



Searching for Regularity in Bounded Functions

Siddharth Iyer  

University of Washington CSE, Seattle, WA, USA

Michael Whitmeyer  

University of Washington CSE, Seattle, WA, USA

Abstract

Given a function f on \mathbb{F}_2^n , we study the following problem. What is the largest affine subspace \mathcal{U} such that when restricted to \mathcal{U} , all the non-trivial Fourier coefficients of f are very small?

For the natural class of bounded Fourier degree d functions $f : \mathbb{F}_2^n \rightarrow [-1, 1]$, we show that there exists an affine subspace of dimension at least $\tilde{\Omega}(n^{1/d}k^{-2})$, wherein all of f 's nontrivial Fourier coefficients become smaller than 2^{-k} . To complement this result, we show the existence of degree d functions with coefficients larger than $2^{-d \log n}$ when restricted to any affine subspace of dimension larger than $\Omega(dn^{1/(d-1)})$. In addition, we give explicit examples of functions with analogous but weaker properties.

Along the way, we provide multiple characterizations of the Fourier coefficients of functions restricted to subspaces of \mathbb{F}_2^n that may be useful in other contexts. Finally, we highlight applications and connections of our results to parity kill number and affine dispersers.

2012 ACM Subject Classification Mathematics of computing

Keywords and phrases regularity, bounded function, Boolean function, Fourier analysis

Digital Object Identifier 10.4230/LIPIcs.ICALP.2023.83

Category Track A: Algorithms, Complexity and Games

Related Version *Full Version*: <https://arxiv.org/abs/2207.13312>

Funding *Siddharth Iyer*: Research supported by NSF grant CCF-2131899.

Michael Whitmeyer: Research supported by NSF grant CCF-2006359.

Acknowledgements We thank Anup Rao for posing the question that launched this project and for his invaluable advice and feedback. We are also grateful Paul Beame for his extremely helpful advice, discussions, and feedback. Finally, we thank Sandy Kaplan for detailed feedback on this writeup.

1 Introduction

The search for structure within large objects is an old one that lies at the heart of Ramsey theory. For example, a famous corollary of Ramsey's theorem is that any graph on n vertices must contain a clique or an independent set of size $\Omega(\log n)$. Another example is Roth's¹ theorem [19] on 3-term arithmetic progressions, which essentially says that *every* subset of $\{1, \dots, n\}$ of density $\delta > \Omega(1/\log \log n)$ must contain a 3-term arithmetic progression.²

Szemerédi's Regularity Lemma is also a well known example of this phenomenon. Roughly speaking, it states that any graph G can be partitioned into $k := M(\delta)$ parts V_1, \dots, V_k , wherein most pairs of parts (V_i, V_j) are δ -regular. In this setting, the δ -regularity of (V_i, V_j) roughly corresponds to saying that the bipartite graph induced across V_i and V_j appears as though its edges were sampled randomly. This powerful statement has found applications

¹ The related Hales-Jewett theorem [10] is also a classic result in Ramsey theory.

² See also the recent quantitative improvement due to Kelley and Meka [13] which gives the same result for all subsets of density at least $\Omega(2^{-\log^{1/11}(n)})$.



© Siddharth Iyer and Michael Whitmeyer;

licensed under Creative Commons License CC-BY 4.0

50th International Colloquium on Automata, Languages, and Programming (ICALP 2023).

Editors: Kousha Etessami, Uriel Feige, and Gabriele Puppis; Article No. 83; pp. 83:1–83:20

Leibniz International Proceedings in Informatics



LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany



in both pure mathematics (e.g., Szemerédi’s [23] generalization of Roth’s result to k -term arithmetic progressions) and theoretical computer science (to test triangle-freeness in dense graphs [20, 1, 22]).

Similar to the definition of regular partitions in Szemerédi’s Regularity Lemma, one can also define a notion of regularity for functions. In particular, for functions $f : \mathbb{F}_2^n \rightarrow \mathbb{R}$, we follow Green [9] and O’Donnell [17] and define a function to be δ -regular if all its nontrivial Fourier coefficients are at most δ in magnitude.³ This definition can be viewed as a pseudorandomness condition; in particular, a randomly chosen Boolean function $f : \mathbb{F}_2^n \rightarrow \{\pm 1\}$ is δ -regular with very high probability, even for $\delta = 2^{-\Omega(n)}$.⁴

The prior works surrounding graph regularity [23, 7, 22] and function regularity [9, 12] have been concerned with obtaining δ -regular *partitions*, which, roughly speaking, are partitions of the object at hand into (mostly) pseudorandom parts. Often, these results have quite poor dependencies on the parameter δ so as not to be practical for any reasonably small value of δ (see Proposition 3 and Proposition 4 for detailed statements). Motivated by this, and by applications in theoretical computer science, we relax our requirement and look to find just *one* δ -regular part. Namely, we seek to understand the following quantity:

$$r(f, \delta) := \min\{\text{codim}(\mathcal{U}) : \mathcal{U} \text{ is an affine subspace such that } f_{\mathcal{U}} \text{ is } \delta\text{-regular}\},$$

where here and throughout this work $f_{\mathcal{U}} : \mathcal{U} \rightarrow \mathbb{R}$ denotes the restriction of f to inputs coming from \mathcal{U} .

Before stating our main results as well as prior work, we make a few remarks about the quantity $r(f, \delta)$. In the special case when $\delta = 0$, the quantity $r(f, 0)$ has been previously studied in the literature, under the name of *parity kill number* [18]. This is the smallest number of parities that need to be fixed in order to make f constant. The value $r(f, 0)$ is also a measure associated with affine dispersers, objects that have received significant attention in the study of pseudorandomness, see e.g. [21, 14, 5, 6, 3]. An affine disperser of dimension k is a coloring of \mathbb{F}_2^n such that no affine subspace of dimension k is monochromatic. If we view an affine disperser as a function $f : \mathbb{F}_2^n \rightarrow \{0, 1, \dots, C\}$, then its dimension is just $n - r(f, 0) + 1$.

Now, we briefly discuss the bounds on $r(f, \delta)$ most relevant to our work. For a general function $f : \mathbb{F}_2^n \rightarrow [-1, 1]$, it is known that $r(f, \delta) \leq 1/\delta$; this follows from a well-known density-increment argument, see [15] (for a short proof of this, see Proposition 5). One might ask if $r(f, \delta)$ is small when we assume f is structured, and a natural example of such functions is the class of functions with low Fourier degree. For general degree d functions $f : \mathbb{F}_2^n \rightarrow [-1, 1]$, the best bound on $r(f, \delta)$ until this work was just the above mentioned bound of $1/\delta$. However, for the class of degree d Boolean functions, we know that $r(f, \delta) \leq r(f, 0) = O(d^3)$; this follows from the polynomial relationship between Fourier degree and decision tree depth, see [16], and [2, 4] for surveys. We emphasize that this result relies crucially on Booleanity (and is independent of δ), and one can ask if the more general class of degree d functions bounded in the interval $[-1, 1]$ also have small $r(f, \delta)$ values. Our main result answers exactly this question, and provides an upper bound for $r(f, \delta)$ in this setting.

► **Theorem 1.** *For any $\delta \in (0, 1)$ and any degree d function $f : \mathbb{F}_2^n \rightarrow [-1, 1]$, we have $r(f, \delta) \leq n - \Omega(n^{1/d}(\log(n/\delta))^{-2})$.*

³ For a formal definition, see Definition 11, and for more background on Fourier analysis, see Section 2.

⁴ See for example [17], Exercise 1.7 and Proposition 6.1.

Note that the general bound $r(f, \delta) \leq 1/\delta$ that we mentioned earlier, is only meaningful when $\delta > 1/n$, however, our theorem allows for δ to be much smaller. The regime of small δ is particularly interesting from the perspective of pseudorandomness. Indeed, in a qualitative sense, we see that by decreasing δ , we are asking for affine subspaces where the restricted function looks increasingly like a random function. Using Theorem 1 together with our connection between $r(f, 0)$ and the dimension of affine disperse, we obtain the following corollary which says that low degree polynomials cannot serve as good affine dispersers.

► **Corollary 2.** *If $f : \mathbb{F}_2^n \rightarrow \{0, \dots, C\}$ has Fourier degree d , then f cannot be an affine disperser of dimension k for any $k \geq \Omega(n^{1/d}(d + \log(nC))^{-2})$.*

Lower Bounds on $r(f, \delta)$. To complement Theorem 1, we present in Table 1 several examples of functions (bounded as well as Boolean) for which $r(f, \delta)$ is large. For each row in the table, we exhibit a class of functions (whose degree and range is as specified), such that for any $\delta' \leq \delta$, no affine subspace of dimension larger than $n - r(f, \delta)$ is δ' -regular.

■ **Table 1** Table of functions with large $r(f, \delta)$ values.

$\delta \leq$	$r(f, \delta) \geq$	$\deg(f)$	$\text{range}(f)$	Ref.
$1/n$	$n/2 - 1$	1	$[-1, 1]$	Lemma 26
$\binom{n}{d}^{-1}$	$n - 2dn^{1/(d-1)}$	d	$[-1, 1]$	Lemma 27
$\Theta(n^{-1/2})$	$\Theta(\sqrt{n})$	n	$\{\pm 1\}$	Lemma 32
$\frac{1}{2} \cdot n^{-d}$ (for $d \leq \frac{\log n}{\log \log n + 1}$)	$n - 2dn^{1/(d-1)}$	$\Omega(n)$	$\{\pm 1\}$	Corollary 30
$1/2^{2^k+1}$ (for integer k)	$\Omega\left(\left(\log \frac{1}{\delta}\right)^{\log_2(3)}\right)$	2^k	$\{\pm 1\}$	Lemma 31

Observe that Lemma 27 provides a somewhat of a converse to Theorem 1. However there is a noticeable gap between the two results, and we conjecture that Lemma 27 is closer to being tight, and that Theorem 1 could be improved. We also note that Lemma 27 and Corollary 30 are not explicit – it would be interesting to find more explicit examples.

1.1 Related Work

To the best of our knowledge, $r(f, \delta)$ has not been explicitly studied before. However, it is closely related to well-studied notions of function regularity as well as the concepts of parity kill number and affine dispersers. In this section, we give a detailed description of both these connections.

