

Pseudorandom Bits for Low-Degree Modular Polynomials

Chin Ho Lee

North Carolina State University

Shachar Lovett*

University of California, San Diego

Tejas Nareddy†

Northeastern University

Emanuele Viola‡

Northeastern University

2026-10-14

Abstract

We revisit the problem of constructing explicit pseudorandom generators which output n bits that look random to n -variate degree- d polynomials modulo m . We give new generators with seed length $2^{O(dm^2)} \log(n/\varepsilon)$, which is optimal for any fixed m and d . Along the way, we prove new Fourier growth bounds for polynomial phase functions on the Boolean cube.

Our results simplify previous constructions, improve their parameters, and extend them to composite moduli. The previous best seed length for odd prime m had a tower dependence on the degree (Lovett, Mukhopadhyay, and Shpilka; FOCS 2010); and for general composite m was at least $2^{\sqrt{\log n}}$ even for degree 2 (Luby, Veličković, and Wigderson; ISTCS 1993).

As an application, we obtain generators with optimal seed length for branching programs over nilpotent groups, extending recent work (Lee and Viola; CCC 2025). We also obtain new results on the so-called coin problem for polynomials modulo composites.

1 Introduction

The explicit construction of pseudorandom generators (PRGs) for multivariate polynomials is a fundamental direction that is also linked to longstanding challenges in circuit lower bounds, via classical simulations of small-depth circuits by polynomials (see, e.g., [Vio26]). In one setting, the generator is allowed to output elements from the entire domain of the polynomial. In this setting, generators are known when the field is small and the degree is at most logarithmic in the number of variables [BV10, Lov09, Vio09], or when the field is large [Bog05, DV22, DGV24, CDG26].

*Supported by Simons Investigator Award #929894 and NSF Award CCF-2425349.

†Supported by NSF grant CCF-2430026.

‡Supported by NSF grant CCF-2430026.

In this paper, we are concerned with a different setting in which the generator is only allowed to output *bits*. The two settings are the same over the field $\mathbb{F}_2$, but over other domains, the latter setting turns out to be more difficult. Even for linear polynomials over $\mathbb{F}_3$, generators with optimal seed length were not known until [LRTV09, MZ09], more than 15 years after optimal generators for $\mathbb{F}_2$ became available [NN93]. The difficulty persists despite significant attention to the problem, motivated in part by the central goal of developing generators for small-depth circuits with modular gates, which naturally take bits as input and can simulate polynomials.

For higher-degree polynomials, the first generator was given by Luby, Veličković, and Wigderson in the early '90s [LVW93]. They achieved a seed length of about $2^{\sqrt{\log n}}$ by a clever use of the Babai–Nisan–Szegedy multiparty lower bound [BNS92]. An alternative proof was given in [Vio07]. This remained the state of the art even for quadratic polynomials over $\mathbb{F}_3$ until a rather technical result by Lovett, Mukhopadhyay, and Shpilka [LMS13]. They achieved a seed length of $O_p(1) \exp^{(O(d))}(1/\varepsilon) \log n$ for degree- d polynomials over the field $\mathbb{F}_p$, where $\exp^{(t)}$ denotes iterated exponentiation.

In this paper, we give new generators that substantially simplify earlier constructions, improve their seed lengths, and extend them to composite moduli. In fact, more generally we obtain generators over domains $[r]^n$, where $[r]$ denotes the set $\{0, \dots, r-1\}$. Let us define for $2 \leq r \leq m$:

$$\text{Poly}(r, m, d) := \{f: [r]^n \rightarrow \mathbb{Z}_m \text{ computable by a polynomial over } \mathbb{Z}_m \text{ of degree at most } d\}.$$

Given a distribution X over $[r]^n$ and a function f on $[r]^n$, we denote by $f(X)$ the distribution of f applied to an input sampled from X , and by $f([r]^n)$ the distribution of f applied to a uniformly chosen element of $[r]^n$. We denote by d_{TV} the statistical distance between two distributions. Our main result is an explicit construction of pseudorandom generators for $\text{Poly}(r, m, d)$ with seed length $O_{m,d}(\log(n/\varepsilon))$.

Theorem 1.1. *Let p be the least prime divisor of m and $2 \leq r \leq p$. There is an explicit distribution X on $[r]^n$ samplable from a seed of length $2^{O(dm^2)} \log(n/\varepsilon)$, such that $d_{\text{TV}}(f(X), f([r]^n)) \leq \varepsilon$ for every degree- d polynomial f in n variables modulo m .*

We obtain several constructions. The best parameters are obtained by bounding the *Fourier growth* of the class $\text{Poly}(r, m, d)$. For a function $f: \mathbb{Z}_r^n \rightarrow \mathbb{Z}_m$, consider its phase $F := e_m(f)$, where $e_m(x) := \exp(\frac{2\pi i}{m}x)$ and i is the imaginary unit. Its Fourier expansion is

$$F(x) = \sum_{S \in \mathbb{Z}_r^n} \widehat{F}(S) \chi_S(x),$$

where $\widehat{F}(S) = \mathbf{E}_{\mathbf{x} \sim \mathbb{Z}_r^n} [F(\mathbf{x}) \overline{\chi_S(\mathbf{x})}]$ and $\chi_S(x) := e_r(\langle S, \mathbf{x} \rangle)$. Let $|S|$ denote the Hamming weight of S , and define

$$L_{1,k}(F) := \sum_{S \in \mathbb{Z}_r^n: |S|=k} |\widehat{F}(S)|,$$

and

$$L_{1,k}(\text{Poly}(r, m, d)) := \max_{f \in \text{Poly}(r, m, d)} L_{1,k}(e_m(f)).$$

A technical component of our main result is showing that the class $\text{Poly}(r, m, d)$ has bounded Fourier growth.

Theorem 1.2. $L_{1,k}(\text{Poly}(r, m, d)) \leq (\frac{r-1}{m} \cdot 2^{m^2d})^k$.

This improves and extends [CHHL19], which proves a weaker bound of $(k2^{3d})^k$ for the special case of $\mathbb{F}_2$ -polynomials of degree d , i.e. the class $\text{Poly}(2, 2, d)$.

As a direct consequence (see, e.g., [Lee19, Fact 9]), Theorem 1.2 gives new bounds on the *coin problem* for polynomials over composite moduli. The coin problem for polynomials was studied in a number of works starting with [SV10]. We denote by B_q^n the Bernoulli distribution over $\{0, 1\}^n$ where each bit is set to 1 independently with probability q and 0 with probability $1 - q$.

Corollary 1.3. $d_{\text{TV}}(f(B_{1/2-\varepsilon/2}^n), f(B_{1/2}^n)) \leq 2^{m^2d}\varepsilon$ for every $f \in \text{Poly}(2, m, d)$.

As another application, we obtain generators with optimal seed length for branching programs over *nilpotent groups*. The earlier work [LV25] obtained such generators for *read-once* branching programs over the special case of nilpotent groups of *prime-power* size. They also observed that, for 2-groups, the result extends to non-read-once programs of unrestricted length. Our techniques also yield a simpler analysis, and extend the latter observation to every finite nilpotent group.

Theorem 1.4. *Let G be a finite nilpotent group. There is a distribution X on $\{0, 1\}^n$ which can be explicitly sampled from a uniformly random seed of length $O_G(\log(n/\varepsilon))$, such that $d_{\text{TV}}(f(X), f(\{0, 1\}^n)) \leq \varepsilon$ for every $f: \{0, 1\}^n \rightarrow G$ that can be written as*

$$f(x_1, x_2, \dots, x_n) = \prod_{i=1}^L g_i^{y_i}$$

for some L , where each $g_i \in G$ and each y_i is a variable x_j .

Related results. The subsequent works [CHL⁺26, Coh26] obtained new pseudorandom generators for Boolean-valued polynomials over arbitrary fields, with polynomial dependence on the degree and poly-logarithmic dependence on n/ε in the seed lengths. These results and the ones in this paper are incomparable. The generators in [CHL⁺26, Coh26] apply to large degree, but their seed lengths are not optimal even for constant degree. Our generators apply to arbitrary integer moduli and achieve optimal seed length $O_{m,d}(\log(n/\varepsilon))$ for fixed modulus and fixed degree; both properties are critical for our applications. The proof techniques are also different.

1.1 Proof Overview

Throughout this discussion, we will assume that both the modulus m and the degree d are constant. As mentioned earlier, we actually obtain several different constructions. We begin with a very simple construction for prime moduli which has seed length $O(\log(n/\varepsilon) \log(1/\varepsilon))$ and contains a key new idea in this work.

1.1.1 Simple Bit-PRGs for polynomials over prime moduli

We start with results on the coin problem ([Corollary 1.3](#)). For this construction, it suffices to use the result of [\[Sri23\]](#) that $O(\varepsilon)$ -biased coins fool low-degree polynomials over $\mathbb{F}_p$ with error ε . Our goal is to *derandomize* the biased coins. For *constant* bias, this was already accomplished in [\[LRTV09, MZ09\]](#) and used in other papers such as [\[LV25\]](#). Their idea is as follows. Observe that a constant-biased *bit* can be sampled by applying a polynomial $\phi: \mathbb{Z}_p \rightarrow \{0, 1\}$ of degree at most $p - 1 = O(1)$ on a uniform field element of $\mathbb{Z}_p$. Hence, to derandomize this biased product distribution against degree- d polynomials, we apply ϕ coordinatewise to the pseudorandom vectors in $\mathbb{Z}_p^n$ that fool polynomials over $\mathbb{F}_p$ of degree $d(p - 1) = O(1)$, that is, the class $\text{Poly}(p, p, d(p - 1))$, using the result [\[Vio09\]](#).

A key new idea in this work is to reduce the bias from constant to ε via an XOR lemma. Specifically, take $t = O(\log(1/\varepsilon))$ independent uniform vectors $V_1, \dots, V_t$ from $\mathbb{Z}_p^n$, apply ϕ to the V_i 's coordinatewise, and output their XOR

$$\phi(V_1) \oplus \dots \oplus \phi(V_t).$$

We again derandomize this construction by replacing the V_i 's with pseudorandom vectors R_i in $\mathbb{Z}_p^n$ that fool low-degree polynomials over $\mathbb{F}_p$ [\[Vio09\]](#) with error $O(\varepsilon/t)$, each of which can be sampled using $O(\log(n/\varepsilon))$ bits.

