Electronic Colloquium on Computational Complexity, Report No. 98 (2023)

Public-Key Encryption, Local Pseudorandom Generators, and the Low-Degree Method

Andrej Bogdanov* Pravesh K. Kothari† Alon Rosen‡

July 5, 2023

Abstract

The low-degree method postulates that no efficient algorithm outperforms low-degree polynomials in certain hypothesis-testing tasks. It has been used to understand computational indistinguishability in high-dimensional statistics.

We explore the use of the low-degree method in the context of cryptography. To this end, we apply it in the design and analysis of a new public-key encryption scheme whose security is based on Goldreich's pseudorandom generator. The scheme is a combination of two proposals of Applebaum, Barak, and Wigderson, and inherits desirable features from both.

1 Introduction

Hypothesis testing is concerned with the computational task of detecting a noisy signal. The question is cast as a distinguishing problem between a pure noise distribution Q and an alternative distribution P that contains a planted signal. The goal is to understand tradeoffs between the "amplitude" θ and the "frequency" m.

Several works [BR13, HWX15, BB20] uncover that such problems exhibit statistical-to-computational gaps: depending on θ, there is a range of frequencies $m \in [m_{\text{stat}}, m_{\text{comp}}]$ for which hypothesis testing is possible, but no efficient algorithm is known.

The *low-degree method* is a heuristic for generating remarkably accurate estimates of the *computational threshold* m_{comp} at which the hypothesis testing problem becomes feasible [HKP+17]. It relies on the observation that for several natural average-case hypothesis testing problems, the optimal polynomial time distinguisher amounts to computing a low-degree polynomial in input samples.

The method was first employed [BHK+19] in constructing lower bound witnesses for the sum-of-squares semidefinite programming hierarchy for the *planted clique problem*. It was later [HKP+17] shown to be powerful enough to capture natural spectral algorithms and in fact used to design new algorithms for certain Bayesian estimation problems [HS17]. Indeed, no efficient algorithm appears to outperform what can be inferred by observing "local statistics".

*University of Ottawa. Email: abogdano@uottawa.ca

†Carnegie Mellon University. Email: praveshk@cs.cmu.edu

‡Bocconi University and Reichman University. Email: alon.rosen@unibocconi.it

ISSN 1433-8092

In [HKP+17], the authors make the *pseudocalibration conjecture* positing that hardness in the low-degree model implies sum-of-squares lower bounds for average-case refutation problems under certain mild niceness conditions. Later works [HS17, Hop18, KWB19] have proposed a stronger variant of the pseudocalibration conjecture positing that thresholds computed in the low-degree method are in fact m_{comp} i.e., a threshold for all polynomial time computable distinguishers. In the past few years, a sequence of works have used the low-degree method to find evidence of gap between computational and statistical thresholds for a number of average-case algorithmic problems.

In this work we will be interested in exploring the applicability of the low-degree method to cryptography, and in particular to the design and analysis of a new public-key cryptosystem. While we do not claim that the method's predictions always coincide with the computational infeasibility threshold m_{comp}, we do believe that it can serve a guideline for sound design, in addition to being a sanity-check for assessing security.

1.1 Goldreich's Pseudorandom Generator

The main object underlying our new public-key encryption scheme is *Goldreich's candidate one-way function* [Gol11]. We will instantiate it in a way that may allow us to conjecture it to be a pseudorandom generator given known attacks.

The function, denoted F_H, maps n bits to m bits. It is described in terms of two main objects:

- A d-hypergraph H on on n vertices and m (ordered) hyperedges, each of size d. See Figure 2 where the vertices are represented by circles (○) and the hyperedges are represented by squares (■).

- A d-ary predicate that is applied to the projection of an n-bit input x on each one of the m hyperedges of H.

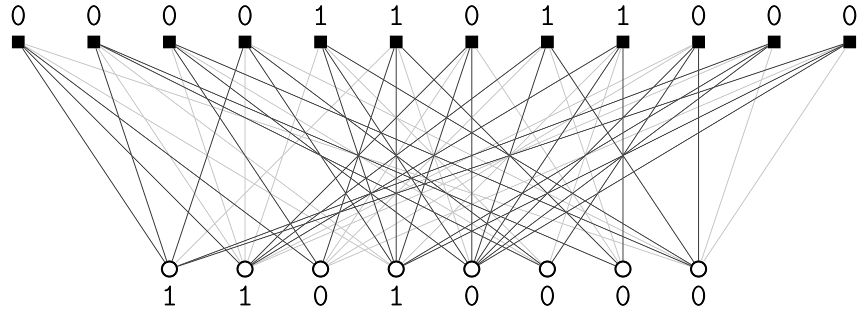

Figure 1: An instance of Goldreich's function with predicate $x_1 \oplus x_2 \oplus x_3 \oplus x_4 x_5$. The non-linear part $x_4 x_5$ is shaded light grey.

For concreteness, we set H to be a 5-hypergraph on n vertices and m hyperedges, let $d = k + 3$, and $k = 2$. In Section 6 we give a more general description parametrized by k. We set the predicate to be $x_1 \oplus x_2 \oplus x_3 \oplus x_4 x_5$ [MST06].

1.2 A New Public-Key Encryption scheme

Recall that a public-key encryption scheme involves three algorithms: Key generation, Encryption, and Decryption (see Section 2). Our scheme has binary message space and we allow both imperfect correctness and security.

Encryption. In our scheme, encryptions of 0 are outputs $y = F_H(x)$ of Goldreich's function applied to a random input x. Encryptions of 1 are random m-bit strings y. Indistinguishability of encryptions of 0 from encryptions of 1 follows from pseudorandomness of F_H given H.

Decryption. Decryption is made possible thanks to logarithmic size hypergraphs (called "hyperloops") that are planted in H in the key generation process. These hyperloops make it effectively possible to distinguish between $y = F_H(x)$ and a random m-bit string y.

Key-generation. A *hyperloop* is a 3-hypergraph in which every vertex has degree two. Let L_0 be a fixed hyperloop with $\ell_0 = O(\log n)$ hyperedges. The public key of our scheme consists of a 5-hypergraph H sampled as follows. Let:

- L be the union of $t = 2^{\Theta(\ell_0)}$ vertex-disjoint copies of L_0,
- Q be a random 3-hypergraph with n vertices and hyperedge probability $O(n^{-3/2-\delta})$,
- $P = Q \cup L$ where L is planted on a random subset of vertices of Q,
- H be obtained by randomly adding 2 vertices to each hyperedge in P.

The public key is the 5-hypergraph H and the secret key is $S_1, \dots, S_t$, where $S_i \subseteq \{1, \dots, m\}$ are the hyperedges corresponding to the i^{th} planted copy of L_0.

