Fourier growth of structured $\mathbb{F}_2$-polynomials and applications

Jarosław Błasiok
Columbia University
jb4451@columbia.edu

Peter Ivanov
Northeastern University
ivanov.p@northeastern.edu

Yaonan Jin
Columbia University
yj2552@columbia.edu

Chin Ho Lee
Columbia University
c.h.lee@columbia.edu

Rocco A. Servedio
Columbia University
rocco@cs.columbia.edu

Emanuele Viola
Northeastern University
viola@ccs.neu.edu

July 22, 2021

Abstract

We analyze the *Fourier growth*, i.e. the L_1 Fourier weight at level k (denoted $L_{1,k}$), of various well-studied classes of "structured" $\mathbb{F}_2$-polynomials. This study is motivated by applications in pseudorandomness, in particular recent results and conjectures due to [CHHL19, CHLT19, CGL+20] which show that upper bounds on Fourier growth (even at level $k = 2$) give unconditional pseudorandom generators.

Our main structural results on Fourier growth are as follows:

- We show that any symmetric degree-d $\mathbb{F}_2$-polynomial p has $L_{1,k}(p) \leq \mathbf{Pr}[p = 1] \cdot O(d)^k$, and this is tight for any constant k. This quadratically strengthens an earlier bound that was implicit in [RSV13].
- We show that any read-Δ degree-d $\mathbb{F}_2$-polynomial p has $L_{1,k}(p) \leq \mathbf{Pr}[p = 1] \cdot (k\Delta d)^{O(k)}$.
- We establish a composition theorem which gives $L_{1,k}$ bounds on disjoint compositions of functions that are closed under restrictions and admit $L_{1,k}$ bounds.

Finally, we apply the above structural results to obtain new unconditional pseudorandom generators and new correlation bounds for various classes of $\mathbb{F}_2$-polynomials.

1 Introduction

1.1 Background: L_1 Fourier norms and Fourier growth

Over the past several decades, Fourier analysis of Boolean functions has emerged as a fundamental tool of great utility across many different areas within theoretical computer science and mathematics. Areas of application include (but are not limited to) combinatorics, the theory of random graphs and statistical physics, social choice theory, Gaussian geometry and the study of metric spaces, cryptography, learning theory, property testing, and many branches of computational complexity such as hardness of approximation, circuit complexity, and pseudorandomness. The excellent book of O'Donnell [O'D14] provides a broad introduction. In this paper we follow the notation of [O'D14], and for a Boolean-valued function f on n Boolean variables and $S \subseteq [n]$, we write $\widehat{f}(S)$ to denote the Fourier coefficient of f on S.

Given the wide range of different contexts within which the Fourier analysis of Boolean functions has been pursued, it is not surprising that many different quantitative parameters of Boolean functions have been analyzed in the literature. In this work we are chiefly interested in the L_1 *Fourier norm at level* k:

Definition 1 (L_1 Fourier norm at level k)**.** The L_1 *Fourier norm of a function* $f\colon \{-1,1\}^n \to \{0,1\}$ *at level* k is the quantity

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

For a function class $\mathcal{F}$, we write $L_{1,k}(\mathcal{F})$ to denote $\max_{f\in\mathcal{F}} L_{1,k}(f)$.

As we explain below, strong motivation for studying the L_1 Fourier norm at level k (even for specific small values of k such as $k = 2$) is given by exciting recent results in unconditional *pseudorandomness*. More generally, the notion of *Fourier growth* is a convenient way of capturing the L_1 Fourier norm at level k for every k:

Definition 2 (Fourier growth)**.** A function class $\mathcal{F} \subseteq \{f\colon \{-1,1\}^n \to \{0,1\}\}$ has *Fourier growth* $L_1(a, b)$ if there exist constants a and b such that $L_{1,k}(\mathcal{F}) \le a \cdot b^k$ for every k.

The notion of Fourier growth was explicitly introduced by Reingold, Steinke, and Vadhan in [RSV13] for the purpose of constructing pseudorandom generators for space-bounded computation (though we note that the Fourier growth of DNF formulas was already analyzed in [Man95], motivated by applications in learning theory). In recent years there has been a surge of research interest in understanding the Fourier growth of different types of functions [Tal17, GSTW16, CHRT18, Lee19, GRZ20, Tal20, SSW20, GRT21]. One strand of motivation for this study has come from the study of quantum computing; in particular, bounds on the Fourier growth of AC^0 [Tal17] were used in the breakthrough result of Raz and Tal [RT19] which gave an oracle separation between the classes BQP and PH. More recently, in order to achieve an optimal separation between quantum and randomized query complexity, several researchers [Tal20, BS21, SSW20] have studied the Fourier growth of decision trees, with the recent work of [SSW20] obtaining optimal bounds. Analyzing the Fourier growth of other classes of functions has also led to separations between quantum and classical computation in other settings [GRT21, GRZ20, GTW21].

Our chief interest in the current paper arises from a different line of work which has established powerful applications of Fourier growth bounds in pseudorandomness. We describe the relevant background, which motivates a new conjecture that we propose on Fourier growth, in the next subsection.

1.2 Motivation for this work: Fourier growth, pseudorandomness, $\mathbb{F}_2$-polynomials, and the [CHLT19] conjecture

Pseudorandom generators from Fourier growth bounds. Constructing explicit, unconditional pseudorandom generators (PRGs) for various classes of Boolean functions is an important goal in complexity theory. In the recent work [CHHL19], Chattopadhyay, Hatami, Hosseini, and Lovett introduced a novel framework for the design of such PRGs. Their approach provides an explicit pseudorandom generator for any class of functions that is closed under restrictions and has bounded Fourier growth:

Theorem 3 (PRGs from Fourier growth: Theorem 23 of [CHHL19])**.** *Let $\mathcal{F}$ be a family of n-variable Boolean functions that is closed under restrictions and has Fourier growth $L_1(a, b)$. Then there is an explicit pseudorandom generator that ϵ-fools $\mathcal{F}$ with seed length $O(b^2 \log(n/\epsilon)(\log\log n + \log(a/\epsilon)))$.*

Building on Theorem 3, in [CHLT19] Chattopadhyay, Hatami, Lovett, and Tal showed that in fact it suffices to have a bound just on $L_{1,2}(\mathcal{F})$ in order to obtain an efficient PRG for $\mathcal{F}$:

Theorem 4 (PRGs from L_1 Fourier norm bounds at level $k = 2$: Theorem 2.1 of [CHLT19])**.** *Let $\mathcal{F}$ be a family of n-variable Boolean functions that is closed under restrictions and has $L_{1,2}(\mathcal{F}) \leq t$. Then there is an explicit pseudorandom generator that ϵ-fools $\mathcal{F}$ with seed length $O((t/\epsilon)^{2+o(1)} \cdot \mathrm{polylog}(n))$.*

Observe that while Theorem 4 requires a weaker structural result than Theorem 3 (a bound only on $L_{1,2}(\mathcal{F})$ as opposed to $L_{1,k}(\mathcal{F})$ for all $k \geq 1$), the resulting pseudorandom generator is quantitatively weaker since it has seed length polynomial rather than logarithmic in the error parameter $1/\epsilon$. Even more recently, in [CGL+20] Chattopadhyay, Gaitonde, Lee, Lovett, and Shetty further developed this framework by interpolating between the two results described above. They showed that a bound on $L_{1,k}$[1] for any $k \geq 3$ suffices to give a PRG, with a seed length whose ϵ-dependence scales with k:

Theorem 5 (PRGs from L_1 Fourier norm bounds up to level k for any k: Theorem 4.3 of [CGL+20])**.** *Let $\mathcal{F}$ be a family of n-variable Boolean functions that is closed under restrictions and has $L_{1,k}(\mathcal{F}) \leq b^k$ for some $k \geq 3$. Then there exists a pseudorandom generator that ϵ-fools $\mathcal{F}$ with seed length* $O\left(\frac{b^{2+\frac{4}{k-2}} \cdot k \cdot \mathrm{polylog}(\frac{n}{\epsilon})}{\epsilon^{\frac{2}{k-2}}}\right)$.

$\mathbb{F}_2$-polynomials and the [CHLT19] conjecture. The works [CHHL19] and [CHLT19] highlighted the challenge of proving $L_{1,k}$ bounds for the class of *bounded-degree $\mathbb{F}_2$-polynomials* as being of special interest. Let

$$\mathsf{Poly}_{n,d} := \text{ the class of all } n\text{-variate } \mathbb{F}_2\text{-polynomials of degree } d.$$

[1] In fact, they showed that a bound on the weaker quantity $M_{1,k}(f) := \max_{x \in \{-1,1\}^n} |\sum_{|S|=k} \widehat{f}(S) x^S|$ suffices.

It follows from Theorem 4 that even proving

$$L_{1,2}(\mathsf{Poly}_{n,\mathrm{polylog}(n)}) \le n^{0.49} \tag{1}$$

would give nontrivial PRGs for $\mathbb{F}_2$-polynomials of polylog(n) degree, improving on [BV10, Lov09, Vio09b]. By the classic connection (due to Razborov [Raz87]) between such polynomials and the class $\mathsf{AC}^0[\oplus]$ of constant-depth circuits with parity gates, this would also give nontrivial PRGs, of seed length n^{1-c}, for $\mathsf{AC}^0[\oplus]$. This would be a breakthrough improvement on existing results, which are poor either in terms of seed length [FSUV13] or in terms of explicitness [CLW20].

The authors of [CHLT19] in fact conjectured the following bound, which is much stronger than Equation (1):

Conjecture 6 ([CHLT19])**.** *For all $d \ge 1$, it holds that $L_{1,2}(\mathsf{Poly}_{n,d}) = O(d^2)$.*

Extending the [CHLT19] conjecture. Given Conjecture 6, and in light of Theorem 5, it is natural to speculate that an even stronger result than Conjecture 6 might hold. We consider the following natural generalization of the [CHLT19] conjecture, extending it from $L_{1,2}(\mathsf{Poly}_{n,d})$ to $L_{1,k}(\mathsf{Poly}_{n,d})$:

Conjecture 7. *For all $d, k \ge 1$, it holds that $L_{1,k}(\mathsf{Poly}_{n,d}) = O(d)^k$.*

We note that if p is the AND function on d bits, then an easy computation shows that $L_{1,k}(p) = \mathbf{Pr}[p = 1] \cdot \binom{d}{k}$. Moreover, in Appendix A we provide a reduction showing that this implies that the upper bound $O(d)^k$ in Conjecture 7 is best possible for any constant k.

For positive results, the work [CHLT19] proved that $L_{1,1}(\mathsf{Poly}_{n,d}) \le 4d$, and already in [CHHL19] it was shown that $L_{1,k}(\mathsf{Poly}_{n,d}) \le (2^{3d} \cdot k)^k$, but to the best of our knowledge no other results towards Conjecture 6 or Conjecture 7 are known.

Given the apparent difficulty of resolving Conjecture 6 and Conjecture 7 in the general forms stated above, it is natural to study $L_{1,2}$ and $L_{1,k}$ bounds for specific subclasses of degree-d $\mathbb{F}_2$-polynomials. This study is the subject of our main structural results, which we describe in the next subsection.

1.3 Our results: Fourier bounds for structured $\mathbb{F}_2$-polynomials

Our main results show that $L_{1,2}$ and $L_{1,k}$ bounds of the flavor of Conjecture 6 and Conjecture 7 indeed hold for several well-studied classes of $\mathbb{F}_2$-polynomials, specifically *symmetric* $\mathbb{F}_2$-polynomials and *read*-Δ $\mathbb{F}_2$-polynomials. We additionally prove a composition theorem that allows us to combine such polynomials (or, more generally, any polynomials that satisfy certain $L_{1,k}$ bounds) in a natural way and obtain $L_{1,k}$ bounds on the resulting combined polynomials.

Before describing our results in detail, we pause to briefly explain why (beyond the fact that they are natural mathematical objects) such "highly structured" polynomials are attractive targets of study given known results. It has been known for more than ten years [BHL12, Lemma 2] that for any degree $d < (1-\epsilon)n$, a *random* $\mathbb{F}_2$-polynomial of degree d (constructed by independently including each monomial of degree at most d with probability 1/2) is extremely unlikely to have bias larger than $\exp(-n/d)$. It follows that as long as d is not too large, a random degree-d polynomial p is overwhelmingly likely to have $L_{1,k}(p) = o_n(1)$, which is much smaller than d^k. (To verify this, consider the polynomials p_S obtained by XORing p with the parity function $\sum_{i \in S} x_i$. Note that the bias of p_S is the Fourier coefficient of $(-1)^p$ on S. Now apply [BHL12, Lemma 2] to each polynomial p_S, and sum the terms.)

Since the conjectures hold true for random polynomials, it is natural to investigate highly structured polynomials.

1.3.1 Symmetric $\mathbb{F}_2$-polynomials

A *symmetric* $\mathbb{F}_2$-polynomial over $x_1, \dots, x_n$ is one whose output depends only on the Hamming weight of its input x. Such a polynomial of degree d can be written in the form

$$p(x) := \sum_{k=0}^{d} c_k \sum_{|S|=k, S \subseteq [n]} \prod_{i \in S} x_i,$$

where $c_0, \dots, c_d \in \{0, 1\}$. While symmetric polynomials may seem like simple objects, their study can sometimes lead to unexpected discoveries; for example, a symmetric, low-degree $\mathbb{F}_2$-polynomial provided a counterexample to the "Inverse conjecture for the Gowers norm" [LMS11, GT09].

We prove the following upper and lower bounds on the L_1 Fourier norm at level k for any symmetric polynomial:

Theorem 8. *Let $p(x_1, \dots, x_n)$ be a symmetric $\mathbb{F}_2$-polynomial of degree d. Then $L_{1,k}(p) \le \mathbf{Pr}[p = 1] \cdot O(d)^k$ for every k.*

Theorem 9. *For every k, there is a symmetric $\mathbb{F}_2$-polynomial $p(x_1, \dots, x_n)$ of degree $d = \Theta(\sqrt{kn})$ such that $L_{1,k}(p) \ge (e^{-k}/2) \cdot \binom{n}{k}^{1/2} = \Omega_k(1) \cdot d^k$.*

Theorem 8 verifies the [CHLT19] conjecture (Conjecture 6), and even the generalized version Conjecture 7, for the class of symmetric polynomials. Theorem 9 complements it by showing that the upper bounds in Theorem 8 are tight for $k = O(1)$ when $d = \Theta(\sqrt{n})$.

Theorem 8 also provides a quadratic sharpening of an earlier bound that was implicit in [RSV13] (as well as providing the "correct" dependence on $\mathbf{Pr}[p = 1]$). In [RSV13] Reingold, Steinke and Vadhan showed that any function f computed by an oblivious, read-once, regular branching program of width w has $L_{1,k}(f) \le (2w^2)^k$. It follows directly from a result of [BGL06] (Lemma 17 below) that any symmetric $\mathbb{F}_2$-polynomial p of degree d can be computed by an oblivious, read-once, regular branching program of width at most $2d$, and hence the [RSV13] result implies that $L_{1,k}(p) \le 8^k d^{2k}$.

1.3.2 Read-Δ $\mathbb{F}_2$-polynomials

For $\Delta \ge 1$, a *read-*Δ $\mathbb{F}_2$-polynomial is one in which each input variable appears in at most Δ monomials. The case $\Delta = 1$ corresponds to the class of *read-once polynomials*, which are simply sums of monomials over disjoint sets of variables; for example, the polynomial $x_1x_2 + x_3x_4$ is read-once whereas $x_1x_2 + x_1x_4$ is read-twice. Read-once polynomials have been studied from the perspective of pseudorandomness [LV20, MRT19, Lee19, DHH20] as they capture several difficulties in improving Nisan's generators [Nis92] for width-4 read-once branching programs.