To analyze correctness, we apply a hybrid argument in which we replace the V_j 's with the R_j 's, one copy at a time. Consider a hybrid distribution

$$H_j := \phi(V_1) \oplus \dots \oplus \phi(V_j) \oplus \phi(R_{j+1}) \oplus \dots \oplus \phi(R_t).$$

Given $f \in \text{Poly}(2, p, d)$, observe that fixing the XOR of all terms except the j -th to w , the resulting test

$$z \mapsto f(\phi(z) \oplus w)$$

is also a low-degree polynomial over $\mathbb{Z}_p$, which is fooled by R_j .

1.1.2 Optimal seed length over prime moduli

To improve the seed length to $O(\log(n/\varepsilon))$ over prime moduli, we use the “pseudorandomness plus noise” framework that has emerged in a series of works [\[AW89, GMR⁺12, HLV18, LV20, CHHL19\]](#) (see also [\[HH24\]](#)). In particular, the result [\[CHHL19\]](#) implies that any function family (that is closed under restrictions and translations) with bounded Fourier growth is ε -fooled by the XOR of a $\text{poly}(\varepsilon/n)$ -biased distribution and constant-biased bits. Crucially, the bias is a constant depending only on p and d , as opposed to $O(\varepsilon)$. As a result, to derandomize the biased bits, it suffices to XOR $O(1)$ many copies of pseudorandom field elements for low-degree polynomials.

Therefore, to apply this framework, we need to show that $\text{Poly}(2, p, d)$ has bounded Fourier growth. We first discuss [Lemma 3.1](#), which proves

$$L_{1,k}(\text{Poly}(p, p, d)) \leq (2p^d)^k = O_{p,d}(1)^k.$$

(Note that the domain of $\text{Poly}(p, p, d)$ is $[p]^n$ instead of $[2]^n$.) Here we focus on $p = 2$, as the idea extends to any p naturally. We will bound $L_{1,k}(\text{Poly}(2, 2, d))$ inductively on the degree d , by refining the following approach taken in [\[CHHL19\]](#).

Fix an $f \in \text{Poly}(2, 2, d)$ and let $F := e_2(f)$ be its phase. The work [CHHL19] bounds $L_{1,k}(F)$ by considering its square:

$$L_{1,k}(F)^2 = \sum_{|A|=|B|=k} |\widehat{F}(A)\widehat{F}(B)| = \sum_{0 \leq |I| \leq k} \sum_{\substack{|A|=|B|=k: \\ A \cap B = I}} |\widehat{F}(A)\widehat{F}(B)|. \quad (1)$$

When $A \cap B = \emptyset$, they upper bound $|\widehat{F}(A)\widehat{F}(B)|$ by the average of $|\widehat{\partial_y F}(A - B)|$, where $\partial_y F$ are the phases of polynomials of degree at most $d - 1$ (see Claim 3.6). However, they handle the overlapping (A, B) 's through a recursion that does not reduce the degree.

Here we reduce the overlapping case to the disjoint case via restrictions of F . Specifically, write $A = S \cup I$ and $B = T \cup I$ where S and T are disjoint. For a vector $z \in \mathbb{Z}_2^I$, let F_z denote the restriction of F obtained by fixing the coordinates in I to z . Then, we have $\widehat{F}(S \cup I) = \mathbf{E}_z[\widehat{F_z}(S)\chi_I(z)]$. As multiplication by $\chi_I(z)$ preserves the magnitude of a coefficient, for any $z, z' \in \mathbb{Z}_2^I$, we can bound $\sum_{S \cap T = \emptyset} |\widehat{F_z}(S)\widehat{F_{z'}}(T)|$ by reducing to its derivatives as discussed before.

Next, we need to account for the number of pairs (A, B) whose intersection is I , where I is any subset of $\{1, \dots, n\}$ of size at most k . Rather than directly counting such pairs, we normalize the absolute level- k Fourier coefficients by $L_{1,k}(F)$ to obtain a distribution $\mathcal{S}$ on k -element sets. Let $\mathbf{S}, \mathbf{T}$ be independent samples from $\mathcal{S}$. For each possible intersection I , we combine the exact expression for $\Pr[\mathbf{S} \cap \mathbf{T} = I]$ with the bound $\Pr[\mathbf{S} \cap \mathbf{T} = I] \leq \Pr[I \subseteq \mathbf{S}]^2$ using their geometric mean to obtain

$$\Pr[\mathbf{S} \cap \mathbf{T} = I] \leq \frac{1}{L_{1,k}(F)} \left(\sum_{\substack{|A|=|B|=k: \\ A \cap B = I}} |\widehat{F}(A)\widehat{F}(B)| \right)^{1/2} \Pr[I \subseteq \mathbf{S}].$$

Summing over I turns the unrestricted sum over subsets in $\{1, \dots, n\}$ in (1) into an expectation over subsets of a random k -element set, giving

$$L_{1,k}(F) \leq \mathbf{E}_{\mathbf{S} \sim \mathcal{S}} \sum_{I \subseteq \mathbf{S}} \left(\sum_{\substack{|A|=|B|=k: \\ A \cap B = I}} |\widehat{F}(A)\widehat{F}(B)| \right)^{1/2}.$$

It remains to upper bound $L_{1,k}(\text{Poly}(2, p, d))$ in terms of $L_{1,k}(\text{Poly}(p, p, d(p - 1)))$. This step is similar in spirit to our way of sampling biased coins using a polynomial over $\mathbb{Z}_p$. In Lemma 3.4, we construct a randomized low-degree polynomial $\Phi_{u,v}$ mapping each coordinate from $[p]$ into $[2]$ such that the $L_{1,k}$ -norm of the phase of $f \in \text{Poly}(2, p, d)$ can be bounded by the $L_{1,k}$ -norm of the phase of f composed with $\Phi_{u,v}$, which has degree at most $d(p - 1)$.

1.1.3 Composite moduli

We now explain how to handle composite moduli. Our approach relies on two complementary reductions. *Modulus reduction* replaces a problem over a composite modulus with one over a smaller modulus, at the cost of increasing the degree. *Domain reduction* replaces a problem over a smaller input alphabet with one over a larger alphabet, again at the cost of increasing the degree. By applying these reductions alternately, we eventually reach polynomials over a prime field, where the techniques for prime moduli apply.

Modulus reduction. Let $m \geq 2$ be composite and p be its least prime divisor. At a very high level, this reduction lets us reduce the question about $\text{Poly}(p, m, d)$ to the question about $\text{Poly}(p, m/p, d(p-1))$.

Our main observation is that the analysis of pseudorandom generators for polynomials modulo 2 in [Vio09] works equally well when the phases of the $\mathbb{F}_2$ -polynomials are multiplied by the phases of polynomials over other moduli. A similar observation was discovered independently in the work of [Kra26] on pseudorandom generators for linear functions modulo m .

To illustrate the idea, let us first consider the special case of $\text{Poly}(2, 6, d)$. By the Chinese Remainder Theorem, their phases can be written as $e_2(g)e_3(h)$ where g and h are integer polynomials of degree at most d . (Recall $e_m(f) = \exp(\frac{2\pi i}{m}f)$.) Let us consider the class of functions

$$\mathcal{F}_t := \{e_2(g)e_3(h) : g \in \text{Poly}(2, 2, t), h \in \text{Poly}(2, 3, d)\}.$$

Let $F \in \mathcal{F}_t$. Observe that its derivative $\partial_{u,v}F := F(x+u)\overline{F(x+v)}$ can be written as $e_2(g')e_3(h')$, where g' has degree at most $t-1$, and h' lies in the span of “shifted polynomials” $\{h(x \oplus u) : u \in \mathbb{Z}_2^n, h \in \text{Poly}(2, 3, d)\}$. (More generally, we only require that h can be written as a linear combination of arbitrary $\mathbb{Z}_3$ -valued functions on $[2]^n$, each depending on at most d coordinates.) This allows us to reduce the degree of the “mod 2” phase without making the “mod 3” phase too complicated.

After taking d derivatives, we reduce the question about $\text{Poly}(2, 6, d)$ to the question about shifted polynomials over $\mathbb{Z}_3$ of degree at most d . For any $x, u \in \{0, 1\}$, the 1-coordinate map $x \mapsto x \oplus u$ can be written as the affine linear function $x + u - 2ux$ over $\mathbb{Z}_3$, and thus the span of shifted polynomials belongs to $\text{Poly}(2, 3, d)$.

Handling moduli m that are divisible by the square of a prime requires some additional care, as taking derivatives of a polynomial phase over a prime power does not necessarily reduce its degree, due to a “wrap-around” issue. As an example, consider the univariate linear function $f(x) := x$ in $\text{Poly}(2, 4, 1)$. Then

$$\partial f(x) := f(x \oplus 1) - f(x) = (x+1) \bmod 2 - x = \begin{cases} 1 & \text{if } x = 0 \\ -1 & \text{if } x = 1, \end{cases}$$

which is not constant. Nevertheless, we have

$$\partial f(x) = 1 - 2x \quad \text{and} \quad e_4(\partial f)(x) = e_4(1 - 2x) = i e_2(x).$$

Thus, the phase of the derivative is a constant phase multiplied by a polynomial phase modulo 2. In particular, we reduce the degree of the first component and the modulus of the second component.

More generally, let $m = p^a m'$ with $\gcd(p, m') = 1$, and take the input domain to be $[p]^n$. Let $F = e_m(f)$, where $f \in \text{Poly}(p, m, d)$. In Lemma 2.5, we show that $\partial_{u,v}F$ can be written as $e_{p^a}(g)e_{m/p}(h)$, where $g \in \text{Poly}(p, p^a, d-1)$ and $h \in \text{Poly}(p, m/p, d(p-1))$.

To summarize, by taking d derivatives, we can reduce our question on $\text{Poly}(p, m, d)$ to a question on $\text{Poly}(p, m/p, d(p-1))$.

We prove our modulus reduction for Fourier growth in Lemma 3.3 and for PRGs in Lemma 4.4.

Domain reduction. The idea of domain reduction already appeared when handling polynomials over prime moduli. For Fourier growth, we reduce bounds on $\text{Poly}(2, p, d)$ to bounds on $\text{Poly}(p, p, O(d))$ via randomized low-degree polynomials mapping the domain from $[p]$ to $[2]$. Likewise, to sample biased elements in the pseudorandomness plus noise framework, we sample these biased bits from uniform elements in $\mathbb{Z}_p$ using a low-degree polynomial mapping $[p]$ to $[2]$.

We extend this idea further as follows. Let p be the least prime divisor of m . We prove in [Lemma 2.3](#) that any map from $[p]$ to $[m]$, when viewed modulo m , can be represented by a polynomial over $\mathbb{Z}_m$ of degree at most $p - 1$. Applying this interpolation coordinatewise reduces questions about $\text{Poly}(r, m, d)$ to questions about $\text{Poly}(p, m, d(p - 1))$.

