Pseudorandom Generators from the Second Fourier Level and Applications to AC0 with Parity Gates

Eshan Chattopadhyay
Department of Computer Science, Cornell University, 107 Hoy Rd, Ithaca, NY, USA
eshanc@cornell.edu

Pooya Hatami[1]
Department of Computer Science, University of Texas at Austin, 2317 Speedway, Austin, TX, USA
pooyahat@gmail.com

Shachar Lovett[2]
Department of Computer Science, University of California San Diego, La Jolla, CA, USA
slovett@ucsd.edu

Avishay Tal[3]
Department of Computer Science, Stanford University, 353 Serra Mall, Stanford, CA, USA
avishay.tal@gmail.com

Abstract

A recent work of Chattopadhyay et al. (CCC 2018) introduced a new framework for the design of pseudorandom generators for Boolean functions. It works under the assumption that the Fourier tails of the Boolean functions are uniformly bounded for all levels by an exponential function. In this work, we design an alternative pseudorandom generator that only requires bounds on the second level of the Fourier tails. It is based on a derandomization of the work of Raz and Tal (ECCC 2018) who used the above framework to obtain an oracle separation between BQP and PH.

As an application, we give a concrete conjecture for bounds on the second level of the Fourier tails for low degree polynomials over the finite field $\mathbb{F}_2$. If true, it would imply an efficient pseudorandom generator for $\mathrm{AC}^0[\oplus]$, a well-known open problem in complexity theory. As a stepping stone towards resolving this conjecture, we prove such bounds for the first level of the Fourier tails.

2012 ACM Subject Classification Theory of computation → Pseudorandomness and derandomization

Keywords and phrases Derandomization, Pseudorandom generator, Explicit construction, Random walk, Small-depth circuits with parity gates

Digital Object Identifier 10.4230/LIPIcs.ITCS.2019.22

Acknowledgements We would like to thank Gil Cohen, Russell Impagliazzo, Valentine Kabanets, James Lee, Ran Raz, Rahul Santhanam, Roy Schwartz and Srikanth Srinivasan for very helpful conversations.

[1] Supported by a Simons Investigator Award #409864, David Zuckerman.
[2] Supported by NSF grant CCF-1614023.
[3] Supported by a Motwani Postdoctoral Fellowship and by NSF grant CCF-1763299.

© Eshan Chattopadhyay, Pooya Hatami, Shachar Lovett, and Avishay Tal;
licensed under Creative Commons License CC-BY
10th Innovations in Theoretical Computer Science (ITCS 2019).
Editor: Avrim Blum; Article No. 22; pp. 22:1–22:15
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

Pseudorandom generators are widely studied in computational complexity theory. The main focus of this paper is a new framework for the design of pseudorandom generators (abbrv. PRGs) based on Fourier tails, introduced recently by Chattopadhyay et al. [2]. We refer to the survey of Vadhan [14] for an introduction to pseudorandomness in complexity theory, and assume basic knowledge with common concepts.

Let $\mathcal{F}$ be a family of n-variate Boolean functions, which is closed under restrictions. Namely, for any $f \in \mathcal{F}$, if we restrict some of the inputs of f to Boolean values, then the restricted function is also in $\mathcal{F}$. Nearly all classes of Boolean functions studied in the literature satisfy this property.

Given an n-variate Boolean function f, its level-k Fourier tails for $k = 1, \ldots, n$ are defined as

$$\mathcal{L}_{1,k}(f) = \sum_{S \subset [n] : |S| = k} |\widehat{f}(S)|.$$

For a function class $\mathcal{F}$ of n-variate Boolean functions define

$$\mathcal{L}_{1,k}(\mathcal{F}) = \max_{f \in \mathcal{F}} \mathcal{L}_{1,k}(f).$$

Chattopadhyay et al. [2] proved a general theorem, which constructs an explicit PRG for functions in $\mathcal{F}$, assuming that $\mathcal{F}$ has bounded k-level Fourier tails for all k. This property is known to hold for many classes of interest (read-once branching programs of bounded width, low-depth circuits, low sensitivity functions, and more; see [2] for details).

▶ **Theorem 1** ([2]). *Let $\mathcal{F}$ be a family of n-variate Boolean functions that is closed under restrictions. Assume that for some $a, b \geq 1$ it holds that*

$$\mathcal{L}_{1,k}(\mathcal{F}) \leq a \cdot b^k \qquad \forall k = 1, \ldots, n.$$

Then for any $\varepsilon > 0$, there exists an explicit PRG for $\mathcal{F}$ with error ε and seed length $s = b^2 \cdot \mathrm{polylog}(an/\varepsilon)$.

Note that for any n-variate Boolean function one can take $a = 1, b = \sqrt{n}$, and hence the quadratic dependence on b in the seed length is optimal.

The main objective of this current work is to investigate whether PRGs can also be obtained from weaker assumptions on the Fourier tail. Specifically, whether it suffices that $\mathcal{L}_{1,k}(\mathcal{F})$ is bounded for a few values of k, instead of for the full regime of $k = 1, \ldots, n$ as was required by [2]. Our main result is that this is indeed the case: it suffices to obtain bounds for the second level of the Fourier tail.

▶ **Theorem 2** (Main result, informal version). *Let $\mathcal{F}$ be a family of n-variate Boolean functions closed under restrictions. Assume that for some $t \geq 1$ it holds that*

$$\mathcal{L}_{1,2}(\mathcal{F}) \leq t.$$

Then for any $\varepsilon > 0$, there exists an explicit PRG for $\mathcal{F}$ with error ε and seed length $\mathrm{poly}(t, \log n, 1/\varepsilon)$.

For a more precise formula for the seed length see Theorem 6. We note that the dependence on the error parameter ε in Theorem 2 is much worse compared to Theorem 1 – polynomial instead of poly-logarithmic. We discuss this in Section 4.

1.1 A potential PRG for $\mathbb{F}_2$-polynomials and $\mathsf{AC}^0[\oplus]$

There are known deep relationships between the ability to construct explicit pseudorandom generators, and the ability to prove correlation bounds, for many classes of Boolean functions. One of the few classes where the latter is known but the former is not is $\mathsf{AC}^0[\oplus]$, which is the classes of constant-depth polynomial-size Boolean circuits with AND, OR, NOT and PARITY gates. Classical works of Razborov [10] and Smolensky [11] prove that this class cannot approximate the MAJORITY function. On the other hand, the problem of constructing explicit PRGs for $\mathsf{AC}^0[\oplus]$ is a well-known open problem in complexity theory. We refer to the survey of Viola [15] for further discussion on this challenge.

We give a concrete (and plausible in our minds) conjecture which, combined with Theorem 2, would imply such a PRG. Let $\mathrm{Poly}_{n,d}$ denote the class of n-variate Boolean functions which are computed by $\mathbb{F}_2$-polynomials of degree at most d. This class is clearly closed under restrictions.

▶ **Conjecture 3.** $\mathcal{L}_{1,2}(\mathrm{Poly}_{n,d}) = O(d^2)$. *That is, if $p : \mathbb{F}_2^n \to \mathbb{F}_2$ is a polynomial of degree d, and $f(x) = (-1)^{p(x)}$, then*

$$\sum_{i,j\in[n],i<j} |\widehat{f}(\{i,j\})| = O(d^2).$$

A corollary of Conjecture 3, when combined with Theorem 2, is the construction of explicit PRGs for degree-d polynomials over $\mathbb{F}_2$ with seed length $\mathrm{poly}(\log n, d, 1/\varepsilon)$. This would be a major breakthrough in complexity theory, as currently no PRGs are known for polynomials of degree $d = \Omega(\log n)$. We note that a similar seed length would follow from a weaker version of Conjecture 3 with the bound $\mathcal{L}_{1,2}(\mathrm{Poly}_{n,d}) \le \mathrm{poly}(\log n, d)$. However, we conjecture that $O(d^2)$ is the correct bound.

