Advice}(z), z) = 1$. $\mathcal{D}$ outputs 0 otherwise. Note that for w, ss, z correctly generated as the output of SS $\mathcal{D}$ always outputs 1. Denote

$$\begin{aligned} A &:= \mathsf{advice}(Z_{n,k}) \\ X &:= \mathcal{W}_{n,k,Z}, \\ Y &:= \mathsf{SS}(\mathcal{W}_{n,k,Z}, A). \end{aligned}$$

By Claim 6:

$$\tilde{\mathrm{H}}^{\epsilon}_{\infty}(X|Y,Z) \leq \tilde{\mathrm{H}}^{\epsilon}_{\infty}(X|Y,Z, X \in \mathsf{GoodSketch}_{Y,A}) + 1.$$

In the above note that A and thus $\mathsf{GoodSketch}_{\mathcal{W}_{n,k},A}$ are computable from the pair $\mathcal{W}_{n,k}, Z$ since Advice is a function. Let X', Y', Z' be a triple of random variables where

$$\Delta((X,Y,Z),(X',Y',Z')) \leq \epsilon.$$

Our goal is to bound the min-entropy of $X', Y', Z'|X' \in \mathsf{GoodSketch}_{Y',\mathsf{Advice}(Z')}$ by Claim 6 the smooth min-entropy without conditioning on this event increases by at most 1. First note that

$$\Pr_{(x,y,z)\leftarrow X',Y',Z'}[\mathcal{D}(x,y,z) = 1 | X' \in \mathsf{GoodSketch}_{Y',\mathsf{Advice}(Z')}] \geq 1 - 2\epsilon.$$

Let $A' := \mathsf{Advice}(Z')$. When $X' \in \mathsf{GoodSketch}_{Y',\mathsf{Advice}(Z')}$ in order for $\mathcal{D}$ to output 1 it must be the case that $\mathtt{Viable}(x', y', \mathsf{Advice}(z'), z') = 1$. That is, the support of x' must be drawn from points in W'_z. For any fixed value of $y \in Y$ and arbitrary random variable A' of length at most ℓ by Corollary 13 the number of such x' is at most $2^{-\frac{\alpha-\ell}{2^k}+1+\mu+2k-\log(\nu)} + \nu + 1$. For any fixed support the min-entropy is maximized by considering the uniform distribution over such points.

$$\begin{aligned} &\tilde{\mathrm{H}}_{\infty}(X'|Y', Z', X' \in \mathsf{GoodSketch}_{Y',\mathsf{Advice}(Z')}, \mathcal{D}(X',Y',Z') = 1) \\ &\quad \leq \log\left(2^{-\frac{\alpha-|\mathsf{advice}|}{2^k}+1+\mu+2k-\log(\nu)} + \nu + 1\right) \\ &\quad \leq 2\max\left\{-\frac{\alpha - |\mathsf{advice}|}{2^k} + 1 + \mu + 2k - \log(\nu), \log(\nu+1)\right\}. \end{aligned}$$

One then has,

$$\begin{aligned} &2^{-\tilde{\mathrm{H}}_{\infty}(X'|Y',Z',X'\in\mathsf{GoodSketch}_{Y',\mathsf{Advice}(Z')})} \\ &\geq \Pr_{x,y,z\leftarrow(X',Y',Z')}[\mathcal{D}(x,y,z) = |X' \in \mathsf{GoodSketch}_{Y',\mathsf{Advice}(Z')}] 2^{-\tilde{\mathrm{H}}_{\infty}(X'|Y',Z',X'\in\mathsf{GoodSketch}_{Y',\mathsf{Advice}(Z')},\mathcal{D}(X',Y',Z')=1)} \\ &+ \Pr_{x,y,z\leftarrow(X',Y',Z')}[\mathcal{D}(x,y,z) = 0|X' \in \mathsf{GoodSketch}_{Y',\mathsf{Advice}(Z')}] 2^{-\tilde{\mathrm{H}}_{\infty}(X'|Y',Z',X'\in\mathsf{GoodSketch}_{Y',\mathsf{Advice}(Z')},\mathcal{D}(X',Y',Z')=0)} \\ &\geq (1-2\epsilon) 2^{-\tilde{\mathrm{H}}_{\infty}(X'|Y',Z',X'\in\mathsf{GoodSketch}_{Y',\mathsf{Advice}(Z')},\mathcal{D}(X',Y',Z')=1)} \end{aligned}$$