Putting the two reductions together. We now explain how to use these ideas to construct PRGs for polynomials modulo composites.

Let us start with constructing a PRG for the class $\text{Poly}(2, 12, d)$. Recall that we need two ingredients: (1) a Fourier growth bound on this class, and (2) sampling biased bits with a polynomial modulo 12. Note that we cannot apply domain reduction directly from $[2]$ to $[12]$, because [Lemma 2.3](#) does not allow arbitrary maps on all of $[12]$ to be interpolated over $\mathbb{Z}_{12}$. The idea, as observed in [\[Kra26\]](#), is that we can first apply modulus reduction to obtain corresponding problems for $\text{Poly}(2, 6, O(d))$, and then once more to reach $\text{Poly}(2, 3, O(d))$. The remaining modulus is now prime. So we can apply domain reduction to reduce further to $\text{Poly}(3, 3, O(d))$.

More generally, consider the problem of bounding the Fourier growth and constructing PRGs for $\text{Poly}(2, m, d)$, where $p_1 < \dots < p_k$ are all the distinct prime factors of m . Let $m = p_1^{a_1} m'$ where $p_1^{a_1}$ is the largest power of p_1 dividing m . We first apply domain reduction to reduce from $\text{Poly}(2, m, d)$ to $\text{Poly}(p_1, m, O(d))$, followed by applying modulus reduction repeatedly to reduce $\text{Poly}(p_1, m, O(d))$ to $\text{Poly}(p_1, m/p_1^{a_1}, O(d))$, where p_1 is no longer a divisor of $m' := m/p_1^{a_1}$. Then we apply domain reduction again to reduce $\text{Poly}(p_1, m', O(d))$ to $\text{Poly}(p_2, m', O(d))$, where $p_2 > p_1$ is the least prime divisor of m' . We apply the two reductions alternately. At the final prime-power modulus $p_k^{a_k}$, we apply modulus reduction only $a_k - 1$ times, reducing the question to $\text{Poly}(p_k, p_k, O(d))$, for which we can either apply [Lemma 3.1](#) or the generator in [\[Vio09\]](#).

Paper organization. In [Section 2](#), we set up some preliminaries and develop some technical tools that will be used in both modulus and domain reductions. We prove our new Fourier growth bounds for modular polynomials ([Theorem 1.2](#)) in [Section 3](#), construct our PRGs ([Theorem 1.1](#)) for polynomials in [Section 4](#), and PRGs for branching programs over nilpotent groups ([Theorem 1.4](#)) in [Section 5](#).

2 Interpolation and Derivative-Graded Classes

In this section, we state some standard facts and definitions in pseudorandomness, followed by several lemmas that will be used in [Sections 3](#) and [4](#).

The following standard claim shows that fooling polynomials in statistical distance is equivalent to fooling the bias of their phases. See, e.g., [\[LV25, Claim 21\]](#) for a proof.

Claim 2.1. Let Y_1 and Y_2 be distributions over $\mathbb{Z}_m$. Define

$$\alpha := \max_{t \in \mathbb{Z}_m \setminus \{0\}} \left| \mathbf{E}_{\mathbf{x} \sim Y_1} [e_m(t\mathbf{x})] - \mathbf{E}_{\mathbf{x} \sim Y_2} [e_m(t\mathbf{x})] \right|.$$

Then:

$$\frac{\alpha}{2} \leq d_{\text{TV}}(Y_1, Y_2) \leq \frac{\sqrt{m-1}}{2} \cdot \alpha.$$

Throughout this paper, we use U_r to denote the uniform distribution on $[r]^n$, and $\mathbb{D} := \{z \in \mathbb{C} : |z| \leq 1\}$ to denote the complex unit disk. We use $e_m(\text{Poly}(r, m, d))$ to denote $\{e_m(f) : f \in \text{Poly}(r, m, d)\}$.

Definition 2.2 (Fooling a function family). Let $r \geq 2$, $\varepsilon \in (0, 1)$, and $\mathcal{F} \subseteq \{f : [r]^n \rightarrow \mathbb{D}\}$ be a family of functions. A distribution X on $[r]^n$ ε -fools $\mathcal{F}$ if

$$|\mathbf{E}[F(X)] - \mathbf{E}[F(U_r)]| \leq \varepsilon \quad \text{for every } F \in \mathcal{F}.$$

We also say X fools $\mathcal{F}$ with error ε .

Lemma 2.3 (Interpolation on a short interval). Let $m \geq 2$, and suppose every prime divisor of m is at least p . Then any map $\phi : [p] \rightarrow \mathbb{Z}_m$ agrees on $[p]$ with a polynomial over $\mathbb{Z}_m$ of degree at most $p-1$.

Note that if m has a prime divisor $q < p$, then the conclusion is false (for example, by considering a map where $\phi(q) \not\equiv \phi(0) \pmod{q}$).

We will apply [Lemma 2.3](#) over both $\mathbb{Z}_m$ and $\mathbb{Z}_\ell$ for divisors ℓ of m .

Proof. The lemma holds because the Lagrange interpolation polynomial only requires inverses of $1, \dots, p-1$ in $\mathbb{Z}_m$, which are available as they are all coprime to m . Alternatively, consider the $p \times p$ Vandermonde matrix

$$M_{i,j} = i^j \in \mathbb{Z}_m, \quad 0 \leq i, j \leq p-1.$$

Its determinant is

$$\det(M) = \prod_{0 \leq i < j \leq p-1} (j - i).$$

Every factor $j - i$ lies in $\{1, \dots, p-1\}$ and is therefore coprime to m . Thus $\det(M)$ is a unit in $\mathbb{Z}_m$, so M is invertible over $\mathbb{Z}_m$. Solving the corresponding interpolation system gives the desired polynomial. Concretely, if we set a vector $v = (\phi(0), \dots, \phi(p-1))$, and solve $Mu = v$ where $u = (u_0, \dots, u_{p-1})$, then the polynomial is $\sum_{i=0}^{p-1} u_i x^i$. $\square$

We will consider functions f mapping $\mathbb{Z}_m^n$ to the torus $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. We will use $e(f)(x)$ to denote $\exp(2\pi i f(x))$.

Definition 2.4 (Derivative-graded classes). Fix an integer $m \geq 2$ and let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ be the torus. For each degree index d and input dimension n , let $\mathcal{F}_d^{[n]}$ be a class of functions $\mathbb{Z}_m^n \rightarrow \mathbb{T}$. Suppressing the dimension index, we call $\{\mathcal{F}_d\}$ derivative-graded if it satisfies the following properties:

1. (Closure under addition) if $f, g \in \mathcal{F}_d$ then $f + g \in \mathcal{F}_d$.
2. (Closure under restriction) $\mathcal{F}_d$ is closed under variable restrictions.
3. (Closure under shifts) if $f \in \mathcal{F}_d$ then $f(x + u) \in \mathcal{F}_d$ for all $u \in \mathbb{Z}_m^n$.
4. (Derivatives reduce degree) For $d \geq 1$, if $f \in \mathcal{F}_d$ then $f(x + u) - f(x + v) \in \mathcal{F}_{d-1}$.

We use $e(\mathcal{F}_d)$ to denote $\{e(f) : f \in \mathcal{F}_d\}$, and $L_{1,k}(\mathcal{F}_d)$ to denote $\sup_{f \in \mathcal{F}_d} L_{1,k}(e(f))$.

We will be working with the following derivative-graded family. First, define

$$\text{Junta}(p, m, d) := \text{span}_{\mathbb{Z}_m} \{h(x_S) : |S| \leq d, h : [p]^S \rightarrow \mathbb{Z}_m\}.$$

In words, $\text{Junta}(p, m, d)$ consists of $\mathbb{Z}_m$ -valued functions on $[p]^n$ that can be expressed as sums of arbitrary functions, each depending on at most d coordinates.

Let m be composite and p be its least prime divisor. Write $m = p^a m'$ with $\gcd(p, m') = 1$. For $0 \leq t \leq d$, define

$$\mathcal{F}_t(p, m, d) := \left\{ \frac{1}{p^a} g(x) + \frac{p}{m} h(x) : g \in \text{Poly}(p, p^a, t) \text{ and } h \in \text{Junta}(p, \frac{m}{p}, d) \right\}.$$

Note that the index t measures the degree of the p^a -modulus component, which decreases under differentiation. The fixed parameter d bounds the coordinate dependence of the lower-modulus component, which need not decrease. When p, m , and d are clear from the context, we will abbreviate it as $\mathcal{F}_t$.

Lemma 2.5. $\{\mathcal{F}_t\}_{t=0}^d$ satisfies [Definition 2.4](#). Moreover, we have

$$\begin{aligned} \mathcal{F}_d &\supseteq \left\{ \frac{1}{m} f : f \in \text{Poly}(p, m, d) \right\} \\ \mathcal{F}_0 &\subseteq \left\{ \frac{k}{p^a} + \frac{p}{m} f : k \in \mathbb{Z}, f \in \text{Poly}(p, \frac{m}{p}, d(p-1)) \right\}. \end{aligned}$$

This lemma lets us reduce the modulus of polynomials by taking derivatives. Specifically, after d derivatives, the phase becomes a polynomial phase modulo m/p of degree at most $d(p-1)$, up to multiplication by a constant unit-modulus factor. It will be used several times throughout the paper.

Proof of Lemma 2.5. It is easy to verify that $\{\mathcal{F}_t\}$ satisfies the first two properties in [Definition 2.4](#). We first show that taking derivatives reduces degree. Closure under shifts then follows because $f(x + u) = f(x) + (f(x + u) - f(x))$ and $\mathcal{F}_t + \mathcal{F}_{t-1} \subseteq \mathcal{F}_t$. For $\mathcal{F}_0$, any shift preserves the constant part and the junta part directly.

