Fractional Pseudorandom Generators from Any Fourier Level

Eshan Chattopadhyay
Department of Computer Science, Cornell University, Ithaca, NY, USA

Jason Gaitonde
Department of Computer Science, Cornell University, Ithaca, NY, USA

Chin Ho Lee
Department of Computer Science, Columbia University, New York City, NY, USA

Shachar Lovett
Department of Computer Science, University of California, San Diego, CA, USA

Abhishek Shetty
Department of Computer Science, University of California, Berkeley, CA, USA

Abstract

We prove new results on the polarizing random walk framework introduced in recent works of Chattopadhyay et al. [4, 6] that exploit L_1 Fourier tail bounds for classes of Boolean functions to construct pseudorandom generators (PRGs). We show that given a bound on the k-th level of the Fourier spectrum, one can construct a PRG with a seed length whose quality scales with k. This interpolates previous works, which either require Fourier bounds on all levels [4], or have polynomial dependence on the error parameter in the seed length [6], and thus answers an open question in [6]. As an example, we show that for polynomial error, Fourier bounds on the first $O(\log n)$ levels is sufficient to recover the seed length in [4], which requires bounds on the entire tail.

We obtain our results by an alternate analysis of fractional PRGs using Taylor's theorem and bounding the degree-k Lagrange remainder term using multilinearity and random restrictions. Interestingly, our analysis relies only on the *level-k unsigned Fourier sum*, which is potentially a much smaller quantity than the L_1 notion in previous works. By generalizing a connection established in [5], we give a new reduction from constructing PRGs to proving correlation bounds. Finally, using these improvements we show how to obtain a PRG for $\mathbb{F}_2$ polynomials with seed length close to the state-of-the-art construction due to Viola [26].

2012 ACM Subject Classification Theory of computation → Pseudorandomness and derandomization

Keywords and phrases Derandomization, pseudorandomness, pseudorandom generators, Fourier analysis

Digital Object Identifier 10.4230/LIPIcs.CCC.2021.10

Funding *Eshan Chattopadhyay*: Supported by NSF grant CCF-1849899 and NSF CAREER award 2045576.
Jason Gaitonde: Supported by NSF grant CCF-1408673 and AFOSR grant FA9550-19-1-0183.
Chin Ho Lee: Supported by the Croucher Foundation and the Simons Collaboration on Algorithms and Geometry.
Shachar Lovett: Supprted by NSF grants CCF-2006443 and DMS-1953928.
Abhishek Shetty: Supported by a JP Morgan Chase Faculty Fellowship.

Acknowledgements We thank the anonymous reviewers for their helpful comments and suggestions.

36th Computational Complexity Conference (CCC 2021).
Editor: Valentine Kabanets; Article No. 10; pp. 10:1–10:24
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

A central pursuit in complexity theory is to understand the need of randomness in efficient computation. Indeed there are important conjectures (such as $\mathbf{P} = \mathbf{BPP}$) in complexity theory which state that one can completely remove the use of randomness without losing much in efficiency. While we are quite far from proving such results, a rich line of work has focused on *derandomizing* simpler models of computation (see [25] for a survey of prior work on derandomization). A key tool for proving such derandomization results is through the notion of a *pseudorandom generator* defined as follows.

▶ **Definition 1.** *Let $\mathcal{F}$ be a class of n-variate Boolean functions. A* pseudorandom generator *(PRG) for $\mathcal{F}$ with error $\varepsilon > 0$ is a random variable $\mathbf{X} \in \{-1,1\}^n$ such that for all $f \in \mathcal{F}$,*

$$\left|\mathbb{E}_{\mathbf{X}}[f(\mathbf{X})] - \mathbb{E}_{\mathbf{U}_n}[f(\mathbf{U}_n)]\right| \le \varepsilon,$$

where $\mathbf{U}_n$ is the uniform distribution on $\{-1,1\}^n$. We also say that $\mathbf{X}$ fools *$\mathcal{F}$ with error ε. If $\mathbf{X} = G(\mathbf{U}_s)$ for some explicit function $G : \{-1,1\}^s \to \{-1,1\}^n$, then $\mathbf{X}$ has* seed length *s.*

There is a long line of research on explicit constructions of PRGs for various classes of Boolean functions in the literature and it is well beyond our scope to survey prior work here. We focus on a recent line of works initiated by Chattopadhyay et al. [4, 6] that provide a framework for constructing pseudorandom generators for any Boolean function classes that exhibit *Fourier tail bounds* (we will define and discuss this in more details in the next subsection; see Section 2.1 for a brief introduction to Fourier analysis of Boolean functions). This provides a unified PRG for several well-studied function classes such as small-depth circuits, low-sensitivity functions, and read-once branching programs that exhibit such Fourier tails.

We now briefly discuss this new framework, and then in Section 1.2 we present our new results, which significantly generalize this approach.

1.1 The Polarizing Random Walk Framework

The *polarizing random walk* framework was introduced by Chattopadhyay, Hatami, Hosseini, and Lovett [4]. The authors showed that for any class of n-variate Boolean functions that is closed under restrictions, one can flexibly construct pseudorandom generators via the following local-to-global principle: it suffices to construct *fractional pseudorandom generators (fractional PRGs)*, a notion that generalizes PRGs to allow the random variable $\mathbf{X}$ (in Definition 1) to be supported on the solid cube $[-1,1]^n$ instead of $\{-1,1\}^n$, while still requiring that $\mathbf{X}$ fools (the multilinear extension) of each Boolean function in the class. Ideally, the variance of each coordinate of $\mathbf{X}$ should be as large as possible. Towards this, we define a fractional PRG $\mathbf{X}$ to be p-noticeable if the variance in each of its coordinates is least p (See Definition 13 for a formal definition of a fractional PRG).

To obtain a genuine pseudorandom generator from a fractional PRG, the authors give a random walk gadget that composes together independent copies of the fractional PRG in a random walk that polarizes $\mathbf{X}$ quickly to take values from the Boolean hypercube $\{-1,1\}^n$. The analysis for how the error accumulates in this process relies on interpreting the intermediate points of $\mathbf{X}$ in this random walk as an average of *random restrictions* of the original Boolean function. As the fractional PRG locally fools the class by definition, this analysis shows that the random walk does not incur much additional error at each intermediate step and the rapid polarization shows that it does not take too many steps. Taken together, these two facts imply that the final random variable (supported on $\{-1,1\}^n$) successfully fools the class.

Through this construction, the design of pseudorandom generators reduces to the easier task of designing fractional pseudorandom generators. It is easier as such random variables need not be Boolean-valued. The authors further construct such fractional pseudorandom generators for any class of functions satisfying *Fourier tail bounds*, that is, every function in the class is such that the L_1 Fourier mass at each level $1 \leq k \leq n$ is at most b^k for some fixed $b \geq 1$. For error ε, their fractional pseudorandom generators have seed length $O(\log \log n + \log(1/\varepsilon))$ and variance $\Theta(b^{-2})$ in each coordinate. Combining this fractional pseudorandom generator with their random walk gadget yields a pseudorandom generator with seed length $b^2 \cdot \mathrm{polylog}(n/\varepsilon)$ for *any* class with such Fourier tail bounds.

As a result, if one can show that a function class admits nontrivial Fourier tail bounds (and is closed under restriction), then the construction in [4] immediately implies a pseudorandom generator for this class. Some examples of Boolean functions that exhibit such tail bounds include $\mathbf{AC}^0$ circuits with the parameter $b = \mathrm{poly}(\log n)$ [13, 23], constant width read-once branching programs with $b = \mathrm{poly}(\log n)$ [7], s-sensitive functions with $b = O(s)$ [11, 10], and product tests [12]. Using these tail bounds, [4] immediately gave PRGs for these function classes. It was also conjectured in [4] that the class of n-variate degree-d polynomials over $\mathbb{F}_2$ satisfy such tail bounds. We discuss this in more detail in Section 1.2.

A natural question is whether the complete control on the entire Fourier tail of a class is necessary to obtain a PRG in this framework. In the subsequent work by Chattopadhyay, Hatami, Lovett, and Tal [6], the authors show how to construct fractional pseudorandom generators using different pseudorandom primitives whose seed length depends on just the *second Fourier level* of the class. They construct their fractional PRGs by derandomizing the celebrated work of Raz and Tal [18], which establishes an oracle separation of **BQP** and **PH**. Raz and Tal show that classes of multilinear functions with small level-two Fourier mass cannot significantly distinguish between a suitable variant of the Forrelation distribution and the uniform distribution.[1] However, this construction incurs exponentially worse dependence on the error parameter in each fractional step to sample sufficiently good approximations to Gaussian random variables. The final seed length given by this construction has the form $O((b^2/\varepsilon)^{2+o(1)}\mathrm{polylog}(n))$, where b^2 is the level-two Fourier mass of the class. This yields exponentially worse dependence on the error compared to the generator of [4], as well as quadratically worse dependence on the level-two mass (though without assumptions on the rest of the Fourier levels).

1.2 Our Contribution

In this paper, we address several open questions in this framework by leveraging a novel connection between polarizing random walk and the classical theory of polynomial approximation. Given these prior works, a very natural question (also explicitly asked in [6]) is whether it is possible to interpolate between these previous constructions by assuming Fourier bounds on an intermediate level. Concretely, can this framework still succeed if one has Fourier control at just level k? If the class further has such Fourier bounds up to and including level k, can one interpolate between the seed lengths of [4] and [6]? Given Fourier bounds from level 1 up to level k, what range of error $\varepsilon > 0$ can the resulting PRG tolerate while maintaining polylogarithmic dependence on $1/\varepsilon$ in the seed length (or equivalently, given a desired error $\varepsilon > 0$, how many levels of Fourier bounds are sufficient to ensure that the seed length remains polylogarithmic in $1/\varepsilon$)?

[1] It turns out that this fact can be interpreted via Itô's Lemma, which shows that the local behavior of a smooth function of Brownian motion is essentially determined by the first two derivatives [28].

Moreover, it was previously not known whether L_1 control of Fourier tails is really necessary for this framework to yield effective PRGs, or whether weaker Fourier quantities would suffice. To this end, define

$$L_{1,k}(f) \triangleq \sum_{S \subseteq [n]:|S|=k} |\hat{f}(S)|$$

to be the *level-k L_1 Fourier mass* of f, and

$$M_k(f) \triangleq \max_{\mathbf{x}\in[-1,1]^n} \left| \sum_{S\subseteq[n]:|S|=k} \hat{f}(S)\mathbf{x}^S \right| = \max_{\mathbf{x}\in\{-1,1\}^n} \left| \sum_{S\subseteq[n]:|S|=k} \hat{f}(S)\mathbf{x}^S \right|.$$

to be the *level-k absolute Fourier sum* of f. For a function class $\mathcal{F}$, we define $L_{1,k}(\mathcal{F})$ and $M_k(\mathcal{F})$ as the maximum of $L_{1,k}(f)$ and $M_k(f)$ taken over $f \in \mathcal{F}$. The recent work by Chattopadhyay, Hatami, Hosseini, Lovett, and Zuckerman [5] considers the weaker quantity of the level-two *unsigned Fourier sum*, defined as the absolute value of the sum of the Fourier coefficients rather than the sum of their absolute values that is considered in [4, 6]. The authors show that the problem of bounding the level-two unsigned Fourier sum corresponds to the problem of bounding the covariance of the function class and the XOR of shifted majority functions. For a class that is closed under negations of the variables, the level-two unsigned Fourier sum is precisely the quantity $M_2(\mathcal{F})$. In particular, using this connection to this weaker object, the authors explicitly ask whether bounding the weaker Fourier quantity $M_2(\mathcal{F})$ (or more generally, $M_k(\mathcal{F})$) yields pseudorandom generators.