Parity Kill Number and Affine Dispersers. As we have already mentioned, $r(f, 0)$ has been studied under the name of *parity kill number*, denoted $C_{\min}^{\oplus}[f]$ (see [18]). Parity kill number can be considered as a further generalization of the *minimum certificate complexity* of f , denoted $C_{\min}[f]$, which is the minimum number of bits one must fix in order to make f constant. In particular, for any $\delta \geq 0$, we have $r(f, \delta) \leq r(f, 0) \leq C_{\min}[f]$. The minimum certificate complexity is one of several natural complexity measures that have been well studied for Boolean functions $f : \mathbb{F}_2^n \rightarrow \{\pm 1\}$ (see [4, 2] for surveys).

As we have already alluded to, the quantity $r(f, 0)$ is also closely related to efficacy of $f : \mathbb{F}_2^n \rightarrow \{0, \dots, C\}$ as an affine disperser. In the case of $C = 1$, Cohen and Tal [6] rule out \mathbb{F}_2 -polynomials of degree d as affine dispersers by showing that any such function satisfies $r(f, 0) \leq n - \Omega(d \cdot n^{1/(d-1)})$. This result resembles our Corollary 2; however, the two results are incomparable for two reasons. First, degree d functions over \mathbb{F}_2 can have very large

Fourier degree; moreover, the corresponding result of [6] applies to functions whose range is \mathbb{F}_2 , while ours applies to functions that take values in the set $\{0, \dots, C\}$, which can have a much larger size. Furthermore, for $f : \mathbb{F}_2^n \rightarrow \{0, \dots, C\}$, a standard argument (analogous to the one in [16]) shows that $r(f, 0) \leq O(Cd^3)$, where d here is the Fourier degree. However, this does not address the case where $C = \Omega(n)$, which is when Corollary 2 becomes useful.

Pseudorandom Partitions. As we have mentioned, much prior work on function regularity has been focused on finding pseudorandom partitions of \mathbb{F}_2^n . To the best of our knowledge, the earliest result in this direction is due to Green [9]; below, the notation $\text{twr}(x)$ refers to an exponential tower of 2's $2^{2^{\dots}}$ of height x .

► **Proposition 3** (Theorem 2.1 in [9]). *For any $f : \mathbb{F}_2^n \rightarrow \{0, 1\}$ and $\delta > 0$, there exists a subspace \mathcal{V} of co-dimension $M(\delta) \leq \text{twr}(\lceil 1/\delta^3 \rceil)$ such that for all but a δ -fraction of the affine subspaces $\mathcal{U} = \alpha + \mathcal{V}$, $f|_{\mathcal{U}}$ is δ -regular.*

In the same paper, Green showed that $M(\delta) \geq \text{twr}(\Omega(\log(1/\delta)))$ was necessary. Subsequently, Hosseini et al. [12] exhibited a better counterexample showing co-dimension $M(\delta) \geq \text{twr}(\lceil 1/16\delta \rceil)$ is required.

In the above upper and lower bound of [9, 12], the partition of \mathbb{F}_2^n is of a specific form – namely, it is every affine shift of a given subspace. Given this observation, one can ask if there is a partition of \mathbb{F}_2^n into affine subspaces of smaller co-dimension so that in most parts f is δ -regular. As the next proposition, due to Girish et al. [8] shows, this is indeed the case.

► **Proposition 4** (Proposition A.1 in [8]). *For any $f : \mathbb{F}_2^n \rightarrow [0, 1]$ and $\delta > 0$, there exists a partition Π of \mathbb{F}_2^n , where every $\pi \in \Pi$ is an affine subspace of co-dimension at most $\frac{1}{\delta^3}$ such that for all but a δ -fraction of the parts, $f|_{\pi}$ is δ -regular.*

The proof of Proposition 4 is based on a simple algorithm that greedily fixes the partitions corresponding to the largest Fourier coefficients; it is included in Appendix A.1 for completeness.

Although, both these results partition \mathbb{F}_2^n into several affine subspaces where f is δ -regular, they are only meaningful when δ is relatively large. Indeed, Proposition 3 is trivial when $\delta < (\log^*(n))^{-1/3}$, and Proposition 4 when $\delta < n^{-1/3}$. As we mentioned earlier, if we relax our requirement to finding just one affine subspace, there is a simple upper bound on $r(f, \delta)$ based on a density-increment argument, which goes back to the works of Roth [19] and Meshulam [15].

► **Proposition 5** (Folklore). *For any $f : \mathbb{F}_2^n \rightarrow [-1, 1]$, we have $r(f, \delta) \leq \frac{1}{\delta}$.*

We provide a proof of Proposition 5 in Appendix A.1 for completeness.

1.2 Techniques

Upper bound on $r(f, \delta)$. We give a brief proof sketch of Theorem 1. The proof proceeds by induction over the Fourier degree. The base case corresponds to degree one functions. Our intuition is derived from the following fact. If we have any k real numbers a_1, \dots, a_k such that the sum of any subset of them has magnitude at most one, then by the pigeonhole principle, there is a non-empty subset $S \subseteq [k]$, and a signing of the numbers in S so that the signed sum has magnitude at most $2^{-\Omega(k)}$. In the degree one case, we partition $\{1, \dots, n\}$ into consecutive disjoint intervals of size $k = O(\log 1/\delta)$. We apply the above intuition to the k Fourier coefficients in each interval, to obtain signed sums that have small magnitude.

Then, by appropriately choosing an affine subspace, \mathcal{U} of dimension $\Omega(n/\log(1/\delta))$, we show that these signed sums are exactly the Fourier coefficients of the function restricted to \mathcal{U} (see Proposition 13 for a more general statement). We give a more detailed description of how this works in Section 3.

At a high level, we reduce the problem for degree d functions to degree $d-1$ by restricting to an affine subspace of dimension $\tilde{\Omega}(n^{1/d})$, where the function is degree d and all Fourier coefficients at the d -th level are extremely small $\ll \delta/n^d$. For a detailed statement, see Lemma 19. When we use the inductive hypothesis for $d-1$, the last constraint ensures that the degree d coefficients cannot increase the new coefficients by more than $O(\delta)$, even if they combine in the most constructive way possible.

Lemma 19 is also obtained by repeatedly applying the pigeonhole principle. However, the key issue now is that several Fourier coefficients could be affected when we apply a restriction, unlike the degree one case. To avoid this, we apply restrictions iteratively so that each one preserves the small Fourier coefficients from past iterations while still ensuring that *several* new Fourier coefficients are also small. The cost of this procedure is that, in each step, we must apply the pigeonhole principle over larger and larger subsets of coordinates.

Lower Bounds. Here, we give a very high level overview of our lower bounds on $r(f, \delta)$. The basic idea is to consider functions f with the property that their Fourier spectrum is concentrated on a small number of Fourier coefficients. It turns out (see Proposition 13) that when we restrict to an affine subspace, say \mathcal{U} , the Fourier coefficients of $f_{\mathcal{U}}$ are simply signed sums of the Fourier coefficients of f . By our choice of f , if the restricted function was δ -regular, then the large coefficients of f involved in the signed sums somehow cancelled each other out. We show that by choosing the vectors corresponding to the large Fourier coefficients in f appropriately, such a cancellation would imply that the co-dimension of \mathcal{U} must be large. For more detailed sketches of the entries in Table 1, see Appendix A.2.

2 Preliminaries

Notation. $\mathbb{1}\{\cdot\}$ denotes an indicator function that takes the value 1 if the clause is satisfied and 0 otherwise. For a set $J \subseteq [n]$, we use $\text{span}(J)$ to denote the subspace spanned by the standard basis vectors corresponding to the elements in J . We refer to the L_1 norm of $\gamma \in \mathbb{F}_2^n$ by $\|\gamma\|_1$. Given a subset $S \subseteq \mathbb{F}_2^n$, we denote $S^{=t} := S \cap \{u : \|u\|_1 = t\}$. Further, we define the degree of a function $f : \mathbb{F}_2^n \rightarrow \mathbb{R}$ to be $\max\{\|\gamma\|_1 : \hat{f}(\gamma) \neq 0\}$. We frequently interpret a linear transformation $M : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$ as a matrix and refer to the linear map obtained by taking the transpose of the matrix as M^T . At several points, we consider the compositions of functions with linear maps. For a function f and a map $M : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$, we denote by $f \circ M$ the composition of the functions f with M . In particular, $f \circ M(x) = f(M(x))$.

Probability. The following basic facts from probability theory are useful for us.

► **Fact 6** (Hoeffding, [11]). *Suppose X_1, \dots, X_n are such that $a \leq X_i \leq b$ for all i . Let $M = \frac{X_1 + \dots + X_n}{n}$. Then,*

$$\Pr\left[|M_n - \mathbf{E}M_n| \geq t\right] \leq 2 \exp\left(\frac{-2t^2n}{|b-a|}\right).$$

► **Definition 7** (Statistical Distance). *Let X and Y be two random variables taking values in a set \mathcal{S} . Then we define the statistical distance between X and Y as*

$$|X - Y| := \max_{\mathcal{T} \subseteq \mathcal{S}} \left| \Pr[X \in \mathcal{T}] - \Pr[Y \in \mathcal{T}] \right| = \frac{1}{2} \sum_{s \in \mathcal{S}} \left| \Pr[X = s] - \Pr[Y = s] \right|.$$

Linear Algebra. We recap two concepts from linear algebra, namely, orthogonal subspaces and direct sum, since they become useful for studying the Fourier spectrum of functions defined over subspaces of \mathbb{F}_2^n . For a subspace \mathcal{A} of \mathbb{F}_2^n , we denote the **orthogonal subspace** of \mathcal{A} as $\mathcal{A}^\perp = \{\gamma \in \mathbb{F}_2^n : \langle \gamma, \gamma' \rangle = 0, \forall \gamma' \in \mathcal{A}\}$. We denote by $\dim(\mathcal{A})$, the dimension of \mathcal{A} and $\text{codim}(\mathcal{A}) = n - \dim(\mathcal{A})$.

We now define the notion of the direct sum of two subspaces.

► **Definition 8** (Independence, Direct Sum). *Two subspaces \mathcal{A}, \mathcal{B} are independent if $a + b \neq 0$ for any non-trivial choice of $a \in \mathcal{A}$ and $b \in \mathcal{B}$. In addition, if $\{a + b : a \in \mathcal{A} \text{ and } b \in \mathcal{B}\} = \mathbb{F}_2^n$, we say that \mathbb{F}_2^n is a direct sum of \mathcal{A} and \mathcal{B} , written as $\mathcal{A} \oplus \mathcal{B} = \mathbb{F}_2^n$.⁵*

If $\mathcal{A} \oplus \mathcal{B} = \mathbb{F}_2^n$, then $\dim(\mathcal{A}) + \dim(\mathcal{B}) = n$. It is also well known that $\dim(\mathcal{A}^\perp) + \dim(\mathcal{A}) = n$. Note, however, that \mathcal{A}^\perp and \mathcal{A} need not be independent,⁶ and often in fact must not be.