We show that the $L_{1,k}$ Fourier norm of read-Δ polynomials is polynomial in d and Δ:

Theorem 10. *Let $p(x_1, \ldots, x_n)$ be a read-Δ polynomial of degree d. Then $L_{1,k}(p) \leq \mathbf{Pr}[p = 1] \cdot O(k)^k \cdot (d\Delta)^{8k}$.*

The work [Lee19] showed that read-once polynomials satisfy an $L_{1,k}$ bound of $O(d)^k$ and this is tight for every k, but we are not aware of previous bounds on even the L_1 Fourier norm at level $k = 2$ for read-Δ polynomials, even for $\Delta = 2$.

As any monomial with degree $\Omega(\log n)$ vanishes under a random restriction with high probability, we have the following corollary which applies to polynomials of any degree.

Corollary 11. *Let $p(x_1, \ldots, x_n)$ be a read-Δ polynomial. Then $L_{1,k}(p) \leq O(k)^{9k} \cdot (\Delta \log n)^{8k}$.*

1.3.3 A composition theorem

The upper bounds of Theorem 8 and Theorem 10 both include a factor of $\mathbf{Pr}[p = 1]$. (We observe that negating p, i.e. adding 1 to it, does not change its $L_{1,2}$ or $L_{1,k}$ and keeps p symmetric (respectively, read-Δ) if it was originally symmetric (respectively, read-Δ), and hence in the context of those theorems we can assume that this $\mathbf{Pr}[p = 1]$ factor is at most $1/2$.) Level-k bounds that include this factor have appeared in earlier works for other classes of functions [OS07, BTW15, CHRT18, Tal20, GTW21], and have been used to obtain high-level bounds for other classes of functions [CHRT18, Tal20, GTW21] and to extend level-k bounds to more general classes of functions [Lee19]. Having these $\mathbf{Pr}[p = 1]$ factors in Theorem 8 and Theorem 10 is important for us in the context of our composition theorem, which we now describe. We begin by defining the notion of a *disjoint composition* of functions:

Definition 12. Let $\mathcal{F}$ be a class of functions from $\{-1, 1\}^m$ to $\{-1, 1\}$ and let $\mathcal{G}$ be a class of functions from $\{-1, 1\}^\ell$ to $\{-1, 1\}$. Define the class $\mathcal{H} = \mathcal{F} \circ \mathcal{G}$ of *disjoint compositions of $\mathcal{F}$ and $\mathcal{G}$* to be the class of all functions from $\{-1, 1\}^{m\ell}$ to $\{-1, 1\}$ of the form

$$h(x^1, \ldots, x^m) = f(g_1(x^1), \ldots, g_m(x^m)),$$

where $g_1, \ldots, g_m \in \mathcal{G}$ are defined on m disjoint sets of variables and $f \in \mathcal{F}$.

As an example of this definition, the class of *block-symmetric* polynomials (i.e. polynomials whose variables are divided into blocks and are symmetric within each block but not overall) are a special case of disjoint compositions where $\mathcal{G}$ is taken to be the class of symmetric polynomials. We remark that block-symmetric polynomials are known to correlate better with parities than symmetric polynomials in certain settings [GKV17].

We prove a composition theorem for upper-bounding the L_1 Fourier norm at level k of the disjoint composition of any classes of functions that are closed under restriction and admit a $L_{1,k}$ bound of the form $\mathbf{Pr}[f = 1] \cdot a \cdot b^k$:

Theorem 13. *Let $g_1, \ldots, g_m \in \mathcal{G}$ and let $f \in \mathcal{F}$, where $\mathcal{F}$ is closed under restrictions. Suppose that for every $1 \leq k \leq K$, we have*

1. *$L_{1,k}(f) \leq \mathbf{Pr}[f = 1] \cdot a_{\mathsf{out}} \cdot b_{\mathsf{out}}^k$ for every $f \in \mathcal{F}$, and*

2. $L_{1,k}(g) \le \mathbf{Pr}[g = 1] \cdot a_{\mathsf{in}} \cdot b_{\mathsf{in}}^k$ *for every* $g \in \mathcal{G}$.

Then for every ± 1*-valued function* $h \in \mathcal{H} = \mathcal{F} \circ \mathcal{G}$*, we have that*

$$L_{1,K}(h) \le \mathbf{Pr}[h = 1] \cdot a_{\mathsf{out}} \cdot (a_{\mathsf{in}} b_{\mathsf{in}} b_{\mathsf{out}})^K.$$

See Theorem 31 for a slightly sharper bound. We remark that Theorem 13 does not assume any $\mathbb{F}_2$-polynomial structure for the functions in $\mathcal{F}$ or $\mathcal{G}$ and thus may be of broader utility.

1.4 Applications of our results

Our structural results imply new pseudorandom generators and correlation bounds.

1.4.1 Pseudorandom generators

Combining our Fourier bounds with the polarizing framework, we obtain new PRGs for read-few polynomials. The following theorem follows from applying Theorem 5 with some $k = \Theta(\log n)$ and the $L_{1,k}$ bound in Corollary 11.

Theorem 14. *There is an explicit pseudorandom generator that* ϵ*-fools read-*Δ $\mathbb{F}_2$*-polynomials with seed length* $\mathrm{poly}(\Delta, \log n, \log(1/\epsilon))$.

For constant ϵ, this improves on a PRG by Servedio and Tan [ST19a], which has a seed length of $2^{O(\sqrt{\log(\Delta n)})}$. (Note that read-$\Delta$ polynomials are also (Δn)-sparse.) We are not aware of any previous PRG for read-2 polynomials with $\mathrm{polylog}(n)$ seed length.

Note that the OR function has L_1 Fourier norm $O(1)$. By expressing a DNF in the Fourier expansion of OR in its terms, it is not hard to see that the same PRG also fools the class of read-Δ DNFs (and read-Δ CNFs similarly) [ST19b].

1.4.2 Correlation bounds

Exhibiting explicit Boolean functions that do not *correlate* with low-degree polynomials is a fundamental challenge in complexity. Perhaps surprisingly, this challenge stands in the way of progress on a striking variety of frontiers in complexity, including circuits, rigidity, and multiparty communication complexity. For a survey of correlation bounds and discussions of these connections we refer the reader to [Vio09a, Vio17, Vio21].

For polynomials of degree larger than $\log_2 n$, the state-of-the-art remains the lower bound proved by Razborov and Smolensky in the 1980s' [Raz87, Smo87], showing that for any degree-d polynomial p and an explicit function h (in fact, majority) we have:

$$\mathbf{Pr}[p(x) = h(x)] \le 1/2 + O(d/\sqrt{n}).$$

Viola [Vio20] recently showed that upper bounds on $L_{1,k}(\mathcal{F})$ imply correlation bounds between $\mathcal{F}$ and an explicit function h_k that is related to majority and is defined as

$$h_k(x) := \mathrm{sgn}\left(\sum_{|S|=k} x^S\right).$$

In particular, proving Conjecture 6 or related conjectures implies new correlation bounds beating Razborov–Smolensky. The formal statement of the connection is given by the following theorem.

Theorem 15 (Theorem 1 in [Vio20]). *For every $k \in [n]$ and $\mathcal{F} \subseteq \{f\colon \{0,1\}^n \to \{-1,1\}\}$, there is a distribution D_k on $\{0,1\}^n$ such that for any $f \in \mathcal{F}$,*

$$\Pr_{x \sim D_k}[f(x) = h_k(x)] \le \frac{1}{2} + \frac{e^k}{2\sqrt{\binom{n}{k}}} L_{1,k}(\mathcal{F}).$$

For example, if $k = 2$ and we assume that the answer to Conjecture 6 is positive, then the right-hand side above becomes $1/2 + O(d^2/n)$, which is a quadratic improvement over the bound by Razborov and Smolensky.

Therefore, Theorems 8 and 10 imply correlation bounds between these polynomials and an explicit function that are better than $O(d/\sqrt{n})$ given in [Raz87, Smo87]. We note that via a connection in [Vio09b], existing PRGs for these polynomials already imply strong correlation bounds between these polynomials and the class of NP. Our results apply to more general classes via the composition theorem, where it is not clear if previous techniques applied. For a concrete example, consider the composition of a degree-(n^α) symmetric polynomial with degree-(n^α) read-(n^α) polynomials. Theorem 13 shows that such polynomial has $L_{1,2} \le n^{O(\alpha)}$. For a sufficiently small $\alpha = \Omega(1)$, we again obtain correlation bounds improving on Razborov–Smolensky.

1.5 Related work

We close this introduction by discussing a recent work of Girish, Tal and Wu [GTW21] on parity decision trees that is related to our results.

Parity decision trees are a generalization of decision trees in which each node queries a parity of some input bits rather than a single input bit. The class of depth-d parity decision trees is a subclass of $\mathbb{F}_2$ degree-d polynomials, as such a parity decision tree can be expressed as a sum of products of sums over $\mathbb{F}_2$, where each product corresponds to a path in the tree (and hence gives rise to $\mathbb{F}_2$-monomials of degree at most d). The Fourier spectrum of parity decision trees was first studied in [BTW15], which obtained a level-1 bound of $O(\sqrt{d})$. This bound was recently extended to higher levels in [GTW21], showing that any depth-d parity decision tree T over n variables has $L_{1,k}(T) \le d^{k/2} \cdot O(k \log n)^k$.

2 Our techniques

We now briefly explain the approaches used to prove our results. We note that each of these results is obtained using very different ingredients, and hence the results can be read independently of each other.

2.1 Symmetric polynomials (Theorems 8 and 9, Section 4)

The starting point of our proof is a result from [BGL06], which says that degree-d symmetric $\mathbb{F}_2$-polynomials only depend on the Hamming weight of their input modulo m for some m (a power of two) which is $\Theta(d)$, and the converse is also true. (See Lemma 17 for the exact statement.) We now explain how to prove our upper and lower bounds given this.

For the upper bound (Theorem 8), since $p(x)$ takes the same value for all strings x with the same weights ℓ mod m, to analyze $L_{1,k}(p)$ it suffices to analyze $\mathbf{E}[(-1)^{x_1+\cdots+x_k}]$ conditioned on x having Hamming weight exactly ℓ mod m.

We bound this conditional expectation by considering separately two cases depending on whether or not $k \le n/m^2$. For the case that $k \le n/m^2$, we use a (slight sharpening of a) result from [BHLV19], which gives a bound of $m^{-k}e^{-\Omega(n/m^2)}$. In the other case, that $k \ge n/m^2$, in Lemma 19 we prove a bound of $O(km/n)^k$. This is established via a careful argument that gives a new bound on the Kravchuk polynomial in certain ranges (see Claim 22), extending and sharpening similar bounds that were recently established in [CPT20] (the bounds of [CPT20] would not suffice for our purposes).

In each of the above two cases, summing over all the $\binom{n}{k}$ coefficients gives the desired bound of $O(m)^k = O(d)^k$.

For the lower bound (Theorem 9), we first consider the symmetric function $h_k(x) := \mathrm{sgn}(\sum_{|S|=k}(-1)^{\sum_{i\in S} x_i})$. (Note that h_1 is the majority function.) As h_k has degree k, it follows from a consequence of the hypercontractive inequality that $L_{1,k}(h_k) \ge e^{-k}\binom{n}{k}^{1/2}$. We define our polynomial p so that on inputs x whose Hamming weight $|x|$ is within $m = \Theta(\sqrt{kn})$ from $n/2$, it agrees with h_k, and on other inputs its value is defined such that p only depends on $|x|$ mod $2m$. By the converse of [BGL06], p can be computed by a symmetric polynomial of degree less than $2m$. Our lower bound then follows from $L_{1,k}(f) \le \binom{n}{k}^{1/2}$ for any Boolean function f, and that by our choice of m, the two functions p and h_k disagree on $e^{-\Omega(k)}$ fraction of the inputs.

2.2 Read-Δ polynomials (Theorem 10, Section 5)

Writing $f := (-1)^p$ for an $\mathbb{F}_2$-polynomial p, we observe that the coefficient $\widehat{f}(S)$ is simply the bias of $p_S(x) := p(x) + \sum_{i\in S} x_i$. Our high-level approach is to decompose the read-few polynomial p_S into many disjoint components, then show that each component has small bias. Since the components are disjoint, the product of these biases gives an upper bound on the bias of p_S.

In more detail, we first partition the variables according to the minimum degree t_i of the monomials containing each variable x_i. Then we start decomposing p_S by collecting all the monomials in p containing x_i to form the polynomial p_i. We observe that the larger t_i is, the more likely p_i is to vanish on a random input, and therefore the closer $p_i + x_i$ is to being unbiased. For most S, we can pick many such p_i's ($i \in S$) from p so that they are disjoint. For the remaining polynomial r, because Δ and d are small, we can further decompose r into many disjoint polynomials r_i. Finally, our upper bound on $|\widehat{f}(S)|$ will be the magnitude of the product of the biases of the p_i's and r_i's. We note that our decomposition of p uses the structure of S; and so the upper bound on $\widehat{f}(S)$ depends on S (see Lemma 27). Summing over each $|\widehat{f}(S)|$ gives our upper bound.

2.3 Composition theorem (Theorem 13, Section 6)

As a warmup, let us first consider directly computing a degree-1 Fourier coefficient $\widehat{h}(\{(i,j)\})$ of the composition. Since the inner functions g_i depend on disjoint variables, by writing the outer function f in its Fourier expansion, it is not hard to see that

$$\widehat{h}(\{(i,j)\}) = \sum_{S\ni i} \widehat{f}(S) \prod_{\ell\in S\setminus\{i\}} \mathbf{E}[g_\ell] \cdot \widehat{g}_i(\{j\}).$$

When the g_i's are balanced, i.e. $\mathbf{E}[g_i] = 0$, we have $\widehat{f}(\{(i,j)\}) = \widehat{f}(\{i\})\widehat{g_i}(\{j\})$, and it follows that $L_{1,1}(h) \leq L_{1,1}(\mathcal{F})L_{1,1}(\mathcal{G})$. To handle the unbalanced case, we apply an idea from [CHHL19] that lets us relate $\sum_{S \ni i} \widehat{f}(S) \prod_{\ell \in S \setminus \{i\}} \mathbf{E}[g_\ell]$ to the average of $\widehat{f_R}(\{i\})$, for some suitably chosen random restriction R on f (see Claim 33). As $\mathcal{F}$ is closed under restrictions, we can apply the $L_{1,1}(\mathcal{F})$ bound on f_R, which in turns gives a bound on $\sum_{S \ni i} \widehat{f}(S) \prod_{\ell \in S \setminus \{i\}} \mathbf{E}[g_\ell]$ in terms of $L_{1,1}(\mathcal{F})$ and $\mathbf{E}[g_i]$.

Bounding $L_{1,k}(h)$ for $k \geq 2$ is more complicated, as each $\widehat{h}(S)$ involves $\widehat{f}(J)$ and $\widehat{g_i}(T)$'s, where the sets J and T have different sizes. We provide more details in Section 6.

3 Preliminaries

Notation. For a string $x \in \{0,1\}^n$ we write $|x|$ to denote its Hamming weight $\sum_{i=1}^n x_i$. We use $\mathcal{X}_w$ to denote $\{x : |x| = w\}$, the set of n-bit strings with Hamming weight w, and $\mathcal{X}_{\ell \bmod m} = \bigcup_{w : w \equiv \ell \bmod m} \mathcal{X}_w = \{x : |x| \equiv \ell \bmod m\}$.

We recall that for an n-variable Boolean function f, the *level-k Fourier L_1 norm of f* is

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

We note that a function f and its negation have the same $L_{1,k}$ for $k \geq 1$. Hence we can often assume that $\mathbf{Pr}[f = 1] \leq 1/2$, or replace the occurrence of $\mathbf{Pr}[f = 1]$ in a bound by $\min\{\mathbf{Pr}[f = 1], \mathbf{Pr}[f = 0]\}$ for a $\{0,1\}$-valued function f (or by $\min\{\mathbf{Pr}[f = 1], \mathbf{Pr}[f = -1]\}$ for a $\{-1,1\}$-valued function). If f is a $\{-1,1\}$-valued function then $\frac{1-|\mathbf{E}[f]|}{2}$ is equal to $\min\{\mathbf{Pr}[f = 1], \mathbf{Pr}[f = -1]\}$, and we will often write $\frac{1-|\mathbf{E}[f]|}{2}$ for convenience.