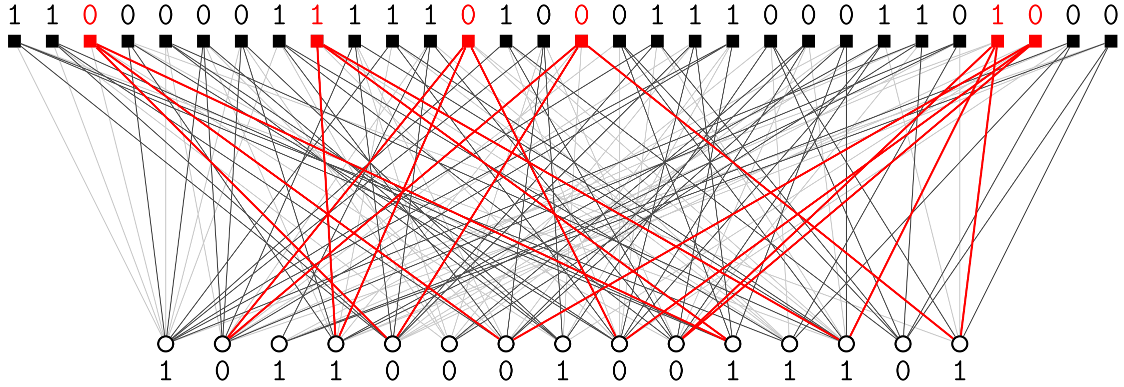

Figure 2: A public key and ciphertext with a single planted hyperloop L_0 (from Figure 3). The secret key and the ciphertext section used in decryption are marked in red.

1.3 Correctness

The hypergraph H is published, enabling anybody to encrypt by evaluating the function F_H on a random input. Knowledge of $S_1, \ldots, S_t$ enables correct decryption, since each planted hyperloop S gives noticeable advantage in distinguishing the output of F_H from a random string. Our decryption function is a majority of parities over hyperloops $S \subseteq \{1, \ldots, m\}$.

Claim 1. *If $y = F_H(x)$ then $z = \oplus_{j \in S} y_j$ has bias $2^{-\ell_0}$.*

Proof. All vertices in S have degree two and so $\oplus_{j \in S}(x_{j_1} \oplus x_{j_2} \oplus x_{j_3}) = 0$. Thus, when $y = F_H(x)$, we have $\oplus_{j \in S} y_j = \oplus_{j \in S} x_{j_4} x_{j_5}$, which has bias $2^{-\ell_0}$. □

Testing whether more than $(1+2^{-\ell_0})t/2$ of $z_1, \ldots, z_t$ are zero distinguishes $y = F_H(x)$ from random, and thus decrypts correctly, with probability $1 - o(1)$.

1.4 Security

For the scheme to be secure it is necessary that the output of the function F_H is pseudorandom and that the public key H hides the planted hyperloops $S_1, \ldots, S_t$. This will in particular be true if:

1. A planted hypergraph $P = Q \cup L$ is indistinguishable from a random one Q.
2. The output of F_H is pseudorandom when H is a random hypergraph Q.

While these two properties may be strictly stronger than required, we will analyze their individual plausibility. Our examination is conducted both in light of the best known asymptotic attacks, and within the low-degree framework.

A distinguisher of non-negligible advantage exists as there is already noticeable probability that F_H contains a constant-sized subset of output bits that always XOR to zero. We will restrict the discussion to efficient distinguishers of *constant* advantage, but arguments remain valid for distinguishing advantage $n^{-o(1)}$. Ruling those out is sufficient to obtain a gap between decryption advantage and distinguishing advantage, which can then be amplified (at some cost in efficiency). On the other hand, the secret key can be inverted by exhaustive search in time $\binom{m}{\ell_0} = n^{O(\log n)}$ so we will restrict the analysis to distinguishers that run in time $n^{o(\log n)}$.

Indistinguishability of P and Q. The distributions P and Q are statistically distinguishable since the planted hyperloops L of size ℓ_0 in P are unlikely to appear in Q. There are two natural distinguishers to try in this context. Exhaustive search for a hyperloop of size ℓ_0 has complexity at least $n^{-O(1)}\binom{m}{\ell_0}$. Another possibility is to look for a discrepancy in the number of hyperedges between P and Q. As long as the $2^{\Theta(\ell_0)}$ hyperedges present are within $o(1)$ of the standard deviation $n^{0.75-\delta/2}$ in the number of edges of Q this discrepancy will not be noticeable.

These two distinguishers are based on counting size-ℓ_0 hyperloops and counting hyperedges. The counts are polynomials of degree ℓ_0 and one, respectively, in the adjacency tensor of H. In Proposition 9 we show that they are close to best possible in the low-degree framework: For every ϵ there exists an δ and a choice of L_0 so that no polynomial of degree at most $(1-\epsilon)\ell_0$ has constant advantage.

Conjecture 2. *For a sufficiently small constant δ, $m = n^{1.5-\delta}$, $\ell_0 = 0.36 \log n$, and $t = n^{0.75-\delta}$, P and Q are $(1 - \Omega(1))$-indistinguishable in $n^{O(1)}$-time.*

Pseudorandomness of F_Q. For a random Q, the output $y = F_Q(x)$ has been conjectured to be computationally indistinguishable from a random $y \in \{0,1\}^m$ given Q. When the hyperedge probability is $O(n^{-3/2-\delta})$ the graph has $m = \Theta(n^{1.5-\delta})$ edges with high probability. The best known distinguisher has complexity $m^{\Theta(n^{2\delta})}$ and is based on a landmark result of Feige, Kim, and Ofek [FKO06]. They prove that H is likely to contain $2^{\Theta(\ell_1)}$ hyperloops each of size $\ell_1 \approx n^{2\delta}$. The distinguisher effectively inverts the secret key and runs our decryption algorithm assuming the "public key" is sampled from the model distribution Q.

In Proposition 14 we show that the advantage of any statistic that depends arbitrarily in Q but has most degree d in y in distinguishing $y = F_Q$ from random is upper bounded by the expected number of hyperloops of size at most d in Q. This expectation is $o(1)$ when $d = o(n^{-2\delta})$, that is for any statistic of degree just too low to "see" the hyperloops in Q. This complements results on the small-bias of F_Q [MST06, ABR16, OW14, AL18].

Conjecture 3. *For every δ, $m = n^{1.5-\delta}$, random $x \in \{0,1\}^n$, $y \in \{0,1\}^m$, $(Q, F_Q(x))$ and (Q, y) are $o(1)$-indistinguishable in $n^{O(1)}$-time.*

1.5 The Low-Degree Method

The low-degree method is a formal framework for arguing *computational* hardness of hypothesis testing. Although the method is, in full generality, neither complete or sound, it correctly predicts the computational threshold m_{comp} for a variety of problems. The method is effective in settings where the computational advantage is non-negligible but vanishing (e.g., $n^{-\Omega(1)}$), and where the model distribution is a high-dimensional product distribution. It is in particular applicable for analyzing the two security claims from Section 1.4 and for detecting vulnerabilities in alternative design choices.