Let $1 \leq t \leq d$. Fix $\frac{1}{p^a} g + \frac{p}{m} h \in \mathcal{F}_t$ where $g \in \text{Poly}(p, p^a, t)$ and $h \in \text{Junta}(p, \frac{m}{p}, d)$. For $u, v \in [p]^n$, let $u +_p v$ denote their coordinatewise addition modulo p . Define

$$\partial_{u,v}^{+p} f(x) := f(x +_p u) - f(x +_p v).$$

It is clear that $\partial_{u,v}^{+p} h \in \text{Junta}(p, \frac{m}{p}, d)$, and we will show that $\frac{1}{p^a} \partial_{u,v}^{+p} g \in \mathcal{F}_{t-1}$. Let $\bar{g}$ be an integer polynomial of degree at most t satisfying $\bar{g}(x) \equiv g(x) \pmod{p^a}$ on $[p]^n$. As $\mathcal{F}_{t-1}$ is

closed under addition, it suffices to consider when $\bar{g}$ is a monomial of degree at most t . For $u, v \in [p]^n$, define

$$\partial_{u,v}\bar{g}(x) := \bar{g}(x+u) - \bar{g}(x+v),$$

where addition is over the integers. Then $\partial_{u,v}\bar{g}$ has degree at most $t-1$. Moreover, for any fixed $w \in [p]^n$, we have $\bar{g}(x+_pw) \equiv \bar{g}(x+w) \pmod{p}$. So $\bar{g}(x+_pw) - \bar{g}(x+w)$ is divisible by p . In addition, as w is fixed, the function

$$R_w(x) := \frac{\bar{g}(x+_pw) - \bar{g}(x+w)}{p}$$

depends only on the same coordinates as $\bar{g}$. So $\partial_{u,v}^{+p}\bar{g} = \partial_{u,v}\bar{g} + p(R_u - R_v)$. Since $m = p^a m'$, we have

$$\frac{1}{p^a}\partial_{u,v}^{+p}g = \frac{1}{p^a}\partial_{u,v}\bar{g} + \frac{1}{p^{a-1}}(R_u - R_v).$$

The polynomial $\partial_{u,v}\bar{g}$ has degree at most $t-1$. Each R_w depends on at most $t \leq d$ coordinates and $\frac{1}{p^{a-1}} = \frac{p}{m}$. So, $m'(R_u - R_v)$, reduced modulo m/p , belongs to $\mathbf{Junta}(p, m/p, d)$. Hence, $\frac{1}{p^a}\partial_{u,v}^{+p}g \in \mathcal{F}_{t-1}$.

It remains to prove the inclusions. For the first one, let $f \in \mathbf{Poly}(p, m, d)$. Define $g := f \bmod p^a$ and $h := f \bmod \frac{m}{p} \in \mathbf{Poly}(p, \frac{m}{p}, d) \subseteq \mathbf{Junta}(p, \frac{m}{p}, d)$. As $\gcd(p, m') = 1$, there are integers A and B such that $1 = Am' + Bp$. Thus, as torus-valued functions, we have $\frac{1}{m}f = \frac{m'}{m}Ag + \frac{p}{m}Bh = \frac{1}{p^a}(Ag) + \frac{p}{m}(Bh)$.

For the second inclusion, note that $\mathbf{Poly}(p, p^a, 0)$ consists of constant functions, and we claim that $\mathbf{Junta}(p, \frac{m}{p}, d) \subseteq \mathbf{Poly}(p, \frac{m}{p}, d(p-1))$. Indeed, every function $h: [p]^S \rightarrow \mathbb{Z}_{m/p}$ can be written as a linear combination of indicators

$$\prod_{i \in S} \mathbb{1}(x_i = a_i),$$

where $\mathbb{1}(x_i = a_i)$ has degree at most $p-1$, by interpolation ([Lemma 2.3](#)). Since $|S| \leq d$, each product indicator has degree at most $d(p-1)$. $\square$

3 Fourier Growth of Modular Polynomials

In this section, we prove [Theorem 1.2](#), restated below. Recall that $\mathbf{Poly}(r, m, d)$ is only defined for $2 \leq r \leq m$.

Theorem 1.2 (Restated). $L_{1,k}(\mathbf{Poly}(r, m, d)) \leq (\frac{r-1}{m} \cdot 2^{m^2d})^k$.

In [Sections 3.1](#) and [3.2](#), we will prove the following reductions.

Lemma 3.1. $L_{1,k}(\mathbf{Poly}(m, m, d)) \leq (2m^d)^k$ for every k .

Lemma 3.2. Let $m \geq 2$ be any integer and p be its least prime divisor. Then for any $r \leq m$,

$$L_{1,k}(\mathbf{Poly}(r, m, d)) \leq \left(\frac{p(r-1)}{p-1}\right)^k L_{1,k}(\mathbf{Poly}(p, m, d(p-1))) \quad \text{for every } k.$$

Lemma 3.3. *Let m be composite, and p be its least prime divisor. Write $m = p^a m'$ with $\gcd(p, m') = 1$. Let $B \geq 1$. Suppose $L_{1,k}(\text{Poly}(p, \frac{m}{p}, d(p-1))) \leq B^k$ for every k . Then*

$$L_{1,k}(\text{Poly}(p, m, d)) \leq \left(\frac{p}{p-1} \cdot p^d B \right)^k \quad \text{for every } k.$$

Proof of Theorem 1.2. For $d = 0$, we have $L_{1,k}(\text{Poly}(r, m, d)) = 1$ for $k = 0$ and 0 for $k > 0$. So, we will assume $d \geq 1$ henceforth. Let $B(r, m, d)$ denote the smallest $B \geq 1$ such that $L_{1,k}(\text{Poly}(r, m, d)) \leq B^k$ for every k . We will upper bound $B(r, m, d)$ using the following reductions:

1. (Lemma 3.1) $B(m, m, d) \leq 2m^d$ for any m .
2. (Lemma 3.2) $B(r, m, d) \leq (r-1) \cdot \frac{p}{p-1} \cdot B(p, m, d(p-1))$ for any $r \leq m$, where p is the least prime divisor of m .
3. (Lemma 3.3) $B(p, m, d) \leq \frac{p}{p-1} \cdot p^d \cdot B(p, \frac{m}{p}, d(p-1))$, where m is composite and p is the smallest prime divisor of m .

We prove the desired bound on $B(r, m, d)$ by induction on the number of prime factors of m , counted with multiplicity, simultaneously for all r and d .

Suppose $m = p$ is prime. Applying Item 2 followed by Item 1 gives

$$B(r, p, d) \leq (r-1) \frac{p}{p-1} B(p, p, d(p-1)) \leq (r-1) \frac{2p}{p-1} p^{d(p-1)} \leq \frac{r-1}{p} 2^{p^2 d}.$$

Suppose m is composite. Write $m = p^a m'$, where p is the smallest prime divisor of m and $\gcd(p, m') = 1$. Applying Item 2 followed by Item 3 gives

$$\begin{aligned} B(r, m, d) &\leq (r-1) \cdot \frac{p}{p-1} \cdot B(p, m, d(p-1)) \\ &\leq (r-1) \cdot \frac{p}{p-1} \cdot \frac{p}{p-1} \cdot p^{d(p-1)} \cdot B(p, \frac{m}{p}, d(p-1)^2) \\ &\leq \frac{r-1}{m} \cdot 2p^2 \cdot p^{d(p-1)} \cdot 2^{m^2 d(1-1/p)^2} \leq \frac{r-1}{m} 2^{m^2 d}, \end{aligned}$$

where the last inequality is because $2p^2 p^{d(p-1)} \leq 2^{dp^2} \leq 2^{\frac{m^2 d}{p}}$. $\square$

We now establish the lemmas that let us reduce the domain r and the degree d of $\text{Poly}(r, m, d)$ when proving their level- k Fourier bounds.

3.1 Domain Reduction via Randomized Polynomials

In this subsection, we prove our domain reduction lemma for Fourier growth.

Lemma 3.4. *Let p be any prime, and $H: [r]^n \rightarrow \mathbb{C}$ be any function. There exists a family of functions $\{G_{u,v}: [p]^n \rightarrow \mathbb{C} : u, v \in [r]^n\}$ such that*

$$L_{1,k}(H) \leq \left(\frac{p(r-1)}{p-1} \right)^k \mathbf{E}_{\mathbf{u}, \mathbf{v} \sim [r]^n} L_{1,k}(G_{\mathbf{u}, \mathbf{v}}).$$

Moreover, suppose every prime divisor of m is at least p . If $H = e_m(f)$ for $f \in \text{Poly}(r, m, d)$, then each $G_{u,v}$ is the phase of some function in $\text{Poly}(p, m, d(p-1))$.

Lemma 3.2 follows from applying the lemma to $\text{Poly}(r, m, d)$.

Proof of Lemma 3.4. We first prove the case $n = 1$, then extend it to general n by applying the argument to one coordinate at a time. Let $h: [r] \rightarrow \mathbb{C}$. We construct functions $g_{a,b}: [p] \rightarrow \mathbb{C}$ for $a, b \in [r]$ with two properties: the zero coefficient $\widehat{h}(0)$ is related to only the zero coefficients $\widehat{g}_{a,b}(0)$, and each nonzero coefficient $\widehat{h}(\alpha)$ is related to only the nonzero coefficients of $g_{a,b}$.

For every $a, b \in [r]$, define the following two-point map $\Phi_{a,b}: [p] \rightarrow [r]$ by

$$\Phi_{a,b}(z) := a + (b - a)\mathbb{1}(z \neq 0) = \begin{cases} a & \text{if } z = 0 \\ b & \text{if } z \neq 0. \end{cases}$$

If every prime divisor of m is at least p , then **Lemma 2.3** shows that both $\mathbb{1}(z \neq 0)$ and $\Phi_{a,b}(z)$ are computable by polynomials over $\mathbb{Z}_m$ of degree at most $p - 1$. Let $g_{a,b}(z) := h \circ \Phi_{a,b}(z)$. Observe that

$$\widehat{g}_{a,b}(0) = \mathbf{E}_{\mathbf{z}}[g_{a,b}(\mathbf{z})] = \frac{1}{p}h(a) + (1 - \frac{1}{p})h(b).$$

And since $g_{a,b}(0) = \sum_{\beta} \widehat{g}_{a,b}(\beta)$, we have

$$\sum_{\beta \neq 0} \widehat{g}_{a,b}(\beta) = g_{a,b}(0) - \widehat{g}_{a,b}(0) = h(a) - \widehat{g}_{a,b}(0) = (1 - \frac{1}{p})(h(a) - h(b)).$$

Taking expectations, we have

$$\begin{aligned} \mathbf{E}_{\mathbf{a}, \mathbf{b}}[\widehat{g}_{\mathbf{a}, \mathbf{b}}(0)] &= \widehat{h}(0) \quad \text{and} \\ \mathbf{E}_{\mathbf{a}, \mathbf{b}}\left[\sum_{\beta \neq 0} \widehat{g}_{\mathbf{a}, \mathbf{b}}(\beta) \mathbf{e}_r(-\alpha \mathbf{a})\right] &= (1 - \frac{1}{p}) \mathbf{E}_{\mathbf{a}}[\mathbf{e}_r(-\alpha \mathbf{a}) h(\mathbf{a})] = (1 - \frac{1}{p}) \widehat{h}(\alpha) \quad \text{for } \alpha \neq 0. \end{aligned}$$

