Code Offset in the Exponent

Luke Demarest
University of Connecticut, Storrs, CT, USA

Benjamin Fuller
University of Connecticut, Storrs, CT, USA

Alexander Russell
University of Connecticut, Storrs, CT, USA

Abstract

Fuzzy extractors derive stable keys from noisy sources. They are a fundamental tool for key derivation from biometric sources. This work introduces a new construction, code offset in the exponent. This construction is the first reusable fuzzy extractor that simultaneously supports structured, low entropy distributions with correlated symbols and confidence information. These properties are specifically motivated by the most pertinent applications – key derivation from biometrics and physical unclonable functions – which typically demonstrate low entropy with additional statistical correlations and benefit from extractors that can leverage confidence information for efficiency.

Code offset in the exponent is a group encoding of the code offset construction (Juels and Wattenberg, CCS 1999). A random codeword of a linear error-correcting code is used as a one-time pad for a sampled value from the noisy source. Rather than encoding this directly, code offset in the exponent encodes by exponentiation of a generator in a cryptographically strong group. We introduce and characterize a condition on noisy sources that directly translates to security of our construction in the generic group model. Our condition requires the inner product between the source distribution and all vectors in the null space of the code to be unpredictable.

2012 ACM Subject Classification Security and privacy → Information-theoretic techniques; Security and privacy → Biometrics

Keywords and phrases fuzzy extractors, code offset, learning with errors, error-correction, generic group model

Digital Object Identifier 10.4230/LIPIcs.ITC.2021.15

Related Version *Full Version*: `Cryptology ePrint Archive` [12]

Funding *Luke Demarest*: Funded in part by NSF Grants No. 1849904 and 1547399.
Benjamin Fuller: Funded in part by NSF Grants No. 1849904 and 1547399. This research is based upon work supported in part by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), via Contract No. 2019-19020700008. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of ODNI, IARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright annotation therein.
Alexander Russell: This material is based upon work supported by the National Science Foundation under Grant No. 1801487.

Acknowledgements The authors give special thanks to reviewer comments and feedback. The authors thank James Bartusek, Ryann Cartor, Fermi Ma, and Mark Zhandry and their helpful discussions of their work.

1 Introduction

Fuzzy extractors [13] permit derivation of a stable key from a noisy *source*. Specifically, given a reading $\mathbf{e}$ from the noisy source, the fuzzy extractor produces a pair $(\mathsf{key}, \mathsf{pub})$, consisting of a derived key and a public value; the public value pub must then permit key to (only) be recovered from any $\mathbf{e}'$ that is sufficiently close to $\mathbf{e}$ in Hamming distance. Fuzzy extractors are the emblematic technique for robust, secure key derivation from biometrics and physical unclonable functions. These applications place special emphasis on the source distribution and for this reason a principal goal of fuzzy extractor design is to precisely identify those distributions over $\mathbf{e}$ for which extraction is possible and, moreover, produce efficient constructions for these distributions.

Despite years of work, existing constructions do not simultaneously secure practical sources while retaining efficient recovery. Canetti et al.'s construction [8, 9] is secure for the widest variety of sources. However, Simhadri et al.'s [31] implementation for the iris estimates only 32 bits of security with algorithms that take ≈ 10 seconds on a 32-core machine.

The fuzzy extraction problem is well-understood in the information-theoretic setting, where the fundamental quantity of interest is the *fuzzy min-entropy* [18, 19] of the distribution of $\mathbf{e}$; this measures the total weight of an arbitrarily centered ball of radius t in the probability distribution over $\mathbf{e}$. While this measure is sufficient for determining the feasibility of information-theoretic fuzzy extraction for a distribution, it doesn't indicate whether it is possible in polynomial time [18, 34]. In the information-theoretic setting, it is not possible to build an information-theoretic fuzzy extractor that simultaneously works for all distributions [18, 17, 19]. That is, a fuzzy extractor exists for each distribution with fuzzy min-entropy but no construction can secure all such distributions.

One can hope to sidestep these limitations by providing only computational security [15, 16]. However, even in this more favorable setting no universal theory has emerged without resorting to general purpose obfuscation. Two known fuzzy extractors use "computational" tools[1] to correct errors, they are:

- Canetti et al.'s [8, 9] construction explicitly places random subsets of $\mathbf{e}$ in a *digital locker* [7] and records the indices used in each subset. To recover, one attempts to open each digital locker with subsets of the value $\mathbf{e}'$. Canetti et al.'s construction is secure when a random subset of locations is hard to predict (Definition 10). However, Simhadri's implementation for the iris provides poor security (32 bits) in order to run in 10 seconds and still requires millions of digital lockers [31].
- Fuller et al. [15, 16] modify the *code-offset* construction [24]. The code-offset construction construction is determined by a linear error-correcting code $\mathbf{A} \in \mathbb{F}_q^{n \times k}$ and a secret, uniformly random $\mathbf{x} \in \mathbb{F}_q^k$; given a sample $\mathbf{e} \in \mathbb{F}_q^n$ from the noisy source, the construction publishes the pair $\mathsf{pub} = (\mathbf{A}, \mathbf{A}\mathbf{x} + \mathbf{e})$. All operations are carried out over the field with q elements. To *reproduce* the value $\mathbf{e}$ note that with a second sample $\mathbf{e}'$ from the source – which we assume has small Hamming distance from $\mathbf{e}$[2] – the difference $(\mathbf{A}\mathbf{x} + \mathbf{e}) - \mathbf{e}' = \mathbf{A}\mathbf{x} + (\mathbf{e} - \mathbf{e}')$ is evidently close to the codeword $\mathbf{A}\mathbf{x}$. By decoding the error correcting code one can recover $\mathbf{x}$ (and $\mathbf{e}$).[3] Security analysis of the code offset treats $\mathbf{A}\mathbf{x}$ as a biased one time pad, proving that $\mathbf{A}\mathbf{x} + \mathbf{e}$ leaks no more than $(n - k) \log q$

[1] Multiple computational fuzzy extractors retain the information-theoretic core and analyze it using standard information-theory techniques [32, 33]; these works are subject to the above limitations.

[2] It is also possible to consider other distances between $\mathbf{e}$ and $\mathbf{e}'$. However the error correction techniques required are different. We consider Hamming error in this work.

[3] Applying a randomness extractor [25] on either $\mathbf{x}$ or $\mathbf{e}$ yields a uniform key.

bits about $\mathbf{e}$. However, many real distributions have entropy less than $(n-k)\log q$, which we call *low entropy*, for which this analysis provides no security guarantee. To support low entropy distributions, Fuller et al. instantiate this construction with $\mathbf{A}$ being randomly distributed and show security whenever the distribution over $\mathbf{e}$ yields a secure learning with errors (LWE) instance. Known LWE error distributions consider i.i.d. symbols (discretized Gaussian [28] and uniform interval [14]).

The digital locker construction supports more distributions (i.i.d. symbols implies that all subsets have entropy). Both constructions use information set decoding [26], that is, repeated selection of random subsets of coordinates with the hope to find a subset with no errors.

The digital locker construction comes with an important drawback. Many physical sources are sampled along with correlated side information that is called *confidence*. Confidence information is a secondary probability distribution $\mathbf{z}$ (correlated with the reading $\mathbf{e}$) that can predict the error rate in a symbol $\mathbf{e}_i$. When $\mathbf{z}_i$ is large this indicates that the symbol of $\mathbf{e}_i$ is less likely to differ. Examples include the magnitude of a convolution in the iris [31] and the magnitude of the difference between two circuit delays in ring oscillator PUFs [22]. By considering bits with high confidence it is possible to reduce the effective error rate from $t = n/10$ to $t = 3n/10^6$ [22]. For a subset size of 128 and $t = n/10$ unlocking with 95% probability requires testing approximately $2 \cdot 10^6$ subsets while $t = 3n/10^6$ requires testing a single subset. This confidence information cannot be used in the digital locker construction as subsets are specified at enrollment time whereas confidence information is determined when $\mathbf{e}$ is drawn. The LWE construction can use this information [23] as it allows on-the-fly testing of all large enough subsets. Confidence information is critical: fuzzy extractors that secure low entropy distributions do not support $t = \Theta(n)$ which is demonstrated in practice, leading to inefficient implementations. Because constructions are used with sources beyond their designed error tolerance any reduction in error rate has a drastic impact on efficiency (see Subsection 3.2).

Our contributions

This work introduces the *code offset in the exponent* construction. Code offset in the exponent yields the first reusable fuzzy extractor that simultaneously

- allows the symbols of $\mathbf{e}$ to be correlated,
- supports structured but low entropy distributions over $\mathbf{e}$ (less than $(n-k)\log q$), and
- allows the use of confidence information for improved efficiency.

This work introduces the *Code Offset in the Exponent* problem:

> Distinguish $r^{\mathbf{Ax}+\mathbf{e}}$, given $(\mathbf{A}, r)$, from a random tuple of group elements, where r is a random generator of a prime order group, $\mathbf{A}$ is a suitable linear code, and $\mathbf{x}$ is a uniform dimension k vector.

A natural fuzzy extractor constructor exists when $r^{\mathbf{Ax}+\mathbf{e}}$ has such pseudorandom properties. We show that when the group effectively limits the adversary to linear operations – by adopting the generic group model – the resulting fuzzy extractor is secure for many low entropy distributions while retaining the ability to use confidence information. This allows code offset in the exponent to benefit from the efficiency gains of using confidence information while remaining secure for a large family of distributions. Specifically, we present three contributions:

Sec 1.1 We define the code offset in the exponent construction and show that it yields a reusable fuzzy extractor if the distribution on $\mathbf{e}$ is *good enough*.

Sec 1.2 We define and describe what constitutes *good enough* in the generic group model with a novel information-theoretic sufficient condition we call MIPURS.

Sec 1.3 We characterize MIPURS, establishing containment relations between MIPURS and the secured distributions in Canetti et al. [8] and Fuller et al. [15] (see Figure 1).

We then review further related work and offer a table of comparisons (Sec 1.4). Section 2 covers definitions and preliminaries including the MIPURS condition. Section 3 details the code offset in the exponent construction. Section 4 characterizes MIPURS distributions.

1.1 Code offset in the exponent

Code offset in the exponent is motivated by the observation that *reproduction* of $\mathbf{e}$ in the LWE construction uses only linear operations. Thus, we explore an adaptation of the code offset construction that effectively limits the adversary to linear operations by translating all relevant arithmetic into a "hard" group. Specifically, we introduce *code offset in the exponent*: If r is a random generator for a cyclic group $\mathbb{G}$ of prime order q, we consider $\mathsf{pub} = (\mathbf{A}, r, r^{\mathbf{Ax}+\mathbf{e}})$ where we adopt the shorthand notation $r^{\mathbf{v}}$, for a vector $\mathbf{v} = (v_1, \ldots, v_n)^\intercal \in (\mathbb{Z}_q)^n$, to indicate the vector $(r^{v_1}, \ldots, r^{v_n})^\intercal$. This construction possesses strong security properties under natural cryptographic assumptions on the group $\mathbb{G}$. We focus on code-offset in the exponent with a random linear code (given by $\mathbf{A}$) and adopt the generic group model [30] to reflect the cryptographic properties of the underlying group. As stated above, the goal is to characterize the distributions on $\mathbf{e}$ for which $r^{\mathbf{Ax}+\mathbf{e}}$ given $(\mathbf{A}, r)$ is pseudorandom. Pseudorandomness suffices to show security of a fuzzy extractor that leaks nothing about $\mathbf{e}$. Analysis of this construction is most natural when $\mathbf{e}$ has symbols over a large alphabet, but binary $\mathbf{e}$ can be amplified (see Section 3.1).

Looking ahead, if one uses a random generator in each enrollment the construction allows multiple (noisy) enrollments of $\mathbf{e}$, known as a reusable fuzzy extractor [6]. The reusability proof uses the details of the generic group proof, while the one time analysis is just based on pseudorandomness. See the full version of this work for details of that proof.

1.2 When is code offset in the exponent hard?

In the generic group model, we establish that distinguishing code offset in the exponent from a random vector of group elements is hard (Theorem 4) for any error distribution $\mathbf{e}$ where the following game is hard to win for any information-theoretic adversary $\mathcal{A}$:[4]

Experiment $\mathbb{E}^{\mathsf{MIPURS}}_{\mathcal{A},\mathbf{e}}(n,k)$:
$\psi \leftarrow \mathbf{e}$; $A \xleftarrow{\$} \mathbb{F}_q^{n\times k}$.
$(b, g) \leftarrow \mathcal{A}(A, \mathbf{e})$.
If $b \in \mathtt{null}(A)$, $b \neq \vec{0}$ and $\langle b, \psi\rangle = g$ output 1.
Output 0.

