Continuous-Source Fuzzy Extractors: Source uncertainty and insecurity

Benjamin Fuller and Lowen Peng

University of Connecticut
Email: benjamin.fuller@uconn.edu, lowen.peng@uconn.edu

***Abstract*—Fuzzy extractors (Dodis et al., Eurocrypt 2004) convert repeated noisy readings of a high-entropy source into the same uniformly distributed key. The functionality of a fuzzy extractor outputs the key when provided with a value close to the original reading of the source. A necessary condition for security, called fuzzy min-entropy, is that the probability of every ball of values of the noisy source is small.**

Many noisy sources are best modeled using continuous metric spaces. To build *continuous-source fuzzy extractors*, prior work assumes that the system designer has a good model of the distribution (Verbitskiy et al., IEEE TIFS 2010). However, it is impossible to build an accurate model of a high entropy distribution just by sampling from the distribution.

Model inaccuracy may be a serious problem. We demonstrate a family of continuous distributions $\mathcal{W}$ that is impossible to secure. No fuzzy extractor designed for $\mathcal{W}$ extracts a meaningful key from an average element of $\mathcal{W}$. This impossibility result is despite the fact that each element $W \in \mathcal{W}$ has high fuzzy min-entropy. We show a qualitatively stronger negative result for secure sketches, which are used to construct most fuzzy extractors.

Our results are for the Euclidean metric and are information-theoretic in nature. To the best of our knowledge all continuous-source fuzzy extractors argue information-theoretic security.

Fuller, Reyzin, and Smith showed comparable negative results for a discrete metric space equipped with the Hamming metric (Asiacrypt 2016). Continuous Euclidean space necessitates new techniques.

I. Introduction

Many physical processes have entropy but exhibit noise between readings of the same process [1]–[9]. When a secret is read multiple times, readings are close (according to some metric dis) but not identical. Wyner [10] and Bennett, Brassard, and Robert [1] identified two fundamental tasks: 1) *Information-reconciliation:* removing noise without leaking information and 2) *Privacy amplification:* converting an entropic secret to uniformly random. We focus on non-interactive protocols that provide information-theoretic security.

In this setting, information reconciliation is performed by a secure sketch [11]. A secure sketch is a pair of algorithms $(\mathsf{SS}, \mathsf{Rec})$. Sketch or $ss \leftarrow \mathsf{SS}(w)$ converts an initial reading w to a nonsecret *helper* value ss. Let t be an error parameter. Then recover or $\mathsf{Rec}(w', ss)$ should output w if $\mathsf{dis}(w, w') \leq t$. The security requirement for a secure sketch is that w is hard to predict given ss.

A fuzzy extractor performs both tasks [11] and consists of two algorithms. The generate algorithm $((\mathsf{key}, \mathsf{pub}) \leftarrow \mathsf{Gen}(w))$ produces a key and nonsecret value pub. The reproduce algorithm $(\mathsf{key} \leftarrow \mathsf{Rep}(w', \mathsf{pub}))$ reproduces key if $\mathsf{dis}(w, w') \leq t$. The value key should be statistically close to uniform knowing pub. Most fuzzy extractors combine a secure sketch and a randomness extractor [12]. We consider sources W over a continuous metric space. We consider n-dimension space with Euclidean distance: $\mathsf{dis}(x, y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$.

The key question in designing a *continuous-source* fuzzy extractor [13] is which distributions W can be "secured." In asking whether a distribution W can be secured, one considers a specific distance tolerance t. Fuller, Reyzin, and Smith [14] introduced a precise notion to measure a noisy distribution's suitability for stable key derivation called *fuzzy min-entropy*. Fuzzy min-entropy codifies the adversary's success when provided with only the *functionality* of a fuzzy extractor (or secure sketch).

For a distribution W, with just *functionality* of Rep (or Rec), the adversary's best strategy is to find w' that maximizes the weight of possible $w \in W$ within distance t of w'. Denote by $B_t(w')$ the closed ball of radius t around w'. Fuzzy min-entropy is formally defined as

$$\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) \stackrel{\mathrm{def}}{=} -\log\left(\max_{w'} \Pr[W \in B_t(w')]\right).$$

Since fuzzy extractors are designed for entropic distributions, the designer only has a model of the underlying physical process. After deployment, the adversary may spend more time modeling, resulting in a more accurate model. As a result fuzzy extractors work for all distributions in a family $\mathcal{W}$. Ensuring security for a whole family is called the *distributional uncertainty* setting.

