Electronic Colloquium on Computational Complexity, Report No. 88 (2009)

Explicit lower bound for fooling polynomials by the sum of small-bias generators

Shachar Lovett* Yoav Tzur†

September 29, 2009

Abstract

Recentry, Viola (CCC'08) showed that the sum of d small-biased distributions fools degree-d polynomial tests; that is, every polynomial expression of degree at most d in the bits of the sum has distribution very close to that induced by this expression evaluated on uniformly selected random bits. We show that this is tight by showing an explicit construction of a small-bias generator (with exponentially small bias), and an explicit degree $d+1$ polynomial, that is distributed almost uniformly on random input, but always takes the value zero when evaluated on the sum of d independent copies of this generator.

1 Introduction

Small-biased distributions, as defined in [NN], are designed to fool all linear tests in the sense that the outputs of all (nontrivial) linear functions of the selected bits are distributed almost uniformly. A natural generalization refers to distributions that fool higher degree polynomials. A result of [Vio], improving over [BV] and [Lov], yields a methodology for obtaining such distributions, using any small-biased distribution: to fool polynomials of degree at most d, take the bitwise sum of d independent samples of any small-biased distribution.

This result has been shown in [BV] to be essentially tight with respect to the number of copies needed: using a counting argument, they show that for fixed bias, any generator with output length ℓ that fools all degree $d+1$ polynomials must have seed length $(d+1)\cdot\log\ell - O(1)$. Thus, for every generator with shorter seed, there *exists* a polynomial expression of degree at most $d+1$ that distinguishes a random output of the generator from truly random bits. For a suitable choice of $\varepsilon = o(1)$, the length of d separate seeds for a standard

*Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel. Email: shachar.lovett@weizmann.ac.il. This research was partially supported by the Israel Science Foundation (grant No. 1300/05)

†Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel. Email: yoav.tzur@weizmann.ac.il. This research was partially supported by the Israel Science Foundation (grant No. 1041/08).

ISSN 1433-8092

construction of an ε-bias generator is still small enough, giving that in general the sum of d small-bias generators does not necessarily fool polynomials of degree $d+1$.

In this work we present an *explicit* polynomial expression of degree $d+1$ in the output bits of the sum of d copies of a specific small-bias generator. Furthermore, this generator can have even an exponentially small bias, whereas the proof of [BV] requires $\varepsilon \geq 1/\mathrm{poly}(\ell)$.

2 Preliminaries

2.1 Definitions

We begin with the definition of small bias, introduced by [NN]:

Definition 1 (Small-biased distribution)**.** *For $\ell \in \mathbb{N}$, $\varepsilon > 0$, a distribution D over $\{0,1\}^\ell$ is called ε-*biased *if for every nonzero $\alpha \in \{0,1\}^\ell$:*

$$\left| \Pr_{x \sim D}[\langle \alpha, x \rangle = 0] - \frac{1}{2} \right| \leq \varepsilon,$$

where $\langle \alpha, x \rangle$ denotes the inner product $\sum_i \alpha_i x_i$ (over $GF(2)$).

The generalization to higher degree polynomials was initiated by [LVW] (in fact, they considered the larger class of depth-2 boolean circuits), and further studied in [Bog] (although there, only super-constant sized fields were considered):

Definition 2 (Fooling polynomials)**.** *For $d, \ell \in \mathbb{N}$, $\varepsilon > 0$, a distribution X over $\{0,1\}^\ell$ is said to ε-*fool degree d polynomials *if for every ℓ-variate polynomial p over $\{0,1\}$ of degree at most d,*

$$\left| \Pr_{x \sim X}[p(x) = 0] - \Pr_{x \sim U}[p(x) = 0] \right| \leq \varepsilon,$$

where U denotes the uniform distribution over $\{0,1\}^\ell$.

