

Fooling Low-Degree Polynomials: PRGs with Optimal Dependence on Input Length and Expander Walks

Gil Cohen*

October 9, 2026

Abstract

In a recent breakthrough, Chattopadhyay, Hatami, Lee, Lovett, Tal, and Viola [CHL⁺26] established exponential correlation bounds for polynomials over $\mathbb{F}_2$ and, as a consequence, obtained a major improvement in PRG constructions for low-degree polynomials over the binary field. In particular, they obtained seed length $\tilde{O}(d^2 \log^2 n)$ for fooling degree- d polynomials in n variables with constant error.

Our main contribution is to achieve the optimal dependence on the number of variables n , at the expense of a polynomial deterioration in the dependence on d . Specifically, we obtain seed length $O(d^4 \log n)$, as well as a bound of $\tilde{O}(d^3 \log n)$. We also obtain nearly linear dependence on $\log(1/\varepsilon)$, improving the quadratic dependence in the Fourier-based PRG of [CHL⁺26].¹

Finally, we show that a random walk on a λ -spectral expander fools degree- d polynomials over $\mathbb{F}_2$ with error $O(d^2 \lambda)$, provided that the labeling is balanced. The bound is independent of the walk length n . To prove this, we extend the Fourier-analytic approach of Cohen, Peri, and Ta-Shma [CPTS21], replacing level-wise bounds on the Fourier L_1 mass with bounds that account for cancellations among coefficients within each level.

*Tel Aviv University. gil@tauex.tau.ac.il. Supported by ERC starting grant 949499 and by the Israel Science Foundation grant 2989/24.

¹The improvement in the dependence on the error parameter was obtained independently by the authors of [CHL⁺26] and will appear in a revised version of their manuscript.

1 Introduction

In a very recent breakthrough, Chattopadhyay, Hatami, Lee, Lovett, Tal, and Viola [CHL⁺26] established exponential correlation bounds for polynomials over $\mathbb{F}_2$ and, as a consequence, obtained a major improvement in PRG constructions for low-degree polynomials over the binary field. In particular, they obtained seed length $\tilde{O}(d^2 \log^2 n)$ for fooling degree- d polynomials in n variables.² This substantially improved upon the celebrated, more than fifteen-year-old result of Viola [Vio09], which gives seed length $O(d \log n + d2^d)$. The improvement already applies to slowly growing degrees, $d \gg \log \log n$. Moreover, their construction is the first to achieve nontrivial seed length for degrees beyond $\log n$.

Viola’s construction is very elegant: it takes the bitwise XOR of d independent samples from a small-bias distribution, that is, a PRG for degree-1 polynomials, with sufficiently small bias. The construction of [CHL⁺26] uses more sophisticated tools. Specifically, it applies the result of Chattopadhyay, Gaitonde, Lee, Lovett, and Shetty [CGL⁺21], which builds on the framework of polarizing random walks introduced by Chattopadhyay, Hatami, Hosseini, and Lovett [CHHL19].

Historical context. The problem of constructing PRGs for low-degree polynomials has a long history. We will not attempt to cover it in full, but briefly recall some of the developments. Broadly speaking, research has followed two lines: constructions over fields of constant size, such as $\mathbb{F}_2$, and constructions over fields whose size grows with the degree d , and sometimes also with the number of variables n . The main reason for that is that over sufficiently large fields, one can employ deep algebraic results such as the Weil bound. For simplicity, all bounds in this discussion are stated for constant error.

In the large-field regime, Bogdanov [Bog05] introduced a technique for constructing PRGs from hitting set generators (HSGs). Subsequent work on HSGs includes constructions by Lu [Lu12], Cohen and Ta-Shma [CT13], and Guruswami and Xing [GX14]. Combined with the latter construction, Bogdanov’s technique yields a PRG with seed length $O(d^4 \log n + \log q)$ over fields of size $q = \Omega(d^6)$.

More recently, Derksen and Viola [DV22] introduced a new approach based on algebraic geometry and invariant theory, with the preservation of indecomposability playing a central role. Their construction achieves near-optimal seed length $\tilde{O}(d \log n) + O(\log q)$ for $q = \Omega((d \log n)^4)$. Another choice of parameters gives optimal seed length $O(d \log n + \log q)$ albeit for $q = \Omega(d^4 n^{0.001})$. Building on Bogdanov’s framework and the insights of Derksen and Viola, Dwivedi, Guo, and Volk [DGV24] removed the dependence on n from the

²The tilde hides polylogarithmic factors in d and $\log n$. For simplicity, unless stated otherwise, we take the error parameter of the PRG to be constant in this introduction.

field-size requirement for an optimal-seed PRG, albeit with an exponential dependence on d . Specifically, they achieve seed length $O(d \log n + \log q)$ for $q = \Omega(d2^d)$, assuming $\text{char}(\mathbb{F}_q) > d(d-1)$. Cohen, Doron, and Goldgraber [CDG26] subsequently reduced the field-size requirement to $q = \Omega((d \log d)^4)$, retaining the optimal seed length and the characteristic assumption.

For arbitrary fields, a different line of work [LVW93, Vio07, BV10, Lov09] led to Viola’s celebrated construction [Vio09], whose seed length is exponential in d . Although the two regimes have often been studied using different techniques, [CDG26] establishes a connection between them. For example, a PRG with seed length $d^{O(1)} \log n$ that works for $q = d^{0.99}$, without any restriction on the characteristic, would yield a PRG over $\mathbb{F}_2$ with seed length $d^{O(1)} \log n$ via a simple reduction using the field trace.

1.1 Our results

The main purpose of this paper is to improve the dependence on n in the seed length of [CHL⁺26], attaining the optimal logarithmic dependence at the expense of a mild polynomial deterioration in the dependence on d . We prove the following.