Fuller, Reyzin, and Smith presented a family of distributions $\mathcal{W}$ where each element $W \in \mathcal{W}$ has fuzzy min-entropy such that no fuzzy extractor $(\mathsf{Gen}_{\mathcal{W}}, \mathsf{Rep}_{\mathcal{W}})$ can simultaneously secure the family $\mathcal{W}$. That is, any fuzzy extractor $(\mathsf{Gen}_{\mathcal{W}}, \mathsf{Rep}_{\mathcal{W}})$ must be insecure for at least one element $W \in \mathcal{W}$ assuming the adversary knows the probability distribution W and the public helper value. Their result is for discrete Hamming space.

Our Contribution and Techniques We show a family of distributions $\mathcal{W}$ where no fuzzy extractor or secure sketch can secure $\mathcal{W}$ (Theorems IV.1 and V.1 respectively). That is, there is a family $\mathcal{W}$ such that for any cryptographic construction designed for $\mathcal{W}$ there exists an adversary that breaks security.

The secure sketch result is qualitatively stronger as it holds even if the secure sketch is allowed to be wrong a constant fraction of the time. The geometry of continuous Euclidean space is more challenging than discrete Hamming space and necessitates new techniques.

To give intuition we consider the case of a secure sketch $(\mathsf{SS}, \mathsf{Rec})$. Our result relies on the following asymmetry: SS sees only w sampled from W while the adversary knows which distribution W was used to sample $w \leftarrow W$. For error tolerance the public output ss must have some information about w. If the family $\mathcal{W}$ is carefully designed we can argue the adversary gains independent knowledge from ss and the distribution W. Together this independent knowledge can be used to break security. In both negative results there are two key components:

1) **Leakage:** Arguing that the public value ss restricts the set of possible w. For SS to be correct for w it must hold that for most nearby w', $\mathsf{Rec}(w', ss) = w$. Denote by C the set of all points w where Rec of nearby points is w. More formally,
$$C = \left\{ w \middle| \Pr_{w' | \mathsf{dis}(w,w') \leq t} [\mathsf{Rec}(w', ss) = w] \geq 1/2 \right\}.$$
The points in C form a Shannon error correcting code. This implies that $\forall x, y \in C$, $\mathsf{dis}(x, y) \geq t/2$. Since C forms a code, one can bound the size of C using packing arguments.
2) **Independence:** Knowing the distribution W provides independent and new information about the point w. To show that learning the distribution W gives fresh information, we consider distributions W that are the set of all points with the same output of a universal hash family [15]. That is, the description of W has two parts, the description of a hash function h and an output y. A distribution $W_{h,y}$ is the uniform distribution over the set $\{w | h(w) = y\}$. Since the hash is universal and h is not known to the SS algorithm, given w, the rest of the support of W is unknown. Thus, the information in $h(y)$ reduces the uncertainty on the point w.

The hash function we use is all points in the coset of a random p-ary lattice with minimum distance t. This hash function and the resulting family of distributions is described in detail in Section III.

For a fuzzy extractor, showing the **Leakage** property is more delicate because a continuous region can map to the same key. The adversary instead partitions the metric space based on the output key and consider only points in the interior of each part. Using volume arguments we can show that many distributions in $\mathcal{W}$ must not have any points in the interior of most parts.

Prior Positive Results for a single distribution Recent work [14], [16] shows that for any discrete distribution W with (super-logarithmic) fuzzy-min entropy there is a secure discrete fuzzy extractor $(\mathsf{Gen}_W, \mathsf{Rep}_W)$. These constructions need to know the probability distribution function of W exactly and are not instantiable in polynomial time. This is called the *precisely known distribution or distribution sensitive setting*. However, these techniques are inherently limited to discrete metric spaces.

Prior continuous-source fuzzy extractors applied quantization or partitioned the input space. As an example, Verbitskiy et al. [17] describe a continuous-source fuzzy extractor in the precisely known distribution model for distributions over $[0, 1]$. However, it is not clear how their technique extends to multiple dimensions. While Verbitskiy et al. [17] extend their construction to the distributional uncertainty setting, they only show security when the statistical distance between the observed distribution $\tilde{W}$ and the actual distribution W is small. The distributions described in this work can not be accurately estimated using a polynomial number of samples.

