Electronic Colloquium on Computational Complexity, Report No. 75 (2007)

Unconditional pseudorandom generators for low degree polynomials

Shachar Lovett[1]

August 9, 2007

Abstract

We give an explicit construction of pseudorandom generators against low degree polynomials over finite fields. We show that the sum of 2^d small-biased generators with error $\epsilon^{2^{O(d)}}$ is a pseudorandom generator against degree d polynomials with error ϵ. This gives a generator with seed length $2^{O(d)} \log{(n/\epsilon)}$. Our construction follows the recent breakthrough result of Bogadnov and Viola [BV07]. Their work shows that the sum of d small-biased generators is a pseudo-random generator against degree d polynomials, assuming the Inverse Gowers Conjecture. However, this conjecture is only proven for $d = 2, 3$. The main advantage of our work is that it does not rely on any unproven conjectures.

1 Introduction

We are interested in explicitly constructing pseudorandom generators (PRG) against low degree polynomials over small finite fields. A pseudorandom generator against a family $\mathbb{T}$ of tests is a function G mapping a small domain into a (much) larger one, such that any test $T \in \mathbb{T}$ cannot distinguish with high probability a random element in the large domain from an application of G on a random element in the small domain. We say a PRG requires R random bits, if the size of the small domain is 2^R.

In our case, let $\mathbb{F}$ be a finite field, and a test is a polynomial $p(x_1, ..., x_n)$ over $\mathbb{F}$. The image of the PRG is a small subset of $\mathbb{F}^n$, and it is pseudorandom against $p(x_1, ..., x_n)$ if the distribution of the outcome of p when applied to a random element in the small subset is close to the distribution of the outcome of p when applied to a uniform element in $\mathbb{F}^n$. We are interested in PRG that are pseudorandom against all degree d polynomials, and use as few random bits as possible.

The case of pseudorandom generators against linear polynomials, usually called small-bias generators (or epsilon-biased generators, a term we will not use in this paper to avoid confusion), was first studied (over $\mathbb{F} = \mathbb{F}_2$) by Naor and Naor ([NN90])

[1] Faculty of Mathematics and Computer Science, The Weizmann Institute of Science, POB 26, Rehovot 76100, Israel. Email: Shachar.Lovett@weizmann.ac.il. This research was supported by grant 1300/05 from the Israel Science Foundation.

ISSN 1433-8092

and later by [AGHP92]. Them and others gave explicit constructions, which were later generalized for any finite field. These constructions have seed length which is up to a constant optimal. The construction of small-bias generators was a major tool in derandomization, PCP's and lower bounds (see [BSVW03] for details and more references regarding small-biased generators). The generalization of the problem for constant degree polynomials was first studied by Luby, Veličković, and Wigderson (see [LVW93]). Their construction required $exp(O(\sqrt{\log n/\epsilon}))$ random bits.

Bogdanov ([Bog05]) constructed a PRG that works when the field size is not too small. He presented a construction which converts a hitting set generator to a pseudorandom generator, which combined with a construction for a hitting set yields a PRG. The minimal field size required for his construction to work is a polynomial in the degree, the required error and log the number of the variables. In this setting (field size not too small), his construction achieves better parameters than ours, where the dependence on d in the seed length is polynomial instead of exponential in our construction. His construction uses techniques and results from Algebraic Geometry.

Recently, Bogdanov and Viola ([BV07]) presented a novel approach for constructing PRG for low degree polynomials over small fields. Their construction is the sum of d small-biased generators. They show, under the Inverse Gowers Conjecture, that such a sum is a PRG for degree d polynomials. However, the Inverse Gowers Conjecture is currently only known to hold for degrees 2 and 3, which means that their construction is provably correct only for quadratic and cubic polynomials.

1.1 Our Contribution

Our work is inspired by the recent work of Bogdanov and Viola [BV07]. We start by describing in high level their work, since our work shares and follows some of the ideas in their work. The analysis of [BV07] depends on analyzing the Gowers Norm of a polynomial. We will start by defining it, before describing their construction and proof technique.

The d-th Gowers Norm of a polynomial $p(\mathbf{x})$ looks at derivatives of p in d random directions, and measure how close these derivatives are to being always 0. In other words, we look on the average value of p on random d-dimensional affine subspaces of $\mathbb{F}^n$. Assume currently for simplicity that $\mathbb{F} = \mathbb{F}_2$, and we shorthand $p(\mathbf{x})$ for $p(x_1, ..., x_n)$. The derivative of $p(\mathbf{x})$ in direction $\mathbf{y}$ is given by:

$$p_{\mathbf{y}}(\mathbf{x}) = p(\mathbf{x} + \mathbf{y}) - p(\mathbf{x})$$

Notice that the degree of $p_{\mathbf{y}}$ is one less than the degree of p. More generally, taking k derivatives in directions $\mathbf{y}_1, ..., \mathbf{y}_k$ reduces the degree by k, where the derivative is given by:

$$p_{\mathbf{y}_1,...,\mathbf{y}_k} = \sum_{S \subset \{1,...,k\}} (-1)^{k-|S|} p(\mathbf{x} + \sum_{i \in S} \mathbf{y}_i)$$

The d-th Gowers Norm of p is defined as:

$$U_d(p) = \left(\mathbb{E}_{\mathbf{x}, \mathbf{y}_1, \ldots \mathbf{y}_d \in \mathbb{F}_2^n} \left[(-1)^{p_{\mathbf{y}_1,...,\mathbf{y}_d}(\mathbf{x})} \right] \right)^{\frac{1}{2^d}}$$

