Electronic Colloquium on Computational Complexity, Report No. 41 (2022)

Boolean functions with small approximate spectral norm

Tsun Ming Cheung * Hamed Hatami † Rosie Zhao ‡ Itai Zilberstein §

Abstract

The sum of the absolute values of the Fourier coefficients of a function $f : \mathbb{F}_2^n \to \mathbb{R}$ is called the spectral norm of f. Green and Sanders' quantitative version of Cohen's idempotent theorem states that if the spectral norm of $f : \mathbb{F}_2^n \to \{0, 1\}$ is at most M, then the support of f belongs to the ring of sets generated by at most $\ell(M)$ cosets, where $\ell(M)$ is a constant that only depends on M.

We prove that the above statement can be generalized to *approximate* spectral norms if and only if the support of f and its complement satisfy a certain arithmetic connectivity condition. In particular, our theorem provides a new proof of the quantitative Cohen's theorem for $\mathbb{F}_2^n$.

1 Introduction

Let $G = \mathbb{F}_2^n$ be the Boolean cube, and $\widehat{G} \cong \mathbb{F}_2^n$ be its Pontryagin dual. For a character $\chi \in \widehat{G}$, the corresponding Fourier coefficient of a function $f : G \to \mathbb{R}$ is defined as

$$\widehat{f}(\chi) := \mathop{\mathbb{E}}_{x \in G}[f(x)\chi(x)].$$

The sum of the absolute values of the Fourier coefficients is called the *algebra norm* or *spectral norm* of f, and is denoted by

$$\|f\|_A := \|\widehat{f}\|_1 = \sum_{\chi \in \widehat{G}} |\widehat{f}(\chi)|.$$

The term *algebra* norm is explained by the inequality $\|fg\|_A \leq \|f\|_A \|g\|_A$. This norm arises naturally in theoretical computer science in connection to learning theory, and it has been studied for several complexity classes of Boolean functions [STV17, KM93, TWXZ13, GTW21, Tal17, MRT19]. These studies are often motivated by the existence of efficient learning algorithms for the classes of Boolean functions that have small algebra norms [KM93]. Furthermore, in recent years, tail bounds in the Fourier L_1 norm have also become essential in constructing pseudo-random generators [CHHL19, RSV13, FK18] and separating quantum and classical computation [RT19, Tal20, BS21].

For a set $A \subseteq G$, let $\mathbf{1}_A$ denote the indicator function of A. The set of Boolean functions that satisfy $\|f\|_A = O(1)$ are fully characterized by an important theorem in harmonic analysis, Cohen's idempotent theorem. The base case of this characterization is described through the following simple proposition that characterizes the set of contractive Boolean functions, i.e. those with $\|f\|_A \leq 1$.

*School of Computer Science, McGill University. tsun.ming.cheung@mail.mcgill.ca.

†School of Computer Science, McGill University. hatami@cs.mcgill.ca. Supported by an NSERC grant.

‡School of Computer Science, McGill University. rosie.zhao@mail.mcgill.ca.

§School of Computer Science, McGill University. itai.zilberstein@mail.mcgill.ca.

ISSN 1433-8092

Proposition 1 (Folklore, see [GS08a, Proposition 1.2]). *A nonempty set $A \subseteq G$ satisfies $\|\mathbf{1}_A\|_A \leq 1$ if and only if A is a coset of a subgroup of G, in which case $\|\mathbf{1}_A\|_A = 1$.*

It is possible to apply set operations to cosets and construct more sophisticated Boolean functions with algebra norm $O(1)$. Recall that a *ring of sets* on G is a subset of $\mathcal{P}(G)$ that includes G, and is closed under complements and (finite) intersections and (finite) unions. We say $A \subseteq G$ has *coset complexity* at most ℓ if it belongs to the ring of sets generated by at most ℓ cosets.

It is straightforward to show that if A has coset complexity at most ℓ, then $\|\mathbf{1}_A\|_A \leq 3^\ell$ (see Lemma 1). The quantitative version of Cohen's idempotent theorem states that the converse is essentially true.

Theorem 1 (Quantitative Cohen's theorem for $\mathbb{F}_2^n$ [Coh60, GS08a]). *If $A \subseteq G$ satisfies $\|\mathbf{1}_A\|_A \leq M$, then A belongs to the ring of sets generated by at most ℓ cosets where $\ell = \ell(M)$ depends only on M.*

The term "idempotent" essentially refers to the assumption that $f = \mathbf{1}_A$ is Boolean, which is equivalent to $f^2 = f$. We should remark that Cohen's original theorem [Coh60] is concerned with locally compact Abelian groups of infinite size. The quantitative version of the theorem, which is also applicable to finite groups, is due to Green and Sanders [GS08b, GS08a]. We will discuss this in more detail in Section 1.1.

Approximate algebra norm: Our goal is to extend Theorem 1 to the set of Boolean functions with small *approximate algebra norms*. For any error parameter $\epsilon > 0$, the *ϵ-approximate algebra norm* of $f : G \to \mathbb{R}$ is defined as

$$\|f\|_{A,\epsilon} := \inf\{\|g\|_A \ : \ \|f - g\|_\infty \leq \epsilon\}.$$

We remind the reader that despite what the notation might suggest, $\|\cdot\|_{A,\epsilon}$ is not a norm. We will always assume $\epsilon \in [0, \frac{1}{2})$ as for Boolean functions, the range $\epsilon \geq \frac{1}{2}$ is trivial and uninteresting.

Approximate norms, in general, are important in the theory of computation as they capture various notions of randomized query and communication complexity. For example, approximate algebra norms are closely related to the notions of randomized parity decision tree complexity and the randomized communication complexity of the so-called XOR-lifts[1]. More precisely, the gaps between these parameters are at most exponential, with no dependencies on the dimension n. We refer the reader to [STV17, KLMY21, HHH21] for more details.

Boolean functions that have small approximate algebra norms have been studied by Méla [MŐ82] and Host, Méla, and Parreau [HMP86] under the concept of *ϵ-quasi-idempotents*. Méla proved in [MŐ82] (see also [GS08a, Proposition 7.1]) that the assertion of Cohen's idempotent theorem remains true under the weaker assumption that $\|\mathbf{1}_A\|_{A,\epsilon} \leq M$ provided that $M \leq c|\log\epsilon|$ where c is a universal constant, and the logarithm, here and throughout the paper, is in base 2.