To the best of our knowledge it is not known how to build a continuous-source fuzzy extractor for each distribution with fuzzy min-entropy. Prior work either considers a small constant number of dimensions [17] or requires dimensions of the input to be uniform or independently distributed [18]. The major open question resulting from this work is whether continuous-source fuzzy extractors exist for each distribution with fuzzy min-entropy.

Correlated random variables A rich line of research views w and w' as samples from a correlated pair of random variables [10], [19]–[24]. Key length is bounded based on mutual information. These works consider the *precisely known distribution* setting.

Organization The remainder of this paper is organized as follows, Section II covers basic notation and mathematical prerequisites, Section III shows the family of distributions $\mathcal{W}$ that is used in both negative results, Section IV shows our fuzzy extractor negative result, and Section V shows our secure sketch negative result. We focus on the fuzzy extractor negative result as it is more challenging.

II. Preliminaries

We use uppercase letters for random variables and corresponding lowercase letters for their samples. Multiple occurrence of the same random variable in an expression signifies the same value of the random variable: for example $(W, \mathsf{SS}(W))$ is a pair of random variables obtained by sampling w according to W and applying the algorithm SS to w. The *statistical distance* between random variables A and B with the same domain is $\mathbf{SD}(A, B) = \frac{1}{2} \sum_a |\Pr[A = a] - \Pr[B = b]|$.

All logarithms in this work are base 2. Let (X, Y) be a pair of random variables. Define *min-entropy* of X as $\mathrm{H}_\infty(X) = -\log\left(\max_x \Pr[X = x]\right)$. The *average (conditional)* min-entropy [11, Section 2.4] of X given Y is
$$\tilde{\mathrm{H}}_\infty(X|Y) = -\log\left(\mathop{\mathbb{E}}_{y \in Y} \max_x \Pr[X = x | Y = y]\right).$$
Fuller, Smith, and Reyzin [14] proposed *fuzzy min-entropy* to measure suitability of a noisy distribution for key extraction. Fuzzy min-entropy captures the adversary's success probability when provided with the functionality of the primitive. We adopt this notion:

Definition II.1. *The t-fuzzy min-entropy of distribution W in a metric space $(\mathcal{M}, \mathsf{dis})$ is:*

$$\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) = -\log\left(\max_{w'} \int_{w\in\mathcal{M}|\mathsf{dis}(w,w')\le t} dw\right).$$

In the above, the measure assigns probability 1 to $\mathcal{M}$ and for any set X assigns probability $|X|/|\mathcal{M}|$.

Fuzzy extractors derive stable keys from noisy sources.

Definition II.2. *[11] An $(\mathcal{M}, \mathcal{W}, \kappa, t, \epsilon)$-*fuzzy extractor *is a pair $(\mathsf{Gen}, \mathsf{Rep})$. Gen on input $w \in \mathcal{M}$ outputs an extracted string $\mathsf{key} \in \{0,1\}^\kappa$ and a helper string $\mathsf{pub} \in \{0,1\}^*$. Rep takes $w' \in \mathcal{M}$ and $\mathsf{pub} \in \{0,1\}^*$ as inputs. $(\mathsf{Gen}, \mathsf{Rep})$ have the following properties:*

1) Correctness: *if* $\mathsf{dis}(w, w') \le t$ *and* $(\mathsf{key}, \mathsf{pub}) \leftarrow \mathsf{Gen}(w)$, $\Pr[\mathsf{Rep}(w', \mathsf{pub}) = \mathsf{key}] = 1$.
2) Security: $\forall W \in \mathcal{W}$, *if* $(\mathsf{Key}, \mathsf{Pub}) \leftarrow \mathsf{Gen}(W)$, $\mathbf{SD}((\mathsf{Key}, \mathsf{Pub}), (U_\kappa, \mathsf{Pub})) \le \epsilon$.

