When are Fuzzy Extractors Possible?

Benjamin Fuller* Leonid Reyzin† Adam Smith‡

December 23, 2014

Abstract

Fuzzy extractors (Dodis et al., Eurocrypt 2004) convert repeated noisy readings of a high-entropy secret into the same uniformly distributed key. A minimum condition for the security of the key is the hardness of guessing a value that is similar to the secret, because the fuzzy extractor converts such a guess to the key. We define *fuzzy min-entropy* to quantify this property of a noisy source of secrets.

High fuzzy min-entropy is necessary for the existence of a fuzzy extractor; moreover, there is evidence that it may be sufficient when only computational security is required. Nevertheless, information-theoretic fuzzy extractors are not known for many practically relevant sources of high fuzzy min-entropy. In this work, we ask: *is fuzzy min-entropy sufficient to build information-theoretic fuzzy extractors?*

We give a positive answer to this question when the fuzzy extractor knows the precise distribution of the physical source. On the other hand, because it is imprudent to assume precise knowledge of a complicated distribution, fuzzy extractors are typically designed to work for families of sources. We show that this uncertainty is an impediment to security by building a family of high fuzzy min-entropy sources for which no fuzzy extractor can exist.

We provide similar but stronger results for secure sketches, whose goal is not to derive a consistent key, but to recover a consistent reading of the secret.

Keywords: Fuzzy extractors, secure sketches, information theory, biometric authentication, error-tolerance, key derivation, error-correcting codes.

*Email: `bfuller@cs.bu.edu`. Boston University and MIT Lincoln Laboratory.

†Email: `reyzin@cs.bu.edu`. Boston University.

‡Email: `asmith@cse.psu.edu`. Pennsylvania State University. This work performed while visiting Boston University's Hariri Institute for Computing.

1 Introduction

Sources of reproducible secret random bits are necessary for many cryptographic applications. In many situations these bits are not explicitly stored for future use, but are obtained by repeating the same process (such as reading a biometric or a physically unclonable function) that generated them the first time. However, bits obtained this way present a problem: noise [Dau04, ZH93, BS00, EHMS00, MG09, MRW02, PRTG02, GCVDD02, TSv$^+$06, SD07, BBR88]. That is, when a secret is read multiple times, readings are close (according to some metric) but not identical. To utilize such sources, it is often necessary to remove noise, in order to derive the same value in subsequent readings.

The same problem occurs in the interactive setting, in which the secret channel used for transmitting the bits between two users is noisy and/or leaky [Wyn75]. Bennett, Brassard, and Robert [BBR88] identify two fundamental tasks. The first, called information reconciliation, removes the noise without leaking significant information. The second, known as privacy amplification, converts the high entropy secret to a uniform random value. In this work, we consider the noninteractive version of these problems, in which these tasks are performed together with a single message.

The noninteractive setting is modeled by a primitive called a fuzzy extractor [DORS08], which consists of two algorithms. The generate algorithm (Gen) takes an initial reading w and produces an output key along with a nonsecret helper value p. The reproduce (Rep) algorithm takes the subsequent reading w' along with the helper value p to reproduce key. The correctness guarantee is that the key is reproduced precisely as long as the distance between w and w' is at most t.

The security requirement for fuzzy extractors is that key is uniform even to a (computationally unbounded) adversary who has observed p. This requirement is harder to satisfy as the allowed error tolerance t increases, because it becomes easier for the adversary to guess key by guessing a w' within distance t of w and running $\mathsf{Rep}(w', p)$.

Fuzzy Min-Entropy We introduce a new entropy notion that precisely measures how hard it is for the adversary to guess a value within distance t of the original reading w, thus subverting the security of key by running Rep. Suppose w is sampled from a distribution W. To have the maximum chance that w' is within distance t of w, the adversary would want to maximize the total probability mass of W within the ball $B_t(w')$ of radius t around w'. We therefore define *fuzzy min-entropy* $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) \stackrel{\mathrm{def}}{=} -\log \max_{w'} \Pr[W \in B_t(w')]$. Observe that this quantity can be bounded in terms of min-entropy: $\mathrm{H}_\infty(W) \geq \mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) \geq \mathrm{H}_\infty(W) - \log|B_t|$.

Superlogarithmic fuzzy min-entropy is *necessary* for nontrivial key extraction (Proposition 2.6 formalizes the above intuition). However, existing constructions do not measure their security in terms of fuzzy min-entropy; instead, their security is shown to be $\mathrm{H}_\infty(W)$ minus some loss that is at least $\log|B_t|$ due to error-tolerance. Since $\mathrm{H}_\infty(W) - \log|B_t| \leq \mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$, it is natural to ask whether this loss is necessary. This question is particularly relevant when the gap between the two sides of the inequality is high. As an example, iris scans appear to have significant $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$ (because iris scans for different people appear to be well-spread in the metric space [Dau06]) but negative $\mathrm{H}_\infty(W) - \log|B_t|$ [BH09, Section 5]. We therefore ask: *is fuzzy min-entropy sufficient for fuzzy extraction?* There is evidence that it may be when the security requirement is computational rather than information-theoretic—see Section 1.2.

Tight Characterization for the Case of a Known Distribution We show that for every source W with superlogarithmic $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$, it is possible to construct a fuzzy extractor with a superlogarithmic length key (Corollary 3.8). We thus show that $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$ is a necessary and sufficient condition for building a fuzzy extractor for a *known distribution* W. It is important to emphasize that these constructions incorporate the knowledge of the complete distribution of W (and, in particular, they are not polynomial-time).

A number of previous works in this known-distribution setting have provided efficient algorithms and tight bounds for specific distributions—generally the uniform distribution or i.i.d. sequences (for example, [JW99, LT03, TG04, HAD06, WRDI11, IW12]). Our characterization may be seen as unifying previous work, and justifies using $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$ as the measure of the quality of a noisy distribution, rather than cruder measures such as $\mathrm{H}_\infty(W) - \log|B_t|$.

Impossibility of Fuzzy Extractors for Families of Distributions Assuming full knowledge of a distribution is often unrealistic. Indeed, high-entropy distributions can never be fully observed directly and must therefore be modeled. It is imprudent to assume that the designer's model of a distribution is completely accurate—the adversary, with greater resources, would likely be able to build a better model. Therefore, fuzzy extractor designs cannot usually be tailored to one particular source. Existing designs work for a family of sources (for example, all sources of min-entropy at least m with at most t errors). Thus, the design is fixed before the distribution is fully known, and the adversary may therefore know more about the distribution than the designer of the fuzzy extractor.

We show that this extra adversarial knowledge can be devastating (Theorem 5.1). Specifically, we show a family of distributions $\mathcal{W}$ such that not even a 2-bit fuzzy extractor can be secure for most distributions in $\mathcal{W}$. We emphasize that each distribution $W \in \mathcal{W}$ has superlogarithmic fuzzy min-entropy—in fact, $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) = \mathrm{H}_\infty(W)$, because all points in W are distance at least t apart. Our proof relies on high dimensionality of W and on perfect correctness of the Rep procedure.

	Known Distribution	Family of Distributions
Secure Sketch	Yes (Corollary 3.8)	No (Theorem 4.1)
Fuzzy Extractor	Yes (Corollary 3.8)	No (Theorem 5.1)

Table 1: Is fuzzy min-entropy sufficient to extract a superlogarithmic length key? Results are information-theoretic.

Stronger Results on Information Reconciliation (Secure Sketches) Traditionally, fuzzy extractors use a secure sketch to perform information reconciliation (mapping w' back to w), followed by randomness extractor [NZ93] to transform w into a uniform key. The security losses incurred in the first of these two steps dominate for typical sources and, indeed, this step is less well understood.[1] Formally, a secure sketch performs non-interactive information reconciliation via pair of algorithms: SS takes w and produces a nonsecret value ss, while Rec takes a value w' within distance t of w and uses SS to output the original reading w.

We show comparable, but stronger, results for secure sketches. Namely, we show in Corollary 3.8 that secure sketches are possible if the distribution W is precisely known. (In fact, we obtain our fuzzy extractors for the case of a known distribution from this result by applying a randomness extractor.)

On the other hand, there is a family of sources with superlogarithmic $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) = \mathrm{H}_\infty(W)$ for which no secure sketch correcting even a few errors is possible (Theorem 4.1). The impossibility result applies

[1]Randomness extractors have matching upper and lower bounds on the security loss: for every extra two bits of output key, they lose one bit of security

even when Rec is allowed to be incorrect with probability up to $1/4$ (as opposed to our fuzzy extractor impossibility result).

1.1 Our Techniques

Techniques for Positive Results for Known Distributions We now explain how to construct a secure sketch for an arbitrary known distribution W (we already explained how to construct a fuzzy extractor from it). We begin with distributions in which all points in the support have the same probability (so-called "flat" distributions). Consider some subsequent reading w'. To achieve correctness, the sketch algorithm must disambiguate which point $w \in W$ within distance t of w' was sketched. Disambiguating multiple points can be accomplished by universal hashing, as long as the size of hash output space is slightly greater than the number of possible points. Thus, our sketch is computed via a universal hash of w. To determine the length of that sketch, consider the heaviest (according to W) ball of radius t. Because the distribution is flat, it is also the ball with the most points of nonzero probability. Thus, the length of the sketch needs to be slightly greater than the logarithm of the number of non-zero probability point in that ball. Since $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$ is determined by the weight of that ball, the number of points cannot be too high and there will be entropy left after the sketch is published.

For an arbitrary distribution, we cannot afford to disambiguate points in the ball with the greatest number of points, because there could be too many low-probability points in a single ball despite a high $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$. We solve this problem by splitting the arbitrary distribution into a number of nearly flat distributions we call "levels." We then write down, as part of the sketch, the level of the original reading w and apply the above construction considering only points in that level. We call this construction *leveled hashing*.

Techniques for Negative Results for Distribution Families We construct a family of distributions $\mathcal{W}$ and prove impossibility for a uniformly random $W \in \mathcal{W}$ (instead of proving impossibility for a worst-case W). We start by observing the following asymmetry: Gen sees only the sample w (obtained via $W \leftarrow \mathcal{W}$ and $w \leftarrow W$), while the adversary knows W. To exploit the asymmetry, we construct $\mathcal{W}$ so that conditioning on the knowledge of W reduces the distribution to a single affine line, but conditioning on w leaves the rest of the distribution uniform on a large fraction of the entire space.

Then we show how the adversary can exploit the knowledge of the affine line to reduce the uncertainty about w (in the secure sketch case) or key (in the fuzzy extractor case). In the secure sketch case, ss can be used to find fixed points of $\mathsf{Rec}(\cdot, ss)$ which, by the correctness requirement of the sketch, must be separated by minimum distance t. This means there aren't too many of them, so few can lie on an average line, permitting the adversary to guess one easily.

