On Parity Decision Trees for Fourier-sparse Boolean Functions

NIKHIL S. MANDE, University of Liverpool, Liverpool, United Kingdom
SWAGATO SANYAL, Indian Institute of Technology Kharagpur, Kharagpur, India

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 previously best-known upper bound on the parity decision tree complexity of Boolean functions (Tsang, Wong, Xie, and Zhang, FOCS'13). 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 $\{\alpha, \beta\}$ of parities in $\mathcal{S}$ such that $\alpha + \beta = \gamma + \delta$. We pose a question in this regard that, if answered in the positive, 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'13, 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 $\{\alpha, \beta\}$ in $\binom{\mathcal{S}}{2}$, there exists another pair $\{\gamma, \delta\}$ 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.

CCS Concepts: • **Theory of computation → Oracles and decision trees;**

Additional Key Words and Phrases: Parity decision trees, log-rank conjecture, analysis of Boolean functions, communication complexity

S. Sanyal is supported by an ISIRD Grant from Sponsored Research and Industrial Consultancy, IIT Kharagpur.
Authors' addresses: N. S. Mande, University of Liverpool, Ashton Building, Ashton Street, Liverpool, United Kingdom, L69 3BX; e-mail: Nikhil.Mande@liverpool.ac.uk; S. Sanyal, Indian Institute of Technology Kharagpur, Department of Computer Science and Engineering, Kharagpur, India; e-mail: swagato@cse.iitkgp.ac.in.

ACM 1942-3454/2024/06-ART9
https://doi.org/10.1145/3647629

ACM Reference Format:
Nikhil S. Mande and Swagato Sanyal. 2024. On Parity Decision Trees for Fourier-sparse Boolean Functions. *ACM Trans. Comput. Theory* 16, 2, Article 9 (June 2024), 26 pages. https://doi.org/10.1145/3647629

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 $F(x, y) = f(x \oplus y)$ for each $(x, y) \in \mathbb{F}_2^n \times \mathbb{F}_2^n$. 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, which is the number of non-zero Fourier coefficients of f (henceforth referred to as k. See Section 2 for a formal introduction to Fourier analysis of Boolean functions) [2]. The natural randomized analogue of the log-rank conjecture is the log-approximate-rank conjecture [5], which was recently refuted by Chattopadhyay et al. [3]. The quantum analogue of the log-rank conjecture was subsequently also refuted by Sinha and de Wolf [11] and Anshu et al. [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 [8, 10, 13] is to design a small-depth **parity decision tree (PDT)** for f and have 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 (PDT(·))** of f thus places an asymptotic upper bound on the communication complexity of F. The work of Hatami et al. [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 $O(\sqrt{\mathrm{rank}}(F) \log \mathrm{rank}(F))$. In particular, this implies that that the deterministic communication complexity of $F = f \circ \oplus$ is $O(\sqrt{k} \log k)$. Tsang et al. [13] showed that $\mathrm{PDT}(f) = O(\sqrt{k})$ (a quantitatively weaker bound was shown in a simultaneous and independent work of Shpilka et al. [10]). 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 [9] 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 2.5). While we

show that it is not a sufficient condition (see Theorem A.1 in Appendix A), it does enable us to prove the above results. In these proofs, we use Proposition 2.5 in conjunction with probabilistic and combinatorial arguments. Finally, we make progress towards establishing the folding property (see Section 3.4). Here, 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 Article

In Section 2, we review some preliminaries and introduce the notation that we use in this article. 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 4–6.

2 NOTATION AND PRELIMINARIES

All logarithms in this article 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 $\{1, 2, \ldots, n\}$. When we use the notation $\mathbb{E}_{x\in X}[\cdot]$, the underlying distribution corresponds to x being sampled uniformly at random from X. 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 $A, B \subseteq \mathbb{F}_2^n$, the set $A + B$ 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 span Γ to be the set of all $\mathbb{F}_2$-linear combinations of vectors in Γ, i.e., span $\Gamma = \{\sum_{\gamma\in\Gamma} c_\gamma \cdot \gamma \mid c_\gamma \in \mathbb{F}_2 \text{ for } \gamma \in \Gamma\}$.

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 $\{\widehat{f}(\alpha) \mid \alpha \in \mathbb{F}_2^n\}$ 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 $f|_A : A \to \{-1, 1\}$ by $f|_A(x) = f(x)$ for all $x \in A$. In other words, $f|_A$ denotes the restriction of f to A.

Throughout this article, for any Boolean function $f : \mathbb{F}_2^n \to \{-1, 1\}$, we denote by $\mathcal{S}$ the Fourier support of f, i.e., $\mathcal{S} = \{\alpha \in \mathbb{F}_2^n \mid \widehat{f}(\alpha) \neq 0\}$. 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 2.1. *Let $f : \mathbb{F}_2^n \to \{-1,1\}$ be any Boolean function. Then, for all $\alpha \in \mathbb{F}_2^n$, the coefficient $\widehat{f}(\alpha)$ is an integral multiple of $1/2^n$.*

We next define plateaued functions.

Definition 2.2 (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 2.3 (Addressing Function). Let k be an even power of 2. The addressing function $\mathrm{ADD}_k : \mathbb{F}_2^{\frac{1}{2}\log k + \sqrt{k}} \to \{-1,1\}$ is defined as

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

where $x \in \mathbb{F}_2^{\frac{1}{2}\log k}$, $y_i \in \mathbb{F}_2$ for $i = 1, \ldots, \sqrt{k}$, and $\mathrm{int}(x)$ is the unique integer in $\{1, \ldots, \sqrt{k}\}$ 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 2.4. 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 $\{\alpha_1, \alpha_2\}$ and $\{\alpha_3, \alpha_4\}$ *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 O_γ the equivalence class of pairs that fold in the direction γ, i.e.,

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

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

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 [12]. 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 2.5. *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, $|O_{\alpha+\beta}| \geq 2$.*

We require the following proposition:

Proposition 2.6. *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 $\widehat{g}(\alpha) > 0, \widehat{g}(\beta) > 0$.*

Proof. If $\widehat{f}(\alpha)\widehat{f}(\beta) > 0$, then the proposition follows by setting $g = f$ if $\widehat{f}(\alpha) > 0$, and $g = -f$ otherwise. Hence, we may assume that $\widehat{f}(\alpha)\widehat{f}(\beta) < 0$. Fix any $y \in \mathbb{F}_2^n$ such that $\chi_{\alpha+\beta}(y) = -1$, i.e., $\chi_\alpha(y) = -\chi_\beta(y)$. Define $h(x) := f(x+y)$. Then,

$$h(x) = \sum_{\delta\in\mathbb{F}_2^n} \widehat{f}(\delta)\chi_\delta(x+y) = \sum_{\delta\in\mathbb{F}_2^n} \widehat{f}(\delta)\chi_\delta(y)\cdot\chi_\delta(x),$$

giving us $\widehat{h}(\delta) = \widehat{f}(\delta)\cdot\chi_\delta(y)$ for all $\delta \in \mathbb{F}_2^n$. In particular, $\widehat{h}(\alpha) = \widehat{f}(\alpha)\cdot\chi_\alpha(y)$ and $\widehat{h}(\beta) = \widehat{f}(\beta)\cdot\chi_\beta(y)$. By the choice of y and the assumption $\widehat{f}(\alpha)\widehat{f}(\beta) < 0$, we have that $\widehat{h}(\alpha)\widehat{h}(\beta) = \widehat{f}(\alpha)\widehat{f}(\beta)\chi_{\alpha+\beta}(y) > 0$. Then, as described before, g can be taken to be h or $-h$, depending on the sign of $\widehat{h}(\alpha)$. Finally, note that $|\widehat{g}(\delta)| = |\widehat{f}(\delta)|$ for each $\delta \in \mathbb{F}_2^n$, implying that the Fourier support of g is $\mathcal{S}$. □

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 2.7. 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 $a_1,\ldots,a_t \in \mathbb{F}_2$ 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 the set $\mathcal{G} := \mathrm{span}\,\Gamma$, and let $\mathcal{C} := \{\mathcal{G}+\beta \mid \beta \in \mathbb{F}_2^n,\ (\mathcal{G}+\beta)\cap\mathcal{S} \neq \emptyset\}$ 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 $\gamma_1(x), \ldots, \gamma_t(x)$. We 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 $\{x \in \mathbb{F}_2^n \mid \gamma_1(x) = b_1, \ldots, \gamma_t(x) = b_t\}$. By substitution into Equation (5), we get

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

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

Example. Let $n = 2$ and f be the AND function of two variables x_1 and x_2, which takes the value -1 if $x_1 = x_2 = 1$, and 1 otherwise. Define $x := (x_1, x_2)$. Define $y_i := (-1)^{x_i}$. Thus, $y_i = -1$ if $x_i = 1$ and $y_i = 1$ if $x_i = 0$. Now, the characters of f are precisely the functions $1, y_1, y_2$, and y_1y_2. The Fourier transform of f is as follows:

$$f(x_1, x_2) = \frac{1}{2} + \frac{1}{2}y_1 + \frac{1}{2}y_2 - \frac{1}{2}y_1y_2.$$

Now define $\gamma_1 := (1, 0)$ (this corresponds to the character $\chi_{\gamma_1} = y_1$) and $\Gamma := \{\gamma_1\}$. Thus, $\mathcal{G} := \operatorname{span}\Gamma = \{(0, 0), (1, 0)\}$. The two cosets of $\mathcal{G}$ are $C_0 = \mathcal{G}$ and $C_1 = \{(0, 1), (1, 1)\}$. Define $\alpha(C_0) := (0, 0)$ and $\alpha(C_1) := (0, 1)$ so $C_0 = \mathcal{G} + \alpha(C_0)$ and $C_1 = \mathcal{G} + \alpha(C_1)$. Thus, we have

$$\chi_{\gamma_1}(x) = y_1, \qquad \chi_{\alpha(C_0)}(x) = 1, \qquad \chi_{\alpha(C_1)}(x) = y_2.$$

We define

$$P_{C_0}(y_1) = P_{C_0}(\gamma_1(x)) := \sum_{\gamma \in \mathcal{G}} \widehat{f}(\alpha(C_0) + \gamma)\chi_\gamma(x_1, x_2) = \frac{1}{2} + \frac{1}{2}y_1,$$

and

$$P_{C_1}(y_1) = P_{C_1}(\gamma_1(x)) := \sum_{\gamma \in \mathcal{G}} \widehat{f}(\alpha(C_1) + \gamma)\chi_\gamma(x_1, x_2) = \frac{1}{2} - \frac{1}{2}y_1.$$