In this work, we positively resolve all of these questions. To do so, we establish novel connections between the polarizing random walk framework and the classical theory of polynomial approximations of Boolean functions. We show that the seed length of a fractional PRG for a given class of functions $\mathcal{F}$ is intimately connected to the uniform error of low-degree approximations of functions on *subcubes* of the form $[-c, c]^n$ for some $c < 1$.

Our main technical result provides an upper bound on this quantity in terms of $M_k(\mathcal{F})$ for every function f in a class $\mathcal{F}$ that is closed under restrictions. For any multilinear polynomial $f : \{-1, 1\}^n \to \mathbb{R}$, define $f_{\geq k}$ to be component of f with monomials of degree at least k. Then our main result asserts the following bound:

▶ **Theorem 2.** *Let $f \in \mathcal{F}$ with $\mathcal{F}$ closed under restrictions. Then for all $c \in (0, 1)$, we have*

$$\max_{\mathbf{x}\in[-c,c]^n} |f_{\geq k}(\mathbf{x})| \leq \left(\frac{c}{1-c}\right)^k M_k(\mathcal{F}).$$

For intuition, recall that by Parseval's identity in Fourier analysis the low-degree Fourier expansion of any Boolean function f is provably the best ℓ_2-approximator on $\{-1, 1\}^n$. Conversely, from elementary analysis, one can show that the best uniform (i.e. ℓ_∞) low-degree approximators of f converge, coefficient-by-coefficient, to the low-degree expansion of f as the domain converges to $\mathbf{0}$. Our main result shows that one can strongly quantify the ℓ_∞ error of the low-degree approximator of Boolean functions on subcubes so long as c is not too close to 1 (compare this bound to when f has degree exactly k).

We complement this result with a corresponding lower bound on the best attainable uniform error for *any* low-degree approximation on these subcubes that will be comparable for sufficiently small values of c (see Theorem 23). These results combined together imply that the low-order expansion of a Boolean function is a reasonable uniform approximation for small domains. Note that the properties of low-degree approximations on subcubes with $c \ll 1$ can be quite different than for $c = 1$; for instance the PARITY function on n bits is well-known to be inapproximable on $\{-1, 1\}^n$ to constant error unless the approximating polynomial has degree $\Omega(n)$, but is trivially approximable for any c bounded away from 1.

From this main result, we can positively resolve the above open questions in the polarizing random walk framework as a nearly immediate corollary. To do so, we provide a new analysis of the fractional pseudorandom generator of [4] that views fractional pseudorandom generators as fooling a low-degree part of a function on $[-c,c]^n$ for some $c<1$, where the high-degree part has small ℓ_∞ norm on $[-c,c]^n$. Recall that the seed length of the final generator depends on the variance of the constituent fractional generator; the connection to the above result is that for a given error ε, the largest subcube on which the above approximation holds can be lower-bounded using just the weaker $M_k(\mathcal{F})$ quantity. Leveraging this insight, our main result in the polarizing random walk framework is the following analysis of a fractional pseudorandom generator:

▶ **Theorem 3.** *Let $\mathcal{F}$ be any class of n-variate Boolean functions that is closed under restrictions. Suppose $M_k(\mathcal{F}) \leq b^k$ for some $b \geq 1$ and $k \geq 1$. Then for any $\varepsilon > 0$, there exists an explicit $\Omega(\varepsilon^{2/k}/b^2)$-noticeable fractional PRG for $\mathcal{F}$ with error ε and seed length $O(k \cdot \log n)$.*[2]

Further, if it holds that $L_{1,i}(\mathcal{F}) \leq b^i$ for all $1 \leq i < k$, then the seed length can be improved to $O(\log\log n + \log k + \log(1/\varepsilon))$.

Using the fractional pseudorandom generator from Theorem 3, we obtain the following consequences almost immediately from the random walk gadget of [4] (see Theorem 14):

1. **Pseudorandom Generators from Fourier Bounds at Level k**: From our fractional pseudorandom generator, we show that the random walk framework yields nontrivial pseudorandom generators assuming Fourier bounds *just at* level k of the associated class, with improvements if we assume bounds from level 1 *up to* level k. The informal statement is the following:

   ▶ **Theorem 4.** *Let $\mathcal{F}$ be any class of n-variate Boolean functions that is closed under restrictions. Suppose that $\mathcal{F}$ satisfies $M_k(\mathcal{F}) \leq b^k$ for some $b \geq 1$ and $k \geq 3$. Then there exists an explicit pseudorandom generator for $\mathcal{F}$ for error ε with seed length $k \cdot b^{2+4/(k-2)}\mathrm{polylog}(n/\varepsilon)/\varepsilon^{2/(k-2)}$. The seed length can be improved if $L_{1,i}(\mathcal{F}) \leq b^i$ for all levels $i \leq k$.*

   See Theorem 27 for the precise statement. One immediate consequence is that if one has a non-trivial bound on $M_3(\mathcal{F})$, then the seed length of our PRG has the same dependence on the error ε as the one in [6]. Further, given $M_4(\mathcal{F}) \leq b^4$, one obtains better seed length than [6]; in particular it has quadratically better dependence on $1/\varepsilon$ in the seed length (as well as polylogarithmic factors in n/ε). More generally, given an appropriate Fourier bound of b^k on just some level $k \leq \mathrm{polylog}(n)$, one obtains a pseudorandom generator with error ε with seed length $O(b^{2+4/(k-2)}\mathrm{polylog}(n/\varepsilon)/\varepsilon^{2/(k-2)})$.

   We note that the fractional PRG from Theorem 3 cannot be converted into a PRG for $k = 1, 2$. Informally, this is because of the following reason: the number of steps one needs to take in the random walk gadget of [4] (with each step using an independent copy of the fractional PRG) scales roughly with the variance of the fractional PRG, and the error adds up in each step. As is clear from Theorem 3, for the variance of the fractional PRG to scale sublinearly with the error, one requires $k > 2$. See Remark 28 for more discussion.

[2] We remark that at this level of generality, this linear dependence on k is essentially necessary. Indeed, any Boolean function on n-variables has L_1 level-n mass at most 1, but one cannot hope to generically fool all Boolean functions simultaneously without using n bits.

2. **Pseudorandom Generators with Polylogarithmic Error Dependence from Up-to-level-k Bounds**: A simple corollary of our fractional pseudorandom generator is that one can recover the polylogarithmic dependence on $1/\varepsilon$ from [4] if $\varepsilon \geq b \cdot \log n \cdot 2^{-O(k)}$ and we have Fourier bounds *up to* level k.

   ▶ **Corollary 5.** *Let $\mathcal{F}$ be any class of n-variate Boolean functions that is closed under restrictions. Suppose that for some level $k \geq 3$ and $b \geq 1$, we have $M_k(\mathcal{F}) \leq b^k$ and $L_{1,i}(\mathcal{F}) \leq b^i$ for $i < k$. Then, for any $\varepsilon \geq b \cdot \log n \cdot 2^{-O(k)}$, there exists an explicit pseudorandom generator for $\mathcal{F}$ with error ε and seed length $O(b^2 \text{polylog}(n/\varepsilon))$.*

   This actually subsumes the analysis of [4] without requiring anything on the full Fourier tail, and addresses an open question of [6] asking how many levels of Fourier bounds one needs control of to regain polylogarithmic dependence on ε. In particular, if one requires error $\varepsilon = 1/\text{poly}(n)$, then it suffices to have Fourier bounds up to level $\Theta(\log n)$ to get the same dependence.

We view this work as a proof of concept that it is indeed possible to interpolate between the two extremes of [4, 6] in the polarizing random walk framework and obtain better results using weakened Fourier assumptions. We prove Theorem 3 in Section 4, from which Theorem 4 and Corollary 5 follow without much difficulty using the existing random walk gadget of [4].

Note that for some Boolean classes of great interest such as the class of low-degree $\mathbb{F}_2$-polynomials, Fourier tail bounds as required by [4] are not yet known and thus Theorem 3 allows us to leverage potentially much weaker bounds proved in [4] to construct a PRG with polylogarithmic dependence on n/ε in the seed length (see Theorem 6). This almost matches the best known PRG due to Viola [26]. In particular, we show the following:

▶ **Theorem 6.** *Let $\mathcal{F}$ be the class of degree-d polynomials over $\mathbb{F}_2$ on n variables. Then there exists an explicit pseudorandom generator for $\mathcal{F}$ with error ε and seed length $2^{O(d)}\text{polylog}(n/\varepsilon)$.*

We present the proof of Theorem 6 in Section 5. While this result does not quite match the current state-of-the-art PRG for this class due to Viola [26] (and therefore fails to give anything nontrivial for $d = \Omega(\log n)$), we view this as a conceptual contribution that the random walk framework can yield an explicit pseudorandom generator with seed length that is polylogarithmic in n/ε, which was not known from previous works [4, 6]. As we discuss below, the results in [4, 6] do not give a PRG for the class of $\mathbb{F}_2$-polynomials with polylogarithmic error dependence using known Fourier tail bounds.

As a concrete application of this approach which would dramatically improve the state-of-the-art PRGs for $\mathbb{F}_2$-polynomials, both [4] and [6] conjecture Fourier bounds on the L_1 mass of the class of degree-d $\mathbb{F}_2$ polynomials. The former conjectures that this class satisfies a tail bound of the form c_d^k for some constant c_d at all levels $1 \leq k \leq n$ (so as to apply their approach), while the latter conjectures just that the level-two L_1 mass is $O(d^2)$. While neither conjecture seems close to being resolved, our work shows that one can instead prove bounds for the smaller quantities $M_k(\mathcal{F})$ for any $k \geq 3$. If one could prove such bounds of the form $(\text{poly}(d, \log n))^k$ for some level $k = \Omega(1)$, or even more optimistically, for some $k = \Omega(\log n)$, this would immediately imply a breakthrough pseudorandom generator for $\mathbf{AC^0}[\oplus]$ using the results of Razborov [19] and Smolensky [21, 22] (see the discussion in [6]).

To our knowledge, our application of $M_k(\mathcal{F})$ bounds is new to the pseudorandomness literature. There are several advantages to proving $M_k(\mathcal{F})$ bounds over $L_{1,k}(\mathcal{F})$ bounds. For one, from the definition we clearly have $M_k(\mathcal{F}) \leq L_{1,k}(\mathcal{F})$ for any class $\mathcal{F}$. This improvement alone potentially gives smaller seed length for any class. From an analytical perspective, we believe that the quantity $M_k(\mathcal{F})$ is easier to estimate. Specifically, for a class $\mathcal{F}$ that is closed under negation of input variables, $M_k(\mathcal{F})$ is precisely an *unsigned Fourier sum* and can be bounded via the recent connections established by Chattopadhyay et al. [5], which reduces $M_2(\mathcal{F})$ bounds to proving correlation bounds against certain resilient functions. We straightforwardly generalize their reduction to $M_k(\mathcal{F})$ bounds in Section 6.