Assume p is a degree $d-1$ polynomial. Taking d derivatives from it returns the zero polynomial, so $p_{\mathbf{y}_1,...,\mathbf{y}_d} \equiv 0$ for any choice of $\mathbf{y}_1, ..., \mathbf{y}_d$ and consequently $U_d(p) = 1$. The Inverse Gowers Conjecture aims to show that if $U_d(p)$ is somewhat large, then there is a degree $d-1$ polynomial that is correlated to p. This can be thought of as a generalization of Fourier analysis, which measures the correlation of a function to a linear function. Actually, the 2-nd Gowers Norm U_2 is tightly related to the Fourier coefficients of the polynomial. The Inverse Gowers Conjecture is currently only proven for U_3, and even there the results are far from what is believed to be true (see [GT05] and [Sam07] for a more detailed discussion on the Gowers norm, and a proof of the Inverse Gowers Conjecture for U_3).

Returning to the argument in [BV07], they analyze the Gowers Norm of a degree-d polynomial $p(\mathbf{x})$, and show a win-win situation for the case when it is small or large. In the first case, when the Gowers Norm is small, they show that the sum of d small-bias generators is pseudorandom against $p(\mathbf{x})$, by relating the distribution of $p(\mathbf{x}_1 + ... + \mathbf{x}_d)$ to the Gowers Norm of p. In the second case, when the Gowers Norm is large, and assuming the Inverse Gowers Conjecture, $p(\mathbf{x})$ is correlated to some degree-$d-1$ polynomial $q(\mathbf{x})$. They use $q(\mathbf{x})$ in order to construct a circuit that computes $p(\mathbf{x})$ for almost all x's. The inputs to this circuit are all degree $d-1$ polynomials, and thus they show that a PRG for degree $d-1$ polynomials with small enough error is also pseudorandom against $p(\mathbf{x})$.

Our construction follows similar lines, however instead of analyzing the Gowers norm of $p(\mathbf{x})$, we analyze its Fourier coefficients. We also divide our treatment into two cases - when p has some large Fourier coefficient, and when all the Fourier coefficients of p are small.

In the first case, when $p(\mathbf{x})$ has no large Fourier coefficients, we look on inputs to p of the form $\mathbf{x} + \mathbf{y}$, where $\mathbf{x}$ and $\mathbf{y}$ are independent. we consider the polynomial

$$\Delta p(\mathbf{x}', \mathbf{x}'', \mathbf{y}', \mathbf{y}'') = p(\mathbf{x}' + \mathbf{y}') - p(\mathbf{x}' + \mathbf{y}'') - p(\mathbf{x}'' + \mathbf{y}') + p(\mathbf{x}'' + \mathbf{y}'')$$

We prove that it is enough to be pseudorandom against Δp in order to be pseudorandom against $p(\mathbf{x} + \mathbf{y})$, and also that in order to do so, it is enough to have $\mathbf{x}$,$\mathbf{x}'$,$\mathbf{x}''$,$\mathbf{y}$,$\mathbf{y}'$ and $\mathbf{y}''$ come from a PRG that is pseudorandom against degree $d-1$ polynomials. The reason is that Δp contains no degree d terms in just one of $\mathbf{x}', \mathbf{x}'', \mathbf{y}'$ or $\mathbf{y}''$. In the second case, when there is some large Fourier coefficient, we know that $p(\mathbf{x})$ is correlated to some linear function. Similarly to the second case in the work of [BV07], we also show in that case, or more generally when $p(\mathbf{x})$ is correlated to some lower degree polynomial, a PRG for degree $d-1$ polynomial with small enough error is also pseudorandom against $p(\mathbf{x})$. However, our proof technique is much more direct than the one used in [BV07], which results in better parameters and also a simpler analysis.

We compare now the parameters obtained by [BV07] and by our construction. Assume we want a PRG against degree d polynomials with error ϵ (for exact definition, see the Preliminaries Section). The construction of [BV07] requires d independent small-bias generators. The required error from the small-biased generators depends on the parameters in which the Gowers Inverse Conjecture can be proven. If we assume optimal results for the Gowers Inverse Conjecture, each of the small-bias generators should have error $\epsilon^{2^{O(d)}}$. So, the PRG has seed length of $d(O(\log n) + 2^{O(d)} \log(1/\epsilon))$.

However, the Gowers Inverse Conjecture is currently only proven for $d = 2$ and $d = 3$. The case of $d = 2$ is exactly Affinity Testing, and near optimal results can be proved using Fourier Analysis (see for example [BLR93]). In the case of $d = 3$, the proven relation between $U_3(p)$ and the proximity of p to quadratic polynomials is far worse than what is believed to be true. This results in making the seed length of the PRG of [BV07] for cubic polynomials much worse than what it might be.

Our PRG construction is the sum of 2^d small-bias generators with error $\epsilon^{2^{O(d)}}$. This gives a PRG with seed length $2^{O(d)} \log(n/\epsilon)$. This is worse than the seed length in the [BV07] construction assuming optimal parameters in the Inverse Gowers Conjecture, but for the parameter range of $\epsilon < 1/poly(n)$ the constructions are equivalent up to a constant.

Summarizing, our construction has the following advantages: First, it is unconditional, and relies on no unproven conjectures. Second, even in the case of U_3 (i.e. PRG against cubic polynomials), the proven parameters for the Inverse Gowers Conjecture are probably far worse than what is believed to be true, which results in much worse seed length of the PRG of [BV07] than the one we get. Third, we present a much simpler analysis for the case when $p(\mathbf{x})$ is correlated to some lower degree polynomial. Notice that this analysis can also be applied for analyzing the construction of [BV07], but it still falls short from proving their construction, because the Inverse Gowers Conjecture must also be proven.