We are in particular interested in the following question: For which planted structures L are the distributions Q and $P = Q \cup L$ computationally indistinguishable? Perhaps the simplest attack is to try and detect a discrepancy in the number of edges. Should the edge numbers be close, could the attacker rely on discrepancies on other fixed-size substructures such as 5-cycles? It turns out that this won't help as long as the planted substructure L is sufficiently small-set expanding (see Proposition 9).

Consider for example an alternative construction $P' = Q \cup L'$ in which L' now consists of a union of $2^{\Theta(\ell_0)}$ *independent* random size-ℓ_0 hyperloops. P' and Q are now distinguishable as L' will induce a significant discrepancy in the number of 4-cliques. These additional structures completely break security of encryption.

The low-degree method is in general incomplete as it does not model algebraic attacks. For example it predicts that random n-bit strings of parities 0 and 1 are degree-$(n-1)$ indistinguishable. Nevertheless we believe that it can be a useful guide in "noisy linear algebra" type constructions with noticeable security error.

One technical difficulty in low-degree analysis is the lack of a triangle inequality. In our case we show that $P = Q \cup L$ is low-degree indistinguishable from Q and that (Q, F_Q) is low-degree indistinguishable from (Q, random). However we cannot compose the two claims to conclude that (P, F_P) is low-degree indistinguishable from (Q, random). Nevertheless, we prove a weaker security claim with an additional assumption on the "distinguisher" in Theorem 16.

1.6 Relation to the ABW Schemes

Applebaum, Barak, and Wigderson [ABW10] proposed two closely related public-key encryption schemes that differ from ours in the choice of planted structure L and predicate used in the underlying pseudorandom generator.

In their first scheme (ABW1) L is a single hyperloop of size $\ell = \Theta(n^{2\delta})$ and the predicate is the randomized function $x_1 \oplus x_2 \oplus x_3 \oplus e$, where e is a noise bit of probability $n^{-2\delta}$. Alternatively, e can be replaced by an AND of $k = \log \ell + O(1)$ input bits. This encryption is not local (although in any reasonable parameter setting a small value of k may suffice.) Their security analysis relies on *statistical* indistinguishability of Q and $P = Q \cup L$ thus obviating the need for additional computational assumptions.

The main difference is that, unlike ABW1, our proposal has constant locality. Another difference is that our construction doesn't use extrinsic noise bits e. The role of the noise is played by the nonlinear part $x_4 x_5$ of our predicate.

In their second scheme (ABW2) L is a single subgraph of size $\ell = \ell_0 = \Theta(\log n)$, with fewer vertices than hyperedges. The predicate in this construction can be arbitrary. To decrypt one checks whether the ℓ-bit part of the ciphertext restricted to L has a preimage. (With a small modification this scheme supports errorless decryption.)

Unlike in ABW2, our secret key consists of *multiple* planted known linear dependencies between the output bits. This endows our scheme with natural leakage-resilience: Even if a small subset of these dependencies becomes public encryption remains secure. Another difference is that our decryption may be of lower complexity in some models as it is a majority of parities, while ABW2 rely on a hardcoded lookup table.

Moreover, we believe that our scheme may be marginally more secure than theirs. A brute-force search for the secret key would have complexities $\binom{m}{\ell_0}$ and $\binom{n}{\ell_0}$ in our and their variant, respectively. The gap is most prominent when $m = n^{3/2-\delta}$ is large, i.e., when δ is small. In the regime where δ approaches $1/2$ lower-degree attacks (based on detecting some substructure present in L) become possible. A more precise low-degree analysis is needed for a fair comparison.

As for security guarantees, Applebaum, Barak, and Widgerson identify a discrepancy in the number of small cycles as a potential vulnerability of their schemes and account for it in parameter setting. The low-degree method systematically rules out all attacks of this type and more. While the low-degree method readily applies to ABW1 and ABW2, its relevance in security analysis is better highlighted in our scheme as it informs choices in the construction (hyperloop sampling in key generation) and parameters (size and density of hyperloops).

1.7 Open Questions

One weakness of our security analysis is that it relies on *computational* indistinguishability of the model hypergraph distribution Q and the planted distribution P that contains $2^{\Theta(\ell_0)}$ copies of the planted hyperloop L_0 with $\ell_0 = \Theta(\log n)$ edges. Might it be possible to argue that the proximity is statistical?

Feige, Kim, and Ofek prove that a random 3-hypegraph with n vertices and hyperedge probability $p = O(n^{-3/2}/\ell_0^{1/2})$ is likely to contain K^{ℓ_0} hyperloops of size ℓ_0 (for any desired constant K). We believe, however, that the number of *disjoint* hyperloops of size ℓ_0 is at most polynomial by the following heuristic argument. In expectation a large fraction of the hyperloop pairs intersect. If we model the intersection graph as a random graph its maximum independent set would have expected size logarithmic in the number of hyperloops 2^{ℓ_0}, namely polynomial in the hyperloop size ℓ_0. Thus it appears that the planted instance P is statistically far apart from the model instance Q.

Nevertheless, most of the intersections between the K^{ℓ_0} hyperloops are small on average. This raises the following question: Is it possible to efficiently sample the collection L of K^{ℓ_0} intersecting size-ℓ_0 hyperloops jointly with the random hypergraph Q? If the answer to this question is positive, the public key can be sampled directly from the model distribution. Although the information bits $z_1, \dots, z_t$ arising from the different hyperloops in the decryption process would be dependent, their correlations are sufficiently small to enable reliable decryption.

Concerning empirical security, it is unclear if the noise Q is needed at all in the construction. Could the scheme be secure even if P consists of nothing but $n^{1.1}$ randomly planted copies of L_0?

2 Public Key Encryption

Our encryption scheme has binary message space, decrypts incorrectly with bounded probability δ, and has noticeable (but still bounded) computational distance ε between the distribution of encryptions of zero and those of one. Assuming both errors are sufficiently small constants they can be amplified to be negligible at a loss of parameters [DNR04].

Definition 4 (Syntax)**.** *A* public key encryption scheme *consists of three algorithms* $(\mathrm{Gen}, \mathrm{Enc}, \mathrm{Dec})$ *such that for* $n \in \mathbb{N}$, $\mathrm{Gen}(1^n)$ *outputs a pair of keys* (sk, pk)*;* $\mathrm{Enc}(pk, b)$ *encrypts a message b with the public key pk and outputs a ciphertext c;* $\mathrm{Dec}(sk, c)$ *decrypts a ciphertext c using the secret key sk and outputs a message b.*

Both key-generation, Gen, and encryption, Enc, are randomized. As mentioned above, we allow the decryption algorithm, Dec, to make errors.