1.3 Overview of Our Approach

To prove Theorem 2, we rely on Taylor's theorem, as well as multilinearity and the random restriction trick of [4]. Recall that Taylor's theorem, when applied to a sufficiently smooth function $h\colon [-1,1] \to \mathbb{R}$, asserts that the Taylor expansion at 0 can be expressed in terms of its first $(k-1)$-th order derivatives at 0 along with a Lagrange error term that depends on its k-th order derivatives at some intermediate point in our domain. In doing so, the higher-order components of the function "collapse" down to the k-th order term. While Taylor's theorem has been extensively applied in the construction of pseudorandom generators, often in tandem with *invariance principles*, we somewhat counterintuitively apply it to the *multilinear expansion of the Boolean functions* themselves.

To apply Taylor's theorem here, we consider one-dimensional restrictions of (the multilinear extension) of a Boolean function $f\colon \{-1,1\}^n \to \{-1,1\}$. While the full Taylor expansion of a polynomial is trivially the same polynomial, the Lagrange error term eliminates the dependence on the high order Fourier coefficients (corresponding to the terms of degree $> k$). Moreover, the low-order terms of the Taylor expansion of f at 0 are precisely the original low-degree part of its Fourier expansion. However, the Lagrange error term requires the derivatives to be evaluated at a point away from 0. While the derivatives of f at a nonzero point are related to the *biased* Fourier coefficients of f, it is not clear how to estimate these quantities. To overcome this difficulty, recall that we are interested in bounds on $|f_{\geq k}(\mathbf{x})|$ for $\mathbf{x} \in \{-c,c\}^n$ where $c < 1$. In Lemma 22, we show that by "recentering" $\mathbf{x}$ using the random restriction technique of [4], we can write the error term as an average of the k-th order derivatives *at 0* of some random restrictions of our original function f, up to a multiplicative factor depending on c.[3] We can then apply multilinearity to bound these error terms using $M_k(\mathcal{F})$ to obtain Theorem 17.

While Theorem 17 shows that the low-order Taylor expansion of a Boolean function is a decent *uniform* approximator on subcubes $[-c,c]^n$ for some sufficiently small c that depends on the class $\mathcal{F}$, it is natural to wonder if one can obtain a better low-order approximation. Using our upper bound along with Chebyshev polynomials on the univariate restrictions, we give a lower bound showing that no low-order approximator can give significantly smaller error over $[-c,c]^n$ for any c less than some quantity depending on the ratio $M_k(\mathcal{F})/M_{k+1}(\mathcal{F})$ for some k. This quantifies the intuition that the low-degree Fourier expansion is a near optimal uniform approximator of f over small enough neighborhoods of $\mathbf{0}$. These arguments are formally carried out in Section 3.

To prove our results in the polarizing random walk framework, we rely on an alternate, simple analysis of fractional pseudorandom generators. The original analysis in [4] assumes control of $L_{1,k}(\mathcal{F})$ at all levels of the Fourier spectrum. We now explain how these assumptions can be weakened using Theorem 17. Consider a candidate fractional PRG $\mathbf{X} \in [-1,1]^n$. We first decompose the multilinear (Fourier) expansion of $f \in \mathcal{F}$ in the same manner as [4]:

$$\left|\mathbb{E}_{\mathbf{X}}[f(\mathbf{X})] - \mathbb{E}_{\mathbf{U}}[f(\mathbf{U})]\right| \leq \underbrace{\sum_{i=1}^{k-1} \sum_{S \subseteq [n]:|S|=i} \left|\hat{f}(S)\right|\left|\mathbb{E}_{\mathbf{X}}[\mathbf{X}^S]\right|}_{\text{low-order terms}} + \underbrace{\left|\mathbb{E}_{\mathbf{X}}[f_{\geq k}(\mathbf{X})]\right|}_{\text{high-order term}}. \tag{1}$$

[3] We note that similar ideas for the $k = 1$ case also appeared in [1] (attributed to Avishay Tal).

[4] requires bounding $L_{1,\ell}(\mathcal{F})$ for all $\ell \geq k$ to give a uniform bound on the high-order term. Using Theorem 17, we can obtain small error in the high-order term so long as we choose $\mathbf{X} \in [-c, c]^n$ for sufficiently small c depending on ε and $M_k(\mathcal{F})$. To handle the low-order terms, we consider two cases: if we further have $L_{1,\ell}(\mathcal{F})$ bounds for $\ell < k$, then we may choose $\mathbf{X}$ to be a scaled $(k-1)$-wise δ-biased distribution to nearly fool each of the low-order terms as in [4]. Otherwise, we may choose $\mathbf{X}$ to be a scaled $(k-1)$-wise independent distribution to incur zero error from the low-order terms. Note that the latter pseudorandom primitives are more expensive in terms of seed length. Finally, to obtain pseudorandom generators, we then simply apply the random walk gadget of [4] to our fractional PRGs as a blackbox. We refer the reader to Section 4 for formal proofs of the ideas in this section.

We immediately leverage this newfound flexibility to construct new pseudorandom generators for $\mathbb{F}_2$-polynomials of degree $d = O(\log n)$. We do this using known $L_{1,k}(\mathcal{F})$ bounds derived in [4]. Previously these bounds were not sufficient to give PRGs with polylogarithmic error dependence as their analysis of fractional PRGs either required control of the entire Fourier tail or could not leverage higher Fourier levels, but they can be employed here due to our more flexible analysis. This result is given in Section 5. Finally, we show how $M_k(\mathcal{F})$ bounds can be obtained using correlation bounds with shifted majority functions in Section 6. This is done by straightforwardly generalizing the analysis of [5], which shows how such correlation bounds can be used to bound the bulk of the terms in the definition of $M_k(\mathcal{F})$.

1.4 Other Related Work

To our knowledge, our use of $M_k(\mathcal{F})$ bounds is new to the derandomization literature. As mentioned earlier, the stronger and better-known $L_{1,k}(\mathcal{F})$ notion has been extensively studied in recent years. In addition to derandomization, a recent line of work [24, 3, 20] has used $L_{1,k}$ bounds for decision trees to obtain an optimal separation of quantum and classical query complexity. Among these works, the work of Bansal and Sinha [3] generalizes the results of Raz and Tal [18] by considering a k-generalization of their Forrelation distribution and bounding the distinguishing advantage of any function with small $L_{1,\ell}$ bounds for $\ell = 1, \ldots, k$. Much as how the results of Chattopadhyay et al. [6] derandomize the result of Raz and Tal, we believe that their construction can be derandomized for pseudorandomness purposes, but appears to give significantly worse seed length, nor obtains bounds in terms of $M_k(\mathcal{F})$. A related work by Girish, Raz, and Zhan [9] establishes a similar result with a different generalization of the Forrelation distribution, but we do not know how to use their construction for pseudorandom generators.

The relationship between $M_k(\mathcal{F})$ and $L_{1,k}(\mathcal{F})$ has been of intense study in the mathematics literature due to renewed interest in *Bohnenblust–Hille* inequalities (see, for instance, the breakthrough work of Defant, Frerick, Ortega-Cerdà, Ounaïes, and Seip [8]). The optimal constant $C_{n,k}$ satisfying $L_{1,k}(f) \leq C_{n,k} M_k(f)$ for any polynomial $f \colon \mathbb{C}^n \to \mathbb{C}$ is known as the *Sidon constant.* It is known that $C_{n,k}$ is, up to small exponential factors in k, proportional to roughly $n^{\frac{k-1}{2}}$, and its tightness is witnessed by a random function with high probability. The quantity $M_k(\mathcal{F})$ also has applications in other areas in theoretical computer science, such as quantum information theory (see for instance the survey of Montanaro [14]) and Boolean function analysis [2].

Subsequent to our work, Viola [27] observed that $M_k(\mathcal{F})$ bounds imply correlation bounds between $\mathcal{F}$ and an explicit function.

2 Preliminaries

As in [4] and [6], we study PRGs for classes $\mathcal{F}$ of n-variate Boolean functions that are closed under restriction (that is, fixing any subset of the input variables of a function in the class yields a function that remains in the class).

2.1 Fourier Analysis

We briefly recall basic Fourier analysis: any Boolean function $f : \{-1,1\}^n \to \{-1,1\}$ admits a unique multilinear expansion, also known as the *Fourier expansion*, given by

$$f(\mathbf{x}) = \sum_{S\subseteq[n]} \hat{f}(S)\mathbf{x}^S, \tag{2}$$

where we write $\mathbf{x}^S \triangleq \prod_{i\in S} x_i$. The Fourier coefficient $\hat{f}(S)$ is given by

$$\hat{f}(S) = \mathbb{E}_{\mathbf{X}\sim\{-1,1\}^n}[f(\mathbf{X})\mathbf{X}^S].$$

For more on Fourier analysis of Boolean functions, see the excellent book by O'Donnell [16]. One may thus extend the domain of f to $[-1,1]^n$, where $f(\mathbf{x})$ for arbitrary $\mathbf{x}$ is evaluated according to the expression in Equation (2). Note that in this case, $f(\mathbf{0}) = \hat{f}(\emptyset) = \mathbb{E}_{\mathbf{U}_n}[f(\mathbf{U}_n)]$. One of the main parameters of interest from the Fourier expansion for this framework is the following:

▶ **Definition 7.** *The* level-k mass of a Boolean function f *is*

$$L_{1,k}(f) \triangleq \sum_{S\subseteq[n]:|S|=k} |\hat{f}(S)|,$$

and the level-k mass of a class $\mathcal{F}$ *is* $L_{1,k}(\mathcal{F}) \triangleq \max_{f\in\mathcal{F}} L_{1,k}(f)$*.*

In this work, we will show how to construct PRGs whose seed length depends on the following, smaller quantity:

▶ **Definition 8.** *For any multilinear polynomial $f : \mathbb{R}^n \to \mathbb{R}$ given by $f(\mathbf{x}) = \sum_{S\subseteq[n]} \hat{f}(S)\mathbf{x}^S$, define the level-$k$ part by*

$$f_k(\mathbf{x}) \triangleq \sum_{S\subseteq[n]:|S|=k} \hat{f}(S)\mathbf{x}^S,$$

and further define $f_{<k}(\mathbf{x}) \triangleq \sum_{i=0}^{k-1} f_i(\mathbf{x})$ and $f_{\geq k}(\mathbf{x}) \triangleq \sum_{i=k}^{n} f_i(\mathbf{x})$. Then we define the level-k absolute Fourier sum of f *by*

$$M_k(f) \triangleq \max_{\mathbf{x}\in[-1,1]^n} \left| \sum_{S\subseteq[n]:|S|=k} \hat{f}(S)\mathbf{x}^S \right| = \max_{\mathbf{x}\in\{-1,1\}^n} \left| \sum_{S\subseteq[n]:|S|=k} \hat{f}(S)\mathbf{x}^S \right|$$

and analogously define $M_k(\mathcal{F}) \triangleq \max_{f\in\mathcal{F}} M_k(f)$ for a class $\mathcal{F}$.

