The sum of d small-bias generators fools polynomials of degree d

Emanuele Viola*

April 9, 2008

Abstract

We prove that the sum of d small-bias generators $L : \mathbb{F}^s \to \mathbb{F}^n$ fools degree-d polynomials in n variables over a prime field $\mathbb{F}$, for any fixed degree d and field $\mathbb{F}$, including $\mathbb{F} = \mathbb{F}_2 = \{0, 1\}$. Our result builds on, simplifies, and improves on both the work by Bogdanov and Viola (FOCS '07) and the beautiful follow-up by Lovett (STOC '08). The first relies on a conjecture that turned out to be true only for some degrees and fields, while the latter considers the sum of 2^d small-bias generators (as opposed to d in our result).

1 Introduction

A *pseudorandom generator* $G\colon \mathbb{F}^s \to \mathbb{F}^n$ for polynomials of degree d over a prime field $\mathbb{F}$ is an efficient procedure that stretches s field elements into $n \gg s$ field elements that *fool* any polynomial of degree d in n variables over $\mathbb{F}$: For every such polynomial p, the statistical distance between $p(U)$, for uniform $U \in \mathbb{F}^n$, and $p(G(S))$, for uniform $S \in \mathbb{F}^s$, is at most a small ϵ.

The fundamental case of linear, i.e. degree-1, polynomials is first studied by Naor and Naor [NN] who give a generator with seed length $s = O(\log_{|\mathbb{F}|} n)$ (for error $\epsilon = 1/n$), which is optimal up to constant factors (cf. [AGHP]).[1] This generator is known as *small-bias generator*, and is one of the most celebrated results in pseudorandomness, with a myriad of applications (see, e.g., the references in [BV]).

The case of higher degree is first addressed by Luby, Veličković, and Wigderson [LVW], and a decade later by Bogdanov [Bog]. However, the generators in [LVW, Bog] have poor seed length or only work over very large fields.

Recently, Bogdanov and the author [BV] introduce a new approach to attack this problem over small fields, which we now describe. The work considers the generator $G_k : \mathbb{F}^s \to \mathbb{F}^n$

*viola@cs.columbia.edu. Supported by grants NSF award CCF-0347282 and NSF award CCF-0523664.

[1] Naor and Naor [NN] only consider the case $\mathbb{F} = \mathbb{F}_2$. However, it has been observed by several researchers that their result extends to any prime field.

that is obtained by summing k copies of a small-bias generator $L : \mathbb{F}^{s'} \to \mathbb{F}^n$ by Naor and Naor [NN], which fools linear (i.e., degree-1) polynomials:

$$G_k(s_1, \ldots, s_k) := L(s_1) + \cdots + L(s_k),$$

where the sum is element-wise. [BV] shows that such a generator can be analyzed using the so-called *Gowers norms*. It unconditionally shows that G_d fools polynomials of degree d for $d \leq 3$. For larger $d > 3$, the work proves a conditional result. Specifically, it introduces a special case of a conjecture known as the Gowers inverse conjecture [GT1, Sam]. This special case is called the "d vs. $d-1$ Gowers inverse conjecture" and we subsequently refer to it as "d-GIC." Under d-GIC, [BV] shows that G_d fools polynomials of degree d for every d. Moreover, a counting argument shows that G_d achieves the optimal dependence of the seed length s on the number of variables n, up to additive terms. (In particular, G_{d-1} does not fool polynomials of degree d.)

Subsequently, Lovett [Lov] unconditionally shows that G_{2^d} fools polynomials of degree d, for every d. Lovett's proof does not use the theory of Gowers norms, but it applies to the sum of an exponential number 2^d of small-bias generators, as opposed to d in [BV].

