Electronic Colloquium on Computational Complexity, Report No. 121 (2020)

Fractional Pseudorandom Generators from the kth Fourier Level

Eshan Chattopadhyay*
Cornell University
eshanc@cornell.edu

Jason Gaitonde†
Cornell University
jsg355@cornell.edu

Abhishek Shetty‡
Cornell University
shetty@cs.cornell.edu

August 3, 2020

Abstract

In recent work by Chattopadhyay et al. [CHHL19, CHLT19], the authors exhibit a simple and flexible construction of pseudorandom generators for classes of Boolean functions that satisfy L_1 Fourier bounds. [CHHL19] show that if a class satisfies such tail bounds at all levels, this implies a PRG whose seed length depends on the quality of these bounds through their innovative *random walk framework* that composes together *fractional PRGs* that polarize quickly to the Boolean hypercube. On the other hand, [CHLT19] show that, by derandomizing the analysis of [RT19], just level-two Fourier bounds suffice to construct a pseudorandom generator using their framework; as this is a much weaker assumption on the class, [CHLT19] naturally obtain exponentially worse dependence on the error in the seed length compared to [CHHL19]. Moreover, this derandomization relies on simulating nearly independent Gaussians for the fractional pseudorandom generator, which necessitates the polynomial dependence on $1/\varepsilon$ in each fractional step.

In this work, we attempt to bridge the gap between these two results. Namely, we partially answer an open question by [CHLT19] that nearly interpolates between them. In particular, we show that if one has bounds up to the level-k L_1 Fourier mass of a closely related class of functions, where $k > 2$, one can obtain improved seed length, the degree to which is determined by how high k can be taken. Our analysis shows that for error $\epsilon = 1/\mathrm{poly}(n)$, one needs control at just level $O(\log n)$ to recover the seed length of [CHHL19], without assumptions on the entire tail. We avoid this by providing a simple, alternate analysis of their fractional PRG that instead relies on Taylor's theorem and p-biased Fourier analysis to avoid assumptions on the weights of the higher-order terms. This further allows us to show that this framework can handle the class of low-degree polynomials over $\mathbb{F}_2$, with slightly worse dependences than the current state-of-the-art, which was not previously known. We hope that this alternate analysis will be fruitful in improving the understanding of this new and powerful framework.

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 $\mathbf{P} = \mathbf{BPP}$, a rich line of work has focused on *derandomizing* simpler models of computation (see Vadhan [Vad12] 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.

*Supported by NSF grant CCF-1849899.
†Supported by NSF grant CCF-1408673 and AFOSR grant F5684A1.
‡Supported by a Cornell University Fellowship and a JP Morgan Chase Faculty Fellowship.

ISSN 1433-8092

Definition 1.1. Let $\mathcal{F}$ be a class of n-variate Boolean functions. Then a *pseudorandom generator* (PRG) for $\mathcal{F}$ with error $\epsilon > 0$ is a random variable $\mathbf{X} \in \{-1,1\}^n$ such that for all $f \in \mathcal{F}$,

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

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.

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. Instead, we focus on a recent line of work initiated by Chattopadhyay et al. [CHHL19, CHLT19] that provides a framework for constructing pseudorandom generators for any Boolean function class that exhibit Fourier tail bounds (we discuss this in more details in the next subsection; see Section 2.3 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 discuss this new framework in Section 1.1, and present our results in Section 1.2.

1.1 The Polarizing Random Walk Framework

We now briefly explain the *polarizing random walks* framework introduced by [CHHL19]. The authors show that for classes of Boolean n-variate functions, that are also closed under restrictions, one can quite flexibly construct pseudorandom generators via a local-to-global principle that works as follows: it is sufficient to construct *fractional pseudorandom generators*, a notion that generalizes a PRG and allows the random variable $\mathbf{X}$ (in Definition 1.1) to be supported on the solid cube $[-1,1]^n$, that can fool the multilinear expansions of each Boolean function in the class (see Definition 2.5 for a formal definition). Ideally, these random variables should be as large as possible while still provably fooling the class.

To construct the full pseudorandom generator, the authors give a random walk gadget that composes together independent copies of such a fractional generator as steps in a random walk that polarizes quickly to the Boolean hypercube. The analysis for how the error accumulates in this process relies on interpreting the intermediate points of this pseudorandom walk as an average of *random restrictions* of the original function; because the fractional generator locally fools the class, this interpretation shows that it does not incur much error at each intermediate step, and the rapid polarization show that it does not take too many steps. Taken together, these two facts imply the resulting random variable successfully fools the class.

The above framework shows that if one can construct non-Boolean random variables, with sufficient variance in each coordinate, and that can locally fool any function in the class, then one immediately obtains a pseudorandom generator using their random walk gadget. As these generators need not be Boolean, the construction of fractional pseudorandom generators is only easier than constructing pseudorandom generators. To that end, [CHHL19] further show how to construct such fractional pseudorandom generators for any class of functions satisfying *Fourier tail bounds*. Namely, they show that if 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 $b > 0$, then one can construct a fractional pseudorandom generator for error ϵ with seed length $O(\log\log n + \log(1/\epsilon))$ 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 $b^2 \cdot \mathrm{polylog}(n/\epsilon)$. As a result, if one can show that a class admits nontrivial Fourier tail bounds, then the [CHHL19] construction immediately implies a pseudorandom generator. Some examples of classes of Boolean functions that exhibit such tail bounds include $\mathbf{AC}^0$ circuits with the parameter $b = \mathrm{poly}(\log n)$ [LMN89, Tal17], constant width read-once branching programs with $b = \mathrm{poly}(\log n)$

[CHRT18], and low-sensitivity functions with $b = O(s)$ [GSW16, Tal17]. Using such Fourier bounds, [CHHL19] immediately gave a polylogarithmic seed length PRG for these function classes. It was also conjectured in [CHHL19] 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.

In subsequent work by [CHLT19], the authors show how to construct fractional pseudorandom generators using far fewer assumptions on the Fourier tails. Building on the analysis of [RT19] in their celebrated work proving the oracle separation of **BQP** and **PH** (which itself relies on the [CHHL19] random walk framework), they show that one can use this same framework to obtain a pseudorandom generator with seed length depending only on bounds for the *second Fourier level* of the class. However, with these weaker assumptions, they require a different fractional PRG. To do this, they essentially derandomize the result of [RT19], which shows that classes of multilinear functions with low level-two Fourier mass cannot nontrivially distinguish between a suitable variant of the Forrelation distribution and the uniform distribution. It turns out that this can be interpreted via the fact that Itô's Lemma shows that the local behavior of a smooth function of Brownian motion is essentially determined by the first two derivatives [Wu20]. [CHLT19] show that one can derandomize this analysis by efficiently constructing fractional PRGs that simulate Gaussians with small covariance using the best-known constructions of codes. However, this construction incurs exponentially worse dependence on the error parameter in each fractional step to nearly sample sufficiently good Gaussian random variables. The final seed length that this framework obtains is of the form $O((b^2/\epsilon)^{2+o(1)}\text{polylog}(n))$, where b^2 is the level-two Fourier mass of the class. Compared to [CHHL19], this yields exponentially worse dependence on the error, as well as quadratically worse dependence on the level-two mass (though [CHLT19] assume nothing about the rest of the Fourier levels).

1.2 Our Contribution

Given these two works, a very natural question (explicitly asked in [CHLT19]) is whether it is possible to interpolate between these constructions by assuming Fourier bounds on an intermediate k. 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 seed lengths of [CHHL19] and [CHLT19]? Given bounds up to level-k, what range of error $\epsilon > 0$ can the resulting PRG tolerate while maintaining polylogarithmic dependence on $1/\epsilon$ (or put contrapositively, given a desired error $\epsilon > 0$, how many levels of Fourier bounds are required to ensure the seed remains polylogarithmic in $1/\epsilon$)?