Definition 5 (δ-correctness)**.** *A public key encryption scheme* $(\mathrm{Gen}, \mathrm{Enc}, \mathrm{Dec})$ *is* correct with probability δ *if*

$$\Pr\left[\mathrm{Dec}(sk, \mathrm{Enc}(pk, b)) = b\right] \geq \delta,$$

where probability is taken over the randomness of Gen *and* Enc*. We call $1 - \delta$ the* decryption error*.*

Security is defined through indistingushability of encryptions [GM84]. To this end, we rely on the notion computational indistinguishability.

Definition 6 (ε-indistinguishability)**.** *Two distributions X, Y are* ε-indistinguishable *if for any probabilistic polynomial time algorithm A:*

$$|\Pr[A(X) = 1] - \Pr[A(Y) = 1]| \leq \varepsilon.$$

Definition 7 (ε-security)**.** *A public key encryption scheme* (Gen, Enc, Dec) *is said to have* security error $\varepsilon \in [0,1]$ *if the distributions* $(pk, \mathrm{Enc}(pk,0))$ *and* $(pk, \mathrm{Enc}(pk,1))$ *are ε-indistinguishable, where probabilities are over the randomness of* Gen *and* Enc*.*

3 The Low-Degree Method

Suppose we want to distinguish distribution P from model distribution Q. One way is to sort the outcomes x in order of decreasing likelihood ratio $L(x) = P(x)/Q(x)$, say "p" if it is large and "q" if it is small. The Neyman-Pearson Lemma says that this test minimizes the false positive error among all tests with a given false negative error.

The likelihood ratio can also be used to argue indistinguishability. For any test T,

$$|P(T) - Q(T)| = |\mathbb{E}_Q[(L-1)\cdot 1_T]| \leq \sqrt{\mathbb{E}_Q[(L-1)^2]\cdot \mathbb{E}_Q[1_T^2]} = \sqrt{\mathrm{Var}_Q[L]\cdot Q(T)}.$$

Therefore the statistical distance is at most the standard deviation of L under Q. Even if the variance is greater than one but bounded, this bound rules out the possibility that $P(T) \to 1$ and $Q(T) \to 0$ so the statistical distance between P and Q must be bounded away from one.

Example Let Q and P consist of n i.i.d. ± 1 bits that are unbiased and ϵ-biased, respectively. The likelihood ratio is $L(x) = \prod(1+\epsilon x_i)$, its variance is $\mathrm{Var}[L] = \prod \mathbb{E}[(1+\epsilon x_i)^2] - 1 = (1+\epsilon^2)^n - 1$. The variance is $o(1)$ as long as $\epsilon \ll 1/\sqrt{n}$, which matches the regime in which we cannot distinguish reliably. If we expand $L(x)$ as a polynomial we get $L(x) = 1 + \epsilon\sum x_i + \epsilon^2 \sum_{i\neq j} x_i x_j + \cdots$. The degree-$d$ part contributes $\binom{n}{d}\epsilon^{2d}$ to the variance so the main contribution comes from the degree-1 part $L^1(x) = 1 + \epsilon\sum x_i$. In fact we can use the value of L^1 to distinguish P and Q when $\mathrm{Var}_Q[L^1]$ is large.

This example suggests using the low-degree projection L^1 or more generally L^d to distinguish P from Q assuming Q is a product distribution over bits. (The theory generalizes to product distributions over other domains.) The advantage of L^d is that it can be computed in size $\binom{n}{d}$. In contrast, L may not be efficiently computable in general. For a number of statistical hypothesis testing problems, the best efficient distinguishers are based on the value of some low-degree polynomial. Among those distinguishers, L^d is optimal in the following sense:

Claim 8. *Among all degree-d polynomials f, L^d maximizes the advantage*

$$a_d = \max_f \frac{\mathbb{E}_P[f] - \mathbb{E}_Q[f]}{\sqrt{\mathrm{Var}_Q[f]}}.$$

Moreover, $a_d = \|L^d - 1\|_Q$.

A degree-d polynomial can capture any "d-local" statistic. For example, if P and Q are graphs (represented by their adjacency matrices) then f can compute the number of copies of any given induced subgraph with d edges. A natural distinguisher in this context is a test of the form $f(x) > t$ for a suitable threshold t. If it happens that $\mathrm{Var}_P[f] = O(\mathrm{Var}_Q[f])$ then f will be concentrated around its means under both P

and Q so a large value of a_d means that $f(P)$ will typically be large while $f(Q)$ will typically be small. If on the other hand $\mathrm{Var}_P[f] \gg \mathrm{Var}_Q[f]$ then it may be reasonable to try $g(x) = (f(x) - \mathbb{E}_Q[f])^2$ as a distinguisher of degree $2d$. Thus small advantage is evidence of failure for all distinguishers of this type.

Proof of Claim 8. The maximum advantage can be rewritten as $\max \mathbb{E}_P[f]$ where f is constrained to have degree d, mean $\mathbb{E}_Q[f] = 0$ and variance $\mathbb{E}_Q[f^2] = 1$. Since f has degree at most d,

$$\mathbb{E}_P[f] = E_Q[f \cdot L] = \mathbb{E}_Q[f \cdot (L^d - \mathbb{E}_Q[L^d])] = \mathbb{E}_Q[f \cdot (L^d - 1)],$$

This expression is maximized when $f = (L^d - 1)/\|L^d - 1\|$. (As the maximum is invariant under scaling and shifting we can also take $f = L^d$.) The advantage is

$$\frac{\mathbb{E}_P[L^d - 1]}{\|L^d - 1\|_Q} = \frac{\mathbb{E}_Q[L^d(L^d - 1)]}{\|L^d - 1\|_Q} = \|L^d - 1\|_Q. \qquad \square$$

If Q is the p-biased product distribution over $\{\pm 1\}^n$ so that $\Pr(X_i = -1) = p$, $\Pr(X_i = 1) = q = 1 - p$. The Fourier basis is given by $\phi_S(x) = \prod_{i \in S} \phi(x_i)$, where $\phi(-1) = -\sqrt{q/p}$ and $\phi(1) = \sqrt{p/q}$. The squared degree-d advantage a_d^2 is

$$a_d^2 = \|L^d - 1\|_Q^2 = \sum_{1 \le |S| \le d} \mathbb{E}_P[\phi_S]^2. \tag{1}$$

4 Planting Hyperloops

A *hyperloop* is a 3-hypergraph in which every vertex has degree two. Let Q be a random 3-hypergraph on n vertices with edge probability p and $P = Q \cup L$ where L is a hyperloop on ℓ edges.

