On Parity Decision Trees for Fourier-Sparse Boolean Functions

Nikhil S. Mande

Georgetown University, Washington, DC, USA
http://nikhil.georgetown.domains/
nikhil.mande@georgetown.edu

Swagato Sanyal

Indian Institute of Technology Kharagpur, India
swagato@cse.iitkgp.ac.in

Abstract

We study parity decision trees for Boolean functions. The motivation of our study is the log-rank conjecture for XOR functions and its connection to Fourier analysis and parity decision tree complexity. Our contributions are as follows. Let $f : \mathbb{F}_2^n \to \{-1, 1\}$ be a Boolean function with Fourier support $\mathcal{S}$ and Fourier sparsity k.

- We prove via the probabilistic method that there exists a parity decision tree of depth $O(\sqrt{k})$ that computes f. This matches the best known upper bound on the parity decision tree complexity of Boolean functions (Tsang, Wong, Xie, and Zhang, FOCS 2013). Moreover, while previous constructions (Tsang et al., FOCS 2013, Shpilka, Tal, and Volk, Comput. Complex. 2017) build the trees by carefully choosing the parities to be queried in each step, our proof shows that a naive sampling of the parities suffices.
- We generalize the above result by showing that if the Fourier spectra of Boolean functions satisfy a natural "folding property", then the above proof can be adapted to establish existence of a tree of complexity polynomially smaller than $O(\sqrt{k})$. More concretely, the folding property we consider is that for most distinct γ, δ in $\mathcal{S}$, there are at least a polynomial (in k) number of pairs (α, β) of parities in $\mathcal{S}$ such that $\alpha + \beta = \gamma + \delta$. We make a conjecture in this regard which, if true, implies that the communication complexity of an XOR function is bounded above by the fourth root of the rank of its communication matrix, improving upon the previously known upper bound of square root of rank (Tsang et al., FOCS 2013, Lovett, J. ACM. 2016).
- Motivated by the above, we present some structural results about the Fourier spectra of Boolean functions. It can be shown by elementary techniques that for any Boolean function f and all (α, β) in $\binom{\mathcal{S}}{2}$, there exists another pair (γ, δ) in $\binom{\mathcal{S}}{2}$ such that $\alpha + \beta = \gamma + \delta$. One can view this as a "trivial" folding property that all Boolean functions satisfy. Prior to our work, it was conceivable that for all $(\alpha, \beta) \in \binom{\mathcal{S}}{2}$, there exists exactly one other pair $(\gamma, \delta) \in \binom{\mathcal{S}}{2}$ with $\alpha + \beta = \gamma + \delta$. We show, among other results, that there must exist several $\gamma \in \mathbb{F}_2^n$ such that there are at least three pairs of parities $(\alpha_1, \alpha_2) \in \binom{\mathcal{S}}{2}$ with $\alpha_1 + \alpha_2 = \gamma$. This, in particular, rules out the possibility stated earlier.

2012 ACM Subject Classification Theory of computation → Oracles and decision trees

Keywords and phrases Parity decision trees, log-rank conjecture, analysis of Boolean functions, communication complexity

Digital Object Identifier 10.4230/LIPIcs.FSTTCS.2020.29

Related Version A full version of the paper is available at https://arxiv.org/abs/2008.00266.

Funding *Swagato Sanyal*: Supported by an ISIRD Grant from Sponsored Research and Industrial Consultancy, IIT Kharagpur.

Acknowledgements We thank Prahladh Harsha, Srikanth Srinivasan, Sourav Chakraborty and Manaswi Paraashar for useful discussions.

1 Introduction

The log-rank conjecture [6] is a fundamental unsolved question in communication complexity that states that the deterministic communication complexity of a Boolean function is polynomially related to the logarithm of the rank (over real numbers) of its communication matrix. The importance of the conjecture stems from the fact that it proposes to characterize communication complexity, which is an interactive complexity measure, by the rank of a matrix which is a traditional and well-understood algebraic measure. In this work we focus on the important and well-studied class of XOR functions. Consider a two-party function $F : \mathbb{F}_2^n \times \mathbb{F}_2^n \to \{-1, 1\}$ whose value on any input (x, y) depends only on the bitwise XOR of x and y, i.e., there exists a function $f : \mathbb{F}_2^n \to \{-1, 1\}$ such that for each $(x, y) \in \mathbb{F}_2^n \times \mathbb{F}_2^n$, $F(x, y) = f(x \oplus y)$. Such a function F is called an XOR function, and is denoted as $F = f \circ \oplus$. The log-rank conjecture and communication complexity of such an XOR function F has interesting connections with the Fourier spectrum of f. For example, it is known that the rank of the communication matrix of F equals the Fourier sparsity of f (henceforth referred to as k) [2]. The natural randomized analogue of the log-rank conjecture is the log-approximate-rank conjecture [5], which was recently refuted by Chattopadhyay, Mande, and Sherif [3]. The quantum analogue of the log-rank conjecture was subsequently also refuted by Sinha and de Wolf [12] and Anshu, Boddu, and Touchette [1]. It is worth noting that an XOR function was used to refute these conjectures.

To design a cheap communication protocol for F, an approach adopted by many works [11, 14, 9] is to design a small-depth parity decision tree (henceforth referred to as PDT) for f, and having a communication protocol simulate the tree; it is easy to see that the parity of a subset of bits of the string $x \oplus y$ can be computed by the communicating parties by interchanging two bits. The parity decision tree complexity (henceforth referred to as $\mathrm{PDT}(\cdot)$) of f thus places an asymptotic upper bound on the communication complexity of F. The work of Hatami, Hosseini and Lovett [4] shows that this approach is polynomially tight; they showed that $\mathrm{PDT}(f)$ is polynomially related to the deterministic communication complexity of F. In light of this, the log-rank conjecture for XOR functions $F = f \circ \oplus$ is readily seen to be equivalent to $\mathrm{PDT}(f)$ being polylogarithmic in k.

However, we are currently very far from achieving this goal. Lovett [7] showed that the deterministic communication complexity of any Boolean function F is bounded above by $O(\sqrt{\mathsf{rank}(F)} \log \mathsf{rank}(F))$. In particular, this implies that that the deterministic communication complexity of $F = f \circ \oplus$ is $O(\sqrt{k} \log k)$. Tsang et al. [14] showed that $\mathrm{PDT}(f) = O(\sqrt{k})$ (a quantitatively weaker bound was shown in a simultaneous and independent work of Shpilka et al. [11]). In addition to bounding $\mathrm{PDT}(f)$ instead of the communication complexity of F, Tsang et al. achieved a quantitative improvement by a logarithmic factor over Lovett's bound for the class of XOR functions. Sanyal [10] showed that the simultaneous communication complexity of F (characterized by the Fourier dimension of f) is bounded above by $O(\sqrt{k} \log k)$, and is tight (up to the $\log k$ factor) for the addressing function.

In this work we derive new understanding about the structure of Fourier spectra of Boolean functions. Aided by this insight we reprove the $O(\sqrt{k})$ upper bound on $\mathrm{PDT}(f)$ (see Sections 3.1 and 3.2). We conditionally improve this bound by a polynomial factor, assuming a "folding property" of the Fourier spectra of Boolean functions (see Section 3.3). To prove these results, we make use of a simple necessary condition for a function to be Boolean (see Proposition 5). While we show that it is not a sufficient condition (see Theorem 27), it does enable us to prove the above results. In these proofs, we use Proposition 5 in conjunction with probabilistic and combinatorial arguments. Finally, we make progress towards establishing the folding property (see Section 3.4). To prove these results, we use the well-known characterization of Boolean functions given by two conditions, namely Parseval's identity (Equation (2)) and a condition attributed to Titsworth (Equation (3)), in conjunction with combinatorial arguments.

1.1 Organization of this paper

In Section 2 we review some preliminaries and introduce the notation that we use in this paper. In this section we also introduce definitions and concepts that are needed to state our results formally. In Section 3 we motivate and formally state our results, and discuss proof techniques. The formal proofs of our main results can be found in Sections 4 and 5 of this paper, and in Sections 5 and 6 of the full version of this paper [8].