We prefer to view distributions as the outputs of pseudorandom generators:

Definition 3 (Small-bias generator)**.** *For $k, \ell \in \mathbb{N}$, $\varepsilon > 0$, a mapping $G : \{0,1\}^k \to \{0,1\}^\ell$ is called an ε-*bias generator *of* stretch $\ell(k)$*, if the distribution induced by $G(s)$ for s selected uniformly in $\{0,1\}^k$ is ε-biased.*

Definition 4 (Pseudorandom generator for polynomials)**.** *For $d, k, \ell \in \mathbb{N}$, $\varepsilon > 0$, a mapping $G : \{0,1\}^k \to \{0,1\}^\ell$ is said to ε-*fool degree d polynomials *if the distribution induced by $G(s)$ for s selected uniformly in $\{0,1\}^k$, ε-fools degree d polynomials. Again, we say that G has* stretch $\ell(k)$*.*

Remark. When discussing pseudorandom generators, it is common to also consider the complexity of the generator itself (as opposed to the complexity of potential distinguishers). This is quite secondary to the current work.

Despite our final interest in distributions over bits, we will also use generators over the larger field $GF(2^n)$ (see Subsections 2.2 and 3.3 for details). When working with distributions over sets larger than $\{0,1\}$, we first need:

Definition 5 (Statistical distance)**.** *For $\varepsilon > 0$, two distributions X, Y are said to be ε-*close (in statistical distance) *if for every event E,*

$$\left|\Pr_X[E] - \Pr_Y[E]\right| \leq \varepsilon.$$

*Conversely, if there exists an event such that $|\Pr_X[E] - \Pr_Y[E]| \geq \varepsilon$, then X and Y are said to be ε-*far (in statistical distance)*.*

We can now define generators over $GF(2^n)$:

Definition 6 ($GF(2^n)$-linear tests resilience)**.** *For $k, \ell \in \mathbb{N}$, $\varepsilon > 0$, a mapping $G : GF(2^n)^k \to GF(2^n)^\ell$ is said to ε-*fool $GF(2^n)$-linear tests*, if for every vector $\alpha \in GF(2^n)^\ell$, and for s chosen uniformly in $GF(2^n)^k$, the distribution of the expression $\sum_{i=1}^{\ell} \alpha_i \cdot G_i(s)$, computed in the arithmetic of $GF(2^n)$, is ε-close to the uniform distribution over $GF(2^n)$.*

Definition 7 ($GF(2^n)$-polynomial tests resilience)**.** *For $d, k, \ell \in \mathbb{N}$, $\varepsilon > 0$, a mapping $G : GF(2^n)^k \to GF(2^n)^\ell$ is said to ε-*fool $GF(2^n)$-polynomials of degree *d if for every polynomial $p \in GF(2^n)[x_1, \ldots, x_\ell]$, the following two distributions are ε-close (in statistical distance):*

- *$p(G(s))$ for s chosen uniformly from $GF(2^n)^k$.*
- *$p(x)$ for x chosen uniformly from $GF(2^n)^\ell$.*

2.2 Representation of $GF(2^n)$

We identify n-bit vectors in $\{0,1\}^n$ with elements of this field in a standard representation scheme. We will explicitly specify, for each variable, whether it is seen as an element of the field $GF(2^n)$ or of the vector-space $\{0,1\}^n = GF(2)^n$.

We will use the linearity properties of this representation scheme:

Fact 8 (linearity of the standard representation)**.** *The following two properties hold for the standard representation scheme of $GF(2^n)$ as vectors in $\{0,1\}^n$:*

- *Addition in the field $GF(2^n)$ corresponds to addition of the respective representations in the vector space $\{0,1\}^n$;*
- *Multiplication of elements in the field $GF(2^n)$ corresponds to a bilinear mapping of the respective representations in the vector space $\{0,1\}^n$. That is, for two vectors $x, y \in \{0,1\}^n$, every bit of the vector representing the multiplication of the two $GF(2^n)$-elements represented by x and y can be written as a $GF(2)$-bilinear form in x and y.*

Fact 8 follows from the representation of $GF(2^n)$ as the quotient $GF(2)[x]/\langle c(x)\rangle$ for $\langle c(x)\rangle$ being the ideal (of the polynomials ring $GF(2)[x]$) generated by some irreducible polynomial $c(x)$ of degree n. For details, see any standard algebra textbook (e.g. [BM]), or Lemma 15 in [Tzu].