In the fuzzy extractor case, the nonsecret value p partitions the metric space into regions that produce a consistent value under Rep (preimages of each key under $\mathsf{Rep}(\cdot, p)$). For each of these regions, the adversary knows that possible w lie t-far from the boundary of the region. However, in the Hamming space, the vast majority of points lie near the boundary (this follows by combining the isoperimetric inequality [Har66] showing that the ball has the smallest boundary and Hoeffding's inequality [Hoe63] for bounding the volume that is t-away from this boundary). This allows the adversary to rule out so many possible w that, combined with the adversarial knowledge of the affine line, many regions become empty, leaving key far from uniform.

The result for fuzzy extractors is delicate. It uses the fact that p partitions the space into nonoverlapping regions, which is implied by perfect correctness. Extending this result to imperfect correctness seems challenging and is an interesting open problem. It also uses the fact that there are few points far from the boundary of every region, which is implied by the geometry of the high-dimensional Hamming space. This fact seems crucial: in contrast, in low-dimensional Euclidean space, which does not have this property, a single fuzzy extractor can work for any distribution with sufficient $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}$. (Such a construction would use quantization or tiling, similar to, for example, [CK03, LT03, CZC04, LC06, BDH+10, VTO+10]. Each sample from W would map to the "tile" containing it, from which the output key would be extracted. A randomly chosen quantizer would have the property that few samples lie near the boundary, giving almost-perfect correctness; if perfect correctness is desired, we can give up on security for those rare samples and simply use a special value of p to indicate that one of them was the input.)

1.2 Related Settings

Other settings with close readings: $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}$ is sufficient The security definition of fuzzy extractors and secure sketches can be weakened to protect only against computationally bounded adversaries [FMR13]. In this computational setting, fuzzy extractors and secure sketches can be constructed for the family of all distributions W with superlogarithmic $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}$ by using virtual grey-box obfuscation for all circuits [BCKP14]. The construction places into p the obfuscated program for testing proximity to w and outputting the appropriate value if the test passes. In addition to relying on strong assumptions for security (namely, the existence of semantically-secure multilinear maps), this construction is not of practical efficiency. Note that if this construction is used for a secure sketch, W will remain unpredictable conditioned on p, but will not have pseudoentropy (see Section 4.1 for details).

Furthermore, the functional definition of fuzzy extractors and secure sketches can be weakened to permit interaction between the party having w and the party having w'. Such a weakening is useful for secure remote authentication [BDK+05]. When both interaction and computational assumptions are allowed, secure two-party computation can produce a key that will be secure whenever the distribution W has fuzzy min-entropy. The two-party computation protocol needs to be secure without assuming authenticated channels; it can be built under the assumptions that collision-resistant hash functions and enhanced trapdoor permutations exist [BCL+11].

Correlated rather than close readings A different model for the problem of key derivation from noisy sources does not explicitly consider the distance between w and w', but rather views w and w' as samples of drawn from a correlated pair of random variables. This model is considered in multiple works, including [Wyn75, CK78, AC93, Mau93]; recent characterizations of when key derivation is possible in this model include [RW05] and [TW14].

Organization The remainder of the paper is organized as follows. In Section 2, we cover preliminaries and fuzzy extractor definitions. In Section 3, we construct a fuzzy extractor for every known distribution with fuzzy min-entropy. In Sections 4 and 5 we provide negative results for families of distributions for secure sketches and fuzzy extractors, respectively.

2 Preliminaries

Usually, we use capitalized letters for random variables and corresponding lowercase letters for their samples. Unless otherwise noted logarithms are base 2. The *min-entropy* of W is $\mathrm{H}_\infty(W) = -\log(\max_w \Pr[W = w])$, and the *average (conditional)* min-entropy of W given P is $\tilde{\mathrm{H}}_\infty(W|P) = -\log(\mathbb{E}_{p\in P} \max_w \Pr[W = w|P = p])$ [DORS08, Section 2.4]. Let $H_0(W)$ be the logarithm of the size of the support of W, that is $H_0(W) = \log |\{w| \Pr[W = w] > 0\}|$. We use an average case of remaining support size $\tilde{H}_0(W|P) = \log(\mathbb{E}_{p\in P} |\{w| \Pr[W = w|P = p] > 0\}|)$.

The *statistical distance* between random variables X and Y with the same domain is $\mathbf{SD}(X,Y) = \frac{1}{2}\sum_x |\Pr[X = x] - \Pr[Y = x]|$. For a distinguisher D we write the *computational distance* between X and Y as $\delta^D(X,Y) = |\mathbb{E}[D(X)] - \mathbb{E}[D(Y)]|$ (we extend it to a class of distinguishers $\mathcal{D}$ by taking the maximum over all distinguishers $D \in \mathcal{D}$). We denote by $\mathcal{D}_s$ the class of randomized circuits which output a single bit and have size at most s.

For a metric space $(\mathcal{M}, \mathsf{dis})$, the *(closed) ball of radius t around w* is the set of all points within radius t, that is, $B_t(w) = \{w'|\mathsf{dis}(w,w') \le t\}$. If the size of a ball in a metric space does not depend on w, we denote by $|B_t|$ the size of a ball of radius t. We consider the Hamming metric over vectors in $\mathcal{Z}^\gamma$, defined via $\mathsf{dis}(w,w') = \{i|w_i \ne w'_i\}$ where $\mathcal{Z}$ is some alphabet. For this metric, $|B_t| = \sum_{i=0}^{t}\binom{\gamma}{i}(|\mathcal{Z}|-1)^i$. U_κ denotes the uniformly distributed random variable on $\{0,1\}^\kappa$. Throughout this work, we consider a sequence of metric spaces $\mathcal{M}_n$ parameterized by n, but we write $\mathcal{M}$ for notational convenience. A *negligible* function $\mathtt{ngl}(n)$ is one that decreases faster than any positive inverse polynomial as $n \to \infty$.

2.1 Fuzzy Extractors and Secure Sketches

In this section, we define fuzzy extractors and secure sketches. Definitions and lemmas are drawn from the work of Dodis et. al. [DORS08, Sections 2.5–4.1] with modifications. First we allow for error, as discussed in [DORS08, Section 8]. Second, in the *family of distributions* setting we consider an arbitrary family $\mathcal{W}$ of distributions instead of families containing all distributions of a given min-entropy. Let $\mathcal{M}$ be a metric space with distance function dis.

Definition 2.1. *An $(\mathcal{M}, \mathcal{W}, \kappa, t, \epsilon)$-*fuzzy extractor *with error δ is a pair of randomized procedures, "generate"* (Gen) *and "reproduce"* (Rep). Gen *on input $w \in \mathcal{M}$ outputs an extracted string* $\mathsf{key} \in \{0,1\}^\kappa$ *and a helper string $p \in \{0,1\}^*$.* Rep *takes $w' \in \mathcal{M}$ and $p \in \{0,1\}^*$ as inputs.* $(\mathsf{Gen}, \mathsf{Rep})$ *have the following properties:*

1. Correctness: *if* $\mathsf{dis}(w,w') \le t$ *and* $(\mathsf{key}, p) \leftarrow \mathsf{Gen}(w)$*, then* $\Pr[\mathsf{Rep}(w',p) = \mathsf{key}] \ge 1-\delta$ *(note that correctness holds for any w' with probability $1-\delta$ over the coins on* Gen *and* Rep*, but w' cannot be a function of p).*

2. Security: *for any distribution $W \in \mathcal{W}$, if* $(\mathsf{Key}, P) \leftarrow \mathsf{Gen}(W)$*, then* $\mathbf{SD}((\mathsf{Key}, P), (U_\kappa, P)) \le \epsilon$.

Fuzzy extractors perform two tasks, information-reconciliation and privacy amplification. The standard construction is *sketch-and-extract:* the uniform key is extracted from w (using a randomness extractor [NZ93]) and the error-tolerance is obtained by using a secure sketch [DORS08, Lemma 4.1]. Secure sketches produce a string ss that minimally decreases the entropy of w, while mapping nearby w' to w:

Definition 2.2. *An $(\mathcal{M}, \mathcal{W}, \tilde{m}, t)$-*secure sketch *with error δ is a pair of randomized procedures, "sketch"* (SS) *and "recover"* (Rec). SS *on input $w \in \mathcal{M}$ returns a bit string $ss \in \{0,1\}^*$.* Rec *takes an element $w' \in \mathcal{M}$ and $ss \in \{0,1\}^*$.* $(\mathsf{SS}, \mathsf{Rec})$ *have the following properties:*

1. Correctness*: $\forall w, w' \in \mathcal{M}$ if* $\mathsf{dis}(w,w') \le t$ *then* $\Pr[\mathsf{Rec}(w', \mathsf{SS}(w)) = w] \ge 1-\delta$ *(note that correctness holds for any w' with probability $1-\delta$ over the coins of* SS *and* Rec*, but w' cannot be a function of* $\mathsf{SS}(w)$*).*

2. Security*: for any distribution $W \in \mathcal{W}$,* $\tilde{\mathrm{H}}_\infty(W|\mathsf{SS}(W)) \ge \tilde{m}$.

In the above definitions, the errors are chosen before ss (resp., p) is known in order for the correctness guarantee to hold: correctness holds for any w' with probability $1-\delta$ over the coins of the algorithms, but w' cannot be a function of the output of $\mathsf{SS}(w)$.

The Case of Known Distribution If in the above definitions we take $\mathcal{W}$ to be a one-element set containing a single distribution W, then the fuzzy extractor/secure sketch is said to be constructed for a *known distribution*. In this case, we need to require correctness only for w that have nonzero probability[2].

Note that we have no requirement that the algorithms are compact or efficient, and so the distribution can be fully known to them. Finding a natural model of specifying distributions that allows for efficient (yet generic) known distribution constructions of sketches and extractors is an interesting problem.

From Secure Sketches to Fuzzy Extractors A fuzzy extractor can be produced from a *secure sketch* and an *average case randomness extractor*:

Definition 2.3. *Let $\mathcal{M}$, χ be finite sets. A function* $\texttt{ext} : \mathcal{M} \times \{0,1\}^d \to \{0,1\}^\kappa$ *a* $(\tilde{m}, \epsilon)$-average case extractor *if for all pairs of random variables X, Y over $\mathcal{M}, \chi$ such that $\tilde{H}_\infty(X|Y) \geq \tilde{m}$, we have* $\mathbf{SD}((\texttt{ext}(X, U_d), U_d, Y), U_\kappa \times U_d \times Y) \leq \epsilon$.

Lemma 2.4. *Assume* $(\mathsf{SS}, \mathsf{Rec})$ *is an $(\mathcal{M}, \mathcal{W}, \tilde{m}, t)$-secure sketch with error δ, and let* $\texttt{ext} : \mathcal{M} \times \{0,1\}^d \to \{0,1\}^\kappa$ *be a $(\tilde{m}, \epsilon)$-average case extractor. Then the following* $(\mathsf{Gen}, \mathsf{Rep})$ *is an $(\mathcal{M}, \mathcal{W}, \kappa, t, \epsilon)$-fuzzy extractor with error δ:*

- $\mathsf{Gen}(w)$: *generate* $x \leftarrow \{0,1\}^d$, *set* $p = (\mathsf{SS}(w), x), r = \texttt{ext}(w; x)$, *and output* (r, p).
- $\mathsf{Rep}(w', (s, x))$: *recover* $w = \mathsf{Rec}(w', s)$ *and output* $r = \texttt{ext}(w; x)$.