On the other hand, Hamming balls of radius 1 show that the requirement $M \leq c|\log\epsilon|$ for some universal constant c is necessary as for $B_k = \{x \in \{0,1\}^k \ : \ \sum_{i=1}^k x_i \leq 1\}$ and $0 < \epsilon < \frac{1}{2}$, we have (see Lemma 3)

$$\|\mathbf{1}_{B_k}\|_A \geq \frac{\sqrt{k}}{2} \qquad \text{and} \qquad \|\mathbf{1}_{B_k}\|_{A,\epsilon} \leq 5|\log\epsilon|.$$

[1]The XOR-lift of a function $f : \mathbb{F}_2^n \to \mathbb{R}$ is the function $f^\oplus : \mathbb{F}_2^n \times \mathbb{F}_2^n \to \mathbb{R}$ defined as $(x, y) \mapsto f(x + y)$.

These bounds show that the functions $\mathbf{1}_{B_k}$ can have arbitrarily large coset complexity, while their approximate algebra norm is uniformly bounded by a constant that depends only on ϵ. Therefore, any potential extension of Cohen's theorem to approximate algebra norm needs to overrule the possibility of containing "affine copies" of arbitrarily large $\mathbf{1}_{B_k}$. This is achieved through the notion of affine connectivity.

Definition 1 (Affine connectivity)**.** *We say that a set $A \subseteq G$ is k-affine connected if $a_0, a_0 + a_1, \ldots, a_0 + a_k \in A$ implies that at least one of the following holds:*

- *The vectors $a_0, a_0 + a_1, \ldots, a_0 + a_k$ are linearly dependent;*
- *There exists $T \subseteq \{1, \ldots, k\}$ with $|T| \geq 2$ such that $a_0 + \sum_{i \in T} a_i \in A$.*

Remark 1. Definition 1 means that no restriction of A to a k-dimensional coset is a copy of B_k. Note also that by the change of variables $b_0 = a_0$ and $b_i = a_0 + a_i$ for $i = 1, \ldots, k$, one can equivalently define k-affine connectivity as the condition that for every linearly independent $b_0, b_1, \ldots, b_k \in A$, there exists $S \subseteq \{0, 1, \ldots, k\}$ such that $|S| > 1$ is odd and $\sum_{i \in S} b_i \in A$.

Our contribution: We prove that if $\|\mathbf{1}_A\|_{A,\epsilon}$ is small, then $\|\mathbf{1}_A\|_A$ is small if and only if both A and A^c are k-affine connected for a small k.

Theorem 2 (Main theorem)**.** *For every $k, M \in \mathbb{N}$ and $\epsilon \in [0, \frac{1}{2})$, there exists $\ell = \ell(k, M, \epsilon) \in \mathbb{N}$ such that the following holds. If $A \subseteq G$ satisfies $\|\mathbf{1}_A\|_{A,\epsilon} \leq M$, then*

(i) *either A or A^c is not k-affine connected, in which case $\|\mathbf{1}_A\|_A \geq \frac{\sqrt{k}}{2}$;*

(ii) *or both A and A^c are k-affine connected, in which case A has coset complexity at most ℓ. In particular, $\|\mathbf{1}_A\|_A \leq 3^\ell$.*

Remark 2. Our proof results in the bound $\ell(k, M, \epsilon) \leq \mathrm{Tower}_2\left(O(\frac{Mk}{1-2\epsilon})\right)$, where $\mathrm{Tower}_2(m)$ denotes the tower of exponentiation with base 2 and height m.

Remark 3. Theorem 2 implies Theorem 1: If $\|\mathbf{1}_A\|_A \leq M$, then by Theorem 2 (i), both A and A^c are $O(M^2)$-affine connected, and thus one can apply Theorem 2 (ii) to conclude Theorem 1. However, the best known upper bound [San19] for Theorem 1 is only $2^{O(M^{3+o(1)})}$, while this proof results in a tower-type bound.

Remark 4. The k-affine connectedness by itself does not imply that $\|\cdot\|_A$ is small, and thus it is also essential that in Theorem 2, we assume $\|\mathbf{1}_A\|_{A,\epsilon} \leq M$. For example, consider the quadratic function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ for even n defined as $f(x) = x_1x_2 + x_3x_4 + \ldots + x_{n-1}x_n$ where the additions are in $\mathbb{F}_2$. Since f is a quadratic function, it satisfies

$$\sum_{S \subseteq [4]} f\left(a_0 + \sum_{i \in S} a_i\right) \equiv 0. \tag{1}$$

for all $a_0, \ldots, a_4 \in \mathbb{F}_2^n$. It follows from Equation (1) that $\mathrm{supp}(f)$ and $\mathrm{supp}(f)^c$ are both 4-affine connected. On the other hand, since f is of high rank, by standard Gauss sum estimates [GW11, Lemma 3.1] or a direct calculation, one can easily verify that $\|f\|_A = \Theta(2^{n/2})$.

1.1 Historical remarks and the general picture

Proposition 1 is a special case of the Kawada-Itô theorem [KI40, Theorem 3], which dates back to 1940. Kawada and Itô characterized idempotent *probability measures* on compact groups as the normalized Haar measures of compact subgroups. This theorem was rediscovered independently by Wendel [Wen54] in the context of harmonic analysis. Later, Rudin [Rud59a, Rud59b], trying to extend this result to all idempotent measures on locally compact Abelian groups, showed that any such measure is concentrated on a compact subgroup. Finally, Cohen [Coh60], building on the works of Helson [Hel53] and Rudin [Rud59a], obtained a full description of idempotent measures on locally compact Abelian groups. Numerous extensions and refinements of Cohen's theorem have been discovered since [Lef72, Hos86, GS08b, Run07, San11, San20, San21].

To state Cohen's original theorem in full generality, we need a few definitions: Let G be a locally compact group, and let $\widehat{G}$ be its Pontryagin dual (which is also a locally compact Abelian group). Let $\mathrm{M}(G)$ be the algebra of all bounded regular Borel measures on G, where multiplication is defined by convolution. Let $B(\widehat{G})$ denote the Fourier–Stieltjes algebra of $\widehat{G}$, which is the set of all $\widehat{\mu} : \widehat{G} \to \mathbb{C}$ for all $\mu \in \mathrm{M}(G)$ endowed with the norm $\|\widehat{\mu}\|_{B(\widehat{G})} := \|\mu\|$. This norm is well-defined since the choice of μ is unique. If $\widehat{G}$ is a *finite* Abelian group, then $B(\widehat{G})$ is the set of all functions on $\widehat{G}$, and $\|\cdot\|_{B(\widehat{G})}$ coincides with the algebra norm: $\|f\|_{B(\widehat{G})} = \|f\|_A$.

