

On ε -Biased Generators in NC^0

Elchanan Mossel*

Amir Shpilka †

Luca Trevisan‡

October 23, 2007

Abstract

Cryan and Miltersen [8] recently considered the question of whether there can be a pseudorandom generator in NC^0 , that is, a pseudorandom generator that maps n -bit strings to m -bit strings such that every bit of the output depends on a constant number k of bits of the seed.

They show that for $k = 3$, if $m \geq 4n + 1$, there is a distinguisher; in fact, they show that in this case it is possible to break the generator with a *linear test*, that is, there is a subset of bits of the output whose XOR has a noticeable bias.

They leave the question open for $k \geq 4$. In fact they ask whether every NC^0 generator can be broken by a statistical test that simply XORs some bits of the input. Equivalently, is it the case that no NC^0 generator can sample an ε -biased space with negligible ε ?

We give a generator for $k = 5$ that maps n bits into cn bits, so that every bit of the output depends on 5 bits of the seed, and the XOR of every subset of the bits of the output has bias $2^{-\Omega(n/c^4)}$. For large values of k , we construct generators that map n bits to $n^{\Omega(\sqrt{k})}$ bits such that every XOR of outputs has bias $2^{-n^{\frac{1}{2\sqrt{k}}}}$.

We also present a polynomial-time distinguisher for $k = 4$, $m \geq 24n$ having constant distinguishing probability. For large values of k we show that a linear distinguisher with a constant distinguishing probability exists once $m \geq \Omega(2^k n^{\lceil k/2 \rceil})$.

Finally, we consider a variant of the problem where each of the output bits is a degree k polynomial in the inputs. We show there exists a degree $k = 2$ pseudorandom generator for which the XOR of every subset of the outputs has bias $2^{-\Omega(n)}$ and which maps n bits to $\Omega(n^2)$ bits.

1 Introduction

A pseudorandom generator is an efficient deterministic procedure that maps a shorter random input into a longer output that is indistinguishable from the uniform distribution by resource-bounded observers.

A formalization of the above informal definition is to consider polynomial-time procedures G mapping n bits into $m(n) > n$ bits such that for every property P computable by a family of polynomial-size circuits we have that the quantity

$$\left| \Pr_{z \in \{0,1\}^{m(n)}} [P(z) = 1] - \Pr_{x \in \{0,1\}^n} [P(G(x))] \right|$$

*Department of Statistics, U.C. Berkeley, CA 94720-3869. Email: mossel@stat.berkeley.edu. Supported by a Miller fellowship

†Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot, Israel. Email: amir.shpilka@weizmann.ac.il. Supported by National Security Agency (NSA) and Advanced Research and Development Activity (ARDA) under Research Office (ARO) contract no. DAAD19-01-1-0506, and by the Koshland fellowship.

‡Computer Science Division, U.C. Berkeley, CA 94720-1776. Email: luca@cs.berkeley.edu. Supported by NSF Grant CCR-9984783/CCR-0406156, US-Israel BSF grant 2002246, a Sloan Research Fellowship and an Okawa Foundation Grant.

goes to zero faster than any inverse polynomial in n . The existence of such a procedure G is equivalent to the existence of one-way functions [15], pseudorandom functions [11] and pseudorandom permutations [23].

What are the minimal computational requirements needed to compute a pseudorandom generator? Linial et al. [20] prove that pseudorandom functions cannot be computed in AC^0 (constant-depth circuits with NOT gates and unbounded fan-in AND and OR gates). To be precise, the results in [20] only rule out security against adversaries running in time $O(n^{(\log n)^{O(1)}})$. Their result does not rule out the possibility that pseudorandom generators could be computed in AC^0 , since the transformation of pseudorandom generators into pseudorandom functions does not preserve bounded-depth.

Kharitonov [19] shows that a pseudorandom generator with superlinear stretch can be computed in NC^1 , that is, it can be computed by a circuit of polynomial size, logarithmic depth, and gates of constant fan-in. (It is known that NC^1 properly contains AC^0 .) Impagliazzo and Naor [17] present a candidate pseudorandom generator in AC^0 . Goldreich [12] suggests a candidate one-way function in NC^0 . Recall that NC^0 is the class of functions computed by bounded-depth circuits with NOT gates and bounded fan-in AND and OR gates. In an NC^0 function, every bit of the output depends on a constant number of bits of the inputs. While it is easy to see that there can be no one-way function such that every bit of the output depends on only two bits of the input (as finding an inverse can be formulated as a 2SAT problem) it still remains open whether there can be a one-way function such that every bit of the output depends on only three bits of the input. Applebaum *et al.* [1] have very recently provided evidence that such one-way functions exist.

Cryan and Miltersen [8] consider the question of whether there can be pseudorandom generators in NC^0 , that is, whether there can be a pseudorandom generator such that every bit of the output depends only on a constant k number of bits of the input. They present a distinguisher in the case $k = 3, m > 4n$, and they observe that their distinguisher is a *linear* distinguisher, that is, it simply XORs a subset of the bits of the output. Cryan and Miltersen ask whether there is any pseudorandom generator in NC^0 when m is superlinear in n . Specifically, they ask whether the following is the case: that for every constant k , and for every generator for which m is super-linear in n and for which every output bit depends on at most k bits of the input, a linear distinguisher exists.

In order to formulate an equivalent version of this problem, we introduce the notion of a ε -biased distribution.

Definition 1. For $\varepsilon > 0$, we say that a random variable $X = (X_1, \dots, X_m)$ ranging over $\{0, 1\}^m$ is ε -biased if for every subset $S \subseteq [m]$ we have $1/2 - \varepsilon \leq \Pr[\bigoplus_{i \in S} X_i = 0] \leq 1/2 + \varepsilon$.

It is known [27, 3] that an ε -biased distribution can be sampled by using only $O(\log(m/\varepsilon))$ random bits, which is tight up to the constant in the big-Oh.

The problem of [8] can therefore be formulated by asking whether there exists any ε -biased generator in NC^0 that samples an m -bit ε -biased distribution starting from, say, $o(m)$ random bits and a negligible ε .

Our Results

We first extend the result of Cryan and Miltersen by giving a (non linear) distinguisher for the case $k = 4, m \geq 24n$.

Theorem 2. Let $G = (g_1, \dots, g_m) : \{0, 1\}^n \rightarrow \{0, 1\}^m$ be a map such that each g_i depends on at most 4 coordinates of the input and $m \geq 24n$. Then there exists a polynomial time algorithm which distinguishes between G and a random string with constant distinguishing probability. More precisely, the algorithm will output “yes” for the output of the generator G with probability $\Omega(1)$, and for a random string with probability $e^{-\Omega(m)}$.

Our distinguisher has a constant distinguishing probability, which we show to be impossible to achieve with linear distinguishers. Our distinguisher uses semidefinite programming and uses an idea similar to the “correlation attacks” used in practice against stream ciphers.

For all k , it is trivial that a distinguisher exists for $m \geq 2^{2^k} \binom{n}{k}$ (the number of functions on k bits), and it is easy to see that a distinguisher exist when $m \geq k \binom{n}{k}$ (as there is a linear dependence among the output bits in this case). We show using a duality lemma proven in [25] that in fact, a distinguisher with a constant distinguishing probability exists once $m \geq \Omega(2^k n^{\lceil k/2 \rceil})$ by proving

Theorem 3. *For every integer $0 < k$ and any $0 < \varepsilon < 2^{-2^{k-1}}$, if $G = (g_1, \dots, g_m)$ is an ε -biased pseudorandom generator, where each of the g_i 's depend on at most k bits, then*

$$m \leq \sum_{t=0}^{\lceil \frac{k}{2} \rceil} \binom{n}{t} 2^{2(k-t)} \leq k 2^{2k} \binom{n}{\lceil \frac{k}{2} \rceil}.$$

Then we present an ε -biased generator mapping n bits into cn bits such that $\varepsilon = 1/2^{\Omega(n/c^4)}$ and every bit of the output depends only on $k = 5$ bits of the seed, i.e., we prove

Theorem 4. *For every c and sufficiently large n , there is a generator in NC_5^0 mapping n bits into cn bits and sampling an ε -biased distribution, where $\varepsilon = 2^{-n/O(c^4)}$.*

The main idea in the construction is to develop a generator with $k = 3$ that handles well linear tests that XOR a *small* number of bits, and then develop a generator with $k = 2$ that handles well linear tests that XOR a *large* number of bits. The final generator outputs the bitwise XOR of the outputs of the two generators, on two independent seeds.

The generator uses a kind of unique-neighbor expander graphs that are shown to exist using the probabilistic method, but that are not known to be efficiently constructible, so the generator is in NC^0 but not in *uniform* NC^0 .

Later we present similar constructions for large values of k . We write $f(n, k) = O_k(g(n))$ if $f(n, k) \leq h(k)g(n)$ for some function h ; similarly we will use the notation o_k .

Theorem 5. *Let k be a positive integer. There exists an ε -biased generator in NC_k^0 from n bits to*

$$\left(\frac{n}{\sqrt{k} - 6} \right)^{\frac{\sqrt{k}}{2} - 3} = n^{\sqrt{k}(\frac{1}{2} - o_k(1))}$$

bits whose bias, ε , is at most

$$\exp \left(- \frac{n^{\frac{1}{2\sqrt{k}}}}{4 \times 2^{\sqrt{k}}} \right).$$

Note the gap for large values of k between our constructions that output $n^{(\sqrt{k}/2)(1-o_k(1))}$ bits, and the bounds showing a distinguisher exists for generators that output $n^{(k/2)(1+o_k(1))}$ bits.