► **Fact 9.** *Let \mathcal{A}, \mathcal{B} be independent subspaces of \mathbb{F}_2^n . Then for all distinct $b, b' \in \mathcal{B}$, the affine subspaces $b + \mathcal{A}$ and $b' + \mathcal{A}$ are mutually disjoint.*

Proof. If $b + a = b' + a'$, then a non-trivial sum of a vector from each \mathcal{A} and \mathcal{B} equals zero, contradicting the fact that $\mathcal{A} \oplus \mathcal{B} = \mathbb{F}_2^n$. ◀

Fourier Analysis. For $f : \mathbb{F}_2^n \rightarrow \mathbb{R}$, we can write f in the Fourier representation as

$$f(x) = \sum_{\gamma \in \mathbb{F}_2^n} \widehat{f}(\gamma) \chi_\gamma(x),$$

where $\chi_\gamma(x) := (-1)^{\langle \gamma, x \rangle}$ and $\widehat{f}(\gamma) = \mathbf{E}_x[f(x) \chi_\gamma(x)]$. We say f has degree d if $\max_{\gamma: \widehat{f}(\gamma) \neq 0} \|\gamma\|_1 = d$, and we refer to the degree d part of f by $f^{=d}(x) := \sum_{\|\gamma\|_1 = d} \widehat{f}(\gamma) \chi_\gamma(x)$. For more on this topic, see [17], which uses notation consistent with ours.

Restrictions. We are ultimately concerned with understanding the Fourier coefficients of a function when it is restricted to some affine subspace of \mathbb{F}_2^n . In the special case where the coordinates in a set $J \subseteq [n]$ are fixed using the vector $b \in \mathbb{F}_2^J$, we denote the restriction of f thus obtained as the function $f_{J \leftarrow b} : \text{span}(\overline{J}) \rightarrow \mathbb{R}$, which can be written as $f_{J \leftarrow b}(x) = f(x + b)$. Next, we recall the formula of the Fourier coefficients of the restricted function. Note that $\{\chi_\gamma(x) := (-1)^{\langle \gamma, x \rangle} : \gamma \in \text{span}(\overline{J})\}$ is a Fourier basis of the restricted function.

► **Fact 10** (Fourier Coefficients of Restricted Functions (see [17], Proposition 3.21)). *For every $\gamma \in \text{span}(\overline{J})$ and $b \in \text{span}(J)$,*

$$\widehat{f_{J \leftarrow b}}(\gamma) = \sum_{\beta \in \text{span}(J)} \widehat{f}(\beta + \gamma) \chi_\beta(b).$$

⁵ Such a subspace \mathcal{B} is sometimes called a **complement** of \mathcal{A} . However, this term can be confused with the orthogonal subspace/complement, so we avoid using this terminology.

⁶ this might be unexpected at first for those used to working over the reals, but it is essentially because the inner product over \mathbb{F}_2 allows self-orthogonal vectors in \mathbb{F}_2^n .

2.1 Fourier Analysis on Subspaces

We move to the general setting of restricting functions to arbitrary affine subspaces.⁷ Let $\mathcal{U} = \mathcal{V} + \alpha$ be an affine subspace of \mathbb{F}_2^n . By the restriction of f to \mathcal{U} , we mean the function $f_{\mathcal{U}} : \mathcal{V} \rightarrow \mathbb{R}$ defined as $f_{\mathcal{U}}(x) = f(x + \alpha)$.

For the remainder of this section (and paper), let \mathcal{W} be such that $\mathcal{W} \oplus \mathcal{V}^\perp = \mathbb{F}_2^n$. For each element $\gamma \in \mathcal{W}$, consider the function $\chi_\gamma : \mathcal{V} \rightarrow \{\pm 1\}$ as $\chi_\gamma(x) = (-1)^{\langle \gamma, x \rangle}$. It is easy to verify that $\{\chi_\gamma : \gamma \in \mathcal{W}\}$ form an orthonormal basis of real-valued functions defined over \mathcal{V} under the inner product given by $\langle p, q \rangle = \mathbf{E}_{x \in \mathcal{V}}[p(x)q(x)]$. We can therefore uniquely associate each vector $\gamma \in \mathcal{W}$ with the function χ_γ , and for $\mathcal{U} = \alpha + \mathcal{V}$, we can write

$$f_{\mathcal{U}}(x) = \sum_{\gamma \in \mathcal{W}} \widehat{f_{\mathcal{U}}}(\gamma) (-1)^{\langle \gamma, x \rangle}. \quad (1)$$

We now state the formal definition of δ -regularity.

► **Definition 11** (δ -regularity). *Let \mathcal{V} be a subspace of \mathbb{F}_2^n and $g : \mathcal{V} \rightarrow \mathbb{R}$. For $\delta \geq 0$, we say g is δ -regular if $\max_{\gamma \neq 0} |\widehat{g}(\gamma)| \leq \delta$.*

In this section, we present three separate formulas (Fact 12, Proposition 13 and Proposition 16) for the Fourier coefficients of $f_{\mathcal{U}}$, each of which is useful in different contexts. With the exception of Fact 12, which is direct, the proofs of the statements in this section can be found in Appendix B.

First, using the above observations, we have the following simple formula for the Fourier coefficients of $f_{\alpha + \mathcal{V}}$, which follows from the orthogonality of the χ_γ we have defined.

► **Fact 12.** *Let \mathcal{V}, \mathcal{W} be subspaces such that $\mathcal{W} \oplus \mathcal{V}^\perp = \mathbb{F}_2^n$ and $\mathcal{U} = \alpha + \mathcal{V}$. For any $\gamma \in \mathcal{W}$, we have that*

$$\widehat{f_{\mathcal{U}}}(\gamma) = \mathbf{E}_{x \in \mathcal{V}}[f(x + \alpha) \cdot (-1)^{\langle \gamma, x \rangle}] = (-1)^{\langle \gamma, \alpha \rangle} \mathbf{E}_{x \in \mathcal{U}}[f(x) \cdot (-1)^{\langle \gamma, x \rangle}].$$

Fact 12 represents a simple and analogous formula for Fourier coefficients of functions restricted to affine subspaces. It also highlights that the magnitude of the Fourier coefficients of a restricted function are unaffected by the choice for shift α as long it corresponds to the same affine subspace.

Our next formula, which shows how the Fourier coefficients of $f_{\mathcal{U}}$ can be written in terms of the Fourier coefficients of f , is an easy consequence of Fact 12.

► **Proposition 13.** *Let \mathcal{V}, \mathcal{W} be subspaces such that $\mathcal{W} \oplus \mathcal{V}^\perp = \mathbb{F}_2^n$ and $\mathcal{U} = \alpha + \mathcal{V}$. For any $\gamma \in \mathcal{W}$, we have*

$$\widehat{f_{\mathcal{U}}}(\gamma) = \sum_{\beta \in \gamma + \mathcal{V}^\perp} \widehat{f}(\beta) \cdot (-1)^{\langle \beta, \alpha \rangle}.$$

For a proof of Proposition 13 as well as proofs for the rest of the statements in this section, see Appendix B. We note that Proposition 13 gives a formula analogous to Fact 10 for restrictions to general affine subspaces. This fact will be useful to construct functions and argue that they never become δ -regular when restricted to any sufficiently large subspace. Before we give our final formula, we highlight one particular choice of \mathcal{W} such that $\mathcal{W} \oplus \mathcal{V}^\perp = \mathbb{F}_2^n$.

⁷ For an arbitrary subspace \mathcal{V} , there is no canonical mapping between vectors and characters when $\mathcal{V} \neq \mathbb{F}_2^n$, and we cannot simply define the vectors χ_γ , for each $\gamma \in \mathcal{V}$, as we did in the case of \mathbb{F}_2^n to be the characters of \mathcal{V} .

► **Definition 14** (M mapping \mathcal{V} to $\text{span}(J)$). Given a k -dimensional subspace \mathcal{V} , let $B = \{\beta_1, \dots, \beta_n\}$ be a basis for \mathbb{F}_2^n such that $\mathcal{V} = \text{span}(\{\beta_1, \dots, \beta_k\})$. For any subset $J \subseteq [n]$ of size k , let $M : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$ be an invertible linear map such that $\{M\beta_i : i \in [k]\} = \{e_j : j \in J\}$.

► **Proposition 15** (Choice of \mathcal{W}). Let \mathcal{V} , M and J be defined as in Definition 14. The subspaces $\mathcal{W} = \{M^\top \gamma : \gamma \in \text{span}(J)\}$ and \mathcal{V}^\perp are independent, and $\mathcal{W} \oplus \mathcal{V}^\perp = \mathbb{F}_2^n$.

Finally, we show that the Fourier coefficients of a function restricted to an affine subspace are the same as the Fourier coefficients of the function $f \circ M$ under a suitable (normal) restriction and for a particular choice of M .

► **Proposition 16**. Let \mathcal{V} , M and J be defined as in Definition 14 and $\mathcal{U} = \alpha + \mathcal{V}$. For any $\gamma \in \text{span}(J)$, we have $|\widehat{f_{\mathcal{U}}}(M^\top \gamma)| = |\widehat{h_{\mathcal{U}'}}(\gamma)|$, where $h = f \circ M^{-1}$ and $\mathcal{U}' = \{Mu : u \in \mathcal{U}\} = M\alpha + \text{span}(J)$ is a standard restriction.

Proposition 16 implies the following important corollary.

► **Corollary 17**. There exists an affine subspace \mathcal{U} of dimension k such that $f_{\mathcal{U}}$ is δ -regular if and only if there exists an invertible linear map $M : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$, a set $J \subseteq [n]$ of size k , and a fixing of coordinates outside J given by $b \in \mathbb{F}_2^{\overline{J}}$ such that the function $h_{\overline{J} \leftarrow b}$ is δ -regular, where $h = f \circ M$.