Note that if $\mu \in \mathrm{M}(G)$ is idempotent (i.e. $\mu * \mu = \mu$), then $\widehat{\mu}^2 = \widehat{\mu}$, so $\widehat{\mu}(\chi) \in \{0, 1\}$ for all $\chi \in \widehat{G}$. Hence the problem of characterizing all idempotent measures in $\mathrm{M}(G)$ is equivalent to finding all subsets $A \subseteq \widehat{G}$ with $\mathbf{1}_A \in B(\widehat{G})$.

We say that a set $A \subseteq G$ has *coset complexity* at most $\ell \in \mathbb{N}$ if it belongs to the ring of sets generated by at most ℓ *open* cosets. The coset complexity of A is defined to be infinite if no such ℓ exists.

Theorem 3 (Cohen's idempotent theorem [Coh60])**.** *Let G be a locally compact Abelian group. A set $A \subseteq G$ satisfies $\mathbf{1}_A \in B(G)$ if and only if the coset complexity of A is finite.*

We refer the interested readers to [Rud90, Chapter 3] for more details. Cohen's theorem left open whether the coset complexity of A is *uniformly* bounded from above by a function of $\|\mathbf{1}_A\|_{B(G)}$. Moreover, it gave no information for finite groups. Five decades later, Green and Sanders [GS08b, GS08a], using modern tools from additive combinatorics, proved a stronger quantitative version of Cohen's theorem that resolved the uniformity question. Their result can be applied to finite groups as well. It states that if $\|\mathbf{1}_A\|_{B(G)} \leq M$, then the coset complexity of A is at most $\ell(M)$, where $\ell(\cdot)$ is a universal function that does not depend on the choice of the underlying group G. They first proved the special case of this theorem for the groups $\mathbb{F}_2^n$, and afterwards generalized it to all locally compact Abelian groups in [GS08b]. The bounds obtained in these two papers were later improved by Sanders [San19, San20].

Cohen's theorem has been generalized to non-Abelian locally compact groups in [Lef72, Hos86], and the quantitative version of the non-Abelian idempotent theorem was also established by Sanders [San11].

It seems conceivable that with a proper generalization of the notion of affine connectivity, Theorem 2 can similarly be generalized to all locally compact Abelian groups, or even all locally compact groups. We defer this to future research.

1.2 Notation

For a positive integer n, we use $[n]$ to denote $\{1, \ldots, n\}$. We denote the complement of a set S by S^c. We will use the standard asymptotic notations of $O(\cdot), \Omega(\cdot), \Theta(\cdot), o(\cdot), \omega(\cdot)$. Sometimes, we shall add subscripts to these notations to indicate that the constants involved depend on these parameters. For example, $O_\epsilon(1)$ means bounded from above by a constant that depends only on ϵ.

For integers $s > 0, t \geq 0$, let $\mathrm{Tower}_s(t)$ be defined recursively as $\mathrm{Tower}_s(t) = s^{\mathrm{Tower}_s(t-1)}$ with the base case $\mathrm{Tower}_s(0) = 1$. For $s > 1$, let $\log_s^*(m)$ be the smallest integer $t \geq 0$ such that $\mathrm{Tower}_s(t) \geq m$.

Throughout the article, G always denotes $\mathbb{F}_2^n$. We consider G as both a group and a vector space over $\mathbb{F}_2$. We denote by $\mathbf{0} \in \mathbb{F}_2^n$ the zero vector. For $i = 1, \ldots, n$, let $\mathbf{e}_i \in G$ denote the i-th standard vector.

The *additive energy* of a set $A \subseteq G$ is defined as

$$\mathrm{E}(A) = |\{(a_1, a_2, a_3, a_4) \in A^4 \ : \ a_1 + a_2 = a_3 + a_4\}| = |G|^3 \sum_{a \in G} |\widehat{\mathbf{1}_A}(a)|^4. \tag{2}$$

For sets $A, B \subseteq G$ and $c \in G$, we define

$$A + c = \{a + c \ : \ a \in A\},$$

and

$$A + B = \{a + b \ : \ a \in A, \ b \in B\}.$$

We often identify $G \cong \widehat{G}$ via the bijection that maps $a \in G$ to the character $\chi_a(x) := (-1)^{a^t x}$. We will use the notation $\widehat{f}(a) := \widehat{f}(\chi_a)$.

The *convolution* of two functions $f_1, f_2 : G \to \mathbb{R}$ is defined as

$$f_1 * f_2(x) = \mathop{\mathbb{E}}_{y \in G}[f_1(x - y) f_2(y)].$$

For a subgroup $W \subseteq G$, we define $\mu_W : G \to \mathbb{R}$ as $\mu_W : x \mapsto \frac{|G|}{|W|} \mathbf{1}_W(x)$ so that

$$f * \mu_W(x) = \mathop{\mathbb{E}}_{y \in G}[f(x - y)\mu_W(y)] = \mathop{\mathbb{E}}_{y \in W + x}[f(y)] = \mathbb{E}[f | W + x].$$

The *annihilator* of W is defined as

$$W^\perp = \{r \in \widehat{G} \mid r^t a = 0 \text{ for all } a \in W\}.$$

Note that convolution with μ_W corresponds to the projection of the Fourier expansion to $W^\perp$:

$$f * \mu_W(x) = \sum_{a \in W^\perp} \widehat{f}(a) \chi_a(x). \tag{3}$$

We call a subset of G a *coset* if it is a coset of some subgroup of G. For a subgroup $W \subseteq G$, we identify the quotient space $G/W \equiv W^\perp$. We denote the cosets of W by $W + a$, and whenever such a notation is used, it is always assumed that W is a subgroup and $a \in G$.

Given a function $f : G \to \mathbb{R}$ and a coset $W + a \subseteq G$, we often identify $f|_{W+a}$ with the function on W, defined as $w \mapsto f(w + a)$. Note that for $w \in W$, we have

$$f(w + a) = \sum_{b \in W} \sum_{c \in W^\perp} \widehat{f}(b + c) \chi_{b+c}(w + a) = \sum_{b \in W} \left(\chi_b(a) \sum_{c \in W^\perp} \widehat{f}(b + c) \chi_c(a) \right) \chi_b(w).$$

Hence with this notation

$$\widehat{f|_{W+a}}(b) = \chi_b(a) \sum_{c \in W^\perp} \widehat{f}(b+c)\chi_c(a) \qquad \text{for all } b \in W. \tag{4}$$

Finally, sometimes it will be more convenient to work with a slight variant of the algebra norm that excludes the principal Fourier coefficient. For a function $g : G \to \mathbb{R}$, define

$$\|g\|_{\mathbb{A}} := \|g - \mathbb{E}[g]\|_A = \sum_{\chi \neq 0} |\widehat{g}(\chi)|.$$

For a function defined on a subgroup $W \subseteq G$, we write $\mathbb{A}(W)$ in the subscript to emphasize the domain of the function.