2 Preliminaries

We work over an arbitrary finite field $\mathbb{F}$. Let $U = U_n$ be the uniform distribution over $\mathbb{F}^n$. We fix $e : \mathbb{F} \to \mathbb{C}$ to be any (non-trivial) character. For example, in a prime field $\mathbb{F}_p$ we can have $e(x) = w^x$ for w root of unity of order p.

When we talk about a degree of a multivariate polynomial, we always mean its total degree. We mark elements of $\mathbb{F}^n$ by $\mathbf{x} = (x_1, ..., x_n)$.

Definition 1. A distribution D over $\mathbb{F}^n$ is said to pseudorandom against a polynomial $p(x_1, ..., x_n)$ with error ϵ if

$$|\mathbb{E}_{\mathbf{x}\in D}[e(p(\mathbf{x}))] - \mathbb{E}_{\mathbf{x}\in U}[e(p(\mathbf{x}))]| < \epsilon$$

Definition 2. A distribution D is said to be pseudorandom against degree d polynomials with error ϵ, if for every degree d polynomial $p(x_1, ..., x_n)$, D is pseudorandom against p with error ϵ.

This notion of pseudorandomness is tightly related to other notions. For example we give the following simple lemma without proof.

Lemma 1. *Let D be a distribution that is pseudorandom against degree d polynomials with error ϵ. Let $p(x_1, ..., x_n)$ be a polynomial of degree at most d. Let $X_D \in \mathbb{F}$ be the random variable of applying p on D, and $X_U \in \mathbb{F}$ similarly the random variable of applying p on U. Then the variation (L_1) distance between X_D and X_U is bounded by:*

$$|X_D - X_U|_1 \leq (|\mathbb{F}| - 1)\epsilon$$

We will use in the paper the following basic properties of characters:

1. For every $x, y \in F$, $e(x)e(y) = e(x+y)$.
2. For every $x \in \mathbb{F}$, $e(-x) = \overline{e(x)}$, the complex conjugate of $e(x)$

We use Fourier analysis in our analysis. We define now the basic facts required for our analysis in the paper.

Definition 3. The Fourier coefficients of a function $f : \mathbb{F}^n \to \mathbb{C}$ are defined by

$$\hat{f}_\alpha = \mathbb{E}_{\mathbf{x} \in U}[f(\mathbf{x})e(-\langle \alpha, \mathbf{x} \rangle)]$$

where $\alpha = (\alpha_{\mathbf{1}}, ..., \alpha_{\mathbf{n}}) \in \mathbb{F}^{\mathbf{n}}$ and $\langle \alpha, \mathbf{x} \rangle = \alpha_1 x_1 + ... + \alpha_n x_n$ is the inner product between α and $\mathbf{x}$. It is well known that set of functions $e(\langle \alpha, \mathbf{x} \rangle)$ is an orthonormal basis over $\mathbb{F}^n$, and that

$$f(\mathbf{x}) = \sum_{\alpha \in \mathbb{F}^{\mathbf{n}}} \hat{f}_\alpha e(\langle \alpha, \mathbf{x} \rangle)$$

For a polynomial $p(\mathbf{x}) \in \mathbb{F}[x_1, ..., x_n]$ we define $\hat{p}_\alpha$ to be the α Fourier coefficient of the function $e(p(\mathbf{x}))$. We will also use the following simple fact, which follows from Parseval's identity and because $|e(p(\mathbf{x}))| = 1$ for all $\mathbf{x} \in \mathbb{F}^n$:

Fact.

$$\sum_{\alpha \in \mathbb{F}^{\mathbf{n}}} |\hat{p}_\alpha|^2 = 1$$

The base of our analysis are PRG for degree-1 polynomials, a.k.a linear polynomials. PRG for this family have been studied extensively, and are usually referred to as small-bias generators or distributions. Formally we define:

Definition 4. A distribution D is called small-bias over $\mathbb{F}^n$ with error δ, if for all linear polynomials $p(\mathbf{x}) = a_1 x_1 + ... + a_n x_n$ we have:

$$|\mathbb{E}_{\mathbf{x} \in D}[e(p(\mathbf{x}))] - \mathbb{E}_{\mathbf{x} \in U}[e(p(\mathbf{x}))]| < \delta \tag{1}$$

Constructions of small-bias distributions have first been studied in [NN90], and optimal up to constant constructions can be found in [AGHP92]. Such constructions require $O(\log(n/\epsilon))$ random bits for the seed (this means they are functions from $\{0,1\}^{O(\log(n/\epsilon))} \to \mathbb{F}^n$). We now state our main theorem.

3 Main theorem

We now state our main theorem:

Theorem 2. *Let D be a small-biased distribution over $\mathbb{F}^n$ with error $O(\epsilon)^{4^d}$. The sum of 2^d independent copies of D is pseudorandom against degree d polynomials with error ϵ. In particular, this gives a pseudorandom generator for degree d polynomials with error ϵ using $2^{O(d)} \log(n/\epsilon)$ random bits for seed.*

Our analysis will basically be a case analysis whether p has some large Fourier coefficient or not. We will show that when a degree d polynomial $p(\mathbf{x})$ has some large Fourier coefficient, then a PRG for degree $d-1$ with a somewhat better error is also pseudorandom against p, while if p are no large Fourier coefficients it is "pseudorandom" in some sense, and then the sum of two PRG for degree $d-1$ is pseudorandom against p.