Observe that the role of the random matrix A in the game above is merely to define a random subspace of (typical) dimension k.

We call this condition on an error distribution MIPURS or *maximum inner product unpredictable over random subspace*. Specifically, a random variable $\mathbf{e}$ over $\mathbb{F}_q^n$ is (k, β)–MIPURS if for all $\mathcal{A}$ (which knows the distribution of $\mathbf{e}$ but not the sampled value ψ), $\Pr[\mathbb{E}^{\mathsf{MIPURS}}_{\mathcal{A},\mathbf{e}}(n,k) = 1] \leq \beta$.

[4] We use boldface to represent random variables, capitals to represent random variables over matrices, and plain letters to represent samples. We use ψ to represent samples from $\mathbf{e}$ to avoid conflict with Euler's number.

When $\mathbf{e}$ is a $(k - \Theta(1), \beta)$–MIPURS distribution for a code with dimension k and $\beta = \mathtt{ngl}(n)$ then code-offset in the exponent yields a fuzzy extractor in the generic group model (Theorem 5). Showing this requires one additional step of key extraction; we use a result of Akavia, Goldwasser, and Vaikuntanathan [1, Lemma 2] which states that dimensions of $\mathbf{x}$ become *hardcore* once there are enough dimensions for LWE to be indistinguishable. This reduction is entirely linear and holds in the generic group setting.

MIPURS is necessary. When $\mathcal{A}$ is information theoretic, for all distributions $\mathbf{e}$ that are not MIPURS one can find a nonzero vector b in the null space of A whose inner product with $\mathbf{e}$ is predictable, thus predicting $\langle b, \mathbf{Ax} + \mathbf{e}\rangle = \langle b, \mathbf{e}\rangle \stackrel{?}{=} g$. This is not the case for a uniform distribution, $\mathbf{U}$: the value $\langle b, \mathbf{U}\rangle$ is uniform (and thus is $\langle b, \mathbf{U}\rangle = g$ with small probability if the size of q is super polynomial). Thus the vector b serves as a way to distinguish $\mathbf{Ax} + \mathbf{e}$ from $\mathbf{U}$.

Beullens and Wee [3] recently introduced the KOALA assumption which roughly assumes that an adversary's only mechanism for distinguishing a vector from a subspace from random is by outputting a vector that is likely to be the null space of the provided vector. This can be seen as specializing [11, Assumption 5] that vectors can only be distinguished by fixed inner products.

The adversary has more power in the MIPURS setting (than in KOALA) in three ways. First, the distribution $\mathbf{e}$ and thus $A\mathbf{x} + \mathbf{e}$ is not linear, second the adversary doesn't have to "nullify" all subspaces – only a single vector, and third, the adversary can predict any inner product, not just 0. One can view MIPURS as an assumption on a group: Whenever an adversary can distinguish the (nonlinear) vector $A\mathbf{x} + \mathbf{e}$ from uniform that there is another adversary that can choose some b and predict $\langle b, A\mathbf{x} + \mathbf{e}\rangle$ (in our setting this choice of b is after seeing A). Theorem 4 can be interpreted as the MIPURS "assumption" holding in the generic group model.

1.3 Supported Distributions

Our technical work characterizes the MIPURS property (summarized in Figure 1). The most involved relationship is showing that all high entropy sources are MIPURS. To provide intuition for our results, we summarize this result here.

For any $d = \mathtt{poly}(n)$ there is an efficiently constructible distribution $\mathbf{e}$ whose entropy is approximately $\log(dq^{n-k-1})$ where the MIPURS game is winnable by an efficient adversary with noticeable probability: For $1 \leq i \leq d$, sample some d random linear spaces $\mathbf{B}_i$ of dimension $n - k - 1$ and define $\mathbf{E}_i$ to be all points in a random coset g_i of $\mathbf{B}_i$. Consider the following distribution $\mathbf{e}$: Pick $i \leftarrow \{1, ..., d\}$ for some polynomial size d then output a random element of $\mathbf{E}_i$. The support size of this distribution is approximately dq^{n-k-1}. Then since $\mathtt{null}(\mathbf{A})$ has dimension at least $n - k$, $\exists b_i \neq \vec{0}$ such that $b_i \in \mathtt{null}(\mathbf{A}) \cap \mathtt{null}(\mathbf{B}_i)$ (since $\dim(\mathtt{null}(\mathbf{A})) + \dim(\mathtt{null}(\mathbf{B}_i)) > n$). The adversary can calculate these b_i's. Then the adversary just picks a random i and predicts (b_i, g_i).[5] This result is nearly tight: all distributions whose entropy is greater than $\log(\mathtt{poly}(n)q^{n-k})$ are MIPURS. Note this is a factor of q away from matching the size of our counterexample for a random code. Informally, this yields the following (see Corollary 25):

▶ **Theorem 1** (Informal). *Let $n, k \in \mathbb{Z}$ be parameters. Let $q = q(n)$ be a large enough prime. For all $\mathbf{e} \in \mathbb{Z}_q^n$ whose minentropy is at least $\omega(\log n) + \log(q^{n-k})$, there exists some $\beta = \mathtt{ngl}(n)$ for which $\mathbf{e}$ is (k, β)–MIPURS.*

[5] If $\mathbf{A}$ is some fixed code (chosen before adversary specifies $\mathbf{e}$), then $\mathbf{E}_i$ can directly be a coset of $\mathbf{A}$ and one can increase the size of $\mathbf{E}$ to dq^{n-k}.

As mentioned above, information theoretic analysis of code offset provides a key of length $\omega(\log n)$ when the initial entropy of $\mathbf{e}$ is at least $\omega(\log n)+(n-k)\log(q)$. However, information theoretic analysis of code offset reduces the entropy of $\mathbf{e}$ which may allow prediction of sensitive attributes. In the generic group analysis no predicate of $\mathbf{e}$ is leaked. The generic group analysis also allows the construction to be safely reused multiple times (with independent generators).

Proof Intuition

Suppose in the above game the adversary generated $\mathbf{e}$ as the span of a linear space $\mathbf{E}$ with the goal that $\mathtt{null}(\mathbf{A}) \cap \mathtt{null}(\mathbf{E}) \supset \{\vec{0}\}$. For a random, independent $\mathbf{B} \stackrel{\text{def}}{=} \mathtt{null}(\mathbf{A})$, the probability of $\mathbf{B}$ and $\mathtt{null}(\mathbf{E})$ overlapping is noticeable only if the sum of the dimensions is more than n (Lemma 19). This creates an upper bound on the dimension of $\mathbf{E}$ of $n-k$ (ignoring the unlikely case when $\mathbf{A}$ is not full rank).

Our proof is dedicated to showing that the general case (where $\mathbf{E}$ is not linear) does not provide the adversary with more power. First we upper bound the size of a set E where each vector is predictable in the MIPURS game. We show for a random sample from E to have a large intersection with a low dimensional space requires E to have size at least that of the low dimensional space (Lemma 18). In Lemma 20, we switch from measuring the size of intersection of a sample of E with respect to the worst case subspace to how "linear" E is with respect to the worst vector in an average case subspace. This result thus controls an "approximate" algebraic structure in the sense of additive combinatorics. We show the adversary can't do much better on a single vector b as long as it is chosen from a random $\mathbf{B}$.

The above argument considers the event that the adversary correctly predicts an inner product of 0; this can be transformed to an arbitrary inner product by a compactness argument which introduces a modest loss in parameters (Theorem 22). Once we have a bound on how large a predictable set E can be, another superlogarithmic factor guarantees that all distributions $\mathbf{e}$ with enough minentropy are not predictable.

1.4 Further Related Work

We have already introduced the work of Canetti et al. [8] and Fuller et al. [16]. Canetti et al. [8] explicitly place some subsets into a digital locker, for security they require that an average subset has average min-entropy, which we call *average subsets have entropy*.

Lemma 11 shows that the MIPURS condition is contained in average subsets have entropy. This containment is proper, we actually show that there are distributions where all subsets have entropy that are not MIPURS. Suppose that $\mathbf{e}$ is a Reed-Solomon code, then all subsets of $\mathbf{e}$ have entropy but as long as the dimension of the code $< n-k-1$ then the null space of $\mathbf{e}$ is likely to intersect with $\mathtt{null}(\mathbf{A})$ (Prop. 14).

There are also MIPURS sources where not all subsets have entropy. Consider a uniform distribution over $n-k$ coordinates with a fixed value in the remaining k coordinates (Prop. 13). Since $\mathtt{null}(\mathbf{A})$ is unlikely to have non zero coordinates only at these fixed k coordinates, predicting the inner product remains difficult. Fortunately, multiplying a binary source where all subsets have entropy by a random vector produces a location source which is contained in MIPURS. It is this transformation we recommend for actual biometrics, see Section 3.1.

One can additionally build a good fuzzy extractor assuming a variant of multilinear maps [5]. Concurrent work of Galbraith and Zobernig [20] introduces a new subset sum assumption to build a secure sketch that is able to handle $t = \Theta(n)$ errors; they conjecture hardness for all securable distributions. A secure sketch is the error correction component in

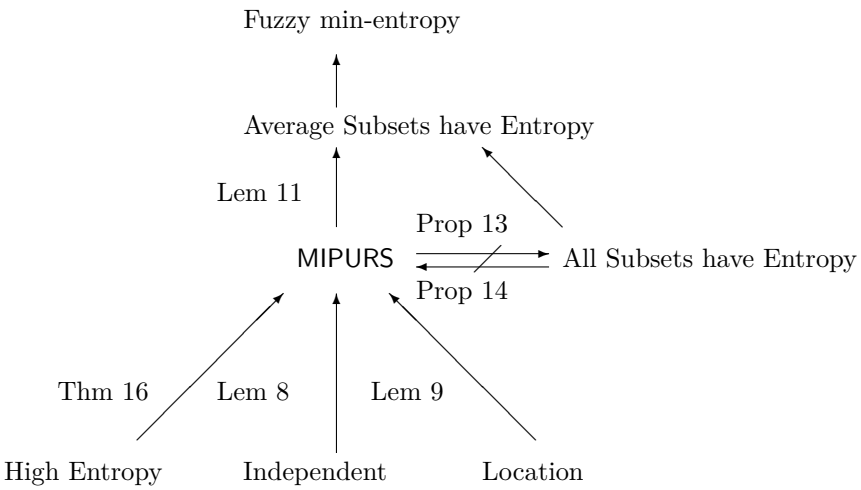

Figure 1 Implications between different types of supported distributions for fuzzy extraction. Arrows are implications. All shown implications are proper. Location sources are those that have random group elements in some locations with zeroes in other locations but it is hard to find a subset of all zero locations. A location source can be produced as the component wise product of a binary source where all subsets have entropy and a random vector of group elements. We consider this type of distribution in Section 3.1.

most fuzzy extractors. Their assumption is security of the cryptographic object and deserves continued study. A line of works [32, 33] use information-theoretic tools for error correction and computational tools to achieve additional properties. Those constructions embed a variant of the code offset. Table 1 summarizes constructions that use computational tools for the "correction" component and the traditional information theoretic analysis of the code offset construction.

2 Notation and Preliminaries

2.1 Notation

We use boldface to represent random variables, capitals to represent random variables over matrices or sets, and corresponding plain letters to represent samples. As one notable exception, we use ψ to represent samples from $\mathbf{e}$ to avoid conflict with Euler's number. We denote the exponential function with $\exp(\cdot)$. When defining ranges for parameters, we use [and] to indicate ranges inclusive of indicated values and (and) to indicate ranges exclusive of the indicated values. For random variables $\mathbf{x}_i$ over some alphabet $\mathcal{Z}$ we denote the tuple by $\mathbf{x} = (\mathbf{x}_1, ..., \mathbf{x}_n)$. For a vector $\mathbf{v}$ we denote the ith entry as $\mathbf{v}_i$. For a set of indices J, $\mathbf{x}_J$ denotes the restriction of $\mathbf{x}$ to the indices in J. For $m \in \mathbb{N}$, we let $[m] = \{1, \ldots, m\}$, so that $[0] = \emptyset$. We use the notation $\mathrm{span}(S)$ to denote the linear span of a set S of vectors and apply the notation to sequences of vectors without any special indication: If $F = (f_1, \ldots, f_m)$ is a sequence of vectors, $\mathrm{span}(F) = \mathrm{span}(\{f_i \mid i \in [m]\})$. The *min-entropy* of a random variable $\mathbf{x}$ is $\mathrm{H}_\infty(\mathbf{x}) = -\log(\max_x \Pr[\mathbf{x} = x])$.