1.3 Proof overview

Before giving an overview of the proof of Theorem 2, we discuss Green and Sanders' proof of Theorem 1. It follows a similar high-level approach as Cohen's proof [Coh60], but uses results from additive combinatorics to obtain effective bounds.

Green and Sanders' proof of Theorem 1: The proof uses a strong induction which requires generalizing the statement from Boolean functions to *almost integer-valued functions*. For $\epsilon \in [0, \frac{1}{2})$, a function $f : G \to \mathbb{R}$ is called ϵ-integer-valued if $\|f - h\|_\infty \leq \epsilon$ for an integer-valued function h.

Let $f : G \to \mathbb{R}$ be an ϵ-integer-valued function with $\|f\|_A \leq M$. By Equation (3), for every subgroup $W \subseteq G$, we have

$$\|f\|_A = \|f * \mu_W\|_A + \|f - f * \mu_W\|_A. \tag{5}$$

The main step of the proof is establishing the existence of a subgroup W and a small $\delta > 0$ such that

(i) $\|f * \mu_W - \sum_{i=1}^{c} \pm \mathbf{1}_{W+a_i}\|_\infty \leq \epsilon + \delta$ where $c = O_{\delta,M,\epsilon}(1)$;

(ii) $\|f * \mu_W\|_A \geq \frac{1}{2}$.

By (i) $f * \mu_W$ is approximated by a sum that involves only $O_{\delta,M,\epsilon}(1)$ cosets. On the other hand, since f and $f * \mu_W$ are ϵ- and $(\epsilon + \delta)$-integer-valued respectively, their difference $f - f * \mu_W$ is $(2\epsilon + \delta)$-integer-valued. Moreover, by (ii) and Equation (5), we have $\|f - f * \mu_W\|_A \leq M - \frac{1}{2}$, and with this decrease in the algebra norm, we can apply the induction hypothesis to $f - f * \mu_W$ to complete the proof.

The $M = O(|\log \epsilon|)$ requirement: In order to decrease M to $M - \frac{1}{2}$, we increased the "error parameter" from ϵ to $2\epsilon + \delta$. Repeating this process inductively for $2M$ steps will decrease the algebra norm to the base case $\|f\|_A \leq \frac{1}{2}$. However, for a meaningful approximation, we need to ensure that the error parameter never exceeds $\frac{1}{2}$. Since the error parameter is more than doubled at each step, it is essential to require $\epsilon = 2^{-\Omega(M)}$ initially.

The requirement that $\epsilon = 2^{-\Omega(M)}$ is not just an artifact of Cohen's proof. Méla [Mél82] constructed an example on a certain Abelian group which illustrates that this requirement is necessary. In Lemma 3 we show that an analogous result holds for $\mathbb{F}_2^n$.

Overview of proof of Theorem 2: Theorem 2 assumes that A and A^c are k-affine connected and $\|\mathbf{1}_A\|_{A,\epsilon} \leq M$. These two assumptions suffice to guarantee the existence of a subgroup W with certain desired properties, similar to those used by Green and Sanders:

- Affine connectivity implies the existence of a coset $V+a$ such that $|V+a| \approx |A| \approx |(V+a)\cap A|$.
- The assumption $\|\mathbf{1}_A\|_{A,\epsilon} \leq M$ allows us to "regularize" V to a large subgroup $W \subseteq V$ such that every coset of W is either almost contained in A or has almost no intersection with A.

These parts of the proof closely follow Green and Sanders' proof of Theorem 1.

The primary issue preventing us from further emulating the proof of Theorem 1 is that $f - f * \mu_W$ is only $(2\epsilon + \delta)$-integer valued. Since we cannot afford a doubling in the error parameter, we depart from Cohen's approach. Instead, we employ a completely new induction that focuses on A's connectivity.

Let us reformulate the definition of affine connectivity in a slightly different language. Let $\mathcal{X} := \{\mathbf{0}\} \subseteq G$, and $r := k + 1$. By Remark 1, A is k-affine connected if and only if for every $x_1, \ldots, x_r \in A \setminus \mathcal{X}$ one of the following holds:

(i) There exists a set $S \subseteq [r]$ such that $\sum_{i\in S} x_i \in \mathcal{X}$.

(ii) There exists a set $S \subseteq [r]$ such that $|S| > 1$ is odd and $\sum_{i\in S} x_i \in A \setminus \mathcal{X}$.

The proof of Theorem 2 is by induction on r and M. Throughout the argument, $\mathcal{X}$ always remains a union of $O(1)$ cosets. The coset complexity of $A \cap \mathcal{X}$ can be shown to be small by applying an induction on M to A's restrictions to each individual coset in $\mathcal{X}$.

The main component of the proof is to establish that it suffices to add $O(1)$ cosets to $\mathcal{X}$ to reduce r. We present the details of this double induction in Section 3.2.

2 Basic facts

In this section, we state a few standard facts that will be used later in the proof of Theorem 2. The following lemma shows that if the coset complexity of A is small, then $\mathbf{1}_A$ can be expressed as a ± 1-linear combination of indicator functions of a few cosets, and thus $\|\mathbf{1}_A\|_A = O(1)$.

Lemma 1. *If $A \subseteq G$ has coset complexity at most ℓ, then $\mathbf{1}_A$ can be expressed as*

$$\mathbf{1}_A = \sum_{i=1}^{t} \epsilon_i \mathbf{1}_{W_i + a_i}, \tag{6}$$

for cosets $W_i + a_i \subseteq G$, $\epsilon_i \in \{-1, 1\}$, and $t \leq 3^\ell$. In particular, $\|\mathbf{1}_A\|_A \leq 3^\ell$.