Theorem 1.1. *For all integers $n \geq 2$ and $1 \leq d \leq n$, there exists an explicit PRG with constant error for the class of n -variate polynomials over $\mathbb{F}_2$ of degree at most d , with seed length*

$$O(d^4 \log n).$$

Theorem 1.1 achieves the optimal dependence on n , at the expense of an additional factor of d^2 in the polynomial dependence on the degree compared with [CHL⁺26]. Their bound, which recall is $\tilde{O}(d^2 \log^2 n)$, has quadratic rather than linear dependence on d , whereas the probabilistic construction gives seed length $O(d \log n)$.

More precisely, we construct a PRG with seed length

$$O\left(d^2 \log n (d + \log \log \log n) (\log d + \log \log n)\right).$$

We obtain **Theorem 1.1** by choosing between this construction and Viola’s PRG, depending on the relation between n and d . When $d \geq \log \log n$, the bound is $O(d^4 \log n)$; otherwise, Viola’s construction gives seed length $O(d \log n)$.

Our proof follows the same framework as [CHL⁺26] and, in particular, builds on their level-wise Fourier bounds for the sign functions associated with low-degree polynomials over $\mathbb{F}_2$ (see **Corollary 2.1**). The only change is in the construction of the fractional PRG (see **Definition 2.2**), where we use a more randomness-efficient approximately k -wise independent distribution.

Improved dependence on the error parameter. We further improve the dependence on the error parameter in the Fourier-based PRG of [CHL⁺26], from quadratic to nearly linear in $\log(1/\varepsilon)$. More precisely, we prove the following.³

Theorem 1.2. *For all integers $n \geq 2$ and $1 \leq d \leq n$, and every $\varepsilon > 0$, there exists a PRG with error ε for the class of n -variate degree-at-most- d polynomials over $\mathbb{F}_2$ with seed length*

$$O\left(d^2 \log n (d + \log L)L\right), \quad L = \log\left(\frac{d^2 \log n}{\varepsilon}\right).$$

The generator is computable in time $\text{poly}(n, d, \log(1/\varepsilon))$.

In fact, already the original version of their paper [CHL⁺26, Theorem A.1] constructs a PRG (based on the Nisan–Wigderson construction) with seed length

$$O\left(d^4 \log^3 n \cdot \log(n/\varepsilon)\right).$$

This construction already achieves linear dependence on $\log(1/\varepsilon)$, though it has worse dependence on n and d . Our refinement combines nearly linear dependence on $\log(1/\varepsilon)$ with the improved dependence on the number of variables.

The seed-length bound itself follows readily by applying the refined reduction of Hatami and Hoza [HH24, Theorem 5.10] to our fractional PRG, in place of the reduction from [CGL⁺21] used above. A direct application of their theorem, however, only guarantees running time $\text{poly}(n, d, 1/\varepsilon)$. To obtain the stated running time, we replace their exhaustive sampling procedure with a dynamic program. For ease of exposition, we first present the simpler construction yielding **Theorem 1.1** in **Section 3** and only then give this refinement in **Section 4**.

Expander random walks fool low-degree polynomials. A PRG constructs a particular distribution that fools a given class of functions. A related question is whether every distribution satisfying some natural structural condition fools that class. A prominent example is Braverman’s theorem that polylogarithmic independence fools AC^0 circuits [Bra10]. Such results deepen our understanding of pseudorandomness and have important applications, particularly when a distribution arises within a construction and we know only that it satisfies certain properties. Indeed, Braverman’s theorem is a key ingredient in the two-source extractor construction of Chattopadhyay and Zuckerman [CZ19].

³This improvement was obtained independently by the authors of [CHL⁺26] and will appear in a revised version of their manuscript (Avishay Tal, personal communication).

Our second main result concerns the fooling power of expander walks. Consider an undirected regular graph G on N vertices, equipped with a balanced labeling by elements of $\mathbb{F}_2$.⁴ Start a random walk at a uniformly random vertex and record the labels of the first n vertices it visits. We show that low-degree polynomials over $\mathbb{F}_2$ cannot distinguish these labels from independent uniform bits, provided that G has sufficiently strong spectral expansion. The broader question of which test functions are fooled by expander walks was explicitly studied in [CPTS21], which considered several classes, including all symmetric functions and functions computed by AC^0 circuits. See also [GK21] for the related sticky random walk model.

Subsequent work extended this study in several directions. Cohen, Minzer, Peleg, Potechin, and Ta-Shma [CMP⁺22] generalized the results for symmetric functions to unbalanced labelings, obtained matching upper and lower bounds, and established limitations on fooling AC^0 circuits. Golowich and Vadhan [GV22] further extended the results for symmetric functions to arbitrary finite alphabets, obtained optimal dependence on the labeling bias in the binary case, and proved bounds for permutation branching programs.

Here we extend this line of results to low-degree polynomials over $\mathbb{F}_2$. In particular, we prove the following.

Theorem 1.3 (Informal). *Let G be a λ -spectral expander on N vertices, equipped with a balanced binary vertex labeling. The labels of the first n vertices visited by a random walk fool all n -variate polynomials over $\mathbb{F}_2$ of degree at most d with error $O(d^2\lambda)$.*

For the formal statement, see [Theorem 5.1](#). The proof combines the Fourier bounds of [CHL⁺26] with the framework of Cohen, Peri, and Ta-Shma [CPTS21]. To make this combination work, we prove a general criterion in terms of the signed level norms, extending the Fourier- L_1 -based criterion used in [CPTS21] for general test functions.

2 Preliminaries

Throughout the paper, logarithms are to base 2 and are truncated below at 1.

