Electronic Colloquium on Computational Complexity, Report No. 109 (2022)

# Searching for Regularity in Bounded Functions

Siddharth Iyer
siyer@cs.washington.edu

Michael Whitmeyer*
mdwhit@cs.washington.edu

July 27, 2022

**Abstract**

Given a function $f : \mathbb{F}_2^n \to [-1, 1]$, this work seeks to find a large affine subspace $\mathcal{U}$ such that $f$, when restricted to $\mathcal{U}$, has small nontrivial Fourier coefficients.

We show that for any function $f : \mathbb{F}_2^n \to [-1, 1]$ with Fourier degree $d$, there exists an affine subspace of dimension at least $\widetilde{\Omega}(n^{1/d!}k^{-2})$, 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/extractors.

## 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's[1] theorem [Rot53] on 3-term arithmetic progressions, which essentially says that *every* subset of $\{1, \ldots, n\}$ of density $\delta > 0$ must contain a 3-term arithmetic progression.

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 $V_1, \ldots, V_k$, wherein most pairs of parts $(V_i, V_j)$ are $\delta$-regular. In this setting, the $\delta$-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 in both pure mathematics (e.g., Szemerédi's [Sze75] genealization of Roth's result to $k$-term arithmetic progressions) and theoretical computer science (to test triangle freeness in dense graphs [RS76; Alo+01; Sha06]).

Inspired by Szemerèdi's regularity lemma, Green [Gre05] gave an analogous regularity lemma for bounded functions of the form $f : \mathbb{F}_2^n \to [0, 1]$. In order to state their result, we briefly recall Fourier analysis over subspaces of $\mathbb{F}_2^n$. Let $\mathcal{V}$ be a subspace of $\mathbb{F}_2^n$ and $\mathcal{U} = \alpha + \mathcal{V}$ be an affine shift of $\mathcal{V}$. For a function $g : \mathcal{U} \to \mathbb{R}$, we can write

$$g(x) = \sum_{\chi} \widehat{g}(\chi) \cdot \chi(x + \alpha),$$

where the sum ranges over all linear maps $\chi : \mathcal{V} \to \{\pm 1\}$, and $\widehat{g}(\chi) := \mathbf{E}_{x \sim \mathcal{U}}[g(x)\chi(x + \alpha)]$[2] is the Fourier coefficient associated with $\chi$ . We refer to the trivial linear map by $\chi_{\mathbf{0}}$, which is simply given by $\chi_{\mathbf{0}}(x) = 1$ for all $x \in \mathcal{V}$. We also need a definition of $\delta$-regularity for functions over $\mathbb{F}_2^n$.

*Research supported by NSF grant CCF-2006359.

[1] The related Hales-Jewett theorem [HJ63] is also a classic result in Ramsey theory.

[2] When $\mathcal{V} = \mathbb{F}_2^n$, we typically associate each $\gamma \in \mathbb{F}_2^n$ with $\chi_\gamma = (-1)^{\langle \gamma, x \rangle}$, as in O'Donnell [ODo21]. However, there is no canonical mapping between vectors and characters when $\mathcal{V} \neq \mathbb{F}_2^n$. For more details on Fourier analysis on subspaces, see Section 2.



ISSN 1433-8092

**Definition 1.1** ($\delta$-regularity[3])**.** *Let $\mathcal{U}$ be an affine subspace of $\mathbb{F}_2^n$ and $g : \mathcal{U} \to \mathbb{R}$. For $\delta \geq 0$, we say $g$ is $\delta$-regular if $\max_{\chi \neq \chi_0} |\widehat{g}(\chi)| \leq \delta$.*

Given a function $f : \mathbb{F}_2^n \to \mathbb{R}$, in the following theorem and throughout this paper, we denote the restriction of $f$ to an affine subspace $\mathcal{U}$ by the function $f_{\mathcal{U}} : \mathcal{U} \to \mathbb{R}$. We now state Green's result; below, the notation $\mathrm{twr}(x)$ refers to an exponential tower of 2's $2^{2^{\cdot^{\cdot^{\cdot}}}}$ of height $x$.

**Theorem 1.2** (Theorem 2.1 in [Gre05][4])**.** *For any $f : \mathbb{F}_2^n \to [0,1]$ and $\delta > 0$, there exists a subspace $\mathcal{V}$ of co-dimension $M(\delta) \leq \mathrm{twr}(\lceil 1/\delta^3 \rceil)$ such that for all but a $\delta$-fraction of the affine subspaces $\mathcal{U} = \alpha + \mathcal{V}$, $f_{\mathcal{U}}$ is $\delta$-regular.*

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

In the above upper and lower bound of [Gre05; Hos+16], 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 might ask if there is a better partition of $\mathbb{F}_2^n$ into affine subspaces of smaller co-dimension so that in most parts $f$ is $\delta$-regular. This is indeed the case.

**Proposition 1.3** (Proposition A.1 in [Gir+21][5])**.** *For any $f : \mathbb{F}_2^n \to [0,1]$ and $\delta > 0$, there exists a partition $\Pi$ 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 $\delta$-fraction of the parts, $f_\pi$ is $\delta$-regular.*

In this work, rather than searching for a partition $\Pi$ such that $f_\pi$ is $\delta$-regular for most $\pi \in \Pi$, we instead look for a single affine subspace $\mathcal{U}$ such that $f_{\mathcal{U}}$ is $\delta$-regular. Namely, our main focus is to understand the quantity

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

A simple upper bound on $\mathsf{r}(f, \delta)$ for functions bounded in the interval $[-1, 1]$ is based on the following folklore claim.

**Proposition 1.4** (Folklore)**.** *For any $f : \mathbb{F}_2^n \to [-1, 1]$, we have $\mathsf{r}(f, \delta) \leq \frac{1}{\delta}$.*

The proofs of Proposition 1.3 and Proposition 1.4 are based on simple algorithms that greedily fix the parities corresponding to the largest Fourier coefficients; they are included in Appendix B for completeness.

## 1.1 Related Work

To the best of our knowledge, $\mathsf{r}(f, \delta)$ has not been explicitly studied in previous work; however, it is connected to several concepts in theoretical computer science. The quantity $\mathsf{r}(f, 0)$ can be interpreted as the minimum number of parities one must fix in order to make $f$ constant, and it has been studied under the name of *parity kill number* (see [ODo+14]).[6] 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 $f$ a constant. In particular, for any $\delta \geq 0$, we have $\mathsf{r}(f, \delta) \leq \mathsf{r}(f, 0) \leq C_{\min}[f]$. The minimum certificate complexity is one of several natural complexity measures that have been well studied for Boolean functions $f : \mathbb{F}_2^n \to \{\pm 1\}$ (see [Bd02; Ben17] for surveys). In particular, for the class of Boolean functions whose Fourier degree is $d$, meaning the only non-zero Fourier coefficients of $f$ correspond to vectors of weight at most $d$, it is known that $C_{\min}[f] \leq 2d^3$ [Mid04].[7] This immediately gives the following bound on $\mathsf{r}(f, \delta)$ for such functions.

[3] It is easy to check that a random $f : \mathbb{F}_2^n \to \{0, 1\}$ will be $2^{-\Omega(n)}$-regular with high probability (see [ODo21], Exercise 1.7 and Proposition 6.1). In this sense, being $\delta$-regular is a type of pseudorandomness condition, analogous to the definition of $\delta$-regularity for graphs in Szemerédi's regularity lemma.

[4] More precisely, an analogue of Theorem 1.2 for $\{0, 1\}$-valued functions is what actually appears in [Gre05]. However, the proof in [Gre05] follows just as well for functions taking values in $[0, 1]$.

[5] This is a simpler version of Proposition A.1 in [Gir+21], where we do not account for multiple functions or condition on any event.

[6] The parity kill number is often denoted $C^{\oplus}_{\min}[f]$ in other works.

[7] Midrijanis showed that $D(f) \leq 2d^3$, but it is obvious that $C_{\min}[f] \leq D(f)$.

**Lemma 1.5.** *For any degree $d$ function $f : \mathbb{F}_2^n :\to \{\pm 1\}$ and $\delta \geq 0$, we have $\mathsf{r}(f, \delta) \leq 2d^3$.*

The quantity $\mathsf{r}(f, 0)$ is also closely related to the dimension of affine dispersers. An affine disperser of dimension $k$ is a coloring of $\mathbb{F}_2^n$ such that no affine subspace of dimension at least $k$ is monochromatic. The search for explicit affine dispersers has received substantial attention in the literature (see e.g. [Sha11; Li16; CGL21; CT15]). If we consider an affine disperser as a function $f : \mathbb{F}_2^n \to \{0, 1, \ldots, C\}$,[8] then as observed by [ODo+14], the dimension of the disperser is simply $n - \mathsf{r}(f, 0) - 1$. In the case of $C = 1$, Cohen and Tal [CT15] rule out $\mathbb{F}_2$-polynomials of degree $d$ as affine dispersers by showing that any such function satisfies $\mathsf{r}(f, 0) \leq n - \Omega(d \cdot n^{1/(d-1)})$.

Affine dispersers are a relaxed variant of affine extractors, which, when viewed as functions, are required to be nearly uniform (rather than nonconstant) on all large affine subspaces (see Definition 5.4 for a formal definition). We observe that if $f$ is a $(k, \delta)$-extractor, then *every* affine subspace $\mathcal{U}$ of dimension at least $k + 1$ is such that $f_{\mathcal{U}}$ is $2C\delta$-regular (see Claim 5.5). We describe our results which give upper and lower bounds on $\mathsf{r}(f, \delta)$.

## 1.2 Our Results

For a function $f : \mathbb{F}_2^n \to [-1, 1]$, this paper asks if we can obtain upper bounds on $\mathsf{r}(f, \delta)$ that are better than Proposition 1.4 or Lemma 1.5. Note that Proposition 1.4 becomes trivial when $\delta < 1/n$, so one might ask if a better dependence of $\mathsf{r}(f, \delta)$ on $\delta$ can be achieved. We focus on the case where the output of $f$ is either bounded in the interval $[-1, 1]$ or takes one of two values $\{-1, 1\}$. Our main result is an upper bound on $\mathsf{r}(f, \delta)$ for functions with Fourier degree at most $d$.

**Theorem 1** (Upper bound on $\mathsf{r}(f, \delta)$ for degree $d$ bounded functions)**.** *For any $\delta \in (0, 1)$ and any degree $d$ function $f : \mathbb{F}_2^n \to [-1, 1]$, we have $\mathsf{r}(f, \delta) \leq n - \Omega\left(n^{1/d!}(\log(n/\delta))^{-2}\right)$.*

For $\delta \ll 1/n$ (where Proposition 1.4 is no longer meaningful), Theorem 1 is equivalent to stating that there is an affine subspace $\mathcal{U}$ of dimension $\widetilde{\Omega}\left(n^{1/d!}(\log 1/\delta)^{-2}\right)$ such that $f_{\mathcal{U}}$ is $\delta$-regular.