Proof. Suppose A belongs to the ring of sets generated by $V_1 + b_1, \ldots, V_\ell + b_\ell$. Each atom of this ring is of the form

$$\bigcap_{i\in S} (V_i + b_i) \cup \bigcap_{j\in S^c} (V_j + b_j)^c,$$

for $S \subseteq [\ell]$. Notice that $\mathbf{1}_{(V_j+b_j)^c} = 1 - \mathbf{1}_{V_j+b_j}$, and the intersection of two cosets is a coset. Therefore we can express the indicator function of such an atom as a sum of ± 1-linear combination

of indicator functions of at most $2^{|S^c|} = 2^{\ell-|S|}$ cosets. Summing over all the atoms in A, we conclude that $\mathbf{1}_A$ can be expressed as such a sum with the number of terms at most

$$\sum_{S\subseteq[\ell]} 2^{\ell-|S|} = (1+2)^\ell = 3^\ell.$$

□

Our next goal is to estimate the algebra norm and the approximate algebra norm of the Hamming ball of radius 1. Our proof of the upper bound on the approximate algebra norm of $\mathbf{1}_{B_k}$ closely follows the argument of Méla [MÉ82].

We first need to state a simple lemma from approximation theory. The proof uses Chebyshev's classical characterization of best approximation by polynomials. Let $C([a,b])$ denote the set of all real-valued continuous functions on the interval $[a,b]$ equipped with the L_∞ norm (i.e., supremum of absolute value). A k-dimensional subspace $V \subseteq C([a,b])$ is said to satisfy *Chebyshev's condition* if every function in V has at most $k-1$ *distinct* zeros in $[a,b]$ (See [Riv90, Definition 2.9]).

Chebyshev's classical theorem states that if V satisfies Chebyshev's condition, and $S \subseteq [a,b]$ is a closed set (e.g., finite), then $v_0 \in V$ is the best L_∞-approximation on S of a given $f \in C(S) \setminus V$ if and only if the following holds: there exist points $x_1 < \ldots < x_{k+1}$ in S such that $|f(x_i) - v_0(x_i)| = \|f - v_0\|_{L_\infty([a,b])}$ for all i, and

$$f(x_i) - v_0(x_i) \qquad \text{for } i = 1, \ldots, k+1,$$

alternate in sign [Riv90, Theorem 2.10].

Lemma 2. *Let $m > 1$ be an integer, and let $\eta_i := \cos\left(\frac{m-i}{2m-1}\pi\right)$ for $i \in [m]$. There exists a function $\sigma : [m] \to \mathbb{R}$ such that*

(i) $\sum_{i=1}^m \eta_i \sigma(i) = 1$,

(ii) $\sum_{i=1}^m \eta_i^{2k-1} \sigma(i) = 0$ *for every* $2 \le k \le m$,

(iii) $\sum_{i=1}^m |\sigma(i)| \le 2m-1$.

Proof. Let V be the linear span of $x^3, x^5, \ldots, x^{2m-1}$ over the reals. Every function in V is an odd polynomial and thus has at most $m-2$ zeros in $(0,\infty)$. Since $\dim(V) = m-1$, V satisfies Chebyshev's condition on any interval $[a,b]$ for $0 < a < b < \infty$.

Let $T_{2m-1}(x) = a_1 x + a_3 x^3 + \ldots + a_{2m-1} x^{2m-1}$ be the Chebyshev polynomial of the first kind of degree $2m-1$. Since $a_1 = (-1)^m(2m-1)$ (see [Riv90, Section 1.1]), the function

$$q(x) := x + \frac{(-1)^{m-1}}{2m-1} T_{2m-1}(x)$$

is in V. We claim that $q(x)$ is the best L_∞-approximation on $S = \{\eta_1, \ldots, \eta_m\}$ of $f(x) = x$ by functions in V. By the trigonometric definition of the Chebyshev polynomial, it can be seen that $0 < \eta_1 < \ldots < \eta_m = 1$ are the extrema points of T_{2m-1} and the signs of $f(x) - q(x) = \frac{(-1)^m}{2m-1} T_{2m-1}(x)$ on these points alternate. Hence, we can apply Chebyshev's theorem and conclude that $q(x)$ is the best L_∞-approximation of $f(x) = x$ on $\{\eta_1, \ldots, \eta_m\}$ by functions in V. Since

$$\max_{i\in[m]} |q(\eta_i) - \eta_i| = \frac{1}{2m-1} \max_{i\in[m]} |T_{2m-1}(\eta_i)| = \frac{1}{2m-1},$$

we conclude that

$$\min_{c_3,c_5,\ldots,c_{2m-1}\in\mathbb{R}} \max_{i\in[m]} \left|c_3\eta_i^3 + c_5\eta_i^5 + \ldots + c_{2m-1}\eta_i^{2m-1} - \eta_i\right| \geq \frac{1}{2m-1}.$$

Hence, by linear programming duality, the solution to the following optimization problem is at least $\frac{1}{2m-1}$, which after normalization yields the desired σ.

$$\begin{aligned} \max \quad & \sum_{i=1}^m \sigma(i)\eta_i \\ \text{subject to} \quad & \sum_{i=1}^m \sigma(i)\eta_i^{2k-1} = 0 \text{ for every } 2 \leq k \leq m \\ & \sum_{i=1}^m |\sigma(i)| \leq 1. \end{aligned}$$

□

In the next lemma, we show a separation in the algebra norm and the approximate algebra norm of the Hamming ball of radius 1. We show that the algebra norm tends to infinity as k grows, while the approximate algebra norm remains bounded by a constant that depends only on ϵ.

Lemma 3. *Let $B_k \subseteq \{0,1\}^k$ be the set of all $x \in \{0,1\}^k$ with $\sum_{i=1}^k x_i \leq 1$, and let $\epsilon \in (0, \frac{1}{2})$. We have*

$$\frac{\sqrt{k}}{2} \leq \|\mathbf{1}_{B_k}\|_A \leq \sqrt{k+1} \qquad \textit{and} \qquad \|\mathbf{1}_{B_k}\|_{A,\epsilon} \leq 5|\log\epsilon|.$$

Proof. The upper bound $\|\mathbf{1}_{B_k}\|_A \leq 2^{k/2}\|\mathbf{1}_{B_k}\|_2 = \sqrt{k+1}$ is immediate from Cauchy-Schwarz inequality and Parseval's identity. To prove the lower bound on $\|\mathbf{1}_{B_k}\|_A$, note that for $y \in G$, we have $\widehat{\mathbf{1}_{B_k}}(y) = 2^{-k}\left(1 + \sum_{i=1}^k (-1)^{y_i}\right)$. Hence

$$\|\mathbf{1}_{B_k}\|_A = \mathop{\mathbb{E}}_{z\in\{-1,1\}^k}\left[\left|1 + \sum_{i=1}^k z_i\right|\right] \geq \mathop{\mathbb{E}}_{z\in\{-1,1\}^k}\left[\left|\sum_{i=1}^k z_i\right|\right] \geq \frac{1}{2}\left(\mathop{\mathbb{E}}_{z\in\{-1,1\}^k}\left[\left|\sum_{i=1}^k z_i\right|^2\right]\right)^{1/2} = \frac{\sqrt{k}}{2},$$

where the first inequality uses the fact that $\sum_{i=1}^k z_i$ is a symmetric random variable, and the second inequality is an application of Khintchine's inequality.