We further note that such PRGs would directly imply PRGs for $\mathsf{AC}^0[\oplus]$.

▶ **Claim 4.** *Assume that Conjecture 3 holds. Then, for any $\varepsilon > 0$ there exists an explicit PRG for the class of $\mathsf{AC}^0[\oplus]$-circuits of size s and depth e on n inputs, with error ε and seed length* $\mathrm{poly}((\log(s/\varepsilon))^e, \log n, 1/\varepsilon)$.

In particular, assuming the conjecture, for any constants e and c, $\mathsf{AC}^0[\oplus]$-circuits of size at most n^c and depth at most e have a PRG with seed length $\mathrm{poly}(\log n, 1/\varepsilon)$ and error ε.

Proof sketch. Let $f : \{0,1\}^n \to \{0,1\}$ be computed by an $\mathsf{AC}^0[\oplus]$ circuit of size s and depth e. Razborov [10] and Smolensky [12] proved that there exists a distribution over polynomials $p : \mathbb{F}_2^n \to \mathbb{F}_2$ of degree $d = O(\log(s/\varepsilon))^e$ such that for each $x \in \{0,1\}^n$, $\mathbf{Pr}_p[p(x) \neq f(x)] \le \varepsilon/3$. Theorem 2 gives a PRG for polynomials of degree d with error $\varepsilon/3$ and seed length $\mathrm{poly}(\log n, d, 1/\varepsilon)$. By the Razborov-Smolensky result, this PRG is also a PRG for f with error ε. ◀

Evidence Supporting The Conjecture

We present three results that supports the validity of Conjecture 3:

1. As a stepping stone towards resolving Conjecture 3, we prove a bound on the first level of the Fourier tail of low degree polynomials over $\mathbb{F}_2$.

   ▶ **Theorem 5.** $\mathcal{L}_{1,1}(\mathrm{Poly}_{n,d}) \le 4d$.

2. We note that as a special case of the main result in [3], any read-once $\mathbb{F}_2$-polynomial f (i.e., a sum of monomials on disjoint sets of variables) has

$$\mathcal{L}_{1,2}(f) \leq O(\log n)^8.$$

Getting a similar $\operatorname{poly}\log(m)$ bound for **general** polynomials with m monomials, would be sufficient for the application in Claim 4.
3. In [2] the following exponential bound was proved on the $\mathcal{L}_{1,2}(\mathrm{Poly}_{n,d})$:

$$\mathcal{L}_{1,2}(\mathrm{Poly}_{n,d}) \leq 4 \cdot 2^{6d} .$$

(as a special case of a bound on $\mathcal{L}_{1,k}$ more generally.) We remark that this bound depends only on d and not on n. Thus, if our conjecture is false, any counter-example to it must have degree $d = \omega(1)$.

Organization

We prove Theorem 2 in Section 2. We prove Theorem 5 in Section 3. We discuss further research in Section 4.

2 PRG from level two Fourier bounds

The main result of this section is an explicit pseudorandom generator for Boolean functions that have bounded Fourier tails on the second level.

▶ **Theorem 6.** *Let $\mathcal{F}$ be a family of n-variate Boolean functions closed under restrictions. Assume that for some $t \geq 1$ it holds that*

$$\mathcal{L}_{1,2}(\mathcal{F}) \leq t.$$

Then for any $\varepsilon > 0$, there exists an explicit PRG for $\mathcal{F}$ with error ε and seed length $O\left((t/\varepsilon)^{2+o(1)} \cdot \operatorname{polylog}(n)\right)$.

The framework to construct the PRG is similar to the one used in [2]. The first step is to construct a fractional PRG for $\mathcal{F}$ that is p-noticeable. Now using the *polarizing* random walk technique used in [2], we convert this fractional PRG into the required standard PRG. Our fractional PRG is based on ideas developed in [9]. We first recall the basic framework of [2].

Pseudorandom generators

Let $f : \{-1,1\}^n \to \{-1,1\}$ be a Boolean function. A PRG for f with error ε is a random variable $X \in \{-1,1\}^n$ such that

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

where U_n is the uniform distribution in $\{-1,1\}^n$. It has seed length s if X can be sampled as

$$X = G(U_s)$$

where $G : \{-1,1\}^s \to \{-1,1\}^n$ is an explicit function[4].
X is a PRG for a class of functions $\mathcal{F}$ if it is a PRG for each function $f \in \mathcal{F}$.

[4] There are various notions of explicitness used in the complexity literature. For our purposes any notion would do.

Fractional pseudorandom generators

Let $f : \{-1,1\}^n \to \{-1,1\}$ be a Boolean function. It has a unique multi-linear extension as $f : [-1,1]^n \to [-1,1]$. A fractional PRG for f with error ε is a random variable $X \in [-1,1]^n$ such that

$$|\,\mathbf{E}[f(X)] - f(\vec{0})| \le \varepsilon.$$

Note that $f(\vec{0}) = \mathbf{E}[f(U_n)]$. It has seed length s if X can be sampled as

$$X = G(U_s)$$

where $G : \{-1,1\}^s \to [-1,1]^n$ is an explicit function. The fractional PRG X is p-noticeable if

$$\mathbf{E}[X_i^2] \ge p \qquad \forall i = 1, \ldots, n.$$

From fractional PRGs to PRGs

The following is the main result of [2], which converts a fractional PRG into a standard PRG.

▶ **Theorem 7** ([2]). *Let $\mathcal{F}$ be a family of n-variate Boolean functions that is closed under restrictions. Let X be a p-noticeable fractional pseudorandom generator for $\mathcal{F}$ with seed length s and error ε. Then, there exists a pseudorandom generator for $\mathcal{F}$ with seed length $O(s \cdot \log(n/\varepsilon)/p)$ and error $O(\varepsilon \cdot \log(n/\varepsilon)/p)$.*

Given, the above theorem, the missing piece to get the desired PRG in Theorem 6 is to construct an appropriate fractional PRG. The following lemma achieves exactly this.

▶ **Lemma 8.** *Let $\mathcal{F}$ be a family of n-variate Boolean functions closed under restrictions. Assume that for some $t \ge 1$ it holds that*

$$\mathcal{L}_{1,2}(\mathcal{F}) \le t.$$

Then for any $\varepsilon > 0$, there exists an explicit p-noticeable fractional PRG for $\mathcal{F}$ with error ε and seed length s where:

$$1/p = O(\log(n/\varepsilon))$$
$$s = O((t/\varepsilon)^{2+o(1)} \cdot \log(n) \cdot \log(n/\varepsilon)).$$

It is direct to obtain Theorem 6 from Theorem 7 and Lemma 8. We prove Lemma 8 in the remainder of this section.

As mentioned before, the fractional PRG is constructed based on ideas developed in [9]. In particular, our fractional PRG can be seen as a derandomization of the main distribution used in [9].

We first abstract and restate one of the main arguments in [9]. This abstraction appeared in a blog post of Boaz Barak and Jarosław Błasiok [1]. Below, we abbreviate a Multi-Variate Gaussian as MVG. Given a random variable $Z \in \mathbb{R}^n$, we denote by $\mathrm{trnc}(Z)$ its truncation to $[-1,1]^n$. That is, $\mathrm{trnc}(Z)_i = \min(1, \max(-1, Z_i))$ for $i \in [n]$.