We consider the Hamming metric. Let $\mathcal{Z}$ be a finite set and consider elements of $\mathcal{Z}^n$; then we define $\mathsf{dis}(x, y) = |\{i \mid x_i \neq y_i\}|$. U_n denotes the uniformly distributed random variable on $\{0,1\}^n$. Logarithms are base 2. We denote the vector of all zero elements as 0. We let $\cdot_c$

Table 1 Comparison of computational techniques for fuzzy extractors. Many schemes [32, 33] use information theoretic techniques for information reconciliation and these are grouped together. These techniques all inherit the information theoretic analysis on the strength of information reconciliation. Reuse is denoted as ◐ if reuse is supported with some assumption about how multiple readings are correlated and ● if no assumption is made. See Figure 1 for relations between supported distributions. The LWE works considered the setting when $k = \Theta(n)$ which leads to $t = \Theta(\log n)$. If one sets $k = \omega(\log n)$ one can achieve error tolerance of $o(n)$ using the analysis in this work, we thus present the more favorable regime for the above comparison.

Construction	Supported low entropy dist.	Reuse	Error rate (t)	Weakness
Code Offset [13]	-	◐	$\Theta(n)$	
LWE [2, 15]	Independent	◐	$o(n)$	
Subset sum [20]	Fuzzy min-ent.	○	$\Theta(n)$	Assumes security
Grey box obf. [5]	Fuzzy min-ent.	○	$\Theta(n)$	Multilinear maps
Digital Locker [8]	Average Subsets have Ent.	●	$o(n)$	No *confidence* info
This work	MIPURS	●	$o(n)$	

denote component-wise multiplication. In our theorems we consider a security parameter γ, when we use the term negligible and super polynomial, we assume other parameters are functions of γ. We elide the notational dependence of other parameters on γ.

2.2 Fuzzy Extractors

Our motivating application is a new fuzzy extractor that performs error correction "in the exponent." A fuzzy extractor is a pair of algorithms designed to extract stable keys from a physical randomness source that has entropy but is noisy. If repeated readings are taken from the source one expects these readings to be close in an appropriate distance metric but not identical. We consider a generic group version of security (computational security is defined in [15], information-theoretic security in [13]).

Before introducing the definition, we review some notation from the generic group model; the model is reviewed in detail in the full version of this work. Let $\mathbb{G}$ be a group of prime order q. For each element $r \in \mathbb{G}$ in the standard game, rather than receiving r, the adversary receives a handle $\sigma(r)$ where σ is a random function with a large range. The adversary is given access to an oracle, which we denote as $\mathcal{O}^{\sigma}_{\mathbb{G}}$, which given $x = \sigma(r_1), y = \sigma(r_2)$ computes $\sigma(\sigma^{-1}(x) + \sigma^{-1}(y))$; when σ can be inferred from context, we write $\mathcal{O}_{\mathbb{G}}$. Since the adversary receives random handles they cannot infer anything about the underlying group elements except using the group operation and testing equality. We assume throughout that the range of σ is large enough that the probability of a collision is statistically insignificant (that is $\ll 1/q$).

We overload the notation $\sigma(\cdot)$ to apply to tuples and, furthermore, adopt the convention that $\sigma(\cdot)$ is the identity on non-group elements; thus, it can be harmlessly applied to all inputs provided to the adversary. Specifically, when $z \stackrel{\text{def}}{=} z_1, \ldots, z_n$ then $\sigma(z)$ only passes z_i through σ if $z_i \in \mathbb{G}_q$. For example, if $z = (r, \mathbf{A}, r^{\mathbf{Ax}+\mathbf{w}})$, then $\sigma(z) = (\sigma(r), \mathbf{A}, \sigma(r^{\mathbf{Ax}+\mathbf{w}}))$.

▶ **Definition 2.** *Let $\mathcal{E}$ be a family of probability distributions over the metric space $(\mathcal{M}, \mathsf{dis})$. A pair of procedures $(\mathsf{Gen} : \mathcal{M} \to \{0,1\}^{\kappa} \times \{0,1\}^{*}, \mathsf{Rep} : \mathcal{M} \times \{0,1\}^{*} \to \{0,1\}^{\kappa})$ is an $(\mathcal{M}, \mathcal{E}, \kappa, t)$-*fuzzy extractor *that is (ϵ_{sec}, m)-hard with error δ if Gen and Rep satisfy the following properties:*

- Correctness*: if* $\mathsf{dis}(\psi, \psi') \leq t$ *and* $(\mathsf{key}, \mathsf{pub}) \leftarrow \mathsf{Gen}(\psi)$*, then*

$$\Pr[\mathsf{Rep}(\psi', \mathsf{pub}) = \mathsf{key}] \geq 1 - \delta.$$

- Security*: for any distribution* $\mathbf{e} \in \mathcal{E}$*, the string* key *is close to random conditioned on* pub *for all* $\mathcal{A}$ *making at most* m *queries to the group oracle* $\mathcal{O}_{\mathbb{G}}$*, that is*

$$\left|\Pr[\mathcal{A}^{\mathcal{O}_{\mathbb{G}}}(\sigma(\mathsf{key}, \mathsf{pub})) = 1] - \Pr[\mathcal{A}^{\mathcal{O}_{\mathbb{G}}}(\sigma(U, \mathsf{pub})) = 1]\right| \leq \epsilon_{sec}.$$

Where the probability of the statement is taken over $\sigma \xleftarrow{\$} \Sigma$ *and* $\mathsf{key}, \mathsf{pub}) \leftarrow \mathsf{Gen}(\mathbf{e})$. We also assume that the adversary receives $\sigma(1)$. The errors are chosen before pub: if the error pattern between ψ and ψ' depends on the output of Gen, then there is no guarantee about the probability of correctness.

2.3 The MIPURS condition

In this section, we introduce our novel *Maximum Inner Product Unpredictable over Random Subspace* (MIPURS) condition.

▶ **Definition 3.** *Let* $\mathbf{e}$ *be a random variable taking values in* $\mathbb{F}_q^n$ *and let* $\mathbf{A}$ *be uniformly distributed over* $\mathbb{F}_q^{n \times k}$ *and independent of* $\mathbf{e}$*. We say that* $\mathbf{e}$ *is a* $(k, \beta) - \mathsf{MIPURS}$ *distribution if for all random variables* $\mathbf{b} \in \mathbb{F}_q^n, \mathbf{g} \in \mathbb{F}_q$ *independent of* $\mathbf{e}$ *(but depending arbitrarily on* $\mathbf{A}$ *and each other)*

$$\mathop{\mathbb{E}}_{\mathbf{A}}\left[\Pr\left[\langle \mathbf{b}, \mathbf{e}\rangle = \mathbf{g} \text{ and } \mathbf{b} \in \mathtt{null}(\mathbf{A}) \setminus \vec{0}\right]\right] \leq \beta\,.$$

To see the equivalence between this definition and the game presented in the introduction, the random variables $\mathbf{b}$ and $\mathbf{g}$ can be seen as encoding the "adversary" and quantifying over all $(\mathbf{b}, \mathbf{g})$ is equivalent to considering all information-theoretic adversaries.

▶ **Theorem 4.** *Let* γ *be a security parameter. Let* q *be a prime and* $n, k \in \mathbb{Z}^+$ *with* $k \leq n \leq q$*. Let* $\mathbf{A} \in \mathbb{F}_q^{n \times k}$ *and* $\mathbf{x} \in \mathbb{F}_q^k$ *be uniformly distributed. Let* $\mathbf{e}$ *be a* $(k, \beta) - \mathsf{MIPURS}$ *distribution. Let* $\mathbf{u} \in (\mathbb{F}_q)^n$ *be uniformly distributed. Let* Σ *be the set of random functions with domain of size* q *and range of size* q^3*. Then for all adversaries* $\mathcal{D}$ *making at most* m *queries*

$$\left|\Pr_{\sigma \xleftarrow{\$} \Sigma}[\mathcal{D}^{\mathcal{O}_{\mathbb{G}}}(\mathbf{A}, \sigma(\mathbf{A}\mathbf{x} + \mathbf{e})) = 1] - \Pr[\mathcal{D}^{\mathcal{O}_{\mathbb{G}}}(\mathbf{A}, \sigma(\mathbf{u})) = 1]\right| < \mu\left(\frac{3}{q} + \beta\right)$$

for $\mu = ((m+n+2)(m+n+1))^2/2$*. If* $1/q = \mathtt{ngl}(\gamma), n, m = \mathtt{poly}(\gamma)$*, and* $\beta = \mathtt{ngl}(\gamma)$ *then the statistical distance between the two cases is* $\mathtt{ngl}(\gamma)$*.*

In the above, the adversary is provided the code directly in the group, not its image in the handle space. The proof of Theorem 4 is a relatively straightforward application of the simultaneous oracle game introduced by Bishop et al. [4, Section 4]; this proof appears in the full version of this work.

3 A Fuzzy Extractor from Hardness of Code Offset in the Exponent

One can directly build a fuzzy extractor out of any $\mathbf{e}$ that satisfies the MIPURS condition. To do so, one instantiates the code-offset construction "in the exponent" and then uses hardcore elements of $\mathbf{x}$ as the key.

▶ **Construction 1.** *Let γ be a security parameter, t be a distance, $k = \omega(\log \gamma)$, $\alpha \in \mathbb{Z}^+$, $\ell \in \mathbb{Z}^+$, let q be a prime and let $\mathbb{G}_q$ be a cyclic group of order q. Let $\mathbb{F}_q$ be the field with q elements. Suppose that $\mathbf{e}$ and $\mathbf{e}' \in \mathbb{F}_q^n$, and let* dis *be the Hamming metric. Define* (Gen, Rep) *as follows:*

$\mathsf{Gen}(\psi = \psi_1, \ldots, \psi_n)$

1. *Sample generator r of $\mathbb{G}_q$.*
2. *Sample $A \leftarrow \mathbb{F}_q^{n \times (k+\alpha)}$, $x \leftarrow \mathbb{F}_q^{k+\alpha}$.*
3. *For $i = 1, ..., n$: set $r^{c_i} = r^{A_i \cdot x + \psi_i}$.*
4. *Set $\mathsf{key} = r^{x_0}, ..., r^{x_{\alpha-1}}$.*
5. *Set $\mathsf{pub} = (r, A, \{r^{c_i}\}_{i=1}^n)$.*
6. *Output* (key, pub).

$\mathsf{Rep}(\psi', \mathbf{pub} = (\boldsymbol{r}, \boldsymbol{A}, \boldsymbol{r^{c_1} \ldots r^{c_n}}))$

1. *For $i = 1, ..., n$, set $r^{c'_i} = r^{c_i} / r^{\psi'_i}$.*
2. *For $i = 1, ..., \ell$:*
   - (i) *Sample $J_i \subseteq \{1, ..., n\}$ where $|J| = k + \alpha$.*
   - (ii) *If $A_{J_i}^{-1}$ does not exist go to 2.*
   - (iii) *Compute $r^s = r^{A_{J_i}^{-1} c'_{J_i}}$.*
   - (iv) *Compute $r^{c''} = r^{As}$.*
   - (v) *If $\mathsf{dis}(r^{c'}, r^{c''}) \leq t$, output $r^{s_0}, ..., r^{s_\alpha}$.*
3. *Output $\perp$.*

▶ **Theorem 5.** *Let c be a constant. Let all parameters be as in Construction 1. Let $\mathcal{E}$ be the set of all (k, β)-*MIPURS *distributions. Suppose that*

- *$k' \stackrel{def}{=} k + \alpha = o(n)$ and $k' = \omega(\log n)$,*
- *t is such that $tk' \leq cn \log n$ for some constant c, which with the above implies $t = o(n)$,*
- *Let $\delta' > 0$ be some value,*
- *Let $\eta > 0$ be some constant and let $\ell = n^{2(1+\eta)c} \log \frac{1}{\delta'}$, and*
- *Let δ be some value such that $\delta \leq \delta' + exp(-\Omega(n))$.*

Then (Gen, Rep) *is a $(\mathbb{F}_q^n, \mathcal{E}, |\mathbb{F}_q^\alpha|, t)$-fuzzy extractor that is (ϵ_{sec}, m)-hard that is correct with probability $1 - \delta$ for all adversaries in the generic group model (making at most m queries) where*

$$\epsilon_{sec} = \left(\frac{((m+n+2)(m+n+1))^2}{2} \right) \left(\frac{3}{q} + \beta \right).$$

The proof of Theorem 5 is shown in the full version of this work [12].