The Fourier transform of f can now be expressed as follows:

$$\begin{aligned} f(x_1, x_2) &= P_{C_0}(\gamma_1(x)) + P_{C_1}(\gamma_1(x))y_2 \\ &= P_{C_0}(\gamma_1(x)) \cdot \chi_{\alpha(C_0)}(x) + P_{C_1}(\gamma_1(x)) \cdot \chi_{\alpha(C_1)}(x). \end{aligned} \tag{7}$$

Now for $b \in \mathbb{F}_2$, consider the affine subspace $H_b = \{x = (x_1, x_2) \in \mathbb{F}_2^2 \mid \gamma_1(x) = x_1 = b\}$. By substitution into Equation (7), we have that

$$f\mid_{H_b}(x) = P_{C_0}(b) \cdot \chi_{\alpha(C_0)}(x) + P_{C_1}(b) \cdot \chi_{\alpha(C_1)}(x).$$

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 2.8 (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 $\mathcal{G} := \operatorname{span}\Gamma$, and let $\mathcal{C}$ denote the set of cosets of $\mathcal{G}$ that

have non-empty intersection with $\mathcal{S}$, that is, $C := \{\mathcal{G} + \beta \mid \beta \in \mathbb{F}_2^n,\ (\mathcal{G} + \beta) \cap \mathcal{S} \neq \emptyset\}$. Define the *bucket complexity* of f with respect to $\mathcal{G}$, denoted $\mathcal{B}(f, \mathcal{G})$, as

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

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

Observation 2.9. *Let Γ and $\mathcal{G}$ be as in Definition 2.8. 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 $f\ |_V$. Then, $k' \leq \mathcal{B}(f, \mathcal{G})$.*

Definition 2.10 (Identification of Characters). For $f, \mathcal{G}$, and C as in Definition 2.8 and any $\beta, \delta \in \mathcal{S}$, we say that *$\beta$ 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 C.

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

Observation 2.11. *Let $f, \mathcal{G}$, and C be as in Definition 2.8. 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 C that contain at least one element from $\overline{L}$. Next, each coset in 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 $|C| \leq (k - h) + \frac{h}{2} = k - \frac{h}{2}$. □

2.2 Folding Properties of Boolean Functions

Definition 2.12. Let $f : \mathbb{F}_2^n \to \{-1, 1\}$ be any Boolean function. We say that f is *(δ, ℓ)-folding* if

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

Proposition 2.5 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 $|O_{\alpha+\beta}|$ is large for many $\beta \in \mathcal{S} \setminus \{\alpha\}$.

Claim 2.13. *Let $f : \mathbb{F}_2^n \to \{-1, 1\}$ be (δ, ℓ)-folding and have Fourier sparsity $k \geq 6$. Define*

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

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

Proof. For each $\alpha \in \mathcal{S}$, define $t(\alpha) := |\{\beta \in \mathcal{S} \setminus \{\alpha\} \mid |O_{\alpha+\beta}| \geq k^\ell + 1\}|$. 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 article. 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 3.2–3.4.

3.1 Low Bucket Complexity Implies Shallow PDTs

A parity function (i.e., a character χ_γ for some $\gamma \in \mathbb{F}_2^n$) or the complement of a parity function has a trivial PDT of complexity 1, namely, the one that queries $\gamma(x)$ and outputs the answer. Thus, from now on in this section, we assume that f is not a parity function or the complement of a parity function. This implies that the absolute value of each Fourier coefficient of f is strictly less than 1.

The following lemma follows from Reference [13, Lemma 28] and Equation (4).

Lemma 3.1 (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 3.1, where $t = O(\sqrt{k})$. Define $\mathcal{G} := \text{span}\,\{\gamma_1, \ldots, \gamma_t\}$. We next observe that $\mathcal{B}(f, \mathcal{G}) \leq k/2$. To see this, note that, since $f\ |_V$ is constant, the constant term of the Fourier transform of $f\ |_V$ is ± 1, and each non-constant term is 0. We thus have from Equation (6) that for each coset $C \in \mathcal{C}$ and any $(b_1, \ldots, b_t) \in \mathbb{F}_2^t$,

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

This implies that each $P_C(\cdot)$ has at least two terms, i.e., each $\beta \in \mathcal{S}$ is identified with some other $\delta \in \mathcal{S}$ with respect to $\mathcal{G}$. To see this, let $C \in \mathcal{C}$ and P_C contains a term with coefficient a. Since a is a Fourier coefficient of f, we have $|a| < 1$ by our assumption about f not being a constant, parity, or complement of a parity. Thus, for P_C to evaluate to 0 or ± 1 there needs to be at least one other term in P_C. Observation 2.11 implies that $\mathcal{B}(f, \mathcal{G}) \leq k/2$. Observation 2.9 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 $O(\sqrt{k} + \sqrt{\frac{k}{2}} + \sqrt{\frac{k}{2^2}} + \cdots) = O(\sqrt{k})$.

3.2 A Random Set of Parities Achieves Low Bucket Complexity

Tsang et al. proved Lemma 3.1 by an iterative procedure in each step of which a single parity is carefully chosen. We show in this article 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 SampleParity(f, p) described in Algorithm 1. Our first result shows that the set $\mathcal{R}$ returned by SampleParity$(f, \frac{1}{\Theta(\sqrt{k})})$ satisfies $\mathcal{B}(f, \text{span}\,\mathcal{R}) \leq k(1 - \Omega(1))$ with high probability.

ALGORITHM 1: SampleParity procedure

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

Theorem 3.2. *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 universal constant $c \in [0, 1)$ such that*

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

With high probability, we have $|\mathcal{R}| = O(\sqrt{k})$. By an argument analogous to the discussion in the previous section, Theorem 3.2 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 3.1, 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 2.5 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. 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 3.2 follows by linearity of expectation over $\alpha \in \mathcal{S}$ and an invocation of Observation 2.11.

We prove Theorem 3.2 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 3.2.

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 $|O_{\alpha_1+\alpha_2}| \geq k^\ell + 1$. This is a weaker assumption on f than it being (δ, ℓ)-folding (see Definition 2.12). Observation 2.11 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 2.9 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, then 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 3.4), 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 2.5.

Theorem 3.3. *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 $p := \frac{4{,}000 \log k}{\delta k^{(1+\ell)/2}}$ and let $\mathcal{R}$ be the random subset of $\mathcal{S}$ that* SampleParity(f, p) *returns. Then, with probability at least $1 - \frac{1}{k}$, $\mathcal{B}(f, \operatorname{span} \mathcal{R}) \leq k - \frac{\delta k}{6}$.*

The proof of Theorem 3.3 proceeds along the lines of that of Theorem 3.2 but is more technical. We prove it in Section 5.

This yields the following corollary:

Corollary 3.4. *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 3.3. Since δ is a constant, $p = \Theta(\frac{\log k}{k^{(1+\ell)/2}})$. By Theorem 3.3, we have $\mathcal{B}(f, \mathrm{span}\ \mathcal{R}) \leq ck$, 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 SampleParity(f, p) satisfies both $|\mathcal{R}| = \widetilde{O}(k^{(1-\ell)/2})$ and $\mathcal{B}(f, \mathrm{span}\ \mathcal{R}) \leq ck$, 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 2.9, 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 resulting tree is at most $\widetilde{O}(k^{(1-\ell)/2} + (ck)^{(1-\ell)/2} + \cdots) = \widetilde{O}(k^{(1-\ell)/2})$. □

Corollary 3.4 naturally raises the question of whether all Boolean functions are $(\Omega(1), \Omega(1))$-folding.

Question 3.5. *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 3.5 in conjunction with Corollary 3.4 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 2.3) is $(1, 1/2 - o(1))$-folding, and not $(\Omega(1), \ell)$-folding for any $\ell \geq \frac{1}{2}$ (see Appendix B). In light of this, one may ask the following more refined question:

Question 3.6. *Does there exist a constant $\delta > 0$ such that every Boolean function $f : \mathbb{F}_2^n \to \{-1, 1\}$ is $(\delta, 1/2 - o(1))$-folding?*

Assuming a positive answer to Question 3.6, Corollary 3.4 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 from Proposition 2.5 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 Appendix (see Theorem A.1) that the conditions in Proposition 2.5 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.

Theorem 3.7. *For any Boolean function $f : \mathbb{F}_2^n \to \{-1, 1\}$ with Fourier sparsity $k > 4$, and every $\alpha \in \mathcal{S}$, there exists $\beta \in \mathcal{S} \setminus \{\alpha\}$ such that $|O_{\alpha+\beta}| \geq 3$.*

To rule out the possibility mentioned above, it suffices to exhibit a single pair $\{\alpha, \beta\} \in \binom{\mathcal{S}}{2}$ with $|O_{\alpha+\beta}| \geq 3$. Theorem 3.7 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 $\mathcal{S}_+ := \{\alpha \in \mathcal{S} \mid \widehat{f}(\alpha) > 0\}$ and $\mathcal{S}_- := \{\alpha \in \mathcal{S} \mid \widehat{f}(\alpha) < 0\}$. We first show that if there exists $\alpha \in \mathcal{S}$ with $|O_{\alpha+\beta}| = 2$ 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, contradicting (1) above. This proof involves a careful analysis of the Fourier coefficients and crucially uses Observation 2.1 and Equation (2).

A natural question raised by Theorem 3.7 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 $|O_{\alpha+\beta}| \geq 3$. The following theorem answers this question in the positive and sheds more light on the structure of such functions:

Theorem 3.8.

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

The proof of Part 2 of Theorem 3.8 follows along the lines of the proof of Theorem 3.7. The proof of Part 1 of Theorem 3.8 constructs such a function by applying a simple modification to the addressing function.

We prove Theorems 3.7 and 3.8 in Section 6.

4 PROOF OF THEOREM 3.2

In this section, we prove our first result, Theorem 3.2.

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 3.2.

Claim 4.1. *Let* $p := \frac{2\sqrt{\log k}}{k^{1/4}}$, *and let* $\mathcal{R}$ *be the set returned by* SampleParity(f, p). *Then*

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

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

Proof. Fix any $\alpha \in \mathcal{S}$. By Proposition 2.5, 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,

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

Case 2: $|\mathcal{D}_\gamma| < k^{1/2}$ for each $\gamma \in \mathcal{S} \setminus \{\alpha\}$.