2 Notation and preliminaries

All logarithms in this paper are taken with base 2. We use the phrase "k is sufficiently large" to mean that there exists a universal constant $C > 0$ such that $k > C$. As is standard, we use the notation $f(n) = \widetilde{O}(h(n))$ $(f(n) = \widetilde{\Theta}(\cdot), f(n) = \widetilde{\Omega}(\cdot))$ to convey that there exists a constant $c \geq 0$ such that that $f(n) = O(h(n) \log^c h(n))$ $(f(n) = \Theta(h(n) \log^c h(n)),$ $f(n) = \Omega(h(n) \log^c h(n))$, respectively). We use the notation $[n]$ to denote the set $\{1, 2, \ldots, n\}$. For any set S, we use the notation $\binom{S}{2}$ to denote the set of all subsets of S of size exactly 2. We abuse notation and denote a generic element of $\binom{S}{2}$ as (a, b) rather than $\{a, b\}$. When we use the notation $\mathbb{E}_{x\in X}[\cdot]$, the underlying distribution corresponds to x being sampled uniformly at random from X. For $a \in \mathbb{F}_2$, we let a^n denote the n-bit string $(a, a, \ldots, a)$. We use the symbol "+" to denote both coordinate-wise addition over $\mathbb{F}_2$ as well as addition over reals; the meaning in use will be clear from context. For sets $A, B \subseteq \mathbb{F}_2^n$, $A + B$ denotes the *sumset* defined by $\{\alpha + \beta \mid \alpha \in A, \beta \in B\}$. For a set $A \subseteq \mathbb{F}_2^n$ and $\gamma \in \mathbb{F}_2^n$, we denote by $A + \gamma$ the set $A + \{\gamma\}$. The above convention also extends to the symbol "$\sum$". For a set of vectors $\Gamma \in \mathbb{F}_2^n$, we define $\mathsf{span}\ \Gamma$ to be the set of all $\mathbb{F}_2$-linear combinations of vectors in Γ, i.e., $\mathsf{span}\ \Gamma = \left\{\sum_{\gamma\in\Gamma} c_\gamma \cdot \gamma \;\middle|\; c_\gamma \in \mathbb{F}_2 \text{ for } \gamma \in \Gamma\right\}$.

Consider the vector space of functions from $\mathbb{F}_2^n$ to $\mathbb{R}$, equipped with the following inner product.

$$\langle f, g\rangle := \mathbb{E}_{x\in\mathbb{F}_2^n}[f(x)g(x)] = \frac{1}{2^n} \sum_{x\in\mathbb{F}_2^n} f(x)g(x).$$

Let $x = (x_1, \ldots, x_n) \in \mathbb{F}_2^n$. For each $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{F}_2^n$, define $\alpha(x) := \sum_{i=1}^n \alpha_i x_i \pmod 2$, and the associated *character* $\chi_\alpha : \mathbb{F}_2^n \to \{-1, 1\}$ by $\chi_\alpha(x) := (-1)^{\alpha(x)}$. Observe that $\chi_\alpha(x)$ is the ± 1-valued parity of the bits $\{x_i \mid \alpha_i = 1\}$; due to this we will also refer to characters as parities. The set of parities $\{\chi_\alpha \mid \alpha \in \mathbb{F}_2^n\}$ forms an orthonormal (with respect to the above inner product) basis for this vector space. Hence, every function $f : \mathbb{F}_2^n \to \mathbb{R}$ can be uniquely written as $f = \sum_{\alpha\in\mathbb{F}_2^n} \widehat{f}(\alpha)\chi_\alpha$, where $\widehat{f}(\alpha) = \langle f, \chi_\alpha\rangle = \mathbb{E}_{x\in\mathbb{F}_2^n}[f(x)\chi_\alpha(x)]$. The coefficients $\left\{\widehat{f}(\alpha) \;\middle|\; \alpha \in \mathbb{F}_2^n\right\}$ are called the Fourier coefficients of f.

For any function $f : \mathbb{F}_2^n \to \{-1, 1\}$ and any set $A \subseteq \mathbb{F}_2^n$, define the function $f\mid_A : A \to \{-1, 1\}$ by $f\mid_A(x) = f(x)$ for all $x \in A$. In other words, $f\mid_A$ denotes the restriction of f to A.

Throughout this paper, for any Boolean function $f : \mathbb{F}_2^n \to \{-1, 1\}$, we denote by $\mathcal{S}$ the Fourier support of f, i.e. $\mathcal{S} = \left\{\alpha \in \mathbb{F}_2^n \;\middle|\; \widehat{f}(\alpha) \neq 0\right\}$. We also denote by k the Fourier sparsity of f, i.e. $k = |\mathcal{S}|$. The dependence of $\mathcal{S}$ and k on f is suppressed and the underlying function will be clear from context.

The representation of Fourier coefficients as an expectation (over $x \in \mathbb{F}_2^n$) immediately yields the following observation about granularity of Fourier coefficients of Boolean functions.

▶ **Observation 1.** *Let $f : \mathbb{F}_2^n \to \{-1, 1\}$ be any Boolean function. Then, for all $\alpha \in \mathbb{F}_2^n$, $\widehat{f}(\alpha)$ is an integral multiple of $1/2^n$.*

We next define plateaued functions.

▶ **Definition 2** (Plateaued functions)**.** *A Boolean function $f : \mathbb{F}_2^n \to \{-1, 1\}$ is said to be* plateaued *if there exists $x \in \mathbb{R}$ such that $\widehat{f}(\alpha) \in \{0, x, -x\}$ for all $\alpha \in \mathbb{F}_2^n$.*

Next we define the addressing function.

▶ **Definition 3** (Addressing function)**.** *Let k be an even power of 2. The addressing function $\mathsf{ADD}_k : \mathbb{F}_2^{\frac{1}{2}\log k + \sqrt{k}} \to \{-1, 1\}$ is defined as*

$$\mathsf{ADD}_k(x, y_1, \dots, y_{\sqrt{k}}) := (-1)^{y_{\mathrm{int}(x)}},$$

where $x \in \mathbb{F}_2^{\frac{1}{2}\log k}$, $y_i \in \mathbb{F}_2$ for $i = 1, \dots, \sqrt{k}$, and $\mathrm{int}(x)$ is the unique integer in $\left\{1, \dots, \sqrt{k}\right\}$ whose binary representation is x.

The Fourier sparsity of ADD_k can be verified to be k. We now define a notion of equivalence on elements of $\binom{\mathcal{S}}{2}$.

▶ **Definition 4.** *For any Boolean function $f : \mathbb{F}_2^n \to \{-1, 1\}$, we say a pair $(\alpha_1, \alpha_2) \in \binom{\mathcal{S}}{2}$ is* equivalent *to $(\alpha_3, \alpha_4) \in \binom{\mathcal{S}}{2}$ if $\alpha_1 + \alpha_2 = \alpha_3 + \alpha_4$.*

In the above definition, if $\alpha_1 + \alpha_2 = \alpha_3 + \alpha_4 = \gamma$, then we say that the pairs (α_1, α_2) and (α_3, α_4) *fold in the direction* γ. We also say that the elements $\alpha_1, \alpha_2, \alpha_3$, and α_4 *participate in the folding direction* γ. It is not hard to verify that the notion of equivalence defined above does indeed form an equivalence relation. We will denote by D_γ the equivalence class of pairs that fold in the direction γ, i.e.,

$$D_\gamma := \left\{(\alpha, \beta) \in \binom{\mathcal{S}}{2} \;\middle|\; \alpha + \beta = \gamma\right\}.$$

We suppress the dependence of D_γ on the underlying function f, which will be clear from context. Unless mentioned otherwise, these are the equivalence classes under consideration throughout this paper.

For any Boolean function $f : \mathbb{F}_2^n \to \{-1, 1\}$, we have for each $x \in \mathbb{F}_2^n$:

$$1 = f^2(x) = \sum_{\gamma \in \mathbb{F}_2^n} \left(\sum_{(\alpha_1, \alpha_2) \in \mathbb{F}_2^n \times \mathbb{F}_2^n : \alpha_1 + \alpha_2 = \gamma} \widehat{f}(\alpha_1)\widehat{f}(\alpha_2) \right) \chi_\gamma(x). \tag{1}$$

Matching the constant term of each side of the above identity we have

$$\sum_{\alpha \in \mathbb{F}_2^n} \widehat{f}(\alpha)^2 = 1, \tag{2}$$

which is commonly referred to as Parseval's identity for Boolean functions. By matching the coefficient of each non-constant χ_γ on each side of Equation (1) we obtain

$$\forall \gamma \neq 0^n, \quad \sum_{(\alpha_1, \alpha_2) \in \mathbb{F}_2^n \times \mathbb{F}_2^n : \alpha_1 + \alpha_2 = \gamma} \widehat{f}(\alpha_1)\widehat{f}(\alpha_2) = 0. \tag{3}$$

Equation (3) is attributed to Titsworth [13]. The following proposition is an easy consequence of Equation (3). It provides a necessary condition for a subset of $\mathbb{F}_2^n$ to be the Fourier support of a Boolean function.

▶ **Proposition 5.** *Let $f : \mathbb{F}_2^n \to \{-1, 1\}$ be a Boolean function. Then, for all $(\alpha, \beta) \in \binom{\mathcal{S}}{2}$, there exists $(\gamma, \delta) \neq (\alpha, \beta) \in \binom{\mathcal{S}}{2}$ such that $\alpha + \beta = \gamma + \delta$. In other words, $|D_{\alpha+\beta}| \geq 2$.*

The Fourier ℓ_1-norm of f is defined as $\|\widehat{f}\|_1 := \sum_{\alpha \in \mathbb{F}_2^n} |\widehat{f}(\alpha)|$. By the Cauchy-Schwarz inequality and Equation (2), we have

$$\|\widehat{f}\|_1 \leq \sqrt{k} \sqrt{\sum_{\alpha \in \mathbb{F}_2^n} \widehat{f}(\alpha)^2} = \sqrt{k}. \tag{4}$$

We next formally define parity decision trees.