3.1 Handling binary sources

In this section we show one way to transform binary sources to a good MIPURS distribution and consider the associated impact on correctness. Assume that the source $\mathbf{e}$ takes binary vlaues and all subsets of $\mathbf{e}$ are hard to predict, one can form a MIPURS distribution by multiplying by an auxiliary random and uniform random variable $\mathbf{r} \in \mathbb{F}_q^n$. This has the effect of placing random errors in the locations where $\mathbf{e}_i = 1$. Since decoding finds a subset without errors (it does not rely on the magnitude of errors) we can augment errors into random errors. We prove that this augmented vector is MIPURS in Section 4.

However, this transform creates a problem with decoding. When bits of $\mathbf{e}$ are 1, denoted $\mathbf{e}_i = 1$ we cannot use location i for decoding as it is a random value (even if $\mathbf{e}'_j = 1$ as well). When one amplifies a binary $\mathbf{e}$, we recommending using another uniform random variable $\mathbf{y} \in \{0,1\}^n$ and check when $\mathbf{y}_i \neq \mathbf{e}_i$ to indicate when to include a random error. Then in reproduction the algorithm should restrict to locations where $\mathbf{y}_i = \mathbf{e}_i$. Using Chernoff bounds one can show this subset is big enough and the error rate in this subset is not much higher than the overall error rate (except with negligible probability). If $k + \alpha$ is just barely $\omega(\log n)$ one can support error rates that are just barely $o(n)$.

To introduce the construction we first need to formalize the required property of the distribution $\mathbf{e}$. We introduce a notion called *all subsets have entropy*:

▶ **Definition 6.** *Let a source* $\mathbf{e} = \mathbf{e}_1, \ldots, \mathbf{e}_n$ *consist of* n*-bit binary strings. For some parameters* k *and* β *we say that the source* $\mathbf{e}$ *is a source where* ***all* k*-subsets have entropy* β** *if* $\mathrm{H}_\infty(\mathbf{e}_{j_1}, \ldots, \mathbf{e}_{j_k}) \geq \beta$ *for any* $1 \leq j_1, \ldots, j_k \leq n$, $j_a \neq j_b$ *for* $a \neq b$.

▶ **Construction 2.** *Let* γ *be a security parameter,* t *be a distance,* $k = \omega(\log \gamma)$, $\alpha \in \mathbb{Z}^+$, q *be a prime and let* $\mathbb{G}_q$ *be some cycle group of order* q*. Let* $\mathbb{F}_q$ *be the field with* q *elements. Let* $\mathcal{E} \in \{0,1\}^n$ *and let* dis *be the Hamming metric. Let* $\tau = \max(0.01, t/n)$*. Define* $(\mathsf{Gen}, \mathsf{Rep})$ *as follows:*

$\mathsf{Gen}(\psi = \psi_1, ..., \psi_n)$

1. *Sample random generator* r *of* $\mathbb{G}_q$.
2. *Sample* $A \leftarrow (\mathbb{F}_q)^{n \times (k+\alpha)}$,
3. *Sample* $x \leftarrow (\mathbb{F}_q)^{k+\alpha}$.
4. *Sample* $y \xleftarrow{\$} \{0,1\}^n$.
5. *For* $i = 1, ..., n$:
   - (i) *If* $\psi_i = y_i$, *set* $r^{\mathsf{c}_i} = r^{A_i \cdot x}$.
   - (ii) *Else set* $r^{\mathsf{c}_i} \xleftarrow{\$} \mathbb{G}_q$.
6. *Set* $\mathsf{key} = r^{x_{0 \ldots \alpha-1}}$.
7. *Set* $\mathsf{pub} = (r, y, A, \{r^{\mathsf{c}_i}\}_{i=1}^n)$.
8. *Output* $(\mathsf{key}, \mathsf{pub})$.

$\mathsf{Rep}(\psi', \mathsf{pub} = (r, y, A, r^{\mathsf{c}_1} \ldots r^{\mathsf{c}_\ell}))$

1. *Let* $\mathcal{I} = \{i | \psi'_i = y_i\}$.
2. *For* $i = 1, ..., \ell$:
   - (i) *Choose random* $J_i \subseteq \mathcal{I}$, *with* $|J_i| = k$.
   - (ii) *If* $A_{J_i}^{-1}$ *does not exist, output* $\perp$.
   - (iii) *Compute* $r^s = r^{A_{J_i}^{-1} \mathsf{c}_{J_i}}$.
   - (iv) *Compute* $r^{\mathsf{c}'} = r^{A(A_{J_i}^{-1} \mathsf{c}_{J_i})}$.
   - (v) *If* $\mathsf{dis}(\mathsf{c}_{\mathcal{I}}, \mathsf{c}'_{\mathcal{I}}) \leq 2|\mathsf{c}_{\mathcal{I}}|\tau$, *output* $r^{s_0}, ..., r^{s_{\alpha-1}}$.
3. *Output* $\perp$.

▶ **Theorem 7.** *Let all parameters be as in Construction 2. Let* $\gamma \in \mathbb{N}$ *and let* $\mathcal{E}$ *be the set of all sources where all* $(k-\gamma)$*-subsets have entropy* β *(Definition 6) over* $\{0,1\}^n$*. Then* $(\mathsf{Gen}, \mathsf{Rep})$ *is a* $(\{0,1\}^n, \mathcal{E}, |\mathbb{F}_q^\alpha|, t)$*-fuzzy extractor that is* (ϵ_{sec}, m)*-hard for all adversaries in the generic group model (making at most* m *queries) where*

$$\epsilon_{sec} = \left(\frac{((m+n+2)(m+n+1))^2}{2}\right)\left(\frac{4}{q} + 2^{-\beta} + \left(\frac{(k-\gamma)\binom{n}{k-\gamma-1}}{q^{\gamma+1}}\right)\right).$$

Furthermore, suppose that

- $k' \stackrel{def}{=} k + \alpha = o(n)$ *and* $k' = \omega(\log n)$,
- t *is such that* $tk' \leq cn \log n$ *for some constant* c, *which with the above implies* $t = o(n)$,
- *Let* $\delta' > 0$ *be some value.*
- *Let* $\eta > 0$ *be some constant.*
- *Let* $\ell = n^{2(1+\eta)c} \log \frac{1}{\delta'}$, *(if* $tk' = o(n \log n)$ *setting* $\ell = n \log 1/\delta'$ *suffices)*

Then there is some function negligible $\mathtt{ngl}(n)$ *such that the* Rep *is correct with probability* $1 - \delta' - \mathtt{ngl}(n)$.

We defer proving Theorem 7 to the full version of this work [12].

3.2 The power of confidence information

Most PUFs and biometrics demonstrate a constant error rate $\tau = t/n$. This is higher than the correction capacity of our construction and Canetti et al.'s digital locker construction [8]. However, existing fuzzy extractors that support constant τ do not support the low entropy distributions found in practice.

While code offset in the exponent is not designed for constant error rates it is efficient for small constant τ. As described in the introduction for the case of PUFs and biometrics, using confidence information can lead to a multiplicative decrease in the effective error rate of bits chosen for information set decoding. The important tradeoff is between the fractional error rate $\tau = t/n$ and the number of required iterations.

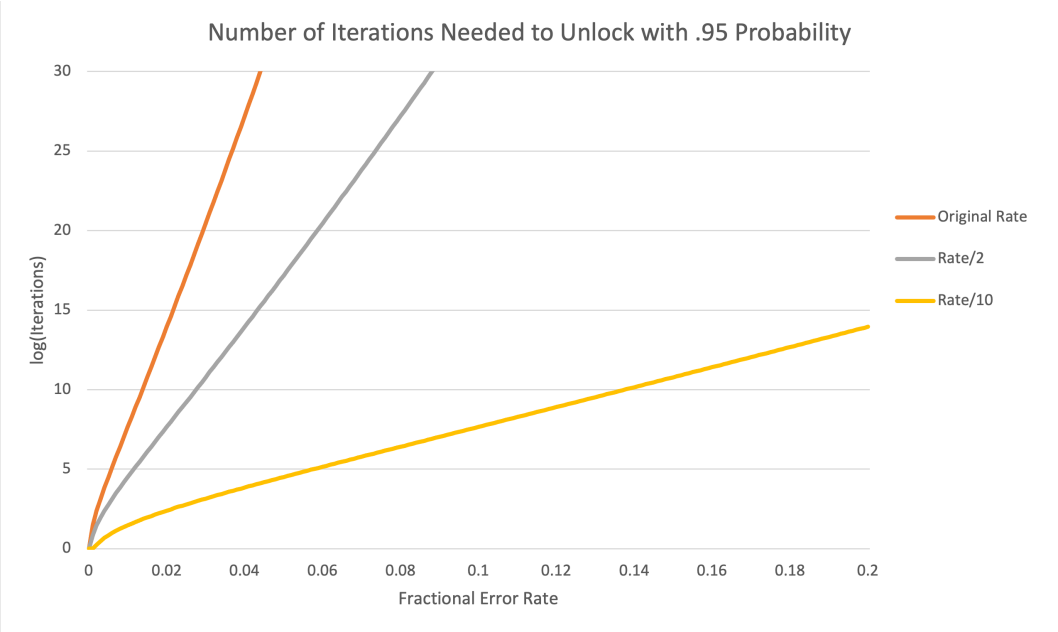

Figure 2 Expected number of iterations ℓ to have Rep output a value with .95 probability across error rate. Three lines represent original error rate t and two reduced error rates of $t/2$ and $t/10$ that may be achievable by using confidence information. Note that the y-axis is in log scale.

We observe, for practical parameters, multiplicative changes in τ lead to exponential changes in the required iterations ℓ. To demonstrate we consider the following parameters: a source of length $n = 1024$ (common for the iris), a subset size of $k = 128$, and an output key of a single group element ($\alpha = 1$). Figure 2 shows how $\log \ell$ increases for different τ. Three lines represent the original error rate and two potential reduced error rates (multiplicative decreases of 2 and 10 respectively). Figure 2 considers τ steps of .001. Between 0 and .06, each step of .001 increases $\log \ell$ by .667 (r^2 value of .999).

As mentioned in the introduction, Canetti et al. [8] digital locker[6] condition for security is that *average* subsets have entropy. A distribution satisfying MIPURS implies that average subsets have entropy (see Section 4.3). Since code offset in the exponent allows the adversary to test any subset, average subsets having entropy does not suffice (see Section 1.4). Section 3.1 showed how to handle distributions where *all* subsets have entropy by multiplying by a random error vector. Unfortunately, as we show in Section 4.4, MIPURS and all subsets have entropy are incomparable notions creating a barrier to removing this random vector in Construction 2.

Reusability

Reusability is the ability to support multiple independent enrollments of the same value, allowing users to reuse the same biometric or PUF, for example, with multiple noncooperating providers. More precisely, the algorithm Gen may be run multiple times on correlated readings $\mathbf{e}^1, ..., \mathbf{e}^\rho$ of a given source. Each time, Gen will produce a different pair of values

[6] Intuitively, a digital locker is a symmetric encryption that is semantically secure even when instantiated with keys that are correlated and only have entropy [10].

$(\mathsf{key}^1, \mathsf{pub}^1), ..., (\mathsf{key}^\rho, \mathsf{pub}^\rho)$. Security for each extracted string key^i should hold even in the presence of all the helper strings $(\mathsf{pub}^1, \ldots, \mathsf{pub}^\rho)$. The reproduction procedure Rep at the ith provider still obtains only a single $\mathbf{e}'$ close to $\mathbf{e}^i$ and uses a single helper string pub_i. Because providers may not trust each other key_i should be secure even when all key_j for $j \neq i$ are also given to the adversary. In the full version of this work [12] we show that Construction 1 is reusable if a random generator is used with each enrollment.

4 Characterizing MIPURS

Definition 3 of MIPURS is admittedly unwieldy. It considers a property of a distribution $\mathbf{e} \in \mathbb{F}_q^n$ with respect to a random matrix. We turn to characterizing distributions that satisfy MIPURS. We begin with easier distributions and conclude with the general entropy case is in Section 4.5. Throughout, we consider a prime order group $\mathbb{G}$ of prime size q, a random linear code $\mathbf{A} \in \mathbb{F}_q^{n \times k}$ and the null space $\mathbf{B} \stackrel{\text{def}}{=} \mathtt{null}(\mathbf{A})$.