Recovering w from w' forms the core of many fuzzy extractor constructions. The primitive that performs just recovery is called a *secure sketch*. We recall the definition from [11, Section 3.1]:

Definition II.3. *An $(\mathcal{M}, \mathcal{W}, \tilde{m}, t)$-*secure sketch *with error δ is a pair $(\mathsf{SS}, \mathsf{Rec})$. 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 $\Pr[\mathsf{Rec}(w', \mathsf{SS}(w)) = w] \ge 1 - \delta$.*
2) Security*: for any $W \in \mathcal{W}$, $\tilde{H}(W|\mathsf{SS}(W)) \ge \tilde{m}$.*

III. The family of distributions $\mathcal{W}$

In this section we describe the family of distributions $\mathcal{W}$ used in our negative results for both fuzzy extractors and secure sketches. We use different properties of this family in the two negative results, however, all of the properties are achieved by the same family of functions. Our negative results are for an average element of $\mathcal{W}$. Thus, instead of thinking of the adversary as receiving the description of W we think of the adversary receiving Z where Z describes the uniform choice of W from $\mathcal{W}$ and we use W_Z to refer to an individual distribution. We define Z to be the restriction of the uniform distribution to points that have a particular output for a specified element of a hash family. Z consists of two components $z = (\mathsf{A}, \mathsf{h})$ that correspond to the description of the hash and its output respectively. We then define $W_z = \{w|\mathsf{Hash}_\mathsf{A}(w) = \mathsf{h}\}$.

The key to both results is showing that it is hard to recover A, h from a single w and that the hash has good geometric properties. The required properties are: 1) universality 2) regularity 3) the set W_z has minimum distance and 4) a large volume is required to cover every possible output of the hash family for every fixed A.

The hash function we use is the coset of the input point with respect to a random p-ary lattice with minimum distance $\ge t$. Let $\mathcal{K}$ be the set of lattices of all p-ary lattices $\Lambda_p(\mathsf{A})$ where $\mathsf{A} \in \mathbb{Z}_p^{n\times m}$ has minimum distance tp defined by $\Lambda_p(\mathsf{A}) = \{y \in \mathbb{Z}_p^n : y = \mathsf{A}s \mod p \text{ for some } s \in \mathbb{Z}^m\}$. Define $\mathsf{Hash}_{\mathsf{A}\in\mathcal{A}} : (\mathbb{R}/\mathbb{Z})^n \to (\mathbb{R}/p\mathbb{Z})^m/\Lambda_p(\mathsf{A})$ be defined by

$$x \mapsto [px]_{\Lambda_p(\mathsf{A})}$$

where we understand $[px]_{\Lambda_p(\mathsf{A})}$ to be a coset containing px with respect to the lattice. Scaling a random lattice to the unit cube is known as Construction A and is well studied in the lattice packing literature (Conway and Sloane [25]). In our presentation we expand the input point rather than compressing the lattice.

Note: The family is stated with respect to the input space $(\mathbb{R}/\mathbb{Z})^n$. This metric space will be used in Section IV. Section V uses the metric $[0,1]^n$. We only use the first three properties in Section V and these properties carry over to $[0,1]^n$.

Theorem III.1. *Let p be some prime and let $n, m \in \mathbb{Z}^+$ such that $m = \mu n$ for some $\mu \in (0, 1/2)$. For some matrix $\mathsf{A} \in \mathbb{Z}_p^{n\times m}$ define the lattice $\Lambda_p(\mathsf{A}) = \{Ax|x \in \mathbb{Z}_p^m\}$. Let $\mathcal{A}$ be the set of all lattices with minimum distance $t' = tp = \tau p\sqrt{n}$ where $\tau = (6p^\mu\sqrt{2e})^{-1}$. Define $\mathsf{Hash}_{\mathsf{A}\in\mathcal{A}}(w) = [pw]_{\Lambda_p(\mathsf{A})}$. If $p \ge (3\sqrt{2e})^{1/(1-\mu)}$ the following are simultaneously achieved:*

1) *is 2^{-a}-universal for $a = (n-m)\log p - 1$, that is $\forall v_1 \ne v_2 \in (\mathbb{R}/\mathbb{Z})^n$,*

$$\Pr_{\mathsf{A}\leftarrow\mathcal{A}}[\mathsf{Hash}_\mathsf{A}(v_1) = \mathsf{Hash}_\mathsf{A}(v_2)] \le 2^{-((n-m)\log p - 1)},$$

2) *is p^m regular, that is*

$$\forall \mathsf{A} \in \mathcal{A}, h \in \mathtt{Range}(\mathsf{Hash}_\mathsf{A}), |\mathsf{Hash}_\mathsf{A}^{-1}(h)| \ge p^m,$$

3) *preimage sets have minimum distance t for $t = \tau\sqrt{n}$, that is $\forall \mathsf{A} \in \mathcal{A}, v_1 \ne v_2$, if*

$$\mathsf{Hash}_\mathsf{A}(v_1) = \mathsf{Hash}_\mathsf{A}(v_2) \text{ then } \mathsf{dis}(v_1, v_2) \ge t,$$

