When are Fuzzy Extractors Possible?

Benjamin Fuller*, Leonid Reyzin†, and Adam Smith†
*University of Connecticut, Email: benjamin.fuller@uconn.edu.
ORC ID: 0000-0001-6450-0088
†Boston University, Email: reyzin@cs.bu.edu, ads22@bu.edu.
ORC IDs: 0000-0002-2052-8203, 0000-0001-9393-1127

***Abstract*—Fuzzy extractors (Dodis et al., SIAM J. Computing 2008) 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 quantify this property in a new notion called *fuzzy min-entropy*. We ask: is fuzzy min-entropy sufficient to build fuzzy extractors? We provide two answers for different settings.**

1) **If the construction is provided a description of the probability distribution W that defines the noisy source then fuzzy min-entropy is a sufficient condition for information-theoretic key extraction from W.**
2) **A more ambitious goal is to design a single extractor that works for all possible sources. This more ambitious goal is impossible: there is a family of sources with high fuzzy min-entropy for which no single fuzzy extractor is secure. This is true in three settings:**
   a) **for standard fuzzy extractors,**
   b) **for fuzzy extractors that are allowed to sometimes be wrong,**
   c) **and for secure sketches, which are the main ingredient of most fuzzy extractor constructions.**

***Keywords*-Fuzzy extractors, secure sketches, authentication, error-tolerance, key derivation, error-correcting codes, entropy**

I. Introduction

Sources of reproducible secret random bits enable cryptographic applications. Often, these bits are implicit, they are obtained by repeating the same process (such as reading a biometric or a physically unclonable function [2]) that generated them the first time. However, bits obtained this way are noisy [3], [4], [5], [6], [7], [8], [9], [2], [10], [11], [12]. 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 [13]. Bennett, Brassard, and Robert [3] identify two fundamental tasks:

1) Information reconciliation removes the noise with minimal leakage to an eavesdropping adversary.
2) 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 [14], 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 when 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 $\mathsf{Rep}(w', p)$. This attack is enabled by the functionality of the fuzzy extractor.

Fuzzy Min-Entropy We introduce a new entropy notion that precisely measures how hard it is for the adversary to execute this attack. It measures the probability of guessing a value within distance t of the original reading w. Suppose w is sampled from a distribution W. To have the maximum chance that w' is within distance t of w, the adversary would choose the point w' that maximizes the total probability mass of W within the ball $B_t(w')$ of radius t around w'. We therefore define *fuzzy min-entropy*

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

A fuzzy extractor's key cannot be longer than the fuzzy min-entropy (Proposition 3.2).

However, existing constructions do not measure their security in terms of fuzzy min-entropy; instead, their security is shown to be the min-entropy of W, denoted $\mathrm{H}_\infty(W)$, minus some loss for error-tolerance. This loss is due to an error-correction component that writes down enough information to determine which point within distance t was the original point. Since there are as many as $|B_t|$ points within radius t it takes $\log |B_t|$ bits to specify which one.[1] This value of $\log |B_t|$ is assumed to be the security loss so the residual security is $\mathrm{H}_\infty(W) - \log |B_t|$. It is easy to show that $\mathrm{H}_\infty(W) - \log |B_t| \leq \mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$, so it is natural to ask whether a loss of $\log |B_t|$ is

This work is based in part on a paper by the same title and authors presented at Asiacrypt 2016 [1].

[1] We omit w in the notation $|B_t|$ since we study metrics where the volume of the ball $B_t(w)$ does not depend on the center w.

necessary. This question is particularly relevant when the gap between the two sides of the inequality is high.[2]

As an example, in the biometric regime, entropy is estimated by comparing the distribution of distances between two different biometrics with a distribution with well understood statistical properties. For example, Daugman [5] shows that the distance between two irises is distributed similarly to a binomial distribution with 249 bits of entropy. The iris is then assumed to have the same entropy as the binomial distribution. The parameter t is determined by the experimental conditions. Daugman recommends setting $t > 205$. In the Hamming metric for strings of length 2048, $|B_{205}| \approx 2^{1024}$ (see Lemma 2.2). Thus, iris scans have

$$\mathrm{H}_\infty(W) - \log|B_t| \approx 249 - 1024 < 0,$$

(see discussion in [15, Section 5]). However, iris scans for different people appear to be well-spread in the metric space [16], the closest observed distance between two different irises is 548. This indicates the distribution of W has $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$.

The current state of fuzzy extractors is unsettling. For many distributions W with min-entropy, we have no known construction and no known impossibility result. We hope the more precise notion of fuzzy min-entropy can rectify this situation. Ideally, one could show a fuzzy extractor exists for every distribution with *enough* fuzzy min-entropy and that they are impossible with less fuzzy min-entropy. That is, we ask: *is fuzzy min-entropy sufficient for fuzzy extraction?* There is evidence that it may be sufficient when the security requirement is computational rather than information-theoretic—see Section I-B. We provide an answer for the case of information-theoretic security in two settings.

Contribution 1: Sufficiency of $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$ for a Precisely Known W We first consider the case when the fuzzy extractor designer has *precise knowledge* of the probability distribution function of W. In this setting, it is possible to construct a fuzzy extractor that extracts a key almost as long as $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$ (Theorem 4.3).[3] Our construction crucially utilizes the probability distribution function of W and, in particular, cannot necessarily be realized in polynomial time (this is similar, for example, to the interactive information-reconciliation feasibility result of [18]). This result shows that $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$ is a necessary and sufficient condition for building a fuzzy extractor for a given distribution W.

A number of previous works in the precise knowledge setting described tight bounds for specific distributions (for example, [19], [20], [21], [22], [23], [24]). These works are summarized in Table I. Our characterization unifies 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 $\mathrm{H}_\infty(W) - \log|B_t|$. Our construction can be viewed as a reference to evaluate the quality of efficient constructions in the precise knowledge setting by seeing how close they get to extracting all of $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$.

[2]For nearly uniform distributions, $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) \approx \mathrm{H}_\infty(W) - \log|B_t|$. In this setting, standard coding based constructions of fuzzy extractors (using almost optimal codes) yield keys of size approximately $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$.

[3]Woodage et al. present an improved version of this theorem [17]. We discuss their work below.

Many works consider i.i.d sources, for example [21]. These works are able to derive a key qualitatively longer than fuzzy min-entropy. This is because one can characterize i.i.d sources in terms of Shannon entropy, denoted $H_1(\cdot)$, instead of min-entropy. Each symbol of W is a separate "draw" from the distribution enabling average case analysis as the dimension n of the metric space increases. Key length for i.i.d. sources asymptotically approaches $H_1(W) - H_1(W|W')$ where W' is the distribution of noisy readings around W [21, Theorem 2]. It is difficult to directly compare with these works as they do not specify concrete losses for a fixed source length and they consider Shannon entropy. Qualitatively, Shannon entropy can be arbitrarily higher than min-entropy. Unfortunately, many biometrics and hardware sources are not i.i.d. (see for example [5]), so this analysis should be used judiciously.

Contribution 2: The Cost of Distributional Uncertainty Assuming precise knowledge of a distribution W is often unrealistic for high-entropy distributions; they can never be fully observed directly and must therefore be modeled. It is unrealistic to assume that the designer's model of a distribution is as accurate as the adversary's model. The adversary may have more resources including time to build a model when the construction is deployed. Existing fuzzy extractors are shown secure for a family of sources (for example, all sources of min-entropy at least m with at most t errors). The attacker may know more about the distribution than the designer. We call this the *distributional uncertainty* setting.

Our second contribution is a set of negative results for the distributional uncertainty setting. We provide two impossibility results for fuzzy extractors. Both demonstrate families $\mathcal{W}$ of distributions over $\{0,1\}^n$ such that each distribution in the family has $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}$ linear in n, but no fuzzy extractor can be secure for most distributions in $\mathcal{W}$. A fuzzy extractor designer who knows only that the distribution comes from $\mathcal{W}$ cannot secure the family, despite the fact that fuzzy extractors can be designed for each distribution in the family individually.

The first impossibility result (Theorem 5.1) assumes that Rep is perfectly correct and rules our fuzzy extractors for entropy rates, defined as $\mu = \mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)/n$, as high as $\mu \approx 0.18$. The second impossibility result (Theorem 6.1), relying on the work of Holenstein and Renner [25], also rules out fuzzy extractors in which Rep is allowed to make a mistake, but applies only to distributions with entropy rates up to $\mu \approx 0.07$.

We also provide a third impossibility result (Theorem 7.2), this time for an important building block called "secure sketch." A secure sketch is a one-round information-reconciliation component (that recovers the original w from the input w'). Secure sketches are used in most fuzzy extractor constructions. The result rules out secure sketches for a family of distributions with entropy rate up to $\mu = 0.5$, even if the secure sketches are allowed to make mistakes. We define secure sketches formally in Section VII. Most fuzzy extractor constructions are analyzed for all families with a certain amount of entropy. Thus, showing impossibility for higher entropy rates raises the lower bound on how much fuzzy min-entropy must be present in the physical distribution for

Work	Metric	Distribution	Residual Entropy	
This Work/[17]	Any discrete	Generic	$\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$	$-2\log 1/\epsilon - \log 1/\delta + 1$
[19]	Hamming	Uniform	$\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$	$-2\log 1/\epsilon - 2$
[20]/[21]	Any with almost perfect codes	i.i.d	$\approx H_1(W) - H_1(W\|W')$	$-o(n)$
[14]	Hamming/Set Difference/Edit	Generic	$\mathrm{H}_\infty(W) - \log\|B_t\|$	$-2\log 1/\epsilon - 2$

TABLE I

COMPARISON OF FUZZY EXTRACTOR CONSTRUCTIONS. HERE, ϵ IS THE STATISTICAL DISTANCE FROM THE UNIFORM DISTRIBUTION, δ IS THE ALLOWED ERROR, AND t IS THE DESIRED ERROR TOLERANCE. NOTE THAT $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) \geq \mathrm{H}_\infty(W) - \log|B_t|$. HERE n IS THE DIMENSION OF THE METRIC SPACE. H_1 REPRESENTS SHANNON ENTROPY AND W' IS THE DISTRIBUTION THAT ADDS NOISE TO THE ENROLLMENT VALUE W.

security to be based on just fuzzy min-entropy. As discussed in Section VIII, another alternative is to assume additional structure about the physical source.

A. Our Techniques

Techniques for Positive Results for a Precisely Known Distribution We now provide intuition for our positive result for a precisely known distribution W with fuzzy min-entropy. We begin with distributions in which all points in the support have the same probability (so-called "flat" distributions). Gen extracts a key from the input w using a randomness extractor [26]. Consider some subsequent reading w'. To achieve correctness, the string p must permit Rep to disambiguate which point $w \in W$ within distance t of w' was given to Gen. 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, Rep includes into the public value p a "sketch" of w computed via a universal hash of w. To determine the length of that sketch, consider the heaviest (according to W) ball B^* of radius t. Because the distribution is flat, B^* 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 points in B^*. Since $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$ is determined by the weight of B^*, the number of points cannot be too high and there will be entropy left after the sketch is published. This remaining entropy suffices to extract a key.

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* (Construction 4.2).

Techniques for Negative Results for Distributional Uncertainty We construct a family of distributions $\mathcal{W}$ and prove impossibility for a uniformly random $W \leftarrow \mathcal{W}$. 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, in our first impossibility result (Theorem 5.1), we construct $\mathcal{W}$ so that conditioning on the knowledge of W reduces the distribution to a small subspace (namely, all points on which a given universal hash function produces a given output), but conditioning on *only* w leaves the rest of the distribution uniform on a large fraction of the entire space. An adversary can exploit the knowledge of the hash value to reduce the uncertainty about key, as follows.

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 at distance at least t from the boundary of the region (else, the fuzzy extractor would have a nonzero probability of error). However, in the Hamming space, the vast majority of points lie near the boundary (this result follows by combining the isoperimetric inequality [27], which shows that the ball has the smallest boundary, with bounds on the volume of the interior of a ball, which show that this boundary is large). This allows the adversary to rule out so many possible w that, combined with the adversarial knowledge of the hash value, many regions become empty, leaving key far from uniform.

For the second impossibility result (Theorem 6.1, which rules out even fuzzy extractors that are allowed a possibility of error), we let the adversary know some fraction of the bits of w. Holenstein and Renner [25] showed that if the adversary knows each bit of w with sufficient probability, and bits of w' differ from bits of w with sufficient probability, then so-called information-theoretic key agreement is impossible. Converting the impossibility of information-theoretic key agreement to impossibility of fuzzy extractors takes a bit of technical work.

B. Related Settings

Other settings with close readings: $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}$ is sufficient The security definition of fuzzy extractors can be weakened to protect only against computationally bounded adversaries [28]. In this computational setting, under the assumption of semantically secure graded encoding, for most distance metrics a single fuzzy extractor can simultaneously secure all possible distributions [29], [30]. This construction is secure when the adversary can rarely learn key with oracle access to the program functionality. The set of distributions with fuzzy min-entropy are exactly those where an adversary learns key with oracle access to the functionality with negligible probability. Bitansky et al.'s [30] construction requires heavy weight and disputed cryptographic tools similar to those used to construct indistinguishability obfuscation [31], [32]. Their result implies that extending our negative result to the computational setting would have negative implications on the existence of certain types of obfuscation.

Furthermore, the functional definition of fuzzy extractors 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 [33]. When both interaction and computational assumptions are allowed, secure two-party computation can produce a good key 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 of collision-resistant hash functions and enhanced trapdoor permutations [34] or oblivious transfer and a variant of the random oracle model [35].

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 realizations of a correlated pair of random variables (W, W'). This model is considered in multiple works, including [13], [36], [37], [38]; recent characterizations of when key derivation is possible in this model include [39] and [40].

Much of the work on correlated pairs considers interactive protocols (as opposed to the noninteractive agreement needed for fuzzy extractors). However, the impossibility results for that setting are directly relevant to our work, because ruling out interactive protocols also rules out noninteractive ones. Recall that our starting point is the observation that the fuzzy extractor's output length (Proposition 3.2) is at most the fuzzy min-entropy of W. Prior work of Tyagi and Watanabe [40], and concurrent (with ours) work of those authors with Viswanath [41], developed general upper bounds on the achievable secret key length for correlated readings (via "conditional independence testing"). We can apply these in our setting by taking W' to be a random string within distance t of W to obtain impossiblity results analogous to Proposition 3.2. Our technique is less general, since it is tailored to the constraint on the distance between w and w'; however, this specificity allows us to give a simple, direct proof.