Therefore,

$$|\widehat{h}(\alpha)| \leq \begin{cases} \mathbf{E}_{\mathbf{a}, \mathbf{b}}|\widehat{g}_{\mathbf{a}, \mathbf{b}}(0)| & \text{if } \alpha = 0 \\ \frac{p}{p-1} \mathbf{E}_{\mathbf{a}, \mathbf{b}} \sum_{\beta \neq 0} |\widehat{g}_{\mathbf{a}, \mathbf{b}}(\beta)| & \text{if } \alpha \neq 0. \end{cases}$$

For general n and $u, v \in [r]^n$, consider

$$G_{u,v}(z_1, \dots, z_n) := H(\Phi_{u_1, v_1}(z_1), \dots, \Phi_{u_n, v_n}(z_n)).$$

Applying the one-coordinate argument to each coordinate of $G_{u,v}$, followed by the triangle inequality, gives

$$|\widehat{H}(S)| \leq \left(\frac{p}{p-1}\right)^{|S|} \mathbf{E}_{\mathbf{u}, \mathbf{v}} \sum_{T: \text{supp}(T) = \text{supp}(S)} |\widehat{G}_{\mathbf{u}, \mathbf{v}}(T)|.$$

Consequently,

$$L_{1,k}(H) \leq \left(\frac{p(r-1)}{p-1}\right)^k \mathbf{E}_{\mathbf{u}, \mathbf{v}} L_{1,k}(G_{\mathbf{u}, \mathbf{v}}).$$

The “moreover” part follows because each coordinate map Φ_{u_i, v_i} has degree at most $p - 1$ and thus each $G_{u,v}$ is the phase of a polynomial of degree at most $d(p - 1)$. $\square$

3.2 Modulus Reduction

Let $\{\mathcal{F}_d\}$ be any derivative-graded family of functions mapping $\mathbb{Z}_m^n \rightarrow \mathbb{T}$ that satisfies [Definition 2.4](#).

Lemma 3.5. *Let $m \geq 2$ and $B \geq 1$. Suppose $L_{1,k}(\mathcal{F}_{d-1}) \leq B^k$ for every k . Then $L_{1,k}(\mathcal{F}_d) \leq (1 + mB)^k$ for every k .*

We first introduce some notation.

Notation. For $S \in \mathbb{Z}_m^n$, let $|S|$ denote the number of nonzero coordinates in S . We also use $\emptyset$ to denote 0^n . For each pair $(S, T) \in (\mathbb{Z}_m^n)^2$, define

- $S \cap T \in \mathbb{Z}_m^n$ coordinatewise by $(S \cap T)_i = S_i$ if $S_i = T_i$ is nonzero, and 0 otherwise;
- $S \subseteq T$ if $S \cap T = S$.
- When S and T have disjoint supports, define $S \cup T := S + T$.

(Note that when $m = 2$, these notations coincide with the usual definitions of set intersection, inclusion, and disjoint union. Also note that for $m > 2$, $S \cap T = \emptyset$ does not imply disjoint supports, as a coordinate with $S_i \neq T_i$, both nonzero, contributes nothing to $S \cap T$.)

Proof of [Lemma 3.5](#). Let $f \in \mathcal{F}_d$, $F(x) := \exp(2\pi i \cdot f(x))$, and $W := L_{1,k}(F)$. If $W = 0$, there is nothing to prove. Otherwise, define a distribution $\mathcal{S}$ on vectors in $\mathbb{Z}_m^n$ of Hamming weight exactly k by

$$\Pr_{\mathbf{S} \sim \mathcal{S}}[\mathbf{S} = S] := \frac{|\widehat{F}(S)|}{W}.$$

Let $\mathbf{S}, \mathbf{T} \sim \mathcal{S}$ be independent. For each $I \in \mathbb{Z}_m^n$ where $|I| \leq k$, let $r := k - |I|$ and

$$J := \text{supp}(I), \quad \overline{J} := \{1, \dots, n\} \setminus J.$$

Define

$$\begin{aligned} \alpha_I &:= \Pr[\mathbf{S} \cap \mathbf{T} = I] = \frac{1}{W^2} \sum_{\substack{A, B \in \mathbb{Z}_m^{\overline{J}}: \\ |A|=|B|=r, A \cap B = \emptyset}} |\widehat{F}(A \cup I) \widehat{F}(B \cup I)| \\ \beta_I &:= \Pr[I \subseteq \mathbf{S}] = \frac{1}{W} \sum_{A \in \mathbb{Z}_m^{\overline{J}}: |A|=r} |\widehat{F}(A \cup I)|. \end{aligned}$$

We will show that

$$\alpha_I \leq \min\left\{\beta_I^2, \frac{(mB)^{2r}}{W^2}\right\} \leq \beta_I \cdot \frac{(mB)^r}{W}, \quad (2)$$

where the last inequality follows from $\min\{x^2, y^2\} \leq xy$.

As every pair (S, T) uniquely determines $S \cap T$, we have

$$\begin{aligned} 1 &= \sum_I \alpha_I \leq \frac{1}{W} \sum_I (mB)^{k-|I|} \beta_I \\ &= \frac{1}{W} \mathbf{E}_{S \sim \mathcal{S}} \sum_{I \subseteq S} (mB)^{k-|I|} = \frac{1}{W} \sum_{i=0}^k \binom{k}{i} (mB)^{k-i} = \frac{(1+mB)^k}{W}. \end{aligned}$$

Therefore $W \leq (1+mB)^k$.

It remains to prove the first inequality in (2). The first term is clear, since $S \cap T = I$ implies $I \subseteq S$ and $I \subseteq T$. To prove the second term, let $x \in \mathbb{Z}_m^{\bar{J}}$, and $z, z' \in \mathbb{Z}_m^J$. Define $F_z(x) := F(x, z)$. Note that $\widehat{F}(A \cup I) = \mathbf{E}_z[\widehat{F}_z(A) \overline{\chi_I(z)}]$. So

$$\alpha_I \leq \frac{1}{W^2} \mathbf{E}_{z, z'} \sum_{|S|=|T|=r: S \cap T = \emptyset} |\widehat{F}_z(S) \widehat{F}_{z'}(T)|.$$

We now show that for every fixed $z, z' \in \mathbb{Z}_m^J$,

$$\sum_{|S|=|T|=r: S \cap T = \emptyset} |\widehat{F}_z(S) \widehat{F}_{z'}(T)| \leq (mB)^{2r}.$$

We begin with the following generalization of [CHHL19, Claim 6.4]. For a function $F: \mathbb{Z}_m^n \rightarrow \mathbb{C}$ and any $y \in \mathbb{Z}_m^{\bar{J}}$, define $\partial_y F_{z, z'}(x) := F_z(x + y) \overline{F_{z'}(x)}$.

Claim 3.6. *Fix a partition $J \sqcup \bar{J}$ of $\{1, \dots, n\}$. For any $z, z' \in \mathbb{Z}_m^J$ and $S, T \in \mathbb{Z}_m^{\bar{J}}$, we have $|\widehat{F}_z(S) \widehat{F}_{z'}(T)| \leq \mathbf{E}_{y \sim \mathbb{Z}_m^{\bar{J}}} |\widehat{\partial_y F_{z, z'}}(S - T)|$.*

In our application, $\partial_y F_{z, z'}$ is the phase of a restriction of a derivative of f , so the assumed Fourier bound applies.

Proof. We have

$$\begin{aligned} |\widehat{F}_z(S) \widehat{F}_{z'}(T)| &= \left| \mathbf{E}_{\mathbf{x}, \mathbf{y}} [F_z(\mathbf{x}) \overline{F_{z'}(\mathbf{y})} \chi_S(\mathbf{x}) \chi_T(\mathbf{y})] \right| \\ &= \left| \mathbf{E}_{\mathbf{x}, \mathbf{y}} [F_z(\mathbf{x} + \mathbf{y}) \overline{F_{z'}(\mathbf{x})} \chi_S(\mathbf{x} + \mathbf{y}) \chi_{-T}(\mathbf{x})] \right| \\ &= \left| \mathbf{E}_{\mathbf{x}, \mathbf{y}} [\partial_y F_{z, z'}(\mathbf{x}) \overline{\chi_{S-T}(\mathbf{x})} \chi_S(\mathbf{y})] \right| \\ &\leq \mathbf{E}_{\mathbf{y}} |\widehat{\partial_y F_{z, z'}}(S - T)|. \quad \square \end{aligned}$$

Fix $D := S - T \in \mathbb{Z}_m^{\bar{J}}$, and set $u := 2r - |D|$. Since $S \cap T = \emptyset$ and $|S| = |T| = r$, there are exactly u coordinates i on which both S_i and T_i are nonzero and different. We first choose those u coordinates from $\text{supp}(D)$. Note that $0 \leq u \leq r$ as $|S| = |T| = r$. For each such coordinate i , there are $m - 2$ pairs of (S'_i, T'_i) such that $S'_i - T'_i = D_i$. From the remaining $2r - 2u$ nonzero coordinates of D , we can choose $r - u$ coordinates i to set $S'_i = D_i$ and

$T'_i = 0$, and we have $S'_i = 0$ and $T'_i = -D_i$ on the rest of the $r - u$ coordinates i . Thus, for any $|D| = 2r - u$, the number of such pairs (S', T') is

$$\binom{2r - u}{u} \binom{2r - 2u}{r - u} (m - 2)^u = \binom{2r - u}{r} \binom{r}{u} (m - 2)^u.$$

Summing over all $0 \leq u \leq r$, by [Claim 3.6](#),

$$\begin{aligned} \sum_{|S|=|T|=r: S \cap T = \emptyset} |\widehat{F_z}(S) \widehat{F_{z'}}(T)| &\leq \mathbf{E}_{\mathbf{y}} \sum_{|S|=|T|=r: S \cap T = \emptyset} |\widehat{\partial_{\mathbf{y}} F_{z, z'}}(S - T)| \\ &= \sum_{u=0}^r \binom{2r - u}{r} \binom{r}{u} (m - 2)^u \mathbf{E}_{\mathbf{y}} \sum_{|D|=2r-u} |\widehat{\partial_{\mathbf{y}} F_{z, z'}}(D)| \\ &\leq B^{2r} \cdot 2^{2r} \sum_{u=0}^r \binom{r}{u} (m - 2)^u \\ &\leq B^{2r} \cdot 2^{2r} (m - 1)^r \leq (mB)^{2r}, \end{aligned}$$

where the last inequality is because $4(m - 1) \leq m^2$ for $m \geq 2$. $\square$