We use Corollary 17 crucially in the proof of Theorem 1, wherein we construct M and b such that $f \circ M_{\overline{[k]} \leftarrow b}$ has small Fourier coefficients. In the proof of this theorem we must understand the Fourier coefficients of $f \circ M$ in terms of the Fourier coefficients of f . The following fact gives an identity relating the Fourier coefficients of the two functions. For completeness, we include the proof in Appendix B.

► **Fact 18** ([17], Exercise 3.1). Let M be an invertible linear transformation, and consider the function $g = f \circ M^{-1} : \mathbb{F}_2^n \rightarrow \mathbb{R}$. Then we have that $\widehat{g}(\gamma) = \widehat{f}(M^\top \gamma)$.

3 Upper Bound on $r(f, \delta)$

Now we prove our main theorem, restated here for convenience.

► **Theorem 1**. For any $\delta \in (0, 1)$ and any degree d function $f : \mathbb{F}_2^n \rightarrow [-1, 1]$, we have $r(f, \delta) \leq n - \Omega(n^{1/d} (\log(n/\delta))^{-2})$.

First, we gain some intuition from degree one functions.

Base Case/Toy Example. Suppose f is a Fourier degree one function. In this case our function has the form

$$f(x) = \widehat{f}(0) + \sum_i \widehat{f}(e_i) (-1)^{x_i}.$$

For a parameter $t \geq 1$ and a subset $S \subseteq [t]$, consider the sum $g_S = \widehat{f}(0) + \sum_{i \in S} \widehat{f}(e_i)$. Note that $g_S = \mathbf{E}[f(x) | x_i = 0 \ \forall i \in S] \in [-1, 1]$. The pigeonhole principle implies that for $t = \Omega(\log 1/\delta)$ there must exist two distinct sets S, S' such that the difference $|g_S - g_{S'}| \leq \delta$. We can further write $g_S - g_{S'} = \sum_{i \in S \Delta S'} \widehat{f}(e_i) (-1)^{|x_i| \cap S'}$.

We now use the set $S \Delta S'$ and the signs to construct an affine subspace where at least one Fourier coefficient will have small magnitude. Assume without loss of generality that $1 \in S \setminus S'$ and $S \Delta S' = [t']$ for some $t' \leq t$. Consider restricting f to the affine subspace \mathcal{U}

defined by the linear equations $x_1 + x_i = b_i$ for each $i \in \{2, \dots, t'\}$, where $b_i = |\{i\} \cap S'|$. We can reason about the Fourier spectrum of $f_{\mathcal{U}}$ by plugging in $x_i = b_i + x_1$. Under this restriction, we see that the Fourier coefficients of $e_{t'+1}, \dots, e_n$ stay the same, and the new Fourier coefficient of e_1 is exactly equal to

$$\widehat{f}(e_1) + \sum_{i=2}^{t'} \widehat{f}(e_i)(-1)^{b_i} = g_S - g_{S'},$$

which we observed has magnitude at most δ . Repeatedly applying this argument roughly $n(\log(1/\delta))^{-1}$ times for the remaining standard basis vectors and fixing remaining coordinates arbitrarily, we obtain an affine subspace of dimension at least $\Omega\left(\frac{n}{\log(1/\delta)}\right)$.

Theorem 1 is proved via induction using the following lemma.

► **Lemma 19.** *For $\tau \in (0, 1)$ and any degree d function $f : \mathbb{F}_2^n \rightarrow [-1, 1]$, there exists an invertible linear map $M : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$, a set $J \subseteq [n]$ with size at least $\frac{d}{4e} \left(\frac{n}{\log 5/\tau}\right)^{1/d}$, and $b \in \text{span}(\bar{J})$ such that $h = f \circ M$ satisfies*

$$\left| \widehat{h_{\bar{J} \leftarrow b}}(\gamma) \right| \leq \begin{cases} \tau & \text{if } \|\gamma\|_1 = d, \\ 0 & \text{for all } \|\gamma\|_1 > d. \end{cases}$$

We now prove Theorem 1 using Lemma 19.

Proof. The proof proceeds by induction over the degree. Our inductive hypothesis is that for any $\delta > 0$ and any degree d function f , there exists an invertible linear map M , a set $I \subseteq [n]$, and $b \in \text{span}(\bar{I})$ such that the following two items hold:

1. $h_{\bar{I} \leftarrow b}$ is δ -regular, where $h = f \circ M$, and
2. for $C_d = \sum_{i=1}^d (i!)^{-1}$, we have

$$|I| \geq \frac{n^{1/d!}}{(8e)^{C_{d-1}} (\log(n/\delta))^{C_d}}.$$

Note that $C_d \leq e - 1 < 2$ for all $d \geq 1$. The existence of the desired affine subspace is then given by Corollary 17, and its dimension is equal to $|I| \geq \Omega\left(n^{1/d!} (\log(n/\delta))^{-2}\right)$.

The base case corresponds to the degree being one. Let us apply Lemma 19 for degree one with $\tau = \delta$ and denote $g = f \circ M$, where M is the linear map M promised by the lemma. Additionally, we have a set J of size at least $\frac{n}{4e \log 5/\delta} \geq \Omega\left(\frac{n}{\log n/\delta}\right)$, and $b \in \text{span}(\bar{J})$ such that

$$\left| \widehat{g_{\bar{J} \leftarrow b}}(\gamma) \right| \leq \begin{cases} \tau & \text{if } \|\gamma\|_1 = 1, \\ 0 & \text{for all } \|\gamma\|_1 > 1. \end{cases} \implies \left| \widehat{g_{\bar{J} \leftarrow b}}(\gamma) \right| \leq \delta, \text{ for all } \gamma \neq 0.$$

Assuming both items hold for some degree $d - 1$, we show them for degree d . Applying Lemma 19 with degree d and $\tau = n^{-d}\delta/3$, we denote $p := (f \circ M)_{\bar{J} \leftarrow b}$, where M , J and b are as promised by the lemma. Note that, by Lemma 19, p has degree at most d , and for any γ with $\|\gamma\|_1 = d$, we have, $|\widehat{p}(\gamma)| \leq \delta/(3n^d)$. Consider the functions $p^{<d}$ and $p^{=d}$, which are the degree at most $d - 1$ part of p and the degree d part of p , respectively. We note that $\frac{p^{<d}}{(1+\delta/3)}$ is bounded in the interval $[-1, 1]$ because for any x ,

$$\left| p^{<d}(x) \right| \leq |p(x)| + |p^{=d}(x)| \leq 1 + \sum_{\gamma: \|\gamma\|_1 = d} |\widehat{p}(\gamma)| \leq 1 + \frac{\delta}{3}.$$

83:10 Searching for Regularity in Bounded Functions

Applying the inductive hypothesis⁸ to $\frac{p^{<d}}{1+\delta/3}$ for the choice of $\delta/3$, we get a linear map M' , a set $I \subseteq J$, and $b' \in \text{span}(J \setminus I)$ such that $\left(\frac{q}{1+\delta/3}\right)_{J \setminus I \leftarrow b'}$ is $\delta/3$ -regular, where $q := p^{<d} \circ M'$. Therefore, for any $\gamma \neq 0$, we have $|\widehat{q_{J \setminus I \leftarrow b'}}(\gamma)| \leq (1 + \frac{\delta}{3}) \frac{\delta}{3} < \frac{2\delta}{3}$. Denoting $p' := p \circ M'$ and $r := p^{=d} \circ M'$, we have for any $\gamma \neq 0$ that

$$\left| \widehat{p'_{J \setminus I \leftarrow b'}}(\gamma) \right| \leq \left| \widehat{q_{J \setminus I \leftarrow b'}}(\gamma) \right| + \left| \widehat{r_{J \setminus I \leftarrow b'}}(\gamma) \right| < 2\delta/3 + \sum_{\beta: \|\beta\|_1=d} |\widehat{g}(\beta)| \leq \delta.$$

This shows that $p'_{J \setminus I \leftarrow b'}$ is δ -regular. Moreover, if we extend M' to act as the identity map on the coordinates in \overline{J} , we can write

$$\begin{aligned} p'_{J \setminus I \leftarrow b'}(x) &= (p \circ M')_{J \setminus I \leftarrow b'}(x) = p(M'(x + b')) \\ &= (f \circ M)_{J \leftarrow b}(M'(x + b')) = f(MM'(x + b' + b)), \end{aligned}$$

which implies that item 1 of the inductive hypothesis is satisfied by applying the linear map MM' and restricting to the set I by fixing the coordinates outside according to $b + b'$.

We now show that the size of I satisfies item 2 above. Note that Lemma 19 promises that $|J| \geq \frac{d}{4e} \left(\frac{n}{\log(15n^d/\delta)}\right)^{1/d}$. Moreover, we have

$$\log(15n^d/\delta) \leq d \log n/\delta + \log 15 \leq 4d \log n/\delta,$$

where the last inequality follows for sufficiently large n . Therefore, $|J| \geq \frac{1}{8e} \left(\frac{n}{\log(n/\delta)}\right)^{1/d}$. Moreover, we assume without loss of generality that $3|J| \leq n$ because, if not, we can arbitrarily fix coordinates in J until it is, which does not affect the crucial property that all remaining degree d Fourier coefficients have small magnitude. Using the bounds on $|J|$ and applying item 2 of the inductive hypothesis for degree $d - 1$, we get

$$\begin{aligned} |I| &\geq \frac{|J|^{1/(d-1)!}}{(8e)^{C_{d-2}} (\log(3|J|/\delta))^{C_{d-1}}} \geq \frac{n^{1/d!}}{(8e)^{C_{d-1}} \log(n/\delta)^{1/d!} (\log(3|J|/\delta))^{C_{d-1}}} \\ &\geq \frac{n^{1/d!}}{(8e)^{C_{d-1}} (\log(n/\delta))^{C_d}}. \end{aligned}$$

This shows item 2 of the inductive hypothesis as desired. \blacktriangleleft

To prove Lemma 19, we need the following claim, which ultimately lets us bound Fourier coefficients in certain affine subspaces.

\triangleright **Claim 20 (Pigeonhole Principle).** Let $f : \mathbb{F}_2^n \rightarrow [-1, 1]$ be degree d . For every $K \subseteq [n]$ of size k such that $n - k \geq \binom{k}{d-1} \log(5/\tau)$, there exists $S \subseteq [n] \setminus K$ and $z \in \{\pm 1\}^S$ such that

1. $\forall \gamma \in \text{span}(K)$ with $\|\gamma\|_1 = d - 1$, we have $\left| \sum_{j \in S} \widehat{f}(\gamma + e_j) \cdot z_j \right| \leq \tau$, and
2. $1 < |S| \leq \binom{k}{d-1} \log(5/\tau)$.