Note that the equality arises by multilinearity, and clearly we have $M_k(f) \leq L_{1,k}(f)$ by the triangle inequality. Without loss of generality, we may further assume that our class is closed under flipping the image, i.e. we may suppose that $f \in \mathcal{F}$ if and only if $-f \in \mathcal{F}$; this transformation does not change either $L_{1,k}(f)$ or $M_k(f)$, and therefore the same bound on the class still holds when completing it to include all such functions. If this is the case, we get the more striking identity:

▶ **Lemma 9.** *Suppose that $\mathcal{F}$ is closed under negation of variables and that $f \in \mathcal{F}$ implies $-f \in \mathcal{F}$. Then*

$$M_k(\mathcal{F}) = \max_{f\in\mathcal{F}} \sum_{S\subseteq[n]:|S|=k} \hat{f}(S) = \max_{f\in\mathcal{F}} f_k(\mathbf{1}).$$

To see why this holds, simply note that if $(f,\mathbf{z}) \in \mathcal{F} \times \{-1,1\}^n$ is a maximizer in the definition of $M_k(\mathcal{F})$ (where we may now assume that the sign is positive), then by replacing the function $f(\mathbf{x})$ with $g(\mathbf{x}) = f(\mathbf{x}\circ\mathbf{z})$, where $\circ$ denotes componentwise multiplication, we have

$$M_k(\mathcal{F}) = \left|\sum_{S\subseteq[n]:|S|=k} \hat{f}(S)\mathbf{z}^S\right| = \sum_{S\subseteq[n]:|S|=k} \hat{g}(S) = \max_{h\in\mathcal{F}} \sum_{S\subseteq[n]:|S|=k} \hat{h}(S).$$

In particular, it suffices to bound the *unsigned level-k Fourier sum* of such a class.

Lastly, we require the following notion:

▶ **Definition 10.** *Let $\mathcal{F}$ be a class of n-variate multilinear polynomials that is closed under restrictions. Define* $\mathrm{conv}(\mathcal{F})$ *as the convex closure of $\mathcal{F}$,*

$$\mathrm{conv}(\mathcal{F}) \triangleq \left\{ \sum_{f\in\mathcal{F}} \lambda_f f \;\middle|\; \sum_{f\in\mathcal{F}} \lambda_f = 1, \lambda_f \geq 0\ \forall f \in \mathcal{F} \right\}.$$

We briefly note the following two elementary facts: first, by the assumption that $\mathcal{F}$ is closed under restrictions, the same is true of $\mathrm{conv}(\mathcal{F})$. The second is the following simple claim:

▶ **Lemma 11.** *For any class $\mathcal{F}$ of Boolean functions, $M_k(\mathcal{F}) = M_k(\mathrm{conv}(\mathcal{F}))$.*

Proof. One direction is obvious: as $\mathcal{F} \subseteq \mathrm{conv}\mathcal{F}$, clearly $M_k(\mathcal{F}) \leq M_k(\mathrm{conv}(\mathcal{F}))$. In the other direction, let $g = \sum_{f\in\mathcal{F}} \lambda_f f$ be an arbitrary element of $\mathrm{conv}(\mathcal{F})$, where $\lambda_f \geq 0$ and $\sum_{f\in\mathcal{F}} \lambda_f = 1$. Then

$$\begin{aligned}
M_k(g) &= \max_{\mathbf{x}\in\{-1,1\}^n} \left|\sum_{S\subseteq[n]:|S|=k} \widehat{g}(S)\mathbf{x}^S\right| \\
&= \max_{\mathbf{x}\in\{-1,1\}^n} \left|\sum_{S\subseteq[n]:|S|=k} \Big(\sum_{f\in\mathcal{F}} \lambda_f \widehat{f}(S)\Big)\mathbf{x}^S\right| \\
&\leq \sum_{f\in\mathcal{F}} \lambda_f \max_{\mathbf{x}\in\{-1,1\}^n} \left|\sum_{S\subseteq[n]:|S|=k} \widehat{f}(S)\mathbf{x}^S\right| \\
&\leq \max_{f\in\mathcal{F}} M_k(f).
\end{aligned}$$

The reverse inequality immediately follows. ◀

2.2 (Fractional) Pseudorandom Generators

We now recall the (well-known) definition of a pseudorandom generator, as well as the generalization of a fractional pseudorandom generator as introduced by [4]:

▶ **Definition 12.** *Let $\mathcal{F}$ be a class of n-variate Boolean functions. Then a* pseudorandom generator *(PRG) for $\mathcal{F}$ with error $\varepsilon > 0$ is a random variable $\mathbf{X} \in \{-1,1\}^n$ such that for all $f \in \mathcal{F}$,*

$$|\mathbb{E}_{\mathbf{X}}[f(\mathbf{X})] - \mathbb{E}_{\mathbf{U}_n}[f(\mathbf{U}_n)]| \leq \varepsilon,$$

where $\mathbf{U}_n$ is the uniform distribution on $\{-1,1\}^n$. If $\mathbf{X} = G(\mathbf{U}_s)$ for some explicit function $G : \{-1,1\}^s \to \{-1,1\}^n$, then $\mathbf{X}$ has seed length *s.*

▶ **Definition 13.** *A* fractional pseudorandom generator *(fractional PRG) for $\mathcal{F}$ with error $\varepsilon > 0$ is a random variable $\mathbf{X} \in [-1,1]^n$ such that for all $f \in \mathcal{F}$ (identifying f with its multilinear expansion)*

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

*where the definition of seed length is the same. A fractional PRG is p-*noticeable *if for each $i \in [n]$, $\mathbb{E}[\mathbf{X}_i^2] \geq p$.*

We now state the main results of [4] and [6] that show how to construct PRGs from suitably combining noticeable fractional PRGs. This is done by the following *amplification theorem*, which roughly composes fractional random variables into a random walk inside the Boolean hypercube:

▶ **Theorem 14.** *Suppose $\mathcal{F}$ is class of n-variate Boolean functions that is closed under restrictions, and that $\mathbf{X}$ is an explicit p-noticeable fractional PRG 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 $O(\varepsilon\log(n/\varepsilon)/p)$.*

Using this result, [4] proved the following theorem that exploits strong L_1 control of each Fourier level:

▶ **Theorem 15.** *Let $\mathcal{F}$ be any class of n-variate Boolean functions that is closed under restrictions. Suppose that $L_{1,k}(\mathcal{F}) \leq b^k$ for some $b \geq 1$ and all $1 \leq k \leq n$. Then for any $\varepsilon > 0$, there exists an explicit PRG for $\mathcal{F}$ with error ε and seed length $b^2 \cdot \mathrm{polylog}(n/\varepsilon)$.*

This is achieved by constructing a fractional PRG that is a scaled version of a $\log(1/\varepsilon)$-wise nearly unbiased distribution. As we will be analyzing a similar fractional PRG, we defer the details to next section. To lessen the requisite assumptions on the Fourier spectrum, Chattopadhyay et al. [6] derandomize a construction of Raz and Tal [18] to prove the following result that requires only level-two control, albeit at a cost of exponentially worse dependence on the error ε, and quadratically worse dependence on the level-two mass:

▶ **Theorem 16.** *Let $\mathcal{F}$ be any class of n-variate Boolean functions that is closed under restrictions. Suppose that $L_{1,2}(\mathcal{F}) \leq b^2$ for some $b \geq 1$. Then for any $\varepsilon > 0$, there exists an explicit PRG for $\mathcal{F}$ with error ε and seed length $O((b^2/\varepsilon)^{2+o(1)}\mathrm{polylog}(n))$.*

3 Low-Degree Polynomial Approximations on Subcubes

Throughout this section, we assume that $\mathcal{F}$ is a class of n-variate Boolean functions closed under restrictions. As mentioned above, the main result from which we derive our improvements in constructing pseudorandom generators is essentially a statement about low-degree polynomial approximations on subcubes $[-c,c]^n$ for $c < 1$. We remark that this setting is equivalent to approximating *noisy* versions $T_c f$ on $[-1,1]^n$, where T_ρ is the ρ-noise operator. This is because for any $\mathbf{y} \in [-c,c]^n$, we can write $\mathbf{y} = c\mathbf{x}$ for some $\mathbf{x} \in [-1,1]^n$ and

$$f(\mathbf{y}) = f(c\mathbf{x}) = \sum_{S\subseteq[n]} \hat{f}(S)(c\mathbf{x})^S = \sum_{S\subseteq[n]} c^{|S|}\hat{f}(S)\mathbf{x}^S = T_c f(\mathbf{x}).$$

In general, given any $k \leq n$, $c \geq 0$, and any $f \in \mathcal{F}$, let $\varepsilon_{c,k}(f)$ be defined by

$$\varepsilon_{c,k}(f) \triangleq \inf_{g:\deg(g)<k}\ \max_{\mathbf{x}\in[-c,c]^n} |f(\mathbf{x}) - g(\mathbf{x})|, \tag{3}$$

and extend the definition to function classes by

$$\varepsilon_{c,k}(\mathcal{F}) \triangleq \max_{f\in\mathcal{F}} \varepsilon_{c,k}(f).$$

Now, given $\varepsilon > 0$, $k \leq n$, and the class $\mathcal{F}$, define $c_k(\mathcal{F},\varepsilon)$ by

$$c_k(\varepsilon,\mathcal{F}) \triangleq \max\{c \geq 0 : \varepsilon_{c,k}(\mathcal{F}) \leq \varepsilon\}.$$

In words, $c_k(\varepsilon,\mathcal{F})$ measures how small a hypercube we must take to ensure that for every function in our class, there exists a degree-$(k-1)$ approximating polynomial that agrees with f up to a uniform ε error on the subcube $[-c,c]^n$; by multilinearity, it actually suffices that this holds at the extreme points $\{-c,c\}^n$. Note that Equation (3) can be formulated as a linear program and its optimal solution is the best low-degree ℓ_∞-approximation to f.

The main technical claim in this section is that we bound $c_k\left(\varepsilon,\mathcal{F}\right)$ in terms of $M_k(\mathcal{F})$. Specifically, we show that for any class $\mathcal{F}$ that is closed under restrictions, truncating the Fourier expansion of a function $f \in \mathcal{F}$ to its first $(k-1)$ levels serves as a good approximation to f on a sufficiently small hypercube around the origin.

▶ **Theorem 17.** *Let $f \in \mathcal{F}$ that is closed under restrictions. Then for all $c \in (0,1)$, we have*

$$\max_{\mathbf{x}\in[-c,c]^n} |f_{\geq k}(\mathbf{x})| \leq \left(\frac{c}{1-c}\right)^k M_k(\mathcal{F}).$$

In particular, it follows that

$$\varepsilon_{c,k}(\mathcal{F}) \leq \left(\frac{c}{1-c}\right)^k M_k(\mathcal{F}).$$

From Theorem 17, one immediately obtains a lower bound on $c_k(\varepsilon,\mathcal{F})$:

▶ **Corollary 18.** *For any class $\mathcal{F}$ that is closed under restrictions, and any $\varepsilon > 0$ and $k \leq n$,*

$$c_k(\varepsilon,\mathcal{F}) = \Omega\left(\left(\frac{\varepsilon}{M_k(\mathcal{F})}\right)^{1/k}\right)$$

Proof. Observe that by setting $c = \Omega\left(\left(\frac{\varepsilon}{M_k(\mathcal{F})}\right)^{1/k}\right)$ in Theorem 17, the right side is bounded by ε. Because $f_{\geq k} = f - f_{<k}$ and $f_{<k}$ has degree strictly less than k, it follows immediately from the definition of $c_k(\varepsilon,\mathcal{F})$ that $c_k(\varepsilon,\mathcal{F})$ is at least c. ◀

We now return to the proof of Theorem 17. To prove this result, we require the following intermediate claims. The first simply shows that we may always bound the contribution of the level-k part of any function in $\mathcal{F}$ by simply rescaling the argument:

▶ **Lemma 19.** *Let $f \in \operatorname{conv}(\mathcal{F})$. Then, for all $c \in (0,1)$ and $\mathbf{x} \in [-c,c]^n$, we have*

$$|f_k(\mathbf{x})| \leq c^k M_k(\mathcal{F}).$$

Proof. Observe that $c^{-1}\mathbf{x} \in [-1,1]^n$ by assumption, and by homogeneity of f_k as a polynomial, we have

$$|f_k(\mathbf{x})| = c^k|f_k(c^{-1}\mathbf{x})| \leq c^k M_k(\operatorname{conv}(\mathcal{F})) = c^k M_k(\mathcal{F}).$$ ◀

The next simple yet powerful claim shows that one can "recenter" functions in $\mathcal{F}$ and they remain in $\mathrm{conv}(\mathcal{F})$ (and therefore, enjoy the same Fourier bounds). This random restriction technique is a key tool in [4].

