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

Emanuele Viola*

April 15, 2009

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 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 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 field $\mathbb{F}$ is an efficient procedure that stretches s field elements into $n \gg s$ field elements and *fools* 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 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 $G_k : \mathbb{F}^s \to \mathbb{F}^n$

*Supported by NSF grant CCF-0845003. This work was partially done while the author was a postdoctoral fellow at Columbia University, supported by grants NSF award CCF-0347282 and NSF award CCF-0523664.
Email: viola@ccs.neu.edu

[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 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 Inverse Conjecture for the Gowers norm [GT2, Sam]. This special case is called the "d vs. $d-1$ Inverse Conjecture for the Gowers norm" and we subsequently refer to it as "d-ICG." Under d-ICG, [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].

Recently, Green and Tao [GT1] prove that d-ICG is true over prime fields *of size bigger than the degree d of the polynomial*. On the negative side, Green and Tao [GT1], and independently Lovett, Meshulam, and Samorodnitsky [LMS], show that d-ICG is *false* in some cases over fields of size smaller than the degree of the polynomial (which in particular falsifies the more general Inverse Conjecture for the Gowers norm [GT2, Sam]). This falsity prevents the analysis in [BV] from going through for small fields, notably over $\mathbb{F}_2 = \{0, 1\}$. Still, it was left open to understand whether, regardless of inverse conjectures, 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 first present them over $\mathbb{F}_2 = \{0, 1\}$ and then discuss extensions to larger fields in Section 4. 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 (Fooling). *We say that a distribution W on $\{0, 1\}^n$ ϵ-*fools *degree-d polynomials in n variables over $\mathbb{F}_2 = \{0, 1\}$ 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 $e[x] := (-1)^x$ for $x \in \{0, 1\}$.

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 immediately expressed in terms of statistical distance, 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] have seed length $O(\log(n/\epsilon))$. Plugging these into 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 $\log n + O(\log(1/\epsilon))$, cf. [NN, Section 3.1.2] and [BV]. 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, cf. [BV].

2 Proof of Theorem 2

The proof of Theorem 2 builds on and somewhat simplifies [BV, Lov]. Following [BV, Lov], the proof 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

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

where Y' is independent from and identically distributed to Y. Now we observe that for every fixed Y and Y', the polynomial $p(x+Y) + p(x+Y') = p(x+Y) - p(x+Y')$ has degree

d in x, though p has degree $d+1$. Since W ϵ_d-fools degree-d polynomials, we can replace W with the uniform distribution $U \in \{0,1\}^n$:

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

At this point, a standard argument given below shows that

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

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

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

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 $\chi_\alpha(x) := e[\sum_i \alpha_i \cdot x_i]$ and $\hat{p}_\alpha = \mathrm{E}_U\, e\left[p(U)+\sum_i \alpha_i\cdot U_i\right]$. We have:

$$\begin{aligned}
&\mathrm{E}_{U,Y,Y'}\, e\left[p(U+Y)+p(U+Y')\right]\\
&= \mathrm{E}_{U,Y,Y'}\left[\left(\sum_{\alpha\in\{0,1\}^n} \hat{p}_\alpha\cdot\chi_\alpha(U+Y)\right)\left(\sum_{\beta\in\{0,1\}^n} \hat{p}_\beta\cdot\chi_\beta(U+Y')\right)\right]\\
&= \mathrm{E}_{U,Y,Y'}\left[\sum_{\alpha,\beta} \hat{p}_\alpha\cdot\hat{p}_\beta\cdot\chi_{\alpha+\beta}(U)\cdot\chi_\alpha(Y)\cdot\chi_\beta(Y')\right]
\end{aligned}$$

Here we use standard manipulations, e.g. $\chi_\alpha(U+Y) = \chi_\alpha(U)\cdot\chi_\alpha(Y)$.

$$= \mathrm{E}_{Y,Y'}\left[\sum_{\gamma=\alpha=\beta} \hat{p}_\gamma^2\cdot\chi_\gamma(Y)\cdot\chi_\gamma(Y')\right]$$

Because $\mathrm{E}_U\, e\left[\chi_{\alpha+\beta}(U)\right]$ equals 0 when $\alpha \neq \beta$, and 1 otherwise.

$$= \mathrm{E}_U\, e\left[p(U)\right]^2 + \sum_{\gamma\neq 0} \hat{p}_\gamma^2\cdot\left(E_Y\left[\chi_\gamma(Y)\right]\right)^2$$

Because $\hat{p}_0 = \mathrm{E}_U\, e\left[p(U)\right]$, and $\chi_0(Y) \equiv 1$.

$$\le \mathrm{E}_U\, e\left[p(U)\right]^2 + \epsilon_1^2\cdot\sum_{\gamma\neq 0}\hat{p}_\gamma^2$$

Because Y ϵ_1-fools degree-1 polynomials such as $\sum_i \gamma_i\cdot Y_i$.

$$\le \mathrm{E}_U\, e\left[p(U)\right]^2 + \epsilon_1^2.$$

Because $\sum_{\gamma\neq 0}\hat{p}_\gamma^2 \le \sum_\gamma \hat{p}_\gamma^2 = 1$ by Parseval's identity. □

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

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

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. □

3 Generators vs. correlation bounds

Although Theorem 2 improves on previous work [BV, Lov], it still gives nothing for degree $d = \log_2 n$. The following simple and general proposition, which does not seem to have appeared in the literature, shows that an explicit generator that fools polynomials of degree $d = \log_2 n$ would solve the long-standing problem of obtaining strong correlation bounds for polynomials of the same degree, see [Vio]. Specifically, this connection follows from the next proposition by letting t range over all polynomials of degree $d = \log_2 n$.