Proposition 9. *Assume that for every $1 \le s \le d$, every set of s hyperedges in L touches at least $(3/2 - \delta)(s + 1) - 2\delta$ vertices. If $p \ge C\sqrt{d}n^{-3/2-\delta}$ and $\ell \le \eta\sqrt{pn^3/Cd^{3/2}}$ for some constant C and sufficiently large n then the degree-d advantage $a_d(P, Q)$ is $\le \eta$.*

The proposition guarantees degree-d indistinguishability as long as L is small-set expanding and the number of planted hyperedges is within $o(\eta/d^{3/4})$ standard deviations of the expected number of hyperedges pn^3 in the host hypergraph Q. Thus in this regime, no low-degree distinguisher can significantly improve over counting hyperedges.

In our intended application L will consist of ℓ/ℓ_0 vertex-disjoint copies of a single hyperloop L_0 with $\ell_0 = O(\log n)$ hyperedges. By Claim 13 a random choice of L_0 will have the desired expansion with constant probability.

Proof. We expand a_d in the Fourier basis as:

$$a_d^2 = \sum_{1 \le |S| \le d} \mathbb{E}_P[\phi_S]^2 = \sum_{1 \le |S| \le d} \left(\frac{1 - p}{p}\right)^{|S|} \Pr[S \subseteq L]^2.$$

A copy of S in L is a map from the vertices of S to the vertices of L that maps edges into edges. Let $C(S, L)$ be the number of such copies. By a union bound

$\Pr[S \subseteq L] \leq C(S,L)/n(n-1)\cdots(n-v(S)+1)$ where $v(S)$ is the number of vertices in S. Assuming that $v(S) \leq 3d = O(\sqrt{n})$ the denominator dominates $n^{v(S)}$ so $\Pr[S \subseteq L] \lesssim C(S,L)n^{-v(S)}$. Therefore $a_d^2 \leq \sum_{1\leq|S|\leq d} f(S)$ where $f(S) = ((1-p)/p)^{|S|}C(S,L)^2 n^{-2v(S)}$.

If the vertex sets of S and S' are disjoint then $f(S\cup S') \leq f(S)f(S')$. Therefore $f(S') \leq \prod_C f(S)$ where the product ranges over the connected components S of S':

$$a_d^2 \leq \sum_{1\leq|S'|\leq d} \prod_{\text{c.c. } S \text{ of } S'} f(S) \leq \Bigg(1 + \sum_{\substack{1\leq|S|\leq d\\ S \text{ connected}}} f(S)\Bigg)^d - 1.$$

To obtain $a_d \leq \eta$ it is therefore sufficient to show that the summation over connected S is at most $\eta^2/2d$.

Claim 10. *If S is connected then $C(S,L) \leq 3|L|\cdot 2^{|S|}$.*

Proof. Let $s = |S|$ and let $e_1,\dots,e_s$ be an ordering of the edges so that e_i is connected to $e_1,\dots,e_{i-1}$ for all i. The first edge $e_1 = \{v_1, v_2, v_3\}$ of S can map into L into at most ℓ ways, and there are $3! = 6$ ways to assign v_1, v_2, v_3 that are consistent with this edge map. Since L has degree two and e_2 intersects e_1, the image of e_2 is fixed by this assignment. There are then at most two ways to assign the vertices of $e_2 \setminus e_1$. By the same argument there are at most two ways to assign the vertices of $e_i \setminus (e_1 \cup \dots \cup e_{i-1})$. Therefore $C(S,L) \leq 3!|L|\cdot 2^{s-1} = 3|L|\cdot 2^s$. □

Applying this bound and disregarding the $(1-p)^{|S|}$ factor (which will be small) we obtain

$$\sum_{\substack{1\leq|S|\leq d\\ S \text{ connected}}} f(S) \leq 9\ell^2 b_d \quad \text{where} \quad b_d = \sum_{1\leq|S|\leq d} \frac{1}{(p/2)^{|S|}n^{2v(S)}}.$$

Let $N(s,v)$ be the number of connected 3-hypergraphs with s edges that span a fixed set of v vertices and appear at least once in L. Then

$$b_d \leq \sum_{s=1}^{d}\sum_{v}\binom{n}{v}N(s,v)(p/2)^{-s}n^{-2v} \leq \sum_{s=1}^{d}\sum_{v}\frac{N(s,v)}{v!}\cdot(p/2)^{-s}n^{-v}. \tag{2}$$

The leading term $s=1$, for which v must equal 3, contributes $O(1/pn^3)$. The objective is to show that it dominates the summation assuming that L is sufficiently expanding. If this is the case then the advantage will be bounded as long as $1/pn^3 = O(\eta^{-2}/d\ell^2)$, or $\ell = O(\eta\sqrt{pn^3/d})$. Owing to some slackness in the calculation we will only show that the dominating term is at most $O(\sqrt{d}/pn^3)$, thereby accounting for the additional $\sqrt{d}$ factor in the statement.

Claim 11. *$N(s,v)/v! = O(c^s s^{s/2})$ for some constant c.*

Proof. Let u be the number of degree-1 vertices. Since all vertices have degree 1 or 2 we must have $v = (3s+u)/2$. There are $\binom{v}{u}$ ways to choose the degree-1 vertices. Once these are fixed we argue that the hypergraph can be chosen in $\Theta(h(s,u))$ ways, where

$$h(s,u) = \frac{(3s)!}{s!\cdot 6^s\cdot 2^{(3s-u)/2}}.$$

Using Stirling's formula we obtain $N(s,v) = O(c'^s s^{2s}/u!(v-u)!)$ for some constant c'. The denominator is minimized when $u = \lfloor v/2 \rfloor$ which gives, again applying Stirling's formula, $N(s,v) = O(c^s s^{2s}/v^v)$. As the maximum degree is 2, v must be at least $3s/2$ and the claim follows.

Let C be the set of $3s$ "clones" consisting two copies of each degree-2 vertex and all the degree-1 vertices. The clones can be partitioned into s hyperedges in $(3s)!/(s! \cdot 6^s)$ ways. Each (s,u)-hypercycle arises from $2^{(3s-u)/2}$ partitions of clones in which no pair of clones is covered by the same hyperedge.

Thus the number of (s,u)-hypercycles is between $qh(s,u)$ and $h(s,u)$, where q is the probability that no pair of clones is covered by the same hyperedge when the partition is chosen at random.