▶ **Lemma 20.** *Let $f \in \mathrm{conv}(\mathcal{F})$, $\mathbf{a} \in [-1,1]^n$ and $\mathbf{b} \in [0,1]$ such that $|a_i| + b_i \leq 1$ for all $i \in [n]$. Define $\tilde{f}$ by $\tilde{f}(\mathbf{x}) = f(\mathbf{a} + \mathbf{b} \circ \mathbf{x})$, where $\circ$ denotes componentwise multiplication. Then, $\tilde{f} \in \mathrm{conv}(\mathcal{F})$.*

Proof. Given $\mathbf{a}$ and $\mathbf{b}$, define a distribution D_i on $Z_i = \{-1, 1, x_i\}$ where x_i is treated as formal variable, such that $\mathbb{E}_{y_i \sim D_i}[y_i] = a_i + b_i x_i$; note that this is possibly by the assumption that $|a_i| + b_i \leq 1$. Let $D = \prod_i D_i$ be the product distribution of the D_i. For any $\mathbf{z} \in \prod_i Z_i$, define $f_{\mathbf{z}}(\mathbf{x})$ as the function obtained by setting $x_i = z_i$ for each i; in particular, each variable gets set to ±1 or remains a formal variable. By our assumption on the closure of $\mathcal{F}$, we clearly have $f_{\mathbf{z}} \in \mathcal{F}$ for any $\mathbf{z}$. By multilinearity and independence of the product distribution, we have $f(\mathbf{a} + \mathbf{b} \circ \mathbf{x}) = \mathbb{E}_{\mathbf{z} \sim D}[f_{\mathbf{z}}(\mathbf{x})]$. Thus $\tilde{f} \in \mathrm{conv}(\mathcal{F})$. ◀

As mentioned before, our approach will be to bound the higher-order terms of the Fourier expansion at the fractional points of the fractional PRG via the error term that arises in Taylor's theorem. Denote by $h^{(k)}$ the k-th derivative of any C^k function $h : \mathbb{R} \to \mathbb{R}$. We then have the following claim:

▶ **Lemma 21.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be multilinear and let $\mathbf{x} \in \mathbb{R}^n$. Define $g : \mathbb{R} \to \mathbb{R}$ by $g(t) = f(t\mathbf{x})$. Then,*

$$g^{(k)}(0) = k! \cdot f_k(\mathbf{x}).$$

Proof. From the definition, it follows that

$$g(t) = \sum_{S \subseteq [n]} t^{|S|} \hat{f}(S) \mathbf{x}^S.$$

Differentiating g with respect to t, we get

$$g^{(k)}(t) = \sum_{S \subseteq [n] : |S| \geq k} \left(\prod_{i=0}^{k-1} (|S| - i) \right) t^{|S|-k} \hat{f}(S) \mathbf{x}^S.$$

Setting $t = 0$ eliminates all of the monomials with $|S| > k$, giving us the required bound. ◀

The last intermediate result we require connects the function defined in the previous part with our assumed Fourier bounds:

▶ **Lemma 22.** *Let $f \in \mathrm{conv}(\mathcal{F})$, $c \in (0,1)$ and $\mathbf{x} \in [-c,c]^n$. Define g as in Lemma 21. Then,*

$$\max_{s \in [0,1]} \left| g^{(k)}(s) \right| \leq \left(\frac{c}{1-c} \right)^k \cdot k! \cdot M_k(\mathcal{F})$$

Proof. Fix $s \in [0,1]$ and let $\lambda = 1 - c \in [0,1]$. Define the auxiliary function $\tilde{f}(\mathbf{y}) = f(s\mathbf{x} + \lambda \mathbf{y})$. Writing $\mathbf{a} = s\mathbf{x}$ and $\mathbf{b} = (\lambda, \ldots, \lambda)$, we clearly have $s|x_i| + \lambda \leq 1$, so we may apply Lemma 20 to see that $\tilde{f} \in \mathrm{conv}(\mathcal{F})$. Now writing $\tilde{g}(t) = \tilde{f}(t\mathbf{x}) = f(s\mathbf{x} + \lambda t\mathbf{x})$, we also have $\tilde{g}(t) = g(s + t\lambda)$. By the chain rule, differentiating both sides k times and then setting $t = 0$, we have

$$\lambda^k g^{(k)}(s) = \tilde{g}^{(k)}(0).$$

On the other hand, by Lemma 21, we have $\tilde{g}^{(k)}(0) = k! \cdot \tilde{f}_k(\mathbf{x})$, and as $\tilde{f} \in \mathrm{conv}(\mathcal{F})$ by Lemma 20, we conclude using Lemma 19 that

$$\left|g^{(k)}(s)\right| = \left|\frac{\tilde{g}^{(k)}(0)}{\lambda^k}\right| \le \left(\frac{c}{1-c}\right)^k \cdot k! \cdot M_k(\mathcal{F}). \qquad ◀$$

With these intermediate claims taken care of, we may now put them together to obtain Theorem 17.

Proof of Theorem 17. The second statement follows immediately from the first by setting $g = f_{<k}$ for any given f, and noticing that $f - g = f_{\ge k}$. Therefore, we focus on the first statement.

Let $f \in \mathcal{F}, \mathbf{x} \in [-c, c]^n$ and define $g(t) = f(t\mathbf{x})$. Then, by Taylor expanding g about $t = 0$ and evaluating g at $t = 1$, we have

$$g(1) = \sum_{i<k} \frac{g^{(i)}(0)}{i!} + R_k, \tag{4}$$

where R_k is the error term and is given in Lagrange form by

$$R_k = \frac{g^{(k)}(s)}{k!}$$

for some $s \in (0, 1)$. By Lemma 21, we easily see that the first term in the right hand side of Equation (4) is precisely $f_{<k}(\mathbf{x})$, and as $g(1) = f(\mathbf{x})$, we clearly then must have $R_k = f_{\ge k}(\mathbf{x})$. Therefore, by Lemma 22, we obtain

$$|f_{\ge k}(\mathbf{x})| = \left|\frac{g^{(k)}(s)}{k!}\right| \le \left(\frac{c}{1-c}\right)^k M_k(\mathcal{F}),$$

as desired. ◀

3.1 Lower Bounds via Chebyshev Polynomials

In this subsection, we show that our bounds on the uniform error of any low-degree polynomial approximator are essentially tight for a reasonable range of $c < 1$. Recall that Theorem 17 shows that the low-degree Fourier expansion is an excellent approximator to the original function for c small enough; we now show that this bound cannot be significantly improved for a reasonable range of c using *any* approximator. Our main result of this section is the following converse:

▶ **Theorem 23.** *Let $\mathcal{F}$ be any class of n-variate multilinear functions that are closed under restrictions. Then for any $c \le \min\left(\frac{1}{3}, 3^{-k}\frac{M_k(\mathcal{F})}{M_{k+1}(\mathcal{F})}\right)$, we have*

$$\varepsilon_{c,k}(\mathcal{F}) \ge \left(\frac{c}{2}\right)^k M_k(\mathcal{F}).$$

Recall that on the interval $[-1, 1]$, the Chebyshev polynomials give the minimum ℓ_∞ norm among all polynomials with same leading coefficient in magnitude:

▶ **Fact 24** (Theorem 1.5.4 of [17]). *If a polynomial $f : \mathbb{R} \to \mathbb{R}$ is monic of degree n, then $\max_{x\in[-1,1]} |f(x)| \ge 2^{-n+1}$, with equality if and only if $f = T_n$, the normalized n-th Chebyshev polynomial.*

Proof of Theorem 23. Let $(f, \mathbf{x})$ attain the maximum in the definition of $M_k(\mathcal{F})$, namely

$$M_k(\mathcal{F}) = \left| \sum_{S\subseteq[n]:|S|=k} \widehat{f}(S)\mathbf{x}^S \right|.$$

First, note that the claim is trivial if every function in $\mathcal{F}$ is of degree at most k, because then $f_{\geq k}$ is a homogeneous polynomial of degree k and this lower bound is trivial. Under this assumption, $M_{k+1}(\mathcal{F}) > 0$. Fix $c \in (0,1)$ and let $p : [-1,1]^n \to \mathbf{R}$ be any multilinear polynomial of degree strictly less than k. Define the univariate function $g : [-1,1] \to \mathbb{R}$ by

$$g(t) = f(tc\mathbf{x}) - p(tc\mathbf{x}).$$

By taking the Fourier expansion of f, it is easy to see that the coefficient of t^ℓ for $\ell \geq k$ is precisely

$$c^\ell \sum_{S\subseteq[n]:|S|=\ell} \widehat{f}(S)\mathbf{x}^S,$$

so that the coefficient of t^k is equal to $c^k M_k(\mathcal{F})$ in magnitude. We then have

$$\begin{aligned}\sup_{\mathbf{z}\in[-c,c]^n} |f(\mathbf{z}) - p(\mathbf{z})| &\geq \max_{\mathbf{z}\in[-c\mathbf{x},c\mathbf{x}]} |f(\mathbf{z}) - p(\mathbf{z})| \\ &= \sup_{t\in[-1,1]} |g(t)| \\ &\geq \sup_{t\in[-1,1]} |g_{\leq k}(t)| - \sup_{t\in[-1,1]} |g_{\geq k+1}(t)|.\end{aligned}$$

By Fact 24, the first term is at least $c^k M_k(\mathcal{F})/2^{k-1}$. On the other hand, the second term can be bounded using Theorem 17 by

$$\sup_{t\in[-1,1]} |g_{\geq k+1}(t)| \leq \left(\frac{c}{1-c}\right)^{k+1} M_{k+1}(\mathcal{F}).$$

Therefore, we obtain

$$\sup_{\mathbf{z}\in[-c,c]^n} |f(\mathbf{z}) - p(\mathbf{z})| \geq 2\left(\frac{c}{2}\right)^k M_k(\mathcal{F}) - \left(\frac{c}{1-c}\right)^{k+1} M_{k+1}(\mathcal{F}).$$

It is straightforward to verify that for $c \leq \min\left(1/3, 3^{-k}\frac{M_k(\mathcal{F})}{M_{k+1}(\mathcal{F})}\right)$, the second term is bounded by half of the first. Because p was an arbitrary low-degree multilinear polynomial, the claim follows. ◀

4 From Polynomial Approximations to PRGs

4.1 From Polynomial Approximations to Fractional PRGs

From Theorem 17, we now show how the construction of fractional PRGs from level-k bounds reduces to efficient polynomial approximation on "large" subcubes.