2.2 Fuzzy Min-Entropy: a Necessary Condition

The value p allows everyone, including the adversary, to find the output of $\mathsf{Rep}(\cdot, p)$ on any input w'. Ideally, p should not provide any useful information beyond this ability, and the outputs of Rep on inputs that are too distant from w should provide no useful information, either. In this ideal scenario, the adversary is limited to trying to guess a w' that is t-close to w. Letting w' be the center of the maximum-weight ball in W would be optimal for the adversary. We therefore measure the quality of a source by (the negative logarithm of) this weight.

Definition 2.5. *The t-fuzzy min-entropy of a distribution W in a metric space* $(\mathcal{M}, \mathsf{dis})$ *is:*

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

Fuzzy min-entropy is a necessary condition for security (proof in Appendix B):

Proposition 2.6. *Let W be a distribution over* $(\mathcal{M}, \mathsf{dis})$ *and let $n = \log|\mathcal{M}|$. If $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) = \Theta(\log n)$ there is no $(\mathcal{M}, W, \kappa, t)$-fuzzy extractor with error $\delta = \mathtt{ngl}(n)$ for $\kappa = \omega(\log n)$.*

There are many distributions with $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}$ with no known fuzzy extractor (or corresponding impossibility result). Our goal is to provide a superlogarithmic length key when provided with superlogarithmic fuzzy min-entropy.

[2]We can extend correctness to all of $\mathcal{M}$ by defining Gen/SS to output the point w as part of p/ss on zero-probability inputs, which will ensure that $\mathsf{Rep}/\mathsf{Rec}$ can always be correct; this does not affect security.

3 Sufficiency of $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$ When the Algorithms Know the Distribution

In this section, we show it is possible to build known-distribution secure sketches (and thus fuzzy extractors through Lemma 2.4) whenever $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) = \omega(\log n)$. We first consider flat distributions and show that hashing maintains fuzzy min-entropy and suffices to disambiguate points. We then turn to arbitrary distributions.

3.1 Flat Distributions

A distribution is flat if all points in its support have the same probability. Let $\mathrm{supp}(W)$ denote the support of W, i.e., the set of points with nonzero probability.

Definition 3.1. *A distribution W is* flat *if for all $w_0, w_1 \in \mathrm{supp}(W)$, $\Pr[W = w_0] = \Pr[W = w_1]$.*

Denote the largest number of points in a ball of radius t in the support of W as $\beta_t = \max_{w' \in \mathcal{M}} |\{w | w \in \mathrm{supp}(W) \wedge \mathsf{dis}(w, w') \le t\}|$. For flat distributions, the weight of this maximum-probability ball (which determines $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$ by Definition 2.5) is proportional to the number of points in it. More precisely,

$$\begin{aligned} \mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) &= -\log\left(\max_{w' \in \mathcal{M}} \left|\{w | w \in \mathrm{supp}(W) \wedge \mathsf{dis}(w, w') \le t\}\right| \cdot \Pr[W = w]\right) \\ &= -\log\left(\max_{w' \in \mathcal{M}} |\{w | w \in \mathrm{supp}(W) \wedge \mathsf{dis}(w, w') \le t\}| \cdot 2^{-\mathrm{H}_\infty(W)}\right) \\ &= \mathrm{H}_\infty(W) - \log \beta_t. \end{aligned} \tag{1}$$

We use universal hashes to construct secure sketches for flat distributions. Skoric et al. constructed secure sketches from universal hashes to correct a polynomial number of error patterns [STGP09].

Definition 3.2 ([CW79])**.** *Let $F : \mathcal{K} \times \mathcal{M} \to R$ be a function. We say that F is* universal *if for all distinct $x_1, x_2 \in \mathcal{M}$:*

$$\Pr_{K \leftarrow \mathcal{K}}[F(K, x_1) = F(K, x_2)] = \frac{1}{|R|} \ .$$

Construction 3.3. *Let $F : \mathcal{K} \times \mathcal{M} \to R$ be a universal hash function. Let W be a distribution. Define $\mathsf{SS}_W, \mathsf{Rec}_W$ as:*

SS_W

1. *Input: w.*
2. *Sample $K \leftarrow \mathcal{K}$.*
3. *Set $p = F(K, w), K$.*

Rec_W

1. *Input: $(w', p = y, K)$*
2. *Let $W^* = \{w \in \mathrm{supp}(W) | \mathsf{dis}(w, w') \le t\}$.*
3. *For $w^* \in W^*$, if $F(K, w^*) = y$, output w^*.*
4. *Output $\perp$.*

Lemma 3.4. *(Proof in Appendix C) Let W be a flat distribution with $\mathrm{H}_\infty(W) \ge m$. Then Construction 3.3 is a $(\mathcal{M}, \{W\}, m - \log|R|, t)$-known distribution secure sketch with error $\delta \le \frac{\beta_t - 1}{|R|}$.*

Corollary 3.5. *Let $n = \log|\mathcal{M}|$. If $|R| \ge |\beta_t| \cdot n^{\omega(1)}$ then Construction 3.3 is correct with overwhelming probability. That is, setting $\log|R| = \log \beta_t + \omega(\log n)$ suffices.*

Construction 3.3 writes down enough information to disambiguate any ball of points. The remaining entropy for this construction is $\tilde{\mathrm{H}}_\infty(W|\mathsf{SS}(W)) = \mathrm{H}_\infty(W) - \log \beta_t - \omega(\log n)$. For a flat distribution this is within a superlogarithmic factor of optimal (see Equation (1)). By choosing δ based on $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$ we build $(\mathsf{SS}, \mathsf{Rec})$ such that $\tilde{\mathrm{H}}_\infty(W|\mathsf{SS}(W)) = \omega(\log n)$.

3.2 Arbitrary Distributions

The worst-case hashing approach does not work for arbitrary sources. The reason is that some balls may have many points but low total weight. For example, let W be a distribution consisting of the following balls. Denote by B_t^1 a ball with $2^{\mathrm{H}_\infty(W)}$ points with probability $\Pr[W \in B_t^1] = 2^{-\mathrm{H}_\infty(W)}$. Let $B_t^2, ..., B_t^{2^{-\mathrm{H}_\infty(W)}}$ be balls with one point each with probability $\Pr[W \in B_t^i] = 2^{-\mathrm{H}_\infty(W)}$. Then the hashing algorithm needs to write down $\mathrm{H}_\infty(W)$ bits to achieve correctness on B_t^1. However, with probability $1 - 2^{-\mathrm{H}_\infty(W)}$ the initial reading is outside of B_t^1, and the hash completely reveals the point.

Dealing with non-flat distributions requires a new strategy. Many solutions for manipulating high entropy distributions leverage a solution for flat distributions and use the fact that high entropy distributions are convex combinations of flat distributions. However, a distribution with high fuzzy min-entropy may be formed from component distributions with little or no fuzzy min-entropy. It is unclear how to leverage the convex combination property in this setting.

The main obstacle in the arbitrary setting is distinguishing between a setting where a ball has a few high probability points and a large number of low probability points. To overcome this problem, we write the probability of $w \in W$ in the sketch output. To ensure this information does not completely reveal w we write down $\lfloor \log \Pr[W = w] \rfloor$. We then use a universal hash whose output length is proportional to the number of close points of the same probability as w. This construction divides the distribution W into probability levels. Each level is nearly flat.

Construction 3.6. *Let $\mathcal{M}$ be a metric space and let $n = \log |\mathcal{M}|$. Let W be a distribution with $\mathrm{H}_\infty(W) = m$. Let $\ell \in \mathbb{Z}^+$ be a parameter. Let $L_i = (2^{-(i+1)}, 2^{-i}]$ for $i = m, ..., m + \ell$. Let $F_i : \mathcal{K}_i \times \mathcal{M} \to R_i$ be a parameterized family of universal hash functions. Define $\mathsf{SS}_W, \mathsf{Rec}_W$ as:*

SS_W

1. *Input:* w.
2. *If* $\Pr[W = w] \le 2^{-(m+\ell)}$. *Set* $p = 0, w$.
3. *Else*
   (a) *Find* i *such that* $\Pr[W = w] \in L_i$.
   (b) *Sample* $K \leftarrow \mathcal{K}_i$.
   (c) *Set* $ss = 1, i, F_i(K, w), K$.

Rec_W

1. *Input:* (w', ss)
2. *If* $ss_0 = 1$, *output* $ss_{1,...,|y|}$.
3. *Else*
   (a) *Parse* $(i, y, K) = ss_{1,...,|y|}$.
   (b) $W^* = \{w \in \mathrm{supp}(W) | \mathsf{dis}(w, w') \le t \wedge \Pr[W = w] \in L_i\}$.
   (c) *For* $w^* \in W^*$, *if* $F_i(K, w^*) = z$, *output* w^*.
   (d) *Output* $\perp$.

We extend our notation for the maximum likelihood ball to the leveled case. Define $\beta_{t,i}$ as the maximum number of points in a ball in level i. That is,

$$\beta_{t,i} = \max_{w' \in \mathcal{M}} \left| \{w | w \in \mathrm{supp}(W) \wedge \mathsf{dis}(w, w') \le t \wedge \Pr[W = w] \in L_i\} \right|.$$

Theorem 3.7. *Let W be a distribution over $\mathcal{M}$ where $n = \log \mathcal{M}$. Let $\delta > 0$ be an function of n. Let $F_i : \mathcal{K}_i \times \mathcal{M} \to R_i$ be a parameterized family of universal hash functions where $|R_i| = (\beta_{t,i} - 1)/\delta$. When $\ell = n$ Construction 3.6 is a $(\mathcal{M}, \{W\}, \tilde{m}, t)$-known distribution secure sketch with error δ for $\tilde{m} = \mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) - \log n - \log 1/\delta - 3$.*

We provide a proof in Appendix D. The main idea is to show that an adversary that knows the level of w cannot perform much better than an adversary without this information.

Corollary 3.8. *Let $\mathcal{M}$ be a metric space where $n = \log |\mathcal{M}|$. For any distribution W over $\mathcal{M}$ with $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) = \omega(\log n)$, there exists a $(\mathcal{M}, \{W\}, \tilde{m}, t)$-known distribution secure sketch with $\tilde{m} = \omega(\log n)$ and $\delta = \mathtt{ngl}(n)$. (Extendible to a fuzzy extractor using Lemma 2.4.)*

In our positive results we considered an arbitrary finite metric space. The relevant parameters were the size of balls and the size of the metric space. Our negative results are specific to the Hamming metric.