We divide the proof into two technical lemmas, dealing with the cases whether p has some large Fourier coefficient, or not.

Lemma 3. *Let $p(x_1, ..., x_n)$ be a degree d polynomial over $\mathbb{F}^n$, such that for all $\alpha \in \mathbb{F}^{\mathbf{n}}$, $|\hat{p}_\alpha| < \epsilon^2/10$. Let D be a distribution that is pseudorandom against degree $d-1$ polynomials with error $\epsilon^4/400$. Then $\boldsymbol{x} + \boldsymbol{y}$, where $\boldsymbol{x}, \boldsymbol{y}$ are independently chosen from D, is pseudorandom against p with error ϵ.*

Lemma 4. *Let $p(x_1, ..., x_n)$ be a degree d polynomial over $\mathbb{F}^n$, such that $|\hat{p}_\alpha| \geq \epsilon^2/10$ for some $\alpha \in \mathbb{F}^{\mathbf{n}}$. Let D be a distribution that is pseudorandom against degree $d-1$ polynomials with error $\epsilon^3/10$. Then D is pseudorandom against $p(\boldsymbol{x})$ with error ϵ.*

Assuming those two lemma, our main theorem now follows directly, by also using the following simple observation. This observation allows us to add "extra" small-bias distributions without harming our PRG construction.

Observation 5. *Let D be a distribution that is pseudorandom against degree d polynomials with error ϵ. Let D' be any other independent distribution. Then $D + D'$ is also pseudorandom against degree d polynomials with error ϵ.*

The remaining of the paper is organized as follows. Lemma 3 is proven in Section 4. Lemma 4 in Section 5.

4 Case I: No large fourier coefficients

In this section we prove Lemma 3. We assume throughout this section that all the Fourier coefficients of $e(p(\mathbf{x}))$ are small, i.e. $|\hat{p}_\alpha| < \epsilon^2/10$ for all $\alpha \in \mathbb{F}^{\mathbf{n}}$.

We start by defining a derivation polynomial.

Definition 5. Let $p(\mathbf{x}) : \mathbb{F}^n \to \mathbb{F}$ be a polynomial. We define its derivation polynomial $\Delta p : (\mathbb{F}^n)^4 \to \mathbb{F}$ as:

$$\Delta p(\mathbf{x}', \mathbf{x}'', \mathbf{y}', \mathbf{y}'') = p(\mathbf{x}' + \mathbf{y}') - p(\mathbf{x}'' + \mathbf{y}') - p(\mathbf{x}' + \mathbf{y}'') + p(\mathbf{x}'' + \mathbf{y}'')$$

The following lemma is crucial in our construction, and uses in part a lemma from [BV07]. We relate the distribution of p on the sum of two independent inputs, to that of Δp.

Lemma 6. *Let $p : \mathbb{F}^n \to \mathbb{F}$. Let D be a distribution over $\mathbb{F}^n$. Let $\boldsymbol{x}, \boldsymbol{y}$ be independently chosen from D, then*

$$|\mathbb{E}_{\boldsymbol{x},\boldsymbol{y} \in D}[e(p(\boldsymbol{x} + \boldsymbol{y}))]|^4 \leq \mathbb{E}_{\boldsymbol{x}',\boldsymbol{x}'',\boldsymbol{y}',\boldsymbol{y}'' \in D}[e(\Delta p(x', x'', y', y''))]$$

where $\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{y}', \boldsymbol{y}''$ are also independent.

Proof. The proof is essentially applying the Cauchy-Schwartz inequality twice. We start by showing:

$$|\mathbb{E}_{\mathbf{x},\mathbf{y}\in D}[e(p(\mathbf{x}+\mathbf{y}))]|^2 \leq \mathbb{E}_{\mathbf{x},\mathbf{y}',\mathbf{y}''\in D}[e(p(\mathbf{x}+\mathbf{y}')-p(\mathbf{x}+\mathbf{y}''))]$$

and then continue to show:

$$|\mathbb{E}_{\mathbf{x},\mathbf{y}\in D}[e(p(\mathbf{x}+\mathbf{y}))]|^4 \leq \mathbb{E}_{\mathbf{x}',\mathbf{x}'',\mathbf{y}',\mathbf{y}''\in D}[e(p(\mathbf{x}'+\mathbf{y}')-p(\mathbf{x}''+\mathbf{y}')-p(\mathbf{x}'+\mathbf{y}'')+p(\mathbf{x}''+\mathbf{y}''))]$$

which is what we want to prove, by the definition of Δp. We prove the first part by applying the Cauchy-Schwartz inequality:

$$\begin{aligned}&|\mathbb{E}_{\mathbf{x},\mathbf{y}\in D}[e(p(\mathbf{x}+\mathbf{y}))]|^2 \leq \mathbb{E}_{\mathbf{x}\in D}|\mathbb{E}_{\mathbf{y}\in D}[e(p(\mathbf{x}+\mathbf{y}))]|^2 = \\ &\mathbb{E}_{\mathbf{x}\in D}\mathbb{E}_{\mathbf{y}'\in D}[e(p(\mathbf{x}+\mathbf{y}'))]\overline{\mathbb{E}_{\mathbf{y}''\in D}[e(p(\mathbf{x}+\mathbf{y}''))]} = \\ &\mathbb{E}_{\mathbf{x},\mathbf{y}',\mathbf{y}''\in D}[e(p(\mathbf{x}+\mathbf{y}')-p(\mathbf{x}+\mathbf{y}''))]\end{aligned}$$

