Tight Chang's-Lemma-Type Bounds for Boolean Functions

Sourav Chakraborty
Indian Statistical Institute, Kolkata, India

Nikhil S. Mande [1]
CWI, Amsterdam, The Netherlands

Rajat Mittal
Indian Institute of Technology, Kanpur, India

Tulasimohan Molli
Tata Institute of Fundamental Research, Mumbai, India

Manaswi Paraashar
Indian Statistical Institute, Kolkata, India

Swagato Sanyal
Indian Institute of Technology, Kharagpur, India

Abstract

Chang's lemma (Duke Mathematical Journal, 2002) is a classical result in mathematics, with applications spanning across additive combinatorics, combinatorial number theory, analysis of Boolean functions, communication complexity and algorithm design. For a Boolean function f that takes values in $\{-1, 1\}$ let $r(f)$ denote its Fourier rank (i.e., the dimension of the span of its Fourier support). For each positive threshold t, Chang's lemma provides a lower bound on $\delta(f) := \Pr[f(x) = -1]$ in terms of the dimension of the span of its characters with Fourier coefficients of magnitude at least $1/t$. In this work we examine the tightness of Chang's lemma with respect to the following three natural settings of the threshold:

- the Fourier sparsity of f, denoted $k(f)$,
- the Fourier max-supp-entropy of f, denoted $k'(f)$, defined to be the maximum value of the reciprocal of the absolute value of a non-zero Fourier coefficient,
- the Fourier max-rank-entropy of f, denoted $k''(f)$, defined to be the minimum t such that characters whose coefficients are at least $1/t$ in magnitude span a $r(f)$-dimensional space.

In this work we prove new lower bounds on $\delta(f)$ in terms of the above measures. One of our lower bounds, $\delta(f) = \Omega\left(r(f)^2/(k(f)\log^2 k(f))\right)$, subsumes and refines the previously best known upper bound $r(f) = O(\sqrt{k(f)}\log k(f))$ on $r(f)$ in terms of $k(f)$ by Sanyal (Theory of Computing, 2019). We improve upon this bound and show $r(f) = O(\sqrt{k(f)\delta(f)}\log k(f))$. Another lower bound, $\delta(f) = \Omega\left(r(f)/(k''(f)\log k(f))\right)$, is based on our improvement of a bound by Chattopadhyay, Hatami, Lovett and Tal (ITCS, 2019) on the sum of absolute values of level-1 Fourier coefficients in terms of $\mathbb{F}_2$-degree. We further show that Chang's lemma for the above-mentioned choices of the threshold is asymptotically outperformed by our bounds for most settings of the parameters involved.

Next, we show that our bounds are tight for a wide range of the parameters involved, by constructing functions witnessing their tightness. All the functions we construct are modifications of the Addressing function, where we replace certain input variables by suitable functions. Our final contribution is to construct Boolean functions f for which our lower bounds asymptotically match $\delta(f)$, and for any choice of the threshold t, the lower bound obtained from Chang's lemma is asymptotically smaller than $\delta(f)$.

Our results imply more refined deterministic one-way communication complexity upper bounds for XOR functions. Given the wide-ranging application of Chang's lemma to areas like additive combinatorics, learning theory and communication complexity, we strongly feel that our refinements of Chang's lemma will find many more applications.

[1] Work mostly done while the author was a postdoc at Georgetown University.

41st IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2021).
Editors: Mikołaj Bojańczyk and Chandra Chekuri; Article No. 10; pp. 10:1–10:22
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

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

Keywords and phrases Analysis of Boolean functions, Chang's lemma, Parity decision trees, Fourier dimension

Digital Object Identifier 10.4230/LIPIcs.FSTTCS.2021.10

Related Version *Full Version*: https://arxiv.org/abs/2012.02335 [6]

Funding R.M. thanks DST (India) for grant DST/INSPIRE/04/2014/001799. S.S. is supported by an ISIRD Grant from SRIC, IIT Kharagpur. N.S.M. is supported by the NWO through QuantERA project QuantAlgo 680-91-034.

Acknowledgements T.M. would like to thank Prahladh Harsha and Ramprasad Saptharishi for helpful discussions.

1 Introduction

Chang's lemma [8, 14] is a classical result in additive combinatorics. Informally, the lemma states that all the large Fourier coefficients of the indicator function of a large subset of an Abelian group reside in a low dimensional subspace. The discovery of this lemma was motivated by an application to improve Frieman's theorem on set additions [8]. The lemma has subsequently found many applications in additive combinatorics and combinatorial number theory. Chang's lemma and the ideas developed in Chang's paper [8] have been used to prove theorems about arithmetic progressions in sumsets [12, 23], structure of Boolean functions with small spectral norm [16], and improved bounds for Roth's theorem on three-term arithmetic progressions in the integers [24, 4, 5]. Green and Ruzsa [15] used the ideas of Chang's lemma to prove a generalization of Frieman's theorem for arbitrary Abelian groups. The lemma is known to be sharp for various settings of parameters for the group $\mathbb{Z}_N$ [13].

In this paper, our focus is a specialization of Chang's lemma for the Boolean hypercube. Let $f : \{-1,1\}^n \to \{-1,1\}$ be a Boolean function. For any positive real number t (which we refer to as the *threshold*) define $\mathcal{S}_t := \{S \subseteq [n] : |\widehat{f}(S)| \geq \frac{1}{t}\}$.[2,3] Viewing elements of $\mathcal{S}_t$ as vectors in $\mathbb{F}_2^n$, Chang's lemma gives a lower bound on $\delta(f) := \Pr[f(x) = -1]$ (called the weight of f), in terms of t and the dimension of the span of $\mathcal{S}_t$ (denoted by $\dim(\mathcal{S}_t)$). Formally, we have the following lemma, referred to as Chang's lemma in this paper. In the literature Chang's lemma is more commonly stated as an upper bound on d in terms of $\delta(f)$ and t. We refer the reader to the full version of our paper for a proof of the equivalence of the two forms, and also for other missing proofs from this paper.

▶ **Lemma 1.1** (Chang's lemma [8])**.** *There exists a universal constant $c > 0$ such that the following is true for every integer $n > 0$. Let $f : \{-1,1\}^n \to \{-1,1\}$ be any function and t be any positive real number. Let $\delta(f) := \Pr_x[f(x) = -1]$ and $d = \dim(\mathcal{S}_t) > 1$. If $\delta(f) < c$, then $\delta(f) = \Omega\left(\frac{\sqrt{d}}{t\sqrt{\log(t^2/d)}}\right)$.*

This lemma has found numerous applications in complexity theory and algorithms [2, 7], analysis of Boolean functions [16, 26], communication complexity [26, 19] and extremal combinatorics [10]. See [20] for a proof of Lemma 1.1.

[2] The function f is implicit in the definition of $\mathcal{S}_t$ and will be clear from context.
[3] We refer the reader to Section A for preliminaries on Fourier analysis.

In this paper, we investigate the tightness of Lemma 1.1 for three natural choices of the threshold t based on the Fourier spectrum of the function (see Section 1.1 for details about these thresholds). We prove additional lower bounds on $\delta(f)$, and compare relative performances of all the bounds under consideration. Our results imply that the bounds given by Chang's lemma for the choices of the threshold that we consider are asymptotically outperformed by one of the bounds we prove for a broad range of the parameters involved. For most regimes of the parameters we are able to construct classes of functions that witness the tightness of our bounds.

Interestingly, for each choice of threshold that we consider, $\dim(\mathcal{S}_t)$ equals the Fourier rank of f (denoted by $r(f)$, see Definition A.12). In particular, setting t to be the Fourier sparsity of f (denoted by $k(f)$) leads to a very natural question about the relationship among $r(f), k(f)$ and $\delta(f)$ for a Boolean function f. The best known upper bound on $r(f)$ in terms of $k(f)$ is $r(f) = O(\sqrt{k(f)}\log k(f))$ [25]. We improve upon this bound by incorporating $\delta(f)$ into it, and show $r(f) = O(\sqrt{k(f)\delta(f)}\log k(f))$. Moreover, we also show that this bound is tight; see Section 1.2 for a detailed discussion.

Throughout this paper, we assume that f is not a constant function or a parity or a negative parity (unless mentioned otherwise). In other words, $k(f), r(f) \geq 2$.

1.1 Thresholds considered for Chang's lemma

For a Boolean function f, let $\mathrm{supp}(f)$ denote the Fourier support of f. In this section, we discuss and motivate the choices of the threshold t considered in this work.

a) The Fourier sparsity of f. It was shown in [11, Theorem 3.3] that for all $S \in \mathrm{supp}(f)$, $|\widehat{f}(S)| \geq \frac{1}{k(f)}$. It follows that $\mathcal{S}_{k(f)} = \mathrm{supp}(f)$ and hence $\dim(\mathcal{S}_{k(f)}) = r(f)$. Moreover, there exist functions (e.g. $f = \mathsf{AND}_n$) for which $\dim(\mathcal{S}_t) = 0$ for $t = o(k(f))$, justifying the choice of threshold $k(f)$.

This choice also leads us to a fundamental structural problem of bounding the weight of a Boolean function f from below, in terms of its Fourier sparsity and Fourier rank. The *uncertainty principle* (see, for example, [17] for a statement and a proof) asserts that $\delta(f) = \Omega\left(\frac{1}{k(f)}\right)$. Chang's lemma with $t = k(f)$ and the fact that $\log\left(k(f)^2/r(f)\right) = \Theta(\log k(f))$ (Lemma A.16 (part 1)) implies that

$$\delta(f) = \Omega\left(\frac{1}{k(f)}\sqrt{\frac{r(f)}{\log k(f)}}\right), \tag{1}$$

thereby subsuming the uncertainty principle (note that $r(f)/\log k(f) \geq 1$) and refining it by incorporating $r(f)$ into the bound.

b) The Fourier max-supp-entropy of f. The next choice of the threshold that we consider is the *Fourier max-supp-entropy* of f, denoted by $k'(f)$, which we define to be $\max_{S\in\mathrm{supp}(f)} \frac{1}{|\widehat{f}(S)|}$ (Definition A.15). By its definition $k'(f)$ is the smallest value of t such that $\mathcal{S}_t = \mathrm{supp}(f)$. Since $k'(f) \leq k(f)$ (see the discussion in the last item), the knowledge of $k'(f)$ can potentially offer us a more fine-grained lower bound on $\delta(f)$ than as in the last item; Chang's lemma with $t = k'(f)$ and $\log\left(k'(f)^2/r(f)\right) = \Theta(\log k'(f))$ (Lemma A.16 (part 2)) implies

$$\delta(f) = \Omega\left(\frac{1}{k'(f)}\sqrt{\frac{r(f)}{\log k'(f)}}\right). \tag{2}$$

Notice that Equation (2) subsumes the bound in Equation (1).

