Locally Computable Arithmetic Pseudorandom Generators

Lior Zichron[*]

August 12, 2017

Abstract

Pseudorandom generators (PRGs) use a short k-bit random seed to generate a longer m-bit pseudorandom string. Locally-computable PRGs are PRGs which enjoy a high level of efficiency: Each of their outputs can be computed based on constant number of inputs. In the last decade, such PRGs were extensively studied. Candidate constructions were suggested, in addition to several interesting applications.

In this work, we initiate the study of local PRGs over large prime field. That is, we view the seed as a sequence of k field elements and the output as a sequence of m field elements. We present two constructions of locally-samplable arithmetic distributions based on Noisy Linear Sparse Mapping and based on Expander Graphs. Both constructions were studied in the binary setting by Alekhnovich (FOCS 2003) and Goldreich (ECCC 2000), respectively. For each of these candidates we present new attacks, and prove lower-bounds against restricted types of adversaries. Our results suggest that, in several aspects, the arithmetic setting seems to be easier to analyze and even more secure than the binary setting. In particular, it seems that security in the arithmetic setting requires modest combinatorial properties than the binary setting. In a follow-up work [ADI$^+$17] our constructions are shown to yield the first protocol for securely computing any arithmetic function with constant computational overhead.

1 Introduction

Pseudorandom generators (PRGs) use a short k-bit random seed to generate a longer m-bit pseudorandom string. Locally-computable PRGs are PRGs which enjoy a high level of efficiency: Each of their outputs can be computed based on constant number of inputs. In the last decade, such PRGs were extensively studied. Candidate constructions were suggested, in addition to several interesting applications in cryptography and computational complexity. (See [App16] for a survey).

In this work, we initiate the study of local PRGs over a large prime field $\mathbb{F}$. That is, we view the seed as a sequence of k field elements and the output as a sequence of m field elements where m is typically polynomially larger than k. The size of the field $\mathbb{F}$ is typically assumed to be super-polynomial or even exponential in the security parameter k. Our goal is to generate a long sequence of m pseudorandom field elements by making only a constant amount of arithmetic operations per output symbol.

We present two candidate constructions based on *Noisy Linear Sparse Mapping* and based on *Expander Graphs.* Both constructions were studied in the binary setting by Alekhnovich [Ale03]

[*] School of Electrical Engineering, Tel-Aviv University, zichronlior@gmail.com. This paper is extracted from the Master Thesis of the second author which was written under the supervision of Professor Benny Applebaum.

and Goldreich [Gol00], respectively. For each of these candidates we present new attacks, and prove lower-bounds against restricted types of adversaries. Our results suggest that, in several aspects, the arithmetic setting seems to be easier to analyze and even more secure than the binary setting. In particular, it seems that security in the arithmetic setting requires modest combinatorial properties than the binary setting. In a follow-up work [ADI$^+$17], joint with Applebaum, Damgård, Ishai and Nielsen, our constructions are shown to yield the first protocol for securely computing any arithmetic function with only constant computational overhead.

1.1 Overview of Our Results

1.1.1 The Security of Sparse Noisy Linear Mapping

We begin by formally defining *Noisy Sparse Linear Mapping*:

Definition 1.1 (Noisy Sparse Linear Mapping over $\mathbb{F}$)**.** *For a field $\mathbb{F}$, an $m \times k$ matrix with at most d non-zero values in each row (hereafter referred to as d-sparse matrix) and noise parameter $\mu \in [0,1]$, we define the* randomized function *$f_{M,\mu} : \mathbb{F}^k \to \mathbb{F}^m$ as follows. Given an input $x \in \mathbb{F}^k$ output the m-long vector $M \cdot x + \alpha$ where $\alpha \in \mathbb{F}^m$ is a "noise vector" that each of its entries α_i is sampled independently as:*

$$\alpha_i = \begin{cases} 0 & \textit{with probability } 1-\mu \\ \textit{uniform field element} & \textit{with probability } \mu \end{cases}$$

We sometimes abuse notation and think of $f_{M,\mu}$ as the probability distribution that corresponds to $f_{M,\mu}(x)$ where $x \overset{R}{\leftarrow} \mathbb{F}^k$.[1]

We relate the pseudorandomness of $f_{M,\mu}$ to the linear-algebraic and combinatorial properties of the corresponding matrix, M. One such key property is the *dual distance* of the matrix (i.e., the distance of the linear code .

Definition 1.2 (Dual Distance)**.** *Let $\mathrm{dd}(M)$ be the maximal integer D for which every subset of M's rows of size at most $i < D$ is linearly independent over $\mathbb{F}$.*[2]

It is not hard to see that a small dual distance can be exploited to distinguish the distribution $f_{M,\mu}$ from the uniform distribution over $\mathbb{F}^m$ with good advantage. Indeed, in order to tell whether $y \in \mathbb{F}^m$ was sampled from the uniform distribution or from $f_{M,\mu}$, we can restrict y to a small subset of linearly dependent rows $L \subset [m]$ and check whether the restricted vector, y_L, belongs to the image of the sub-matrix M_L. A vector from $f_{M,\mu}$ passes the test with probability $(1-\mu)^{|L|}$, whereas a uniform vector passes the test with probability of at most $1/|\mathbb{F}|$ which is tiny for a large field.

It follows that there is always an efficient (in fact linear) distinguisher with distinguishing advantage of roughly $(1-\mu)^{\mathrm{dd}(M)}$. We conjecture that one cannot do much better. Namely, that, for sparse matrix M, the distribution $f_{M,\mu}$ ε-fools efficient adversaries with $\varepsilon = \exp(-\Omega(c_\mu \mathrm{dd}(M)))$

[1] The above distribution is locally samplable. Also, it is non-trivially pseudorandom in the sense that, for constant μ and $m > \omega(k)$, the sampling complexity is smaller than the length of the resulting pseudorandom sequence. Still, strictly speaking, it does not yield a locally-computable PRG since it is not clear how to sample $f_{M,\mu}$ by a procedure which is both local and economic (in its use of randomness). This point will not bother us, since the construction as is, turns to be useful for the applications presented in [ADI$^+$17].

[2] The name "dual distance" comes from the fact that the left null space of M is a linear code of distance $\mathrm{dd}(M)$.

where c_μ depends on the noise rate μ.[3] Indeed, for limited classes of attacks, we can prove that our conjecture holds. In Section 3.2 we prove the following theorem (a fuller statement appears there):

Theorem 1.3. *For every prime order field $\mathbb{F}$ and every d-sparse matrix $M \in \mathbb{F}^{m\times k}$ with $r = \mathrm{dd}(M)$ and noise parameter $\mu \in (0,1)$, the noisy mapping $f_{M,\mu}$ satisfies the following properties:*

1. *Its output is r-wise independent (i.e., it perfectly fools adversaries that can compute an arbitrary function that depends on r elements).*
2. *It ε-fools linear adversaries (that can be compute by degree-1 polynomials) with $\varepsilon = e^{-\mu r}$.*
3. *It ε-fools adversaries that can be computed by degree-b polynomials with $\varepsilon = 16e^{\frac{-\mu r}{b2^{b-1}}}$.*
4. *It ε-fools (n,t)-product tests where $n < r$ and $\varepsilon = t(1-\mu)^{\Omega(r^2/m)}$.*

Roughly speaking, the last item refers to adversaries that partition the vector $y \in \mathbb{F}^m$ to t disjoint blocks of size n each, applies an arbitrary function to each block separately and take the product (over the complex numbers) of all the results. (See Section 2 for formal definition.)

How to sample matrices with large dual distance? We suggest to sample a d-sparse matrix $M \in \mathbb{F}^{m\times k}$ in two steps. First, choose the locations of the non-zero entries of the matrix (e.g., by selecting a random set of d entries per row), and then fill them with random field elements. The outcome of the first step can be viewed as a d-sparse $m \times k$ zero-one matrix G. To analyze the process, we relate the dual distance of the final matrix M to the expansion properties of the matrix G which can be naturally viewed as a d-uniform hypergraph over the vertex set $[k]$ with m hyperedges. (Hereafter referred to as (m,k,d)-hypergraph.)

It is well known that if every set S of at most r hyperedges in G expands by a factor $\alpha > d/2$ (i.e., the hyperedges in S "touch" more than $\alpha|S|$ vertices) then M will have (with probability 1) a dual distance of at least r. Interestingly, we show that, over large field $\mathbb{F}$, it suffices to require a much weaker expansion factor of $(1+\varepsilon)$ that is independent of the sparsity parameter d. This is essentially optimal since a shrinking set of hyperedges (which expands by a factor smaller than 1) induce a linearly-dependent set of rows in M. By analyzing the expansion of a random sparse hypergraphs (using standard tools) we derive estimation for the dual distance of M (which are better than the ones available for small fields).

Formally, let us denote by $\mathcal{M}(G,\mathbb{F})$ the outcome of the second step of the process applied to some (m,k,d)-hypergraph G, and let $\mathcal{M}(m,k,d,\mathbb{F})$ denote the distribution obtained by selecting the (m,k,d)-hypergraph G at random and then sampling from $\mathcal{M}(G,\mathbb{F})$. In Section 3.3, we prove the following lemma.

Lemma 1.4. *Suppose that G is a (m,k,d)-hypergraph which is $(r,1+\varepsilon)$-expanding and $|\mathbb{F}|^\varepsilon > m$. Then,*

$$\Pr_{M\leftarrow\mathcal{M}(G,\mathbb{F})}[\mathrm{dd}(M) < r] < |\mathbb{F}|^{-1}\sum_{t=1}^{r}\left(\frac{m}{|\mathbb{F}|^\varepsilon}\right)^t$$

Consequently, a random $M \leftarrow \mathcal{M}(m,k,d,\mathbb{F})$, with $|\mathbb{F}|^\varepsilon > m$ and $m = \Delta k$ has dual-distance of $r = \frac{k}{\Delta^{\frac{1}{d-2.1}}}$ with probability of at least $1 - O(\left|\mathbb{F}^{-1}\right|) - o(1)$.

[3] We say that a distribution $\mathcal{D}$ over the domain $\mathbb{F}^m$, ε-fools a class of adversaries $\mathcal{A}$ if for every adversary $A \in \mathcal{A}$ the statistical distance between $A(\mathcal{D})$ and $A(U_m)$ is at most ε, where U_m is the uniform distribution over $\mathbb{F}^m$. (Note that the output of A may not be binary).

In Section 7, we calculate concrete security parameters based on the above results. For example, we show that for a noisy-mapping $f_{M,\mu} : \mathbb{F}^k \to \mathbb{F}^{k^2}$ over prime field of size $\log(|\mathbb{F}|) \geq 512$, with noise parameter $\mu = 0.25$ and input sizes $k = 156$ or $k = 249$, the best known attack run in time at least 2^{80} or 2^{100}, correspondingly.

1.1.2 Results Regarding the Security of Goldreich's functions

We begin by formally defining *Goldreich's functions* over a finite field $\mathbb{F}$:

Definition 1.5 (Goldreich's functions)**.** *Let s denote some d-tuple with distinct elements in $[k]$ and for every $x \in \mathbb{F}^k$ let $x|_s$ denote the restriction of x to s. For a d-ary polynomial $P : \mathbb{F}^d \to \mathbb{F}$, and a tuple $G = (s_1, \dots, s_m)$, where each s_i is a d-tuple of distinct elements in $[k]$, we let $f_{G,P} : \mathbb{F}^k \to \mathbb{F}^m$ denote the d-local function,*

$$f_{G,P} = \big(P(x|_{s_1}), \dots, P(x|_{s_m})\big)$$

We sometimes view G as a hypergraph over k vertices with m ordered hyperedges $s_1, \dots, s_m$ each of cardinality d. Correspondingly, we refer to G as the input-output dependency hypergraph of $f_{G,P}$. We define a probability distribution $\mathcal{G}_{m,k,d}$ over hypergraphs $G = (s_1, \dots, s_m)$ by choosing each s_i independently and uniformly at random. Finally, for d-ary polynomial P, we let $\mathcal{F}_{m,k,P}$ denote the probability distribution over functions $f : \mathbb{F}^k \to \mathbb{F}^m$ obtained by selecting $G \overset{R}{\leftarrow} \mathcal{G}_{m,k,d}$ and letting $f = f_{G,P}$.

The sum-product function. We mainly focus on the *sum-product polynomial* that, for parameters a, b (and arity $d = a + b$), is defined by

$$\mathsf{SP}_{a,b}(w_1, \dots, w_{a+b}) = (w_1 + \cdots + w_a) + (w_{a+1} \cdot \cdots \cdot w_{a+b}).$$

When analyzing the function $f_{G,\mathsf{SP}_{a,b}}$ it will be convenient to "split" the dependency (m, k, d)-hypergraph G into an a-uniform addition hypergraph whose hyperedges contain the first a entries of the hyperedges of G, and to a b-uniform *product hypergraph* whose hyperedges contain the last b entries of the hyperedges of G. We refer to the addition hypergraph as G_Σ and to the product hypergraph as G_Π.

The sum-product function over binary field. Over the binary field, the pseudorandomness of the sum-product function was extensively studied. In particular, it was shown that for output lengths of $m = k^{3/2-\varepsilon}$ and parameters $a = 3, b = 2$ a random sum-product function $f \overset{R}{\leftarrow} \mathcal{F}_{k,m,\mathsf{SP}_{a,p}}$ is likely to fool linear distinguishers [ABR12, OW14]. It was conjectured that this result can be extended to arbitrary polynomial output length by taking the parameters a and b to be sufficiently large constants. This conjecture was refuted by [AL16] who showed a linear attack for sum-product functions with output length of k^2.