Using Theorem 1 and the connection between the dimension of affine dispersers and $\mathsf{r}(f, 0)$, we show that low Fourier degree functions make poor affine dispersers.

**Theorem 1.6.** *If $f : \mathbb{F}_2^n \to \{0, \ldots, C\}$ has Fourier degree $d$, then $\mathsf{r}(f, 0) = n - \Omega\left(n^{1/d!}(d + \log(nC))^{-2}\right)$.*

The above result resembles that of [CT15]; 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 [CT15] applies to functions whose range is $\mathbb{F}_2$, while ours applies to functions that take values in the set $\{0, \ldots, C\}$, which can have a much larger size. Furthermore, for $f : \mathbb{F}_2^n \to \{0, \ldots, C\}$, an argument similar to the one in Lemma 1.5 shows that $\mathsf{r}(f, 0) \leq O(Cd^3)$, where $d$ here is the Fourier degree. However, this does not address the case where $C = \Omega(n)$, which is when Theorem 1.6 becomes useful.

To complement these results, we present several examples of functions (bounded, Boolean, explicit, and existential) for which $\mathsf{r}(f, \delta)$ is large.

| $\delta$ | $\mathsf{r}(f, \delta) \geq$ | Explicit? | Boolean/bounded | Ref. |
|---|---|---|---|---|
| $1/n$ | $n/2 - 1$ | Yes | bounded | Lemma 4.2 |
| $\binom{n}{d}^{-1}$ | $n - 2dn^{1/(d-1)}$ | No | bounded | Lemma 4.3 |
| $\Theta(n^{-1/2})$ | $\Theta(\sqrt{n})$ | Yes | Boolean | Lemma 4.11 |
| $\frac{1}{2} \cdot n^{-d}$ (for $d \leq \frac{\log n}{\log\log n + 1}$) | $n - 2dn^{1/(d-1)}$ | No | Boolean | Corollary 4.7 |
| $1/2^{2^k+1}$ (for integer $k$) | $\Omega\left((\log \frac{1}{\delta})^{\log_2(3)}\right)$ | Yes | Boolean | Lemma 4.8 |

Table 1: Table of functions with large $\mathsf{r}(f, \delta)$ values.

We next give a brief overview of the techniques used to prove our results.

[8] Affine dispersers can also be defined as requiring that $\mathrm{supp}(f_{\mathcal{U}}) \geq (1 - \varepsilon)C$ for some parameter $\varepsilon$. Or, the output space can alternatively be defined as $\{0, 1\}^m$ for some $m$ (e.g. [Sha11; Li16; CGL21]) or $\mathbb{F}_q$ for some (prime power) $q$ (e.g. [CT15]).

## 1.3 Techniques

### 1.3.1 Upper bound on $\mathsf{r}(f, \delta)$

We give a brief proof sketch of Theorem 1; the central idea here is an application of the pigeonhole principle. The proof proceeds by induction over the Fourier degree. The base case corresponds to degree one functions, and 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 $g_S - g_{S'} = \sum_{i \in S \triangle S'} \widehat{f}(e_i)(-1)^{|\{i\} \cap S'|}$.

We now use the set $S \triangle S'$ 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 $S \triangle S' = [t']$ 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 $i \in \{2, \ldots, t'\}$, 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 $e_{t'+1}, \ldots, e_n$ 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 $\delta$. Repeatedly applying this argument roughly $\frac{n}{\log(1/\delta)}$ times for the remaining standard basis vectors, we obtain an affine subspace of dimension at least $\Omega\left(\frac{n}{\log(1/\delta)}\right)$.

At a high level, we reduce the problem for degree $d$ functions to degree $d-1$ by restricting to an affine subspace of dimension $\Omega\left(\left(\frac{n}{\log(n/\delta)}\right)^{1/d}\right)$, 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 3.1. 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 3.1 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.

### 1.3.2 Lower bounds on $\mathsf{r}(f, \delta)$

With the exception of the last entry in Table 1 (Lemma 4.8), all lower bounds share the same high-level template and are based on variants of a linear algebraic argument (see Claim 4.1 and Lemma 4.14) that gives upper bounds on the number of low weight vectors in low dimensional affine subspaces. We sketch this high level template and then explain why such upper bounds imply lower bounds on $\mathsf{r}(f, \delta)$.

To start, we analyze the Fourier spectrum of a carefully chosen function $f$ when restricted to an arbitrary affine subspace. Crucially, if we restrict $f$ to the affine subspace $\alpha + \mathcal{V}$, the resulting Fourier coefficients of $f_{\alpha+\mathcal{V}}$ are simply signed sums of the Fourier coefficients of $f$ corresponding to vectors in a shift of $\mathcal{V}^\perp$ (see Fact 2.7). Moreover, if $f$ has several large ($\gg \delta$) Fourier coefficients but $f_{\alpha+\mathcal{V}}$ is $\delta$-regular, then any signed sum involving a large coefficient must also involve other nonzero coefficients in order for the sum to evaluate to at most $\delta$.

Furthermore, we choose functions with the property that their large Fourier coefficients all correspond to low weight vectors. By the above discussion, if $f_{\alpha+\mathcal{V}}$ is $\delta$-regular, then many affine shifts of $\mathcal{V}^\perp$ must contain multiple low weight vectors. However, if $\dim(\mathcal{V}^\perp) = \operatorname{codim}(\mathcal{V})$ is small, we can show by a simple

linear algebraic argument that some affine shift of $\mathcal{V}^\perp$ must contain few low weight vectors, which in turn prevents some Fourier coefficient of $f_{\alpha+\mathcal{V}}$ from being small. This gives a contradiction, and we conclude that $\operatorname{codim}(\mathcal{V})$ cannot be too large.

We now describe the functions used in each lower bound claim and how the linear algebraic argument applies to them.

**Sketch of Lemma 4.2.** 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 Claim 4.1, which we use to show that if the dimension of $\operatorname{codim}(\mathcal{V}) < n/2$, then at least one shift of $\mathcal{V}^\perp$ must contain *exactly* one standard basis vector. By the preceding discussion, this implies that $f_{\alpha+\mathcal{V}}$ has a non-trivial Fourier coefficient with magnitude exactly $1/n > \delta$.

We remark that Lemma 4.2 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 4.3 and Corollary 4.7.** To achieve Lemma 4.3, 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 $r(f,0) \le n/2$. 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 $\mathbf{z}_\gamma$ is a random sign. A simple argument, again utilizing Claim 4.1, 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 $\mathbf{z}_\gamma$'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 $\mathsf{r}(f,\delta)$; however, Corollary 4.7 rules this out. The proof of this claim is based on a simple lemma of [Hos+16] (Lemma 4.5), which uses the probabilistic method to convert a bounded function that is not $\delta$-regular in large affine subspaces to a Boolean function with the same property. Applying this lemma to the lower bound from Lemma 4.3 achieves the result.

**Sketch of Lemma 4.11.** 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 Lemma 4.14). Then, we use properties of the Fourier spectrum of the majority function (see Claim 4.12) to show that the signed sum of the Fourier coefficients of majority corresponding to vectors in this affine subspace, is on the order of $\left|\widehat{f}(e_1)\right| = \Omega(n^{-1/2})$. 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 $\left|\widehat{f}(e_1)\right|/2$. We also note that Lemma 4.11 is tight up to constant factors via Proposition 1.4. Conversely, Lemma 4.11 implies that for $\delta \ge n^{-1/2}$, the majority function on $O(1/\delta^2)$ variables is an explicit Boolean function for which $\mathsf{r}(f,\delta) \ge \Omega(1/\delta)$.

**Rationale for Lemma 4.8.** The last entry in the table corresponds to Lemma 4.8 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 [Tal13; ODo+14]) to achieve the bound. The function itself is the same one considered in [ODo+14], and we use their main theorem (Theorem 4.9) crucially to obtain our lower bound. We present a slightly generalized version of the main theorem of [ODo+14], so we include a proof of Theorem 4.9 in Appendix A.

We make some final comments about the lower bounds from Corollary 4.7. 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 4.3. 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 $\mathsf{r}(f, \delta)$.

## 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 $\mathsf{span}(J)$ to denote the subspace spanned by the standard basis vectors corresponding to the elements in $J$. We refer to the $\mathrm{L}_1$ 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 $f : \mathbb{F}_2^n \to \mathbb{R}$ to be $\max\{\|\gamma\|_1 : \widehat{f}(\gamma) \neq 0\}$. We frequently interpret a linear transformation $M : \mathbb{F}_2^n \to \mathbb{F}_2^n$ as a matrix and refer to the linear map obtained by taking the transpose of the matrix as $M^{\mathsf{T}}$. At several points, we consider the compositions of functions with linear maps. For a function $f$ and a map $M : \mathbb{F}_2^n \to \mathbb{F}_2^n$, 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 2.1** (Hoeffding, [Hoe63])**.** *Suppose $X_1, \ldots, X_n$ are such that $a \leq X_i \leq b$ for all $i$. Let $M = \frac{X_1 + \ldots X_n}{n}$. Then,*

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

**Definition 2.2** (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*

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

**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 $\mathrm{codim}(\mathcal{A}) = n - \dim(\mathcal{A})$.

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

**Definition 2.3** (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 $\mathcal{A} \oplus \mathcal{B} = \mathbb{F}_2^n$.*[9]

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,[10] and often in fact must not be.

**Fact 2.4.** *Let $\mathcal{A}, \mathcal{B}$ be subspaces of $\mathbb{F}_2^n$ such that $\mathcal{A} \oplus \mathcal{B} = \mathbb{F}_2^n$. 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 $\mathcal{A} \oplus \mathcal{B} = \mathbb{F}_2^n$. □

[9] 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 it.

[10] this 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 $\mathbb{F}_2^n$.

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

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

where $\chi_\gamma(x) = (-1)^{\langle \gamma, x \rangle}$ and $\widehat{f}(\chi_\gamma) = \mathbf{E}_x[f(x)\chi_\gamma(x)]$. For shorthand, we refer to $\widehat{f}(\chi_\gamma)$ simply as $\widehat{f}(\gamma)$. We denote $f^{=d}(x) := \sum_{\|\gamma\|_1 = d} \widehat{f}(\gamma)\chi_\gamma(x)$ as the degree $d$ part of $f$. For more on this topic, see [ODo21], 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 $f_{J \leftarrow b} : \mathsf{span}(\overline{J}) \to \mathbb{R}$, 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 $\{\chi_\gamma(x) := (-1)^{\langle \gamma, x \rangle} : \gamma \in \mathsf{span}(\overline{J})\}$ is the Fourier basis of the restricted function.

**Fact 2.5** (Fourier Coefficients of Restricted Functions (see [ODo21], Proposition 3.21))**.** *For every $\gamma \in \mathsf{span}(\overline{J})$ and $b \in \mathsf{span}(J)$,*

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

## 2.1 Fourier Analysis on Subspaces

We move to the general setting of restricting functions to arbitrary affine subspaces. 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 $f_\mathcal{U} : \mathcal{V} \to \mathbb{R}$ 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 $\chi_\gamma : \mathcal{V} \to \{\pm 1\}$ 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 $\chi_\gamma$, and for $\mathcal{U} = \alpha + \mathcal{V}$, we can write

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

In this section, we present three separate formulas for the Fourier coefficients of $f_\mathcal{U}$, each of which is useful in different contexts.