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 $\mathcal{F}$.

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

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 $\alpha + \delta = \delta_1 + \delta_2 \in \text{span}\ \mathcal{R}$, i.e., α is identified with δ with respect to 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 2.11. □

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

We now proceed to prove Theorem 3.2 by refining the ideas developed in Section 4.1.

We require the following inequality:

Proposition 4.2. *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. The proof proceeds via induction on d.

Base case: d=0. The statement can be easily verified to be true; each side evaluates to 1.

Inductive step: Assume that the statement is true for $d \geq 0$ and all $p \in [0,1]$ such that $pd \leq 1$. We now show that the hypothesis holds for $d+1$ and all $p \in [0, \frac{1}{d+1}]$. We have

$$\begin{aligned}
(1-p)^{d+1} &= (1-p)\cdot(1-p)^d \\
&\leq (1-p)\left(1 - \frac{1}{2}pd\right) && \text{by inductive hypothesis, since } pd \leq p(d+1) \leq 1 \\
&= 1 - \left(\frac{1}{2}p + \frac{1}{2}pd\right) - \frac{1}{2}p + \frac{1}{2}p^2 d \\
&\leq 1 - \frac{1}{2}p(d+1). && \text{since } pd \leq 1
\end{aligned}$$

□

To focus on the flow of the arguments at a high level, in the following proof, we state and use some claims without proofs. We present the proofs of the claims in the next subsection. We also refer the reader to Figure 1 for a pictorial representation of most of the notations defined in this proof.

Proof of Theorem 3.2. 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 $\mathbb{E}[\mathcal{B}(f, \text{span}\ \mathcal{R}')] \leq ck$.

We will show that for each $\alpha \in \mathcal{S}$, there is a constant probability of it being identified with an element of $\mathcal{S} \setminus \{\alpha\}$. The theorem will follow by a linearity of expectation over all possible choices

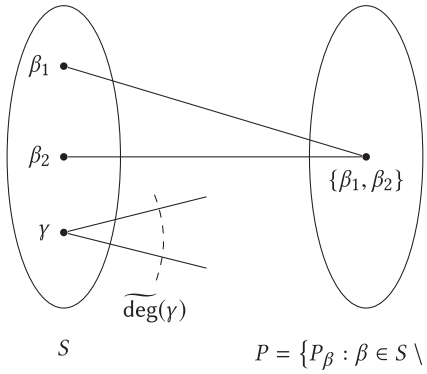

Fig. 1. Figure depicting objects dealt with in Proof of Theorem 3.2. Fix $\alpha \in S$. The left set of vertices in the bipartite graph is S. The right set of vertices is $P = \{P_\beta : \beta \in S \setminus \{\alpha\}\}$. Here, $P_\beta = \{\beta_1, \beta_2\}$ is as obtained from Proposition 2.5. A vertex $\{\beta_1, \beta_2\}$ on the right has an edge to β_1 and β_2 on the left. These define all the edges in the graph. The degree of a vertex γ on the left is denoted by $\widetilde{\deg}(\gamma)$. LowDeg is defined to be the subset of the vertices γ on the left that satisfy $\widetilde{\deg}(\gamma) < 4k^{1/2}$. GoodNbr is defined to be the subset of vertices $\{\beta_1, \beta_2\}$ on the right for which at least one of its neighbors (i.e., at least one of β_1, β_2) is in LowDeg.

of α, and Observation 2.11. To this end, fix an $\alpha \in \mathcal{S}$. Proposition 2.5 invoked with the pair $\{\alpha, \beta\}$, for an arbitrary $\beta \in \mathcal{S} \setminus \{\alpha\}$, implies existence of $\beta_1, \beta_2 \in \mathcal{S} \setminus \{\alpha, \beta\}$ such that $\alpha + \beta = \beta_1 + \beta_2$. Define $P_\beta := \{\beta_1, \beta_2\}$. Note that by definition, $\alpha \notin P_\beta$ for each $\beta \in \mathcal{S} \setminus \{\alpha\}$. Define $P := \{P_\beta \mid \beta \in \mathcal{S} \setminus \{\alpha\}\}$. Thus, P is a set of pairs of elements of $\mathcal{S}$.

The pairs P_β for distinct $\beta \in \mathcal{S} \setminus \{\alpha\}$ are distinct. To see this, suppose that for $\beta' \neq \beta$, $P_{\beta'} = P_\beta = \{\beta_1, \beta_2\}$. Then, $\beta_1 + \beta_2 = \alpha + \beta = \alpha + \beta'$ giving us $\beta' = \beta$. We thus have that $|P| = |\mathcal{S} \setminus \{\alpha\}| = k - 1$.

We define a bipartite graph G with vertex set $\mathcal{S} \cup P$ and bipartition $(\mathcal{S}, P)$. For $\gamma \in \mathcal{S}$ and $P_\beta = \{\beta_1, \beta_2\} \in P$, there is an edge between γ and P_β if and only if $\gamma \in P_\beta$, i.e., $\gamma = \beta_1$ or $\gamma = \beta_2$. Note that $\alpha \notin P_\beta$ for any $\beta \in \mathcal{S} \setminus \{\alpha\}$. Hence, $\alpha \in \mathcal{S}$ is an isolated vertex in G.

For $\gamma \in \mathcal{S}$, define $\widetilde{\deg}(\gamma)$ to be the degree of γ in G. In other words, $\widetilde{\deg}(\gamma) := |\{\beta \in \mathcal{S} \setminus \{\alpha\} \mid \gamma \in P_\beta = \{\beta_1, \beta_2\}\}|$. The next claim shows that there is a small number of vertices in $\mathcal{S}$ with large degree. Define HighDeg $:= \{\gamma \in \mathcal{S} \mid \widetilde{\deg}(\gamma) \geq 4k^{1/2}\}$.

Claim 4.3. $|\text{HighDeg}| \leq k^{1/2}/2$.

Define LowDeg $:= \overline{\text{HighDeg}} := \mathcal{S} \setminus \text{HighDeg}$. We next show, using Claim 4.3, that there are many vertices $P_\beta = \{\beta_1, \beta_2\}$ in P with at least one low-degree neighbor (i.e., $\beta_1 \in$ LowDeg or $\beta_2 \in$ LowDeg). Define GoodNbr $:= \{P_\beta = \{\beta_1, \beta_2\} \in P \mid \beta_1 \in \text{LowDeg or } \beta_2 \in \text{LowDeg}\}$.

Claim 4.4. *For large enough k,* $|\text{GoodNbr}| \geq \frac{2k}{3}$.

The rest of the proof will focus on low-degree vertices in S (i.e., the set LowDeg), and vertices in P with at least one low-degree neighbor (i.e., the set GoodNbr of vertices $P_\beta = \{\beta_1, \beta_2\}$ such that $\beta_1 \in$ LowDeg or $\beta_2 \in$ LowDeg).

Fix an ordering σ on $\mathcal{S}$ such that all elements of LowDeg appear before all elements of HighDeg. By the defining property of σ, for each $P_\beta \in$ GoodNbr, the first element (according to σ) in the pair P_β is from LowDeg. For $\gamma \in$ LowDeg, let $\mathcal{E}(\gamma)$ be the event that there exists $\beta \in \mathcal{S} \setminus \{\alpha\}$ such that:

(1) γ is the first element in P_β (which, in particular, implies that $P_\beta \in$ GoodNbr, since $\gamma \in$ LowDeg), and
(2) $\beta \in \mathcal{R}_1$ (i.e., the first run of SampleParity(f, p') chooses β).

The next important component in our proof, presented in Claim 4.5, is to show that with high probability, $\mathcal{E}(\gamma)$ holds for a large number of $\gamma \in \mathsf{LowDeg}$.

Define the random set $A := \{\gamma \in \mathsf{LowDeg} \mid \mathcal{E}(\gamma) \text{ occurs}\}$.

Claim 4.5. $\Pr_{\mathcal{R}_1}[|A| \geq \frac{k^{1/2}}{24}] \geq 0.9$.

Thus, with probability at least 0.9, for a large number of elements γ of LowDeg, the first run of SampleParity(f, p') samples a β such that γ is the first element of P_β. Since there is a large number of such elements $\gamma \in \mathsf{LowDeg}$ by Claim 4.3, the next independent run of SampleParity(f, p') picks up at least one of them with constant probability as, shown below.

$$\Pr_{\mathcal{R}_1, \mathcal{R}_2}\left[A \cap \mathcal{R}_2 \neq \emptyset \;\middle|\; |A| \geq \frac{k^{1/2}}{24}\right] \geq 1 - (1-p')^{k^{1/2}/24} \geq 1 - e^{-p' \cdot \frac{k^{1/2}}{24}} = 1 - e^{-\frac{1}{96}}. \tag{8}$$

Define $c_1 := 1 - e^{-\frac{1}{96}}$. By Claim 4.5 and Equation (8), the probability of the event $\mathcal{E} := \{|A| \geq \frac{k^{1/2}}{24}\} \cap \{A \cap \mathcal{R}_2 \neq \emptyset\}$ is at least $0.9c_1$. Suppose that the event $\mathcal{E}$ occurs, and let $\gamma \in A \cap \mathcal{R}_2$. By the definitions of A and $\mathcal{E}(\gamma)$, we have that there exists $\beta \in \mathcal{S} \setminus \{\alpha\}$ such that (1) γ is the first element of P_β, and (2) $\beta \in \mathcal{R}_1$. Let $P_\beta = \{\gamma, \delta\}$. 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 span $\mathcal{R}'$. In summary, we have shown that for an arbitrary $\alpha \in \mathcal{S}$,

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

By linearity of expectation over all possible choices of $\alpha \in \mathcal{S}$,

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

Observation 2.11 then implies that

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

where $c := (1 - \frac{0.9c_1}{2})$. □

Missing proofs

Proof of Claim 4.3. For each $\beta \in \mathcal{S} \setminus \{\alpha\}$, $|P_\beta| = 2$. Hence, the degree of each vertex in P is 2. Thus, the total number of edges in G is $2(k-1)$. This implies that the average degree of a vertex in $\mathcal{S}$ is $2(k-1)/k \leq 2$. In other words, $\mathbb{E}_{\gamma \sim \mathcal{S} \setminus \{\alpha\}}[\widetilde{\deg}(\gamma)] = 2$. By Markov's inequality, $|\mathsf{HighDeg}| \leq k \cdot \frac{2}{4k^{1/2}} = k^{1/2}/2$. □