Proof of [Lemma 3.1](#). Note that $\mathcal{F}_d := \{\frac{1}{m}f : f \in \text{Poly}(m, m, d)\}$ satisfies the properties defined in [Section 2](#). We have $L_{1,k}(\text{Poly}(m, m, 0)) \leq 1$ for every k . Applying [Lemma 3.5](#) d times gives

$$L_{1,k}(\text{Poly}(m, m, d)) \leq \left(\sum_{j=0}^d m^j \right)^k \leq (2m^d)^k. \quad \square$$

Proof of [Lemma 3.3](#). For $t \leq d$, define

$$\mathcal{F}_t := \left\{ \frac{1}{p^a} g(x) + \frac{p}{m} h(x) : g \in \text{Poly}(p, p^a, t) \text{ and } h \in \text{Junta}(p, \frac{m}{p}, d) \right\}.$$

Recall that by [Lemma 2.5](#) we have

$$\left\{ \frac{1}{m} f : f \in \text{Poly}(p, m, d) \right\} \subseteq \mathcal{F}_d \quad \text{and} \quad \mathcal{F}_0 \subseteq \left\{ \frac{k}{p^a} + \frac{p}{m} f : k \in \mathbb{Z}, f \in \text{Poly}(p, \frac{m}{p}, d(p - 1)) \right\}.$$

Suppose $L_{1,k}(\text{Poly}(p, \frac{m}{p}, d(p - 1))) \leq B^k$ for every k . Applying [Lemma 3.5](#) iteratively d times, we have

$$L_{1,k}(\text{Poly}(p, m, d)) \leq L_{1,k}(\mathcal{F}_d) \leq \left(\sum_{j=0}^{d-1} p^j + p^d B \right)^k \leq \left(\frac{p}{p-1} p^d B \right)^k. \quad \square$$

4 Pseudorandom Generators for Modular Polynomials

In this section, we construct our pseudorandom generators for polynomials modulo arbitrary moduli, proving [Theorem 1.1](#), restated below.

Theorem 1.1 (Restated). *Let p be the least prime divisor of m and $2 \leq r \leq p$. There is an explicit distribution X on $[r]^n$ samplable from a seed of length $2^{O(dm^2)} \log(n/\varepsilon)$, such that $d_{\text{TV}}(f(X), f([r]^n)) \leq \varepsilon$ for every degree- d polynomial f in n variables modulo m .*

We identify $[m]^n$ with $\mathbb{Z}_m^n$ whenever addition or Fourier analysis over $\mathbb{Z}_m$ is used. A distribution Z on $[m]^n$ is called λ -biased if

$$|\mathbf{E}[e_m(\langle \xi, Z \rangle)]| \leq \lambda \quad \text{for every nonzero } \xi \in \mathbb{Z}_m^n.$$

We use two known constructions: small-bias distributions over $\mathbb{Z}_m$ for arbitrary m [JM21], and pseudorandom generators for low-degree polynomials over finite fields [BV10, Lov09, Vio09].

Lemma 4.1 (Small-bias spaces over $\mathbb{Z}_m$ [JM21]). *For every m and every $\lambda \in (0, 1)$, there is an explicit λ -biased distribution on $[m]^n$ with seed length $O(\log n + \log(1/\lambda) + \log \log m)$.*

Theorem 4.2 ([Vio09]). *There are distributions X over $\mathbb{F}_p^n$ that can be explicitly sampled from a uniform seed of $O(d2^d \log(p/\varepsilon) + d \log n)$ bits such that for any polynomial f in n variables of degree d over $\mathbb{F}_p$ we have $d_{\text{TV}}(f(X), f(U_p)) \leq \varepsilon$.*

For vectors $x, y \in [r]^n$, write $x +_r y$ for coordinatewise addition modulo r . In [Sections 4.1](#) and [4.2](#), we will establish the following reductions:

Lemma 4.3 (Domain reduction). *Let $m \geq 2$, p be the least prime divisor of m , and $2 \leq r \leq p$. Set $B := (p-1)2^{m^2(p-1)^d}$, and let $t \geq \log_2(2B)$ be an integer. Let D be an $(\varepsilon/2)$ -biased distribution over $\mathbb{Z}_r^n$, and X be a distribution over $\mathbb{Z}_p^n$ that $\varepsilon/(2t)$ -fools $e_m(\text{Poly}(p, m, (p-1)d))$ and is independent of D . Let $X_1, \dots, X_t$ be independent copies of X , set $Y_i := X_i \bmod r$, and define*

$$\phi(Y_1, \dots, Y_t) := Y_1 +_r \dots +_r Y_t.$$

Then $D +_r \phi(Y_1, \dots, Y_t)$ ε -fools $e_m(\text{Poly}(r, m, d))$.

Lemma 4.4 (Modulus reduction). *Let m be composite and p be its least prime divisor. Write $m = p^a m'$ where $\gcd(p, m') = 1$. (Thus, p^a is the largest power of p dividing m .) Set $\lambda := (\varepsilon/9)^{2^d}$. Let $Y, S_1, \dots, S_d$ be independent distributions on $[p]^n$ where Y λ -fools $e_{m/p}(\text{Poly}(p, \frac{m}{p}, d(p-1)))$, and $S_1, \dots, S_d$ are λ -biased. Then*

$$X := Y +_p S_1 +_p \dots +_p S_d$$

ε -fools $e_m(\text{Poly}(p, m, d))$.

We now use these lemmas to prove [Theorem 1.1](#).

Proof of Theorem 1.1. Let $s(r, m, d, \varepsilon)$ denote the seed length required to fool phases of $\text{Poly}(r, m, d)$ with error ε . It suffices to show that $s(r, m, d, \varepsilon) \leq 2^{O(dm^2)} \log(n/\varepsilon)$ and then appeal to [Claim 2.1](#). The case $d = 0$ is trivial as the tests are constant. So henceforth we assume $d \geq 1$.

We will use the following reductions to upper bound $s(r, m, d, \varepsilon)$. Let $\lambda_D(\eta) := (\eta/9)^{2^D}$.

1. ([Lemma 4.1](#)) $s(m, m, 1, \varepsilon) \leq O(\log(\frac{mn}{\varepsilon}))$.
2. ([Theorem 4.2](#)) $s(p, p, d, \varepsilon) \leq O(d2^d \log(p/\varepsilon) + d \log n)$ where p is prime and $d \geq 1$.

3. (**Lemma 4.3**) $s(r, m, d, \varepsilon) \leq t \cdot s(p, m, d(p-1), \frac{\varepsilon}{2t}) + s(r, r, 1, \frac{\varepsilon}{2})$, where p is the least prime divisor of m , $r \leq p$, and $t = O(m^2 d(p-1))$.
4. (**Lemma 4.4**) $s(p, m, d, \varepsilon) \leq s(p, \frac{m}{p}, d(p-1), \lambda) + d \cdot s(p, p, 1, \lambda)$, where m is composite, p is the least prime divisor of m , and $\lambda = \lambda_d(\varepsilon)$.

Using these reductions, we prove the theorem by induction on the number of prime factors of m , counted with multiplicity, simultaneously for all r and d . Let $m \geq 2$ and p be its least prime divisor. We will show that for a sufficiently large constant C ,

$$s(r, m, d, \varepsilon) \leq 2^{Cdm^2} \log(n/\varepsilon) \quad \text{for every } 2 \leq r \leq p.$$

The bounds for $s(m, m, 1, \varepsilon)$ and $s(p, p, d, \varepsilon)$ follow directly from Items 1 and 2, respectively.

Suppose m is prime. Then by domain reduction (Item 3), Items 1 and 2, we have that for some $t = (md)^{O(1)}$,

$$\begin{aligned} s(r, m, d, \varepsilon) &\leq t \cdot s\left(m, m, d(p-1), \frac{\varepsilon}{2t}\right) + O\left(\log\left(\frac{nm}{\varepsilon}\right)\right) \\ &\leq (md)^{O(1)} \cdot 2^{d(m-1)} \cdot \log\left(\frac{n}{\varepsilon}\right) \leq 2^{Cdm^2} \log\left(\frac{n}{\varepsilon}\right). \end{aligned}$$

Suppose m is composite. We first apply domain reduction (Item 3) and then modulus reduction (Item 4). This reduces the modulus to $m' = m/p$, and the degree to $d' = d(p-1)^2$. Note that the input alphabet size p is at most the least prime divisor of m' , so we can apply the induction hypothesis. Thus, for $\lambda^* := \lambda_{d(p-1)}(\varepsilon/(2t)) = (\varepsilon/(18t))^{2^{d(p-1)}}$ and some $t = (md)^{O(1)}$,

$$\begin{aligned} s(r, m, d, \varepsilon) &\leq t \cdot s\left(p, m, d(p-1), \frac{\varepsilon}{2t}\right) + s\left(r, r, 1, \frac{\varepsilon}{2}\right) \\ &\leq t \cdot s\left(p, m', d', \lambda^*\right) + t \cdot d(p-1) \cdot s\left(p, p, 1, \lambda^*\right) + s\left(r, r, 1, \frac{\varepsilon}{2}\right) \\ &\leq (md)^{O(1)} \cdot 2^{d(p-1)} \cdot 2^{C d' m'^2} \cdot \log\left(\frac{n}{\varepsilon}\right). \end{aligned}$$

Note that $d' m'^2 \leq dm^2(1 - 1/p)^2 \leq dm^2 - dm$. So, the above is at most $2^{Cdm^2} \log(\frac{n}{\varepsilon})$. $\square$

4.1 Domain Reduction

In this subsection, we prove **Lemma 4.3**.