In [18] an equivalent statement of the well-known sensitivity conjecture was presented in terms of $k'(f)$.[4] Granularity is another widely-studied measure that is closely associated with Fourier max-supp-entropy.

c) The Fourier max-rank-entropy of f. Our final choice of the threshold is the *Fourier max-rank-entropy of* f, denoted by $k''(f)$, which we define to be the smallest positive real number t such that $\dim(\mathcal{S}_t) = r(f)$ (Definition A.15). We have that $k''(f) \leq k'(f) \leq k(f)$ by their definitions. Amongst all settings of the threshold t for which $\dim(\mathcal{S}_t) = r(f)$, the value $t = k''(f)$ yields the best lower bound from Chang's lemma. Chang's lemma with $t = k''(f)$ implies

$$\delta(f) = \Omega\left(\frac{1}{k''(f)}\sqrt{\frac{r(f)}{\log\left(k''(f)^2/r(f)\right)}}\right), \tag{3}$$

which subsumes the bounds in Equations (2) and (1).

1.2 Our contributions

We prove the following results regarding the three natural instantiations of the threshold t (mentioned in the preceding section) for Chang's lemma.

a) The Fourier sparsity of f. Recall that Chang's lemma with threshold $t = k(f)$ (Equation (1)) implies that $\delta(f) = \Omega\left(\frac{1}{k(f)}\sqrt{\frac{r(f)}{\log k(f)}}\right)$. It was shown in [1] that $\delta(f) = \Omega\left(\frac{1}{k(f)}\left(\frac{r(f)}{\log k(f)}\right)\right)$, improving upon this bound asymptotically (note that $r(f)/\log k(f) \geq 1$). In this work we improve their bound further.

▶ **Theorem 1.2.** *Let $f : \{-1,1\}^n \to \{-1,1\}$ be any function such that $k(f) > 1$. Then* $\delta(f) = \Omega\left(\frac{1}{k(f)}\left(\frac{r(f)}{\log k(f)}\right)^2\right)$.

Observe that the statement of Theorem 1.2 is equivalent to $r(f) = O(\sqrt{k(f)\delta(f)}\log k(f))$. This bound subsumes the bound $r(f) = O(\sqrt{k(f)}\log k(f))$ shown by Sanyal [25]. We prove Theorem 1.2 by incorporating $\delta(f)$ in Sanyal's arguments and thereby refining his proof. See Section 2.1 for the proof of Theorem 1.2.

We also show that Theorem 1.2 is tight. For nearly all admissible values of ρ and κ we construct many Boolean functions f such that $k(f) = O(\kappa)$, $r(f) = O(\rho)$ and $\delta(f) = O\left(\frac{1}{\kappa}\left(\frac{\rho}{\log\kappa}\right)^2\right)$ (Theorem 1.4 and Claim B.2). For a comparison with Sanyal's bound see Section 1.3.

b) The Fourier max-supp-entropy of f. Recall from Section 1.1 that the Fourier max-supp-entropy of f, denoted $k'(f)$, is defined as $k'(f) = \max_{S \in \mathrm{supp}(f)} \frac{1}{|\widehat{f}(S)|}$. It can be shown that $\sqrt{k(f)} \leq k'(f) \leq k(f)/2$ (Lemma A.16 (part 2)). We prove the following lower bound.

▶ **Theorem 1.3.** *Let $f : \{-1,1\}^n \to \{-1,1\}$ be any function such that $k(f) > 1$. Then,* $\delta(f) = \Omega\left(\max\left\{\frac{1}{k(f)}\left(\frac{r(f)}{\log k(f)}\right)^2, \frac{k(f)}{k'(f)^2}\right\}\right)$.

[4] In [18] $\log(k'(f)^2)$ is called the Fourier max-entropy while we refer to $k'(f)$ as the Fourier max-supp-entropy.

As is evident from the statement, Theorem 1.3 presents two lower bounds, one of which is Theorem 1.2. The other lower bound $\delta(f) \geq \frac{k(f)}{k'(f)^2}$ is Claim A.17.

Chang's lemma with the threshold t set to $k'(f)$ (Equation (2)), together with the observation that $\log k(f) = \Theta(\log k'(f))$, implies $\delta(f) = \Omega\left(\frac{1}{k'(f)}\sqrt{\frac{r(f)}{\log k(f)}}\right)$. Theorem 1.3 subsumes this bound since

$$\delta(f) = \Omega\left(\frac{1}{k(f)}\left(\frac{r(f)}{\log k(f)}\right)^2 \cdot \frac{k(f)}{k'(f)^2}\right)^{1/2} = \Omega\left(\frac{1}{k'(f)}\sqrt{\frac{r(f)}{\log k(f)}}\right),$$

where the equality follows from $r(f)/\log k(f) \geq 1$.

In addition, observe from the last equality above that the bound of Theorem 1.3 is asymptotically larger than the bound obtained from Chang's lemma for $t = k'(f)$ (Equation (2)) except when $r(f)/\log k(f) = \Theta(1)$. Theorem 1.4 complements Theorem 1.3 by showing that for nearly all admissible values of $r(f), k(f)$ and $k'(f)$, there exists a function for which the larger of the two bounds presented in Theorem 1.3 is tight.

▶ **Theorem 1.4.** *For all $\rho, \kappa, \kappa' \in \mathbb{N}$ such that κ is sufficiently large, for all constants $\epsilon > 0$ such that $\log\kappa \leq \rho \leq \kappa^{\frac{1}{2}-\epsilon}$ and $\kappa^{\frac{1}{2}} \leq \kappa' \leq \kappa$, there exists a Boolean function $f_{\rho,\kappa,\kappa'}$ such that $r(f_{\rho,\kappa,\kappa'}) = \Theta(\rho)$, $k(f_{\rho,\kappa,\kappa'}) = \Theta(\kappa)$, $k'(f_{\rho,\kappa,\kappa'}) = \Theta(\kappa')$ and*

$$\delta(f_{\rho,\kappa,\kappa'}) = \Theta\left(\max\left\{\frac{1}{\kappa}\left(\frac{\rho}{\log\kappa}\right)^2, \frac{\kappa}{\kappa'^2}\right\}\right).$$

The range of parameters considered in Theorem 1.4 is justified by Lemma A.16. We prove Theorem 1.4 in two parts. Fix any ρ, κ such that $\log\kappa \leq \rho \leq \kappa^{\frac{1}{2}-\epsilon}$ for some constant $\epsilon > 0$. First, for each value of $\kappa' \in [\frac{\kappa\log\kappa}{\rho}, \kappa]$ we construct a function f for which the first lower bound on $\delta(f)$ from Theorem 1.3 is tight (Claim B.2). Next, for each value of $\kappa' \in [\kappa^{\frac{1}{2}}, \frac{\kappa\log\kappa}{\rho}]$ we construct a function f for which the second lower bound on $\delta(f)$ from Theorem 1.3 is tight (Claim B.3). See Figure 1 for a graphical visualization of the bounds in Theorem 1.3 for any fixed values of ρ and κ.

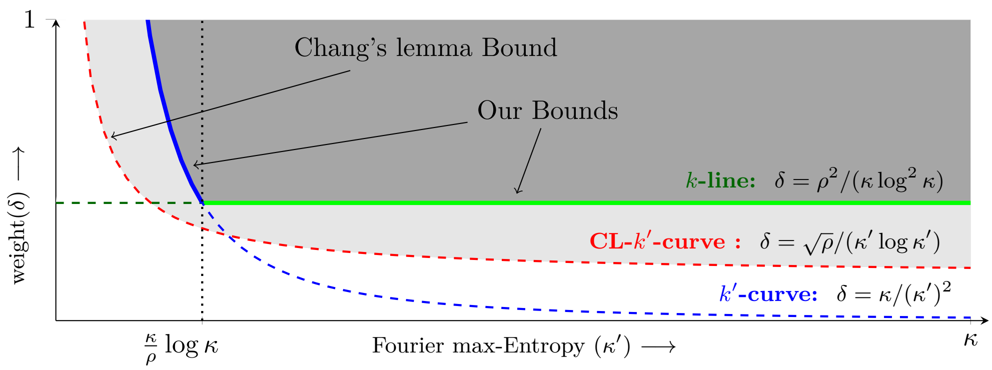

Figure 1 This plot is constructed for any fixed values of ρ, κ for which $\log\kappa \leq \rho \leq \sqrt{\kappa}$, and depicts the relationship between $\delta(f)$ and $k'(f)$ for functions f with $r(f) = \Theta(\rho)$ and $k(f) = \Theta(\kappa)$. For any fixed values of ρ, κ, we will refer to this plot as the (ρ, κ)-k'-plot. Chang's lemma implies that Boolean functions lie above the CL-k'-curve. Theorem 1.3 improves upon Chang's lemma and shows that Boolean functions lie above both the k-line and the k'-curve, highlighted by the dark grey region in the figure. Roughly speaking, Theorem 1.4 exhibits functions that lie on the boundary of the dark grey region described by the k-line and the k'-curve.

c) The Fourier max-rank-entropy of f. Recall from Section 1.1 that the Fourier max-rank-entropy of f, denoted $k''(f)$, is the smallest positive real number t such that $\dim(\mathcal{S}_t) = r(f)$. It can be shown that $\max\left\{\sqrt{r(f)}, \frac{r(f)}{\log k(f)}\right\} \leq k''(f) \leq k(f)$ (Lemma A.16 (part 2)). We prove the following lower bound.

▶ **Theorem 1.5.** *Let $f : \{-1,1\}^n \to \{-1,1\}$ be any function such that $k(f) > 1$. Then,*

$$\delta(f) = \Omega\left(\max\left\{\frac{1}{k(f)}\left(\frac{r(f)}{\log k(f)}\right)^2, \frac{r(f)}{k''(f)\log k(f)}\right\}\right).$$

Theorem 1.5 yields a better lower bound than Chang's lemma with the threshold $t = k''(f)$ (Equation (3)), except when $r(f) < (\log k(f))^2$ (see the caption of Figure 2). Theorem 1.5 presents two lower bounds: the first one is Theorem 1.2, and the second one is Lemma 2.5. Lemma 2.5 is proven by strengthening a bound due to [9] on the sum of absolute values of level-1 Fourier coefficients of a Boolean function in terms of its $\mathbb{F}_2$-degree. A proof of Theorem 1.5 can be found in Section 2.2.

We also show that for nearly all admissible values of $r(f), k(f)$ and $k''(f)$, there exist functions for which the larger of the two bounds presented in Theorem 1.5 is nearly tight.