We extend these results in several ways. First, we close the gap between the lower-bound and the upper-bound and show that over the binary field (and any other prime-order field) sum-product fools linear attacks for almost quadratic output lengths. In particular, we prove the following theorem in Section 4.3.

Theorem 1.6. *For any $\delta > 0$, $a > 4$,$b \geq 8$ output length $m = O(k^{2-\delta})$, and prime field $\mathbb{F}$, with high probability over the choice of a randomly sampled $G \leftarrow \mathcal{G}(m, k, a+b)$-hypergraph*[4]*, the function $f_{G,\mathsf{SP}_{a,b}} : \mathbb{F}^k \to \mathbb{F}^m$ ε-fools any linear test with $\varepsilon = \exp(\Omega(k^{\min(\delta/2, \delta - \frac{2}{a-2.1})}))$.*

We move on to analyze the security of sum-product over large fields. First, we analyze (in Section 4.1) the linear attack of [AL16] against the sum-product function over a large field $\mathbb{F}$ with quadratic output length and show that it achieves a distinguishing advantage of $\frac{|\mathbb{F}|-1}{|\mathbb{F}|^2} = \Omega(1/|\mathbb{F}|)$. Moreover, we show that this is essentially tight even if the output length is larger than the input by an arbitrary polynomial.

Theorem 1.7. *Let G be a (m, k, d)-hypergraph with the property that every pair of hyperedges intersect in less than b vertices. Let $P : \mathbb{F}^d \to \mathbb{F}$ be any multilinear polynomial of degree at least b. Then, $f_{G,P} : \mathbb{F}^k \to \mathbb{F}^m$ ε-fools linear tests with $\varepsilon = \frac{k-1}{|\mathbb{F}|}$.*

It follows that when the field is large (super-polynomial in the security parameter) one can achieve negligible bias ε even for an arbitrary polynomial output length $m = k^s$ by using the sum-product polynomial with sufficiently large parameters a and b. In particular, for $b > s/2$, a randomly chosen $f \overset{R}{\leftarrow} \mathcal{F}_{k^s,k,\mathsf{SP}_{a,b}}$ is likely to $\frac{k-1}{|\mathbb{F}|}$-fools linear tests.

Security against low-degree polynomials. We can partially extend our results to the case of low degree polynomials. Specifically, we show that no such polynomial can distinguish the image of $f_{G,\mathsf{SP}_{a,b}}$ from the uniform distribution with *1-sided error*. Put differently, there is no low degree non-trivial polynomial $Q : \mathbb{F}^m \to \mathbb{F}$ that *annihilates* the sum-product function. (Recall that $Q : \mathbb{F}^m \to \mathbb{F}$ annihilates the function $f : \mathbb{F}^k \to \mathbb{F}^m$ if $Q \circ f : \mathbb{F}^k \to \mathbb{F}$ is the zero polynomial.) The following theorem is proved in Section 5.2.

Theorem 1.8. *Consider the function $f_{G,\mathsf{SP}_{a,b}} : \mathbb{F}^k \to \mathbb{F}^m$ and assume that G_Π is $(2r+1, b/2+1)$-expanding. Then, $f_{G,\mathsf{SP}_{a,b}}$ cannot be annihilated by any non-trivial polynomial of degree smaller than r.*

We note that if there exists a degree ℓ annihilating polynomial then it can be found by making only $O(m^{3\ell})$ arithmetic operations (see Section 5.1). It should be emphasized that the complexity (number of arithmetic operations) of such attack is independent of the field size.[5] Assuming that one cannot do better, and plugging in the expansion parameters of a randomly chosen graph (Claim 2.10), we get $\Omega\left(\frac{k}{\Delta^{\frac{1}{b-2.1}}}\right)$ bits of security against that attack - for $f_{G,\mathsf{SP}_{a,b}} : \mathbb{F}^k \to \mathbb{F}^{\Delta k}$.

We do not know how to prove that the function fools general (low degree) polynomials. (This is open even for degree-2 polynomials over the binary field.) We bound, however, the distinguishing advantage achievable by polynomial tests $T : \mathbb{F}^m \to \mathbb{F}$, of degree at most q, where each of their inputs appears in at most t monomials. We show that, for slightly super-linear output lengths, the sum-product function fools such (q, t) *sparse* polynomials. (Our analysis here is limited for fixed-size fields.)

[4] Throughout the paper, an event happens with high probability if it happens with probability which converges to 1 when k goes to infinity.

[5] Without the latter restriction, one can easily get an attack of complexity $|\mathbb{F}|^r$. Given $y \in \mathbb{F}^m$ check if the restriction of y to the set S of shrinking hyperedges has a preimage x. Since S is shrinking, a random y passes the test with probability at most $1/|\mathbb{F}|$. Moreover, the test can be implemented by trying all assignments for the inputs of x that participates in the hyperedges of S, which takes at most $\mathbb{F}^r$ operations.

Theorem 1.9. *For every $a, b \geq 4$ and $0 < \rho < 0.5\frac{a-2}{a-1}$, $m = k^{1+\rho}$ and every constants q and t, with high probability over $G \leftarrow \mathcal{G}(m, k, a+b)$, the function $f_{G,\mathsf{SP}_{a,b}} : \mathbb{F}^k \to \mathbb{F}^m$ ε-fools any (q,t) sparse polynomial $T : \mathbb{F}^m \to \mathbb{F}$ with $\varepsilon = \exp(-\Omega(k^{1-2\rho\frac{a-1}{a-2}}))$ where the constant in the Ω depends on a, b, q, t and the field size $\mathbb{F}$.*

See Section 6 for proof and details.

Concrete parameters. Finally, taking into account the known attacks, we suggest (in Section 7) some parameters for instantiation. For example, we show that over prime field of size $\log(|\mathbb{F}|) \geq 128$, the $(3,4)$-product-sum polynomial $\mathsf{SP}_{3,4}$ with input length of $k \geq 250$ and output length of $m = k^2$, a random function $f \overset{R}{\leftarrow} \mathcal{F}_{k,m,\mathsf{SP}_{a,b}}$ is likely to $k2^{-128}$-fools linear test.

2 Preliminaries

2.1 List of Relevant Attacks

In the following, $\mathbb{F}$ is taken to be some finite field.

Definition 2.1 (Statistical Distance)**.** *Let $\mathcal{A}$ and $\mathcal{B}$ be two distributions over $\mathbb{F}$ (the term could be similarly applied to any alphabet) then:*

$$\mathbf{SD}(\mathcal{A}, \mathcal{B}) = \frac{1}{2}\sum_{i \in \mathbb{F}} \left| \Pr_{x \leftarrow \mathcal{A}}[x = i] - \Pr_{x \leftarrow \mathcal{B}}[x = i] \right|$$

We proceed by defining the notion of ε-fooling (possibly non-Boolean) distinguishers.

Definition 2.2 (ε-Fooling)**.** *A distribution W over $\mathbb{F}^m$ is ε-fooling a function family $\mathcal{H}$ if for any $h \in \mathcal{H}$ we have that $\mathbf{SD}(h(W), h(U_m)) < \varepsilon$ where U_m is the uniform distribution over $\mathbb{F}^m$. Similarly, a function $f : \mathbb{F}^k \to \mathbb{F}^m$ is ε-fooling a function family $\mathcal{H}$ if the distribution $f(U_k)$ ε-fools it. Finally, we say that an ensemble $\mathcal{F} = \{\mathcal{F}_n\}_{n \in \mathbb{N}}$ ε-fools function family $\mathcal{H}$* with high probability *if*

$$\Pr_{f \leftarrow \mathcal{F}_n}[f \ \varepsilon\text{-fools } \mathcal{H}] \geq 1 - o(1).$$

Definition 2.3 (Degree q Polynomial Test)**.** *A degree q polynomial test is a function family of all polynomial functions $T : \mathbb{F}^m \to \mathbb{F}$ such that any of T's monomials is of degree at most q. An important special case is the* linear test. *Specifically, for any linear test $T : \mathbb{F}^m \to \mathbb{F}$ there exists some $\alpha \in \mathbb{F}^m$ such that for all $y \in \mathbb{F}^m$ we have that $T(y) \equiv \langle \alpha, y \rangle$ where $\langle \cdot, \cdot \rangle$ stands for inner product.*

Definition 2.4 ((n,t)-Product Test)**.** *A t-function product test is a function family of all functions $h : \mathbb{F}^m \to \mathbb{C}_1$ ($\mathbb{C}_1$ is the complex unit disc) that can be written as the product of t disjoint input functions. Meaning that there exists $g_1, \ldots, g_t : \mathbb{F}^n \to \mathbb{C}_1$ functions and a division of $[m]$ into t disjoint sets $\{H_i\}_{i=1}^t$ each of size $|H_i| = n$ such that $h(x) = \prod_{i=1}^t g_i(x|_{H_i})$.*

The notion of *Annihilating Polynomial* can be viewed as a distinguisher with one-sided error.

Definition 2.5 (Annihilating Polynomial)**.** *Let $f : \mathbb{F}^n \to \mathbb{F}^m$ be some degree d polynomial. We say that a non-trivial polynomial $Q : \mathbb{F}^m \to \mathbb{F}$ is f-annihilating if $\forall x \in \mathbb{F}^n$ we have that $Q\big(f(x)\big) = 0$.*

We note that a low-degree annihilating polynomial yields a good distinguisher.

Remark 2.6 (Annihilating Polynomial is a 'Good Attacker'). *Let $f : \mathbb{F}^n \to \mathbb{F}^m$ be any function and let $Q : \mathbb{F}^m \to \mathbb{F}$ be an f-annihilating polynomial. By Schwartz-Zippel Lemma we have that Q distinguish $y \notin \operatorname{Im}(f)$ from any output of f with probability at least $1 - \frac{\deg(Q)}{|\mathbb{F}|}$.*

2.2 Graphs and Matrices Properties

In this paper we show that many resilience properties of the PRG's results from the structure of their dependencies graph. Specifically we would show that they relate to the graph expansion. Informally, the expansion refers to the ratio between a subgroup of the output and the input variables that affect it. More formally:

Definition 2.7 (Boundary Set). *For any matrix $M_{m\times k}$ and subset of the columns $S \subset [k]$ we define the* Boundary Set *of S:*

$$B(S) \doteq \{i \in [m] \textit{such that } \exists j \in S \textit{with } M_{i,j} \neq 0\}$$

meaning the subset of the rows matching non-zero entries of S's columns. The term is similarly defined for any subset of the rows.

Equipped with this notation we can proceed to defining the set expansion:

Definition 2.8 (Expanding / Shrinking Sets). *For any matrix $M_{m\times k}$ and subset of its rows $S \subset [m]$ we note $\alpha = \frac{|B(S)|}{|S|}$. If $\alpha > 1$ we say that S is an* expanding set *with expansion parameter α otherwise we say that it is a* shrinking set *with parameter α. Correspondingly, for a hypergraph $G_{m,k,d}$ we say that a subset S of its hyperedges is expanding or shrinking if the same subset of the rows in its adjacency matrix is expanding or shrinking, respectively.*

Definition 2.9 (Minimal non α-Expanding Set). *For any $\alpha > 0$ and matrix M we let $\rho_\alpha(M)$ denote the size of the minimal $S \subset [m]$ such that S is not α-expanding, i.e., $\frac{|B(S)|}{|S|} < \alpha$. We use similar notation, $\rho_\alpha(G)$,for the minimal non α-expanding subset of G's hyperedges*

We introduce the following useful relations for analyzing graph expansion:

Claim 2.10. *A random $G \leftarrow \mathcal{G}_{m,k,d}$-hypergraph, with $m = \Delta k$, $a \geq 3$, $d > a/2$ and $r = \frac{k}{\Delta^{\frac{2}{a-2.1}}}$ is, with high probability, $(r, d - \frac{a}{2})$-expanding. That is,*

$$\rho_{d-\frac{a}{2}}(G) = \Omega\left(\frac{k}{\Delta^{\frac{2}{a-2.1}}}\right)$$

Which in turn immediately yields the following corollary:

Corollary 2.11. *A random $G \leftarrow \mathcal{G}_{m,k,d}$-hypergraph, with $m = \Delta k$, $d > 2$, with high probability, contains no shrinking set of size $O\left(\frac{k}{\Delta^{\frac{1}{d-2.1}}}\right)$. (i.e., $\rho_1(G) = \Omega\left(\frac{k}{\Delta^{\frac{1}{d-2.1}}}\right)$)*

And it is supplemented by the following claim:

Claim 2.12. *For every $\alpha \in (0,1)$, any (m,k,d)-hypergraph, with $m = k^c$, has an α-shrinking set of hyperedges of size*

$$t = \Theta\left(\alpha^{\frac{-1}{d-1}} k^{\frac{d-c}{d-1}}\right)$$

In particular,

$$\rho_\alpha(G) = \Theta\left(\alpha^{\frac{-1}{d-1}} k^{\frac{d-c}{d-1}}\right)$$

Proof of Claim 2.10. For a graph $G \leftarrow \mathcal{G}_{\Delta k,k,d}$ let N_ℓ denote the size of the set:

$$\{I \subset [m]; |I| = \ell, |\cup_{i \in I} s_i| \leq (d - \frac{a}{2})\ell\}$$