To prove the upper bound on $\|\mathbf{1}_{B_k}\|_{A,\epsilon}$, for $a \in \mathbb{F}_2^k$ and $s \in [-1,1]$, define

$$0 \leq \widehat{g}_s(a) := 2^{-k}\left(\prod_{i=1}^k (1 + s(-1)^{a_i})\right) = 2^{-k}\sum_{x\in\mathbb{F}_2^k} s^{|x|}\chi_a(x),$$

where $|x|$ denotes the Hamming weight of x. Let $g_s(x) = \sum_{a\in\mathbb{F}_2^k} \widehat{g}_s(a)\chi_a(x)$. It is also straightforward to verify that $g_s(x) = s^{|x|}$. By positivity of the Fourier coefficients $\widehat{g}_s(a)$, we have

$$\|g_s\|_A = g_s(0) = 1 \qquad \text{for all } s \in [-1,1]. \tag{7}$$

Moreover, substituting $s = \epsilon$ gives

$$\left\| \frac{g_\epsilon - (1-\epsilon)\mathbf{1}_{\{\mathbf{0}\}}}{\epsilon} - \mathbf{1}_{B_k} \right\|_\infty \leq \epsilon.$$

This shows $\|\mathbf{1}_{B_k}\|_{A,\epsilon} \leq 2/\epsilon$, but this upper bound can be further strengthened.

For $s \in [0,1]$, define $h_s : \mathbb{F}_2^k \to \mathbb{R}$ as

$$h_s(x) := \frac{g_s(x) - g_{-s}(x)}{2} = \begin{cases} s^{|x|} & \text{if } |x| \text{ is odd} \\ 0 & \text{if } |x| \text{ is even} \end{cases}. \tag{8}$$

By Equation (7),

$$\|h_s\|_A \leq \frac{\|g_s\|_A + \|g_{-s}\|_A}{2} = 1 \qquad \text{for all } s \in [0,1]. \tag{9}$$

Let $\sigma : [m] \to \mathbb{R}$ and $\eta_i = \cos\left(\frac{m-i}{2m-1}\pi\right)$ for $i \in [m]$ be as defined in Lemma 2, where $m = \lceil |\log \epsilon| \rceil$. Define $h : \mathbb{F}_2^k \to \mathbb{R}$ as

$$h(x) = 2\sum_{i=1}^{m} \sigma(i) h_{\frac{\eta_i}{2}}(x).$$

Note that

- If $|x|$ is even, then by Equation (8), $h(x) = 0$.
- If $|x| = 1$, then $h_s(x) = s$, and thus by Lemma 2 (i), $h(x) = 2\sum_{i=1}^{m} \frac{\eta_i \sigma(i)}{2} = 1$.
- If $|x| \leq 2m-1$ is odd, then by Lemma 2 (ii),

$$h(x) = 2\sum_{i=1}^{m} \sigma(i) \left(\frac{\eta_i}{2}\right)^{|x|} = 0.$$

- If $|x| \geq 2m+1$ is odd, then by Lemma 2 (iii) and the triangle inequality,

$$h(x) = 2\sum_{i=1}^{m} \sigma(i) \left(\frac{\eta_i}{2}\right)^{|x|} \leq 2^{1-|x|} \sum_{i=1}^{m} |\sigma(i)| \leq 2^{-2m}(2m-1) \leq 2^{-m} \leq \epsilon.$$

Hence $\left\|(h + \mathbf{1}_{\{\mathbf{0}\}}) - \mathbf{1}_{B_k}\right\|_\infty \leq \epsilon$. Finally note that by Equation (9), we have

$$\|h\|_A \leq 2\sum_{i=1}^{m} |\sigma(i)| \|h_{\frac{\eta_i}{2}}(x)\|_A \leq 2\sum_{i=1}^{m} |\sigma(i)| \leq 2(2m-1) \leq 4|\log \epsilon|.$$

We conclude that

$$\|\mathbf{1}_{B_k}\|_{A,\epsilon} \leq 1 + \|h\|_A \leq 5|\log \epsilon|.$$

□

Lemma 3, combined with Lemma 1, shows that the coset complexity of $\mathbf{1}_{B_k}$ is $\Omega(\log k)$, while its ϵ-approximate algebra norm is at most $5|\log \epsilon|$. This illustrates that the assertion of Theorem 1 is not necessarily true under the weaker assumption that $\|\mathbf{1}_A\|_{A,\epsilon} \leq M$ if $M \geq 5|\log(\epsilon)|$.

3 Proof of Theorem 2

We present the proof of Theorem 2 in two parts. In Section 3.1, we prove the existence of a subgroup W that satisfies the properties that will be used in the main induction. The main induction is presented in Section 3.2.

3.1 Part I: Finding a "good" subgroup W

We first prove two lemmas (Lemma 4 and Lemma 5) that establish the existence of a coset $V + a$ such that $|V + a| \approx |A| \approx |(V + a) \cap A|$.

Lemma 4. *Suppose $A \subseteq G$ has coset complexity at most ℓ. There exists a coset $V + a \subseteq G$ such that*

$$|V + a| \geq 2^{-\ell}|A| \qquad \text{and} \qquad V + a \subseteq A.$$

Proof. The proof is by a simple induction on ℓ. The base case $\ell = 1$ is trivial as A must be a coset or the complement of a coset in this case. For $\ell > 1$, suppose that A belongs to the ring generated by $V_1 + b_1, \ldots, V_\ell + b_\ell$. Note that both $A \cap (V_\ell + b_\ell)$ and $A \cap (V_\ell + b_\ell)^c$ have coset complexity at most $\ell - 1$, and one of them has size larger than $\frac{|A|}{2}$. Applying the induction hypothesis to this set completes the proof. □

Recall that $\mathrm{E}(A)$ denotes the additive energy of A. The following lemma is essentially from [San19]. Its proof is based on several fundamental results in additive combinatorics.