Finally, we begin a study of the question of whether there are pseudorandom generators with superlinear stretch such that each bit of the output is a function of the seed expressible as a degree- k polynomial over $GF(2)$, where k is a constant. This is a generalization of the main question addressed in this paper, since a function depending on only k inputs can always be expressed as a degree- k polynomial. Furthermore, low-degree polynomials are a standard class of “low complexity” functions from an algebraic perspective. In our NC_5^0 construction of an ε -biased generator with exponentially small ε and superlinear stretch, every bit of the output is a degree-2 polynomial. We show that

Theorem 6. $\forall 1 \leq m \leq n$ there exists an ε -biased generator $G = (g_1, \dots, g_t) : \{0, 1\}^n \mapsto \{0, 1\}^t$, $t = \lfloor \frac{n}{2} \rfloor \cdot m$, such that g_i is a degree 2 polynomial, and the bias of any non trivial linear combination of the g_i 's is at most $2^{-\frac{n-2m}{4}}$.

Later Results and Open Questions

Applebaum *et al.* [1] have recently made substantial progress on the main questions left open by our work about the cases $k = 3, 4$.

In the case $k = 3$, Applebaum *et al.* [1] present a construction of an ε -biased generator with $m = (1 + \alpha) \cdot n$, where $\alpha > 0$ is an absolute constant. They also show that under relatively general assumptions, there are one-way functions such that every bit of the output depends on only 3 bits of the input.

In the case $k = 4$, Applebaum *et al.* [1] present a construction of a pseudorandom generator with $m = n + n^\alpha$, where α can be chosen to be any constant smaller than 1. The generator is secure under the assumption that there exists pseudorandom generators in $\bigoplus L/\text{poly}$, which is a fairly general assumption.

It remains open whether a cryptographically strong generator can be realized in the case $k = 3$, whether a cryptographically strong generator with linear stretch can be realized in the case $k = 4$, and whether a cryptographically strong generator with superlinear stretch can be realized in the case $k = 5$.

Another important open problem which may be more accessible is to understand the right asymptotic for ε -biased generators for large k . It is tempting to conjecture that either the upper bound $n^{O(k)}$ or the lower bound $n^{\Omega(\sqrt{k})}$ is actually tight.

Organization

In section 2 we review the analysis for the case $k = 3$ of [8]. In section 3 we give a distinguisher for the case $k = 4$. In section 4 we prove an upper bound on the length of the output of an ε -biased generator in NC_k^0 .

In section 5 we construct an ε -biased generator for the cases $k = 4, 5$. The results for larger k are discussed in section 6. In section 7 we explicitly construct an ε -biased generator such that every bit of the output is a polynomial of degree 2.

An extended abstract reporting on the results here appeared in [26].

2 Review of the Case $k = 3$

In this section we summarize the main result of [8]. We also generalize some of the arguments of [8] that are needed for our results.

2.1 Preliminaries

We say that a function $g : \{0, 1\}^n \rightarrow \{0, 1\}$ is *balanced* if $\Pr_x[g(x) = 1] = 1/2$. We say that a function $g : \{0, 1\}^n \rightarrow \{0, 1\}$ is *unbiased* towards a function $f : \{0, 1\}^n \rightarrow \{0, 1\}$ if $\Pr_x[g(x) = f(x)] = 1/2$, and that it is *biased* towards f (or *correlated* with f) otherwise. A function $g : \{0, 1\}^n \rightarrow \{0, 1\}$ is *affine* if there are values $a_0, \dots, a_n \in \{0, 1\}$ such that $g(x_1, \dots, x_n) = a_0 \oplus a_1 x_1 \oplus \dots \oplus a_n x_n$, it is *non-affine* otherwise.

The following lemma was proved by case analysis for $k = 3$ in [8], and the case $k = 4$ could also be derived from a case analysis appearing in [8] (but it is not explicitly stated). The proof of the general case follows using the Fourier representation of boolean functions.

The Fourier representation is easier to work with when considering functions from $\{\pm 1\}^n \rightarrow \{\pm 1\}$. For a boolean function $f : \{0, 1\}^k \rightarrow \{0, 1\}$ we write F for the function $F : \{\pm 1\}^k \rightarrow \{\pm 1\}$ defined as

$$F((-1)^{x_1}, \dots, (-1)^{x_k}) = (-1)^{f(x_1, \dots, x_k)}. \quad (1)$$

For the boolean functions f, g, h discussed in this section, the functions F, G, H will be the corresponding mappings to $\{\pm 1\}$. For a set $S \subseteq [k]$, we let $U_S : \{\pm 1\}^k \rightarrow \{\pm 1\}$ be defined as $U_S(X) = \prod_{i \in S} X_i$, that is U_S is the character corresponding to S . It is well known that $\{U_S\}_{S \subseteq [k]}$ is an orthonormal basis for the space of functions from $\{\pm 1\}^k$ to \mathbb{R} with respect to the inner product

$$\langle F, G \rangle = \frac{1}{2^k} \sum_{x \in \{0,1\}^k} F(x) \cdot G(x).$$

We write $F(X) = \sum_S \hat{F}(S) U_S(X)$ for the representation of F in the basis $\{U_S\}$. Because of orthonormality, the coefficients $\hat{F}(S)$ satisfy the relation $\hat{F} = \langle F, U_S \rangle$.

Note that if f, g are boolean functions and F, G are defined as in (1), then $\Pr[f(x) = g(x)] = \Pr[F(x) = G(x)] = 1/2 + 1/2 \langle F, G \rangle$. In particular, f and g are correlated if and only if $\langle F, G \rangle \neq 0$.

Lemma 7. *Let $g : \{0, 1\}^n \rightarrow \{0, 1\}$ be a non-affine function that depends on only k variables. Then*

- *There exists an affine function on at most $k - 2$ variables that is correlated with g .*
- *Let l be the affine function that is biased towards g and that depends on a minimal number of variables. That is, for some d , l depends on d variables, $\Pr_x[g(x) = l(x)] > 1/2$, and g is unbiased towards affine functions that depend on less than d variables.*

Then $\Pr_x[g(x) = l(x)] \geq 1/2 + 2^{d-k}$.

Proof. • Let $f : \{0, 1\}^k \rightarrow \{0, 1\}$ be a non-affine function. We prove that there exists a set S of size at most $k - 2$ such that $\hat{F}(S) \neq 0$. This implies that F is correlated with U_S and therefore that f is correlated with $\bigoplus_{i \in S} x_i$ as needed.

Look at the function $h(x_1, \dots, x_k) = f(x_1, \dots, x_k) \oplus \bigoplus_{i=1}^k x_i$. Since f is non-affine, h is not a constant function. Let H be the $\{\pm 1\}$ representation of h . As the $\{\pm 1\}$ representation of $\bigoplus_{i=1}^k x_i$ is $U_{[k]}$, we get that H has the Fourier representation

$$H = U_{[k]} \cdot F = U_{[k]} \cdot \sum_{S \subseteq [k]} \hat{F}(S) U_S = \sum_{S \subseteq [k]} \hat{F}(S) U_{[k] \setminus S} = \sum_S \hat{F}([k] \setminus S) U_S.$$

It therefore suffices to prove that $U_{[k]} \cdot F$ has a coefficient $\hat{F}(S) \neq 0$ with $|S| \geq 2$. We will prove that any function which depends on more than one bit, has a non-zero coefficient with $|S| \geq 2$. This will prove the first part, since if h depends on at most one bit then f is affine.

Indeed, assume the contradiction

$$F = a_0 + \sum_i a_i U_{\{i\}}$$

For a \pm vector X , write X^i for the vector where the i 'th coordinate of X is multiplied by -1 . Note that for all i and all X , it holds that $2a_i = F(X) - F(X^i) \in \{0, \pm 2\}$, which implies that $a_i \in \{0, \pm 1\}$. Parseval's inequality implies that $\sum a_i^2 = 1$. We therefore conclude that $F(X)$ depends on one bit as needed. This completes the proof of the first claim.

- Note that f is correlated with $\bigoplus_{i \in S} x_i$ if and only if $\hat{F}(S) \neq 0$. Moreover,

$$\Pr[f(x) = \bigoplus_{i \in S} x_i] = \frac{1 + \hat{F}(S)}{2}.$$

The claim will therefore follow once we prove that if $F = \sum_{|S| \geq d} \hat{F}(S)U_S$, and $\hat{F}(S) \neq 0$ for a set S of size d , then $|\hat{F}(S)| \geq 2^{d+1-k}$.

By looking at $U_{[k]}F$ instead of F , it suffices to prove that if

$$F = \sum_{|S| \leq k-d} \hat{F}(S)U_S, \quad (2)$$

and S' is a set of size $k-d$ such that $\hat{F}(S') \neq 0$, then $|\hat{F}(S')| \geq 2^{d-k+1}$. In order to prove the last claim, define

$$A(X) = \sum_{T \subseteq S'} (-1)^{|T|} F(X^T) = \sum_{T \subseteq S'} (-1)^{|T|} \sum_{S \subseteq [k]} \hat{F}(S)U_S(X^T) = \sum_{S \subseteq [k]} \hat{F}(S) \sum_{T \subseteq S'} (-1)^{|T|} U_S(X^T),$$

where X^T is X where the coordinates at T are flipped (multiplied by -1). It is then clear that A obtains an *even* integer value in the interval $[-2^{k-d}, 2^{k-d}]$.

On the other hand, if S does not contain S' and $j \in S' \setminus S$, then for all X

$$\begin{aligned} \sum_{T \subseteq S'} (-1)^{|T|} U_S(X^T) &= \sum_{T \subseteq S', j \notin T} (-1)^{|T|} U_S(X^T) + \sum_{T \subseteq S', j \in T} (-1)^{|T|} U_S(X^T) \\ &= \sum_{T \subseteq S', j \notin T} U_S(X^T) ((-1)^{|T|} + (-1)^{|T|+1}) = 0. \end{aligned}$$

Since $\hat{F}(S) = 0$ for all S strictly containing S' , it follows that

$$A(X) = \hat{F}(S') \sum_{T \subseteq S'} (-1)^{|T|} u_{S'}(X^T) = 2^{k-d} \hat{F}(S') u_{S'}(X).$$

We therefore conclude that $\hat{F}(S')$ is of the form $\frac{2i}{2^{k-d}}$, for some integer $i \in [-2^{k-d-1}, 2^{k-d-1}]$. In particular, since $\hat{F}(S') \neq 0$, it follows that $|\hat{F}(S')| \geq 2^{-d+k+1}$ as needed. □

For example, for $k = 3$, a non-affine function g is either unbalanced, or it is biased towards one of its inputs; in the latter case it agrees with an input bit (or with its complement) with probability at least $3/4$.