A *parity decision tree* (PDT) is a binary tree whose leaf nodes are labeled in $\{-1, 1\}$, each internal node is labeled by a parity χ_α and has two outgoing edges, labeled -1 and 1. On an input $x \in \mathbb{F}_2^n$, the tree's computation proceeds from the root down as follows: compute $\chi_\alpha(x)$ as indicated by the node's label and following the edge indicated by the value output, and continue in a similar fashion until a reaching a leaf, at which point the value of the leaf is output. When the computation reaches a particular internal node, the PDT is said to *query* the parity label of that node. The PDT is said to compute a function $f : \mathbb{F}_2^n \to \{-1, 1\}$ if its output equals the value of f for all $x \in \mathbb{F}_2^n$. The parity decision tree complexity of f, denoted $\mathrm{PDT}(f)$ is defined as

$$\mathrm{PDT}(f) := \min_{T : T \text{ is a PDT computing } f} \mathrm{depth}(T).$$

2.1 Restriction to an affine subspace

In this section we discuss the effect of restricting a function $f : \mathbb{F}_2^n \to \mathbb{R}$ to an affine subspace, on the Fourier spectrum of f.

▶ **Definition 6** (Affine subspace). *A set $V \subseteq \mathbb{F}_2^n$ is called an affine subspace if there exist linearly independent vectors $\ell_1, \ldots, \ell_t \in \mathbb{F}_2^n$ and elements $a_1, \ldots, a_t \in \mathbb{F}_2$ such that $V = \{x \in \mathbb{F}_2^n \mid \ell_i(x) = a_i \ \forall i \in \{1, \ldots, t\}\}$. t is called the co-dimension of V.*

Consider a set $\Gamma := \{\gamma_1, \ldots, \gamma_t\}$ of vectors in $\mathbb{F}_2^n$. Define $\mathcal{G} := \mathsf{span}\ \Gamma$, and $\mathcal{C} := \{\mathcal{G} + \beta \mid \beta \in \mathbb{F}_2^n,\ (\mathcal{G} + \beta) \cap \mathcal{S} \neq \emptyset\}$ to be the cosets of $\mathcal{G}$ that have non-trivial intersection with $\mathcal{S}$. For each $C \in \mathcal{C}$, let $\alpha(C)$ denote an arbitrary but fixed element in $C \cap \mathcal{S}$. In light of this, we write the Fourier transform of f as

$$f(x) = \sum_{C \in \mathcal{C}} \left(\sum_{\gamma \in \mathcal{G}} \widehat{f}(\alpha(C) + \gamma) \chi_\gamma(x) \right) \chi_{\alpha(C)}(x), \tag{5}$$

For any such fixed C, the value of the sum $\sum_{\gamma \in \mathcal{G}} \widehat{f}(\alpha(C) + \gamma) \chi_\gamma(x)$ that appears in Equation (5) is determined by the values $\gamma_1(x), \ldots, \gamma_t(x)$. Denote this sum by $P_C(\gamma_1(x), \ldots, \gamma_t(x))$. For $\mathbf{b} := (b_1, \ldots, b_t) \in \mathbb{F}_2^t$, let $H_{\mathbf{b}}$ be the affine subspace $\{x \in \mathbb{F}_2^n \mid \gamma_1(x) = b_1, \ldots, \gamma_t(x) = b_t\}$. It follows immediately that the Fourier transform of $f\,|_{H_{\mathbf{b}}}$ is given by

$$f\,|_{H_{\mathbf{b}}}(x) = \sum_{C \in \mathcal{C}} P_C(b_1, \ldots, b_t) \chi_{\alpha(C)}(x). \tag{6}$$

In particular, for each $\mathbf{b}$, the Fourier sparsity of $f\,|_{H_{\mathbf{b}}}$ is bounded above by $|\mathcal{C}|$.

We note here that each element in $\mathcal{S}$ is mapped to a unique element in $\mathcal{C}$. The elements of $\mathcal{C}$ can thus be thought of as buckets that form a partition of $\mathcal{S}$. Keeping this view in mind we define the following.

▶ **Definition 7** (Bucket complexity). *Let $f : \mathbb{F}_2^n \to \{-1, 1\}$ be any Boolean function. Consider a set of vectors $\Gamma = \{\gamma_1, \ldots, \gamma_t\}$ in $\mathbb{F}_2^n$. Let $\mathcal{G} := \mathsf{span}\,\Gamma$, and let $\mathcal{C}$ denote the set of cosets of $\mathcal{G}$ that have non-empty intersection with $\mathcal{S}$, that is, $\mathcal{C} := \{\mathcal{G} + \beta \mid \beta \in \mathbb{F}_2^n,\ (\mathcal{G} + \beta) \cap \mathcal{S} \neq \emptyset\}$. Define the* bucket complexity *of f with respect to $\mathcal{G}$, denoted $\mathcal{B}(f, \mathcal{G})$, as*

$$\mathcal{B}(f, \mathcal{G}) = |\mathcal{C}|.$$

We now make the following useful observation, which follows from Equation (6).

▶ **Observation 8.** *Let Γ and $\mathcal{G}$ be as in Definition 7. Let $\mathbf{b} = (b_1, \ldots, b_t) \in \mathbb{F}_2^t$ be arbitrary. Let V be the affine subspace $\{x \in \mathbb{F}_2^n \mid \gamma_1(x) = b_1, \ldots, \gamma_t(x) = b_t\}$. Let k' be the Fourier sparsity of $f\mid_V$. Then $k' \leq \mathcal{B}(f, \mathcal{G})$.*

▶ **Definition 9** (Identification of characters). *For $f, \mathcal{G}$, and $\mathcal{C}$ as in Definition 7 and any $\beta, \delta \in \mathcal{S}$, we say that β* and *δ* are identified with respect to *$\mathcal{G}$ if $\beta + \delta \in \mathcal{G}$, or equivalently, if β and δ belong to the same coset in $\mathcal{C}$.*

The following observation plays a key role in the results discussed in this paper.

▶ **Observation 10.** *Let $f, \mathcal{G}$ and $\mathcal{C}$ be as in Definition 7. If there exists a set $L \subseteq \mathcal{S}$ of size h such that each $\beta \in L$ is identified with some other $\delta \in \mathcal{S}$ with respect to $\mathcal{G}$, then $\mathcal{B}(f, \mathcal{G}) \leq k - \frac{h}{2}$.*

Proof. Since $|\overline{L}| = k - h$, there are at most $k - h$ cosets in $\mathcal{C}$ that contain at least one element from $\overline{L}$. Next, each coset in $\mathcal{C}$ that contains only elements from L has at least 2 elements (by the hypothesis). Hence, the number of cosets containing only elements from L is at most $h/2$. Combining the above two, we have that $|\mathcal{C}| \leq (k - h) + \frac{h}{2} = k - \frac{h}{2}$. ◀

2.2 Folding properties of Boolean functions

▶ **Definition 11.** *Let $f : \mathbb{F}_2^n \to \{-1, 1\}$ be any Boolean function. We say that f is (δ, ℓ)-*folding *if*

$$\left|\left\{(\alpha, \beta) \in \binom{\mathcal{S}}{2} \;\middle|\; |D_{\alpha+\beta}| \geq k^\ell + 1\right\}\right| \geq \delta\binom{k}{2}.$$

Proposition 5 implies that any Boolean function is $(1, 0)$-folding.

We next show by a simple averaging argument that if f has "good folding properties", then there are many $\alpha \in \mathcal{S}$, such that $|D_{\alpha+\beta}|$ is large for many $\beta \in \mathcal{S} \setminus \{\alpha\}$.

▷ Claim 12. Let $f : \mathbb{F}_2^n \to \{-1, 1\}$ be (δ, ℓ)-folding and $k \geq 6$. Define

$$U := \left\{\alpha \in \mathcal{S} \;\middle|\; \text{there exist at least } \delta k/2 \text{ many } \beta \in \mathcal{S} \setminus \{\alpha\} \text{ with } |D_{\alpha+\beta}| \geq k^\ell + 1\right\}.$$

Then $|U| \geq \frac{\delta k}{3}$.

Proof. For each $\alpha \in \mathcal{S}$, define $t(\alpha) := \left|\left\{\beta \in \mathcal{S} \setminus \{\alpha\} \;\middle|\; |D_{\alpha+\beta}| \geq k^\ell + 1\right\}\right|$. By the hypothesis, $\sum_{\alpha \in \mathcal{S}} t(\alpha) \geq \delta k(k - 1)$. We have

$$|U| \cdot k + (k - |U|) \cdot \frac{\delta k}{2} \geq \sum_{\alpha \in \mathcal{S}} t(\alpha) \geq \delta k(k - 1)$$

$$\implies |U|\left(k - \frac{\delta k}{2}\right) \geq \delta k^2 - \delta k - \frac{\delta k^2}{2} \implies |U| \geq \frac{\delta(k - 2)}{2 - \delta},$$

implying $|U| \geq \frac{\delta k}{3}$ for $k \geq 6$.