The conditional independence testing framework also applies to what we call the "unknown distribution" case (dubbed "correlated information at the eavesdropper" in the correlated readings literature). Given a joint distribution on W, W', Z, where Z is held by the eavesdropping adversary, the conditional testing framework gives an upper bound on the achievable key length (and hence an impossibility result when that bound is small). The framework does not show how one can construct distributions for which this bound is small. In particular, it is an open question whether one can derive versions of our impossibility results for unknown distributions (Theorems 5.1 and 6.1) using the framework. Even using the specific constructions of W, W', Z that arise in our proofs, it is open whether the conditional testing framework provides a good bound on the key length.

C. Concurrent and subsequent work

A construction that is very similar to our positive result in the known distribution setting (Construction 4.2) was used independently in a concurrent work of Hayashi, Tyagi, and Watanabe [42], who used the term "spectrum slicing" to describe it. They also extended this technique to the case of distributional uncertainty, using it in an interactive protocol, with one side telling the other to keep increasing the length of the sketch until Rep could succeed. This interactive approach was used in subsequent works, as well (e.g., [43], [41], [44]).

Li and Anantharam [44] consider correlated readings in the known distribution setting. They show that maximum expected key length for interactive protocols that are allowed to output variable-lengths keys is closely related to the mutual information between W and W'.

Woodage et al. [17] showed a clever extension to our leveled hashing construction (Construction 4.2). They observed the level information does not have to be explicitly included as part of the sketch. The construction uses leveled hashing but with two important changes:

1) The level is not written. Denote by h_ℓ the hash with the greatest number of output bits. If the used hash, h_i, has $|h_i| < h_\ell$ then the output is extended with random bits to length $|h_\ell|$. Instead of looking for an exact match the Rep algorithm finds the close point that matches the stored string at the longest prefix. This can be seen as considering all possible levels of the original hash.
2) In the our construction, the hash output is determined by how many points are in the neighborhood of points with that probability. This may lead to some levels with short hash outputs. This is a problem for Woodage et al.'s construction, if there are multiple short levels, the longest prefix may be the wrong level with noticeable probability. To address this problem, Woodage et al. use the (negative log of the) probability of a point to compute the length of the hash output. This ensures that all levels have length of at least the min-entropy of the distribution making collisions unlikely. Importantly, this change requires a change in the security argument, essentially arguing that all sketches are equally likely regardless of the starting level. This change requires an augmentation to the hash function called strong universality [45].

Fuller and Peng [46] extended our negative results to sources that are drawn from continuous metric spaces equipped with the Euclidean metric. There are two main differences between the Euclidean space and our setting:

1) Fewer points lie near the boundary of a ball in Euclidean space.
2) The use of continuous spaces requires volume techniques. So rather than showing that the hash value leaves few possibilities for w, they show that the "volume" of distributions is larger than the interior of parts. Thus, any choice for the interior of parts must not contain a fraction of distributions.

These changes necessitate the use of a different family that is derived from all cosets of random lattices with sufficient minimum distance (known as construction A).

II. Preliminaries

Random Variables We use uppercase letters for random variables and corresponding lowercase letters for their samples. A repeated occurrence of the same random variable 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

$$\begin{aligned}\mathbf{SD}(A,B) &= \frac{1}{2}\sum_a |\Pr[A=a] - \Pr[B=b]| \\ &= \max_S \Pr[A \in S] - \Pr[B \in S].\end{aligned}$$

Entropy Let $\log$ denote the base 2 logarithm. Let (X,Y) be a pair of random variables. Define *min-entropy* of X as $\mathrm{H}_\infty(X) = -\log(\max_x \Pr[X=x])$, and the *average (conditional)* min-entropy [14, Section 2.4] of X given Y as

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

Define Hartley entropy $H_0(X)$ to be the logarithm of the size of the support of X, that is $H_0(X) = \log|\{x|\Pr[X=x]>0\}|$. Define average-case Hartley entropy by averaging the support size:

$$\tilde{H}_0(X|Y) = \log(\mathop{\mathbb{E}}_{y\leftarrow Y}|\{y|\Pr[X=x|Y=y]>0\}|).$$

For $0<a<1$, the binary entropy is $h_2(p) = -p\log p - (1-p)\log(1-p)$, which corresponds to the Shannon entropy of any random variable that is 0 with probability p and 1 with probability $1-p$.

Randomness Extractors We use randomness extractors [26], as defined for the average case in [14, Section 2.5].

Definition 2.1: Let $\mathcal{M}$, χ be finite sets. A function $\mathtt{ext} : \mathcal{M}\times\{0,1\}^d \to \{0,1\}^\kappa$ 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

$$\mathbf{SD}((\mathtt{ext}(X,U_d),U_d,Y),U_\kappa\times U_d\times Y) \leq \epsilon.$$

Metric Spaces and Balls Let $\mathcal{M}$ be some finite space and let the function $\mathsf{dis} : \mathcal{M}\times\mathcal{M}\to\mathbb{R}^+\cup\{0\}$ be a distance metric (identity of indiscernibles, symmetric, and triangle inequality). 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, $B_t(w) = \{w'|\mathsf{dis}(w,w') \leq t\}$. We consider the Hamming metric over vectors in $\mathcal{Z}^n$ for some finite alphabet $\mathcal{Z}$, defined via $\mathsf{dis}(w,w') = |\{i|w_i \neq w'_i\}|$. In this space, the size of a ball in a metric space does not depend on w, so we denote by $|B_t|$ the size of a ball (centered arbitrarily) of radius t. U_κ denotes the uniformly distributed random variable on $\{0,1\}^\kappa$. We use the bounds on $|B_t|$ in $\{0,1\}^n$, see [47, Lemma 4.7.2, equation 4.7.5, p. 115] for proofs.

Lemma 2.2: Let $\tau = t/n$. The volume $|B_t|$ of the ball of radius in t in the Hamming space $\{0,1\}^n$ satisfies

$$\frac{1}{\sqrt{8n\tau(1-\tau)}}\cdot 2^{nh_2(\tau)} \leq |B_t| \leq 2^{nh_2(\tau)}.$$

We modify the definition of fuzzy extractors slightly from the work of Dodis et al. [14, Sections 3.2]. First, we allow for error as discussed in [14, Section 8]. Second, in the *distributional uncertainty* setting we consider a general 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 2.3: 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 $\mathsf{dis}(w,w') \leq t$ and $(\mathsf{key},p)\leftarrow\mathsf{Gen}(w)$, then $\Pr[\mathsf{Rep}(w',p)=\mathsf{key}] \geq 1-\delta$.
2) *Security:* for any distribution $W\in\mathcal{W}$, if $(\mathsf{Key},P)\leftarrow\mathsf{Gen}(W)$, then $\mathbf{SD}((\mathsf{Key},P),(U_\kappa,P)) \leq \epsilon$.

In the above definition, the value of w' must be chosen before p is known in order for the correctness guarantee to hold (alternatively, w' can be sampled from a probability distribution that is independent of p).

The Case of a Precisely Known Distribution If in the above definition we take $\mathcal{W}$ to be a one-element set containing a single distribution W, then the fuzzy extractor is said to be for a *precisely known distribution*. In this case, we need to require correctness only for w that have nonzero probability. We specify no requirement that the algorithms are compact or efficient, and so the distribution can be fully known to them.

III. New Notion: Fuzzy Min-Entropy

The fuzzy extractor helper string 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. We measure the quality of a source by (the negative logarithm of) the success of this attack.

Definition 3.1: The t-fuzzy min-entropy of a distribution W in a metric space $(\mathcal{M},\mathsf{dis})$ is:

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

Fuzzy min-entropy measures the functionality provided to the adversary by Rep (since p is public), and thus is a necessary condition for security. We formalize this statement in the following proposition.

Proposition 3.2: Let W be a distribution over $(\mathcal{M},\mathsf{dis})$ with $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) = m$. Let $(\mathsf{Gen},\mathsf{Rep})$ be a $(\mathcal{M},\{W\},\kappa,t,\epsilon)$-fuzzy extractor with error δ. Then

$$2^{-\kappa} \geq 2^{-m} - \delta - \epsilon.$$

If $\delta = \epsilon = 2^{-\kappa}$, then κ cannot exceed $m+2$.

Proof: Let W be a distribution where $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) = m$. This means that there exists a point $w'\in\mathcal{M}$ such that $\sum_{w\in\mathcal{M}|\mathsf{dis}(w,w')\leq t}\Pr[W=w] = 2^{-m}$. Consider the following function $D_{w'}$:

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

Clearly, $\Pr[D_{w'}(\mathsf{Key}, P) = 1] \geq 2^{-m} - \delta$, while $\Pr[D_{w'}(U_\kappa, P) = 1] = 1/2^\kappa$. Thus,

$$\begin{aligned}&\mathbf{SD}((\mathsf{Key}, P), (U_\kappa, P))\\ &\quad \geq \Pr[D_{w'}(\mathsf{Key}, P) = 1] - \Pr[D_{w'}(U_\kappa, P) = 1]\\ &\quad \geq 2^{-m} - \delta - 2^{-\kappa}.\end{aligned}$$

■

Proposition 3.2 extends to the settings of computational security and interactive protocols if the definition gives the adversary access to the true Key. We explore properties of fuzzy min-entropy below. These properties are included to demonstrate the utility of fuzzy min-entropy and are not necessary to complete the proofs in this work. Conditioning on an event p of probability $\Pr[P = p]$ decreases fuzzy min-entropy by a factor of at most $\log 1/\Pr[P = p]$.

Lemma 3.3: $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W|P = p) \geq \mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) + \log \Pr[P = p]$.

Proof:

$$\begin{aligned}&\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W|P = p)\\ &\quad = -\log\left(\max_{w'} \sum_{w|\mathsf{dis}(w,w')\leq t} \Pr[W = w|P = p]\right)\\ &\quad = -\log\left(\max_{w'} \sum_{w|\mathsf{dis}(w,w')\leq t} \frac{\Pr[W = w \wedge P = p]}{\Pr[P = p]}\right)\\ &\quad \geq -\log\left(\max_{w'} \sum_{w|\mathsf{dis}(w,w')\leq t} \frac{\Pr[W = w]}{\Pr[P = p]}\right)\\ &\quad = \mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) + \log \Pr[P = p].\end{aligned}$$

The second line follows from the first using Bayes rule, the third follows from the second using the monotonicity of probability. The last line follows by factoring $1/\Pr[P = p]$ from the sum, and noting the sum then represents $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$. ■

Conditional fuzzy min-entropy Properly defined, fuzzy min-entropy obeys a chain rule. We start by defining a conditional notion of fuzzy min-entropy for a random variable P.

Definition 3.4: For distributions W, P, the t-conditional fuzzy min-entropy of $W|P$ in a metric space $(\mathcal{M}, \mathsf{dis})$ is:

$$\begin{aligned}&\tilde{\mathrm{H}}^{\mathtt{fuzz}}_{t,\infty}(W|P)\\ &\quad = -\log\left(\mathop{\mathbb{E}}_{p\leftarrow P} \max_{w'} \sum_{w|\mathsf{dis}(w,w')\leq t} \Pr[W = w|P = p]\right).\end{aligned}$$

Then a chain rule analogous to average min-entropy [14, Lemma 2.2b] applies:

Lemma 3.5: $\tilde{\mathrm{H}}^{\mathtt{fuzz}}_{t,\infty}(W|P) \geq \mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) - H_0(P)$.

Proof:

$$\begin{aligned}&\tilde{\mathrm{H}}^{\mathtt{fuzz}}_{t,\infty}(W|P)\\ &= -\log\left(\mathop{\mathbb{E}}_{p\leftarrow P} \max_{w'} \sum_{w|\mathsf{dis}(w,w')\leq t} \Pr[W = w|P = p]\right)\\ &= -\log\left(\sum_p \max_{w'} \sum_{w|\mathsf{dis}(w,w')\leq t} \Pr[W = w|P = p]\Pr[P = p]\right)\\ &= -\log\left(\sum_p \max_{w'} \sum_{w|\mathsf{dis}(w,w')\leq t} \Pr[W = w \wedge P = p]\right)\\ &\geq -\log\left(\sum_p \max_{w'} \sum_{w|\mathsf{dis}(w,w')\leq t} \Pr[W = w]\right)\\ &\geq -\log\left(2^{H_0(P)}\left(\max_{w'} \sum_{w|\mathsf{dis}(w,w')\leq t} \Pr[W = w]\right)\right)\\ &\geq \mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) - H_0(P).\end{aligned}$$

Here the second line follows from the first using the definition of expectation. The third follows using Bayes rule. The fourth follows using monotonicity of expectation. By definition, there are at $2^{H_0(P)}$ possibilities for p, yielding the fifth line. The last line results by recognizing $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$ and converting to entropy. ■

IV. $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$ is Sufficient in the Precise Knowledge Setting

In this section, we build fuzzy extractors that extract almost all of $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$ for any distribution W. These constructions assume precise knowledge of W and are not efficient. We begin with flat distributions and then turn to arbitrary distributions.

Let $\mathrm{supp}(W) = \{w|\Pr[W = w] > 0\}$ denote the support of a distribution W. A distribution W is *flat* if all elements of $\mathrm{supp}(W)$ have the same probability. Our construction for this case is quite simple: to produce p, Gen outputs a hash of its input point w and an extractor seed; to produce key, Gen applies the extractor to w. Given w', Rep looks for $w \in \mathrm{supp}(W)$ that is near w' and has the correct hash value, and applies the extractor to this w to get key.

The specific hash function we use is *universal*. (We note that universal hashing has a long history of use for information reconciliation, for example [3], [18], and [48]. This construction is not novel; rather, we present it as a stepping stone for the case of general distributions.)