▶ **Theorem 1.6.** *For all $\rho, \kappa, \kappa'' \in \mathbb{N}$ such that κ is sufficiently large, for all $\epsilon > 0$ such that $\log\kappa \leq \rho \leq \kappa^{\frac{1}{2}-\epsilon}$ and $\rho \leq \kappa'' \leq \kappa$ there exists a Boolean function $f_{\rho,\kappa,\kappa''}$ such that $r(f_{\rho,\kappa,\kappa''}) = \Theta(\rho)$, $k(f_{\rho,\kappa,\kappa''}) = \Theta(\kappa)$, $k''(f_{\rho,\kappa,\kappa''}) = \Theta(\kappa'')$ and*

$$\delta(f_{\rho,\kappa,\kappa''}) = \Theta\left(\max\left\{\frac{1}{\kappa}\left(\frac{\rho}{\log\kappa}\right)^2, \frac{\rho}{\kappa''\log(\kappa''/\rho)}\right\}\right).$$

The range of parameters considered in Theorem 1.6 is justified by Lemma A.16. Theorem 1.6 is proved in two parts. Fix any ρ, κ such that $\log\kappa \leq \rho \leq \kappa^{\frac{1}{2}-\epsilon}$ for some constant $\epsilon > 0$. First, for each value of $\kappa'' \in [\frac{\kappa\log\kappa}{\rho}, \kappa]$ we construct a function f for which the first lower bound on $\delta(f)$ from Theorem 1.5 is tight (Claim B.5). In fact these are the same functions that are used to prove the first bound in Theorem 1.4. Next, for each value of $\kappa'' \in [e\rho, \frac{\kappa\log\kappa}{\rho}]$ we construct a function f for which $\delta(f) = \Theta(\frac{\rho}{\kappa''\log(\kappa''/\rho)})$ (Claim B.4). From the above discussion one may verify that for every ρ, κ that we consider and for every $\kappa'' \geq \rho\cdot\kappa^{\Omega(1)}$, the function that we construct witnesses tightness of the lower bound in Theorem 1.5.

In general, for all settings of ρ, κ and κ'' that we consider, the upper bound on $\delta(f)$ from Theorem 1.6 is off by a factor of at most $O(\log\kappa)$ from the lower bound in Theorem 1.5. See Figure 2 for a graphical visualization of the bounds in Theorem 1.5 for any fixed values of ρ and κ.

Dominating Chang's lemma for all thresholds. Our final contribution is to show that there exists a function for which: our lower bounds (Theorem 1.3 and 1.5) asymptotically match the weight, but for any choice of the threshold the lower bound obtained from Chang's lemma (Lemma 1.1) is asymptotically smaller than the weight. See [6, Section 7] in the full version of our paper for a proof of the below claim.

▷ Claim 1.7 (Beating Chang's lemma for all thresholds). For any integer $t > 4$ there exists a function $f : \{-1,1\}^{\log t} \times \{-1,1\}^{t\log t} \to \{-1,1\}$ such that

- $\delta(f) = \frac{1}{t}$.
- For all real $x > 0$ for which $\dim(\mathcal{S}_x) > 1$, we have $\frac{\sqrt{\dim(\mathcal{S}_x)}}{x\sqrt{\log(x^2/\dim(\mathcal{S}_x))}} = O\left(\frac{1}{t^{3/2}}\right)$.
- $$\frac{1}{k(f)}\left(\frac{r(f)}{\log k(f)}\right)^2 = \Omega\left(\frac{1}{t}\right), \quad \frac{k(f)}{k'(f)^2} = \Omega\left(\frac{1}{t}\right) \quad \text{and} \quad \frac{r(f)}{k''(f)\log k(f)} = \Omega\left(\frac{1}{t}\right).$$

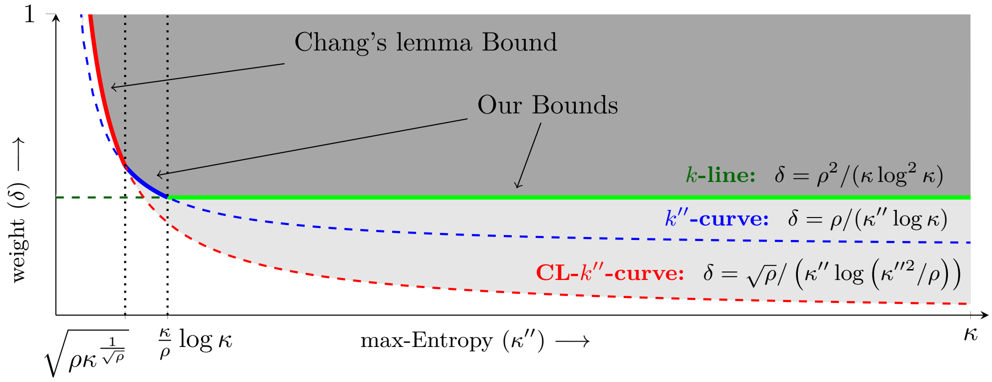

Figure 2 This plot is constructed for any fixed values of ρ, κ for which $\log \kappa \leq \rho \leq \sqrt{\kappa}$, and depicts the relationship between $\delta(f)$ and $k''(f)$ for functions f with $r(f) = \Theta(\rho)$ and $k(f) = \Theta(\kappa)$. For any fixed values of ρ, κ, we will refer to this plot as (ρ, κ)-k''-plot. Chang's lemma implies that Boolean functions lie above the CL-k''-curve. Theorem 1.5 improves upon Chang's lemma and shows that Boolean functions lie above both the k-line and the k''-curve, highlighted by the dark grey region in the figure. Although the picture indicates that the CL-k''-curve is better than the k''-curve for certain ranges of κ'', this is actually only possible for certain values of ρ and κ. This is because the CL-k''-curve and the k''-curve intersect at $\sqrt{\rho\kappa^{1/\sqrt{\rho}}}$, which is less than $\sqrt{\rho}$ if $\rho \geq (\log \kappa)^2$. By Lemma A.16 we know that for any function f on this plot, the range of $k''(f)$ is between $\max\{\sqrt{\rho}, \rho/\log \kappa\}$ and κ. Thus our bounds in Theorem 1.5 dominate those given by the CL-k''-curve in all (ρ, κ)-k'' plots where $\rho \geq \log^2 \kappa$.

In particular, Claim 1.7 shows that our bounds can be strictly stronger than those given by Chang's lemma, in the following sense.

- All the lower bounds on $\delta(f)$ from Theorems 1.3 and 1.5 are tight, as witnessed by f from Claim 1.7.
- For the function f from Claim 1.7, no matter what threshold x is chosen in Lemma 1.1, the best possible lower bound on $\delta(f)$ that we get can get from Lemma 1.1 is $\Omega\left(\frac{1}{t^{3/2}}\right)$. This is polynomially smaller than $1/t$, the actual weight of f.

1.3 Applications of our results

An application of our result is an enhanced understanding of the bound $r(f) = O(\sqrt{k(f)} \log k(f))$ proven by Sanyal [25]. This bound is a special case of Theorem 1.2 for $\delta(f) = \Theta(1)$. It is not known whether the $\log k(f)$ term is required in Sanyal's upper bound on $r(f)$ (when f equals the Addressing function, $r(f) = \Omega(\sqrt{k(f)})$, see Definition A.10 and Observation A.19). For all the functions we construct witnessing the tightness of the bound in Theorem 1.2, $\delta(f) = o(1)$. We prove Theorem 1.2 by generalizing Sanyal's proof. As stated before, our bound is tight in this generality, i.e. the logarithmic factor is required in the upper bound on $r(f)$. This sheds light on the presence of the logarithmic term in the bound $r(f) = O(\sqrt{k(f)} \log k(f))$.

Also, Fourier sparsity and Fourier rank of f have intimate connections with the communication complexity of functions of the form $F := f \circ \mathsf{XOR}$. The Fourier sparsity of f equals the real rank ($\mathsf{rank}(M_F)$) of the communication matrix M_F of F, and the Fourier rank of f equals the deterministic (and even exact quantum) one-way communication complexity of F [22]. Theorem 1.2 thus implies an improved upper bound of $O(\sqrt{k(f)\delta(f)} \log k(f))$ on the one-way communication complexity of F in these models, which asymptotically beats the best known upper bound of $O(\sqrt{\mathsf{rank}(M_F)})$ even for two-way protocols [26, 21], for the special case of functions of this form (when $\delta(f) = o(1/\log k)$).

Given the wide-ranging application of Chang's lemma to areas like additive combinatorics, learning theory and communication complexity, we strongly feel that our refinements of Chang's lemma will find many more applications.

2 Lower bound proofs

For lower bounds on $\delta(f)$ of a Boolean function f, we need to prove two theorems: Theorems 1.3 and 1.5. The proof of Theorem 1.3 is given in Section 2.1 and the proof of Theorem 1.5 is given in Section 2.2.

2.1 Proof of Theorem 1.3 (and Theorem 1.2)

Remember that we defined the *Fourier max-supp-entropy* of a Boolean function f, denoted by $k'(f)$, to be $\max_{S\in\mathrm{supp}(f)} \frac{1}{|\widehat{f}(S)|}$.

The main aim of this section is to give a lower bound on $\delta(f)$ with respect to $k'(f)$ for a Boolean function f (Theorem 1.3).

We first prove Theorem 1.2 which implies Theorem 1.3 (together with Claim A.17).

Theorem 1.2 can be viewed as an upper bound of $O(\sqrt{k(f)\delta(f)}\log k(f))$ on the Fourier rank of f. In order to prove Theorem 1.2, we give an algorithm (Algorithm 1) which takes a Boolean function f as input and outputs a set of $O(\sqrt{\delta(f)k(f)}\log k(f))$ parities such that any assignment of these parities makes the function constant. From Observation A.14, this implies an upper bound of $O(\sqrt{\delta(f)k(f)}\log k(f))$ on Fourier rank of the function. We start by formally describing this algorithm. The central ingredient in the algorithm is a lemma in [26, Lemma 28].

▶ **Lemma 2.1** ([26]). *Let $f : \{-1,1\}^n \to \{-1,1\}$ a function. There is an affine subspace $V \subseteq \{-1,1\}^n$ of co-dimension at most $3\sqrt{\delta(f)k(f)}$ such that f is constant on V.*

Recall that for a function $f : \{-1,1\}^n \to \{-1,1\}$, a set of parities Γ and an assignment $b \in \{-1,1\}^\Gamma$, we define the restriction $f|_{(\Gamma,b)} := f|_{\{x\in\{-1,1\}^n:\chi_\gamma(x)=b_\gamma \text{ for all } \gamma\in\Gamma\}}$. Also let $\mathcal{B}_\Gamma := \{b \in \{-1,1\}^\Gamma : f|_{(\Gamma,b)}$ *is not constant*$\}$.

Algorithm 1

Input: A function $f : \{-1,1\}^n \to \{-1,1\}$.
Output: A set Γ of parities whose evaluation determines f.
Initialization: $f_{\min} \leftarrow f$, $\Gamma \leftarrow \emptyset$.
while $\mathcal{B}_\Gamma$ *is non-empty* **do**

- **(a) Update Γ:** Let Γ' be the smallest set of parities, such that, there exists $b \in \{-1,1\}^{\Gamma'}$ for which $f_{\min}|_{(\Gamma',b)}$ is constant,

$$\Gamma \leftarrow \Gamma \cup \Gamma'.$$

- **(b) Update $f_{\min}$:** Define $b^* := \mathrm{argmin}_{b\in\mathcal{B}_\Gamma} \left\{\frac{\delta(f|_{(\Gamma,b)})}{k(f|_{(\Gamma,b)})}\right\}$, and update

$$f_{\min} \leftarrow f|_{(\Gamma,b^*)}.$$

end
Return Γ.

Since number of parities are finite and we fix at least one parity at each iteration of Step a of the **while** loop, the algorithm terminates. The termination condition implies that the algorithm outputs a set of parities Γ such that for any assignment $b \in \{-1,1\}^{\Gamma}$ of Γ, the restricted function $f_{(\Gamma,b)}$ becomes constant.