Very recently, Green and Tao [GT2] prove that d-GIC is true *when the field size* $|\mathbb{F}|$ *is bigger than the degree* d *of the polynomial*. Thus, in this case, the approach in [BV] works and in particular one has that G_d fools polynomials of degree d. On the negative side, Green and Tao [GT2], and independently Lovett, Meshulam, and Samorodnitsky [LMS], show that d-GIC is *false* when the field size is much smaller than the degree of the polynomial (which in particular falsifies the more general Gowers inverse conjecture [GT1, Sam]). This falsity prevents the analysis in [BV] to go through for small fields, notably over $\mathbb{F}_2 = \{0, 1\}$. Still, it was left open to understand whether, regardless of the Gowers inverse conjecture, the generator G_d in [BV] fools polynomials of degree d over small fields such as $\mathbb{F}_2$. In this work we answer this question in the affirmative.

1.1 Our results

In this section we state our results. We state them over $\mathbb{F}_2 = \{0, 1\}$ for simplicity, though they hold over any prime field (the necessary details appear in [BV]). Also, we state them for distributions rather than generators; the translation into the language of generators is immediate. Let us start by formalizing the standard notion of *fooling*.

Definition 1 (Fool)**.** *We say that a distribution W on $\{0,1\}^n$* ϵ-fools *degree-d polynomials in n variables over $\mathbb{F}_2$ if for every such polynomial p we have:*

$$|\mathrm{E}_W\, e\,[p(W)] - \mathrm{E}_U\, e\,[p(U)]| \leq \epsilon,$$

where U is the uniform distribution over $\{0,1\}^n$ and $e[x] := (-1)^x$.

The requirement in Definition 1 informally means that degree-d polynomials have advantage at most ϵ in distinguishing a pseudorandom input W from a truly random input U.

This requirement can be equivalently expressed in terms of statistical distance (cf. [BV]), but the above formulation is more convenient for our purposes.

The following is our main theorem.

Theorem 2 (The sum of d small-bias generators fools degree-d polynomials)**.** *Let $Y_1, \ldots, Y_d \in \{0,1\}^n$ be d independent distributions that ϵ-fool degree-1 polynomials in n variables over $\mathbb{F}_2 = \{0,1\}$. Then the distribution $W := Y_1 + \cdots + Y_d$ ϵ_d-fools degree-d polynomials in n variables over $\mathbb{F}_2$ where*

$$\epsilon_d := 16 \cdot \epsilon^{1/2^{d-1}}.$$

Standard constructions of small-bias generators [NN, AGHP] has seed length $O(\log n/\epsilon)$. Plugging these in Theorem 2 gives an explicit generator $\mathbb{F}_2^s \to \mathbb{F}_2^n$ whose output distribution (over random input) ϵ-fools degree-d polynomials with seed length $s = O(d \cdot \log n + d \cdot 2^d \cdot \log(1/\epsilon))$. Folklore constructions of small-bias generators have the more refined seed length $\log n + O(\log(1/\epsilon))$. Plugging these in Theorem 2 gives a generator whose output distribution ϵ-fools degree-d polynomials with seed length $s = d \cdot \log n + O(d \cdot 2^d \cdot \log(1/\epsilon))$, which for fixed d and ϵ is optimal in n up to an additive constant [BV].

Although Theorem 2 improves on previous work [BV, Lov], it still gives nothing for degree $d = \log_2 n$. Whether this barrier can be broken is an interesting open problem that is reminiscent of the analogous open problem in the literature on correlation bounds (cf. [VW]).

2 Proof of Theorem 2

The proof of Theorem 2 builds on and somewhat simplifies [BV, Lov]. Following [BV, Lov], the proofs goes by induction on d. However, it differs in the inductive step. The inductive step in [BV] is a case analysis based on the *Gowers norm* of the polynomial p to be fooled, while the one in [Lov] is a case analysis based on the *Fourier coefficients* of p. The inductive step in this work is in hindsight natural: It is a case analysis based on the *bias* of p, which is the quantity

$$\mathrm{E}_{U \in \{0,1\}^n} e\left[p(U)\right] \in [-1,1].$$

The next Lemma 3 deals with polynomials whose bias is close to 0, whereas Lemma 4 deals with polynomials whose bias is far from 0. The analysis in the case of bias close to 0 (Lemma 3) is the main contribution of this work and departure from [BV, Lov]. The simplification of the inductive step, mentioned above, is less crucial in the sense that one could plug Lemma 3 in the analysis in [Lov] to obtain Theorem 2 with a slightly worse error bound.