2.1 Fourier analysis and low-degree polynomials

Let $P : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$ be any function. We will mainly be interested in functions whose degree over $\mathbb{F}_2$ is at most d , but the following discussion is general. Our construction and its

⁴This assumption is made mainly for simplicity. We expect our results to extend to unbalanced labelings, with independent bits of the corresponding bias as the reference distribution.

analysis use Fourier analysis over the reals, so it is convenient to switch from $\mathbb{F}_2$ to $\{-1, 1\}$ notation for both the inputs and the output. To this end, define

$$f(x) = (-1)^{P((1-x_1)/2, \dots, (1-x_n)/2)}, \quad x \in \{-1, 1\}^n.$$

The function f depends on P , of course, but we omit this dependence from the notation. We identify f with its unique real multilinear extension:

$$f(x) = \sum_{S \subseteq [n]} \hat{f}(S) x_S, \quad x_S = \prod_{i \in S} x_i.$$

The coefficients are given by

$$\hat{f}(S) = \mathbb{E}_x[f(x) x_S],$$

where x is uniform over $\{-1, 1\}^n$. In particular, $f(0) = \hat{f}(\emptyset) = \mathbb{E}_x[f(x)]$.

For $j \in \{0, 1, \dots, n\}$, denote the j -th homogeneous part of f by

$$f_j(x) = \sum_{|S|=j} \hat{f}(S) x_S,$$

and define

$$M_j(f) \triangleq \max_{x \in \{-1, 1\}^n} |f_j(x)|.$$

We also write $f_{<k} = \sum_{j < k} f_j$ and $f_{\geq k} = \sum_{j \geq k} f_j$, with both sums restricted to $j \in \{0, \dots, n\}$.

We will use the following recently established bound.

Corollary 2.1 (Theorem B.1 of [CHL⁺26]). *Let $P : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$ have degree at most d , where $d \geq 1$, and let f be the corresponding sign function, identified with its multilinear extension. Then, for every $1 \leq j \leq n$,*

$$M_j(f) \leq (2ed)^j.$$

We will also use the older bound

$$L_{1,j}(f) \triangleq \sum_{|S|=j} |\hat{f}(S)| \leq (j \cdot 2^{3d})^j, \quad 1 \leq j \leq n, \quad (2.1)$$

from Theorem 6.1 of [CHHL19], also recorded as Proposition 29 of [CGL⁺21]. These are very different quantities: $M_j(f)$ allows cancellations among Fourier coefficients, whereas $L_{1,j}(f)$ does not.

2.2 PRGs and fractional PRGs

Chattopadhyay, Hatami, Hosseini, and Lovett [CHHL19] introduced a paradigm for constructing PRGs by composing simpler primitives called *fractional PRGs*.

Definition 2.2. *Let $\mathcal{F}$ be a class of functions from $\{-1, 1\}^n$ to $\{-1, 1\}$. A fractional pseudorandom generator (fractional PRG) for $\mathcal{F}$ with error ε is a random variable $\mathbf{X}$ supported on $[-1, 1]^n$ such that, for every $f \in \mathcal{F}$, identified with its multilinear extension,*

$$|\mathbb{E}[f(\mathbf{X})] - f(0)| \leq \varepsilon.$$

Following Definition 13 of [CGL⁺21], we define a fractional PRG as a random variable. Equivalently, one may describe a fractional PRG with seed length r by a function $G : \{0, 1\}^r \rightarrow [-1, 1]^n$, with $\mathbf{X} = G(\mathbf{U}_r)$, where $\mathbf{U}_r$ is uniform over $\{0, 1\}^r$. This formulation makes it convenient to discuss explicitness, meaning that G is efficiently computable, and seed length. An ordinary PRG is the special case in which the range is $\{-1, 1\}^n$ rather than the solid cube $[-1, 1]^n$.

Several reductions from constructing PRGs to constructing fractional PRGs are known. Their cost, particularly in seed length, is governed by how much second moment the fractional PRG has in each coordinate. Formally, a fractional PRG is p -noticeable if $\mathbb{E}[\mathbf{X}_i^2] \geq p$ for every $i \in [n]$. We use the following reduction.

Theorem 2.3 (Theorem 14 of [CGL⁺21]). *Let $\mathcal{F}$ be a class of n -variate Boolean functions that is closed under restrictions, and let $\mathbf{X}$ be an explicit p -noticeable fractional PRG for $\mathcal{F}$ with error ε and seed length s . Then there exists an explicit PRG for $\mathcal{F}$ with seed length $O(s \log(n/\varepsilon)/p)$ and error $O(\varepsilon \log(n/\varepsilon)/p)$.*

3 Logarithmic dependence on input length

In this section, we prove Theorem 1.1. We begin by constructing the key ingredient: a fractional PRG.

Proposition 3.1. *For all integers $n \geq 2$ and $1 \leq d \leq n$, and every $\varepsilon > 0$, there exists an explicit p -noticeable fractional PRG with error ε for the sign functions of n -variate polynomials over $\mathbb{F}_2$ of degree at most d , where $p = \Omega(d^{-2})$, with seed length*

$$O\left(\log \log n + \left(d + \log \log \frac{1}{\varepsilon}\right) \log \frac{1}{\varepsilon}\right).$$

Proof. Let $k = \lceil 2 + \log(1/\varepsilon) \rceil$, and let $\mathbf{Y} \in \{-1, 1\}^n$ be a distribution with bias at most η for every nonempty parity of fewer than k coordinates. That is,

$$|\mathbb{E}[\mathbf{Y}_S]| \leq \eta \quad \text{whenever } 1 \leq |S| < k.$$

We will set η shortly. Standard constructions [NN93, AGHP92] allow such a distribution to be sampled efficiently with seed length

$$O(\log \log n + \log k + \log(1/\eta)). \quad (3.1)$$

Define our fractional PRG by

$$\mathbf{X} = \rho \mathbf{Y}, \quad \rho = \frac{1}{16d}.$$

In particular, $2ed\rho = e/8 < 1/2$.