4) *and has $p^{-\mu n}$-preimage volume, that is $\forall \mathsf{A} \in \mathcal{A}, V \subseteq (\mathbb{R}/\mathbb{Z})^n$,*

$$\Pr_{h\leftarrow\mathtt{Range}(\mathsf{Hash}_\mathsf{A})}\left[\mathsf{Hash}_\mathsf{A}^{-1}(h) \cap V \ne \emptyset\right] \le \frac{\mathsf{Vol}(V)}{p^{-\mu n}}.$$

The proof of this theorem is delayed to the full version [26].

IV. No fuzzy extractor can secure $\mathcal{W}$

We now prove it is impossible to build a fuzzy extractor that secures $\mathcal{W}$. As discussed in the introduction we use $(\mathbb{R}/\mathbb{Z})^n$ as the input space equipped with the Euclidean metric.

Theorem IV.1. *Let $\gamma \ge 1$ be a constant. Let $\mathcal{M} = (\mathbb{R}/\mathbb{Z})^n$. Let $\mu \in [0, 1/2)$ be a constant and define $m = \mu n$. Then there is a $\mathcal{W}$ of distributions (see Section III) with parameters*

1) *Let p be a prime such that $p \ge (3\sqrt{2e})^{1/(1-\mu)}$,*
2) *Noise rate $\tau = \frac{1}{6p^\mu\sqrt{2e}}$, and*
3) *For all $W \in \mathcal{W}$, $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty} = \mathrm{H}_\infty(W) \ge m$,*

such that, for any $(\mathsf{Gen}, \mathsf{Rep})$ *that is a* $(\mathcal{M}, \mathcal{W}, \kappa, t, \epsilon)$*-fuzzy-extractor with* noise rate $\tau \stackrel{d}{=} t/\sqrt{n}$ *if*

$$\kappa \geq 1 + \max\left\{0, \log\gamma + n\left(\log p^{\mu} - \log(\frac{6\sqrt{e}+\sqrt{\pi}}{6\sqrt{e}})\right)\right\},$$

then $\epsilon \geq \frac{1}{2} - \frac{e}{2\sqrt{2\pi\gamma}}$.

We show the impossibility for an average member of $\mathcal{W}$. Recall that we think of the distribution $W \in \mathcal{W}$ as being described by an auxiliary variable Z that is a pair (A, h) where $W_z = \{w | \mathsf{Hash}_{\mathsf{A}}(w) = \mathsf{h}\}$. The hash function we use is $\mathsf{Hash}_{\mathsf{A}\in\mathcal{A}} : (\mathbb{R}/\mathbb{Z})^n \to (\mathbb{R}/p\mathbb{Z})^m/\Lambda_p(\mathsf{A})$ be defined by

$$x \mapsto [px]_{\Lambda_p(\mathsf{A})}$$

where $[px]_{\Lambda_p(\mathsf{A})}$ the coset of the input point with respect to A. The conditions of Theorem IV.1 implies those of Theorem III.1 and thus we can use Theorem III.1. For this proof we need the regularity, minimum distance, and preimage volume conditions. By the 2^m-regularity and minimum distance properties of Hash, $\forall z \in Z, \mathrm{H}_{\infty}(W_z) = \mathrm{H}^{\mathsf{fuzz}}_{t,\infty}(W_z) = m$.

We now want to show that for a random $z \leftarrow Z$, if $(\mathsf{key}, \mathsf{pub})$ is the output of $\mathsf{Gen}(W_z)$, then key can be easily distinguished from uniform in the presence of pub and z. The outline for the proof is as follows:

- Without knowing z, the value w is uniform.
- The value pub partitions the key space.
- Each part is the partition created by pub is bounded in size.
- Valid w come from the interior of a part (by correctness of Rep, for every candidate input w to Gen all of its neighbors w' produce the same output of $\mathsf{Rep}(w', \mathsf{pub})$).
- The volume of the interior of a part is smaller than the volume of a part.
- Many parts have interior volume smaller than the preimage volume of the lattice (the volume of the Voronoi region of the lattice).
- Many elements $W \in \mathcal{W}$ have no point in the interior of the part.
- By averaging across parts, the average distribution W has no points in the interior of many parts.
- It is possible to distinguish a random key from one produced by Gen by checking if it comes from a part whose interior has no preimage in W.