The only remaining step is to show is that the number of parities fixed in Algorithm 1 is $O(\sqrt{\delta(f)k(f)}\log k(f))$. For this we define an equivalence relation and observe a few properties of restricted functions (restricted according to an assignmen t of a set of parities).

Equivalence relation for a set of parities

Let f be the input to Algorithm 1, first we define an equivalence relation given a set of parities over the variables of f. Given a set of parities Γ, define the following equivalence relation among parities in $\mathrm{supp}(f)$.

$$\forall \gamma_1, \gamma_2 \in \mathrm{supp}(f), \gamma_1 \equiv \gamma_2 \text{ iff } \gamma_1 + \gamma_2 \in \mathrm{span}(\Gamma). \tag{4}$$

Let ℓ be the number of equivalence classes according to the equivalence relation for Γ. For $j \in [\ell]$, let k_j be the size of the j-th equivalence class. Since the equivalence classes form a partition of $\mathrm{supp}(f)$, we have

▶ **Observation 2.2.** *Following the notation of the paragraph above,* $\sum_{j=1}^{\ell} k_j = k(f)$.

Let $\beta_1, \ldots, \beta_\ell \in \mathrm{supp}(f)$ be some representatives of the equivalence classes. For $j \in [\ell]$, let $\beta_j + \alpha_{j,1}, \ldots, \beta_j + \alpha_{j,k_j}$ be the elements of the j-th equivalence class. This notation gives a compact representation of f in terms of these equivalence classes. For all $x \in \{-1,1\}^n$,

$$f(x) = \sum_{j=1}^{\ell} P_j(x)\chi_{\beta_j}(x), \tag{5}$$

where

$$P_j(x) = \sum_{r=1}^{k_j} \widehat{f}(\beta_j + \alpha_{j,r}) \cdot \chi_{\alpha_{j,r}}(x). \tag{6}$$

Note that P_j are non-zero multilinear polynomials and depend only on the parities in Γ. So, fixing parities in Γ collapses all the parities in an equivalence class to their representative, thereby making P_j's constant.

We will denote Γ after the i-th iteration of the **while** loop by $\Gamma^{(i)}$ (so $\Gamma^{(0)} = \emptyset$). Let $f^{(i)}_{\min}$ be the selected function $f_{\min}$ after the i-th iteration (thus $f^{(0)}_{\min} = f$).

With the above properties of restricted functions we are ready to prove the main technical lemma needed to show Theorem 1.2.

▶ **Lemma 2.3.** *Let* $f : \{-1,1\}^n \to \{-1,1\}$ *a function. Suppose* Γ *be a set of parities and* ℓ *be the number of equivalence classes of* $\mathrm{supp}(f)$ *under the equivalence relation defined by in Equation* (4)*, Then, there exists a* $b \in \{-1,1\}^{\Gamma}$ *such that* $f|_{(\Gamma,b)}$ *is non-constant and* $\frac{\delta(f|_{(\Gamma,b)})}{k(f|_{(\Gamma,b)})} \leq \frac{4k(f)\delta(f)}{\ell^2}$.

Proof. For the sake of succinctness, when Γ is clear from the context, let $V_b = \{x \in \{-1,1\}^n : \forall \gamma \in \Gamma, x_\gamma = b_\gamma\}$, for all $b \in \{-1,1\}^{\Gamma}$, and $f|_b = f|_{\{x : x \in V_b\}}$.

Since we are interested in a non-constant $f|_b$, define $k_{\{\emptyset\}^c}(f)$ to be the number of non-zero non-empty monomials in Fourier representation of f. We first need to prove the following two bounds on the expected values of $\delta(f|_b)$ and $k_{\{\emptyset\}^c}(f|_b)$.

- $\mathbb{E}_b\left[\delta(f|_b)\right] = \delta(f)$,
- $\mathbb{E}_b\left[k_{\{\emptyset\}^c}(f|_b)\right] \geq \frac{\ell^2}{4k(f)}$.

Expected value of $\delta(f|_b)$. Since $\left\{V_b : b \in \{-1,1\}^{\Gamma^{(i)}}\right\}$ form a partition on $\{-1,1\}^n$ and all partitions are of the same size, we get the expected value of $\delta(f|_b)$.

$$\mathbb{E}_b\left[\delta(f|_b)\right] = \delta(f). \tag{7}$$

Expected value of $k_{\{\emptyset\}^c}(f|_b)$. From Equation (5), for all $b \in \{-1,1\}^{\Gamma}$ and for all $x \in \{-1,1\}^n$,

$$f|_b(x) = \sum_{j=1}^{\ell} P_j(b)\chi_{\beta_j}(x). \tag{8}$$

For each $j \in [\ell]$ and $b \in \{-1,1\}^{\Gamma}$, let $I_j(b)$ be the indicator function for $P_j(b) \neq 0$,

$$I_j(b) = \begin{cases} 1 & \text{if } P_j(b) \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

From Equation (6), each P_j is a polynomial having monomials $\{\chi_{\alpha_{j,r}} : r \in [k_j]\}$ with Fourier sparsity of P_j being equal to k_j. Since each P_j is a non-zero polynomial, by Lemma A.2

$$\mathbb{E}_b\left[I_j(b)\right] = \Pr_{b\sim\{-1,1\}^{\Gamma}}\left[P_j(b) \neq 0\right] \geq \frac{1}{k_j}. \tag{9}$$

We calculate the expectation of $k_{\{\emptyset\}^c}(f|_b)$.

$$\begin{aligned}
\mathbb{E}_b\left[k_{\{\emptyset\}^c}(f|_b)\right] &= \mathbb{E}_b\left[\sum_{j=1}^{\ell-1} I_j(b)\right] && \text{by Equation (8)}\\
&= \sum_{j=1}^{\ell-1}\mathbb{E}_b\left[I_j(b)\right] && \text{by linearity of expectation}\\
&\geq \sum_{j=1}^{\ell-1}\frac{1}{k_j} && \text{by Equation (9)}\\
&\geq \frac{(\ell-1)^2}{\sum_{j=1}^{\ell-1} k_j} && \text{by Cauchy-Schwarz inequality}\\
&\geq \frac{\ell^2}{4k(f)}. && \text{by Observation 2.2}
\end{aligned}$$

To finish the proof of the theorem, we use bounds on the two expected values,[5]

$$\begin{aligned}
&\frac{\mathbb{E}_b\left[\delta(f|_b)\right]}{\mathbb{E}_b\left[k_{\{\emptyset\}^c}(f|_b)\right]} \leq \frac{4k(f)\delta(f)}{\ell^2}\\
\iff &\mathbb{E}_b\left[\delta(f|_{V_b}) - \frac{4k(f)\delta(f)}{\ell^2}k_{\{\emptyset\}^c}(f|_{V_b})\right] \leq 0. && \text{by linearity of expectation}
\end{aligned}$$

[5] this part of our proof is inspired by a proof of the Cheeger's inequality in spectral graph theory. See, for example, the proof of Fact 2 in `https://people.eecs.berkeley.edu/~luca/expanders2016/lecture04.pdf`.

If $\delta(f|_{V_b}) - \frac{4k(f)\delta(f)}{\ell^2}k_{\{\emptyset\}^c}(f|_{V_b}) = 0$ for all b, then pick any non-constant $f|_b$. Otherwise, there exists a b_0 such that $\delta(f|_{V_{b_0}}) - \frac{4k(f)\delta(f)}{\ell^2}k_{\{\emptyset\}^c}(f|_{V_{b_0}}) < 0$. Since this equation can only be satisfied when $k_{\{\emptyset\}^c}(f|_{V_{b_0}}) > 0$, $f|_{V_{b_0}}$ is not constant. Dividing by $k_{\{\emptyset\}^c}(f|_{V_{b_0}})$,

$$\frac{\delta(f|_{b_0})}{k(f|_{b_0})} \leq \frac{\delta(f|_{b_0})}{k_{\{\emptyset\}^c}(f|_{b_0})} \leq \frac{4k(f)\delta(f)}{\ell^2},$$

and $f|_{b_0}$ is non-constant. ◀

Lemma 2.3 allows us to bound the number of parities fixed in the i-th iteration (in terms of the decrease in number of equivalence classes).

▶ **Lemma 2.4.** *Suppose f is given as input to Algorithm 1. Consider the i-th iteration of Algorithm 1. Let q_i be the be number of parities fixed in Step a of the i-th iteration of the* **while** *loop, and ℓ_i be the number of equivalence classes after Step a of the i-th iteration. Then*

$$\frac{q_i}{(\ell_{i-1} - \ell_i)} \leq \frac{6\sqrt{\delta(f)k(f)}}{\ell_{i-1}}.$$

Proof. Recall that $\Gamma = \Gamma^{(i)}$ after the i-th of Step a of Algorithm 1. Again, for the sake of succinctness, let $V_b = \{x \in \{-1,1\}^n : \forall \gamma \in \Gamma^{(i)}, x_\gamma = b_\gamma\}$, for all $b \in \{-1,1\}^{\Gamma^{(i)}}$, and $f|_b = f|_{\{x : x \in V_b\}}$. Let $f_{\min}$ be the function chosen after the i-th iteration of Step b of Algorithm 1. Since Step b of Algorithm 1 chooses $f_{\min}$ to be a non-constant function such that weight-to-sparsity ratio is minimized, from Lemma 2.3 we have,

$$\frac{\delta(f_{\min})}{k(f_{\min})} \leq \frac{4k(f)\delta(f)}{\ell_{i-1}^2}. \tag{10}$$

Write every $f|_b$ as in Equation (5), and define $\mathcal{S}^{(i)} := \bigcup_{b \in \{-1,1\}^{\Gamma^{(i)}}} \mathrm{supp}(f|_b)$. We now prove that $|\mathcal{S}^{(i)}| = \ell_i$.

- $|\mathcal{S}^{(i)}| \leq \ell_i$: Follows from the representation in Equation (5), since each $\mathrm{supp}(f|_b)$ is a subset of $\{\chi_{\beta_j^{(i)}} \mid j \in [\ell_i]\}$.
- $|\mathcal{S}^{(i)}| \geq \ell_i$: Since $P_j^{(i)}$ is a non-zero polynomial, there exists an assignment to parities in $\Gamma^{(i)}$, such that, $P_j^{(i)}$ is non-zero. Thus, for all $j \in [\ell_i]$, we have $\chi_{\beta_j^{(i)}} \in \mathcal{S}^{(i)}$.

Since $|\mathcal{S}^{(i)}| = \ell_i$, Lemma 2.1 guarantees that $q_i \leq 3\sqrt{k(f_{\min})\delta(f_{\min})}$. Since $f_{\min}$ becomes constant after fixing these q_i parities, every parity in $\mathrm{supp}(f_{\min})$ is paired with at least one other parity in $\mathrm{supp}(f_{\min})$ for the equivalence class with respect to $\Gamma^{(i)}$.[6] This implies that $\ell_{i-1} - \ell_i \geq \frac{k(f_{\min})}{2}$ Combining the two inequalities in the last paragraph we have,

$$\frac{q_i}{(\ell_{i-1} - \ell_i)} \leq 6\sqrt{\frac{\delta(f_{\min})}{k(f_{\min})}}.$$

From Equation (10),

$$\frac{q_i}{(\ell_{i-1} - \ell_i)} \leq \frac{6\sqrt{\delta(f)k(f)}}{\ell_{i-1}}. \tag{11}$$

◀

[6] There is a boundary case ($k(f) = 1$) which can be dealt with separately, as in [25, Lemma 3.4]. For readability, we assume $k(f) \geq 2$.