3 The construction

We will give an explicit small-bias generator and show that the sum of d independent copies of this generator does *not* fool an explicit polynomial of degree $d+1$.

While we are interested in distributions and polynomial tests over bits, our distribution and polynomial will be defined over $GF(2^n)$. Using the linearity of the representation, we will then obtain a distribution and a polynomial over bits (see Subsection 3.3).

3.1 The generator

We use the following generator, considered in [Tzu], which is related to a well known construction from [AGHP]:

Construction 9 (The geometric generator)**.** *For $n, \ell \in \mathbb{N}$, define a mapping $F : GF(2^n) \times GF(2^n) \to GF(2^n)^{\ell+1}$ by letting the i-th output element for input elements a, b be $f_i(a,b) = a \cdot b^i$, for $i = 0, ..., \ell$ (using the arithmetic of $GF(2^n)$).*

The following is Proposition 7 in [Tzu]:[1]

Proposition 10. *The geometric generator $\frac{\ell}{2^n}$-fools $GF(2^n)$-linear tests.*

For our purposes, any $\ell \geq 2d+1$ would suffice. To get a final bias of ε over bits (see subsection 3.3), we choose $n = \log \ell + \log \frac{1}{\varepsilon}$ (and note that ε can be $2^{-\Omega(\ell)}$ for $n = \Omega(\ell)$).

The element-wise sum of d instances of F gives the generator $G : GF(2^n)^{2d} \to GF(2^n)^{\ell+1}$ defined as $g_i(a_1, b_1, \ldots, a_d, b_d) = \sum_{j=1}^{d} a_j {b_j}^i$, using the arithmetic of $GF(2^n)$.

3.2 The distinguishing polynomial

We now present a polynomial D over $GF(2^n)$ of the first $2d+1$ output elements of G, denoted $g_0, ..., g_{2d}$, that has degree $d+1$, and show that while the output of this polynomial is close to uniform on uniform input, it always takes the value zero when applied to an output of G.

[1] For self containment, we give a quick outline of the proof: observe that every fixed $GF(2^n)$-linear combination $\bar{\alpha}$ in the output of the geometric generator $F(a,b)$ looks like $a \cdot q(b)$ where q is a polynomial (determined by $\bar{\alpha}$) of degree at most ℓ over $GF(2^n)$. If b is not one of the (at most) ℓ roots of q, then $a \cdot q(b)$ is distributed uniformly when a is selected uniformly. Thus the expression is distributed $\frac{\ell}{2^n}$-close to uniform over $GF(2^n)$.

The polynomial $D(g_0, \ldots, g_{2d})$ will be defined as the determinant of the following $(d+1) \times (d+1)$ Hankel matrix:

$$A_g^{(d)} = \begin{pmatrix} g_0 & g_1 & \cdots & g_d \\ g_1 & g_2 & \cdots & g_{d+1} \\ \vdots & \vdots & & \vdots \\ g_d & g_{d+1} & \cdots & g_{2d} \end{pmatrix}$$

(that is, the (i,k)-th entry of $A_g^{(d)}$ is g_{i+k}).

Indeed, this is a polynomial over $GF(2^n)$ of degree $d+1$ in the output blocks of G. We first claim that it is close to uniform when applied to uniform input:

Lemma 11. *Let M be a random $m \times m$ Hankel matrix over a finite field $\mathbb{F}$ (i.e., the Hankel matrix defined by $M_{i,k} = y_{i+k}$ for $y_0, \ldots, y_{2m-2}$ chosen uniformly at random from $\mathbb{F}$). Then, the distribution of the determinant of M is $\frac{m-1}{|\mathbb{F}|}$-close to uniform (in statistical distance).*