Lemma 3 (Fooling polynomials with bias close to 0)**.** *Let $W \in \{0,1\}^n$ be a distribution that ϵ_d-fools degree-d polynomials, and let $Y \in \{0,1\}^n$ be a distribution that ϵ_1-fools degree-1 polynomials. Let p be a polynomial of degree $d+1$ in n variables over $\mathbb{F}_2$. Then*

$$\left|\mathrm{E}_{W,Y}\, e\left[p(W+Y)\right] - \mathrm{E}_U\, e\left[p(U)\right]\right| \le 2 \cdot \left|\, \mathrm{E}_U\, e\left[p(U)\right] \right| + \epsilon_1 + \sqrt{\epsilon_d}.$$

Proof of Lemma 3. We start by an application of the Cauchy-Schwarz inequality which gives

$$\mathrm{E}_{W,Y}\, e\left[p(W+Y)\right]^2 \leq E_W\left[\mathrm{E}_Y\, e\left[p(W+Y)\right]^2\right] = \mathrm{E}_{W,Y,Y'}\, e\left[p(W+Y)+p(W+Y')\right], \quad (1)$$

where Y' is independent from and identically distributed to Y. Now we observe that for every fixed Y and Y', the polynomial $p(U+Y)+p(U+Y')$ has degree d in U, though p has degree $d+1$. Since W ϵ_d-fools degree-d polynomials, we can replace W with the uniform distribution $U \in \{0,1\}^n$:

$$\mathrm{E}_{W,Y,Y'}\, e\left[p(W+Y)+p(W+Y')\right] \leq \mathrm{E}_{U,Y,Y'}\, e\left[p(U+Y)+p(U+Y')\right] + \epsilon_d. \quad (2)$$

At this point, a standard argument shows that

$$\mathrm{E}_{U,Y,Y'}\, e\left[p(U+Y)+p(U+Y')\right] \leq \mathrm{E}_{U,U'}\, e\left[p(U)+p(U')\right] + \epsilon_1^2 = \mathrm{E}_U\, e\left[p(U)\right]^2 + \epsilon_1^2. \quad (3)$$

Therefore, chaining Equations (1), (2), and (3), we have that

$$\begin{aligned}|\mathrm{E}_{W,Y}\, e\left[p(W+Y)\right] - \mathrm{E}_U\, e\left[p(U)\right]| &\leq |\mathrm{E}_{W,Y}\, e\left[p(W+Y)\right]| + |\,\mathrm{E}_U\, e\left[p(U)\right]\,| \leq \\ &\sqrt{\mathrm{E}_U\, e\left[p(U)\right]^2 + \epsilon_1^2 + \epsilon_d} + |\,\mathrm{E}_U\, e\left[p(U)\right]\,| \leq 2\cdot|\,\mathrm{E}_U\, e\left[p(U)\right]\,| + \epsilon_1 + \sqrt{\epsilon_d},\end{aligned}$$

which concludes the proof of the lemma.

For completeness, we include a derivation of Equation (3) next. This equation makes no assumption on p and can be thought of as a form of the so-called expander mixing lemma. The derivation we present uses the Fourier expansion of p: $e(p(x)) = \sum_{\alpha\in\{0,1\}^n} \hat{p}_\alpha \cdot \chi_\alpha(x)$, where $\chi_\alpha(x) := e(\sum_i \alpha_i \cdot x_i)$ is the inner product between α and x. We have:

$$\begin{aligned}
&\mathrm{E}_{U,Y,Y'}\, e\left[p(U+Y)+p(U+Y')\right] \\
=\;& \mathrm{E}_{U,Y,Y'}\left[\left(\sum_{\alpha\in\{0,1\}^n} \hat{p}_\alpha\cdot\chi_\alpha(U+Y)\right)\left(\sum_{\beta\in\{0,1\}^n}\hat{p}_\beta\cdot\chi_\beta(U+Y')\right)\right] \\
=\;& \mathrm{E}_{U,Y,Y'}\left[\sum_{\alpha,\beta}\hat{p}_\alpha\cdot\hat{p}_\beta\cdot\chi_{\alpha+\beta}(U)\cdot\chi_\alpha(Y)\cdot\chi_\beta(Y')\right] \\
&\quad \text{Here we use standard manipulations, e.g. } \chi_\alpha(U+Y) = \chi_\alpha(U)\cdot\chi_\alpha(Y). \\
=\;& \mathrm{E}_{Y,Y'}\left[\sum_{\gamma=\alpha=\beta}\hat{p}_\gamma^2\cdot\chi_\gamma(Y)\cdot\chi_\gamma(Y')\right] \\
&\quad \text{Because } \mathrm{E}_U\, e\left[\chi_{\alpha+\beta}(U)\right] \text{ equals 0 when } \alpha\neq\beta\text{, and 1 otherwise.} \\
=\;& \mathrm{E}_U\, e\left[p(U)\right]^2 + \sum_{\gamma\neq 0}\hat{p}_\gamma^2\cdot\left(E_Y\left[\chi_\gamma(Y)\right]\right)^2 \\
&\quad \text{Because } \hat{p}_0 = \mathrm{E}_U\, e\left[p(U)\right]\text{, and } \chi_0(Y)\equiv 1. \\
\leq\;& \mathrm{E}_U\, e\left[p(U)\right]^2 + \epsilon_1^2\cdot\sum_{\gamma\neq 0}\hat{p}_\gamma^2 \\
&\quad \text{Because } Y\ \epsilon_1\text{-fools degree-1 polynomials such as } \textstyle\sum_i \gamma_i\cdot Y_i. \\
\leq\;& \mathrm{E}_U\, e\left[p(U)\right]^2 + \epsilon_1^2. \\
&\quad \text{Because } \textstyle\sum_{\gamma\neq 0}\hat{p}_\gamma^2 \leq \sum_\gamma \hat{p}_\gamma^2 = 1 \text{ by Parseval's identity.}
\end{aligned}$$

□

We now move to the case of bias far from 0. This case was solved both in [BV] and more compactly in [Lov]. We present a stripped-down version of the solution in [Lov] which is sufficient for our purposes and achieves slightly better parameters.

Lemma 4 (Fooling polynomials with bias far from 0). *Let W be a distribution that ϵ_d-fools degree-d polynomials. Let p be a polynomial of degree $d+1$. Then*

$$|\mathrm{E}_W\, e\,[p(W)] - \mathrm{E}_U\, e\,[p(U)]| \leq \frac{\epsilon_d}{|\mathrm{E}_U\, e\,[p(U)]|}.$$

Proof of Lemma 4. We have the following derivation

$$\begin{aligned} & |\mathrm{E}_W\, e\,[p(W)] - \mathrm{E}_U\, e\,[p(U)]| \cdot |\mathrm{E}_U\, e\,[p(U)]| \\ = \; & |\mathrm{E}_{W,U'}\, e\,[p(W)+p(U')] - \mathrm{E}_{U,U'}\, e\,[p(U)+p(U')]| \\ = \; & |\,\mathrm{E}_{W,U'}\, e\,[p(W)+p(W+U')] - \mathrm{E}_{U,U'}\, e\,[p(U)+p(U+U')]\,| \\ & \qquad \text{Because } U' \text{ is uniformly distributed over } \{0,1\}^n. \\ \leq \; & \mathrm{E}_{U'}\,|\,\mathrm{E}_W\, e\,[p(W)+p(W+U')] - \mathrm{E}_U\, e\,[p(U)+p(U+U')]\,| \leq \epsilon_d, \end{aligned}$$

where in the last inequality we use that for every fixed U' the polynomial $p(x)+p(x+U')$ has degree d in x, though p has degree $d+1$, and that W ϵ_d-fools degree-d polynomials. □

To conclude, we work out the parameters for the proof of Theorem 2.