4.1 Independent Sources $\subset$ MIPURS

In most versions of LWE, each error coordinate is independently distributed and contributes entropy. Examples include the discretized Gaussian introduced by Regev [28, 29], and a uniform interval introduced by Döttling and Müller-Quade [14]. We show that these distributions fit within our MIPURS characterization.

▶ **Lemma 8.** *Let $\mathbf{e} = \mathbf{e}_1, \ldots, \mathbf{e}_n \in \mathbb{F}_q^n$ be a distribution where each $\mathbf{e}_i$ is independently sampled. Let $\alpha = \min_{1 \leq i \leq n} \mathrm{H}_\infty(\mathbf{e}_i)$. For any $k \leq n$, $\mathbf{e}$ is a $(k, \beta) -$ MIPURS distribution for $\beta = 2^{-\alpha}$.*

Proof of Lemma 8. Consider a fixed element $b \neq 0$ in $\mathbf{B}$. Since the components of $\mathbf{e}$ are independent, predicting $\langle b, \mathbf{e} \rangle$ is at least as hard as predicting $\mathbf{e}_i$ for each i such that $\mathbf{b}_i \neq 0$. This can be seen by fixing b and $\mathbf{e}_j$ for $j \neq i$ and noting that the value of $\mathbf{e}_i$ then uniquely determines $\langle b, \mathbf{e} \rangle$. Since $b \neq 0$ there exists at least one such i. Thus,

$$\Pr_{\mathbf{B}}\left[\max_{g} \max_{b \in \mathbf{B} \setminus 0} \Pr_{\mathbf{e}}[\langle b, \mathbf{e} \rangle = g]\right] \leq 2^{-\alpha} \stackrel{\text{def}}{=} \beta. \qquad \blacktriangleleft$$

4.2 Location Sources $\subset$ MIPURS

Next, we consider $\mathbf{e}'$ given by the coordinatewise product of a uniform vector $\mathbf{r} \in \mathbb{F}_q^n$ and a "selection vector" $\mathbf{e} \in \{0,1\}^n$: that is, $\mathbf{e}'_i = \mathbf{r}_i \cdot_c \mathbf{e}_i$ where all large enough subsets of $\mathbf{e}$ are unpredictable ($\cdot_c$ is component-wise multiplication). Location sources are important for applications (see Section 3).

▶ **Lemma 9.** *Let $\gamma \in \mathbb{N}$ and $k \in \mathbb{Z}^+$. Let $\mathbf{e} \in \{0,1\}^n$ be a distribution where all $(k-\gamma)$-subsets have entropy α. Define the distribution $\mathbf{e}'$ as the coordinatewise product of a uniform vector $\mathbf{r} \in \mathbb{F}_q^n$ and $\mathbf{e}$: that is, $\mathbf{e}'_i = \mathbf{e}_i \cdot_c \mathbf{r}_i$. Then the distribution $\mathbf{e}'$ is a (k, β)-*MIPURS *distribution for $\beta = 2^{-\alpha} + ((k-\gamma)\binom{n}{k-\gamma-1})/q^{\gamma+1}$.*

Proof of Lemma 9. We use $\mathbf{A} \in \mathbb{F}_q^{n \times k}$ to represent the random matrix from the definition of a MIPURS distribution and let $\mathbf{B} \in \mathbb{F}_q^{n \times n-k}$ represent its null space. We start by bounding the "minimum distance" of $\mathbf{B}$, that is, the minimum weight of a non-zero element of $\mathbf{B} = \mathtt{null}(\mathbf{A})$. Observe that the number of vectors in $\mathbb{F}_q^n$ of weight less than $k - \gamma$ is

$$\sum_{j=0}^{k-\gamma-1} \binom{n}{j} q^j \leq (k-\gamma)\binom{n}{k-\gamma-1} q^{k-\gamma-1}.$$

The probability that any fixed, nonzero vector lies x in $\mathbf{B}$ is q^{-k}, as it must annihilate k independent, uniform linear equations. That is, $\sum_i x_i \mathbf{A}_{is} = 0$ for each $1 \leq s \leq k$. Thus

$$\mathbb{E}[|\{w \in \mathtt{null}(\mathbf{A}) \setminus 0 \mid \mathsf{wt}(w) < k - \gamma\}|] \leq (k-\gamma)\binom{n}{k-\gamma-1} q^{-\gamma-1}. \tag{1}$$

By Markov's inequality, the probability that there is at least one such small weight vector in $\mathtt{null}(\mathbf{A})$ is no more than the expected number of such vectors. Hence

$$\Pr[\exists w \in \mathtt{null}(\mathbf{A}) \setminus 0, \mathsf{wt}(w) < k - \gamma] \leq (k-\gamma)\binom{n}{k-\gamma-1} q^{-\gamma-1}.$$

For some $\mathbf{b}$ in the span of $\mathbf{B}$ with weight at least $k - \gamma$, consider the product $\langle \mathbf{b}, \mathbf{e}' \rangle = \sum_{i=1}^{n} \mathbf{b}_i \cdot \mathbf{e}_i \cdot \mathbf{r}_i$. Define $\mathcal{I}$ as the set of nonzero coordinates in $\mathbf{b}$. With probability at least $1 - 2^{-\alpha}$ there is some nonzero coordinate in $\mathbf{e}_{\mathcal{I}}$. Conditioned on this fact this means that at least one value $\mathbf{r}_i$ is included in the inner product. Thus, the entropy of the inner product is bounded below by the entropy of $\mathbf{e}_i \cdot \mathbf{r}_i$ which since $\mathbf{e}_i \neq 0$ is bounded by the entropy of $\mathbf{r}_i$. In this case, the prediction probability of an inner product (and therefore a single element of) is $1/q$. The argument concludes by assuming perfect predictability when there exists $\mathbf{b}$ in $\mathbf{B}$ with weight of at most $k - \gamma - 1$. ◀

4.3 MIPURS ⊂ Average Subasets Have Entropy

As mentioned in the Introduction, Canetti et al. [8] showed a fuzzy extractor construction for all sources where an average subset has entropy:[7]

▶ **Definition 10** ([8] average subsets have entropy)**.** *Let the source* $\mathbf{e} = \mathbf{e}_1, \ldots, \mathbf{e}_n$ *consist of strings of length* n *over some arbitrary alphabet* $\mathcal{Z}$*. We say that the source* $\mathbf{e}$ *is a source with a* ***k-average subsets have entropy*** β *if*

$$\mathop{\mathbb{E}}_{j_1,\ldots,j_k \xleftarrow{\$} [1,\ldots,n], j_\alpha \neq j_\gamma} \left(\max_z \{\Pr[(\mathbf{e}_{j_1}, \ldots, \mathbf{e}_{j_k}) = z \mid j_1, \ldots, j_k]\} \right) \leq \beta.$$

We now show that a MIPURS distribution also has that average subsets have entropy.

▶ **Lemma 11.** *Let* $\mathbf{e} = \mathbf{e}_1, \ldots, \mathbf{e}_n$ *be a source over alphabet* $\mathcal{Z}$ *such that* $\mathbf{e}$ *is* $(k, \beta) -$ MIPURS*. Then* $\mathbf{e}$ *has* (k', β')*-entropy samples for any* k' *and*

$$\beta' = \frac{\beta}{\left(1 - \frac{\left(q^{k'-(k+1)}\right)}{\left(2^{k'}\binom{n}{k'}\right)}\right)}.$$

Proof of Lemma 11. We proceed by contradiction, that is suppose that $\mathbf{e}$ does not have k', β' entropy samples. That is,

$$\mathop{\mathbb{E}}_{j_1,\ldots,j'_k \xleftarrow{\$} [1,\ldots,n], j_\alpha \neq j_\gamma} \left(\max_z \{\Pr[(\mathbf{e}_{j_1}, \ldots, \mathbf{e}_{j_k}) = z \mid j_1, \ldots, j'_k]\} \right) > \beta'.$$

We consider the following definition of $\mathbf{b}, \mathbf{g}$ in the MIPURS game:

1. Receive input $\mathbf{A}$, compute $\mathbf{B} = \mathtt{null}(\mathbf{A})$.

[7] We make a small modification to their definition changing to sampling without replacement.

2. Select random $\mathbf{b} \in \mathbf{B}$ such that $\mathsf{wt}(\mathbf{b}) \leq k'$, $\mathbf{b} \neq \mathbf{0}$. If no such $\mathbf{b}$ exists output $\mathbf{b} = \mathbf{0}, \mathbf{g} = 0$.
3. Define $\mathcal{I}$ as the set of nonzero locations in $\mathbf{b}$. If $|\mathcal{I}| < k'$ insert random distinct locations until $|\mathcal{I}| = k'$.
4. Compute $z = \arg\max_z \{\Pr[\mathbf{e}_{\mathcal{I}} = z \mid \mathcal{I}]\}$.
5. Output $\mathbf{g} = \langle \mathbf{b}, z \rangle$.

If z is the correct prediction for $\mathbf{e}_{\mathcal{I}}$ then $\mathbf{g} = \langle \mathbf{b}, z \rangle = \langle \mathbf{b}, \mathbf{e} \rangle$. As noted above, the probability of any particular value nonzero $\mathbf{b}$ being in $\mathbf{B}$ is q^{-k}. Thus, conditioned on finding a good $\mathbf{b}$, the distribution of the random variable $\mathbf{b}$ is exactly that of a uniform weight k' value. This implies that $\mathbb{E}_{\mathbf{b}}\left[\max_z \{\Pr[(\mathbf{e}_{\mathcal{I}}) = z \mid \mathcal{I}\}\right] > \beta'$. It remains to analyze the probability that $\mathbf{B}$ contains no vectors of weight k'. Here we derive an elementary bound, asymptotic formulations exist in the information theory literature [21, Theorem 1.1].

▶ **Lemma 12.** *Let V denote a random subspace of $\mathbb{F}_q^n$ of dimension κ. Let W_ℓ denote the subset of $\mathbb{F}_q^n$ consisting of all vectors with weight ℓ, then*

$$\Pr[V \cap W_\ell = 0] \leq \frac{(q^n - 1)}{\binom{n}{\ell}(q-1)^{\ell-1}(q^\kappa - 1))}.$$

Proof of Lemma 12. We begin by noting that $|W_\ell| = \binom{n}{\ell}(q-1)^\ell$. For a vector $\vec{v} \in W_\ell$, define $X_{\vec{v}} = 1$ if $\vec{v} \in V$ and 0 otherwise. Then

$$\mathbb{E}\left[\sum_{\vec{v} \in W_\ell} X_{\vec{v}}\right] = \binom{n}{\ell}(q-1)^\ell \frac{q^\kappa - 1}{q^n - 1}.$$

We wish to compute the second moment of the sum $\sum X_{\vec{v}}$. We have

$$\begin{aligned}
\mathbb{E}\left[\sum_{\vec{v},\vec{w} \in W_\ell} X_{\vec{v}} X_{\vec{w}}\right] &= \mathbb{E}\left[\sum_{\substack{\vec{v},\vec{w} \in W_\ell \\ \vec{v},\vec{w} \text{ independent}}} X_{\vec{v}} X_{\vec{w}}\right] + \mathbb{E}\left[\sum_{\substack{\vec{v},\vec{w} \in W_\ell \\ \vec{v},\vec{w} \text{ dependent}}} X_{\vec{v}} X_{\vec{w}}\right] \\
&\leq \binom{n}{\ell}(q-1)^\ell \left(\binom{n}{\ell}(q-1)^\ell - (q-1)\right) \max_{\text{indep. } \vec{v},\vec{w}} \Pr[\vec{v}, \vec{w} \in V] \\
&\quad + \binom{n}{\ell}(q-1)^{\ell+1} \max_{\text{dependent } \vec{v},\vec{w}} \Pr[\vec{v}, \vec{w} \in V] \\
&\leq \underbrace{\left(\binom{n}{\ell}(q-1)^\ell\right)^2 \frac{(q^\kappa - 1)(q^{\kappa-1} - 1)}{(q^n - 1)(q^{n-1} - 1)}}_{(\ddagger)} + \binom{n}{\ell}(q-1)^{\ell+1} \frac{q^\kappa - 1}{q^n - 1}.
\end{aligned}$$