We analyze the error by considering the high Fourier levels first, followed by the nonconstant low levels. By [Corollary 2.1](#), for every $y \in \{-1, 1\}^n$,

$$|f_{\geq k}(\rho y)| \leq \sum_{j=k}^n \rho^j M_j(f) \leq \sum_{j \geq k} (2ed\rho)^j \leq \sum_{j \geq k} 2^{-j} = 2^{1-k} \leq \frac{\varepsilon}{2},$$

where the last inequality follows from our choice of k .

For the low Fourier levels, [Equation \(2.1\)](#) and the guarantee on $\mathbf{Y}$ give

$$|\mathbb{E}[f_{<k}(\rho \mathbf{Y})] - f(0)| \leq \eta \sum_{j=1}^{k-1} \rho^j L_{1,j}(f) \leq \eta k (k2^{3d})^k \leq \eta (k2^{3d})^{2k},$$

Set

$$\eta = \frac{\varepsilon}{2(k2^{3d})^{2k}}.$$

The contribution of the low levels is then at most $\varepsilon/2$. Together with the bound on the high levels, this proves the required error bound.

As for the seed length,

$$\begin{aligned} \log \frac{1}{\eta} &= \log \frac{1}{\varepsilon} + O(k(d + \log k)) \\ &= O\left(\left(d + \log \log \frac{1}{\varepsilon}\right) \log \frac{1}{\varepsilon}\right). \end{aligned}$$

Substituting this into [Equation \(3.1\)](#) gives the claimed seed length. Finally, every coordinate satisfies $\mathbf{X}_i^2 = \rho^2 = \Theta(d^{-2})$, so $\mathbf{X}$ has the required noticeability. $\square$

Corollary 3.2. *For all integers $n \geq 2$ and $1 \leq d \leq n$, and every $\varepsilon > 0$, there exists an explicit PRG with error ε for the class of n -variate polynomials over $\mathbb{F}_2$ of degree at most d , with seed length*

$$O\left(d^2 \log\left(\frac{n}{\varepsilon}\right) \left(d + \log \log \frac{1}{\varepsilon} + \log \log \log n\right) \left(\log d + \log \frac{1}{\varepsilon} + \log \log n\right)\right).$$

Proof. We apply [Theorem 2.3](#) to the fractional PRG from [Proposition 3.1](#). To help the reader with the somewhat daunting parameter substitution, let δ denote the error of the fractional PRG. Its seed length is

$$s = O\left(\log \log n + \left(d + \log \log \frac{1}{\delta}\right) \log \frac{1}{\delta}\right), \quad (3.2)$$

and its noticeability is $p = \Theta(d^{-2})$. By [Theorem 2.3](#), the resulting PRG has error at most

$$\varepsilon = O\left(d^2 \delta \log \frac{n}{\delta}\right)$$

and seed length

$$O\left(d^2 s \log \frac{n}{\delta}\right).$$

For $n \geq 2$ and $0 < \delta < 1$, we have the cruder but convenient estimate

$$\delta \log \frac{n}{\delta} = O(\sqrt{\delta} \log n).$$

Thus,

$$\log \frac{1}{\delta} = O\left(\log \frac{1}{\varepsilon} + \log d + \log \log n\right).$$

Since $d \leq n$, we also have

$$\log \frac{n}{\delta} = O\left(\log \frac{n}{\varepsilon}\right).$$

Moreover,

$$d + \log \log \frac{1}{\delta} = O\left(d + \log \log \frac{1}{\varepsilon} + \log \log \log n\right),$$

where the possible $\log \log d$ term is absorbed into d . Substituting these estimates into [Equation \(3.2\)](#) and the seed-length bound from the reduction yields the claimed expression.

For completeness, we remark that the hypotheses of [Theorem 2.3](#) hold here exactly as in previous applications to low-degree polynomials. The class is closed under Boolean restrictions, since fixing variables does not increase the degree over $\mathbb{F}_2$. Fractional recentering need not preserve the class itself, but it preserves its convex hull: for $a \in [-1, 1]^n$, the function

$$x \mapsto f\left(a + (1 - |a|) \odot x\right)$$

is an average of Boolean restrictions of f . Indeed, independently for each coordinate i , fix that coordinate to $\text{sgn}(a_i)$ with probability $|a_i|$, and leave it free otherwise. Multilinearity gives the above function upon averaging. Fractional fooling extends to these averages by linearity. This is the recentering argument of Claim 3.3 in [\[CHHL19\]](#) and Lemma 20 in [\[CGL⁺21\]](#). $\square$

We are now in a position to deduce [Theorem 1.1](#).

Proof of [Theorem 1.1](#). At constant error, [Corollary 3.2](#) gives seed length

$$O\left(d^2 \log n (d + \log \log \log n)(\log d + \log \log n)\right).$$

We obtain a bound of $O(d^4 \log n)$ by choosing between this generator and Viola's generator [[Vio09](#)], whose seed length at constant error is $O(d \log n + d2^d)$.

If $d < \log \log n$, then $2^d < \log n$, and Viola's generator has seed length $O(d \log n)$. If $d \geq \log \log n$, then both $d + \log \log \log n$ and $\log d + \log \log n$ are $O(d)$. Our generator therefore has seed length $O(d^4 \log n)$, as required. $\square$

4 Improved dependence on the error

The construction from [Section 3](#) has a nearly quadratic dependence on $\log(1/\varepsilon)$. In this section we improve this dependence using the refined reduction from fractional PRGs to PRGs given by Hatami and Hoza [[HH24](#), Theorem 5.10]. Their reduction shortens the polarized random walk by replacing its final rounding step with a more careful sampling procedure. As stated, the reduction has running time polynomial in $1/\varepsilon$. We show that its sampling procedure can be implemented by dynamic programming, establishing [Theorem 1.2](#).