▶ **Theorem 9** ([9], restated). *Let $n, t \ge 1, \delta \in (0,1)$. Let $Z \in \mathbb{R}^n$ be a zero-mean MVG random variable with the following two properties:*

(i) *For $i \in [n]$: $\boldsymbol{Var}[Z_i] \le \frac{1}{8\ln(n/\delta)}$.*
(ii) *For $i, j \in [n], i \ne j$: $|\,\boldsymbol{Cov}[Z_i, Z_j]| \le \delta$.*

Let $\mathcal{F}$ be a class of n-variate Boolean functions which is closed under restrictions. Assume that $\mathcal{L}_{1,2}(\mathcal{F}) \le t$. Then for any $f \in \mathcal{F}$ it holds that $|\,\boldsymbol{E}[f(\mathrm{trnc}(Z))] - f(\vec{0})| \le O(\delta \cdot t)$.

For completeness, we prove Theorem 9 in the appendix – the proof basically repeats the argument in [9] but for a general multivariate Gaussian distribution, instead of the Forrelation distribution considered there. We now show how to use Theorem 9 to construct a p-noticeable fractional PRG for $\mathcal{F}$ with error ε and seed length s, where $1/p = O(\log(n/\varepsilon))$ and $s = \mathrm{poly}(t, \log(n), 1/\varepsilon)$.

I. We show that a MVG distribution with the parameters needed in Theorem 9 can be of rank $\ell = \mathrm{poly}(\log(n), t, 1/\varepsilon)$. That is, we sample ℓ independent $\mathcal{N}(0,1)$ random variables and apply an explicit linear transformation $T : \mathbb{R}^\ell \to \mathbb{R}^n$ to get a random variable in $\mathbb{R}^n$ that satisfies the two conditions of Theorem 9.

II. We discretized the above process.

Step I: Dimension Reduction

Let $\delta = \varepsilon/t$ and let ℓ be a parameter to be determined soon. Let $\mathcal{C}$ be a code on $\{0,1\}^\ell$ with at least n codewords, such that $\mathcal{C}$ is δ-balanced. Namely, every codeword in $\mathcal{C}$ has Hamming weight between $(\frac{1}{2} - \delta)\ell$ and $(\frac{1}{2} + \delta)\ell$. Such a code can be obtained from explicit constructions of small-biased spaces over $\{0,1\}^\ell$. The best known construction is by Ta-Shma [13] which achieves $\ell = (\log n)/\delta^{2+o(1)}$.

Set $p = 1/(8\ln(n/\delta))$. Let $c^1, \ldots, c^n \in \mathcal{C}$ be distinct codewords. Define an $n \times \ell$ matrix A given by

$$A_{i,j} = \sqrt{\frac{p}{\ell}} \cdot (-1)^{c^i_j},$$

where $c^i = (c^i_1, \ldots, c^i_\ell)$. Let Y be a random vector in $\mathbb{R}^\ell$ where each Y_i is an independent $\mathcal{N}(0,1)$ Gaussian. Define

$$Z = AY.$$

It is straightforward to verify from the construction that Z is a multivariate Gaussian distribution over $\mathbb{R}^n$ with mean zero which satisfies that $\mathbf{Var}[Z_i] = p$ for all $i \in [n]$, and $|\mathbf{Cov}[Z_i, Z_j]| \le \delta$ for all distinct $i, j \in [n]$.

Step II: Discretizing the Randomness

We prove the following lemma, which allows to approximately sample a standard MVG $Y \in \mathbb{R}^\ell$ as needed above using a few random bits.

▶ **Lemma 10.** *For any $\ell, \eta > 0$ there exists $s = O(\ell \cdot \log(\ell/\eta))$ and an explicit generator $G : \{0,1\}^s \to \mathbb{R}^\ell$ such that the following holds.*

Let $f : [-1,1]^n \to [-1,1]$ be a multi-linear function, $A \in [-1,1]^{n\times\ell}$ and Y be a random variable over $\mathbb{R}^\ell$ where each Y_i is an independent $\mathcal{N}(0,1)$ Gaussian. Then

$$|\boldsymbol{E}[f(\mathrm{trnc}(AY))] - \boldsymbol{E}[f(\mathrm{trnc}(AG(U_s)))]| \le \eta(n+2).$$

We say that a random variable $W \in \mathbb{R}$ is a λ-approximate Gaussian if there is a correlated standard Gaussian $W' \sim \mathcal{N}(0,1)$ such that $\mathbf{Pr}[|W - W'| > \lambda] < \lambda$. We will use the following lemma of Kane [6] which shows how to approximate a Gaussian in a randomness efficient way.

▶ **Lemma 11** ([6]). *There is an explicit construction of a λ-approximate Gaussian random variable using $O(\log(1/\lambda))$ bits of randomness.*

The generator G would simply be ℓ independent copies of a λ-approximate Gaussian given by the above lemma, with $\lambda = \frac{\eta}{\ell}$. We denote by $Y' = G(U_s)$ and by Y the coupled standard MVG in $\mathbb{R}^\ell$.

Let $\mathcal{E}$ denote the event that $\|Y - Y'\|_\infty \leq \lambda$. By a union bound, $\mathbf{Pr}(\mathcal{E}) \geq 1 - \eta$. Conditioned on $\mathcal{E}$ it is easy to check that $\|\mathrm{trnc}(AY) - \mathrm{trnc}(AY')\|_\infty \leq \eta$. Finally, we use the multi-linearity and boundedness of f in the following lemma to finish the proof.

▶ **Lemma 12.** *Let $f : [-1,1]^n \to [-1,1]$ be a multi-linear function. Then for every $x, y \in [-1,1]^n$ we have $|f(x) - f(y)| \leq n \cdot \|x - y\|_\infty$.*

Proof. For every $i \in \{0, 1, \ldots, n\}$ define $z^{(i)} := (x_1, ..., x_i, y_{i+1}, ..., y_n)$, note that $z^{(n)} = x$ and $z^{(0)} = y$. We have

$$f(x) - f(y) = \sum_{i=1}^{n} f(z^{(i)}) - f(z^{(i-1)}).$$

Now note that since f is a multilinear function, for every i,

$$|f(z^{(i)}) - f(z^{(i-1)})| = |h_i(x_i) - h_i(y_i)| \leq |x_i - y_i|,$$

where $h_i(z) = f(x_1, \ldots, x_{i-1}, z, y_{i+1}, \ldots, y_n)$. The above inequality holds as h_i is an affine function mapping $[-1,1]$ to $[-1,1]$. We thus obtain that

$$|f(x) - f(y)| \leq \sum_{i=1}^{n} |x_i - y_i| \leq n \cdot \|x - y\|_\infty. \qquad ◀$$

Using Lemma 12 and condition on the event $\mathcal{E}$ we have

$$|f(\mathrm{trnc}(AY)) - f(\mathrm{trnc}(AY'))| \leq \eta n.$$

As f is bounded in $[-1,1]$ we obtain the bound

$$|\,\mathbf{E}[f(\mathrm{trnc}(AY))] - \mathbf{E}[f(\mathrm{trnc}(AY'))]| \leq \eta n + 2\,\mathbf{Pr}[\neg\mathcal{E}] \leq \eta(n+2).$$

Completing the proof

We put things together to finish the proof of Lemma 8. Recall that $\delta = \varepsilon/t$ and $\ell = (\log n)/\delta^{2+o(1)} = (t/\varepsilon)^{2+o(1)} \cdot \log(n)$ from Step I. Set $p = 1/(8\ln(n/\delta)) = 1/(8\ln(nt/\varepsilon))$. Let $A \in [-1,1]^{n\times\ell}$ be the matrix constructed in step I. Set $\eta = \varepsilon/(n+2)$ and let $G : \{0,1\}^s \to \mathbb{R}^\ell$ be the generator constructed in step II. We take

$$X = AG(U_s).$$

The arguments above show that X is a fractional PRG for $\mathcal{F}$ with error $O(\varepsilon)$. In addition, X is p-noticeable. To conclude we compute the seed length s:

$$s = O(\ell \cdot \log(\ell/\eta)) = O\left((t/\varepsilon)^{2+o(1)} \cdot \log n \cdot \log(n/\varepsilon)\right).$$

3 Level one Fourier bounds for polynomials

In this section we bound the level one Fourier tail of low degree polynomials over $\mathbb{F}_2$.