That is, N_ℓ counts the number of ℓ-size subsets of the hyperedges which fail to be $(d - \frac{a}{2})$-expanding. Then, we have that:

$$\begin{aligned}
\Pr[G \textit{ has non } (d - \frac{a}{2})\textit{-exspanding } \ell\textit{-size subset}] &\leq \mathsf{E}[N_\ell] & (1)\\
&= \binom{\Delta k}{\ell}\binom{k}{\ell(d-\frac{a}{2})}\binom{\ell(d-\frac{a}{2})}{d}^{\ell}\binom{k}{d}^{-\ell} & (2)\\
&\leq (\frac{e\Delta k}{\ell})^{\ell}(\frac{ek}{\ell(d-\frac{a}{2})})^{\ell(d-\frac{a}{2})}(\frac{e\ell(d-\frac{a}{2})}{d})^{d\ell}(\frac{ek}{d})^{-d\ell} & (3)\\
&= \left[c_{d,a}\Delta(\frac{\ell}{k})^{\frac{a-2}{2}}\right]^{\ell} & (4)\\
&= \left[\ell\frac{\Delta^{\frac{2}{a-2}}}{\alpha_{d,a}k}\right]^{\frac{a-2}{2}\ell} & (5)
\end{aligned}$$

Where the transition from probability to expectation shown in (1) follows from union bound, and the next transition follows from standard bound of binomial coefficient. And for $r = \frac{k}{\Delta^{\frac{2}{a-2.1}}}$ we have that:

$$
\begin{aligned}
\Pr[G \textit{ is not } (r, d-\frac{a}{2})\textit{-exspanding }] &\leq \mathsf{E}[\sum_{\ell=1}^{r} N_\ell] && (6)\\
&\leq \sum_{\ell=1}^{r} \mathsf{E}[N_\ell] && (7)\\
&\leq \sum_{\ell=1}^{r} \left[\ell \frac{\Delta^{\frac{2}{a-2}}}{\alpha_{d,a} k}\right]^{\frac{a-2}{2}\ell} && (8)\\
&\leq \sum_{\ell=1}^{r} \left[r \frac{\Delta^{\frac{2}{a-2}}}{\alpha_{d,a} k}\right]^{\frac{a-2}{2}\ell} && (9)\\
&\leq \sum_{\ell=1}^{r} \left[\frac{\Delta^{\frac{2}{a-2}-\frac{2}{a-2.1}}}{\alpha_{d,a}}\right]^{\frac{a-2}{2}\ell} && (10)\\
&\leq \sum_{\ell=1}^{\infty} \left[\frac{\Delta^{\frac{-0.2}{(a-2)(a-2.1)}}}{\alpha_{d,a}}\right]^{\frac{a-2}{2}\ell} && (11)\\
&= o(1) && (12)
\end{aligned}
$$

□

Proof of Claim 2.12. Let G be some (m, k, d)-hypergraph and $T \subset [k]$ some size t subset of the input. We say that an hyperedge e is *contained in T* if all of its vertices belong to it. If we sample a uniformly chosen T the probability that an hyperedge e will be contained in T is

$$\binom{k-d}{t-d}\binom{k}{t}^{-1} = \frac{t \cdot (t-1) \cdots (t-d)}{k \cdot (k-1) \cdots (k-d)} = \left(\frac{t}{k}\right)^d \gamma_{k,t,d}$$

where for constant d and $k, t \to \infty$ we have that $\gamma_{k,t,d} \to 1$. We let N be random variable representing the number of hyperedges contained in T and by linearity of expectation we have that:

$$\mathsf{E}[N] = m \cdot \Pr[e \textit{ is contained in } T] = k^c \cdot \left(\frac{t}{k}\right)^d \gamma_{k,t,d}$$

and for $t = (\gamma_{k,t,d})^{\frac{-1}{d}} \alpha^{\frac{-1}{d-1}} k^{\frac{d-c}{d-1}}$ we get that:

$$\mathsf{E}[N] = t \left(\frac{t}{k}\right)^{d-1} k^{c-1} = \alpha^{-1} t$$

meaning that $\rho_\alpha(G) = \Theta\left(\alpha^{\frac{-1}{d-1}} k^{\frac{d-c}{d-1}}\right)$ □

Finally we mention the following properties of random graphs:

Claim 2.13. *A randomly chosen $(\Delta k, k, d)$-hypergraph, with $\Delta = k^\varepsilon$ for some constant $\varepsilon > 0$, satisfies the following properties with probability $1 - o(1)$:*

1. *Each vertex is member of at most $2\Delta d$ hyperedges. Meaning that, $\forall v \in V$ we have that $|\{e \in E; v \in e\}| \leq 2\Delta d$*
2. *An upper bound on the number of hyperedge that intersect each single hyperedge: $\forall e \in E$ we have that $|\{e' \in E; e' \cap e \neq \emptyset\}| \leq 2\Delta d^2$*
3. *If $\varepsilon < 1$ then, $\forall e', e \in E$ we have that $e' \cap e \leq 3$*

Proof. We let $G \leftarrow \mathcal{G}_{\Delta k,k,d}$ be some randomly sampled hypergraph for $\Delta = k^{\varepsilon}$ and prove each of these statements separately:

1. The first item follows directly from upper bounding the probability of that event using *Chernoff's Bound.* We first note that $\mathsf{E}\,|\{e \in E; v \in e\}| = \Delta d$ and proceed by:

$$\begin{aligned}
\Pr[\exists v \text{ s.t. } |\{e \in E; v \in e\}| \geq 2\Delta d] &\leq k \Pr[|\{e \in E; v \in e\}| \geq 2\Delta d] && (13)\\
&= k \Pr[|\{e \in E; v \in e\}| \geq 2\,\mathsf{E}\,|\{e \in E; v \in e\}|] && (14)\\
&\leq k \exp[-\frac{1}{3}\,\mathsf{E}\,|\{e \in E; v \in e\}|] && (15)\\
&\leq k \exp[-\frac{d}{3}k^{\varepsilon}] && (16)
\end{aligned}$$

Where the transition to 15 results from Chernoff's Bound and the term obviously goes to zero.

2. The second item is an immediate corollary: each hyperedge consists of exactly d vertices and each vertex is contained in at most $2\Delta d$ hyperedges. Hence, a hyperedge share a vertex with at most $d2\Delta d = 2\Delta d^2$ different hyperedges.

3. The last item can be obtained using the union bound on the probability that two different hyperedges share 4 vertices or more. Assume $\varepsilon < 1$ then:

$$\begin{aligned}
\Pr[\exists e', e \in E; e' \cap e \geq 4] &\leq m^2 \Pr[e_1 \cap e_2 \geq 4] && (17)\\
&\leq m^2 \binom{k}{2d-4}\binom{2d-4}{4}\binom{k}{d}^{-2} && (18)\\
&\leq O\left(m^2 k^{-4}\right) && (19)\\
&\leq O\left(k^{2\varepsilon-2}\right) = o(1) && (20)
\end{aligned}$$

Where the transition in 17 results from taking union bound.

□

2.3 Characteristic Distance and the Gowers' Norm

We begin by defining the *Characteristic Distance* of a random variable:

Definition 2.14 (Characteristic Distance)**.** *For a random variables $A, B : \Omega \to F_p$ we shall define the 'Characteristic Distance' of A from B by:*

$$\mathbf{CD}(A,B) \doteq |E[\omega^A] - E[\omega^B]|$$

where ω is the p'th root of 1.

We use the *Characteristic Distance* and *Statistical Distance* interchangeably due to the following claim:

Claim 2.15. *For every pair of random variables X and X', taking values in $\mathbb{F}$ we have*

$$\mathbf{CD}(X, X') \leq 2\mathbf{SD}(X, X') \leq \sqrt{|\mathbb{F}| - 1} \cdot \mathbf{CD}(X, X')$$

See [BV10] for proof. Moreover, in this section we introduce the functional *Gowers Uniformity.* The usefulness of the Gowers Uniformity stems from its ability to eliminate lower order terms and its relation to the *Characteristic Distance* (in particular the zero degree Gowers Uniformity of a function equals to its Characteristic Distance from the uniform distribution).

Definition 2.16 (Gowers Uniformity). *For every function $f : \mathbb{F}^k \to \mathbb{F}$ and every positive integer q, the degree-q Gowers Uniformity $U_q(f)$ is defined as:*

$$U_q(f) \doteq \mathop{\mathsf{E}}_{x, y_1, \ldots, y_q \leftarrow F^n} [\omega^{\Delta_{y_1, \ldots, y_q} f(x)}]$$

where $\Delta_{y_1, \ldots, y_q}$ is a degree-q directional derivative and $\omega \doteq exp(2\pi i/|\mathbb{F}|)$.

We cite the following proposition from [**?**] which discuss properties of the Gowers Uniformity:

Proposition 2.17 (Proposition 2.7 of [**?**]). *The Gowers Uniformity has the following properties:*

1. $|U_q(f)| \leq U_{q+1}(f)^{1/2}$
2. $U_{q+1}(f - p) = U_{q+1}(f)$ *for any degree-q polynomial $p : \mathbb{F}^k \to \mathbb{F}$*
3. *Let $\left\{f_i : \mathbb{F}^k \to \mathbb{F}\right\}_{i=1}^t$ be a set of functions and let $S_i \subset [k]$ denote that variables that f_i depends on, then, if $\{S_i\}_{i=1}^t$ are disjoint then:*

$$U_q\left(\sum_{i=1}^t f_i\right) = \prod_{i=1}^t U_q(f_i)$$

The first two items appear in [**?**], since the last item is a slight generalization we shell prove it:

Proof of item 3. Let, for any $i \in [t]$, $f_i : \mathbb{F}^k \to \mathbb{F}$ be a function with the property that $\{f_1(U_k), \ldots, f_t(U_k)\}$ is a set of independently distributed random variables, then:

$$U_q\left(\sum_{i=1}^t f_i\right) = \mathop{\mathsf{E}}_{x, y_1, \ldots, y_q \leftarrow F^n}\left[\omega^{\Delta_{y_1, \ldots, y_q} \sum_{i=1}^t f_i(x)}\right] \tag{21}$$

$$= \mathop{\mathsf{E}}_{x, y_1, \ldots, y_q \leftarrow F^n}\left[\omega^{\sum_{i=1}^t \Delta_{y_1, \ldots, y_q} f_i(x)}\right] \tag{22}$$

$$= \mathop{\mathsf{E}}_{x, y_1, \ldots, y_q \leftarrow F^n}\left[\prod_{i=1}^t \omega^{\Delta_{y_1, \ldots, y_q} f_i(x)}\right] \tag{23}$$

$$= \prod_{i=1}^t \mathop{\mathsf{E}}_{x, y_1, \ldots, y_q \leftarrow F^n}\left[\omega^{\Delta_{y_1, \ldots, y_q} f_i(x)}\right] \tag{24}$$

$$= \prod_{i=1}^t U_q(f_i) \tag{25}$$

Where 22 results from the linearity of derivation and the transition to 24 results from the fact that, since each f_i depends on different input variables, $\left\{\Delta_{y_1,\ldots,y_q} f_i(x)\right\}$ are independently distributed and the expectation of a product equals the product of the expectations. □

Finally, we introduce the following lemma which we used in the proof of theorem 4.10 and 6.1:

Lemma 2.18. *Let $f_1, \ldots, f_t : \mathbb{F}^d \to \mathbb{F}$ be some degree-b functions and $S_1, \ldots, S_t \subset [k]$ distinct subsets of size $|S_i| = d$. Let $R : \mathbb{F}^k \to \mathbb{F}$ be some function of total degree of at most $c < b$ in $S = \bigcup_i S_i$ variables (i.e., for each of R's monomials the sum of the degrees of variables $\{x_i\}_{i \in S}$ is at most c). Then:*

$$\mathbf{CD}\left(\sum_{i=1}^{t} f_i(x|_{S_i}) + R(x), U\right) \le \left(\prod_{i=1}^{t} U_{c+1}(f_i)\right)^{\frac{1}{2^{c+1}}}$$

Proof. Let $f_1, \ldots, f_t : \mathbb{F}^d \to \mathbb{F}$ be some degree-b functions and $S_1, \ldots, S_t \subset [k]$ disjoint subsets of size $|S_i| = d$. Let $R : \mathbb{F}^k \to \mathbb{F}$ be some function of total degree of at most $c < b$ in $S = \bigcup_i S_i$ variables and denote

$$f(x) = \sum_{i=1}^{t} f_i(x|_{S_i}) + R(x)$$

By triangle inequality:

$$\mathbf{CD}(f(x), U) = \left| \mathop{\mathsf{E}}_{x \leftarrow \mathbb{F}^n} [\omega^{f(x)}] \right| \le \mathop{\mathsf{E}}_{x_i, i \notin S} \left| \mathop{\mathsf{E}}_{x_i, i \in S} [\omega^{f(x)}] \right|$$