Efficient sampling from a polarized product distribution. The final step of the Hatami–Hoza reduction approximately samples a product distribution that places almost all of its mass in a small Hamming ball. Their sampling procedure enumerates this ball [[HH24](#), Lemma 5.12]. The following lemma achieves the same randomness bound without this enumeration.

Lemma 4.1. *Let $a_1, \dots, a_n \in [0, 1/2]$ be rational numbers whose numerators and denominators have at most B bits. Let $\mathbf{E} \in \{0, 1\}^n$ have independent coordinates with $\Pr[\mathbf{E}_i = 1] = a_i$. Suppose that $0 \leq K \leq n$ and $0 < \gamma \leq 1/2$ satisfy*

$$\Pr[|\mathbf{E}| > K] \leq \gamma,$$

where $|\mathbf{E}|$ denotes Hamming weight. There is a sampler whose output distribution is within total variation distance 2γ of $\mathbf{E}$, using

$$O\left(K \log n + \log(1/\gamma)\right)$$

uniform random bits and running in time $\text{poly}(n, B, \log(1/\gamma))$.

Proof. Write

$$\mathcal{B}_K = \{e \in \{0, 1\}^n : |e| \leq K\}, \quad N_K = |\mathcal{B}_K|,$$

and let Q be the distribution of $\mathbf{E}$ conditioned on $\mathbf{E} \in \mathcal{B}_K$. This conditioning changes the distribution by total variation distance at most γ . We therefore only need to sample Q to within distance γ .

For $i \in \{1, \dots, n+1\}$ and $0 \leq j \leq K$, define

$$F_i(j) = \Pr \left[\sum_{\ell=i}^n \mathbf{E}_\ell \leq j \right].$$

These quantities satisfy

$$F_i(j) = (1 - a_i)F_{i+1}(j) + a_i F_{i+1}(j-1),$$

with $F_{n+1}(j) = 1$ for $j \geq 0$ and $F_i(j) = 0$ for $j < 0$. Hence, all of them can be computed using $O(nK)$ arithmetic operations. Moreover, for every prefix $u \in \{0, 1\}^t$, its probability under Q is

$$Q[\text{prefix } u] = \frac{\left(\prod_{i=1}^t a_i^{u_i} (1 - a_i)^{1-u_i} \right) F_{t+1}(K - |u|)}{F_1(K)}.$$

The denominator is positive since $\Pr[\mathbf{E} = 0^n] \geq 2^{-n}$.

Order the strings in $\mathcal{B}_K$ lexicographically and partition $[0, 1)$ into consecutive half-open intervals of lengths $Q(e)$, one for each $e \in \mathcal{B}_K$.

Given a point $r \in [0, 1)$, we locate its interval by determining the corresponding string one coordinate at a time. Initially, the prefix is empty and its interval is $[0, 1)$. If the current prefix is u and its interval starts at a , we split this interval at $a + Q[\text{prefix } u0]$. We append 0 if r lies below this splitting point, and 1 otherwise, updating the interval accordingly. The prefix probabilities computed above allow us to perform each step efficiently, without enumerating the strings in $\mathcal{B}_K$.

Set

$$m = \left\lceil \log \frac{N_K}{\gamma} \right\rceil.$$

Choose R uniformly from $\{0, \dots, 2^m - 1\}$ and output the string whose interval contains $R/2^m$. For each interval, the fraction of these grid points that it contains differs from its length by at most 2^{-m} . Thus, the output distribution is within total variation distance at most

$$\frac{N_K}{2} 2^{-m} \leq \gamma$$

of Q . Together with the cost of conditioning, this gives the claimed error. Since $\log N_K = O(K \log n)$, the randomness bound follows.

Finally, all computations can be performed exactly with rational arithmetic. A common denominator for the product probabilities is the product of the denominators of the a_i , which has $O(nB)$ bits. The dynamic program and the prefix probabilities therefore have polynomial bit length, as do the comparisons with $R/2^m$. Computing $N_K = \sum_{j=0}^K \binom{n}{j}$ also takes polynomial time. This proves the running-time bound. $\square$

Applying the refined reduction. We now explain how to use this sampler in the Hatami–Hoza construction. Their refined reduction transforms a p -noticeable fractional PRG with error δ and seed length s into a PRG with seed length

$$O\left((s + \log(1/\delta))p^{-1} \log n\right) \quad (4.1)$$

and error $O(\delta p^{-1} \log n)$, for classes closed under Boolean restrictions and input shifts. Here an input shift in sign notation means replacing $f(x)$ by $f(x \odot z)$ for a fixed $z \in \{-1, 1\}^n$. As mentioned, degree-at-most- d polynomials over $\mathbb{F}_2$ have both closure properties.

Proof of Theorem 1.2. Apply Proposition 3.1 with fractional error

$$\delta = \frac{c\varepsilon}{d^2 \log n},$$

where $c > 0$ is a sufficiently small absolute constant. This gives noticeability $p = \Theta(d^{-2})$ and seed length

$$s = O\left(\log \log n + \left(d + \log \log \frac{1}{\delta}\right) \log \frac{1}{\delta}\right).$$

We have $\log(1/\delta) = \Theta(L)$, and hence

$$s + \log(1/\delta) = O\left((d + \log L)L\right).$$

Substituting into Equation (4.1) gives the stated seed length, and the final error is at most ε after choosing c sufficiently small.