Unless otherwise indicated, we will use the letters p, q, r, etc. to denote $\mathbb{F}_2$-polynomials (with inputs in $\{0,1\}^n$ and outputs in $\{0,1\}$) and the letters f, g, h, etc. to denote general Boolean functions (where the inputs may be $\{0,1\}^n$ or $\{-1,1\}^n$ and the outputs may be $\{0,1\}$ or $\{-1,1\}$ depending on convenience). We note that changing from $\{0,1\}$ outputs to $\{-1,1\}$ outputs only changes $L_{1,k}$ by a factor of 2.

We use standard multilinear monomial notation as follows: given a vector $\beta = (\beta_1, \ldots, \beta_n)$ and a subset $T \subseteq [n]$, we write β^T to denote $\prod_{j \in T} \beta_j$.

4 $L_{1,k}$ bounds for symmetric polynomials

The main result of this section is Theorem 16, which gives an upper bound on $L_{1,k}(p)$ for any symmetric $\mathbb{F}_2$-polynomial p of degree d, covering the entire range of parameters $1 \leq k, d \leq n$:

Theorem 16 (Restatement of Theorem 8)**.** *Let $p \colon \{0,1\}^n \to \{0,1\}$ be a symmetric $\mathbb{F}_2$-polynomial of degree d. For every $1 \leq k, d \leq n$,*

$$L_{1,k}(p) := \sum_{|S|=k} |\widehat{p}(S)| \leq \mathbf{Pr}[p(x) = 1] \cdot O(d)^k.$$

Proof idea. As the polynomial p is symmetric, its Fourier coefficient $\widehat{p}(S)$ only depends on $|S|$, the size of S. Hence to bound $L_{1,k}$ it suffices to analyze the coefficient $\widehat{p}(\{1, \ldots, k\}) = \mathbf{E}_{x \sim \{0,1\}^n}[p(x)(-1)^{x_1+\cdots+x_k}]$.

Our proof uses a result from [BGL06] (Lemma 17 below), which says that degree-d symmetric $\mathbb{F}_2$-polynomials only depend on the Hamming weight of their input modulo m for some $m = O(d)$. Given this, since $p(x)$ takes the same value for strings x with the same weights $\ell \bmod m$, we can in turn bound each $\mathbf{E}[(-1)^{x_1+\cdots+x_k}]$ conditioned on x having Hamming weight exactly $\ell \bmod m$, i.e. $x \in \mathcal{X}_{\ell \bmod m}$. We consider two cases depending on whether or not $k \leq n/m^2$. If $k \leq n/m^2$, we can apply a (slight sharpening of a) result from [BHLV19], which gives a bound of $m^{-k}e^{-\Omega(n/m^2)}$. If $k \geq n/m^2$, in Lemma 19 we prove a bound of $O(km/n)^k$. In each case, summing over all the $\binom{n}{k}$ coefficients gives the desired bound of $O(m)^k = O(d)^k$.

We now give some intuition for Lemma 19, which upper bounds the magnitude of the ratio

$$\mathop{\mathbf{E}}_{x \sim \mathcal{X}_{\ell \bmod m}}[(-1)^{x_1+\cdots+x_k}] = \frac{\sum_{x \in \mathcal{X}_{\ell \bmod m}} (-1)^{x_1+\cdots+x_k}}{|\mathcal{X}_{\ell \bmod m}|} \tag{2}$$

by $O(km/n)^k$. Let us first consider $k = 1$ and $m = \Theta(\sqrt{n})$. As most strings x have Hamming weight within $[n/2 - \Theta(\sqrt{n}), n/2 + \Theta(\sqrt{n})]$, it is natural to think about the weight $|x|$ in the form of $n/2 + m\mathbb{Z} + \ell'$. It is easy to see that the denominator is at least $\Omega(2^n/\sqrt{n})$, so we focus on bounding the numerator. Consider the quantity $\sum_{x \in \mathcal{X}_{n/2+s}} \mathbf{E}[(-1)^{x_1}]$ for some s. As we are summing over all strings of the same Hamming weight, we can instead consider $\sum_{x \in \mathcal{X}_{n/2+s}} \mathbf{E}_{i \sim [n]}[(-1)^{x_i}]$ (see Equation (5)). For any string of weight $n/2 + s$, it is easy to see that

$$\mathop{\mathbf{E}}_{i \sim [n]}[(-1)^{x_i}] = (1/2 - s/n) - (1/2 + s/n) = -2s/n. \tag{3}$$

Therefore, in the $k = 1$ case we get that

$$\left| \mathop{\mathbf{E}}_{x \sim \mathcal{X}_{\ell \bmod m}}[(-1)^{x_1+\cdots+x_k}] \right| \leq 2 \sum_c \binom{n}{n/2 + cm + \ell'} \frac{|cm + \ell'|}{n}.$$

Using the fact that $\binom{n}{n/2+cm+\ell'}$ is exponentially decreasing in $|c|$, in Claim 20 we show that this is at most $O(2^n/n)$. So the ratio in (2) is at most $O(1/\sqrt{n})$, as desired, when $k = 1$.

However, already for $k = 2$, a direct (but tedious) calculation shows that

$$\mathop{\mathbf{E}}_{i<j}[(-1)^{x_i+x_j}] = \frac{4s^2 - 2ns + n}{n(n-1)}, \tag{4}$$

which no longer decreases in s like in (3). Nevertheless, we observe that this is bounded by $O(1/n + (|s|/n)^2)$, which is sufficient for bounding the ratio by $O(1/n)$. Building on this, for any k we obtain a bound of $2^{O(k)}((k/n)^{k/2} + (|s|/n))^k$ in Claim 22, and by a more careful calculation we are able to obtain the desired bound of $O(km/n)^k$ on Equation (2).

4.1 Proof of Theorem 16

We now prove the theorem. We will use the following result from [BGL06], which says that degree-d symmetric $\mathbb{F}_2$-polynomials only depend on their input's Hamming weight modulo $O(d)$.

Lemma 17 (Corollaries 2.5 and 2.7 in [BGL06], $p = 2$)**.** *Let $p\colon \{0,1\}^n \to \{0,1\}$ be any function and m be a power of two. If p is a symmetric $\mathbb{F}_2$-polynomial of degree d, where $m/2 \le d < m$, then $p(x)$ only depends on $|x| \bmod m$. Conversely, if $p(x)$ only depends on $|x| \bmod m$, then it can be computed by a symmetric $\mathbb{F}_2$-polynomial of degree less than m.*

We will also use two bounds on the biases of parities under the uniform distribution over $\mathcal{X}_{\ell \bmod m}$, one holds for $k \le n/(2d)^2 \le n/m^2$ (Claim 18) and the other for $k \ge n/(2d)^2 \ge n/(4m^2)$ (Lemma 19). Claim 18 is essentially taken from [BHLV19]. However, the statement in [BHLV19] has a slightly worse bound, so below we explain the changes required to give the bound of Claim 18. The proof of Lemma 19 involves bounding the magnitude of Kravchuk polynomials. As it is somewhat technical we defer its proof to Section 4.2.

Claim 18 (Lemma 10 in [BHLV19])**.** *For every $1 \le k \le n/m^2$ and every integer ℓ,*

$$2^{-n}\Bigg|\sum_{x\in\mathcal{X}_{\ell \bmod m}} (-1)^{x_1+\cdots+x_k}\Bigg| \le m^{-(k+1)}e^{-\Omega(n/m^2)},$$

while for $k = 0$,

$$\Big|2^{-n}|\mathcal{X}_{\ell \bmod m}| - 1/m\Big| \le m^{-1}e^{-\Omega(n/m^2)}.$$

Lemma 19. *For $k \ge n/(4m^2)$, we have*

$$\binom{n}{k} \cdot \max_{\ell}\left|\frac{\sum_{x\in\mathcal{X}_{\ell \bmod m}}(-1)^{x_1+\cdots+x_k}}{|\mathcal{X}_{\ell \bmod m}|}\right| \le O(m)^k.$$

We now use Claim 18 and Lemma 19 to prove Theorem 16.

Proof of Theorem 16. As p is symmetric, all the level-k coefficients are the same, so it suffices to give a bound on $\widehat{p}(\{1,2,\ldots,k\})$. Let $\tilde{p}\colon \{0,\ldots,n\} \to \{0,1\}$ be the function defined by $\tilde{p}(|x|) := p(x_1,\ldots,x_n)$. By Lemma 17, we have $\tilde{p}(\ell) = \tilde{p}(\ell \bmod m)$ for some $d < m \le 2d$ where m is a power of 2. Using the definition of $\widehat{p}(\{1,\ldots,k\})$, we have

$$\begin{aligned}
|\widehat{p}(\{1,\ldots,k\})| &= \left|\mathop{\mathbf{E}}_{x\sim\{0,1\}^n}\left[p(x)(-1)^{x_1+\cdots+x_k}\right]\right| \\
&= \left|\sum_{\ell=0}^{m-1} \tilde{p}(\ell)\frac{|\mathcal{X}_{\ell \bmod m}|}{2^n}\cdot\frac{\sum_{x\in\mathcal{X}_{\ell \bmod m}}(-1)^{x_1+\cdots+x_k}}{|\mathcal{X}_{\ell \bmod m}|}\right| \\
&\le \mathbf{E}[p]\cdot\max_{0\le\ell\le m-1}\left|\frac{\sum_{x\in\mathcal{X}_{\ell \bmod m}}(-1)^{x_1+\cdots+x_k}}{|\mathcal{X}_{\ell \bmod m}|}\right|,
\end{aligned}$$

where we use the shorthand $\mathbf{E}[p] = \mathbf{E}_{x\sim\{0,1\}^n}[p(x)]$ in the last step.

When $k \le n/(2d)^2 \le n/m^2$, by Claim 18 (using the first bound for the numerator and the second $k = 0$ bound for the denominator) we have

$$\max_{0\le \ell \le m-1} \left| \frac{\sum_{x\in\mathcal{X}_{\ell \bmod m}} (-1)^{x_1+\cdots+x_k}}{|\mathcal{X}_{\ell \bmod m}|} \right| \le \frac{m^{-(k+1)} e^{-\Omega(n/m^2)}}{m^{-1}(1-e^{-\Omega(n/m^2)})} \le O(1)\cdot m^{-k} e^{-\Omega(n/m^2)},$$

where the last inequality holds because $1 \le k \le n/m^2$ and hence the $(1-e^{-\Omega(n/m^2)})$ factor in the denominator of the left-hand side is $\Omega(1)$. Hence, summing over all the $\binom{n}{k}$ level-k coefficients, we get that

$$L_{1,k}(p) \le \mathbf{E}[p]\cdot \binom{n}{k} \cdot O(1)\cdot m^{-k} e^{-\Omega(n/m^2)} \le \mathbf{E}[p]\cdot O(1)\cdot m^k \left(\frac{ne}{km^2}\right)^k e^{-\Omega(n/m^2)} \le \mathbf{E}[p]\cdot O(m)^k,$$

where the last inequality is because for constant c, the function $(x/k)^k e^{-cx}$ is maximized when $x = k/c$, and is $O(1)^k$.

When $k \ge n/(2d)^2 \ge n/(4m^2)$, by Lemma 19 we have

$$L_{1,k}(p) \le \mathbf{E}[p]\cdot \binom{n}{k} \max_{0\le \ell\le m-1} \left| \frac{\sum_{x\in\mathcal{X}_{\ell \bmod m}} (-1)^{x_1+\cdots+x_k}}{|\mathcal{X}_{\ell \bmod m}|} \right| \le \mathbf{E}[p]\cdot O(m)^k. \qquad \square$$

It remains to prove Claim 18 and Lemma 19.

Proof of Claim 18. The second inequality (i.e. when $k = 0$) is explicitly shown in [BHLV19, Lemma 10]. For the first inequality (i.e. when $k \ge 1$), following the proof of [BHLV19, Lemma 10] but without using the last inequality in the proof of their Claim 11, we have

$$\left| \sum_{x\in\mathcal{X}_{\ell \bmod m}} (-1)^{x_1+\cdots+x_k} \right| \le \frac{2^n}{m} \sum_{j=0}^{m-1} \left|\cos\frac{j\pi}{m}\right|^{n-k} \left|\sin\frac{j\pi}{m}\right|^k.$$

For $0 < \theta < \pi/2$, we have $\cos\theta \le e^{-\theta^2/2}$ and $\sin\theta \le \theta$. So

$$\begin{aligned} \sum_{j=0}^{m-1} \left|\cos\frac{j\pi}{m}\right|^{n-k} \left|\sin\frac{j\pi}{m}\right|^k &= 2 \sum_{0<j<m/2} \left(\cos\frac{j\pi}{m}\right)^{n-k} \left(\sin\frac{j\pi}{m}\right)^k \\ &\le 2 \sum_{0<j<m/2} e^{-\frac{n-k}{2}\left(\frac{j\pi}{m}\right)^2} \left(\frac{j\pi}{m}\right)^k. \end{aligned}$$

We now observe that each term in the above summation is at most half of the previous one. This is because for every $j \ge 1$ we have

$$\begin{aligned} \frac{\exp\left(-\frac{(n-k)(j+1)^2\pi^2}{2m^2}\right)\left(\frac{(j+1)\pi}{m}\right)^k}{\exp\left(-\frac{(n-k)j^2\pi^2}{2m^2}\right)\left(\frac{j\pi}{m}\right)^k} &\le \exp\left(-\frac{\pi^2(n-k)}{m^2} j\right)\cdot \left(1+\frac{1}{j}\right)^k \\ &\le \exp\left(-\frac{\pi^2(n-k)}{m^2}\right)\cdot 2^k \\ &\le \exp\left(-\frac{3\pi^2}{4} k\right)\cdot 2^k \end{aligned}$$

$$
\begin{aligned}
&\leq \exp\Big(-\frac{3\pi^2}{4}\Big)\cdot 2\\
&\leq \frac{1}{2},
\end{aligned}
$$

where the third step holds because w.l.o.g. $m \geq 2$ and $1 \leq k \leq n/m^2 \leq n/4$, and hence $(n-k)/m^2 \geq 3k/4$. For these reasons, we can deduce that

$$
\sum_{j=0}^{m-1}\left|\cos\frac{j\pi}{m}\right|^{n-k}\left|\sin\frac{j\pi}{m}\right|^{k} \leq 2e^{-\frac{\pi^2(n-k)}{2m^2}}\Big(\frac{\pi}{m}\Big)^{k}\sum_{1\leq j<m/2}2^{1-j} \leq m^{-k}\exp\big(-\Omega(n/m^2)\big). \quad \square
$$

4.2 Proof of Lemma 19

In this section, we prove Lemma 19. Let $\mathcal{K}(n,k,w) := \sum_{x\in\mathcal{X}_w}(-1)^{x_1+\cdots+x_k}$ be the *Kravchuk polynomial.* We recall the "symmetry relation" for this polynomial (see e.g. [Wik21]):

$$
\binom{n}{k}\mathcal{K}(n,k,w) = \sum_{y\in\mathcal{X}_k}\sum_{x\in\mathcal{X}_w}(-1)^{\langle y,x\rangle} = \sum_{x\in\mathcal{X}_w}\sum_{y\in\mathcal{X}_k}(-1)^{\langle x,y\rangle} = \binom{n}{w}\mathcal{K}(n,w,k). \tag{5}
$$

Using Equation (5), we have

$$
\begin{aligned}
\binom{n}{k}\max_{0\leq\ell\leq m-1}\left|\frac{\sum_{x\in\mathcal{X}_{\ell \bmod m}}(-1)^{x_1+\cdots+x_k}}{|\mathcal{X}_{\ell \bmod m}|}\right| &= \max_{\ell}\frac{\left|\sum_{w\equiv\ell \bmod m}\binom{n}{k}\mathcal{K}(n,k,w)\right|}{\sum_{w\equiv\ell \bmod m}\binom{n}{w}}\\
&= \max_{\ell}\frac{\left|\sum_{w\equiv\ell \bmod m}\binom{n}{w}\mathcal{K}(n,w,k)\right|}{\sum_{w\equiv\ell \bmod m}\binom{n}{w}}\\
&\leq \max_{\ell}\frac{\sum_{c}\binom{n}{\lfloor n/2\rfloor+cm+\ell}|\mathcal{K}(n,\lfloor n/2\rfloor+cm+\ell,k)|}{\sum_{c}\binom{n}{\lfloor n/2\rfloor+cm+\ell}}\\
&\leq 2\max_{\ell}\frac{\sum_{c\geq 0}\binom{n}{\lfloor n/2\rfloor+cm+\ell}|\mathcal{K}(n,\lfloor n/2\rfloor+cm+\ell,k)|}{\sum_{c\geq 0}\binom{n}{\lfloor n/2\rfloor+cm+\ell}},
\end{aligned}
\tag{6}
$$

where the indices of c are integers that iterate over $\lfloor n/2\rfloor+cm+\ell\in\{0,\ldots,n\}$, and the last line uses the symmetry of $\binom{n}{w}\mathcal{K}(n,w,k)$ around $w=n/2$.