▶ **Theorem 13.** *Let $p : \mathbb{F}_2^n \to \mathbb{F}_2$ be a polynomial of degree d, and let $f(x) = (-1)^{p(x)}$. Then*

$$\mathcal{L}_{1,1}(f) = \sum_{i=1}^{n} |\widehat{f}(i)| \leq 4d.$$

Proof. We assume for simplicity that n is even; the proof is analogous for odd n. We have

$$\sum_{i=1}^{n} |\widehat{f}(i)| = \sum_{i=1}^{n} s_i \cdot \mathop{\mathbf{E}}_{x} \left[f(x)(-1)^{x_i} \right],$$

where $s_i = \mathrm{sign}(\widehat{f}(i))$. We may assume without loss of generality that $s_i = 1$ for all i, by replacing x_i with $1 - x_i$ whenever $s_i = -1$. Thus, it suffices to upper bound

$$E := \mathop{\mathbf{E}}_{x} \left[f(x) \sum_{i=1}^{n} (-1)^{x_i} \right].$$

For $t = 1, \ldots, n/2$ define the functions $T_t : \{0,1\}^n \to \{-1, 0, 1\}$ as follows:

$$T_t(x) := \begin{cases} -1 & \text{if } \sum x_i \geq n/2 + t \\ 1 & \text{if } \sum x_i \leq n/2 - t \\ 0 & \text{otherwise} \end{cases}.$$

Then

$$E = 2 \sum_{t=1}^{n/2} \mathop{\mathbf{E}}_{x} \left[f(x) T_t(x) \right].$$

We need a few more definitions. Let $U_t := \{x \in \{0,1\}^n : |\sum x_i - n/2| \geq t\}$. Define $M_t : U_t \to \mathbb{F}_2$ as $M_t(x) = 0$ if $\sum x_i \geq n/2 + t$, and $M_t(x) = 1$ if $\sum x_i \leq n/2 - t$. Note that $T_t(x) = (-1)^{M_t(x)}$ for $x \in U_t$, and $T_t(x) = 0$ for $x \notin U_t$. Let $A_t := \{x \in U_t : p(x) = M_t(x)\}$. Then

$$e_t := \mathop{\mathbf{E}}_{x} \left[f(x) T_t(x) \right] = \frac{2|A_t| - |U_t|}{2^n}.$$

We next apply a dimension argument similar to that used by Razborov [10] and Smolensky [11] (we adopt a Kopparty's presentation of the argument [7, Lemma 6]). Consider the space of functions $g : A_t \to \mathbb{F}_2$. On the one hand, its dimension is $|A_t|$. On the other hand, any function $g : U_t \to \mathbb{F}_2$ can be decomposed as

$$g(x) = g_1(x) M_t(x) + g_2(x),$$

where g_1, g_2 are polynomials over $\mathbb{F}_2$ of degree $\leq n/2 - t$. Thus, any function $g : A_t \to \mathbb{F}_2$ can be expressed as a polynomial $g(x) = g_1(x)p(x) + g_2(x)$ which is of degree $\leq n/2 - t + d$. Thus, we can bound $|A_t|$ by the dimension of this linear space of polynomials,

$$|A_t| \leq \sum_{i=0}^{n/2-t+d} \binom{n}{i}.$$

Using the fact that $|U_t| = 2\sum_{i=0}^{n/2-t}\binom{n}{i}$ we can upper bound e_t by

$$e_t \le \frac{2\sum_{i=1}^{d}\binom{n}{n/2-t+i}}{2^n}.$$

We thus can bound E by

$$E = 2\sum_{t=1}^{n/2} e_t \le 4\sum_{t=1}^{n/2}\sum_{i=1}^{d}\frac{\binom{n}{n/2-t+i}}{2^n} = 4\sum_{i=1}^{d}\sum_{j=0}^{n/2-1}\frac{\binom{n}{j+i}}{2^n} \le 4d. \qquad \blacktriangleleft$$

3.1 Level two Fourier bounds from level one bounds

We present a simple argument showing that for any family $\mathcal{F}$ of n-variate Boolean functions that is closed under restrictions, a bound of $\mathcal{L}_{1,1}(\mathcal{F}) \le t$ implies $\mathcal{L}_{1,2}(\mathcal{F}) \le O(t \cdot \sqrt{n\log n})$. Using this connection, we get that polynomials of degree polylog(n) have $\mathcal{L}_{1,2}(\cdot)$ at most $\sqrt{n}\cdot$ polylog(n). Recall that we conjecture that the right bound should be polylog(n) (i.e., exponentially smaller). Nevertheless, even improving this bound slightly to $n^{1/2-o(1)}$ would imply a non-trivial PRG fooling polylog(n)-degree $\mathbb{F}_2$-polynomials and $\mathsf{AC}^0[\oplus]$ circuits with seed-length $n^{1-o(1)}$. In comparison, the current state of the art PRG for $\mathsf{AC}^0[\oplus]$ circuits has seed-length $n - n/\text{polylog}(n)$ [4].

▶ **Claim 14.** *Let $\mathcal{F}$ be a class of n-variate Boolean functions that is closed under restrictions. Let $t \ge 1$. Assume that $\mathcal{L}_{1,1}(\mathcal{F}) \le t$. Then, $\mathcal{L}_{1,2}(\mathcal{F}) \le t \cdot O(\sqrt{n\log n})$.*

Proof. Let $f : \{-1,1\}^n \to \{-1,1\}$ be some Boolean function in $\mathcal{F}$. We bound $\mathcal{L}_{1,2}(f) = \sum_{i<j}|\hat{f}(i,j)|$. We begin by partitioning the set of coordinates of f into two disjoint parts $[n] = X \cup Y$ and summing only the cross-terms $L_1(X,Y) = \sum_{i\in X}\sum_{j\in Y}|\hat{f}(i,j)|$. We note that there exists a partition $[n] = X \cup Y$ such that $L_1(X,Y) \ge \mathcal{L}_{1,2}(f)/2$. This holds since a random partition has on expectation

$$\mathop{\mathbf{E}}_{X,Y}[L_1(X,Y)] = \sum_{i<j}|\hat{f}(i,j)| \cdot (\mathbf{Pr}[i \in X, j \in Y] + \mathbf{Pr}[i \in Y, j \in X]) = \mathcal{L}_{1,2}(f) \cdot \frac{1}{2}.$$