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 $\chi_\gamma$ we have defined.

**Fact 2.6.** *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*

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

Fact 2.6 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 $\alpha$ 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 2.6.

**Fact 2.7.** *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*

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

*Proof.* Using Fact 2.6, we can write

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

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

We note that Fact 2.7 gives a formula analogous to Fact 2.5 for restrictions to general affine subspaces. Fact 2.5 will be useful to construct functions and argue that they never become $\delta$-regular when restricted to any sufficiently large subspace. Before we give our final formula, we highlight one particular choice of $\mathcal{W}$ such that $\mathcal{W}\oplus\mathcal{V}^{\perp} = \mathbb{F}_2^n$.

**Definition 2.8** ($M$ mapping $\mathcal{V}$ to $\mathsf{span}(J)$)**.** *Given a $k$-dimensional subspace $\mathcal{V}$, let $B = \{\beta_1,\dots,\beta_n\}$ be a basis for $\mathbb{F}_2^n$ such that $\mathcal{V} = \mathsf{span}(\{\beta_1,\dots,\beta_k\}$. For any subset $J\subseteq[n]$ of size $k$, let $M:\mathbb{F}_2^n\to\mathbb{F}_2^n$ be an invertible linear map such that $\{M\beta_i : i\in[k]\} = \{e_j : j\in J\}$.*

**Fact 2.9** (Choice of $\mathcal{W}$)**.** *Let $\mathcal{V}$, $M$ and $J$ be defined as in Definition 2.8. The subspaces $\mathcal{W} = \{M^{\mathsf{T}}\gamma : \gamma\in\mathsf{span}(J)\}$ and $\mathcal{V}^{\perp}$ are independent, and $\mathcal{W}\oplus\mathcal{V}^{\perp} = \mathbb{F}_2^n$.*

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

$$0 = \langle v, M^{\mathsf{T}}\gamma + u\rangle = \langle v, M^{\mathsf{T}}\gamma\rangle = \langle Mv,\gamma\rangle,$$

which is impossible unless $\gamma = 0$ since this implies $Mv\in\mathsf{span}(J)^{\perp} = \mathsf{span}(\overline{J})$ 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 $\dim(\mathcal{W}\oplus\mathcal{V}^{\perp}) = \dim(\mathcal{W}) + \dim(\mathcal{V}^{\perp}) = k + n - k = n$. □

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 $M$.

**Fact 2.10.** *Let $\mathcal{V}$, $M$ and $J$ be defined as in Definition 2.8 and $\mathcal{U} = \alpha+\mathcal{V}$. For any $\gamma\in\mathsf{span}(J)$, we have,*

$$\left|\widehat{f_{\mathcal{U}}}(M^{\mathsf{T}}\gamma)\right| = \left|\widehat{h_{\mathcal{U}'}}(\gamma)\right|,$$

*where $h = f\circ M^{-1}$ and $\mathcal{U}' = \{Mu : u\in\mathcal{U}\} = M\alpha + \mathsf{span}(J)$ is a standard restriction.*

*Proof.* Repeatedly using Fact 2.6, we have that

$$\left|\widehat{f_{\mathcal{U}}}(M^{\mathsf{T}}\gamma)\right| = \left|\mathop{\mathbf{E}}_{x\in\mathcal{U}}\left[f(x)(-1)^{\langle M^{\mathsf{T}}\gamma,x\rangle}\right]\right| = \left|\mathop{\mathbf{E}}_{x\in\mathcal{U}}\left[f(x)(-1)^{\langle\gamma,Mx\rangle}\right]\right| = \left|\mathop{\mathbf{E}}_{z\in\mathcal{U}'}\left[f(M^{-1}z)(-1)^{\langle\gamma,z\rangle}\right]\right| = \left|\widehat{g_{\mathcal{U}'}}(\gamma)\right|.$$

□

Fact 2.10 implies the following important corollary.

**Corollary 2.11.** *There exists an affine subspace $\mathcal{U}$ of dimension $k$ such that $f_{\mathcal{U}}$ is $\delta$-regular if and only if there exists an invertible linear map $M:\mathbb{F}_2^n\to\mathbb{F}_2^n$, 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 $\delta$-regular, where $h = f\circ M$.*

We use Corollary 2.11 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 2.12** ([ODo21], Exercise 3.1)**.** *Let $M$ be an invertible linear transformation, and consider the function $g = f\circ M^{-1}:\mathbb{F}_2^n\to\mathbb{R}$. Then we have that $\widehat{g}(\gamma) = \widehat{f}(M^{\mathsf{T}}\gamma)$.*

## 3 Proof of Theorem 1

The theorem is proved via the following lemma, which allows us to carry out our induction.

**Lemma 3.1.** *For $\tau \in (0,1)$ and any degree $d$ function $f : \mathbb{F}_2^n \to [-1,1]$, there exists an invertible linear map $M : \mathbb{F}_2^n \to \mathbb{F}_2^n$, a set $J \subseteq [n]$ with size at least $\frac{d}{4e}\left(\frac{n}{\log 5/\tau}\right)^{1/d}$, and $b \in \mathsf{span}(\overline{J})$ such that $h = f \circ M$ satisfies*

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

We now prove Theorem 1 using Lemma 3.1.

**Theorem 3.2** (Theorem 1 restated)**.** *For any $\delta \in (0,1)$ and any degree $d$ function $f : \mathbb{F}_2^n \to [-1,1]$, we have $\mathsf{r}(f,\delta) \le n - \Omega\left(n^{1/d!}(\log(n/\delta))^{-2}\right)$.*

*Proof.* We show, by induction over the degree, that there exists an invertible linear map $M$, a set $I \subseteq [n]$, and $b \in \mathsf{span}(\overline{I})$ such that for $\delta > 0$ and any degree $d$ function $f$,

1. $h_{I\leftarrow b}$ is $\delta$-regular, where $h = f \circ M$, and
2. for $C_d = \sum_{i=1}^{d}(i!)^{-1}$, we have
$$|I| \ge \frac{n^{1/d!}}{(8e)^{C_{d-1}}\left(\log(n/\delta)\right)^{C_d}}.$$

Note that $C_d \le e - 1 < 2$ for all $d \ge 1$. The existence of the desired affine subspace is then given by Corollary 2.11, and its dimension is equal to $|I| \ge \Omega\left(n^{1/d!}\left(\log(n/\delta)\right)^{-2}\right)$.

The base case corresponds to the degree being one. Let us apply Lemma 3.1 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 $\frac{n}{\log 5/\delta} \ge \Omega\left(\frac{n}{\log n/\delta}\right)$, and $b \in \mathsf{span}(\overline{J})$ such that

$$|\widehat{g_{\overline{J}\leftarrow b}}(\gamma)| \le \begin{cases} \tau & \text{if } \|\gamma\|_1 = 1, \\ 0 & \text{for all } \|\gamma\|_1 > 1. \end{cases} \implies |\widehat{g_{\overline{J}\leftarrow b}}(\gamma)| \le \delta, \text{ for all } \gamma \ne 0.$$

Assuming both items hold for some degree $d-1$, we show them for degree $d$. Applying Lemma 3.1 with degree $d$ and $\tau = n^{-d}\delta/3$, we denote $p := (f \circ M)_{\overline{J}\leftarrow b}$, where $M$, $J$ and $b$ are as promised by the lemma. Note that, by Lemma 3.1, $p$ has degree at most $d$, and for any $\gamma$ with $\|\gamma\|_1 = d$, we have, $|\widehat{p}(\gamma)| \le \delta/(3n^d)$. 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 $x$,

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

Applying the inductive hypothesis[11] 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 $b' \in \mathsf{span}(J \setminus I)$ 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 $\gamma \ne 0$, we have $\left|\widehat{q_{J\setminus I\leftarrow b'}}(\gamma)\right| \le \left(1+\frac{\delta}{3}\right)\frac{\delta}{3} < \frac{2\delta}{3}$. Denoting $p' := p \circ M'$ and $r := p^{=d} \circ M'$, we have for any $\gamma \ne 0$ that

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

[11] Technically, $\frac{p^{<d}}{1+\delta/3} : \mathsf{span}(J) \to [-1,1]$. 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.

This shows that ${p'}_{J\setminus I\leftarrow b'}$ is $\delta$-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 3.1 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

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

This shows item 2 of the inductive hypothesis as desired. □

To prove Lemma 3.1, we need the following claim, which ultimately lets us bound Fourier coefficients in certain affine subspaces.

**Claim 3.3** (Pigeonhole Principle)**.** *Let $f:\mathbb{F}_2^n \to [-1,1]$ 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 $z\in\{\pm1\}^S$ such that*

1. *$\forall \gamma \in \mathsf{span}(K)$ 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

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

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

Now, divide the interval $[-1,1]$ into $2/\tau$ intervals of length $\tau$. For a fixed $U\subseteq T$ of even size, consider putting the values of $a_U(\gamma)$ for all $\gamma\in\mathsf{span}(K)^{=d-1}$ 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

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

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

We can now prove Lemma 3.1.

*Proof of Lemma 3.1.* 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 $\mathsf{span}(K)$ has magnitude at most $\tau$. We build $K$ one coordinate at a time by repeatedly invoking Claim 3.3 and arguing that the quantities guaranteed to be small by Claim 3.3 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 $d$. 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

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

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 3.3 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

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

We can also assume that $d+i \in S$ and $z_{d+i} = 1$. Now consider the invertible linear transformation $M_i : \mathbb{F}_2^n \to \mathbb{F}_2^n$ 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 $b_i \in \mathsf{span}(J_i)$, 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 $j \in J_i$.

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 $\left|\widehat{h}(\gamma)\right| \leq \tau$ for all $\gamma \in \mathsf{span}([d+i])$ with $\|\gamma\|_1 = d$. We start by analyzing the Fourier coefficients of $h$, for which by Fact 2.5 we have

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

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 $\gamma$,

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

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 $\gamma$, Equation (3) reduces to

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

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 $\gamma\in\mathsf{span}(J\setminus J_i)$, 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 $\|\gamma\|_1\ge d+1$, 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 $\gamma\in\mathsf{span}([d+i])$ with $\|\gamma\|_1=d$, it must be that $|\widehat{h}(\gamma)|\le\tau$. Applying Equation (5) for such $\gamma$, we note that

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

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

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 2.12, we know that $\widehat{g\circ M_i}(\gamma)=\widehat{g\circ M_i^{-1}}(\gamma)=\widehat{g}(M_i^{\mathsf{T}}\gamma)$. Since the rows of $M_i^{\mathsf{T}}$ are the same as the columns of $M_i$, we have

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

Therefore, we can write $M_i^{\mathsf{T}}\gamma=\sum_{j\ne d+i}\gamma_j e_j+e_{d+i}\langle v+e_{d+i},\gamma\rangle=\gamma+e_{d+i}\langle v,\gamma\rangle$, 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

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

Simplifying the binomial sum, we get

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

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,

$$|K| = \frac{d}{e}\left(\frac{n-d+1}{\log 5/\tau}\right)^{1/d} - d + d - 1$$
$$\geq \frac{d}{e}\left(\frac{n}{\log 5/\tau}\right)^{1/d}\left(1-\frac{d-1}{n}\right)^{1/d} - 1 \geq \frac{d}{e}\left(\frac{n}{\log 5/\tau}\right)^{1/d}\left(\frac{1}{d}\right)^{1/d} - 1 \geq \frac{d}{4e}\left(\frac{n}{\log 5/\tau}\right)^{1/d}.$$

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 $\mathsf{span}(K)$, which do not change under the restriction. □

# 4 Lower Bounds on $\mathsf{r}(f,\delta)$

In this section, we prove lower bounds on $\mathsf{r}(f,\delta)$. 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.