And thus,

$$\tilde{\mathrm{H}}_{\infty}(X'|Y', Z', X' \in \mathsf{GoodSketch}_{Y',\mathsf{Advice}(Z')}) \leq -\log(1-2\epsilon) + \tilde{\mathrm{H}}_{\infty}(X'|Y', X' \in W_Z).$$

Define

$$\tilde{m} := -\log(1-2\epsilon_{\mathsf{SS}}) + 1 + 2\max\left\{-\frac{\alpha + -|\mathsf{advice}|}{2^k} + 1 + \mu + 2k - \log(\nu), \log(\nu+1)\right\}.$$

This implies that $\tilde{\mathrm{H}}^{\epsilon}_{\infty}(\mathcal{W}_{n,k,Z}|\mathsf{SS}(\mathcal{W}_{n,k,Z}, A), Z_{n,k}) \leq \tilde{m}$. This completes the Proof of Lemma 14. □

We now proceed to the proof of Theorem 11. Let

$$\tilde{m} := -\mathsf{log}(1-2\epsilon_{\mathsf{SS}}) + 1 + 2\max\left\{-\frac{\alpha + -|\mathsf{advice}|}{2^k} + 1 + \mu + 2k - \mathsf{log}(\nu)\,, \mathsf{log}(\nu+1)\right\}.$$

Restating Lemma 14 one has that

$$\tilde{\mathrm{H}}_\infty^{\epsilon_{\mathsf{SS}}}(\mathcal{W}_{n,k,Z}|\mathsf{SS}(\mathcal{W}_{n,k,Z}, A), Z_{n,k}) \leq \tilde{m}.$$

Define

$$\begin{aligned} A' &:= \mathsf{Advice}(Z_{n,k,t,\gamma}), \\ SS' &:= \mathsf{SS}(\mathcal{W}_{n,k,t,\gamma,Z}, A'), \\ \epsilon_{\mathcal{W}_{n,k,Z}} &:= \left(e2^{\gamma-\beta}\right)^{2^{k-\gamma}} 2^n. \end{aligned}$$

By Lemma 7, $\Delta(\mathcal{W}_{n,k,Z}, \mathcal{W}_{n,k,t,\gamma,Z}) \leq \epsilon_{\mathcal{W}_{n,k,Z}}$ and thus $\tilde{\mathrm{H}}_\infty^{\epsilon_{\mathsf{SS}}-\epsilon_{\mathtt{PCode}}}(\mathcal{W}_{n,k,t,\gamma,Z}|SS', Z_{n,k,t,\gamma}) \leq \tilde{m}$. Suppose not, then there exists some E, F, G where

$$\Delta((E,F,G), (\mathtt{PCode}^*_{n,k,t,\alpha}, \mathsf{SS}(\mathtt{PCode}^*_{n,k,t,\alpha}), Z_{\mathtt{PCode}^*_{n,k,t,\alpha}})) \leq \epsilon_{\mathsf{SS}} - \epsilon_{\mathtt{PCode}}.$$

and $\tilde{\mathrm{H}}_\infty(E|F,G) > \tilde{m}$. Thus,

$$\begin{aligned} &\Delta((E,F,G), (\mathcal{W}_{n,k,Z}, \mathsf{SS}(\mathcal{W}_{n,k,Z}, A), Z_{n,k})) \\ &\quad \leq \Delta((E,F,G), ((Z_{n,k,t,\gamma}, SS', \mathcal{W}_{n,k,t,\gamma,Z})) + \Delta(Z_{n,k}, Z_{n,k,t,\gamma}) \\ &\quad \leq \epsilon_{\mathsf{SS}} - \epsilon_{\mathtt{PCode}} + \epsilon_{\mathtt{PCode}} = \epsilon_{\mathsf{SS}}. \end{aligned}$$

This contradicts the fact that $\tilde{\mathrm{H}}_\infty^{\epsilon_{\mathsf{SS}}}(\mathcal{W}_{n,k,Z}|\mathsf{SS}(\mathcal{W}_{n,k}), Z_{n,k})) \leq \tilde{m}$. Finally, Theorem 11 follows by application of Lemma 5 with setting $\zeta = \chi$ and noting that $\chi \geq 1$.