In this work, we make progress on all of these questions. We do so by providing a new analysis of the fractional pseudorandom generator of [CHHL19]. Informally, what we show is that by only assuming a bound on the level-k L_1 mass of an *associated class of functions*, the (possibly slightly modified, depending on the precise assumptions on the class) [CHHL19] fractional pseudorandom generator will still be valid. Formally, our main result is the following analysis of a fractional pseudorandom generator:

Theorem 1.1. *Let $\mathcal{F}$ be any class of n-variate Boolean functions closed under restrictions. Suppose that $\overline{\mathcal{F}} \circ \mathsf{AND}_2$ satisfies $L_{1,k}(\overline{\mathcal{F}} \circ \mathsf{AND}_2) \leq b^k$ for some $b > 0$ and $k > 2$. Then for any $\epsilon > 0$, there exists a $\Omega(\epsilon^{2/k}/b^2)$-noticeable explicit fractional PRG for $\mathcal{F}$ with error ϵ and seed length $O(k \cdot \log(n))$.*[1]

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

[1] 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.

We formally define the class $\overline{\mathcal{F}} \circ \mathsf{AND}_2$ in Section 2. Roughly, it is the set of functions obtainable from $\mathcal{F}$ by possibly negating some variables, and then replacing each variable with the AND_2 of two new variables. For many natural classes (such as $\mathbf{AC^0}$ circuits, read-once branching programs, low-sensitivity functions, and low-degree polynomials over $\mathbb{F}_2$) this associated class is contained in the $2n$-variate analogue of the original class; we give a detailed discussion in Section 4. As mentioned in Section 1.1, the classes of $\mathbf{AC^0}$ circuits, read-once branching programs, and low-sensitivity functions are known to satisfy L_1 Fourier tail bounds on the entire spectrum, and our work in fact shows that one could obtain similar seed length to [CHHL19] *even if one only had these bounds up to some level k*. For the class of $\mathbb{F}_2$ polynomials, such Fourier tail bounds are not yet known and Theorem 1.1 allows us to leverage weaker bounds proved in [CHHL19] to construct a PRG with polylogarithmic dependence on n/ϵ (see Theorem 1.3), almost matching the best known PRG due to Viola [Vio09]. As we discuss below, the results in [CHHL19, CHLT19] are not good enough to use known Fourier tail bounds to obtain such a PRG for the class of $\mathbb{F}_2$ polynomials.

Using the fractional pseudorandom generator from Theorem 1.1, we obtain the following consequences almost immediately from the random walk gadget (see Theorem 2.2):

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 *up to* level-k. The informal statement is the following:

   Theorem 1.2. *Let $\mathcal{F}$ be any class of n-variate Boolean functions closed under restrictions. Suppose that $\overline{\mathcal{F}} \circ \mathsf{AND}_2$ satisfies $L_{1,k}(\overline{\mathcal{F}} \circ \mathsf{AND}_2) \leq b^k$ for some $b > 0$ and $k > 2$. Then there exists an explicit pseudorandom generator for $\mathcal{F}$ for error ϵ with seed length $k \cdot b^{2+4/(k-2)}\mathrm{polylog}(n/\epsilon)/\epsilon^{2/(k-2)}$. The seed length can be improved if such bounds also hold on $\mathcal{F}$ up to level-k.*

   See Theorem 3.6 for the precise statement. One immediate consequence of this is that if one has comparable Fourier bounds even just at level $k = 4$, one obtains quadratically better dependence on the error in the seed length (as well as polylogarithmic factors in n/ϵ) compared to [CHLT19]. In particular, given a Fourier bound of b^k on just the associated class for some level $k \leq \mathrm{polylog}(n)$, one obtains an pseudorandom generator with error ϵ with seed length $O(b^{2+4/(k-2)}\mathrm{polylog}(n/\epsilon)/\epsilon^{2/(k-2)})$.

2. **Pseudorandom Generators with Polylogarithmic Error Dependence for $\epsilon \geq b \cdot \log(n) \cdot 2^{-\Omega(k)}$ from up to Level-k Bounds**: A simple corollary of our fractional pseudorandom generator is that one can recover the polylogarithmic dependence on $1/\epsilon$ from [CHHL19] if $\varepsilon \geq b \cdot \log(n) \cdot 2^{-\Omega(k)}$ and we have Fourier bounds *up to* level-k.

   Corollary 1.1. *Let $\mathcal{F}$ be any class of n-variate Boolean functions closed under restrictions. Suppose that $L_{1,i}(\mathcal{F} \cup \overline{\mathcal{F}} \circ \mathsf{AND}_2) \leq b^i$ for some $b > 0$ and all $1 \leq i \leq k$ for some k. Then, for any $\epsilon \geq b \cdot \log(n) \cdot 2^{-\Omega(k)}$, there exists an explicit pseudorandom generator for $\mathcal{F}$ for error ϵ with seed length $O(b^2\mathrm{polylog}(n/\epsilon))$.*

   This actually covers the analysis of [CHHL19] without requiring anything on the full Fourier tail, and addresses an open question of [CHLT19] asking how many levels of Fourier bounds one needs control of to regain polylogarithmic dependence on ε. In particular, if one requires error $\epsilon = 1/\mathrm{poly}(n)$, then it suffices to have Fourier bounds up to level $\Theta(\log(n))$ to get the same dependence.

We view this work mostly as a proof-of-concept that it is indeed possible to interpolate between these two extremes in the polarizing random walk framework and obtain better results using weakened Fourier assumptions, albeit of a mildly different class. Moreover, given the flexibility of this framework, we also view it as evidence that with alternate Fourier analysis, one can further improve the analysis and applicability of this random walk paradigm. We prove Theorem 1.1 in Section 3, from which Theorem 1.2 and Corollary 1.1 follow without much difficulty using the existing random walk technique of [CHHL19].

As a concrete possible application of this approach which would provably improve on these works, both [CHHL19] and [CHLT19] conjecture Fourier bounds on the L_1 mass of the class of $\mathbb{F}_2$ polynomials of degree at most d. 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 yet to being resolved, one can imagine that should the latter be proved, it may be feasible to extend the analysis to achieve a bound of $(\mathrm{poly}(d))^k$ for $k = \Omega(1)$, or even more optimistically, $k = \Omega(\log n)$. The pseudorandom generator implied by the analysis here would thus apply and yield significantly improved seed length compared to [CHLT19], though we note that our generator does not actually apply if such a bound only holds at level $k = 2$. Such a result would imply a pseudorandom generator with improved seed length compared to [CHLT19] for $\mathbf{AC^0}[\oplus]$ (see the discussion in [CHLT19], using results of Razborov [Raz87] and Smolensky [Smo87, Smo93]).

Nonetheless, our analysis here, coupled with weaker Fourier tail bounds obtained in [CHHL19], immediately implies the following:

Theorem 1.3. *Let $\mathcal{F}$ be the class of degree at most 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 $2^{O(d)}\mathrm{polylog}(n/\epsilon)$.*

See Section 4 for the precise dependences in the seed length. While this result does not quite match the current state-of-the-art PRG for this class due to Viola [Vio09] (and similarly, fails to give anything nontrivial for $d = \Omega(\log n)$), we view this as a conceptual contribution that the random walks framework can yield an explicit pseudorandom generator with error dependence that is polylogarithmic in n/ϵ, which was not previously known from [CHHL19] or [CHLT19]. The Fourier tail bounds that are sufficient for our analysis are too weak to be employed in [CHHL19], and leads to quadratic error dependence in the level-two fractional pseudorandom generator [CHLT19]. We hope this further suggests that the random walks framework can be fruitfully employed to study more classes than previously considered. We present the proof of Theorem 1.3 in Section 4.