Proof. Consider any subset of $T \subseteq \overline{K}$ of size $\binom{k}{d-1} \log(5/\tau)$. For any $U \subseteq T$, consider the sum

$$a_U(\gamma) := \widehat{f}(\gamma) + \sum_{j \in U} \widehat{f}(\gamma + e_j).$$

⁸ Technically, $\frac{p^{<d}}{1+\delta/3} : \text{span}(J) \rightarrow [-1, 1]$. However, we can abuse notation slightly and consider it as a function from \mathbb{F}_2^J to $[-1, 1]$ in order to apply the inductive hypothesis.

We must have that $a_U(\gamma) \in [-1, 1]$ since it is exactly equal to the Fourier coefficient corresponding to γ if we restricted everything in U to be one. This follows because f is degree d .

Now, divide the interval $[-1, 1]$ into $2/\tau$ intervals of length τ . For a fixed $U \subseteq T$ of even size, consider putting the values of $a_U(\gamma)$ for all $\gamma \in \text{span}(K)^{=d-1}$ into a vector v_U of length $\binom{k}{d-1}$. First, note that the number of even subsets of T is at least $2^{\binom{k}{d-1} \log(5/\tau) - 1} > (2/\tau)^{\binom{k}{d-1}}$. Moreover, the number of possible interval vectors is at most $(2/\tau)^{\binom{k}{d-1}}$. Therefore, by the pigeonhole principle, there must be two distinct sets $U, U' \subseteq T$ such that $\|v_U - v_{U'}\|_\infty \leq \tau$.

Thus, we have that

$$\|v_U - v_{U'}\|_\infty \leq \tau \iff \sum_{i \in U \Delta U'} (-1)^{|\{i\} \cap U'|} \widehat{f}(\gamma + e_i) \leq \tau \quad \forall \gamma \in \text{span}(K)^{=d-1}.$$

Since U, U' have even size and are not equal, $U \Delta U'$ has even size as well, so we can set our $S = U \Delta U' \subseteq T$ and $z_i = (-1)^{|\{i\} \cap U'|}$, and the claim follows. \triangleleft

We can now prove Lemma 19.

Proof of Lemma 19. We build the map M , the set J , and the vector b iteratively. Throughout the iterations, we seek to maintain a set K of coordinates for which (under a suitable linear transformation M) every Fourier coefficient corresponding to a vector of weight d in $\text{span}(K)$ has magnitude at most τ . We build K one coordinate at a time by repeatedly invoking Claim 20 and arguing that the quantities guaranteed to be small by Claim 20 are exactly the (new) Fourier coefficients. When we can no longer add more coordinates to K , we fix any remaining coordinates (outside of K that are still alive), and we are left with a function, over only the coordinates in K , that has the desired property.

Note that we can start with K being an arbitrary subset of size $d - 1$ (w.l.o.g. let it be $[d - 1]$) since any such subset has no Fourier coefficients of degree d . Therefore, we can assume without loss of generality that $\tau \geq 5 \cdot 2^{-n/(4e)^d}$, since otherwise $\frac{d}{4e} \left(\frac{n}{\log(5/\tau)} \right)^{1/d} < d$ and the lemma becomes trivial. In each iteration, we maintain the following invariant for M , J and b . In iteration i , there exists some $K \subseteq J$ of size $d + i - 1$ such that the function $g = (f \circ M)_{\overline{J} \leftarrow b}$ satisfies

$$|\widehat{g}(\gamma)| \leq \begin{cases} \tau & \text{if } \gamma \in \text{span}(K) \text{ and } \|\gamma\|_1 = d, \\ 0 & \text{for all } \|\gamma\|_1 > d. \end{cases}$$

Assume without loss of generality that $J = [j]$ for some $j \leq n$ and $K = [d + i - 1] \subseteq J$. Since g has degree d , we can apply Claim 20 to g and obtain a subset $S \subseteq J \setminus K$ of size at most $\binom{d+i-1}{d-1} (\log(5/\tau))$ and a sign vector $z \in \{\pm 1\}^S$ so that

$$\left| \sum_{j \in S} \widehat{g}(\gamma + e_j) \cdot z_j \right| \leq \tau, \quad \text{for all } \gamma \in \text{span}([d + i - 1]) \text{ such that } \|\gamma\|_1 = d - 1. \quad (2)$$

We can also assume that $d + i \in S$ and $z_{d+i} = 1$. Now consider the invertible linear transformation $M_i : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$ that maps e_{d+i} to $\sum_{j \in S} e_j$ and behaves as the identity map on the remaining standard basis vectors. Further, denote $J_i := S \setminus \{d + i\}$ and let $b_i \in \text{span}(J_i)$, where $(b_i)_j := (1 - z_j)/2$ for each $j \in J_i$. Intuitively, applying the linear transformation M_i and then fixing the coordinates in J_i to b_i corresponds to restricting the affine subspace described by the equations $x_j + x_{d+i} = (1 - z_j)/2$ for all $j \in J_i$.

83:12 Searching for Regularity in Bounded Functions

After this iteration, we show that if we set $M' \leftarrow MM_i$, $J' \leftarrow J \setminus J_i$ and $b' \leftarrow b + b_i$, the invariant holds with $K' \leftarrow K \cup \{d + i\}$. For these choices, we have

$$\begin{aligned} (f \circ M')_{\overline{J'} \leftarrow b'}(x) &= f \circ M(M_i(x + b')) = f \circ M(M_i(x + b_i + b)) \\ &= f \circ M(M_i(x + b_i) + b) \\ &= g \circ M_i(x + b_i) = (g \circ M_i)_{J_i \leftarrow b_i}(x), \end{aligned}$$

and it therefore suffices to show that $(g \circ M_i)_{J_i \leftarrow b_i}$ – denoted by h henceforth, for shorthand – is degree d and $|\widehat{h}(\gamma)| \leq \tau$ for all $\gamma \in \text{span}([d + i])$ with $\|\gamma\|_1 = d$. We start by analyzing the Fourier coefficients of h , for which by Fact 10 we have

$$\widehat{h}(\gamma) = \sum_{\beta \in \text{span}(J_i)} \widehat{g \circ M_i}(\gamma + \beta)(-1)^{\langle \beta, b_i \rangle}. \quad (3)$$

Next, we observe the following relation between the Fourier coefficients of $g \circ M_i$ and those of g , which we use to simplify Equation (3). Denoting $v := \sum_{j \in J_i} e_j$, we claim that, for any γ ,

$$\widehat{g \circ M_i}(\gamma) = \widehat{g}(\gamma + e_{d+i} \langle \gamma, v \rangle). \quad (4)$$

Before proving Equation (4), we use it to prove that h has the desired properties. Note that since g is degree d , Equation (4) implies that if $\widehat{g \circ M_i}(\gamma) \neq 0$, then $\|\gamma + e_{d+i} \langle \gamma, v \rangle\|_1 \leq d$, which in turn implies that $\|\gamma\|_1 \leq d + 1$. This immediately tells us that $g \circ M_i$ has degree at most $d + 1$; therefore, h also has degree at most $d + 1$ since the degree cannot increase under restrictions. Now, for any γ , Equation (3) reduces to

$$\begin{aligned} \widehat{h}(\gamma) &= \sum_{\substack{\beta \in \text{span}(J_i), \\ \|\beta\|_1 \leq d+1-\|\gamma\|_1}} \widehat{g \circ M_i}(\gamma + \beta)(-1)^{\langle \beta, b_i \rangle} \\ &= \widehat{g \circ M_i}(\gamma) + \sum_{\substack{\beta \in \text{span}(J_i), \\ 0 < \|\beta\|_1 \leq d+1-\|\gamma\|_1}} \widehat{g \circ M_i}(\gamma + \beta)(-1)^{\langle \beta, b_i \rangle} \\ &= \widehat{g}(\gamma + e_{d+i} \langle \gamma, v \rangle) + \sum_{\substack{\beta \in \text{span}(J_i), \\ 0 < \|\beta\|_1 \leq d+1-\|\gamma\|_1}} \widehat{g}(\gamma + \beta + e_{d+i} \langle \gamma + \beta, v \rangle)(-1)^{\langle \beta, b_i \rangle}, \end{aligned} \quad (5)$$

where, in the first equality, we used the fact that if $\|\beta\|_1 > d + 1 - \|\gamma\|_1$, then $\|\beta + \gamma\|_1 > d + 1$ and the corresponding Fourier coefficient in $g \circ M_i$ is just zero, and in the last equality, we used Equation (4). Moreover, for any $\gamma \in \text{span}(J \setminus J_i)$, we have $\langle \gamma, v \rangle = 0$, which means that $\widehat{g}(\gamma + e_{d+i} \langle \gamma, v \rangle) = \widehat{g}(\gamma)$. We can now conclude that h has degree at most d . Indeed, if $\|\gamma\|_1 \geq d + 1$, then Equation (5) implies that $\widehat{h}(\gamma) = \widehat{g}(\gamma) = 0$ since g has degree at most d .

Next, we show that for any $\gamma \in \text{span}([d + i])$ with $\|\gamma\|_1 = d$, it must be that $|\widehat{h}(\gamma)| \leq \tau$. Applying Equation (5) for such γ , we note that

$$\begin{aligned} \widehat{h}(\gamma) &= \widehat{g}(\gamma) + \sum_{j \in J_i} \widehat{g}(\gamma + e_j + e_{d+i} \langle \gamma + e_j, v \rangle)(-1)^{\langle e_j, b_i \rangle} \\ &= \widehat{g}(\gamma) + \sum_{j \in J_i} \widehat{g}(\gamma + e_j + e_{d+i})z_j. \end{aligned}$$

We now consider two cases. First, when $\gamma_{d+i} = 0$, the above equation implies that $\widehat{h}(\gamma) = \widehat{g}(\gamma)$ since $\|\gamma + e_{d+i} + e_j\|_1 = d + 2$ for every $j \in J_i$, and g has degree at most d . Therefore, in

this case, $|\widehat{h}(\gamma)| = |\widehat{g}(\gamma)| \leq \tau$ by the inductive hypothesis. Otherwise, $\gamma_{d+i} = 1$, and now using both Equation (2) and the fact that $\gamma + e_{d+i} \in \text{span}(\{e_1, \dots, e_{d+i-1}\})$, we conclude that $|\widehat{h}(\gamma)| = \left| \sum_{j \in S} \widehat{g}((\gamma + e_{d+i}) + e_j) z_j \right| \leq \tau$.