Definition 4.1 ([45]): Let $F : \mathcal{K} \times \mathcal{M} \to R$ be a function. We say that F is *universal* if for all distinct $x_1, x_2 \in \mathcal{M}$:

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

In our case, the hash output length needs to be sufficient to disambiguate elements of $\mathrm{supp}(W) \cap B_t(w')$ with high probability. Observe that there are at most $2^{\mathrm{H}_\infty(W) - \mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)}$ such elements when W is flat, so output length slightly greater (by $\log 1/\delta$) than $\mathrm{H}_\infty(W) - \mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$ will suffice. Thus, the

output key length will be $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) - \log 1/\delta - 2\log 1/\epsilon + 2$ (by using average-case leftover hash lemma, per [14, Lemma 2.2b, Lemma 2.4]). As this construction is only a warm-up, so we do not state it formally and proceed to general distributions.

A. Fuzzy Extractor for Arbitrary Distributions

The above hashing approach does not work for arbitrary sources. Consider a distribution W consisting of the following balls: B^1_t is a ball with $2^{\mathrm{H}_\infty(W)}$ points with total probability $\Pr[W \in B^1_t] = 2^{-\mathrm{H}_\infty(W)}$, $B^2_t, ..., B^{2^{-\mathrm{H}_\infty(W)}}_t$ are balls with one point each with probability $\Pr[W \in B^i_t] = 2^{-\mathrm{H}_\infty(W)}$. The above hashing algorithm writes down $\mathrm{H}_\infty(W)$ bits to achieve correctness on B^1_t. However, with probability $1 - 2^{-\mathrm{H}_\infty(W)}$ the initial reading is outside of B^1_t, and the hash completely reveals the point.

Instead, we use a layered approach: we separate the input distribution W into nearly-flat layers, write down the layer from which the input w came (i.e., the approximate probability of w) as part of p, and rely on the construction from the previous part for each layer. In other words, the hash function output is now variable-length, longer if probability of w is lower. Thus, p now reveals a bit more about w. To limit this information and the resulting security loss, we limit number of layers. As a result, we lose only $1 + \log H_0(W)$ more bits of security compared to the previous section.

The main idea is that providing the level information makes the distribution look nearly flat (the probability of points differs by at most a factor of two, which increases the entropy loss as compared to the flat case by only one bit). The level information itself increases the entropy loss by $\log H_0(W)$ bits, because there are only $H_0(W)$ levels that contain enough weight to matter. In subsequent work, Woodage et al. show that level information does not have to be leaked [17, Theorem 3]. We now present a formal description of our construction.

Construction 4.2: Let W be a distribution over a metric space $\mathcal{M}$ with $\mathrm{H}_\infty(W) = m$.

- Let $\delta \le \frac{1}{2}$ be the error parameter.
- Let $\ell = m + H_0(W) - 1$; round ℓ down so that $\ell - m$ is an integer (i.e., set $\ell = m + \lfloor(\ell - m)\rfloor$).
- For each $i = m, m+1, \ldots, \ell-1$, let $L_i = (2^{-(i+1)}, 2^{-i}]$ and let $F_i : \mathcal{K}_i \times \mathcal{M} \to R_i$ be a family of universal hash functions with $\log |R_i| = i + 1 - \mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) + \log 1/\delta$. Let $L_\ell = (0, 2^{-\ell}]$.
- Let $\mathtt{ext}$ be an $(\tilde{m}, \epsilon)$-average-case extractor for $\tilde{m} = \mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) - \log H_0(W) - \log 1/\delta - 1$ with output length κ.

Define $\mathsf{Gen}_W, \mathsf{Rep}_W$ as in Figure 1.

We instantiate this construction with the extractor parameters given by a universal hash (namely, $\kappa = \tilde{m} - 2\log 1/\epsilon + 2$):

Theorem 4.3: For any metric space $\mathcal{M}$, distribution W over $\mathcal{M}$, distance t, error $\delta > 0$, and security $\epsilon > 0$, there exists a $(\mathcal{M}, \{W\}, \kappa, t, \epsilon)$-known distribution fuzzy extractor with error δ for $\kappa = \mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) - \log H_0(W) - \log 1/\delta - 2\log 1/\epsilon + 1$.

Proof of Theorem 4.3: We first argue **correctness**. Fix some w, w' within distance t. When $\Pr[W = w] \in L_\ell$, then Rep is always correct, so let's consider only the case when $\Pr[W = w] \notin L_\ell$. The algorithm Rep will never output $\perp$ since at least the correct w will match the hash. Thus, an error happens when another element $w^* \in W^*$ has the same hash value $F(K_i, w^*)$ as $F(K_i, w)$. Observe that the total probability mass of W^* is greater than $|W^*| \cdot 2^{-(i+1)}$ but less than or equal to the maximum probability mass in a ball of radius t, $2^{-\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)}$. Therefore, $|W^*| \le 2^{i+1-\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)}$. Each element of W^* has the same hash as $F(K, w)$ with probability at most $1/|R_i|$, and thus correctness with error $|W^*|/|R| \le \delta$ follows by the union bound.

Security: We now argue security of the construction. Let $W_i = \{w | \Pr[W = w] \in L_i\}$. For ease of notation, let us make the special case of $i = \ell$ as part of the general case, as follows: define $\mathcal{K}_\ell = \{0\}$, $F_\ell(0, w) = w$, and $R_\ell = W_\ell$. Also, denote by SS the randomized function that maps w to ss. First, we set up the analysis by levels:

$$
\begin{aligned}
2^{-\tilde{\mathrm{H}}_\infty(W|\mathsf{SS}(W))} &= \mathop{\mathbb{E}}_{ss} \max_w \Pr[W = w \,|\, \mathsf{SS}(W) = ss] \\
&= \sum_{ss} \max_w \Pr[W = w \wedge \mathsf{SS}(W) = ss] \\
&= \sum_{i=m}^{\ell} \sum_{K \in \mathcal{K}_i} \sum_{y \in R_i} \max_w \Pr[W = w \wedge \mathsf{SS}(W) = (i, y, K)] \\
&\le \sum_{i=m}^{\ell} \sum_{K \in \mathcal{K}_i} \sum_{y \in R_i} \max_{w \in W_i} \Pr\left[\begin{array}{c} W = w \wedge F_i(K, w) = y \\ \wedge \;\; K \text{ output by } \mathsf{Gen} \end{array}\right].
\end{aligned}
$$

We now pay the penalty of $|R_i|$ for the presence of y (observe that removing the condition that $F_i(K, w) = y$ from the conjunction cannot reduce the probability):

$$
\begin{aligned}
&2^{-\tilde{\mathrm{H}}_\infty(W|\mathsf{SS}(W))} \\
&\le \sum_{i=m}^{\ell} \sum_{K \in \mathcal{K}_i} \sum_{y \in R_i} \max_{w \in W_i} \Pr[W = w \wedge K \text{ is chosen by } \mathsf{SS}] \\
&= \sum_{i=m}^{\ell} \sum_{K \in \mathcal{K}_i} |R_i| \cdot \max_{w \in W_i} \Pr\left[W = w \wedge K \text{ is chosen by } \mathsf{SS}\right].
\end{aligned}
$$

We now get rid of the key, because it is independent:

$$
\begin{aligned}
2^{-\tilde{\mathrm{H}}_\infty(W|\mathsf{SS}(W))} &\le \sum_{i=m}^{\ell} \sum_{K \in \mathcal{K}_i} |R_i| \cdot \max_{w \in W_i} \Pr[W = w] \cdot \frac{1}{|\mathcal{K}_i|} \\
&= \sum_{i=m}^{\ell} |R_i| \cdot \max_{w \in W_i} \Pr[W = w] \\
&\le |R_\ell| \cdot 2^{-\ell} + \sum_{i=m}^{\ell-1} |R_i| \cdot 2^{-i}.
\end{aligned}
$$

Finally, we add everything up, recalling that $|R_i|$ for $i < \ell$ is $2^{i+1-\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) + \log 1/\delta}$.

Gen_W
1) Input: w.
2) Find i such that $\Pr[W = w] \in L_i$.
3) If $i = \ell$ then set $ss = (i, w, 0)$.
4) Else sample $K \leftarrow \mathcal{K}_i$ and set $ss = (i, F_i(K, w), K)$
5) Sample a uniform extractor seed $\mathtt{seed}$
6) Output key $= \mathtt{ext}(w, seed)$, $p = (ss, \mathtt{seed})$.

Rep_W
1) Input: $(w', p = (ss, \mathtt{seed}))$
2) Parse ss as (i, y, K)
3) If $i = \ell$ then set $w^* = y$.
4) Else
   a) Let $W^* = \{w^* | \mathsf{dis}(w^*, w') \le t \wedge \Pr[W = w^*] \in L_i\}$.
   b) Find any $w^* \in W^*$ such that $F_i(K, w^*) = y$; if none exists, set $w^* = \perp$.
5) Output $\mathtt{ext}(w^*, \mathtt{seed})$.

Fig. 1. Fuzzy extractor construction for known distribution W.

$$
\begin{aligned}
2^{-\tilde{\mathrm{H}}_\infty(W|\mathsf{SS}(W))} &\le 2^{H_0(W)} \cdot 2^{-\ell} \\
&\quad + (\ell - m) \cdot 2^{1-\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)+\log 1/\delta} \\
&\qquad \text{(next line uses } \ell > m + H_0(W) - 2) \\
&< 2^{2-m} + (\ell - m) \cdot 2^{1-\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)+\log 1/\delta} \\
&\quad \text{(next line uses } m \ge \mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) \text{ and } \log 1/\delta \ge 1) \\
&\le (\ell - m + 1) \cdot 2^{1-\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)+\log 1/\delta} \\
&\qquad \text{(next line uses } \ell \le m + H_0(W) - 1) \\
&\le H_0(W) \cdot 2^{1-\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)+\log 1/\delta}.
\end{aligned}
$$

Taking the negative logarithm of both sides, we obtain $\tilde{m} \stackrel{\text{def}}{=} \tilde{\mathrm{H}}_\infty(W|\mathsf{SS}(W)) = \mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) - \log H_0(W) - \log 1/\delta - 1$. Applying the $(\tilde{m}, \epsilon)$ randomness extractor gives us the desired result. ∎

V. Impossibility of Fuzzy Extractors 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 building fuzzy extractors when the distribution W is precisely known. However, it is usually infeasible to characterize a high-entropy distribution W. Traditionally, algorithms deal with this *distributional uncertainty* by providing security for a family of distributions $\mathcal{W}$. In this section, we show that distributional uncertainty comes at a real cost.

We demonstrate an example over the binary Hamming metric in which every $W \in \mathcal{W}$ has linear $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W)$ (which is in fact equal to $\mathrm{H}_\infty(W)$), and yet there is some $W \in \mathcal{W}$ where even for 3-bit keys, the key distribution is far from uniform, $\epsilon = \frac{1}{4}$. In fact, we show that the adversary need not work hard: even a uniformly random choice of distribution W from $\mathcal{W}$ will thwart the security of any $(\mathsf{Gen}, \mathsf{Rep})$. The one caveat is that, for this result, we require Rep to be always correct (i.e., $\delta = 0$). As mentioned in the introduction, this perfect correctness requirement is removed in Sections VI and VII at a cost of lower entropy rate and stronger primitive, respectively.

The result is based on the following reasoning: Gen sees only a random sample w from a random $W \in \mathcal{W}$, but not W. The adversary knows the distribution W but not which particular value w was sampled. Because Gen does not know which W the input w came from, Gen must produce p that works for many distributions W that contain w in their support. Such p necessarily reveals a lot of information. The adversary can combine information gleaned from p with information about W to narrow down the possible choices for w and thus distinguish key from uniform.

Theorem 5.1: There exists a family of distributions $\mathcal{W}$ over $\{0,1\}^n$ equipped with the Hamming metric such that for each element $W \in \mathcal{W}$, $\mathrm{H}^{\mathtt{fuzz}}_{t,\infty}(W) = \mathrm{H}_\infty(W) \ge m$, and yet any $(\{0,1\}^n, \mathcal{W}, \kappa, t, \epsilon)$-fuzzy extractor with error $\delta = 0$ has $\epsilon > 1/4$.

This holds as long as $\kappa \ge 3$ and under the following conditions on the entropy rate $\mu = m/n$, noise rate $\tau = t/n$, and n:

- any $0 \le \tau < \frac{1}{2}$,
- any $\mu > 0$ such that $\mu < 1 - h_2(\tau)$ and $\mu < 1 - h_2\left(\frac{1}{2} - \tau\right)$, and
- any $n \ge \max\left(\frac{2}{1-h_2(\tau)-\mu}, \frac{5}{1-h_2\left(\frac{1}{2}-\tau\right)-\mu}\right)$.

The conditions on μ and τ imply the result applies to any entropy rate $\mu \le .18$ as long as τ is set appropriately and n is sufficiently large (for example, the result applies to $n \ge 1275$ and $\tau = .6\sqrt{\mu}$ when $0.08 \le \mu \le .18$). The τ vs. μ tradeoff is depicted in Figure 2.

Here we first provide short intuition, followed by the proof. The overall goal of the proof is show a lower bound on the value of ϵ which is the quality of the output key.

- One can partition the input metric space according to what value of key is output by $\mathsf{Rep}(w, p)$.
- The value of p reduces the set of possible w because, by correctness of Rep, every candidate input w to Gen must be such that all of its neighbors w' of distance at most t produce the same output of $\mathsf{Rep}(w', p)$.
- The isoperimetric inequality then shows for most parts, almost all points are not in the interior (Lemma 5.2).
- The above gives a bound on the residual entropy of w conditioned on p for most values of key. The second part of the proof incorporates the adversary's knowledge of the distribution $W \in \mathcal{W}$.
- We show the theorem holds for an average member of $\mathcal{W}$. Let Z denote a uniform choice of W from $\mathcal{W}$ and denote by W_z the choice specified by a particular value of z.