Proof. We proceed by induction on m. For $m = 1$, $\det(M)$ is exactly the only element of M, chosen uniformly from $\mathbb{F}$. Now fix $m > 1$, and let x be the first (top-left) element of M, and $\bar{y} = (y_1, \ldots, y_{2m-2})$ denote the rest of the elements (on the top row and rightmost column). Denote the submatrix resulting from removing the first row and column by M', and note that it only contains the elements $y_2, \ldots, y_{2m-2}$. We develop the determinant of M by the first row, and write $\det(M) = x \cdot \det(M') + f(\bar{y})$, for some function f of $\bar{y}$. By the induction hypothesis, $\det(M')$ is distributed $\frac{m-2}{|\mathbb{F}|}$-close to uniform, so $\Pr[\det(M') = 0] \leq \frac{1}{|\mathbb{F}|} + \frac{m-2}{|\mathbb{F}|} = \frac{m-1}{|\mathbb{F}|}$. For any fixed nonzero value of $\det(M') \neq 0$ and for any fixed value of $\bar{y}$, the function $\det(M)$ is a (nonconstant) affine function of the uniformly chosen x, implying that, conditioned on $\det(M') \neq 0$, the determinant of M is distributed uniformly in $\mathbb{F}$. □

Corollary 12. *For $g_0, \ldots, g_{2d}$ chosen uniformly at random from $GF(2^n)$, the distribution of $D(g_0, \ldots, g_{2d})$ is $\frac{d}{2^n}$-close to uniform (in statistical distance).*

On the other hand, the polynomial D is identically zero on the output of G:

Proposition 13. *For every seed $\bar{s} = (a_1, b_1, \ldots, a_d, b_d) \in GF(2^n)^{2d}$, the expression $D(G(\bar{s}))$ evaluates to zero.*

Proof. We will show that the matrix $A_g^{(d)}$ is singular for every seed, and thus the polynomial $D(g_0, \ldots, g_{2d}) = \det(A_g^{(d)})$, will always take the value zero when evaluated on an output of G. To show that $A_g^{(d)}$ is singular for any seed, we show that its columns are always linearly dependent. More specifically, we show that for any $b_1, \ldots, b_d \in GF(2^n)$ there exist $\lambda_0, \ldots, \lambda_d \in GF(2^n)$, not all zero, such that for all $0 \leq k \leq d$, it holds that $\sum_{i=0}^{d} \lambda_i g_{i+k} = 0$. Letting $\bar{c}_i = (g_i, \ldots, g_{i+d})^T$ denote the i-th column of $A_g^{(d)}$, this means that $\sum_{i=0}^{d} \lambda_i \bar{c}_i$ is the zero vector of $GF(2^n)^{d+1}$.

Consider the polynomial $\Lambda(x) = \prod_{j=1}^{d}(x - b_j)$, the degree d polynomial with roots $b_1, \ldots b_d$, and set each λ_i to be the coefficient of x^i in $\Lambda(x)$. Note that always $\lambda_d = 1$. Then, using the definition of G (i.e., $g_i = \sum_j a_j {b_j}^i$), we get for every $0 \leq k \leq \ell - d$:

$$\begin{aligned}
\sum_{i=0}^{d} \lambda_i g_{i+k} &= \sum_{i=0}^{d} \lambda_i \sum_{j=1}^{d} a_j {b_j}^{i+k} \\
&= \sum_{j=1}^{d} a_j {b_j}^k \cdot \sum_{i=0}^{d} \lambda_i {b_j}^i \\
&= \sum_{j=1}^{d} a_j {b_j}^k \cdot \Lambda(b_j),
\end{aligned}$$

which is 0 as the b_j's are all roots of $\Lambda(x)$. □

We have thus obtained, using the event $D = 0$ in Definition 5:

Theorem 14 (D distinguishes G from random)**.** *For D the determinant of $A_g^{(d)}$, the distributions $D(U_{\ell+1})$ and $D(G(U_{2d}))$ are $(1 - \frac{d+1}{2^n})$-far (in statistical distance), where U_k denotes the uniform distribution over $GF(2^n)^k$.*

3.3 A distribution over bits

Let $F' : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^{(\ell+1)\cdot n}$ decode the $2n$ input bits to two elements $a, b \in GF(2^n)$, and output the concatenation of the representations of the elements $f_0(a,b) \ldots f_\ell(a,b)$, where $f_i(a,b)$ is the i-th output block of the geometric generator of Construction 9.