For $k = 4$, a function g either is affine, or it is unbalanced, or it has agreement at least $5/8$ with an affine function that depends on only one input bit, or it has agreement at least $3/4$ with an affine function that depends on only two input bits.

2.2 The Case $k = 3$

Let $G : \{0, 1\}^n \rightarrow \{0, 1\}^m$ be a generator and let $g_i : \{0, 1\}^n \rightarrow \{0, 1\}$ be the i -th bit of the output of the generator. Suppose each g_i depends on only three bits of the input.

Suppose that one of the g_i is not a balanced function. Then we immediately have a distinguisher.

Suppose that more than n of the g_i are affine. Then one of them is linearly dependent on the others, and we also have a distinguisher.

It remains to consider the case where at least $m - n$ of the functions g_i are balanced and not affine. Let I be the set of i for which g_i is as above. Then, by lemma 7, for each such g_i there is a affine function l_i that depends on

only one bit, such that g_i agrees with l_i on a $3/4$ fraction of the inputs. By replacing g_i with $g_i \oplus 1$ when needed, we may assume that each such g_i has correlation at least $3/4$ with one of the bits of its input. The following lemma now implies a constant distinguishing probability once $m \geq 4n + 1$. While the above analysis uses the same ideas as in [8], it is slightly better because we achieve constant bias instead of inverse polynomial bias. We first prove a very general lemma that will be also used in later sections, and then we derive the conclusion that we need for the case of $k = 3$.

Lemma 8. *For every $\delta > 0$ there are constants $c_\delta \leq \lceil \frac{1}{\delta^2} \rceil - 1 \leq \frac{1}{\delta^2}$ and $\varepsilon_\delta \geq \frac{3\delta^2}{4}$ such that the following holds. Let $G : \{0, 1\}^n \rightarrow \{0, 1\}^m$, and let $G(x) = (g_1(x), \dots, g_m(x))$. Let L be a set of functions and suppose that each function $g_i(x)$ agrees with an element of L or with its complement with probability at least $1/2 + \delta$. In other words, for every g_i there exists $f \in L$ such that*

$$\Pr_x[g_i(x) = f(x)] \geq \frac{1}{2} + \delta \text{ or}$$

$$\Pr_x[g_i(x) \neq f(x)] \geq \frac{1}{2} + \delta.$$

Assume that $m \geq 1 + c_\delta |L|$. Then there are $i \neq j$ such that $g_i \oplus g_j$ has bias at least ε_δ . Moreover, $c_{1/4} \leq 3$ and $c_{1/8} \leq 9$.

Proof. By the pigeonhole principle there is a function $f \in L$ and a set of indices $C \subseteq [m]$, such that $|C| \geq \lceil \frac{m}{|L|} \rceil$, and for every $i \in C$, g_i or $1 - g_i$ is correlated with f . Assume w.l.o.g. that for every $i \in C$, g_i is correlated with f (otherwise replace g_i with $1 - g_i$).

Define the random variable

$$Z(x) = |\#\{i \in C : g_i(x) = 0\} - \#\{i \in C : g_i(x) = 1\}|.$$

Consider the expectation of $Z(x)$ (where x is uniformly chosen from $\{0, 1\}^n$). We have that

$$\begin{aligned} \mathbf{E}[Z(x)] &= \mathbf{E}[|\#\{i \in C : g_i(x) = f(x)\} - \#\{i \in C : g_i(x) \neq f(x)\}|] \\ &\geq \mathbf{E}[\#\{i \in C : g_i(x) = f(x)\}] - \mathbf{E}[\#\{i \in C : g_i(x) \neq f(x)\}] \geq |C| \cdot \left(\left(\frac{1}{2} + \delta \right) - \left(\frac{1}{2} - \delta \right) \right) = 2\delta|C|. \end{aligned}$$

We conclude that for $|C| = O(\delta^{-2})$, the restriction of the generator to C has constant statistical distance from the uniform distribution over $|C|$ bits, for which that average value of Z is $O(\sqrt{|C|})$. By the Vazirani XOR lemma [31] (see [10] for an excellent exposition of the XOR lemma), it also follows that the XOR of some subset of the bits of C has bias $\Omega(2^{-|C|}) = 2^{-O(\delta^{-2})}$. However we would like to obtain a better dependence between δ and ε .

For $i, j \in C$ define $Z_{i,j}(x)$ to be 1 if $g_i(x) = g_j(x)$ and -1 otherwise. Note that $E[Z_{i,j}]$ equals twice the bias of $g_i \oplus g_j$. Clearly $Z_{i,i} = 1$. We have that $Z(x)^2 = \sum_{i,j} Z_{i,j}$. In particular we get that

$$\mathbf{E} \left[\sum_{i,j} Z_{i,j}(x) \right] = \mathbf{E} [Z(x)^2] \geq \mathbf{E}[Z(x)]^2 \geq 4\delta^2|C|^2.$$

Hence for $|C| = \lceil \frac{1}{\delta^2} \rceil$ we get that

$$\mathbf{E} \left[\sum_{i,j} Z_{i,j}(x) \right] \geq 4|C|.$$

As $\mathbf{E}[\sum_i Z_{i,i}] = |C|$, it follows that $\mathbf{E}[\sum_{i \neq j} Z_{i,j}] \geq 3|C|$, and so there must be $i \neq j \in C$ such that

$$\mathbf{E}[Z_{i,j}] \geq \frac{3|C|}{|C| \cdot (|C| - 1)} \geq \frac{3\delta^2}{2}.$$

In other words, $g_i \oplus g_j$ has a $\frac{3\delta^2}{4}$ bias. Thus taking $m = 1 + |L| \cdot (\lceil \frac{1}{\delta^2} \rceil - 1)$ we obtain $c_\delta = \lceil \frac{1}{\delta^2} \rceil - 1$.

We now consider two special cases.

Let $|C| = 4, \delta = \frac{1}{4}$. By the above argument we get that $\mathbf{E}[Z(x)] \geq 2 \times \frac{1}{4} \times |C| = 2$. On the other hand, for the uniform distribution on 4 bits the average of $Z(x)$ is

$$\frac{2}{16} \left(2 \times \binom{4}{1} + 4 \times \binom{4}{0} \right) = \frac{3}{2} < 2 = 2.$$

Thus, if $|C| = 4$ we get by Vazirani's XOR lemma that some subset of the g_i 's has some constant bias, so we can set $c_{1/4} = 3$.

Similarly, when $|C| = 10$ the average of $Z(x)$ for the uniform distribution is

$$\frac{2}{2^{10}} \sum_{i=0}^4 (10 - 2i) \binom{10}{i} = \frac{2520}{1024} < 2 \times \frac{1}{8} \cdot 10,$$

so we can set $c_{1/8} = 9$. □

To conclude the case of $k = 3$ we note that if $m \geq 1 + 4n$, and the output of the generator contains at most n affine functions then at least $1 + 3n$ output bits that are not affine and so we can apply Lemma 8, where $L = \{\pi_1, \dots, \pi_n\}$ is the set of n "projection" functions $\pi_i()$ such that $\pi_i(x_1, \dots, x_n) = x_i$. The consequence of Lemma 8 is that two of the output bits are correlated.

3 Distinguisher for the Case $k = 4$

In this section we construct a distinguisher for $k = 4$. We restate Theorem 2.

Theorem. *Let $G = (g_1, \dots, g_m) : \{0, 1\}^n \rightarrow \{0, 1\}^m$ be a map such that each g_i depends on at most 4 coordinates of the input and $m \geq 24n$. Then there exists a polynomial time algorithm which distinguishes between G and a random string with constant distinguishing probability. More precisely, the algorithm will output "yes" for the output of the generator G with probability $\Omega(1)$, and for a random string with probability $e^{-\Omega(m)}$.*

- The first case we consider is where there are more than $0.001m$ of the g_i that are unbalanced. Suppose that g_1, \dots, g_p are unbalanced and $p \geq 0.001m$. Then there exist fixed bits b_1, \dots, b_p such that $\Pr[g_i = b_i] \geq 9/16$. Thus by Markov's inequality:

$$\Pr_{z \in \{0,1\}^n} \left[\frac{|\{i \mid g_i(z) = b_i\}|}{p} \geq \frac{17}{32} \right] \geq \frac{1}{32}.$$

On the other hand, if r_1, \dots, r_p are chosen uniformly at random, then

$$\Pr \left[\frac{|\{i \mid r_i = b_i\}|}{p} \geq \frac{17}{32} \right] \leq e^{-\Omega(m)}$$

by Chernoff's inequality.

- The second case is where more than $n + 0.001m$ of the g_i are linear. In this case we can write at least $0.001m$ independent linear combinations in the output bits of the generator that hold with probability 1. The probability that these combinations hold for truly random bits is $2^{-0.001m}$. Thus the statement of the theorem follows in this case as well.
- If one of the g_i is biased towards one of the bits of its input, then it follows from Lemma 7 that it must agree with that bit or its complement with probability at least $5/8$. Suppose that more than $c_{1/8}n = 9n + 0.001m$ of the functions g_i have bias towards one bit. Then by the proof of Lemma 8, there exists at least $p \geq 0.0001m$ disjoint sets S_1, \dots, S_p of the g_i 's such that $|S_r| \leq 10$ and $\bigoplus_{i \in S_r} g_i$ has bias at least 2^{-10} bias towards a constant bit b_r for all $1 \leq r \leq p$. Thus, as in the first case,

$$\Pr_{z \in \{0,1\}^n} \left[\frac{|\{r \mid \bigoplus_{i \in S_r} g_i(z) = b_r\}|}{p} \geq \frac{1}{2} + 2^{-11} \right] \geq 2^{-11}$$

and from Chernoff's bound it follows that if r_i are truly random then

$$\Pr \left[\frac{|\{r \mid \bigoplus_{i \in S_r} r_i = b_r\}|}{p} \geq \frac{1}{2} + 2^{-11} \right] \leq e^{-\Omega(m)}.$$

Thus, the proof follows in this case as well.