We prove the second part by applying the Cauchy-Schwartz inequality again:

$$\begin{aligned}&|\mathbb{E}_{\mathbf{x},\mathbf{y}\in D}[e(p(\mathbf{x}+\mathbf{y}))]|^4 \leq |\mathbb{E}_{\mathbf{x},\mathbf{y}',\mathbf{y}''\in D}[e(p(\mathbf{x}+\mathbf{y}')-p(\mathbf{x}+\mathbf{y}''))]|^2 \leq \\ &\mathbb{E}_{\mathbf{y}',\mathbf{y}''\in D}|\mathbb{E}_{\mathbf{x}\in D}[e(p(\mathbf{x}+\mathbf{y}')-p(\mathbf{x}+\mathbf{y}''))]|^2 = \\ &\mathbb{E}_{\mathbf{x}',\mathbf{x}'',\mathbf{y}',\mathbf{y}''\in D}[e(p(\mathbf{x}'+\mathbf{y}')-p(\mathbf{x}'+\mathbf{y}'')-p(\mathbf{x}''+\mathbf{y}')+p(\mathbf{x}''+\mathbf{y}''))]\end{aligned}$$

□

In particular the following correlation follows:

Corollary 7. $\mathbb{E}_{\boldsymbol{x}',\boldsymbol{x}'',\boldsymbol{y}',\boldsymbol{y}''\in D}[e(\Delta p(\boldsymbol{x}',\boldsymbol{x}'',\boldsymbol{y}',\boldsymbol{y}''))] \geq 0$

We analyze the expression $\mathbb{E}_{\mathbf{x}',\mathbf{x}'',\mathbf{y}',\mathbf{y}''\in D}[e(\Delta p(\mathbf{x}',\mathbf{x}'',\mathbf{y}',\mathbf{y}''))]$ in two cases, when $D = U$ is the uniform distribution, and when D is a PRG for degree $d-1$ polynomials. We show that in both cases it is at most $\epsilon/2$. Combining this with Lemma 6 yields the required result. We start our analysis in the uniform case.

We begin by showing the (well-known) connection between the average value of Δp to the Fourier coefficients of p, regarding Δp as a linearity-test for p. See for example [BLR93] for a similar analysis in more depth.

Lemma 8.

$$\mathbb{E}_{\boldsymbol{x}',\boldsymbol{x}'',\boldsymbol{y}',\boldsymbol{y}''\in U}[e(p(\boldsymbol{x}'+\boldsymbol{y}')-p(\boldsymbol{x}'+\boldsymbol{y}'')-p(\boldsymbol{x}''+\boldsymbol{y}')+p(\boldsymbol{x}''+\boldsymbol{y}''))] = \sum_{\alpha\in\mathbb{F}^{\mathbf{n}}}|\hat{p}_\alpha|^4$$

Proof. We can write $e(p(\mathbf{x}))$ in the Fourier basis as:

$$e(p(\mathbf{x})) = \sum_{\alpha\in\mathbb{F}^{\mathbf{n}}}\hat{p}_\alpha e(\langle\alpha,\mathbf{x}\rangle)$$

Notice that:

$$e(-p(\mathbf{x})) = \overline{e(p(\mathbf{x}))} = \sum_{\alpha\in\mathbb{F}^{\mathbf{n}}}\overline{\hat{p}_\alpha}e(-\langle\alpha,\mathbf{x}\rangle)$$

We expand now all four terms of p in

$$e(p(\mathbf{x}'+\mathbf{y}')-p(\mathbf{x}'+\mathbf{y}'')-p(\mathbf{x}''+\mathbf{y}')+p(\mathbf{x}''+\mathbf{y}''))$$

This is equal to:

$$\sum_{\alpha_{\mathbf{1}},\alpha_{\mathbf{2}},\alpha_{\mathbf{3}},\alpha_{\mathbf{4}}\in\mathbb{F}^{\mathbf{n}}} \hat{p}_{\alpha_{\mathbf{1}}} e(\langle\alpha_{\mathbf{1}},\mathbf{x}'+\mathbf{y}'\rangle)\overline{\hat{p}_{\alpha_{\mathbf{2}}}}e(-\langle\alpha_{\mathbf{2}},\mathbf{x}'+\mathbf{y}''\rangle) \\ \overline{\hat{p}_{\alpha_{\mathbf{3}}}}e(-\langle\alpha_{\mathbf{3}},\mathbf{x}''+\mathbf{y}'\rangle)\hat{p}_{\alpha_{\mathbf{4}}}e(\langle\alpha_{\mathbf{4}},\mathbf{x}''+\mathbf{y}''\rangle)$$

Remember we are interested in the expected value over uniform $\mathbf{x}',\mathbf{x}'',\mathbf{y}',\mathbf{y}''\in\mathbb{F}^n$, i.e. in:

$$\mathbb{E}_{\mathbf{x}',\mathbf{x}'',\mathbf{y}',\mathbf{y}''\in U}[e(p(\mathbf{x}'+\mathbf{y}')-p(\mathbf{x}'+\mathbf{y}'')-p(\mathbf{x}''+\mathbf{y}')+p(\mathbf{x}''+\mathbf{y}''))]$$