▶ **Theorem 25.** *Let $\mathcal{F}$ be closed under restrictions. Then there exists a fractional PRG for $\mathcal{F}$ with error ε and seed length $O(k\log n)$ that is $(c_k(\varepsilon/2,\mathcal{F}))^2$-noticeable. In particular, if $M_k(\mathcal{F}) = b^k$, there exists such a fractional PRG that is $\Omega\left(\frac{\varepsilon^{2/k}}{b^2}\right)$-noticeable with seed length $O(k\log n)$.*

Proof. The second statement follows immediately from the first using Corollary 18, so we focus on the first statement.

Fix $f \in \mathcal{F}$, $\varepsilon > 0$, and let $\mathbf{X}$ be a $(k-1)$-wise independent random variable over $\{-1,1\}^n$ such that $|\mathbf{X}_i| = c \le 1/2$ for all $i \in [n]$ for some $c > 0$ we specify momentarily. It is well-known that $\mathbf{X}$ can be sampled efficiently with seed length $O(k \log n)$ [25]. By definition of $c := c_k(\varepsilon/2, \mathcal{F})$, there exists a degree-$(k-1)$ multilinear polynomial $\widetilde{f}$ which ε-approximates f on the subcube $[-c,c]^n$, i.e.

$$\max_{y \in [-c,c]^n} \left| f(y) - \widetilde{f}(y) \right| \le \varepsilon/2. \tag{5}$$

Then we have, via the Fourier expansion of f,

$$\begin{aligned} \left|\mathbb{E}_{\mathbf{X}}[f(\mathbf{X})] - f(\mathbf{0})\right| &\le \frac{\varepsilon}{2} + \left|\mathbb{E}_{\mathbf{X}}[f(\mathbf{X})] - \widetilde{f}(\mathbf{0})\right| \\ &= \frac{\varepsilon}{2} + \left|\mathbb{E}_{\mathbf{X}}\left[f(\mathbf{X}) - \widetilde{f}(\mathbf{X})\right]\right| \\ &\le \frac{\varepsilon}{2} + \mathbb{E}_{\mathbf{X}}\left[\left|f(\mathbf{X}) - \widetilde{f}(\mathbf{X})\right|\right] \\ &\le \varepsilon. \end{aligned}$$

The first inequality applies Equation (5) at the point $\mathbf{x} = \mathbf{0}$, and the second uses the fact that $\mathbf{X}$ is $(k-1)$-wise independent and $\widetilde{f}$ has degree at most $k-1$. The final inequality holds because of (5) and the fact that $\mathbf{X} \in [-c,c]^n$. Therefore, $\mathbf{X}$ satisfies the definition of a fractional PRG. Note that by construction, $\mathbf{X}$ is c^2-noticeable since it takes values in $\{-c,c\}^n$. ◀

Although it does not fit so neatly in this approximation framework, one can essentially recover the improved seed length of [4] (which we recall assumes $L_{1,i}(\mathcal{F})$ bounds for $i = 1,\dots,n$) if one further has $L_{1,i}(\mathcal{F})$ bounds just up to level $k-1$:

▶ **Theorem 26.** *Let $\mathcal{F}$ be closed under restrictions, and suppose that $M_k(\mathcal{F}) \le b^k$ for some $b \ge 1$, $k \ge 2$. If it further holds that $L_{1,i}(\mathcal{F}) \le b^i$ for all $1 \le i < k$, then there exists a $\Theta(\varepsilon^{2/k}/b^2)$-noticeable fractional pseudorandom generator for $\mathcal{F}$ with error ε and seed length $O(\log\log n + \log k + \log(1/\varepsilon))$.*

Proof. Fix $f \in \mathcal{F}$, and let $\mathbf{X}$ be a random variable such that $|\mathbf{X}_i| = c$ for all $i \in [n]$ for some $c > 0$ we specify momentarily. Then we have, via the Fourier expansion of f,

$$\left|\mathbb{E}_{\mathbf{X}}[f(\mathbf{X})] - f(\mathbf{0})\right| = \left|\mathbb{E}_{\mathbf{X}}\left[\sum_{S \subseteq [n]: 1 \le |S| \le k-1} \hat{f}(S)\mathbf{X}^S\right]\right| + \left|\mathbb{E}_{\mathbf{X}}[f_{\ge k}(\mathbf{X})]\right|.$$

We first deal with the second term on the right hand side. By Theorem 17 we have

$$\left|\mathbb{E}_{\mathbf{X}}[f_{\ge k}(\mathbf{X})]\right| \le \left(\frac{c}{1-c}\right)^k M_k(\mathcal{F}).$$

By assumption, $M_k(\mathcal{F}) \le b^k$ for some $b \ge 1$; therefore, by taking $c = \Theta(\varepsilon^{1/k}/b)$, this term is at most $\varepsilon/2$. To deal with the first term, we take the same approach as [4]. Under the assumption $L_{1,i}(\mathcal{F}) \le b^i$ for all $i < k$, one may apply their analysis by letting $\mathbf{X} = c \cdot \mathbf{Y}'$, where $\mathbf{Y}'$ is an $(k-1)$-wise $(\varepsilon/2)$-biased independent random variable over $\{-1,1\}^n$. It is clear that $\mathbf{X}$ is $c^2 = \Theta(\varepsilon^{2/k}/b^2)$-noticeable. Moreover, exactly as in [4], we have

$$\left|\mathbb{E}_{\mathbf{X}}\left[\sum_{S \subseteq [n]: 1 \le |S| \le k-1} \hat{f}(S)\mathbf{X}^S\right]\right| \le \sum_{i=1}^{k-1} c^i \sum_{S \subseteq [n]: |S| = i} |\hat{f}(S)| |\mathbb{E}[\mathbf{Y}'^S]| \le (\varepsilon/2) \sum_{i=1}^{k-1} (bc)^i \le \varepsilon/2,$$

because by our choice of c we have $bc \leq 1/2$. By standard constructions, $\mathbf{Y}'$ can be efficiently sampled with seed length $O(\log\log n + \log k + \log(1/\varepsilon))$ [15]. Combining these two errors proves the theorem. ◀

4.2 From Fractional PRGs to PRGs

Using Theorem 25 and Theorem 26 in tandem with Theorem 14, it is fairly immediate to obtain PRGs that rely only on a bound on some k-th Fourier level. Similarly, bounds on levels up to k can be leveraged to get an improved seed length.

▶ **Theorem 27** (Theorem 4, restated). *Let $\mathcal{F}$ be any class of n-variate Boolean functions that is closed under restrictions. Suppose that $M_k(\mathcal{F}) \leq b^k$ for some $b \geq 1$ and $k \geq 3$. Then for any $\varepsilon > 0$, there exists an explicit PRG for $\mathcal{F}$ with error ε with seed length*

$$O\left(\frac{b^{2+\frac{4}{k-2}} \cdot k \log n \cdot \log^{1+\frac{2}{k-2}}(n/\varepsilon)}{\varepsilon^{\frac{2}{k-2}}}\right).$$

If it further holds that $L_{1,i}(\mathcal{F}) \leq b^i$ for all $1 \leq i < k$, then the seed length can be improved to

$$O\left(\frac{b^{2+\frac{4}{k-2}} \cdot (\log\log n + \log k + \log(b/\varepsilon)) \cdot \log^{1+\frac{2}{k-2}}(n/\varepsilon)}{\varepsilon^{\frac{2}{k-2}}}\right).$$

Proof. By Theorem 14, given an explicit p-noticeable fractional PRG for $\mathcal{F}$ with error δ and seed length s, one immediately obtains an explicit PRG for $\mathcal{F}$ with error $O(\delta \log(n/\delta)/p)$ and seed length $O(s \log(n/\delta)/p)$.

For the first statement, by our assumption and using the fractional PRG guaranteed by Theorem 25, for any $\delta > 0$, we immediately obtain an explicit PRG for $\mathcal{F}$ with error $O(b^2\delta^{1-2/k}\log(n/\delta))$ and seed length $O(b^2 k \log(n)\log(n/\delta)/\delta^{2/k})$. To get the error below ε, we set

$$\delta = \Theta\left(\left(\frac{\varepsilon}{b^2\log(n/\varepsilon)}\right)^{\frac{k}{k-2}}\right)$$

(the astute reader may notice we implicitly use $b \leq n$ here). This yields a PRG with error ε and seed length

$$O\left(\frac{b^{2+\frac{4}{k-2}} \cdot k \log n \cdot \log^{1+\frac{2}{k-2}}(n/\varepsilon)}{\varepsilon^{\frac{2}{k-2}}}\right).$$

The second statement follows in an identical manner from the improved seed length given in the second part of Theorem 26 in the case that one has control on the L_1 Fourier mass on the lower levels. ◀

Corollary 5 is now an immediate consequence of Theorem 27; for any desired $\varepsilon > b \cdot \log(n) \cdot 2^{-O(k)}$, one can simply apply Theorem 27 using level $k = \Theta(\log(b\log(n)/\varepsilon))$ to obtain a PRG for $\mathcal{F}$ with error at most ε with seed length

$$O(b^2 \cdot \log(b\log(n)/\varepsilon) \cdot \log(n/\varepsilon)).$$

Note that for error $\varepsilon = 1/\mathrm{poly}(n)$, one needs bounds only up to level $\Theta(\log n)$ (again, using the fact that $b \leq n$). This also partially answers an open question of [6], which asks how many levels of Fourier bounds suffice to recover polylogarithmic dependence in $1/\varepsilon$.

▶ Remark 28. Note that this Taylor's theorem approach does not yield anything nontrivial given bounds just on the second level, unlike the fractional PRG in [6]. This is actually a necessary byproduct of combining this approach with the random walk gadget of [4]. Given only level-two bounds, this approach attempts to use j-wise independence for $j < k = 2$ and smallness to deal with errors on the high degree terms ($k \geq 2$). However, the trivial random variable that is $\pm\mathbf{1}$ with equal probability is trivially 1-wise independent, as each component is a uniform random bit, albeit completely correlated. No matter how we scale them, one can show that composing arbitrarily many independent copies of this random variable via the random walk gadget must necessarily polarize to $\pm\mathbf{1}$ at termination, which clearly cannot fool any nontrivial functions.

5 Low-degree Polynomials over $\mathbb{F}_2$

Our analysis recovers all the existing applications of [4] (among them, $\mathbf{AC}^0$ circuits, low-sensitivity functions, and read-once branching programs); indeed, all the classes considered there satisfy L_1 Fourier bounds on the entire tail. To our knowledge, our new analysis does not immediately improve the seed lengths obtained there, though it shows that (i) *the seed lengths there can potentially be improved using stronger bounds on M_k*, and (ii) *the PRGs there would still have fooled those classes had these Fourier bounds been known only up to some level k.*

However, the generality afforded to us by this new analysis allows us to obtain a new PRG for low-degree polynomials over $\mathbb{F}_2$, which addresses an open question of [4] by showing that this framework can handle this class. Indeed, let $\mathcal{F}$ be the set of n-variate, degree-d polynomials over $\mathbb{F}_2$. As a preliminary step towards deriving Fourier tail bounds that would imply a nontrivial PRG for this class using their framework, [4] proves the following Fourier bounds:

▶ **Proposition 29** (Theorem 6.1 of [4]). *Let $p\colon \mathbb{F}_2^n \to \mathbb{F}_2$ be a degree-d polynomial, and let $f(\mathbf{x}) = (-1)^{p(\mathbf{x})}$. Then $L_{1,k}(f) \leq (k \cdot 2^{3d})^k$.*

Note that this result cannot be applied to their original analysis, for they require a nontrivial bound at all levels, while this bound is trivial for $k = \Omega(\sqrt{n})$ and any d. While Theorem 16 can yield a nontrivial PRG by just applying the level-two bound, the dependence on $1/\varepsilon$ is at least quadratic.[4] However, using our new, more flexible analysis, one can obtain a nontrivial PRG with polylogarithmic dependence on the error parameter. Our formal result is the following:

▶ **Theorem 30.** *Let $\mathcal{F}$ be the class of degree-d polynomials over $\mathbb{F}_2$ on n variables. Then there exists an explicit pseudorandom generator for $\mathcal{F}$ with error ε and seed length*

$$2^{O(d)} \cdot \log^3(\log(n)/\varepsilon) \cdot \log(n/\varepsilon).$$

Proof. Fix $\varepsilon > 0$ and let $k = \Theta(\log(\log(n)/\varepsilon))$. By Proposition 29, we have that for all $j \leq k$,

$$L_{1,j}(\mathcal{F}) \leq \Theta\big(\log(\log(n)/\varepsilon) \cdot 2^{3d}\big)^j.$$

[4] By applying this Fourier bound at level-two, one can use the fractional PRG of [6] to obtain seed length $2^{O(d)}\mathrm{polylog}(n)/\varepsilon^{2+o(1)}$ using the random walks framework. This gives exponentially worse error dependence compared to our approach.