Proof of Claim 4.4. Define the set $\mathsf{BadNbr} := \overline{\mathsf{GoodNbr}} := P \setminus \mathsf{GoodNbr} = \{P_\beta \in P \mid P_\beta \subseteq \mathsf{HighDeg}\}$. It has been observed before that the pairs P_β for distinct $\beta \in \mathcal{S} \setminus \{\alpha\}$ are distinct. Since both of the neighbors of an element in BadNbr are in HighDeg, we have $|\mathsf{BadNbr}| \leq \binom{|\mathsf{HighDeg}|}{2} \leq \frac{k}{8}$ by Claim 4.3. Consequently, $|\mathsf{GoodNbr}| = |P| - |\mathsf{BadNbr}| \geq k - 1 - \frac{k}{8} \geq \frac{2k}{3}$, where the last inequality holds for large enough k. □

Proof of Claim 4.5. For each $\gamma \in \mathsf{LowDeg}$, we define $\mathrm{d}(\gamma)$ to be the number of pairs $P_\beta \in \mathsf{GoodNbr}$ such that γ is the first element in P_β. By Claim 4.4, we have that

$$\sum_{\gamma \in \mathsf{LowDeg}} \mathrm{d}(\gamma) = |\mathsf{GoodNbr}| \geq \frac{2k}{3}, \tag{9}$$

where the last inequality holds for sufficiently large k. We give a lower bound on the probability of $\mathcal{E}(\gamma)$ as follows:

$$\Pr_{\mathcal{R}_1}[\mathcal{E}(\gamma)] = 1 - (1-p')^{\mathrm{d}(\gamma)} \geq \frac{p' \cdot \mathrm{d}(\gamma)}{2}, \tag{10}$$

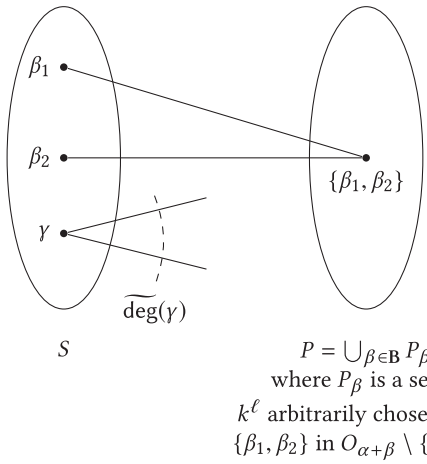

Fig. 2. Figure depicting objects dealt with in Proof of Theorem 3.3. Fix $\alpha \in S$. The left set of vertices in the bipartite graph is S. From Claim 2.13, α must participate in many (at least $\delta k/2$) large (size at least $k^\ell + 1$) folding directions. Define $\mathbf{B}$ to be a set of $\delta k/2$ arbitrarily chosen such directions. For each $\beta \in \mathbf{B}$, arbitrarily choose a set of k^ℓ distinct pairs $\{\beta_1, \beta_2\}$ in $O_{\alpha+\beta} \setminus \{\{\alpha, \beta\}\}$, and denote this set by P_β. The right set of vertices is $P = \bigcup_{\beta \in \mathbf{B}} P_\beta$. A vertex $\{\beta_1, \beta_2\}$ on the right has an edge to β_1 and β_2 on the left. These define all the edges in the graph. The degree of a vertex γ on the left is denoted by $\widetilde{\deg}(\gamma)$. LowDeg is defined to be the subset of the vertices γ on the left that satisfy $\widetilde{\deg}(\gamma) \leq 5k^{(1+\ell)/2}/2$. GoodNbr is defined to be the subset of vertices $\{\beta_1, \beta_2\}$ on the right for which at least one of its neighbors (i.e., at least one of β_1, β_2) is in LowDeg. Define $\mathbf{B}_{\text{good}} := \{\beta \in \mathbf{B} \mid |P_\beta \cap \text{GoodNbr}| \geq k^\ell/8\}$.

where the last inequality follows by Proposition 4.2. Here, Proposition 4.2 is applicable, since $\mathrm{d}(\gamma) \leq \widetilde{\deg}(\gamma) \leq 4k^{1/2}$ (since $\gamma \in \text{LowDeg}$), and $p' = \frac{1}{4k^{1/2}}$.

Let us now first bound the expected size of A from below. Since $|A| = \sum_{\gamma \in \text{LowDeg}} \mathbf{1}_{\mathcal{E}(\gamma)}$, we have by the linearity of expectation that

$$\mathbb{E}_{\mathcal{R}_1}[|A|] = \sum_{\gamma \in \text{LowDeg}} \Pr_{\mathcal{R}_1}[\mathcal{E}(\gamma)] \geq \sum_{\gamma \in \text{LowDeg}} \frac{p' \cdot \mathrm{d}(\gamma)}{2} \quad \text{by linearity of expectation and Equation (10)}$$

$$\geq \frac{1}{2} \cdot \frac{1}{4k^{1/2}} \cdot \frac{2k}{3} \geq \frac{k^{1/2}}{12}. \quad \text{by Equation (9), and substituting the value of } p'$$

Next, we observe that for distinct $\gamma \in \text{LowDeg}$, the sets $\{P_\beta \in \text{GoodNbr} \mid \gamma \text{ is the first element of } P_\beta\}$ are disjoint. Thus, the events $\mathcal{E}(\gamma)$ are independent. Since $|A| = \sum_{\gamma \in \text{LowDeg}} \mathbf{1}_{\mathcal{E}(\gamma)}$, we have by a Chernoff bound that, $\Pr_{\mathcal{R}_1}[|A| \geq \frac{k^{1/2}}{24}] \geq 0.9$. □

5 PROOF OF THEOREM 3.3

In this section, we prove Theorem 3.3. Similar to the proof of Theorem 3.2, to focus on the flow of arguments at a high level, we state several claims and use them without proofs. Their proofs are presented in the next subsection. The proof of Theorem 3.3 proceeds along the same lines of that of Theorem 3.2. We have made an effort to use similar notation to the extent possible. Furthermore, many of the claims used have their analogs in the proof of Theorem 3.2. As in the earlier proof, we refer the reader to Figure 2 for a pictorial representation of most of the notations defined in the following proof.

Proof of Theorem 3.3. Since f is (δ, ℓ)-folding, Claim 2.13 implies that many $\alpha \in \mathcal{S}$ participate in many large folding directions. More precisely, if we define

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

then $|U| \geq \delta k/3$.

As in the proof of Theorem 3.2, for technical reasons, we consider a slightly different probabilistic procedure than SampleParity(f, p). Define $p_1 := \frac{4 \log k}{5e\delta k^{(1+\ell)/2}}$ and $p_2 := \frac{2{,}000 \log k}{\delta k^{(1+\ell)/2}}$. Let $\mathcal{R}_1$ and $\mathcal{R}_2$ be the sets returned by independent runs of SampleParity(f, p_1) and SampleParity(f, p_2), respectively, 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 $1 - (1 - p_1)(1 - p_2) = p_1 + p_2 - p_1 p_2 < 2p_2 = p$. Hence, it suffices to prove that with probability at least $1 - \frac{1}{k}$, we have $\mathcal{B}(f, \operatorname{span} \mathcal{R}') \leq k - \frac{\delta k}{6}$.

We will show that with probability at least $1 - \frac{1}{k}$, each element of U is identified with some other element of $\mathcal{S}$ w.r.t. span $\mathcal{R}'$. The theorem would then follow by Observation 2.11. To this end, fix an $\alpha \in U$. Define $\mathbf{B}$ to be a set of arbitrarily chosen $\delta k/2$ elements $\beta \in \mathcal{S} \setminus \{\alpha\}$ such that $|O_{\alpha+\beta}| \geq k^\ell + 1$. For each $\beta \in \mathbf{B}$, there are at least k^ℓ distinct pairs $\{\beta_1, \beta_2\} \in O_{\alpha+\beta} \setminus \{\{\alpha, \beta\}\}$. Let P_β be a set of k^ℓ arbitrarily chosen such pairs. We remark here that in the proof of Theorem 3.2, P_β represented a single pair $\{\beta_1, \beta_2\}$. Here, however, P_β is a set of pairs. Define $P := \bigcup_{\beta \in \mathbf{B}} P_\beta$.

For any $\beta \neq \beta' \in \mathbf{B}$, we have $P_\beta \cap P_{\beta'} = \emptyset$. To see this, suppose that $\{\beta_1, \beta_2\} \in P_\beta \cap P_{\beta'}$. Then, $\beta_1 + \beta_2 = \alpha + \beta = \alpha + \beta'$, giving us $\beta = \beta'$. We thus have that

$$|P| = \sum_{\beta \in \mathbf{B}} |P_\beta| = \frac{\delta k}{2} \cdot k^\ell = \delta k^{1+\ell}/2. \tag{11}$$

Much in the spirit of the proof of Theorem 3.2, we define a bipartite graph G on vertex set $\mathcal{S} \cup P$, with bipartition $(\mathcal{S}, P)$. There is an edge between $\gamma \in \mathcal{S}$ and $\{\beta_1, \beta_2\} \in P$ if and only if $\gamma \in \{\beta_1, \beta_2\}$, i.e., $\gamma = \beta_1$ or $\gamma = \beta_2$. For each $\gamma \in \mathcal{S}$, define $\widetilde{\deg}(\gamma)$ to be the degree of γ in G. This is precisely the total number of pairs in P that γ appears in. Note that $\mathcal{S} \ni \alpha \notin \{\beta_1, \beta_2\}$ for any pair $\{\beta_1, \beta_2\} \in P$; thus, α is an isolated vertex in G and $\widetilde{\deg}(\alpha) = 0$.

Analogous to Claim 4.3 in the proof of Theorem 3.2, we next show that there is a small number of vertices in $\mathcal{S}$ that have high degree. Define HighDeg $:= \{\gamma \in \mathcal{S} \mid \widetilde{\deg}(\gamma) > 5k^{(1+\ell)/2}/2\}$.

Claim 5.1. $|\mathsf{HighDeg}| < \frac{2\delta k^{(1+\ell)/2}}{5}$.