□

4.1 Analysis of parameters

We assume that $\epsilon \leq 1/8$ and $\delta < 1/4$. As before for $\left(e2^{\gamma-\beta}\right)^{2^{k-\gamma}} 2^n$ to be negligible it suffices that

Condition 1 That $\gamma \leq \beta - \mathsf{log}(2e)$.[3]

Condition 2 Let $0 < c_k < 1$ be some arbitrary constant and suppose that $k = \gamma + c_k n$ which implies that $k \geq \gamma + \mathsf{log}(n + \omega(\mathsf{log}(n)))$.

These two conditions imply that $-\mathsf{log}(1-2\epsilon) \leq 1$ and $\epsilon' \geq 1/8 - \mathtt{ngl}(\lambda)$.

Condition 3 That $\nu = 1$.

Define

$$\chi := -\frac{\alpha - \ell}{2^k} + 1 + \mu + 2k - \mathsf{log}(\nu)$$

We now turn to our analysis of χ. Recall that $(n/k)^k \leq \binom{n}{k} < ((ne)/k)^k$. Recalling parameters:

$$\begin{aligned} \mu &\leq \frac{(n(1-h_2(t/n)) + h_2(2\delta))}{(2\delta)}, \\ \alpha &= \mathsf{log}\left(\binom{2^n}{2^k}\right) \geq \mathsf{log}\left(2^{n2^k}/2^{k2^k}\right) = (n-k)2^k, \end{aligned}$$

[3] As in Section 3.3 the additional constraint that $\gamma \leq \beta - \mathsf{log}(2e)$ imposes an additive $\mathsf{log}(2e)$ impact on the maximal fuzzy min-entropy that can be supported.

This implies that

$$\begin{aligned}\chi &= -\frac{\alpha-\ell-k}{2^k}+\mu+2k+1\\ &\leq -\frac{\alpha+\mathsf{log}(\nu)-\ell-k}{2^k}+\frac{n(1-h_2(t/n))+h_2(2\delta)}{1-2\delta}+2k+1,\\ &\leq -\frac{(n-k)2^k+\mathsf{log}(\nu)-\ell-k}{2^k}+\frac{n(1-h_2(t/n))+h_2(2\delta)}{1-2\delta}+2k+1,\\ &\leq -\frac{(n-3k)2^k-\ell-k}{2^k}+\frac{n(1-h_2(t/n))+h_2(2\delta)}{1-2\delta}+1.\end{aligned}$$

We consider two settings for δ one when $\delta < 1/4$ and another when $\delta = 0$.

Constant error, $\delta < 1/4$ As long as for constants c_k, c_ℓ one has

$$\begin{aligned}&\ell \leq 3c_\ell n 2^k,\\ &\delta < 1/4,\\ 0 \leq \frac{\gamma}{n} &\leq \min\left\{(1-h_2(t/n))+\frac{\mathsf{log}(n)+1-2\mathsf{log}(2e)}{2n}, \frac{2}{3}h_2(t/n)-\frac{1}{3}-\frac{4c_k+c_\ell}{3}-\frac{2}{3n}\right\}.\end{aligned}$$

then

$$\begin{aligned}\chi &\leq -\frac{(n-3k)2^k-\ell-k}{2^k}+\frac{n(1-h_2(t/n))+h_2(2\delta)}{1-2\delta}+1\\ &\leq -\frac{(n-3k)2^k-\ell-k}{2^k}+\frac{n(1-h_2(t/n))+h_2(2\delta)}{1-2\delta}+1\\ &\leq -\frac{(n-3k)2^k-\ell-k}{2^k}+2n(1-h_2(t/n))+2\\ &\leq -(n-(4c_k+c_\ell)n-3\gamma)+2n(1-h_2(t/n))+2\\ &\leq -n+(4c_k+c_\ell)n+3\gamma+2n(1-h_2(t/n))+2\\ &\leq n+(4c_k+c_\ell)n+3\gamma-2nh_2(t/n))+2\leq 0\end{aligned}$$

then $\tilde{m} \leq 3+2\max\{\chi, \mathsf{log}(2)\} \leq 5$ which implies that $1/32$ of the distributions have no secure sketch.

No error, $\delta = 0$ One has

$$\begin{aligned}&\ell \leq 3c_\ell n 2^k,\\ &\delta < 1/4,\\ 0 \leq \frac{\gamma}{n} &\leq \min\left\{(1-h_2(t/n))+\frac{\mathsf{log}(n)+1-2\mathsf{log}(2e)}{2n}, \frac{1}{3}h_2(t/n)-\frac{4c_k+c_\ell}{3}-\frac{2}{3n}\right\}.\end{aligned}$$

yielding $\tilde{m} \leq 5$ which implies that $1/32$ of the distributions have no secure sketch.