Proposition 5 (Generator implies correlation bound)**.** *Let $G : \{0,1\}^s \to \{0,1\}^n$ be any given map. Define the function $f : \{0,1\}^n \to \{0,1\}$ by $f(x) = 1$ iff $x = G(y)$ for some $y \in \{0,1\}^s$. Consider a random $D \in \{0,1\}^n$ that with probability $1/2$ is a uniform $D = U$, and with probability $1/2$ is $D = G(S)$ for a uniform $S \in \{0,1\}^s$.*

If a function $t : \{0,1\}^n \to \{0,1\}$ correlates with f with respect to D, i.e.

$$\mathrm{E}_D\, e\left[t(D) + f(D)\right] \geq \epsilon,$$

then t distinguishes G from random, i.e.

$$\mathrm{E}_U\, e\left[t(U)\right] - \mathrm{E}_S\, e\left[t(G(S))\right] \geq 2\epsilon - 2^{s-n+1}.$$

Proof. We have:

$$\begin{aligned}\epsilon &\leq \frac{1}{2} \cdot \mathrm{E}_U\, e\left[t(U) + f(U)\right] + \frac{1}{2} \cdot \mathrm{E}_S\, e\left[t(G(S)) + f(G(S))\right] \\ &\leq \frac{1}{2}\left(\mathrm{E}_U\, e\left[t(U)\right] + 2^{s-n+1}\right) - \frac{1}{2} \cdot \mathrm{E}_S\, e\left[t(G(S))\right]. \quad \square\end{aligned}$$

4 Generators over larger fields

In this section we explain how our results extend to any finite field $\mathbb{F}$ of size $|\mathbb{F}| > 2$. In this more general case we require our definition of fooling (Definition 1) to hold for every character $e : \mathbb{F} \to \mathbb{C}$. This definition is equivalent to the definition in terms of statistical distance mentioned at the beginning of Section 1, up to a multiplicative loss of $|\mathbb{F}|$ [BV]. As also pointed out in [BV], if $|\mathbb{F}|$ is prime then it is enough to consider the fixed character $e(x) := e^{2 \cdot \pi \cdot i \cdot x / |\mathbb{F}|}$ where in the latter expression $i = \sqrt{-1}$ and e denotes the constant $2.7182\ldots$ The main results, Theorem 2 and Lemmas 3 and 4, continue to hold as stated for any fixed character. The only changes in the proof of Lemma 3 are that Equation (1) becomes

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

note the appearance of the minus sign allowing for the subsequent degree reduction, and that Equation (3) is proved via Fourier analysis over the larger domain. Lemma 4 can be seen to extend to larger fields by multiplying by $|\mathrm{E}_U\, e\left[-p(U)\right]| = |\mathrm{E}_U\, e\left[p(U)\right]|$ in the first step of the proof.

Acknowledgments. We thank Omid Etesami for pointing out an inconsistency in the notation used in an earlier version of this work, Parikshit Gopalan for asking about non-prime fields, Salil Vadhan for helpful comments on the write-up, and Avi Wigderson for useful conversations. We are also grateful to the anonymous referees for useful feedback.

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 *37th Annual Symposium on Theory of Computing (STOC)*, pages 21–30. ACM, 2005. 1

[BV] A. Bogdanov and E. Viola. Pseudorandom bits for polynomials. In *48th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 41–51. IEEE, 2007. 1, 2, 3, 5, 6

[GT1] B. Green and T. Tao. The distribution of polynomials over finite fields, with applications to the Gowers norms, 2007. arXiv:0711.3191v1. 2

[GT2] B. Green and T. Tao. An inverse theorem for the Gowers $U^3(G)$ norm. *Proceedings of the Edinburgh Mathematical Society (Series 2)*, 51(01):73–153, 2008. 2

[Lov] S. Lovett. Unconditional pseudorandom generators for low degree polynomials. In *40th Annual Symposium on the Theory of Computing (STOC)*, pages 557–562. ACM, 2008. 2, 3, 5, 6

[LMS] S. Lovett, R. Meshulam, and A. Samorodnitsky. Inverse Conjecture for the Gowers norm is false. In *40th Annual Symposium on the Theory of Computing (STOC)*, pages 547–556. ACM, 2008. 2

[LVW] M. Luby, B. Veličković, and A. Wigderson. Deterministic approximate counting of depth-2 circuits. In *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. *SIAM J. Comput.*, 22(4):838–856, 1993. 1, 2, 3

[Sam] A. Samorodnitsky. Low-degree tests at large distances. In *39th Annual Symposium on Theory of Computing (STOC)*, pages 506–515. ACM, 2007. 2

[Vio] E. Viola. Correlation bounds for polynomials over $\{0,1\}$. *SIGACT News*, 2009. 6