We are now ready to prove Theorem 1.2.

Proof of Theorem 1.2. We only need to show that the number of parities fixed in Algorithm 1 is $O(\sqrt{\delta(f)k(f)}\log k(f))$ (Observation A.14). Suppose the while loop runs for t iterations. Let q_i be the number of queries made in Step a of Algorithm 1 in the i-th iteration. From Lemma 2.3, we have

$$q_i \leq \frac{6\sqrt{\delta(f)k(f)}}{\ell_{i-1}}(\ell_{i-1} - \ell_i)$$

Thus when Algorithm 1 is run of f, the total number of queries made by the algorithm is

$$\begin{aligned}\sum_{i=1}^{t} q_i &\leq 6\sqrt{\delta(f)k(f)}\sum_{i=1}^{t}\frac{(\ell_{i-1} - \ell_i)}{\ell_{i-1}}\\ &\leq 6\sqrt{\delta(f)k(f)}\sum_{i=1}^{t}\left(\frac{1}{\ell_{i-1}} + \frac{1}{\ell_{i-1}-1}\dots + \frac{1}{\ell_i + 1}\right)\\ &\leq 6\sqrt{\delta(f)k(f)}\sum_{i=1}^{\ell_0}\frac{1}{i}\\ &\leq 6\sqrt{\delta(f)k(f)}\log \ell_0 = 6\sqrt{\delta(f)k(f)}\log k(f).\end{aligned}$$

Observation A.14 implies $r(f) = O(\sqrt{\delta(f)k(f)}\log k(f))$. ◀

Along with Theorem 1.2, this proves Theorem 1.3.

Proof of Theorem 1.3. The bound $\delta(f) = \Omega\left(\frac{1}{k(f)}\left(\frac{r(f)}{\log k(f)}\right)^2\right)$ follows from Theorem 1.2 and the bound $\delta(f) = \Omega\left(\frac{k(f)}{(k'(f))^2}\right)$ from Claim A.17. ◀

2.2 Proof of Theorem 1.5

Recall that we defined *max-rank-entropy* of a Boolean function f, denoted by $k''(f)$, to be $\text{argmin}_t\{\dim(\mathcal{S}_t)\} = r(f)$. The main aim of this section is to give a lower bound on $\delta(f)$ with respect to $k''(f)$ for a Boolean function f (Theorem 1.5). The second bound of Theorem 1.5 is given by the following lemma.

▶ **Lemma 2.5.** *Let $f : \{-1,1\}^n \to \{-1,1\}$ be any function such that $k(f) > 1$. Then, $\delta(f) = \Omega\left(\frac{r(f)}{k''(f)\log k(f)}\right)$.*

Together with Theorem 1.2 proved in Section 2.1, Lemma 2.5 implies Theorem 1.5. We now give the proof of Lemma 2.5.

Lemma 2.5 gives a lower bound of $\Omega\left(\frac{r(f)}{k''(f)\log k(f)}\right)$ on $\delta(f)$. The crucial ingredient for this lower bound is Lemma 2.7, which is a refinement of the following theorem.

▶ **Theorem 2.6** ([9, Theorem 13])**.** *Let $f : \{-1,1\}^n \to \{-1,1\}$ be any function such that $\deg_{\mathbb{F}_2}(f) = d$. Then, $\sum_{i\in[n]}|\widehat{f}(\{i\})| \leq 4d$.*

▶ **Lemma 2.7.** *For any Boolean function f, $\sum_{i=1}^{n}|\widehat{f}(i)| = O(\delta(f)\deg_{\mathbb{F}_2}(f))$.*

The proof of Lemma 2.7 for a Boolean function f essentially applies Theorem 2.6 on the xor of disjoint copies of f. The only difference in the statement of Lemma 2.7 and Theorem 2.6 is that the right hand side becomes $O(\delta(f)\cdot\deg_{\mathbb{F}_2}(f))$ instead of $4\deg_{\mathbb{F}_2}(f)$.

Proof of Lemma 2.7. Assume $\delta(f) \leq 1/4$ (otherwise Theorem 2.6 implies $\sum_{i=1}^{n} |\widehat{f}(i)| = O(\delta(f)d)$). Define $F : \{-1,1\}^{nt} \to \{-1,1\}$ to be $F(x^{(1)},\dots,x^{(t)}) := f(x^{(1)}) \times \dots \times f(x^{(t)})$, where t is a parameter to be fixed later, and $x^{(i)} \in \{-1,1\}^n$ for all $i \in [t]$. Since $\deg_{\mathbb{F}_2}(F) = \deg_{\mathbb{F}_2}(f)$, Theorem 2.6 implies

$$\sum_{\substack{S\subseteq[nt]\\|S|=1}} |\widehat{F}(S)| = O(d). \tag{12}$$

Since $(1-x)^{1/x}$ is a decreasing function in x for $x \in (0,1/2]$, we have

$$(1-x)^{1/x} \geq 1/4 \quad \text{for all } x \in (0,1/2]. \tag{13}$$

Expressing the Fourier coefficients of F in terms of the Fourier coefficients of f,

$$\begin{aligned}
\sum_{\substack{S\subseteq[nt]\\|S|=1}} |\widehat{F}(S)| &= t \cdot \widehat{f}(\emptyset)^{t-1} \sum_{i=1}^{n} |\widehat{f}(i)| \\
&= \left(1+\frac{1}{2\delta(f)}\right) \cdot (1-2\delta(f))^{\frac{1}{2\delta(f)}} \sum_{i=1}^{n} |\widehat{f}(i)| \\
&\qquad\qquad \text{Choosing } t = 1+\tfrac{1}{2\delta(f)}, \text{ and since } \widehat{f}(\emptyset) = 1-2\delta(f) \\
&\geq \left(1+\frac{1}{2\delta(f)}\right) \cdot \left(\frac{1}{4}\right) \sum_{i=1}^{n} |\widehat{f}(i)| \qquad \text{by Equation (13)} \\
&\geq \frac{1}{8\delta(f)} \cdot \sum_{i=1}^{n} |\widehat{f}(i)|.
\end{aligned}$$

Now, Equation (12) implies the desired bound, $\sum_{i=1}^{n} |\widehat{f}(i)| = O(\delta(f)d)$. ◀

We would like to extend the upper bound of Lemma 2.7 to any basis of span(supp(f)) instead of just the standard basis of the set of parities.

▶ **Corollary 2.8.** *Let $f : \{-1,1\}^n \to \{-1,1\}$ be any function with $\deg_{\mathbb{F}_2}(f) = d$. Suppose $\mathcal{S} \subseteq \mathrm{supp}(f)$ is a basis of $\mathrm{span}(\mathrm{supp}(f))$, then*

$$\sum_{S\in\mathcal{S}} |\widehat{f}(S)| = O(\delta(f)d) = O(\delta(f)\log k(f)).$$

Proof. The main idea of the proof is to do a basis change on parities and construct another function h, the corollary will follow by applying Lemma 2.7 on h.

Recall that we denote both a subset of $[n]$ and the corresponding indicator vector in $\mathbb{F}_2^n$, by the same notation.

Let $\mathcal{S} = \{S_1,\dots,S_{r(f)}\}$, extend $\mathcal{S}$ to $\mathcal{S}' = \{S_1,\dots,S_{r(f)},S_{r(f)+1},\dots,S_n\}$, a complete basis of $\mathbb{F}_2^n$. Observe that $\widehat{f}(S_i) = 0$, for $i \in \{r(f)+1,\dots,n\}$ (since $\mathcal{S}$ spans supp(f)). Fix the change of basis matrix $B \in \mathbb{F}_2^{n\times n}$ with i-th column as S_i, $i \in [n]$.

Consider the function $h : \{-1,1\}^n \to \mathbb{R}$ satisfying $\widehat{h}(\alpha) = \widehat{f}(B\alpha)$, for all $\alpha \in \mathbb{F}_2^n$. By Claim A.4, h is Boolean and $\deg_{\mathbb{F}_2}(h) = \deg_{\mathbb{F}_2}(f)$. Using Lemma 2.7, $\sum_{i\in[n]} |\widehat{h}(\{i\})| = O(\delta(f)d)$. From the definition of h, $\widehat{h}(e_i) = \widehat{f}(S_i)$ for $i \in [r(f)]$ and $\widehat{h}(e_i) = 0$ for $i \in \{r(f)+1,\dots,n\}$, we have $\sum_{S\in\mathcal{S}} |\widehat{f}(S)| = O(\delta(f)d)$. The second equality in the statement of the lemma follows from Lemma A.3. ◀

Proof of Lemma 2.5. Observe that every term on the left hand side of Corollary 2.8 is bigger than $1/k''(f)$, giving the required lower bound on $\delta(f)$ and finishing the proof of Lemma 2.5. ◀

Proof of Theorem 1.5. From Lemma 2.5 we have $\delta(f) = \Omega\left(\frac{r(f)}{k''(f)\log k(f)}\right)$, and from Theorem 1.2 we have $\delta(f) = \Omega\left(\frac{r(f)^2}{k(f)\log^2 k(f)}\right)$. ◀

The following corollary combines the lower bounds on $\delta(f)$ from Theorem 1.5 and Lemma 1.1 by setting $k''(f)$ as the threshold.

▶ **Corollary 2.9.** *Let $f : \{-1,1\}^n \to \{-1,1\}$ be any function such that $k(f) > 1$. Then,*

$$\delta(f) = \Omega\left(\max\left\{\frac{r(f)^2}{k(f)\log^2 k(f)}, \frac{r(f)}{k''(f)\log k(f)}, \frac{\sqrt{r(f)}}{k''(f)\log(k''(f)^2/r(f)}\right\}\right).$$

3 Proof techniques for upper bound results

In this section we give the overview of our two upper bound results, Theorems 1.4 and 1.6. For presenting the overview of the proofs of these theorems we will use (ρ,κ)-k'-plots (Figure 1) and (ρ,κ)-k''-plots (Figure 2), respectively. In an (ρ,κ)-k'-plot ((ρ,κ)-k''-plot, respectively) we will refer to the "intersection point" as the point of intersection between the k-line and k'-curve (the point of intersection between the k-line and k''-curve, respectively). Which intersection point we are referring to should be clear from the context.

3.1 Proof techniques for Theorem 1.4

To prove Theorem 1.4, we split our goal into two natural parts: constructing functions on the k-line and constructing functions on the k'-curve. Both the classes of functions are modifications of the Addressing function (Definition A.10). In these modifications, all or some of the target variables of the Addressing function are replaced with an AND function or a Bent function or a combination of them. We first provide a description of some functions that lie on the intersection point. While we do not require this, we choose to describe these functions in order to provide more intuition.