◁

3 Our contributions

In this section we give a high-level account of our contributions in this paper. In Section 3.1 we discuss the PDT construction of Tsang et al. We motivate, state our results, and briefly discuss proof ideas in Sections 3.2, 3.3, and 3.4.

3.1 Low bucket complexity implies shallow PDTs

The following lemma follows from [14, Lemma 28] and Equation (4).

▶ **Lemma 13** (Tsang, Wong, Xie, and Zhang). *Let $f : \mathbb{F}_2^n \to \{-1, 1\}$ be any Boolean function. Then there exists an affine subspace V of $\mathbb{F}_2^n$ of co-dimension $O(\sqrt{k})$ such that f is constant on V.*

Let $V = \{x \in \mathbb{F}_2^n \mid \gamma_1(x) = b_1, \ldots, \gamma_t(x) = b_t\}$ be the affine subspace V obtained from Lemma 13, where $t = O(\sqrt{k})$. Define $\mathcal{G} := \mathsf{span}\,\{\gamma_1, \ldots, \gamma_t\}$. We next observe that $\mathcal{B}(f, \mathcal{G}) \leq k/2$. To see this, note that since $f\,|_V$ is constant, we have from Equation (6) that for each coset $C \in \mathcal{C}$ and any $(b_1, \ldots, b_t) \in \mathbb{F}_2^t$,

$$P_C(b_1, \ldots, b_t) = \begin{cases} \pm 1 & \text{if } 0^n \in C \\ 0 & \text{otherwise.} \end{cases}$$

Since f is a non-constant function, this implies that each $P_C(\cdot)$ has at least 2 terms, i.e., each $\beta \in \mathcal{S}$ is identified with some other $\delta \in \mathcal{S}$ with respect to $\mathcal{G}$. Observation 10 implies that $\mathcal{B}(f, \mathcal{G}) \leq k/2$. Observation 8 implies that the Fourier sparsity of the restriction of f to each coset of V is at most $k/2$.

This immediately leads to a recursive construction of a PDT for f of depth $O(\sqrt{k})$ as follows. The first step is to query the parities $\gamma_1, \ldots, \gamma_t$. After this step, each leaf of the partial tree obtained is a restriction of f to some coset of V. Next we recursively compute each leaf. Since after each batch of queries, the sparsity reduces by a factor of 2, the depth of the tree thus obtained is $O\left(\sqrt{k} + \sqrt{\frac{k}{2}} + \sqrt{\frac{k}{2^2}} + \cdots\right) = O(\sqrt{k})$.

3.2 A random set of parities achieves low bucket complexity

Tsang et al. proved Lemma 13 by an iterative procedure in each step of which a single parity is carefully chosen. We show in this paper that a randomly sampled set of parities achieves the desired bucket complexity upper bound with high probability. More specifically, for a parameter $p \in [0, 1]$, consider the procedure $\textsc{SampleParity}(f, p)$ described in Algorithm 1. Our first result shows that the set $\mathcal{R}$ returned by $\textsc{SampleParity}\left(f, \frac{1}{\Theta(\sqrt{k})}\right)$

Algorithm 1

```
procedure SampleParity (f, p)
    R ← ∅;
    for each α ∈ S do
        independently with probability p, R ← R ∪ {α};
    end for
    Return R;
end procedure
```

satisfies $\mathcal{B}(f, \mathsf{span}\ \mathcal{R}) \leq (1 - \Omega(1))k$ with high probability.

▶ **Theorem 14.** *Let $f : \mathbb{F}_2^n \to \{-1, 1\}$ be a Boolean function and k be sufficiently large. Let $p = \frac{1}{2k^{1/2}}$ and $\mathcal{R}$ be the random set of parities returned by* SampleParity*(f, p). There exists a constant $c \in [0, 1)$ such that*

$$\mathbb{E}[\mathcal{B}(f, \mathsf{span}\ \mathcal{R})] \leq ck.$$

With high probability we have $|\mathcal{R}| = O(\sqrt{k})$. By an argument analogous to the discussion in the previous section, Theorem 14 recovers the $O(\sqrt{k})$ upper bound on $\mathrm{PDT}(f)$. An additional insight that our work provides is that a PDT of depth $O(\sqrt{k})$ can be obtained by a naive sampling procedure applied iteratively.

We note here that while Tsang et al. prove a bucket complexity upper bound of $k/2$ via Lemma 13 which restricts the function to a constant, we derive a bucket complexity upper bound of $(1 - \Omega(1))k$ by analyzing the procedure SampleParity.

Proof idea

Fix any $\alpha \in \mathcal{S}$. Proposition 5 implies that for every $\beta \in \mathcal{S} \setminus \{\alpha\}$, there exists $(\gamma, \delta) \in \binom{\mathcal{S}}{2} \setminus \{(\alpha, \beta)\}$ such that $\alpha + \beta = \gamma + \delta$. Observe that if two parities in the set $A := \{\beta, \gamma, \delta\}$ are chosen in $\mathcal{R}$, then α is identified with the third parity in A w.r.t. $\mathsf{span}\ \mathcal{R}$. Now, the expected number of $\beta \in \mathcal{S} \setminus \{\alpha\}$ for which the aforementioned identification occurs is seen by linearity of expectation to be $\Omega(kp^2)$, which is $\Omega(1)$ by the choice of p. The crux of the proof is in strengthening this bound on expectation to conclude that with constant probability, there exists at least one $\beta \in \mathcal{S} \setminus \{\alpha\}$ such that the above identification occurs. Theorem 14 follows by linearity of expectation over $\alpha \in \mathcal{S}$, and an invocation of Observation 10.

We prove Theorem 14 in Section 4.2. In Section 4.1 we prove a weaker statement that admits a simpler proof, and yet contains some key ideas that go into the proof of Theorem 14.

3.3 Good folding yields better PDTs

Assume that for any Boolean function f there exist $\alpha_1, \alpha_2 \in \mathcal{S}$ such that $|D_{\alpha_1+\alpha_2}| \geq k^\ell + 1$. This is a weaker assumption on f than it being (δ, ℓ)-folding. Observation 10 implies that $\mathcal{B}(f, \{0^n, \alpha_1 + \alpha_2\}) \leq k - k^\ell - 1 \leq k(1 - k^{-(1-\ell)})$. This suggests the following PDT for f. First the parity $\alpha_1 + \alpha_2$ is queried at the root. Observation 8 implies that the Fourier sparsity of f restricted to the affine subspace (of co-dimension 1) corresponding to each outcome of this query is at most $k(1 - k^{-(1-\ell)})$. Repeating this heuristic recursively for each leaf leads to a PDT of depth $O(k^{1-\ell} \log k)$.

We have now set up the backdrop to introduce our next contribution. In the preceding discussion we had assumed the following about any Boolean function f: there exists a pair in $\binom{\mathcal{S}}{2}$ with a large equivalence class. One implication of our next result is that if we instead assume that any Boolean function is $(\Omega(1), \ell)$-folding, the procedure SampleParity with p set to $1/\widetilde{\Theta}(k^{(1+\ell)/2})$ achieves a bucket complexity upper bound of $k(1 - \Omega(1))$ with high probability. By an argument analogous to the discussion in Section 3.1 (also see Corollary 16), this yields a PDT with depth $\widetilde{O}(k^{(1-\ell)/2})$. This is a quadratic improvement over the $\widetilde{O}(k^{1-\ell})$ bound discussed in the last paragraph. Besides, it can be seen to recover (up to a logarithmic factor) our first result by setting $\ell = 0$, since any Boolean function is $(1, 0)$-folding by Proposition 5.

▶ **Theorem 15.** *Let $0 \leq \ell \leq 1 - \Omega(1)$ and $\delta \in (0, 1]$. Let $f : \mathbb{F}_2^n \to \{-1, 1\}$ be (δ, ℓ)-folding with k sufficiently large. Set $p := \frac{4000 \log k}{\delta k^{(1+\ell)/2}}$ and let $\mathcal{R}$ be the random subset of $\mathcal{S}$ that* SampleParity*(f, p) returns. Then with probability at least $1 - \frac{1}{k}$, $\mathcal{B}(f, \mathsf{span}\ \mathcal{R}) \leq k - \frac{\delta k}{6}$.*

The proof of Theorem 15 proceeds along the lines of that of Theorem 14, but is more technical. A proof of it can be found in Section 5 of the full version of our paper [8].

This yields the following corollary.