Therefore it suffices to upper bound the inner expectation for any fixing of the variables $\{x_i : i \notin S\}$. Let x^* denote a random vector in $\mathbb{F}^n$ whose S entries are uniformly chosen from $\mathbb{F}^S$ and its $\bar{S}$-entries are fixed to some arbitrary value in $\mathbb{F}^{\bar{S}}$. Then:

$$\left| \mathop{\mathsf{E}}_{x^*} \left[\omega^{f(x^*)}\right] \right| = \left| \mathop{\mathsf{E}}_{x^*} \left[\omega^{\sum_{i=1}^{t} f_i(x|_{S_i}) + R(x^*)}\right] \right| \tag{26}$$

$$= U_1 \left(\sum_{i=1}^{t} f_i(x|_{S_i}) + R(x^*)\right) \tag{27}$$

$$\le U_{c+1} \left(\sum_{i=1}^{t} f_i(x|_{S_i}) + R(x^*)\right)^{\frac{1}{2^{c+1}}} \tag{28}$$

$$= U_{c+1} \left(\sum_{i=1}^{t} f_i(x|_{S_i})\right)^{\frac{1}{2^{c+1}}} \tag{29}$$

$$= \left(\prod_{i=1}^{t} U_{c+1}(f_i)\right)^{\frac{1}{2^{c+1}}} \tag{30}$$

Where the transition to (27) results from the definition of the degree-1 Gowers Uniformity. The transition to (28) results from the first item of Proposition 2.17 and the transition to (29) results from the second item of 2.17. Finally, we use the third item of 2.17, combined with the fact that S_i are disjoint sets, to obtain (30). □

3 Noisy Sparse Linear Mapping

3.1 Linear Dependency Attack

High level approach In this section we introduce a successful attack against Noisy Sparse Linear Mapping. The general approach of the attack is that for any linearly dependent subset of the row $R \subset [m]$ and "pseudorandom" $y \leftarrow f_{M,\mu}(U_k)$ if $y|_R$ is non-noisy then the equation $y|_R = M|_R x$ has a solution. We later show that, with high probability, it is not the case for a random $y \leftarrow \mathbb{F}^m$. Hence the method attempts to go through exponentially "many" linearly dependent subsets hoping to find a non-noisy one. There are several ways to choose the linearly-dependent sets R.[6]

In particular, our approach is to find a subset $S \subset [m]$ which is $(1-\mu(1+\varepsilon))$-shrinking, to go through all of its $|S|\,(1-\mu(1+\varepsilon))$-size subsets R. Since the expected number of noisy rows in S is $|S|\,\mu$, we would have (with high probability) a non-noisy $|S|\,(1-\mu(1+\varepsilon))$-size subsets R. This set is shrinking (due to the large shrinkage of S), and therefore the corresponding rows are linearly dependent.

We proceed with a formal description of our attack (Algorithm 1).

```
input  : $M$, $y \in \mathbb{F}^m$, noise parameter $\mu$, success parameter $\varepsilon > 0$ and a
         $(1-\mu(1-\varepsilon))$-shrinking subset $S$
for $R$ subset of $S$ of size $|R| = |S|\left(1-\mu(1+\varepsilon)\right)$ do
    try solving $[M|_R; y|_R]$ using gaussian elimenation;
    if Succeed then
        output: 1
    end
end
output: 0
```

Algorithm 1: Attacking the Noisy Mapping

3.1.1 Analysis

In the following, for any $\mu \in [0,1]$ we let $H(\mu) = -\mu\log(\mu) - (1-\mu)\log(1-\mu)$ denote the binary entropy function. We further mention the relation $\binom{M}{\mu M} \approx 2^{MH(\mu)}$ and in particular $\log\binom{M}{\mu M} \leq MH(\mu)$.

Claim 3.1. *Algorithm 1 runs in time* $O\left(|S|^3 \cdot 2^{|S|H(\mu(1+\varepsilon))}\right)$

Proof. The algorithm goes through all $\left(1-\mu(1+\varepsilon)\right)|S|$ size subsets of the output, meaning that it goes through the Gaussian Elimination Process $\binom{|S|}{\left(1-\mu(1+\varepsilon)\right)|S|}$ times which, approximately, equals $2^{|S|H(\mu(1+\varepsilon))}$ and since the Gaussian Elimination Process runs in time $O(|B(S)|^3) \leq O(|S|^3)$ (since S is shrinking) the desired result follows immediately. □

[6] Since a set R of linearly-dependent outputs is clean of noise with probability $(1-\mu)^{|R|} < (1-\mu)^{\mathrm{dd}(M)}$, the expected number of sets that need to be examined (in order to get constant distinguishing advantage) is $\Omega(\frac{1}{1-\mu}^{\mathrm{dd}(M)})$. Hence, the expected complexity of this family of attacks is (at best) exponential in $\Omega\left(\mathrm{dd}(M)\log\frac{1}{1-\mu}\right)$.

k	$\log_k(m)$	d	$\log\lvert\mathbb{F}\rvert$	Noise μ	ε used	log-complexity
200	2	7	128	0.25	0.5	99
400	2	7	128	0.25	0.5	171
800	2	7	128	0.25	0.5	295
1600	2	7	128	0.25	0.5	512

Table 1: This table shows the expected calculated complexity, meaning a bound on the best known attack against *Noisy Mapping*. In particular, for input size k, output length described by $\log_k(m)$, matrix degree parameter-d and field size $|\mathbb{F}|$, and noise parameter μ the table presents a lower bound on the log-complexity.

Claim 3.2. *If y was sampled according to $f_{M,\mu}(U_k)$, then Algorithm 1 output 1 with probability at least* $\exp\left(\frac{-\varepsilon^2\cdot\mu|S|}{3}\right)$.

Proof. We first note that for any subset R that the algorithm goes through, if the rows of R are linearly dependent and no noise was added to them then the algorithm outputs 1. We proceed by mentioning that any such subset of the rows is linearly dependent since $|B(R)| \leq |B(S)|$ and $|R| = |S|\left(1-\mu(1+\varepsilon)\right) = |B(S)|\,(1+\delta)$ hence it is a shrinking set with parameter $1+\delta$. Finally, we note that the probability that the total number of noisy output variables exceeds $\mu(1+\varepsilon)$ can be upper bounded, using *Chernoff's Bound*, by $\exp\left(\frac{-\varepsilon^2\cdot\mu|S|}{3}\right)$ and our claim follows immediately. □

Claim 3.3. *For $y \overset{R}{\leftarrow} \mathbb{F}^m$, Algorithm 1 outputs 1 with probability at most* $2^{-\log|\mathbb{F}|+H(\mu(1+\varepsilon))|S|}$.

Proof. For any $R \subset S$ we have that R is a shrinking set therefore $Im\left\{M|_R\right\}$ has a rank of at most $|R|-1$. Hence, $|Im\left\{M|_R\right\}| \leq |\mathbb{F}|^{|R-1|}$ meaning that the probability that $y \leftarrow \mathbb{F}^{|R|}$ will be in $Im\left\{M|_R\right\}$ is $\frac{1}{|\mathbb{F}|}$. By union bound the probability that that event would occur for any $R \subset B(S)$ is at most:

$$\frac{1}{|\mathbb{F}|}\cdot\binom{|S|}{(1-\mu(1+\varepsilon))\,|S|} \approx \frac{1}{|\mathbb{F}|}\cdot 2^{|S|H(\mu(1+\varepsilon))} = 2^{-\log|\mathbb{F}|+H(\mu(1+\varepsilon))|S|}$$

□

Corollary 3.4. *For any* Noisy Sparse Linear Mapping *and $0 < \varepsilon < \frac{1}{\mu}-1$ we denote $\rho = \rho_{(1-\mu(1+\varepsilon))}(M)$ and there exists an attack that runs for time exponential in $O\left(\rho H(\mu(1+\varepsilon))\right)$ and distinguishing advantage of*

$$1 - 2^{-\log|\mathbb{F}|+\rho H(\mu(1+\varepsilon))} - \exp(\frac{-\varepsilon^2\mu\rho}{3})$$

Based on corollary 3.4 and using techniques introduced in Section 7 we can now proceed to assessing the affectiveness of our attack. We first note that for a matrix sampled by the distribution $M \leftarrow \mathcal{M}(m,k,d,\mathbb{F})$ (see definition 3.8) we can get actual parameters. For instance, in Table 1 we examine parameters settings that guarantee distinguishing advantage of at least 0.8 and calculate the expected shrinking set sizes to calculate the attacks complexity.
More generally, Claim 2.12 yields the following corollary:

Corollary 3.5. *Any (m,k,d)-matrix, for $m = k^c$, contains $(1-\mu(1+\varepsilon))$-shrinking set of size*

$$\Theta\left((1-\mu(1+\varepsilon))^{\frac{-1}{d-1}} \cdot k^{\frac{d-c}{d-1}}\right)$$

Meaning that the complexity of the attack is exponential in:

$$O\left((1-\mu(1+\varepsilon))^{\frac{-1}{d-1}} \cdot k^{\frac{d-c}{d-1}} \cdot H(\mu(1+\varepsilon))\right)$$

Remark 3.6. *We note that the attack can be easily extended to working over any finite-field (not only "large field") by demanding a slightly smaller shrinking parameter - we focus on large field for simplicity of notation.*

3.2 Large Dual-Distance Implies ε-Fooling

In this section we prove the following main theorem:

Theorem 3.7. *For every $M \in \mathcal{M}_{m,k,d}$ with $r = \mathrm{dd}(M)$ and noise parameter $\mu \in [0,1]$, the noisy mapping $f_{M,\mu}$ satisfies the following properties:*

1. *Its output is r-wise independent.*
2. *It ε-fools linear distinguishers with $\varepsilon = e^{-\mu r}$.*
3. *It fools degree b distinguishers with $\varepsilon = 16e^{\frac{-\mu r}{b2^{b-1}}}$.*
4. *It ε-fools (n,t)-product tests where $n < r$ and $\varepsilon = t(1-\mu)^{\Omega(r^2/m)}$.*

The last item makes sense only for sub-quadratic output length $m < k^{2-\delta}$ (since r is sub-linear in k).

Proof. We will prove each of the statements separately:

1. $\mathrm{dd}(M) = r$ meaning that any subset of the rows $s \subset [m]$ of size $|s| = t \leq r$ is linearly independent. Hence, the $t \times m$ submatrix M_s, induced by s, has image dimension of exactly t and kernel dimension of $k - t$. Therefore, for any $y \in \mathbb{F}^t$, $\Pr_x[y = M_s x] = |ker\{M_s\}| / \left|\mathbb{F}^k\right| = \left|\mathbb{F}^{-t}\right|$. Thus, the random variable MU_k is r-wise independent. Since the noise is chosen independently from $x \overset{R}{\leftarrow} U_k$, the random variable $f_{M,\mu}(x)$ is also r-wise independent.

2. For any $\beta \in \mathbb{F}^m$ we let $s = \{i \in [m]; \beta_i \neq 0\}$. Based on the first item, if $|s| \leq r$ then $\mathbf{SD}(\langle \beta, f_{M,\mu}\rangle, U) = 0$. Hence we can assume $|s| > r$. However,

$$\mathbf{SD}(\langle \beta, f_{M,\mu}\rangle, U) \leq \Pr[\forall i \in s, \alpha_i = 0] \leq (1-\mu)^{|s|} \leq e^{-\mu \cdot r}$$

3. [Vio09] showed that the sum of b, identically independently distributed, distributions that ε-fools linear test, fools degree d polynomial test. Therefore, it suffices to show that the distribution $\{f_{M,\mu}(U_k)\}$ can be represented as the sum of d such distributions. We do so, by showing that it is actually the sum of b *Noisy Mappings* defined by the same matrix, M, and smaller noise parameter, μ'. We first note that for any μ' the sum of b *iid* copies of the noisy vector, $\alpha^{\mu'}, \mu' \in [01]$, distributed as in Definition 1.1, is a noisy vector α^{μ} for

$\mu = 1 - (1-\mu')^b \leq b\mu$. Hence there exists some $\mu' \geq \mu/b$ such that $\sum_{j=1}^{b} \alpha^{\mu',j} = \alpha^{\mu}$ where each $\alpha^{\mu',j}$ is an *iid* copy of $\alpha^{\mu'}$. Therefore,

$$\{f_{M,\mu}(U_k)\} \equiv \{\sum_{j=1}^{b} M \cdot U_k^{(j)} + \sum_{j=1}^{b} \alpha^{\mu',j}\} \equiv \{\sum_{j=1}^{b} f_{M,\mu'}(U_k^{(j)})\}$$

The main theorem of [Vio09] states that the sum of b *iid* Generators that ε-fools linear test ε_b-fools degree b test with $\varepsilon_b \doteq 16 \cdot \varepsilon^{1/2^{b-1}}$. Hence, the Noisy Linear Mapping ε-fools degree b-test with $\varepsilon = 16e^{\mu' r/2^{b-1}} \leq 16e^{\mu r/(b \cdot 2^{b-1})}$.

4. **Based on Theorem 5 in [HLV16]:** Theorem 5 in [HLV16] states that the sum of two distributions $\mathcal{D} + \mathcal{E}$ where $\mathcal{D}$ is r-wise independence and

$$\mathcal{E}_i = \begin{cases} 0 & \text{with probability } 1-\mu \\ \text{uniform field element} & \text{with probability } \mu \end{cases}.$$

ε-fools (n,t)-product test where $n < r$ with $\varepsilon = t(1-\mu)^{\Omega(r^2/m)}$.
Let $g_1, \ldots, g_t : \mathbb{F}^n \to \mathbb{C}_1$, for some n, t where $\mathbb{C}_1$ is the complex unit disk. We note that $f_{M,\mu}(U_k)$, can be written as the sum of two distributions $D, E \in \mathbb{F}^m$ where $D \equiv M \cdot U_k$ is r-wise independent (based on the first item) and for any $i \in [m]$:

$$E_i = \begin{cases} 0 & \text{with probability } 1-\mu \\ \text{uniform field element} & \text{with probability } \mu \end{cases}.$$

by definition. For $j \in [t]$ we let $H_j \subset [m]$ of size $|H_j| = n$ represent a division of $[m]$ into t disjoint n-size sets. Therefore, based on the Theorem 5 of [HLV16], we have that if $r \geq n$ then the t product test defined by $\{g_j, H_j\}_{j=1}^{t}$ is ε-fooled by the generator with $\varepsilon = t(1-\mu)^{\Omega(r^2/m)}$. Hence, the test is ε-fooled by $f_{M,\mu}$.