Construction of functions at the intersection point in any (ρ,κ)-k'-plot. Note that a function lies at the intersection point when

$$k'(f) = \frac{k(f)\log(k(f))}{r(f)}. \tag{14}$$

Thus, we want to construct a function f with $k(f) = \Theta(\kappa)$, $r(f) = \Theta(\rho)$, $k'(f) = \Theta\left(\frac{\kappa\log\kappa}{\rho}\right)$ and $\delta(f) = \rho^2/\kappa(\log^2\kappa)$. In particular, we want to construct functions for all ρ,κ satisfying $\log\kappa \le \rho \le \kappa^{\frac{1}{2}}$. Note that, the Addressing function $\mathsf{AD}_t : \{-1,1\}^{\log t + t} \to \{-1,1\}$ has sparsity t^2, rank $(t+\log t)$, max-supp-entropy t and weight $1/2$ (Observation A.19) and thus, AD_t satisfies Equation (14). This only gives functions on the intersection point on all (ρ,κ)-k'-plots where $\rho = \Theta(\sqrt{\kappa})$, while we have to exhibit such functions for all (ρ,κ)-k'-plots where $\log\kappa \le \rho = O(\sqrt{\kappa})$.

Our next step is to tweak AD_t in such a way that the rank of the new function f does not change significantly while the sparsity and max-supp-entropy both increase by the same multiplicative factor. This would ensure that the resulting function satisfies Equation (14). If the resulting function's weight decreases to the required value, we would have a function at the intersection point.

In order to tweak AD_t, we consider a special kind of composed function $f := \mathsf{AD}_t \circ_{\text{target}} g$ (see Definition A.11 for a precise definition) obtained by replacing each target variable in the addressing function with a function g where each copy of g acts on a set of new variables. We prove a *composition lemma* (Lemma B.1) that gives the properties of such composed functions. Due to the structure of the Fourier spectrum of the Addressing function, Lemma B.1 gives us $r(f) \approx t \cdot r(g)$, $k(f) \approx t^2 \cdot k(g)$, $k'(f) = t \cdot k'(g)$ and $\delta(f) = \delta(g)$.

So, if g is a function on a small number of variables (say $\log t'$) with near-maximal sparsity and max-supp-entropy ($\Theta(t')$), then the resulting function satisfies Equation (14). The AND function is a natural choice for g. We denote the resulting function by $\mathsf{AD}_{t,t'}$, and this is a function at the intersection point for all plots by suitably varying t and t'.

Constructing functions on the k-line. We start with $\mathsf{AD}_{t,t'}$, the function at the intersection point in (ρ,κ)-k'-plots. We modify $\mathsf{AD}_{t,t'}$ in such a way that its sparsity, rank and weight do not change much, while the max-supp-entropy increases. We replace a single $\mathsf{AND}_{\log t'}$ in $\mathsf{AD}_{t,t'}$ by $\mathsf{AND}_{\log a}$ for some suitable $a > t$, denote the new function by $\mathsf{AD}_{t,t',a}$. A suitable setting of the parameters t, t' and a yields functions on the k-line for all plots (Claim B.2).

Constructing functions on the k'-curve of the (ρ,κ)-k'-plot. We start with $\mathsf{AD}_{t,t'}$ at the intersection point on $(\rho,\kappa/\ell)$-k'-plot (for some parameter $\ell > 0$). We modify $\mathsf{AD}_{t,t'}$ in such a way that its rank and weight do not change, the sparsity increases by a multiplicative factor of ℓ and the max-supp-entropy increases by a factor of $\sqrt{\ell}$. The new function f will be on the k'-curve in the (ρ,κ)-k'-plot because $\frac{k(f)}{k'(f)^2} = \frac{k(\mathsf{AD}_{t,t'})}{k'(\mathsf{AD}_{t,t'})^2} = \delta(\mathsf{AD}_{t,t'}) = \delta(f)$. Note that $k'(f) \approx \frac{\kappa \log(\kappa)}{\rho\sqrt{\ell}}$, thus making ℓ suitably large yields functions on the k'-curve for all $\rho \leq \kappa' \leq \frac{\kappa\log(\kappa)}{\rho}$ for all plots.

We now change $\mathsf{AD}_{t,t'}$ to have the properties mentioned above. We modify each $\mathsf{AND}_{\log t'}$ in $\mathsf{AD}_{t,t'}$ as follows: replace a single variable x by $x \cdot B$, where B is a bent function on $\log \ell$ new variables. We denote this new inner function by $\mathsf{AB}_{t',\ell}$, and $\mathsf{AD}_t \circ_{\text{target}} \mathsf{AB}_{t',\ell}$ by $\mathsf{AAB}_{t,t',\ell}$. The effect of changing $\mathsf{AND}_{\log t'}$ to $\mathsf{AB}_{t',\ell}$ keeps its rank and weight roughly the same, while increasing its sparsity by a factor of ℓ and increasing its max-supp-entropy by a factor of $\sqrt{\ell}$. We show, using our composition lemma (Lemma B.1), that the properties of $\mathsf{AD}_t \circ_{\text{target}} \mathsf{AND}_{\log t'}$ and $\mathsf{AD}_t \circ_{\text{target}} \mathsf{AB}_{t',\ell}$ change in a similar fashion. Thus, a suitable setting of the parameters t, t', ℓ yields functions on the k'-curve for all plots (Claim B.3).

3.2 Proof techniques for Theorem 1.6

We split our goal into two parts: constructing functions on the k-line when $\frac{\kappa}{\rho}\log\kappa \leq \kappa'' \leq \kappa$, and constructing functions on the k''-curve when $\kappa \leq \frac{\kappa}{\rho}\log\kappa$. To construct functions on the k-line, we use the functions $\mathsf{AD}_{t,t',a}$ constructed for the proof of Theorem 1.4, since $k'(\mathsf{AD}_{t,t',a}) = k''(\mathsf{AD}_{t,t',a})$.

For constructing functions on the k''-curve, we need to construct functions f such that

$$\delta(f) = \Theta\left(\frac{r(f)}{k''(f)\log\left(k''(f)/r(f)\right)}\right). \tag{15}$$

We will use a similar technique as in our construction of functions on the k'-curve in Theorem 1.4. We start from the function $\mathsf{AD}_{t,t'}$ at the intersection point. Note that $\mathsf{AD}_{t,t'}$ satisfies Equation (15). We modify $\mathsf{AD}_{t,t'}$ such that the rank, weight and max-rank-entropy changes very little but the sparsity increases by a multiplicative parameter 2^p. We achieve this by replacing a variable (say x) in $\mathsf{AD}_{t,t'}$ with $x \cdot \mathsf{AND}(y_1, \ldots, y_p)$, where x and y_is are all variables in $\mathsf{AD}_{t,t'}$, but for any i, x and y_i do not appear in the same monomial (Claim B.4). The new function f still satisfies Equation (15). This places f on the k''-curve in a plot corresponding to the same rank as that of $\mathsf{AD}_{t,t'}$, but where the sparsity increases by a factor of 2^p. By suitably setting p, t and t', we obtain functions on the k''-curve for all plots. This proves the second bound in Theorem 1.6.

4 Conclusions

In this paper, for Boolean functions f, we study the relationship between weight and other Fourier-analytic measures namely rank, sparsity, max-supp-entropy and max-rank-entropy. For a threshold $t > 0$, Chang's lemma gives a lower bound on the weight of a Boolean function f in terms of $\dim\left(\left\{S \subseteq [n] : |\widehat{f}(S)| \geq \frac{1}{t}\right\}\right)$. We consider three natural thresholds t in Chang's lemma, namely $k(f)$, $k'(f)$ and $k''(f)$, yielding three lower bounds on weight in terms of these measures. We prove new lower bounds on weight in Theorems 1.3 and 1.5, and our bounds dominate all the above-mentioned bounds from Chang's lemma for a wide range of parameters.

When $\log k(f) = \Theta(r(f))$, the function $f = \mathsf{AND}$ already shows that all the above lower bounds are tight. To consider all other feasible relationships between $k(f)$ and $r(f)$, we divide our investigation of these lower bounds into two different parts. In the first part, we vary over all feasible settings of $r(f)$, $k(f)$ and $k'(f)$, and construct functions that witness tightness of our lower bounds in Theorem 1.3 for nearly all such feasible settings (Theorem 1.4). In the second part, we vary over all feasible settings of $r(f)$, $k(f)$ and $k''(f)$, and construct functions that witness near-tightness of our lower bounds in Theorem 1.5 for nearly all such feasible settings (Theorem 1.6). These functions are constructed by carefully composing the Addressing function with suitable inner functions. We show a composition lemma (Lemma B.1), which relates the properties of the composed function with those of the inner functions; this allows us to come up with functions that match our lower bounds.

We also construct functions for which our lower bounds are asymptotically stronger than the lower bounds obtained from Chang's lemma for all choices of threshold (see Claim 1.7). All functions that we construct in this work might be of independent interest.

Open Problems. Since our proof of Theorem 1.2 is a generalization of the proof of the upper bound $r(f) = O(\sqrt{k(f)}\log k(f))$ due to Sanyal [25], it sheds light on the presence of the $\log k$ factor in Sanyal's upper bound. This still leaves the following question open: do there exist Boolean functions f for which $r(f) = \omega(\sqrt{k(f)})$?

There are some ranges of parameters where we were not able to construct functions with upper bounds matching our lower bounds from Theorem 1.5. It will be interesting to see if our techniques can be extended to cover these ranges as well.

All thresholds t considered for Chang's lemma in this work satisfy $\dim(\{S \subseteq [n] : |\widehat{f}(S)| \geq \frac{1}{t}\}) = r(f)$. It is an interesting problem to obtain Chang's-lemma-type bounds for thresholds for which this dimension is strictly less than $r(f)$.

References

1 Srinivasan Arunachalam, Sourav Chakraborty, Troy Lee, Manaswi Paraashar, and Ronald de Wolf. Two new results about quantum exact learning. In *46th International Colloquium on Automata, Languages, and Programming, ICALP 2019*, volume 132 of *LIPIcs*, pages 16:1–16:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019.

2 Eli Ben-Sasson, Noga Ron-Zewi, Madhur Tulsiani, and Julia Wolf. Sampling-based proofs of almost-periodicity results and algorithmic applications. In *41st International Colloquium on Automata, Languages, and Programming ICALP 2014*, volume 8572 of *Lecture Notes in Computer Science*, pages 955–966. Springer, 2014.

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

4 Thomas F Bloom. A quantitative improvement for Roth's theorem on arithmetic progressions. *Journal of the London Mathematical Society*, 93(3):643–663, 2016.

5 Thomas F Bloom and Olof Sisask. Breaking the logarithmic barrier in Roth's theorem on arithmetic progressions. *arXiv preprint*, 2020. arXiv:2007.03528.

6 Sourav Chakraborty, Nikhil S. Mande, Rajat Mittal, Tulasimohan Molli, Manaswi Paraashar, and Swagato Sanyal. Tight chang's-lemma-type bounds for boolean functions. *CoRR*, abs/2012.02335, 2020. arXiv:2012.02335.