It remains to show Equation (4). We start by observing that $M_i = M_i^{-1}$, which can be verified by noting that $M_i^{-1} e_{d+i} = M_i^{-1}(e_{d+i} + v + v) = e_{d+i} + v$ and M_i^{-1} acts as the identity map on the remaining standard basis vectors. From Fact 18, we know that $\widehat{g \circ M_i^\top}(\gamma) = \widehat{g \circ M_i^{-1}}(\gamma) = \widehat{g}(M_i^\top \gamma)$. Since the rows of M_i^\top are the same as the columns of M_i , we have

$$(M_i^\top \gamma)_j = \begin{cases} \langle v + e_{d+i}, \gamma \rangle & \text{if } j = d+i, \\ \gamma_j & \text{otherwise.} \end{cases}$$

Therefore, we can write $M_i^\top \gamma = \sum_{j \neq d+i} \gamma_j e_j + e_{d+i} \langle v + e_{d+i}, \gamma \rangle = \gamma + e_{d+i} \langle v, \gamma \rangle$, as claimed.

We conclude the argument by calculating how many times we can repeat the above procedure. Note that, in the i -th iteration, we fixed at most $\binom{d+i-1}{d-1} \log 5/\tau - 1$ coordinates and we added exactly one coordinate to K . We can thus continue this process until iteration t for the largest value of t such that

$$\log(5/\tau) \cdot \left(\sum_{i=1}^t \binom{d+i-1}{d-1} \right) \leq n - d + 1.$$

Simplifying the binomial sum, we get

$$\begin{aligned} \sum_{i=1}^t \binom{d+i-1}{d-1} &= \sum_{i=1}^t \binom{d+i-1}{i} = \sum_{i=1}^t \binom{d+i-1}{i} + \binom{d}{0} - 1 \\ &= \binom{d+t}{t} - 1 < \left(\frac{e(d+t)}{d} \right)^d, \end{aligned}$$

where the last equality follows by repeatedly using the identity $\binom{a}{i} + \binom{a}{i-1} = \binom{a+1}{i}$. Thus, we can set $t = \frac{d}{e} \left(\frac{n-d+1}{\log 5/\tau} \right)^{1/d} - d$. Adding in the initial $d-1$ coordinates, at the end of the t iterations, we can bound $|K|$ as,

$$\begin{aligned} |K| &= \frac{d}{e} \left(\frac{n-d+1}{\log 5/\tau} \right)^{1/d} - d + d - 1 \\ &\geq \frac{d}{e} \left(\frac{n}{\log 5/\tau} \left(1 - \frac{d-1}{n} \right) \right)^{1/d} - 1 \\ &\geq \frac{d}{e} \left(\frac{n}{\log 5/\tau} \cdot \frac{1}{d} \right)^{1/d} - 1 \\ &\geq \frac{d}{2e} \left(\frac{n}{\log 5/\tau} \right)^{1/d} - 1 && (d^{1/d} \leq 2 \quad \forall d \geq 1) \\ &= \frac{d}{4e} \left(\frac{n}{\log 5/\tau} \right)^{1/d} + \frac{d}{4e} \left(\frac{n}{\log 5/\tau} \right)^{1/d} - 1 \\ &\geq \frac{d}{4e} \left(\frac{n}{\log 5/\tau} \right)^{1/d} + d - 1 && (\text{since } \tau \geq 5 \cdot 2^{-n/(4e)^d}) \\ &\geq \frac{d}{4e} \left(\frac{n}{\log 5/\tau} \right)^{1/d}. && (d \geq 1) \end{aligned}$$

83:14 Searching for Regularity in Bounded Functions

At the end of t iterations, we can fix any coordinates outside the set K arbitrarily to ensure that the only non-zero Fourier coefficients with L_1 norm d in the resulting function must correspond to vectors in $\text{span}(K)$, which do not change under the restriction. ◀

4 Applications

We now present an application of Theorem 1 that shows a tradeoff between the dimension of a disperser and its Fourier degree, and a connection to extractors, as well. First, we introduce a definition that generalizes Boolean functions and helps us reason about the Fourier spectrum of dispersers.

► **Definition 21.** We say a function $f : \mathbb{F}_2^n \rightarrow \mathbb{R}$ is G -granular if for every $x \in \mathbb{F}_2^n$, we have that $f(x)$ is an integer multiple of G .

▷ **Claim 22.** If a degree d function $f : \mathbb{F}_2^n \rightarrow \mathbb{R}$ is G -granular, then for every $\gamma \in \mathbb{F}_2^n$, we have that $\tilde{f}(\gamma)$ is an integer multiple of $2^{-d} \cdot G$.

We defer the proof of Claim 22 to the full version. Assuming the claim, we now show that low degree granular functions cannot have a large parity kill number. As a consequence, we get that low-degree affine dispersers cannot have small dimension (Corollary 2).

► **Lemma 23.** Every degree d function $f : \mathbb{F}_2^n \rightarrow [-1, 1]$ that is G -granular satisfies

$$C_{\min}^{\oplus}[f] \leq n - \Omega\left(n^{1/d}(d + \log n/G)^{-2}\right).$$

Proof. If f is G -granular and degree d , then from Claim 22 we know that all its Fourier coefficients must be integer multiples of $2^{-d} \cdot G$. Moreover, a Fourier coefficient of f in any affine subspace is simply a signed sum of the Fourier coefficients of f and therefore it must also be an integer multiple of $2^{-d} \cdot G$. This shows that if f is δ -regular in some affine subspace \mathcal{U} with $\delta < 2^{-d} \cdot G$, then $f_{\mathcal{U}}$ must be constant. The lemma follows by using Theorem 1 for $\delta = 2^{-d-1} \cdot G$. ◀

Proof of Corollary 2. Using f , we can construct a degree d function $h : \mathbb{F}_2^n \rightarrow [-1, 1]$ as $h(x) = 1 - \frac{2f(x)}{C}$. Noting that h is $2/C$ -granular and using the above lemma, it follows that

$$C_{\min}^{\oplus}[f] = C_{\min}^{\oplus}[h] \leq n - \Omega\left(n^{1/d}(d + \log(nC))^{-2}\right),$$

which shows that there is some affine subspace of dimension at least $\Omega\left(n^{1/d}(2d + \log(nC))^{-2}\right)$ where f is constant. ◀

► **Remark 24.** Affine dispersers can be viewed as a relaxation of affine extractors, objects that have been well studied in the pseudorandomness literature. In a similar vein, we observe a connection between the notion of δ -regularity and affine extractors. In particular, we note that affine extractors can be viewed as functions that are δ -regular in *all* affine subspaces of sufficiently large dimension. We detail this connection in the full version of this work.

5 Lower Bounds on $r(f, \delta)$

In this section, we prove lower bounds on $r(f, \delta)$. With the exception of Lemma 26, we defer the proofs of all the statements to the appendices. We start with lower bounds for functions f that are bounded in the interval $[-1, 1]$; in the subsequent section, we give lower bounds for Boolean functions. For detailed sketches of all the results in this section, see Appendix A.2.

5.1 Bounded Functions

We begin with a simple bound on the number of standard basis vectors in low-dimensional affine subspaces, which is crucial in the analysis of the lower bounds.

► **Observation 25.** *For a subspace $\mathcal{V} \subseteq \mathbb{F}_2^n$ of co-dimension C , and \mathcal{W} such that $\mathcal{W} \oplus \mathcal{V}^\perp = \mathbb{F}_2^n$, there exists a set $S \subseteq \mathcal{W}$ of size at least $n - C$ such that for every $u \in S$,*

$$|(u + \mathcal{V}^\perp)^{=1}| \geq 1.$$

Moreover, there exists a subset $S_1 \subseteq S$ of size at least $n - 2C$ whose corresponding shifts contain exactly one standard basis vector.

The proof of Observation 25 can be found in Appendix C. Using this observation, we provide an example of a degree one function f that witnesses large values for $r(f, \delta)$.

► **Lemma 26.** *There is a degree one function $f : \mathbb{F}_2^n \rightarrow [-1, 1]$ for which $r(f, \delta) \geq n/2$, for all $\delta < 1/n$.*

Proof. The counterexample is given by the function $f(x) = \frac{1}{n} \cdot \sum_i (-1)^{e_i \cdot x}$. Let \mathcal{V} be a subspace of \mathbb{F}_2^n of co-dimension C , and suppose we restrict the function to the affine subspace $\mathcal{U} = \alpha + \mathcal{V}$. By Observation 25, if $C \leq n/2 - 1$, there exists at least two vectors $\gamma, \gamma' \in \mathcal{W}$ (where \mathcal{W} is such that $\mathcal{W} \oplus \mathcal{V}^\perp = \mathbb{F}_2^n$) such that $|(\gamma + \mathcal{V}^\perp)^{=1}| = |(\gamma' + \mathcal{V}^\perp)^{=1}| = 1$. Assume without loss of generality that $\gamma \neq 0$. Then, by Proposition 13, we have that

$$|\widehat{f_{\mathcal{U}}}(\gamma)| = \left| \sum_{\eta \in \mathcal{U} + \mathcal{V}^\perp} \widehat{f}(\eta) (-1)^{\langle \eta, \alpha \rangle} \right| = \frac{1}{n} > \delta,$$

which follows by observing that exactly one of the summands in the last sum corresponds to a weight one vector and is non-zero. Therefore, $r(f, \delta) \geq n/2$. ◀

We next show that Lemma 26 can be generalized to degree d bounded functions.

► **Lemma 27.** *For $d > 2$ and $\delta < \binom{n}{d}^{-1}$, there exists a degree d function $f : \mathbb{F}_2^n \rightarrow [-1, 1]$ for which $r(f, \delta) \geq n - 2dn^{1/(d-1)}$.*

The proof of Lemma 27 can also be found in Appendix C. We note that unlike the degree one case, the functions achieving the lower bound in the above lemma are not explicit. Additionally, when $d = 2$, we see that Lemma 27 is trivial; it would be interesting to obtain a tighter result in this case.

5.2 Boolean Functions

This section has two parts. The first gives non-explicit lower bounds on $r(f, \delta)$ for Boolean functions, and the second gives explicit lower bounds.