□

3.3 Over Large Field: Random Matrix has Large Dual-Distance

This subsection serves as another demonstration to the power of *Noisy Mappings over Large Fields.* In particular, we show that, over large field, a randomly sampled matrix $M \leftarrow \mathcal{M}(m,k,d,\mathbb{F})$ will have, with high probability, high *Dual Distance.* And hence, by the previous subsection, matrices sampled by this distribution would be resistance against various attacks. We again emphasize the fact that the above appears only for large fields (in a sense that we will now define).

Definition 3.8. *For a field $\mathbb{F}$ and (m,k,d)-hypergraph G we define a probability distribution $\mathcal{M}(G,\mathbb{F})$ over $m \times k$ matrices as follows: Take $M_{i,j}$ to be a fresh random non-zero field element if j appears in the i-th hyperedge of G; otherwise, set $M_{i,j}$ to zero. The uniform distribution over d-sparse $m \times k$ matrices over $\mathbb{F}$, denoted by $\mathcal{M}(m,k,d,\mathbb{F})$, corresponds to the experiment where G is sampled uniformly from $\mathcal{G}_{m,k,d}$ and then M is sampled according to $\mathcal{M}(G,\mathbb{F})$.*

Lemma 3.9. *Suppose that G is a (m, k, d)-hypergraph which is $(r, 1+\varepsilon)$-expanding and $|\mathbb{F}|^{\varepsilon} > m$. Then,*

$$\Pr_{M \leftarrow \mathcal{M}(G,\mathbb{F})}[\mathrm{dd}(M) < r] < |\mathbb{F}|^{-1} \sum_{t=1}^{r} \left(\frac{m}{|\mathbb{F}|^{\varepsilon}} \right)^{t}$$

Consequently, a random $M \leftarrow \mathcal{M}(m, k, d, \mathbb{F})$, with $|\mathbb{F}|^{\varepsilon} > m$ and $m = \Delta k$ has dual-distance of $r = \frac{k}{\Delta^{\frac{1}{d-2.1}}}$ with probability of at least $1 - O(\left|\mathbb{F}^{-1}\right|) - o(1)$. (see section 7) for concrete parameters.

We use the following claim:

Claim 3.10. *There exists $v \in \mathbb{F}^m$ with r non-zero values and $v'M = 0 \iff$ There exists $v \in \mathbb{F}^m$ with r non-zero values, $v'M = 0$ and $\sum_i u_i = 1$*

Proof. The first direction is trivial. Let $v \in \mathbb{F}^m$ be some vector with r non-zero values and assume that $v'M = 0$ we denote $u = \left(\sum_i v_i\right)^{-1} v$. We have that $u'M = 0$, u has at most r non-zero values and $\sum_i u_i = 1$. □

Proof of Lemma 3.9. Let $G \in \mathcal{G}_{m,k,d}$ be some $(r, 1+\varepsilon)$-expander. We upper bound the probability that, over a random sampling $M \leftarrow \mathcal{M}(G, \mathbb{F})$, there exists a non zero $v \in \mathbb{F}^m$ s.t. $v'M = 0$ and $|v| \leq r$. For any subset $s \subset [m]$ we denote $V_s = \{v; v_i \neq 0 \Leftrightarrow i \in s \sum_i v_i = 1\}$. Therefore,

$$\begin{aligned}
\Pr[\exists v' s.t\ v'M = 0\ \&\ |v| \leq r] &= \Pr[\exists v' s.t\ v'M = 0\ \&\ |s| \leq r\ \&\ v \in V_s] && (31)\\
&\leq \sum_{t=1}^{r} \sum_{V_s; |s|=t} \sum_{v \in V_s} \Pr_M[v'M = 0] && (32)\\
&\leq \sum_{t=1}^{r} \binom{m}{t} |V_s| \Pr_M[v'M = 0] && (33)\\
&\leq \sum_{t=1}^{r} \binom{m}{t} |\mathbb{F}|^{t-1} |\mathbb{F}|^{-t(1+\varepsilon)} && (34)\\
&\leq |\mathbb{F}|^{-1} \sum_{t=1}^{r} (\frac{m}{|\mathbb{F}|^{\varepsilon}})^{t} && (35)
\end{aligned}$$

Where the first transition follows from Claim 3.10, the transition in 33 follows from simply counting the number of options for choosing size t distinct output elements, and the transition in 34 results from calculating the size of V_s and the probability that the product $v'M$ would be zero and in case $|\mathbb{F}|^{\varepsilon} < m$ we obtain $O(|\mathbb{F}|^{-1})$. We further note that by Claim 2.10 a random $M \leftarrow \mathcal{M}(m, k, d, \mathbb{F})$ with $m = \Delta k$ is, with probability $1 - o(1)$, $(\frac{k}{\Delta^{\frac{1}{d-2.1}}}, 1)$-expanding. Therefore, in case $|\mathbb{F}|^{\varepsilon} > m$, such $M \leftarrow \mathcal{M}(m, k, d, \mathbb{F})$ has dual-distance of $r = \frac{k}{\Delta^{\frac{1}{d-2.1}}}$ with probability of at least $1 - O(\left|\mathbb{F}^{-1}\right|) - o(1)$. □

Corollary 3.11. *Let $M \leftarrow \mathcal{M}(m, k, d, \mathbb{F})$ then with probability at least $1 - \delta$, the noisy mapping $f_{M,\mu}$ is r-wise independent and ε-fools degree-d distinguishers with $r = \left(\frac{k}{\Delta^{\frac{1}{d-2.1}}}\right)$, $\varepsilon = 16e^{\frac{-\mu r}{b2^{b-1}}}$,and $\delta = O(\left|\mathbb{F}^{-1}\right|) + o(1)$.*

4 The Bias of Goldreich's functions

4.1 Linear Attack for Quadratic Stretch

We consider the following linear attack against Goldreich's functions with $m = \Omega(k^2)$ (which generalizes the attack of [AL16]):

```
input : a sum-product Goldreich Function $G = G_\Sigma \bigcup G_\Pi$ where $G_\Sigma$ and $G_\Pi$ are
        $(m, k, a)$ and a $(m, k, b)$ hypergraphs, respectively. A vector $y \in \mathbb{F}^m$
output: 1 iff $y \in Im\left\{f_{G,\mathsf{SP}_{a,b}}\right\}$
Let $M$ be the $(m, k, b)$-matrix where $M_{i,j} = 1$ if the $i$-th hyperedge of $G_\Sigma$ contains the
  $j$-th vertex;
Find $i \in [k]$ which participates in at least $k + 1$ hyperedges in $G_\Pi$;
Let $L_i \subset [m]$ denote the hyperedges of $G_\Pi$ that contain the $i$-th vertex;
Solve $v \cdot M|_{L_i} = 0$ for $v \in \mathbb{F}^{|L_i|}$;
if $\langle y|_{L_i}, v\rangle = 0$ then
    output: 1
else
    output: 0
end
```

Algorithm 2: Linear Attack for Goldreich's functions

Claim 4.1. *The running time of Algorithm 2 is $O\left(k^3\right)$ (arithmetic operations)*

Proof. Results immediately from the size of the matrix M, $m \times k$, and the complexity of preforming Gaussian Elimination. □

Claim 4.2. *For $y \leftarrow \mathbb{F}^m$ Algorithm 2 output 1 with probability $\frac{1}{|\mathbb{F}|}$.*

Proof. The inner product of a non-zero vector $v \in \mathbb{F}^m$ and a randomly distributed $y \leftarrow \mathbb{F}^m$ is randomly distributed over $\mathbb{F}^m$ □

Claim 4.3. *If $m = \Omega\left(k^2\right)$ and $y \in Im\left\{f_{G,\mathsf{SP}_{a,b}}\right\}$ then Algorithm 2 output 1 with probability at least $\frac{2|\mathbb{F}|-1}{|\mathbb{F}|^2}$*

Proof. $m = \Omega\left(k^2\right)$ and hence, there exists a vertex of G_Π which is contained in at least $k + 1$ different hyperedges and so the L_i columns of M are linearly dependent. Therefore, the algorithm would find $v \in \mathbb{F}^{|L_i|}$ such that $v \cdot M|_{L_i} = 0$ meaning that:

$$\Pr\left[\langle y_{L_i}, v\rangle = 0\right] = \Pr\left[\sum_{l\in L_i} v_l \left[\sum_{j\in S_l} x_j + \prod_{j\in T_l} x_j\right] = 0\right] \tag{36}$$

$$= \Pr\left[\left[\sum_{l\in L_i} v_l \sum_{j\in S_l} x_l\right] + \left[\sum_{l\in L_i} v_l \prod_{j\in T_l} x_j\right] = 0\right] \tag{37}$$

$$= \Pr\left[\sum_{l\in L_i} v_l \prod_{j\in T_l} x_j = 0\right] \tag{38}$$

$$= \Pr\left[x_i \sum_{l\in L_i} v_l \prod_{j\in T_l\setminus x_i} x_j = 0\right] \tag{39}$$

$$= \Pr[x_i = 0] + \Pr[x_i \neq 0]\Pr\left[x_i \sum_{l\in L_i} v_l \prod_{j\in T_l\setminus x_i} x_j = 0 \middle| x_i \neq 0\right] \tag{40}$$

$$\geq \frac{1}{|\mathbb{F}|} + (1 - \frac{1}{|\mathbb{F}|})\frac{1}{|\mathbb{F}|} = \frac{2\,|\mathbb{F}| - 1}{|\mathbb{F}|^2} \tag{41}$$

The transition to 38 follows from the fact that $v \cdot M|_{L_i} = 0$, and the transition to 39 follows from the fact that $v_l \neq 0$ only for $l \in L_i$ meaning that each T_l contains the vertex i. The last inequality follows by noting that, for $x_i = \alpha \neq 0$, the polynomial $\alpha \sum_{l=1}^{m} v_l \prod_{j\in T_l\setminus x_i} x_j$ is a multilinear polynomial, and therefore it takes the value zero with probability at least $1/|\mathbb{F}|$. (This will be proved later in Claim 4.8.) □

Corollary 4.4. *For any $f_{G,\mathsf{SP}_{a,b}} : \mathbb{F}^k \to \mathbb{F}^m$, such that $m = \Omega(k^2)$ there exist a linear attack with distinguishing advantage of $\frac{2|\mathbb{F}|-1}{|\mathbb{F}|^2}$*

4.2 Small Bias Over Large Field

We prove the following theorem:

Theorem 4.5. *Let G be a (n, m, d) hypergraph with the property that every pair of hyperedges intersect in less than b vertices. Let $P : \mathbb{F}^d \to \mathbb{F}$ be any multilinear polynomial of degree at least b. Then, $f_{G,P} : \mathbb{F}^k \to \mathbb{F}^m$ ε-fools linear tests with $\varepsilon = \frac{k-1}{|\mathbb{F}|}$.*

Remark 4.6. *Taking P to be the (a, b)-sum product polynomial, and choosing G to be a random (m, k, d)-graph with $m = k^{b/2}$, we derive a low-bias generator (with $\varepsilon = \frac{k-1}{|\mathbb{F}|}$) with high probability. This can be improved to $m = O(k^{b/2+1/2})$ by sampling G at random and then removing "bad hyperedges" which intersect with other hyperedges in more than b vertices. It is not hard to see that the number of removed hyperedges is likely to be small $o(n)$.*

Fix some non-trivial linear test $\alpha \in \mathbb{F}^m$, and observe that the polynomial $g(x) = \langle \alpha, f_{G,P}(x)\rangle$ is a non-trivial multi linear polynomial of degree b. (Since each pairs of outputs of f share at most $b - 1$ variables, degree-b monomials do not cancel out.) Theorem 4.5 therefore follows from the following lemma.

Lemma 4.7. *Let $Q : \mathbb{F}^n \to \mathbb{F}$ be multi-linear polynomial then* $\mathbf{SD}(Q(U_n), U) \leq \frac{n-1}{|\mathbb{F}|}$.

Before proving the lemma, we need the following claim.

Claim 4.8. *Let $Q : \mathbb{F}^n \to \mathbb{F}$ be some non-trivial multi-linear polynomial over a field $\mathbb{F}$ of cardinality p. Then $\forall z \in \mathbb{F}$ we have that*

$$\Pr[Q(U_n) = z] \geq \frac{1}{p} \left(\frac{p-1}{p} \right)^{n-1}.$$

The claim is tight for the degree n multilinear monomial $\prod_{i \in [n]} x_i$.