1.3 Overview of Our Approach

To prove our results, we rely on an alternate, simple analysis of the fractional pseudorandom generator considered by [CHHL19] when they assume control on the entire Fourier tail, and then use their gadget construction to obtain the full pseudorandom generator. As this latter part can be done in an entirely black-box fashion, we need only focus on the former. For this first part, their approach is the following: consider a random variable $\mathbf{X}$. By writing out f in the multilinear (Fourier) expansion, one has

$$|\mathbb{E}_{\mathbf{X}}[f(\mathbf{X})] - \mathbb{E}_{\mathbf{U}_n}[f(\mathbf{U}_n)]| = \Big|\mathbb{E}_{\mathbf{X}}\Big[\sum_{\emptyset \neq S \subseteq [n]} \hat{f}(S)\mathbf{X}^S\Big]\Big| \tag{1}$$

$$\leq \underbrace{\sum_{i=1}^{k-1} \sum_{S \subseteq [n]:|S|=i} |\hat{f}(S)||\mathbb{E}_{\mathbf{X}}[\mathbf{X}^S]|}_{\text{low-order terms}} + \underbrace{\sum_{i=k}^{n} \sum_{S \subseteq [n]:|S|=i} |\hat{f}(S)||\mathbb{E}_{\mathbf{X}}[\mathbf{X}^S]|}_{\text{high-order terms}}, \tag{2}$$

where the second line is the triangle inequality. [CHHL19] use two different methods to control each component; the first part is handled by sufficiently strong independence in the coordinates, while the latter is handled by physical smallness of the random variable. To control the low-degree terms, they require $\mathbf{X}$ to be $O(\epsilon)$-almost k-wise independent. By definition, this forces all the Fourier characters to be small in the low-order component. To deal with the higher-order component, they use the fact that $\mathbf{X}$ *need not be Boolean*; that is, they can scale any such distribution down. Because they assume that each Fourier level is bounded by b^k, it suffices to scale down a $\{\pm 1\}^n$ valued random variable by $\Theta(1/b)$, forcing the higher-order error terms to form a geometric series. One can show that to balance these terms, one may take the threshold at $k = \Theta(\log(1/\epsilon))$ to ensure that this yields ϵ error and seed length $O(\log\log n + \log(1/\epsilon))$ using known explicit constructions of almost k-wise independent distributions.

Our analysis of this fractional PRG is instead driven by Taylor's theorem (see Theorem 2.1). The general approach of using Taylor's theorem in the construction of PRGs has been quite fruitful. The typical way in which it is used is to write some smooth function to be fooled as a low-degree polynomial, which is relatively easy to fool, plus an error term bounded by the next derivatives of the function which ideally is negligible; this approach is often tied to *invariance principles*. We do the same decomposition, which initially agrees with [CHHL19]:

$$|\mathbb{E}_{\mathbf{X}}[f(\mathbf{X})] - \mathbb{E}_{\mathbf{U}}[f(\mathbf{U})]| \leq \underbrace{\sum_{i=1}^{k-1} \sum_{S\subseteq[n]:|S|=i} |\hat{f}(S)||\mathbb{E}_{\mathbf{X}}[\mathbf{X}^S]|}_{\text{low-order terms}} + \underbrace{|\mathbb{E}_{\mathbf{X}}[R_k(\mathbf{X})]|}_{\text{high-order term}}, \tag{3}$$

where our higher order term arises from Taylor's theorem. Naïvely, if one wanted to analyze the quality of a fractional PRG for the multilinear expansion of f using Taylor's theorem, this might do precisely nothing; indeed, the multilinear expansion of f is by definition a polynomial, so is equal to its Taylor series. Therefore, "using Taylor's theorem" would give exactly nothing but the original [CHHL19] approach! To get anything different than [CHHL19], we must truncate the Taylor series up to some level $k-1$ (which exactly agrees up to level $k-1$ with the multilinear expansion) and then study the error term, which is given by some kth order derivatives of the Fourier expansion. The hope in doing this is that implicitly, Taylor's Theorem collapses the higher-order terms into level-k derivatives. One might then hope that there are significant sign cancellations that thus avoid the use of the complete Fourier tail assumption and triangle inequality as in (2). In particular, we only pay the triangle inequality at level-k by doing so, though we now must study these new error terms.

If these derivatives in the error term were evaluated at the origin $\mathbf{0} \in \mathbb{R}^n$, then we would be able to bound the entire error term by any level-k L_1-bounds on our class $\mathcal{F}$. This is because these derivatives at $\mathbf{0}$ are precisely the level-k Fourier coefficients. However, the remainder term is evaluated at some intermediate point on the line between $\mathbf{0}$ and the realization of the supposed fractional PRG $\mathbf{X}$. It turns out, though, that these derivatives have a natural interpretation as $\mathbf{p}$-biased Fourier coefficients for some $\mathbf{p} \in [0,1]^n$ depending on where the derivative is evaluated.

In general, the map $\mathbf{p} \mapsto \widehat{f^{(\mathbf{p})}}(S)$ need not be particularly well-behaved, where $\widehat{f^{(\mathbf{p})}}(S)$ is any $\mathbf{p}$-biased Fourier coefficient. To give control over the error, we exploit multilinearity again. We will be able to first push $\mathbf{p}$ towards $\{1/4, 3/4\}^n$ and only increase the error term. Furthermore, by potentially negating some of our variables and passing to a new function f', we will be able to assume $\mathbf{p} = \frac{1}{4}\cdot\mathbf{1}$ (i.e. all biases are the same and equal to $1/4$). At this point, we use a reduction by [Kel12] that upper-bounds the level-k *biased* Fourier mass by the level-k *unbiased* Fourier mass of some new function g that is obtained by simulating the biased version of f' by replacing each variable of f' with the AND_2 of two new variables. By utilizing the assumed level-k bounds on the set of functions that can be obtained in this way, one can then apply the original [CHHL19] approach of using some version of

$k-1$-wise independence to deal with the lower order terms, and then scale the fractional PRG down by an appropriate factor to fool the error term.

To obtain pseudorandom generators, we then need only apply the random walk gadget of [CHHL19]. In the applications we consider, we need only show that the associated class we need Fourier bounds on are implied by Fourier bounds on the original class. We refer the reader to Section 3 for formal proofs of the ideas sketched in this section.

2 Preliminaries

As in [CHHL19] and [CHLT19], we study PRGs for classes $\mathcal{F}$ of n-variate Boolean functions that are closed under restriction (that is, fixing any subset of the variables yields a function that remains in the class).

2.1 Taylor's Theorem

In order to state the multivariable *Taylor's Theorem*, let us set up some notation. We write $\alpha \in \mathbb{N}^n$ to denote multi-indices, so that

$$\begin{aligned} |\alpha| &\triangleq \sum_{i=1}^{n} \alpha_i \\ \alpha! &\triangleq \alpha_1! \dots \alpha_n! \\ \partial^\alpha &\triangleq \prod_{i=1}^{n} \frac{\partial^{\alpha_i}}{\partial x_i^{\alpha_i}}. \end{aligned}$$

The theorem then asserts the following.