The proof of theorem is delayed to the full version [26].

Parameter discussion: There are settings of $\mu, \tau, \kappa = \Theta(1)$ such that the statistical distance ϵ is a constant. Taking $\log p^{-\mu} \leq \log(\frac{6\sqrt{e}+\sqrt{\pi}}{6\sqrt{e}}) \approx -.1334$ implies that κ only needs to satisfy $\kappa \geq 1 + \log\gamma$. Substituting $p \geq (3\sqrt{2e})^{1/(1-\mu)}$ and ignoring factors due to finding a prime p this condition holds when $\mu \leq .045$. When $\gamma = 4$ then $\epsilon \geq .35$ (when $\kappa \geq 3$). The full setting of achievable parameter ranges for a constant κ, ϵ are in Figure 1.

V. No Secure Sketch can secure $\mathcal{W}$

We now show no secure sketch can be secure for an average member of $\mathcal{W}$. We consider the metric space $[0,1]^n$ but we can embed other bounded, continuous spaces of finite dimension into the unit cube.

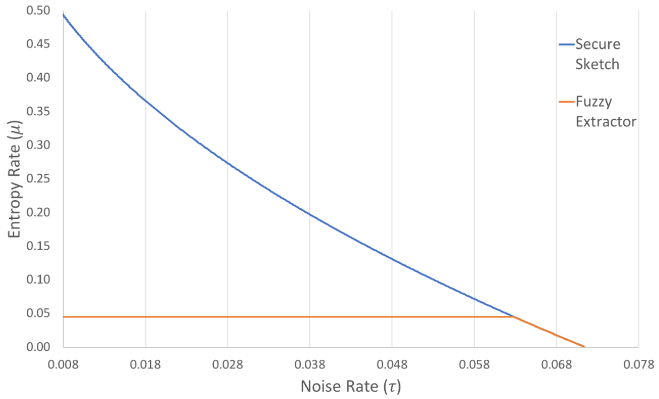

Fig. 1. Tradeoff between entropy rate μ and noise rate τ for both fuzzy extractors (red) and secure sketches (blue). Illustration of parameters in Theorem IV.1 and Theorem V.1. In this analysis we assume that there always exists a prime of size exactly $3(2e)^{1/(1-\mu)}$. The allowed noise rate τ may be reduced to find such a prime. Recall that Bertrand's postulate states that a prime exists between n and $2n$ for any integer $n > 1$.

Theorem V.1. *Let* $\mathcal{M} = [0,1]^n$ *with the Euclidean metric* dis *where* n *is even positive integer. Let* $\mu \in [0, \frac{1}{2})$ *be a constant and define* $m \stackrel{\mathrm{d}}{=} \mu n$. *There is a family* $\mathcal{W}$ *where for all* $W \in \mathcal{W}$, $\mathrm{H}^{\mathsf{fuzz}}_{t,\infty}(W) = \mathrm{H}_{\infty}(W) \geq m$, *such that for any* $(\mathcal{M}, \mathcal{W}, \tilde{m}, t)$*-secure sketch with error* δ, *we have* $\tilde{m} \leq 3$ *provided the following conditions hold:*

1) $n \geq 2(h_2(2\delta) + \log e)$. *Note that* $n \geq 6$ *suffices.*
2) *Let* p *be a prime integer parameter such that* $p \geq (3\sqrt{2e})^{\frac{1}{(1-\mu)}}$
3) *Define the* noise rate $\tau \stackrel{\mathrm{d}}{=} t/\sqrt{n}$ *where* $\tau = \frac{1}{6p^{\mu}\sqrt{2e}}$.
4) *The error parameter* δ *satisfies*

$$\delta \in \left[0, \frac{1}{2} - \frac{\log(1/\tau)}{2(1-\mu)\log p}\right)$$

Parameters As an example, $\mu = .3$ implies that $p \geq 129$, so if we consider $p = 131$, we have that $\tau \leq .017$ and $\delta \leq .08$. That is, there is a family with constant fuzzy entropy, constant error rate, and constant error where no good secure sketch exists. The trade-off between μ and τ is illustrated in Figure 1.

Interpreting the result: A secure sketch "discretizes" the input space into regions that produce a consistent value. Thus it is not surprising that a continuous secure sketch is not always possible. Note, the geometry of the Euclidean metric is more challenging than the Hamming metric due to the slower growth of volume. Proof of this theorem is delayed to the full version [26].