Proof. The proof is by induction on the number of variables n. The basis is trivial since the only relevant polynomial is $x_1 + c$ for a constant $c \in \mathbb{F}$. We prove the induction step. Assume, without loss of generality, that Q depends on x_n. Therefore, Q can be written as $Q(x_1, \ldots, x_n) = x_n Q_1(x_1, \ldots x_{n-1}) + Q_2(x_1, \ldots x_{n-1})$ where $Q_1, Q_2 : \mathbb{F}^{n-1} \to \mathbb{F}$ are multi-linear polynomials and Q_1 is non-trivial. Observe that, for any fixing of $(x_1, \ldots, x_{n-1})$ which does not zero Q_1, the distribution of the polynomial $Q(x)$ (induced by the choice of x_n) is uniform over $\mathbb{F}$. It follows that for any $z \in \mathbb{F}$ we have that

$$\Pr[Q = z] \geq \frac{1}{p}(1 - \Pr[Q_1 = 0]).$$

By the induction hypothesis, $\Pr[Q_1 = 0] \leq 1 - \frac{p-1}{p}(\frac{p-1}{p})^{n-2} = 1 - (\frac{p-1}{p})^{n-1}$, and so $\Pr[Q = z]$ is at least $\frac{1}{p}(\frac{p-1}{p})^{n-1}$, as required. □

We proceed with a proof of Lemma 4.7.

Proof. Let $p = |\mathbb{F}|$. Observe that

$$\begin{aligned} \mathbf{SD}(Q(U_n), U) &= \sum_{z : Pr[U=z] \geq \Pr[Q(x)=z]} |\Pr[U = z] - \Pr[Q(x) = z]| \\ &\leq p \cdot \max_{z \in \mathbb{F}} \left(\frac{1}{p} - \Pr[Q(U_n) = z] \right) \end{aligned}$$

By Claim 4.8, the weight of z under $Q(U_n)$ is at least

$$\frac{1}{p} \left(\frac{p-1}{p} \right)^{n-1} \geq \frac{1}{p} \left(1 - \frac{n-1}{p} \right).$$

Hence, the statistical distance is upper-bounded by

$$p \cdot \frac{1}{p} \left(\frac{n-1}{p} \right) = \left(\frac{n-1}{p} \right),$$

and the lemma follows.

□

Combining the results above with the known relation between statistical distance and characteristic distance [BV10], we derive the following statement (which will be useful later).

Corollary 4.9. *Let $Q : \mathbb{F}^d \to \mathbb{F}$ be multi linear polynomial then* $\mathbf{CD}\left(Q(U_n), U\right) \leq \min\{\frac{2d}{|\mathbb{F}|}, 1 - e^{-\frac{d}{|\mathbb{F}|}}\}$.

4.3 Achieving Low Bias for Subquadratic Stretch over every field

Theorem 4.10. *For any $\delta > 0$, $a > 4$,$b \geq 8$ output length $m = O(k^{2-\delta})$, and prime field $\mathbb{F}$, with high probability over the choice of a randomly sampled $G \leftarrow \mathcal{G}(m, k, a+b)$-hypergraph*[7]*, the function $f_{G,\mathsf{SP}_{a,b}} : \mathbb{F}^k \to \mathbb{F}^m$ ε-fools any linear test with $\varepsilon = \exp(\Omega(k^{\min(\delta/2,\delta-\frac{2}{a-2.1})}))$.*

Our proof makes use of the following definition.

Definition 4.11 ((α, β)-ruled)**.** *A (k, m, d)-hypergraph is* (α, β)-ruled *if any subset of the hyperedges $K \subset [m]$ of size at least $|K| \geq \alpha$ has a* ruling subset *$R \subset K$ of size at least $|R| \geq \beta$ such that:*

1. *No two hyperedge $e_1, e_2 \in R$ intersect.*
2. *Any hyperedge $e \in K \setminus R$ share at most $d-1$ vertices with all the vertices that touch R, i.e., $\bigcup_{e' \in R} e'$.*

If the second item is substituted with the condition that for each $e \in K \setminus R$ and $e_1, e_2 \in R$ if $e \cap e_1 > 0$ then $e \cap e_2 = 0$, we say that the graph is (α, β)-strongly ruled *and that R is a* strongly ruling subset*.*

The following lemma (whose proof appears in Section 4.3.2) shows that a random hypergraph is likely to satisfy the above definition.

Lemma 4.12. *For parameters $d \geq 8$, $\delta > 0$ and $m = O(\Delta k)$ with $\Delta = k^{1-\delta}$, we let $G \overset{R}{\leftarrow} \mathcal{G}(m, k, d)$. Then, with probability $1 - o(1)$, the graph is (α, β)-ruled for any α, β which satisfy $\beta \geq \min\left\{k^{\delta/2}, \frac{\alpha}{2\Delta d^2}\right\}$.*

The proof of Theorem 4.10 is based on the following key lemma whose proof is deferred to Section 4.3.1.

Lemma 4.13. *Let $f_{G,\mathsf{SP}_{a,b}} : \mathbb{F}^k \to \mathbb{F}^m$ be the sum-product Goldreich's functions and $\mathbb{F}$ any prime field. Assume that G_Σ and G_Π are (m, k, a) and (m, k, b)-hypergraphs with the following properties:*

1. *G is $(r, d - a/2)$-expanding, i.e., any set of hyperedges, of size $s < r$, contains at least $s \cdot (d - a/2)$ distinct vertices.*
2. *G_Π is (r, β)-ruled.*

Then the following holds:

1. *For any $\alpha \in \mathbb{F}^m$ of Hamming weight at most r, the random variable $\langle \alpha, f_{G,\mathsf{SP}_{a,b}}(x) \rangle$ is uniformly distributed over $\mathbb{F}$.*
2. *The generator $f_{G,\mathsf{SP}_{a,b}}$ $4b^2 c_b^{\frac{\beta}{2^b}-1}$-fools every linear test where*

$$c_b \leq \min\left\{\frac{3b^2}{|\mathbb{F}|}, 1 - e^{-\frac{b^2}{|\mathbb{F}|}}\right\} \leq 1 - e^{-b^2/2} < 1.$$

[7] Throughout the paper, an event happens with high probability if it happens with probability which converges to 1 when k goes to infinity.

Theorem 4.10 follows immediately from Lemmas 4.12 and 4.13.

Proof of Theorem 4.10 (assuming Lemmas 4.12 and 4.13). Let $f_{G,\mathsf{SP}_{a,b}} : \mathbb{F}^k \to \mathbb{F}^m$ be the sum-prod Goldreich Function, where G_Σ and G_Π are (m,k,a) and (m,k,b) hypergraphs and $a > 4, b \geq 8$. We wish to show that $f_{G,\mathsf{SP}_{a,b}}$ meets the conditions of Lemma 4.13. By Claim 2.10, with high probability, G is $(r, d{-}a/2)$-expanding for $r = k^{1-\frac{2}{d-2.1}}$. Also, by Lemma 4.12, with high probability, G_Π is (r,β)-ruled, for

$$\beta \geq \min\left\{k^{\delta/2}, \frac{r}{2\Delta d}\right\} \geq \min\left\{k^{\delta/2}, \frac{k^{1-\frac{2}{d-2}}}{2k^{1-\delta}d}\right\} = \Omega(k^{\min(\delta/2,\delta-\frac{2}{a-2.1})})$$

Therefore, by Lemma 4.13 it fools every linear test with bias exponential in β. □

4.3.1 Proof of Lemma 4.13

Let $f_{G,\mathsf{SP}_{a,b}} : \mathbb{F}^k \to \mathbb{F}^m$ be a *sum-product* Goldreich's functions, with the following properties:

1. G is $(r, d-a/2)$-expanding.

2. G_Π is (r,β)-ruled.

The first item of our proof was shown in [ABR12] meaning that for any $\alpha \in \mathbb{F}^m$ such that α has an Hamming weight of at most r we have that $\langle \alpha, f_{G,\mathsf{SP}_{a,b}}(x)\rangle$ is uniformly distributed over $\mathbb{F}$. Hence, we may assume that the set of indexes where $\alpha_i \neq 0$, denoted K, is of size at least $|K| \geq r$.

Consider the multiplication hypergraph, G_Π. It is (r,β)-ruled and therefore it has a ruling subset $R \subset K$ of size at least β. Let S_i denote the vertices of the i'th hyperedge and $S = \bigcup_{i\in R} S_i$ the set of vertices which touch some hyperedge in R. Since R is a ruling subset $\{S_i\}_{i\in R}$ are disjoint subsets of size $|S_i| = d$. We can now write the linear test as:

$$\langle \alpha, f_{G,\mathsf{SP}_{a,b}}(x)\rangle = \sum_{i\in K} \alpha_i x^{S_i} + L(x)$$

where x^T stands for the monomial $\prod_{i\in T} x_i$ and $L(x)$ is some linear function (obtained from the additive part of the sum-product polynomial). The above can be partitioned into

$$\sum_{i\in R} \alpha_i x^{S_i} + \sum_{i\in K\setminus R} \alpha_i x^{S_i} + L(x).$$

Let $T(x) = \sum_{i\in K\setminus R} \alpha_i x^{S_i} + L(x)$. The degree of T is at most $b-1$ in S's variables (since hyperedges outside R touch at most $b-1$ vertices in S). Hence, by Lemma 2.18 we have that:

$$\mathbf{CD}\left(\langle \alpha, f_{G,\mathsf{SP}_{a,b}}(x)\rangle, U\right) \leq U_b(\prod_b)^{\frac{\beta}{2^b}}$$

where $U_b(\prod_b)$ stands for the degree b Gowers uniformity of the degree b monomial $\prod_b$. Since the b'th directional derivative of $\prod_b$ is a multi-linear polynomial of degree b in $(b{+}1)b$ variables. $U_b(\prod_b)$

can be written as the degree-1 Gowers Uniformity of a multi-linear polynomial in $b^2 + b$ variables and hence using Corollary 4.9 it can be upper bounded by

$$U_b(\prod_b) \leq \min\left\{1 - e^{-\frac{b^2}{|\mathbb{F}|}}, \frac{3 \cdot b^2}{|\mathbb{F}|}\right\} \leq 1 - e^{-b^2/2}$$

Finally, we note that based on Claim 2.15 the relation between the Statistical Distance to the Characteristic Distance corresponds to $\sqrt{|\mathbb{F}|}$. Therefore:

1. In case $|\mathbb{F}| \geq 16b^4$:

$$\mathbf{SD}\left(\langle \alpha, f_{G,\mathsf{SP}_{a,b}}(x)\rangle, U\right) \leq \sqrt{|\mathbb{F}|} U_b(\prod_b)^{\frac{\beta}{2^b}} \leq \frac{3}{4}\left(\frac{3 \cdot b^2}{|\mathbb{F}|}\right)^{\frac{\beta}{2^b}-1}$$

2. In case $|\mathbb{F}| \geq 16b^4$:

$$\mathbf{SD}\left(\langle \alpha, f_{G,\mathsf{SP}_{a,b}}(x)\rangle, U\right) \leq \sqrt{|\mathbb{F}|} U_b(\prod_b)^{\frac{\beta}{2^b}} \leq 4b^2(1 - e^{-b^2/2})$$

And in both cases $\mathbf{SD}\left(\langle \alpha, f_{G,\mathsf{SP}_{a,b}}(x)\rangle, U\right) \leq 4b^2 c_b^{\frac{\beta}{2^b}-1}$ Where:

$$c_b \leq \min\left\{1 - e^{-\frac{b^2}{|\mathbb{F}|}}, \frac{3 \cdot b^2}{|\mathbb{F}|}\right\} \leq 1 - e^{-b^2/2}$$

. □

4.3.2 Proof of Lemma 4.12

Let $G \overset{R}{\leftarrow} \mathcal{G}(m, k, d)$ where $d \geq 8$, $\delta > 0$ and $m = O(\Delta k)$ with $\Delta = k^{1-\delta}$. We condition on the event that:

1. Any pair of hyperedges in G intersect in at most 3 vertices;
2. Each hyperedge intersects with at most $2\Delta d^2$ other hyperedges;
3. The graph is $(r, d - \frac{5}{2})$-expanding for $r = \frac{k}{\Delta} = k^\delta$

By Claims 2.13 and 2.10, this holds with high probability over the choice of G. For any subset K of G's hyperedges we identify a ruling subset R using Algorithm 3.

We proceed by proving that R is indeed a ruling-set and analyzing the halt point of Algorithm 3 in order to lower bound the size of $|R|$.

Claim 4.14. *R is a ruling subset of K.*

Proof. At the initialization R is a (trivial) ruling subset of K. We argue that this invariant is preserved. Indeed, let e^* be an hyperedge which is chosen at the beginning of an iteration. Since e^* is not a member of E_1, it does not intersect with any other $e \in R$. To see that e^* does not violate the second property of a ruling subset, consider some $e' \notin R$ and let $i < d$ be the size of the intersection between e' and the set of nodes which participate in R. Either $i \geq d - 3$, and so e^* and e' do not intersect (since $e^* \notin E_2$), or $i < d - 3$ and, by assumption, $|e^* \cap e'| \leq 3$. In both cases, the intersection between e' and $e^* \cup (\bigcup_{e \in R} e)$ is at most $d - 1$ as required. □

```
input  : Some $(m, k, d)$-hypergraph, $G$, and a subset of the hyperedges $K \subset [m]$
output: A ruling subset $R \subset K$
Initialize $E_1, E_2, R = \emptyset$;
while $K \setminus (E_1 \cup E_2)$ is non-empty do
    Choose some $e \in K \setminus (E_1 \cup E_2)$ and add it to R;
    Add to $E_1$ any hyperedge that intercuts $e$ ;
    Add to $E_2$ any hyperedge which intersects with an hyperedge that shares $d - 3$ or
      more vertices with $R$'s hyperedges;
end
output: R
```

Algorithm 3: Finding Ruling Subset

Since we assumed that each hyperedge intersects with at most $2\Delta d^2$ other hyperedges, we conclude that the size of E_1 is at most $|E_1| < |R| \, 2\Delta d^2$. The next claim shows that unless R is large, E_2 cannot be too large.