- It remains to consider the case where at least $0.997m - 10n$ of the functions are balanced, non-linear, and unbiased towards single bits. Following [8], we call such functions *problematic*. It follows from Lemma 7 that for each problematic g there is an affine function l of two variables that agrees with g on a $3/4$ fraction of the inputs. Again, by replacing g_i by $g_i \oplus 1$, when needed, we may assume that all the problematic g_i 's have $3/4$ agreement probability with some linear function.

Let P be the set of i such that g_i is problematic. For each such i we denote by l_i the linear function of two inputs that agrees with g_i on a $3/4$ fraction of the inputs. In the next section we show how if $p = |P| \geq 0.997m - 10n \geq 13.9n$, then one can "break" the generator using correlation attack. Correlation attacks are often used in practice to break pseudorandom generators. The distinguisher below is an interesting example where one can actually prove that correlation attack results in a polynomial time distinguisher.

3.1 The Distinguisher Based on Semidefinite Programming

Given a string $(r_1, \dots, r_p) \in \{0, 1\}^p$, consider the following linear system over $GF(2)$ with two variables per equation.

$$\forall i \in P \quad l_i(x) = r_i. \tag{3}$$

We will argue that the fraction of satisfied equations in the system (3) is distributed differently if r_1, \dots, r_p is uniform or if it is the output of G . Since the expected number of equations (3) satisfied when $r_i = g_i$ is at least $3p/4$, it follows by Markov's inequality that

Lemma 9. *If r_1, \dots, r_p are the output of g_1, \dots, g_p , respectively (where the g_i 's are problematic), then, for every $\varepsilon > 0$, there is a probability of at least ε that at least $3/4 - \varepsilon$ fraction of the equations in (3) are satisfiable. More formally*

$$\Pr_{z \in \{0,1\}^n} \left[\frac{|\{i \mid g_i(z) = l_i(z)\}|}{p} \geq \frac{3}{4} - \varepsilon \right] \geq \varepsilon.$$

Lemma 10. *If r_1, \dots, r_p are chosen uniformly at random from $\{0, 1\}^p$, and $p > (1/2\delta^2)(\ln 2)(n + c)$, then the probability that there is an assignment that satisfies more than a $1/2 + \delta$ fraction of the equations of (3) is at most 2^{-c} .*

Proof. Fix an assignment z ; then, by Chernoff’s inequality, the probability that a fraction at least $1/2 + \delta$ of the r_i agree with $l_i(z)$ is at most $e^{-2\delta^2 p} \leq 2^{-c-n}$. By a union bound, there is at most a probability 2^{-c} that such a z exists. \square

Given a system of linear equations over $GF(2)$ with two variables per equation, it is NP-hard to determine the largest number of equations that can be satisfied, but the problem can be approximated to within a .878 factor using semidefinite programming [13]. We now prove theorem 2.

Proof of Theorem 2: Let $\delta = .158$, $\varepsilon = 10^{-4}$. Thus, $.878(3/4 - \varepsilon) > 1/2 + \delta$. Given a string (r_1, \dots, r_m) , which is either random in $\{0, 1\}^m$ or from the distribution $G(z)$ (where z is random), we consider the system (3). Using semidefinite programming [13] we get a polynomial time algorithm that is successful if a $3/4 - \varepsilon$ fraction of the equations hold, and fails if no more than $0.878(3/4 - \varepsilon) > 1/2 + \delta$ of the equations hold. Let $c = 0.0005n$. By lemma 10 if $p > 13.89n > (1/2\delta^2)(\ln 2)(n + c)$, then the probability that more than $1/2 + \delta$ of the equations are satisfied, when r_1, \dots, r_m are chosen randomly, is at most $2^{-c} = \exp(-n)$. On the other hand, when $(r_1, \dots, r_m) = G(z)$, for some z , then the probability that at least $3/4 - \varepsilon$ fraction of the equations are satisfied is at least ε , so the theorem follows. \square

3.2 Correlation Attacks

In this section we discuss how our distinguisher for the case $k = 4$ can be seen as a “correlation attack.”

Correlation attacks are a class of attacks that are often attempted in practice against candidate pseudorandom generators. Pseudorandom generators are called “stream ciphers” in the applied cryptography literature, see e.g. the introduction of [18] for an overview.

The basic idea is as follows. Given a candidate generator $G : \{0, 1\}^n \rightarrow \{0, 1\}^m$, where $G(x) = g_1(x), \dots, g_m(x)$, we first try and find linear relations between input bits and output bits that are satisfied with non-trivial probability. For example, suppose we find coefficients $a_{i,j}$, $b_{i,j}$ and c_j such that each of the equations

$$\begin{aligned} \sum_{i=1}^n a_{i,1}x_i + \sum_{i=1}^m b_{i,1}g_i(x) &= c_1 \pmod{2} \\ \sum_{i=1}^n a_{i,2}x_i + \sum_{i=1}^m b_{i,2}g_i(x) &= c_2 \pmod{2} \\ \dots \\ \sum_{i=1}^n a_{i,t}x_i + \sum_{i=1}^m b_{i,t}g_i(x) &= c_t \pmod{2} \end{aligned} \tag{4}$$

is satisfied with probability bounded away from $1/2$.

Now we want to use this system of equations in order to build a distinguisher. The distinguisher is given a sample $\mathbf{z} = (z_1, \dots, z_m)$ and has to decide whether \mathbf{z} is uniform or is the output of G . The distinguisher substitutes z_i in place of $g_i(x)$ in (4) and then tries to find an \mathbf{x} that maximizes the number of satisfied equations. The hope is that, if $\mathbf{z} = G(\mathbf{x})$, then we will find \mathbf{x} as a solution of the optimization problem.

Unfortunately, maximizing the number of satisfied equations in a linear system over $GF(2)$ is an NP-hard problem, and, in fact, it is NP-hard to achieve an approximation factor better than $1/2$ [14]. In practice, one uses belief-propagation algorithms that often work, although the method is typically not amenable to a formal analysis.

In Section 3.1, we were able to derive a formal analysis of a related method because we ended up with a system of equations having only two variables per equation, a class of instances for which good approximation algorithms

are known. Furthermore, we did not try to argue that, when the method is applied to the output of the generator, we are likely to recover the seed; instead, we argued that just being able to approximate the largest fraction of satisfiable equations gives a way to distinguish samples of the generators from random strings.

4 $O(n^{k/2})$ upper bound

In this section we prove the following theorem which gives an upper bound on the maximal stretch of an ε -biased generator in NC_k^0 . We restate Theorem 3.

Theorem. *For every integer $0 < k \leq n$ and any $0 \leq \varepsilon < 2^{-2k-1}$, if $G = (g_1, \dots, g_m)$ is an ε -biased pseudorandom generator, where each of the g_i 's depend on at most k bits, then*

$$m \leq \sum_{t=0}^{\lceil \frac{k}{2} \rceil} \binom{n}{t} 2^{2(k-t)} \leq k 2^{2k} \binom{n}{\lceil \frac{k}{2} \rceil}. \quad (5)$$

The proof uses the following lemma from [25].

Lemma 11 ([25]). *Let $f : \{0, 1\}^k \rightarrow \{0, 1\}$ then for all r*

- *Either f is a polynomial of degree at most r over $GF(2)$, or*
- *f is biased towards an affine function of at most $k - r$ variables.*

Proof of Theorem 3: For $0 \leq t \leq n$, write $B(t) = \sum_{i=0}^t \binom{n}{i}$. Set $s = \lfloor k/2 \rfloor$, $r = k - s$. By Lemma 11 every g_i is either a degree $\leq r$ polynomial, or is biased towards an affine function of at most s variables. Let p be the number of degree $\leq r$ polynomials among the g_i 's, and b_t be the number of g_i 's biased towards an affine function of exactly t variables (but not towards an affine function with less than t variables). Clearly, $m \leq p + \sum_{t=0}^s b_t$. Note that the $B(r)$ monomials of degree $\leq r$ on the variables x_1, \dots, x_n form a basis for the vector space of all degree $\leq r$ polynomials in x_1, \dots, x_n . Therefore if $p > B(r)$, there is a linear dependency between the g_i 's. We therefore conclude that

$$p \leq B(r). \quad (6)$$

On the other hand, note that by Lemma 7, if g is biased towards an affine function of $t \leq s$ variables (but not towards an affine function with less than t variables) then there exists an affine function ℓ of t variables such that $\Pr[g = \ell] \geq 1/2 + 2^{t-k}$. Moreover, there are exactly $\binom{n}{t}$ linear functions on t variables. For $t \leq s$ let L_t be the set of linear functions on t variables. Lemma 8 implies that if

$$b_t \geq 1 + |L_t| \cdot c_{2^{t-k}} = 1 + \binom{n}{t} \cdot (2^{2(k-t)} - 1)$$

then there is a \oplus of two of the g_i 's that has at least a $\frac{3}{4} 2^{2t-2k} > 2^{-2k-1}$ bias. It therefore follows that

$$b_t \leq \binom{n}{t} (2^{2(k-t)} - 1). \quad (7)$$

Combining (7) and (6) we obtain that

$$m \leq B(r) + \sum_{t=0}^{\lceil \frac{k}{2} \rceil} \binom{n}{t} (2^{2(k-t)} - 1) \leq \sum_{t=0}^{\lceil \frac{k}{2} \rceil} \binom{n}{t} 2^{2(k-t)} \leq k 2^{2k} \binom{n}{\lceil \frac{k}{2} \rceil}$$

as needed.