Proof of Lemma 4.3. Let $V_1, \dots, V_t$ be independent uniform vectors in $[p]^n$, and set $R_i := V_i \bmod r$. For $j = 0, \dots, t$, define the hybrid random variable

$$H_j := D +_r Y_1 +_r \dots +_r Y_j +_r R_{j+1} +_r \dots +_r R_t.$$

Fix $f \in \text{Poly}(r, m, d)$ and let $F := e_m(f)$. To compare H_{j-1} and H_j , factor out the sum $W_j \in [r]^n$ of all terms except the j -th, so that $H_{j-1} = R_j +_r W_j$ and $H_j = Y_j +_r W_j$. Condition on $W_j = w$. For each coordinate $i \in \{1, \dots, n\}$, the map

$$x_i \mapsto ((x_i \bmod r) +_r w_i) \quad x_i \in [p],$$

is a map from $[p]$ to $[r] \subset \mathbb{Z}_m$. By [Lemma 2.3](#), it is represented on $[p]$ by a (univariate) polynomial over $\mathbb{Z}_m$ of degree at most $p - 1$. Consequently, the map

$$x \mapsto f((x \bmod r) +_r w)$$

lies in $\text{Poly}(p, m, (p - 1)d)$. The fooling property of X therefore gives

$$|\mathbf{E}[F(H_j)] - \mathbf{E}[F(H_{j-1})]| \leq \mathbf{E}_{\mathbf{w} \sim W_j} |\mathbf{E}[F(Y_j +_r \mathbf{w})] - \mathbf{E}[F(R_j +_r \mathbf{w})]| \leq \frac{\varepsilon}{2t}.$$

Summing over the hybrids yields

$$|\mathbf{E}[F(H_t)] - \mathbf{E}[F(H_0)]| \leq \frac{\varepsilon}{2}.$$

It remains to compare $H_0 = D +_r R_1 +_r \dots +_r R_t$ with U_r . Let μ be the one-coordinate distribution obtained by reducing a uniform element of $[p]$ modulo r . For $a \in \mathbb{Z}_r$, its Fourier coefficient is

$$\widehat{\mu}(a) = \frac{1}{p} \sum_{x=0}^{p-1} e_r(ax).$$

Write $p = qr + s$, where $0 \leq s < r$. For any $r \leq p$, as $\sum_{x=0}^{r-1} e_r(ax) = 0$ for $a \neq 0$, we have

$$\rho := \max_{a \neq 0} |\widehat{\mu}(a)| \leq \frac{\min\{s, r - s\}}{p} \leq \frac{r}{2p} \leq 1/2.$$

The t -fold convolution μ^{*t} therefore satisfies

$$\max_{a \neq 0} |\widehat{\mu^{*t}}(a)| \leq \rho^t \leq \frac{1}{2B}.$$

By [Theorem 1.2](#), we have $L_{1,k}(F) \leq B^k$ for every k . Therefore, letting $\chi_S(x) := e_r(\langle S, x \rangle)$,

$$\begin{aligned} |\mathbf{E}[F(H_0)] - \mathbf{E}[F(U_r)]| &\leq \sum_{k=1}^n \sum_{|S|=k} |\widehat{F}(S)| |\mathbf{E}[\chi_S(D)]| |\mathbf{E}[\chi_S(R_1 +_r \dots +_r R_t)]| \\ &\leq \frac{\varepsilon}{2} \sum_{k=1}^n L_{1,k}(F) \cdot \left(\frac{1}{2B}\right)^k \\ &\leq \frac{\varepsilon}{2} \sum_{k=1}^n 2^{-k} \leq \frac{\varepsilon}{2}. \end{aligned}$$

Combining the two bounds proves the lemma. □

4.2 Modulus Reduction

In this subsection, we prove [Lemma 4.4](#). Throughout this subsection, all additions of input vectors are coordinatewise modulo p .

Proof of Lemma 4.4. For $t \leq d$, define

$$\mathcal{F}_t := \left\{ \frac{1}{p^a} g(x) + \frac{p}{m} h(x) : g \in \text{Poly}(p, p^a, t) \text{ and } h \in \text{Junta}(p, \frac{m}{p}, d) \right\}.$$

Recall that by Lemma 2.5 we have

$$\left\{ \frac{1}{m} f : f \in \text{Poly}(p, m, d) \right\} \subseteq \mathcal{F}_d \quad \text{and} \quad \mathcal{F}_0 \subseteq \left\{ \frac{k}{p^a} + \frac{p}{m} f : k \in \mathbb{Z}, f \in \text{Poly}(p, \frac{m}{p}, d(p-1)) \right\}.$$

Let $X_t := Y + S_1 + \dots + S_t$. We prove by induction on t that X_t fools $e(\mathcal{F}_t)$ with error $\varepsilon_t := 9\lambda^{2^{-t}}$. The base case $t = 0$ follows since Y fools $e(\mathcal{F}_0)$ with error $\lambda \leq \varepsilon_0 = 9\lambda$.

For the inductive step, fix $f \in \mathcal{F}_t$, and let $F(x) := \exp(2\pi i \cdot f(x))$. Set

$$\alpha := \mathbf{E}[F(U_p)], \quad \beta := \varepsilon_t/3.$$

By induction, X_{t-1} fools $e(\mathcal{F}_{t-1})$ with error $\varepsilon_{t-1} = \beta^2$. We split into two cases, depending on whether F is biased or not.

Case 1 - biased case: Assume $|\alpha| \geq \beta$. For $w \in [p]^n$, define

$$\partial_w F(x) := F(x + w) \overline{F(x)}.$$

Observe that for every fixed $w, z \in [p]^n$, the function

$$x \mapsto \partial_w F(x + z)$$

belongs to $e(\mathcal{F}_{t-1})$. Translation invariance of U_p gives, for every $x \in [p]^n$,

$$\mathbf{E}_{w \sim U_p} [\partial_w F(x)] = \mathbf{E}_{w \sim U_p} [F(x + w) \overline{F(x)}] = \alpha \overline{F(x)}.$$

Therefore,

$$\begin{aligned} |\mathbf{E}[F(X_t)] - \alpha| &= |\mathbf{E}[\overline{F(X_t)}] - \overline{\alpha}| = |\mathbf{E}[\overline{F(X_t)}] - \mathbf{E}[\overline{F(U_p)}]| \\ &\leq \frac{1}{|\alpha|} \mathbf{E}_{w \sim U_p, s \sim S_t} \left| \mathbf{E}_{z \sim X_{t-1}} [\partial_w F(z + s)] - \mathbf{E}_{z \sim U_p} [\partial_w F(z + s)] \right| \\ &\leq \frac{\varepsilon_{t-1}}{\beta} = \beta < \varepsilon_t. \end{aligned}$$

Case 2 - unbiased case: Assume $|\alpha| < \beta$. By Cauchy-Schwarz,

$$|\mathbf{E}[F(X_t)]|^2 \leq \mathbf{E}_{z \sim X_{t-1}} \left| \mathbf{E}_{s \sim S_t} [F(z + s)] \right|^2 = \mathbf{E}_{s_1, s_2 \sim S_t} \mathbf{E}_{z \sim X_{t-1}} [F(z + s_1) \overline{F(z + s_2)}].$$

For fixed s_1, s_2 , the inner function of z belongs to $e(\mathcal{F}_{t-1})$. Hence

$$|\mathbf{E}[F(X_t)]|^2 \leq \mathbf{E}_{z \sim U_p} \left| \mathbf{E}_{s \sim S_t} [F(z + s)] \right|^2 + \varepsilon_{t-1}.$$

Let

$$\chi_\xi(x) := e_p(\langle \xi, x \rangle), \quad \widehat{F}(\xi) := \mathbf{E}_{x \sim U_p} [F(x) \overline{\chi_\xi(x)}].$$

Parseval's identity and the λ -bias of S_t give

$$\begin{aligned} \mathbf{E}_{\mathbf{z} \sim U_p} \left| \mathbf{E}_{\mathbf{s} \sim S_t} [F(\mathbf{z} + \mathbf{s})] \right|^2 &= \sum_{\xi \in \mathbb{F}_p^n} |\widehat{F}(\xi)|^2 |\mathbf{E}[\chi_\xi(S_t)]|^2 \\ &\leq |\widehat{F}(0)|^2 + \lambda^2 \sum_{\xi \neq 0} |\widehat{F}(\xi)|^2 \\ &\leq \beta^2 + \beta^2, \end{aligned}$$

where we used $\widehat{F}(0) = \alpha$, $\sum_{\xi} |\widehat{F}(\xi)|^2 = \mathbf{E}[|F(U_p)|^2] = 1$, and $\lambda \leq \beta$. Thus

$$|\mathbf{E}[F(X_t)]| \leq \sqrt{3} \beta,$$

and consequently

$$|\mathbf{E}[F(X_t)] - \mathbf{E}[F(U_p)]| \leq (1 + \sqrt{3})\beta < \varepsilon_t.$$

This completes the proof. $\square$

5 Polynomial Simulation of Branching Programs over Nilpotent Groups

In this section, we prove [Theorem 1.4](#), restated below.

Theorem 1.4 (Restated). *Let G be a finite nilpotent group. There is a distribution X on $\{0, 1\}^n$ which can be explicitly sampled from a uniformly random seed of length $O_G(\log(n/\varepsilon))$, such that $d_{\text{TV}}(f(X), f(\{0, 1\}^n)) \leq \varepsilon$ for every $f: \{0, 1\}^n \rightarrow G$ that can be written as*

$$f(x_1, x_2, \dots, x_n) = \prod_{i=1}^L g_i^{y_i}$$

for some L , where each $g_i \in G$ and each y_i is a variable x_j .

Proof. If G is trivial, then there is nothing to prove. Henceforth we will assume $N := |G| > 1$. Identify G with the direct product of its Sylow subgroups

$$G = \prod_{p|N} G_p,$$

where each G_p is a p -group of order p^{k_p} for some integer $k_p \geq 0$. We first construct, for each $\alpha \in G$, a polynomial $Q_\alpha \in \text{Poly}(2, m, d)$ for $m = \prod_{p|N} p$ and some $d = O_G(1)$ such that

$$Q_\alpha(x) = 1 \bmod m \iff f(x) = \alpha.$$

We will use the following polynomial simulation for p -groups [\[LV25\]](#).¹

¹The proof in the conference version contains a minor mistake which has been corrected in the revised version on ECCC.