Theorem 2.1 (Taylor Approximation Theorem, see for example [Apo74])**.** *Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is C^∞. Then, we have for any $\mathbf{x}, \mathbf{h} \in \mathbb{R}^n$ and any $k \geq 0$,*

$$f(\mathbf{x}+\mathbf{h}) = \sum_{\alpha \in \mathbb{N}^n : |\alpha| \leq k} \frac{\partial^\alpha f}{\alpha!}(\mathbf{x})\mathbf{h}^\alpha + R_{\mathbf{x},k+1}(\mathbf{h}), \tag{4}$$

where the remainder term $R_{\mathbf{x},k+1}$ is given by

$$R_{\mathbf{x},k+1}(\mathbf{h}) = \sum_{\alpha \in \mathbb{N}^n : |\alpha| = k+1} \frac{\partial^\alpha f}{\alpha!}(\mathbf{x}+c\mathbf{h})\mathbf{h}^\alpha \tag{5}$$

for some $c \in [0,1]$.[2]

2.2 Boolean Functions

We will need the following definitions in our analysis:

Definition 2.1. Let $\mathcal{F}$ be a class of n-variate Boolean functions. Then the *closure of $\mathcal{F}$ under negations*, denoted $\overline{\mathcal{F}}$, is the set of functions $g : \{-1,1\}^n \to \{-1,1\}$ such that there exists $f \in \mathcal{F}$ and some sign vector $\boldsymbol{\epsilon} \in \{-1,1\}^n$ such that $g(\mathbf{x}) = f(\boldsymbol{\epsilon} \circ \mathbf{x})$ for all $\mathbf{x} \in \{-1,1\}^n$, where $\boldsymbol{\epsilon} \circ \mathbf{x} \triangleq (\epsilon_1 x_1, \dots, \epsilon_n x_n)$.

[2] Note that this is sometimes referred to as the *Lagrange form* of the error term.

In words, $\overline{\mathcal{F}}$ is the set of functions that can be constructed by possibly negating some subset of variables in some function in $\mathcal{F}$. Many natural classes of functions are such that they are already closed under negations, i.e. $\mathcal{F} = \overline{\mathcal{F}}$; for instance, if $\mathcal{F}$ is defined as the set of degree-d polynomial threshold functions, clearly this holds. Similarly, if $\mathcal{F}$ is the set of degree-d polynomials considered over $\mathbb{F}_2$, this holds as well.

Definition 2.2. Let $\mathcal{F}$ be a class of n-variate Boolean functions. Then we define $\mathcal{F} \circ \mathsf{AND}_2$ as the set of all functions $g : \{-1,1\}^{2n} \to \{-1,1\}$ such that there exists $f \in \mathcal{F}$ such that for all $\mathbf{x}, \mathbf{y} \in \{-1,1\}^n$,

$$g(\mathbf{x}, \mathbf{y}) = f(\mathsf{AND}_2(x_1, y_1), \ldots, \mathsf{AND}_2(x_n, y_n)). \tag{6}$$

Here, we define $\mathsf{AND}_2 : \{-1,1\}^2 \to \{-1,1\}$ by $\mathsf{AND}_2(x, y) = -1$ if and only if $x = y = -1$. In the case where we instead consider the domain of our Boolean functions as $\mathbb{F}_2^n$, we make the standard identification $+1 \mapsto 0$ and $-1 \mapsto 1$, where the images are viewed as elements of $\mathbb{F}_2$.

That is, $\mathcal{F} \circ \mathsf{AND}_2$ is the set of functions obtained by taking each function in $\mathcal{F}$, and replacing each argument with the AND_2 of two distinct bits; note that such functions are defined on $2n$ variables. By construction, this means that clearly we cannot have $\mathcal{F} = \mathcal{F} \circ \mathsf{AND}_2$. However, for many natural classes, we will be able to exhibit a natural class $\mathcal{G}$ of $2n$-variate Boolean functions containing this set that will itself be similar to $\mathcal{F}$; when this happens, one hopes that level-k Fourier bounds on $\mathcal{F}$ will, with some small loss, lift directly to $\mathcal{G}$. Again, we show how this occurs for each of the original classes studied in [CHHL19] in Section 4.

2.3 Fourier Analysis

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

$$f(\mathbf{x}) = \sum_{S \subseteq [n]} \hat{f}(S) \mathbf{x}^S, \tag{7}$$

where we write $\mathbf{x}^S \triangleq \prod_{i \in S} x_i$. The (unbiased) 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 [O'D14]. 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 (7). Note that in this case, $f(\mathbf{0}) = \hat{f}(\emptyset) = \mathbb{E}_{\mathbf{U}_n}[f(\mathbf{U}_n)]$. The main parameter of interest from the Fourier expansion for our purposes is the following:

Definition 2.3. 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

$$L_{1,k}(\mathcal{F}) \triangleq \max_{f \in \mathcal{F}} L_{1,k}(f).$$

2.3.1 p-biased Analysis

Our analysis will require considering Fourier expansions with respect to a biased product measure. Let $L^2(\{-1,1\}^n, \pi_{p_1} \otimes \ldots \otimes \pi_{p_n})$ be the space of square-integrable functions on the Boolean hypercube with respect to the biased product measure $\mu_{\mathbf{p}} \triangleq \pi_{p_1} \otimes \ldots \otimes \pi_{p_n}$, where π_{p_i} is the biased measure on $\{-1,1\}$ that gives probability p_i to -1 and $q_i \triangleq 1-p_i$ to 1. Then the (unique, up to sign) orthonormal Fourier basis is given by, for each $S \subseteq [n]$

$$\phi_S^{(\mathbf{p})}(\mathbf{x}) \triangleq \prod_{i \in S} \frac{x_i - \mu_i}{\sigma_i}, \tag{8}$$

where $\mu_i \triangleq q_i - p_i$ and $\sigma_i \triangleq 2\sqrt{p_i q_i}$.

Given a vector of biases $\mathbf{p} \in (0,1)^n$, the $\mathbf{p}$-biased Fourier expansion of $f : \{-1,1\}^n \to \{-1,1\}$ is given by

$$f(\mathbf{x}) = \sum_{S \subseteq [n]} \widehat{f^{(\mathbf{p})}}(S) \phi_S^{(\mathbf{p})}(\mathbf{x}), \tag{9}$$

where $\widehat{f^{(\mathbf{p})}}(S) \triangleq \mathbb{E}_{\mathbf{X} \sim \mu_{\mathbf{p}}}[f(\mathbf{x})\phi_S^{(\mathbf{p})}(\mathbf{x})]$. Note that this agrees with the usual Fourier transform, which is taken with biases $1/2$ in each coordinate.

We need the following fact that connects derivatives of the multilinear expansion evaluated at nonzero points and biased Fourier coefficients.

Fact 2.1 (Exercise 8.25 of O'Donnell [O'D14])**.** *Let f be any n-variate Boolean function and identify it with the multilinear expansion. Then for any $\mathbf{p} \in (0,1)^n$ and $S \subseteq [n]$,*

$$\widehat{f^{(\mathbf{p})}}(S) = \Big(\prod_{i \in S} \sigma_i\Big) \partial^S f(\mu_1, \ldots, \mu_n).$$

For completeness, we give a short proof of this fact.