Later in this section, we will prove (in Claim 22) that

$$
|\mathcal{K}(n,\lfloor n/2\rfloor+s,k)| \leq 2^{O(k)}\left(\Big(\frac{n}{k}\Big)^{\frac{k}{2}}+\Big(\frac{|s|}{k}\Big)^{k}\right).
$$

Plugging this bound into Equation (6), we have

$$
\begin{aligned}
(6) &\leq 2^{O(k)}\max_{\ell}\left(\frac{\sum_{c\geq 0}\binom{n}{\lfloor n/2\rfloor+cm+\ell}\big((n/k)^{k/2}+((cm+\ell)/k)^{k}\big)}{\sum_{c\geq 0}\binom{n}{\lfloor n/2\rfloor+cm+\ell}}\right)\\
&= O(n/k)^{\frac{k}{2}}+\frac{2^{O(k)}}{k^k}\max_{\ell}\left(\frac{\sum_{c\geq 0}\binom{n}{\lfloor n/2\rfloor+cm+\ell}(cm+\ell)^{k}}{\sum_{c\geq 0}\binom{n}{\lfloor n/2\rfloor+cm+\ell}}\right).
\end{aligned}
\tag{7}
$$

We will prove the following claim which bounds the ratio of the two summations.

Claim 20. *For any $\ell \in \{0, \ldots, m-1\}$ and $k \geq n/(2m)^2$, we have*

$$\frac{\sum_{c\geq 0} \binom{n}{\lfloor n/2\rfloor + cm+\ell}(cm+\ell)^k}{\sum_{c\geq 0}\binom{n}{\lfloor n/2\rfloor+cm+\ell}} \leq O(mk)^k.$$

Therefore, as $(n/k)^{1/2} \leq 2m$, we have

$$(7) \leq O(n/k)^{\frac{k}{2}} + \frac{2^{O(k)}}{k^k} \cdot O(mk)^k = O(m)^k.$$

This completes the proof of Lemma 19. $\square$

It remains to prove Claims 20 and 22. We begin with Claim 20.

Proof of Claim 20. Fix any $\ell \in \{0, \ldots, m-1\}$, we first bound

$$\frac{\sum_{c\geq 0} \binom{n}{\lfloor n/2\rfloor + cm+\ell}(cm+\ell)^k}{\sum_{c\geq 0}\binom{n}{\lfloor n/2\rfloor+cm+\ell}}$$

with some quantity that does not depend on ℓ. First, observe that the denominator is at least $\binom{n}{\lfloor n/2\rfloor+\ell}$. Next, we have

$$\begin{aligned}
\frac{\binom{n}{\lfloor n/2\rfloor+\ell+cm}}{\binom{n}{\lfloor n/2\rfloor+\ell}} &= \prod_{i=1}^{cm}\left(\frac{n-\lfloor n/2\rfloor-\ell-cm+i}{\lfloor n/2\rfloor+\ell+i}\right)\\
&= \prod_{i=1}^{cm}\left(1-\frac{2\ell+cm+\lfloor n/2\rfloor-\lceil n/2\rceil}{\lfloor n/2\rfloor+\ell+i}\right)\\
&\leq \left(1-\frac{2\ell+cm-1}{n}\right)^{cm}\\
&\leq \exp\left(-\frac{cm(cm-1)}{n}\right)\\
&\leq e\cdot\exp(-(cm)^2/n) && (e^{cm/n}\leq e)\\
&\leq e\cdot\exp(-(c^2/4k)) && (k\geq n/(2m)^2).
\end{aligned}$$

Finally, we have $\ell \leq m$ and $(cm+\ell) \leq (c+1)m \leq 2cm$ for $c \geq 1$. Therefore

$$\begin{aligned}
\frac{\sum_{c\geq 0}(cm+\ell)^k\binom{n}{\lfloor n/2\rfloor+cm+\ell}}{\sum_{c\geq 0}\binom{n}{\lfloor n/2\rfloor+cm+\ell}} &\leq \frac{\sum_{c\geq 0}(cm+\ell)^k\binom{n}{\lfloor n/2\rfloor+cm+\ell}}{\binom{n}{\lfloor n/2\rfloor+\ell}}\\
&\leq m^k + e\sum_{c\geq 1}(2cm)^k e^{-\frac{c^2}{4k}}\\
&= m^k + e(2m)^k\sum_{c\geq 1}c^k e^{-\frac{c^2}{4k}}.
\end{aligned}$$

In Claim 21 below we will show that the summation in the last line is bounded by $O(k^k)$. Applying this bound completes the proof. $\square$

Claim 21. $\sum_{c\geq 1} c^k e^{-\frac{c^2}{4k}} \leq O(k^k)$ *for any $k \geq 1$.*

Proof. Consider the function

$$\lambda_a(x) := x^k \cdot e^{-ax^2},$$

for some parameter $a > 0$. Its derivative is $\lambda_a'(x) = (k - 2ax^2)x^{k-1}e^{-ax^2}$; so for $x \geq 0$, $\lambda_a(x)$ is increasing when $x \leq \sqrt{k/(2a)}$, and is decreasing when $x \geq \sqrt{k/(2a)}$. For such a (nonnegative) *first-increasing-then-decreasing* function, it can be seen that

$$\sum_{c\geq 1} \lambda_a(c) \leq \underbrace{\max\{\lambda_a(x) : x \geq 0\}}_{\text{A}} + \underbrace{\int_0^\infty \lambda_a(x)dx}_{\text{B}}.$$

The first term (A) is exactly equal to

$$\text{(A)} = \lambda_a\left(\sqrt{k/(2a)}\right) = \left(\frac{k}{2ea}\right)^{k/2}.$$

Using the substitution $u = ax^2$, the second term (B) is equal to

$$\text{(B)} = \int_0^\infty x^k e^{-ax^2} dx = \frac{1}{2} a^{-\frac{k+1}{2}} \cdot \Gamma\left(\frac{k+1}{2}\right).$$

where $\Gamma(k) := \int_0^\infty x^{k-1}e^{-x}dx$ is the Gamma function. It is known (see, for example, [OLBC10, Equation (5.6.1)]) that $\Gamma(\frac{k+1}{2}) \leq \sqrt{2\pi} \cdot e^{\frac{1}{12}} \cdot (\frac{k}{2e})^{k/2}$ for each integer $k \geq 1$. Plugging this upper bound into the above formula gives

$$\text{(B)} \leq e^{\frac{1}{12}} \cdot \sqrt{\frac{\pi}{2a}} \cdot \left(\frac{k}{2ea}\right)^{k/2} \leq \frac{2}{\sqrt{a}} \cdot \left(\frac{k}{2ea}\right)^{k/2}.$$

By setting $a = 1/(4k)$, we deduce that

$$\text{(A)} + \text{(B)} \leq \left(1 + 4\sqrt{k}\right) \cdot \left(2k^2/e\right)^{k/2} = \frac{1 + 4\sqrt{k}}{(e/2)^{k/2}} \cdot k^k = O(1) \cdot k^k,$$

proving the claim. $\square$

It remains to prove Claim 22, which gives a bound on $\mathcal{K}(n, w, k) := \sum_{x\in\mathcal{X}_k}(-1)^{x_1+\cdots+x_w}$.

Claim 22. $|\mathcal{K}(n, \lfloor n/2 \rfloor + s, k)| \leq 2^{O(k)}((n/k)^{k/2} + (|s|/k)^k)$ *for any* $k \geq 1$.

A similar claim was also made in [CPT20, Claim 4.10], but the bounds we obtain here are more general and sharper for larger $|s|$, which is crucial for our application. We start with a preliminary claim.

Claim 23. *For any positive integer* s,

$$\sum_{i=0}^{\lfloor k/2 \rfloor} \frac{n^i}{i!} \frac{s^{k-2i}}{(k-2i)!} \leq \begin{cases} (2e)^k \left(\frac{n}{k}\right)^{k/2} & \text{if } s \leq \sqrt{nk} \\ (2e)^k \left(\frac{s}{k}\right)^k & \text{if } s \geq \sqrt{nk}. \end{cases}$$

Proof. Write $s = (cn)^{1/2}$ for some $c > 0$. Then $n^i s^{k-2i} = (cn)^{k/2}/c^i$, and

$$\sum_{i=0}^{\lfloor k/2 \rfloor} \frac{n^i}{i!} \frac{s^{k-2i}}{(k-2i)!} = (cn)^{\frac{k}{2}} \sum_{i=0}^{\lfloor k/2 \rfloor} \frac{1}{c^i i!(k-2i)!}. \tag{8}$$

We can write

$$\frac{1}{i!(k-2i)!} = \frac{(2i)!}{k!i!} \frac{k!}{(2i)!(k-2i)!} = \frac{(2i)!}{k!i!} \binom{k}{2i}.$$

So

$$\begin{aligned}
(8) &= \frac{(cn)^{\frac{k}{2}}}{k!} \sum_{i=0}^{\lfloor k/2 \rfloor} \frac{(2i)!}{c^i i!} \binom{k}{2i} \\
&\leq \frac{(cn)^{\frac{k}{2}}}{k!} \max_{0 \leq i \leq \lfloor \frac{k}{2} \rfloor} \left(\frac{2i}{c}\right)^i \cdot \sum_{i=0}^{\lfloor k/2 \rfloor} \binom{k}{2i} \\
&\leq 2^k \cdot \frac{(cn)^{\frac{k}{2}}}{k!} \max_{0 \leq i \leq \lfloor \frac{k}{2} \rfloor} \left(\frac{2i}{c}\right)^i.
\end{aligned}$$

If $c \leq k$, then for $0 \leq i \leq \lfloor k/2 \rfloor$, we have $2i/c \leq k/c$ and thus $(2i/c)^i \leq (k/c)^i \leq (k/c)^{k/2}$ because $k/c \geq 1$. Hence,

$$2^k \cdot \frac{(cn)^{\frac{k}{2}}}{k!} \max_{0 \leq i \leq \lfloor \frac{k}{2} \rfloor} \left(\frac{2i}{c}\right)^i \leq 2^k \frac{(cn)^{\frac{k}{2}}}{k!} \left(\frac{k}{c}\right)^{\frac{k}{2}} \leq (2e)^k \left(\frac{n}{k}\right)^{k/2}.$$

If $c \geq k$, then for $0 \leq i \leq \lfloor k/2 \rfloor$, we have $(2i/c)^i \leq 1$, and thus

$$2^k \cdot \frac{(cn)^{\frac{k}{2}}}{k!} \max_{0 \leq i \leq \lfloor \frac{k}{2} \rfloor} \left(\frac{2i}{c}\right)^i \leq (2e)^k \left(\frac{cn}{k^2}\right)^{\frac{k}{2}} = (2e)^k \left(\frac{s}{k}\right)^k.$$

□

Proof of Claim 22. Observe that

$$(1-x)^w (1+x)^{n-w} = \sum_{k=0}^{n} \mathcal{K}(n, w, k) x^k,$$

Our goal is to bound above its k-th coefficient. We will assume n is even, otherwise we can apply the bound for $n+1$, which affects its magnitude by at most a factor of 2. By the symmetry of $\mathcal{K}(n, u, k)$ around $u = n/2$, for Claim 22 we may take $w = n/2 - s$, and for this choice of w we have

$$\sum_{k=0}^{n} \mathcal{K}(n, n/2 - s, k) x^k = (1-x)^{\frac{n}{2}-s} (1+x)^{\frac{n}{2}+s} = (1-x^2)^{\frac{n}{2}} \left(\frac{1+x}{1-x}\right)^s. \tag{9}$$

We bound from above the coefficients of the power series expansion of the two terms on the right hand side. The i-th coefficient of $(1-x^2)^{n/2}$ is $\binom{n/2}{i/2}$ if i is even and is 0 otherwise (because the function is even). For $(\frac{1+x}{1-x})^s$, we will assume that s is nonnegative as this only affects the sign of its coefficients. Since $\frac{1+x}{1-x} = 1 + \sum_{j \geq 1} 2x^j$

and $\frac{1}{1-2x} = 1 + \sum_{j\ge 1}(2x)^j$ for $|x| < 1/2$ in their power series expansion, the j-th coefficient of $(\frac{1+x}{1-x})^s$ is at most the j-th coefficient of $(1-2x)^{-s}$, which is

$$2^j \cdot \frac{s\cdot(s+1)\cdots(s+j-1)}{1\cdot 2\cdots j} = \binom{s+j-1}{j}2^j.$$

Therefore, the magnitude of the k-th coefficient of Equation (9) is at most

$$\sum_{i=0}^{\lfloor k/2\rfloor}\binom{n/2}{i}\binom{|s|+k-2i-1}{k-2i}2^{k-2i}. \tag{10}$$

Using $\binom{n}{k}\le n^k/k!$, we have

$$(10) \le \sum_{i=0}^{\lfloor k/2\rfloor}\frac{n^i}{i!}\frac{(|s|+k-2i-1)^{k-2i}}{(k-2i)!}2^{k-2i}.$$

Note that $(s+t)^j \le 2^j\max\{s^j,t^j\}$ and so

$$\frac{(|s|+k-2i-1)^{k-2i}}{(k-2i)!} \le 2^{k-2i}\left(\frac{\max\{|s|^{k-2i},(k-2i-1)^{k-2i}\}}{(k-2i)!}\right) \le 2^{k-2i}\left(\frac{\max\{|s|^{k-2i},k^{k-2i}\}}{(k-2i)!}\right).$$

Therefore, (10) is bounded by

$$2^{2k}\left(\sum_{i=0}^{\lfloor k/2\rfloor}\frac{n^i}{i!}\frac{\max\{|s|^{k-2i},k^{k-2i}\}}{(k-2i)!}\right).$$

The claim then follows from Claim 23. $\square$

4.3 Proof of Theorem 9

In this subsection, we give examples of symmetric polynomials p with $L_{1,k}(p) = \Omega_k(1)\cdot d^k$ for $d = \Theta(\sqrt{n})$, matching our upper bound up to a constant factor when $k = O(1)$. We restate our theorem for convenience.