7 Siu On Chan, James R. Lee, Prasad Raghavendra, and David Steurer. Approximate constraint satisfaction requires large LP relaxations. *J. ACM*, 63(4):34:1–34:22, 2016.

8 Mei-Chu Chang. A polynomial bound in Freiman's theorem. *Duke mathematical journal*, 113(3):399–419, 2002.

9 Eshan Chattopadhyay, Pooya Hatami, Shachar Lovett, and Avishay Tal. Pseudorandom generators from the second Fourier level and applications to AC0 with parity gates. In *10th Innovations in Theoretical Computer Science Conference, ITCS*, pages 22:1–22:15, 2019. doi:10.4230/LIPIcs.ITCS.2019.22.

10 Ehud Friedgut, Jeff Kahn, Gil Kalai, and Nathan Keller. Chvátal's conjecture and correlation inequalities. *J. Comb. Theory, Ser. A*, 156:22–43, 2018.

11 Parikshit Gopalan, Ryan O'Donnell, Rocco A. Servedio, Amir Shpilka, and Karl Wimmer. Testing Fourier dimensionality and sparsity. *SIAM J. Comput.*, 40(4):1075–1100, 2011. doi:10.1137/100785429.

12 Ben Green. Arithmetic progressions in sumsets. *Geometric and Functional Analysis GAFA*, 12(3):584–597, 2002.

13 Ben Green. Some constructions in the inverse spectral theory of cyclic groups. *Combinatorics, Probability and Computing*, 12(2):127–138, 2003.

14 Ben Green. Spectral structure of sets of integers. In *Fourier analysis and convexity*, pages 83–96. Springer, 2004.

15 Ben Green and Imre Z. Ruzsa. Freiman's theorem in an arbitrary abelian group. *Journal of the London Mathematical Society*, 75(1):163–175, 2007.

16 Ben Green and Tom Sanders. Boolean functions with small spectral norm. *Geometric and Functional Analysis GAFA*, 18:144–162, 2008.

17 Tom Gur and Omer Tamuz. Testing Booleanity and the uncertainty principle. *Chic. J. Theor. Comput. Sci.*, 2013, 2013.

18 Pooya Hatami, Raghav Kulkarni, and Denis Pankratov. Variations on the sensitivity conjecture. *Theory of Computing*, 4:1–27, 2011. doi:10.4086/toc.gs.2011.004.

19 Kaave Hosseini, Shachar Lovett, and Grigory Yaroslavtsev. Optimality of linear sketching under modular updates. In *34th Computational Complexity Conference, CCC 2019*, volume 137 of *LIPIcs*, pages 13:1–13:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019.

20 Russell Impagliazzo, Cristopher Moore, and Alexander Russell. An entropic proof of Chang's inequality. *SIAM J. Discret. Math.*, 28(1):173–176, 2014.

21 Shachar Lovett. Communication is bounded by root of rank. *Journal of the ACM (JACM)*, 63(1):1–9, 2016.

22 Ashley Montanaro and Tobias Osborne. On the communication complexity of XOR functions. *CoRR*, abs/0909.3392, 2009. arXiv:0909.3392.

23 Tom Sanders. Additive structures in sumsets. *Mathematical Proceedings of the Cambridge Philosophical Society*, 144(2):289–316, 2008.

24 Tom Sanders. On Roth's theorem on progressions. *Annals of Mathematics*, 174:619–636, 2011.

25 Swagato Sanyal. Fourier sparsity and dimension. *Theory of Computing*, 15(11):1–13, 2019. doi:10.4086/toc.2019.v015a011.

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

A Preliminaries

All logarithms in this paper are taken to be base 2. We use the notation $[n]$ to denote the set $\{1, 2, \ldots, n\}$. When necessary, we assume t is a power of 2. We use the notation 1^n (respectively, $(-1)^n$) to denote the n-bit string $(1, 1, \ldots, 1)$ (respectively, $(-1, -1, \ldots, -1)$).

For a function $f : \{-1, 1\}^n \to \{-1, 1\}$, its $\mathbb{F}_2$-degree, denoted by $\deg_{\mathbb{F}_2}(f)$, is the degree of its unique $\mathbb{F}_2$-polynomial representation. Throughout this paper, we often identify subsets of $[n]$ with their corresponding characteristic vectors in $\mathbb{F}_2^n$. Thus when we refer to linear algebraic measures of a collection of subsets of $[n]$, we mean the measure on the corresponding subset of $\mathbb{F}_2^n$ (where $\mathbb{F}_2^n$ is viewed as an $\mathbb{F}_2$-vector space).

Throughout this paper, we assume that f is not a constant function or a parity or a negative parity, unless mentioned otherwise.

A.1 Fourier analysis of Boolean functions

Consider the vector space of functions from $\{-1, 1\}^n$ to $\mathbb{R}$ equipped with the following inner product.

$$\langle f, g \rangle := \frac{1}{2^n} \sum_{x \in \{-1,1\}^n} f(x)g(x).$$

For a set $S \subseteq [n]$, define a *parity* function (which we also refer to as *characters*) $\chi_S : \{-1, 1\}^n \to \{-1, 1\}$ by $\chi_S(x) = \prod_{i \in S} x_i$. The set of parity functions $\{\chi_S : S \subseteq [n]\}$ forms an orthonormal basis for this vector space. Hence, every function $f : \{-1, 1\}^n \to \mathbb{R}$ has a unique representation as

$$f = \sum_{S \subseteq [n]} \widehat{f}(S)\chi_S,$$

where $\widehat{f}(S) = \langle f, \chi_S \rangle$ for all $S \subseteq [n]$. The coefficients $\left\{\widehat{f}(S) : S \subseteq [n]\right\}$ are called the *Fourier coefficients* of f. Define the Fourier ℓ_1-norm of a function $f : \{-1, 1\}^n \to \mathbb{R}$ by $\|\widehat{f}\|_1 := \sum_{S \subseteq [n]} |\widehat{f}(S)|$. The Fourier support of f, denoted by $\mathrm{supp}(f)$, is defined as

$$\mathrm{supp}(f) = \left\{S \subseteq [n] : \widehat{f}(S) \neq 0\right\}.$$

▶ Remark A.1. In the literature, Fourier support is generally denoted by $\mathrm{supp}(\widehat{f})$. For ease of notation we drop the hat symbol above f. A similar convention has been adopted in the remaining parts of the paper.

Let $f : \{-1, 1\}^n \to \mathbb{R}$ be any function. The Fourier sparsity of f, denoted by $k(f)$, is defined as $k(f) = |\mathrm{supp}(f)|$. For simplicity we assume that $k(f) \geq 2$ for all Boolean functions f considered in this paper (unless explicitly mentioned otherwise). We often simply refer to the Fourier sparsity as *sparsity*. For ease of notation, we sometimes abuse notation and say that the elements of the Fourier support of f are the characters $\left\{\chi_S : S \subseteq [n], \widehat{f}(S) \neq 0\right\}$, rather than the corresponding sets.

We require the following lemma (see, for example, [17]).

▶ **Lemma A.2** (Uncertainty Principle). *Let $f : \{-1,1\}^n \to \mathbb{R}$ be a polynomial and let U_n denote the uniform distribution on $\{-1,1\}^n$. Then,*

$$\Pr_{x \sim U_n}[f(x) \neq 0] \geq \frac{1}{k(f)}.$$

We also require the following lemma relating the $\mathbb{F}_2$-degree of a Boolean function and its Fourier sparsity (see, for example, [3]).

▶ **Lemma A.3.** *Let $f : \{-1,1\}^n \to \{-1,1\}$ be any function with $k(f) > 1$. Then,*

$$\deg_{\mathbb{F}_2}(f) \leq \log k(f).$$

The next claim shows that $\deg_{\mathbb{F}_2}(f)$ does not change under a change of basis over the Fourier domain.

▷ Claim A.4. Let $f : \{-1,1\}^n \to \{-1,1\}$ be any function and let $B \in \mathbb{F}_2^{n \times n}$ be an invertible matrix. Define the function $f_B : \{-1,1\}^n \to \mathbb{R}$ as

$$\widehat{f_B}(\alpha) = \widehat{f}(B\alpha) \quad \text{for all } \alpha \in \mathbb{F}_2^n,$$

Then f_B is Boolean valued and $\deg_{\mathbb{F}_2}(f_B) = \deg_{\mathbb{F}_2}(f)$.

The following corollary follows from [9, Theorem 13] and Lemma A.3.

▶ **Corollary A.5.** *Let $f : \{-1,1\}^n \to \{-1,1\}$ be any function, and let $\mathcal{S} \subseteq \mathrm{supp}(f)$ be a basis of $\mathrm{span}(\mathrm{supp}(f))$. Then,*

$$\sum_{S \in \mathcal{S}} |\widehat{f}(S)| \leq 4 \log k(f).$$

We now define notions of restriction of a function $f : \{-1,1\}^n \to \{-1,1\}$ to a subset $A \subseteq \{-1,1\}^n$.

▶ **Definition A.6** (Restriction). *Let $f : \{-1,1\}^n \to \{-1,1\}$ and $A \subseteq \{-1,1\}^n$. The restriction of f to A is the function $f|_A : A \to \{-1,1\}$ defined as $f|_A(x) = f(x)$ for all $x \in A$.*

▶ **Definition A.7** (Affine Restriction). *Let $f : \{-1,1\}^n \to \{-1,1\}$, let Γ be a set of parities and $b \in \{-1,1\}^\Gamma$ be an assignment to these parities. Define the function $f|_{(\Gamma,b)}$ to be the restriction of f to the affine subspace obtained by fixing parities in Γ according to b. That is,*

$$f|_{(\Gamma,b)} := f|_{\{x \in \{-1,1\}^n : \chi_\gamma(x) = b_\gamma \text{ for all } \gamma \in \Gamma\}}.$$

A.2 Fourier expansions and properties of some standard functions

For any integer $n > 0$, define the function $\mathsf{AND}_n : \{-1,1\}^n \to \{-1,1\}$ by $\mathsf{AND}_n(x) = -1$ if $x = (-1)^n$, and 1 otherwise. We drop the subscript n when it is clear from the context.

▶ **Definition A.8** (Bent functions). *A function $f : \{-1,1\}^n \to \{-1,1\}$ is said to be a bent function if $|\widehat{f}(S)| = |\widehat{f}(T)|$ for all $S, T \subseteq [n]$.*

▶ **Definition A.9** (Indicator function). *For any integer $n \geq 1$ and $b \in \{-1,1\}^n$, define the function $\mathbb{I}_b : \{-1,1\}^n \to \{0,1\}$ by*

$$\mathbb{I}_b(x) = \begin{cases} 1 & x = b, \\ 0 & \textit{otherwise}. \end{cases}$$

▶ **Definition A.10** (Addressing function). *For any integer $t \geq 2$, define the Addressing function* $\mathsf{AD}_t : \{-1,1\}^{\log t} \times \{-1,1\}^t \to \{-1,1\}$ *by*

$$\mathsf{AD}_t(x, y) = y_{\mathrm{bin}(x)},$$

where $x \in \{-1,1\}^{\log t}$ and $y \in \{-1,1\}^t$, and $\mathrm{bin}(x)$ *denotes the integer in $[t]$ whose binary representation is given by x (where -1's are viewed as 1 in the string x, and 1's are viewed as 0). We refer to the x-variables as* addressing variables*, and the y-variables as* target variables*.*

We next define a way of modifying the Addressing function that is of use to us. In this modification, we replace target variables by functions, each acting on disjoint variables.