Lemma 5.1 ([LV25, Lemma 13]). *Let G be a group of order p^k . There is a one-to-one mapping $\pi: G \rightarrow \mathbb{F}_p^k$ and a constant C_G with the following property. For every L and $g_1, \dots, g_L \in G$, there exist polynomials $h_1, \dots, h_k \in \mathbb{F}_p[y_1, \dots, y_L]$ with $\deg(h_i) \leq C_G$ for every i such that for every $y \in \{0, 1\}^L$,*

$$\pi\left(\prod_{i=1}^L g_i^{y_i}\right) = (h_1(y), h_2(y), \dots, h_k(y)).$$

Fix a Sylow subgroup G_p of G . Let $\rho_p: G \rightarrow G_p$ be the projection map, and define

$$f_p := \rho_p \circ f, \quad \alpha_p := \rho_p(\alpha).$$

Write $y_i = x_{j_i}$ in the group program and define $y(x) := (x_{j_1}, \dots, x_{j_L})$. We apply [Lemma 5.1](#) to the fixed elements $\rho_p(g_1), \dots, \rho_p(g_L)$. Let $\pi_p: G_p \rightarrow \mathbb{F}_p^{k_p}$ denote the one-to-one mapping in [Lemma 5.1](#). We can write

$$\pi_p(f_p(x)) = (h_{p,1}(y(x)), \dots, h_{p,k_p}(y(x))).$$

Write $\pi_p(\alpha_p) = (\alpha_{p,1}, \dots, \alpha_{p,k_p})$. Observe that the polynomial

$$Q_{p,\alpha}(x) = \prod_{j=1}^{k_p} \left(1 - (h_{p,j}(y(x)) - \alpha_{p,j})^{p-1}\right)$$

satisfies $Q_{p,\alpha}(x) = \mathbb{1}(f_p(x) = \alpha_p)$ and has degree at most $k_p(p-1)C_{G_p}$ in $y(x)$, where C_{G_p} is the constant in [Lemma 5.1](#). So we can choose $d = \max_{p|N} k_p(p-1)C_{G_p} = O_G(1)$. Since each $y_i(x)$ is a coordinate of x , substitution does not increase degree.

Applying the Chinese Remainder Theorem to the polynomials $Q_{p,\alpha}$ coefficientwise gives

$$\begin{aligned} Q_\alpha(x) = 1 \pmod{m} &\iff Q_{p,\alpha}(x) = 1 \pmod{p} \text{ for every } p \mid N \\ &\iff f_p(x) = \alpha_p \text{ for every } p \mid N \\ &\iff f(x) = \alpha. \end{aligned}$$

We now apply [Theorem 1.1](#) with error $\eta := \varepsilon/|G|$, and let X be its output. For every $\alpha \in G$, we have

$$\begin{aligned} |\Pr[f(X) = \alpha] - \Pr[f(\{0, 1\}^n) = \alpha]| &= |\Pr[Q_\alpha(X) = 1] - \Pr[Q_\alpha(\{0, 1\}^n) = 1]| \\ &\leq d_{\text{TV}}(Q_\alpha(X), Q_\alpha(\{0, 1\}^n)) \leq \eta. \end{aligned}$$

Therefore, $d_{\text{TV}}(f(X), f(\{0, 1\}^n)) \leq \frac{1}{2} \sum_{\alpha \in G} \eta \leq \varepsilon$. Since both m and d only depend on G , the seed length is $O_G(\log(n/\varepsilon))$. □

Acknowledgments

We thank Prashanth Amireddy and Salil Vadhan for early discussions related to this work. Part of this work was done while most of the authors were at the Simons Institute for the Theory of Computing, attending the Fall 2026 Pseudorandomness & High-Dimensional Expansion program, which provided an ideal working environment.

AI methodology

For our PRGs for polynomials over prime moduli, we initially had a more complicated proof with weaker bounds, obtained without AI assistance. (In a nutshell, this proof simplified the inductive argument in [LMS13] by working with phases throughout, and obtained stronger parameters using a lemma from [CHH⁺20].) Later, ChatGPT found a simpler proof with improved bounds, which we refined. It also suggested the proof for the improved bounds for $L_{1,k}(\text{Poly}(2, 2, d))$ and its extension.

The results over composite moduli were obtained without AI assistance.

References

- [AW89] Miklós Ajtai and Avi Wigderson. Deterministic simulation of probabilistic constant depth circuits. *Adv. Comput. Res.*, 5:199–222, 1989. (cit. on page 4.)
- [BNS92] László Babai, Noam Nisan, and Mátio Szegedy. Multiparty protocols, pseudorandom generators for logspace, and time-space trade-offs. *J. Comput. System Sci.*, 45(2):204–232, 1992. (cit. on page 2.)
- [Bog05] Andrej Bogdanov. Pseudorandom generators for low degree polynomials. In *37th Annual Symposium on Theory of Computing (STOC)*, pages 21–30, 2005. (cit. on page 1.)
- [BV10] Andrej Bogdanov and Emanuele Viola. Pseudorandom bits for polynomials. *SIAM J. Comput.*, 39(6):2464–2486, 2010. (cit. on pages 1 and 16.)
- [CDG26] Gil Cohen, Dean Doron, and Noam Goldgraber. Optimal PRGs for low-degree polynomials over polynomial-size fields. Report TR26-016, Electron. Colloquium Comput. Complex. (ECCC), 2026. (cit. on page 1.)
- [CHH⁺20] Eshan Chattopadhyay, Pooya Hatami, Kaave Hosseini, Shachar Lovett, and David Zuckerman. XOR lemmas for resilient functions against polynomials. In *52nd Annual Symposium on Theory of Computing (STOC)*, pages 234–246, 2020. (cit. on page 22.)
- [CHHL19] Eshan Chattopadhyay, Pooya Hatami, Kaave Hosseini, and Shachar Lovett. Pseudorandom generators from polarizing random walks. *Theory Comput.*, 15(10):1–26, 2019. (cit. on pages 3, 4, 5, and 14.)
- [CHL⁺26] Eshan Chattopadhyay, Pooya Hatami, Chin Ho Lee, Shachar Lovett, Avishay Tal, and Emanuele Viola. Exponential correlation bounds for polynomials. Report TR26-208, Electron. Colloquium Comput. Complex. (ECCC), 2026. Revision 1, October 11, 2026. (cit. on page 3.)
- [Coh26] Gil Cohen. Fooling low-degree polynomials: PRGs with optimal dependence on input length and expander walks. Report TR26-243, Electron. Colloquium Comput. Complex. (ECCC), 2026. (cit. on page 3.)

- [DGV24] Ashish Dwivedi, Zeyu Guo, and Ben Lee Volk. Optimal pseudorandom generators for low-degree polynomials over moderately large fields. In *28th International Workshop on Randomization and Approximation Techniques in Computer Science (RANDOM)*, pages 44:1–44:19, 2024. (cit. on page 1.)
- [DV22] Harm Derksen and Emanuele Viola. Fooling polynomials using invariant theory. In *63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 399–406, 2022. (cit. on page 1.)
- [GMR⁺12] Parikshit Gopalan, Raghu Meka, Omer Reingold, Luca Trevisan, and Salil Vadhan. Better pseudorandom generators from milder pseudorandom restrictions. In *53rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 120–129, 2012. (cit. on page 4.)
- [HH24] Pooya Hatami and William Hoza. Paradigms for unconditional pseudorandom generators. *Found. Trends Theor. Comput. Sci.*, 16(1–2):1–210, 2024. (cit. on page 4.)
- [HLV18] Elad Haramaty, Chin Ho Lee, and Emanuele Viola. Bounded independence plus noise fools products. *SIAM J. Comput.*, 47(2):493–523, 2018. (cit. on page 4.)
- [JM21] Akhil Jalan and Dana Moshkovitz. Near-optimal Cayley expanders for abelian groups. In *41st Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*, pages 24:1–24:23, 2021. (cit. on page 16.)
- [Kra26] Gonen Krak. Optimal bitwise pseudorandom generators for modular sums modulo every fixed integer. Report TR26-130, Electron. Colloquium Comput. Complex. (ECCC), 2026. Revision 5. (cit. on pages 6 and 7.)
- [Lee19] Chin Ho Lee. Fourier bounds and pseudorandom generators for product tests. In *34th Computational Complexity Conference (CCC)*, pages 7:1–7:25, 2019. (cit. on page 3.)
- [LMS13] Shachar Lovett, Partha Mukhopadhyay, and Amir Shpilka. Pseudorandom generators for $CC^0[p]$ and the Fourier spectrum of low-degree polynomials over finite fields. *Comput. Complexity*, 22(4):679–725, 2013. (cit. on pages 2 and 22.)
- [Lov09] Shachar Lovett. Unconditional pseudorandom generators for low-degree polynomials. *Theory Comput.*, 5(3):69–82, 2009. (cit. on pages 1 and 16.)
- [LRTV09] Shachar Lovett, Omer Reingold, Luca Trevisan, and Salil Vadhan. Pseudorandom bit generators that fool modular sums. In *13th International Workshop on Randomization and Approximation Techniques in Computer Science (RANDOM)*, pages 615–630, 2009. (cit. on pages 2 and 4.)
- [LV20] Chin Ho Lee and Emanuele Viola. More on bounded independence plus noise: pseudorandom generators for read-once polynomials. *Theory Comput.*, 16(7):1–50, 2020. (cit. on page 4.)

- [LV25] Chin Ho Lee and Emanuele Viola. Pseudorandom bits for non-commutative programs. In *40th Computational Complexity Conference (CCC)*, pages 9:1–9:22, 2025. Corrected full version: ECCC TR25-071, Revision 2, July 17, 2026. (cit. on pages 3, 4, 7, 20, and 21.)
- [LVW93] Michael Luby, Boban Veličković, and Avi Wigderson. Deterministic approximate counting of depth-2 circuits. In *2nd Israel Symposium on Theory and Computing Systems (ISTCS)*, pages 18–24, 1993. (cit. on page 2.)
- [MZ09] Raghu Meka and David Zuckerman. Small-bias spaces for group products. In *13th International Workshop on Randomization and Approximation Techniques in Computer Science (RANDOM)*, pages 658–672, 2009. (cit. on pages 2 and 4.)
- [NN93] Joseph Naor and Moni Naor. Small-bias probability spaces: efficient constructions and applications. *SIAM J. Comput.*, 22(4):838–856, 1993. (cit. on page 2.)
- [Sri23] Srikanth Srinivasan. A robust version of Hegedűs’s lemma, with applications. *TheoretiCS*, 2:5:1–5:38, 2023. (cit. on page 4.)
- [SV10] Ronen Shaltiel and Emanuele Viola. Hardness amplification proofs require majority. *SIAM J. Comput.*, 39(7):3122–3154, 2010. (cit. on page 3.)
- [Vio07] Emanuele Viola. Pseudorandom bits for constant-depth circuits with few arbitrary symmetric gates. *SIAM J. Comput.*, 36(5):1387–1403, 2007. (cit. on page 2.)
- [Vio09] Emanuele Viola. The sum of d small-bias generators fools polynomials of degree d . *Comput. Complexity*, 18(2):209–217, 2009. (cit. on pages 1, 4, 6, 7, and 16.)
- [Vio26] Emanuele Viola. Mathematics of the impossible. Book manuscript, 2026. (cit. on page 1.)