Theorem 24 (Restatement of Theorem 9)**.** *For every k, there is a symmetric $\mathbb{F}_2$-polynomial $p(x_1,\dots,x_n)$ of degree $d = \Theta(\sqrt{kn})$ such that $L_{1,k}(p) \ge (e^{-k}/2)\cdot\binom{n}{k}^{1/2} = \Omega_k(1)\cdot d^k$.*

Proof. Let $m := C\sqrt{kn}$ for some constant $4 \le C \le 8$ such that m is a power of two. Consider the function $h_k\colon \{0,1\}^n \to \{-1,1\}$ defined by

$$h_k(x) = \operatorname{sgn}\left(\sum_{|S|=k}(-1)^{\sum_{i\in S}x_i}\right).$$

Let $\widetilde{h_k}\colon \{0,\dots,n\} \to \{-1,1\}$ be the symmetrization of h_k so that $\widetilde{h_k}(|x|) := h_k(x)$. Now consider the periodic function $\widetilde{p}\colon \{0,\dots,n\} \to \{-1,1\}$ that on input $z \in \{0,\dots,n\}$ such that

$$z + 2tm \in [n/2-m, n/2+m) \text{ for some } t \in \mathbb{Z},$$

evaluates to $\widetilde{p}(z) := \widetilde{h}_k(z + 2tm)$. We define the $\mathbb{F}_2$-polynomial p by $p(x_1, \ldots, x_n) := \tilde{p}(|x|)$. Since $p(x)$ depends only on $|x| \bmod 2m$, by Lemma 17, it can be computed by a symmetric $\mathbb{F}_2$-polynomial of degree $d < 2m$. Moreover, $p(x)$ agrees with $h_k(x)$ when $|x| \in [n/2 - m, n/2 + m)$. Let $f = (-1)^p$. We have (note that from here on we switch from $x \in \{0,1\}^n$ to $x \in \{-1,1\}^n$)

$$
\begin{aligned}
L_{1,k}(f) \geq \left| \sum_{|S|=k} \widehat{f}(S) \right| &= \left| \sum_{|S|=k} \mathop{\mathbf{E}}_x \Big[f(x) x^S \Big] \right| \\
&= \left| \sum_{|S|=k} \mathop{\mathbf{E}}_x \Big[h_k(x) x^S \Big] + \mathop{\mathbf{E}}_x \Big[\big(f(x) - h_k(x) \big) x^S \Big] \right| \\
&= \left| \mathop{\mathbf{E}}_x \left[h_k(x) \sum_{|S|=k} x^S \right] + \mathop{\mathbf{E}}_x \left[\big(f(x) - h_k(x) \big) \sum_{|S|=k} x^S \right] \right| \\
&\geq \left| \left| \mathop{\mathbf{E}}_x \left[h_k(x) \sum_{|S|=k} x^S \right] \right| - \left| \mathop{\mathbf{E}}_x \left[\big(f(x) - h_k(x) \big) \sum_{|S|=k} x^S \right] \right| \right| .
\end{aligned}
$$

We proceed to lower bound the first term and upper bound the second term. For the first term we use the same lower bound of $e^{-k}\binom{n}{k}^{1/2}$ in the proof of [Vio20, Theorem 1]; for completeness we include the argument here. One can verify that $\mathbf{E}_x[(\sum_{|S|=k} x^S)^2] = \binom{n}{k}$. Applying [O'D14, Theorem 9.22], which states that $\mathbf{E}[|h(x)|] \geq e^{-k}\, \mathbf{E}[h(x)^2]^{1/2}$ for any $h\colon \{-1,1\}^n \to \mathbb{R}$ of degree k, we get

$$
e^{-k}\binom{n}{k}^{1/2} = e^{-k} \mathop{\mathbf{E}}_x \left[\left(\sum_{|S|=k} x^S \right)^2 \right]^{1/2} \leq \mathop{\mathbf{E}}_x \left[\left| \sum_{|S|=k} x^S \right| \right] \leq \mathop{\mathbf{E}}_x \left[\left(\sum_{|S|=k} x^S \right)^2 \right]^{1/2} = \binom{n}{k}^{1/2} .
$$

So for the first term we have

$$
\mathop{\mathbf{E}}_x \left[h_k(x) \sum_{|S|=k} x^S \right] = \mathop{\mathbf{E}}_x \left[\left| \sum_{|S|=k} x^S \right| \right] \geq e^{-k} \binom{n}{k}^{1/2} .
$$

We now bound above the second term. As $|f(x) - h_k(x)| \leq 2$ for every x, by the Cauchy–Schwarz inequality, we get

$$
\begin{aligned}
\left| \mathop{\mathbf{E}}_x \left[\big(f(x) - h_k(x) \big) \sum_{|S|=k} x^S \right] \right| &\leq 2 \cdot \mathop{\mathbf{Pr}}_x \left[2 \Big| \sum_i x_i \Big| \geq m \right] \mathop{\mathbf{E}}_x \left[\left(\sum_{|S|=k} x^S \right)^2 \right]^{1/2} \\
&= 2 \cdot \mathop{\mathbf{Pr}}_x \left[2 \Big| \sum_i x_i \Big| \geq m \right] \binom{n}{k}^{1/2} \\
&\leq (e^{-k}/2) \cdot \binom{n}{k}^{1/2} ,
\end{aligned}
$$

where the last inequality follows from the Chernoff bound recalling our choice of $m = \Theta(\sqrt{kn})$. Therefore $L_{1,k}(f) \geq (e^{-k}/2) \cdot \binom{n}{k}^{1/2}$, proving the theorem. □

5 $L_{1,k}$ bounds for read-Δ polynomials

In this section, we give our $L_{1,k}$ bounds on read-few polynomials, which are restated below for convenience.

Theorem 25 (Restatement of Theorem 10)**.** *Let $p(x_1, \dots, x_n)$ be any read-Δ degree-d polynomial. For every $1 \le k \le n$,*

$$L_{1,k}(p) \le \mathbf{Pr}[p = 1] \cdot O(k)^k \cdot (\Delta d)^{8k}.$$

By observing that any degree-$\Omega(\log n)$ monomial vanishes under a random restriction with high probability, we can use Theorem 25 to obtain an upper bound for read-Δ polynomials that is independent of d:

Corollary 26 (Restatement of Corollary 11)**.** *Let $p(x_1, \dots, x_n)$ be a read-Δ polynomial. For any $1 \le k \le n$,*

$$L_{1,k}(p) \le O(k)^{9k} \cdot (\Delta \log n)^{8k}.$$

Proof. Let R be a ρ-random restriction that independently for each x_i with probability $1-\rho$ sets x_i to a uniform random bit and keeps x_i alive with the remaining ρ probability. For any $f\colon \{-1,1\}^n \to \{-1,1\}$, we have $\mathbf{E}_R[\widehat{f_R}(S)] = \rho^{|S|}\widehat{f}(S)$. Thus,

$$L_{1,k}(f) = \sum_{|S|=k} |\widehat{f}(S)| = \sum_{|S|=k} |\rho^{-k} \mathop{\mathbf{E}}_R[\widehat{f_R}(S)]| \le \rho^{-k} \mathop{\mathbf{E}}_R[L_{1,k}(f_R)].$$

Let $p(x_1, \dots, x_n)$ be any read-Δ polynomial. Note that $L_{1,k}(f) \le n^{k/2}$ for any f. Hence, by Theorem 25, for any $d_{\max}$ and any ρ, we have

$$L_{1,k}(p) \le \rho^{-k} \cdot O(k)^k(\Delta d_{\max})^{8k} + \rho^{-k} \cdot \mathop{\mathbf{Pr}}_R[\deg(p_R) \ge d_{\max}] \cdot n^{k/2}.$$

We now upper bound $\mathbf{Pr}_R[\deg(p_R) \ge d_{\max}]$. Each monomial with degree greater than $d_{\max}$ survives R with probability at most $\rho^{d_{\max}}$, and there are at most $(\Delta n)/d_{\max}$ such monomials in p. It follows by a union bound that $\mathbf{Pr}_R[\deg(p_R) \ge d_{\max}] \le \rho^{d_{\max}}(\Delta n)/d_{\max}$. So setting $\rho = 1/2$ and $d_{\max} = 2k \log n$, we have

$$\begin{aligned} L_{1,k}(p) &\le O(k)^k(\Delta \cdot d_{\max})^{8k} + \frac{\rho^{d_{\max}}\Delta \cdot n^{1+k/2}}{\rho^k d_{\max}} \\ &\le O(k)^{9k} \cdot (\Delta \log n)^{8k}. \end{aligned}$$

$\square$

Proof idea. We first observe that for $f = (-1)^p$, the Fourier coefficient $\widehat{f}(S)$ is simply the bias of the $\mathbb{F}_2$-polynomial $p_S(x) := p(x) + \sum_{i \in S} x_i$. Assuming that p_S depends on all n variables, by a simple greedy argument we can collect $n/\mathrm{poly}(\Delta, d)$ polynomials in p_S so that each of them depends on disjoint variables, and it is not hard to show that the product of the biases of these polynomials upper bounds the bias of p_S. From this is easy to see that any read-Δ degree-d polynomial has bias $\exp(2^{-d}n/\mathrm{poly}(\Delta, d))$. However, this quantity is too large to sum over $\binom{n}{k}$ coefficients.

Our next idea (Lemma 27) is to give a more refined decomposition of the polynomial p by inspecting the variables $x_i : i \in S$ more closely. Suppose the variables $x_i : i \in S$ are far apart in their dependency graph (see the definition of G_p below), as must indeed be the case for most of the $\binom{n}{k}$ size-k sets S. Then we can collect all the monomials containing each x_i to form a polynomial p_i, and these p_i's will depend on disjoint variables. Moreover, if every monomial in p_i has high degree (see the definition of $V_t(p)$ below), then $p_i = 0$ with high probability and therefore $p_i + x_i$ is almost unbiased. Therefore, we can first collect these p_i and x_i from p_S; then, for the remaining $m \geq |S| \cdot \text{poly}(\Delta, d)$ monomials in p_S, as before we collect $m/\text{poly}(\Delta, d)$ polynomials r_i so that they depend on disjoint variables, but this time we collect these monomials using the variables in $V_t(p)$, and give an upper bound in terms of the size $|V_t(p)|$. Multiplying the biases of the $p_i + x_i$'s and the bias of r gives our refined upper bound on $\widehat{f}(S)$ in Lemma 27.

We now proceed to the actual proof. We first define some notions that will be used throughout our arguments. For a read-Δ degree-d polynomial p, we define $V_t(p) : t \in [d]$ and G_p as follows.

For every $t \in [d]$, define

$$V_t(p) := \{i \in [n] : \text{the minimum degree of the monomials in } p \text{ containing } x_i \text{ is } t\}.$$

Note that the sets $V_1(p), \ldots, V_d(p)$ form a partition of the input variables p depends on.

Define the undirected graph G_p on $[n]$, where $i, j \in [n]$ are adjacent if x_i and x_j both appear in the same monomial in q. Note that G_p has degree at most Δd. For $S \subseteq [n]$, we use $N_{=d}(S)$ to denote the indices that are at distance exactly i to S in G_p, and use $N_{\leq d}(S)$ to denote $\bigcup_{j=0}^{d} N_{=j}(S)$.

We first state our key lemma, which gives a refined bound on each $\widehat{f}(S)$ stronger than the naive bound sketched in the first paragraph of the "Proof Idea" above, and use it to prove Theorem 25. We defer its proof to the next section.

Lemma 27 (Main lemma for read-Δ polynomials)**.** *Let $p(x_1, \ldots, x_n)$ be a read-Δ degree-d polynomial. Let $S \subseteq [n], |S| \geq \ell$ be a subset containing some ℓ indices $i_1, \ldots, i_\ell \in S$ whose pairwise distances in G_p are at least 4, and let $t_1, \ldots, t_\ell \in [d]$ be such that each $i_j \in V_{t_j}(p)$. Let $f = (-1)^p$. Then*

$$|\widehat{f}(S)| \leq O(1)^{|S|} \cdot \Delta^\ell \prod_{j \in [\ell]} \left(2^{-t_j} \exp\left(-\frac{2^{-t_j} |V_{t_j}(p)|}{\ell \cdot (\Delta d)^4} \right) \right).$$

Proof of Theorem 25. Using a reduction given in the proof of [CMM+20, Lemma 2.2], it suffices to prove the same bound without the acceptance probability factor, i.e. to prove that for every $1 \leq k \leq n$,

$$L_{1,k}(p) \leq O(k)^k \cdot (\Delta d)^{8k}.$$

As [CMM+20] did not provide an explicit statement of the reduction, for completeness we provide a self-contained statement and proof in Lemma 36 in Appendix A.

For every subset $S \subseteq [n]$ of size k, there exists an $\ell \leq k$ and $i_1, \ldots, i_\ell \in S$ such that their pairwise distances in G_p are at least 4, each $i_j \in V_{t_j}(p)$ for some $t_j \in [d]$, and each of the remaining $k - \ell$ indices in S is within distance at most 3 to some i_j.

Fix any $i_1, \ldots, i_\ell$, and let us bound the number of subsets $S \subseteq [n]$ of size k that can contain $i_1, \ldots, i_\ell$. Because $|N_{\leq 3}(j)| \leq (\Delta d)^3 + (\Delta d)^2 + \Delta d + 1 \leq 4(\Delta d)^3$ for every

$j \in [n]$, the remaining $k-\ell$ indices of S can appear in at most

$$\begin{aligned}\sum_{j_1+\cdots+j_\ell=k-\ell}\prod_{b\in[\ell]}\binom{4(\Delta d)^3}{j_b} &= \binom{4\ell(\Delta d)^3}{k-\ell}\\ &\le (4(\Delta d)^3)^k\cdot e^{k-\ell}\left(\frac{\ell}{k-\ell}\right)^{k-\ell}\\ &\le (e\Delta d)^{3k}\end{aligned}$$

different ways, where the equality uses the Vandermonde identity, the first inequality uses $\binom{n}{k} \le (en/k)^k$, and the last one uses $(\frac{\ell}{k-\ell})^{k-\ell} \le (1+\frac{\ell}{k-\ell})^{k-\ell} \le e^\ell$ and $4e < e^3$. Therefore, by Lemma 27,

$$\begin{aligned}\sum_{S:|S|=k}|\widehat{f}(S)| &\le \sum_{\ell=1}^{k}\sum_{t\subseteq[d]^\ell}\left[\left(\prod_{j\in[\ell]}|V_{t_j}(p)|\right)\cdot(e\Delta d)^{3k}\cdot O(1)^k\Delta^\ell\prod_{j'\in[\ell]}\left(2^{-t_{j'}}\exp\left(-\frac{2^{-t_{j'}}|V_{t_{j'}}(p)|}{\ell(\Delta d)^4}\right)\right)\right]\\ &\le O(1)^k\cdot(\Delta d)^{3k}\sum_{\ell=1}^{k}\Delta^\ell\sum_{t\subseteq[d]^\ell}\prod_{j\in[\ell]}\left(2^{-t_j}|V_{t_j}(p)|\exp\left(-\frac{2^{-t_j}|V_{t_j}(p)|}{\ell(\Delta d)^4}\right)\right)\\ &\le O(1)^k\cdot(\Delta d)^{3k}\sum_{\ell=1}^{k}\Delta^\ell\cdot d^\ell\cdot(\ell(\Delta d)^4)^\ell\\ &\le O(k)^k\cdot(\Delta d)^{3k}\cdot(\Delta d)^{5k}\\ &= O(k)^k\cdot(\Delta d)^{8k},\end{aligned}$$

where the third inequality is because the function $x \mapsto xe^{-x/c}$ is maximized when $x = c$. This completes the proof. □

5.1 Proof of Lemma 27

First we need a simple lemma that lower bounds the acceptance probability of a sparse polynomial.