□

5 Constructions for $k = 5$ and $k = 4$

5.1 Overview

In this section we prove Theorem 4. We will also give a construction of a $k = 4$ generator with inverse-polynomial bias. In both cases, we will construct a generator mapping $2n$ bits into cn bits. It is helpful to think of c as a large constant, although the results for $k = 5$ hold also if c is a function of n .

We will construct two generators: one will be good against linear tests that involve a small number of output bits (we call them *small tests*), and another is good against linear tests that involve a large number of output bits (we call them *large tests*). The final generator will be obtained by computing the two generators on independent seeds, and then XOR-ing their output bit by bit. In this way, we fool every possible test.

The generator that is good against large tests is such that every bit of the output is just the product of two bits of the seed. We argue that the sum (modulo 2) of t output bits of the generator has bias exponentially small in t/c^2 , where c , as above, is the stretch of the generator.

Then we describe a generator that completely fools linear tests of size up to about n/c^2 , and such that every bit of the output is the sum of three bits of the seed. Combined with the generator for large tests, we get a generator in NC_5^0 such that every linear test has bias $2^{-O(n/c^4)}$.

5.2 The Generator for Large Tests

Let us call the bits of the seed y_1, \dots, y_n .

Let K be an undirected graph formed by $n/(2c + 1)$ disjoint cliques each with $2c + 1$ vertices (we assume for simplicity that $n/(2c + 1)$ is an integer). K has n vertices that we identify with the elements of $[n]$. K has $cn = m$ edges. Fix some ordering of the edges of K , and let (a_j, b_j) be the j -th edge of K . Define the functions q_1, \dots, q_m as $q_j(y_1, \dots, y_n) = y_{a_j} y_{b_j}$.

Lemma 12. *For every subset $S \subseteq [m]$, the function $q_S(\mathbf{y}) = \sum_{j \in S} q_j(\mathbf{y})$ is such that*

$$|\Pr_{\mathbf{y}}[q_S(\mathbf{y}) = 0] - \frac{1}{2}| \leq \left(\frac{1}{2}\right)^{1+|S|/(2c^2+c)}.$$

The proof relies on the following two standard lemmas. The first one from [8] is a special case of the Schwartz-Zippel lemma [29, 32].

Lemma 13 ([8]). *Let p be a non-constant degree-2 multilinear polynomial over $GF(2)$. Then $1/4 \leq \Pr[p(x) = 0] \leq 3/4$.*

It is well known and easy to prove by induction that

Lemma 14. *Let X_1, \dots, X_t be independent 0/1 random variables, and suppose that for every i we have $\delta \leq \Pr[X_i = 0] \leq 1 - \delta$. Then*

$$\frac{1}{2} - \frac{1}{2}(1 - 2\delta)^t \leq \Pr \left[\bigoplus_i X_i = 0 \right] \leq \frac{1}{2} + \frac{1}{2}(1 - 2\delta)^t.$$

We can now prove lemma 12.

Proof of Lemma 12.: We can think of S as a subset of the edges of K . Each connected component of K has $2c^2 + c$ edges, so S contains edges coming from at least $|S|/(2c^2 + c)$ different connected components. Let t be the number

of connected components. If we decompose the summation $\sum_{j \in S} q_j(y_1, \dots, y_n)$ into terms depending on each of the connected components, then each term is a non-trivial degree-2 polynomial, and the t terms are independent random variables when y_1, \dots, y_n are picked at random. We can then apply lemma 14, where the X_i are the values taken by each of the t terms in the summation, $\delta = 1/4$, and $t \geq |S|/(2c^2 + c)$. \square

In particular it follows that if we define $G_1(y_1, \dots, y_n) = (q_1, \dots, q_m)$ then any linear combination of at least $\Omega(n)$ coordinates of the output of G has an exponentially small bias.

5.3 The Generator for Small Tests

Let $A \in \{0, 1\}^{n \times m}$ be a matrix such that every row is a vector in $\{0, 1\}^n$ with exactly three non-zero entries, and also assume that every set of $\sigma - 1$ rows of A is linearly independent. Let A_1, \dots, A_m be the rows of A . We define the linear functions l_1, \dots, l_m as $l_i(\mathbf{x}) = A_i \cdot \mathbf{x}$. Note that each of these linear functions depends on only three bits of the input.

Proposition 15. *For every subset $S \subseteq [m]$, $|S| < \sigma$, the function $l_S(\mathbf{x}) = \sum_{j \in S} l_j(\mathbf{x})$ is balanced.*

Proof. We have $l_S(\mathbf{x}) = (\sum_{j \in S} A_j) \cdot \mathbf{x}$, and since $\sum_{j \in S} A_j$ is a non-zero element of $\{0, 1\}^n$ (as $\{A_i\}_{i \in S}$ are linearly independent), it follows that $l_S(\cdot)$ is a non-trivial linear function, and therefore it is balanced. \square

Lemma 16. *For every $c = c(n) = o(\sqrt{n}/(\log n)^{3/4})$ and for sufficiently large n there is a 0/1 matrix A with cn rows and n columns such that every row has exactly three non-zero entries and such that every set of $\sigma - 1 = n/(4e^2c^2(n)) - 1$ rows are linearly independent.*

Proof. We shall construct the matrix A as the adjacency matrix of a bi-partite expander graph. We begin by showing a relation between an expansion of bi-partite graphs and linear independence of related linear functions.

Let $G = (L, R, E)$ be a bi-partite graph such that $|R| = n$. G has the *b - right unique neighbor* property, if for any set $V \subseteq L$, $|V| \leq b$ there exists a vertex $u \in R$ such that $|N(u) \cap V| = 1$. Assign the n input variables to the different vertices in R . For every vertex $v \in L$ the corresponding output is the linear function

$$\ell_v(X) = \sum_{i \in N(v)} x_i$$

Lemma 17. *If G has the b-right unique neighbor property then for any set B such that $|B| < b$, the linear combination $\ell = \sum_{v \in B} \ell_v$ is nonzero.*

Proof. We have that

$$\ell = \sum_{v \in B} \ell_v = \sum_{i: |N(i) \cap B| = \text{odd}} x_i$$

The right unique neighbor property guarantees that there is an input variable that belongs to exactly one output. Therefore ℓ is not zero. \square

Note that we actually need the odd-neighbor property (i.e. that for any set of size less than b there is a neighbor with odd number of neighbors in the set), but our calculations show that the graphs that we use have the stronger unique-neighbor property. The problem of constructing explicit expanders with the unique neighbor property was extensively studied in recent years and many new constructions were found [2, 7, 9, 22]. However, none of these give the parameters we need here. Thus we only prove the existence of such a graph instead of giving an explicit construction. Our proof actually show that if we pick a random graph (with the correct parameters) then w.h.p. it will have the unique-neighbor property.

The existence of graphs with the unique neighbor property will follow from the existence of certain expanders. We say that a bipartite graph (L, R, E) is (σ, α) -expanding if for every subset $S \subseteq L$ of vertices on the left, if $|S| \leq \sigma$ then $|N(S)| > \alpha \cdot |S|$, where (as before) $N(S)$, defined as

$$N(S) = \{v \in R : \exists u \in S \text{ such that } (u, v) \in E\},$$

is the neighborhood of S .

Lemma 18. *Suppose that the degrees of all vertices in L are bounded by Δ . If $|N(S)| > \Delta|S|/2$ for all sets $S \subseteq L$ of size at most σ , then G has the σ -right unique neighbor property.*

Proof. If there is no unique neighbor, then by counting edges $|N(S)| \leq \Delta|S|/2$. □

The following lemma shows the existence of a bi-partite expander graph with the required properties.

Lemma 19. *For every $c(n) = o(\sqrt{n}/(\log n)^{3/4})$ and sufficiently large n there is a $(\sigma, 3/2)$ -expanding graph $([c(n) \cdot n], [n], E)$ with $\sigma = n/(4e^4 c^2(n))$ such that every vertex on the left has degree 3.*

Proof. We construct the graph at random by connecting each vertex on the left to three distinct randomly chosen vertices on the right. (For different left vertices the random choices are independent.)

Fix a size s , $2 \leq s \leq n/(2e^2 c)$, and consider the probability that there is a subset $S \subseteq [cn]$ of s vertices on the right (i.e. $S \subset R$) whose neighborhood is contained in a set $T \subseteq [n]$ of $3s/2$ vertices on the left. Clearly, this probability is less than $(\frac{3s}{2n})^{3s}$. The number of possible choices for S is $\binom{cn}{s}$ and the number of possible choices for T is $\binom{n}{3s/2}$, and, by a union bound, the probability that the construction fails to satisfy the required property is at most

$$\sum_{s=2}^{\sigma} \binom{cn}{s} \cdot \binom{n}{3s/2} \left(\frac{3s}{2n}\right)^{3s}. \quad (8)$$

Using the inequality $\binom{n}{k} \leq \left(\frac{ne}{k}\right)^k$ we can see that Expression (8) is at most

$$\sum_{s=2}^{\sigma} \left(\frac{ecn}{s}\right)^s \cdot \left(\frac{2en}{3s}\right)^{3s/2} \cdot \left(\frac{3s}{2n}\right)^{3s} \leq \sum_{s=2}^{\sigma} \left(\frac{2e^3 c \sqrt{s}}{\sqrt{n}}\right)^s \quad (9)$$

$$= O\left(\left(\frac{c}{\sqrt{n}}\right)^2 + \left(\frac{c}{\sqrt{n}}\right)^3 + \left(\frac{c}{\sqrt{n}}\right)^4 \cdot (\log n)^3\right) = o(1), \quad (10)$$

where the last line can be verified by breaking the second sum in expression (9) up into the the term $s = 2$ which is $O((c/\sqrt{n})^2)$, $s = 3$, which is $O((c/\sqrt{n})^3)$, the terms $s = 4, \dots, 2 \log n$, each of which is at most $O(c\sqrt{\log n}/\sqrt{n})^4$, and the remaining terms, each of which is at most $1/n^2$. □

We now finish the proof of lemma 16. Consider the graph G constructed in Lemma 19 and let A be the $|L| \times |R|$ matrix such that $A_{v,u} = 1$ if and only if (v, u) is an edge of G . Note that every row of A has exactly 3 non-zero entries. By Lemma 18, G has the σ -right unique neighbor property. Therefore by Lemma 17 the linear functions corresponding to any subset of σ rows are linearly independent. The proof follows. □

In particular we get that if we define $G_2(\mathbf{x}) = (A_1 \cdot \mathbf{x}, \dots, A_m \cdot \mathbf{x})$ (where A is the matrix guaranteed by lemma 16) then any linear combination of at most $n^2/4e^2 c^2 - 1$ coordinates of the output of G_2 is unbiased.