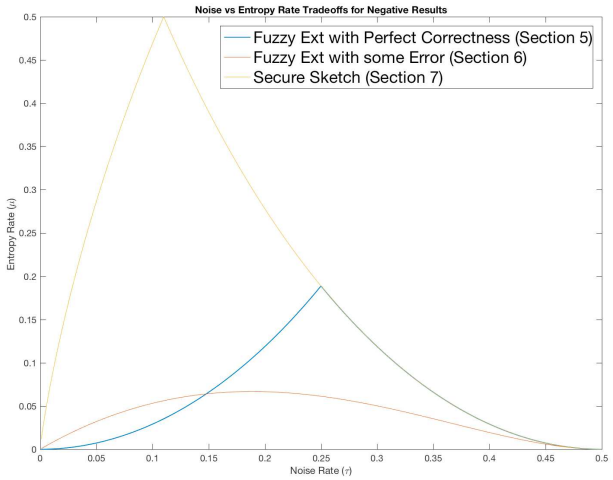

Fig. 2. The region of τ (x-axis) and μ (y-axis) pairs for which the three negative results apply (Theorems 5.1, 6.1 and 7.2). The Section 5 and Section 7 curves overlap starting at $\tau = .25$.

- Let $\{\mathsf{Hash}_\mathsf{k}\}_{\mathsf{k}\in\mathcal{K}}$ be a family of hash functions with domain $\mathcal{M} = \{0,1\}^n$ that is universal (small collision probability for any two points across the hash key), regular (large preimage size for any output value), and has preimage sets with high minimum distance. Then define $z = (\mathsf{k}, \mathsf{h})$ and define $W_{(\mathsf{k},\mathsf{h})}$ as the uniform distribution over the set $\{w|\mathsf{Hash}(w,\mathsf{k}) = \mathsf{h}\}_{(\mathsf{k},\mathsf{h})}$.
- The hash function we use is the output of a parity check matrix for a random code with high distance. Thus, each distribution W_z is a coset of some randomly chosen code C with good distance (that is not known by the construction). This family has the required properties (see Lemma 5.4).
- Since z is regular and preimage sets have minimum distance each W_z has high fuzzy min-entropy.
- The hash is universal, so learning the value of z reduces the set of possible values by another factor (Lemma 5.3).
- With this additional loss, for the average W_z, the interior of many parts contain no points from W_z. One can now build a distinguisher for a key derived from W_z from a random key. If a key comes from a part whose interior is empty the distinguisher outputs *random*, otherwise it outputs *real*.

We now proceed with the full proof.

Proof of Theorem 5.1: We show the impossibility for an average member of $\mathcal{W}$. We defer describing the family $\mathcal{W}$ until after a new bound on the preimage set size of most keys for a fuzzy extractor. The following lemma shows that the knowledge of p and key reduces the entropy of w.

Lemma 5.2: Let $\mathcal{M} = \{0,1\}^n$ equipped with the Hamming metric, $\kappa \geq 2$, $0 \leq t \leq n/2$, and $\epsilon \geq 0$. Suppose $(\mathsf{Gen}, \mathsf{Rep})$ is a $(\mathcal{M}, \mathcal{W}, \kappa, t, \epsilon)$-fuzzy extractor with error $\delta = 0$, for some distribution family $\mathcal{W}$ over $\mathcal{M}$. Let $\tau = t/n$. For any fixed p that is a possible output of Gen, there is a set $\mathsf{GdKey}_\mathsf{p} \subseteq \{0,1\}^\kappa$ of size at least $2^{\kappa-1}$ such that for every $\mathsf{key} \in \mathsf{GdKey}_\mathsf{p}$,

$$\log|\{v \in \mathcal{M}|(\mathsf{key}, p)\}| \leq n \cdot h_2\left(\frac{1}{2} - \tau\right) \leq n \cdot \left(1 - \frac{2}{\ln 2} \cdot \tau^2\right),$$

and, therefore, for any distribution $D_\mathcal{M}$ on $\mathcal{M}$,

$$H_0(D_\mathcal{M}|\mathsf{Gen}(D_\mathcal{M}) = (\mathsf{key}, p)) \leq n \cdot h_2\left(\frac{1}{2} - \tau\right) \leq n \cdot \left(1 - \frac{2}{\ln 2} \cdot \tau^2\right).$$

Proof: The set $\mathsf{GdKey}_\mathsf{p}$ consists of all keys for which $H_0(\mathcal{M}|\mathsf{Rep}(\mathcal{M}, p) = \mathsf{key}) \leq 2^{n-\kappa+1}$. The intuition is as follows. By perfect correctness of Rep, the input w to Gen has the following property: for all w' within distance t of w, $\mathsf{Rep}(w', p) = \mathsf{Rep}(w, p)$. Thus, if we partition $\mathcal{M}$ according to the output of Rep, the true w is t away from the interior of a part. Interior sets are small, which means the set of possible w values is small. (We note that by perfect correctness, Rep has a deterministic output even if the algorithm is randomized, so this partition is well-defined.)

To formalize this intuition, fix p and partition $\mathcal{M}$ according to the output of $\mathsf{Rep}(\cdot, p)$ as follows: let $Q_{p,\mathsf{key}} = \{w' \in \mathcal{M}|\mathsf{Rep}(w', p) = \mathsf{key}\}$. Note that there are 2^κ keys and thus 2^κ parts $Q_{p,\mathsf{key}}$. Let $\mathsf{GdKey}_\mathsf{p}$ by the set of keys for which these parts are not too large: $\mathsf{key} \in \mathsf{GdKey}_\mathsf{p} \Leftrightarrow |Q_{p,\mathsf{key}}| \leq 2 \cdot \mathcal{M}/2^\kappa = 2^{n-\kappa+1}$. Observe that $\mathsf{GdKey}_\mathsf{p}$ contains at least half the keys: $|\mathsf{GdKey}_\mathsf{p}| \geq 2^{\kappa-1}$ (if not, then $\cup_\mathsf{key}|Q_{p,\mathsf{key}}| > |\mathcal{M}|$). For the remainder of the proof we focus on elements in $\mathsf{GdKey}_\mathsf{p}$.

As explained above, if w is the input to Gen, then every point w' within distance t of w must be in the same part $Q_{p,\mathsf{key}}$ as w, by correctness of Rep. Thus, w must come from the interior of some $Q_{p,\mathsf{key}}$, where interior is defined as

$$\mathsf{Inter}(Q_{p,\mathsf{key}}) = \left\{ w \in Q_{p,\mathsf{key}} \;\middle|\; \begin{array}{c} \forall w' \text{ s.t. } \mathsf{dis}(w,w') \leq t, \\ w' \in Q_{p,\mathsf{key}} \end{array} \right\}.$$

We now use the isoperimetric inequality to bound the size of $\mathsf{Inter}(Q_{p,\mathsf{key}})$. Define a *near-ball*[4] centered at x to be any set S that is contained in a ball of some radius η and contains the ball of radius $\eta - 1$ around x. The inequality of [49, Theorem 1] (the original result is due to Harper [27]) says that for any sets $A, B \subset \{0,1\}^n$, there are near-balls X and Y centered at 0^n and 1^n, respectively, such that $|A| = |X|$, $|B| = |Y|$, and $\min_{a\in A, b\in B} \mathsf{dis}(a,b) \leq \min_{x\in X, y\in Y} \mathsf{dis}(x,y)$.

Letting A be the $\mathsf{Inter}(Q_{p,\mathsf{key}})$ and B be the complement of $Q_{p,\mathsf{key}}$ and applying this inequality, we get a near-ball $S_{p,\mathsf{key}}$ centered at 0^n and a near-ball D centered at 1^n, such that $|S_{p,\mathsf{key}}| = |\mathsf{Inter}(Q_{p,\mathsf{key}})|$, $|D| = 2^n - |Q_{p,\mathsf{key}}|$, and $\forall s \in S_{p,\mathsf{key}}, d \in D$, $\mathsf{dis}(s,d) > t$. Note that since $\mathsf{key} \in \mathsf{GdKey}_\mathsf{p}$ and $\kappa \geq 2$, we have $|Q_{p,\mathsf{key}}| \leq 2^{n-\kappa+1}$, and therefore $|D| \geq 2^{n-1}$.

Thus, D includes all the strings of Hamming weight $\lceil n/2 \rceil$ (because it is centered at 1^n and takes up at least half the

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

space), which means that the maximum Hamming weight of an element of $S_{p,\mathsf{key}}$ is $\lceil n/2 \rceil - t - 1 \leq n/2 - t$ (because each element of $S_{p,\mathsf{key}}$ is at distance more than t from D). We can now use binary entropy to bound the size of $S_{p,\mathsf{key}}$ by Lemma 2.2:

$$\begin{aligned}|\mathsf{Inter}(Q_{p,\mathsf{key}})| &= |S_{p,\mathsf{key}}| \\ &\leq |\{x | \mathsf{dis}(x,0) \leq n/2 - t\}| \\ &\leq 2^{n \cdot h_2\left(\frac{1}{2} - \frac{t}{n}\right)}.\end{aligned}$$

The theorem statement follows by taking the logarithm of both sides and by observing (using Taylor series expansion at $\tau = 0$ and noting that the third derivative is negative) that $h_2\left(\frac{1}{2} - \tau\right) \leq 1 - \frac{2}{\ln 2} \cdot \tau^2$. ∎

We now introduce the family $\mathcal{W}$. Let $\{\mathsf{Hash}_\mathsf{k}\}_{\mathsf{k} \in \mathcal{K}}$ be a family of hash function with domain $\mathcal{M}$ and the following properties:

- 2^{-a}-universality: for all $v_1 \neq v_2 \in \mathcal{M}$, $\Pr_{\mathsf{k} \leftarrow \mathcal{K}}[\mathsf{Hash}_\mathsf{k}(v_1) = \mathsf{Hash}_\mathsf{k}(v_2)] \leq 2^{-a}$, where $a = n \cdot h_2\left(\frac{1}{2} - \tau\right) + 3$.
- 2^m-regularity: for each $\mathsf{k} \in \mathcal{K}$ and h in the range of Hash_k, $|\mathsf{Hash}_\mathsf{k}^{-1}(\mathsf{h})| = 2^m$, where $m \geq \mu n$.
- preimage sets have minimum distance $t + 1$: for all $\mathsf{k} \in \mathcal{K}$, if $v_1 \neq v_2$ but $\mathsf{Hash}_\mathsf{k}(v_1) = \mathsf{Hash}_\mathsf{k}(v_2)$, then $\mathsf{dis}(v_1, v_2) > t$.

We demonstrate the existence of such a hash family in Lemma 5.4. Let Z be the random variable consisting of pairs (k, h), where k is uniform in $\mathcal{K}$ and h is uniform in the range of Hash_k. Let W_z for $z = (\mathsf{k}, \mathsf{h})$ be the uniform distribution on $\mathsf{Hash}_\mathsf{k}^{-1}(\mathsf{h})$. By the 2^m-regularity and minimum distance properties of Hash, $\mathrm{H}_\infty(W_z) = \mathrm{H}^{\mathbf{fuzz}}_{t,\infty}(W_z) = m$. Let $\mathcal{W} = \{W_z\}$.

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

First, view the sequence of events that we are trying to analyze as a game. The adversary chooses a uniform $\mathsf{k} \in \mathcal{K}$ and uniform h in the range of Hash_k. A uniform w from $\mathcal{M}$ s.t. $\mathsf{Hash}_\mathsf{k}(w) = \mathsf{h}$ then gets chosen, $(\mathsf{key}, p) = \mathsf{Gen}(w)$ gets computed, and the adversary receives p. The output of this game is $(\mathsf{k}, \mathsf{h}, w, p, \mathsf{key})$.

Consider now an alternative game. A uniform w gets chosen from $\mathcal{M}$ and uniform key k gets chosen from $\mathcal{K}$. $(\mathsf{key}, p) = \mathsf{Gen}(w)$ gets computed. The adversary receives $(\mathsf{k}, \mathsf{h} = \mathsf{Hash}_\mathsf{k}(w), p)$. The output of the game is $(\mathsf{k}, \mathsf{h}, w, p, \mathsf{key})$.

The distributions of the adversary's views and the outputs in the two games are identical: indeed, in both games, three random variable are uniform and independent (i.e., w is uniform in $\mathcal{M}$, k is uniform in $\mathcal{K}$, and the random coins of Gen are uniform in their domain), and the rest are determined fully by these three. However, the second game is easier to analyze, which is what we now do.

In this game, the value w is uniform on $\mathcal{M}$ (in the absence of knowledge about w). Knowledge of p reduced the set of possible w from 2^n to $2^{n \cdot h_2\left(\frac{1}{2} - \tau\right)}$, (Lemma 5.2). We know show that knowledge of z reduces the set of possible w by another factor of 2^a. Let K denote the uniform distribution on $\mathcal{K}$.