Fix a partition (X,Y) for which $L_1(X,Y) \ge \mathcal{L}_{1,2}(f) \cdot \frac{1}{2}$. In the remainder, we bound

$$L_1(X,Y) = \sum_{i\in X, j\in Y}|\hat{f}(\{i,j\})| = \sum_{i\in X, j\in Y} s_{i,j} \cdot \hat{f}(\{i,j\})$$

for some sign matrix $s \in \{-1,1\}^{X\times Y}$. For any fixed $x \in \{-1,1\}^X$ we denote by $f_x : \{-1,1\}^Y \to \{-1,1\}$ the function defined by $f_x(y) = f(x,y)$. Note that f_x is a restriction of f thus by our assumption, its $\mathcal{L}_{1,1}$ is at most t. We get

$$\begin{aligned}
L_1(X,Y) &= \mathop{\mathbf{E}}_{\substack{x\in\{-1,1\}^X,\\ y\in\{-1,1\}^Y}}\left[\sum_{i\in X, j\in Y} s_{i,j} \cdot f(x,y) \cdot x_i \cdot y_j\right] \\
&= \mathop{\mathbf{E}}_{x\in\{-1,1\}^X}\left[\sum_{i\in X, j\in Y} s_{i,j} \cdot x_i \cdot \mathop{\mathbf{E}}_{y\in\{-1,1\}^Y}[f(x,y)\cdot y_j]\right] \\
&= \mathop{\mathbf{E}}_{x\in\{-1,1\}^X}\left[\sum_{i\in X, j\in Y} s_{i,j} \cdot x_i \cdot \hat{f}_x(\{j\})\right] = \mathop{\mathbf{E}}_{x\in\{-1,1\}^X}\left[\sum_{j\in Y}\left(\hat{f}_x(\{j\}) \cdot \sum_{i\in X} s_{i,j}\cdot x_i\right)\right] \\
&\le \mathop{\mathbf{E}}_{x\in\{-1,1\}^X}\left[\sum_{j\in Y}|\hat{f}_x(\{j\})| \cdot \left|\sum_{i\in X} s_{i,j}\cdot x_i\right|\right] \le \mathop{\mathbf{E}}_{x\in\{-1,1\}^X}\left[t \cdot \max_{j\in Y}\left|\sum_{i\in X} s_{i,j}\cdot x_i\right|\right]
\end{aligned}$$

By Chernoff's bounds, the expectation of $\max_{j \in Y} |\sum_{i \in X} s_{i,j} \cdot x_i|$ is at most $O(\sqrt{n \log n})$. Thus overall $\mathcal{L}_{1,2}(f) \leq 2 \cdot L_1(X, Y) \leq 2t \cdot O(\sqrt{n \cdot \log n})$. ◀

4 Further research

A clear advantage of Theorem 2 over Theorem 1 is that only bounds on the second level of the Fourier tails are needed, instead of bounds for all levels. However, we pay a price, as the dependence on the error parameter ε is polynomial instead of poly-logarithmic. This raises a natural problem: can a better dependency on ε be obtained if the Fourier tails are assumed to be bounded for several levels k? In particular, information on how many levels is needed in order to obtain poly-logarithmic dependency on the error ε? We leave these questions to future work.

References

1. Boaz Barack and Jarosław Błasiok. On the Raz-Tal oracle separation of BQP and PH, 2018. `https://windowsontheory.org/2018/06/17/on-the-raz-tal-oracle-separation-of-bqp-and-ph/`.
2. Eshan Chattopadhyay, Pooya Hatami, Kaave Hosseini, and Shachar Lovett. Pseudorandom Generators from Polarizing Random Walks. In *33rd Computational Complexity Conference, CCC 2018*, pages 1:1–1:21, 2018.
3. Eshan Chattopadhyay, Pooya Hatami, Omer Reingold, and Avishay Tal. Improved pseudorandomness for unordered branching programs through local monotonicity. In *STOC*, pages 363–375. ACM, 2018.
4. Bill Fefferman, Ronen Shaltiel, Christopher Umans, and Emanuele Viola. On Beating the Hybrid Argument. *Theory of Computing*, 9:809–843, 2013.
5. Leon Isserlis. On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika*, 12(1/2):134–139, 1918.
6. Daniel M. Kane. A Polylogarithmic PRG for Degree 2 Threshold Functions in the Gaussian Setting. In *Conference on Computational Complexity*, volume 33 of *LIPIcs*, pages 567–581, 2015.
7. Swastik Kopparty. On the complexity of powering in finite fields. In *STOC*, pages 489–498. ACM, 2011.
8. Peter Mörters and Yuval Peres. *Brownian motion*, volume 30. Cambridge University Press, 2010.
9. Ran Raz and Avishay Tal. Oracle Separation of BQP and PH. *Electronic Colloquium on Computational Complexity (ECCC)*, 25:107, 2018.
10. 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, 1987.
11. R. Smolensky. On representations by low-degree polynomials. In *FOCS*, pages 130–138, 1993.
12. Roman Smolensky. Algebraic Methods in the Theory of Lower Bounds for Boolean Circuit Complexity. In *STOC*, pages 77–82. ACM, 1987.
13. Amnon Ta-Shma. Explicit, almost optimal, epsilon-balanced codes. In *STOC*, pages 238–251, 2017.
14. Salil P. Vadhan. Pseudorandomness. *Foundations and Trends in Theoretical Computer Science*, 7(1-3):1–336, 2012.
15. Emanuele Viola. Guest Column: correlation bounds for polynomials over {0 1}. *ACM SIGACT News*, 40(1):27–44, 2009.

A Proof of Theorem 9

Throughout this section we take G to be a multivariate Gaussian distribution with zero mean, covariances at most δ and variances at most 1. That is, if $G = (G_1, \ldots, G_n)$ then $\mathbf{E}[G_i] = 0, \mathbf{E}[G_i^2] \leq 1$ and $|\mathbf{E}[G_i G_j]| \leq \delta$ for $i \neq j$.

A.1 Preliminaries

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a multi-linear function, defined by

$$f(z) = \sum_{S \subseteq [n]} \widehat{f}(S) \cdot \prod_{i \in S} z_i, \tag{1}$$

where $\widehat{f}(S) \in \mathbb{R}$. We bound the difference between $\mathbf{E}_{z \sim G}[f(\mathrm{trnc}(pz))]$ and $\mathbf{E}_{z \sim G}[f(pz)]$ for a small $p \in (0,1)$. Note that whenever $z' \in [-1,1]^n$, there is no difference between $f(z')$ and $f(\mathrm{trnc}(z'))$, and we only need to bound the difference when z' is outside $[-1,1]^n$. The next claim bounds the value of $|f(z')|$ when z' is outside $[-1,1]^n$.

▶ **Claim 15** ([9, Claim 5.1]). *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a multi-linear function that maps $\{-1,1\}^n$ to $[-1,1]$. Let $z' \in \mathbb{R}^n$. Then, $|f(z')| \leq \prod_{i=1}^n \max(1, |z'_i|)$.*

For $\alpha \in (0,1), z \in \mathbb{R}^n$, we get that the value of $|f(\alpha z)|$ is bounded by $\prod_i \max(1, |\alpha z_i|)$. The following claim bounds the latter times the indicator that $\alpha z \neq \mathrm{trnc}(\alpha z)$.

▶ **Claim 16.** *Let $\alpha \in (0, 1/\sqrt{4n}]$. Then,*

$$\mathop{\boldsymbol{E}}_{z \sim G}\Big[\prod_{i=1}^n \max(1, |\alpha z_i|) \cdot \mathbb{1}_{\alpha z \neq \mathrm{trnc}(\alpha z)}\Big] \leq \sum_{k=1}^{\infty} e^{-k/(4\alpha^2 n)} \cdot n^k.$$

Proof. For $i \in [n]$ and $a_i \in \mathbb{N}$, we consider the event

$$a_i \leq |\alpha \cdot z_i| < a_i + 1,$$

denoted by $\mathcal{E}_{i,a_i}$. Since each z_i is a Gaussian with mean 0 and variance at most 1, we have $\mathbf{Pr}[\mathcal{E}_{i,a_i}] \leq e^{-a_i^2/(2\alpha^2)}$. Using Claim 15 we have

$$\begin{aligned}
(*) &= \mathop{\mathbf{E}}_{z \sim G}\Big[\prod_{i=1}^n \max(1, |\alpha z_i|) \cdot \mathbb{1}_{\alpha z \neq \mathrm{trnc}(\alpha z)}\Big] \\
&\leq \sum_{\vec{a} \in \mathbb{N}^n, \vec{a} \neq 0^n} \mathbf{Pr}[\wedge_{i=1}^n \mathcal{E}_{i,a_i}] \cdot \prod_{i=1}^n (1 + a_i) \\
&\leq \sum_{\vec{a} \in \mathbb{N}^n, \vec{a} \neq 0^n} \min_{i \in [n]}\{\mathbf{Pr}[\mathcal{E}_{i,a_i}]\} \cdot \prod_{i=1}^n (1 + a_i) \\
&\leq \sum_{\vec{a} \in \mathbb{N}^n, \vec{a} \neq 0^n} \prod_{i=1}^n \mathbf{Pr}[\mathcal{E}_{i,a_i}]^{1/n} \cdot \prod_{i=1}^n (1 + a_i) \qquad (2)
\end{aligned}$$