4 Impossibility of Secure Sketches for a Family with $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}$

In the previous section, we showed the sufficiency of $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$ for known distribution algorithms. Unfortunately, it is unrealistic to assume that W is completely known. Traditionally, algorithms deal with this uncertainty by providing security for a family of distributions $\mathcal{W}$.

In this section, we show uncertainty of W comes at a real cost. The security game of a fuzzy extractor can be thought of as a three stage process: 1) the challenger specifies $(\mathsf{SS}, \mathsf{Rec})$, 2) the adversary sees $(\mathsf{SS}, \mathsf{Rec})$ and specifies $W \in \mathcal{W}$ 3) the adversary wins if $\tilde{\mathrm{H}}_\infty(W|\mathsf{SS}(W)) < \tilde{m}$. We prove impossibility in a game that is harder for the adversary to win: 1) the challenger specifies $(\mathsf{SS}, \mathsf{Rec})$ 2) the adversary samples a random distribution from $W \leftarrow \mathcal{W}$ 3) the adversary wins if $\tilde{\mathrm{H}}_\infty(W|\mathsf{SS}(W)) < \tilde{m}$.

Let V be the process of uniformly sampling $W \leftarrow \mathcal{W}$ and then sampling $w \leftarrow W$. Let the random variable Z indicate which W was sampled. The view of the challenger is V, while the view of the adversary is a distribution $V|Z$. Our results rule out security for an average member of $\mathcal{W}$. It may be possible to improve parameters by ruling out only a worst case W. In Appendix A, we show that providing security for a family $\mathcal{W}$ is equivalent to providing security for all distributions over that family.

We now show a family of distributions $\mathcal{W}$ that does not admit a secure sketch.

Theorem 4.1. *Let n be a security parameter. There exists a family of distributions $\mathcal{W}$ such that for each element $W \in \mathcal{W}$, $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) = \omega(\log n)$, and yet for any $(\mathcal{M}, \mathcal{W}, \tilde{m}, t)$-secure sketch $(\mathsf{SS}, \mathsf{Rec})$ with error $\delta < 1/4$ and distance $t \geq 4$, the remaining entropy $\tilde{m} < 2$.*

Furthermore, this is true on average. Let V be process of uniformly sampling $W \leftarrow \mathcal{W}$ and sampling $w \leftarrow W$, and let Z indicate which W is sampled. Then

$$\tilde{\mathrm{H}}_\infty(V|\mathsf{SS}(V), Z) < 2.$$

Proof. We prove the stronger average case statement. We first describe a family $\mathcal{W}$. Let $\mathbb{F}$ be some field of size $q = \omega(\mathtt{poly}(n))$. Let $\mathcal{W}$ be the set of all distributions of the form

$$W = \begin{pmatrix} \vec{1} \\ a_2 \\ \vdots \\ a_\gamma \end{pmatrix} W_1 + \begin{pmatrix} 0 \\ b_2 \\ \vdots \\ b_\gamma \end{pmatrix}$$

where W_1 is uniform and $W_i = a_i W_1 + b_i$ for $2 \le i \le \gamma$ and $a_i, b_i \in \mathbb{F}, a_i \neq 0$. This type of distribution is an affine line in space $\mathbb{F}^\gamma$. Define V as the process of uniformly choosing $W \leftarrow \mathcal{W}$ and then sampling from $w \leftarrow W$. The adversary sees $\mathsf{SS}(V)$ and Z. Z is the description of the line $Z = a_2, b_2, ..., a_\gamma, b_\gamma$. The algorithms $\mathsf{SS}, \mathsf{Rec}$ never see Z. Fix some $4 \le t < \gamma$. We show the following (in Appendix E):

- Proposition E.2: for all $W \in \mathcal{W}$, $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) = \omega(\log n)$. That is, $\forall z, \mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(V|Z = z) = \omega(\log n)$.
- Proposition E.3: the distribution V is uniform.
- Lemma E.4: for any secure sketch on V, the support size of $V|\mathsf{SS}(V)$ decreases significantly. Here we show the minimum distance of $V|\mathsf{SS}(V)$ is at least t.
- Lemma E.5: for most lines Z, the intersection of the support of $V|Z$ and $V|\mathsf{SS}(V)$ is small.
That is, $\tilde{H}_0(V|\mathsf{SS}(V), Z) < 2$.

□

Note: There is a tradeoff between the size of $\mathbb{F}$ and the error tolerance required for the counter example. By increasing t it is possible to show a counter example for a smaller $\mathbb{F}$.

4.1 Implications for Computational Secure Sketches

Fuller et al. showed that computational secure sketches that provide pseudoentropy imply information-theoretic secure sketches with almost the same parameters [FMR13, Corollary 3.8]. The definition of Fuller et al. uses a weak version of pseudoentropy [HILL99] due to Gentry and Wichs [GW11].

Definition 4.2. *Let (W, S) be a pair of random variables. W has* relaxed HILL entropy *at least $\tilde{m}$ conditioned on S, denoted $H^{\mathtt{HILL\text{-}rlx}}_{\epsilon_{sec}, s_{sec}}(W|S) \ge \tilde{m}$ if there exists a joint distribution (X, Y), such that $\tilde{H}_\infty(X|Y) \ge \tilde{m}$ and $\delta^{\mathcal{D}_{s_{sec}}}((W, S), (X, Y)) \le \epsilon_{sec}$.*

A *HILL secure sketch* is obtained by replacing the security condition in Definition 2.2 with HILL entropy (adding parameters ϵ_{sec}, s_{sec} to the definition, representing the distinguishing circuit). By the contrapositive of [FMR13, Corollary 3.8], no sketch can retain HILL entropy for the same family of distributions:

Corollary 4.3. *Let n be a security parameter and let $\mathcal{M} = |\mathbb{F}|^\gamma$. There exists a family of distributions $\mathcal{W}$ over $\mathcal{M}$ such that for each element $W \in \mathcal{W}$, $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) = \omega(\log n)$ and for any $(\mathcal{M}, \mathcal{W}, \tilde{m}, t)$-HILL secure sketch* $(\mathsf{SS}, \mathsf{Rec})$ *that is $(s_{sec}, \epsilon_{sec})$-hard and error δ. If $s_{sec} \ge t(|\mathsf{Rec}| + \gamma \log |\mathbb{F}|)$, $t \ge 4$, and $\epsilon_{sec} + t\delta < 1/16$, then $\tilde{m} < 4$.*

Secure sketches that provide computational unpredictability are implied by virtual-grey box obfuscation of proximity functions [BCKP14]. Our impossibility result says nothing about this weaker form of a secure sketch. Extraction from unpredictability entropy can be done using an extractor with a reconstruction property [BSW03, HLR07]; however, in the computational setting, the obfuscated function can simply hide a randomly generated key, and therefore extraction is not necessary to obtain a fuzzy extractor.

5 Impossibility of Fuzzy Extractors for a Family with $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}$

In the previous section, we showed a family of distributions that does not admit a secure sketch. We provide a similar result for fuzzy extractors.

Theorem 5.1. *Let n be a security parameter. There exists a family of distributions $\mathcal{W}$ over $\{0,1\}^n$ satisfying the following conditions. For each element $W \in \mathcal{W}$, $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) = \omega(\log n)$. Let $\kappa \geq 2$ and $t = \omega(n^{1/2} \log n)$. Any $(\mathcal{M}, \mathcal{W}, \kappa, t, \epsilon)$-fuzzy extractor with error $\delta = 0$ has $\epsilon > 1/8 - \mathtt{ngl}(n)$.*

Furthermore, this is true on average. Let V be process of uniformly sampling $W \leftarrow \mathcal{W}$ and sampling $w \leftarrow W$ and let Z indicate which W is sampled. Let $(\mathsf{Key}, P) \leftarrow \mathsf{Gen}(V)$. Then,

$$\mathbf{SD}((\mathsf{Key}, P, Z), (U_\kappa, P, Z)) > 1/8 - \mathtt{ngl}(n)\,.$$

Proof Outline. We prove the stronger average case statement. Let $\nu = \omega(\log n)$ and $\nu = o(n^{1/2}/\log n)$. Let $t = 4\nu n^{1/2}$ and note that $n/\nu > t$.

Our counterexample uses a slightly different family of distributions $\mathcal{W}$ than the counterexample for secure sketches. We will work over a binary alphabet (we used a large alphabet in our counterexample for secure sketches). A property of the binary Hamming space is that a large fraction of any set of bounded size is the near "boundary" of that set. This will be crucial in our proof. We will embed the larger alphabet we used into the binary Hamming metric. Let $x_1, ..., x_\nu \in \{0,1\}^\nu$. Let $\mathbb{F}$ denote the field of size 2^ν. Let $a_2, ..., a_{n/\nu} \in \mathbb{F}$ such that $a_i \neq 0$ and let $b_2, ..., b_{n/\nu} \in \mathbb{F}$. Interpret $x_1, ..., x_\nu$ as a element $x \in \mathbb{F}$ and let

$$w = \begin{pmatrix} \vec{1} \\ a_2 \\ \vdots \\ a_{n/\nu} \end{pmatrix} x + \begin{pmatrix} 0 \\ b_2 \\ \vdots \\ b_{n/\nu} \end{pmatrix}.$$

The multiplication is in $\mathbb{F}$. Define a distribution W as the uniform distribution over values of x for a particular value of $a_2, ..., a_{n/\nu}$, $b_2, ..., b_{n/\nu}$. Let $\mathcal{W}$ be the set of all such W.

Define V as the process of uniformly choosing $W \leftarrow \mathcal{W}$ and then sampling from $w \leftarrow W$. The adversary sees $\mathsf{SS}(V)$ and Z, where Z is the description of the line $Z = a_2, ..., a_{n/\nu}, b_2, ..., b_{n/\nu}$.

We now present an outline of the proof (proofs in Appendix F):

- Proposition F.1: for all $W \in \mathcal{W}$, $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) = \omega(\log n)$. That is, $\forall z, \mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(V|Z = z) = \omega(\log n)$.
- Proposition F.2: the distribution V is uniform.
- Lemma F.3: In expectation across Z, a large subset of keys that are not possible. In more detail,
  - Half the keys have at most $2^{n-\kappa}$ pre images in the metric space (this is at most half the metric space). Denote this set as R_{small}.
  - Consider some $\mathsf{key} \in R_{small}$. Consider the set of $V_{\mathsf{key}} = \{w | \mathsf{Rep}(w, p) = \mathsf{key}\}$. All points in $V | \mathsf{SS}(V)$ are distance t from a boundary of V_{key} (the functionality of Rep guarantees that for the true w all nearby points map to the same key). We show that most of V_{key} is near a boundary. A result of Frankel and Füredi says that the boundary of a region is minimized by a ball containing the same number of points [FF81]. Hoeffding's inequality says that most of a ball lies near its boundary [Hoe63]. Together these two results imply that V_{key} is small.