We now use the Fourier expansion, and group elements by their related values. After doing so, the above expectation is equal to:

$$\sum_{\alpha_{\mathbf{1}},\alpha_{\mathbf{2}},\alpha_{\mathbf{3}},\alpha_{\mathbf{4}}\in\mathbb{F}^{\mathbf{n}}} \hat{p}_{\alpha_{\mathbf{1}}}\overline{\hat{p}_{\alpha_{\mathbf{2}}}\hat{p}_{\alpha_{\mathbf{3}}}}\hat{p}_{\alpha_{\mathbf{4}}}\mathbb{E}_{x'\in U}[e(\langle\alpha_{\mathbf{1}}-\alpha_{\mathbf{2}},\mathbf{x}'\rangle)]\mathbb{E}_{x''\in U}[e(\langle\alpha_{\mathbf{4}}-\alpha_{\mathbf{3}},\mathbf{x}''\rangle)] \\ \mathbb{E}_{y'\in U}[e(\langle\alpha_{\mathbf{1}}-\alpha_{\mathbf{3}},\mathbf{y}'\rangle)]\mathbb{E}_{y''\in U}[e(\langle\alpha_{\mathbf{4}}-\alpha_{\mathbf{2}},\mathbf{y}''\rangle)]$$

The term inside the sum for $\alpha_{\mathbf{1}},...,\alpha_{\mathbf{4}}$ is zero unless $\alpha_{\mathbf{1}}=\alpha_{\mathbf{2}}=\alpha_{\mathbf{3}}=\alpha_{\mathbf{4}}=\alpha$, and in that case its contribution is $|\hat{p}_\alpha|^4$. This finishes the proof of the lemma. □

We now use this relation between between Δp and the Fourier coefficients of p to show that the expected value of Δp is small.

Lemma 9. $|\mathbb{E}_{\boldsymbol{x}',\boldsymbol{x}'',\boldsymbol{y}',\boldsymbol{y}''\in U}[e(\Delta p(\boldsymbol{x}',\boldsymbol{x}'',\boldsymbol{y}',\boldsymbol{y}''))]|<\epsilon^4/100$

Proof. We use Lemma 8. We have:

$$\mathbb{E}_{\mathbf{x}',\mathbf{x}'',\mathbf{y}',\mathbf{y}''\in U}[e(p(\mathbf{x}'+\mathbf{y}')-p(\mathbf{x}'+\mathbf{y}'')-p(\mathbf{x}''+\mathbf{y}')+p(\mathbf{x}''+\mathbf{y}''))]=\sum_{\alpha\in\mathbb{F}^{\mathbf{n}}}|\hat{p}_\alpha|^4$$

We combine now the fact that $\sum_{\alpha\in\mathbb{F}^{\mathbf{n}}}|\hat{p}_\alpha|^2=1$ and our assumption that $|\hat{p}_\alpha|<\epsilon^2/10$ for all $\alpha\in\mathbb{F}^{\mathbf{n}}$, to yield the required bound. □

Combining Lemmas 6 and 9 we get that:

$$|\mathbb{E}_{\mathbf{x},\mathbf{y}\in U}[e(p(\mathbf{x}+\mathbf{y}))]|<\left(\frac{\epsilon^4}{100}\right)^{1/4}<\epsilon/2$$

We now turn to handle the pseudorandom case. We start by the following observation:

Observation 10. *The polynomial $\Delta p(\boldsymbol{x}',\boldsymbol{x}'',\boldsymbol{y}',\boldsymbol{y}'')$ has total degree d, but has no degree d terms which have variables from only one of $\boldsymbol{x}',\boldsymbol{x}'',\boldsymbol{y}',\boldsymbol{y}''$. So, the total degrees of variables from $\boldsymbol{x}'$ in each term is at most $d-1$. The same is true also for $\boldsymbol{x}'',\boldsymbol{y}'$ and $\boldsymbol{y}''$.*

We now show that if D is a distribution that is pseudorandom against degree $d-1$ polynomials, then it also pseudorandom against Δp. We use a hybrid argument similar to the one in [BV07].

Lemma 11. *Let D be a distribution that is pseudorandom against degree $d-1$ polynomials with error δ. Then:*

$$|\mathbb{E}_{\boldsymbol{x}',\boldsymbol{x}'',\boldsymbol{y}',\boldsymbol{y}''\in D}[e(\Delta p(\boldsymbol{x}',\boldsymbol{x}'',\boldsymbol{y}',\boldsymbol{y}''))]- \\ \mathbb{E}_{\boldsymbol{x}',\boldsymbol{x}'',\boldsymbol{y}',\boldsymbol{y}''\in U}[e(\Delta p(\boldsymbol{x}',\boldsymbol{x}'',\boldsymbol{y}',\boldsymbol{y}''))]| < 4\delta$$

Proof. We will change the inputs $\mathbf{x}',\mathbf{x}'',\mathbf{y}'$ and $\mathbf{y}''$ one by one from U to D, and we will show that the expected value of $e(\Delta p)$ changes by at most δ in each step, accumulating to a total of at most 4δ. Formally, let H_k $(k = 0..4)$ be the joint distribution of $\mathbf{x}',\mathbf{x}'',\mathbf{y}',\mathbf{y}''$, when the first k are taken from D and the last $4-k$ taken from U. For example, H_1 is the distribution where $\mathbf{x}' \in D$ and $\mathbf{x}'',\mathbf{y}',\mathbf{y}'' \in U$, where $\mathbf{x}',\mathbf{x}'',\mathbf{y}',\mathbf{y}''$ are independent.

We prove that the distance between $e(\Delta p)$ under H_{k-1} and H_k is at most δ, for all $k = 1,2,3,4$. For the sake of clarity, we will focus on the proof for $k = 1$. The proof for the other three cases is essentially identical.