The following is Corollary 8 in [Tzu] (and follows from our Proposition 10 by Corollary 6 in [Tzu]):

Claim 15. *F' is an $\frac{\ell}{2^n}$-bias generator.*

Analogously, let $G' : \{0,1\}^{2d\cdot n} \to \{0,1\}^{(\ell+1)\cdot n}$ decode its $2dn$ input bits as $2d$ elements $a_1, b_1, \ldots, a_d, b_d \in GF(2^n)$, and output the concatenation of the bit strings representing the output elements of $G(a_1, b_1, \ldots, a_d, b_d)$.

Viola's result [Vio] implies that G' fools polynomials of degree d; we will show an explicit polynomial of degree $d+1$ that distinguishes a random output of G' from a random element of $\{0,1\}^{(\ell+1)\cdot n}$. Having shown that the polynomial D, over $GF(2^n)$, acts significantly differently on an output of G than on random input, we will derive the explicit polynomial in the output bits of G'.

Lemma 16. *Fix an ℓ-variate polynomial $D : GF(2^n)^\ell \to GF(2^n)$ of degree d, and define the mapping $D' : \{0,1\}^{\ell\cdot n} \to \{0,1\}^n$ to treat its input as the representation of ℓ elements $x_1, \ldots, x_\ell \in GF(2^n)$, and output the vector representing $D(x_1, \ldots, x_\ell)$. Then, each of the n output bits of D' is a polynomial of degree at most d in the $\ell \cdot n$ input bits.*

Proof. We will show the claim for a polynomial consisting of a single monomial; the general claim follows from the fact that addition in the field $GF(2^n)$ is exactly bitwise addition in the vector space $\{0,1\}^n$ (Fact 8). We proceed by induction on the degree d. For $d = 0$ the claim is immediate since D is constant. Now fix $d > 0$, and assume without loss of generality that $D(x_1, \ldots, x_\ell) = x_1 \cdot \ldots \cdot x_d$. By Fact 8, the representation of $D(x_1, \ldots, x_\ell)$ is a bilinear expression in the bits of the two vectors x_1 and y_1, where y_1 is the vector representing the multiplication $x_2 \cdot \ldots \cdot x_d$. By the induction hypothesis, every bit in y_1 is a polynomial of degree at most $d-1$ in the bits of $x_2, \ldots x_d$, so each bit of a bilinear form in x_1 and y_1 is a polynomial of degree at most d in the bits of $x_1, \ldots, x_d$. □

Finally, by combining Theorem 14 with Lemma 16, letting D' be the binary version of D (as in Lemma 16), and setting D'_1 to the first bit (say) of D', we obtain our main result:

Theorem 17 (D'_1 distinguishes G' from random)**.** *The polynomial $D'_1 : \{0,1\}^{(\ell+1)\cdot n} \to \{0,1\}$ has degree at most $d+1$ and satisfies:*

$$\left| \Pr_{s \in \{0,1\}^{2d \cdot n}} [D'_1(G'(s)) = 0] - \Pr_{x \in \{0,1\}^{(\ell+1)\cdot n}} [D'_1(x) = 0] \right| \geq \frac{1}{2} - \frac{d}{2^n}.$$

Proof. By Lemma 16, D'_1 indeed has degree at most $d+1$.

By Corollary 12, the distribution of $D(x)$ is $\frac{d}{2^n}$-close to uniform over $GF(2^n)$ when x is chosen uniformly from $GF(2^n)^{\ell+1}$, and thus (by definition) the distribution of $D'(x)$ is $\frac{d}{2^n}$-close to uniform over $\{0,1\}^n$, where x is chosen uniformly from $\{0,1\}^{(\ell+1)\cdot n}$. Specifically, considering the event "first bit is zero" in Definition 5, we have

$$\Pr_{x \in \{0,1\}^{(\ell+1)\cdot n}} [D'_1(x) = 0] \leq \frac{1}{2} + \frac{d}{2^n}.$$

On the other hand, by Proposition 13, $D(G(s)) = 0$ for every $s \in GF(2^n)^{2d}$, giving that $D'(G'(s)) = 0^n$ for every $s \in \{0,1\}^{2d \cdot n}$, and specifically

$$\Pr_{s \in \{0,1\}^{2d \cdot n}} [D'_1(G'(s)) = 0] = 1.$$

The theorem follows. □

3.4 On using larger prime fields

All the above results (including the small-bias generator of Construction 9) generalize naturally to larger prime fields, as does the result of [Vio]. The analogue of Definition 1 will now be (see, e.g., [Eve] or [GW]):