Lemma 28. *Let $r(x_1,\ldots,x_n)$ be a polynomial with at most Δ monomials, each of which has degree at least t, and $r(0)=0$. Then $\mathbf{Pr}[r(x)=0] \ge \max\{\frac{1}{\Delta+1}, 1-2^{-t}\Delta\}$.*

Proof. The lower bound of $1/(\Delta+1)$ follows from [KL93, Corollary 1]. For the other bound, note that each monomial evaluates to 1 with probability at most 2^{-t}; so by a union bound $\mathbf{Pr}[r(x)=1] \le 2^{-t}\Delta$. □

We use G_p to partition $[n]$ as follows. For each $j \in [\ell]$, let $T_j := N_{=1}(i_j)$, the set of distance-1 neighbors of i_j. Let $R := [n] \setminus \bigcup_j(\{i_j\} \cup T_j)$ be the remaining variables (note that these are the variables whose distance from every i_j is at least two). As the i_j's are at distance at least 4 apart, then for $j \neq j'$ each index in $\{i_j\} \cup T_j$ has distance at least 2 to each index in $\{i_{j'}\} \cup T_{j'}$; so the variables in $\{i_j\} \cup T_j$ appear in disjoint monomials from the variables in $\{i_{j'}\} \cup T_{j'}$. We now partition the monomials in p. For each $j \in [\ell]$, let p_j, q_j be two polynomials such that p_j is composed of all the monomials in p that contain x_{i_j}, and q_j is composed of all monomials that contain some variable in x_{T_j} but not x_{i_j}. Let r be the polynomial composed of the remaining monomials in p. We can write p as

$$\sum_{j=1}^{\ell}\big(p_j(x_{i_j}, x_{T_j}) + q_j(x_{T_j}, x_R)\big) + r(x_R).$$

Now, consider the polynomial

$$\begin{aligned} p_S(x_1,\ldots,x_n) &:= p(x_1,\ldots,x_n) + \sum_{i\in S} x_i \\ &= \sum_{j=1}^{\ell} s_j(x_{i_j}, x_{T_j}, x_R) + r'(x_R), \end{aligned} \tag{11}$$

where

$$s_j(x_{i_j}, x_{T_j}, x_R) := p_j(x_{i_j}, x_{T_j}) + q_j(x_{T_j}, x_R) + x_{i_j} + \sum_{k\in S\cap T_j} x_k \tag{12}$$

$$r'(x_R) := r(x_R) + \sum_{k\in S\cap R} x_k.$$

The following claim gives an upper bound on the bias of s_j; we defer its proof till later. (To interpret the claim it may be helpful to recall that $i_j \in V_{t_j}(p)$, and hence the minimum degree of the monomials in p containing x_{i_j} is t_j.)

Claim 29. *For each possible outcome of x_R, we have $|\mathbf{E}_{x_{i_j}, x_{T_j}}[(-1)^{s_j(x_{i_j}, x_{T_j}, x_R)}]| \le \Delta 2^{-(t_j-1)}$ for every $j \in [\ell]$.*

We have that

$$\begin{aligned} |\widehat{f}(S)| &= \left|\mathbf{E}\Big[(-1)^{p_S(x)}\Big]\right| \\ &= \left|\mathop{\mathbf{E}}_{x_R}\left[(-1)^{r'(x_R)} \cdot \prod_{j\in[\ell]} \mathop{\mathbf{E}}_{x_{T_j}, x_{i_j}}\Big[(-1)^{s_j(x_{i_j}, x_{T_j}, x_R)}\Big]\right]\right| && (13) \\ &\le (2\Delta)^{\ell} \cdot \left(\prod_{j\in[\ell]} 2^{-t_j}\right) \cdot \left|\mathop{\mathbf{E}}_{x_R}\Big[(-1)^{r'(x_R)}\Big]\right|, && (14) \end{aligned}$$

where Equation (13) is by independence of the different (x_{T_j}, x_{i_j})'s and Equation (14) is by Claim 29. We now claim that

$$|V_t(r')| \ge |V_t(p)| - \ell(\Delta d)^2 - |S| \quad \text{for every } t \in [d]. \tag{15}$$

This is because a variable x_j belongs to $V_t(p) \setminus V_t(r')$ only if it belongs to or is adjacent to the at most $\ell\Delta d$ restricted variables $\{x_{i_j}, x_{T_j} : j \in [\ell]\}$ in G_p, or it is one of the $\{x_i : i \in S\}$. In Lemma 30 below we prove that for every fixed assignment to $x_{i_j}, x_{T_j} : j \in [\ell]$, and any $t \in [d]$, we have that

$$\left|\mathop{\mathbf{E}}_{x_R}\Big[(-1)^{r'(x_R)}\Big]\right| \le \exp\left(-\frac{2^{-t}|V_t(r')|}{(\Delta d)^4}\right). \tag{16}$$

Continuing from above, combining Equations (14), (15) and (16), we have

$$
\begin{aligned}
|\widehat{f}(S)| &\le (2\Delta)^{\ell}\left(\prod_{j\in[\ell]} 2^{-t_j}\right)\cdot \min\left\{\exp\left(-\frac{2^{-t}(|V_t(p)|-\ell(\Delta d)^2-|S|)}{(\Delta d)^4}\right) : t\in[d]\right\}\\
&\le O(1)^{|S|}\cdot(2\Delta)^{\ell}\left(\prod_{j\in[\ell]} 2^{-t_j}\right)\cdot \min\left\{\exp\left(-\frac{2^{-t}(|V_t(p)|)}{(\Delta d)^4}\right) : t\in[d]\right\}\\
&= O(1)^{|S|}\cdot(2\Delta)^{\ell}\left(\prod_{j\in[\ell]} 2^{-t_j}\min\left\{\exp\left(-\frac{2^{-t}(|V_t(p)|)}{\ell\cdot(\Delta d)^4}\right) : t\in[d]\right\}\right)\\
&= O(1)^{|S|}\cdot(2\Delta)^{\ell}\prod_{j\in[\ell]}\left(2^{-t_j}\exp\left(-\frac{2^{-t_j}|V_{t_j}(p)|}{\ell\cdot(\Delta d)^4}\right)\right),
\end{aligned}
$$

where second inequality is because $\exp\left(\frac{\ell(\Delta d)^2+|S|}{(\Delta d)^4}\right)\le e^{\ell+|S|}\le O(1)^{|S|}$ because $\ell\le|S|$. This proves the lemma. $\square$

We now prove Claim 29 and Lemma 30.

Proof of Claim 29. We may assume $t_j\ge 2$ as otherwise the conclusion is trivial. Since $i_j\in V_{t_j}(p)$, recalling Equations (11) and (12), every monomial in p_j containing x_{i_j} has degree at least t_j, and by collecting these monomials, we can write

$$
\begin{aligned}
s_j(x_{i_j},x_{T_j},x_R) &= x_{i_j}+p_j(x_{i_j},x_{T_j})+\sum_{k\in S\cap T_j}x_k+q_j(x_{T_j},x_R)\\
&= x_{i_j}(1+u_j(x_{T_j}))+v_j(x_{T_j},x_R)
\end{aligned}
$$

for some polynomials u_j and v_j, where every monomial in $u_j(x_{T_j})$ has degree at least $t_j-1\ge 1$, and thus $u_j(0)=0$. Now, if an outcome of x_{T_j} is such that $u_j(x_{T_j})=0$, then the expectation of $(-1)^{s_j}$ is zero because v_j does not depend on x_{i_j} and $\mathbf{E}[(-1)^{x_{i_j}}]=0$. Hence we have that for every outcome of x_R,

$$
\left|\mathop{\mathbf{E}}_{x_{i_j},x_{T_j}}\left[(-1)^{s_j(x_{i_j},x_{T_j},x_R)}\right]\right|\le \mathop{\mathbf{E}}_{x_{T_j}}\left|\mathop{\mathbf{E}}_{x_{i_j}}\left[(-1)^{s_j(x_{i_j},x_{T_j},x_R)}\right]\right|\le \mathbf{Pr}[u_j(x_{T_j})=1]\le\Delta 2^{-(t_j-1)},
$$

$\square$

where the final inequality is by Lemma 28.

Lemma 30. *Let $q(x_1,\dots,x_n)$ be a read-Δ degree-d polynomial. For every $t\in[d]$ we have*

$$
\left|\mathbf{E}\left[(-1)^q\right]\right|\le\exp\left(-\frac{2^{-t}|V_t(q)|}{(\Delta d)^4}\right).
$$

Proof. We partition $V_t(q)$ according to G_q using the following greedy procedure:

1. Set $W_1=[n]$ and $i=1$.
2. Pick a monomial x^{S_i} of degree exactly t containing some variable $x_j : j\in W_i\cap V_t(q)$.
3. Let $T_i:=N_{=1}(S_i)$, and set $W_{i+1}=W_i\setminus N_{\le 3}(S_i)$.
4. Repeat steps 2 and 3 until we cannot pick such an $S_{\ell+1}$. Let $R=[n]\setminus\bigcup_{i\in[\ell]}(S_i\cup T_i)$ be the remaining variables.

By construction, the pairwise distances between the $(S_i \cup T_i)$'s are at least 2, so the variables $x_{S_i \cup T_i}$'s must only appear in disjoint monomials. Since each time we remove at most $t \cdot (\Delta d)^3$ elements from $V_t(q)$, we have $\ell \geq |V_t(q)|/(t(\Delta d)^3)$.

We now partition the monomials in q. For each $i \in [\ell]$, let p_i, q_i be two polynomials such that p_i is composed of all monomials in q that contain a variable in x_{S_i}, and q_i is composed of all monomials in q that contain a variable in x_{T_i} but none in x_{S_i}. Let r be the polynomial composed of the remaining monomials in q. We can write $q(x)$ as

$$\sum_{i=1}^{\ell} \big(p_i(x_{S_i}, x_{T_i}) + q_i(x_{T_i}, x_R)\big) + r(x_R).$$

By collecting the monomials containing x^{S_i}, we can write

$$p_i(x_{S_i}, x_{T_i}) + q_i(x_{T_i}, x_R) = x^{S_i}(1 + u_i(x_{T_i})) + v_i(x_{S_i}, x_{T_i}, x_R).$$

for some polynomials u_i and v_i, where (1) u_i consists of the at most $\Delta - 1$ monomials in p_i that contain x^{S_i}, and $u_i(0) = 0$, and (2) x^{S_i} does not appear in any monomial in v_i. Let $\delta := \mathbf{Pr}[u_i(x_{T_i}) = 0]$. By Lemma 28 we have $\delta \geq 1/\Delta$. Therefore, for every assignment to x_R and every $i \in [\ell]$,

$$\begin{aligned}
&\mathop{\mathbf{E}}_{x_{T_i}}\left|\mathop{\mathbf{E}}_{x_{S_i}}\left[(-1)^{p_i(x_{S_i}, x_{T_i}) + q_i(x_{T_i}, x_R)}\right]\right| \\
&= (1-\delta)\left|\mathop{\mathbf{E}}_{x_{S_i}}\left[(-1)^{v_i(x_{S_i}, x_{T_i}, x_R)}\right]\right| + \delta \mathop{\mathbf{E}}_{x_{T_i}: u_i(x_{T_i})=0}\left[\left|\mathop{\mathbf{E}}_{x_{S_i}}\left[(-1)^{x^{S_i} + v_i(x_{S_i}, x_{T_i}, x_R)}\right]\right|\right] \\
&\leq (1-\delta) + \delta(1 - 2^{-t}) \\
&\leq 1 - 2^{-t}/\Delta,
\end{aligned}$$

where the first inequality is because for any choice of x_R and x_{T_i}, the polynomial $x^{S_i} + v_i(x_{S_i}, x_{T_i}, x_R)$ has degree exactly t, and therefore its bias is at most $1 - 2^{-t}$. Also, for every fixed outcome of x_R, the restricted polynomials $p_i(x_{S_i}, x_{T_i}) + q_i(x_{T_i}, x_R)$ depend on disjoint variables. Therefore,

$$\left|\mathop{\mathbf{E}}_{x_{T_i}, x_{S_i}: i \in [\ell]}\left[(-1)^{q(x)}\right]\right| \leq \left(1 - \frac{2^{-t}}{\Delta}\right)^{\ell} \leq \exp\left(-\frac{\ell \cdot 2^{-t}}{\Delta}\right) \leq \exp\left(-\frac{2^{-t}|V_t(q)|}{(\Delta d)^4}\right),$$

because $\ell \geq |V_t(q)|/(t\Delta^3 d^3) \geq |V_t(q)|/(\Delta^3 d^4)$. $\square$

6 $L_{1,k}$ bounds for disjoint compositions

In this section we give $L_{1,k}$ bounds on *disjoint compositions* of functions, which we define below.

Let $\mathcal{F}$ and $\mathcal{G}$ be two families consisting of functions mapping $\{-1,1\}^m$ and $\{-1,1\}^\ell$ to $\{-1,1\}$ respectively. Let $f \in \mathcal{F}$ and $g_1, \ldots, g_m \in \mathcal{G}$. We define $h\colon \{-1,1\}^{m\ell} \to \{-1,1\}$, the *disjoint composition* of f and $g_1, \ldots, g_m$, to be

$$h(x_{1,1}, \ldots, x_{1,\ell}, \ldots, x_{m,1}, \ldots, x_{m,\ell}) := f(g_1(x_{1,1}, \ldots, x_{1,\ell}), \ldots, g_m(x_{m,1}, \ldots, x_{m,\ell})).$$

Theorem 31 (Sharper version of Theorem 13)**.** *Let $g_1, \ldots, g_m \in \mathcal{G}$, and $f \in \mathcal{F}$, where $\mathcal{G}, \mathcal{F}$ are as above and $\mathcal{F}$ is closed under restrictions. Suppose for every $1 \leq k \leq K$,*

1. $L_{1,k}(f) \le \frac{1-|\mathbf{E}[f]|}{2} \cdot a_{\mathsf{out}} \cdot b_{\mathsf{out}}^k$ *for every* $f \in \mathcal{F}$, *and*
2. $L_{1,k}(g) \le \frac{1-|\mathbf{E}[g]|}{2} \cdot a_{\mathsf{in}} \cdot b_{\mathsf{in}}^k$ *for every* $g \in \mathcal{G}$.

Then

$$L_{1,K}(h) \le \frac{1-|\mathbf{E}[h]|}{2} \cdot a_{\mathsf{out}} \cdot b_{\mathsf{in}}^K \cdot \frac{a_{\mathsf{in}} b_{\mathsf{out}}}{2} \left(1 + \frac{a_{\mathsf{in}} b_{\mathsf{out}}}{2}\right)^{K-1}.$$

In particular, when $a_{\mathsf{in}} b_{\mathsf{out}} \ge 2$ *or* $K = 1$ *we have* $L_{1,K}[h] \le \frac{1-|\mathbf{E}[h]|}{2} \cdot a_{\mathsf{out}} \cdot \frac{(a_{\mathsf{in}} b_{\mathsf{in}} b_{\mathsf{out}})^K}{2}$.

Proof idea. Before proving Theorem 31, we briefly describe the main ideas of the proof. For a subset $J \subseteq [m]$, let $\partial_J f$ denote the *J-th derivative of f*, which can be expressed as

$$\partial_J f(x_1, \dots, x_m) := \sum_{T \supseteq J} \widehat{f}(T) x^{T \setminus J}.$$

Note that $\widehat{f}(J) = \partial_J f(\vec{0})$.

Let us begin by considering the task of bounding $L_{1,1}(h) = \sum_{(i,j)\in[m]\times[\ell]} |\widehat{h}\{(i,j)\}|$. Let $\beta = (\beta_1, \dots, \beta_m)$, where $\beta_i := \mathbf{E}[g_i]$. Using the Fourier expansion of f, we have

$$\widehat{h}\{(i,j)\} = \sum_{S \subseteq [m]} \widehat{f}(S)\, \mathbf{E}\left[\prod_{k \in S} g_k(x_k) \cdot x_{i,j}\right].$$

If $S \not\ni i$, then the expectation is zero, because $\prod_{k\in S} g_k(x_k)$ and $x_{i,j}$ are independent and $\mathbf{E}[x_{i,j}] = 0$. So, we have

$$\widehat{h}\{(i,j)\} = \sum_{S \ni i} \widehat{f}(S) \beta^{S \setminus \{i\}} \cdot \widehat{g}_i(\{j\}) = \partial_i f(\beta) \cdot \widehat{g}_i(\{j\}).$$

If the functions g_i are balanced, i.e. $\mathbf{E}[g_i] = 0$ for all i, then we would have $\beta = \vec{0}$, and

$$\widehat{h}\{(i,j)\} = \partial_i f(\vec{0}) \cdot \widehat{g}_i(\{j\}) = \widehat{f}(\{i\}) \widehat{g}_i(\{j\}).$$

So in this case we have

$$L_{1,1}(h) = \sum_{i\in[m], j\in[\ell]} \left|\widehat{h}(\{(i,j)\})\right| = \sum_{i\in[m]} \sum_{j\in[\ell]} |\widehat{f}(\{i\}) \widehat{g}_i(\{j\})| = \sum_{i\in[m]} |\widehat{f}(\{i\})| \sum_{j\in[\ell]} |\widehat{g}_i(\{j\})|$$

and we can apply our bounds on $L_{1,1}(\mathcal{F})$ and $L_{1,1}(\mathcal{G})$ to $\sum_{i\in[m]} \widehat{f}\{i\}$ and $\sum_{j\in[\ell]} \widehat{g}_i\{j\}$ respectively. Specializing to the case $g_1 = \cdots = g_m$, we have

Claim 32. *Suppose* $g_1 = g_2 = \cdots = g_m =: g$ *and* $\mathbf{E}[g] = 0$. *Then* $L_{1,1}(h) = L_{1,1}(f) L_{1,1}(g)$.