– As before, there are many possible values for z_1, z_2 for the side information Z (and these possible values are equally likely). Furthermore, the distributions $V|Z = z_1$ and $V|Z = z_2$ have disjoint support outside of v.
– For most values of possible Z, the intersection between the viable pre images of $V|Z$ and V_{key} contains at most one point (the received point v). Checking if $V|Z \cap V_{\mathsf{key}}$ is nonempty is an effective distinguisher.

□

Note: As stated in Section 1.2, using strong computational assumptions it is possible to avoid this result. We note that the specific family $\mathcal{W}$, Canetti et al. [CFP+14, Construction 5.3] construct computational fuzzy extractors for this family of distributions when $\mathbb{F}$ is large enough under weaker assumptions. (Their construction is stated with imperfect correctness. A construction with perfect correctness is obtained by using a code that corrects t bidirectional errors instead of a code that corrects t unidirectional errors.)

Comparison with Theorem 4.1 The parameters in this result are weaker than those in Theorem 4.1. This result requires: 1) higher error tolerance $t = \omega(n^{1/2} \log n)$ 2) the fuzzy extractor must have perfect correctness. The secure sketch counter example needs $t = 4$ and allows the Rec to be wrong almost $1/4$ of the time.

Acknowledgements

The authors are grateful to Gene Itkis and Yevgeniy Dodis for helpful discussions. The work of Benjamin Fuller is sponsored in part by US NSF grants 1012910 and 1012798 and the United States Air Force under Air Force Contract FA8721-05-C-0002. Opinions, interpretations, conclusions and recommendations are those of the authors and are not necessarily endorsed by the United States Government. Leonid Reyzin is supported in part by US NSF grants 0831281, 1012910, 1012798, and 1422965. Adam Smith is supported in part by NSF awards 0747294 and 0941553.

References

[AC93] Rudolf Ahlswede and Imre Csiszár. Common randomness in information theory and cryptography - I: secret sharing. *IEEE Transactions on Information Theory*, 39(4):1121–1132, 1993.

[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 - 34th Annual Cryptology Conference, Santa Barbara, CA, USA, August 17-21, 2014, Proceedings, Part II*, 2014.

[BCL+11] Boaz Barak, Ran Canetti, Yehuda Lindell, Rafael Pass, and Tal Rabin. Secure computation without authentication. *J. Cryptology*, 24(4):720–760, 2011.

[BDH$^+$10] Ileana Buhan, Jeroen Doumen, Pieter H. Hartel, Qiang Tang, and Raymond N. J. Veldhuis. Embedding renewable cryptographic keys into noisy data. *Int. J. Inf. Sec.*, 9(3):193–208, 2010.

[BDK$^+$05] Xavier Boyen, Yevgeniy Dodis, Jonathan Katz, Rafail Ostrovsky, and Adam Smith. Secure remote authentication using biometric data. In *EUROCRYPT*, pages 147–163. Springer, 2005.

[BH09] Marina Blanton and William MP Hudelson. Biometric-based non-transferable anonymous credentials. In *Information and Communications Security*, pages 165–180. Springer, 2009.

[BS00] Sacha Brostoff and M.Angela Sasse. Are passfaces more usable than passwords?: A field trial investigation. *People and Computers*, pages 405–424, 2000.

[BSW03] Boaz Barak, Ronen Shaltiel, and Avi Wigderson. Computational analogues of entropy. In *11th International Conference on Random Structures and Algorithms*, pages 200–215, 2003.

[CFP$^+$14] Ran Canetti, Benjamin Fuller, Omer Paneth, Leonid Reyzin, and Adam Smith. Key derivation from noisy sources with more errors than entropy. *In submission*, 2014.

[CK78] Imre Csiszár and János Körner. Broadcast channels with confidential messages. *IEEE Transactions on Information Theory*, 24(3):339–348, 1978.

[CK03] L. Csirmaz and G.O.H. Katona. Geometrical cryptography. In *Proc. International Workshop on Coding and Cryptography*, 2003.

[CW79] Larry Carter and Mark N. Wegman. Universal classes of hash functions. *J. Comput. Syst. Sci.*, 18(2):143–154, 1979.

[CZC04] Yao-Jen Chang, Wende Zhang, and Tsuhan Chen. Biometrics-based cryptographic key generation. In *ICME*, pages 2203–2206. IEEE, 2004.

[Dau04] John Daugman. How iris recognition works. *Circuits and Systems for Video Technology, IEEE Transactions on*, 14(1):21 – 30, January 2004.

[Dau06] J. Daugman. Probing the uniqueness and randomness of iriscodes: Results from 200 billion iris pair comparisons. *Proceedings of the IEEE*, 94(11):1927–1935, 2006.

[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.

[EHMS00] Carl Ellison, Chris Hall, Randy Milbert, and Bruce Schneier. Protecting secret keys with personal entropy. *Future Generation Computer Systems*, 16(4):311–318, 2000.

[FF81] Peter Frankl and Zoltán Füredi. A short proof for a theorem of Harper about Hamming-spheres. *Discrete Mathematics*, 34(3):311–313, 1981.

[FMR13] Benjamin Fuller, Xianrui Meng, and Leonid Reyzin. Computational fuzzy extractors. In *Advances in Cryptology-ASIACRYPT 2013*, pages 174–193. Springer, 2013.

[GCVDD02] Blaise Gassend, Dwaine Clarke, Marten Van Dijk, and Srinivas Devadas. Silicon physical random functions. In *Proceedings of the 9th ACM conference on Computer and communications security*, pages 148–160. ACM, 2002.

[GW11] Craig Gentry and Daniel Wichs. Separating succinct non-interactive arguments from all falsifiable assumptions. *STOC. ACM, New York*, pages 99–108, 2011.

[HAD06] Feng Hao, Ross Anderson, and John Daugman. Combining crypto with biometrics effectively. *Computers, IEEE Transactions on*, 55(9):1081–1088, 2006.

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

[HILL99] Johan Håstad, Russell Impagliazzo, Leonid A. Levin, and Michael Luby. A pseudorandom generator from any one-way function. *SIAM Journal on Computing*, 28(4):1364–1396, 1999.

[HLR07] Chun-Yuan Hsiao, Chi-Jen Lu, and Leonid Reyzin. Conditional computational entropy, or toward separating pseudoentropy from compressibility. In *EUROCRYPT*, pages 169–186, 2007.

[Hoe63] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American statistical association*, 58(301):13–30, 1963.

[IW12] Tanya Ignatenko and Frans M.J. Willems. Biometric security from an information-theoretical perspective. *Foundations and Trends® in Communications and Information Theory*, 7(2–3):135–316, 2012.

[JW99] Ari Juels and Martin Wattenberg. A fuzzy commitment scheme. In *Sixth ACM Conference on Computer and Communication Security*, pages 28–36. ACM, November 1999.

[LC06] Qiming Li and Ee-Chien Chang. Robust, short and sensitive authentication tags using secure sketch. In *ACM Multimedia Security Workshop*, 2006.

[LT03] J.-P. M. G. Linnartz and P. Tuyls. New shielding functions to enhance privacy and prevent misuse of biometric templates. In *AVBPA*, pages 393–402, 2003.

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

[MG09] Rene Mayrhofer and Hans Gellersen. Shake well before use: Intuitive and secure pairing of mobile devices. *IEEE Transactions on Mobile Computing*, 8(6):792–806, 2009.

[MRW02] Fabian Monrose, Michael K Reiter, and Susanne Wetzel. Password hardening based on keystroke dynamics. *International Journal of Information Security*, 1(2):69–83, 2002.

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

[PRTG02] Ravikanth Pappu, Ben Recht, Jason Taylor, and Neil Gershenfeld. Physical one-way functions. *Science*, 297(5589):2026–2030, 2002.

[RW05] Renato Renner and Stefan Wolf. Simple and tight bounds for information reconciliation and privacy amplification. In Bimal K. Roy, editor, *Advances in Cryptology - ASIACRYPT 2005, 11th International Conference on the Theory and Application of Cryptology and Information Security, Chennai, India, December 4-8, 2005, Proceedings*, volume 3788 of *Lecture Notes in Computer Science*, pages 199–216. Springer, 2005.

[SD07] G. Edward Suh and Srinivas Devadas. Physical unclonable functions for device authentication and secret key generation. In *Proceedings of the 44th annual Design Automation Conference*, pages 9–14. ACM, 2007.

[STGP09] Boris Skoric, Pim Tuyls, Jorge Guajardo, and Bart Preneel. An efficient fuzzy extractor for limited noise. *Foundations*, 2009.

[TG04] Pim Tuyls and Jasper Goseling. Capacity and examples of template-protecting biometric authentication systems. In *Biometric Authentication*, volume 3087 of *Lecture Notes in Computer Science*, pages 158–170. Springer, 2004.

[TSv+06] Pim Tuyls, Geert-Jan Schrijen, Boris Škoriá, Jan Geloven, Nynke Verhaegh, and Rob Wolters. Read-proof hardware from protective coatings. In *Cryptographic Hardware and Embedded Systems - CHES 2006*, pages 369–383. 2006.

[TW14] Himanshu Tyagi and Shun Watanabe. Converses for secret key agreement and secure computing. *CoRR*, abs/1404.5715, 2014.

[vN28] John von Neumann. Zur theorie der gesellschaftsspiele. *Mathematische Annalen*, 100(1):295–320, 1928.

[VTO+10] Evgeny A. Verbitskiy, Pim Tuyls, Chibuzo Obi, Berry Schoenmakers, and Boris Skoric. Key extraction from general nondiscrete signals. *IEEE Transactions on Information Forensics and Security*, 5(2):269–279, 2010.

[WRDI11] Ye Wang, Shantanu Rane, Stark C. Draper, and Prakash Ishwar. A theoretical analysis of authentication, privacy and reusability across secure biometric systems. *CoRR*, abs/1112.5630, 2011.

[Wyn75] Aaron D Wyner. The wire-tap channel. *Bell System Technical Journal, The*, 54(8):1355–1387, 1975.

[ZH93] Moshe Zviran and William J. Haga. A comparison of password techniques for multilevel authentication mechanisms. *The Computer Journal*, 36(3):227–237, 1993.

A A Definitional Equivalence

As described in Section 4, our negative results rule out security for an average member of $\mathcal{W}$. It may be possible to significantly improve parameters by only ruling out security for a single member W.