Define LowDeg $:= \overline{\mathsf{HighDeg}} := \mathcal{S} \setminus$ HighDeg. Now, define GoodNbr $\subseteq P$ to be the set of vertices in P that has at least one low-degree neighbor. Formally, define

$$\mathsf{GoodNbr} := \{\{\mu, \nu\} \in P \mid \mu \in \mathsf{LowDeg} \text{ or } \nu \in \mathsf{LowDeg}\}.$$

The next claim gives a lower bound on the size of GoodNbr. The proof uses Claim 5.1.

Claim 5.2. $|\mathsf{GoodNbr}| > \frac{\delta k^{1+\ell}}{4}$.

Thus, GoodNbr is a large collection of pairs. Now GoodNbr has contributions from P_β for various $\beta \in \mathbf{B}$. Recall that $|\mathbf{B}| = \delta k/2$ and for each $\beta \in \mathbf{B}$, $|P_\beta| = |P_\beta \cap P| = k^\ell$. We are interested in proving a lower bound on the number of β in $\mathbf{B}$ such that P_β contributes significantly to GoodNbr, i.e., $|P_\beta \cap \mathsf{GoodNbr}|$ is large. Define $\mathbf{B}_{\text{good}} := \{\beta \in \mathbf{B} \mid |P_\beta \cap \mathsf{GoodNbr}| \geq k^\ell/8\}$. The next claim shows that $\mathbf{B}_{\text{good}}$ has $\Omega(\delta k)$ elements.

Claim 5.3. $|\mathbf{B}_{\text{good}}| \geq \frac{3\delta k}{14}$.

The rest of the proof focuses on the vertices in LowDeg, the vertices in set GoodNbr $\subseteq P$ with at least one low-degree neighbor, and the set $\mathbf{B}_{\text{good}}$ of elements $\beta \in \mathbf{B}$ such that P_β has a large contribution to GoodNbr.

For $\gamma \in \text{LowDeg}$ and $\beta \in \mathbf{B}_{\text{good}}$, we say that β **hits** γ if γ appears in a pair in P_β (which is also a pair in GoodNbr, since $\gamma \in \text{LowDeg}$). Our goal, accomplished in Claim 5.7, is to show that, with high probability, there exists a $\beta \in \mathbf{B}_{\text{good}}$ and a $\gamma \in \text{LowDeg}$ such that

(1) β hits γ, i.e., there is a pair $\{\gamma, \mu\} \in P_\beta$ for some $\mu \in \mathcal{S} \setminus \{\alpha\}$, and
(2) $\beta \in \mathcal{R}_1$ and $\gamma \in \mathcal{R}_2$. In particular, both β and γ are selected by some run of SampleParity.

We show later that this implies that α is identified with μ with respect to span $\mathcal{R}'$, and the proof will follow by Observation 2.11 and a union bound over all α.

We now build up towards showing our aforementioned goal. Claim 5.3 shows that there are many elements in $\mathbf{B}_{\text{good}}$. Hence, with high probability, a large number of elements of $\mathbf{B}_{\text{good}}$ are selected in the first run of SampleParity(f, p_1). This is proven in the following claim. Define $\mathbf{B}_{\text{good}}^{\text{sampled}} := \mathbf{B}_{\text{good}} \cap \mathcal{R}_1$.

Claim 5.4. *For k sufficiently large,* $\Pr_{\mathcal{R}_1}[|\mathbf{B}_{\text{good}}^{\text{sampled}}| < \frac{k^{(1-\ell)/2}\log k}{25}] < \frac{1}{3k^2}$.

For a fixed $\gamma \in \text{LowDeg}$, let $\mathcal{E}(\gamma)$ denote the event that the run of SampleParity(f, p_1) chooses at least $4\log k/\delta$ different elements in $\mathbf{B}_{\text{good}}$ that hit γ. Formally,

$$\mathcal{E}(\gamma) := \left\{ \left| \left\{ \beta \in \mathbf{B}_{\text{good}} \;\middle|\; \beta \in \mathcal{R}_1 \text{ and } \beta \text{ hits } \gamma \right\} \right| \geq \frac{4\log k}{\delta} \right\}.$$

We next show that the probability that $\mathcal{E}(\gamma)$ happens for even one γ in LowDeg is low.

Claim 5.5. $\Pr_{\mathcal{R}_1}[\bigcup_{\gamma \in \text{LowDeg}} \mathcal{E}(\gamma)] \leq \frac{1}{3k^2}$.

Define $\mathcal{F}$ to be the event $\{\overline{\bigcup_{\gamma \in \text{LowDeg}} \mathcal{E}(\gamma)}\} \cap \{|\mathbf{B}_{\text{good}}^{\text{sampled}}| \geq \frac{1}{25} \cdot k^{(1-\ell)/2}\log k\}$. In other words, $\mathcal{F}$ is the event that none of the "bad" events tackled by Claims 5.4 and 5.5 occur. By these two claims and a union bound,

$$\Pr_{\mathcal{R}_1}[\mathcal{F}] > 1 - \frac{2}{3k^2}. \tag{12}$$

Define $Y := \{\gamma \in \text{LowDeg} \mid \gamma \text{ is hit by some } \beta \in \mathbf{B}_{\text{good}}^{\text{sampled}}\}$. We now show that conditioned on $\mathcal{F}$, the set Y is large.

Claim 5.6. *Conditioned on $\mathcal{F}$, $|Y| \geq \frac{\delta k^{(1+\ell)/2}}{800}$.*

We are now equipped to state and prove Claim 5.7.

Claim 5.7. *For k sufficiently large,* $\Pr[\exists \beta \in \mathcal{R}_1 \text{ and } \gamma \in \mathcal{R}_2 \text{ such that } \gamma \text{ is hit by } \beta] \geq 1 - \frac{1}{k^2}$.

Proof. First, condition on $\mathcal{F}$. Claim 5.6 gives us that $|Y| \geq \frac{\delta k^{(1+\ell)/2}}{800}$. Since Y is large, the second run of SampleParity picks up an element of Y with high probability:

$$\Pr_{\mathcal{R}_1, \mathcal{R}_2}[\mathcal{R}_2 \cap Y \neq \emptyset \mid \mathcal{F}] \geq 1 - (1-p_2)^{\frac{\delta k^{(1+\ell)/2}}{800}} \geq 1 - e^{-p_2 \cdot \frac{\delta k^{(1+\ell)/2}}{800}} \geq 1 - \frac{1}{k^3}. \tag{13}$$

Let the event $\mathcal{R}_2 \cap Y \neq \emptyset$ occur, and let $\gamma \in \mathcal{R}_2 \cap Y$. By the definition of Y, we have that γ is hit by some $\beta \in \mathbf{B}_{\text{good}}^{\text{sampled}} = \mathcal{R}_1 \cap \mathbf{B}_{\text{good}} \subseteq \mathcal{R}_1$. Hence, the event in the statement of the claim occurs.

The claim follows by Equations (13) and (12), a union bound and the easy observation that $\frac{1}{k^3} + \frac{2}{3k^2} \leq \frac{1}{k^2}$ for large enough k. □

Suppose that the event in the statement of Claim 5.7 holds, and β and γ be as in the statement of the claim. By the definition of γ being hit by β, this means γ appears in a pair in P_β, say, $\{\gamma, \mu\}$. Thus, $\alpha+\beta = \gamma+\mu$, which implies $\alpha+\mu = \beta+\gamma$. Since $\beta \in \mathcal{R}_1$ and $\gamma \in \mathcal{R}_2$, we have $\alpha+\mu \in \operatorname{span}\mathcal{R}'$. Thus, α is identified with μ with respect to span $\mathcal{R}'$.

Recall that $\alpha \in U$ was chosen arbitrarily in the beginning of the proof. To finish the proof of Theorem 3.3, we first take a union bound over all $\alpha \in U$. Thus, the event in Claim 5.7 holds for each $\alpha \in U$ with probability at least $1 - |U| \cdot \frac{1}{k^2} \geq 1 - \frac{1}{k}$. The theorem follows by invoking Observation 2.11 and recalling that $|U| \geq \frac{\delta k}{3}$. □

Missing proofs

Proof of Claim 5.1. Each vertex $\{\beta_1, \beta_2\} \in P$ has two neighbors, namely, β_1 and β_2. Hence, the total number of edges in G is $2 \cdot |P| = \delta k^{1+\ell}$ (Equation (11)). We thus have that the average degree of a vertex in $\mathcal{S}$ is $\delta k^{1+\ell}/k = \delta k^{\ell}$. In other words, $\mathbb{E}_{\gamma \in \mathcal{S}}[\widetilde{\deg}(\gamma)] = \delta k^{\ell}$. By Markov's inequality, $|\mathsf{HighDeg}| < \delta k^{\ell} \cdot \frac{2}{5k^{(1+\ell)/2}} \cdot k = \frac{2\delta k^{(1+\ell)/2}}{5}$. □

Proof of Claim 5.2. The number of pairs $\{\beta_1, \beta_2\} \in P$ such that β_1 and β_2 are both in HighDeg is bounded above by $\binom{|\mathsf{HighDeg}|}{2}$. By Equation (11) and Claim 5.1, we have that,

$$|\mathsf{GoodNbr}| \geq |P| - \binom{|\mathsf{HighDeg}|}{2} \geq \frac{\delta k^{1+\ell}}{2} - \frac{4\delta^2 k^{1+\ell}}{50} = \delta k^{1+\ell}\left(\frac{1}{2} - \frac{2\delta}{25}\right) > \frac{\delta k^{1+\ell}}{4}.$$ □

Proof of Claim 5.3. By the definition of $\mathbf{B}_{\mathrm{good}}$, we have

$$|\mathsf{GoodNbr}| \leq |\mathbf{B}_{\mathrm{good}}| \cdot k^{\ell} + \left(\frac{\delta k}{2} - |\mathbf{B}_{\mathrm{good}}|\right) \cdot \frac{k^{\ell}}{8}.$$

Along with Claim 5.2, we obtain

$$\begin{aligned} & |\mathbf{B}_{\mathrm{good}}| \cdot k^{\ell} + \left(\frac{\delta k}{2} - |\mathbf{B}_{\mathrm{good}}|\right) \cdot \frac{k^{\ell}}{8} \geq \frac{\delta k^{1+\ell}}{4} \\ \implies & |\mathbf{B}_{\mathrm{good}}| \cdot \frac{7k^{\ell}}{8} \geq \frac{3\delta k^{1+\ell}}{16} \\ \implies & |\mathbf{B}_{\mathrm{good}}| \geq \frac{3\delta k}{14}. \end{aligned}$$ □