It remains to lower bound q by a constant. The random partition can be sampled by randomly arranging the $3s$ clones and assigning clones $3j$, $3j+1$, and $3j+2$ to the j-th hyperedge. After arranging the u degree-one vertices and the first clone of the remaining $(3s-u)/2$ vertices, no pair is covered by the same hyperedge as long as each of the second clones is separated by the corresponding first clone by at least two other clones when its position in the arrangement is chosen. For any given second clone, this happens with probability at least $1-4/((3s+u)/2)$ (as there are at most four forbidden positions). Thus q is at least $(1-4/((3s+u)/2))^{(3s-u)/2} \geq (1-4/(3s/2))^{3s/2} \geq e^{-4}$. □

Plugging into (2) we obtain

$$b_d \lesssim \sum_{s=1}^{d} \sum_{v} (p/2c\sqrt{s})^{-s} n^{-v} \lesssim \sum_{s=1}^{d} (p/2c\sqrt{d})^{-s} n^{-v(s)},$$

where $v(s) = \min_{S \subseteq L, S \text{ connected}, |S|=s} v(S)$.

Assume $p/2c\sqrt{d} \geq n^{-3/2-\delta}$ for some constant $\delta > 0$. Then the summation is dominated by the term $s = 1$ as long as $(3/2+\delta)s - v(s) < -3/2 + \delta$, or

$$v(s) \geq (3/2+\delta)(s+1) - 2\delta \text{ for every } s \leq d. \tag{3}$$

□

4.1 Expansion of 3-regular graphs

Assume L consists of ℓ/ℓ_0 vertex-disjoint copies of a single hyperloop L_0 with ℓ_0 edges. If L_0 satisfies (3) so will L. It will be more convenient to analyze the dual object L_0^* of L_0 obtained by transposing the incidence matrix of L_0. Then L_0^* is a simple 3-regular graph with ℓ_0 vertices and $3\ell_0/2$ edges. Equation (3) then any set of $s \leq d$ vertices in L_0^* must touch at least $(3/2+\delta)(s+1) - 2\delta$ edges.

Claim 12. *If a set S of size s touches e edges then the cut $(S, \overline{S})$ has size at least $2e - 3s$.*

Proof. We can write $e = in + out$ where in and out is the number of edges inside S and leaving S, respectively. Since every vertex (in S) has degree 3, $2in + out = 3s$. Therefore $out = 2e - 3s$. □

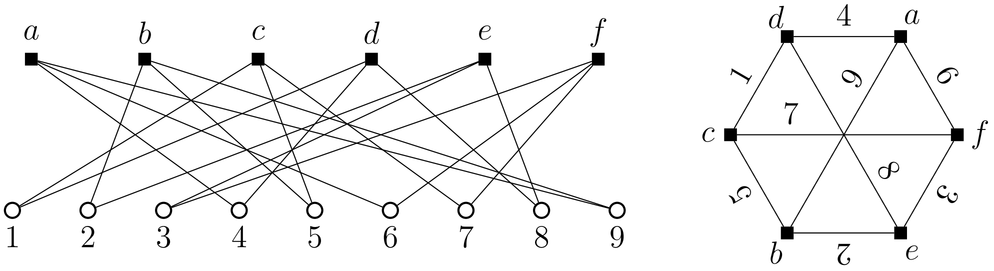

Figure 3: (a) A hyperloop L and (b) its dual representation L^*

It is therefore sufficient that the cut $(S, \overline{S})$ has size at least $2\delta(s-1)+3$ for every set S of s vertices in L_0^*. If L_0^* has sufficiently high girth and high spectral expansion this would hold for sets up to size linear in ℓ_0. However this type of analysis would likely give poor concrete parameters: Even if L_0^* is Ramanujan its spectral expansion would be at most $1 - 2\sqrt{2}/3 \approx 0.06$, which merely guarantees that $|(S, \overline{S})| \geq 0.17s$. To obtain the desired expansion for small sets the girth would need to be at least 18 resulting in a prohibitively large L_0^*.

In terms of concrete parameters there exists a hyperloop L_0 on 14 vertices that satisfies (3) with $d = 9$ and $\delta = 1/8$ (see Figure 4). A random construction also works well asymptotically:

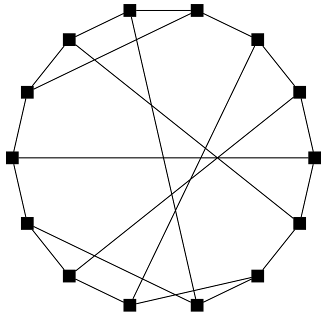

Figure 4: A size 14 hyperloop L_0^* with $\delta = 1/8$ for $d = 9$.

Claim 13. *For every $\epsilon > 0$ there exists a $\delta > 0$ so that for sufficiently large ℓ_0 and for a random L_0 (3) is satisfied with constant probability up to $d = (1-\epsilon)\ell_0$.*

Proof. We sample L_0 from the configuration model in which vertices are cloned thrice and then the clones are randomly matched. We consider three parameter ranges.

If $\epsilon\ell_0 < s \leq (1-\epsilon)\ell_0$, with probability approaching one as $\ell_0 \to \infty$, L_0 is a edge-expander [HLW06] so $|(S, \overline{S})| \geq \alpha \min\{|S|, |\overline{S}|\}$ for every S for some absolute expansion constant α. This is at least $2\delta(|S|-1)+3$ for all $|S| = s$ in the desired range as long as $\delta \leq \alpha\epsilon/2 - 3/2(\ell_0 - 1)$.

If $4 \leq s \leq \epsilon\ell_0$, the probability that there exists a set of s vertices that touches at most $v = \alpha s$ edges is at most

$$\begin{aligned}\binom{\ell_0}{s}\binom{3\ell_0/2}{v} \cdot \frac{2v}{3\ell_0} \cdot \frac{2v-1}{3\ell_0 - 1} \cdot \frac{2v-3s+1}{3\ell_0 - 3s+1} &\leq \left(\frac{e\ell_0}{s}\right)^s \left(\frac{3e\ell}{2v}\right)^v \left(\frac{2v}{3\ell_0}\right)^{3s} \\ &= \left(c(\alpha)\left(\frac{s}{\ell_0}\right)^{2-\alpha}\right)^s,\end{aligned}$$

where $c(\alpha) = (8e\alpha^3/27)(3e/2\alpha)^\alpha$. Setting $\alpha = 1/8 - 3\delta/4$ we obtain that (3) can be satisfied for all $4 \leq d \leq \ell_0$ with probability that approaches one as ϵ approaches zero.