Proof. First, observe that as both the unbiased Fourier expansion (7) and the $\mathbf{p}$-biased expansion (9) are polynomials of degree one in each variable that agree on $\{\pm 1\}^n$, they agree as formal polynomials and therefore in $\mathbb{R}^n$. To extract $\widehat{f^{(\mathbf{p})}}(S)$, it suffices to first apply ∂^S to (9) to keep only the summands for subsets containing S, normalize by multiplying by $\prod_{i \in S} \sigma_i$, and then set $\mathbf{x} = (\mu_1, \ldots, \mu_n)$ to kill off every summand that *strictly* contains S. Applying these operations to (7) via this equality gives the desired claim. □

2.4 (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 [CHHL19]:

Definition 2.4. Let $\mathcal{F}$ be a class of n-variate Boolean functions. Then a *pseudorandom generator* (PRG) for $\mathcal{F}$ with error $\epsilon > 0$ is a random variable $\mathbf{X} \in \{-1,1\}^n$ such that for all $f \in \mathcal{F}$,

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

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

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

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 [CHHL19] and [CHLT19] 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 2.2. *Suppose $\mathcal{F}$ is class of n-variate Boolean functions closed under restrictions, and that $\mathbf{X}$ is a p-noticeable fractional PRG with error ϵ and seed length s. Then there exists an explicit PRG for $\mathcal{F}$ with seed length $O(s \log(n/\epsilon)/p)$ and error $O(\epsilon \log(n/\epsilon)/p)$.*

Using this result, [CHHL19] proved the following theorem that exploits strong L_1 control of each Fourier level:

Theorem 2.3. *Let $\mathcal{F}$ be any class of n-variate Boolean functions closed under restrictions. Suppose that $L_{1,k}(\mathcal{F}) \leq b^k$ for some $b > 0$ and all $1 \leq k \leq n$. Then for any $\epsilon > 0$, there exists an explicit PRG for $\mathcal{F}$ with error ϵ and seed length $b^2 \cdot \mathrm{polylog}(n/\epsilon)$.*

This is achieved by constructing a fractional PRG that is a scaled version of a nearly $\log(1/\epsilon)$-wise independent 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, [CHLT19] derandomize a construction of [RT19] 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 2.4. *Let $\mathcal{F}$ be any class of n-variate Boolean functions closed under restrictions. Suppose that $L_{1,2}(\mathcal{F}) \leq b^2$ for some $b > 0$. Then for any $\epsilon > 0$, 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 Fractional PRGs from the kth Fourier Level

We now turn to the proof of our main result yielding a fractional pseudorandom generator from level-k bounds, Theorem 1.1. We restate the result here:

Theorem 3.1 (Theorem 1.1, restated)**.** *Let $\mathcal{F}$ be any class of n-variate Boolean functions closed under restrictions. Suppose that $\overline{\mathcal{F}} \circ \mathsf{AND}_2$ satisfies $L_{1,k}(\overline{\mathcal{F}} \circ \mathsf{AND}_2) \leq b^k$ for some $b > 0$ and $k > 2$. Then for any $\epsilon > 0$, there exists a $\Omega(\epsilon^{2/k}/b^2)$-noticeable explicit fractional PRG for $\mathcal{F}$ with error ϵ and seed length $O(k \cdot \log(n))$.*

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 $O(\log\log(n) + \log k + \log(1/\epsilon))$.

To set up the proof of this theorem, we require some auxilliary results. Recall from the proof overview above that through our use of Taylor's theorem (Theorem 2.1), we will encounter *biased* Fourier coefficients. Ideally, we would like to relate these coefficients to unbiased coefficients we better understand. We do so using the following reduction of Keller [Kel12], which gives an upper bound of biased Fourier coefficients using the unbiased Fourier coefficients of the natural $2n$-variate function that simulates the biased bits:

Theorem 3.2 (Theorem 1.2 in [Kel12]). *Consider the hypercube $\{-1,1\}^n$ with bias $p = t/2^m$. For any function $f : \{-1,1\}^n \to \mathbb{R}$, there exists a function $g : \{-1,1\}^{mn}$ where*

$$g(y_1, \ldots, y_{mn}) = f(h(y_1, \ldots, y_m), \ldots, h(y_{m(n-1)+1}, \ldots y_{mn})),$$

where h is some function on m bits, such that for any $k \leq n$,

$$\sum_{S \subseteq [n] : |S| = k} |\hat{f}^{(p)}(S)| \leq \left(\sqrt{\frac{1-p}{p \log^2(1/p)}}\right)^k \sum_{T \subseteq [mn] : |T| = k} |\hat{g}^{(1/2)}(T)|.$$

Note that the Fourier coefficients on the left are in the p-biased orthonormal basis (where all biases are the same in each component), while the Fourier coefficients on the right are with respect to the unbiased measure. In the case that $p = 1/2^2$, we may take $h = \mathsf{AND}_2$.[3]

However, the use of Taylor's theorem will lead to somewhat arbitrary biases in each coordinate separately, making the reduction afforded in the previous result difficult to directly apply. To overcome this, we exploit two simple, but useful facts. The first simply asserts that, when controlling the L_1 Fourier masses that arise, possibly with respect to a biased product measure, one may safely assume that all biases are at most $1/2$ by modifying the functions we encounter by negating some variables. Intuitively, this is clear: by negating a variable and then flipping the bias about $1/2$, the biased Fourier coefficients with respect to the new biases will not change (except for possibly a sign). This is encoded by the following lemma.

Lemma 3.3. *Let $\mathbf{p} \in (0,1)^n$ be a vector of biases for the $\mathbf{p}$-biased hypercube. Fix $i \in [n]$ and consider $\mathbf{p}'$ given by $\mathbf{p}$ but setting $\mathbf{p}'_i = 1 - \mathbf{p}_i$. Then for any function $f : \{-1,1\}^n \to \mathbb{R}$, writing $f'(x_1, \ldots, x_n) := f(x_1, \ldots, -x_i, \ldots, x_n)$, we have for any $S \subseteq [n]$*

$$|\widehat{f^{(\mathbf{p})}}(S)| = |\widehat{f'^{(\mathbf{p}')}}(S)|. \tag{10}$$

Proof. Suppose $i \in S$. From [O'D14, page 212], we can write (recalling $q_j = 1 - p_j$),

$$\begin{aligned}
\widehat{f^{(\mathbf{p})}}(S) &= \mathbb{E}_{\mathbf{x} \sim \mu_{\mathbf{p}}}\left[f(x_1, \ldots, x_n)\left(\prod_{j \in S} \frac{x_j - (q_j - p_j)}{2\sqrt{p_j(1-p_j)}}\right)\right] \\
&= \mathbb{E}_{\mathbf{x} \sim \mu_{\mathbf{p}'}}\left[f(x_1, \ldots, -x_i, \ldots, x_n)\left(\prod_{j \in S \setminus \{i\}} \frac{x_j - (q_j - p_j)}{2\sqrt{p_j(1-p_j)}}\right)\left(\frac{-x_i - (q_i - p_i)}{2\sqrt{p_i(1-p_i)}}\right)\right] \\
&= -\mathbb{E}_{\mathbf{x} \sim \mu_{\mathbf{p}'}}\left[f(x_1, \ldots, -x_i, \ldots, x_n)\left(\prod_{j \in S \setminus \{i\}} \frac{x_j - (q_j - p_j)}{2\sqrt{p_j(1-p_j)}}\right)\left(\frac{x_i - (p_i - q_i)}{2\sqrt{p_i(1-p_i)}}\right)\right] \\
&= -\mathbb{E}_{\mathbf{x} \sim \mu_{\mathbf{p}'}}\left[f(x_1, \ldots, -x_i, \ldots, x_n)\left(\prod_{j \in S \setminus \{i\}} \frac{x_j - (q_j - p_j)}{2\sqrt{p_j(1-p_j)}}\right)\left(\frac{x_i - (q'_i - p'_i)}{2\sqrt{p_i(1-p_i)}}\right)\right] \\
&= -\mathbb{E}_{\mathbf{x} \sim \mu_{\mathbf{p}'}}\left[f(x_1, \ldots, -x_i, \ldots, x_n)\left(\prod_{j \in S} \frac{x_j - (q'_j - p'_j)}{2\sqrt{p'_j(1-p'_j)}}\right)\right] \\
&= -\mathbb{E}_{\mathbf{x} \sim \mu_{\mathbf{p}'}}\left[f'(x_1, \ldots, x_n)\left(\prod_{j \in S} \frac{x_j - (q'_j - p'_j)}{2\sqrt{p'_j(1-p'_j)}}\right)\right] \\
&= -\widehat{f'^{(\mathbf{p}')}}(S),
\end{aligned}$$