We bound

$$\mathbf{Pr}[\mathcal{E}_{i,a_i}]^{1/n} \cdot (1 + a_i) \leq e^{-a_i^2/(2\alpha^2 n)} \cdot (1 + a_i) \leq e^{-a_i^2/(4\alpha^2 n)}$$

since $1 + a_i \le e^{a_i} \le e^{a_i^2/(4\alpha^2 n)}$ for $\alpha^2 \le 1/4n$. We plug this estimate in Equation (2):

$$\begin{aligned}
(*) &\le \sum_{\vec{a}\in\mathbb{N}^N, \vec{a}\neq 0^n} e^{-\sum_i a_i^2/(4\alpha^2 n)} \\
&\le \sum_{k=1}^{\infty} e^{-k/(4\alpha^2 n)} \cdot \left|\left\{\vec{a}\in\mathbb{N}^n : \sum_i a_i = k\right\}\right| \\
&= \sum_{k=1}^{\infty} e^{-k/(4\alpha^2 n)} \cdot \binom{n+k-1}{k} \le \sum_{k=1}^{\infty} e^{-k/(4\alpha^2 n)} \cdot n^k.
\end{aligned}$$

◀

▶ **Claim 17.** *Let $p \le 1/4\sqrt{n}$. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a multi-linear function that maps $\{-1,1\}^n$ to $[-1,1]$. Let $v \in [-1/2, 1/2]^n$. Then,*

$$\mathop{\boldsymbol{E}}_{z\sim G}[|f(\mathrm{trnc}(v + p\cdot z)) - f(v + p\cdot z)|] \le 2 \cdot \sum_{k=1}^{\infty} e^{-k/(16p^2 n)} \cdot n^k.$$

Proof. Let $\mathcal{E}$ be the event that $(\mathrm{trnc}(v + p\cdot z) \neq v + p\cdot z)$. Note that $\mathcal{E}$ implies the event $2pz \neq \mathrm{trnc}(2pz)$ since $v \in [-1/2, 1/2]^n$. Using Claim 15, we get

$$\begin{aligned}
\mathop{\mathbf{E}}_{z\sim G}[|f(\mathrm{trnc}(v + p\cdot z)) - f(v + p\cdot z)|] &\le \mathop{\mathbf{E}}_{z\sim G}[(1 + |f(v + p\cdot z)|)\cdot \mathbb{1}_{\mathcal{E}}] \\
&\le \mathop{\mathbf{E}}_{z\sim G}[(1 + |f(v + p\cdot z)|)\cdot \mathbb{1}_{2pz\neq \mathrm{trnc}(2pz)}] \\
&\le \mathop{\mathbf{E}}_{z\sim G}\left[\left(1 + \prod_{i=1}^{n} \max(1, |v_i + p\cdot z_i|)\right)\cdot \mathbb{1}_{2pz\neq \mathrm{trnc}(2pz)}\right] \\
&\le \mathop{\mathbf{E}}_{z\sim G}\left[2 \cdot \prod_{i=1}^{n} \max(1, |v_i + p\cdot z_i|)\cdot \mathbb{1}_{2pz\neq \mathrm{trnc}(2pz)}\right].
\end{aligned}$$

However, $\prod_{i=1}^n \max(1, |v_i + p\cdot z_i|) \le \prod_{i=1}^n \max(1, 1/2 + p|z_i|) \le \prod_{i=1}^n \max(1, 2p|z_i|)$. Using Claim 16 with $\alpha = 2p$, we get

$$\begin{aligned}
\mathop{\mathbf{E}}_{z\sim G}[|f(\mathrm{trnc}(v + p\cdot z)) - f(v + p\cdot z)|] &\le \mathop{\mathbf{E}}_{z\sim G}\left[2 \cdot \prod_{i=1}^{n} \max(1, 2p|z_i|)\cdot \mathbb{1}_{2pz\neq \mathrm{trnc}(2pz)}\right] \\
&\le 2 \cdot \mathop{\mathbf{E}}_{z\sim G}\left[\prod_{i=1}^{n} \max(1, \alpha|z_i|)\cdot \mathbb{1}_{\alpha z\neq \mathrm{trnc}(\alpha z)}\right] \\
&\le 2 \cdot \sum_{k=1}^{\infty} e^{-k/(4\alpha^2 n)} \cdot n^k.
\end{aligned}$$

◀

▶ **Claim 18** (Application of Isserlis' Theorem)**.** *Let G be a MVG distribution over $\mathbb{R}^n$ with zero-mean and covariances at most δ. For $S \subseteq [n]$, let $\widehat{G}(S) = \boldsymbol{E}_{Z\sim G}[\prod_{i\in S} Z_i]$. Then,*

1. *$\widehat{G}(S) = 0$ if $|S|$ is odd.*
2. *$|\widehat{G}(S)| \le (k-1)!! \cdot \delta^{k/2}$ if $|S| = k$ is even.*

Proof. Both items rely on Isserlis' Theorem [5] (See also `http://en.wikipedia.org/wiki/Isserlis'_theorem`) that gives a formula for the moments of any zero-mean multi Gaussian distribution. Isserlis' Theorem [5] states that in a zero-mean multivariate Gaussian distribution $Z_1, \ldots, Z_n$, for a sequence of indices $(i_1, \ldots, i_k) \in [n]$, we have $\mathbf{E}[Z_{i_1} \cdots Z_{i_k}] = 0$ if k is odd and $\mathbf{E}[Z_{i_1} \cdots Z_{i_k}] = \sum\prod \mathbf{E}[Z_{i_r} Z_{i_\ell}]$, where the notation $\sum\prod$ means summing over all distinct ways of partitioning $Z_{i_1}, \ldots, Z_{i_k}$ into pairs and each summand is the product of the $k/2$ pairs. If $|S| = k$ is even, since the covariance of each pair in G is at most δ in absolute value and there are at most $(k-1)!!$ partitions to pairs, we get $|\widehat{G}(S)| \le (k-1)!! \cdot \delta^{k/2}$. ◀

The next claim expresses the difference of a multi-linear function f on two vectors, v and $v+z$, as the expected difference of random restrictions of f on 0 and $2z$, provided that $v \in [-1/2, 1/2]^n$. Applying this lemma when the entries of z are infinitesimally small means that bounded variation of random restrictions of f around 0 implies bounded variation of f around any $v \in [-1/2, 1/2]^n$.

▶ **Claim 19** ([2, Claim 3.3], restated in [1]). *Let f be a multi-linear function on $\mathbb{R}^n$ and $v \in [-1/2, 1/2]^n$. There exists a distribution over random restrictions ρ such that for any $z \in \mathbb{R}^n$,*

$$f(v+z) - f(v) = \mathop{\boldsymbol{E}}_{\rho}[f_\rho(2 \cdot z) - f_\rho(\vec{0})].$$

Proof. Given $v \in [-1/2, 1/2]^n$, we define a distribution $\mathcal{R}_v$ over restrictions $\rho \in \{-1, 1, *\}^n$, as follows. For each entry $i \in [n]$ independently, we set $\rho_i = 1$ with probability $1/4 + v_i/2$, $\rho_i = -1$ with probability $1/4 - v_i/2$, and $\rho_i = *$ with probability $1/2$. Note that since $v \in [-1/2, 1/2]^n$ all these probabilities are indeed non-negative.