5.2.1 Non-explicit Lower Bounds on $r(f, \delta)$

We can turn our lower bounds on $r(f, \delta)$ for bounded functions into (non-explicit) lower bounds for Boolean functions. To do so, we use the following simple but powerful lemma of [12], which states that given a bounded function with a large $r(f, \delta)$, there must exist some Boolean function g with similarly a large $r(g, 2\delta)$.

► **Proposition 28** ([12], Claim 1.2). *Let $\tau > 0$ and $f : \mathbb{F}_2^n \rightarrow [-1, 1]$. There exists a Boolean function $g : \mathbb{F}_2^n \rightarrow \{\pm 1\}$ satisfying, for every affine subspace \mathcal{U} such that $|\mathcal{U}| \geq \frac{4n^2}{7^2}$ and any $\gamma \in \mathbb{F}_2^n$, that $\left| \widehat{f}_{\mathcal{U}}(\gamma) - \widehat{g}_{\mathcal{U}}(\gamma) \right| \leq \tau$.*

Using Proposition 28, we have the following lemma.

► **Lemma 29.** *For all $d \geq 3$ and $\delta < \frac{1}{2} \cdot \binom{n}{d}^{-1}$, there exists a Boolean function f with*

$$r(f, \delta) \geq n - \max \left\{ 2d \cdot n^{1/(d-1)}, \log(16n^2/\delta^2) \right\}.$$

Proof. By Lemma 27, there exists a bounded f that is not δ -regular in any affine subspace of dimension at least $2dn^{1/(d-1)}$ for all $\delta < \binom{n}{d}^{-1}$. Proposition 28 tells us that there exists a Boolean function g whose Fourier coefficients agree up to an additive error $\delta/2$ with the Fourier coefficients of f on all affine subspaces of dimension at least $\log(16n^2/\delta^2)$. Therefore, if f is not δ -regular on all of these affine subspaces, then g is also not $\delta/2$ -regular on any of these subspaces. ◀

We can plug some parameters into Lemma 29 and achieve the following more parsable corollary.

► **Corollary 30.** *For every $3 \leq d \leq \frac{\log n}{\log \log n + 1}$ and $\delta = \frac{1}{2} \cdot n^{-d}$, there exists a Boolean function f with $r(f, \delta) \geq n - 2d \cdot n^{1/(d-1)}$.*

We include the proofs of Proposition 28 and Corollary 30 in Appendix C.

5.2.2 Explicit Lower Bounds on $r(f, \delta)$

We now turn to lower bounds on $r(f, \delta)$ given by explicit Boolean functions. Our first example comes from analyzing certain Boolean functions studied by [18] that provide lower bounds for $r(f, 0)$.

► **Lemma 31** (Related to Corollary 1.1 in [18]). *For each $\delta > 0$, there exists an explicit Boolean function $f : \mathbb{F}_2^n \rightarrow \{0, 1\}$ with $r(f, \delta) = \Omega\left(\left(\log \frac{1}{\delta}\right)^{\log_2(3)}\right)$.*

Our second example is the majority function; we show that the majority function, denoted by MAJ_n , has a large $r(f, \delta)$ value when $\delta = O(1/\sqrt{n})$. This in particular rules out the possibility of proving a bound of the form $r(f, \delta) \leq \text{poly}(\log(1/\delta))$ for Boolean f .

► **Lemma 32.** *There is an absolute constant $C > 0$, such that for all sufficiently large n , $r(\text{MAJ}_n, \delta) \geq \Omega(n^{1/2})$ for any $\delta \leq C/\sqrt{n}$.*

The proofs of Lemma 31 and Lemma 32 can be found in the full version of this work.

References

- 1 Noga Alon, Eldar Fischer, Michael Krivelevich, and Mario Szegedy. Efficient testing of large graphs. *Combinatorica*, 20, May 2001. doi:10.1007/s004930070001.
- 2 Shalev Ben David. *Quantum speedups in query complexity*. PhD thesis, Massachusetts Institute of Technology, 2017.
- 3 Eli Ben-Sasson and Swastik Kopparty. Affine dispersers from subspace polynomials. *SIAM J. Comput.*, 41(4):880–914, 2012.
- 4 Harry Buhrman and Ronald de Wolf. Complexity measures and decision tree complexity: a survey. *Theoretical Computer Science*, 288(1):21–43, 2002. doi:10.1016/S0304-3975(01)00144-X.
- 5 Eshan Chattopadhyay, Jesse Goodman, and Jyun-Jie Liao. Affine extractors for almost logarithmic entropy. In *FOCS*, pages 622–633. IEEE, 2021.

- 6 Gil Cohen and Avishay Tal. Two structural results for low degree polynomials and applications. In *APPROX-RANDOM*, volume 40 of *LIPICs*, pages 680–709. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2015.
- 7 Alan M. Frieze and Ravi Kannan. The regularity lemma and approximation schemes for dense problems. In *FOCS*, pages 12–20. IEEE Computer Society, 1996.
- 8 Uma Girish, Justin Holmgren, Kunal Mittal, Ran Raz, and Wei Zhan. Parallel repetition for the GHZ game: A simpler proof. In *APPROX-RANDOM*, volume 207 of *LIPICs*, pages 62:1–62:19. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021.
- 9 Ben Green. A szemerédi-type regularity lemma in abelian groups. *Geometric and Functional Analysis*, 15:340–376, January 2005. doi:10.1007/s00039-005-0509-8.
- 10 A. W. Hales and R. I. Jewett. Regularity and positional games. *Transactions of the American Mathematical Society*, 106(2):222–229, 1963.
- 11 Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963. URL: <http://www.jstor.org/stable/2282952>.
- 12 Kaave Hosseini, Shachar Lovett, Guy Moshkovitz, and Asaf Shapira. An improved lower bound for arithmetic regularity. *Mathematical Proceedings of the Cambridge Philosophical Society*, 161(2):193–197, 2016. doi:10.1017/S030500411600013X.
- 13 Zander Kelley and Raghu Meka. Strong bounds for 3-progressions, 2023. arXiv:2302.05537.
- 14 Xin Li. Improved two-source extractors, and affine extractors for polylogarithmic entropy. In *FOCS*, pages 168–177. IEEE Computer Society, 2016.
- 15 Roy Meshulam. On subsets of finite abelian groups with no 3-term arithmetic progressions. *J. Comb. Theory Ser. A*, 71(1):168–172, July 1995. doi:10.1016/0097-3165(95)90024-1.
- 16 Gatis Midrijanis. Exact quantum query complexity for total boolean functions, 2004. doi:10.48550/arXiv.quant-ph/0403168.
- 17 Ryan O’Donnell. Analysis of boolean functions. *CoRR*, abs/2105.10386, 2021. arXiv:2105.10386.
- 18 Ryan O’Donnell, John Wright, Yu Zhao, Xiaorui Sun, and Li-Yang Tan. A composition theorem for parity kill number. In *Computational Complexity Conference*, pages 144–154. IEEE Computer Society, 2014.
- 19 K. F. Roth. On certain sets of integers. *Journal of the London Mathematical Society*, s1-28(1):104–109, 1953. doi:10.1112/jlms/s1-28.1.104.
- 20 I. Ruzsa and E. Szemerédi. Triple systems with no six points carrying three triangles. *Combinatorica*, 18, January 1976.
- 21 Ronen Shaltiel. Dispersers for affine sources with sub-polynomial entropy. In *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, pages 247–256, 2011. doi:10.1109/FOCS.2011.37.
- 22 Asaf Shapira. *Graph Property Testing and Related Problems*. University of Tel-Aviv, 2006.
- 23 Endre Szemerédi. Regular partitions of graphs. Technical report, Stanford Univ Calif Dept of Computer Science, 1975.
- 24 Avishay Tal. Properties and applications of boolean function composition. In *ITCS*, pages 441–454. ACM, 2013.

A Section 1 Omissions

A.1 Omitted Proofs

In this section we provide the proofs of Proposition 4 and Proposition 5. We first begin with a corollary of Proposition 13 which will be useful in the analysis of the claims.

► **Corollary 33.** *If \mathcal{V} has dimension $n - 1$ and $\mathcal{V}^\perp = \text{span}(\{\gamma\})$, we have that $\widehat{f_{\alpha+\mathcal{V}}}(0) = \widehat{f}(0) \pm \widehat{f}(\gamma)$. Moreover, there exists a choice of α such that $|\widehat{f_{\alpha+\mathcal{V}}}(0)| = |\widehat{f}(0)| + |\widehat{f}(\gamma)|$.*

83:18 Searching for Regularity in Bounded Functions

Proof. By Proposition 13, we have

$$\widehat{f_{\alpha+\nu}}(0) = (-1)^{\langle 0, \alpha \rangle} \widehat{f}(0) + (-1)^{\langle \gamma, \alpha \rangle} \widehat{f}(\gamma) = \widehat{f}(0) + (-1)^{\langle \gamma, \alpha \rangle} \widehat{f}(\gamma).$$

This immediately implies both parts of the corollary. \blacktriangleleft

Proof of Proposition 4. Given some $f : \mathbb{F}_2^n \rightarrow [-1, 1]$, consider the following simple procedure:

- While at least δ fraction of $\pi \in \Pi$ have some γ_π such that $|\widehat{f_\pi}(\gamma_\pi)| > \delta$, further partition each π into $\pi \cap \{x : \langle \gamma_\pi, x \rangle = 0\}$ and $\pi \cap \{x : \langle \gamma_\pi, x \rangle = 1\}$.

We would like to show that we cannot perform the above partitioning action more than $\frac{1}{3\delta}$ times. Towards this end, define the potential function $\Phi(\Pi) := \mathbf{E}_{\pi \in \Pi} \widehat{f_\pi}(0)^2 = \mathbf{E}_{\pi \in \Pi} [(\mathbf{E} f_\pi)^2] \in [0, 1]$. Whenever we partition further, by Corollary 33 each $|\widehat{f_\pi}(0)|$ is updated to either $|\widehat{f_\pi}(0) + \widehat{f_\pi}(\gamma_\pi)|$ or $|\widehat{f_\pi}(0) - \widehat{f_\pi}(\gamma_\pi)|$. Therefore, the contribution of π to Φ in one step of the partitioning process is

$$\frac{1}{2} \left((\widehat{f_\pi}(0) + \widehat{f_\pi}(\gamma_\pi))^2 + (\widehat{f_\pi}(0) - \widehat{f_\pi}(\gamma_\pi))^2 \right) - \widehat{f_\pi}(0)^2 = \widehat{f_\pi}(\gamma_\pi)^2.$$

Since we assume at least δ fraction of $\pi \in \Pi$ had some γ_π such that $|\widehat{f_\pi}(\gamma_\pi)| > \delta$, at each step of the refinement Φ must increase by at least δ^3 , completing the proof. \blacktriangleleft

Proof of Proposition 5. Suppose without loss of generality, $\mathbf{E} f \geq 0$. Start with the trivial subspace, $\pi_0 = \mathbb{F}_2^n$. While there exists γ such that $|\widehat{f_{\pi_t}}(\gamma)| > \delta$, by Corollary 33 we can fix the parity corresponding to γ in such a way that ensures that $|\widehat{f_{\pi_{t+1}}}(0)| = \left| \widehat{f_{\pi_t}}(0) + |\widehat{f_{\pi_t}}(\gamma)| \right| > \widehat{f_{\pi_t}}(\gamma) + \delta$. Since $\widehat{f_\pi}(0) \leq 1$ for all π , this process can happen at most $\frac{1}{\delta}$ times. \blacktriangleleft

A.2 Omitted Sketches

We give the main ideas behind the lower bounds in Table 1.