[3] See the bottom of page 6 of [Kel12]. Keller does the computation with the usual, squared Fourier weight, but by inspecting the proof, the same result holds with the L^1 mass, up to a squareroot in the bias factor.

where we use the fact that $p'_j = p_j$ for all $j \neq i$, the fact that the denominators are invariant under flipping biases, and then the definition of f'. The case where $i \notin S$ is simpler; the Fourier character does not change under flipping biases, so the desired equation holds even without the absolute values. □

Corollary 3.4. *Let $\mathbf{p} \in (0,1)^n$ be a vector of biases for the $\mathbf{p}$-biased hypercube. Then for any function $f : \{-1,1\}^n \to \mathbb{R}$, there exists some function f' such that for some fixed signs $\epsilon \in \{-1,1\}^n$, $f'(x_1, \ldots, x_n) = f(\epsilon_1 x_1, \ldots, \epsilon_n x_n)$, with the property that for any $S \subseteq [n]$*

$$|\widehat{f^{(\mathbf{p})}}(S)| = |\widehat{f'^{(\mathbf{p}')}}(S)|,$$

where $\mathbf{p}'$ is the result of flipping each bias in $\mathbf{p}$ so that it is at most $1/2$.

Proof. Apply the previous proposition each time one flips an extra bias about $1/2$, noting the absolute values are preserved each time, until each bias is at most $1/2$. □

The second obvious property we exploit is multilinearity of the unbiased Fourier expansion, which extends to every partial derivative as well. This will enable us to "push" the biases in Taylor's theorem towards points that will be amenable to using the [Kel12] result. The technical claim we require is the following:

Lemma 3.5. *Let $\mathbf{X} \in \mathbb{R}^n$ be any random variable such that $|\mathbf{X}_i| \leq x$ for some $x \geq 0$ and all i. Let $\{h_S\}$ be arbitrary multilinear functions where the index ranges over $\{S \subseteq [n] : |S| = k\}$, and suppose that $\mathbf{C}(\mathbf{X}) \in \mathbb{R}$ is any random variable depending on $\mathbf{X}$ that takes values in $[0,1]$. Then for any $y \geq x$,*

$$\left|\mathbb{E}_{\mathbf{X}}\left[\sum_{S\subseteq[n]:|S|=k} h_S(\mathbf{C}(\mathbf{X})\cdot\mathbf{X})\mathbf{X}^S\right]\right| \leq x^k \max_{\mathbf{z}\in\{-y,y\}^n} \sum_{S\subseteq[n]:|S|=k} |h_{\mathbf{S}}(\mathbf{z})|. \tag{11}$$

Proof.

$$\begin{aligned}
\left|\mathbb{E}_{\mathbf{X}}\left[\sum_{S\subseteq[n]:|S|=k} h_S(\mathbf{C}(\mathbf{X})\cdot\mathbf{X})\mathbf{X}^S\right]\right| &\leq \mathbb{E}_{\mathbf{X}}\left[\left|\sum_{S\subseteq[n]:|S|=k} h_S(\mathbf{C}(\mathbf{X})\cdot\mathbf{X})\mathbf{X}^S\right|\right] \\
&\leq \mathbb{E}_{\mathbf{X}}\left[\sup_{\mathbf{z}\in[-y,y]^n}\left|\sum_{S\subseteq[n]:|S|=k} h_S(\mathbf{z})\mathbf{X}^S\right|\right] \\
&= \mathbb{E}_{\mathbf{X}}\left[\mathbb{E}\left[\sup_{\mathbf{z}\in[-y,y]^n}\left|\sum_{S\subseteq[n]:|S|=k} h_S(\mathbf{z})\mathbf{X}^S\right|\middle|\mathbf{X}\right]\right] \\
&= \mathbb{E}_{\mathbf{X}}\left[\mathbb{E}\left[\max_{\mathbf{z}\in\{-y,y\}^n}\left|\sum_{S\subseteq[n]:|S|=k} h_S(\mathbf{z})\mathbf{X}^S\right|\middle|\mathbf{X}\right]\right] \\
&\leq x^k\mathbb{E}_{\mathbf{X}}\left[\mathbb{E}\left[\max_{\mathbf{z}\in\{-y,y\}^n}\sum_{S\subseteq[n]:|S|=k} |h_S(\mathbf{z})|\middle|\mathbf{X}\right]\right] \\
&= x^k \max_{\mathbf{z}\in\{-y,y\}^n}\sum_{S\subseteq[n]:|S|=k} |h_S(\mathbf{z})|.
\end{aligned}$$

The second equality holds because for any fixing of $\mathbf{X}$, the sum becomes a linear combination of multilinear functions in the argument $\mathbf{z}$, which remains multilinear. Therefore, the maximizer lies in $\{-y,y\}^n$ (if the maximum is attained with a coordinate in $(-y,y)$, fixing the rest of the coordinates gives a linear function in this coordinate, which can thus be pushed to $\{-y,y\}$ and not decrease the absolute value of the expression). At this point, we may apply the triangle inequality, yielding the claim. □

With these technical results in place, we now have everything we need to prove Theorem 1.1.

Proof of Theorem 1.1. Fix $f \in \mathcal{F}$. Then for any random variable $\mathbf{X}$, we have by Taylor's theorem (Theorem 2.1) and the fact that the multilinear expansion up to level-k agrees with the Taylor series that

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

The first term will be treated as in [CHHL19] depending on the assumptions on the lower levels of the Fourier spectrum, so we turn to bounding the latter term. As f is multilinear, all partial derivatives with a repeated partial derivative in any component is identically zero. Therefore,

$$\mathbb{E}_{\mathbf{X}}[R_{\mathbf{0},k}(\mathbf{X})] = \mathbb{E}_{\mathbf{X}}\left[\sum_{T\subseteq[n]:|T|=k} \partial^T f(\mathbf{C}(\mathbf{X})\cdot\mathbf{X})\mathbf{X}^T\right] \tag{12}$$

where $\mathbf{C}(\mathbf{X})$ is a random variable depending on $\mathbf{X}$ in $[0,1]$ by the remainder part of Taylor's Theorem. Because each of the partial derivatives of a multilinear function remains multilinear, we may apply Lemma 3.5 to deduce

$$|\mathbb{E}_{\mathbf{X}}[R_{\mathbf{0},k}(\mathbf{X})]| \le x^k \max_{\mathbf{z}\in\{-1/2,1/2\}^n} \sum_{T\subset[n]:|T|=k} |\partial^T f(\mathbf{z})|, \tag{13}$$

where we will ensure that we construct $\mathbf{X}$ such that $|\mathbf{X}_i| \le x \le 1/2$.