Acknowledgements

This work was supported by NSF Grants #2232813 and #2141033 and the Office of Naval Research. The author thanks anonymous reviewers, Luke Demarest, and Alexander Russell for their important feedback on the manuscript.

References

[ABC+18] Quentin Alamélou, Paul-Edmond Berthier, Chloé Cachet, Stéphane Cauchie, Benjamin Fuller, Philippe Gaborit, and Sailesh Simhadri. Pseudoentropic isometries: A new framework for fuzzy extractor reusability. In *Proceedings of the 2018 on Asia Conference on Computer and Communications Security*, pages 673–684, 2018.

[ACEK17] Daniel Apon, Chongwon Cho, Karim Eldefrawy, and Jonathan Katz. Efficient, reusable fuzzy extractors from LWE. In *International Conference on Cyber Security Cryptography and Machine Learning*, pages 1–18. Springer, 2017.

[Ash65] Robert B Ash. *Information theory*. Dover Publications, New York, first edition, 1965.

[BBC+14] Boaz Barak, Nir Bitansky, Ran Canetti, Yael Tauman Kalai, Omer Paneth, and Amit Sahai. Obfuscation for evasive functions. In *Theory of Cryptography Conference*, pages 26–51. Springer, 2014.

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

[BCKP14] Nir Bitansky, Ran Canetti, Yael Tauman Kalai, and Omer Paneth. On virtual grey box obfuscation for general circuits. In *Advances in Cryptology - CRYPTO 2014*, 2014.

[BCKP17] Nir Bitansky, Ran Canetti, Yael Tauman Kalai, and Omer Paneth. On virtual grey box obfuscation for general circuits. *Algorithmica*, 79(4):1014–1051, 2017.

[BDK+11] Boaz Barak, Yevgeniy Dodis, Hugo Krawczyk, Olivier Pereira, Krzysztof Pietrzak, François-Xavier Standaert, and Yu Yu. Leftover hash lemma, revisited. In *Annual Cryptology Conference*, pages 1–20. Springer, 2011.

[CFP+21] Ran Canetti, Benjamin Fuller, Omer Paneth, Leonid Reyzin, and Adam Smith. Reusable fuzzy extractors for low-entropy distributions. *Journal of Cryptology*, 34(1):1–33, 2021.

[CW77] J Lawrence Carter and Mark N Wegman. Universal classes of hash functions. In *Proceedings of the ninth annual ACM symposium on Theory of computing*, pages 106–112, 1977.

[DFR21] Luke Demarest, Benjamin Fuller, and Alexander Russell. Code offset in the exponent. In *2nd Conference on Information-Theoretic Cryptography*, 2021.

[DORS08] Yevgeniy Dodis, Rafail Ostrovsky, Leonid Reyzin, and Adam Smith. Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. *SIAM journal on computing*, 38(1):97–139, 2008.

[FMR20] Benjamin Fuller, Xianrui Meng, and Leonid Reyzin. Computational fuzzy extractors. *Information and Computation*, 275:104602, 2020.

[FP19] Benjamin Fuller and Lowen Peng. Continuous-source fuzzy extractors: source uncertainty and insecurity. In *2019 IEEE International Symposium on Information Theory (ISIT)*, pages 2952–2956. IEEE, 2019.

[FRS16] Benjamin Fuller, Leonid Reyzin, and Adam Smith. When are fuzzy extractors possible? In *International Conference on the Theory and Application of Cryptology and Information Security*, pages 277–306. Springer, 2016.

[FRS20] Benjamin Fuller, Leonid Reyzin, and Adam Smith. When are fuzzy extractors possible? *IEEE Transactions on Information Theory*, 66(8):5282–5298, 2020.

[GW11] Craig Gentry and Daniel Wichs. Separating succinct non-interactive arguments from all falsifiable assumptions. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pages 99–108, 2011.

[GZ19] Steven D Galbraith and Lukas Zobernig. Obfuscated fuzzy hamming distance and conjunctions from subset product problems. In *Theory of Cryptography Conference*, pages 81–110. Springer, 2019.

[Har66] Lawrence H Harper. Optimal numberings and isoperimetric problems on graphs. *Journal of Combinatorial Theory*, 1(3):385–393, 1966.