Let $\rho \sim \mathcal{R}_v$. For any vector $z \in \mathbb{R}^n$, we define a vector $\tilde{z} = \tilde{z}(z, \rho) \in \mathbb{R}^n$, as follows:

$$\tilde{z}_i = \begin{cases} \rho_i & \text{if } \rho_i \in \{-1, 1\} \\ 2 \cdot z_i & \text{otherwise} \end{cases}$$

Thus, for a fixed $z \in \mathbb{R}^n$, the vector $\tilde{z}$ is a random variable that depends on ρ. We show that for any fixed $z \in \mathbb{R}^n$, the distribution of the random variable $\tilde{z}$ is a product distribution (over inputs in $\mathbb{R}^n$), and the expectation of $\tilde{z}$ is the vector $v + z$. Indeed, each coordinate $\tilde{z}_i$ is independent of the other coordinates, and its expected value is

$$\mathop{\mathbf{E}}_{\rho \sim \mathcal{R}_v}[\tilde{z}_i] = v_i + z_i.$$

Hence, since f is multi-linear and $\tilde{z}$ has a product distribution, by Equation (1), $\mathbf{E}_{\rho \sim \mathcal{R}_v}[f(\tilde{z})] = f(v+z)$. We get

$$f(v+z) - f(v) = \mathop{\mathbf{E}}_{\rho \sim \mathcal{R}_v}[f(\tilde{z}(z, \rho))] - \mathop{\mathbf{E}}_{\rho \sim \mathcal{R}_v}[f(\tilde{z}(\vec{0}, \rho))] = \mathop{\mathbf{E}}_{\rho \sim \mathcal{R}_v}[f(\tilde{z}(z, \rho)) - f(\tilde{z}(\vec{0}, \rho))]$$

However, for any fixed ρ, we have $f(\tilde{z}(z, \rho)) = f_\rho(2z)$, where f_ρ is attained from f by fixing the coordinates that were fixed in ρ, according to ρ. Thus,

$$f(v+z) - f(v) = \mathop{\mathbf{E}}_{\rho \sim \mathcal{R}_v}[f_\rho(2 \cdot z) - f_\rho(\vec{0})].$$ ◀

A.2 The Proof

▶ **Claim 20.** *Let $f : \{-1, 1\}^n \to \{-1, 1\}$ be a Boolean function with $\mathcal{L}_{1,2}(f) \le t$. Let $p \le 1/2n$. Then,*

$$\left| \mathop{\boldsymbol{E}}_{Z \sim G}[f(pZ)] - f(\vec{0}) \right| \le p^2 \cdot t \cdot \delta + O(p^4 \cdot n^4 \cdot \delta^2) .$$

Proof. By Equation (1) and since $f(\vec{0}) = \widehat{f}(\emptyset)$,

$$\begin{aligned} \left| \mathop{\mathbf{E}}_{Z \sim G}[f(pZ)] - f(\vec{0}) \right| &= \left| \mathop{\mathbf{E}}_{Z \sim G}\left[\sum_{\emptyset \neq S \subseteq [n]} \widehat{f}(S) \cdot \prod_{i \in S}(p \cdot Z_i) \right] \right| \\ &= \left| \sum_{\emptyset \neq S \subseteq [n]} \widehat{f}(S) \cdot p^{|S|} \cdot \mathop{\mathbf{E}}_{Z \sim G}\left[\prod_{i \in S} Z_i \right] \right| \\ &\le \sum_{k=1}^{n} p^k \cdot \left(\max_{S : |S| = k} |\widehat{G}(S)| \right) \cdot \sum_{S \subseteq [n], |S| = k} |\widehat{f}(S)| \end{aligned}$$

For odd k, Claim 18 gives $\max_{S:|S|=k} |\widehat{G}(S)| = 0$. For even k, we have $\max_{S:|S|=k} |\widehat{G}(S)| \leq (k-1)!! \cdot \delta^{k/2}$ by Claim 18. Plugging these bounds in the above expression gives

$$\begin{aligned}
\Big| \mathop{\mathbf{E}}_{Z\sim G}[f(pZ)] - f(\vec{0})\Big| &\leq p^2 \cdot \delta \cdot \sum_{S:|S|=2} |\widehat{f}(S)| + \sum_{k\geq 4, k \text{ even}} p^k \cdot \delta^{k/2} \cdot (k-1)!! \cdot \sum_{S:|S|=k} |\widehat{f}(S)| \\
&\leq p^2 \cdot \delta \cdot t + \sum_{k\geq 4, k \text{ even}} p^k \cdot \delta^{k/2} \cdot (k-1)!! \cdot \binom{n}{k} \\
&\qquad\qquad (\mathcal{L}_{1,2}(f) \leq t \text{ and } \forall S : |\widehat{f}(S)| \leq 1) \\
&\leq p^2 \cdot \delta \cdot t + \sum_{k\geq 4, k \text{ even}} p^k \cdot \delta^{k/2} \cdot n^k \\
&\leq p^2 \cdot \delta \cdot t + O(p^4 \cdot n^4 \cdot \delta^2) \qquad (p \leq 1/2n)
\end{aligned}$$

◀

▶ **Theorem 21** (Theorem 9, restated). *Let $n \in \mathbb{N}$, $\delta, \sigma \in (0,1)$. Let G be a zero-mean multivariate Gaussian distribution over $\mathbb{R}^n$ where $Z \sim G$ has the following two properties:*

1. *For $i \in [n]$: $\boldsymbol{Var}[Z_i] \leq \sigma^2$*
2. *For $i, j \in [n], i \neq j$: $|\boldsymbol{Cov}[Z_i, Z_j]| \leq \delta$.*

Let $\mathcal{F}$ be a class of n-variate Boolean functions which is closed under restrictions. Assume that $\mathcal{L}_{1,2}(\mathcal{F}) \leq t$. Then for any $f \in \mathcal{F}$ it holds that $|\boldsymbol{E}[f(\mathrm{trnc}(Z))] - f(\vec{0})| \leq 4\delta \cdot t + 4n \cdot e^{-1/8\sigma^2}$.

Proof. Let $m \in \mathbb{N}$ be sufficiently large (in particular $m \geq (4n)^4$) and $p = 1/\sqrt{m}$. Let $Z^{(1)}, \ldots, Z^{(m)} \sim G$. We define $m+1$ hybrids $H_0, \ldots, H_m$. Let $H_0 = \vec{0}$. For $i = 1, \ldots, m$, let $H_i = p \cdot (Z^{(1)} + \ldots + Z^{(i)})$. We observe that $H_m \sim G$. This is true since H_m is a multivariate Gaussian with the same expectation and the same covariance matrix as $Z \sim G$. We can think of $H_0, H_1, \ldots, H_m$ as a n-dimensional random walk. We bound

$$\big|\mathbf{E}[f(\mathrm{trnc}(H_m))] - f(\vec{0})]\big|$$

by considering two cases depending on whether or not at some point in the random walk we stepped outside of $[-1/2, 1/2]^n$.

For $i \in \{0, \ldots, m\}$, let $\mathcal{E}_i$ be the event that $H_i \in [-1/2, 1/2]^n$. We show that $\mathcal{E}_i$ happens with high probability. In fact, we show that $\mathcal{E} = \mathcal{E}_1 \wedge \mathcal{E}_2 \wedge \ldots \wedge \mathcal{E}_m$ happens with high probability, with no dependency on the number of steps m. The claim follows from known properties of Brownian motions. For $j \in [n]$, let $\mathcal{D}^{(j)}$ be the event that there exists an $i \in [m]$ with $|(H_i)_j| > 1/2$. Clearly $\neg\mathcal{E} \equiv \mathcal{D}^{(1)} \vee \mathcal{D}^{(2)} \vee \ldots \vee \mathcal{D}^{(n)}$.