It remains to justify the running time. We recall the part of the refined reduction relevant to this issue. Set

$$\ell = \left\lceil \frac{\log(1/\delta)}{\log n} \right\rceil.$$

The fractional output is multiplied coordinatewise by an independent ℓ -wise uniform sign vector. In our case every coordinate of the fractional output has the same fixed magnitude, so this makes its coordinates ℓ -wise independent while preserving both noticeability and fractional fooling. The latter follows from closure under input shifts. This modification costs $O(\log(1/\delta))$ additional seed bits: for $\ell = 1$ a single uniform sign suffices, for $\ell \geq 2$ use the standard $O(\ell \log n)$ construction, and for $\ell > n$ use a uniform sign vector. Global sign symmetry, if needed, costs one additional bit.

After $t = O(p^{-1} \log n)$ steps, the refined polarization analysis [HH24, Lemma 5.11] gives the following guarantee. Except with probability δ , the resulting point $y \in [-1, 1]^n$ satisfies

$$\Pr[\text{dist}_H(\mathbf{W}, \text{sign}(y)) > K] \leq \delta, \quad K = \min\{n, 2\ell\},$$

where $\mathbf{W}$ has independent sign coordinates of means y_i . We take $\text{sign}(0) = 1$.

For a fixed such y , put $a_i = (1 - |y_i|)/2$. Then $\mathbf{W}_i = \text{sign}(y_i)(-1)^{\mathbf{E}_i}$ for independent Bernoulli variables $\mathbf{E}_i$ of means a_i . Apply Lemma 4.1 with $\gamma = \delta$. This replaces the exhaustive sampler in [HH24, Lemma 5.12] with a polynomial-time sampler, preserving its $O(\delta)$ error. The exceptional choices of y contribute only another $O(\delta)$ to the error, so the preceding analysis still applies. The sampler uses

$$O(K \log n + \log(1/\delta)) = O(\log n + \log(1/\delta))$$

random bits, which are absorbed by Equation (4.1).

To account for bit complexity, use the rational scaling $\rho = 1/(16d)$ in Proposition 3.1; its proof and parameters remain valid with this choice. Every step then has coordinates $\pm\rho$. The walk starts at $y^{(0)} = 0$ and updates by

$$y^{(j+1)} = y^{(j)} + (1 - |y^{(j)}|) \odot \mathbf{X}^{(j+1)},$$

where the absolute value is taken coordinatewise. After t steps, each coordinate has denominator dividing $(16d)^t$ and therefore has $O(t \log d)$ bits. These bit lengths are polynomial in n and d . The fractional generator, the limited-independence distribution, the walk, and the final sampler are consequently all computable in time $\text{poly}(n, d, \log(1/\varepsilon))$. $\square$

5 Expander walks fool low-degree polynomials

Let G be a regular undirected graph on N vertices, and let M be its transition matrix. Write u for the normalized all-ones vector, whose entries are all $N^{-1/2}$, and let $J = uu^\top$. Recall that G is a λ -spectral expander if $\|M - J\|_2 \leq \lambda$.

Let $b : V(G) \rightarrow \{-1, 1\}$ be a balanced labeling, meaning that

$$\mathbb{E}_{v \sim V(G)}[b(v)] = 0.$$

Start a random walk $\mathbf{V}_1, \dots, \mathbf{V}_n$ at a uniformly random vertex, and set

$$\mathbf{X} = (b(\mathbf{V}_1), \dots, b(\mathbf{V}_n)).$$

Theorem 5.1. *Let $d \geq 1$ and $0 \leq \lambda \leq 1/(128e^2d^2)$. For every polynomial $P : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$ of degree at most d , let*

$$f(x) = (-1)^{P((1-x_1)/2, \dots, (1-x_n)/2)}, \quad x \in \{-1, 1\}^n,$$

be its associated sign function, identified with its real multilinear extension. Then

$$|\mathbb{E}[f(\mathbf{X})] - \mathbb{E}[f(\mathbf{U}_n)]| \leq 64e^2d^2\lambda.$$

Equivalently, if $\mathbf{Y}_i = (1 - \mathbf{X}_i)/2$ and $\mathbf{W}$ is uniform over $\mathbb{F}_2^n$, then

$$|\Pr[P(\mathbf{Y}) = 1] - \Pr[P(\mathbf{W}) = 1]| \leq 32e^2d^2\lambda.$$

In particular, the bound is independent of the walk length n .

The main step is to represent the walk's label sequence $\mathbf{X}$ as another distribution subjected to independent noise. This allows us to use bounds on signed Fourier-level sums, preserving cancellations within each level.

Lemma 5.2 (Noise representation). *Suppose $0 < \lambda \leq 1/8$, and set $\rho = \sqrt{8\lambda}$. There exists a random vector $\mathbf{Z} \in \{-1, 1\}^n$ such that*

$$\mathbf{X} \stackrel{\text{law}}{=} \mathbf{Z} \odot \xi,$$

where $\odot$ denotes coordinatewise multiplication, and $\xi_1, \dots, \xi_n$ are independent signs of mean ρ , independent of $\mathbf{Z}$. Moreover, $\mathbb{E}[\mathbf{Z}_i] = 0$ for every $i \in [n]$.

Proof. To define $\mathbf{Z}$ consider the function $q : \{-1, 1\}^n \rightarrow \mathbb{R}$ which is defined by

$$q(z) = 2^{-n} \mathbb{E}_{\mathbf{X}} \left[\prod_{i=1}^n \left(1 + \frac{z_i \mathbf{X}_i}{\rho} \right) \right].$$

Note that summing over z gives $\sum_z q(z) = 1$. We first show that $q(z) \geq 0$ for every z .

Let $B = \text{diag}(b)$, so that $\|B\|_2 = 1$ and $\langle u, Bu \rangle = 0$. Fix z , and define

$$D_i = I + \frac{z_i B}{\rho}, \quad v_0 = u, \quad v_i = MD_i v_{i-1}.$$

The matrix expression for the walk expectation gives

$$2^n q(z) = u^\top D_1 MD_2 M \cdots MD_n u.$$

Since M and the D_i are symmetric, transposing this scalar expression reverses the order of the factors. Using also $u^\top M = u^\top$, we obtain

$$2^n q(z) = u^\top D_n M \cdots MD_1 u = \langle u, v_n \rangle.$$

For $n = 1$, the product is simply $u^\top D_1 u$.