Lemma 5.3: Let L be a distribution. Let $\{\mathsf{Hash}_\mathsf{k}\}_{\mathsf{k} \in \mathcal{K}}$ be a family of 2^{-a}-universal hash functions on the support of L. Assume k is uniform in $\mathcal{K}$ and independent of L. Then

$$\begin{aligned}\tilde{H}_0(L | \mathsf{K}, \mathsf{Hash}_\mathsf{K}(L)) &< \log(1 + |\operatorname{supp}(L)| \cdot 2^{-a}) \\ &\leq \max(1, 1 + H_0(L) - a).\end{aligned}$$

Proof: Let U_L denote the uniform distribution on the support of L.

$$\begin{aligned}&2^{\tilde{H}_0(L | \mathsf{K}, \mathsf{Hash}_\mathsf{K}(L))} \\ &= \mathop{\mathbb{E}}_{\mathsf{k} \leftarrow \mathsf{K}, \mathsf{h}} |\{v \in L | \mathsf{Hash}_\mathsf{k}(v) = \mathsf{h}\}| \\ &= \mathop{\mathbb{E}}_{\mathsf{k} \leftarrow \mathsf{K}} \sum_\mathsf{h} \Pr[\mathsf{Hash}_\mathsf{k}(L) = \mathsf{h}] \cdot |\{v \in L | \mathsf{Hash}_\mathsf{k}(v) = \mathsf{h}\}| \\ &= \mathop{\mathbb{E}}_{\mathsf{k} \leftarrow \mathsf{K}} \sum_\mathsf{h} |L| \cdot \Pr[\mathsf{Hash}_\mathsf{k}(L) = \mathsf{h}] \cdot \Pr[\mathsf{Hash}_\mathsf{k}(U_L) = \mathsf{h}] \\ &= |\operatorname{supp}(L)| \cdot \mathop{\mathbb{E}}_{\mathsf{k} \leftarrow \mathsf{K}} \Pr_{v_1 \leftarrow L, v_2 \leftarrow U_L} \Pr[\mathsf{Hash}_\mathsf{k}(v_1) = \mathsf{Hash}_\mathsf{k}(v_2)] \\ &= |\operatorname{supp}(L)| \cdot \Pr_{v_1 \leftarrow L, v_2 \leftarrow U_L, \mathsf{k} \leftarrow \mathsf{K}} \Pr[\mathsf{Hash}_\mathsf{k}(v_1) = \mathsf{Hash}_\mathsf{k}(v_2)] \\ &\leq |\operatorname{supp}(L)| \cdot \begin{pmatrix} \Pr_{v_1 \leftarrow L, v_2 \leftarrow U_L}[v_1 = v_2] + \\ \Pr_{v_1 \leftarrow L, v_2 \leftarrow U_L}[v_1 \neq v_2] \cdot 2^{-a} \end{pmatrix} \\ &< 1 + |\operatorname{supp}(L)| \cdot 2^{-a}.\end{aligned}$$

This completes the proof of Lemma 5.3. ∎

Let M denote the uniform distribution on $\mathcal{M}$. By Lemma 5.2, for any p, $H_0(\mathsf{M} | \mathsf{Gen}(\mathsf{M}) = (\mathsf{key}, p)$ such that $\mathsf{key} \in \mathsf{GdKey}_\mathsf{p}) \leq n \cdot h_2\left(\frac{1}{2} - \frac{t}{n}\right) + \kappa$ (because there are most 2^κ keys in $\mathsf{GdKey}_\mathsf{p}$). Applying Lemma 5.3 (and recalling that $\kappa \geq 3$), we get that for any p,

$$\begin{aligned}&\tilde{H}_0(\mathsf{M} | \mathsf{Gen}(\mathsf{M}) \\ &= (\mathsf{key}, p) \text{ s.t. } \mathsf{key} \in \mathsf{GdKey}_\mathsf{p}, \mathsf{K}, \mathsf{Hash}_\mathsf{K}(\mathsf{M})) \\ &< \max\left(1, 1 + n \cdot h_2\left(\frac{1}{2} - \frac{t}{n}\right) + \kappa - a\right) \leq \kappa - 2.\end{aligned}$$

(Note carefully the somewhat confusing conditioning notation above, because we are conditioning on both events and variables. The event is $\mathsf{key} \in \mathsf{GdKey}_\mathsf{p}$ and the variables are k and $\mathsf{Hash}_\mathsf{k}(\mathsf{M})$.)

By correctness, for a fixed p, $\mathsf{Rep}(w, p)$ can produce only one key—the same one that was produces during $\mathsf{Gen}(w)$. Since applying a deterministic function (in this case, Rep) cannot increase H_0, we get that for each p,

$$\begin{aligned}\tilde{H}_0(\mathsf{key} | \mathsf{Gen}(\mathsf{M}) = (\mathsf{key}, p) \text{ s.t. } \mathsf{key} \in \mathsf{GdKey}_\mathsf{p}, \mathsf{K}, \mathsf{Hash}_\mathsf{K}(\mathsf{M})) \\ < \kappa - 2.\end{aligned}$$

Thus, on average over $z = (\mathsf{k}, \mathsf{h})$, over half the keys in $\mathsf{GdKey}_\mathsf{p}$ (i.e., over a quarter of all possible 2^κ keys) cannot be produced. Let $\mathsf{Implaus}$ be the set of triples $(\mathsf{key}, p, z = (\mathsf{k}, \mathsf{h}))$ such that $\Pr[\mathsf{Gen}(W_z) = (\mathsf{key}, p)] = 0$. Triples drawn by sampling w from W_z and computing $(p, \mathsf{key}) = \mathsf{Gen}(w)$ never come from this set. On other hand, random triples come $\mathsf{Implaus}$ at over quarter of the time. Thus, by definition of statistical distance, $\epsilon > \frac{1}{4}$. It remains to show that the hash family with the desired properties exists.

Lemma 5.4: For any $0 \leq \tau < \frac{1}{2}$, $\mu > 0$, α, and n such that $\mu \leq 1 - h_2(\tau) - \frac{2}{n}$ and $\mu \leq 1 - \alpha - \frac{2}{n}$, there exists a family of

hash functions $\{\mathsf{Hash}_\mathsf{k}\}_{\mathsf{k}\in\mathcal{K}}$ on $\{0,1\}^n$ that is 2^{-a}-universal for $a = \alpha n$, 2^m regular for $m \geq \mu n$, and whose preimage sets have minimum distance $t+1$ for $t = \tau n$.

Proof: Let $\mathcal{C}$ be the the set of all binary linear codes of rate μ (to be precise, dimension $m = \lceil \mu n \rceil$), length n, and minimum distance $t+1$:

$$\mathcal{C} = \left\{ C \;\middle|\; \begin{array}{l} C \text{ is a linear subspace of } \{0,1\}^n, \\ \dim(C) = m, \min_{c\in C-\{0^n\}} \mathsf{dis}(c,0^n) > t \end{array} \right\}.$$

For each $C \in \mathcal{C}$, fix H_C, an $(n-m)\times n$ parity check matrix for C, such that $C = \ker H_C$. For $v \in \{0,1\}^n$, let the syndrome $\mathsf{syn}_C(v) = H_C \cdot v$. Let $\{\mathsf{Hash}_\mathsf{k}\}_{\mathsf{k}\in\mathcal{K}} = \{\mathsf{syn}_C\}_{C\in\mathcal{C}}$.

2^m regularity follows from the fact that for each $h \in \{0,1\}^{n-\mu n}$, $\mathsf{Hash}_\mathsf{k}^{-1}(h)$ is a coset of C, which has size 2^m. The minimum distance property is also easy: if $v_1 \neq v_2$ but $\mathsf{syn}_C(v_1) = \mathsf{syn}_C(v_2)$, then $H_C(v_1 - v_2) = 0^n$, hence $v_1 - v_2 \in C - \{0^n\}$ and hence $\mathsf{dis}(v_1,v_2) = \mathsf{dis}(v_1-v_2,0) > t$.

We show 2^{-a}-universality by first considering a slightly larger hash family. Let $\mathcal{K}'$ be the set of *all* m-dimensional subspaces of $\{0,1\}^n$; for each $C' \in \mathcal{K}'$, choose a parity check matrix $H_{C'}$ such that $C' = \ker H_{C'}$, and let $\mathsf{syn}_{C'}(v) = H_{C'} \cdot v$. Let $\{\mathsf{Hash}'_{\mathsf{k}'}\}_{\mathsf{k}'\in\mathcal{K}'} = \{\mathsf{syn}_{C'}\}_{C'\in\mathcal{K}'}$. This family is 2^{m-n}-universal: for $v_1 \neq v_2$, $\Pr_{C'\in\mathcal{K}'}[H_{C'}\cdot v_1 = H_{C'}\cdot v_2] = \Pr_{C'\in\mathcal{K}'}[v_1 - v_2 \in \ker H_{C'} = C'] = \frac{2^m}{2^n}$, because C' is a random m-dimensional subspace. Note that this family is not much bigger than our family $\{\mathsf{Hash}_\mathsf{k}\}_{\mathsf{k}\in\mathcal{K}}$, because, as long as $\mu < 1 - h_2(\tau)$, almost every subspace of $\{0,1\}^n$ of dimension m has minimum distance $t+1$ for a sufficiently large n. Formally,

$$\begin{aligned}
&\Pr_{C'\in\mathcal{K}'}[C' \notin \mathcal{C}] \\
&= \Pr_{C'\in\mathcal{K}'}[\exists v_1 \neq v_2 \in C' \text{ s. t. } \mathsf{dis}(v_1,v_2) \leq t] \\
&= \Pr_{C'\in\mathcal{K}'}[\exists v_1 \neq v_2 \in C' \text{ s. t. } \mathsf{dis}(v_1-v_2,0^n) \leq t] \\
&= \Pr_{C'\in\mathcal{K}'}[\exists v \in C' - \{0^n\} \text{ s. t. } \mathsf{dis}(v,0^n) \leq t] \\
&\leq \sum_{v\in B_t(0^n)-\{0^n\}} \Pr_{C'\in\mathcal{K}'}[v \in C'] \leq 2^{nh_2(\tau)} \cdot \frac{2^m}{2^n} \leq \frac{1}{2}
\end{aligned}$$

(the penultimate inequality follows by Lemma 2.2 and the last one from $m \leq \mu n + 1$ and $\mu \leq 1 - h_2(\tau) - \frac{2}{n}$).

Since this larger family is universal and at most factor of two bigger than our family, our family is also universal:

$$\begin{aligned}
&\Pr_{C\in\mathcal{C}}[\mathsf{syn}_C(v_1) = \mathsf{syn}_C(v_2)] \\
&= \frac{|\{C \in \mathcal{C} | \mathsf{syn}_C(v_1) = \mathsf{syn}_C(v_2)\}|}{|\mathcal{C}|} \\
&\leq \frac{|\{C \in \mathcal{K}' | \mathsf{syn}_C(v_1) = \mathsf{syn}_C(v_2)\}|}{|\mathcal{K}'|} \cdot \frac{|\mathcal{K}'|}{|\mathcal{C}|} \leq 2^{m-n+1}
\end{aligned}$$

Thus, we obtain the desired result as long as $m - n + 1 \leq -a$, which is implied by the condition $\mu \leq 1 - \alpha - \frac{2}{n}$ and the fact that $m \leq \mu n + 1$. This completes the proof of Lemma 5.4. ■

Applying Lemma 5.4 with $\alpha = h_2\left(\frac{1}{2} - \tau\right) + \frac{3}{n}$, we see that the largest possible μ is

$$\max_\tau \min\left(1 - h_2(\tau), 1 - h_2\left(\frac{1}{2} - \tau\right)\right) \approx 0.1887.$$

Using the quadratic approximation to $h_2\left(\frac{1}{2} - \tau\right)$ (see Lemma 5.2), we can let μ be a free variable and set $\tau = .6\sqrt{\mu}$, in which case both constraints will be satisfied for all $0 < \mu \leq .18$ and sufficiently large n, as in the theorem statement. This concludes the proof of Theorem 5.1. ■

VI. Impossibility in the Case of Imperfect Correctness

The impossibility result in the previous section applies only to fuzzy extractors with perfect correctness. In this section, we build on the work of Holenstein and Renner [25] to show the impossibility of fuzzy extractors even when they are allowed to make mistakes a constant fraction δ (as much as 4%) of the time. However, the drawback of this result, as compared to the previous section, is that we can show impossibility only for a relatively low entropy rate of at most 7%. In Section VII, we rule out stronger primitives called secure sketches with nonzero error (which are used in most fuzzy extractor constructions), even for entropy rate as high as 50%.

Theorem 6.1: Let $\mathcal{M}$ denote the Hamming space $\{0,1\}^n$. There exists a family of distributions $\mathcal{W}$ over $\mathcal{M}$ such that for each element $W \in \mathcal{W}$, $\mathrm{H}^{\mathbf{fuzz}}_{t,\infty}(W) = \mathrm{H}_\infty(W) \geq m$, and yet any $(\mathcal{M},\mathcal{W},\kappa,t,\epsilon)$-fuzzy extractor with error $\delta \leq \frac{1}{25}$ has $\epsilon > \frac{1}{25}$.

This holds for any $\kappa > 0$ under the following conditions on the entropy rate $\mu = m/n$, noise rate $\tau = t/n$, and n:

- any $0 \leq \tau \leq \frac{1}{2}$ and μ such that
$$\mu < 4\tau(1-\tau)\left(1 - h_2\left(\frac{1}{4-4\tau}\right)\right)$$
- any sufficiently large n (as a function of τ and μ)

Note that the conditions on μ and τ imply that the result applies to any entropy rate $\mu \leq \frac{1}{15}$ as long as τ is set appropriately and n is sufficiently large. The τ vs. μ tradeoff is depicted in Figure 2.

The core structure of the proof is the same as Theorem 5.1. We construct a family $\mathcal{W}$ where knowing the element z (specifying $W_z \in \mathcal{W}$) reveals substantial information. However, this proof uses a different family and different techniques. The outline proceeds as follows:

1) The Rep algorithm (with p specified) can be used as a decoding algorithm for a binary symmetric channel (BSC) with error probability $(1-\alpha)/2$. To show this, we just need to argue that for a BSC with error probability $(1-\alpha)/2$, the probability of more than t errors is small.
2) The adversary (in the auxiliary knowledge Z) will receive $\{\perp, 0, 1\}$ for each position in W. The value $\perp$ indicates an erasure, and 0 or 1 indicates the true bit of W. So the family W_Z fixes some bits of W. These bits are known to the adversary but not the construction. Let $1-\beta$ denote the probability of receiving a $\perp$. The adversary's view corresponds to a classic erasure channel.
3) As long as β is not too large, one can show that W_Z has high fuzzy min-entropy as it corresponds to the uniform distribution over a restricted set of bits. We need to cap the number of bits received by the adversary for this to be true for all elements of W_Z. However, by standard

tail bounds, this removes very few distributions from the family.

4) We use a result of Holenstein and Renner [25, Theorem 4] that says the Shannon capacity of a β-erasure channel is greater than the capacity of a $(1-\alpha)/2$-binary symmetric channel.
5) From this theorem we can argue that the key has less Shannon entropy to the adversary than to Rep with a valid input.
6) The remainder of the proof is technical and converts this gap in Shannon entropy to a deficiency of the resulting key.

We now proceed with the full proof.

Proof: Similarly to the proof of Theorem 5.1, we will prove that any fuzzy extractor fails for an average element of $\mathcal{W}$: letting Z denote a choice of W from $\mathcal{W}$, we will show that $\mathbf{SD}((\mathsf{Key}, P, Z), (U_\kappa, P, Z)) > \frac{1}{25}$.