Recall the security game of a fuzzy extractor: 1) the challenger specifies $(\mathsf{SS}, \mathsf{Rec})$, 2) the adversary specifies a source $W \in \mathcal{W}$ 3) The challenger wins if $\tilde{\mathrm{H}}_\infty(W|\mathsf{SS}(W)) \geq \tilde{m}$. Instead of just thinking of the uniform distribution over $\mathcal{W}$, consider an arbitrary distribution V over elements of $\mathcal{W}$. The minimax theorem says we can reverse which of these actions is announced first [vN28] if $\mathcal{A}$ announces V instead of a single element W. That is, the following two player games have the same equilibrium:

Experiment $\mathbf{Exp}_1^{\mathcal{W}}(\mathcal{A},\mathcal{C},\tilde{m})$:	Experiment $\mathbf{Exp}_2^{\mathcal{W}}(\mathcal{A},\mathcal{C},\tilde{m})$:
$(\mathsf{SS},\mathsf{Rec}) \leftarrow \mathcal{C}(\mathcal{W})$	$V \leftarrow \mathcal{A}(\mathcal{W})$
$W \leftarrow \mathcal{A}(\mathcal{W},\mathsf{SS},\mathsf{Rec})$	$(\mathsf{SS},\mathsf{Rec}) \leftarrow \mathcal{C}(V,\mathcal{W})$
If $W \notin \mathcal{W}$, $\mathcal{C}$ wins.	$W \leftarrow V$
If $\tilde{\mathrm{H}}_\infty(W \vert \mathsf{SS}(W)) \geq \tilde{m}$, $\mathcal{C}$ wins.	If $W \notin \mathcal{W}$, $\mathcal{C}$ wins.
Else $\mathcal{A}$ wins.	If $\tilde{\mathrm{H}}_\infty(W \vert \mathsf{SS}(W)) \geq \tilde{m}$, $\mathcal{C}$ wins.
	Else $\mathcal{A}$ wins.

This means that showing security for a family of distributions $\mathcal{W}$ is equivalent to showing security for all distributions V when the distribution is known to the algorithms V. In our negative results, the adversary uses the uniform distribution V over $\mathcal{W}$. However, it may be possible to improve parameters by using a different V. This would just rule out some member of $\mathcal{W}$ not an average member. This is true for fuzzy extractors as well and is resilient to changes in parameters including imperfect correctness.

B Proof of Proposition 2.6

Proof. Let W be a distribution where $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) = \Theta(\log n)$. This means that there exists a point $w' \in \mathcal{M}$ such that $\Pr_{w \in W}[\mathsf{dis}(w, w') \leq t] \geq 1/\mathtt{poly}(n)$. Consider the following distinguisher D:

- Input r, p.
- If $\mathsf{Rep}(w', p) = r$, output 1.
- Else output 0.

Clearly, $\Pr[D(\mathsf{Key}, P) = 1] \geq 1/\mathtt{poly}(n) - \delta$, while $\Pr[D(U_\kappa, P) = 1] = 1/2^{-\kappa}$. Thus, when $\kappa = \omega(\log n)$:

$$\delta^D((R,P),(U_\kappa,P)) \geq \frac{1}{\mathtt{poly}(n)} - \delta - \frac{1}{2^{-\kappa}} = 1/\mathtt{poly}(n).$$

Note that D only provides an input and looks at the output, thus it extends to an interactive protocol. Also, D is of size $\max|\mathcal{M}| + |\mathsf{Rep}|$ where $\max|\mathcal{M}|$ is the longest description of an item in the metric space. Thus, D is also a distinguisher in the computational setting. □

C Proof of Lemma 3.4

Proof. We first argue security. Fix some $W \in \mathcal{W}$. Since $\mathcal{K}$ and W are independent $\tilde{\mathrm{H}}_\infty(W|\mathcal{K}) = \mathrm{H}_\infty(W) = m$. Then by [DORS08, Lemma 2.2b], $\tilde{\mathrm{H}}_\infty(W|\mathcal{K}, F(\mathcal{K}, W)) \geq \mathrm{H}_\infty(W) - \log|F(\mathcal{W}, W)| \geq m - \log|R|$.

We now argue correctness. Fix some w, w'. Let W^* denote the set of elements in W within distance t of w'. The size of W^* is at most β_t. Since w, w' are independent of SS this set is independent of the choice of $\mathcal{K}$. The algorithm Rec will never output $\perp$ as the correct w will match the hash. The probability that another element w^* collides is:

$$\begin{aligned}\Pr[\exists w^* \in W^* | w^* \neq w \wedge F(K, w^*) = F(K, w)] &\leq \sum_{w^* \in W^* | w^* \neq w} \Pr[F(K, w^*) = F(K, w)] \\ &= \sum_{w^* \in W^* | w^* \neq w} \frac{1}{|R|} \leq \frac{\beta_t - 1}{|R|}\end{aligned}$$

The inequality proceeds by union bound. The first equality proceeds by the universality of F and the second inequality proceeds by noting the number of wrong neighbors is bounded by $\beta_t - 1$. This completes the proof. $\square$

D Proof of Theorem 3.7

Proof. Throughout the proof we assume that $\ell = n$ is the number of levels. The proof can be carried out for an arbitrary ℓ but it leads to a complicated theorem statement.

Correctness: Fix some w, w'. If $\Pr[W = w] \le 2^{-(m+\ell)} = 2^{-(m+n)}$, then w is simply transmitted to Rec and correctness is clear. When $\Pr[W = w] > 2^{-(m+n)}$ let L_i^* be the level of $\Pr[W = w]$.

Let W^* denote the set of elements of W in L_i within distance t of w'. The size of W^* is at most $\beta_{t,i}$. The choice of w, w' is independent of SS, so this set is independent of $\mathcal{K}_i$ (it does effect the value of i but not the particular outcome from $\mathcal{K}_i$). The probability that another element w^* matches the hash is:

$$\begin{aligned}\Pr[\exists w^* \in W^* | w^* \neq w \wedge F(K, w^*) = F(K, w)] &\le \sum_{w^* \in W^* | w^* \neq w} \Pr[F(K, w^*) = F(K, w)] \\ &= \sum_{w^* \in W^* | w^* \neq w} \frac{1}{|R_i|} \le \frac{\beta_{t,i} - 1}{|R_i|} = \delta\end{aligned}$$

The inequality is by union bound. The first equality follows from the universality of F. The second inequality follows since the number of neighbors is bounded by $\beta_{t,i}$.

Ideal Adversary with access to Level Information: To aid in the argument in security, we show the level information on its own is not too harmful.

The best strategy for an adversary that receives i as is to guess a point that has the most nearby weight in that level. The adversary chooses

$$w^* = \arg\max_{w' \in \mathcal{M}} \Pr_{w \in W | 2^{-(i+1)} < \Pr[W = w] \le 2^{-i} \wedge \mathsf{dis}(w, w^*)} [W = w].$$

The success of this adversary is at least $2^{-(i+1)}\beta_{t,i}$ as there at $\beta_{t,i}$ nearby points in that layer each with probability at least $2^{-(i+1)}$. There are n outcomes for i. The overall success of such an adversary is at most n better than an adversary without such input (by [DORS08, Lemma 2.2]). That is,

$$\begin{aligned}\mathop{\mathbb{E}}_{i | m \le i \le m+n} 2^{-(i+1)}\beta_{t,i} &\le \mathop{\mathbb{E}}_{i | m \le i \le n+m} \left(\max_{w^* \in W} \sum_{w \in W | 2^{-(i+1)} < \Pr[W = w] \le 2^{-i} \wedge \mathsf{dis}(w, w^*) \le t} \Pr[W = w] \right) \\ &\le n \left(\max_{w^* \in W} \sum_{w \in W | \mathsf{dis}(w, w^*) \le t} \Pr[W = w] \right) \\ &= n 2^{-\mathrm{H}^{\mathsf{fuzz}}_{t,\infty}(W)} \qquad (2)\end{aligned}$$

Security: We now argue security. First note that the total weight of points whose probability is less than $2^{-(n+m)}$ is at most 2^{-m} (there are at most 2^n points in the distribution). Let 1_{low} be the indicator random variable for $\Pr[W = w] \le 2^{-(n+m)}$. Then

$$\begin{aligned}\tilde{\mathrm{H}}_\infty(W | \mathsf{SS}(W)) &= -\log \left(\Pr[1_{\mathrm{low}} = 1] * 1 + \Pr[1_{\mathrm{low}} = 0] 2^{-\tilde{\mathrm{H}}_\infty(W | \mathsf{SS}(W) \wedge 1_{\mathrm{low}} = 0)} \right) \\ &\quad - \log \left(2^{-m} + (1 - 2^{-m}) 2^{-\tilde{\mathrm{H}}_\infty(W | \mathsf{SS}(W) \wedge 1_{\mathrm{low}} = 0)} \right)\end{aligned}$$

For the remainder of the proof, we seek a bound on

$$2^{-\tilde{\mathrm{H}}_\infty(W|\mathsf{SS}(W)\wedge 1_{\mathrm{low}}=0} = \max_{w\in W|2^{-(n+m)}<\Pr[W=w]} \Pr[W=w|\mathsf{SS}(W)].$$

We separate out this quantity into levels:

$$\begin{aligned}
\max_{w\in W|\Pr[W=w]>2^{-(m+n)}} (\Pr[W=w|\mathsf{SS}(W)]) &= \mathop{\mathbb{E}}_{i|m\le i\le m+n}\left(\max_{w\in W|\Pr[W=w]\in L_i} \Pr[W=w|\mathsf{SS}(W),i]\right)\\
&= \mathop{\mathbb{E}}_{i|m\le i\le m+n}\left(\max_{w\in W|\Pr[W=w]\in L_i} \Pr[W=w]\cdot 2^{|\mathsf{SS}(W)|i|}\right)\\
&\le \mathop{\mathbb{E}}_{i|m\le i\le m+n}\left(\max_{w\in W|\Pr[W=w]\in L_i} \Pr[W=w]\cdot 2^{H_0(\mathsf{SS}(W)|i)}\right)\\
&\le \mathop{\mathbb{E}}_{i|m\le i\le m+n}\left(2^{-i} * \beta_{t,i}/\delta\right)\\
&\le \frac{\mathbb{E}_{i|m\le i\le m+n}\left(2^{-(i+1)}\cdot\beta_{t,i}\right)}{2\delta}\\
&= \frac{n2^{-\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)}}{2\delta}.
\end{aligned}$$

Where the last line follows by Equation (2). Combining both cases we have:

$$\begin{aligned}
\tilde{\mathrm{H}}_\infty(W|\mathsf{SS}(W)) &= -\log\left(2^{-m} + \frac{(1-2^{-m})(n)2^{-\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)}}{2\delta}\right)\\
&\ge -\log\min\{2^{-m}, \frac{(1-2^{-m})n2^{-\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)}}{2\delta}\}) - 1\\
&\ge \mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) - \log n + \log\delta - \log(1-2^{-m}) - 2\\
&\ge \mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) - \log n + \log\delta - 3
\end{aligned}$$

Where the third line follows from the second because $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) \le \mathrm{H}_\infty(W) = m$. The last line follows from the fourth because if $m \ge 1$ then $\log(1-2^{-m}) \le 1$ and if $m < 1$ the entire bound is vacuous as $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) < 1$. □

E Proof of Theorem 4.1

Let $c' \leftarrow \mathsf{Neigh}_t(c)$ sample a uniform point within distance t of c. The proof of Theorem 4.1 uses the definition of a Shannon code:

Definition E.1. *Let C be a set over space $\mathcal{M}$. We say that C is an (t,δ)-*Shannon code *if there exists a procedure* Rec *such that for all $t' \le t$ and for all $c \in C$, $\Pr[c' \leftarrow \mathsf{Neigh}_t(c) \wedge \mathsf{Rec}(c') \ne c] \le \delta$.*

We now prove item in the outline of Theorem 4.1.