For $k = 1$, we want to show that:

$$|\mathbb{E}_{\mathbf{x}',\mathbf{x}'',\mathbf{y}',\mathbf{y}''\in U}[e(\Delta p(\mathbf{x}',\mathbf{x}'',\mathbf{y}',\mathbf{y}''))] - \mathbb{E}_{\mathbf{x}'\in D,\mathbf{x}'',\mathbf{y}',\mathbf{y}''\in U}[e(\Delta p(\mathbf{x}',\mathbf{x}'',\mathbf{y}',\mathbf{y}''))]| < \delta$$

The joint distribution of $\mathbf{x}'',\mathbf{y}',\mathbf{y}''$ is identical in both terms, so we have:

$$|\mathbb{E}_{\mathbf{x}',\mathbf{x}'',\mathbf{y}',\mathbf{y}''\in U}[e(\Delta p(\mathbf{x}',\mathbf{x}'',\mathbf{y}',\mathbf{y}''))] - \mathbb{E}_{\mathbf{x}'\in D,\mathbf{x}'',\mathbf{y}',\mathbf{y}''\in U}[e(\Delta p(\mathbf{x}',\mathbf{x}'',\mathbf{y}',\mathbf{y}''))]| \leq \\ \mathbb{E}_{\mathbf{x}'',\mathbf{y}',\mathbf{y}''\in U}|\mathbb{E}_{\mathbf{x}'\in U}[e(\Delta p(\mathbf{x}',\mathbf{x}'',\mathbf{y}',\mathbf{y}''))] - \mathbb{E}_{\mathbf{x}'\in D}[e(\Delta p(\mathbf{x}',\mathbf{x}'',\mathbf{y}',\mathbf{y}''))]|$$

Now, for any fixing of values for $\mathbf{x}'' = a$,$\mathbf{y}' = b$,$\mathbf{y}'' = c$, $\Delta p(\mathbf{x}',a,b,c)$ is a polynomial just in $\mathbf{x}'$. Observation 10 tells us it is a polynomial of degree at most $d-1$. Since D is pseudorandom against degree $d-1$ polynomials, the inequality follows for every fixing of $\mathbf{x}'',\mathbf{y}',\mathbf{y}''$, and so also for the expected value.

□

If we take D to be PRG for degree $d-1$ polynomials with error $\epsilon^4/400$, and combine this with Lemmas 6 and 9, we get that:

$$|\mathbb{E}_{\mathbf{x}',\mathbf{x}'',\mathbf{y}',\mathbf{y}''\in D}[e(\Delta p(\mathbf{x}',\mathbf{x}'',\mathbf{y}',\mathbf{y}''))]| < \frac{\epsilon^4}{100} + 4\frac{\epsilon^4}{400} = \frac{\epsilon^4}{50}$$

and so using Lemma 6 we get that:

$$|\mathbb{E}_{\mathbf{x},\mathbf{y}\in D}[e(p(\mathbf{x}+\mathbf{y}))]| < (\frac{\epsilon^4}{50})^{1/4} < \epsilon/2$$

This finishes the proof of Lemma 3.

5 Case II: There is some large Fourier coefficient

In this section we prove Lemma 4. We assume throughout this section that p has some large Fourier coefficient of p. To be precise, there is some $\alpha \in \mathbb{F}^{\mathbf{n}}$ s.t:

$$|\hat{p}_\alpha| \geq \epsilon^2/10$$

Let $l(\mathbf{x})$ be the corresponding linear function, i.e. $l(\mathbf{x}) = \langle \mathbf{x}, \alpha \rangle$. Define:

$$\eta = \overline{\hat{p}_\alpha} = \mathbb{E}_{\mathbf{x} \in U}[e(l(\mathbf{x}) - p(\mathbf{x}))]$$

η is a measure for the approximation of $p(\mathbf{x})$ by $l(\mathbf{x})$. By our assumption on $\hat{p}_\alpha$, we know that $|\eta| \geq \epsilon^2/10$.

For any constant $\mathbf{a} \in \mathbb{F}^n$ define the polynomial:

$$q_{\mathbf{a}}(\mathbf{x}) = p(\mathbf{x}) - p(\mathbf{x} + \mathbf{a}) + l(\mathbf{x} + \mathbf{a})$$

Notice that $q_{\mathbf{a}}(\mathbf{x})$ has degree at most $d-1$, because $l(\mathbf{x}+\mathbf{a})$ is linear (and so of degree less than d), and the degree-d terms in $p(\mathbf{x})$ and $p(\mathbf{x}+\mathbf{a})$ cancel out.

We can think of $q_{\mathbf{a}}(\mathbf{x})$ as using $l(\mathbf{x})$, which approximates $p(\mathbf{x})$ non-uniformly, and the derivative of $p(\mathbf{x})$ in a random direction $\mathbf{a}$, to build a random degree $d-1$ polynomial which approximates $p(\mathbf{x})$ uniformly. In order to show this formally, we define $\nu_{\mathbf{x}}(\mathbf{a}) = \frac{1}{\eta} e(q_{\mathbf{a}}(\mathbf{x}))$. We will see that $\nu_{\mathbf{x}}(\mathbf{a})$, taken on a random $\mathbf{a} \in \mathbb{F}^n$ value, is exactly $e(p(\mathbf{x}))$.