5.4 Putting Everything Together: Proof of theorem 4

In order to obtain the generator, recall that $m = cn$ and take $G_1 : \{0, 1\}^n \rightarrow \{0, 1\}^m$, and $G_2 : \{0, 1\}^n \rightarrow \{0, 1\}^m$ be the generators defined above (with the parameter c). Then we take $G : \{0, 1\}^{2n} \rightarrow \{0, 1\}^m$ defined by $G(x, y) = G_1(x) \oplus G_2(y)$. We get that by lemma 12 any combination of more than σ outputs of G has bias at most $2^{-\sigma/(c^2+c)}$, and that by lemma 16, any combination of at most $\sigma = n/(4e^2c^2)$ of the outputs of G is unbiased. This completes the proof of the theorem.

5.5 Generator for $k = 4$

When $k = 4$ we want to replace the generator for small sets by a generator which depends only on two bits. The construction is essentially the one in [8].

Let H be an undirected graph with n vertices, that we identify with $[n]$, having cn edges and girth γ . Fix some ordering of the edges of H , and let (a_j, b_j) be the j -th edge of H . We define the linear functions l_1, \dots, l_m as $l_j(x_1, \dots, x_n) = x_{a_j} + x_{b_j}$.

Proposition 20. *For every subset $S \subseteq [m]$, $|S| < \gamma$, the function $l_S(\mathbf{x}) = \sum_{j \in S} l_j(\mathbf{x})$ is balanced.*

Proof. Since $|S| < \gamma$, the subgraph of H induced by the edges of S is a forest. Therefore $l_S(\mathbf{x})$ is a non-zero linear function, and hence balanced. \square

The explicit construction of expanders by Lubotzky-Phillips-Sarnak [21] has high girth:

Lemma 21 ([21]). *For every c and for sufficiently large n there are explicitly constructible graphs H with n vertices, cn edges, and girth $\Omega((\log n)/(\log c))$.*

We thus obtain.

Theorem 22. *For every c and sufficiently large n , there is a generator in uniform NC_4^0 mapping n bits into cn bits and sampling an ε -biased distribution, where $\varepsilon = n^{-1/O(c^2 \log c)}$.*

6 ε -biased generator for large k

In this section we construct an ε -biased generator in NC_k^0 , for large k , that outputs $n^{\Omega(\sqrt{k})}$ bits. More precisely we prove Theorem 5:

Theorem. *Let k be a positive integer. There exists an ε -biased generator in NC_k^0 from n bits to*

$$\left(\frac{n}{\sqrt{k} - 6} \right)^{\frac{\sqrt{k} - 3}{2}} = n^{\sqrt{k}(\frac{1}{2} - o_k(1))}$$

bits whose bias ε is at most

$$\exp \left(- \frac{1}{4 \times 2^{\sqrt{k}}} \right).$$

6.1 The Generator for Large Tests

In this section we prove the following Lemma.

Lemma 23. Let $n = p^2$ and let d be an integer. Then there exists a generator $G_1 : (g_1, \dots, g_m) : \{0, 1\}^n \rightarrow \{0, 1\}^m$, where $m = \binom{p}{d}$ such that for all $J \subseteq [m]$ the bias of $g = \bigoplus_{j \in J} g_j$ is at most

$$\exp\left(-\frac{|J|^{\frac{1}{d}}}{2^d}\right). \quad (11)$$

Proof. Consider the following bi-partite graph $G = (L, R, E)$ where $|L| = p$ (left vertices), $|R| = \binom{p}{d}$ (right vertices). Identify the vertices of L with the numbers $1, \dots, p$ and the vertices of R with $\binom{[p]}{d}$, the set of all subsets of $[p] \triangleq \{1, \dots, p\}$ of size d . The edges of G are all pairs (i, S) such that $i \in [p]$, $S \in \binom{[p]}{d}$ and $i \in S$. For a set of vertices, V , we denote with $N(V)$ the set of neighbors of V :

$$N(V) = \{u \in L \cup R : \exists v \in V \text{ such that } (u, v) \in E\}.$$

For a vertex i let $\deg(i) = |N(\{i\})|$.

Proposition 24. For any set of right vertices $V \subseteq R$ we have that $|N(V)| \geq \frac{d|V|^{\frac{1}{d}}}{e}$.

Proof. Note that for any set of t left vertices, L' , there are (exactly) $\binom{t}{d}$ right vertices, R' , such that $N(R') = L'$. The result follows from the inequality

$$|V| \leq \binom{|N(V)|}{d} \leq \left(\frac{e|N(V)|}{d}\right)^d.$$

□

Our construction will assign a monomial of degree d , in the input variables, to each edge. We think about the vertices of L as representing disjoint subsets of the input variables (each of size p) and each edge leaving such input set as corresponding to a monomial in its variables. The right vertices, R , correspond to the output bits. Each output is the sum of the monomials that label the edges that fan into it. We now give the formal construction.

Let $X = \bigsqcup_{i=1}^p X_i$ be a partition of $X = \{x_1, \dots, x_n\}$ into p disjoint sets each of size p . We assign the set X_i to the i -th vertex of L . Let M_i be the set of all multilinear monomials of degree d in the variables of X_i . We have that

$$|M_i| = \binom{p}{d} > \binom{p-1}{d-1} = \deg(i)$$

Therefore we can assign to each edge leaving i a different monomial from M_i . Denote by M_e the monomial corresponding to the edge e . Each right vertex corresponds to an output bit. For a right vertex j the j 'th output, which we denote by g_j , is the sum of all monomials that were assigned to the edges adjacent to j :

$$g_j = \sum_{e:j \in e} M_e.$$

Thus each output is the sum of d monomials each of degree d . Hence each output depends on d^2 input variables. We now show that any large linear combination of the output bits has a small bias by proving (11). Let $g = \bigoplus_{j \in J} g_j$. The proof is essentially the same as the proof of lemma 12 and follows from the following easy propositions.

Proposition 25. Let $g = \bigoplus_{j \in J} g_j$, then g can be written as the sum of at least $N(J)$ polynomials of degree d , each in a different set of variables.

Proof. The set of outputs J , has $N(J)$ left neighbors. The edges connecting the set J to a neighbor $i \in N(J)$ are labeled with polynomials of degree d in X_i . □

From the Schwartz-Zippel lemma [29, 32] we get

Proposition 26. *For any polynomial g of degree d we have*

$$\frac{1}{2^d} \leq \Pr[g = 0] \leq 1 - \frac{1}{2^d}.$$

Thus according to lemma 14 we get that the bias of g is at most

$$\frac{1}{2} \left(1 - \frac{2}{2^d}\right)^{N(J)} \leq \frac{1}{2} \cdot \exp\left(\frac{-2N(J)}{2^d}\right) \leq \exp\left(\frac{-|J|^{\frac{1}{d}}}{2^d}\right)$$

This finishes the proof of Lemma 23. □

6.2 The Generator for Small Tests

Similar to the $k = 4, 5$ cases this generator will output only linear functions. We will have the property that any small set of these linear functions is linearly independent. This is a standard construction that follows from unique neighbor property of expanding graphs.

Lemma 27. *Let t be positive integer t and $\Delta = 10t$. There exists a mapping from n bits to n^t bits such that every output depends linearly on Δ input variables, and such that any linear combination of at most \sqrt{n} outputs is non-zero and therefore unbiased.*

Proof. As in the proof of lemma 16, we shall construct a linear mapping from an expander bi-partite graph with the unique neighbor property.

Lemma 28. *Let t be a positive integer and $\Delta = 10t$. Then there exists a family of bi-partite graphs $G_n = (L, R, E)$ with $|L| = n^t$, $|R| = n$, $\forall v \in L \deg(v) = \Delta$, such that G_n is a $(\sigma = \lceil \sqrt{n} \rceil, 5t)$ expanding graph.*

Proof. Let $|R| = n$, $|L| = n^t$. Connect every vertex in L to a randomly chosen multi set of size Δ of distinct right vertices. We continue as in Lemma 19. Fix a size s , $2 \leq s \leq \sigma = \lceil \sqrt{n} \rceil$, and consider the probability that there is a subset $S \subseteq [n^t]$ of s vertices on the right whose neighborhood is contained in a set $T \subseteq [n]$ of $\Delta s/2$ vertices on the left. This probability is less than $\left(\frac{\Delta s}{2n}\right)^{\Delta s}$. The number of possible choices for S is $\binom{n^t}{s}$ and the number of possible choices for T is $\binom{n}{\Delta s/2}$. Therefore applying the union bound and recalling that $\Delta = 10t$ the probability that the construction fails to satisfy the required property is at most

$$\sum_{s=2}^{\sigma} \left(\frac{en^t}{s}\right)^s \cdot \left(\frac{2en}{\Delta s}\right)^{\Delta s/2} \cdot \left(\frac{\Delta s}{2n}\right)^{\Delta s} \leq \sum_{s=2}^{\sigma} \left(\frac{(e\Delta s)^\Delta}{n^{4t}}\right)^s = o(1).$$

□

We now finish the proof of lemma 27. Let G be the graph constructed in Lemma 28. Label each vertex on the right by one of the variables x_i and each vertex on the left by the linear combination of the variables adjacent to it. By Lemma 18, G has the σ -right unique neighbor property. Therefore by Lemma 17 every set consisting of $\sigma - 1$ linear functions (corresponding to left vertices) is linearly independent. The proof follows. □