We start by describing the family of distributions. In this case, Z will not be uniform but rather binomial (with tails cut off). Essentially, Z will contain each bit of w with (appropriately chosen) probability β; given $Z = z$, the remaining bits of w will be uniform and independent.

For a string $z \in \{0,1,\perp\}^n$, denote by $info(z)$ the number of entries in z that are not $\perp$: $info(z) = |\{i \text{ s.t } z_i \neq \perp\}|$. Let W_z be the uniform distribution over all strings in $\{0,1\}^n$ that agree with z in positions that are not $\perp$ in z (i.e., all strings $w \in \{0,1\}^n$ such that for $1 \le i \le n$, either $z_i = \perp$ or $w_i = z_i$).

Let $0 \le \beta' \le 1$ be a parameter (we will set it at the end of the proof). Let Z' denote the distribution on strings in $\{0,1,\perp\}^n$ in which each symbol is, independently of other symbols, $\perp$ with probability $1-\beta'$, 0 with probability $\beta'/2$, and 1 with probability $\beta'/2$. Let $\beta = \beta' + \frac{1.4}{\sqrt{n}}$. Consider two distribution families: $\mathcal{W}' = \{W_z\}_{z \leftarrow Z'}$ and a smaller family $\mathcal{W} = \{W_z\}_{z \leftarrow Z}$, where $Z = Z' | info(Z') \le \beta n$ (the second family is smaller because, although on average $info(Z') = \beta' n$, there is a small chance that $info(Z')$ is higher than even βn).

We will use $\mathcal{W}$ to prove the theorem statement. First, we will show that every distribution $W_z \in \mathcal{W}$ has sufficient $\mathrm{H}^{\mathbf{fuzz}}_{t,\infty}$. Indeed, z constrains $info(z)$ coordinates out of n and leaves the rest uniform. Thus, $\mathrm{H}^{\mathbf{fuzz}}_{t,\infty}(W_z)$ is the same as $\mathrm{H}^{\mathbf{fuzz}}_{t,\infty}$ of the uniform distribution on the space $\{0,1\}^{n-info(z)}$. Let $a = n - info(z)$. By Lemma 2.2

$$\begin{aligned}\mathrm{H}^{\mathbf{fuzz}}_{t,\infty}(W_z) &\ge a\left(1 - h_2\left(\frac{t}{a}\right)\right)\\ &\ge n(1-\beta)\left(1 - h_2\left(\frac{t}{n(1-\beta)}\right)\right)\\ &= n(1-\beta)\left(1 - h_2\left(\frac{\tau}{1-\beta}\right)\right).\end{aligned}$$

and therefore

$$\mu = (1-\beta)\left(1 - h_2\left(\frac{\tau}{1-\beta}\right)\right). \qquad (1)$$

Note that smaller β gives a higher fuzzy entropy rate.

Second, we now want to show, similarly to the proof of Theorem 5.1, that $\mathbf{SD}((\mathsf{Key}, P, Z), (U_\kappa, P, Z)) > \frac{1}{25}$. We will do so by considering the family $\mathcal{W}$. Observe that by triangle inequality

$$\begin{aligned}&\mathbf{SD}((\mathsf{Key}, P, Z), (U_\kappa, P, Z))\\ &\quad\ge \mathbf{SD}((\mathsf{Key}, P, Z'), (U_\kappa, P, Z'))\\ &\quad- \mathbf{SD}((\mathsf{Key}, P, Z'), (\mathsf{Key}, P, Z))\\ &\quad- \mathbf{SD}((U_\kappa, P, Z), (U_\kappa, P, Z'))\\ &\quad\ge \mathbf{SD}((\mathsf{Key}, P, Z'), (U_\kappa, P, Z')) - 2\cdot\mathbf{SD}(Z', Z)\\ &\quad\ge \mathbf{SD}((\mathsf{Key}, P, Z'), (U_\kappa, P, Z')) - \frac{1}{25}.\end{aligned}$$

The last line follows by Hoeffding's inequality [50],

$$\begin{aligned}\mathbf{SD}(Z', Z) &= \Pr[info(Z') > \beta n]\\ &\le \exp\left(-2n\left(\frac{1.4}{\sqrt{n}}\right)^2\right) < \frac{1}{50}.\end{aligned}$$

Denote $\mathbf{SD}((\mathsf{Key}, P, Z'), (U_\kappa, P, Z'))$ by ϵ'. To bound ϵ', we recall a result of Holenstein and Renner [25, Theorem 4] (we will use the version presented in [51, Lemma 4.4]). For a random variable W with a values in $\{0,1\}^n$, let W^{noisy} denote a noisy copy of W: namely, the random variable obtained by passing W through a binary symmetric channel with error rate $\frac{1-\alpha}{2}$ (that is, $W^{noisy}_i = W_i$ with probability $\frac{1+\alpha}{2}$ and $W^{noisy}_i = 1 - W_i$ with probability $\frac{1-\alpha}{2}$, independently for each position i). Holenstein and Renner show that if $\alpha^2 \le \beta$, then Shannon entropy of Key conditioned on P and W^{noisy} is greater than Shannon entropy of Key conditioned on Z and W^{noisy}. Intuitively, this means that the Rep, when given P and W^{noisy}, knows less about Key than the adversary (who knows P and Z).

Recall the definitions of Shannon entropy $H_1(X) \stackrel{\text{def}}{=} \mathbb{E}_{x \leftarrow X} - \log \Pr[X = x]$ and conditional Shannon entropy $H_1(X|Y) \stackrel{\text{def}}{=} \mathbb{E}_{y \leftarrow Y} H_1(X|Y = y)$.

Theorem 6.2 ([25, Theorem 4]; [51, Lemma 4.4]): Suppose that (P, Key) is a pair of random variables derived from W. If $\alpha^2 \le \beta'$, then

$$H_1(\mathsf{Key}|P, Z') \le H_1(\mathsf{Key}|P, W^{noisy})$$

where H_1 denotes Shannon entropy, W^{noisy} is W passed through a binary symmetric channel with error rate $\frac{1-\alpha}{2}$, and Z' is W passed through a binary erasure channel with erasure rate $1-\beta'$.

(For a reader interested in how our statement of Lemma 6.2 follows from [51, Lemma 4.4], note that what we call $\mathsf{Key}, P, W^{noisy}$, and Z' are called U, V, Y, and Z, respectively, in [51]. Note also that we use only the part of the lemma that says that secret key rate $S_\rightarrow = 0$ when $\alpha^2 \le \beta$, and the definition [51, Definition 3.1] of the notion $S_\rightarrow$ in terms of Shannon entropy.)

We now need to translate this bound on Shannon entropy to the language of statistical distance ϵ of the key from uniform, reliability δ of the procedure Rep, and key length κ, as used in the definition of fuzzy extractors. First, we will do this translation for the case of noisy rather than worst-case input to Rep.

Lemma 6.3: Let (W, W^{noisy}, Z') be a triple of correlated random variables such that

- W and W^{noisy} are uniform over $\{0,1\}^n$,
- W^{noisy} is W passed through a binary symmetric channel with error rate $\frac{1-\alpha}{2}$ (that is, each bit position of W agrees with corresponding bit position of W^{noisy} with probability $\frac{1+\alpha}{2}$), and
- Z' is W passed through a binary erasure channel with erasure rate $1-\beta'$ (that is, each bit position of Z' agrees with the corresponding bit position of W with probability β' and is equal to $\perp$ otherwise).

Suppose $\mathsf{Gen}(W)$ produces (Key, P) with Key of length κ. Suppose $\Pr[\mathsf{Rep}(W^{noisy}, P) = \mathsf{Key}] = 1-\delta']$. Suppose further that $\mathbf{SD}((\mathsf{Key}, P, Z'), (U_\kappa, P, Z')) = \epsilon'$. If $\alpha^2 \le \beta'$, then

$$\kappa \le \frac{h_2(\epsilon') + h_2(\delta')}{1-\epsilon'-\delta'}.$$

In other words, if $\alpha^2 \le \beta'$, $\epsilon' \le \frac{1}{12}$, and $\delta' \le \frac{1}{12}$, then even a 1-bit Key is impossible to obtain.

(We note that a similar result follows from [51, Theorem 3.17] if we set the variables $S_\rightarrow$, γ, and m in that theorem to $0, \delta$, and κ, respectively. However, we could not verify the correctness of that theorem due to its informal treatment of what "ϵ-close to uniform" means; it seems that the small correction term $-h_2(\epsilon)$, just like in our result, is needed on the right-hand side to make that theorem correct.)

Proof of Lemma 6.3: Reliability allows us to bound the entropy of the key. By Fano's inequality [52, Section 6.2, p. 187],

$$H_1(\mathsf{Key}|P, W^{noisy}) \le \kappa\delta' + h_2(\delta').$$

Hence, by Theorem 6.2 (and the assumption that $\alpha^2 > \beta'$), we have

$$H_1(\mathsf{Key}|P, Z') \le \kappa\delta' + h_2(\delta'). \tag{2}$$

We now need the following lemma, which shows that near-uniformity implies high entropy.

Lemma 6.4: For a pair of random variables (A, B) such that the statistical distance between (A, B) and $U_\kappa \times B$ is ϵ, then $H_1(A|B) \ge (1-\epsilon)\kappa - h_2(\epsilon)$.

Proof: Let E denote a binary random variable correlated with (A, B) as follows: when $A = a$ and $B = b$, then $E = 0$ with probability

$$\max(\Pr[(A,B) = (a,b)] - \Pr[U_\kappa \times B = (a,b)], 0).$$

Similarly, let F denote a binary random variable correlated with $U_\kappa \times B$ as follows: when $U_\kappa = a$ and $B = b$, then $F = 0$ with probability

$$\max(\Pr[U_\kappa \times B = (a,b)] - \Pr[(A,B) = (a,b)], 0).$$

Note that $\Pr[E=0] = \Pr[F=0] = \epsilon$, by definition of statistical distance. Note also that $(A, B|E=1)$ is the same distribution as $(U_\kappa \times B|F=1)$. Since conditioning cannot increase Shannon entropy (by a simple argument — see, e.g., [47, Theorem 1.4.4]), we get

$$\begin{aligned} H_1(A|B) &\ge H_1(A|B,E) \\ &= \Pr[E=1]H_1(A|B,E=1) \\ &\quad + \Pr[E=0]H_1(A|B,E=0) \\ &\ge (1-\epsilon)H_1(A|B,E=1) \\ &= (1-\epsilon)H_1(U_\kappa|B,F=1). \end{aligned}$$

To bound this latter quantity, note that (the first line follows from the chain rule $H_1(X) \le H_1(X,Y) = H_1(X|Y) + H_1(Y)$ [47, Theorem 1.4.4])

$$\begin{aligned} \kappa &= H_1(U_\kappa|B) \\ &\le \quad H_1(U_\kappa|B,F) \qquad\quad + H_1(F) \\ &= (1-\epsilon)H_1(U_\kappa|B,F=1) + \epsilon \cdot H_1(U_\kappa|B,F=0) + h_2(\epsilon) \\ &\le (1-\epsilon)H_1(U_\kappa|B,F=1) + \epsilon\cdot\kappa + h_2(\epsilon) \end{aligned}$$

Rearranging terms, we get

$$H_1(U_\kappa|B,F=1) \ge \kappa - h_2(\epsilon)/(1-\epsilon),$$

and thus

$$H_1(A|B) \ge (1-\epsilon)\kappa - h_2(\epsilon).$$

This concludes the proof of Lemma 6.4. ∎

Combining (2) and Lemma 6.4 (applied to $A = \mathsf{Key}$, $B = (P, Z')$, and $\epsilon = \epsilon'$), we get the claimed bound. This concludes the proof of Lemma 6.3. ∎

Next, we translate this result from the noisy-input-case to the worst-case input case. Set $\alpha = \sqrt{\beta'}$. Suppose $t \ge n\left(\frac{1-\sqrt{\beta'}}{2} + \frac{1.4}{\sqrt{n}}\right)$. By Hoeffding's inequality [50],

$$\Pr[\mathsf{dis}(W, W^{noisy}) > t] \le \exp\left(-2n\left(\frac{1.4}{\sqrt{n}}\right)^2\right) < \frac{1}{50}.$$

Thus, a fuzzy extractor that corrects t errors with reliability δ implies that $\Pr[\mathsf{Rep}(W^{noisy}, P) = \mathsf{Key}] \ge 1-\delta']$ for $\delta' = \delta + \frac{1}{50}$. Since $\delta \le 1/25$, we have $\delta' < 1/12$ and Lemma 6.3 applies to gives us $\epsilon' > 1/12$ and $\epsilon > 1/12 - 1/25 > 1/25$ as long as $\kappa > 0$.

Finally, we work out the relationship between μ and τ and eliminate β, as follows. Recall that $\beta = \beta' + \frac{1.4}{\sqrt{n}}$; therefore $\sqrt{\beta} \le \sqrt{\beta'} + \frac{1.2}{n^{1/4}}$, and it suffices to take $\tau \ge \frac{1-\sqrt{\beta}}{2} + \frac{2}{\sqrt[4]{n}}$. Thus, we can set any $\tau > \frac{1-\sqrt{\beta}}{2}$ as long as n is sufficiently large. Solving for β (that is, taking any $\beta > (1-2\tau)^2$) and substituting into Equation 1, we can get any $\mu < 4\tau(1-\tau)\left(1 - h_2\left(\frac{1}{4-4\tau}\right)\right)$ for a sufficiently large n. ∎

VII. Stronger Impossibility Result for Secure Sketches

Most fuzzy extractor constructions share the following feature with our Construction 4.2: p includes information that is needed to recover w from w'; both Gen and Rep simply apply an extractor to w. The recovery of w from w', known as information-reconciliation, forms the core of many fuzzy extractor constructions. The primitive that performs this information reconciliation is called *secure sketch*. In this section we show stronger impossibility results for secure sketches. First,

we recall their definition from [14, Section 3.1] (modified slightly, in the same way as Definition 2.3).

Definition 7.1: 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 $\mathsf{dis}(w, w') \leq t$ then $\Pr[\mathsf{Rec}(w', \mathsf{SS}(w)) = w] \geq 1 - \delta$.
2) *Security*: for any distribution $W \in \mathcal{W}$, $\tilde{\mathrm{H}}_\infty(W|\mathsf{SS}(W)) \geq \tilde{m}$.