▶ **Corollary 16.** *Let* $0 \leq \ell \leq 1 - \Omega(1)$ *and* $\delta = \Omega(1)$. *Suppose all Boolean functions* $f : \mathbb{F}_2^n \to \{-1, 1\}$ *with sufficiently large* k *are* (δ, ℓ)*-folding. Then,*

$$\mathrm{PDT}(f) = \widetilde{O}(k^{(1-\ell)/2}).$$

Proof. Fix any Boolean function $f : \mathbb{F}_2^n \to \{-1, 1\}$ with sufficiently large k. Let p and $\mathcal{R}$ be as in the statement of Theorem 15. Since δ is a constant, $p = \Theta\left(\frac{\log k}{k^{(1+\ell)/2}}\right)$. By Theorem 15, we have $\mathcal{B}(f, \mathsf{span}\ \mathcal{R}) \leq ck$, for some $c = (1 - \Omega(1))$, with probability strictly greater than 1/2. By a Chernoff bound $|\mathcal{R}| = \widetilde{O}(k^{(1-\ell)/2})$ with probability strictly greater than 1/2. Finally, by a union bound, we have that with non-zero probability the set $\mathcal{R}$ returned by $\textsc{SampleParity}(f, p)$ satisfies both $|\mathcal{R}| = \widetilde{O}(k^{(1-\ell)/2})$ and $\mathcal{B}(f, \mathsf{span}\ \mathcal{R}) \leq ck$, for some $c = (1 - \Omega(1))$. Choose such an $\mathcal{R}$ and consider the following PDT for f, whose construction closely follows the discussion in Section 3.1.

First, query all parities in $\mathcal{R}$. Now, let V be the affine subspace corresponding to an arbitrary leaf of this partial tree. By the properties of $\mathcal{R}$ and Observation 8, we have that the Fourier sparsity of $f|_V$ is at most ck. Repeat the same process inductively for each leaf. The depth of the resultant tree is at most $\widetilde{O}(k^{(1-\ell)/2} + (ck)^{(1-\ell)/2} + \cdots) = \widetilde{O}(k^{(1-\ell)/2})$. ◀

Corollary 16 naturally raises the question of whether all Boolean functions are $(\Omega(1), \Omega(1))$-folding.

▶ **Question 17.** *Do there exist constants* $\ell, \delta \in (0, 1]$ *such that every Boolean function* $f : \mathbb{F}_2^n \to \{-1, 1\}$ *is* (δ, ℓ)*-folding?*

An affirmative answer to Question 17 in conjunction with Corollary 16 and the discussion in Section 1 implies an upper bound on the communication complexity of XOR functions $F = f \circ \oplus$ that is polynomially smaller than the best known bound of $O(\sqrt{\mathsf{rank}(F)})$.

What is the largest ℓ for which all Boolean functions are $(\Omega(1), \ell)$-folding? The addressing function ADD_k (see Definition 3) is $(1, 1/2 - o(1))$-folding, and not $(\Omega(1), \ell)$-folding for any $\ell \geq \frac{1}{2}$ (see [8, Appendix B]). In light of this, we make the following conjecture.

▶ **Conjecture 18.** *There exists a constant* $\delta > 0$ *such that any Boolean function* $f : \mathbb{F}_2^n \to \{-1, 1\}$ *is* $(\delta, 1/2 - o(1))$*-folding.*

Assuming Conjecture 18, Corollary 16 would imply an upper bound of $\widetilde{O}(\mathsf{rank}^{1/4+o(1)}(F))$ on the communication complexity of XOR functions $F = f \circ \oplus$.

3.4 Boolean functions have non-trivial folding properties

Recall that Conjecture 18 states that any Boolean function is (δ, ℓ)-folding with $\delta = \Omega(1)$ and $\ell = 1/2 - o(1)$. Also recall from Proposition 5 that a necessary condition for a function to be Boolean valued is that it is (δ, ℓ)-folding with $\delta = 1$ and $\ell = 0$. We show in the Section 6 (see Theorem 27) that the conditions in Proposition 5 are not sufficient for a function to be Boolean valued.

To the best of our knowledge, it was not known prior to our work whether any better bound than this was known for Boolean functions (in terms of ℓ, for *any* non-zero δ). In particular, it was consistent with prior knowledge that there exist functions for which each equivalence class of $\binom{\mathcal{S}}{2}$ contains exactly 2 elements. We rule out this possibility, and our contribution is a step towards Conjecture 18.

▶ **Theorem 19.** *For any Boolean function $f : \mathbb{F}_2^n \to \{-1, 1\}$ with $k > 4$, and every $\alpha \in \mathcal{S}$, there exists $\beta \in \mathcal{S} \setminus \{\alpha\}$ such that $|D_{\alpha+\beta}| \geq 3$.*

In order to rule out the possibility mentioned above, it suffices to exhibit a single pair $(\alpha, \beta) \in \binom{\mathcal{S}}{2}$ with $|D_{\alpha+\beta}| \geq 3$. Theorem 19 further shows that *every* element $\alpha \in \mathcal{S}$ participates in such a pair.

Proof idea

We prove this via a series of arguments. Define $\mathcal{S}_+ := \left\{\alpha \in \mathcal{S} \;\middle|\; \widehat{f}(\alpha) > 0\right\}$ and $\mathcal{S}_- := \left\{\alpha \in \mathcal{S} \;\middle|\; \widehat{f}(\alpha) < 0\right\}$. We first show that if there exists $\alpha \in \mathcal{S}$ with $|D_{\alpha+\beta}| = 2$ for all $\beta \in \mathcal{S} \setminus \{\alpha\}$, then both of the following hold.

1. Either $|\mathcal{S}_+|$ or $|\mathcal{S}_-|$ is odd.
2. The function f must be plateaued.

The proofs use Equation (3). Next, we show that for plateaued Boolean functions, both $|\mathcal{S}_+|$ and $|\mathcal{S}_-|$ are even, yielding a contradiction in view of the first bullet above. This proof involves a careful analysis of the Fourier coefficients and crucially uses Observation 1 and Equation (2).

A natural question raised by Theorem 19 is whether there exists a Boolean function f and $\alpha \in \mathcal{S}$ such that there exists only one element $\beta \in \mathcal{S} \setminus \{\alpha\}$ with $|D_{\alpha+\beta}| \geq 3$. The following theorem answers this question in the positive, and sheds more light on the structure of such functions.

▶ **Theorem 20.**

1. *There exists a Boolean function $f : \mathbb{F}_2^n \to \{-1, 1\}$ and $(\alpha, \beta) \in \binom{\mathcal{S}}{2}$ such that $|D_{\alpha+\gamma}| = 2$ for all $\gamma \in \mathcal{S} \setminus \{\alpha, \beta\}$.*
2. *Let $f : \mathbb{F}_2^n \to \{-1, 1\}$ be any Boolean function. If there exists $(\alpha, \beta) \in \binom{\mathcal{S}}{2}$ such that $|D_{\alpha+\gamma}| = 2$ for all $\gamma \in \mathcal{S} \setminus \{\alpha, \beta\}$, then $|D_{\alpha+\beta}| = k/2$.*

The proof of Part 2 of Theorem 20 follows along the lines of the proof of Theorem 19. The proof of Part 1 of Theorem 20 constructs such a function by considering any Boolean function and applying a simple modification to it.

We prove Theorem 19 in Section 5 and Theorem 20 in [8, Section 6].

4 Proof of Theorem 14

In this Section, we prove our first result, Theorem 14.

4.1 Warm up: sampling $\widetilde{O}(k^{3/4})$ parities.

In this section we prove a quantitatively weaker statement. This admits a simpler proof and introduces many key ideas that go into our proof of Theorem 14.

▷ Claim 21. Let $p := \frac{2\sqrt{\log k}}{k^{1/4}}$, and let $\mathcal{R}$ be the set returned by SampleParity(f, p). Then

$$\Pr[\mathcal{B}(f, \operatorname{span} \mathcal{R}) \leq k/2] \geq 1 - \frac{1}{k^{1/3}}.$$

By a Chernoff bound, with high probability, $|\mathcal{R}| = \widetilde{O}(k^{3/4})$.

Proof. Fix any $\alpha \in \mathcal{S}$. By Proposition 5 we have that for each $\beta \in \mathcal{S}\setminus\{\alpha\}$, there exist $\beta_1, \beta_2 \in \mathcal{S} \setminus \{\alpha, \beta\}$ such that $\alpha + \beta + \beta_1 + \beta_2 = 0$. Fix any such β_1, β_2, and define $Q_\beta := \{\beta, \beta_1, \beta_2\}$. Note that the sets Q_β are not necessarily distinct. Define the multiset of unordered triples $\mathcal{F} := \{Q_\beta \mid \beta \in \mathcal{S} \setminus \{\alpha\}\}$. For each $\gamma \in \mathcal{S} \setminus \{\alpha\}$, define $\mathcal{D}_\gamma := \{\beta \in \mathcal{S} \setminus \{\alpha\} \mid \gamma \in Q_\beta\}$. We now show that with high probability there exists $F \in \mathcal{F}$ such that $|F \cap \mathcal{R}| \geq 2$. We consider two cases below.