Fixing $\mathbf{z}^* \in \{-1/2, 1/2\}^n$ as the maximizer of the right hand side, we may now apply Fact 2.1. This yields that for some $\mathbf{p} \in \{1/4, 3/4\}^n$

$$|\mathbb{E}_{\mathbf{X}}[R_{\mathbf{0},k}(\mathbf{X})]| \le \left(\frac{2\sqrt{3}x}{3}\right)^k \sum_{T\subset[n]:|T|=k} |\widehat{f^{(\mathbf{p})}}(S)|, \tag{14}$$

as $\sigma_i = \sqrt{3}/2$ for each $i \in [n]$ with these biases. Now applying Corollary 3.4, by changing f to some $h \in \overline{\mathcal{F}}$ by possibly negating some variables, we can assume $\mathbf{p} = (1/4, \dots, 1/4)$, so that

$$|\mathbb{E}_{\mathbf{X}}[R_{\mathbf{0},k}(\mathbf{X})]| \le \left(\frac{2\sqrt{3}x}{3}\right)^k \sum_{T\subset[n]:|T|=k} |\widehat{h^{(\mathbf{1/4})}}(S)| \tag{15}$$

As $\mathbf{p}_i \le 1/2$, we may apply Theorem 3.2, finally obtaining

$$|\mathbb{E}_{\mathbf{X}}[R_{\mathbf{0},k}(\mathbf{X})]| \le \big(Cx\big)^k \sum_{T\subset[n]:|T|=k} |\widehat{g}(S)|, \tag{16}$$

where $g : \{-1,1\}^{2n} \to \{-1,1\}$ is obtained by simulating each $1/4$-biased variable in h using two unbiased bits via AND_2, and C is some absolute constant. In particular, by following these transformations, $g \in \overline{\mathcal{F}} \circ \mathsf{AND}_2$, and so by our assumption on this class, we have

$$|\mathbb{E}_{\mathbf{X}}[R_{\mathbf{0},k}(\mathbf{X})]| \le \big(Cbx\big)^k \tag{17}$$

for some constant $C > 0$. Therefore, if $x = \Theta(\epsilon^{1/k}/b)$, this part of the error is at most $\epsilon/2$.

To deal with the lower order terms, we take the same approach as [CHHL19]. If no assumptions are made on the level-j mass for $j < k$, then one may take $\mathbf{X} = x \cdot \mathbf{Y}$, where $\mathbf{Y} \in \{-1, 1\}^n$ is $k-1$-wise independent to incur zero error on the lower order terms. This has seed length $O(k \log(n))$ [Vad12] and gives no additional error.

If one assumes a bound of b^i for all even $i < k$, then one may apply their same analysis and instead let $\mathbf{X} = x \cdot \mathbf{Y}'$, where $\mathbf{Y}'$ is an $\epsilon/2$-almost $k-1$-wise independent distribution; the same computation as done there shows that the error incurred by this distribution on the low order terms is bounded by $\epsilon/2$, with seed length $O(\log\log(n) + \log k + \log(1/\epsilon))$ [NN93]. Combining these two errors yields the claim. □

3.1 From Fractional Generators to PRGs

Using Theorem 1.1 and Theorem 2.2, it is fairly immediate to obtain PRGs that rely only on level-k bounds. Similarly, bounds on levels up to k can be leveraged to get an improved seed length.

Theorem 3.6 (Theorem 1.2, restated)**.** *Let $\mathcal{F}$ be any class of n-variate Boolean functions closed under restrictions. Suppose that $\overline{\mathcal{F}} \circ \mathsf{AND}_2$ satisfies $L_{1,k}(\overline{\mathcal{F}} \circ \mathsf{AND}_2) \leq b^k$ for some $b > 0$ and $k > 2$. Then for any $\epsilon > 0$, there exists an explicit PRG for $\mathcal{F}$ with error ϵ with seed length*

$$s = O\left(\frac{b^{2+\frac{4}{k-2}} \cdot k \cdot \log(n) \log^{1+\frac{2}{k-2}}(n/\epsilon)}{\epsilon^{\frac{2}{k-2}}}\right). \tag{18}$$

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

$$s = O\left(\frac{b^{2+\frac{4}{k-2}} \cdot (\log\log n + \log k + \log(b/\epsilon)) \cdot \log^{1+\frac{2}{k-2}}(n/\epsilon)}{\epsilon^{\frac{2}{k-2}}}\right). \tag{19}$$

Proof. By Theorem 2.2, 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 $O(s \log(n/\epsilon)/p)$.

For the first statement, by our assumption and using the fractional PRG guaranteed by Theorem 1.1, 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

$$\delta = \Theta\left(\left(\frac{\epsilon}{b^2 \log(n/\epsilon)}\right)^{\frac{k}{k-2}}\right) \tag{20}$$

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

$$s = O\left(\frac{b^{2+\frac{4}{k-2}} \cdot k \cdot \log(n) \log^{1+\frac{2}{k-2}}(n/\epsilon)}{\epsilon^{\frac{2}{k-2}}}\right). \tag{21}$$

The second statement follows in an identical manner from using the improved seed length from the second part of Theorem 1.1 in the case that one has control on the Fourier mass on the lower levels. □

Corollary 1.1 is now an immediate consequence of Theorem 3.6; for any desired $\epsilon > b \cdot \log(n) \cdot 2^{-\Omega(k)}$, one can simply apply Theorem 3.6 using level $\ell = \Theta(\log(b \log(n)/\epsilon))$ to obtain a PRG for $\mathcal{F}$ with error at most ϵ with seed length

$$s = O(b^2 \cdot \log(b \log(n)/\epsilon) \cdot \log(n/\epsilon)). \tag{22}$$

Note that if $\epsilon = 1/\mathrm{poly}(n)$, then this means that 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 [CHLT19], which asks how many levels of Fourier bounds suffice to recover polylogarithmic dependence in $1/\varepsilon$.

Remark 3.7. *Note that this Taylor approach does not yield anything nontrivial given just level-two bounds, unlike the fractional generator in [CHLT19]. This is actually a necessary byproduct of combining this approach with the random walk gadget of [CHHL19]. 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 trivially 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 will necessarily polarize to $\pm\mathbf{1}$ at termination, which clearly cannot fool nontrivial functions.*

4 Applications

In this section, we demonstrate how for many natural classes, level-k bounds on $\overline{F} \circ \mathsf{AND}_2$ can often be reduced to level-k bounds on a suitable $2n$-variate version of the original function class. As a result, level-k bounds on $\mathcal{F}$ can often lift nearly directly to level-k bounds on the classes we need in our fractional PRG analysis with some small loss. For most of the classes we consider, the class satisfies L_1 Fourier tail bounds at all levels, and thus our analysis gives nothing new over [CHHL19]. Rather, what we emphasize in this section is that our analysis permits pseudorandom generators in the random walks framework even if these bounds were only known at or up to some fixed level-k, *not on all levels*. However, for the case of low-degree polynomials over $\mathbb{F}_2$, our analysis will actually imply a new nontrivial PRG with polylogarithmic error dependence, thus partially resolving an open question of [CHHL19].

4.1 Low-Degree Polynomials over $\mathbb{F}_2$

Let $\mathcal{F}$ be the set of n-variate, degree-d polynomials over $\mathbb{F}_2$. Note that considered with domain $\mathbb{F}_2^2$, the function AND_2 is nothing but the usual product of the inputs. In particular, because clearly $\mathcal{F} = \overline{\mathcal{F}}$, it follows that $\overline{\mathcal{F}} \circ \mathsf{AND}_2 \subseteq \mathcal{G}$, where we write $\mathcal{G}$ for the set of $2n$-variate polynomials over $\mathbb{F}_2$ with degree-$2d$. As a preliminary step towards deriving Fourier tail bounds that would imply a nontrivial PRG for this class using their framework, [CHHL19] prove the following Fourier bounds:

Proposition 4.1 (Theorem 6.1 of [CHHL19])**.** *Let $p : \mathbb{F}_2^n \to \mathbb{F}_2$ 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 $k = \Omega(\sqrt{n})$. While Theorem 2.4 can yield a nontrivial PRG by just applying the level-two bound, the dependence on $1/\epsilon$ is at least quadratic.[4] However, using our new, more flexible analysis, one can obtain a nontrivial PRG with polylogarithmic dependence on the seed length. Our formal result is the following:

Theorem 4.2. *Let $\mathcal{F}$ be the class of degree at most 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*

$$s = O(2^{O(d)} \cdot \log^3(\log(n)/\varepsilon) \cdot \log(n/\varepsilon)). \tag{23}$$

Proof. Fix $\varepsilon > 0$. Let $\mathcal{G}$ be the class of degree-$2d$ polynomials on $2n$ variables, and let $k = \Theta(\log(\log(n)/\varepsilon))$. By Proposition 4.1, we have that for all $j \leq k$,

$$L_{1,k}(\mathcal{G}) \leq (\Theta(\log(\log(n)/\varepsilon) \cdot 2^{6d}))^j. \tag{24}$$

[4] By applying this Fourier bound at level-two, one can use the fractional PRG of [CHLT19] to obtain seed length $2^{O(d)}\mathrm{polylog}(n)/\epsilon^{2+o(1)}$ using the random walks framework. This gives exponentially worse error dependence compared to our approach.