## 4.1 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.

**Claim 4.1.** *For any 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 $u \in S$,*

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

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

*Proof.* Let $S = \{u : u \in \mathcal{W} \text{ and } |u+\mathcal{V}^\perp|^{=1} \geq 1\}$. 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 $\dim(\mathsf{span}(S \cup \mathcal{V}^\perp)) = n$. However, we also know that $\dim(\mathsf{span}(S \cup \mathcal{V}^\perp)) \leq |S| + C$, and rearranging we get $|S| \geq n - C$. Next, let $S_1 = \{u \in S : |u+\mathcal{V}^\perp|^{=1} = 1\}$. By Fact 2.4, for any $u, u' \in S$, we have $u + \mathcal{V}^\perp \neq u' + \mathcal{V}^\perp$. Therefore,

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

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

**Lemma 4.2.** *There is a degree one function $f : \mathbb{F}_2^n \to [-1,1]$ for which $\mathsf{r}(f,\delta) \geq n/2$, 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 Claim 4.1, 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 Fact 2.7, we have that

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

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, $\mathsf{r}(f,\delta) \geq n/2$. □

We next show how to generalize Lemma 4.2 to degree $d$ bounded functions.

**Lemma 4.3.** *For $d > 2$ and $\delta < \binom{n}{d}^{-1}$, there exists a degree $d$ function $f : \mathbb{F}_2^n \to [-1,1]$ for which $\mathsf{r}(f,\delta) \geq n - 2dn^{1/(d-1)}$.*

*Proof.* The counterexample is obtained using a probabilistic argument. We consider the homogeneous degree $d$ polynomial with random signs $f_{\mathbf{z}} : \mathbb{F}_2^n \to [-1, 1]$, defined as

$$f_{\mathbf{z}}(x) = \sum_{\gamma : \|\gamma\|_1 = d} \frac{\mathbf{z}_\gamma \cdot (-1)^{\langle \gamma, x \rangle}}{\binom{n}{d}},$$

where each $\mathbf{z}_\gamma \sim \{\pm 1\}$ is a uniformly random sign.

Let $\mathcal{V}$ be a subspace of $\mathbb{F}_2^n$ of co-dimension $C$, and suppose we restrict $f_{\mathbf{z}}$ to an affine subspace $\mathcal{U} = \alpha + \mathcal{V}$. By Claim 4.1, we have a $S \subseteq \mathcal{W}$ (where $\mathcal{W}$ is such that $\mathcal{W} \oplus \mathcal{V}^\perp = \mathbb{F}_2^n$) of size at least $k := n - C$ such that $|(u + \mathcal{V}^\perp)^{=1}| \geq 1$ for each $u \in S$. Moreover, by Fact 2.4, for every $v, v' \in \mathsf{span}(S)$ we have that $v + \mathcal{V}^\perp \neq v' + \mathcal{V}^\perp$. Therefore, there is a set $T \subseteq \mathcal{W}$ of size at least $\binom{k}{d}$ such that for every $u \in T$, we have $|(u + \mathcal{V}^\perp)^{=d}| \geq 1$. By Fact 2.7, for each $u \in T$, we have

$$|\widehat{f_{\mathcal{U}}}(u)| = \Big| \sum_{\eta \in u + \mathcal{V}^\perp} \widehat{f}(\eta)(-1)^{\langle \eta, \alpha \rangle} \Big|.$$

We now observe that if $(u + \mathcal{V}^\perp)^{=d}$ has odd size, then $\left|\widehat{f_{\mathcal{U}}}(u)\right| \geq \binom{n}{d}^{-1}$. Therefore, if $f_{\mathcal{U}}$ was $\delta$-regular, then for each $u \in T$, it must be that the set $(u + \mathcal{V}^\perp)^{=d}$ has even size, and, in particular, that $|(u + \mathcal{V}^\perp)^{=d}| \geq 2$.

Let $\mathcal{V}$ be a subspace such that each non-trivial affine subspace of $\mathcal{V}^\perp$ has an even number of weight $d$ vectors. For a given affine subspace $\mathcal{U} = \alpha + \mathcal{V}$ and a random choice of the signs $\mathbf{z}_\gamma$'s, the probability that $f_{\mathcal{U}}$ is $\delta$-regular is therefore at most $2^{-\binom{k}{d}}$. Let $\mathcal{B}$ (for "$\mathcal{B}$ad") be the event that there is an affine subspace $\mathcal{U}$ where $f_{\mathcal{U}}$ is $\delta$-regular. We can simply union bound over all possible affine subspaces of dimension at least $k$ to bound the probability of $\mathcal{B}$. For any $k$, observe that the number of affine subspaces of dimension $k$ is at most $2^{n(k+1)}$. Thus, we have

$$\mathbf{Pr}[\mathcal{B}] \leq \sum_{j=k}^{n} 2^{n(j+1)} \cdot 2^{-\binom{j}{d}} \leq \sum_{j=k}^{n} 2^{n(j+1) - \left(\frac{j}{d}\right)^d}.$$

Note that $h(x) = n(x+1) - \left(\frac{x}{d}\right)^d$ is concave in $[0, \infty)$; moreover, a quick calculation shows that it is maximized when $x = d \cdot n^{1/(d-1)}$. Setting $k = 2dn^{\frac{1}{d-1}}$, our desired probability is at most

$$\begin{aligned}
\mathbf{Pr}[\mathcal{B}] &\leq (n-k) \cdot 2^{n(2dn^{\frac{1}{d-1}}+1) - 2^d \cdot n^{\frac{d}{d-1}}} && \text{(Every term is smaller than the first.)} \\
&\leq (n-k) \cdot 2^{(2d+1-2^d) n^{1+\frac{1}{d-1}}} \\
&\leq o(1). && (2d + 1 - 2^d \leq -1 \quad \forall d \geq 3.)
\end{aligned}$$

Therefore, there exists a signing $\mathbf{z}_\gamma$ such that for any affine subspace of dimension at least $2dn^{1/(d-1)}$, the restriction of $f_z$ is not $\delta$-regular. □

*Remark* 4.4. Note that Lemma 4.3 is trivial when $d = 2$; it would be interesting to obtain a tighter result in this case.

## 4.2 Boolean Functions

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

### 4.2.1 Non-explicit Lower Bounds on $\mathsf{r}(f, \delta)$

We can turn our lower bounds on $\mathsf{r}(f, \delta)$ for bounded functions into (non-explicit) lower bounds for Boolean functions. To do so, we use the following simple but powerful lemma of [Hos+16], which states that given a bounded function with a large $\mathsf{r}(f, \delta)$, there must exist some Boolean function $g$ with similarly a large $\mathsf{r}(g, 2\delta)$.

**Lemma 4.5** ([Hos+16], Claim 1.2)**.** *Let $\tau > 0$ and $f : \mathbb{F}_2^n \to [-1,1]$. There exists a Boolean function $g : \mathbb{F}_2^n \to \{\pm 1\}$ satisfying, for every affine subspace $\mathcal{U}$ such that $|\mathcal{U}| \geq \frac{4n^2}{\tau^2}$ and any $\gamma \in \mathbb{F}_2^n$, that*

$$\left|\widehat{f_{\mathcal{U}}}(\gamma) - \widehat{g_{\mathcal{U}}}(\gamma)\right| \leq \tau.$$

*Proof.* Let $g(x)$ equal 1 with probability $\frac{1+f(x)}{2}$, and $-1$ otherwise. Let $\mathcal{U} = \alpha + \mathcal{V}$ for some subspace $\mathcal{V}$.

By Fact 2.6 we can write

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

Consider the random variable

$$\widehat{g_{\mathcal{U}}}(\gamma) = (-1)^{\langle \gamma, \alpha \rangle} \mathop{\mathbf{E}}_{y \in \mathcal{U}}[g(y) \cdot (-1)^{\langle \gamma, y \rangle}].$$

Observe that $\mathbf{E}_g \widehat{g_{\mathcal{U}}}(\gamma) = \widehat{f_{\mathcal{U}}}(\gamma)$. Moreover, every term in the summation is in $[-1,1]$, so by a Hoeffding bound (see Fact 2.1), the probability $\left|\widehat{g_{\mathcal{U}}}(\gamma) - \widehat{f_{\mathcal{U}}}(\gamma)\right| \geq \tau$ is at most $2\exp\left(-\tau^2|\mathcal{U}|^2/2\right) \leq 2^{-2n^2+1}$.

On the other hand, there are at most $2^{n^2}$ affine subspaces of $\mathbb{F}_2^n$, and at most $2^n$ choices for $\gamma$. Therefore, by a union bound, the probability that $g$ has the property we desire is at least $1 - 2^{n^2+n-2n^2+1} > 0$, and the claim follows. □

Using Lemma 4.5, we have the following lemma.

**Lemma 4.6.** *For all $d \geq 3$ and $\delta < \frac{1}{2} \cdot \binom{n}{d}^{-1}$, there exists a Boolean function $f$ with*

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

*Proof.* By Lemma 4.3, there exists a bounded $f$ that is not $\delta$-regular in any affine subspace of dimension at least $2dn^{1/(d-1)}$ for all $\delta < \binom{n}{d}^{-1}$. Lemma 4.5 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 $\log\left(16n^2/\delta^2\right)$. Therefore, if $f$ is not $\delta$-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 4.6 and achieve the following more parsable corollary.