6.3 Putting things together: Proof of theorem 5

Let $\kappa = (\lfloor \sqrt{k} \rfloor - 5)^2$, $\nu = \lfloor \sqrt{\frac{n}{2}} \rfloor^2$. We have that

$$k > \kappa + 10\sqrt{\kappa}, \quad \kappa > k - 12\sqrt{k}, \quad \frac{n}{2} \geq \nu > \frac{n}{2} - \sqrt{2n}.$$

Let $X = \{x_1, \dots, x_\nu\}$, $Y = \{y_1, \dots, y_\nu\}$. Let $f_1(X), \dots, f_{\binom{p}{d}}(X)$ be the outputs of the generator against large tests with the parameters $p = \sqrt{\nu}$, $d = \sqrt{\kappa}$. Let $h_1(Y), \dots, h_{\nu^\kappa}(Y)$ be the outputs of the generator for small tests on Y , given the parameter $t = \sqrt{\kappa}$. Note that

$$\nu^\kappa > \binom{\sqrt{\nu}}{\sqrt{\kappa}} = \binom{p}{d}.$$

Our generator G will output the functions

$$\forall 1 \leq i \leq \binom{p}{d} \quad g_i(X, Y) = f_i(X) + h_i(Y).$$

Notice that as we have more h_i 's than f_i 's we do not use most of the h_i 's. Clearly, each output of the generator depends on $\kappa + 10\sqrt{\kappa} < k$ input variables. From lemmas 23,27 we get that the bias of any non trivial linear combination of the outputs is at most

$$\exp\left(-\frac{\nu^{\frac{1}{d}}}{2^d}\right) \leq \exp\left(-\frac{n^{\frac{1}{2\sqrt{k}}}}{4 \times 2^{\sqrt{k}}}\right).$$

Our generator takes $2\nu \leq n$ inputs and outputs

$$\binom{p}{d} \geq \left(\frac{e^2\nu}{\kappa}\right)^{\frac{\sqrt{\kappa}}{2}} \geq \left(\frac{n}{\sqrt{k}-6}\right)^{\frac{\sqrt{k}}{2}-3} = n^{\sqrt{k}(\frac{1}{2}-o_k(1))}$$

as needed.

7 A degree 2 generator

In this section we consider a variant of the problem presented in the paper. Suppose that we require that every output bit is a degree k polynomial in the input bits. It is clear that if we want the output to be ε -biased, then the number of output bits m is at most the dimension of the space of degree k polynomials in n variables, which is $\sum_{i=0}^k \binom{n}{i} = O(n^k)$ (as otherwise there will be a linear dependence among the output bits). Clearly this is a relaxation of the problem described above. In particular any upper bound here will imply an upper bound for NC_k^0 . The problem is also of independent interest, as low degree generators are ‘‘simple’’ in an intuitive sense.

We now show how to construct a generator of ε -biased set such that every output is a polynomial of degree 2 in the input variables. We show that unlike the NC_0^0 case we can output $\Omega(n^2)$ bits. In particular we prove Theorem 6:

Theorem. $\forall 1 \leq m \leq n$ there exists an ε -biased generator $G = (g_1, \dots, g_t) : \{0, 1\}^n \mapsto \{0, 1\}^t$, $t = \lfloor \frac{n}{2} \rfloor \cdot m$, such that g_i is a degree 2 polynomial, and the bias of any non trivial linear combination of the g_i 's is at most $2^{-\frac{n-2m}{4}}$.

We begin by studying the bias of a degree 2 polynomial, over $GF(2)$. In this section we will only consider degree 2 polynomials P such that $P(0) = 0$. Below we denote with x^T and A^T the transpose of the vector x and the matrix A , respectively.

7.1 The Bias of Degree 2 polynomials

Let $P(x_1, \dots, x_n)$ be a degree 2 polynomial. P is also called a quadratic form over $GF(2)$. We say that a matrix A represents P with respect to a basis of $GF(2)^n$, $\{v_i\}_{i=1}^n$, if for every vector $v = \sum_{i=1}^n x_i \cdot v_i$ we have that $P(v) = x^T A x$. Notice that we can always find an upper triangular matrix that represents P ; let

$$P(a_1, \dots, a_n) = \sum_{1 \leq i \leq j \leq n} \alpha_{i,j} a_i a_j$$

Define

$$A(P)_{i,j} = \begin{cases} \alpha_{i,j} & i \leq j \\ 0 & i > j \end{cases}$$

Clearly $P(\sum_{i=1}^n e_i \cdot x_i) = x^T A(P)x$ and $A(P)$ represents P with respect to the standard basis.

The bias of a quadratic form is bounded by the rank of the matrix representing it as follows.

Theorem 29. *The bias of a degree 2 polynomial P is at most*

$$2^{-\left(1 + \frac{\text{rank}(A+A^T)}{4}\right)}$$

for any matrix A that represents P .

Theorem 29 shows that in order to output m polynomials of degree 2, such that any non trivial linear combination of them is almost unbiased, it suffices to find matrices A_1, \dots, A_m such that for any non trivial combination of them, $B = \sum_{i=1}^m \alpha_i A_i$ ($\alpha_i \in GF(2)$), we have that $\text{rank}(B + B^T)$ is high.

7.1.1 Proof of theorem 29

The following claim is trivial.

Proposition 30. *$P \equiv 0$ iff there exists a symmetric matrix that represents P w.r.t. some basis iff any matrix that represents P is symmetric.*

The proof of theorem 29 will follow from the following lemmas.

Lemma 31. *For any quadratic form P on n variables, there exists a basis of $GF(2)^n$ e_i, f_i $i = 1, \dots, r$ and g_j $j = 1, \dots, s$ such that $2r + s = n$ and n elements in $GF(2)$, a_i, b_i $i = 1, \dots, r$, c_j $j = 1, \dots, s$, such that for*

$$v = \sum_{i=1}^r x_i e_i + \sum_{i=1}^r x_{r+i} f_i + \sum_{j=1}^s x_{2r+j} g_j$$

we have

$$P(v) = \sum_{i=1}^r (a_i x_i^2 + x_i x_{r+i} + b_i x_{r+i}^2) + \sum_{j=1}^s c_j x_{2r+j}^2 = \sum_{i=1}^r (a_i x_i + x_i x_{r+i} + b_i x_{r+i}) + \sum_{j=1}^s c_j x_{2r+j}. \quad (12)$$

Such a basis is called “a canonical basis for P ”.

Proof. See the proof of theorem 5.1.7 in [16]. □

Lemma 32. Let P be a quadratic form on n variables. Let A represent P with respect to the standard basis (in particular, A is upper triangular) and D represent P with respect to the canonical basis. Then

$$\text{rank}(D) \geq \frac{\text{rank}(A + A^T)}{2}$$

Proof. Let B be the matrix whose columns are $e_1, \dots, e_r, f_1, \dots, f_r, g_1, \dots, g_s$ written w.r.t. the standard basis. We have that

$$\forall x \in GF(2)^n \quad x^T D x = x^T B^T A B x.$$

In other words

$$\forall x \in GF(2)^n \quad x^T (D - B^T A B) x = 0.$$

Therefore there exists a symmetric matrix S such that

$$D - B^T A B = S,$$

or

$$D = B^T (A + (B^{-1})^T S (B^{-1})) B.$$

As $(B^{-1})^T S (B^{-1})$ is a symmetric matrix we get by the next lemma (lemma 33) that

$$\text{rank}(D) = \text{rank}(A + (B^{-1})^T S (B^{-1})) \geq \frac{\text{rank}(A + A^T)}{2}.$$

□

Lemma 33. For an upper triangular matrix A and any symmetric matrix S we have that

$$\text{rank}(A + S) \geq \frac{\text{rank}(A + A^T)}{2}.$$

Proof. Let $r = \text{rank}(A + S) = \text{rank}((A + S)^T) = \text{rank}(A^T + S)$. Then

$$\text{rank}(A + A^T) = \text{rank}(A + S + S + A^T) \leq \text{rank}(A + S) + \text{rank}(A^T + S) = 2r.$$

□

PROOF OF THEOREM 29. Clearly the bias of P does not change if we calculate it w.r.t. to a canonical basis, $\{v_i\}_{i=1}^n$. Let $v = \sum_{i=1}^n x_i \cdot v_i$, we have that

$$P(v) = \sum_{i=1}^r (a_i x_i + x_i x_{r+i} + b_i x_{r+i}) + \sum_{j=1}^s c_j x_{2r+j}.$$

Note that if for some $1 \leq j \leq s$ $c_j \neq 0$ then P is unbiased. Otherwise, we get by proposition 26 that for every i the bias of $(a_i x_i^2 + x_i x_{r+i} + b_i x_{r+i}^2)$ is at most $\frac{1}{4}$. Therefore according to Lemma 14 we get that the bias of P is at most $(\frac{1}{2})^{r+1}$. As we assumed that $\forall j \quad c_j = 0$ we see that

$$r \geq \frac{\text{rank}(D)}{2}.$$

The theorem now follows from lemma 32. □

7.2 The generator

In this subsection we give a construction of a linear space of matrices with the property that for every non zero matrix in the space, A , we have that $\text{rank}(A + A^T)$ is high. Such a construction was first given by Roth [28], and later simplified by Meshulam [24] (see also [30]). For completeness we give the construction here.

Theorem 34. *For any positive natural numbers $n \geq m$ there exist $t = \lfloor \frac{n}{2} \rfloor \cdot m$ matrices $A_1, \dots, A_t \in M_n(GF(2))$ such that for every non trivial combination $B = \sum_{i=1}^t \alpha_i A_i$ we have that*

$$\text{rank}(B + B^T) \geq n - 2m$$

Proof. Denote with $\mathbb{F} = GF(2^n)$ the field with 2^n elements. \mathbb{F} is a linear space over $GF(2)$ of dimension n . We will abuse notation and think about each $y \in \mathbb{F}$ both as a field element and as a vector in $GF(2)^n$. Fix a basis for \mathbb{F} over $GF(2)$ of the form $1, x, x^2, \dots, x^{n-1}$ for some $x \in \mathbb{F}$. Each element, $y \in \mathbb{F}$, can be viewed as a linear transformation of \mathbb{F} over $GF(2)$ in the following manner:

$$\forall z \in \mathbb{F} \quad y(z) = y \cdot z.$$

Thus for every $y \in \mathbb{F}$ there is a corresponding matrix $A_y \in M_n(GF(2))$, that represents y over the basis we chose. We denote $A = A_x$ (the same x as in the basis).