Note that $(m-t)/(n-t) < m/n$ assuming that $t \leq m < n$ and hence that that

$$(\ddagger) = \left(\binom{n}{\ell}(q-1)^\ell\right)^2 \frac{(q^\kappa - 1)(q^\kappa - q)}{(q^n - 1)(q^n - q)} \leq \left(\binom{n}{\ell}(q-1)^\ell\right)^2 \frac{(q^\kappa - 1)^2}{(q^n - 1)^2} \leq \mathbb{E}\left[\sum X_{\vec{v}}\right]^2.$$

It follows that

$$\mathsf{Var}\left[\sum X_{\vec{v}}\right] = \mathbb{E}\left[\left(\sum X_{\vec{v}}\right)^2\right] - \mathbb{E}\left[\sum X_{\vec{v}}\right]^2 \leq \binom{n}{\ell}(q-1)^{\ell+1} \frac{q^\kappa - 1}{q^n - 1}.$$

Then using Chebyshev's inequality with a constant of

$$\alpha = \sqrt{\left(\binom{n}{\ell}(q-1)^\ell \frac{q^\kappa - 1}{q^n - 1}\right)^2 \Big/ \left(\binom{n}{\ell}(q-1)^{\ell+1} \frac{q^\kappa - 1}{q^n - 1}\right)}$$

one finds:

$$\Pr\left[\sum X_{\vec{v}} = 0\right] \leq \frac{\mathsf{Var}[\sum X_{\vec{v}}]}{\mathbb{E}\left[\sum X_{\vec{v}}\right]^2} \leq \frac{\binom{n}{\ell}(q-1)^{\ell+1}\frac{q^\kappa-1}{q^n-1}}{\left(\binom{n}{\ell}(q-1)^{\ell}\frac{q^\kappa-1}{q^n-1}\right)^2} \leq \frac{(q^n-1)}{\left(\binom{n}{\ell}(q-1)^{\ell-1}(q^\kappa-1)\right)}.$$

This completes the proof of Lemma 12. ◀

Thus, for $\dim(\mathbf{B}) \geq n-k$ it is true that for any k':

$$\Pr[\mathbf{B} \cap W_{k'} = 0] \leq \frac{(q^n-1)}{\binom{n}{k'}(q-1)^{k'-1}(q^{n-k}-1)} \leq \frac{q^n}{2^{k'}\binom{n}{k'}q^{n-k+k'-1}} = \frac{q^{k'-(k+1)}}{2^{k'}\binom{n}{k'}}.$$

We note that the overall success of prediction of $\mathbf{b}, \mathbf{g}$ in the MIPURS game is bounded below by $\Pr[\mathbf{B} \cap W_{k'} = 0] * 0 + (1 - \Pr[\mathbf{B} \cap W_{k'} = 0]) * \beta' = \beta$. This completes the proof of Lemma 11. ◀

4.4 MIPURS and all subsets have entropy

We now consider the relationship between MIPURS and all subsets have entropy. Recall, that we showed that for a distribution $\mathbf{e}$ where all subsets have entropy multiplying by a random vector produced a MIPURS distribution. With two simple examples, we show that MIPURS is not contained by all subsets have entropy and all subsets have entropy is not contained by MIPURS.

▶ **Proposition 13** (MIPURS $\not\to$ all subsets have entropy)**.** *Define $\mathbf{e} \in \mathbb{F}_q^n$ as the distribution that is fixed in the first k positions and uniform in all other positions. Clearly for any $\beta > 0$ it does not hold that all k-subsets have entropy. Furthermore, $\mathbf{e}$ is $(k, \beta) -$ MIPURS for $\beta \geq (1 - \frac{1}{q^k} - \frac{k}{q^{n-k}}) \log q$.*

To show the above proposition, assume perfect predictability in the MIPURS game in the case when $\mathbf{A}$ is not full rank or when $1^k||0^{n-k}$ is in $\mathtt{null}(\mathbf{A})$. Otherwise, full entropy results from the same argument as Lemma 8.

For the second direction we assume that $\mathbf{e}$ is a Reed-Solomon [27] code (the counterexample is similar to the one presented in Section 1.3). For the field $\mathbb{F}_q$ of size q, a message length k, and code length n, such that $k \leq n \leq q$, define the Vandermonde matrix $\mathbf{V}$ where the ith row, $\mathbf{V}_i = [i^0, i^1, ..., i^k]$. The Reed Solomon Code $\mathbb{RS}(n, k, q)$ is the set of all points $\mathbf{Vx}$ where $\mathbf{x} \in \mathbb{F}_q^k$.

▶ **Proposition 14** (all subsets have entropy $\not\to$ MIPURS)**.** *Let $k < n/2$ and let $\mathbf{e}$ be the uniform distribution over $\mathbb{RS}(n, n-k-1, q)$ then all k subsets of $\mathbf{e}$ have entropy $k \log q$. Furthermore, $\mathbf{e}$ is **not** $(k, \beta) -$ MIPURS for any $\beta < 1$.*

Note $\dim(\mathtt{null}(\mathbf{A})) \geq n-k$ and thus $\mathtt{null}(\mathbf{A})$ and $\mathtt{null}(\mathbb{RS}(n, n-k-1, q)|$ are guaranteed to a have a nontrivial intersection. The result follows by picking some $\mathbf{b}$ in this intersection and setting $g = 0$.

4.5 High entropy $\subset$ MIPURS

We now turn to the general entropy condition: MIPURS is hard for all distribution where the min-entropy exceeds $\log q^{n-k}$ (by a super logarithmic amount). For conciseness, we introduce $\kappa \stackrel{\text{def}}{=} n-k$.

The adversary is given a generating matrix of the code, $\mathbf{A}$; this determines $\mathbf{B} = \mathtt{null}(\mathbf{A})$. Our proof is divided into three parts. Denote by E a set of possible error vectors.

1. Theorem 16: We show that the number of vectors $\psi \in E$ that are likely to have 0 inner product with an adversarially chosen vector in $\mathbf{B}$ is small. Intuitively, we show that this set is "not much larger than a κ-dimensional subspace."
2. Theorem 22: We then show it is difficult to predict the value of the inner product: even if the adversary may select arbitrarily coupled $\mathbf{b}$ and $\mathbf{g}$, it is difficult to achieve $\langle \mathbf{b}, \psi \rangle = \mathbf{g}$.
3. Lemma 24: We show that any distribution $\mathbf{e}$ with sufficient entropy cannot lie in the set of *predictable* error vectors E with high probability.

We codify the set of possible adversarial strategies by introducing a notion of *κ-induced random variables*. For the moment, we assume that $\mathbf{B}$ is a uniformly selected subspace of dimension exactly κ; at the end of the proof we remove this restriction to apply these results when $\mathbf{B}$ has the distribution given by $\texttt{null}(\mathbf{A})$ (Corollary 25).

▶ **Definition 15.** *Let $\mathbf{b}$ be a random variable taking values in $\mathbb{F}_q^n$. We say that $\mathbf{b}$ is* κ-induced *if there exists a (typically dependent) random variable $\mathbf{B}$, uniform on the collection of κ-dimensional subspaces of $\mathbb{F}_q^n$, so that $\mathbf{b} \in \mathbf{B}$ and $\mathbf{b} \neq \vec{0}$ with certainty: $\Pr\left[\mathbf{b} \in \mathbf{B} \wedge \mathbf{b} \neq \vec{0}\right] = 1$. Note, in fact, that the random variables $\mathbf{B}$ and $\mathbf{b}$ are necessarily dependent unless $n = \kappa$.*

It suffices to consider the maximum probability in Definition 3 with respect to κ-induced random variables. This is because for any $\mathbf{b}$ that is not κ-induced we can find another $\mathbf{b}$ that is κ induced that does no worse in the game in Definition 3. For example when $\mathbf{b}$ is not in $\mathbf{B}$ or is the zero vector, one can replace $\mathbf{b}$ with a random element in the span of $\mathbf{B}$.

We now show that if the set E is large enough there is no strategy for $\mathbf{b}$ that guarantees $\langle \mathbf{b}, \psi \rangle = 0$ with significant probability. The next theorem (Thm. 22) will, more generally, consider prediction of the inner product itself. For a κ induced random variable $\mathbf{b}$, define

$$E_\epsilon^{(\mathbf{b},0)} = \left\{ f \in_p^n \;\middle|\; \Pr_{\mathbf{b}}[\langle \mathbf{b}, f \rangle = 0] \geq \epsilon \right\}.$$

When $\mathbf{b}$ can be inferred from context, we simply refer to this set as E_ϵ. Then define $P_{\kappa,\epsilon} = \max_{\mathbf{b}} |E_\epsilon^{(\mathbf{b},0)}|$ where the maximum is over all κ-induced random variables in $\mathbb{F}_q^n$.

▶ **Theorem 16.** *Let q be a prime and let $d > 1$, $\kappa, m, \eta \in \mathbb{Z}^+$ be parameters for which $\kappa \leq n$. Then assuming $P_{\kappa,\epsilon} > d \cdot q^\kappa$ we must have*

$$\epsilon \leq \binom{\kappa+\eta}{m} + \binom{m}{\kappa}\left(\binom{m}{\eta}\left(\frac{1}{d}\right)^\eta + \left(\frac{2}{q}\right)\right).$$

Before proving Theorem 16, we introduce and prove two combinatorial lemmas (18 and 20). We then proceed with the proof of Theorem 16. The major challenge is that the set E_ϵ (for a particular $\mathbf{b}$) is typically not a linear subspace; these results show that is has reasonable "approximate linear" structure. We begin with the notion of *linear density* to measure, intuitively, how close the set is to linear.

▶ **Definition 17.** *The* ℓ-linear density *of a sequence of vectors $F = (f^1, \ldots, f^m)$, with each $f^i \in \mathbb{F}_q^n$, is the maximum number of entries that are covered by a subspace of dimension ℓ. Formally,*

$$\Delta^\ell(F) = \max_{V, \dim(V)=\ell} |\{i \mid f^i \in V\}|.$$

▶ **Lemma 18.** *Let q be a prime and let $n, \ell \in \mathbb{Z}^+$ satisfy $\ell \leq n$. Let $E \subset \mathbb{F}_q^n$ satisfy $|E| \geq q^\ell$ and let $\mathbf{F} = (\mathbf{f}^1, \ldots, \mathbf{f}^m)$ be a sequence of uniformly and independently chosen elements of E. Define d so that $|E| = dq^\ell$; then for any $\eta \geq 0$,*

$$\Pr_{\mathbf{F}}[\Delta^\ell(\mathbf{F}) \geq \ell + \eta] \leq \binom{m}{\ell}\binom{m-\ell}{\eta}\left(\frac{1}{d}\right)^\eta.$$

Proof of Lemma 18. By the definition of linear density, if $\Delta^\ell(\mathbf{F}) \geq \ell + \eta$ there must be at least one subset of $\ell + \eta$ indices $I \subset [m]$ so that $\{\mathbf{f}^i \mid i \in I\}$ is contained in a subspace of dimension ℓ. In order for a subset I to have this property, there must be a partition of I into a disjoint union $S \cup L$, where S has cardinality ℓ and T indexes the remaining η "lucky" vectors that lie in the span of the vectors given by S. Formally, $\forall t \in T, \mathbf{f}^t \in \text{span}(\{\mathbf{f}^s \mid s \in S\})$.

Fix, for the moment, ℓ indices of $\mathbf{F}$ to identify a candidate subset of vectors to play the role of S and η indices of $\mathbf{F}$ to identify a candidate set T. The probability that each of the η vectors indexed by T lie in the space spanned by S is clearly no more than $(q^\ell/|E|)^\eta \leq (1/d)^\eta$. Taking the union bound over these choices of indices completes the argument: The probability of a sequence is no more than $\binom{m}{\ell}\binom{m-\ell}{\eta}d^{-\eta}$, as desired. ◀

Before introducing our second combinatorial lemma (Lem 20), we need a Lemma bounding the probability of a fixed subspace having a nontrivial intersection with a random subspace.