Definition 18. *For $\ell \in \mathbb{N}$, $\varepsilon > 0$ and a prime q, a distribution X over $GF(q)^\ell$ is called ε-*biased *if for every nonzero $\alpha \in GF(q)^\ell$:*

$$\left| \mathbb{E}_{x \sim X} [e^{\langle x, \alpha \rangle \cdot 2\pi i / q}] \right| \leq \varepsilon,$$

where $\langle \alpha, x \rangle$ denotes the inner product $\sum_i \alpha_i \cdot x_i$ over $GF(q)$, and the multiplication by $2\pi i/q$ is then done over the complex field $\mathbb{C}$.

Standard arguments give that in this case,

$$\left| \Pr_{x \sim D}[\langle x, \alpha \rangle = 0] - \frac{1}{q} \right| \leq \sqrt{q-1} \cdot \varepsilon/2.$$

(see, e.g., Appendix B in [BV]).

Generalizing D'_1 of Theorem 17 to $D'^{(q)}_1$, we obtain:

Theorem 19. *For every prime q, the polynomial $D'^{(q)}_1 : GF(q)^{(\ell+1)\cdot n} \to GF(q)$ has degree at most $d+1$ and satisfies:*

$$\left| \Pr_{s \in GF(q)^{2d \cdot n}} \left[D'^{(q)}_1(G'^{(q)}(s)) = 0 \right] - \Pr_{x \in GF(q)^{(\ell+1)\cdot n}} \left[D'^{(q)}_1(x) = 0 \right] \right| \geq 1 - \frac{1}{q} - \frac{d}{q^n}.$$

Acknowledgement

Y. T. would like to thank his advisor, Oded Goldreich, for his help and support.

References

[AGHP] N. Alon, O. Goldreich, J. Hastad and R. Peralta, "Simple Constructions of Almost k-wise Independent Random Variables", *Random Structures and Algorithms*, vol. 3, pp. 289–304, 1992.

[BM] G. Birkhoff and S. MacLane, *A Survey of Modern Algebra*, third edition, MacMillan, New York, 1965.

[Bog] A. Bogdanov, "Pseudorandom generators for low degree polynomials", In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing*, pp. 21–30, 2005.

[BV] A. Bogdanov and E. Viola, "Pseudorandom bits for polynomials", In *48th Annual Symposium on Foundations of Computer Science*, pp. 41–51, 2007.

[Eve] G. Even, "Construction of Small Probabilistic Spaces for Deterministic Simulation", M.Sc. (in computer science) thesis, submitted to the Senate of the Technion (Israel Intitute of Technology) in Aug. 1991.

[GW] O. Goldreich and A. Wigderson, "Tiny Families of Functions with Random Properties: A Quality–Size Trade–off for Hashing", *Journal of Random structures and Algorithms*, vol. 11, No. 4, pp. 315–343, 1997.

[Lov] S. Lovett, "Unconditional pseudorandom generators for low degree polynomials", In *40th Annual Symposium on the Theory of Computing*, pp. 557–562, 2008.

[LVW] M. Luby, B. Velickovic and A. Wigderson, "Deterministic approximate counting of depth-2 circuits", In *Proceedings of the 2nd Israeli Symposium on Theoretical Computer Science*, pp. 18–24, 1993.

[NN] J. Naor and M. Naor, "Small-Bias Probability Spaces: Efficient Constructions and Applications", *SIAM Journal on Computing*, vol. 22, pp. 838–856, 1993.

[Tzu] Y. Tzur, "$GF(2^n)$-Linear Tests versus $GF(2)$-Linear Tests", *Electronic Colloquium on Computational Complexity*, TR 09-018, 2009.

[Vio] E. Viola, "The sum of d small-bias generators fools polynomials of degree d", In *23rd IEEE Conference on Computational Complexity*, 2008.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de