Proposition E.2. *For each $W \in \mathcal{W}$, $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) = \omega(\log n)$.*

Proof. Consider some $W \in \mathcal{W}$. The value w_1 is uniform in a field of size $\omega(\texttt{poly}(n))$, so $\mathrm{H}_\infty(W) = \omega(\log n)$. We now show that for any $w, w' \in W$, $\mathsf{dis}(w, w') = \gamma > t$. This shows that $\mathrm{H}^{\texttt{fuzz}}_{t,\infty}(W) = \mathrm{H}_\infty(W)$. Fix some $w, w' \in W$. Clearly, $w_1 \neq w'_1$, for any i, $w_i = a_i w_1 + b_i$ and $w'_i = a_i w'_1 + b_i$. Since $a_i \neq 0$, $a_i w_1 \neq a_i w'_1$ and thus $a_i w_1 + b_i \neq a_i w'_1 + b_i$. That is, $\mathsf{dis}(w, w') = \gamma$. □

Proposition E.3. *V is the uniform distribution over $\mathbb{F}^\gamma$.*

Proof. Consider some $w \in V$. Then w was drawn from some intermediate distribution W with coefficients $a_2, b_2, ..., a_\gamma, b_\gamma$. The value w_1 is uniformly random and w_i are uniformly random since $b_2, ..., b_\gamma$ are uniformly random. □

Lemma E.4. *Fix some* $\mathsf{SS}, \mathsf{Rec}$ *algorithm with error* $\delta < 1/4$*, then* $\tilde{H}_0(V|\mathsf{SS}(V)) \leq (\gamma - t + 1)\log|\mathbb{F}| + 1$.[3]

Proof. We assume that Rec is deterministic in our analysis. Any randomness necessary for the Rec algorithm can be provided by SS. This is the same as considering Rep that outputs any coin it flips. Since w, w' are independent of p this does not effect correctness. Security is defined based on the output of Rec so outputting the coins of Rep does not effect security. By the definition of correctness for $(\mathsf{SS}, \mathsf{Rec})$,

$$\forall w, w', \Pr_{ss \leftarrow \mathsf{SS}(w)}[\mathsf{Rec}(w', ss) \neq w] < 1/4.$$

Fix some w. By Markov's inequality, there exists a set A_{ss} such that $\Pr[ss \in A_{ss}] \geq 1/2$ and $\forall ss \in A_{ss}$,

$$\{w' | \mathsf{dis}(w', w) \leq t \wedge \mathsf{Rec}(w', p) \neq w\} \leq 2\delta < 1/2.$$

Consider some $ss^* \in A_{ss}$. We now show that $H_0(V|\mathsf{SS}(V) = ss^*) \leq (\gamma - t + 1)\log|\mathbb{F}|$. For the sketched value w, $\{w' | \mathsf{dis}(w, w') \leq t \wedge \mathsf{Rec}(w', p) \neq w] \leq 2\delta$.

For every value in $V|\mathsf{SS}(V) = ss^*$ this is also true. This makes the support of $V|\mathsf{SS}(V) = ss^*$ a $(t, 2\delta)$-Shannon code (see Definition E.1). This implies that for all $w_1, w_2 \in V|\mathsf{SS}(V) = ss^*$, $\mathsf{dis}(w_1, w_2) \geq t$ (since $2\delta < 1/2$). That is $V|\mathsf{SS}(V) = ss^*$ is a set with minimum distance at least t.

By the Singleton bound, this implies that $H_0(V|\mathsf{SS}(V) = ss^*) \leq (\gamma - t + 1)|\mathbb{F}|$. Averaging over $\mathsf{SS}(V) = ss^*$ one has that $\tilde{H}_0(V|P) \leq (\gamma - t + 1)\log|\mathbb{F}| + 1$. □

Lemma E.5. $\tilde{\mathrm{H}}_\infty(V|\mathsf{SS}(V), Z) < 2$.

Proof. Recall that Z consists of 2γ coefficients and there are $(|\mathbb{F}| - 1)^{\gamma-1}|\mathbb{F}|^{\gamma-1}$ equally likely values for Z. As described above, the view of $\mathsf{SS}, \mathsf{Rec}$ is a uniform distribution V. The only information seen by SS algorithm is in the point $V = v$. The length of this point is $|\mathbb{F}|^\gamma$. Conditioned on this information there are still many possible values for Z. That is,

$$\forall v, H_0(Z|V = v) = \log\left(\frac{(|\mathbb{F}| - 1)^{\gamma-1}|\mathbb{F}|^{\gamma-1}}{|\mathbb{F}|^\gamma}\right) = \log\left((|\mathbb{F}| - 1)^{\gamma-1}/|\mathbb{F}|\right).$$

Consider two possible z_1, z_2 that are possible values of Z (having seen v). The distributions $V|Z = z_1$ and $V|Z = z_2$ intersect at one point (namely v).

We now show for any sketch algorithm there are few possible values of $V|Z$ in the support of $V|\mathsf{SS}(V)$. The distributions $V|Z = z_1$ and $V|Z = z_2$ for possible z_1, z_2 (having seen v) overlap only at the point

[3] This result is an extension of lower bounds from [DORS08, Appendix C]. Dealing with imperfect correctness makes the bound more complicated. In particular, we can only argue about the average remaining support size.

v. This means for any $v^* \in V|\mathsf{SS}(V)$ (other than the true v) there is at most one z such that $v^* \in V|\mathsf{SS}(V), Z = z$.

The optimum strategy is to include these values uniformly from different Z values. We show this across different sketch values. Consider some fixed sketch value s and let $h_s = H_0(V|\mathsf{SS}(V) = s)$. Recall that

$$\tilde{H}_0(V|\mathsf{SS}(V)) = \log \mathop{\mathbb{E}}_{s\in\mathsf{SS}(V)} 2^{H_0(V|\mathsf{SS}(V)=s)} = \log \mathop{\mathbb{E}}_{s\in\mathsf{SS}(V)} 2^{h_s}$$

Conditioned on seeing the point V there are $(|\mathbb{F}|-1)^{\gamma-1}/|\mathbb{F}|$ possible values for Z with disjoint support outside of the sketched point. Consider these possible values for Z as containers to be filled with the $2^{h_{ss}}$ items (possible values of $V|\mathsf{SS}(V) = ss$). Each container receives automatically receives one free point (all the distributions share v). The average number of items in each container is maximized when the containers are filled equally. That is, the average number of items in each container is bounded by the number of items divided by the number of container. That is,

$$\begin{aligned}\tilde{H}_0(V|Z, \mathsf{SS}(V) = ss) &\le \log\left(\frac{\#\text{ items} + \#\text{ containers}}{\#\text{ containers}}\right)\\ &= \log\left(\frac{2^{h_{ss}}|\mathbb{F}|}{(|\mathbb{F}|-1)^{\gamma-1}} + 1\right)\end{aligned}$$

Then averaging over the possible values of s, we have the following as long as $t \ge 4$ (using Lemma E.6, which appears below):

$$\begin{aligned}\tilde{H}_0(V|Z,\mathsf{SS}(V)) &= \log \mathop{\mathbb{E}}_{s\in\mathsf{SS}(V)} 2^{\tilde{H}_0(V|\mathsf{SS}(V)=ss,(Z|\mathsf{SS}(V)=ss))}\\ &= \log \mathop{\mathbb{E}}_{s\in\mathsf{SS}(V)} \left(\frac{2^{h_s}|\mathbb{F}|}{(|\mathbb{F}|-1)^{\gamma-1}} + 1\right)\\ &\le \max\left\{\log\left(\frac{|\mathbb{F}|}{(|\mathbb{F}|-1)^{\gamma-1}} \mathop{\mathbb{E}}_{s\in\mathsf{SS}(V)} 2^{h_s}\right) + 1, 1\right\}.\end{aligned}$$

Where the inequality follows because $\log x + 1 \le \max\{1 + \log x, 1\}$ for $x \ge 0$. The left operand to max is bounded by 2 (bounding the max by 2):

$$\begin{aligned}\log\left(\frac{|\mathbb{F}|}{(|\mathbb{F}|-1)^{\gamma-1}} \mathop{\mathbb{E}}_{s\in\mathsf{SS}(V)} 2^{h_s}\right) + 1 &= \log|\mathbb{F}| - (\gamma-1)\log(|\mathbb{F}|-1) + \log\left(\mathop{\mathbb{E}}_{s\in\mathsf{SS}(V)} 2^{h_s}\right) + 1\\ &= \log|\mathbb{F}| - (\gamma-1)\log(|\mathbb{F}|-1) + \tilde{H}_0(V|\mathsf{SS}(V)) + 1\\ &\le \log|\mathbb{F}| - (\gamma-1)\log(|\mathbb{F}|-1) + (\gamma - t + 1)\log|\mathbb{F}| + 2\\ &\le (\gamma - t + 2)\log|\mathbb{F}| - (\gamma-1)\log(|\mathbb{F}|-1) + 2\\ &< (\gamma - t + 2)\log|\mathbb{F}| - (\gamma-2)\log|\mathbb{F}| + 2 \quad \text{(by Lemma E.6)}\\ &\le (4-t)\log|\mathbb{F}| + 2 < 2\,.\end{aligned}$$

□

Lemma E.6. *For any real numbers $\alpha \le \eta$ with $\eta \ge e + 1$ (in particular, $\eta \ge 4$ suffices), the following holds: $\alpha\log(\eta - 1) > (\alpha - 1)\log\eta$.*

Proof. Because $\eta - 1$ is positive, and $1 + x < e^x$ for positive x,

$$1 + \frac{1}{\eta - 1} < e^{\frac{1}{\eta-1}}\,.$$

Therefore,

$$\left(1 + \frac{1}{\eta - 1}\right)^{\alpha-1} < e^{\frac{\alpha-1}{\eta-1}} \leq e < \eta - 1$$

(since $\alpha \leq \eta$). Multiplying both sides by $(\eta - 1)^{\alpha-1}$, we obtain

$$\eta^{\alpha-1} < (\eta - 1)^{\alpha}\,.$$

Taking the logarithm of both sides yields the statement of the lemma. □

F Proof of Theorem 5.1

Proposition F.1. *For each $W \in \mathcal{W}$, $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) = \omega(\log n)$.*