In general the g_i's may not all be the same and may not be balanced, and so it seems unclear how we can apply our $L_{1,1}(\mathcal{F})$ bound on $\sum_{i\in[m]} \partial_i f(\beta_1, \dots, \beta_m)$ when $\beta \ne \vec{0}$. To deal with this, in Claim 33 below we apply a clever idea introduced in [CHHL19] that lets us relate $f(\beta)$ at a nonzero point β to the average of $f_{R_\beta}(\vec{0})$, where f_{R_β} is f with some of its inputs fixed by a random restriction R_β. As $\mathcal{F}$ is closed under restrictions, we have that $f_{R_\beta} \in \mathcal{F}$ and we can apply the $L_{1,1}(\mathcal{F})$ bound on $\sum_i \partial_i f_{R_\beta}(\vec{0})$, which in turn gives a bound on $\sum_{i\in[m]} \partial_i f(\beta_1, \dots, \beta_m)$.

Bounding $L_{1,K}(h)$ for $K \ge 2$ is more complicated, as now each $\widehat{h}(S)$ involves many $\widehat{f}(J)$ and $\widehat{g}_i(T)$'s, where the sets J and T have different sizes. So one has to group the coefficients carefully.

6.1 Useful notation

For a set $S \subseteq [m] \times [\ell]$, let $S|_f := \{i \in [m] : (i,j) \in S \text{ for some } j \in [\ell]\}$ be the "set of first coordinates" that occur in S, and let $S|_i := \{j \in [\ell] : (i,j) \in S\}$. Note that if $(i,j) \in S$, then $i \in S|_f$ and $j \in S|_i$. Let β denote the vector $(\beta_1, \ldots, \beta_m)$, where $\beta_i := \mathbf{E}[g_i]$ for each $i \in [m]$. For a set $J = \{i_1, \ldots, i_{|J|}\} \subseteq [m]$ and $f = f(y_1, \ldots, y_m)$, we write $\partial_J f$ to denote $\frac{\partial^{|J|} f}{\partial y_{i_1} \cdots \partial y_{i_{|J|}}}$. Since $\partial_J y^T = \mathbb{1}(T \supseteq J) y^{T \setminus J}$, by the multilinearity of f we have that

$$\partial_J f(\beta) = \sum_{T \supseteq J} \widehat{f}(T) \beta^{T \setminus J}. \tag{17}$$

6.2 The random restriction R_β

Given $\beta \in [-1,1]^m$, let R_β be the random restriction which is the randomized function from $\{-1,1\}^m$ to $\{-1,1\}^m$ whose i-th coordinate is (independently) defined by

$$R_\beta(y)_i := \begin{cases} \operatorname{sgn}(\beta_i) & \text{with probability } |\beta_i| \\ y_i & \text{with probability } 1 - |\beta_i|. \end{cases}$$

Note that we have

$$\mathop{\mathbf{E}}_{R_\beta, y}[R_\beta(y)_i] = \mathop{\mathbf{E}}_{R_\beta}[R_\beta(\vec{0})_i] = \beta_i.$$

Define $f_{R_\beta}(y)$ to be the (randomized) function $f(R_\beta(y))$. By the multilinearity of f and independence of the $R_\beta(y)_i$ we have

$$\mathop{\mathbf{E}}_{R_\beta, y}[f_{R_\beta}(y)] = \mathop{\mathbf{E}}_{R_\beta}[f_{R_\beta}(\vec{0})] = f(\beta).$$

The following claim relates the two derivatives $\partial_S f(\beta)$ and $\partial_S f_{R_\beta}(\vec{0}) = \widehat{f_{R_\beta}}(S)$.

Claim 33.

$$\partial_S f(\beta) = \prod_{i \in S} \frac{1}{1 - |\beta_i|} \cdot \mathop{\mathbf{E}}_{R_\beta}[\partial_S f_{R_\beta}(\vec{0})] = \prod_{i \in S} \frac{1}{1 - |\beta_i|} \cdot \mathop{\mathbf{E}}_{R_\beta}[\widehat{f_{R_\beta}}(S)].$$

Proof. The second equality follows from (17). To prove the first one, it suffices to show

$$\partial_{\{i\}} f(\beta) = \frac{1}{1 - |\beta_i|} \mathop{\mathbf{E}}_{R_\beta}[\partial_{\{i\}} f_{R_\beta}(\vec{0})].$$

The rest follows by a straightforward induction. By definition of derivatives, and noting $\mathbf{E}_{R_\beta}[R_\beta(\frac{1}{1-|\beta_i|} z \cdot e_i)_j]$ equal $\beta_i + z \cdot e_i$ if $i = j$ and β_j otherwise,

$$\begin{aligned} \frac{\partial f}{\partial y_i}(\beta) &= \lim_{z \to 0} \frac{f(\beta + z \cdot e_i) - f(\beta)}{z} \\ &= \mathop{\mathbf{E}}_{R_\beta}\left[\lim_{z \to 0} \frac{f_{R_\beta}\left(\frac{1}{1-|\beta_i|} \cdot z \cdot e_i\right) - f_{R_\beta}(\vec{0})}{z}\right] \\ &= \frac{1}{1 - |\beta_i|} \mathop{\mathbf{E}}_{R_\beta}\left[\frac{\partial f_{R_\beta}}{\partial y_i}(\vec{0})\right]. \end{aligned}$$

□

We now use Claim 33 to express each coefficient of h in terms of the coefficients of f and g_i.

Lemma 34. *For $S \subseteq [m]\times[\ell]$, we have* $\widehat{h}(S) = \prod_{i\in S|_f} \widehat{g}_i(S|_i)\cdot\prod_{i\in S|_f} \frac{1}{1-|\beta_i|}\cdot\mathbf{E}_{R_\beta}[\widehat{f_{R_\beta}}(S|_f)]$.

Proof. For $S \subseteq [m]\times[\ell]$ and $T \subseteq [m]$, decompose T into $(T \cap S|_f) \cup (T \setminus S|_f)$ and $S|_f$ into $(T \cap S|_f) \cup (S|_f \setminus T)$. Expanding h using the Fourier expansion of f, by the definitions of $S|_f$ and $S|_i$, we have

$$\begin{aligned}\widehat{h}(S) &= \sum_{T\subseteq[m]} \widehat{f}(T)\,\mathop{\mathbf{E}}_x\left[\prod_{i\in T} g_i(x)\cdot x^S\right]\\ &= \sum_{T\subseteq[m]} \widehat{f}(T)\left(\left(\prod_{i\in S|_f\setminus T} \mathbf{E}\Big[\prod_{j\in S|_i} x_{i,j}\Big]\right)\cdot\left(\prod_{i\in T\cap S|_f} \mathbf{E}\Big[g_i(x)\prod_{j\in S|_i} x_{i,j}\Big]\right)\cdot\left(\prod_{i\in T\setminus S|_f}\mathbf{E}[g_i(x)]\right)\right).\end{aligned}$$

When $S|_f \not\subseteq T$, we have $S|_f \setminus T \neq \varnothing$ and thus $\prod_{i\in S|_f\setminus T}\prod_{j\in S|_i}\mathbf{E}[x_{i,j}] = 0$. Hence,

$$\prod_{i\in S|_f\setminus T}\mathbf{E}\Big[\prod_{j\in S|_i} x_{i,j}\Big] = \mathbb{1}(S|_f \subseteq T).$$

Note that $\mathbf{E}[g_i(x)\prod_{j\in S} x_{i,j}] = \widehat{g}_i(S)$ for every $S \subseteq [\ell]$. Using $\mathbf{E}[g_i(x)] = \beta_i$, Equation (17) and Claim 33, we have

$$\begin{aligned}\widehat{h}(S) &= \sum_{T\subseteq[m]:T\supseteq S|_f} \widehat{f}(T)\prod_{i\in S|_f}\widehat{g}_i(S|_i)\prod_{i\in T\setminus S|_f}\mathbf{E}[g_i(x)]\\ &= \prod_{i\in S|_f}\widehat{g}_i(S|_i)\cdot\left(\sum_{T\subseteq[m]:T\supseteq S|_f}\widehat{f}(T)\prod_{i\in T\setminus S|_f}\beta_i\right) && (\mathbf{E}[g_i(x)]=\beta_i)\\ &= \prod_{i\in S|_f}\widehat{g}_i(S|_i)\cdot\partial_{S|_f} f(\beta_1,\dots,\beta_m) && (\text{Equation (17)})\\ &= \prod_{i\in S|_f}\widehat{g}_i(S|_i)\cdot\prod_{i\in S|_f}\frac{1}{1-|\beta_i|}\cdot\mathop{\mathbf{E}}_{R_\beta}\big[\widehat{f_{R_\beta}}(S|_f)\big]. && (\text{Claim 33})\end{aligned}$$

□

6.3 Proof of Theorem 31

By Lemma 34, $L_{1,K}(h)$ is equal to

$$\sum_{S\subseteq[m]\times[\ell]:|S|=K}|\widehat{h}(S)| = \sum_{S\subseteq[m]\times[\ell]:|S|=K}\left|\prod_{i\in S|_f}\widehat{g}_i(S|_i)\cdot\prod_{i\in S|_f}\frac{1}{1-|\beta_i|}\cdot\mathop{\mathbf{E}}_{R_\beta}[\widehat{f_{R_\beta}}(S|_f)]\right|.$$

We enumerate all the subsets $S \subseteq [m]\times[\ell]$ of size K in the following order: For every $|J| = k \in [K]$ out of the m blocks of ℓ coordinates, we enumerate all possible combinations of the (disjoint) nonempty subsets $\{S_i : i \in J\}$ in those k blocks whose sizes sum to K. Rewriting the summation above in this order, we obtain

$$\sum_{S\subseteq[m]\times[\ell]:|S|=K}|\widehat{h}(S)| = \sum_{k=1}^{K}\sum_{\substack{J\subseteq[m]\\|J|=k}}\sum_{\substack{w\subseteq[\ell]^J\\ \sum_{i\in J}w_i=K}}\sum_{\substack{\{S_i\}_{i\in J}\subseteq[\ell]^J:\\ \forall i\in J:|S_i|=w_i}}\left|\prod_{i\in J}\widehat{g}_i(S_i)\prod_{i\in J}\frac{1}{1-|\beta_i|}\mathop{\mathbf{E}}_{R_\beta}\big[\widehat{f_{R_\beta}}(J)\big]\right|$$

$$\leq \sum_{k=1}^{K} \sum_{\substack{J\subseteq[m]\\|J|=k}} \sum_{\substack{w\subseteq[\ell]^J\\ \sum_{i\in J} w_i=K}} \sum_{\substack{\{S_i\}_{i\in J}\subseteq[\ell]^J:\\ \forall i\in J:|S_i|=w_i}} \prod_{i\in J}|\widehat{g}_i(S_i)| \prod_{i\in J} \frac{1}{1-|\beta_i|} \left| \mathop{\mathbf{E}}_{R_\beta}\left[\widehat{f_{R_\beta}}(J)\right]\right|. \tag{18}$$

Since $L_{1,w_i}(g_i) \leq \frac{1-|\beta_i|}{2} \cdot a_{\mathsf{in}} \cdot b_{\mathsf{in}}^{w_i}$, for every $\{w_i\}_{i\in J}$ such that $\sum_{i\in J} w_i = K$, we have

$$\sum_{\substack{\{S_i\}_{i\in J}\subseteq[\ell]^J:\\ \forall i\in J:|S_i|=w_i}} \prod_{i\in J}|\widehat{g}_i(S_i)| = \prod_{i\in J} L_{1,w_i}(g_i) \leq \prod_{i\in J}\Big(\frac{1-|\beta_i|}{2} a_{\mathsf{in}} b_{\mathsf{in}}^{w_i}\Big) = b_{\mathsf{in}}^K a_{\mathsf{in}}^{|J|} \prod_{i\in J} \frac{1-|\beta_i|}{2}.$$

Plugging the above into (18), we get that

$$\begin{aligned}
\sum_{S\subseteq[m]\times[\ell]:|S|=K} |\widehat{h}(S)| &\leq b_{\mathsf{in}}^K \sum_{k=1}^{K} a_{\mathsf{in}}^k \sum_{\substack{J\subseteq[m]\\|J|=k}} \sum_{\substack{w\subseteq[\ell]^J\\ \sum_{i\in J} w_i=K}} \prod_{i\in J}\left(\frac{1-|\beta_i|}{2}\cdot\frac{1}{1-|\beta_i|}\cdot\left|\mathop{\mathbf{E}}_{R_\beta}\left[\widehat{f_{R_\beta}}(J)\right]\right|\right)\\
&= b_{\mathsf{in}}^K \sum_{k=1}^{K} \left(\frac{a_{\mathsf{in}}}{2}\right)^k \sum_{\substack{J\subseteq[m]\\|J|=k}} \left|\mathop{\mathbf{E}}_{R_\beta}\left[\widehat{f_{R_\beta}}(J)\right]\right| \sum_{\substack{w\subseteq[\ell]^J\\ \sum_{i\in J} w_i=K}} 1\\
&\leq b_{\mathsf{in}}^K \sum_{k=1}^{K} \left(\frac{a_{\mathsf{in}}}{2}\right)^k \binom{K-1}{k-1} \sum_{\substack{J\subseteq[m]\\|J|=k}} \left|\mathop{\mathbf{E}}_{R_\beta}\left[\widehat{f_{R_\beta}}(J)\right]\right|,
\end{aligned} \tag{19}$$

where the last inequality is because for every subset $J \subseteq [m]$, the set $\{w \subseteq [\ell]^J : \sum_{i\in J} w_i = K\}$ has size at most $\binom{K-1}{|J|-1}$. We now bound $|\mathbf{E}_{R_\beta}[\widehat{f_{R_\beta}}(J)]|$. Since for every restriction R_β, we have $f_{R_\beta} \in \mathcal{F}$ (by the assumption that $\mathcal{F}$ is closed under restrictions), it follows that

$$L_{1,k}(f_{R_\beta}) \leq \frac{1-|\mathbf{E}_y[f_{R_\beta}(y)]|}{2} a_{\mathsf{out}} b_{\mathsf{out}}^k \leq \frac{1-\mathbf{E}_y[f_{R_\beta}(y)]}{2} a_{\mathsf{out}} b_{\mathsf{out}}^k.$$

So

$$\begin{aligned}
\sum_{J\subseteq[m],|J|=k} \left|\mathop{\mathbf{E}}_{R_\beta}\left[\widehat{f_{R_\beta}}(J)\right]\right| &\leq \mathop{\mathbf{E}}_{R_\beta}[L_{1,k}(f_{R_\beta})]\\
&\leq \frac{1-\mathbf{E}_{R_\beta,y}[f_{R_\beta}(y)]}{2} a_{\mathsf{out}} b_{\mathsf{out}}^k\\
&= \frac{1-\mathbf{E}[h]}{2} a_{\mathsf{out}} b_{\mathsf{out}}^k.
\end{aligned}$$

Continuing from (19), we get

$$\begin{aligned}
\sum_{S\subseteq[m]\times[\ell]:|S|=K} |\widehat{h}(S)| &\leq \frac{1-\mathbf{E}[h]}{2} \cdot b_{\mathsf{in}}^K \cdot \sum_{k=1}^{K} \left(\frac{a_{\mathsf{in}}}{2}\right)^k \cdot \binom{K-1}{k-1} \cdot a_{\mathsf{out}} b_{\mathsf{out}}^k\\
&= \frac{1-\mathbf{E}[h]}{2} \cdot a_{\mathsf{out}} \cdot b_{\mathsf{in}}^K \cdot \frac{a_{\mathsf{in}} b_{\mathsf{out}}}{2} \left(1 + \frac{a_{\mathsf{in}} b_{\mathsf{out}}}{2}\right)^{K-1}
\end{aligned}$$

where the last step used the binomial theorem. Applying the same argument to $-h$ lets us replace $\frac{1-\mathbf{E}[h]}{2}$ with $\frac{1-|\mathbf{E}[h]|}{2}$, concluding the proof of Theorem 31. $\square$

6.4 Bounding $L_{1,1}(p)$ for balanced polynomials

We give a quick application of Claim 32 to show that $L_{1,1}(p) \le d$ for balanced degree-d polynomials. (Note that [CHLT19] showed that $L_{1,1}(p) \le 4d$ for general degree-d polynomials.)