Lemma 12. *For every $\boldsymbol{x} \in F^n$, $\mathbb{E}_{\boldsymbol{a} \in U}[\nu_{\boldsymbol{x}}(\boldsymbol{a})] = e(p(\boldsymbol{x}))$.*

Proof. $\mathbb{E}_{\mathbf{a} \in U}[\nu_{\mathbf{x}}(\mathbf{a})] = \frac{1}{\eta} e(p(\mathbf{x})) \mathbb{E}_{\mathbf{a} \in U}[e(l(\mathbf{x}+\mathbf{a}) - p(\mathbf{x}+\mathbf{a}))] = e(p(\mathbf{x}))$ □

Effectively, we have shown that $p(\mathbf{x})$ can be approximated uniformly by a (random) degree $d-1$ polynomial $q_{\mathbf{a}}(\mathbf{x})$. We can now use this to show that a distribution that is pseudorandom against degree $d-1$ polynomials, is also pseudorandom against p. First, we prove the following lemma:

Lemma 13. *Let D be a distribution that is pseudorandom against degree $d-1$ polynomials with error δ. For every $a \in F^n$:*

$$|\mathbb{E}_{\boldsymbol{x} \in D}[\nu_{\boldsymbol{x}}(\boldsymbol{a})] - \mathbb{E}_{\boldsymbol{x} \in U}[\nu_{\boldsymbol{x}}(\boldsymbol{a})]| < \frac{\delta}{|\eta|}$$

Proof. $|\mathbb{E}_{\mathbf{x} \in D}[\nu_{\mathbf{x}}(\mathbf{a})] - \mathbb{E}_{\mathbf{x} \in U}[\nu_{\mathbf{x}}(\mathbf{a})]| = \frac{1}{|\eta|} |\mathbb{E}_{\mathbf{x} \in D}[e(q_{\mathbf{a}}(\mathbf{x}))] - \mathbb{E}_{\mathbf{x} \in U}[e(q_{\mathbf{a}}(\mathbf{x}))]| < \frac{\delta}{|\eta|}$, where we use the fact that $q_{\mathbf{a}}$ is a polynomial of degree at most $d-1$, and so D is pseudorandom against $q_{\mathbf{a}}$ with error δ. □

We now conclude by proving Lemma 4

Proof. (of Lemma 4) Let D be a distribution that is pseudorandom against degree $d-1$ polynomials with error $\epsilon^3/10$. Then:

$$\begin{aligned}&|\mathbb{E}_{\mathbf{x}\in D}[e(p(\mathbf{x}))] - \mathbb{E}_{\mathbf{x}\in U}[e(p(\mathbf{x}))]| = \\ &|\mathbb{E}_{\mathbf{x}\in D}\mathbb{E}_{\mathbf{a}\in U}[\nu_{\mathbf{x}}(\mathbf{a})] - \mathbb{E}_{\mathbf{x}\in U}\mathbb{E}_{\mathbf{a}\in U}[\nu_{\mathbf{x}}(\mathbf{a})]| \leq \\ &\mathbb{E}_{a\in U}|\mathbb{E}_{\mathbf{x}\in D}[\nu_{\mathbf{x}}(\mathbf{a})] - \mathbb{E}_{\mathbf{x}\in U}[\nu_{\mathbf{x}}(\mathbf{a})]| < \frac{\epsilon^3/10}{|\eta|} \leq \epsilon\end{aligned}$$

□

Acknowledgement. I thank my supervisor, Omer Reingold, for useful comments and for his constant support and interest in the work. I thank Andrej Bogdanov and Emanuele Viola for making an early version of their paper available and for insightful comments. In particular I thank Bogdanov for an observation that somewhat simplifies the analysis in the case where all the Fourier coefficients are small, which allowed to reduce the number of small-bias terms from 3^d to 2^d. I thank Zeev Dvir, Dana Moshkovitz and Ariel Gabizon for helpful comments.

References

[AGHP92] N. Alon, O. Goldreich, J. Hastad, and R. Peralta. Simple constructions of almost k-wise independent random variables. *Random Structures & Algorithms*, 3(3):289-304, 1992

[BLR93] M. Blum, M. Luby and R. Rubinfeld. Self-Testing/Correcting with Applications to Numerical Problems. *Journal of Computer and System Sciences*, 47(3):549–595, 1993

[Bog05] A. Bogdanov, Pseudorandom generators for low degree polynomials. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing (STOC 2005)*, pages 21-30, New York, 2005. ACM.

[BSVW03] E. Ben-Sasson, M. Sudan, S. Vadhan, and A. Wigderson. Randomness-efficient low degree tests and short PCPs via epsilon-biased sets. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing (STOC 2003)*, pages 612-621 (electronic),2003.

[BV07] A. Bogdanov and E. Viola. Pseudorandom bits for polynomials via the Gowers norm, accepted to *the 48th Annual Symposium on Foundations of Computer Science (FOCS 2007).*

[GT05] B. Green and T. Tao. An inverse theorem for the Gowers U3 norm, in *Proceedings of the Edinburgh Mathematical Society, to appear.*

[LVW93] 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.*

[NN90] 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 (STOC 1990)*, pages 213-223, 1990.*

[Sam07] A. Samorodnitsky, Low-degree tests at large distances, In Proceedings of the 39th Annual ACM Symposium on Theory of computing (STOC 2007)*, pages 506–515, 2007.*

[ST06] A. Samorodnitsky and L. Trevisan. Gowers uniformity, influence of variables, and PCPs. In Proceedings of the 38th Annual ACM Symposium on Theory of Computation (STOC 2006)*, pages 11–20, 2006.*

[Vio06] E. Viola. New correlation bounds for GF(2) polynomials using Gowers uniformity. Electronic Colloquium on Computational Complexity*, Technical Report TR06–097, 2006. http://www.eccc.uni-trier.de/eccc.*

[Vio07] E. Viola. Pseudorandom Bits for Constant-Depth Circuits with Few Arbitrary Symmetric Gates. SIAM Journal on Computing*, 36(5):1387-1403, 2007.*

[VW07] 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.*

ECCC ISSN 1433-8092

http://eccc.hpi-web.de/