VI. Conclusion

Our two results show that model inaccuracy may be a major hurdle to constructing a continuous source fuzzy extractor. There are three ways to overcome our results:

1) Our results use distributions W that have fuzzy min-entropy at most $.5n$ and algorithms that correct $t \approx$

$.07\sqrt{n}$ errors (see Figure 1). One may be able to avoid these results when more fuzzy min-entropy is present or less error tolerance is required.

2) One could use properties of a distribution beyond fuzzy min-entropy. For example, Li et al. [18] assumed that dimensions were independently distributed.
3) Some discrete fuzzy extractors provide computational security [27]–[33]. One could provide computational security instead of information-theoretic security. We are not aware of any prior continuous-source fuzzy extractors that argue computational security.

For secure sketches we considered the metric space $[0,1]^n$. Our results can be extended to other bounded subsets of $\mathbb{R}^n$. Our fuzzy extractor result instead considers $(\mathbb{R}/\mathbb{Z})^n$. This is due to a technical limitation of the proof technique. We show that the volume of the interior of a part is smaller than the volume of the part. Roughly, maximum security drops by a factor proportional to the ratio between these volumes. To get a good bound on key length (reducing by a factor proportional to n) this ratio must be exponential in the dimension n. In the metric space $[0,1]^n$ most parts can be on a boundary of the unit cube. In the worst case these objects can be 1-dimension so their interior volume is only a constant factor smaller than their total volume.

We consider this to be an artifact of working with the unit cube. If a fuzzy extractor only secures points on the boundary then the data does not simultaneously vary in n dimensions. Since extraneous dimensions complicate error-correction, a system designer would first reduce dimensionality (see for example [34]) to find a representation that varies across all dimensions. This transformed distribution would be used for stable key derivation. In the "mod" space there are no boundary points, the entire region is "n-dimensional."

REFERENCES

[1] C. H. Bennett, G. Brassard, and J.-M. Robert, "Privacy amplification by public discussion," *SIAM Journal on Computing*, vol. 17, no. 2, pp. 210–229, 1988.

[2] S. Brostoff and M. Sasse, "Are passfaces more usable than passwords?: A field trial investigation," *People and Computers*, pp. 405–424, 2000.

[3] J. Daugman, "How iris recognition works," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 14, no. 1, pp. 21 – 30, January 2004.

[4] C. Ellison, C. Hall, R. Milbert, and B. Schneier, "Protecting secret keys with personal entropy," *Future Generation Computer Systems*, vol. 16, no. 4, pp. 311–318, 2000.

[5] B. Gassend, D. Clarke, M. Van Dijk, and S. Devadas, "Silicon physical random functions," in *Proceedings of the 9th ACM conference on Computer and communications security*. ACM, 2002, pp. 148–160.

[6] R. Mayrhofer and H. Gellersen, "Shake well before use: Intuitive and secure pairing of mobile devices," *IEEE Transactions on Mobile Computing*, vol. 8, no. 6, pp. 792–806, 2009.

[7] R. Pappu, B. Recht, J. Taylor, and N. Gershenfeld, "Physical one-way functions," *Science*, vol. 297, no. 5589, pp. 2026–2030, 2002.

[8] G. E. Suh and S. Devadas, "Physical unclonable functions for device authentication and secret key generation," in *Proceedings of the 44th annual Design Automation Conference*. ACM, 2007, pp. 9–14.

[9] P. Tuyls, G. J. Schrijen, B. Skoric, J. van Geloven, N. Verhaegh, and R. Wolters, "Read-proof hardware from protective coatings," in *Cryptographic Hardware and Embedded Systems - CHES 2006*, ser. Lecture Notes in Computer Science, L. Goubin and M. Matsui, Eds., vol. 4249. Springer, 2006, pp. 369–383.

[10] A. D. Wyner, "The wire-tap channel," *Bell System Technical Journal, The*, vol. 54, no. 8, pp. 1355–1387, 1975.

[11] Y. Dodis, R. Ostrovsky, L. Reyzin, and A. Smith, "Fuzzy extractors: How to generate strong keys from biometrics and other noisy data," *SIAM Journal on Computing*, vol. 38, no. 1, pp. 97–139, 2008.