Case 1: There exists $\gamma \in \mathcal{S} \setminus \{\alpha\}$ such that $|\mathcal{D}_\gamma| \geq k^{1/2}$.
Consider the multiset of unordered pairs $\mathcal{A} := \{Q_\beta \setminus \{\gamma\} \mid \beta \in \mathcal{D}_\gamma\}$. Each pair in $\mathcal{A}$ can repeat at most thrice. Hence there are at least $k^{1/2}/3$ distinct pairs in $\mathcal{A}$. Moreover the distinct pairs in $\mathcal{A}$ are disjoint. This can be inferred from the observation that the sum of the two elements in each pair in $\mathcal{A}$ equals $\alpha + \gamma$. Thus

$$\Pr\left[\forall A \in \mathcal{A}, A \not\subseteq \mathcal{R}\right] \leq (1-p^2)^{k^{1/2}/3} = \left(1 - \frac{4\log k}{k^{1/2}}\right)^{k^{1/2}/3} \leq \frac{1}{k^{4/3}}.$$

Case 2: For each $\gamma \in \mathcal{S} \setminus \{\alpha\}$, $|\mathcal{D}_\gamma| < k^{1/2}$.
In this case each triple in $\mathcal{F}$ has non-empty intersection with at most $3k^{1/2}$ sets in $\mathcal{F}$. Thus one can greedily obtain a collection $\mathcal{T}$ of at least $\frac{k-1}{3k^{1/2}}$ disjoint triples in $\mathcal{F}$.

$$\Pr\left[\forall T \in \mathcal{T}, |T \cap \mathcal{R}| < 2\right] \leq (1-p^2)^{\frac{k-1}{3k^{1/2}}} = \left(1 - \frac{4\log k}{k^{1/2}}\right)^{\frac{k-1}{3k^{1/2}}},$$

which is at most $\frac{1}{k^{4/3}}$ for sufficiently large k.

From the above two cases it follows that with probability at least $1 - \frac{1}{k^{4/3}}$, there exists a triple $F \in \mathcal{F}$ such that $|F \cap \mathcal{R}| \geq 2$. Assume existence of such a triple F, and let $\delta_1, \delta_2 \in F \cap \mathcal{R}$. Let $\delta := F \setminus \{\delta_1, \delta_2\}$. Since $\alpha + \delta_1 + \delta_2 + \delta = 0^n$, we have that $\alpha + \delta = \delta_1 + \delta_2 \in \mathsf{span}\ \mathcal{R}$, i.e., α is identified with δ with respect to $\mathsf{span}\ \mathcal{R}$. By a union bound over all $\alpha \in \mathcal{S}$ it follows that with probability at least $1 - \frac{1}{k^{1/3}}$, for every $\alpha \in \mathcal{S}$ there exists a $\delta \in \mathcal{S} \setminus \{\alpha\}$ such that α is identified with δ w.r.t. $\mathcal{R}$. The claim follows by Observation 10. ◁

4.2 Sampling $O(k^{1/2})$ parities

We now proceed to prove Theorem 14 by refining the ideas developed in Section 4.1. Recall that by a Chernoff bound, $|\mathcal{R}| = O(\sqrt{k})$ with high probability (where $\mathcal{R}$ is as in Theorem 14). We require the following inequality whose proof can be found in [8, Section 4.2].

▶ **Proposition 22.** *For any non-negative integer d, and $p \in [0, 1]$ be such that $pd \leq 1$. Then,*

$$(1-p)^d \leq 1 - \frac{1}{2}pd.$$

Proof of Theorem 14. For technical reasons we instead consider a two-step probabilistic procedure. Define $p' := \frac{1}{4k^{1/2}}$. Let $\mathcal{R}_1$ and $\mathcal{R}_2$ be the sets returned by two independent runs of SampleParity(f, p'), and let $\mathcal{R}' := \mathcal{R}_1 \cup \mathcal{R}_2$. Each $\alpha \in \mathcal{S}$ is independently included in $\mathcal{R}'$ with probability equal to $1 - (1-p')^2 < 2p' = p$. Hence it suffices to prove that there exists a constant $c \in (0, 1]$ such that $\mathbb{E}[\mathcal{B}(f, \mathsf{span}\ \mathcal{R}')] \leq ck$.

Fix any $\alpha \in \mathcal{S}$ and let Q_β and $\mathcal{F}$ be as in the proof of Claim 21. For $\gamma \in \mathcal{S} \setminus \{\alpha\}$, define $\widetilde{\mathsf{deg}}(\gamma) := |\{\beta \in \mathcal{S} \setminus \{\alpha\} \mid \gamma \in Q_\beta \setminus \{\beta\}\}|$. Clearly, $\mathbb{E}_{\gamma \sim \mathcal{S}\setminus\{\alpha\}}[\widetilde{\mathsf{deg}}(\gamma)] = 2$. Define $A := \left\{\gamma \in \mathcal{S} \setminus \{\alpha\} \;\middle|\; \widetilde{\mathsf{deg}}(\gamma) \geq 4k^{1/2}\right\}$. By Markov's inequality, $|A| \leq k^{1/2}/2$. Fix an ordering σ on $\mathcal{S} \setminus \{\alpha\}$ such that all elements of $\overline{A} := (\mathcal{S} \setminus \{\alpha\}) \setminus A$ appear before all elements of A.

Define $T := \{\beta \in \mathcal{S} \setminus \{\alpha\} \mid Q_\beta \setminus \{\beta\} \subseteq A\}$. Observe that the pairs $Q_\beta \setminus \{\beta\}$ for distinct $\beta \in \mathcal{S} \setminus \{\alpha\}$ are distinct. This can be inferred from the observation that the sum (with respect to coordinate-wise addition in $\mathbb{F}_2$) of the two elements of $Q_\beta \setminus \{\beta\}$ equals $\alpha + \beta$. This gives us the following bound on the size of T:

$$|T| \leq \binom{|A|}{2} \leq \frac{k}{8}. \tag{7}$$

Define $\overline{T} := (\mathcal{S} \setminus \{\alpha\}) \setminus T$. For each $\beta \in \overline{T}$, the first character (according to σ) in the pair $Q_\beta \setminus \{\beta\}$ is from $\overline{A}$. For each $\gamma \in \overline{A}$, define $\mathsf{d}(\gamma)$ to be the number of $\beta \in \overline{T}$ such that γ is the first element in $Q_\beta \setminus \{\beta\}$. By Equation (7) we have

$$\sum_{\gamma \in \overline{A}} \mathsf{d}(\gamma) = |\overline{T}| \geq k - 1 - \frac{k}{8} \geq \frac{2k}{3} \tag{8}$$

where the last inequality holds for sufficiently large k.

For $\gamma \in \overline{A}$, let $\mathcal{E}(\gamma)$ be the event that there exists $\beta \in \overline{T} \cap \mathcal{R}_1$ such that γ is the first element in $Q_\beta \setminus \{\beta\}$. We have

$$\Pr_{\mathcal{R}_1}[\mathcal{E}(\gamma)] = 1 - (1 - p')^{\mathsf{d}(\gamma)} \geq \frac{p' \cdot \mathsf{d}(\gamma)}{2}, \tag{9}$$

where the last inequality follows by Proposition 22. Here Proposition 22 is applicable since $\mathsf{d}(\gamma) \leq \widetilde{\mathsf{deg}}(\gamma) \leq 4k^{1/2}$ (since $\gamma \in \overline{A}$), and $p' = \frac{1}{4k^{1/2}}$. Define the random set $B := \left\{\gamma \in \overline{A} \;\middle|\; \mathcal{E}(\gamma) \text{ occurs}\right\}$. We have

$$\begin{aligned} \mathbb{E}_{\mathcal{R}_1}[|B|] &= \sum_{\gamma \in \overline{A}} \Pr_{\mathcal{R}_1}[\mathcal{E}(\gamma)] \geq \sum_{\gamma \in \overline{A}} \frac{p' \cdot \mathsf{d}(\gamma)}{2} && \text{by linearity of expectation and Equation (9)} \\ &\geq \frac{1}{2} \cdot \frac{1}{4k^{1/2}} \cdot \frac{2k}{3} \geq \frac{k^{1/2}}{12}. && \text{by Equation (8), and substituting the value of } p' \end{aligned}$$

Furthermore, the events $\mathcal{E}(\gamma)$ are independent. By a Chernoff bound, $\Pr_{\mathcal{R}_1}\left[|B| \geq \frac{k^{1/2}}{24}\right] \geq 0.9$. Now,

$$\begin{aligned} \Pr_{\mathcal{R}_1, \mathcal{R}_2}\left[B \cap \mathcal{R}_2 \neq \emptyset \;\middle|\; |B| \geq \frac{k^{1/2}}{24}\right] &\geq 1 - (1 - p')^{k^{1/2}/24} \geq 1 - e^{-p' \cdot \frac{k^{1/2}}{24}} = 1 - e^{-\frac{1}{96}} \\ &= c_1, \text{ say.} \end{aligned}$$

Thus, the probability of the event $\mathcal{E} := \left\{|B| \geq \frac{k^{1/2}}{24}\right\} \wedge \{B \cap \mathcal{R}_2 \neq \emptyset\}$ is at least $0.9c_1$. Suppose the event $\mathcal{E}$ occurs, and let $\gamma \in B \cap \mathcal{R}_2$. By the definitions of B and $\mathcal{E}(\gamma)$, there exists $\beta \in \overline{T} \cap \mathcal{R}_1$ such that γ is the first element of $Q_\beta \setminus \{\beta\}$. Let $\delta := Q_\beta \setminus \{\beta, \gamma\}$. Then, $\alpha + \delta = \beta + \gamma$. Since $\beta \in \mathcal{R}_1$ and $\gamma \in \mathcal{R}_2$, α is identified with δ with respect to $\mathsf{span}\ \mathcal{R}'$. In summary, we have shown that for any $\alpha \in \mathcal{S}$,

$$\Pr_{\mathcal{R}_1, \mathcal{R}_2}[\alpha \text{ is identified with some } \delta \in \mathcal{S} \setminus \{\alpha\} \text{ w.r.t. } \mathsf{span}\ \mathcal{R}'] \geq 0.9c_1.$$