Secure sketches are more demanding than fuzzy extractors (secure sketches can be converted to fuzzy extractors by using a randomness extractors like in our Construction 4.2 [14, Lemma 4.1]). We prove a stronger impossibility result for them. Specifically, in the case of secure sketches, we can extend the results of Theorems 5.1 and 6.1 to cover imperfect correctness (that is, $\delta > 0$) and entropy rate μ up to $\frac{1}{2}$. We stress that most fuzzy extractor constructions rely on secure sketches.

Theorem 7.2: Let $\mathcal{M}$ denote the Hamming space $\{0,1\}^n$. There exists a family of distributions $\mathcal{W}$ over $\mathcal{M}$ such that for each element $W \in \mathcal{W}$, $\mathrm{H}^{\mathbf{fuzz}}_{t,\infty}(W) = \mathrm{H}_\infty(W) \geq m$, and yet any $(\mathcal{M}, \mathcal{W}, \tilde{m}, t)$-secure sketch with error δ has $\tilde{m} \leq 2$.

This holds under the following conditions on δ, the entropy rate $\mu = m/n$, noise rate $\tau = t/n$, and n:

- any $0 \leq \tau < \frac{1}{2}$ and $\mu > 0$ such that $\mu < h_2(\tau)$ and $\mu < 1 - h_2(\tau)$
- any $n \geq \max\left(\frac{.5 \log n + 4\delta n + 4}{h_2(\tau) - \mu}, \frac{2}{1 - h_2(\tau) - \mu}\right)$

Note that the result holds for any $\mu < 0.5$ as long as $\delta < (h_2(\tau) - \mu)/4$ and n is sufficiently large. The τ vs. μ tradeoff is depicted in Figure 2. Any fuzzy extractor that uses secure sketch (part of its output is $\mathsf{SS}(w)$) is subject to these bounds. In addition, any fuzzy extractor where the *true* input point w can be computed from key is subject to this bound as well (called an almost injective invertible fuzzy extractor by Yasanuga and Yuzawa [53]).

Before starting the formal proof we note that the overall strategy is the same as Theorem 5.1. The only substantive difference is that the functionality of secure sketches allow us to prove a stronger upper bound on the number of possible w's when $\mathsf{SS}(w)$ is known (Lemma 7.3). The core of this proof is arguing that the set of possible $\{v | \Pr_{v'|\mathsf{dis}(v,v') \leq t}[\mathsf{Rec}(v', \mathsf{SS}(v)) = v] \geq 1/2\}$ form a good error correcting code.

The family used is the same as in Theorem 5.1 with more parameter flexibility as more entropy is lost in Lemma 7.3.

Proof of Theorem 7.2: Similarly to the proof of Theorem 5.1, we will prove that any secure sketch algorithm fails for an average element of $\mathcal{W}$: letting Z denote a uniform choice of W from $\mathcal{W}$, we will show that $\tilde{\mathrm{H}}_\infty(W_Z|\mathsf{SS}(W_Z), Z) \leq 2$. The overall proof strategy is the same as for Theorem 5.1. We highlight only the changes here. Recall that $|B_t|$ denotes the volume of the ball of radius t in the space $\{0,1\}^n$. The parameters of the hash family are the same, except for universality: we require 2^{-a}-universality for $a = (n - \log|B_t| + h_2(2\delta))/(1 - 2\delta)$.

We postpone the question of the existence of such a hash family until the end of the proof. We can now state an analogue of Lemma 5.2. This result is an extension of lower bounds from [14, Appendix C], which handles only the case of perfect correctness. It shows that the value of the sketch reduces the entropy of a uniform point by approximately $\log|B_t|$.

Lemma 7.3: Let $\mathcal{M}$ denote the Hamming space $\{0,1\}^n$ and $|B_t|$ denote the volume of a Hamming ball of radius t in $\{0,1\}^n$. Suppose $(\mathsf{SS}, \mathsf{Rec})$ is a $(\mathcal{M}, \mathcal{W}, \tilde{m}, t)$ secure sketch with error δ, for some distribution family $\mathcal{W}$ over $\mathcal{M}$. Then for every $v \in \mathcal{M}$ there exists a set GdSS_v such that $\Pr[\mathsf{SS}(v) \in \mathsf{GdSS}_v] \geq 1/2$ and for any fixed ss,

$$\log|\{v \in \mathcal{M} | ss \in \mathsf{GdSS}_v\}| \leq \frac{n - \log|B_t| + h_2(2\delta)}{1 - 2\delta},$$

and, therefore, for any distribution $D_{\mathcal{M}}$ over $\mathcal{M}$,

$$H_0(D_{\mathcal{M}} | ss \in \mathsf{GdSS}_{D_{\mathcal{M}}}) \leq \frac{n - \log|B_t| + h_2(2\delta)}{1 - 2\delta}.$$

Proof: For any $v \in M$, define $\mathsf{Neigh}_t(v)$ be the uniform distribution on the ball of radius t around v and let

$$\mathsf{GdSS}_v = \{ss | \Pr_{v' \leftarrow \mathsf{Neigh}_t(v)}[\mathsf{Rec}(v', ss) \neq v] \leq 2\delta]\}.$$

We prove the lemma by showing two propositions.

Proposition 7.4: For all $v \in \mathcal{M}$, $\Pr[\mathsf{SS}(v) \in \mathsf{GdSS}_v] \geq 1/2$.

Proof: Let the indicator variable $1_{v',ss}$ be 1 if $\mathsf{Rec}(v', ss) = v$ and 0 otherwise. Let q_{ss} be the quality of the sketch on the ball $B_t(v)$:

$$q_{ss} = \Pr_{v' \xleftarrow{\$} \mathsf{Neigh}_t(v)}[\mathsf{Rec}(v', ss) = v] = \mathop{\mathbb{E}}_{v' \xleftarrow{\$} \mathsf{Neigh}_t(v)} 1_{v',ss}.$$

By the definition of correctness for $(\mathsf{SS}, \mathsf{Rec})$, for all $v' \in B_t(v)$,

$$\Pr_{ss \leftarrow \mathsf{SS}(v)}[\mathsf{Rec}(v', ss) = v] \geq 1 - \delta.$$

Hence, $\mathbb{E}_{ss \leftarrow \mathsf{Gen}(v)} 1_{v',ss} \geq 1 - \delta$. Therefore,

$$\mathop{\mathbb{E}}_{ss \leftarrow \mathsf{Gen}(v)} q_{ss} = \mathop{\mathbb{E}}_{ss} \mathop{\mathbb{E}}_{v'} 1_{v',ss} = \mathop{\mathbb{E}}_{v'} \mathop{\mathbb{E}}_{ss} 1_{v',ss} \geq \mathop{\mathbb{E}}_{v'}(1 - \delta) = 1 - \delta.$$

Therefore, applying Markov's inequality to $1 - q_{ss}$, we get $\Pr[q_{ss} \geq 1 - 2\delta] = \Pr[1 - q_{ss} \leq 2\delta] \leq 1/2$. ∎

To finish the proof of Lemma 7.3, we will show that the set $\{v \in \mathcal{M} | ss \in \mathsf{GdSS}_v\}$ forms a kind of error-correcting code, and then bound the size of the code.

Definition 7.5: We say that a set C is an (t, δ)-*Shannon code* if there exists a (possibly randomized) function Decode such that for all $c \in C$,

$$\Pr_{c' \leftarrow \mathsf{Neigh}_t(c)}[\mathsf{Decode}(c') \neq c] \leq \delta.$$

The set $\{v \in \mathcal{M} | ss \in \mathsf{GdSS}_v\}$ forms $(t, 2\delta)$ Shannon code if we set $\mathsf{Decode}(y) = \mathsf{Rec}(y, ss)$. We now bound the size of such a code.

Proposition 7.6: If $C \subseteq \{0,1\}^n$ is a (t, δ)-Shannon code, then

$$\log|C| \leq \frac{n - \log|B_t| + h_2(\delta)}{1 - \delta}.$$

Proof: Let the pair of random variables (X, Y) be obtained as follows: let X be a uniformly chosen element of C and Y be a uniformly chosen element of the ball of radius t around Y. By the existence of Decode and Fano's inequality [52, Section 6.2, p. 187], $H_1(X|Y) \le h_2(\delta)+\delta \log |C|$. At the same time, $H_1(X|Y) = H_1(X)-H_1(Y)+H_1(Y|X)$ (because $H_1(X,Y) = H_1(X)+H_1(Y|X) = H_1(Y)+H_1(X|Y)$), and therefore $H_1(X|Y) \ge \log |C|-n+\log |B_t|$ (because $H_1(Y) \le n$). Therefore, $\log |C| - n + \log |B_t| \le h_2(\delta) + \delta \log |C|$, and the lemma follows by rearranging terms. ∎

Lemma 7.3 follows from Proposition 7.6. ∎

We now show that entropy drops further when the adversary learns $\mathsf{Hash}_\mathsf{k}(w)$. Let M denote the uniform distribution on $\mathcal{M}$ and K denote the uniform distribution on $\mathcal{K}$. Applying Lemma 5.3 to Lemma 7.3, we get that for any ss,

$$\begin{aligned}\tilde{H}_0(\mathsf{M}|ss \in \mathsf{GdSS}_\mathsf{M}, \mathsf{K}, \mathsf{Hash}_\mathsf{K}(\mathsf{M})) \\ < \max\left(1, 1 + \frac{n - \log |B_t| + h_2(2\delta)}{1-2\delta} - a\right). \end{aligned} \quad (3)$$

To complete the proof, we will use this bound on $\tilde{H}_0$ as a bound on $\tilde{\mathrm{H}}_\infty$, as justified by the following lemma:

Lemma 7.7: For any random variables X and Y, $\tilde{\mathrm{H}}_\infty(X|Y) \le \tilde{H}_0(X|Y)$.

Proof: Starting with the definition of $\tilde{\mathrm{H}}_\infty$, recall that $-\log a = \log 1/a$, and apply Jensen's inequality to get

$$\begin{aligned}&\log \frac{1}{\mathbb{E}_{y \leftarrow Y} \max_x \Pr[X = x|Y = y])} \\ &\le \log \mathop{\mathbb{E}}_{y \leftarrow Y} \frac{1}{\max_x \Pr[X = x|Y = y])} \\ &\le \log \mathop{\mathbb{E}}_{y \leftarrow Y} |\{x| \Pr[X = x|Y = y] > 0\}|.\end{aligned}$$

∎

We need just one more lemma before we can complete the result, an analogue of [14, Lemma 2.2b] for conditioning on a single value $Z = z$ rather than with Z on average (we view conditioning on a single value as equivalent to conditioning on an event).

Lemma 7.8: For any pair of random variables (X, Y) and event η that is a (possibly randomized) function of (X, Y), $\tilde{\mathrm{H}}_\infty(X|\eta, Y) \ge \tilde{\mathrm{H}}_\infty(X|Y) - \log 1/\Pr[\eta]$.

Proof: The intuition is that to guess X given Y, the adversary can guess that η happened and fail if the guess is wrong. Formally,

$$\begin{aligned}&\tilde{\mathrm{H}}_\infty(X|Y) \\ &\quad= -\log \mathop{\mathbb{E}}_{y \leftarrow Y} \max_x \Pr[X = x|Y = y] \\ &\quad= -\log \mathop{\mathbb{E}}_{y \leftarrow Y} \max_x \frac{\Pr[X = x \wedge Y = y]}{\Pr[Y = y]} \\ &\quad\le -\log \mathop{\mathbb{E}}_{y \leftarrow Y} \max_x \frac{\Pr[X = x \wedge Y = y \wedge \eta]}{\Pr[Y = y]} \\ &\quad= -\log \mathop{\mathbb{E}}_{y \leftarrow Y} \max_x \frac{\Pr[X = x \wedge Y = y|\eta] \Pr[\eta]}{\Pr[Y = y]} \\ &\quad= \log \frac{1}{\Pr[\eta]} - \log \sum_{y \leftarrow Y} \max_x \Pr[X = x \wedge Y = y|\eta] \\ &\quad= \log \frac{1}{\Pr[\eta]} - \log \mathop{\mathbb{E}}_{y \leftarrow Y|\eta} \max_x \frac{\Pr[X = x \wedge Y = y|\eta]}{\Pr[Y = y|\eta]} \\ &\quad= \log \frac{1}{\Pr[\eta]} - \log \mathop{\mathbb{E}}_{y \leftarrow Y|\eta} \max_x \Pr[X = x|\eta \wedge Y = y] \\ &\quad= \log \frac{1}{\Pr[\eta]} + \tilde{\mathrm{H}}_\infty(X|\eta, Y).\end{aligned}$$

∎

Combining Lemmas 7.8 and 7.7 with Equation 3, we get

$$\begin{aligned}&\tilde{\mathrm{H}}_\infty(W_Z|Z, \mathsf{SS}(W_Z)) = \tilde{\mathrm{H}}_\infty(\mathsf{M}|\mathsf{SS}(\mathsf{M}), \mathsf{K}, \mathsf{Hash}_\mathsf{K}(\mathsf{M})) \\ &\le \log \frac{1}{\Pr[\mathsf{SS}(\mathsf{M}) \in \mathsf{GdSS}_\mathsf{M}]} + \\ &\quad \tilde{\mathrm{H}}_\infty(\mathsf{M}|ss \text{ s.t. } ss = \mathsf{SS}(\mathsf{M}) \wedge ss \in \mathsf{GdSS}_\mathsf{M}, \mathsf{K}, \mathsf{Hash}_\mathsf{K}(\mathsf{M})) \\ &\le \log \frac{1}{\Pr[\mathsf{SS}(\mathsf{M}) \in \mathsf{GdSS}_\mathsf{M}]} + \\ &\quad \tilde{H}_0(\mathsf{M}|ss \text{ s.t. } ss = \mathsf{SS}(\mathsf{M}) \wedge ss \in \mathsf{GdSS}_\mathsf{M}, \mathsf{K}, \mathsf{Hash}_\mathsf{K}(\mathsf{M})) \\ &< \log \frac{1}{\Pr[\mathsf{SS}(\mathsf{M}) \in \mathsf{GdSS}_\mathsf{M}]} \\ &\quad + \max\left(1, 1 + \frac{n - \log |B_t| + h_2(2\delta)}{1-2\delta} - a\right).\end{aligned}$$

We can have shown that $\tilde{\mathrm{H}}_\infty(W_Z|Z, \mathsf{SS}(W_Z)) \le 2$, because the first term of the above sum is at most 1 by Proposition 7.4 and the second term is 1 by our choice of a as $a = (n - \log |B_t| + h_2(2\delta))/(1-2\delta)$.