Write $v_i = a_i u + w_i$, where $w_i \perp u$. We claim that, for every i ,

$$a_i \geq 0, \quad \|w_i\|_2 \leq \frac{\rho}{2} a_i. \quad (5.1)$$

This holds for $i = 0$. Suppose it holds for $v = au + w$, and write the next iterate as

$$v' = M \left(I + \frac{sB}{\rho} \right) v = a' u + w', \quad s \in \{-1, 1\},$$

where $w' \perp u$. Since $\langle u, Bu \rangle = 0$,

$$a' = a + \frac{s}{\rho} \langle u, Bw \rangle \geq a - \frac{\|w\|_2}{\rho} \geq \frac{a}{2}.$$

Also,

$$w' = (M - J) \left(w + \frac{s}{\rho} B(au + w) \right),$$

and hence

$$\begin{aligned} \|w'\|_2 &\leq \lambda \left(\|w\|_2 + \frac{a + \|w\|_2}{\rho} \right) \leq \lambda a \left(\frac{\rho}{2} + \frac{1}{\rho} + \frac{1}{2} \right) \\ &\leq \frac{2\lambda a}{\rho} = \frac{\rho a}{4} \leq \frac{\rho a'}{2}. \end{aligned}$$

Here we used $0 < \rho \leq 1$ and $\rho^2 = 8\lambda$. This proves Equation (5.1), and therefore $q(z) = 2^{-n} a_n \geq 0$.

With this, let $\mathbf{Z}$ have probability mass function q . Expanding the definition of q gives

$$q(z) = 2^{-n} \sum_{S \subseteq [n]} \rho^{-|S|} \mathbb{E}[\mathbf{X}_S] z_S.$$

By orthogonality of the Fourier characters, it follows that, for every $S \subseteq [n]$,

$$\mathbb{E}[\mathbf{Z}_S] = \rho^{-|S|} \mathbb{E}[\mathbf{X}_S].$$

Choose ξ as in the statement. Independence gives

$$\mathbb{E}[(\mathbf{Z} \odot \xi)_S] = \rho^{|S|} \mathbb{E}[\mathbf{Z}_S] = \mathbb{E}[\mathbf{X}_S].$$

Since all Fourier moments agree, $\mathbf{Z} \odot \xi$ and $\mathbf{X}$ have the same distribution. Finally, stationarity and balance give $\mathbb{E}[\mathbf{X}_i] = 0$, and hence

$$\mathbb{E}[\mathbf{Z}_i] = \rho^{-1} \mathbb{E}[\mathbf{X}_i] = 0$$

for every $i \in [n]$. □

We are now ready to prove [Theorem 5.1](#).

Proof of [Theorem 5.1](#). Assume that $\lambda > 0$ (otherwise the statement is trivial), and set $\rho = \sqrt{8\lambda}$. By [Lemma 5.2](#), we may write $\mathbf{X}$ in distribution as $\mathbf{Z} \odot \xi$. Conditional on $\mathbf{Z}$, independence of the noise coordinates gives

$$\mathbb{E}_\xi[f(\mathbf{Z} \odot \xi)] = \sum_{k=0}^n \rho^k f_k(\mathbf{Z}).$$

The constant term is $f_0 = \mathbb{E}[f(\mathbf{U}_n)]$, and the linear term has expectation zero because $\mathbb{E}[\mathbf{Z}_i] = 0$ for every i . Consequently, by [Corollary 2.1](#),

$$\begin{aligned} |\mathbb{E}[f(\mathbf{X})] - \mathbb{E}[f(\mathbf{U}_n)]| &= \left| \sum_{k=2}^n \rho^k \mathbb{E}[f_k(\mathbf{Z})] \right| \\ &\leq \sum_{k=2}^n \rho^k M_k(f) \leq \sum_{k=2}^n (2ed\rho)^k. \end{aligned}$$

Set $\alpha = 2ed\rho$. Our assumption on λ gives

$$\alpha^2 = 32e^2 d^2 \lambda \leq \frac{1}{4}.$$

Therefore,

$$\sum_{k=2}^n \alpha^k \leq \frac{\alpha^2}{1 - \alpha} \leq 2\alpha^2 = 64e^2 d^2 \lambda.$$

The bound on distinguishing probabilities follows from $\mathbf{Pr}[P = 1] = (1 - \mathbb{E}[(-1)^P])/2$. $\square$

A cancellation-sensitive bound. The same argument gives, for every real-valued function $f : \{-1, 1\}^n \rightarrow \mathbb{R}$,

$$|\mathbb{E}[f(\mathbf{X})] - \mathbb{E}[f(\mathbf{U}_n)]| \leq \sum_{k=2}^n (8\lambda)^{k/2} M_k(f), \quad 0 < \lambda \leq 1/8.$$

The quantity appearing here is the signed level norm $M_k(f)$, rather than the Fourier L_1 mass $\sum_{|S|=k} |\hat{f}(S)|$ used in the corresponding criterion of [\[CPTS21\]](#). Thus, the bound can also be useful in other settings where cancellations yield stronger bounds on individual Fourier levels.

The unbalanced case. Balance is used twice in the argument: in the positivity proof, through $\langle u, Bu \rangle = 0$, and in eliminating the linear term when comparing with the uniform distribution. We believe that a suitable modification of the proof extends to unbalanced labelings, with independent labels of the same bias replacing the uniform reference distribution. The resulting bound may depend on the labeling bias.