By linearity of expectation,

$$\mathbb{E}_{\mathcal{R}_1, \mathcal{R}_2}[|\{\alpha \in \mathcal{S} \mid \alpha \text{ is identified with some } \delta \in \mathcal{S} \setminus \{\alpha\} \text{ w.r.t. } \mathsf{span}\ \mathcal{R}'\}|] \geq k \cdot 0.9c_1.$$

Observation 10 then implies

$$\mathbb{E}_{\mathcal{R}_1, \mathcal{R}_2}[\mathcal{B}(f, \mathsf{span}\ \mathcal{R}')] \leq k - \frac{k \cdot 0.9c_1}{2} = ck,$$

where $c = \left(1 - \frac{0.9c_1}{2}\right)$. ◀

5 Proof of Theorem 19

In this section we prove Theorem 19, which states that for any Boolean function $f : \mathbb{F}_2^n \to \{-1, 1\}$ and $\alpha \in \mathcal{S}$, there exists at least one $\beta \in \mathcal{S}$ with $|D_{\alpha+\beta}| \geq 3$.

We first recall and introduce some notation. Recall from Proposition 5 that for any Boolean function $f : \mathbb{F}_2^n \to \{-1, 1\}$ and every $\gamma \in (\mathcal{S} + \mathcal{S}) \setminus \{0^n\}$, we have $|D_\gamma| \geq 2$. For any γ with $|D_\gamma| > 2$, we say that γ is a *non-trivial folding direction.* Hence, Theorem 19 can be rephrased to say that for any Boolean function $f : \mathbb{F}_2^n \to \{-1, 1\}$, every element $\alpha \in \mathcal{S}$ must participate in at least one non-trivial folding direction. For any Boolean function $f : \mathbb{F}_2^n \to \{-1, 1\}$, define $\mathcal{S}_+ := \left\{\alpha \in \mathcal{S} \;\middle|\; \widehat{f}(\alpha) > 0\right\}$, and $\mathcal{S}_- := \left\{\alpha \in \mathcal{S} \;\middle|\; \widehat{f}(\alpha) < 0\right\}$. For any set S, we use the notation $\binom{S}{3}$ to denote the set of all subsets of S of size exactly 3. We abuse notation and denote a generic element of $\binom{S}{3}$ as (a, b, c) rather than $\{a, b, c\}$.

We require the following proposition. For a proof, refer to [8, Section 6].

▶ **Proposition 23.** *Let $f : \mathbb{F}_2^n \to \{-1, 1\}$ be a Boolean function with Fourier support $\mathcal{S}$ with $k = |\mathcal{S}| \geq 2$. Let α, β be two distinct parities in $\mathcal{S}$. Then, there exists a Boolean function $g : \mathbb{F}_2^n \to \{-1, 1\}$ with Fourier support $\mathcal{S}$ and $\widehat{g}(\alpha) > 0, \widehat{g}(\beta) > 0$.*

We next state a preliminary claim.

▷ Claim 24. Let $f : \mathbb{F}_2^n \to \{-1, 1\}$ be any Boolean function. Suppose there exists $\alpha \in \mathcal{S}$ such that $|D_{\alpha+\beta}| = 2$ for all $\beta \in \mathcal{S} \setminus \{\alpha\}$. Then, either $|\mathcal{S}_+|$ is odd or $|\mathcal{S}_-|$ is odd.

Proof. Fix any set $\alpha \in \mathcal{S}$ such that $|D_{\alpha+\beta}| = 2$ for all $\beta \in \mathcal{S} \setminus \{\alpha\}$. Assume $\alpha \in \mathcal{S}_+$ (else run this argument with $\mathcal{S}_+$ and $\mathcal{S}_-$ interchanged). Consider the set of unordered triples

$$T = \left\{(\beta, \gamma, \delta) \in \binom{\mathcal{S} \setminus \{\alpha\}}{3} \;\middle|\; \alpha + \beta + \gamma + \delta = 0^n\right\}.$$

Let T_+ denote the set of triples in T that contain at least one element $\beta \in \mathcal{S}_+$, i.e.,

$$T_+ := \left\{(\beta, \gamma, \delta) \in T \;\middle|\; \text{at least one of } \widehat{f}(\beta), \widehat{f}(\gamma), \widehat{f}(\delta) \text{ is positive}\right\}.$$

Since $|D_{\alpha+\beta}| = 2$ for all $\beta \in \mathcal{S} \setminus \{\alpha\}$, this implies that any $\beta \in \mathcal{S}$ (in particular any $\beta \in \mathcal{S}_+$) appears in exactly one triple. For any $\beta \in \mathcal{S}_+$, say this triple is $(\beta, \beta_1, \beta_2)$. Equation (3) implies that

$$\widehat{f}(\alpha)\widehat{f}(\beta) + \widehat{f}(\beta_1)\widehat{f}(\beta_2) = 0.$$

Since α and β are both in $\mathcal{S}_+$, exactly one of β_1, β_2 is in $\mathcal{S}_+$ and the other is in $\mathcal{S}_-$.

Thus each triple in T_+ contains exactly two elements of $\mathcal{S}_+$, and none of these elements appears in any other triple. Moreover each element of $\mathcal{S}_+$ appears in some triple in T_+. Accounting for α being in $\mathcal{S}_+$, we conclude that if $|T_+| = t$, then $|\mathcal{S}_+| = 2t + 1$, which is odd. ◁

We state another claim that we require.

▷ Claim 25. Let $f : \mathbb{F}_2^n \to \{-1, 1\}$ be any Boolean function. If there exists $\alpha \in \mathcal{S}$ such that $|D_{\alpha+\beta}| = 2$ for all $\beta \in \mathcal{S} \setminus \{\alpha\}$, then f is plateaued.

Proof. Fix any $\alpha \in \mathcal{S}$ such that $|D_{\alpha+\beta}| = 2$ for all $\beta \in \mathcal{S} \setminus \{\alpha\}$. Towards a contradiction, suppose f is not plateaued. This implies existence of $\gamma \in \mathcal{S}$ such that $|\widehat{f}(\alpha)| \neq |\widehat{f}(\gamma)|$. Proposition 5 implies existence of $\mu, \nu \in \mathcal{S}$ be such that $\alpha + \gamma = \mu + \nu$. We also have that

$$\alpha + \nu = \mu + \gamma, \qquad \alpha + \mu = \gamma + \nu.$$

Arrange α, γ, μ and ν in non-increasing order of the absolute values of their Fourier coefficients. Let the resultant sequence be $\delta_1, \delta_2, \delta_3, \delta_4$. Thus,

$$|\widehat{f}(\delta_1)| \geq |\widehat{f}(\delta_2)| \geq |\widehat{f}(\delta_3)| \geq |\widehat{f}(\delta_4)|.$$

Since $|\widehat{f}(\alpha)| \neq |\widehat{f}(\gamma)|$, at least one of these inequalities must be strict, which in particular implies that $|\widehat{f}(\delta_1)||\widehat{f}(\delta_2)| > |\widehat{f}(\delta_3)||\widehat{f}(\delta_4)|$. Now by the hypothesis, for all $1 \leq i < j \leq 4$, and $\{k, m\} := \{1,2,3,4\} \setminus \{i, j\}$ we have that $|D_{\delta_i+\delta_j}| = |D_{\delta_k+\delta_m}| = 2$. Thus, by Equation (3) we have that $\widehat{f}(\delta_1)\widehat{f}(\delta_2) = -\widehat{f}(\delta_3)\widehat{f}(\delta_4)$, implying that $|\widehat{f}(\delta_1)||\widehat{f}(\delta_2)| = |\widehat{f}(\delta_3)||\widehat{f}(\delta_4)|$, which is a contradiction. ◁

The next claim shows that Theorem 19 holds true if f is a plateaued function.

▷ Claim 26. Let $f : \mathbb{F}_2^n \to \{-1, 1\}$ be any plateaued Boolean function with $k > 4$. Then, for any $\alpha \in \mathcal{S}$, there exists $\beta \in \mathcal{S} \setminus \{\alpha\}$ such that $|D_{\alpha+\beta}| \geq 3$.

Proof. Towards a contradiction, let $\alpha \in \mathcal{S}$ be such that $|D_{\alpha+\beta}| = 2$ for all $\beta \in \mathcal{S} \setminus \{\alpha\}$. Let $s = |\mathcal{S}_+|$ and $t = |\mathcal{S}_-|$. We now prove that s and t must both be even.