It remains to show that the desired hash family exists. Note in that (because $\delta < .25$) setting any $\alpha \ge 1 - h_2(\tau) + \frac{.5 \log n + 4\delta n + 2}{n}$ and choosing an αn-universal hash function will be sufficient, because, by Lemma 2.2, $\log |B_t| \ge n h_2(\tau) - \frac{1}{2} \log n - 1$, and so

$$\begin{aligned}a &= \frac{n - \log |B_t| + h_2(2\delta)}{1-2\delta} \\ &\le n \cdot \frac{1 - h_2(\tau) + (.5 \log n + 1 + h_2(2\delta))/n}{1-2\delta} \\ &< n \cdot (1 - h_2(\tau) + \frac{.5 \log n + 1 + h_2(2\delta)}{n} + 4\delta) \\ &\le n \cdot \left(1 - h_2(\tau) + \frac{.5 \log n + 4\delta n + 2}{n}\right) \\ &\le n \cdot \alpha\end{aligned}$$

(the second inequality is true because for any $x < 1$ and $0 < y < .5$, $x/(1-y) < x+2y$, because $x < (x+2y)(1-y)$, because $0 < y(2-x-2y)$; the third inequality follows from

$h_2(2\delta) < 1$). Such a hash family exists by Lemma 5.4 as long as $\mu \leq 1 - \alpha - 2/n \leq h_2(\tau) - (.5 \log n + 4\delta n + 4)/n$ and $\mu \leq 1 - h_2(\tau) - 2/n$). ∎

VIII. Conclusion

This work introduces fuzzy min-entropy as a new metric for measuring the suitability of deriving keys from a noisy probability distribution. This condition is sufficient for security if the distribution is exactly known. This setting is comparable to the traditional setting when Eve receives no auxiliary information.

Our negative results show that providing security simultaneously for a family of sources is impossible for all distributions with fuzzy min-entropy. The core of all of these proofs is constructing a family of distributions $\mathcal{W}$ where the description of the element $W \in \mathcal{W}$ provides the adversary with information independent of what is (necessarily) leaked by the fuzzy extractor (or secure sketch). Our three results require a careful tuning between the information leaked by the fuzzy extractor and the independent information in the description. This state of affairs seems somewhat bleak, however, there are several ways to avoid these negative results:

1) Focus on providing security for high entropy distributions only. However, many noisy distributions come from nature and system designers cannot effectively adjust their parameters,
2) Assume some additional structure from the distribution such as independence between dimensions [21] or that random subsets of dimensions have high entropy [54].
3) Restrict the adversary, for example, assuming the adversary runs in polynomial time. Recently, constructions have shown fuzzy extractors secure against bounded adversaries relying on hardness of discrete log [55], [56], decoding random codes [28], [57], security of hash functions [54], [58], and general cryptographic primitives [59]. Fuller, Meng, and Reyzin showed that computationally-secure sketches imply information-theoretic ones, so computationally-secure sketches are subject to the negative results in this work [28, Theorem 1]. A comparable theorem is not known for computationally-secure fuzzy extractors.

Acknowledgements

The authors are grateful to Gene Itkis and Yevgeniy Dodis for helpful discussions and to Thomas Holenstein for clarifying the results of [51], [25]. The work of Benjamin Fuller was sponsored in part by US NSF grants 1012910, 1012798, and 1849904 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 and The Institute of Science and Technology, Austria, where part of this work was performed. Adam Smith's work was supported in part by NSF awards 0747294, 0941553 and 1447700 and was performed partly while at Boston University's Hariri Institute for Computing and RISCS Center, and the Harvard Center for Research on Computation & Society.

Authors

Benjamin Fuller (benjamin.fuller@uconn.edu) is an Assistant Professor of Computer Science at University of Connecticut. Prior to joining the University of Connecticut, he was a technical staff member atthe Massachusetts Institute of Technology (MIT) Lincoln Laboratory from 2007 to 2016. He received a B.S. degree from Rensselaer Polytechnic Institute in 2006 and M.A. and Ph.D. degrees from Boston University in 2011 and 2015, respectively. His research focuses on theoretic and applied cryptography with a past focus on cryptographic key derivation. In particular, his research focuses on cryptography with noise.

Leonid Reyzin received his A.B. from Harvard College in 1996 and his Ph.D. from the Massachusetts Institute of Technology in 2001, advised by Professor Silvio Micali. Since receiving his Ph.D., he has been a Professor of Computer Science at Boston University. Leo Reyzin's research is in cryptography and network security, with work ranging from foundational to applied. In particular, he has made contributions to the development of leakage-resilient cryptography, secure key derivation, authenticated data structures, moderately hard functions, and cryptographic proof systems. The applications of his work have ranged from cryptocurrencies to internet routing.

Leo Reyzin has held visiting positions at MIT, UCLA, and IST Austria, as well as industry consulting positions. He has served as the general chair of Crypto 2020 and the program co-chair of the 2018 Theory of Cryptography Conference. His teaching was recognized by Boston University's Neu Family Award for Excellence in Teaching. His research was recognized with the 2014 Applied Networking Research Prize, the 2015 Theory of Cryptography Test of Time Award, and the 2017 Eurocrypt Best Paper Award, and the 2019 IACR Test of Time Award.

Adam Smith received his B.Sc. from McGill University in 1999 and his Ph.D. from the Massachusetts Institute of Technology in 2004, advised by Professor Madhu Sudan. He was on the faculty of the Pennsylvania State University from 2007 to 2017, and joined Boston University as a Professor of Computer Science in 2017. He has held visiting positions at the Weizmann Institute of Science, UCLA, Boston University, and Harvard. His work was awarded a Theory of Cryptography Test of Time Award in 2016, the 2017 Gödel Prize, and the 2019 Eurocrypt Test of Time Award.

References

[1] B. Fuller, L. Reyzin, and A. Smith, "When are fuzzy extractors possible?" in *Advances in Cryptology – ASIACRYPT*. Springer, 2016, pp. 277–306.
[2] R. Pappu, B. Recht, J. Taylor, and N. Gershenfeld, "Physical one-way functions," *Science*, vol. 297, no. 5589, pp. 2026–2030, 2002.
[3] 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.
[4] S. Brostoff and M. Sasse, "Are passfaces more usable than passwords?: A field trial investigation," *People and Computers*, pp. 405–424, 2000.

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

[6] 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.

[7] 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.

[8] 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.

[9] F. Monrose, M. K. Reiter, and S. Wetzel, "Password hardening based on keystroke dynamics," *International Journal of Information Security*, vol. 1, no. 2, pp. 69–83, 2002.

[10] 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.

[11] 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.

[12] M. Zviran and W. J. Haga, "A comparison of password techniques for multilevel authentication mechanisms," *The Computer Journal*, vol. 36, no. 3, pp. 227–237, 1993.

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

[14] 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.

[15] M. Blanton and W. M. Hudelson, "Biometric-based non-transferable anonymous credentials," in *Information and Communications Security*. Springer, 2009, pp. 165–180.

[16] J. Daugman, "Probing the uniqueness and randomness of iriscodes: Results from 200 billion iris pair comparisons," *Proceedings of the IEEE*, vol. 94, no. 11, pp. 1927–1935, 2006.

[17] 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.

[18] R. Renner and S. Wolf, "The exact price for unconditionally secure asymmetric cryptography," in *Advances in Cryptology - EUROCRYPT 2004*, ser. Lecture Notes in Computer Science, C. Cachin and J. Camenisch, Eds., vol. 3027. Springer, 2004, pp. 109–125.

[19] A. Juels and M. Wattenberg, "A fuzzy commitment scheme," in *Sixth ACM Conference on Computer and Communication Security*. ACM, Nov. 1999, pp. 28–36.

[20] J.-P. M. G. Linnartz and P. Tuyls, "New shielding functions to enhance privacy and prevent misuse of biometric templates," in *International Conference on Audio-and Video-Based Biometric Person Authentication*, 2003, pp. 393–402.

[21] P. Tuyls and J. Goseling, "Capacity and examples of template-protecting biometric authentication systems," in *Biometric Authentication*, ser. Lecture Notes in Computer Science, vol. 3087. Springer, 2004, pp. 158–170.

[22] F. Hao, R. Anderson, and J. Daugman, "Combining crypto with biometrics effectively," *Computers, IEEE Transactions on*, vol. 55, no. 9, pp. 1081–1088, 2006.

[23] Y. Wang, S. Rane, S. C. Draper, and P. Ishwar, "A theoretical analysis of authentication, privacy and reusability across secure biometric systems," *ArXiv CoRR report*, vol. 1112.5630, 2011.

[24] T. Ignatenko and F. M. Willems, "Biometric security from an information-theoretical perspective," *Foundations and Trends in Communications and Information Theory*, vol. 7, no. 2–3, pp. 135–316, 2012.

[25] T. Holenstein and R. Renner, "One-way secret-key agreement and applications to circuit polarization and immunization of public-key encryption," in *Advances in Cryptology - CRYPTO 2005*, ser. Lecture Notes in Computer Science, V. Shoup, Ed., vol. 3621. Springer, 2005, pp. 478–493.

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

[27] L. H. Harper, "Optimal numberings and isoperimetric problems on graphs," *Journal of Combinatorial Theory*, vol. 1, no. 3, pp. 385–393, 1966.

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

[29] N. Bitansky, R. Canetti, Y. T. Kalai, and O. Paneth, "On virtual grey box obfuscation for general circuits," in *Advances in Cryptology - CRYPTO 2014*, 2014.

[30] ——, "On virtual grey box obfuscation for general circuits," *Algorithmica*, vol. 79, no. 4, pp. 1014–1051, 2017.

[31] S. Garg, C. Gentry, S. Halevi, M. Raykova, A. Sahai, and B. Waters, "Candidate indistinguishability obfuscation and functional encryption for all circuits," *Proc. of FOCS*, 2013.

[32] ——, "Candidate indistinguishability obfuscation and functional encryption for all circuits," *SIAM Journal on Computing*, vol. 45, no. 3, pp. 882–929, 2016.

[33] X. Boyen, Y. Dodis, J. Katz, R. Ostrovsky, and A. Smith, "Secure remote authentication using biometric data," in *EUROCRYPT*. Springer, 2005, pp. 147–163.

[34] B. Barak, R. Canetti, Y. Lindell, R. Pass, and T. Rabin, "Secure computation without authentication," *J. Cryptology*, vol. 24, no. 4, pp. 720–760, 2011.

[35] P.-A. Dupont, J. Hesse, D. Pointcheval, L. Reyzin, and S. Yakoubov, "Fuzzy password-authenticated key exchange," in *Advances in Cryptology – EUROCRYPT*. Springer, 2018, pp. 393–424.

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

[37] 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.

[38] 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.

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

[40] 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.

[41] H. Tyagi, P. Viswanath, and S. Watanabe, "Interactive communication for data exchange," *IEEE Trans. Information Theory*, vol. 64, no. 1, pp. 26–37, 2018. [Online]. Available: https://doi.org/10.1109/TIT.2017.2769124

[42] 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.

[43] H. Tyagi and S. Watanabe, "Universal multiparty data exchange and secret key agreement," *IEEE Transactions on Information Theory*, vol. 63, no. 7, pp. 4057–4074, 2017.

[44] C. T. Li and V. Anantharam, "One-shot variable-length secret key agreement approaching mutual information," *CoRR*, vol. abs/1809.01793, 2018. [Online]. Available: http://arxiv.org/abs/1809.01793

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

[46] B. Fuller and L. Peng, "When are continuous-source fuzzy extractors possible?" in *IEEE International Symposium on Information Theory*, 2019.

[47] R. Ash, *Information Theory*. Intersciene Publishers, 1965.

[48] B. Skoric and P. Tuyls, "An efficient fuzzy extractor for limited noise," Cryptology ePrint Archive, Report 2009/030, 2009.

[49] P. Frankl and Z. Füredi, "A short proof for a theorem of Harper about Hamming-spheres," *Discrete Mathematics*, vol. 34, no. 3, pp. 311–313, 1981.

[50] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *Journal of the American statistical association*, vol. 58, no. 301, pp. 13–30, 1963.

[51] T. Holenstein, "Strengthening key agreement using hard-core sets," Ph.D. dissertation, ETH Zurich, May 2006, reprint as vol. 7 of *ETH Series in Information Security and Cryptography*, ISBN 3-86626-088-2, Hartung-Gorre Verlag, Konstanz, 2006.

[52] R. Fano, *Transmission of Information: A Statistical Theory of Communications*, ser. MIT Press Classics. M.I.T. Press, 1961.

[53] K. Yasunaga and K. Yuzawa, "On the limitations of computational fuzzy extractors," *work*, vol. 3, no. 1, pp. 7–9, 2018.

[54] 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.

[55] Y. Wen, S. Liu, and S. Han, "Reusable fuzzy extractor from the decisional diffie–hellman assumption," *Designs, Codes and Cryptography*, vol. 86, no. 11, pp. 2495–2512, 2018.

[56] Y. Wen and S. Liu, "Robustly reusable fuzzy extractor from standard assumptions," in *Advances in Cryptology – ASIACRYPT*. Springer, 2018, pp. 459–489.

[57] C. Herder, L. Ren, M. van Dijk, M.-D. 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, 2016.

[58] 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 *Proceedings of the 2018 on Asia Conference on Computer and Communications Security*. ACM, 2018, pp. 673–684.

[59] Y. Wen, S. Liu, and D. Gu, "Generic constructions of robustly reusable fuzzy extractor," in *IACR International Workshop on Public Key Cryptography*. Springer, 2019, pp. 349–378.