Let $\varphi : \mathbb{F} \mapsto \mathbb{F}$ be the Frobenius transformation, that is $\varphi(y) = y^2$. Let $\varphi^{(k)} = \varphi \circ \varphi \dots \circ \varphi$, k times. That is $\varphi^{(k)}(y) = y^{2^k}$. It is easy to see that φ is a linear transformation of \mathbb{F} over $GF(2)$. We denote with B the matrix that represents φ over our basis. That is, by abusing notations,

$$\forall y \in \mathbb{F} \quad By = y^2$$

Let $V \subseteq M_n(GF(2))$ be the linear space spanned by the matrices

$$V = \text{span}\{ A^i \cdot B^j \mid i = 0, \dots, n-1, j = 0, \dots, m-1 \}$$

Lemma 35. *V is a linear space of matrices of dimension nm such that for any $0 \neq E \in V$ we have that*

$$\text{rank}(E) > n - m$$

Proof. Let $0 \neq E \in V$. We want to calculate $\dim(\ker(E))$. For any $y \in \mathbb{F}$ we think about Ey also as an element of $GF(2)^n$. It is clear that

$$\begin{aligned} Ey &= \left(\sum_{i=0}^{n-1} \sum_{j=0}^{m-1} \alpha_{i,j} A^i \cdot B^j \right) y \\ &= \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} \alpha_{i,j} A^i (y^{2^j}) = \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} \alpha_{i,j} x^i y^{2^j}. \end{aligned}$$

That is, Ey is a polynomial of degree 2^{m-1} in y . Therefore it has at most 2^{m-1} roots. As E is a linear transformation, we get that its roots are a linear space. Since there are at most 2^{m-1} roots, the dimension of $\ker(E)$ is at most $m-1$. Hence $\text{rank}(E) \geq n - m + 1$. \square

We now finish the proof of the theorem. Let V be the space guaranteed by lemma 35 in $M_{\lfloor \frac{n}{2} \rfloor}(GF(2))$ of dimension $t = \lfloor \frac{n}{2} \rfloor \cdot m$. Let E_1, \dots, E_t be a basis for V . Let A_i be a $n \times n$ matrix of the following form

$$A_i = \begin{pmatrix} 0 & E_i \\ 0 & 0 \end{pmatrix}$$

Where the 0 stands for the all zero matrix in $M_{\lfloor \frac{n}{2} \rfloor}(GF(2))$. For any non trivial combination $B = \sum_{i=1}^t \alpha_i A_i$ we get

$$B = \sum_{i=1}^t \alpha_i A_i = \begin{pmatrix} 0 & \sum_{i=1}^t \alpha_i E_i \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & E \\ 0 & 0 \end{pmatrix}$$

where $E = \sum_{i=1}^t \alpha_i E_i$. Since $\{E_i\}$ is a basis and not all the α_i 's are zero then $0 \neq E \in V$. Therefore $\text{rank}(E) \geq \lfloor \frac{n}{2} \rfloor - m + 1$. We get that

$$\begin{aligned} & \text{rank}(B + B^t) \\ &= \text{rank} \begin{pmatrix} 0 & E \\ E & 0 \end{pmatrix} \\ &= 2 \cdot \text{rank}(E) \\ &\geq 2 \left(\lfloor \frac{n}{2} \rfloor - m + 1 \right) \\ &\geq n - 2m \end{aligned}$$

□

Proof of Theorem 6: Let A_1, \dots, A_t be the matrices guaranteed by theorem 34. Define $g_i(x) = x^T A_i x$. Consider any non trivial linear combination

$$g(x) = \sum_{i=1}^t \alpha_i g_i(x) = x^T \left(\sum_{i=1}^t \alpha_i A_i \right) x.$$

According to theorem 34, we have that $\text{rank}(g) \geq n - 2m$. Theorem 29 shows that the bias of g is at most $2^{\frac{n-2m}{4}}$. □

Acknowledgements

We wish to thank David Wagner for suggesting the relevance of correlation attacks. A.S. would also like to thank Avi Wigderson for helpful discussions. The authors would like to thank the anonymous referees for numerous valuable comments that greatly improved the presentation of the paper.

References

- [1] B. Applebaum, Y. Ishai, and E. Kushilevitz. Cryptography in NC^0 . In *Proceedings of the 45th IEEE Symposium on Foundations of Computer Science*, pages 166-175, 2004.
- [2] N. Alon, M. Capalbo. Explicit Unique-Neighbor Expanders. In *Proceedings of the 43rd IEEE Symposium on Foundations of Computer Science*, pages 73-79, 2000.
- [3] N. Alon, O. Goldreich, J. Håstad, and R. Peralta. Simple constructions of almost k -wise independent random variables. *Random Structures and Algorithms*, 3(3):289–304, 1992.
- [4] P. Beame, R. Karp, T. Pitassi, and M. Saks. On the complexity of unsatisfiability proofs for random k -cnf formulas. In *Proceedings of the 30th ACM Symposium on Theory of Computing*, pages 561-571, 1998.
- [5] A. Bogdanov, K. Obata, and L. Trevisan. A lower bound for testing 3-colorability in bounded degree graphs. In *Proceedings of the 43rd IEEE Symposium on Foundations of Computer Science*, pages 93–102, 2002.
- [6] E. Ben-Sasson and A. Wigderson. Short proofs are narrow: Resolution made simple. *Journal of the ACM*, 48(2), 2001.
- [7] M. Capalbo. Explicit Constant-Degree Unique-Neighbor Expanders, 2001.
- [8] M. Cryan and P. B. Miltersen. On pseudorandom generators in NC^0 . In *Proceedings of 26th Mathematical Foundations of Computer Science*, pages 272-284, 2001.
- [9] M. Capalbo, O. Reingold, S. Vadhan and A. Wigderson. Randomness Conductors and Constant-Degree Expansion Beyond the Degree/2 Barrier. In *Proceedings of the 34th ACM Symposium on the Theory of Computing*, 659-668, 2000.
- [10] Oded Goldreich. Three XOR-Lemmas - An Exposition. *Electronic Colloquium on Computational Complexity (ECCC)* 2(56): (1995)
- [11] O. Goldreich, S. Goldwasser, and S. Micali. How to construct random functions. *Journal of the ACM*, 33(4):792–807, 1986.
- [12] O. Goldreich. Candidate one-way functions based on expander graphs. *Electronic Colloquium on Computational Complexity (ECCC)* TR00-090, 2000.
- [13] M.X. Goemans and D.P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, 42(6):1115–1145, 1995.
- [14] J. Håstad. Some optimal inapproximability results. In *Proceedings of the 29th ACM Symposium on Theory of Computing*, pages 1-10, 1997.
- [15] J. Håstad, R. Impagliazzo, L. Levin, and M. Luby. A pseudorandom generator from any one-way function. *SIAM Journal on Computing*, 28(4):1364–1396, 1999.
- [16] J. W. P. Hirschfeld, *Projective Geometries over Finite Fields*, Oxford University Press, 1979.
- [17] R. Impagliazzo and M. Naor. Efficient cryptographic schemes provably as secure as subset sum. *Journal of Cryptology*, 9(4):199–216, 1996.

- [18] T. Johansson and F. Jönsson. Improved fast correlation attacks on stream ciphers via convolutional codes. In *Proceedings of International Conference on the Theory and Application of Cryptographic Techniques (EUROCRYPT)*, pages 347-362, 1999.
- [19] M. Kharitonov. Cryptographic hardness of distribution-specific learning. In *Proceedings of 25th ACM Symposium on Theory of Computing*, pages 372-381, 1993.
- [20] N. Linial, Y. Mansour, and N. Nisan. Constant depth circuits, fourier transform and learnability. *Journal of the ACM*, 40(3):607–620, 1993.
- [21] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8:261–277, 1988.
- [22] C. J. Lu and O. Reingold and S. Vadhan and A. Wigderson. Extractors: Optimal Up to Constant Factors. In *proceedings of the 35th ACM Symposium on the Theory of Computing*, pages 602-611, 2003.
- [23] M. Luby and C. Rackoff. How to construct pseudorandom permutations from pseudorandom functions. *SIAM Journal on Computing*, 2(17):373–386, 1988.
- [24] R. Meshulam. Spaces of Hankel matrices over finite fields, *Linear Algebra and its Applications* 218:73-76, 1995.
- [25] E. Mossel, R. O’Donnell and R. Servedio. Learning Juntas. In *proceedings of the 35th ACM Symposium on the Theory of Computing*, pages 206–212, 2003.
- [26] E. Mossel, A. Shpilka and L. Trevisan. On ϵ -biased generators in NC^0 . In *Proceeding of the 44th IEEE Symposium on Foundations of Computer Science*, pages 136–145, 2003.
- [27] J. Naor and M. Naor. Small-bias probability spaces: efficient constructions and applications, *SIAM Journal on Computing*, 22(4):838–856, 1993.
- [28] R. Roth. Maximum rank array codes and their application to crisscross error correction, *IEEE Transactions on Information Theory* 37:328–336, 1991.
- [29] J. T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *Journal of the ACM*, 27(4):701–717, 1980.
- [30] A. Shpilka. On the rigidity of matrices. Manuscript, 2002.
- [31] U. Vazirani. *Randomness, Adversaries and Computation*. PhD thesis, University of California, Berkeley, 1986.
- [32] R. Zippel. Probabilistic algorithms for sparse polynomials. In *Symbolic and algebraic computation (EUROSAM)*, pages 216–226, 1979.