▶ **Lemma 19.** *Let q be a prime and $\kappa, n \in \mathbb{N}$ with $\kappa \leq n$. Let $\mathbf{V}$ be a random variable uniform on the set of all κ-dimensional subspaces of $\mathbb{F}_q^n$. Let W be a fixed subspace of dimension ℓ. Then*

$$\Pr[\mathbf{V} \cap W \neq \{0\}] \leq q^{\kappa+\ell-(n+1)} \cdot \left(\frac{q}{q-1}\right).$$

Proof of Lemma 19. Let $\mathcal{L}$ denote the set of all 1-dimensional subspaces in W. Each 1-dimensional subspace is described by an equivalence class of $q-1$ vectors under the relation $x \sim y \Leftrightarrow \exists \lambda \in \mathbb{F}_q^*, \lambda x = y$. Thus $|\mathcal{L}| = (q^\ell - 1)/(q-1) \leq q^{\ell-1}(q/(q-1))$. Then

$$\Pr[\mathbf{V} \cap W \neq \{\vec{0}\}] \leq \sum_{L \in \mathcal{L}} \Pr[L \subset \mathbf{V}] \leq |\mathcal{L}| \max_{v \in \mathbb{F}_q^n \setminus \{\vec{0}\}} \Pr[v \in \mathbf{V}] \leq q^{\kappa+\ell-(n+1)} \left(\frac{q}{q-1}\right),$$

where we recall the fact that for any particular fixed nonzero vector v, $\Pr[v \in \mathbf{V}] = \frac{q^\kappa - 1}{q^n - 1} \leq q^{\kappa-n}$. ◀

▶ **Lemma 20.** *Let q be a prime, let $\ell, \kappa, n \in \mathbb{Z}^+$ satisfy $\ell, \kappa \leq n$. Let $F = (f^1, \ldots, f^m)$ be a sequence of elements of $\mathbb{F}_q^n$ with $\dim(\text{span}(F)) \geq \ell$. Then, for any κ-induced random variable $\mathbf{b}$ taking values in $\mathbb{F}_q^n$,*

$$\Pr_{\mathbf{b}}\left[|\{i \mid \langle \mathbf{b}, f^i \rangle = 0\}| \geq \Delta^\ell(F)\right] \leq \binom{m}{\ell} q^{\kappa-\ell-1} \left(\frac{q}{q-1}\right) \leq 2\binom{m}{\ell} q^{\kappa-\ell-1}.$$

Proof of Lemma 20. Let $\mathcal{V}_F$ denote the collection of all ℓ-dimensional subspaces of $\mathbb{F}_q^n$ spanned by subsets of elements in the sequence F. That is,

$$\mathcal{V}_F = \{V \mid V = \text{span}(\{f^i \mid i \in I\}), I \subset [m], \dim(V) = \ell\}.$$

Then $|\mathcal{V}_F| \leq \binom{m}{\ell}$, as each such subspace is spanned by at least one subset of F of size ℓ. As $\dim(\text{span}(F)) \geq \ell$, the set $\mathcal{V}_F$ is nonempty.

Observe that if $I \subset [m]$ has cardinality at least $\Delta^\ell(F)$ then, by definition, $\dim(\text{span}(\{f^i \mid i \in I\})) \geq \ell$; otherwise, an additional element of F could be added to the set indexed by I to yield a set of size exceeding $\Delta^\ell(F)$ which still lies in a subspace of dimension ℓ (contradicting the definition of Δ^ℓ). Note in the case that $m = \ell$ (and there is no element to add) then $\Delta^\ell(F) = \ell = \dim(\text{span}(\{f^i \mid i \in I\}))$. Thus, if $I \subset [m]$ has cardinality at least $\Delta^\ell(F)$, there must be some $V \in \mathcal{V}_F$ for which $V \subset \text{span}(\{f^i \mid i \in I\})$. In particular

$$\begin{aligned}\Pr_{\mathbf{b}}\left[|\{f^i \in F \mid \langle \mathbf{b}, f^i \rangle = 0\}| \geq \Delta^\ell(F)\right] &\leq \Pr_{\mathbf{b}}\left[\exists V \in \mathcal{V}_F, \forall v \in V, \langle v, \mathbf{b} \rangle = 0\right] \\ &\leq \sum_{V \in \mathcal{V}_F} \Pr_{\mathbf{b}}\left[\forall v \in V, \langle v, \mathbf{b} \rangle = 0\right] = \sum_{V \in \mathcal{V}_F} \Pr_{\mathbf{b}}\left[\mathbf{b} \in V^\perp\right],\end{aligned}$$

where we have adopted the notation $V^\perp = \{w \mid \forall v \in V, \langle v, w\rangle = 0\}$. Recall that when V is a subspace of dimension ℓ, $V^\perp$ is a subspace of dimension $n - \ell$. To complete the proof, we recall that $\mathbf{b}$ is κ-induced, so that there is an associated random variable $\mathbf{B}$, uniform on dimension κ subspaces, for which $\mathbf{b} \in \mathbf{B}$ with certainty; applying Lemma 19 we may then conclude

$$\sum_{V \in \mathcal{V}_F} \Pr_{\mathbf{b}}[\mathbf{b} \in V^\perp] \leq \sum_{V \in \mathcal{V}_F} \Pr_{\mathbf{B}}[\mathbf{B} \cap V^\perp \neq \{\vec{0}\}] \leq \binom{m}{\ell} q^{\kappa-\ell-1} \left(\frac{q}{q-1}\right).$$

This completes the proof of Lemma 20. ◀

Proof of Theorem 16. Now we analyze the relationship between our two paramenters of interest: ϵ and d. Fix some $\epsilon > 0$. Let $\mathbf{b}$ be a κ-induced random variable for which $|E_\epsilon^{(\mathbf{b},0)}| = P_{\kappa,\epsilon}$ and let $\mathbf{B}$ be the coupled variable, uniform on subspaces, for which $\mathbf{b} \in \mathbf{B}$.

For the purposes of analysis we consider a sequence of m vectors chosen independently and uniformly from $E_\epsilon = E_\epsilon^{(\mathbf{b},0)}$ with replacement; we let $\mathbf{F} = (\mathbf{f}^1, \ldots, \mathbf{f}^m)$ denote the set of vectors so chosen. We study the expectation of the number of vectors in $\mathbf{F}$ that are orthogonal to $\mathbf{b}$. We first give an immediate lower bound by linearity of expectation and the definition of E_ϵ: $\mathbb{E}_{\mathbf{b},\mathbf{F}}[|\{\mathbf{f}^i \in F \mid \langle \mathbf{b}, \mathbf{f}^i\rangle = 0\}|] \geq \epsilon \cdot m$.

We now infer an upper bound on this expectation using Lemmas 18 and 20. We say that the samples $\mathbf{F}$ from E_ϵ are *compact* if $\Delta^\kappa(\mathbf{F}) \geq \kappa + \eta$. The probability of this *compact* event is no more than $\binom{m}{\kappa}\binom{m-\kappa}{\eta}\left(\frac{1}{d}\right)^\eta$ by Lemma 18. For *compact* selections, we crudely upper bound the expectation by m; for *spread* selections we further split the expectation based on the random variable $\mathbf{B}$. We say that $\mathbf{B}$ is *susceptible* (for a fixed $F = (f^1, ..., f^m)$) if there exists some $b \in \mathbf{B}$ such that $|\{f^i \in F \mid \langle b, f^i\rangle = 0\}| \geq \Delta^\kappa(F)$. Otherwise, $\mathbf{B}$ is *resistant*. The probability of a *susceptible* selection of $\mathbf{B}$ is bounded above by $(2/q)\binom{m}{\kappa}$ in light of Lemma 20 (applied with $\ell = \kappa$). In the pessimistic case (that $\mathbf{B}$ is *susceptible*), we again upper bound the expectation by m. Then if the experiment is neither *compact* nor *susceptible*, we may clearly upper bound the expectation by $\kappa + \eta$. So, for any $\eta > 0$ we conclude that

$$\mathop{\mathbb{E}}_{\mathbf{b},\mathbf{B},F}[|\{f_i \in F \mid \langle \mathbf{b}, f_i\rangle = 0\}|] \leq (\kappa + \eta) + m\left(\binom{m}{\kappa}\binom{m-\kappa}{\eta}\left(\frac{1}{d}\right)^\eta + \frac{2}{q}\binom{m}{\kappa}\right)$$

and hence that

$$\epsilon \leq \left(\frac{\kappa+\eta}{m}\right) + \binom{m}{\kappa}\left(\binom{m}{\eta}\left(\frac{1}{d}\right)^\eta + \frac{2}{q}\right).$$

This completes the proof of Theorem 16. ◀

▶ **Corollary 21.** *Let κ and n be parameters satisfying $1 \leq \kappa < n$ and let q be a prime such that $q \geq 2^{4\kappa}$. Then for $\epsilon \geq 5eq^{-1/(2(\kappa+1))}$ we have $P_{\kappa,\epsilon} \leq 5eq^\kappa/\epsilon$. In particular, for such ϵ and any κ-induced $\mathbf{b}$, the set $|E_\epsilon^{(\mathbf{b},0)}| \leq 5eq^\kappa/\epsilon$.*

Proof of Corollary 21. Consider parameters for Theorem 16 that satisfy the following:

$$1 < d \leq q^{1/(2(\kappa+1))}, \qquad m = \frac{d\eta}{2e}, \qquad \text{and} \qquad \eta = \log q\,.$$

First note that $\kappa < 4\kappa \leq \log q = \eta$ (as $q \geq 2^{4\kappa}$). Then, consider a set $E_\epsilon^{(\mathbf{b},0)}$ for some $\mathbf{b}$. We have

$$\epsilon \leq \left(\frac{\kappa+\eta}{m}\right) + \binom{m}{\kappa}\left(\left(\frac{me}{\eta d}\right)^\eta + \frac{2}{q}\right) \leq \left(\frac{2\eta}{m}\right) + 3\binom{m}{\kappa}q^{-1} \leq \underbrace{\left(\frac{4e}{d}\right) + 3\binom{d\eta/2e}{\kappa}q^{-1}}_{(\dagger)}.$$

Since $q \geq 2^{4\kappa}$, we may write $q = 2^{2\alpha\kappa}$ for some $\alpha \geq 2$ and it follows that $\left(\frac{\log q}{\kappa}\right)^{\kappa} = (2\alpha)^{\kappa} \leq (2^{\alpha})^{\kappa} = \sqrt{q}$ because $2\alpha \leq 2^{\alpha}$ for all $\alpha \geq 2$. In light of this, consider the second term in the expression (†) above:

$$3\binom{d\eta/2e}{\kappa} q^{-1} \leq 3\left(\frac{d\eta}{2\kappa}\right)^{\kappa} q^{-1} \leq \frac{3}{2}\left(\frac{d\eta}{\kappa}\right)^{\kappa} q^{-1} \leq \frac{3}{2}\left(\frac{d^{\kappa}}{\sqrt{q}}\right)\left(\left(\frac{\log q}{\kappa}\right)^{\kappa}\frac{1}{\sqrt{q}}\right) \leq \frac{3}{2d} \leq \frac{e}{d}\,.$$

We conclude that for any $1 < d \leq q^{1/(2(\kappa+1))}$, $P_{\kappa,\epsilon} \geq dq^{k} \implies \epsilon \leq 5e/d$. Observe then that for any $\epsilon > 5e/q^{1/(2(\kappa+1))}$ we may apply the argument above to $P_{\kappa,\epsilon}$ with $d = 5e/\epsilon$ and conclude that $P_{\kappa,\epsilon} \leq 5eq^{\kappa}/\epsilon$. ◀

Predicting Arbitrary Values

We now show that the adversary cannot due much better than Theorem 16 even if the task is predicting and arbitrary inner product (not just zero).