**Corollary 4.7.** *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 $\mathsf{r}(f, \delta) \geq n - 2d \cdot n^{1/(d-1)}$.*

*Proof.* The function is the same as in Lemma 4.6. We argue that by our choice of parameters, $k$ is always maximized by the first term. We first note that $\frac{1}{2} \cdot \binom{n}{d}^{-1} > \frac{1}{2} \cdot n^{-d} = \delta$, so our choice for $\delta$ is valid. Next, we have that

$$\log(16n^2/\delta^2) = 5 + 2\log n + 2d\log n \leq 3d\log n \leq 3\frac{\log^2 n}{\log\log n},$$

where we used the fact that $d \geq 3$ and $n$ is sufficiently large. On the other hand, note that the function $h(x) = 2xn^{1/(x-1)}$ is decreasing when $x \leq \frac{\log n}{\log\log n + 1}$. Therefore, we have that

$$2dn^{1/(d-1)} \geq 2 \cdot \frac{\log n}{\log\log n} \cdot n^{\frac{\log\log n + 1}{\log n}} = 4 \cdot \frac{\log^2 n}{\log\log n}.$$

Therefore, the first term is the larger term in Lemma 4.6, as desired. □

### 4.2.2 Explicit Lower Bounds on $\mathsf{r}(f,\delta)$

**Lemma 4.8** (Related to Corollary 1.1 in [ODo+14]). *For each $\delta > 0$, there exists an explicit Boolean function $f : \mathbb{F}_2^n \to \{0,1\}$ with $\mathsf{r}(f,\delta) = \Omega\left((\log \frac{1}{\delta})^{\log_2(3)}\right)$.*

The proof of Lemma 4.8 is based on Theorem 4.9, which appeared in a slightly weaker form in [ODo+14].

**Theorem 4.9** ([ODo+14]). *Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$, and $g : \mathbb{F}_2^m \to \mathbb{F}_2$. We have that*

$$C^{\oplus}_{\min}[f \circ g] \geq C^{\oplus}_{\min}[f] + C_{\min}[f] \cdot B_g,$$

*where $B_g = \max\{\log C^{\oplus}_{\min}[g] - 1, 1\}$.*

In fact, in [ODo+14] Theorem 4.9 appeared as

$$C^{\oplus}_{\min}[f \circ g] \geq C^{\oplus}_{\min}[f] + C_{\min}[f],$$

but they assumed only that $C^{\oplus}_{\min}[g] \geq 2$. Therefore, the above result is strictly stronger for any $g$ such that $C^{\oplus}_{\min}[g] > 4$. We include a proof of this slightly stronger fact in Appendix A.

We require the following corollary of Theorem 4.9.

**Corollary 4.10.** *We have that*

$$C^{\oplus}_{\min}[f^{\circ k}] \geq B_g \cdot \frac{C_{\min}[f]^k - C_{\min}[f]}{C_{\min}[f] - 1} + C^{\oplus}_{\min}[f] \geq B_g \cdot C_{\min}[f]^{k-1},$$

*where $B_g = \max\{\log C^{\oplus}_{\min}[g] - 1, 1\}$.*

*Proof that Theorem 4.9 implies Corollary 4.10.* Let $f = f^{\circ(k-1)}$ and $g = f$. We have by the theorem that

$$\begin{aligned}
C^{\oplus}_{\min}[f^{\circ k}] &\geq C^{\oplus}_{\min}[f^{\circ(k-1)}] + C_{\min}[f^{\circ(k-1)}] \cdot B_g \\
&\geq C^{\oplus}_{\min}[f^{\circ(k-1)}] + C_{\min}[f]^{k-1} \cdot B_g && \text{(Supermultiplicativity of } C_{\min}\text{, see [Tal13])} \\
&\geq B_g \cdot \sum_{i=1}^{k-1} C_{\min}[f]^i + C^{\oplus}_{\min}[f] \\
&= B_g \cdot \frac{C_{\min}[f]^k - C_{\min}[f]}{C_{\min}[f] - 1} + C^{\oplus}_{\min}[f] \geq B_g \cdot C_{\min}[f]^{k-1}. && \square
\end{aligned}$$

For our application, we make the following crucial observation: if $f$ has Fourier coefficients that are all of equal magnitude $\delta$, then any restriction to an affine subspace results in Fourier coefficients of the restricted function that are integer multiples of $\delta$. Hence, if $f$ is $\delta'$-regular, for any $\delta' < \delta$, then $f$ is in fact, constant. In this scenario, finding a subspace in which $f$ is $\delta$-regular is equivalent to finding a subspace where it is constant.

*Proof of Lemma 4.8.* Consider the following function $g : \mathbb{F}_2^n \to \{\pm 1\}$:

$$g(x_1,x_2,x_3,x_4) = \frac{1}{2}(-1)^{x_1+x_3} + \frac{1}{2}(-1)^{x_2+x_3} + \frac{1}{2}(-1)^{x_1+x_4} - \frac{1}{2}(-1)^{x_2+x_4}.$$

Define the function $f := \frac{1-g}{2}$ so that $f : \mathbb{F}_2^n \to \mathbb{F}_2$. In other words, f is equal to $x_1 + x_3$ if $x_1 = x_2$ and $x_1 + x_4 + \mathbb{1}\{x_1 = 0\}$ otherwise. Note that $f$ is a degree 2 function, where all non-zero Fourier coefficients have the same magnitude. We also claim that $C^{\oplus}_{\min}[f] = 2$. Indeed, we can fix $x_1 + x_2 = 0$ and $x_1 + x_3 = 0$, and we know $f$ equals 0. On the other hand, we have that $C_{\min}[f] = 3$.

We examine $f^{\circ k} = f(f_1, f_2, f_3, f_4)$, where the $f_i$'s are copies of $f^{\circ(k-1)}$ over disjoint sets of inputs. We claim by induction that $\deg(f^{\circ k}) = 2^k$. This is clearly true when $k = 1$, and for the inductive step we can write

$$f(f_1,f_2,f_3,f_4) = \begin{cases} f_1 + f_3 & \text{if } f_1 = f_2 \\ f_4 + f_1 & \text{if } f_1 \neq f_2. \end{cases}$$

Since $\mathbb{1}\{f_1 = f_2\} = f_1 + f_2 + 1$, we can write

$$f(f_1, f_2, f_3, f_4) = (f_1 + f_3)(f_1 + f_2 + 1) + (f_1 + f_4)(f_1 + f_2).$$

Therefore, by the inductive hypothesis and the fact that the $f_i$'s are supported over disjoint variables, we have that $\deg f^{\circ k} = 2 \cdot \deg f^{\circ(k-1)} = 2^k$. Therefore,[12] all the Fourier coefficients are integer multiples of $1/2^{2^k}$. So, to make $f$ $\delta$-regular for $\delta < 1/2^{2^k}$, it must be fixed to a constant. Suppose we set $\delta = 1/2^{2^k+1}$. By Corollary 4.10, we have that

$$C^{\oplus}_{\min}[f^{\circ k}] \geq C_{\min}[f]^{k-1} = 3^{k-1} = \frac{1}{3} \cdot (2^k)^{\log_2(3)} = \frac{1}{3} \cdot \frac{1}{2^{\log_2(3)}} \log(1/\delta)^{\log_2(3)} = \frac{1}{9} \cdot \log(1/\delta)^{\log_2(3)}. \quad \square$$

We now show that the majority function, denoted by $\mathsf{MAJ}_n$, also has a large $\mathsf{r}(f, \delta)$ value when $\delta = O(1/\sqrt{n})$.

**Lemma 4.11.** *There is an absolute constant $C > 0$, such that for all sufficiently large $n$, $\mathsf{r}(\mathsf{MAJ}_n, \delta) \geq \Omega(n^{1/2})$ for any $\delta \leq C/\sqrt{n}$.*

We need the following three claims to prove this lemma.

**Claim 4.12** (Fourier Spectrum of $\mathsf{MAJ}_n$, Corollary of Theorem 5.19 in [ODo21])**.** *Consider $f = \mathsf{MAJ}_n$. Each of the following hold.*

1. *For each $t \in \mathbb{N}$ and $\gamma \in \mathbb{F}_2^n$ with $\|\gamma\|_1 = t$,*

$$\left|\widehat{f}(\gamma)\right| \leq \begin{cases} \left(\frac{t}{n}\right)^{\frac{t-1}{2}} \left|\widehat{f}(e_1)\right|, & \text{if } t \text{ is odd,} \\ 0 & \text{otherwise.} \end{cases}$$

2. *For any $\gamma$ with $\|\gamma\|_1 = 1$, $\left|\widehat{f}(\gamma)\right| \geq \sqrt{\frac{2}{\pi n}}$.*

3. *For any $\gamma, \gamma'$ such that $\|\gamma\|_1 + \|\gamma'\|_1 = n + 1$, it holds that $\left|\widehat{f}(\gamma)\right| = \left|\widehat{f}(\gamma')\right|$.*

**Claim 4.13** ([htt])**.** *Let $\mathcal{U} = \alpha + \mathcal{W}$ be any affine subspace of $\mathbb{F}_2^n$. For every $t \in [n]$, let $t^* := \min\{t, n - t\}$. Then, it holds that*

$$|\mathcal{U}^{=t}| \leq \binom{\dim(\mathcal{W}) + 1}{\leq t^*}.$$

**Lemma 4.14.** *Let $\mathcal{V}$ be a subspace of $\mathbb{F}_2^n$ of co-dimension $C$ and $\mathcal{W}$ be such that $\mathcal{W} \oplus \mathcal{V}^\perp = \mathbb{F}_2^n$. For each $\ell \leq C + 1$, there exists $S_\ell \subseteq \mathcal{W}$ such that $|S_\ell| \geq n - C(\ell + 1)$, and for each $\gamma \in S_\ell$ the following two hold:*

1. $\left|\left(\gamma + \mathcal{V}^\perp\right)^{=1}\right| = 1$ *and*

2. $\left|\left(\gamma + \mathcal{V}^\perp\right)^{=t}\right| \leq 2 \cdot \binom{2C+1}{t-1}$, *for each $t \leq \ell$.*

Claim 4.12 and Claim 4.13 are powerful enough by themselves to achieve a weaker form of Lemma 4.11: one can use them to show that $\mathsf{MAJ}_n$ is not $\Omega(n^{-1/2})$-regular in any subspace of co-dimension $O(n^{1/3})$.[13] We now use Claim 4.12, Claim 4.13, and Lemma 4.14, the proofs of which are deferred to Appendix B, to prove Lemma 4.11.

[12] See [ODo21], Exercise 1.9 or Claim 5.2.

[13] Following the proof sketch in Section 1.3.2, the reason the analysis breaks if we try to use only Claim 4.13 and set $C = n^{1/3+\varepsilon}$ for any $\varepsilon > 0$ is as follows. By Claim 4.13, there could be on the order of $C^3 = n^{1+3\varepsilon}$ weight three vectors in our signed sum corresponding to the new Fourier coefficient. Since $|\widehat{\mathsf{MAJ}_n}(\gamma)| = \Theta(n^{-3/2})$ when $\|\gamma\|_1 = 3$, these coefficients could combine constructively to a magnitude of $\approx n^{1+3\varepsilon} \cdot n^{-3/2} \gg n^{-1/2}$, thus potentially cancelling out the (single) level one coefficient, which has magnitude $|\widehat{\mathsf{MAJ}_n}(e_1)| = \Theta(n^{-1/2})$.