By setting $b = \Theta(\log(\log(n)/\varepsilon) \cdot 2^{3d})$, we may apply Theorem 27 for $\mathcal{F}$ and error ε. Note that $\varepsilon^{-\Theta(1/\log(1/\varepsilon))} = O(1)$, so plugging in this value of b, we immediately obtain the desired pseudorandom generator. ◀

For comparison, the best known construction by Viola [26], obtained by summing d independent copies of a sufficiently good small-bias space, attains seed length $d \cdot \log n + O(d \cdot 2^d \log(1/\varepsilon))$, which for constant ε and d is within a constant factor of the optimal seed length. The generator implied by our analysis recovers this polylogarithmic dependence in n/ε, although with slightly worse dependence on $\log n$ and polynomially worse dependence in $\log(1/\varepsilon)$. Neither generator can handle superlogarithmic degree. While this result clearly falls short of the state-of-the-art, we emphasize that this generator is conceptually distinct from the existing constructions, and yet belongs to this generic random walk framework.

Our analysis allows us to exploit known Fourier bounds that are too weak for the existing analyses to obtain polylogarithmic error dependence. In particular, to get a nontrivial pseudorandom generator for polynomials of superlogarithmic degree with nontrivial seed length, our work shows that the following weaker conjecture would suffice to break the logarithmic degree barrier and still achieve polylogarithmic (in n) seed length for $\varepsilon = 1/\text{poly}(n)$:

▶ **Conjecture 31.** *Let $\mathcal{F}$ be the class of degree-d polynomials over $\mathbb{F}_2$ on n variables. Then*

$$M_k(\mathcal{F}) \leq (\text{poly}(k, \log n) \cdot 2^{o(d)})^k$$

for $k \leq O(\log n)$.

In fact, we observe that to break the logarithmic degree barrier, it actually suffices that this holds just at level $k = 3$, though with poor dependence on ε. Note that this is a significantly weaker conjecture than positing that the same bounds hold for $L_{1,k}(\mathcal{F})$. Moreover, as we explain in the next section, $M_k(\mathcal{F})$ can be controlled using correlation bounds, which are much better studied than L_1 Fourier bounds.

6 Bounds on $M_k(\mathcal{F})$ via Correlation with Shifted Majorities

As we have seen, our new analysis lets one construct PRGs from the weaker quantity $M_k(\mathcal{F})$. In this section, we extend the argument of Chattopadhyay, Hatami, Hosseini, Lovett, and Zuckerman [5] to show how bounds on $M_k(\mathcal{F})$ follow from covariance bounds with certain resilient functions (in particular, shifted majorities). In their paper, they deal with the case of $k = 2$; we rather straightforwardly generalize this argument, but stress that the approach is the same as in Section 6 of their paper. To that end, for convenience and consistency with their argument, we adopt their conventions and requisite definitions just for this section. We will now consider Boolean functions written as $f : \{0,1\}^n \to \{0,1\}$. Translating to this notation, for any such Boolean function f, let $e(f)(\mathbf{x}) \triangleq (-1)^{f(\mathbf{x})}$. Then, letting $F = e(f)$, we now have $\hat{F}(S) = \mathbb{E}_{\mathbf{x}}[F(\mathbf{x})e(\sum_{i \in S} x_i)]$.

▶ **Definition 32.** *The* covariance between f and g, *where f, g are Boolean is*

$$\text{cov}(f, g) \triangleq \left|\mathbb{E}[e(f(\mathbf{x}))e(g(\mathbf{x}))] - \mathbb{E}[e(f(\mathbf{x}))]\mathbb{E}[e(g(\mathbf{x}))]\right|.$$

The covariance between a function f and a class $\mathcal{G}$ is defined as $\text{cov}(f, \mathcal{G}) \triangleq \max_{g \in \mathcal{G}} \text{cov}(f, g)$.

For any $\mathbf{x} \in \{0,1\}^n$, we write $|\mathbf{x}|$ for its Hamming weight, i.e. $\sum_{i=1}^n x_i$. For any $a \in \{0,1,\ldots,n\}$, [5] defines Maj_a by

$$\mathrm{Maj}_a(\mathbf{x}) \triangleq \begin{cases} 1 & \text{if } |\mathbf{x}| > a \\ 0 & \text{otherwise,} \end{cases}$$

as well as the following associated functions for any $\theta \in [n/2]$:

$$\mathrm{Thr}_\theta(x) \triangleq \begin{cases} (-1)^{\mathrm{Maj}_{n/2}(\mathbf{x})} & \text{if } \big||\mathbf{x}| - n/2\big| > \theta \\ 0 & \text{otherwise.} \end{cases}$$

We now prove the following lemma relating $M_k(\mathcal{F})$ with covariance bounds against the k-XORs of these functions:

▶ **Lemma 33** (Lemma 6.1 of [5], adapted). *Let $\mathcal{F}$ be any family of (kn)-variate Boolean functions that is closed under relabeling and negation of input variables. Suppose that for any $a_1,\ldots,a_k$ such that $|a_i - n/2| = O(\sqrt{kn\log n})$ for all $i \in [k]$, and all $f \in \mathcal{F}$, we have for some $t \geq 1$*

$$\mathrm{cov}\big(f, \oplus_{i=1}^k \mathrm{Maj}_{a_i}\big) \leq \left(\sqrt{\frac{t}{n}}\right)^k,$$

where $\oplus$ denotes the XOR *function. Then,*

$$M_k(\mathcal{F}) \leq O\big(\sqrt{tk\log n}\big)^k.$$

To prove this lemma, [5] uses the following sequence of claims.

▶ **Fact 34** (Claim 6.2 in [5]). *For any $f \in \mathcal{F}$, let $F(\mathbf{x}_1,\ldots,\mathbf{x}_k) = e(f(\mathbf{x}_1,\ldots,\mathbf{x}_k))$. Under the hypotheses of Lemma 33, for any $1 \leq a_1,\ldots,a_k \leq O(\sqrt{kn\log n})$,*

$$\left|\mathbb{E}_{\mathbf{x}_1,\ldots,\mathbf{x}_k}\left[\big(F(\mathbf{x}_1,\ldots,\mathbf{x}_k) - \mathbb{E}[F]\big)\prod_{i=1}^k \mathrm{Thr}_{a_i}(\mathbf{x}_i)\right]\right| \leq \left(\sqrt{\frac{t}{n}}\right)^k.$$

▶ **Fact 35** (Claim 6.3 of [5]). *For any $\mathbf{x} \in \{0,1\}^n$, $\sum_{i=1}^n e(\mathbf{x}_i) = 2\sum_{1\leq a\leq n/2} \mathrm{Thr}_a(\mathbf{x})$.*

▶ **Fact 36** (Claim 6.4 of [5], adapted). *For any Boolean function $f : \{0,1\}^{kn} \to \{0,1\}$, there exists a k-equipartition of $[kn]$ into disjoint sets $S_1,\ldots,S_k$ such that*

$$\left|\sum_{S\subseteq[kn]:|S|=k} \hat{f}(S)\right| \leq C^k \left|\sum_{i_j\in S_j\ \forall j\in[k]} \hat{f}(\{i_1,\ldots,i_k\})\right|$$

for some absolute constant $C > 0$.

As this fact is not quite identical to that in [5], we give an argument here:

Proof. We use the probabilistic method: let $\mathcal{P}$ be the set of k-equipartitions of $[kn]$. Let $T \subseteq [kn]$ of size k be arbitrary; without loss of generality, suppose $T = [k]$. Consider a uniformly random k-equipartition $P = S_1 \sqcup \cdots \sqcup S_k \in \mathcal{P}$. The probability that each $i \in T$ belongs to a distinct S_j is easily seen to be

$$\prod_{i=1}^{k-1} \frac{(k-i)\cdot n}{kn-i} \geq \frac{(k-1)!\, n^{k-1}}{(kn)^{k-1}} = \frac{(k-1)!}{k^{k-1}} = e^{-O(k)},$$

where the last equality uses Stirling's approximation. By symmetry, let $\alpha \in \mathbb{N}$ be the number of k-equipartitions that any arbitrary subset T is in. Then we have

$$\begin{aligned}\alpha\left|\sum_{S\subseteq[kn]:|S|=k}\hat{f}(S)\right| &= \left|\sum_{P\in\mathcal{P}}\sum_{i_j\in S_j\ \forall j\in[k]}\hat{f}(\{i_1,\dots,i_k\})\right|\\ &\le \sum_{P\in\mathcal{P}}\left|\sum_{i_j\in S_j\ \forall j\in[k]}\hat{f}(\{i_1,\dots,i_k\})\right|\\ &\le |\mathcal{P}|\max_{P\in\mathcal{P}}\left|\sum_{i_j\in S_j\ \forall j\in[k]}\hat{f}(\{i_1,\dots,i_k\})\right|. \qquad \blacktriangleleft\end{aligned}$$

The first line follows from simple counting, while the second is the triangle inequality. Rearranging, we deduce that (writing T as a generic subset of size k)

$$\begin{aligned}\left|\sum_{S\subseteq[kn]:|S|=k}\hat{f}(S)\right| &\le \frac{|\mathcal{P}|}{\alpha}\max_{P\in\mathcal{P}}\left|\sum_{i_j\in S_j\ \forall j\in[k]}\hat{f}(\{i_1,\dots,i_k\})\right|\\ &= \Pr_{P\sim\mathcal{P}}(T\in P)^{-1}\max_{P\in\mathcal{P}}\left|\sum_{i_j\in S_j\ \forall j\in[k]}\hat{f}(\{i_1,\dots,i_k\})\right|\\ &\le e^{O(k)}\max_{P\in\mathcal{P}}\left|\sum_{i_j\in S_j\ \forall j\in[k]}\hat{f}(\{i_1,\dots,i_k\})\right|.\end{aligned}$$

The last fact that is needed can be deduced from the Chernoff bound:

▶ **Fact 37** (Claim 6.5 of [5], adapted). *For any $a \ge \Omega(\sqrt{kn\log n})$, $\mathbb{E}[|\mathrm{Thr}_a|] \le O(1/n^k)$.*

With these facts, we can now prove Lemma 33 in an entirely analogous fashion to [5]:

Proof of Lemma 33. Fix $f \in \mathcal{F}$, and again write $F(\mathbf{x}_1,\dots,\mathbf{x}_k) = e(f(\mathbf{x}_1,\dots,\mathbf{x}_k))$. Let $F' = F - \mathbb{E}[F]$. Let $U_j = \{i : (j-1)n+1 \le i \le jn\}$. Then, possibly after relabelling variables, we have by Fact 36 that

$$\left|\sum_{S\subseteq[kn]:|S|=k}\hat{f}(S)\right| \le C^k\left|\sum_{i_j\in U_j,\forall j\in[k]}\hat{f}(\{i_1,\dots,i_k\})\right|,$$

so we may turn to bounding this latter term. We have

$$\begin{aligned}\left|\sum_{i_j\in U_j,\forall j\in[k]}\hat{f}(\{i_1,\dots,i_k\})\right| &= \left|\sum_{i_j\in U_j,\forall j\in[k]}\mathbb{E}\Big[F'(\mathbf{x}_1,\dots,\mathbf{x}_k)\prod_{j=1}^{k}e\Big((\mathbf{x}_j)_{i_j}\Big)\Big]\right|\\ &= \left|\mathbb{E}\Big[F'(\mathbf{x}_1,\dots,\mathbf{x}_k)\prod_{j=1}^{k}\Big(\sum_{i_j\in U_j}e\Big((\mathbf{x}_j)_{i_j}\Big)\Big)\Big]\right|\\ &\le 2^k\sum_{1\le a_i\le n/2,\forall i\in[k]}\left|\mathbb{E}\Big[F'(\mathbf{x}_1,\dots,\mathbf{x}_k)\prod_{i=1}^{k}\mathrm{Thr}_{a_i}(\mathbf{x}_i)\Big]\right|\\ &\le 2^k\left(\sum_{1\le a_i\le O(\sqrt{kn\log n}),\forall i\in[k]}\left|\mathbb{E}\Big[F'(\mathbf{x}_1,\dots,\mathbf{x}_k)\prod_{i=1}^{k}\mathrm{Thr}_{a_i}(\mathbf{x}_i)\Big]\right|+O(1)\right)\\ &\le 2^k\cdot O\Big(\sqrt{kn\log n}\Big)^k\cdot\left(\sqrt{\frac{t}{n}}\right)^k\\ &= O\Big(\sqrt{tk\log n}\Big)^k.\end{aligned}$$

The first inequality follows from Fact 35, the second from Fact 37, and the last from Fact 34. Because we assumed that $\mathcal{F}$ is closed under negations of input variables and $f \in \mathcal{F}$ was arbitrary, we obtain the desired claim from Lemma 9 after absorbing the constant C above into the implicit constant in this bound. ◀

7 Discussion and Open Questions

In this work, we have given a nearly complete interpolation between the previous PRGs obtained in the polarizing random walk framework by exploiting level-k bounds on the class of functions, thus answering an open question from [6]. We do so by exploiting an alternate Fourier analysis via Taylor's theorem and utilizing multilinearity and random restrictions. This new analysis enables us to construct PRGs from bounds on the potentially much smaller and better-understood Fourier quantity $M_k(\mathcal{F})$, for any $k \geq 3$. By generalizing the connection established in [5], this reduces the problem of constructing PRGs in this framework to proving correlation bounds. Further, we show how to get a PRG with an improved seed length if we have bounds on $L_{1,i}(\mathcal{F})$, for all $i \leq k$, where $k \geq 3$. A natural open question along these lines is to obtain the improved seed length using bounds on $M_i(\mathcal{F})$ (instead of $L_{1,i}(\mathcal{F})$) for all $i \leq k$. Another natural question is to construct a PRG using bounds on just M_2 (recall that [6] gives such a construction using bounds on $L_{1,2}(\mathcal{F})$ and our analysis only gives a non-trivial PRG from bounds on $M_k(\mathcal{F})$ when $k \geq 3$).

Finally, exploiting known level-k bounds for $\mathbb{F}_2$ polynomials, our approach shows that the polarizing random walk framework can yield pseudorandom generators for the class of $\mathbb{F}_2$ polynomials that is competitive with the state of the art. As mentioned, we hope this paper gives evidence that stronger Fourier control (perhaps via proving the required correlation bounds) can give better PRGs using this framework, and can also handle classes that were previously not known to be possible. In particular, we emphasize that proving Conjecture 31 even for the case of $k = 3$ will lead to PRGs for $\mathbb{F}_2$-polynomials with degree $\omega(\log n)$, a longstanding problem in complexity theory.

References

1 Rohit Agrawal. Coin theorems and the Fourier expansion. *Chicago Journal of Theoretical Computer Science*, 2020(4), August 2020.

2 Srinivasan Arunachalam, Sourav Chakraborty, Michal Koucký, Nitin Saurabh, and Ronald de Wolf. Improved bounds on Fourier entropy and min-entropy. In Christophe Paul and Markus Bläser, editors, *37th International Symposium on Theoretical Aspects of Computer Science, STACS 2020, March 10-13, 2020, Montpellier, France*, volume 154 of *LIPIcs*, pages 45:1–45:19. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPIcs.STACS.2020.45.

3 Nikhil Bansal and Makrand Sinha. k-forrelation optimally separates quantum and classical query complexity. *CoRR*, abs/2008.07003, 2020. arXiv:2008.07003.

4 Eshan Chattopadhyay, Pooya Hatami, Kaave Hosseini, and Shachar Lovett. Pseudorandom generators from polarizing random walks. *Theory of Computing*, 15(10):1–26, 2019. doi:10.4086/toc.2019.v015a010.

5 Eshan Chattopadhyay, Pooya Hatami, Kaave Hosseini, Shachar Lovett, and David Zuckerman. XOR lemmas for resilient functions against polynomials. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2020, page 234–246, New York, NY, USA, 2020. Association for Computing Machinery. doi:10.1145/3357713.3384242.

6 Eshan Chattopadhyay, Pooya Hatami, Shachar Lovett, and Avishay Tal. Pseudorandom Generators from the Second Fourier Level and Applications to AC0 with Parity Gates. In Avrim Blum, editor, *10th Innovations in Theoretical Computer Science Conference (ITCS 2019)*, volume 124 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 22:1–22:15, Dagstuhl, Germany, 2019. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. doi:10.4230/LIPIcs.ITCS.2019.22.

7 Eshan Chattopadhyay, Pooya Hatami, Omer Reingold, and Avishay Tal. Improved pseudorandomness for unordered branching programs through local monotonicity. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 363–375, 2018. doi:10.1145/3188745.3188800.

8 Andreas Defant, Leonhard Frerick, Joaquim Ortega-Cerdà, Myriam Ounaïes, and Kristian Seip. The Bohnenblust-Hille inequality for homogeneous polynomials is hypercontractive. *Annals of mathematics*, pages 485–497, 2011.

9 Uma Girish, Ran Raz, and Wei Zhan. Lower bounds for XOR of forrelations. *CoRR*, abs/2007.03631, 2020. arXiv:2007.03631.

10 Parikshit Gopalan, Rocco A. Servedio, Avishay Tal, and Avi Wigderson. Degree and sensitivity: tails of two distributions, 2016. arXiv:1604.07432.

11 Parikshit Gopalan, Rocco A. Servedio, and Avi Wigderson. Degree and sensitivity: Tails of two distributions. In Ran Raz, editor, *31st Conference on Computational Complexity, CCC 2016, May 29 to June 1, 2016, Tokyo, Japan*, volume 50 of *LIPIcs*, pages 13:1–13:23. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016. doi:10.4230/LIPIcs.CCC.2016.13.

12 Chin Ho Lee. Fourier bounds and pseudorandom generators for product tests. In Amir Shpilka, editor, *34th Computational Complexity Conference, CCC 2019, July 18-20, 2019, New Brunswick, NJ, USA*, volume 137 of *LIPIcs*, pages 7:1–7:25. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. doi:10.4230/LIPIcs.CCC.2019.7.

13 Nathan Linial, Yishay Mansour, and Noam Nisan. Constant depth circuits, Fourier transform, and learnability. In *30th Annual Symposium on Foundations of Computer Science*, pages 574–579. IEEE, 1989.

14 Ashley Montanaro. Some applications of hypercontractive inequalities in quantum information theory. *Journal of Mathematical Physics*, 53(12):122206, 2012.

15 Joseph Naor and Moni Naor. Small-bias probability spaces: Efficient constructions and applications. In Harriet Ortiz, editor, *Proceedings of the 22nd Annual ACM Symposium on Theory of Computing, May 13-17, 1990, Baltimore, Maryland, USA*, pages 213–223. ACM, 1990. doi:10.1145/100216.100244.

16 Ryan O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, 2014. doi:10.1017/CBO9781139814782.

17 Qazi Ibadu Rahman and Gerhard Schmeisser. *Analytic theory of polynomials*, volume 26 of *London Mathematical Society Monographs. New Series*. The Clarendon Press, Oxford University Press, Oxford, 2002.

18 Ran Raz and Avishay Tal. Oracle separation of BQP and PH. In Moses Charikar and Edith Cohen, editors, *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019, Phoenix, AZ, USA, June 23-26, 2019*, pages 13–23. ACM, 2019. doi:10.1145/3313276.3316315.

19 Alexander A. Razborov. Lower bounds on the size of bounded depth circuits over a complete basis with logical addition. *Mathematical notes of the Academy of Sciences of the USSR*, 41(4):333–338, April 1987. doi:10.1007/BF01137685.

20 Alexander A. Sherstov, Andrey A. Storozhenko, and Pei Wu. An optimal separation of randomized and quantum query complexity. *Electron. Colloquium Comput. Complex.*, 27:128, 2020. URL: https://eccc.weizmann.ac.il/report/2020/128.

21 Roman Smolensky. Algebraic methods in the theory of lower bounds for Boolean circuit complexity. In *Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing*, STOC '87, page 77–82, New York, NY, USA, 1987. Association for Computing Machinery. doi:10.1145/28395.28404.

22 Roman Smolensky. On representations by low-degree polynomials. In *Proceedings of the 1993 IEEE 34th Annual Foundations of Computer Science*, SFCS '93, page 130–138, USA, 1993. IEEE Computer Society. doi:10.1109/SFCS.1993.366874.

23 Avishay Tal. Tight bounds on the Fourier spectrum of AC0. In Ryan O'Donnell, editor, *32nd Computational Complexity Conference, CCC 2017, July 6-9, 2017, Riga, Latvia*, volume 79 of *LIPIcs*, pages 15:1–15:31. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017. doi:10.4230/LIPIcs.CCC.2017.15.

24 Avishay Tal. Towards optimal separations between quantum and randomized query complexities. *Electron. Colloquium Comput. Complex.*, 26:179, 2019. URL: https://eccc.weizmann.ac.il/report/2019/179.

25 Salil P. Vadhan. Pseudorandomness. *Foundations and Trends in Theoretical Computer Science*, 7(1–3):1–336, 2012.

26 Emanuele Viola. The sum of d small-bias generators fools polynomials of degree d. *Computational Complexity*, 18(2):209–217, 2009. doi:10.1007/s00037-009-0273-5.

27 Emanuele Viola. Fourier conjectures, correlation bounds, and majority. *Electron. Colloquium Comput. Complex.*, 27:175, 2020. URL: https://eccc.weizmann.ac.il/report/2020/175.

28 Xinyu Wu. A stochastic calculus approach to the oracle separation of BQP and PH. *CoRR*, abs/2007.02431, 2020. arXiv:2007.02431.