If $1 \leq s \leq 3$ we will argue that $v(1) = 3$, $v(2) = 5$, and $v(3) = 7$, namely the graph has no parallel edges, no self-loops, and has girth at least five, with probability $\Omega(1)$. Consider the following procedure for sampling the graph. Start with the integer sequence $s = (1, 2, \dots, 3\ell_0/2)$. Now insert another copy of each integer at a random position in the sequence. In the resulting sequence of length $3\ell_0$ identify the integers with edges and the "clones" at positions $3j, 3j+1, 3j+2$ with vertex j. We describe a sequence of events $G_1, \dots, G_{3\ell/2}$ where G_i is measurable in the filtration obtained by exposing the j-th insertion, each G_i has probability at least $1 - O(1/\ell_0)$ conditioned on $G_1, \dots, G_{i-1}$, and the conjunction $G_1 \cap \dots \cap G_{3\ell/2}$ implies the desired properties.

The property $v(1) = 3$ (no parallel edges or self-loops) will be satisfied as long as the two copies of every integer are spaces at least three items apart. It is clearly sufficient that this holds at the time of insertion as subsequent insertions can only increase the distance. The corresponding event G_j has clearly the desired properties as at the time of each insertion there are only four forbidden positions out of at least $3\ell_0/2$.

Similarly, $v(2) = 5$ and $v(3) = 7$ (the girth is at least five) is satisfied as long at when i is inserted it does not land two slots within any number that is within "two hops" to the copy of i that is already present, where a hop between i and i' is allowed if they appear within two positions of each other. This specifies at most 160 forbidden positions so the corresponding event G_j still has probability $1 - O(1/\ell_0)$. □

5 Low-degree Security of Goldreich's Function

We show that Goldreich's function on a random hypergraph is secure with respect to low-degree tests. We consider tests f that receive as input a hypergraph H and a string y that is either an output F_H of Goldreich's function or a random string R. The test f may depend arbitrarily on H but must have degree at most d in y.

Proposition 14. *The squared low-degree advantage of f is at most the expected number of projections of F_H on nonempty subsets of size at most d that have nonzero bias.*

In particular, if the predicate is of the form $X_1 + X_2 + X_3 + g(Y)$ then all nonzero bias subsets must come from hyperloops induced by the X-variables. Therefore a_d^2 is at most the expected number of hyperloops of size at most d in a random 3-hypergraph. Any hyperloop that spans a specific set of v vertices must have at least $2v/3$ hyperedges, so in the $H(n,p)$ model the expected number of hyperloops that span some set of v vertices is at most

$$\binom{n}{v}\binom{\binom{v}{3}}{2v/3}p^{2v/3} \leq \left(\frac{en}{v}\right)^v \left(\frac{ev^2p}{4}\right)^{2v/3}.$$

Assuming $d \leq 0.4p^{-2}n^{-3} = \tilde{\Omega}(n^{2\delta})$, the expectation is dominated by the first term $v = 6$ for which it has value $\tilde{O}(n^{-4\delta})$.

Proof of Proposition 14. As in the proof of Claim 8, the advantage is $\max_f \mathbb{E}[f(H, F_H)]$ where f is constrained to have zero mean and unit variance under the model distribution (H, R). We can write $\mathbb{E}_{H,F_H}[f] = \mathbb{E}_{H,R}[f \cdot L]$ where L is the joint likelihood ratio

$$L(h, r) = \frac{\Pr(H = h, F_H = r)}{\Pr(H = h, r = R)} = \frac{\Pr(F_H = r | H = h)}{\Pr(r = R)}.$$

Namely, $L(h, r)$ equals the conditional likelihood ratio $L(r|h)$. Thus the optimal choice of f is the conditional degree-d projection $L^d(r|h)$ and the squared advantage is $\mathrm{Var}[L^d]$. By the total variance formula, $\mathrm{Var}[L^d] = \mathbb{E}\,\mathrm{Var}[L^d|H] + \mathrm{Var}\,\mathbb{E}[L^d|H]$. For fixed h, L^d has the Fourier expansion

$$L^d(\,\cdot\,|h) = \sum_{|S|\le d} \mathbb{E}[L(R|h)\chi_S(R)]\chi_S = \sum_{|S|\le d} \mathbb{E}[\chi_S(F_H)]\chi_S,$$

In particular, $\mathbb{E}[L^d(\cdot|h)] = 1$ for every h and $\mathrm{Var}\,\mathbb{E}[L^d|H] = 0$. It follows that the advantage is

$$\mathrm{Var}[L^d] = \mathbb{E}\,\mathrm{Var}[L^d|H] = \mathbb{E} \sum_{1\le|S|\le d} \mathbb{E}[\chi_S(F_H)|H]^2.$$

As $\chi_S(F_H)$ is nonzero only when F_H is nonuniform and it is at most one otherwise, the right hand side is at most the expected number of biased subsets of F_H. □

6 The Encryption Scheme

We present a general construction that exhibits a tradeoff between the parameter k that governs the locality of encryption and the size of the hyperloop ℓ_0.

We will assume that the vertices in a hyperedge are ordered. Let

- Q be a random 3-hypergraph with n vertices and hyperedge probability $C\sqrt{d}n^{-3/2-\delta}$,
- L_0 be a fixed 3-hypergraph on $\ell_0 = 0.09 \cdot 2^k \log n$ vertices satisfying Claim 13,
- L consists of $t = O(1/\beta^2 \log 1/\delta)$ vertex-disjoint copies of L_0, $\beta = (1 - 2^{-k+1})^{\ell_0}$,
- P be the m-edge hypergraph union of Q and a copy of L planted on a random subset of $3\ell/2$ vertices of Q,
- H be $(k + 3)$-hypergraph obtained by extending each hyperedge of P with k random vertices,
- $F\colon \{0,1\}^n \to \{0,1\}^m$ be the function obtained by evaluating the $(k + 3)$-ary predicate $x_1 \oplus x_2 \oplus x_3 \oplus (y_1 \wedge \cdots \wedge y_k)$ on all sequences of input bits indexed by hyperedges in H.

The key generation procedure outputs H as the public key and disjoint ℓ_0-subsets $S_1, \ldots, S_t$ of $\{1, \ldots, m\}$ indexing the copies of L_0 in P as the secret key.

To encrypt a 0, output $y = F(x)$ for a random x. To encrypt a 1, output a random string of length m.

To decrypt y, calculate the parities $z_i = \oplus_{j\in S_i} y_j$ for all $1 \le j \le t$. If more than $(1 + \beta)t/2$ of them are zero output 0, otherwise output 1.

Call the public key good if all extensions of the hyperedges in L are pairwise disjoint. By a union bound the public key is good except with probability $O(\ell^2k^2/n) = n^{-\Omega(1)}$.

Claim 15. *Assuming H is good, decryption is correct except with probability δ.*

Proof. For an encryption of 1, the bits $z_1, \dots, z_t$ are independent random so the probability that more than $(1+\beta)t/2$ of them are zero is at most δ by Chernoff bounds.