[HILL93] Johan Håstad, Russell Impagliazzo, Leonid A Levin, and Michael Luby. Construction of a pseudo-random generator from any one-way function. In *SIAM Journal on Computing*. Citeseer, 1993.

[HMSS12] Matthias Hiller, Dominik Merli, Frederic Stumpf, and Georg Sigl. Complementary ibs: Application specific error correction for PUFs. In *IEEE International Symposium on Hardware-Oriented Security and Trust (HOST)*, pages 1–6. IEEE, 2012.

[HTW14] Masahito Hayashi, Himanshu Tyagi, and Shun Watanabe. Secret key agreement: General capacity and second-order asymptotics. In *2014 IEEE International Symposium on Information Theory*, pages 1136–1140. IEEE, 2014.

[HTW16] Masahito Hayashi, Himanshu Tyagi, and Shun Watanabe. Secret key agreement: General capacity and second-order asymptotics. *IEEE Transactions on Information Theory*, 62(7):3796–3810, 2016.

[LA18] Cheuk Ting Li and Venkat Anantharam. One-shot variable-length secret key agreement approaching mutual information. *CoRR*, abs/1809.01793, 2018.

[Mau93] Ueli M. Maurer. Secret key agreement by public discussion from common information. *IEEE Transactions on Information Theory*, 39(3):733–742, 1993.

[MTV09] Roel Maes, Pim Tuyls, and Ingrid Verbauwhede. Low-overhead implementation of a soft decision helper data algorithm for SRAM PUFs. In *Cryptographic Hardware and Embedded Systems-CHES 2009*, pages 332–347. Springer, 2009.

[MW96] Ueli M. Maurer and Stefan Wolf. Towards characterizing when information-theoretic secret key agreement is possible. In Kwangjo Kim and Tsutomu Matsumoto, editors, *Advances in Cryptology - ASIACRYPT '96, International Conference on the Theory and Applications of Cryptology and Information Security, Kyongju, Korea, November 3-7, 1996, Proceedings*, volume 1163 of *Lecture Notes in Computer Science*, pages 196–209. Springer, 1996.

[NZ93] Noam Nisan and David Zuckerman. Randomness is linear in space. *Journal of Computer and System Sciences*, pages 43–52, 1993.

[Rey11] Leonid Reyzin. Some notions of entropy for cryptography: (invited talk). In *Information Theoretic Security: 5th International Conference, ICITS 2011, Amsterdam, The Netherlands, May 21-24, 2011. Proceedings 5*, pages 138–142. Springer, 2011.

[RW05] Renato Renner and Stefan Wolf. Simple and tight bounds for information reconciliation and privacy amplification. In *International conference on the theory and application of cryptology and information security*, pages 199–216. Springer, 2005.

[SSF19] Sailesh Simhadri, James Steel, and Benjamin Fuller. Cryptographic authentication from the iris. In *International Conference on Information Security*, pages 465–485. Springer, 2019.

[ST09] Boris Skoric and Pim Tuyls. An efficient fuzzy extractor for limited noise. *Cryptology ePrint Archive*, 2009.

[TVW18] Himanshu Tyagi, Pramod Viswanath, and Shun Watanabe. Interactive communication for data exchange. *IEEE Trans. Information Theory*, 64(1):26–37, 2018.

[TW17] Himanshu Tyagi and Shun Watanabe. Universal multiparty data exchange and secret key agreement. *IEEE Transactions on Information Theory*, 63(7):4057–4074, 2017.

[WCD+17] Joanne Woodage, Rahul Chatterjee, Yevgeniy Dodis, Ari Juels, and Thomas Ristenpart. A new distribution-sensitive secure sketch and popularity-proportional hashing. In *Advances in Cryptology – CRYPTO*, pages 682–710. Springer, 2017.

[WL18] Yunhua Wen and Shengli Liu. Robustly reusable fuzzy extractor from standard assumptions. In *International Conference on the Theory and Application of Cryptology and Information Security*, pages 459–489. Springer, 2018.

[WLG19] Yunhua Wen, Shengli Liu, and Dawu Gu. Generic constructions of robustly reusable fuzzy extractor. In *IACR International Workshop on Public Key Cryptography*, pages 349–378. Springer, 2019.

[YD10] Meng-Day Yu and Srinivas Devadas. Secure and robust error correction for physical unclonable functions. *IEEE Design & Test of Computers*, 27(1):48–65, 2010.