Claim 4.15. *If $|R| < k^{\delta/2}$ and $d \geq 8$ then the size of E_2 is at most $2\Delta d^2 \, |R|$.*

Proof. Assume that $|R| < k^{\delta/2}$ and let W be the set of hyperedges that share more than $d - 3$ vertices with R (*i.e.,* the set E_2 is the set of hyperedges that intersect with some hyperedge in W). We upper-bound W using the expansion properties of the graph. Observe that the set $T = W \cup R$ of hyperedges touches only $d \, |R| + 3 \, |W|$ vertices. Therefore, if W is larger than R, then T does not expand by a factor of $(d + 3)/2 \leq d - \frac{5}{2}$. (The inequality follows since $d \geq 8$). Since $|R| < k^{\delta/2}$ this violates the expansion properties of the hypergraph (property 3 above). Finally, based on our second assumption, it follows that $|E_2| \leq 2\Delta d^2 \, |W| \leq 2\Delta d^2 \, |R|$. □

Overall, R is a ruling set of size at least $|R| \geq min \left\{ k^{\delta/4}, \frac{|K|}{2\Delta d^2} \right\}$. The proof of Lemma 4.12 follows. That is, with probability $1 - o(1)$, the graph is (α, β)-ruled for any α, β which satisfy $\beta \geq \min \left\{ k^{\delta/2}, \frac{\alpha}{2\Delta d^2} \right\}$. □

5 Goldreich's functions has no low degree annihilating polynomial

5.1 Finding Annihilating Polynomial for Shrinking Set

We begin by noting that over a large field an annihilating polynomial forms a good distinguisher.

Remark 5.1. *Let $f : \mathbb{F}^n \to \mathbb{F}^m$ be a function and let $Q : \mathbb{F}^m \to \mathbb{F}$ be an f-annihilating polynomial. By Schwartz-Zippel Lemma,* $\Pr_x[Q(f(x)) = 0] - \Pr_y[Q(y) = 0] \geq 1 - \frac{\deg(Q)}{|\mathbb{F}|}$.

Later, in the next subsection, we will later show that as long as G is a good expander $f_{P,G}$ has no low degree, i.e., $O(1)$, annihilating polynomial. In this subsection we further investigate the feasibility of attacks which are based on annihilating polynomials. In particular, we describe a simple general algorithm that in time $\text{poly}(m^\ell)$ finds, with high probability, an annihilating polynomial of degree ℓ if such exists.

```
input : A degree parameter $\ell$, number of iterations $L$
output: $Q : \mathbb{F}^m \to \mathbb{F}$ of degree $\ell$ such that $Q \circ f = 0$ or abort
for $i = 1$ to $L$ do
    Sample $x^i \leftarrow \mathbb{F}^k$;
    Compute $y^i = f(x^i)$;
    Compute the vector $v_i$ which is all the monomials of $y^i$ of degree at most $\ell$;
    i.e., $v$ is of length $N = \sum_{\ell' < \ell} \binom{m+\ell'}{\ell'}$;
end
Let $V$ be the $L \times N$ matrix whose $i$-th row is $v_i$;
if $V$'s columns are independent then
    abort
else
    Solve $Vu = 0$ and $u \neq 0$ using Gaussian Elimination;
    output: $Q : \mathbb{F}^m \to \mathbb{F}$ the degree $\ell$ polynomial whose coefficients are $u$;
end
```

Algorithm 4: Finding Annihilating Polynomial for a degree-b function $f : \mathbb{F}^k \to \mathbb{F}^m$.

Claim 5.2. *If there exist an annihilating polynomial of degree at most ℓ then the algorithm outputs such an annihilating polynomial, except with probability $|\mathbb{F}|^{m^\ell} \cdot \left(\frac{\ell b}{|\mathbb{F}|}\right)^L$.*

Proof. Suppose that Q is an annihilating polynomial of degree ℓ. Then $Q(y^i) = 0$ for every $i \in [L]$, and so the coefficient vector of Q forms a valid non-zero solution for the system $Vu = 0$. We proceed by upper bounding the probability that the algorithm outputs a polynomial $Q' \neq 0$ which is not an annihilating polynomial. Indeed, by union bound, the probability that there exists a degree-ℓ non-zero polynomial Q' which is not f-annihilating but is zeroed over all y^i is at most

$$\begin{aligned}\sum_{Q'} \Pr_{x=(x^i)_{i\in[L]}}[Q'(f(x^i)) = 0 \textit{ for all } i] &\leq \sum_{Q'} \prod_{i=1}^{L} \Pr_{x^i}[Q'(f(x^i)) = 0] \\ &\leq |\mathbb{F}|^{m^\ell} \cdot \left(\frac{\ell b}{|\mathbb{F}|}\right)^L,\end{aligned}$$

where the sum ranges over all degree-ℓ non-zero polynomials Q' which are not f-annihilating, and the last inequality follows from the Schwartz-Zippel Lemma by using the fact that $Q' \circ f$ is just a degree-ℓb polynomial in the x's. □

Assuming that $\ell b \leq |\mathbb{F}|$, we can take L to be $m^\ell + 1$ and get an error of $1/|\mathbb{F}|$. For this setting (and assuming that f is computable in polynomial time) the complexity of the algorithm is dominated by the Gaussian elimination step whose cost is $O(m^{3\ell})$ arithmetic operations.

Corollary 5.3. *If a polynomial-time function $f : \mathbb{F}^k \to \mathbb{F}^m$ of degree b has annelaiting annihilating polynomial of degree ℓ where $\ell b < |\mathbb{F}|$, then in time $O(m^{3\ell})$ one can find such an annealing polynomial with high probability.*

5.2 Lower-bounds on the degree of Annihilating Polynomials

Theorem 5.4. *Let $f_{G,\mathsf{SP}_{a,b}} : \mathbb{F}^k \to \mathbb{F}^m$ be a* sum-product *Goldreich's functions and assume that G_Π is $(2r+1, b/2+1)$-expanding. Then there is no $f_{G,\mathsf{SP}_{a,b}}$-annihilating polynomial of degree smaller than r.*

The proof is based on the following more general lemma (which is based on an argument of Bogdanov [Bog05]).

Lemma 5.5. *Let H is some $m \times k$-matrix over $\mathbb{F}$ with $\mathrm{dd}(H) = 2 \cdot r + 1$. For $i \in [m]$, let $L_i : \mathbb{F}^k \to \mathbb{F}$ be a polynomial of degree smaller than $\sum_{j=1}^m H_{i,j}$. Consider the function $f : \mathbb{F}^k \to \mathbb{F}^m$ whose i-th output is defined by $f(x)_i = \prod_{j=1}^m x_j^{H_{i,j}} + L_i(x)$. Then there is no f-annihilating polynomial of degree smaller than r.*

Proof of Theorem 5.4 based on Lemma 5.5. Consider the function $f_{G,\mathsf{SP}_{a,b}} : \mathbb{F}^k \to \mathbb{F}^m$ and assume that G_Π is $(2r+1, b/2+1)$-expanding. We let H represent the dependency matrix of G_Π and obviously it is $(2r+1, b/2+1)$-expanding. We can write the i-th output of $f_{G,\mathsf{SP}_{a,b}}$ as

$$\prod_{j=1}^m x_j^{H_{\Pi,i,j}} + L_i(x),$$

where L_i is linear. We also have that $\mathrm{dd}(H) = 2r+1$ since it is $(2r+1, b/2+1)$-expanding and that:

$$deg(L_i) = 1 < b = \sum_{j=1}^m H_{i,j}$$

and therefore by Lemma 5.5 it has no degree r annihilating polynomial. □

Proof of Lemma 5.5. We follow the line of the proof presented in [Bog05]. Consider some degree-r monomial A over the output variables $y = (y_1, \ldots, y_m)$, that is $A = y_1^{\alpha_1} \cdots y_m^{\alpha_m}$ such that $\sum_{i=1}^m \alpha_i \leq r$. Preform the substitution $y_j \rightsquigarrow f(x)_j$ and consider the resulting polynomial A' over the x-variables. The polynomial A' contains the monomial $\prod_{i=1}^n x_i^{\sum_{j=1}^m H_{i,j}\alpha_j}$. Moreover, since $deg(L_i) < \sum_{j=1}^m H_{i,j}$, it is the highest degree monomial that A' contains.

Now assume towards a contradiction that $T : \mathbb{F}^m \to \mathbb{F}$ is a degree-r annihilating polynomial. Among all the monomials of T, take $A = y_1^{\alpha_1} \cdots y_m^{\alpha_m}$ to be the monomial for which the resulting A' has the highest total degree. Since $T \circ f$ is the zero polynomial, T must contain some monomial, $B = y_1^{\beta_1} \cdots y_m^{\beta_m}$ with $\sum_{i=1}^m \beta_i \leq r$ such that under the described substitution, B' contains the monomial $\prod_{i=1}^n x_i^{\sum_{j=1}^m H_{i,j}\alpha_j}$. However, since A' has the highest total degree we obtain

$$\sum_{i,j} \beta_j deg(L_i) < \sum_{i,j} \beta_j H_{i,j} \leq \sum_{i,j} \alpha_j H_{i,j}$$

meaning that it is possible only if for any $i \in [n]$ we have that:

$$\sum_{j=1}^m H_{i,j}(\alpha_j - \beta_j) = 0$$

Define a vector $v \in \mathbb{F}^m$ by $v_j = \alpha_j - \beta_j (\mod |\mathbb{F}|)$. By construction, $v \neq 0$, moreover, v can't have Hamming weight of more than $2 \cdot r$. However, $v' \cdot H = 0$ which contradicts the fact that $\mathrm{dd}(H) = 2 \cdot r + 1$. □

6 Fooling Sparse Polynomials

In this section we show that the sum-product Goldreich's functions also fools sparse polynomials. Recall that a polynomial $T : \mathbb{F}^m \to \mathbb{F}$ is (q,t) *sparse* if it has a degree q and each of its inputs appears in at most t monomials.

Theorem 6.1. *For every $a, b \geq 4$ and $0 < \rho < 0.5\frac{a-2.1}{a-1.1}$, $m = k^{1+\rho}$ and every constants q and t, with high probability over $G \leftarrow \mathcal{G}(m,k,a+b)$, the function $f_{G,\mathsf{SP}_{a,b}} : \mathbb{F}^k \to \mathbb{F}^m$ ε-fools any (q,t) sparse polynomial $T : \mathbb{F}^m \to \mathbb{F}$ with $\varepsilon = \gamma^{\Omega(k^{1-2\rho\frac{a-1.1}{a-2.1}})}$ where $\gamma < 1$ stands for the maximal bias of a non-trivial polynomial over $\mathbb{F}$ in qd variables of degree at most qb and where each variable is of degree at most q.*

Remark 6.2. *We note that for constant size fields γ is a constant (which depends on a, b, q and t). We believe that the bias can be upper-bounded for large fields as well, though our proof falls short of proving this.*

Roughly speaking, we show that a sparse polynomial can be viewed as a linear combination of distinct variables bounded degree polynomials. Formally, fix some (q,t) sparse polynomial T which outputs the sum of m_T monomials $M_1, \ldots, M_{m_T}$. Consider the function $T' : \mathbb{F}^m \to \mathbb{F}^{m_T}$ which maps $y \in \mathbb{F}^m$ to outputs $(M_1(y), \ldots, M_{m_T}(y))$ and let $f'(x) = T'(f_{G,\mathsf{SP}_{a,b}}(x))$. Now we can view the test T as a linear test over f'. Hence, it suffices to argue that f' fools the corresponding linear tests. To prove this we combine the arguments from the previous two sections.

For a $(m, k, d = a+b)$ hypergraph G, we consider the dependency hypergraph, G_T, of the function $f'(x) = T'(f_G, \mathsf{SP}_{a,b}(x))$. Namely, G_T contains k vertices and m_T hyperedges each of cardinality of at most qd. We similarly define $G_{\Pi T}$ as dependency hypergraph that corresponds to the function $T'(f_{G_\Pi,\Pi_b}(x))$. We begin with few simple observations regarding the structure of G_T.

Claim 6.3. *With high probability over $G \leftarrow \mathcal{G}(m,k,a+b)$, for any (q,t)-sparse polynomial $T : \mathbb{F}^m \to \mathbb{F}$ the hypergraph G_T satisfies the following properties:*

1. *$m_T \leq tm$.*

2. *Each hyperedge of G_T contains at most qd vertices. Correspondingly, each hyperedge of $G_{\Pi,T}$ contains at most qb hyperedges.*

3. *Each vertex of G_T participates in at most $t2k^\rho d$ hyperedges. Correspondingly, each vertex of $G_{\Pi,T}$ participates in at most $t2k^\rho b$ hyperedges.*

4. *Every hyperedge of G_T intersects with at most $tq2k^\rho d^2$ hyperedges. Correspondingly, every hyperedge of $G_{\Pi,T}$ intersects with at most $tq2k^\rho b^2$ hyperedges.*

Proof. The claim follows immediately from Claim 2.13 however we elaborate on each of these items separately:

1. Since each of G's hyperedges appear at most t times the overall number of hyperdeges in G_T must be upper bounded by m_T

2. Trivial.

3. Based on the second item of Claim 2.13 each of G's vertices appear in at most $2k^{\rho}d$ hyperedges in G. Since each of G's hyperedges appear in at most t of G_T's hyperedges the claim follows.