For an encryption of 0, each bit z_i evaluates to an ℓ_0-XOR of disjoint k-ANDs so it has bias β. As $z_1, \dots, z_\ell$ are independent the probability that fewer than $(1+\beta)t/2$ are zero is at most δ again. □

Theorem 16. *If f has degree less than $(1-\epsilon)\ell_0$ and bounded 4-norm, the distinguishing advantage $\mathbb{E}[f(P, F_P)] - \mathbb{E}[f(Q, R)]$ is $n^{-\Omega(1)}$.*

We do not know if a bounded variance assumption on f would suffice.

Proof. We may assume $\mathbb{E}[f(Q, R)] = 0$. By Proposition 9, for every g of degree at most $d = (1-\epsilon)\ell_0$, $\mathbb{E}[g(P)] - \mathbb{E}[g(Q)] \le n^{-\Omega(1)}\sqrt{\mathrm{Var}[g(Q)]}$. Given f of degree d let $g(G) = \mathbb{E}[f(G, F_G)|G]$. Then g has the same degree as f and

$$\mathbb{E}[f(P, F_P)] - \mathbb{E}[f(Q, F_Q)] \le n^{-\Omega(1)}\sqrt{\mathrm{Var}[f(Q, F_Q)]}.$$

By Proposition 14 applied to f^2,

$$|\mathbb{E}[f(Q, F_Q)^2] - \mathbb{E}[f(Q, R)^2]| \le n^{-\Omega(1)}\sqrt{\mathrm{Var}[f(Q, R)^2]}.$$

By the boundedness of the 4-norm of f,

$$\mathrm{Var}[f(Q, F_Q)^2] \le \mathrm{Var}[f(Q, R)^2] + n^{-\Omega(1)}$$

so $\mathbb{E}[f(P, F_P)] - \mathbb{E}[f(Q, F_Q)] = n^{-\Omega(1)}$. By Proposition 14 applied to f this time,

$$\mathbb{E}[f(Q, F_Q)] - \mathbb{E}[f(Q, R)] = n^{-\Omega(1)}\sqrt{\mathrm{Var}[f(Q, R)]} = O(n^{-\Omega(1)}).$$

The claim follows by the triangle inequality. □

Acknowledgements

We thank Caicai Chen, Yuval Ishai, and Chris Jones for their advice and feedback. Part of this work done when the first and second authors visited Bocconi University. The first named author was supported by RGC GRF CUHK14209920 and the third named author was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 101019547) and Cariplo CRYPTONOMEX grant. The second named author is supported by NSF CAREER Award #2047933, Alfred P. Sloan Fellowship and a Google Research Scholar Award.

References

[ABR16] Benny Applebaum, Andrej Bogdanov, and Alon Rosen. A dichotomy for local small-bias generators. *J. Cryptology*, 29(3):577–596, July 2016.

[ABW10] Benny Applebaum, Boaz Barak, and Avi Wigderson. Public-key cryptography from different assumptions. In *Proceedings of the Forty-Second ACM Symposium on Theory of Computing*, STOC '10, page 171–180, New York, NY, USA, 2010. Association for Computing Machinery.

[AL18] Benny Applebaum and Shachar Lovett. Algebraic attacks against random local functions and their countermeasures. *SIAM Journal on Computing*, 47(1):52–79, 2018.

[BB20] Matthew S. Brennan and Guy Bresler. Reducibility and statistical-computational gaps from secret leakage. In Jacob D. Abernethy and Shivani Agarwal, editors, *Conference on Learning Theory, COLT 2020, 9-12 July 2020, Virtual Event [Graz, Austria]*, volume 125 of *Proceedings of Machine Learning Research*, pages 648–847. PMLR, 2020.

[BHK$^+$19] Boaz Barak, Samuel B. Hopkins, Jonathan A. Kelner, Pravesh K. Kothari, Ankur Moitra, and Aaron Potechin. A nearly tight sum-of-squares lower bound for the planted clique problem. *SIAM J. Comput.*, 48(2):687–735, 2019.

[BR13] Quentin Berthet and Philippe Rigollet. Complexity theoretic lower bounds for sparse principal component detection. In Shai Shalev-Shwartz and Ingo Steinwart, editors, *Proceedings of the 26th Annual Conference on Learning Theory*, volume 30 of *Proceedings of Machine Learning Research*, pages 1046–1066, Princeton, NJ, USA, 12–14 Jun 2013. PMLR.

[DNR04] Cynthia Dwork, Moni Naor, and Omer Reingold. Immunizing encryption schemes from decryption errors. In *EUROCRYPT*, 2004.

[FKO06] Uriel Feige, Jeong Han Kim, and Eran Ofek. Witnesses for non-satisfiability of dense random 3cnf formulas. In *2006 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS'06)*, pages 497–508, 2006.

[GM84] Shafi Goldwasser and Silvio Micali. Probabilistic encryption. *J. Comput. Syst. Sci.*, 28(2):270–299, 1984.

[Gol11] Oded Goldreich. *Candidate One-Way Functions Based on Expander Graphs*, pages 76–87. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

[HKP$^+$17] Samuel B. Hopkins, Pravesh K. Kothari, Aaron Potechin, Prasad Raghavendra, Tselil Schramm, and David Steurer. The power of sum-of-squares for detecting hidden structures. In Chris Umans, editor, *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, pages 720–731. IEEE Computer Society, 2017.

[HLW06] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bull. Amer. Math. Soc.*, 43(04):439–562, August 2006.

[Hop18] Samuel Hopkins. *Statistical Inference and the Sum of Squares Method.* PhD thesis, Cornell University, 2018.

[HS17] Samuel B. Hopkins and David Steurer. Efficient bayesian estimation from few samples: Community detection and related problems. In Chris Umans, editor, *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, pages 379–390. IEEE Computer Society, 2017.

[HWX15] Bruce Hajek, Yihong Wu, and Jiaming Xu. Computational lower bounds for community detection on random graphs. In *Proceedings of The 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 899–928, Paris, France, 03–06 Jul 2015. PMLR.

[KWB19] Dmitriy Kunisky, Alexander S. Wein, and Afonso S. Bandeira. Notes on computational hardness of hypothesis testing: Predictions using the low-degree likelihood ratio, 2019.

[MST06] Elchanan Mossel, Amir Shpilka, and Luca Trevisan. On epsilon-biased generators in nc^{0}. *Random Struct. Algorithms*, 29(1):56–81, 2006.

[OW14] Ryan O'Donnell and David Witmer. Goldreich's prg: Evidence for near-optimal polynomial stretch. *2014 IEEE 29th Conference on Computational Complexity (CCC)*, pages 1–12, 2014.