Lemma 5. *If $A \subseteq G$ satisfies $\mathrm{E}(A) \geq \epsilon|A|^3$, then there exists a coset $V + a$ with*

$$|V + a| \geq 2^{-O(|\log \epsilon|^{3+o(1)})}|A| \qquad \text{and} \qquad |A \cap (V + a)| \geq 2^{-O(|\log \epsilon|^{1+o(1)})}|V + a|.$$

Proof. By the Balog-Szemerédi-Gowers theorem [TV10, Theorem 2.31] there is a subset $A' \subseteq A$ such that $|A'| \geq \epsilon^{O(1)}|A|$ and $|A' + A'| \leq \epsilon^{-O(1)}|A'|$. Now we can apply [San19, Proposition 2.2] to conclude the existence of the desired coset $V + a$. □

The following lemma is an adaptation of [GS08a, Lemma 3.4]. It says that if $\|f\|_A \leq M$, then every subgroup V can be regularized to a slightly smaller subgroup W such that f has small variance on all cosets of W.

Lemma 6. *Suppose $f : G \to \mathbb{R}$ satisfies $\|f\|_A \leq M$, and let $V \subseteq G$ be a subgroup and $\delta > 0$ be a parameter. There exists a subgroup $W \subseteq V$ such that $\dim(W) \geq \dim(V) - \frac{M}{\delta}$ and $\mathrm{Var}[f|W + c] \leq \delta M$ for every c.*

Proof. If a function $g : \mathbb{F}_2^k \to \mathbb{R}$ satisfies $\|g\|_A \leq M$ and $|\widehat{g}(r)| \leq \delta$ for all $r \neq 0$, then

$$\mathrm{Var}[g] = \sum_{r \neq 0} |\widehat{g}(r)|^2 \leq \delta \|g\|_A \leq \delta M.$$

By [GS08a, Lemma 3.4], there exists a subgroup $W \subseteq V$ such that $\dim(W) \geq \dim(V) - \frac{M}{\delta}$, and for every $r \notin W^\perp$,

$$\sum_{r' \in W^\perp + r} |\widehat{f}(r')| \leq \delta. \tag{10}$$

For $c \in W^\perp$, define $g : W \to \mathbb{R}$ as $g(y) := f(y+c)$, and note that by Equation (4), for every $r \in W$,

$$\widehat{g}(r) = \sum_{t \in W^\perp} \widehat{f}(r+t)\chi_t(c).$$

In particular, we have $\|g\|_{A(W)} \le \|f\|_A \le M$, and moreover by Equation (10), we have $|\widehat{g}(r)| \le \delta$ for every $r \neq 0$. It follows that $\mathrm{Var}[f|W+c] = \mathrm{Var}[g] \le \delta M$ as desired. □

The following corollary is the conclusion of this section. It shows that if the assumptions of Lemma 4 or Lemma 5 hold, then one can find the desired subgroup W with the properties that are needed in the proof of Theorem 2.

Corollary 1. *Let $M \ge \frac{1}{2}, \epsilon_1, \epsilon_2 > 0$, $\epsilon \in [0, \frac{1}{2})$, and $\delta < \min\{1/2, \epsilon_2\}$ be parameters. Suppose $A \subseteq G$, and $g : G \to \mathbb{R}$ satisfies $\|\mathbf{1}_A - g\|_\infty \le \epsilon$ and $\|g\|_A \le M$. If there exists a coset $V + a$ with*

$$|V + a| \ge \epsilon_1 |A| \qquad \text{and} \qquad |A \cap (V + a)| \ge \epsilon_2 |V|,$$

then there exists a subgroup $W \subseteq V$ such that

(i) *$\mathbb{E}[\mathbf{1}_A|W+c] \le \delta$ or $\mathbb{E}[\mathbf{1}_A|W+c] \ge 1-\delta$ for every c.*

(ii) *The set*

$$\mathcal{F}_W = \{c \in W^\perp : \ \mathbb{E}[\mathbf{1}_A|W+c] \ge 1-\delta\}$$

satisfies $1 \le |\mathcal{F}_W| \le 2^{\frac{5M^2}{(1-2\epsilon)^2\delta}}/\epsilon_1$.

(iii) *If $\mathcal{F}_W \neq W^\perp$, then $\|g|_{W+c}\|_{\mathbb{A}(W)} \le \|g\|_{\mathbb{A}} - \frac{1-2\epsilon-2\delta}{2}$ for every c.*

Proof. By Lemma 6, there exists a subgroup $W \subseteq V$ such that $\dim(W) \ge \dim(V) - \frac{4M^2}{(1-2\epsilon)^2\delta}$ and $\mathrm{Var}[g|W+c] \le \frac{(1-2\epsilon)^2\delta}{4}$ for every c. We prove that W is the desired subgroup.

We first prove that (i) is satisfied. Let $\alpha = \mathbb{E}[g|W+c]$. If $\alpha \le \frac{1}{2}$, then since $\|\mathbf{1}_A - g\|_\infty \le \epsilon$, we have

$$\mathrm{Var}[g|W+c] = \mathop{\mathbb{E}}_{x \in W+c}[|g(x)-\alpha|^2] \ge \Pr_{x \in W+c}[x \in A](1-\epsilon-\alpha)^2 \ge \Pr_{x \in W+c}[x \in A]\left(\frac{1}{2}-\epsilon\right)^2,$$

which shows

$$\mathbb{E}[\mathbf{1}_A|W+c] \le \left(\frac{1}{2}-\epsilon\right)^{-2} \cdot \mathrm{Var}[g|W+c] \le \delta.$$

Similarly if $\alpha \ge \frac{1}{2}$, then $\mathbb{E}[1-\mathbf{1}_A|W+c] \le \delta$.