4. Each hyperedge of G_T consists of at most q hyperedges in G, each contains at d vertices. Combined with the previous item the claim follows.

□

Moreover, we note that by Claim 2.10 we have that:

Claim 6.4. *the following hold with high probability:*

1. G_{Π} is $(2q+1, b/2+1)$-expanding.

2. G is $(r, d-\frac{a}{2})$-expanding for $r = k^{1-\frac{2\rho}{a-2.1}}$.

6.1 Proof of Theorem 6.1

In the following we condition on the event that G_T and G satisfy the properties listed in Claims 6.3 and 6.4. Let m_T be the number of monomials in T. We first note that if $m_T \leq k^{1-\frac{2\rho}{a-2.1}}/q$ then T depends on at most $r = k^{1-\frac{2\rho}{a-2.1}}$ output variables of $f_{G,\mathsf{SP}_{a,b}}$. In this case, the function $f_{G,\mathsf{SP}_{a,b}}$ perfectly fools T. Indeed, by Claim 6.3 G is $(r, d-\frac{a}{2})$-expanding, and so, by Lemma 4.13, $f_{G,\mathsf{SP}_{a,b}}$ is r-wise independent.

Let us therefore assume from now on that $m_T = \Omega(k^{1-\frac{2\rho}{a-2.1}})$. Let $M_i(y)$ be the i-th monomial of T and let $M'_i(x)$ denote the polynomial obtained by substituting y_j by the j-th output of $f = f_{G,\mathsf{SP}_{a,b}}(x)$. The idea is to find a subset of input variables $S \subset [k]$ such that after fixing all variables outside S, the test $T(f(x))$ decomposes into $\ell \geq \Omega(k^{1-2\rho\frac{a-1.1}{a-2.1}})$ non-zero polynomials $T_i(x_{S_i})$ over pair-wise disjoint sets of x-variables $S_1, \ldots, S_\ell$. Over a fixed-size field, the bias of such a sum is exponentially small in ℓ.

We construct the polynomials T_i and the set of variables S via the following greedy process. Start with an empty S and a list L of all the monomials of T. At the i-th step, choose a monomial $M(y) \in L$ of maximal degree (hereafter referred to as the i-th leader), let e be the hyperedge that corresponds to M in G_T. We collect all hyperedges of distance at most 2 of e. That is, let $J_i \subset [m_T]$ be the set

$$\left\{ j \in [m_T] : e_j \cap e \neq \emptyset \bigvee \left(\exists e' \text{ s.t. } e' \cap e_j \neq \emptyset \bigwedge e' \cap e \neq \emptyset \right) \right\},$$

where e_j is the i-th hyperedge of G_T. Let $T_i(x) = M'(x) + \sum_{j \in J_i} M'_j(x)$ and remove from the list L the monomial M and all the monomials $M_j, j \in J_i$. Denote e by S_i and update $S = S \cup S_i$. Let ℓ denote the total number of polynomials T_i. By construction, the sets S_i's are pairwise disjoint. Also, in each step, we remove at most $2(tq2k^{\rho}d^2)^2$ (due to the properties of G_T), hence

$$\ell = \Omega(m_T/(tq2k^{\rho}d^2)^2) = \Omega(k^{1-\frac{2\rho}{a-2.1}-2\rho}) = \Omega(k^{1-2\rho\frac{a-1.1}{a-2.1}})$$

. We proceed with an argument similar to the one in Theorem 4.10. Observe that the polynomial

$$T(f(x)) = \sum_{i=1}^{\ell} T_i(x).$$

Recall that, by the triangle inequality, in order to upper-bound $\mathbf{CD}(T \circ f, U)$ it suffices to upper-bound

$$\mathop{\mathsf{E}}_{x_i, i \in S}\left[\omega^{T(f(x))}\right],$$

for any fixing of the variables $\{x_i : i \notin S\}$. Fix the variables $x_{\bar{S}}$ arbitrarily. Now, $T(f(x_S))$ decomposes to ℓ polynomials over pairwise-disjoint sets of variables x_{S_i}. We claim that these polynomials are non-zero.

Claim 6.4.1. *For every $i \in [\ell]$, for every fixing α of the variables $x_{\bar{S}}$ the polynomial $T_i(x)$ is not the zero polynomial.*

Proof. Fix some i and for now let us not fix the $x_{\bar{S}}$ variables. Let $M(y)$ be the leader of T_i (the i-th leader) and let $q' \le q$ denote its degree (in the y variables). Recall that the monomials $M_j, j \in J_i$ have degree at most q'. Therefore the polynomial $T_i(x) = M'(x) + \sum_{j \in J_i} M'_j(x)$ is of degree of at most $q'b$. Moreover, there exists a monomial $K(x)$ in $M'(x)$ of degree exactly $q'b$ and all its variables are S variables. (This is due to the "product part" of the sum-product polynomial.) The proof now follows by noting that: (1) since G_Π is $(2q+1, b/2+1)$-expanding a degree $q'b$ monomial in $M'(x)$ cannot be canceled due to the addition of $\sum_{j \in J_i} M'_j(x)$ (as shown in the proof of Theorem 6.1 any maximal degree monomial have at least a single variable that is not shared by any other maximal degree monomial); and (2) After fixing the $x_{\bar{S}}$ variables, the monomial $K(x)$ remains unchanged. Therefore, even after fixing the variables $x_{\bar{S}}$ to α, the polynomial $T_i(x)$ is not the zero polynomial. □

Let x^* denote a random variable whose S entries are uniformly distributed over $\mathbb{F}^S$ and its $\bar{S}$ entries are fixed to α. Since the T_i are defined over disjoint variables, we can apply Lemma 2.18 we have that:

$$\mathbf{CD}\left(T(f(x^*))\right) \le \prod_i U_1\left(T_i(x^*)\right) \le \gamma^\ell$$

where γ stands for the maximal bias of a non-trivial polynomial in qd variables of degree at most qb and where each variable is of degree at most q and,

$$\ell = O(k^{1-\frac{2\rho}{a-2.1}-2\rho}) = O(k^{1-2\rho\frac{a-1.1}{a-2.1}})$$

. □

7 Concrete Parameters

In this section we estimate the security of our constructions for some concrete choices of parameters. Estimates for combinatorial properties of random hypergraphs and matrices were computed as in the preliminaries section (e.g., Claim 2.10) where binomial coefficiecnts were approximated using standard approximation formulas for $\ln(n!)$.

m	d	$\log \lvert\mathbb{F}\rvert$	Required Bias	Minimal k
250000	10	64	80	415
250000	10	2048	80	308
360000	10	1024	80	322
360000	7	512	80	584
360000	10	32	80	575
360000	10	64	80	436
360000	10	128	80	373
490000	7	256	80	651
490000	10	32	80	607
490000	10	2048	80	332
250000	10	256	120	476
360000	10	128	120	540
490000	10	128	120	562

Table 2: The bias of noisy linear mapping with noise rate $\mu = 0.4$. For a given output size m, sparsity parameter d, prime-order field $\mathbb{F}$, and desired bias 2^{-b}, we calculated the required input size k such that, with probability 0.999 over a randomly chosen d-sparse matrix $M \in \mathbb{F}^{m \times k}$, the function $f_{M,\mu}$ cannot be distinguished by linear adversaries with advantage better than 2^{-b}. It is calculated based on the first item of Theorem 3.7, regarding the bias of Noisy Sparse Linear Mapping, and supplemented by the estimation of the size of the expanding set in Claim 2.10

Required Security	$\log_k(m)$	d	Computed Input
70	2	7	229
70	2	10	132
80	2	7	271
80	2	10	153
90	2	7	315
90	2	10	177
100	2	7	360
100	2	10	200
160	2	7	651
160	2	10	345
200	2	7	860
200	2	10	446
300	2	7	1431
300	2	10	712

Table 3: Estimated input for security against Attack 1. For noise parameter $\mu = 0.25$, sparsity parameter d, and desired attack complexity $\Omega\left(2^{\text{Required Security}}\right)$, we calculated the required input size k such that, with probability 0.999 over a randomly chosen d-sparse matrix $M \in \mathbb{F}^{m \times k}$, the running time of Attack 1, for the function $f_{M,\mu}$, is at least $\Omega\left(2^{\text{Required Security}}\right)$. It is calculated based on Claim 3.1, regarding the complexity of Attack 1, and supplemented by the estimation of the size of the expanding set in Claim 2.10

k	$\log_k(m)$	d	Computed Sec025	Computed Sec05
100	1.5	7	56	124
100	1.5	10	73	148
100	2	7	36	82
100	2	10	55	113
150	1.5	7	80	176
150	1.5	10	106	215
150	2	7	50	112
150	2	10	78	160
200	1.5	7	105	227
200	1.5	10	139	280
200	2	7	63	141
200	2	10	100	205
250	1.5	7	128	276
250	1.5	10	170	345
250	2	7	75	167
250	2	10	121	249
300	1.5	7	150	325
300	1.5	10	201	407
300	2	7	86	194

Table 4: Security estimation based on the linear-dependency attack described in Algorithm 1. For an input length k, output length m (described via the stretch degree $\log_k(m)$), sparsity parameter d, and field $\mathbb{F}$, we calculate a lower bound on the expected log-complexity of the attack when applied to a randomly chosen d-sparse matrix $M \in \mathbb{F}^{m \times k}$. We consider two cases for the noise rate: 0.25 an 0.5, and denote the corresponding bounds by "computed Sec0.5" and "computed Sec0.25", respectively. We base on the analysis of the attack in Section 3.1.

$\log_k(m)$	α	d	Min k for $r = 80$	Min k for $r = 120$
1	1	7	88	133
1	1	10	83	124
1	1	20	81	120
1	1	30	81	120
1.25	1	7	137	211
1.25	1	10	107	165
1.25	1	20	88	135
1.25	1	30	85	129
1.5	1	7	185	295
1.5	1	10	128	200
1.5	1	20	96	146
1.5	1	30	88	135
2	1	7	364	617
2	1	10	183	297
2	1	20	109	169
2	1	30	94	146

Table 5: Shrinking set estimation. For a given stretch degree $\log_k(m)$, sparsity parameter d, prime-order field $\mathbb{F}$, and desired shrinking-set lower-bound r, we calculated the minimal input size k such that, with probability 0.999 over a randomly chosen d-sparse matrix $M \in \mathbb{F}^{m\times k}$, M will have no shrinking set of size smaller than r.

$\log_k(m)$	α	d	Min k for $r = 80$	Min k for $r = 120$
1	3.5	7	2501	3784
1	5	10	2329	3523
1	10	20	2899	4384
1	15	30	3680	5561
1.25	3.5	7	6332	10054
1.25	5	10	4134	6444
1.25	10	20	3808	5837
1.25	15	30	4434	6762
1.5	3.5	7	19406	32625
1.5	5	10	7744	12460
1.5	10	20	5013	7791
1.5	15	30	5334	8210
2	3.5	7	529652	1057341
2	5	10	35590	61928
2	10	20	8938	14286
2	15	30	7757	12151

Table 6: Lossless expansion of random graphs. For a given stretch degree $\log_k(m)$, degree parameter d, and desired expansion parameter (α, r) where α is set to $d/2$, we calculated the minimal input size k such that, with probability 0.999, a randomly chosen (m, k, d)-hypergraph G will be (α, r)-expanding.

References

[ABR12] Benny Applebaum, Andrej Bogdanov, and Alon Rosen. A dichotomy for local small-bias generators. In *Theory of Cryptography Conference*, pages 600–617. Springer, 2012.

[ADI$^+$17] Benny Applebaum, Ivan Damgård, Yuval Ishai, Michael Nielsen, and Lior Zichron. Secure arithmetic computation with constant computational overhead. To appear in Crypto 2017., 2017. Availabale at `http://eprint.iacr.org/2017/617`.

[AL16] Benny Applebaum and Shachar Lovett. Algebraic attacks against random local functions and their countermeasures. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing*, pages 1087–1100. ACM, 2016.

[Ale03] Michael Alekhnovich. More on average case vs approximation complexity. In *Foundations of Computer Science, 2003. Proceedings. 44th Annual IEEE Symposium on*, pages 298–307. IEEE, 2003.

[App16] Benny Applebaum. Cryptographic hardness of random local functions. *computational complexity*, 25(3):667–722, 2016. Available at `http://eccc.hpi-web.de/report/2015/027`.

[Bog05] Andrej Bogdanov. Pseudorandom generators for low degree polynomials. In *Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, pages 21–30. ACM, 2005.

[BV10] Andrej Bogdanov and Emanuele Viola. Pseudorandom bits for polynomials. *SIAM Journal on Computing*, 39(6):2464–2486, 2010.

[Gol00] Oded Goldreich. Candidate one-way functions based on expander graphs. *IACR Cryptology ePrint Archive*, 2000:63, 2000.

[HLV16] Elad Haramaty, Chin Ho Lee, and Emanuele Viola. Bounded independence plus noise fools products. 2016.

[OW14] Ryan O'Donnell and David Witmer. Goldreich's PRG: evidence for near-optimal polynomial stretch. In *Proc. of IEEE 29th Conference on Computational Complexity, CCC*, pages 1–12, 2014.

[Vio09] Emanuele Viola. The sum of d small-bias generators fools polynomials of degree d. *Computational Complexity*, 18(2):209–217, 2009.