AI Usage

This disclosure is included as part of the TCS community’s initial efforts to adapt to AI use, and in particular in response to the recommendations of the September 2026 report *AI and TCS: The Next Six Months*.

GPT 5.6 Sol and GPT 6 Astra were used in preparing this manuscript, although almost all ideas were developed without AI assistance. For the most part, AI was used for tasks the author already knew how to carry out. Beyond copy-editing, AI was particularly helpful in developing the idea of the noise representation and in working out the technical details of the dynamic programming approach used in the proof of [Theorem 1.2](#). The author takes full responsibility for the correctness and presentation of the paper.

References

- [AGHP92] Noga Alon, Oded Goldreich, Johan Håstad, and René Peralta. Simple constructions of almost k -wise independent random variables. *Random Structures & Algorithms*, 3(3):289–304, 1992.
- [Bog05] Andrej Bogdanov. Pseudorandom generators for low degree polynomials. In *Annual ACM Symposium on Theory of Computing (STOC)*, page 21–30. ACM, 2005.
- [Bra10] Mark Braverman. Polylogarithmic independence fools AC^0 circuits. *Journal of the ACM*, 57(5):28:1–28:10, 2010.
- [BV10] Andrej Bogdanov and Emanuele Viola. Pseudorandom bits for polynomials. *SIAM Journal on Computing*, 39(6):2464–2486, 2010.
- [CDG26] Gil Cohen, Dean Doron, and Noam Goldgraber. Optimal PRGs for low-degree polynomials over polynomial-size fields. *arXiv preprint arXiv:2602.10030*, 2026.
- [CGL⁺21] Eshan Chattopadhyay, Jason Gaitonde, Chin Ho Lee, Shachar Lovett, and Abhishek Shetty. Fractional pseudorandom generators from any Fourier level. In *36th Computational Complexity Conference*, volume 200 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 10, 24. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2021.

- [CHHL19] Eshan Chattopadhyay, Pooya Hatami, Kaave Hosseini, and Shachar Lovett. Pseudorandom generators from polarizing random walks. *Theory Comput.*, 15:Paper No. 10, 26, 2019.
- [CHL⁺26] Eshan Chattopadhyay, Pooya Hatami, Chin Ho Lee, Shachar Lovett, Avishay Tal, and Emanuele Viola. Exponential correlation bounds for polynomials. *arXiv preprint arXiv:2609.28839*, 2026.
- [CMP⁺22] Gil Cohen, Dor Minzer, Shir Peleg, Aaron Potechin, and Amnon Ta-Shma. Expander random walks: the general case and limitations. In *49th EATCS International Conference on Automata, Languages, and Programming*, volume 229 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 43, 18. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2022.
- [CPTS21] Gil Cohen, Noam Peri, and Amnon Ta-Shma. Expander random walks: a Fourier-analytic approach. In *STOC '21—Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 1643–1655. ACM, New York, [2021] ©2021.
- [CT13] Gil Cohen and Amnon Ta-Shma. Pseudorandom generators for low degree polynomials from algebraic geometry codes. In *Electronic Colloquium on Computational Complexity (ECCC)*, 2013.
- [CZ19] Eshan Chattopadhyay and David Zuckerman. Explicit two-source extractors and resilient functions. *Annals of Mathematics*, 189(3):653–705, 2019.
- [DGV24] Ashish Dwivedi, Zeyu Guo, and Ben Lee Volk. Optimal pseudorandom generators for low-degree polynomials over moderately large fields. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM)*, pages 44:1–44:19. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2024.
- [DV22] Harm Derksen and Emanuele Viola. Fooling polynomials using invariant theory. In *Annual Symposium on Foundations of Computer Science (FOCS)*, pages 399–406. IEEE, 2022.
- [GK21] Venkatesan Guruswami and Vinayak M. Kumar. Pseudobinominality of the sticky random walk. In *12th Innovations in Theoretical Computer Science Conference (ITCS 2021)*, volume 185 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 48:1–48:19. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021.

- [GV22] Louis Golowich and Salil Vadhan. Pseudorandomness of expander random walks for symmetric functions and permutation branching programs. In *37th Computational Complexity Conference*, volume 234 of *LIPICs. Leibniz Int. Proc. Inform.*, pages Art. No. 27, 13. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2022.
- [GX14] Venkatesan Guruswami and Chaoping Xing. Hitting sets for low-degree polynomials with optimal density. In *Conference on Computational Complexity (CCC)*, pages 161–168. IEEE, 2014.
- [HH24] Pooya Hatami and William Hoza. Paradigms for unconditional pseudorandom generators. *Foundations and Trends in Theoretical Computer Science*, 16(1–2):1–210, 2024.
- [Lov09] Shachar Lovett. Unconditional pseudorandom generators for low-degree polynomials. *Theory of Computing*, 5(1):69–82, 2009.
- [Lu12] Chi-Jen Lu. Hitting set generators for sparse polynomials over any finite fields. In *Computational Complexity Conference (CCC)*, pages 280–286. IEEE, 2012.
- [LVW93] Michael Luby, Boban Velickovic, and Avi Wigderson. Deterministic approximate counting of depth-2 circuits. In *Israel Symposium on Theory and Computing Systems (ISTC)*, pages 18–24. IEEE, 1993.
- [NN93] Joseph Naor and Moni Naor. Small-bias probability spaces: efficient constructions and applications. *SIAM J. Comput.*, 22(4):838–856, 1993.
- [Vio07] Emanuele Viola. Pseudorandom bits for constant-depth circuits with few arbitrary symmetric gates. *SIAM Journal on Computing*, 36(5):1387–1403, 2007.
- [Vio09] Emanuele Viola. The sum of d small-bias generators fools polynomials of degree d . *computational complexity*, 18(2):209–217, 2009.