We show that for each $j \in [n]$, $\mathbf{Pr}[\mathcal{D}^{(j)}] \leq 4 \cdot e^{-8/\sigma^2}$ and then apply a union bound. Each $\{(H_i)_j\}_{i=0}^m$ is a random walk with m steps, which can be viewed as a discretization of a one-dimensional Brownian motion. A standard one-dimensional Brownian motion (or Wiener process) is a random process $\{B(t)\}_{t\geq 0}$ with the properties: (1) $B(0) = 0$, (2) for all $t, s \geq 0$, $B(t+s) - B(t)$ is independent of the past $\{B(t')\}_{t'\leq t}$ (3) for all $t, s \geq 0$, $B(t+s) - B(t) \sim \mathcal{N}(0, s)$. Let $\sigma_j^2 := \mathbf{Var}[z_j]$. We observe that if B is a standard one-dimensional Brownian motion, then $\{B(\sigma_j^2 \cdot i/m)\}_{i=0}^m$ is distributed exactly as $\{(H_i)_j\}_{i=0}^m$. Let $M(t) = \sup_{0\leq s\leq t} B(s)$ and $M'(t) = \inf_{0\leq s\leq t} B(s)$. It suffices to show that $M(\sigma_j^2) \leq 1/2$ and $M'(\sigma_j^2) \geq -1/2$ with high probability. Known results on Brownian motions state that $\mathbf{Pr}[M'(t) < -1/2] = \mathbf{Pr}[M(t) > 1/2] = \mathbf{Pr}[|B(t)| > 1/2]$ (cf. [8, Theorem 2.21]). The latter is at most $e^{-1/8t}$ since $B(t) \sim \mathcal{N}(0, t)$. Overall, we get

$$\mathbf{Pr}[\neg\mathcal{E}] \leq \sum_{j=1}^n \mathbf{Pr}[\mathcal{D}^{(j)}] \leq \sum_{j=1}^n \Big(\mathbf{Pr}[M'(\sigma_j^2) < -1/2] + \mathbf{Pr}[M(\sigma_j^2) > 1/2]\Big) \leq 2n \cdot e^{-1/8\sigma^2}$$

Next, we bound $|\mathbf{E}_{Z^{(i+1)}}[f(\mathrm{trnc}(H_i + p\cdot Z^{(i+1)})) - f(\mathrm{trnc}(H_i))]|$ conditioned on the event $\mathcal{E}_i$, for $i = 0, 1, \ldots, m-1$. Let $v = H_i$. Condition on the event $\mathcal{E}_i$, and in fact condition on the entire history in the first i steps, which in particular fixes v. By Claim 19, we have

$$\begin{aligned}\left|\mathop{\mathbf{E}}_{Z^{(i+1)}}[f(v + p\cdot Z^{(i+1)}) - f(v)]\right| &= \left|\mathop{\mathbf{E}}_{Z^{(i+1)}}\mathop{\mathbf{E}}_{\rho}[f_\rho(2p\cdot Z^{(i+1)}) - f_\rho(\vec{0})]\right| \\ &\le \mathop{\mathbf{E}}_{\rho}\left|\mathop{\mathbf{E}}_{Z^{(i+1)}}[f_\rho(2p\cdot Z^{(i+1)}) - f_\rho(\vec{0})]\right|.\end{aligned}$$

By Claim 20 we have that the latter is at most $(2p)^2 t\delta + O(p^4n^4\delta^2)$ as long as $2p \le 1/2n$. We wish to show a similar bound on the truncated version of $H_i + p\cdot Z^{(i+1)}$. Note that conditioned on $\mathcal{E}_i$, we have $H_i = \mathrm{trnc}(H_i)$, but this is not necessarily the case for $H_i + p\cdot Z^{(i+1)}$. Using Claim 17 we get $\left|\mathbf{E}_{Z^{(i+1)}}[f(p\cdot Z^{(i+1)} + v) - f(\mathrm{trnc}(p\cdot Z^{(i+1)} + v))]\right| \le 2\cdot\sum_{k=1}^{\infty} e^{-k/(16p^2n)}\cdot n^k$. By the triangle inequality we get

$$\begin{aligned}&\left|\mathop{\mathbf{E}}_{Z^{(i+1)}}[f(\mathrm{trnc}(v + p\cdot Z^{(i+1)})) - f(\mathrm{trnc}(v))]\right| \\ &\le \left|\mathop{\mathbf{E}}_{Z^{(i+1)}}[f(\mathrm{trnc}(v + p\cdot Z^{(i+1)})) - f(v + p\cdot Z^{(i+1)})]\right| \\ &\qquad + \left|\mathop{\mathbf{E}}_{Z^{(i+1)}}[f(v + p\cdot Z^{(i+1)}) - f(\mathrm{trnc}(v))]\right| \\ &\le \left(2\cdot\sum_{k=1}^{\infty} e^{-k/(16p^2n)}\cdot n^k\right) + \left(4p^2\cdot\delta\cdot t + O(p^4n^4\delta^2)\right). \qquad (3)\end{aligned}$$

To finish the proof, using triangle inequality we have

$$\left|\mathbf{E}[f(\mathrm{trnc}(H_m)) - f(\vec{0})]\right| \le \left|\mathbf{E}[f(\mathrm{trnc}(H_m))\cdot \mathbb{1}_{\mathcal{E}} - f(\vec{0})]\right| + |\mathbf{E}[f(\mathrm{trnc}(H_m))\cdot \mathbb{1}_{\neg\mathcal{E}}]|$$

We bound the second summand by $\mathbf{Pr}[\neg\mathcal{E}]$ since f is bounded in $[-1, 1]$ on truncated vectors, whereas the first summand is bounded using a telescopic sum of the $m+1$ hybrids:

$$\begin{aligned}&\left|\mathbf{E}[f(\mathrm{trnc}(H_m))\cdot \mathbb{1}_{\mathcal{E}}) - f(\vec{0})]\right| \\ &\le \sum_{i=0}^{m-1} |\mathbf{E}[f(\mathrm{trnc}(H_{i+1}))\cdot \mathbb{1}_{\mathcal{E}_1\wedge\ldots\wedge\mathcal{E}_{i+1}} - f(\mathrm{trnc}(H_i))\cdot \mathbb{1}_{\mathcal{E}_1\wedge\ldots\wedge\mathcal{E}_i}]| \\ &\le \sum_{i=0}^{m-1} |\mathbf{E}[f(\mathrm{trnc}(H_{i+1}))\cdot \mathbb{1}_{\mathcal{E}_1\wedge\ldots\wedge\mathcal{E}_i} - f(\mathrm{trnc}(H_i))\cdot \mathbb{1}_{\mathcal{E}_1\wedge\ldots\wedge\mathcal{E}_i})]| \\ &\qquad + |\mathbf{E}[f(\mathrm{trnc}(H_{i+1}))\cdot(\mathbb{1}_{\mathcal{E}_1\wedge\ldots\wedge\mathcal{E}_{i+1}} - \mathbb{1}_{\mathcal{E}_1\wedge\ldots\wedge\mathcal{E}_i})]| \\ &\le \sum_{i=0}^{m-1}\left(2\cdot\sum_{k=1}^{\infty} e^{-k/(16p^2n)}\cdot n^k\right) + \left(4p^2\cdot\delta\cdot t + O(p^4n^4\delta^2)\right) + \mathbf{E}[|\mathbb{1}_{\mathcal{E}_1\wedge\ldots\wedge\mathcal{E}_{i+1}} - \mathbb{1}_{\mathcal{E}_1\wedge\ldots\wedge\mathcal{E}_i}|] \\ &\qquad\qquad\qquad