Sketch of Lemma 26. The proof of this claim is based on the homogeneous degree-one function $f(x) = \frac{1}{n} \sum_i (-1)^{x_i}$. Its key idea comes from Observation 25, which we use to show that if the dimension of $\text{codim}(\mathcal{V}) < n/2$, then at least one shift of \mathcal{V}^\perp must contain *exactly* one standard basis vector. By the preceding discussion, this implies that $f_{\alpha+\nu}$ has a non-trivial Fourier coefficient with magnitude exactly $1/n > \delta$.

We remark that Lemma 26 is tight. The function f is symmetric, and for any such function, we can fix $n/2$ parities to obtain an affine subspace where every vector has weight $n/2$, which in turn fixes the function.

Sketch of Lemma 27 and Corollary 30. To achieve Lemma 27, one might expect to extend the above argument to the homogeneous degree d function $f(x) = \binom{n}{d}^{-1} \sum_{\gamma: \|\gamma\|_1=d} (-1)^{\langle \gamma, x \rangle}$. Unfortunately, this function is symmetric, and we have $r(f, 0) \leq n/2$. We therefore consider a random homogeneous degree d function $f_{\mathbf{z}}(x) = \binom{n}{d}^{-1} \sum_{\gamma: \|\gamma\|_1=d} \mathbf{z}_\gamma \cdot (-1)^{\langle \gamma, x \rangle}$, where each \mathbf{z}_γ is a random sign. A simple argument, again utilizing Observation 25, shows that there must be at least $\binom{k}{d}$ affine subspaces of \mathcal{V}^\perp with at least one vector of weight d . By our earlier reasoning, each of those subspaces must in fact contain at least two vectors of weight d so that the restricted function would have a non-trivial Fourier coefficient with magnitude $\binom{n}{d}^{-1} > \delta$. Moreover, the probability (over the signs \mathbf{z}_γ 's) that each of the $\binom{k}{d}$ signed sums cancels is at most $2^{-\binom{k}{d}}$, and a union bound over all the possible affine subspaces of dimension $k = \Theta(dn^{1/(d-1)})$ completes the argument.

If we restrict our attention to Boolean functions, we might hope to obtain strong upper bounds for $r(f, \delta)$; however, Corollary 30 rules this out. The proof of this claim is based on a simple lemma of [12] (Proposition 28), which uses the probabilistic method to convert a bounded function that is not δ -regular in large affine subspaces to a Boolean function with the same property. Applying this lemma to the lower bound from Lemma 27 achieves the result.

Sketch of Lemma 32. This lower bound is based on the majority function. Its key idea is that there exists a non-trivial affine subspace of \mathcal{V}^\perp containing *exactly* one weight-1 vector and relatively few vectors of higher weight (see the full version of the work for details on this). Then, we use properties of the Fourier spectrum of the majority function to show that the signed sum of the Fourier coefficients of majority corresponding to vectors in this affine subspace, is on the order of $|\widehat{f}(e_1)| = \Omega(n^{-1/2})$. Specifically, we argue that even if the coefficients coming from higher weight vectors in the aforementioned sum combined in the most constructive way possible, they cannot combine to more than $|\widehat{f}(e_1)|/2$. We also note that Lemma 32 is tight up to constant factors via Proposition 5. Conversely, Lemma 32 implies that for $\delta \geq n^{-1/2}$, the majority function on $O(1/\delta^2)$ variables is an explicit Boolean function for which $r(f, \delta) \geq \Omega(1/\delta)$.

Rationale for Lemma 31. The last entry in the table corresponds to Lemma 31 and is based on a simple function f on 4 inputs that is composed with itself k times. We use key properties of the composition of Boolean functions (from [24, 18]) to achieve the bound. The function itself is the same one considered in [18], and we use their main theorem crucially to obtain our lower bound. We present a slightly generalized version of the main theorem of [18], so we include a proof this in the full version of this work.

We make some final comments about the lower bounds from Corollary 30. The Boolean functions that achieve the lower bounds share the property that the magnitudes of their Fourier coefficients are extremely close to their bounded counterparts in Lemma 27. However, even though the bounded functions themselves have low degree, the Boolean functions are very far from being low-degree functions; in fact, *almost all* their Fourier mass comes from the high-degree terms. Notably, these functions are also non-explicit affine dispersers with small dimension, and it would be interesting to find explicit Boolean functions with similar strong lower bounds on the $r(f, \delta)$.

B Section 2 Omitted Proofs

Proof of Proposition 13. Using Fact 12, we can write

$$\begin{aligned} \widehat{fu}(\gamma) &= \mathbf{E}_{x \in \mathcal{V}} [f(x + \alpha) \cdot (-1)^{\langle \gamma, x \rangle}] = \mathbf{E}_{x \in \mathcal{V}} \sum_{\beta} \widehat{f}(\beta) (-1)^{\langle \beta, x + \alpha \rangle} (-1)^{\langle \gamma, x \rangle} \\ &= \sum_{\beta} \widehat{f}(\beta) (-1)^{\langle \beta, \alpha \rangle} \mathbf{E}_{x \in \mathcal{V}} [(-1)^{\langle \beta + \gamma, x \rangle}] \\ &= \sum_{\beta \in \gamma + \mathcal{V}^\perp} \widehat{f}(\beta) (-1)^{\langle \beta, \alpha \rangle}, \end{aligned}$$

where the last equality follows by observing that $\mathbf{E}_{x \in \mathcal{V}} [(-1)^{\langle \gamma + \beta, x \rangle}] = 1$ if $\beta \in \gamma + \mathcal{V}^\perp$, and zero otherwise. ◀

83:20 Searching for Regularity in Bounded Functions

Proof of Proposition 15. We first show that \mathcal{W} and \mathcal{V}^\perp are independent. Suppose that $M^\top \gamma + u = 0$, where $\gamma \in \text{span}(J)$ and $u \in \mathcal{V}^\perp$. For any $v \in \mathcal{V}$ such that $v \neq 0$, we have

$$0 = \langle v, M^\top \gamma + u \rangle = \langle v, M^\top \gamma \rangle = \langle Mv, \gamma \rangle,$$

which is impossible unless $\gamma = 0$ since this implies $Mv \in \text{span}(J)^\perp = \text{span}(\bar{J})$ and $Mv \neq 0$. This in turn implies that $u = 0$ and therefore that \mathcal{W} and \mathcal{V}^\perp are independent. The claim follows by noting that $\dim(\mathcal{W} \oplus \mathcal{V}^\perp) = \dim(\mathcal{W}) + \dim(\mathcal{V}^\perp) = k + n - k = n$. ◀

Proof of Proposition 16. Repeatedly using Fact 12, we have that

$$\begin{aligned} |\widehat{f}_u(M^\top \gamma)| &= \left| \mathbf{E}_{x \in \mathcal{U}} \left[f(x) (-1)^{\langle M^\top \gamma, x \rangle} \right] \right| = \left| \mathbf{E}_{x \in \mathcal{U}} \left[f(x) (-1)^{\langle \gamma, Mx \rangle} \right] \right| \\ &= \left| \mathbf{E}_{z \in \mathcal{U}'} \left[f(M^{-1}z) (-1)^{\langle \gamma, z \rangle} \right] \right| = |\widehat{g}_{u'}(\gamma)|. \quad \blacktriangleleft \end{aligned}$$

Proof of Fact 18. We have that

$$\begin{aligned} \widehat{g}(\gamma) &= \mathbf{E}_x [g(x) \chi_\gamma(x)] = \mathbf{E} [f(Mx) \chi_\gamma(x)] \\ &= \mathbf{E}_y [f(y) \chi_\gamma(M^{-1}y)] \\ &= \mathbf{E}_y [f(y) \chi_{M^{-\top} \gamma}(y)] = \widehat{f}(M^{-\top} \gamma), \end{aligned}$$

where we have used the fact that $\chi_\gamma(M^{-1}y) = (-1)^{\langle \gamma, M^{-1}y \rangle} = (-1)^{\langle M^{-\top} \gamma, y \rangle}$. ◀

C Proofs from Section 5

C.1 Proofs from Section 5.1

Proof of Observation 25. Let $S = \{u : u \in \mathcal{W} \text{ and } |(u + \mathcal{V}^\perp)^{\neq 1}| \geq 1\}$. Since every standard basis vector can be expressed as $u + v$ for some $u \in S$ and $v \in \mathcal{V}^\perp$, we have that $\dim(\text{span}(S \cup \mathcal{V}^\perp)) = n$. However, we also know that $\dim(\text{span}(S \cup \mathcal{V}^\perp)) \leq |S| + C$, and rearranging we get $|S| \geq n - C$. Next, let $S_1 = \{u \in S : |(u + \mathcal{V}^\perp)^{\neq 1}| = 1\}$. By Fact 9, for any $u, u' \in S$, we have $u + \mathcal{V}^\perp \neq u' + \mathcal{V}^\perp$. Therefore,

$$n = \sum_{u \in S} |(u + \mathcal{V}^\perp)^{\neq 1}| = \sum_{u \in S_1} |(u + \mathcal{V}^\perp)^{\neq 1}| + \sum_{u \in S \setminus S_1} |(u + \mathcal{V}^\perp)^{\neq 1}| \geq |S_1| + 2(|S| - |S_1|),$$

and rearranging, we get $|S_1| \geq 2|S| - n \geq n - 2C$. ◀