*Proof of Lemma 4.11.* Let $\mathcal{V}$ be a subspace of $\mathbb{F}_2^n$ of co-dimension $C = \frac{\sqrt{n}}{10e} - 1$, and suppose we restrict $\mathsf{MAJ}_n$ to the affine subspace $\mathcal{U} = \alpha + \mathcal{V}$. Applying Lemma 4.14 with $\ell = C + 1$, we get a subset $S \subseteq \mathcal{W}$ (where $\mathcal{W}$ is such that $\mathcal{W} \oplus \mathcal{V}^\perp = \mathbb{F}_2^n$) of size at least 3 such that each element $\gamma \in S$ satisfies both items in the lemma. In particular, there must be $u \in S$ that satisfies both properties as well as, $\mathbf{0} \notin u + \mathcal{V}^\perp$ and $\mathbf{1} \notin u + \mathcal{V}^\perp$.[14] For notational ease, let us denote $E := u + \mathcal{V}^\perp$. By Fact 2.7, we have

$$\begin{aligned}\left|\widehat{f_{\mathcal{U}}}(u)\right| = \left|\sum_{\eta \in E} \widehat{f}(\eta)(-1)^{\eta,\alpha}\right| &\geq \left|\widehat{f}(e_1)\right| - \sum_{t>1}^{n-1} |E^{=t}| \cdot \left|\widehat{f}\left(\sum_{i=1}^{t} e_i\right)\right| \\ &= \left|\widehat{f}(e_1)\right| - \sum_{t>1}^{\frac{n+1}{2}} \left(|E^{=t}| + |E^{=n-t+1}|\right) \cdot \left|\widehat{f}\left(\sum_{i=1}^{t} e_i\right)\right|. \end{aligned} \tag{6}$$

In the second to last step, we used the facts that majority is a symmetric function and that $|E^{=1}| = 1$ and $|E^{=n}| = 0$. In the last step, we used item 3 of Claim 4.12. Next, we claim that

$$|E^{=t}| + |E^{=n-t+1}| \leq \begin{cases} (t+1)\binom{2C+1}{t-1} & \text{when } t \leq \ell, \\ 2^{C+1} & \text{otherwise.} \end{cases} \tag{7}$$

For the first case, when $t \leq \ell$, by item 2 of Lemma 4.14, we have $|E^{=t}| \leq 2\binom{2C+1}{t-1}$. Furthermore, from Claim 4.13, we have $|E^{=n-t+1}| \leq \binom{C+1}{\leq t-1} \leq \binom{2C+1}{\leq t-1} \leq (t-1) \cdot \binom{2C+1}{t-1}$ for all $1 < t \leq \frac{\sqrt{n}}{10e}$. When $t > \ell$, we note that both $|E^{=t}|$ and $|E^{=n-t+1}|$ are at most $2^C$ since the dimension of $\mathcal{V}^\perp$ is $C$, and this is tighter when $t > C + 1$.

Using Equation (7) and Claim 4.12, we can estimate the sum in Equation (6) as

$$\left|\widehat{f_{\mathcal{U}}}(u)\right| \leq \left|\widehat{f}(e_1)\right| \left( 1 - \underbrace{\sum_{t=3}^{\ell} \binom{2C+1}{t-1} \left(\frac{t}{n}\right)^{\frac{t-1}{2}} (t+1)}_{A} - \underbrace{\sum_{t>\ell}^{\frac{n+1}{2}} 2^{C+1} \left(\frac{t}{n}\right)^{\frac{t-1}{2}}}_{B} \right).$$

We complete the argument by showing an upper bound on both the above sums. Starting with $A$, and recalling that $C = \frac{\sqrt{n}}{10e} - 1$, we see that

$$\begin{aligned} A &\leq \sum_{t=3}^{\frac{\sqrt{n}}{10e}} \left(\frac{\sqrt{n}}{5(t-1)}\right)^{t-1} \left(\frac{t}{n}\right)^{\frac{t-1}{2}} \cdot (t+1) && \left(\tbinom{n}{k} \leq \left(\tfrac{en}{k}\right)^k .\right) \\ &\leq \sum_{t=3}^{\frac{\sqrt{n}}{10e}} \left(\frac{\sqrt{t}}{5(t-1)}\right)^{t-1} \cdot (t+1) \\ &\leq \sum_{t=3}^{\frac{\sqrt{n}}{10e}} \left(\frac{(t+1)^{\frac{1}{2}}}{2(t+1)}\right)^{t-1} \cdot (t+1) && \left(5(t-1) \geq 2(t+1) \quad \forall t \geq 3.\right) \\ &= \sum_{t=3}^{\frac{\sqrt{n}}{10e}} \left(\frac{1}{2(t+1)^{1/2}}\right)^{t-1} \cdot (t+1) \leq \sum_{i=2}^{\infty} \left(\frac{1}{2}\right)^i \leq 1/2. \end{aligned}$$

In the penultimate inequality, we used the fact that for the first term, when $t = 3$, we have $\left(\frac{1}{2(t+1)^{1/2}}\right)^{t-1} \cdot (t+1) = \left(\frac{1}{4}\right)^2 \cdot 4 = 1/4$, and the ratio of the summands (for $t \geq 3$) is

$$\frac{(2(t+1)^{1/2})^{t-1}}{(2(t+2)^{1/2})^t} \cdot \frac{t+2}{t+1} \leq \frac{1}{2(t+1)^{1/2}} \cdot 2 \leq \frac{1}{(t+1)^{1/2}} \leq \frac{1}{2}.$$

[14] It is vital that $\mathbf{1} \notin M\alpha^* + \mathcal{V}^\perp$ since $|\widehat{\mathsf{MAJ}_n}(\mathbf{1})| = |\widehat{\mathsf{MAJ}_n}(e_i)|$, so they could cancel each other out.

To bound $B$, we note that the function $h(x) = \left(\frac{x}{n}\right)^{(x-1)/2}$ is strictly convex, which means its maximum occurs either at $t = C$ or $t = \frac{n+1}{2}$. Again, setting $C = \frac{\sqrt{n}}{10e} - 1$, a quick calculation shows that the maximum is achieved for the first term, and this term is at most

$$\begin{aligned} 2^{\frac{\sqrt{n}}{10e}} \cdot \left(\frac{1}{10e\sqrt{n}}\right)^{\frac{\sqrt{n}}{20e}} &\leq 2^{\frac{\sqrt{n}}{10e}} \cdot 2^{-\log n \cdot \frac{\sqrt{n}}{40e}} \\ &\leq 2^{\frac{\sqrt{n}}{10e} - \left(\frac{\sqrt{n}}{10e} + 2\log n\right)} && \left(\tfrac{\sqrt{n}}{40e}\log n \geq \tfrac{\sqrt{n}}{10e} + 2\log n \text{ for large enough } n.\right) \\ &= 2^{-2\log n} = \frac{1}{n^2}. \end{aligned}$$

This implies that $B \leq n \cdot \frac{1}{n^2} \leq o(1)$. Using item 2 of Claim 4.12, we conclude that there is a non-trivial Fourier coefficient

$$\left|\widehat{f_{\mathcal{U}}}(u)\right| \geq \left|\widehat{f}(e_1)\right| (1 - 1/2 - o(1)) = \Omega(n^{-1/2}). \quad \square$$

# 5 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 5.1.** *We say a function $f : \mathbb{F}_2^n \to \mathbb{R}$ is $G$-granular if for every $x \in \mathbb{F}_2^n$, we have that $f(x)$ is an integer multiple of $G$.*

**Claim 5.2.** *If a degree $d$ function $f : \mathbb{F}_2^n \to \mathbb{R}$ is $G$-granular, then for every $\gamma \in \mathbb{F}_2^n$, we have that $\widehat{f}(\gamma)$ is an integer multiple of $2^{-d} \cdot G$.*

*Proof.* Note that if we associate $\mathbb{F}_2$ with $\{0,1\}$, any $f : \mathbb{F}_2^n \to \mathbb{R}$ has a real multilinear polynomial representation $q : \{0,1\}^n \to \mathbb{R}$, where $q(x) = f(x)$ for all $x \in \{0,1\}^n$ (see Exercise 1.9 in [ODo21]). In particular, we can write $q$ as a sum of its indicators:

$$q(x) = \sum_{a \in \{0,1\}^n} \mathbb{1}\{x = a\} \cdot q(a).$$

Noting that $\mathbb{1}\{x = a\} = \prod_i (1 - a_i - x_i)(1 - 2a_i)$, we see that every coefficient of $q$ is an integer multiple of $G$.

However, we can also associate $f$ with a real multilinear polynomial, $p : \{\pm 1\}^n \to \mathbb{R}$, such that $f(x) = p((-1)^x)$ for all $x \in \mathbb{F}_2^n$. Note then, that $p(x) = q((1 - x_1)/2, \ldots, (1 - x_n)/2))$, so if $p$ has degree $d$, then all its coefficients are integer multiples of $G \cdot 2^{-d}$. Finally, note that $f$ and $p$ have the same Fourier coefficients (and therefore degree), which implies the result. $\square$

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 (Theorem 1.6).

**Lemma 5.3.** *Every degree $d$ function $f : \mathbb{F}_2^n \to [-1, 1]$ that is $G$-granular satisfies*

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

*Proof.* If $f$ is $G$-granular and degree $d$, then from Claim 5.2 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 $\delta$-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 $\delta = 2^{-d-1} \cdot G$. $\square$

*Proof of Theorem 1.6.* Using $f$, we can construct a degree $d$ function $h : \mathbb{F}_2^n \to [-1, 1]$ as $h(x) = 1 - \frac{2f(x)}{C}$. Noting that $h$ is $2/C$-granular and using the above lemma, it follows that

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

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. □

Last, we give a connection between the notion of $\delta$-regularity and affine extractors. Formally, we define affine extractors as follows.

**Definition 5.4** (Affine Extractor)**.** *A function $f : \mathbb{F}_2^n \to \{0, \ldots, C\}$ is said to be a* $(k, \delta)$-affine extractor *if for all affine subspaces $\mathcal{U}$ of dimension at least $k$, we have that*

$$|f_{\mathcal{U}} - \mathsf{Unif}_C| \leq \delta,$$

*where $\mathsf{Unif}_C$ is the uniform distribution over $\{0, \ldots, C\}$.*

**Claim 5.5.** *If $f$ is a $(k, \delta)$-extractor, then $f$ becomes $2C\delta$-regular when restricted to any affine subspace of dimension at least $k + 1$.*

*Proof.* Note that if $f : \mathbb{F}_2^n \to \{0, \ldots, C\}$ is a $(k, \delta)$-extractor then in any affine subspace $\mathcal{U}$, of dimension at least $k$, we have,

$$\left|\widehat{f_{\mathcal{U}}}(\chi_{\mathbf{0}}) - \frac{C}{2}\right| = \left|\sum_c c\left(\mathop{\mathbf{Pr}}_{x\in\mathcal{U}}[f(x) = c] - \frac{1}{C+1}\right)\right| \leq C\sum_c \left|\mathop{\mathbf{Pr}}_{x\in\mathcal{U}}[f(x) = c] - \frac{1}{C+1}\right| \leq 2C\delta.$$

Suppose $f$ is a $(k, \delta)$-affine extractor. Let us assume to a contradiction that $\mathcal{U}$ is an affine subspace of dimension at least $k+1$, where $f_{\mathcal{U}}$ has a Fourier coefficient with magnitude larger than $2C\delta$. By Corollary B.1, we can fix the parity corresponding to this Fourier coefficient in such a way that the bias of the function increases by $2C\delta$, which gives the desired contradiction. □

# 6 Future Directions

We highlight two open problems that offer particularly interesting research directions. First, there is a tantalizing, and large, gap between our Theorem 1 and Lemma 4.3 for bounded degree $d$ functions. We suspect that Lemma 4.3 is closer to being tight and ask the following question.