[12] N. Nisan and D. Zuckerman, "Randomness is linear in space," *Journal of Computer and System Sciences*, pp. 43–52, 1993.

[13] I. Buhan, J. Doumen, P. Hartel, and R. Veldhuis, "Fuzzy extractors for continuous distributions," in *Proceedings of the 2nd ACM symposium on Information, computer and communications security*. ACM, 2007, pp. 353–355.

[14] B. Fuller, L. Reyzin, and A. Smith, "When are fuzzy extractors possible?" in *Advances in Cryptology–ASIACRYPT*. Springer, 2016, pp. 277–306.

[15] L. Carter and M. N. Wegman, "Universal classes of hash functions," *J. Comput. Syst. Sci.*, vol. 18, no. 2, pp. 143–154, 1979.

[16] J. Woodage, R. Chatterjee, Y. Dodis, A. Juels, and T. Ristenpart, "A new distribution-sensitive secure sketch and popularity-proportional hashing," in *Advances in Cryptology – CRYPTO*. Springer, 2017, pp. 682–710.

[17] E. A. Verbitskiy, P. Tuyls, C. Obi, B. Schoenmakers, and B. Skoric, "Key extraction from general nondiscrete signals," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 2, pp. 269–279, 2010.

[18] Q. Li, Y. Sutcu, and N. Memon, "Secure sketch for biometric templates," in *Advances in Cryptology – ASIACRYPT*. Springer, 2006, pp. 99–113.

[19] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 339–348, 1978.

[20] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography - I: secret sharing," *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1121–1132, 1993.

[21] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 733–742, 1993.

[22] R. Renner and S. Wolf, "Simple and tight bounds for information reconciliation and privacy amplification," in *Advances in Cryptology - ASIACRYPT*, ser. LNCS, B. K. Roy, Ed., vol. 3788. Springer, 2005, pp. 199–216.

[23] H. Tyagi and S. Watanabe, "Converses for secret key agreement and secure computing," *IEEE Transactions on Information Theory*, vol. 61, no. 9, pp. 4809–4827, 2015.

[24] M. Hayashi, H. Tyagi, and S. Watanabe, "Secret key agreement: General capacity and second-order asymptotics," in *2014 IEEE International Symposium on Information Theory*. IEEE, 2014, pp. 1136–1140.

[25] J. H. Conway and N. J. A. Sloane, *Sphere packings, lattices and groups*. Springer Science & Business Media, 2013, vol. 290.

[26] B. Fuller and L. Peng, "When are continuous-source fuzzy extractors possible?" Cryptology ePrint Archive, Report 2018/461, 2018, https://eprint.iacr.org/2018/461.

[27] B. Fuller, X. Meng, and L. Reyzin, "Computational fuzzy extractors," in *Advances in Cryptology-ASIACRYPT*. Springer, 2013, pp. 174–193.

[28] R. Canetti, B. Fuller, O. Paneth, L. Reyzin, and A. Smith, "Reusable fuzzy extractors for low-entropy distributions," in *Advances in Cryptology–EUROCRYPT*. Springer, 2016, pp. 117–146.

[29] C. Herder, L. Ren, M. van Dijk, M.-D. M. Yu, and S. Devadas, "Trapdoor computational fuzzy extractors and stateless cryptographically-secure physical unclonable functions," *IEEE Transactions on Dependable and Secure Computing*, vol. 14, no. 1, pp. 65–82, 2017.

[30] D. Apon, C. Cho, K. Eldefrawy, and J. Katz, "Efficient, reusable fuzzy extractors from LWE," in *International Conference on Cyber Security Cryptography and Machine Learning*. Springer, 2017, pp. 1–18.

[31] Q. Alamélou, P.-E. Berthier, C. Cachet, S. Cauchie, B. Fuller, P. Gaborit, and S. Simhadri, "Pseudoentropic isometries: A new framework for fuzzy extractor reusability," in *AsiaCCS*, 2018.

[32] Y. Wen, S. Liu, and S. Han, "Reusable fuzzy extractor from the decisional Diffie–Hellman assumption," *Designs, Codes and Cryptography*, pp. 1–18, 2018.

[33] Y. Wen and S. Liu, "Robustly reusable fuzzy extractor from standard assumptions," in *Advances in Cryptology – ASIACRYPT*, 2018.

[34] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "A survey of multilinear subspace learning for tensor data," *Pattern Recognition*, vol. 44, no. 7, pp. 1540–1551, 2011.