Since f is plateaued, Equation (2) implies that $|\widehat{f}(\gamma)| = 1/\sqrt{k}$ for all $\gamma \in \mathcal{S}$. By Observation 1 we know that $1/\sqrt{k} = c/2^n$ for some $c \in \mathbb{Z}$. This implies that $k = 2^{2n}/c^2$. Since k is an integer, c must be a power of 2, and hence $k = 2^{2h}$ for some $h > 1$ (since we assumed $k > 4$).

Assume $f(0^n) = 1$ (else run the same argument with f replaced by $-f$). This implies

$$\sum_{\gamma \in \mathcal{S}_+} |\widehat{f}(\gamma)| - \sum_{\delta \in \mathcal{S}_-} |\widehat{f}(\delta)| = 1.$$

That is, $(s-t)/\sqrt{k} = 1$. Since $s + t = k$, this implies $s = \frac{k}{2} + \frac{\sqrt{k}}{2}$ and $t = \frac{k}{2} - \frac{\sqrt{k}}{2}$. Since $k = 2^{2h}$ for some $h > 1$ (since we assumed $k > 4$), s and t are both even. This is a contradiction in view of Claim 24. ◁

We next use Claim 25 to remove the assumption of f being plateaued in the previous claim, which proves Theorem 19.

Proof of Theorem 19. Towards a contradiction, suppose there exists $\alpha \in \mathcal{S}$ such that $|D_{\alpha+\beta}| = 2$ for all $\beta \in \mathcal{S} \setminus \{\alpha\}$. Claim 25 implies that f must be plateaued. Next, Claim 26 implies that there must exist $\gamma \in \mathcal{S}$ such that $|D_{\alpha+\gamma}| \geq 3$, which is a contradiction. ◀

6 Ruling out sufficiency of Proposition 5

In this section, we prove that the conditions in Proposition 5 are not sufficient for a function to be Boolean. To the best of our knowledge, ours is the first work to show this.

▶ **Theorem 27.** *There exists a set $\mathcal{S} \subseteq \mathbb{F}_2^n$ such that $|D_{\alpha+\beta}| \geq 2$ for all $(\alpha, \beta) \in \binom{\mathcal{S}}{2}$, but $\mathcal{S}$ is not the Fourier support of any Boolean function $f : \mathbb{F}_2^n \to \{-1, 1\}$.*

For sets $A, B \subseteq [n]$, let $A \triangle B$ denote the symmetric difference of the sets A and B. For $x \in \mathbb{R} \setminus \{0\}$, define $\mathrm{sgn}(x) := -1$ if $x < 0$, and $\mathrm{sgn}(x) := 1$ if $x > 0$.

Proof. For the purpose of this proof, we require the natural equivalence between elements of $\mathbb{F}_2^n$ and subsets of $[n]$. Under this equivalence, the sum of two elements in $\mathbb{F}_2^n$ corresponds to the symmetric difference of the corresponding sets in $[n]$. The following is a property of symmetric difference. For any sets $A, B, C, D \subseteq [n]$,

$$A \triangle B = C \triangle D \iff A \triangle C = B \triangle D. \tag{10}$$

Hence it suffices to exhibit a collection $\mathcal{S}$ of subsets of $[n]$ such that for all $(S,T) \in \binom{\mathcal{S}}{2}$, there exist $(U,V) \neq (S,T) \in \binom{\mathcal{S}}{2}$ with $S \triangle T = U \triangle V$, and $\mathcal{S}$ is not the Fourier support of any Boolean function $f : \mathbb{F}_2^n \to \{-1,1\}$. To this end, consider the set

$$\mathcal{S} = \{\{1\}, \ldots, \{n\}, \{1,2,n\}, \ldots, \{1,n-1,n\}\}.$$

Below we list out all equivalence classes of $\binom{\mathcal{S}}{2}$. For any distinct $i, j \in \{2,3,\ldots,n-1\}$ we have $\{i\} \triangle \{j\} = \{1,i,n\} \triangle \{1,j,n\}$. Thus

$$D_{\{i\}\triangle\{j\}} = \{(\{i\},\{j\}), (\{1,i,n\},\{1,j,n\})\} \quad \forall i,j \in \{2,3,\ldots,n-1\}. \tag{11}$$

For any $i \in \{2,3,\ldots,n-1\}$ we have

$$\begin{aligned}
\{1\} \triangle \{i\} &= \{n\} \triangle \{1,i,n\}, \\
\{n\} \triangle \{i\} &= \{1\} \triangle \{1,i,n\}.
\end{aligned}$$

We also have

$$\{1\} \triangle \{n\} = \{i\} \triangle \{1,i,n\} \quad \text{for all } i \in \{2,3,\ldots,n-1\}.$$

Along with Equation (10), these establish the fact that $|D_{\alpha+\beta}| \geq 2$ for all $(\alpha,\beta) \in \binom{\mathcal{S}}{2}$. We now provide a proof of the fact that $\mathcal{S}$ cannot be the Fourier support of any Boolean function. Consider the following six sets.

$$S_1 = \{2\}, S_2 = \{3\}, S_3 = \{4\}, S_4 = \{1,2,n\}, S_5 = \{1,3,n\}, S_6 = \{1,4,n\}.$$

If $\mathcal{S}$ is the support of a Boolean function, then Equation (3) holds true. Equation (11) then implies

$$\begin{aligned}
\widehat{f}(S_1)\widehat{f}(S_2) + \widehat{f}(S_4)\widehat{f}(S_5) &= 0, \\
\widehat{f}(S_1)\widehat{f}(S_3) + \widehat{f}(S_4)\widehat{f}(S_6) &= 0, \\
\widehat{f}(S_2)\widehat{f}(S_3) + \widehat{f}(S_5)\widehat{f}(S_6) &= 0.
\end{aligned}$$

Let $s_i = \operatorname{sgn}(\widehat{f}(S_i))$ for $i \in [6]$. Thus,

$$\begin{aligned}
s_1 s_2 &= -s_4 s_5 \\
s_1 s_3 &= -s_4 s_6 \\
s_2 s_3 &= -s_5 s_6.
\end{aligned}$$

Multiplying out the left hand sides and right hand sides of the above, we obtain $1 = -1$, which is a contradiction. Hence $\mathcal{S}$ cannot be the support of any Boolean function. ◀

References

1 Anurag Anshu, Naresh Goud Boddu, and Dave Touchette. Quantum log-approximate-rank conjecture is also false. In *60th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2019, Baltimore, Maryland, USA, November 9-12, 2019*, pages 982–994, 2019. doi:10.1109/FOCS.2019.00063.

2 Anna Bernasconi and Bruno Codenotti. Spectral analysis of boolean functions as a graph eigenvalue problem. *IEEE Trans. Computers*, 48(3):345–351, 1999. doi:10.1109/12.755000.

3 Arkadev Chattopadhyay, Nikhil S. Mande, and Suhail Sherif. The log-approximate-rank conjecture is false. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019, Phoenix, AZ, USA, June 23-26, 2019*, pages 42–53, 2019. doi:10.1145/3313276.3316353.

4 Hamed Hatami, Kaave Hosseini, and Shachar Lovett. Structure of protocols for XOR functions. *SIAM J. Comput.*, 47(1):208–217, 2018. doi:10.1137/17M1136869.

5 Troy Lee and Adi Shraibman. Lower bounds in communication complexity. *Foundations and Trends in Theoretical Computer Science*, 3(4):263–398, 2009. doi:10.1561/0400000040.

6 László Lovász and Michael E. Saks. Lattices, möbius functions and communication complexity. In *29th Annual Symposium on Foundations of Computer Science, White Plains, New York, USA, 24-26 October 1988*, pages 81–90, 1988. doi:10.1109/SFCS.1988.21924.

7 Shachar Lovett. Communication is bounded by root of rank. *J. ACM*, 63(1):1:1–1:9, 2016. doi:10.1145/2724704.

8 Nikhil S. Mande and Swagato Sanyal. On parity decision trees for Fourier-sparse Boolean functions. *CoRR*, abs/2008.00266, 2020. arXiv:2008.00266.

9 Ashley Montanaro and Tobias Osborne. On the communication complexity of XOR functions. *CoRR*, abs/0909.3392, 2009. URL: http://arxiv.org/abs/0909.3392.

10 Swagato Sanyal. Fourier sparsity and dimension. *Theory of Computing*, 15(1):1–13, 2019.

11 Amir Shpilka, Avishay Tal, and Ben lee Volk. On the structure of boolean functions with small spectral norm. *Comput. Complex.*, 26(1):229–273, 2017. doi:10.1007/s00037-015-0110-y.

12 Makrand Sinha and Ronald de Wolf. Exponential separation between quantum communication and logarithm of approximate rank. In *60th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2019, Baltimore, Maryland, USA, November 9-12, 2019*, pages 966–981, 2019. doi:10.1109/FOCS.2019.00062.

13 Robert C Titsworth. *Correlation properties of cyclic sequences.* PhD thesis, California Institute of Technology, 1962.

14 Hing Yin Tsang, Chung Hoi Wong, Ning Xie, and Shengyu Zhang. Fourier sparsity, spectral norm, and the log-rank conjecture. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013, 26-29 October, 2013, Berkeley, CA, USA*, pages 658–667, 2013. doi:10.1109/FOCS.2013.76.