**Direction 6.1.** *Can the upper bound on $\mathsf{r}(f, \delta)$ in Theorem 1 be improved?*

Moreover, it would be interesting to find explicit Boolean and bounded functions with large $r(f, \delta)$ values.

**Direction 6.2.** *Find (explicit) examples of functions $f : \mathbb{F}_2^n \to [-1, 1]$ with $r(f, \delta)$ values comparable to those obtained in Lemma 4.3. Similarly, find (explicit) Boolean functions with similar $r(f, \delta)$ values.*

# 7 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.

# References


[Alo+01] Noga Alon, Eldar Fischer, Michael Krivelevich, and Mario Szegedy. "Efficient Testing of Large Graphs". In: *Combinatorica* 20 (May 2001). DOI: 10.1007/s004930070001.

[Bd02] Harry Buhrman and Ronald de Wolf. "Complexity measures and decision tree complexity: a survey". In: *Theoretical Computer Science* 288.1 (2002), pp. 21–43. DOI: https://doi.org/10.1016/S0304-3975(01)00144-X.

[Ben17] Shalev Ben David. "Quantum speedups in query complexity". PhD thesis. Massachusetts Institute of Technology, 2017.

[CGL21] Eshan Chattopadhyay, Jesse Goodman, and Jyun-Jie Liao. "Affine Extractors for Almost Logarithmic Entropy". In: *FOCS*. IEEE, 2021, pp. 622–633.

[CT15] Gil Cohen and Avishay Tal. "Two Structural Results for Low Degree Polynomials and Applications". In: *APPROX-RANDOM*. Vol. 40. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2015, pp. 680–709.

[Gir+21] Uma Girish, Justin Holmgren, Kunal Mittal, Ran Raz, and Wei Zhan. "Parallel Repetition for the GHZ Game: A Simpler Proof". In: *APPROX-RANDOM*. Vol. 207. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021, 62:1–62:19.

[Gre05] Ben Green. "A Szemerédi-type regularity lemma in abelian groups". In: *Geometric and Functional Analysis* 15 (Jan. 2005), pp. 340–376. DOI: 10.1007/s00039-005-0509-8.

[HJ63] A. W. Hales and R. I. Jewett. "Regularity and Positional Games". In: *Transactions of the American Mathematical Society* 106.2 (1963), pp. 222–229. ISSN: 00029947.

[Hoe63] Wassily Hoeffding. "Probability Inequalities for Sums of Bounded Random Variables". In: *Journal of the American Statistical Association* 58.301 (1963), pp. 13–30. ISSN: 01621459. URL: http://www.jstor.org/stable/2282952 (visited on 06/27/2022).

[Hos+16] Kaave Hosseini, Shachar Lovett, Guy Moshkovitz, and Asaf Shapira. "An improved lower bound for arithmetic regularity". In: *Mathematical Proceedings of the Cambridge Philosophical Society* 161.2 (2016), pp. 193–197. DOI: 10.1017/S030500411600013X.

[htt] Fedor Petrov (https://mathoverflow.net/users/4312/fedor-petrov). *Low-Hamming weight vectors in low-dimensional subspaces of* $\mathbb{F}_p^n$. MathOverflow. URL: https://mathoverflow.net/q/389026.

[Li16] Xin Li. "Improved Two-Source Extractors, and Affine Extractors for Polylogarithmic Entropy". In: *FOCS*. IEEE Computer Society, 2016, pp. 168–177.

[Mid04] Gatis Midrijanis. *Exact quantum query complexity for total Boolean functions*. 2004. DOI: 10.48550/ARXIV.QUANT-PH/0403168. URL: https://arxiv.org/abs/quant-ph/0403168.

[ODo+14] Ryan O'Donnell, John Wright, Yu Zhao, Xiaorui Sun, and Li-Yang Tan. "A Composition Theorem for Parity Kill Number". In: *Computational Complexity Conference*. IEEE Computer Society, 2014, pp. 144–154.

[ODo21] Ryan O'Donnell. "Analysis of Boolean Functions". In: *CoRR* abs/2105.10386 (2021).

[Rot53] K. F. Roth. "On Certain Sets of Integers". In: *Journal of the London Mathematical Society* s1-28.1 (1953), pp. 104–109. DOI: https://doi.org/10.1112/jlms/s1-28.1.104.

[RS76] I. Ruzsa and E. Szemerédi. "Triple systems with no six points carrying three triangles". In: *Combinatorica* 18 (Jan. 1976).

[Sha06] Asaf Shapira. *Graph Property Testing and Related Problems*. University of Tel-Aviv, 2006.

[Sha11] Ronen Shaltiel. "Dispersers for Affine Sources with Sub-polynomial Entropy". In: *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*. 2011, pp. 247–256. DOI: 10.1109/FOCS.2011.37.

[Sze75] Endre Szemerédi. *Regular partitions of graphs*. Tech. rep. Stanford Univ Calif Dept of Computer Science, 1975.

[Tal13] Avishay Tal. "Properties and applications of boolean function composition". In: *ITCS*. ACM, 2013, pp. 441–454.

# A Short Proof of the Parity Kill Number Theorem ([ODo+14])

We present a more concise and slightly improved version of the main theorem of [ODo+14], which appears as Theorem 4.9 above.

The following proposition suffices to prove the theorem.

**Proposition A.1.** *Let $f' : \mathbb{F}_2^n \times \mathbb{F}_2 \to \mathbb{F}_2$ and $g : \mathbb{F}_2^k \to \mathbb{F}_2$. We let $f : \mathbb{F}_2^n \times \mathbb{F}_2^k \to F_2$ be defined as*

$$f(x, y) = f'(x, g(y)).$$

*Then for any affine subspace $H \subseteq \mathbb{F}_2^n \times \mathbb{F}_2^k$ on which $f$ is constant, there exists some $H' \subseteq \mathbb{F}_2^n \times \mathbb{F}_2$ on which $f'$ is constant such that either:*

1. $\operatorname{codim}(H') \leq \operatorname{codim}(H) - B_g$, *where* $B_g = \max\{1, \log C^{\oplus}_{\min}[g] - 1\}$, *as before.*
2. *The $(n+1)$-st coordinate (so $g(y)$) is irrelevant in $H'$ and* $\operatorname{codim}(H') \leq \operatorname{codim}(H)$.

*Furthermore, among the first $n$ coordinates, any coordinate that was irrelevant in $H$ remains irrelevant in $H'$.*

Before proving Proposition A.1, let's see how it implies Theorem 4.9. Note that $f \circ g = f(g(x_1), ..., g(x_n))$, so we will apply Proposition A.1 $n$ times. The crucial observation is that we must fall into the first case of Proposition A.1 at least $C_{\min}[f]$ times. This is because if $f(x, y)$ is constant on $H$, then $H$ must depend on at least $C_{\min}[f]$ coordinates.

Let then $H \subseteq \mathbb{F}_2^{n \cdot m}$ be a minimum co-dimension subspace on which $f \circ g$ is constant, so that $\operatorname{codim}(H) = C^{\oplus}_{\min}[f \circ g]$. Applying Proposition A.1 $n$ times, we derive $H'$ on which $f$ is constant.

$$\begin{aligned} C^{\oplus}_{\min}[f] &\leq \operatorname{codim}(H') \\ &\leq \operatorname{codim}(H) - B_g \cdot C_{\min}[f] \\ &= C^{\oplus}_{\min}[f \circ g] - B_g \cdot C_{\min}[f]. \end{aligned}$$

Rearranging gives the theorem.

Finally, before we prove Proposition A.1, we need the following lemma, the proof of which is not complicated but we will omit and can be found in [ODo+14].

**Lemma A.2** ([ODo+14], Lemma 3.3)**.** *Let $H \subseteq \mathbb{F}_2^k \times \mathbb{F}_2^n$ be an affine subspace. Then there exists an invertible linear transformation $L$ on $\mathbb{F}_2^k \times \mathbb{F}_2^n$ such that, after applying this linear transformation, the constraints of $H$ can be partitioned into*

- *$\mathcal{B}_{x,y}$, which contain constraints of the form $x_i + y_i = \sigma_i$, for $1 \leq i \leq t$.*
- *$\mathcal{B}_x$, which contain constraints of the form $x_j = \sigma_j$, for $t + 1 \leq j \leq t'$.*
- *$\mathcal{B}_y$, which contain constraints of the form $y_k = \sigma_k$, for $t' + 1 \leq k \leq t''$.*

*and $t + (t' - t) + (t'' - t') = \operatorname{codim}(H)$.*

The takeaway of the above lemma is that since parity kill number is invariant under affine transformations, we can "canonize" any affine subspace in a way that minimizes the interactions between coordinates.

*Proof of Proposition A.1.* WLOG suppose that $H$ is of the form given in Lemma A.2.

1. **Easy Case: $|\mathcal{B}_{x,y}| = 0$.**

   Let's denote $C_y$ as the set of all $y$ that satisfy the constraints in $\mathcal{B}_y$, and let $C_x$ (analogously) be the set of $x$ that satisfy the constraints of $\mathcal{B}_x$.

a) **Subcase 1:** Suppose that $g(y) = b$ for all $y \in C_y$. Then we can let

$$H' = \{(x, z) \mid x \in C_x, z = b\}.$$

$f'$ is clearly constant on $H'$. Note that

$$\text{codim}(H') = |\mathcal{B}_x| + 1 = \text{codim}(H) - |\mathcal{B}_y| + 1 \leq \text{codim}(H) - C^{\oplus}_{\min}[g] + 1 \leq \text{codim}(H) - B_g,$$

as desired for the first case of Proposition A.1.

b) **Subcase 2:** Suppose that $g$ is not constant on the inputs in $C_y$. In this case, we claim that

$$H' = \{(x, z) \mid x \in C_x\}$$

makes $f'$ constant. Indeed, suppose it doesn't. Then there are two inputs $(x, z)$ and $(x', z')$ such that $f'(x, z) \neq f'(x', z')$. But then, we can pick $y, y' \in C_y$ such that $g(y) = z$ and $g(y') = z'$, and this results in $(x, y), (x', y') \in H$ such that $f(x, y) \neq f(x', y')$, a contradiction.

Finally, note that $\text{codim}(H') = |\mathcal{B}_x| = \text{codim}(H) - |\mathcal{B}_y| \leq \text{codim}(H) - B_g$. In fact, we don't even need this to be true in order to fall into the second case of the proposition (since $H'$ does not depend on its last coordinate), but it is nonetheless true.