Proof of Theorem 2. Let ϵ_d be the error for polynomials of degree d, i.e. the maximum over polynomials p of degree d of the quantity

$$|\mathrm{E}_W\, e\,[p(W)] - \mathrm{E}_U\, e\,[p(U)]|.$$

We claim that for every $d > 0$ we have

$$\epsilon_{d+1} \leq 4 \cdot \sqrt{\epsilon_d}. \qquad (\star)$$

Indeed, let p be an arbitrary polynomial of degree $d+1$. If $|\mathrm{E}_U\, e\,[p(U)]| \leq \sqrt{\epsilon_d}$ we have by Lemma 3 that

$$|\mathrm{E}_W\, e\,[p(W)] - \mathrm{E}_U\, e\,[p(U)]| \leq 2 \cdot \sqrt{\epsilon_d} + \epsilon + \sqrt{\epsilon_d} \leq 4 \cdot \sqrt{\epsilon_d},$$

which confirms $(\star)$ in this case. Otherwise, if $|\mathrm{E}_U\, e\,[p(U)]| \geq \sqrt{\epsilon_d}$ we have by Lemma 4 that

$$|\mathrm{E}_W\, e\,[p(W)] - \mathrm{E}_U\, e\,[p(U)]| \leq \frac{\epsilon_d}{\sqrt{\epsilon_d}} = \sqrt{\epsilon_d} \leq 4 \cdot \sqrt{\epsilon_d},$$

which again confirms $(\star)$ in this case.

Finally, from $(\star)$ it follows that

$$\epsilon_d \leq 4^{\sum_{i=0}^{d-2} 2^{-i}} \cdot \epsilon^{1/2^{d-1}} \leq 16 \cdot \epsilon^{1/2^{d-1}}$$

for every d, and thus the theorem is proved. □

Acknowledgments. We thank Avi Wigderson for useful conversations, Omid Etesami for pointing out an inconsistency in the notation used in an earlier version of this work, and Salil Vadhan for helpful comments on the write-up.

References

[AGHP] N. Alon, O. Goldreich, J. Håstad, and R. Peralta. Simple constructions of almost k-wise independent random variables. *Random Structures & Algorithms*, 3(3):289–304, 1992. 1, 3

[Bog] A. Bogdanov. Pseudorandom generators for low degree polynomials. In *STOC'05: Proceedings of the 37th Annual ACM Symposium on Theory of Computing*, pages 21–30, New York, 2005. ACM. 1

[BV] A. Bogdanov and E. Viola. Pseudorandom bits for polynomials. In *48th Annual Symposium on Foundations of Computer Science*. IEEE, Oct. 2007. 1, 2, 3, 5

[GT1] B. Green and T. Tao. An inverse theorem for the Gowers U^3 norm, 2005. arXiv.org:math/0503014. 2

[GT2] B. Green and T. Tao. The distribution of polynomials over finite fields, with applications to the Gowers norms, 2007. arXiv:0711.3191v1. 2

[Lov] S. Lovett. Pseudorandom generators for low degree polynomials. In *Proceedings of the 40th Annual ACM Symposium on the Theory of Computing (STOC)*, Victoria, Canada, 17–20 May 2008. 2, 3, 5

[LMS] S. Lovett, R. Meshulam, and A. Samorodnitsky. Inverse Conjecture for the Gowers norm is false. In *Proceedings of the 40th Annual ACM Symposium on the Theory of Computing (STOC)*, Victoria, Canada, 17–20 May 2008. 2

[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 (ISTCS)*, pages 18–24, 1993. 1

[NN] J. Naor and M. Naor. Small-bias probability spaces: efficient constructions and applications. In *Proceedings of the 22nd Annual ACM Symposium on the Theory of Computing*, pages 213–223, 1990. 1, 2, 3

[Sam] A. Samorodnitsky. Low-degree tests at large distances. In *Proceedings of the 39th Annual ACM Symposium on Theory of Computing, San Diego, CA USA*, 2007. 2

[VW] E. Viola and A. Wigderson. Norms, XOR lemmas, and lower bounds for GF(2) polynomials and multiparty protocols. In *Proceedings of the 22nd Annual Conference on Computational Complexity*. IEEE, June 13–16 2007. To appear in the journal *Theory of Computing*. 3