Proof of Claim 5.5. For $\gamma \in \mathsf{LowDeg}$, define

$$\mathrm{d}(\gamma) := |\left\{\beta \in \mathbf{B}_{\mathrm{good}} \;\middle|\; \beta \text{ hits } \gamma\right\}|.$$

We have $\mathrm{d}(\gamma) \leq \widetilde{\deg}(\gamma) \leq 5k^{(1+\ell)/2}/2$, where the first inequality holds, since γ appears in at most one pair in any P_β, and the second inequality follows from the definition of LowDeg. Thus, for any fixed $\gamma \in \mathsf{LowDeg}$, we have

$$\begin{aligned} \Pr_{\mathcal{R}_1}[\mathcal{E}(\gamma)] &\leq \binom{\mathrm{d}(\gamma)}{(4\log k)/\delta} p_1^{(4\log k)/\delta} \leq \left(\frac{e\delta \cdot \mathrm{d}(\gamma) \cdot p_1}{4\log k}\right)^{(4\log k)/\delta} \\ &\leq \left(\frac{1}{2}\right)^{(4\log k)/\delta} && \text{since } p_1 = \frac{4\log k}{5e\delta k^{(1+\ell)/2}} \\ &\leq \frac{1}{3k^3}. && \text{for sufficiently large } k \end{aligned}$$

Since $|\mathsf{LowDeg}| \leq k$, we have by a union bound that $\Pr_{\mathcal{R}_1}[\bigcup_{\gamma \in \mathsf{LowDeg}} \mathcal{E}(\gamma)] \leq \frac{1}{3k^2}$. □

Proof of Claim 5.4. Recall from Claim 5.3 that $|\mathbf{B}_{\text{good}}| \geq 3\delta k/14$. By our choice of p_1, the expected number of elements in $\mathbf{B}_{\text{good}}$ that are included in $\mathcal{R}_1$ is at least $\frac{3\delta k}{14} \cdot \frac{4\log k}{5e\delta k^{(1+\ell)/2}} = \frac{6}{35e} \cdot k^{(1-\ell)/2} \log k$. By a Chernoff bound, the number of elements β in $\mathbf{B}_{\text{good}}$ that are included in $\mathcal{R}_1$ is less than $\frac{1}{25} \cdot k^{(1-\ell)/2} \log k$ with probability at most $\exp(-\Omega(k^{(1-\ell)/2} \log k))$. Since $\ell \leq 1 - \Omega(1)$, this probability is at most $\exp(-(k^{\Omega(1)}))$, which is at most $\frac{1}{3k^2}$ for sufficiently large k. □

Proof of Claim 5.6. Recall that $\gamma \in \text{LowDeg}$ is said to be hit by $\beta \in X$ if $\gamma \in P_\beta = \{\beta_1, \beta_2\}$, i.e., $\gamma = \beta_1$ or $\gamma = \beta_2$. Hence, the number of low-degree vertices hit by elements in X is exactly the number of distinct low-degree vertices present in the pairs of $\cup_{\beta \in X}(\text{GoodNbr} \cap P_\beta)$. By the definition of $\mathbf{B}_{\text{good}}$, for each element $\beta \in X \subseteq \mathbf{B}_{\text{good}}$, $|\text{GoodNbr} \cap P_\beta| \geq k^\ell/8$, and each pair in $\text{GoodNbr} \cap P_\beta$ contains at least one low-degree vertex (by the definition of GoodNbr). We thus have, since the P_β's are disjoint, that $|\cup_{\beta \in X}(\text{GoodNbr} \cap P_\beta)| \geq |X| \cdot (k^\ell/8)$. Now, conditioned on $\mathcal{F}$, we have that $\mathcal{E}(\gamma)$ does not hold for any $\gamma \in \text{LowDeg}$. This, together with the observation that any $\gamma \in \text{LowDeg}$ can appear in at most one pair in any P_β, implies that no low-degree vertex appears in at least $4\log k/\delta$ pairs in $\cup_{\beta \in X}(\text{GoodNbr} \cap P_\beta)$. It follows that $|Y| \geq \frac{|X| \cdot (k^\ell/8)}{(4\log k)/\delta} \geq \frac{\delta k^{(1+\ell)/2}}{800}$. □

6 PROOFS OF THEOREMS 3.7 AND 3.8

In this section, we prove Theorems 3.7 and 3.8. Theorem 3.7 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 $|O_{\alpha+\beta}| \geq 3$. Theorem 3.8 consists of two parts; the first part asserts existence of a function $f : \mathbb{F}_2^n \to \{-1, 1\}$ and $\alpha \in \mathcal{S}$ for which there exists only one $\beta \in \mathcal{S}$ with $|O_{\alpha+\beta}| \geq 3$, and the second part gives additional structure on such functions.

Organization of section: We first review and define some preliminaries. Towards proving Theorem 3.7, we first show in Claims 6.1 and 6.2 that if the theorem is not true, then f must be plateaued. We then show in Claim 6.3 that the theorem holds true for plateaued functions, concluding the proof of the theorem. We then show the second part of Theorem 3.8 along the lines of the proof of Theorem 3.7 and finally end the section with a proof of the first part of Theorem 3.8 by using a simple variant of the Addressing function.

We first recall and introduce some notation. Recall from Proposition 2.5 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 $|O_\gamma| \geq 2$. For any γ with $|O_\gamma| > 2$, we say that γ is a *non-trivial folding direction*. Hence, Theorem 3.7 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 $\mathcal{S}_+ := \{\alpha \in \mathcal{S} \mid \widehat{f}(\alpha) > 0\}$, and $\mathcal{S}_- := \{\alpha \in \mathcal{S} \mid \widehat{f}(\alpha) < 0\}$. 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 next state a preliminary claim.

Claim 6.1. *Let $f : \mathbb{F}_2^n \to \{-1, 1\}$ be any Boolean function. Suppose there exists $\alpha \in \mathcal{S}$ such that $|O_{\alpha+\beta}| = 2$ 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 $|O_{\alpha+\beta}| = 2$ 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 $|O_{\alpha+\beta}| = 2$ 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 appear 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 6.2. *Let $f : \mathbb{F}_2^n \to \{-1, 1\}$ be any Boolean function. If there exists $\alpha \in \mathcal{S}$ such that $|O_{\alpha+\beta}| = 2$ for all $\beta \in \mathcal{S} \setminus \{\alpha\}$, then f is plateaued.*

Proof of Claim 6.2. Fix any $\alpha \in \mathcal{S}$ such that $|O_{\alpha+\beta}| = 2$ 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 2.5 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 resulting 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 $|O_{\delta_i+\delta_j}| = |O_{\delta_k+\delta_m}| = 2$. 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 3.7 holds true if f is a plateaued function.

Claim 6.3. *Let $f : \mathbb{F}_2^n \to \{-1, 1\}$ be any plateaued Boolean function with Fourier sparsity $k > 4$. Then, for any $\alpha \in \mathcal{S}$, there exists $\beta \in \mathcal{S} \setminus \{\alpha\}$ such that $|O_{\alpha+\beta}| \geq 3$.*

Proof. Towards a contradiction, let $\alpha \in \mathcal{S}$ be such that $|O_{\alpha+\beta}| = 2$ 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 2.1, 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 6.1. □

We next use Claim 6.2 to remove the assumption of f being plateaued in the previous claim, which proves Theorem 3.7.

Proof of Theorem 3.7. Towards a contradiction, suppose there exists $\alpha \in \mathcal{S}$ such that $|O_{\alpha+\beta}| = 2$ for all $\beta \in \mathcal{S} \setminus \{\alpha\}$. Claim 6.2 implies that f must be plateaued. Next, Claim 6.3 implies that there must exist $\gamma \in \mathcal{S}$ such that $|O_{\alpha+\gamma}| \geq 3$, which is a contradiction. □

We have shown that for any Boolean function $f : \mathbb{F}_2^n \to \{-1, 1\}$, each $\alpha \in \mathcal{S}$ participates in at least one non-trivial folding direction. We next investigate if an element $\alpha \in \mathcal{S}$ can participate in exactly one non-trivial folding and first prove Part 2 of Theorem 3.8. This states that if an element $\alpha \in \mathcal{S}$ only participates in one non-trivial folding direction, say δ, then all elements of $\mathcal{S}$ must participate in the folding direction δ.

Proof of Part 2 of Theorem 3.8. Suppose there exist $\{\alpha, \beta\} \in \binom{\mathcal{S}}{2}$ such that $|O_{\alpha+\gamma}| = 2$ for all $\gamma \in \mathcal{S} \setminus \{\alpha, \beta\}$. Let $|O_{\alpha+\beta}| = q > 2$ (Theorem 3.7 shows q cannot equal 2), and say $O_{\alpha+\beta} = \{\{\alpha, \beta\}, \{\zeta_{1,1}, \zeta_{1,2}\}, \{\zeta_{2,1}, \zeta_{2,2}\}, \ldots, \{\zeta_{q-1,1}, \zeta_{q-1,2}\}\}$. Let $\mathcal{P} := \{\zeta \in \mathcal{S} \mid \zeta \text{ appears in a pair in } O_{\alpha+\beta}\}$. Since the pairs are all disjoint, we have $|\mathcal{P}| = 2q$. Assume without loss of generality (justified by Proposition 2.6) that $\widehat{f}(\alpha) > 0$ and $\widehat{f}(\beta) > 0$.