2. **(Slightly) Harder Case:** $|\mathcal{B}_{x,y}| \neq 0$**.**

a) **Subcase 1:** $g$ becomes a junta on $y_1, ..., y_t$ when restricted to $C_y$. In this case, let $I = \{i_1, ..., i_s\} \subseteq [t]$ be the junta variables, so that $g(y) = h(y_{i_1}, ..., y_{i_s})$ for all $y \in C_y$. Then we claim that $f'$ is constant on

$$H' = \{(x, z) | x \in C_x, x_i = 0 \, \forall i \in [t] \setminus I, z = h(x_{i_1} \oplus \sigma_{i_1}, ...x_{i_s} \oplus \sigma_{i_s})\}.$$

Indeed, suppose it is not, so that $f'(x, z) \neq f'(x', z')$. Take $y \in C_y$ such that $y_i = \sigma_i \, \forall i \in [t] \setminus I$, and $y_j = \sigma_j \oplus x_j \, \forall k \in I$. Then we have that $g(y) = h(y_{i_1}, ..., y_{i_s}) = z$. Similarly, we can find $y' \in C_y$ such that $g(y') = z'$. We end up at a contradiction though, since $(x, y)$ and $(x', y')$ are both in $H$, but are such that $f(x, y) \neq f(x', y')$.

Finally, note that the codimension of $H'$ is exactly $|\mathcal{B}_x| + |\mathcal{B}_{x,y}| - s + 1 = \text{codim}(H) - (|\mathcal{B}_y| + s) + 1$. Next, we claim that $|\mathcal{B}| + s \geq \log C^{\oplus}_{\min}[g]$. To see why this is the case, note that we can fix $g$ by fixing at most $|\mathcal{B}_y| + 2^s$ parities/variables. This implies that $|\mathcal{B}_y| + 2^s \geq C^{\oplus}_{\min}[g]$ which implies that $|\mathcal{B}_y| + s \geq \log C^{\oplus}_{\min}[g]$. Thus, we have that

$$\text{codim}(H') \leq \text{codim}(H) - B_g,$$

as desired.

b) **Subcase 2:** There exists some $b_1, ..., b_t$ such that $g(y)$ is *not* constant on

$$C'_y := \left\{ y \,\middle|_{y_i = b_i \;\forall\, 1 \leq i \leq t}^{y \in C_y} \right\}.$$

In this case, let

$$H' = \{(x, z) \mid x \in C_x, \;\; x_j = b_j \oplus \sigma_j \;\; \forall \, 1 \leq j \leq t\}.$$

First, we claim that $f'$ is constant on $H'$. As before, suppose it is not, so that $f'(x, z) \neq f'(x', z')$. Then by definition, there exists $y, y'$ such that $y_i = y'_i = b_i$ for all $i \in [t]$, such that $g(y) = z$ and $g(y') = z'$. In this case, $(x, y)$ and $(x', y')$ are both in $H$, but are such that $f(x, y) \neq f(x', y')$, a contradiction.

Finally, note that $\text{codim}(H') \leq \text{codim}(H)$, but that $H'$ is independent of its last coordinate $z$, so that we fall into the second case of Proposition A.1.

□

# B Omitted Proofs

*Proof of Fact 2.12.* We have that

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

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

## B.1 Proofs of Proposition 1.3 and Proposition 1.4

In this section we provide the proofs of Proposition 1.3 and Proposition 1.4. We first begin with a corollary of Fact 2.7 which will be useful in the analysis of the claims.

**Corollary B.1.** *When $\mathcal{V}$ has dimension $n-1$, this corresponds to fixing a single parity $\sum_{i:\gamma_i=1} x_i$ to $b \in \{0,1\}$. Then $\mathcal{V}^\perp$ is simply $\mathsf{span}(\{\gamma\})$ and $\alpha$ is any vector such that $\langle \gamma, \alpha\rangle = b$.*

*Then for all $\gamma' \neq \gamma$ we have by Fact 2.7 that*

$$\widehat{f_{\alpha+\mathcal{V}}}(\chi_{\gamma'}) = (-1)^{\langle \gamma', \alpha\rangle}\widehat{f}(\gamma') + (-1)^{\langle \gamma+\gamma', \alpha\rangle} = (-1)^{\langle \gamma', \alpha\rangle}\left(\widehat{f}(\gamma') + (-1)^b \cdot \widehat{f}(\gamma+\gamma')\right).$$

*In particular, there exists a choice of $b$ such that*

$$\left|\widehat{f_{\alpha+\mathcal{V}}}(\chi_{\mathbf{0}})\right| = \left|\widehat{f}(\mathbf{0}) + \widehat{f}(\gamma)\right|.$$

*Proof of Proposition 1.3.* Given some $f : \mathbb{F}_2^n \to [-1,1]$, consider the following simple procedure:

- While at least $\delta$ fraction of $\pi \in \Pi$ have some $\gamma_\pi$ such that $|\widehat{f_\pi}(\gamma_\pi)| > \delta$, further partition each $\pi$ 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 that $\frac{1}{\delta^3}$ 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 B.1 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 $\pi$ to $\Phi$ 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 $\delta$ fraction of $\pi \in \Pi$ had some $\gamma_\pi$ such that $|\widehat{f_\pi}(\gamma_\pi)| > \delta$, at each step of the refinement $\Phi$ must increase by at least $\delta^3$, completing the proof. □

*Proof of Proposition 1.4.* Suppose without loss of generality, $\mathbf{E}\, f \geq 0$. Start with the trivial subspace, $\pi_0 = \mathbb{F}_2^n$. While there exists $\gamma$ such that $|\widehat{f_{\pi_t}}(\gamma)| > \delta$, by Corollary B.1 we can fix the parity corresponding to $\gamma$ 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 $\pi$, this process can happen at most $\frac{1}{\delta}$ times. □

## B.2 Claim 4.12

*Proof of Claim 4.12.* Theorem 5.19 in [ODo21] gives the following formula for the Fourier coefficients of the Majority function:

$$\left|\widehat{\mathsf{MAJ}_n}(\gamma)\right| = \frac{\binom{\frac{n-1}{2}}{\frac{t-1}{2}}}{\binom{n-1}{t-1}} \cdot \frac{2}{2^n}\binom{n-1}{\frac{n-1}{2}},$$

which holds for all $\gamma$ such that $\|\gamma\|_1 = t$ is odd. Otherwise, $\widehat{\mathsf{MAJ}_n}(\gamma) = 0$. By the above equation, we have that

$$\begin{aligned}\frac{\widehat{\mathsf{MAJ}_n}(\gamma)}{\widehat{\mathsf{MAJ}_n}(e_1)} &= \frac{\binom{\frac{n-1}{2}}{\frac{t-1}{2}}}{\binom{n-1}{t-1}} \\ &= \frac{\left(\frac{n-1}{2}\right)! \cdot (t-1)! \cdot (n-t)!}{\left(\frac{t-1}{2}\right)! \cdot \left(\frac{n-t}{2}\right)! \cdot (n-1)!} \\ &= \frac{(t-2)!! \cdot (n-t-1)!!}{(n-2)!!} \\ &= \frac{(t-2) \cdot (t-4) \cdots 1}{(n-2) \cdot (n-4) \cdots (n-t+1)} \\ &\leq \left(\frac{t}{n}\right)^{\frac{t-1}{2}}.\end{aligned}$$

□

## B.3 Claim 4.13

*Proof of Claim 4.13.* First, consider

$$\widetilde{\mathcal{U}} = \begin{cases} \mathcal{U} & \text{if } t \leq n/2 \\ \mathbf{1} + \mathcal{U} & \text{if } t > n/2. \end{cases}$$

Note that $\dim(\widetilde{\mathcal{U}}) \leq \dim(\mathcal{V}) + 1$. Moreover, note that $|\mathcal{U}^{=t}| = |\widetilde{\mathcal{U}}^{=t^*}|$. Using Gaussian elimination, we can find a basis find a basis $b_1, \ldots, b_k$ for $\mathcal{U}$ such that $N(b_1) < N(b_2) < \ldots < N(b_k)$, where $k = \dim(\mathcal{U})$ and $N(b) := \min_i\{i : b_i \neq 0\}$. Moreover (again via Gaussian elimination), we can ensure that $b_i$ is the only basis vector with a 1 in entry $N(b_i)$. Therefore, any vector in $\mathcal{U}$ involving more than $t^*$ basis vectors must have more than $t^*$ nonzero entries. Therefore, we have that

$$|\mathcal{U}^{=t}| \leq |\mathcal{U}^{\leq t}| = |\widetilde{\mathcal{U}}^{\leq t^*}| \leq \sum_i^{t^*} \binom{\dim(\widetilde{\mathcal{U}})}{i} \leq \binom{\dim(\mathcal{V}) + 1}{\leq t^*}. \quad \square$$

## B.4 Lemma 4.14

*Proof of Lemma 4.14.* We will prove the statement by induction on $\ell$. Setting $S_1$ to be $\{\gamma_1, \ldots, \gamma_{n-2C}\}$ guaranteed by Claim 4.1 such that $|(\gamma + \mathcal{V}^\perp)^{=1}| = 1$ for all $\gamma \in S_1$ proves the base case when $\ell = 1$.

Now suppose we have some $S_{\ell-1}$ that satisfies the conditions in the lemma. We will pick $S_\ell \subseteq S_{\ell-1}$ that satisfies condition (2) for $t = \ell$, and argue that the number that do not satisfy the condition is at most $C$. Indeed, suppose towards a contradiction that $|S_{\ell-1} \setminus S_\ell| \geq C + 1$. Let $J \subseteq S_{\ell-1} \setminus S_\ell$ be any subset of size $C + 1$ and $H := \bigcup_{\gamma \in J} (\gamma + \mathcal{V}^\perp)^{=\ell}$. Since $S_\ell \subseteq \mathcal{W}$, we can say by Fact 2.4 that the sets $(\gamma + \mathcal{V}^\perp)_{\gamma \in S_\ell}$ are all mutually disjoint and therefore,

$$|H| = \sum_{\gamma \in J} |(\gamma + \mathcal{V}^\perp)^{=\ell}| > 2(C+1)\binom{2C+1}{\ell-1} = \ell\binom{2C+2}{\ell}.$$

However, $H \subseteq (\mathsf{span}(\{\gamma : \gamma \in J\} \cup V^\perp))^{=\ell}$. Since, $\dim(\mathsf{span}(\{\gamma : \gamma \in J\} \cup V^\perp)) \leq |J| + C = 2C + 1$, by Claim 4.13 it must be that $|H| \leq \binom{2C+2}{\leq \ell} \leq \ell\binom{2C+2}{\ell}$, where the inequality holds for all $\ell \leq (2C+2)/2 = C+1$. This is a contradiction, and we conclude that $|S_{\ell-1} \setminus S_\ell| \leq C$. □



ECCC ISSN 1433-8092
https://eccc.weizmann.ac.il