▶ **Theorem 22.** *Let* $\mathbf{b}$ *be a* κ*-induced random variable in* $\mathbb{F}_q^n$ *and let* $\mathbf{g}$ *be a random variable over* $\mathbb{F}_q$ *(arbitrarily dependent on* $\mathbf{b}$*). For* $\epsilon > 0$ *we generalize the notation above so that*

$$E_{\epsilon}^{(\mathbf{b},\mathbf{g})} = \left\{ f \in_q^n \;\middle|\; \Pr_{\mathbf{b},\mathbf{g}}[\langle \mathbf{b}, f\rangle = \mathbf{g}] \geq \epsilon \right\}. \qquad \textit{then} \qquad |E_{\epsilon^2/8}^{(\mathbf{b},0)}| \geq \frac{\epsilon^2}{8}|E_{\epsilon}^{(\mathbf{b},\mathbf{g})}|\,.$$

Proof of Theorem 22. For an element $\psi \in E_{\epsilon}^{(\mathbf{b},\mathbf{g})}$, define $F_{\psi} = \{(x, \langle x, \psi\rangle) \mid x \in \mathbb{F}_q^n\}$. Note that $\Pr_{\mathbf{b},\mathbf{g}}[(\mathbf{b},\mathbf{g}) \in F_{\psi}] \geq \epsilon$ by assumption. For any $\delta < \epsilon$, there is a subset $F^* \subset E_{\epsilon}^{(\mathbf{b},\mathbf{g})}$ for which $|F^*| \leq 1/\delta$ and for any $\psi \in E_{\epsilon}^{(\mathbf{b},\mathbf{g})}$, $\Pr_{\mathbf{b},\mathbf{g}}\left[(\mathbf{b},\mathbf{g}) \in \left(F_{\psi} \cap \left(\bigcup_{f' \in F^*} F_{f'}\right)\right)\right] \geq \epsilon - \delta$. To see this, consider incrementally adding elements of $E_{\epsilon}^{(\mathbf{b},\mathbf{g})}$ into F^* so as to greedily increase $\Pr_{\mathbf{b},\mathbf{g}}\left[(\mathbf{b},\mathbf{g}) \in \bigcup_{f' \in F^*} F_{f'}\right]$. If this process is carried out until no $\psi \in E_{\epsilon}^{(\mathbf{b},\mathbf{g})}$ increases the total probability by more than δ, then it follows that every F_{ψ} intersects with the set with probability mass at least $\epsilon - \delta$, as desired. Note also that this termination condition is achieved after including no more than $1/\delta$ sets. It follows that for any $\psi \in E_{\epsilon}^{(\mathbf{b},\mathbf{g})}$,

$$\mathop{\mathbb{E}}_{f' \in F^*} \Pr_{\mathbf{b}}[\langle \mathbf{b}, \psi\rangle = \langle \mathbf{b}, f'\rangle] \geq (\epsilon - \delta)\delta \;\text{ and }\; \mathop{\mathbb{E}}_{f' \in F^*} \mathop{\mathbb{E}}_{\psi \in E_{\epsilon}^{(\mathbf{b},\mathbf{g})}} \Pr_{\mathbf{b}}[\langle \mathbf{b}, \psi\rangle = \langle \mathbf{b}, f'\rangle] \geq (\epsilon - \delta)\delta\,.$$

Then there exists an f^* for which

$$\mathop{\mathbb{E}}_{\psi \in E_{\epsilon}^{(\mathbf{b},\mathbf{g})}} \Pr[\langle \mathbf{b}, \psi\rangle = \langle \mathbf{b}, f^*\rangle] \geq (\epsilon - \delta)\delta\,.$$

Setting $\delta = \epsilon/2$ and we see that

$$\mathop{\mathbb{E}}_{\psi \in E_{\epsilon}^{(\mathbf{b},\mathbf{g})}} \Pr[\langle \mathbf{b}, \psi\rangle = \langle \mathbf{b}, f^*\rangle] = \Pr_{\mathbf{b}, \psi \in E_{\epsilon}^{(\mathbf{b},\mathbf{g})}}[\langle \mathbf{b}, \psi\rangle = \langle \mathbf{b}, f^*\rangle] \geq \frac{\epsilon^2}{4}\,.$$

Using this expectation (of a probability), we bound the probability it is greater than $1/2$ its mean. As the inner product is bi-linear,

$$\Pr_{\psi \in E_{\epsilon}^{(\mathbf{b},\mathbf{g})}}\left[\Pr_{\mathbf{b}}[\langle \mathbf{b}, \psi - f^*\rangle = 0] \geq \frac{\epsilon^2}{8}\right] \geq \frac{\epsilon^2}{8}\,.$$

Thus, an $\epsilon^2/8$ fraction of the set $\{\psi - f^* \mid \psi \in E_{\epsilon}^{(\mathbf{b},\mathbf{g})}\}$ must be a subset of $E_{\epsilon^2/8}^{(\mathbf{b},0)}$: The claim of the theorem follows, that $|E_{\epsilon^2/8}^{(\mathbf{b},0)}| \geq (\epsilon^2/8)|E_{\epsilon}^{(\mathbf{b},\mathbf{g})}|$. ◀

With the language and settings of this last Theorem, applying Corollary 21 to appropriately control $|E^{(\mathbf{b},0)}_{\epsilon^2/8}|$ yields the following bound on $|E^{(\mathbf{b},\mathbf{g})}_{\epsilon}|$.

▶ **Corollary 23.** *Let κ and n be parameters satisfying $1 \leq \kappa < n$ and let q be a prime such that $q \geq 2^{4\kappa}$. Let $\mathbf{b}$ be any κ-induced random variable in $\mathbb{F}_q^n$ and $\mathbf{g}$ any random variable in $\mathbb{F}_q$. Then for any $\epsilon \geq 11q^{-1/(4(\kappa+1))}$ it holds that $|E^{(\mathbf{b},\mathbf{g})}_{\epsilon}| \leq (320eq^{\kappa})/(\epsilon^4)$.*

This implies all high min-entropy distributions are not predictable in the above game.

▶ **Lemma 24.** *Let $\mathbf{b}$ be a κ-induced random variable in $\mathbb{F}_q^n$. Let $\mathbf{g}$ be an arbitrary random variable in $\mathbb{F}_q$. Let $\mathbf{e}$ be a random variable with $\mathrm{H}_\infty(\mathbf{e}) = s$. Let $E^{(\mathbf{b},\mathbf{g})}_{\epsilon}$ be as defined in Theorem 22. Then for $\epsilon > 0$, $\Pr_{\psi \leftarrow \mathbf{e},\mathbf{b},\mathbf{g}}[\langle \mathbf{b}, \psi \rangle = \mathbf{g}] \leq 2^{-s}|E^{(\mathbf{b},\mathbf{g})}_{\epsilon}| + \epsilon$.*

Proof of Lemma 24. Our predictable set $E_\epsilon = E^{(\mathbf{b},\mathbf{g})}_{\epsilon}$ gives us no guarantee on the instability of the inner product. If $\psi \in E_\epsilon$ then we upper bound the probability by 1. Because $\mathbf{e}$ has min-entropy s, we know that no element is selected with probability greater than 2^{-s}, thus the probability of a lying inside a set of size $|E_\epsilon|$ is at most $|E_\epsilon|/2^s$. Outside of our predictable set, we know that the probability of a stable inner product cannot be greater than ϵ by definition of E_ϵ. Therefore if ψ does not fall in the predictable set, we bound the probability by ϵ (for simplicity, we ignore the multiplicative term less than 1). ◀

▶ **Corollary 25.** *Let k and n be parameters with $n > k$ and let q be a prime such that $q \geq 2^{4(n-k)}$. Let $\epsilon \geq 11q^{-1/(4(n-k+1))}$ be a parameter. Then for all distributions $\mathbf{e} \in \mathbb{F}_q^n$ such that $\mathrm{H}_\infty(\mathbf{e}) \geq \log\left(320eq^{n-k}\epsilon^{-5}\right)$, it holds that (for any $\mathbf{b}$ and $\mathbf{g}$ above)*

$$\Pr_{\mathbf{b},\mathbf{g},\mathbf{e}}[\langle \mathbf{b}, \mathbf{e} \rangle = \mathbf{g}] \leq 2\epsilon + k/q^{n-k}$$

and thus $\mathbf{e}$ is $(k, 2\epsilon + k/q^{n-k}) - \mathsf{MIPURS}$.

The additional k/q^{n-k} term is due to the probability that $\mathbf{A}$ may not be full rank, all of the above analysis was conditioned on $\mathbf{A}$ being full rank. The corollary then follows by replacing $\kappa = n - k$.

References

1. Adi Akavia, Shafi Goldwasser, and Vinod Vaikuntanathan. Simultaneous hardcore bits and cryptography against memory attacks. In *Theory of cryptography conference*, pages 474–495. Springer, 2009.
2. Daniel Apon, Chongwon Cho, Karim Eldefrawy, and Jonathan Katz. Efficient, reusable fuzzy extractors from LWE. In *International Conference on Cyber Security Cryptography and Machine Learning*, pages 1–18. Springer, 2017.
3. Ward Beullens and Hoeteck Wee. Obfuscating simple functionalities from knowledge assumptions. In *IACR International Workshop on Public Key Cryptography*, pages 254–283. Springer, 2019.
4. Allison Bishop, Lucas Kowalczyk, Tal Malkin, Valerio Pastro, Mariana Raykova, and Kevin Shi. A simple obfuscation scheme for pattern-matching with wildcards. In *Annual International Cryptology Conference*, pages 731–752. Springer, 2018.
5. Nir Bitansky, Ran Canetti, Yael Tauman Kalai, and Omer Paneth. On virtual grey box obfuscation for general circuits. *Algorithmica*, 79(4):1014–1051, 2017.
6. Xavier Boyen. Reusable cryptographic fuzzy extractors. In *Proceedings of the 11th ACM conference on Computer and Communications Security*, pages 82–91, 2004.

7 Ran Canetti and Ronny Ramzi Dakdouk. Obfuscating point functions with multibit output. In *Advances in Cryptology–EUROCRYPT 2008*, pages 489–508. Springer, 2008.

8 Ran Canetti, Benjamin Fuller, Omer Paneth, Leonid Reyzin, and Adam Smith. Reusable fuzzy extractors for low-entropy distributions. In *Advances in Cryptology – EUROCRYPT*, pages 117–146. Springer, 2016.

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

10 Ran Canetti, Yael Tauman Kalai, Mayank Varia, and Daniel Wichs. On symmetric encryption and point obfuscation. In *Theory of Cryptography Conference*, pages 52–71. Springer, 2010.

11 Ran Canetti, Guy N Rothblum, and Mayank Varia. Obfuscation of hyperplane membership. In *Theory of Cryptography Conference*, pages 72–89. Springer, 2010.

12 Luke Demarest, Benjamin Fuller, and Alexander Russell. Code offset in the exponent. *Cryptology ePrint archive*, 2018. URL: `https://eprint.iacr.org/2018/1005`.

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

14 Nico Döttling and Jörn Müller-Quade. Lossy codes and a new variant of the learning-with-errors problem. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 18–34. Springer, 2013.

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

16 Benjamin Fuller, Xianrui Meng, and Leonid Reyzin. Computational fuzzy extractors. *Information and Computation*, page 104602, 2020.

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

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

19 Benjamin Fuller, Leonid Reyzin, and Adam Smith. When are fuzzy extractors possible? *IEEE Transactions on Information Theory*, 2020.

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

21 Jing Hao, Han Huang, Galyna Livshyts, and Konstantin Tikhomirov. Distribution of the minimum distance of random linear codes. In *2020 IEEE International Symposium on Information Theory (ISIT)*, pages 114–119. IEEE, 2020.

22 Charles Herder, Ling Ren, Marten van Dijk, Meng-Day Yu, and Srinivas Devadas. Trapdoor computational fuzzy extractors and stateless cryptographically-secure physical unclonable functions. *IEEE Transactions on Dependable and Secure Computing*, 2016.

23 Chenglu Jin, Charles Herder, Ling Ren, Phuong Ha Nguyen, Benjamin Fuller, Srinivas Devadas, and Marten Van Dijk. Fpga implementation of a cryptographically-secure puf based on learning parity with noise. *Cryptography*, 1(3):23, 2017.

24 Ari Juels and Martin Wattenberg. A fuzzy commitment scheme. In *Proceedings of the 6th ACM conference on Computer and communications security*, pages 28–36, 1999.

25 Noam Nisan and David Zuckerman. Randomness is linear in space. *Journal of Computer and System Sciences*, 52(1):43–52, 1996.

26 Eugene Prange. The use of information sets in decoding cyclic codes. *IRE Transactions on Information Theory*, 8(5):5–9, 1962.

27 Irving S Reed and Gustave Solomon. Polynomial codes over certain finite fields. *Journal of the society for industrial and applied mathematics*, 8(2):300–304, 1960.

28 Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. *Journal of the ACM (JACM)*, 56(6):1–40, 2009.

29 Oded Regev. The learning with errors problem. *Invited survey in CCC*, 7(30):11, 2010.

30 Victor Shoup. Lower bounds for discrete logarithms and related problems. In *International Conference on the Theory and Applications of Cryptographic Techniques*, pages 256–266. Springer, 1997.

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

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

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

34 Joanne Woodage, Rahul Chatterjee, Yevgeniy Dodis, Ari Juels, and Thomas Ristenpart. A new distribution-sensitive secure sketch and popularity-proportional hashing. In *Annual International Cryptology Conference*, pages 682–710. Springer, 2017.