If $\mathcal{P} = \mathcal{S}$, then $|O_{\alpha+\beta}| = k/2$, which proves the theorem. Otherwise, pick any $\delta \in \mathcal{S} \setminus \mathcal{P}$. Since $|O_{\alpha+\delta}| = 2$ by assumption, let $O_{\alpha+\delta} = \{\{\alpha, \delta\}, \{\mu, \nu\}\}$, i.e., $\alpha + \delta = \mu + \nu$. Next note that $\mu \neq \beta$, since otherwise $\alpha + \beta = \delta + \nu$, which implies that $\{\delta, \nu\} \in O_{\alpha+\beta}$, contradicting our choice of δ. Similarly, $\nu \neq \beta$. Thus, we also obtain

$$O_{\alpha+\mu} = \{\{\alpha, \mu\}, \{\delta, \nu\}\}, \qquad O_{\alpha+\nu} = \{\{\alpha, \nu\}, \{\delta, \mu\}\}. \tag{14}$$

Arrange α, δ, μ, and ν in non-increasing order of the magnitudes of the corresponding Fourier coefficients; let the resulting ordering be $\delta_1, \delta_2, \delta_3, \delta_4$. Thus, we have, $|\widehat{f}(\delta_1)| \geq |\widehat{f}(\delta_2)| \geq |\widehat{f}(\delta_3)| \geq |\widehat{f}(\delta_4)|$. By Equations (3) and (14), we have $|\widehat{f}(\delta_1)||\widehat{f}(\delta_2)| = |\widehat{f}(\delta_3)||\widehat{f}(\delta_4)|$, from which we conclude $|\widehat{f}(\delta_1)| = |\widehat{f}(\delta_2)| = |\widehat{f}(\delta_3)| = |\widehat{f}(\delta_4)|$. Hence,

$$|\widehat{f}(\alpha)| = |\widehat{f}(\delta)| \text{ for all } \delta \in \mathcal{S} \setminus \mathcal{P}. \tag{15}$$

Next, consider any $\{\zeta_{i,1}, \zeta_{i,2}\} \in O_{\alpha+\beta} \setminus \{\{\alpha, \beta\}\}$. Since $|O_{\alpha+\zeta_{i,1}}| = 2$ by our hypothesis, we have $O_{\alpha+\zeta_{i,1}} = \{\{\alpha, \zeta_{i,1}\}, \{\beta, \zeta_{i,2}\}\}$. Similarly, we have $O_{\alpha+\zeta_{i,2}} = \{\{\alpha, \zeta_{i,2}\}, \{\beta, \zeta_{i,1}\}\}$. From Equation (3), we have

$$\widehat{f}(\alpha)\widehat{f}(\zeta_{i,1}) = -\widehat{f}(\beta)\widehat{f}(\zeta_{i,2}), \qquad \widehat{f}(\alpha)\widehat{f}(\zeta_{i,2}) = -\widehat{f}(\beta)\widehat{f}(\zeta_{i,1}). \tag{16}$$

Multiplying each side of these equalities, and, since we have assumed $\widehat{f}(\alpha) > 0$ and $\widehat{f}(\beta) > 0$, we obtain

$$\widehat{f}(\alpha) = \widehat{f}(\beta) =: a, \tag{17}$$

and

$$|\widehat{f}(\zeta_{i,1})| = |\widehat{f}(\zeta_{i,2})| =: a_i, \text{ for all } i \in [q-1]. \tag{18}$$

By substituting $\gamma = \alpha + \beta$ in Equation (3), we have

$$\widehat{f}(\alpha)\widehat{f}(\beta) + \sum_{i \in [q-1]} \widehat{f}(\zeta_{i,1})\widehat{f}(\zeta_{i,2}) = 0. \tag{19}$$

We also conclude from Equations (16) and (17) that

$$\widehat{f}(\zeta_{i,1}) = -\widehat{f}(\zeta_{i,2}) \qquad \forall i \in [q-1]. \tag{20}$$

Hence, Equations (19), (17), and (18) imply

$$a^2 - \sum_{i \in [q-1]} a_i^2 = 0. \tag{21}$$

We next use Parseval's identity to deduce the value of a. We have

$$\begin{aligned}
1 &= \sum_{\delta \in \mathcal{S}} \widehat{f}(\delta)^2 && \text{by Equation (2)} \\
&= \sum_{\delta \notin \mathcal{P}} \widehat{f}(\delta)^2 + \sum_{\delta \in \mathcal{P}} \widehat{f}(\delta)^2 \\
&= (k-2q)a^2 + \sum_{\delta \in \mathcal{P}} \widehat{f}(\delta)^2 && \text{since } |\mathcal{P}| = 2q, \text{ and by Equations (15) and (17)} \\
&= (k-2q+2)a^2 + 2\sum_{i \in [q-1]} a_i^2 && \text{by Equations (17) and (18)} \\
&= (k-2q+4)a^2. && \text{by Equation (21)}
\end{aligned}$$

Hence,

$$a = \frac{1}{\sqrt{k-2q+4}}. \tag{22}$$

By Observation 2.1, we have $\frac{1}{\sqrt{k-2q+4}} = \frac{c}{2^n}$ for some $c \in \mathbb{Z}$. Thus, $k-2q+4 = 2^{2n}/c^2$. Since each side of this equality is an integer, c must be a power of 2, and hence

$$k = 2^{2h} + 2q - 4 \text{ for some } h \in \mathbb{Z}. \tag{23}$$

Define $s := |\mathcal{S}_+|$ and $t := |\mathcal{S}_-|$. Since the Fourier sparsity of f equals k, we have $s + t = k$. We next show that $s - t = f(0^n)/a$. We have

$$f(0^n) = \sum_{\delta \in \mathcal{S}_+} |\widehat{f}(\delta)| - \sum_{\delta \in \mathcal{S}_-} |\widehat{f}(\delta)|. \tag{24}$$

From Equation (20), for each $i \in [q-1]$ exactly one element of the pair $\{\zeta_{i,1}, \zeta_{i,2}\}$ is in $\mathcal{S}_+$ and the other is in $\mathcal{S}_-$. Furthermore, $|\widehat{f}(\zeta_{i,1})| = |\widehat{f}(\zeta_{i,2})|$. Thus, the contribution of parities in $\mathcal{P} \setminus \{\alpha, \beta\}$ to the above sum is 0. Hence, it is enough to account for the contribution from $(\mathcal{S} \setminus \mathcal{P}) \cup \{\alpha, \beta\}$, which contains $s-(q-1)$ elements of $\mathcal{S}_+$ and $t-(q-1)$ elements of $\mathcal{S}_-$. Now, recall from Equations (15) and (17) that $\widehat{f}(\alpha), \widehat{f}(\beta)$ and the absolute values of coefficients of parities in $\mathcal{S} \setminus \mathcal{P}$ are all equal to a. Thus, we have from Equation (24),

$$\begin{aligned}
& f(0^n) = (s-(q-1))a - (t-(q-1))a \\
\implies\ & s - t = \frac{f(0^n)}{a}.
\end{aligned}$$

Thus, we have that

$$\begin{aligned}
s &= \frac{s+t}{2} + \frac{s-t}{2} \\
&= \frac{k}{2} + \frac{f(0^n)}{2a} \\
&= \frac{k}{2} + \frac{f(0^n)\sqrt{k-2q+4}}{2} && \text{by Equation (22)} \\
&= 2^{2h-1} + q - 2 + f(0^n) \cdot 2^{h-1}, && (25)
\end{aligned}$$

where the last equality follows from Equation (23). Now, if $h = 1$, then $|O_{\alpha+\beta}| = q = k/2$ (by Equation (23)), which proves the theorem. The case $h = 0$ is ruled out by Equation (23) along with the fact that $q \leq k/2$. In the rest of the proof, we assume that $h > 1$ and derive a contradiction. By Equation (25) the parity of s equals the parity of q.

Next, by a different counting argument, we show that the parity of s does not equal the parity of q, which will yield the desired contradiction. This counting mimics the argument in the proof of Claim 6.1. We recall here that by our assumption $\alpha, \beta \in \mathcal{S}_+$.

Consider the set of unordered triples

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

As in the proof of Claim 6.1, define

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

First consider any parity $\nu \in \mathcal{S} \setminus \mathcal{P}$. Since $|O_{\alpha+\nu}| = 2$ by our definition of $\mathcal{P}$, say $O_{\alpha+\nu} = \{\{\alpha, \nu\}, \{\zeta, \xi\}\}$. Thus, ν appears in exactly one triple in T_+, namely, $\{\nu, \zeta, \xi\}$. Since $\alpha \in \mathcal{S}_+$, Equation (3) implies that each such triple contributes to either zero or two elements of $\mathcal{S}_+$. Hence, these triples, along with the elements α and β, account for an even number of elements of $\mathcal{S}_+$. Now, by our earlier discussion, each pair $\{\zeta_{i,1}, \zeta_{i,2}\} \in O_{\alpha+\beta} \setminus \{\{\alpha, \beta\}\}$ for $i \in [q-1]$ contributes exactly one parity to $\mathcal{S}_+$. Hence, the parity of s equals the parity of $(q-1)$, which is a contradiction. □

One might expect that a stronger statement than the one in Part 2 of Theorem 3.8 is possible. For example, given any Boolean function $f : \mathbb{F}_2^n \to \{-1, 1\}$ and $\alpha \in \mathcal{S}$, it is feasible that there must exist at least two parities $\beta \in \mathcal{S}$ with $|O_{\alpha+\beta}| > 2$. We rule out this possibility in the next proof, hence showing that the hypothesis in Part 2 of Theorem 3.8 can indeed by satisfied. We exhibit an explicit function that witnesses this.

Proof of Part 1 of Theorem 3.8. Recall from Definition 2.3 that ADD_k denotes the addressing function on $n := ((\log k)/2 + \sqrt{k})$ many input bits. Let $(\mathbf{x}, \mathbf{y})$ be a generic input to ADD_k, where $\mathbf{x} \in \mathbb{F}_2^{(\log k)/2}$ and $\mathbf{y} \in \mathbb{F}_2^{\sqrt{k}}$.

Define $g : \mathbb{F}_2^{n+2} \to \{-1, 1\}$ by

$$g(z_1, z_2, \mathbf{x}, \mathbf{y}) := \begin{cases} (-1)^{z_1} & \text{if } \mathrm{ADD}_k(\mathbf{x}, \mathbf{y}) = 1 \\ (-1)^{z_2} & \text{if } \mathrm{ADD}_k(\mathbf{x}, \mathbf{y}) = -1. \end{cases} \tag{26}$$

Define $b_1 = (-1)^{z_1}$ and $b_2 = (-1)^{z_2}$. Now,

$$\begin{aligned} g(z_1, z_2, \mathbf{x}, \mathbf{y}) &= b_1 \left(\frac{1 + \mathrm{ADD}_k(\mathbf{x}, \mathbf{y})}{2} \right) + b_2 \left(\frac{1 - \mathrm{ADD}_k(\mathbf{x}, \mathbf{y})}{2} \right) \\ &= \frac{b_1}{2} + \frac{b_2}{2} + \frac{b_1 \cdot \mathrm{ADD}_k(\mathbf{x}, \mathbf{y})}{2} - \frac{b_2 \cdot \mathrm{ADD}_k(\mathbf{x}, \mathbf{y})}{2}. \end{aligned}$$

For the rest of this proof, we view elements of $\mathbb{F}_2^{n+2}$ as subsets of the input variables, via the natural correspondence. Under this equivalence, addition of elements in $\mathbb{F}_2^{n+2}$ corresponds to the symmetric difference of their respective sets.