As $\mathcal{F}$ clearly embeds in $\mathcal{G}$ trivially with the same Fourier coefficients by simply omitting variables, this bound extends to $\mathcal{F}$ as well. In particular, setting $b = \Theta(\log(\log(n)/\varepsilon) \cdot 2^{6d})$, we may apply Theorem 3.6 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 [Vio09], obtained by summing d independent copies of a sufficiently good biased space, attains seed length $d \cdot \log(n) + O(d \cdot 2^d \log(1/\varepsilon))$, which for constant ε and d is within an additive constant of the optimal possible seed. 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 walks 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 polylogarithmic seed length, our work shows that the following weaker conjecture would suffice to break the logarithmic degree barrier and still obtain polylogarithmic error dependence:

Conjecture 4.3. *Let $\mathcal{F}$ be the class of degree-d polynomials over $\mathbb{F}_2$ on n variables. Then*

$$L_{1,k}(\mathcal{F}) \leq (\mathrm{poly}(k, \log n) \cdot 2^{o(d)})^k \tag{25}$$

for all $1 \leq k \leq n$.

4.2 Other Families with Level-k Fourier Bounds

For the remaining applications of [CHHL19], we do not obtain anything new since bounds are known for all the Fourier levels. However, we include these examples below to show that Fourier bounds for $\mathcal{F}$ often immediately apply to Fourier bounds on $\overline{\mathcal{F}} \circ \mathsf{AND}_2$.

Bounded-sensitivity functions Let $\mathcal{F}$ be the set of n-variate functions with sensitivity at most s. It is easy to see that $\overline{\mathcal{F}} = \mathcal{F}$. Moreover, it is simple to observe that $\mathcal{F} \circ \mathsf{AND}_2$ consists of functions of sensitivity at most $2s$ (see, for instance, [Tal13]). It is known that $L_{1,k}(\mathcal{F}) \leq O(s)^k$ [GSW16]. As our reduction maps such functions to functions with bounded sensitivity, this bound extends to $\mathcal{F} \circ \mathsf{AND}_2$ with a small change in the implied constant.

Note that as a consequence of the recent proof of the Sensitivity Conjecture [Hua19], it is immediate that s^2-wise independence completely fools functions with sensitivity at most s.

Small-Depth Circuits Similarly, it is known that for the class $\mathcal{F}$ of size m, depth d $\mathbf{AC^0}$ circuits with n input variables, we have $L_{1,k}(\mathcal{F}) \leq O(\log^{d-1}(m))^k$ [LMN89, Tal17]. It is easy to observe that the class $\overline{\mathcal{F}} \circ \mathsf{AND}_2$ is a subset of the class of size $m + 2n$, depth $d + 1$ $\mathbf{AC^0}$ circuits with $2n$ input variables and thus follow similar level-k Fourier bounds. As a result, we can construct a PRGs for $\mathbf{AC^0}$ using Theorem 1.2.

Read-Once Branching Programs For the class of read-once branching programs of length n and width w, it is known that $L_{1,k}(\mathcal{F}) \leq O(\log^w n)^k$ [CHRT18]. It is easy to observe that the class $\overline{\mathcal{F}} \circ \mathsf{AND}_2$ is a subset of the class of read-once branching programs of length $2n$ and width w, and hence also follows the required Fourier level bounds to be applicable in Theorem 1.2.

5 Discussion

In this work, we have given a partial interpolation between the PRGs obtained in the polarizing random walks framework by exploiting level-k bounds on a similar class of functions, thus partially answering an open question from [CHLT19]. We do so by exploiting Taylor's Theorem and an alternate Fourier analysis, passing through p-biased coefficients to obtain the result. As mentioned, we hope this paper gives more evidence that intermediate levels of Fourier control can yield improved PRGs using this new flexible framework.

One immediate question is, given just Fourier bounds on level-k of some class $\mathcal{F}$, can one deduce similar results, without needing bounds on the related class $\overline{\mathcal{F}} \circ \mathsf{AND}_2$? It would be ideal to obtain comparable results without needing to go through this auxiliary class, though in the known applications this can be done with little loss. Along these lines, [GRZ20] have now generalized the analysis of [RT19] to level-$2k$, which was itself the basis for the level-two results in [CHLT19]. However, we are currently unaware both if their analysis can be derandomized or how to do so to obtain a PRG.

A related question is, given such Fourier bounds, can one directly obtain Fourier bounds for $\overline{\mathcal{F}} \circ \mathsf{AND}_2$? We are unaware of how to do this in general; indeed, for the spectrally extremely simple parity function, this transformation can transform it to the inner product function, which is bent. However, we stress that for most reasonable, computational classes, this image of this transformation will typically lie in an appropriate $2n$-variate version of the class, which one expects should admit comparable level-k bounds (as explained in Section 4).

References

[Apo74] Tom M. Apostol. *Mathematical analysis*. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., second edition, 1974.

[CHHL19] Eshan Chattopadhyay, Pooya Hatami, Kaave Hosseini, and Shachar Lovett. Pseudorandom generators from polarizing random walks. *Theory of Computing*, 15(10):1–26, 2019.

[CHLT19] 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.

[CHRT18] 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.

[GRZ20] Uma Girish, Ran Raz, and Wei Zhan. Lower bounds for XOR of Forrelations. *arXiv preprint arXiv:2007.03631*, 2020.

[GSW16] 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.

[Hua19] Hao Huang. Induced subgraphs of Hypercubes and a proof of the Sensitivity Conjecture. *Annals of Mathematics*, 190(3):949–955, 2019.

[Kel12] Nathan Keller. A simple reduction from a biased measure on the discrete cube to the uniform measure. *European Journal of Combinatorics*, 33(8):1943 – 1957, 2012.

[LMN89] 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.

[NN93] Joseph Naor and Moni Naor. Small-bias probability spaces: Efficient constructions and applications. *SIAM Journal on Computing*, 22(4):838–856, 1993.

[O'D14] Ryan O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, 2014.

[Raz87] A. 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, Apr 1987.

[RT19] Ran Raz and Avishay Tal. Oracle separation of BQP and PH. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, page 13–23, New York, NY, USA, 2019. Association for Computing Machinery.

[Smo87] R. 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.

[Smo93] R. 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.

[Tal13] Avishay Tal. Properties and applications of Boolean function composition. In *Proceedings of the 4th Conference on Innovations in Theoretical Computer Science*, ITCS '13, page 441–454, New York, NY, USA, 2013. Association for Computing Machinery.

[Tal17] Avishay Tal. Tight bounds on the Fourier spectrum of AC0. In *32nd Computational Complexity Conference (CCC 2017)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2017.

[Vad12] Salil P Vadhan. Pseudorandomness. *Foundations and Trends in Theoretical Computer Science*, 7(1–3):1–336, 2012.

[Vio09] Emanuele Viola. The sum of d small-bias generators fools polynomials of degree d. *Computational Complexity*, 18(2):209–217, 2009.

[Wu20] Xinyu Wu. A stochastic calculus approach to the oracle separation of BQP and PH. *arXiv preprint arXiv:2007.02431*, 2020.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il