Proof. Consider some fixed $W \in \mathcal{W}$. The bits $w_{1,...,\nu}$ are uniform, so $\mathrm{H}_\infty(W) = \omega(\log n)$. Recall that $t = o(n/\nu)$. Fix some $w, w' \in W$. Denote by x, x' the values that produce w, w' respectively. Clearly, $x \neq x'$. Thus, for any i, $a_i x + b_i \neq a_i x' + b_i$. This implies that $w_{i\nu+1,....,(i+1)\nu} \neq w'_{i\nu+1,...,(i+1)\nu}$. That is, at least one of the bits in each block differs between w and w', and so $\mathsf{dis}(w, w') \geq n/\nu$. Since no two values in the support of W lie in the same ball of radius t, we have $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) = \mathrm{H}_\infty(W) = \omega(\log n)$. □

Proposition F.2. *V is the uniform distribution over $\mathbb{F}^\gamma$.*

Proof. Consider some $w \in V$ over $\{0,1\}^n$. Then $w \leftarrow W$ with coefficients $a_2, b_2, ..., a_\gamma, b_\gamma$. The value $w_{1,...,\nu} = x$ is uniformly random and $w_{i\nu+1,...,(i+1)\nu}$ are uniformly random since $b_2, ..., b_\gamma$ are random. □

Lemma F.3. *Fix some* $(\mathsf{Gen}, \mathsf{Rep})$ *algorithm with $\kappa \geq 2$. There exists an information-theoretic distinguisher between (R, P, Z) and (U_κ, P, Z) with advantage $\epsilon = 1/8 - \mathtt{ngl}(n)$.*

Proof. As in the proof of Theorem 4.1, we assume that Rep is deterministic. Denote by $(\mathsf{Key}, P) \leftarrow \mathsf{Gen}(V)$. By Markov's inequality, there exists a set A_p such that $\Pr[p \in A_p] \geq 1/2$ and $\forall p \in A_p$,

$$(\mathsf{Key}|P = p, P = p) \approx_{2\epsilon} (U_\kappa, P = p).$$

Consider some $p^* \in A_p$. The distribution $\mathsf{Key}|P = p^*$ is the set of possible keys. The distribution $\mathsf{Key}|P = p^*$ induces a partition on the metric space. That is, for every $w \in \mathcal{M}$, there exists a unique value key such that $\mathsf{Rep}(w, p^*) = \mathsf{key}$. Denote this partition by $Q_{p^*,\mathsf{key}} = \{w | \mathsf{Rep}(w, p^*) = \mathsf{key}\}$.

There exists a set R_{small} where $|R_{small}| \geq 2^{\kappa-1}$ such that for all $\mathsf{key} \in R_{small}$, $|Q_{p^*,r}| \leq \mathcal{M}/2^\kappa = 2^{n-\kappa}$. If not, then $\cup_{\mathsf{key}} |Q_{p^*,\mathsf{key}}| > |\mathcal{M}|$. For the remainder of the proof we restrict ourselves to elements in R_{small}. Only points that are distance t from points outside of $Q_{p^*,r}$ are viable points in the metric space. These are the interior of $Q_{p^*,r}$:

$$\mathsf{Inter}(Q_{p^*,\mathsf{key}}) = \{w | \mathsf{Rep}(w, p^*) = \mathsf{key} \wedge \forall w', \mathsf{dis}(w, w') \leq t \wedge \mathsf{Rep}(w', p^*) = \mathsf{key}\},$$

We will use the term deficient ball[4]:

[4] In most statements of the isoperimetric inequality, this type of set is simply called a ball. We use the term deficient ball for emphasis.

Definition F.4. *A set S is a η-deficient ball if there exists a point x such that $B_{\eta-1}(x) \subseteq S \subseteq B_\eta(x)$.*

Consider some $\mathsf{key}^* \in R_{small}$. We now proceed to show that the interior of each Q_{p^*,key^*} is small:

Lemma F.5. $|\mathsf{Inter}(Q_{p^*,\mathsf{key}^*})| \leq 2^{n-4\nu}$.

Proof. By the isoperimetric inequality on the Hamming space (we use a version due to [FF81, Theorem 1], the original result is due to Harper [Har66]), there exists a η-deficient ball S_{p^*,key^*} centered at 0 and a set D such that $|S_{p^*,\mathsf{key}^*}| = |\mathsf{Inter}(Q_{p^*,\mathsf{key}^*})|$, $|D| = |Q^{\complement}_{p^*,\mathsf{key}^*}|$ and $\forall s \in S_{p^*,\mathsf{key}^*}, d \in D$, $\mathsf{dis}(s,d) \geq t$ (alternatively, the distance between the sets is t). Furthermore, note that $S_{p^*,\mathsf{key}^*} \cup D$ is a deficient ball (and its radius is $\eta + t$). We now find bound the size of S_{p^*,key^*}.

Recall that $|S_{p^*,\mathsf{key}^*} \cup D| = |Q_{p^*,\mathsf{key}^*}| \leq 2^{n-\kappa} \leq |\mathcal{M}|/2$. Since this set contains less than half the points in the metric space we know its radius at most $n/2$. This means that $|S_{p^*,\mathsf{key}^*}|$ is a deficient sphere of radius at most $n/2 - t$. Let X denote a uniform string on $\{0,1\}^n$. We use Hoeffding's inequality [Hoe63]:

$$|S_{p^*,\mathsf{key}^*}| \leq \{x | \mathsf{dis}(x,0) \leq n-t\} = 2^n \Pr_{X \leftarrow \{0,1\}^n}[wt(X) \leq (1/2 - t/n)n] \leq 2^n e^{-n((t/n)^2)} = 2^n e^{-4\nu} \leq 2^{n-4\nu}$$

□

We have shown that $|\mathsf{Inter}(Q_{p^*,\mathsf{key}^*})| \leq 2^{n-4\nu}$. To complete the proof it suffices to show that for most values of the auxiliary information Z there are many parts Q_{p^*,key^*} that do not receive any points. Recall that Z consists of $2n/\nu$ coefficients and there are $(2^{n/\nu}-1)^{\nu-1}2^{n-\nu}$ equally likely values for Z. As described above, the view of $\mathsf{Gen}, \mathsf{Rep}$ is a uniform distribution V. We know show there are many possible values for $Z|P = p^*$. The only information about Z is contained in the point $V = v$. The length of this point is 2^n. Conditioned on this information there are still many possible values for Z. That is,

$$\begin{aligned}
\forall v, H_0(Z|V=v) &= \log\left(\frac{(2^{n/\nu}-1)^{\nu-1}2^{n-\nu}}{2^n}\right) \\
&= \log \frac{(2^{n/\nu}-1)^{\nu-1}}{2^\nu} \\
&> \log \frac{(2^{n/\nu})^{\nu-2}}{2^\nu} \qquad \text{(by Lemma E.6)} \\
&= \log \frac{2^{(n-2\nu))}}{2^\nu} = n - 3\nu.
\end{aligned}$$

Consider two possible z_1, z_2 that are possible values of Z. The distributions $V|Z = z_1$ and $V|Z = z_2$ intersect at one point (namely v).

This means that the Gen algorithm may include points for possible Z values into parts Q_{p^*,key^*} (other than v) and these values are disjoint. The optimum strategy is to include these values uniformly from different Z values. Consider the set of all preimages of R_{small} denoted $Q_{small} = \cup_{\mathsf{key} \in R_{small}} \mathsf{Inter}(Q_{\mathsf{key},p^*})$. Note that $Q_{small} \leq 2^{n-4\nu}|R_{small}|$. We now show that the intersection between Q_{key,p^*} is small for most possible values z. As before each container (the values of z) receives one item for free (the point v).

$$\begin{aligned}\mathop{\mathbb{E}}_{z} |Q_{small} \cap (V|P = p^* \wedge Z = z)| &\leq \left(\frac{\#\text{ items} + \#\text{ containers}}{\#\text{ containers}}\right) \\ &\leq \frac{2^{n-4\nu}|R_{small}|}{2^{n-3\nu}} + 1 \\ &= \frac{|R_{small}|}{2^{\nu}} + 1\end{aligned}$$

In expectation across Z,

$$\frac{\frac{|R_{small}|}{2^{\nu}} + 1}{|R_{small}|} \leq \frac{1}{2^{\nu}} + \frac{1}{|R_{small}|}$$

fraction of R_{small} receive any support. We now present a distinguisher D_{p^*} for a particular p^*:

1. On input x, z.
2. Compute $V|P = p^* \wedge Z = z$ and $Q_{p^*,x}$.
3. If $(Q_{p^*,x} \cap V|P = p^* \wedge Z = z) = \emptyset$ output $b = 0$.
4. Else output $b = 1$.

The distinguisher $D(x, p, z)$ is formed by calling $D_p(x, z)$ when $p \in A_p$ and outputting a random bit otherwise. The advantage of D is

$$\begin{aligned}&\Pr[D(\mathsf{Key}, P, Z) = 1] - \Pr[D(U, P, Z) = 1] \\ &= (\Pr[D(\mathsf{Key}, P, Z) = 1|P \in A_p] - \Pr[D(U, P, Z) = 1|P \in A_p]) \Pr[P \in A_p] \\ &\geq \sum_{p^* \in A_p} \Pr[P = p^*] (1 - \Pr[D_{p^*}(U, Z) = 1]) \\ &\geq \sum_{p^* \in A_p} \Pr[P = p^*] (1 - \Pr[D_{p^*}(U, Z) = 1|U \in R_{small}] \Pr[U \in R_{small}] - \Pr[U \notin R_{small}]) \\ &\geq \sum_{p^* \in A_p} \Pr[P = p^*] \left(1 - \left(\left(\frac{1}{|R_{small}|} + \frac{1}{2^{\nu}}\right) \Pr[U \in R_{small}]\right) - \Pr[U \notin R_{small}]\right) \\ &\geq \sum_{p^* \in A_p} \Pr[P = p^*] \left(1 - \frac{1}{2^{\nu}} - \frac{1}{2} \Pr[U \in R_{small}] - \Pr[U \notin R_{small}]\right) \\ &\geq \sum_{p^* \in A_p} \Pr[P = p^*] \left(1 - \frac{1}{2^{\nu}} - \frac{1}{2} \Pr[U \in R_{small}] - \Pr[U \notin R_{small}]\right) \\ &\geq \sum_{p^* \in A_p} \Pr[P = p^*] \left(1 - \frac{1}{2^{\nu}} - 1 + \frac{1}{2} \Pr[U \in R_{small}]\right) \\ &\geq \sum_{p^* \in A_p} \Pr[P = p^*] (1/4 - \mathtt{ngl}(n)) \geq \frac{1}{8} - \mathtt{ngl}(n).\end{aligned}$$

The sixth line follows since $R_{small} \geq 2^{\kappa-1} \geq 2$. The eighth line follows because $\Pr[U \in R_{small}] \geq 1/2$. The last inequality proceeds because $\Pr[P \in A_p] \geq 1/2$. This completes the proof of Lemma F.3. □