▶ **Definition A.11** (Composed addressing functions). *Let $t \geq 2$, $\ell_1, \ldots, \ell_t \geq 1$ be any integers. Let $g_i : \{-1,1\}^{\ell_i} \to \{-1,1\}$ be any functions for $i \in [t]$. Define the function $\mathsf{AD}_t \circ_{\mathrm{target}} (g_1, \ldots, g_t) : \{-1,1\}^{\log t} \times \{-1,1\}^{\ell_1 + \cdots + \ell_t} \to \{-1,1\}$ by*

$$\mathsf{AD}_t \circ_{\mathrm{target}} (g_1, \ldots, g_t)(x, y_1, \ldots, y_t) = \mathsf{AD}_t(x, g_1(y_1), \ldots, g_t(y_t)),$$

where $x \in \{-1,1\}^{\log t}$ and $y_i \in \{-1,1\}^{\ell_i}$ for all $i \in [t]$.

For any function $g : \{-1,1\}^s \to \{-1,1\}$, we use the notation $\mathsf{AD}_t \circ_{\mathrm{target}} g$ to denote the function $\mathsf{AD}_t \circ_{\mathrm{target}} (g, g, \ldots, g) : \{-1,1\}^{\log t} \times \{-1,1\}^{ts} \to \{-1,1\}$.

A.3 Fourier-analytic measures of Boolean functions

We now introduce a few Fourier-analytic measures on Boolean functions that we use throughout the rest of the paper, and state some important relationships between them. Recall that we use the notation $\dim(S)$ to denote the dimension of the span of the set S.

▶ **Definition A.12** (Fourier rank). *Let $f : \{-1,1\}^n \to \{-1,1\}$ be any function. Define the Fourier rank of f, denoted $r(f)$, by*

$$r(f) = \dim(\mathrm{supp}(f)).$$

We often refer to Fourier rank as simply *rank*. Sanyal [25] showed the following upper bound on the rank of Boolean functions in terms of their sparsity.

▶ **Theorem A.13** ([25, Theorem 1.2]). *Let $f : \{-1,1\}^n \to \{-1,1\}$ be any function. Then*

$$r(f) = O(\sqrt{k(f)} \log k(f)).$$

We require the following observation which gives a simple upper bound on the rank of a Boolean function.

▶ **Observation A.14.** *Let $f : \{-1,1\}^n \to \{-1,1\}$ be any function and Γ be a set of parities. If for all $b \in \{-1,1\}^\Gamma$ the restricted function $f|_{(\Gamma,b)}$ is constant then $r(f) \leq |\Gamma|$.*

Recall that for any function $f : \{-1,1\}^n \to \{-1,1\}$ and any real $t > 0$, we define $\mathcal{S}_t := \{S \subseteq [n] : |\widehat{f}(S)| \geq 1/t\}$ (we suppress the dependence of $\mathcal{S}_t$ on f as the underlying function will be clear from context).

▶ **Definition A.15.** *Let $f : \{-1,1\}^n \to \{-1,1\}$ be any function. Define the Fourier max-supp-entropy of f, denoted $k'(f)$, by*

$$k'(f) := \operatorname*{argmin}_t \{\mathcal{S}_t = \operatorname{supp}(f)\}.$$

Equivalently,

$$k'(f) := \max_{S \in \operatorname{supp}(f)} \left\{ \frac{1}{|\widehat{f}(S)|} \right\}.$$

Define the Fourier max-rank-entropy of f, denoted $k''(f)$, by

$$k''(f) := \operatorname*{argmin}_t \{\dim(\mathcal{S}_t) = r(f)\}.$$

We often refer to the Fourier max-supp-entropy and Fourier max-rank-entropy as simply *max-supp-entropy* and *max-rank-entropy*, respectively.

▶ **Lemma A.16** (Relationships between parameters)**.** *Let $f : \{-1,1\}^n \to \{-1,1\}$ be any function. Then the following inequalities hold.*

1. $\log k(f) \leq r(f) = O(\sqrt{k(f)} \log k(f))$.
2. $\sqrt{k(f)} \leq k'(f) \leq k(f)/2$.
3. $\max\left\{\sqrt{r(f)}, r(f)/(4 \log k(f))\right\} \leq k''(f) \leq k'(f)$.

▷ Claim A.17. Let $f : \{-1,1\}^n \to \{-1,1\}$ a function with $k(f) \geq 2$. Then

$$\delta(f) = \Omega\left(\frac{k(f)}{k'(f)^2}\right).$$

▷ Claim A.18. Let $f : \{-1,1\}^n \to \{-1,1\}$ be any function. Then

$$\|\widehat{f}\|_1 \leq 3\sqrt{k(f)\delta(f)}.$$

We require the following observation about the rank, sparsity, max-supp-entropy, max-rank-entropy and weight of the addressing function, AD_t, which follows immediately from definitions and first principles. We omit its proof.

▶ **Observation A.19.** *Let $t \geq 2$ be any positive integer. Then the rank, sparsity, max-supp-entropy, max-rank-entropy and weight of AD_t are $(t + \log t)$, t^2, t, t and $1/2$, respectively.*

B Upper bound proofs

The following lemma is a useful tool for our upper bounds. We refer the reader to [6, Section 6.2.1] in the full version of our paper for a proof.

▶ **Lemma B.1** (Composition lemma)**.** *Let $t \geq 2, m \geq 1$ be any positive integers, and let $g : \{-1,1\}^m \to \{-1,1\}$ be a non-constant function such that there exists a non-empty set $S \subseteq [m]$ with $0 \neq |\widehat{g}(S)| \leq |\widehat{g}(\emptyset)|$. Let $f : \{-1,1\}^{\log t + mt} \to \{-1,1\}$ be defined as*

$$f = \mathsf{AD}_t \circ_{\text{target}} g.$$

Then

$$r(f) = t \cdot r(g) + \log t, \tag{16}$$
$$k(f) = 1 + t^2(k(g) - 1), \tag{17}$$
$$k'(f) = t \cdot k'(g), \tag{18}$$
$$k''(f) = t \cdot k''(g), \tag{19}$$
$$\delta(f) = \delta(g). \tag{20}$$

B.1 Setting parameters in our constructed functions

In this section we state the main claims that go into proving Theorems 1.4 and 1.6. Recall that these theorems require us to exhibit functions which achieve certain bounds. Claims B.2 and B.3 correspond to the bounds in Theorem 1.4. Claims B.4 and B.5 correspond to the bounds in Theorem 1.6. All functions referred to below are informally defined in Section 3. See the full version of our paper [6, Section 6.1] for formal definitions and [6, Sections 6.2, 6.3] for proofs of these claims.

▷ Claim B.2. For all $\rho, \kappa, \kappa' \in \mathbb{N}$ such that κ is sufficiently large, for all $\epsilon > 0$ such that $\log \kappa \le \rho \le \kappa^{\frac{1}{2}-\epsilon}$ and $\frac{\kappa \log \kappa}{\rho} \le \kappa' \le \kappa$, for $t = \frac{2\rho}{\log \kappa}$, $t' = \frac{\kappa \log^2 \kappa}{\rho^2}$ and $a = \frac{2\kappa' \log \kappa}{\rho}$,

- $\Omega(\epsilon\rho) = r(\mathsf{AD}_{t,t',a}) = O(\rho)$.
- $k(\mathsf{AD}_{t,t',a}) = \Theta(\kappa)$.
- $k'(\mathsf{AD}_{t,t',a}) = \Theta(\kappa')$.
- $\delta(\mathsf{AD}_{t,t',a}) = \Theta\left(\frac{1}{\kappa}\left(\frac{\rho}{\log \kappa}\right)^2\right)$.

▷ Claim B.3. For all $\rho, \kappa, \kappa' \in \mathbb{N}$ such that κ is sufficiently large, for all constants $\epsilon > 0$, such that $\kappa^{1/2} \le \kappa' \le (\kappa \log \kappa)/\rho$ and $\log \kappa \le \rho \le \kappa^{\frac{1}{2}-\epsilon}$ for $t = \frac{2\rho}{\log \kappa}$, $t' = \frac{4\kappa'^2}{\kappa}$ and $\ell = 2\left(\frac{\kappa \log \kappa}{\kappa' \rho}\right)^2$,

- $\Omega(\epsilon\rho) = r(\mathsf{AAB}_{t,t',\ell}) = O(\rho)$.
- $k(\mathsf{AAB}_{t,t',\ell}) = \Theta(\kappa)$.
- $k'(\mathsf{AAB}_{t,t',\ell}) = \Theta(\kappa')$.
- $\delta(f) = O\left(\frac{\kappa}{\kappa'^2}\right)$.

▷ Claim B.4. For all $\rho, \kappa, \kappa'' \in \mathbb{N}$ such that κ is sufficiently large, for all constants $\epsilon > 0$ such that $\log \kappa \le \rho \le \kappa^{1/2-\epsilon}$, $e\rho \le \kappa'' \le \frac{\kappa \log \kappa}{\rho}$, for $t = \frac{2\rho}{\log(\kappa''/\rho)}$, $t' = \frac{\kappa''}{\rho} \log(\kappa''/\rho)$, $p = \log\left(\frac{4\kappa}{\kappa''}\right)$,

- $r(\mathsf{mAD}_{t,t',p}) = \Theta(\rho)$.
- $\Omega(\kappa) = k(\mathsf{mAD}_{t,t',p}) = O(\kappa/\epsilon)$.
- $k''(\mathsf{mAD}_{t,t',p}) = \Theta(\kappa'')$.
- $\delta(\mathsf{mAD}_{t,t',p}) = \frac{\rho}{\kappa'' \log(\kappa''/\rho)}$.

▷ Claim B.5. For all $\rho, \kappa, \kappa' \in \mathbb{N}$ such that κ is sufficiently large, for all $\epsilon > 0$ such that $\log \kappa \le \rho \le \kappa^{\frac{1}{2}-\epsilon}$ and $\frac{\kappa \log \kappa}{\rho} \le \kappa'' \le \kappa$, there exists a constant $c \ge 1$ such that the following holds for $t = \frac{2\rho}{\log \kappa}$, $t' = \frac{c\kappa \log^2 \kappa}{\rho^2}$ and $a = \frac{2c\kappa'' \log \kappa}{\rho}$.

- $\Omega(\epsilon\rho) = r(\mathsf{AD}_{t,t',a}) = O(\rho)$.
- $k(\mathsf{AD}_{t,t',a}) = \Theta(\kappa)$.
- $k''(\mathsf{AD}_{t,t',a}) = \Theta(\kappa'')$.
- $\delta(\mathsf{AD}_{t,t',a}) = \Theta\left(\frac{1}{\kappa}\left(\frac{\rho}{\log \kappa}\right)^2\right)$.