Let $\mathcal{S}$ denote the Fourier support of ADD_k. Let $\mathcal{S}'$ denote the Fourier support of g, and let $\alpha = \{z_1\}$ and $\beta = \{z_2\}$. Since f is non-constant, we have

$$\mathcal{S}' = \{\alpha, \beta\} \cup \{\alpha \cup T \mid T \in \mathcal{S}\} \cup \{\beta \cup T \mid T \in \mathcal{S}\}.$$

By Equation (29), every element in $\mathcal{S}$ contains exactly one variable from $\mathbf{y}$. Hence, for any $T_1, T_2 \in \mathcal{S}$, the set $T_1 \triangle T_2$ contains either no variables from $\mathbf{y}$ or exactly 2 variables from $\mathbf{y}$. Thus, for any $T, T_1, T_2 \in \mathcal{S}$, we have $T \neq T_1 \triangle T_2$. It now follows that for any $T \in \mathcal{S}$,

$$\begin{aligned} O_{\alpha \triangle (\alpha \cup T)} = O_T &= \{\{\alpha, \alpha \cup T\}, \{\beta, \beta \cup T\}\}, \\ O_{\alpha \triangle (\beta \cup T)} = O_{\alpha \cup \beta \cup T} &= \{\{\alpha, \beta \cup T\}, \{\beta, \alpha \cup T\}\}. \end{aligned}$$

Hence, $|O_{\alpha \triangle \gamma}| = 2$ for all $\gamma \in \mathcal{S}' \setminus \{\alpha, \beta\}$, which proves the theorem. □

APPENDICES

A RULING OUT SUFFICIENCY OF PROPOSITION 2.5

In this section, we prove that the conditions in Proposition 2.5 are not sufficient for a function to be Boolean. To the best of our knowledge, ours is the first work to show this.

THEOREM A.1. *There exists a set $\mathcal{S} \subseteq \mathbb{F}_2^n$ such that $|O_{\alpha+\beta}| \geq 2$ 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 $\text{sgn}(x) := -1$ if $x < 0$, and $\text{sgn}(x) := 1$ 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 $A, B, C, D \subseteq [n]$,

$$A \triangle B = C \triangle D \iff A \triangle C = B \triangle D. \tag{27}$$

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,

$$O_{\{i\} \triangle \{j\}} = \{\{\{i\}, \{j\}\}, \{\{1, i, n\}, \{1, j, n\}\}\} \quad \forall i, j \in \{2, 3, \ldots, n-1\}. \tag{28}$$

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

$$\{1\} \triangle \{i\} = \{n\} \triangle \{1, i, n\},$$
$$\{n\} \triangle \{i\} = \{1\} \triangle \{1, i, n\}.$$

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 (27), these establish the fact that $|O_{\alpha+\beta}| \geq 2$ 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 (28) then implies

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

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

$$s_1 s_2 = -s_4 s_5$$
$$s_1 s_3 = -s_4 s_6$$
$$s_2 s_3 = -s_5 s_6.$$

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. □

B FOLDING PROPERTIES OF THE ADDRESSING FUNCTION

In this section, we analyze folding properties of the addressing function. In the context of Question 3.5, this function demonstrates that it is not true that all Boolean functions are $(\Omega(1), 1/2)$-folding.

Claim B.1. *The function* $\mathrm{ADD}_k : \mathbb{F}_2^{\frac{1}{2}\log k + \sqrt{k}} \to \{-1, 1\}$ *as defined in Definition 2.3 is*
(1) $(1, 1/2 - o(1))$*-folding, and*
(2) *not* $(\Omega(1), \ell)$*-folding for any* $\ell \geq 1/2$.

For this proof, we view elements of $\mathbb{F}_2^{\frac{1}{2}\log k + \sqrt{k}}$ as subsets of the set of variables, via the natural equivalence. Addition over $\mathbb{F}_2^n$ corresponds to symmetric difference of the respective sets. Recall that for a string $x \in \mathbb{F}_2^{\frac{1}{2}\log k}$, $\mathrm{int}(x)$ is the unique integer in $\{1, \ldots, \sqrt{k}\}$ whose binary representation is x.

Proof. For $a \in [\sqrt{k}]$, define the function $\mathbf{1}_a : \mathbb{F}_2^{\frac{1}{2}\log k} \to \{0, 1\}$ by

$$\mathbf{1}_a(x) = \begin{cases} 1 & \mathrm{int}(x) = a \\ 0 & \text{otherwise.} \end{cases}$$

The Fourier expansion of ADD_k is given by

$$\sum_{a \in [\sqrt{k}]} \mathbf{1}_a(x)(-1)^{y_a}. \tag{29}$$

It can be verified that for any $a \in [\sqrt{k}]$, the Fourier support of $\mathbf{1}_a$ consists of all subsets of $X := \{x_1, \ldots, x_{\frac{1}{2}\log k}\}$. In light of this, the Fourier support $\mathcal{S}$ of ADD_k is given by

$$\mathcal{S} = \bigcup_{a \in [\sqrt{k}], M \subseteq X} \{M \cup \{y_a\}\}. \tag{30}$$

It is not hard to verify that $|\mathcal{S}| = k$. Let α, β be two distinct elements of $\mathcal{S}$. We now determine $O_{\alpha+\beta}$.

Case 1 : $\alpha = M_1 \cup \{y_a\}$ *and* $\beta = M_2 \cup \{y_a\}$ *for distinct* $M_1, M_2 \subseteq X$, *and any* $a \in [\sqrt{k}]$.
In this case,

$$O_{\alpha \triangle \beta} = \left\{ \{M_1' \cup \{y_b\}, M_2' \cup \{y_b\}\} \;\middle|\; M_1' \triangle M_2' = M_1 \triangle M_2, b \in [\sqrt{k}] \right\}.$$

Hence, $|O_{\alpha \triangle \beta}| = k$.

Case 2 : $\alpha = M_1 \cup \{y_a\}$ *and* $\beta = M_2 \cup \{y_b\}$ *for distinct* $a, b \in [\sqrt{k}]$, *and any* $M_1, M_2 \subseteq X$.
In this case,

$$O_{\alpha \triangle \beta} = \left\{ \{M_1' \cup \{y_a\}, M_2' \cup \{y_b\}\} \;\middle|\; M_1' \triangle M_2' = M_1 \triangle M_2 \right\}.$$

Hence, $|O_{\alpha \triangle \beta}| = \sqrt{k}$.

From both cases above, we conclude that for all $\{\alpha, \beta\} \in \binom{S}{2}$, we have $|O_{\alpha \triangle \beta}| \geq \sqrt{k} = k^{\ell} + 1$ for $\ell = 1/2 - o(1)$. This immediately proves the first part of the claim.

It is easy to verify that the number of pairs $\{\alpha, \beta\} \in \binom{S}{2}$ that fall under the second case above is $(1 - o(1))\binom{k}{2}$. For all these pairs, we have $|O_{\alpha \triangle \beta}| = \sqrt{k} < k^{1/2} + 1$. Hence, the number of pairs $\{\alpha, \beta\} \in \binom{S}{2}$ with $|O_{\alpha \triangle \beta}| \geq k^{1/2} + 1$ is $o(1) \cdot \binom{k}{2}$. This proves the second part of the claim. □

ACKNOWLEDGMENTS

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

REFERENCES

[1] Anurag Anshu, Naresh Goud Boddu, and Dave Touchette. 2019. Quantum log-approximate-rank conjecture is also false. In *60th IEEE Annual Symposium on Foundations of Computer Science (FOCS'19)*. 982–994. DOI: https://doi.org/10.1109/FOCS.2019.00063

[2] Anna Bernasconi and Bruno Codenotti. 1999. Spectral analysis of Boolean functions as a graph eigenvalue problem. *IEEE Trans. Comput.* 48, 3 (1999), 345–351. DOI: https://doi.org/10.1109/12.755000

[3] Arkadev Chattopadhyay, Nikhil S. Mande, and Suhail Sherif. 2020. The log-approximate-rank conjecture is false. *J. ACM* 67, 4 (2020), 23:1–23:28. DOI: https://doi.org/10.1145/3396695

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

[5] Troy Lee and Adi Shraibman. 2009. Lower bounds in communication complexity. *Found. Trends Theoret. Comput. Sci.* 3, 4 (2009), 263–398. DOI: https://doi.org/10.1561/0400000040

[6] László Lovász and Michael E. Saks. 1988. Lattices, Möbius functions and communication complexity. In *29th Annual Symposium on Foundations of Computer Science*. 81–90. DOI: https://doi.org/10.1109/SFCS.1988.21924

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

[8] Ashley Montanaro and Tobias Osborne. 2009. On the communication complexity of XOR functions. *CoRR* abs/0909.3392 (2009).

[9] Swagato Sanyal. 2019. Fourier sparsity and dimension. *Theor. Comput.* 15, 1 (2019), 1–13.

[10] Amir Shpilka, Avishay Tal, and Ben Lee Volk. 2017. On the structure of Boolean functions with small spectral norm. *Comput. Complex.* 26, 1 (2017), 229–273. DOI: https://doi.org/10.1007/s00037-015-0110-y

[11] Makrand Sinha and Ronald de Wolf. 2019. Exponential separation between quantum communication and logarithm of approximate rank. In *60th IEEE Annual Symposium on Foundations of Computer Science (FOCS'19)*. 966–981. DOI: https://doi.org/10.1109/FOCS.2019.00062

[12] Robert C. Titsworth. 1962. *Correlation Properties of Cyclic Sequences*. Ph. D. Dissertation. California Institute of Technology.

[13] Hing Yin Tsang, Chung Hoi Wong, Ning Xie, and Shengyu Zhang. 2013. Fourier sparsity, spectral norm, and the log-rank conjecture. In *54th Annual IEEE Symposium on Foundations of Computer Science (FOCS'13)*. 658–667. DOI: https://doi.org/10.1109/FOCS.2013.76

Received 20 May 2022; revised 8 February 2024; accepted 8 February 2024