Claim 35. *For every degree-d polynomial $p\colon \{0,1\}^n \to \{0,1\}$ with $\mathbf{E}[p] = 1/2$, we have $L_{1,1}(p) \le d$.*

Proof. Suppose not, and suppose that there exists an $\epsilon > 0$ such that $L_{1,1}(p) = (1+\epsilon)d$ for some degree-d balanced polynomial p on n variables. Consider the composition p' on disjoint sets of variables with the inner functions $g_1 = g_2 = \dots g_n$ and the outer function all equal to p. Note that p' is a balanced degree d^2 polynomial with $L_{1,1}(p') = (1+\epsilon)^2 d^2$.

We can repeat this iteratively by replacing p in p' in the composition. After k repetitions, we get a polynomial q of degree d^{2^k} that has $L_{1,1}(q) = (1+\epsilon)^{2^k} d^{2^k}$. For large enough k this is greater than $4d^{2^k}$, but this contradicts the $4d$ upper bound of [CHLT19], a contradiction. □

Acknowledgements

We thank Shivam Nadimpalli for stimulating discussions at the early stage of the project. J.B. is supported by a Junior Fellowship from the Simons Society of Fellows. P.I. is supported by NSF awards CCF-1813930 and CCF-2114116. Y.J. is supported by NSF IIS-1838154, NSF CCF-1703925, NSF CCF-1814873 and NSF CCF-1563155. C.H.L. is supported by the Croucher Foundation and the Simons Collaboration on Algorithms and Geometry. R.A.S. is supported by NSF grants CCF-1814873, IIS-1838154, CCF-1563155, and by the Simons Collaboration on Algorithms and Geometry. E.V. is supported by NSF awards CCF-1813930 and CCF-2114116.

References

[BGL06] Nayantara Bhatnagar, Parikshit Gopalan, and Richard J. Lipton. Symmetric polynomials over $\mathbb{Z}_m$ and simultaneous communication protocols. *J. Comput. System Sci.*, 72(2):252–285, 2006. 1.3.1, 2.1, 2.1, 4, 4.1, 17

[BHL12] Ido Ben-Eliezer, Rani Hod, and Shachar Lovett. Random low-degree polynomials are hard to approximate. *Comput. Complexity*, 21(1):63–81, 2012. 1.3, 1.3

[BHLV19] Ravi Boppana, Johan Håstad, Chin Ho Lee, and Emanuele Viola. Bounded independence versus symmetric tests. *ACM Trans. Comput. Theory*, 11(4):Art. 21, 27, 2019. 2.1, 4, 4.1, 4.1, 18, 4.1, 4.1

[BS21] Nikhil Bansal and Makrand Sinha. K-forrelation optimally separates quantum and classical query complexity. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2021, New York, NY, USA, 2021. 1.1

[BTW15] Eric Blais, Li-Yang Tan, and Andrew Wan. An inequality for the fourier spectrum of parity decision trees, 2015. 1.3.3, 1.5

[BV10] Andrej Bogdanov and Emanuele Viola. Pseudorandom bits for polynomials. *SIAM J. Comput.*, 39(6):2464–2486, 2010. 1.2

[CGL+20] Eshan Chattopadhyay, Jason Gaitonde, Chin Ho Lee, Shachar Lovett, and Abhishek Shetty. Fractional Pseudorandom Generators from Any Fourier Level, 2020. (document), 1.2, 5

[CHHL19] Eshan Chattopadhyay, Pooya Hatami, Kaave Hosseini, and Shachar Lovett. Pseudorandom generators from polarizing random walks. *Theory Comput.*, 15:Paper No. 10, 26, 2019. (document), 1.2, 3, 1.2, 1.2, 2.3, 6

[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 *10th Innovations in Theoretical Computer Science*, volume 124. 2019. (document), 1.2, 1.2, 4, 1.2, 1.2, 1.2, 6, 1.2, 1.2, 1.2, 1.3.1, 6.4, 6.4

[CHRT18] Eshan Chattopadhyay, Pooya Hatami, Omer Reingold, and Avishay Tal. Improved pseudorandomness for unordered branching programs through local monotonicity. In *STOC'18—Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 363–375. ACM, New York, 2018. 1.1, 1.3.3, 1.3.3

[CLW20] Lijie Chen, Xin Lyu, and R. Ryan Williams. Almost-everywhere circuit lower bounds from non-trivial derandomization. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1–12, 2020. 1.2

[CMM+20] Sourav Chakraborty, Nikhil S. Mande, Rajat Mittal, Tulasimohan Molli, Manaswi Paraashar, and Swagato Sanyal. Tight chang's-lemma-type bounds for boolean functions. *CoRR*, abs/2012.02335, 2020. 5, 5

[CPT20] Gil Cohen, Noam Peri, and Amnon Ta-Shma. Expander random walks: A fourier-analytic approach. *Electron. Colloquium Comput. Complex.*, 27:163, 2020. 2.1, 2.1, 4.2

[DHH20] Dean Doron, Pooya Hatami, and William M. Hoza. Log-seed pseudorandom generators via iterated restrictions. In *35th Computational Complexity Conference*, volume 169. 2020. 1.3.2

[FSUV13] Bill Fefferman, Ronen Shaltiel, Christopher Umans, and Emanuele Viola. On beating the hybrid argument. *Theory Comput.*, 9:809–843, 2013. 1.2

[GKV17] Frederic Green, Daniel Kreymer, and Emanuele Viola. Block-symmetric polynomials correlate with parity better than symmetric. *Comput. Complexity*, 26(2):323–364, 2017. 1.3.3

[GRT21] Uma Girish, Ran Raz, and Avishay Tal. Quantum Versus Randomized Communication Complexity, with Efficient Players. In James R. Lee, editor, *12th Innovations in Theoretical Computer Science Conference (ITCS 2021)*, volume 185, 2021. 1.1, 1.1

[GRZ20] Uma Girish, Ran Raz, and Wei Zhan. Lower bounds for XOR of forrelations, 2020. 1.1, 1.1

[GSTW16] Parikshit Gopalan, Rocco A. Servedio, Avishay Tal, and Avi Wigderson. Degree and sensitivity: tails of two distributions. *Electron. Colloquium Comput. Complex.*, 23:69, 2016. 1.1

[GT09] Ben Green and Terence Tao. The distribution of polynomials over finite fields, with applications to the Gowers norms. *Contrib. Discrete Math.*, 4(2):1–36, 2009. 1.3.1

[GTW21] Uma Girish, Avishay Tal, and Kewen Wu. Fourier Growth of Parity Decision Trees. In Valentine Kabanets, editor, *36th Computational Complexity Conference (CCC 2021)*, volume 200, 2021. 1.1, 1.3.3, 1.3.3, 1.5, 1.5

[KL93] Marek Karpinski and Michael Luby. Approximating the number of zeroes of a GF[2] polynomial. *Journal of Algorithms*, 14(2):280–287, 1993. 5.1

[Lee19] Chin Ho Lee. Fourier bounds and pseudorandom generators for product tests. In *34th Computational Complexity Conference*, volume 137. 2019. 1.1, 1.3.2, 1.3.2, 1.3.3

[LMS11] Shachar Lovett, Roy Meshulam, and Alex Samorodnitsky. Inverse conjecture for the Gowers norm is false. *Theory Comput.*, 7:131–145, 2011. 1.3.1

[Lov09] Shachar Lovett. Unconditional pseudorandom generators for low-degree polynomials. *Theory Comput.*, 5:69–82, 2009. 1.2

[LV20] Chin Ho Lee and Emanuele Viola. More on bounded independence plus noise: pseudorandom generators for read-once polynomials. *Theory Comput.*, 16:Paper No. 7, 50, 2020. 1.3.2

[Man95] Yishay Mansour. An $O(n^{\log \log n})$ learning algorithm for DNF under the uniform distribution. volume 50, pages 543–550. 1995. Fifth Annual Workshop on Computational Learning Theory (COLT) (Pittsburgh, PA, 1992). 1.1

[MRT19] Raghu Meka, Omer Reingold, and Avishay Tal. Pseudorandom generators for width-3 branching programs. In *STOC'19—Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pages 626–637. ACM, New York, 2019. 1.3.2

[Nis92] Noam Nisan. Pseudorandom generators for space-bounded computation. *Combinatorica*, 12(4):449–461, 1992. 1.3.2

[O'D14] Ryan O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, 2014. 1.1, 1.1, 4.3

[OLBC10] Frank WJ Olver, Daniel W Lozier, Ronald F Boisvert, and Charles W Clark. *NIST handbook of mathematical functions hardback and CD-ROM*. Cambridge university press, 2010. 4.2

[OS07] Ryan O'Donnell and Rocco A. Servedio. Learning monotone decision trees in polynomial time. *SIAM J. Comput.*, 37(3):827–844, 2007. 1.3.3

[Raz87] Alexander A. Razborov. Lower bounds on the dimension of schemes of bounded depth in a complete basis containing the logical addition function. *Mat. Zametki*, 41(4):598–607, 623, 1987. 1.2, 1.4.2, 1.4.2

[RSV13] Omer Reingold, Thomas Steinke, and Salil P. Vadhan. Pseudorandomness for regular branching programs via fourier analysis. In *RANDOM 2013*, volume 8096 of *Lecture Notes in Computer Science*, pages 655–670, 2013. (document), 1.1, 1.3.1, 1.3.1, 1.3.1

[RT19] Ran Raz and Avishay Tal. Oracle separation of BQP and PH. In *STOC'19—Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pages 13–23. ACM, New York, 2019. 1.1

[Smo87] Roman Smolensky. Algebraic methods in the theory of lower bounds for boolean circuit complexity. In *Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing*, STOC '87, page 77–82, New York, NY, USA, 1987. Association for Computing Machinery. 1.4.2, 1.4.2

[SSW20] Alexander A. Sherstov, Andrey A. Storozhenko, and Pei Wu. An Optimal Separation of Randomized and Quantum Query Complexity, 2020. 1.1, 1.1, 1.1

[ST19a] Rocco A. Servedio and Li-Yang Tan. Improved pseudorandom generators from pseudorandom multi-switching lemmas. In *Approximation, randomization, and combinatorial optimization. Algorithms and techniques*, volume 145. 2019. 1.4.1

[ST19b] Rocco A. Servedio and Li-Yang Tan. Pseudorandomness for read-k DNF formulas. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 621–638. SIAM, Philadelphia, PA, 2019. 1.4.1

[Tal17] Avishay Tal. Tight bounds on the Fourier spectrum of $\mathbf{AC^0}$. In *32nd Computational Complexity Conference*, volume 79. 2017. 1.1, 1.1

[Tal20] Avishay Tal. Towards optimal separations between quantum and randomized query complexities. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 228–239, 2020. 1.1, 1.1, 1.3.3, 1.3.3

[Vio09a] Emanuele Viola. On the power of small-depth computation. *Foundations and Trends in Theoretical Computer Science*, 5(1):1–72, 2009. 1.4.2

[Vio09b] Emanuele Viola. The sum of d small-bias generators fools polynomials of degree d. *Comput. Complexity*, 18(2):209–217, 2009. 1.2, 1.4.2

[Vio17] Emanuele Viola. Challenges in computational lower bounds. *SIGACT News, Open Problems Column*, 48(1), 2017. 1.4.2

[Vio20] Emanuele Viola. Fourier conjectures, correlation bounds, and majority. *Electron. Colloquium Comput. Complex.*, 27:175, 2020. 1.4.2, 15, 4.3

[Vio21] Emanuele Viola. New lower bounds for probabilistic degree and AC0 with parity gates. *Theory of Computing*, 2021. Available at http://www.ccs.neu.edu/home/viola/. 1.4.2

[Wik21] Wikipedia contributors. Kravchuk polynomials. `https://en.wikipedia.org/wiki/Kravchuk_polynomials`, accessed April 25, 2021. 4.2

A Reduction to bound without acceptance probability

In this section, we show that given any $L_{1,k}$ Fourier norm bound on a class of functions that is closed under XOR on disjoint variables, such a bound can be automatically "upgraded" to a refined bound that depends on the acceptance probability:

Lemma 36. *Let $\mathcal{F}$ be a class of $\{-1,1\}$-valued functions such that for every $f \in \mathcal{F}$, the XOR of disjoint copies of f (over disjoint sets of variables) also belongs to $\mathcal{F}$. If $L_{1,k}(\mathcal{F}) \le b^k$, then for every $f \in \mathcal{F}$ it holds that $L_{1,k}(f) \le 2e \cdot \frac{1-|\mathbf{E}[f]|}{2} \cdot b^k$.*

Proof. Suppose not, and let $f \in \mathcal{F}$ be such that $L_{1,k}(f) > 2e \cdot \frac{1-|\mathbf{E}[f]|}{2} \cdot b^k$. We first observe that since $L_{1,k}(\mathcal{F}) \le b^k$, it must be the case that $1 - |\mathbf{E}[f]| \le 1/e$. Let $\alpha := \frac{1-|\mathbf{E}[f]|}{2} \in [0, \frac{1}{2e}]$ so that $|\mathbf{E}[f]| = 1 - 2\alpha \ge 1 - 1/e$. Let $f^{\oplus t}$ be the XOR of t disjoint copies of f on tn variables, where the integer t is to be determined below. By our assumption, we have $f^{\oplus t} \in \mathcal{F}$ and thus

$$\begin{aligned} L_{1,k}(f^{\oplus t}) &\ge \binom{t}{1} \cdot L_{1,0}(f)^{t-1} \cdot L_{1,k}(f) && \text{(by disjointness)} \\ &= t \cdot (1-2\alpha)^{t-1} \cdot L_{1,k}(f) && (L_{1,0}(f) = \mathbf{E}[f]) \\ &> t \cdot (1-2\alpha)^{t-1} \cdot 2e \cdot \alpha \cdot b^k =: \Lambda(t). \end{aligned}$$

We note that if $\alpha = 0$ then $|\mathbf{E}[f]| = 1$, so all the Fourier weight of f is on the constant coefficient, and hence the claimed inequality holds trivially. So we subsequently assume that $0 < \alpha \le \frac{1}{2e}$. Let $t^* := \frac{1}{-\ln(1-2\alpha)} > 0$. It is easy to verify that $\Lambda(t)$ is increasing when $t \le t^*$, and is decreasing when $t \ge t^*$.

We choose $t = \lceil t^* \rceil$. Since $\alpha \le \frac{1}{2e} < \frac{e-1}{2e} \approx 0.3161$, we have $t^* > 1$ and thus

$$\begin{aligned} L_{1,k}(f^{\oplus t}) > \Lambda(\lceil t^* \rceil) \ge \Lambda(t^*+1) &= \left(\frac{1}{-\ln(1-2\alpha)} + 1\right) \cdot (1-2\alpha)^{\frac{1}{-\ln(1-2\alpha)}} \cdot 2e \cdot \alpha \cdot b^k \\ &= \left(\frac{2\alpha}{-\ln(1-2\alpha)} + 2\alpha\right) \cdot b^k \ge b^k, \end{aligned}$$

where the last inequality holds for every $\alpha \in (0, \frac{e-1}{2e}]$ and can be checked via elementary calculations. This contradicts $L_{1,k}(\mathcal{F}) \le b^k$, and the lemma is proved. $\square$