For (ii), we first prove the lower bound by contradiction. Suppose the contrary that $|\mathcal{F}_W| = 0$, then by (i), we have

$$|A \cap (W+c)| \le \delta |W| \tag{11}$$

for every $c \in W^\perp$. By our choice of δ, summing Equation (11) over all cosets $W+c$ in $V+a$ gives $|A \cap (V+a)| \le \delta|V| < \epsilon_2|V|$, which is a contradiction. For the upper bound on $|\mathcal{F}_W|$, as $\mathbb{E}[\mathbf{1}_A|W+c] \ge 1-\delta$ for any $c \in \mathcal{F}_W$, we have

$$|A| \ge \sum_{c \in \mathcal{F}_W} |A \cap (W+c)| \ge |\mathcal{F}_W| \cdot (1-\delta)|W|,$$

which yields the desired upper bound

$$|\mathcal{F}_W| \le \frac{|A|}{(1-\delta)|W|} \le \frac{1}{1-1/2} \cdot \frac{|A|}{|V|} \cdot \frac{|V|}{|W|} \le \frac{2 \times 2^{\frac{4M^2}{(1-2\epsilon)^2\delta}}}{\epsilon_1} \le \frac{2^{\frac{5M^2}{(1-2\epsilon)^2\delta}}}{\epsilon_1},$$

where the last inequality uses the fact that $M \ge \delta$.

To prove (iii), note that by Equation (4), we have

$$\|g|_{W+c}\|_{\mathbb{A}(W)} = \sum_{b \in W \setminus \{0\}} \left| \sum_{r \in W^\perp} \widehat{g}(b+r)\chi_r(c) \right|.$$

By the triangle inequality, we obtain the following inequality relating $\|g\|_{\mathbb{A}}$ and $\|g|_{W+c}\|_{\mathbb{A}(W)}$:

$$\|g|_{W+c}\|_{\mathbb{A}(W)} \le \sum_{b \in W \setminus \{0\}} \sum_{r \in W^\perp} |\widehat{g}(b+r)| = \sum_{\substack{b \in W, r \in W^\perp \\ (b,r) \neq (0,0)}} |\widehat{g}(b+r)| - \sum_{r \in W^\perp \setminus \{0\}} |\widehat{g}(r)| = \|g\|_{\mathbb{A}} - \sum_{r \in W^\perp \setminus \{0\}} |\widehat{g}(r)|.$$

It remains to show that the last sum is at least $\frac{1-2\epsilon-2\delta}{2}$. By (i) and (ii), assuming $\mathcal{F}_W \neq W^\perp$, there exist $c_1, c_2 \in W^\perp$ such that $\mathbb{E}[g|W+c_1] \ge 1-\epsilon-\delta$ and $\mathbb{E}[g|W+c_2] \le \epsilon+\delta$. Therefore, by the triangle inequality,

$$1-2\epsilon-2\delta \le \mathbb{E}[g|W+c_1] - \mathbb{E}[g|W+c_2] = \sum_{r \in W^\perp} \widehat{g}(r)[\chi_r(c_1) - \chi_r(c_2)] \le 2 \sum_{r \in W^\perp \setminus \{0\}} |\widehat{g}(r)|.$$

This completes the proof of (iii). □

Remark 5. Corollary 1 (iii) is one of the new ideas in the proof of Theorem 2. Switching from $\|\cdot\|_A$ to $\|\cdot\|_{\mathbb{A}}$ guarantees a significant decrease in the norm on every coset of W.

3.2 Part II: Induction

In this section, we finish the proof of Theorem 2 by presenting the main inductive argument, which is the principal novelty of our proof. We start by strengthening the induction hypothesis through the following definition.

Definition 2. *Let $m, k \in \mathbb{N}$, and $0 \le \epsilon < \frac{1}{2}$ be parameters. We say that $A \subseteq G$ and $g : G \to \mathbb{R}$ satisfy the property $\mathcal{P}_{k,\epsilon}(m)$ if*

(i) *both A and A^c are k-affine connected, and*

(ii) $\|\mathbf{1}_A - g\|_\infty \le \epsilon$ *and* $\|g\|_{\mathbb{A}} \le \left(\frac{1-2\epsilon}{4}\right) m$.

Moreover, if $t, r \in \mathbb{N}$ with $r \le k+1$, and $|A| \le \frac{|G|}{2}$, then we say that A and g satisfy the property $\mathcal{P}'_{k,\epsilon}(m,r,t)$ if (i) and (ii) hold, and additionally there exists $\mathcal{X} = \bigcup_{i=1}^{t} (W_i + a_i)$ where every $W_i + a_i$ is a coset in G and the following conditions are satisfied:

(iii) $\|g|_{W_i+c}\|_{\mathbb{A}(W_i)} \le \left(\frac{1-2\epsilon}{4}\right)(m-1)$ *for every $i \in [t]$ and every $c \in G$.*

(iv) *For every $x_1, \ldots, x_r \in A \setminus \mathcal{X}$, either*

(a) *there exists a set* $S \subseteq [r]$ *such that* $|S| > 1$ *is odd and* $\sum_{i \in S} x_i \in A \setminus \mathcal{X}$*; or,*

(b) *there exists a nonempty* $S \subseteq [r]$ *such that* $\sum_{i \in S} x_i \in \mathcal{X}$*.*

Remark 6. Note that if A and g satisfy $\mathcal{P}_{k,\epsilon}(m)$ and $|A| \leq \frac{|G|}{2}$, then taking $\mathcal{X} = \{\mathbf{0}\}$ shows that A and g satisfy $\mathcal{P}'_{k,\epsilon}(m, k+1, 1)$. Indeed, with these parameters, (iii) is trivially satisfied as $\|g|_{\{c\}}\|_{\mathbb{A}(\{\mathbf{0}\})} = 0$ for all $c \in G$, and (iv) is equivalent to the assumption that A is k-affine connected.

Similarly, if $|A| > \frac{|G|}{2}$, then A^c and $1 - g$ satisfy $\mathcal{P}'_{k,\epsilon}(m, k+1, 1)$.

The following lemma is the core of the proof of Theorem 2.

Lemma 7 (Main lemma)**.** *Let* ϵ, m, r, k, t *be as in Definition 2. If* $A \subseteq G$ *and* $g : G \to \mathbb{R}$ *satisfy* $\mathcal{P}_{k,\epsilon}(m)$*, then the coset complexity of* A *is at most*

$$\ell_{k,\epsilon}(m) := \mathrm{Tower}_4\left((m-1)k + 1 + \log_4^*\left(\frac{1}{1-2\epsilon}\right) + O(1)\right).$$

If A *and* g *satisfy* $\mathcal{P}'_{k,\epsilon}(m, r, t)$*, then the coset complexity of* A *is at most*

$$\ell_{k,\epsilon}(m, r, t) := \begin{cases} 1 & m = 1, r \leq k \\ \ell_{k,\epsilon}(m) & r = k+1, t = 1 \\ \mathrm{Tower}_4\left(r + \log_4^* \max\left\{t, \ell_{k,\epsilon}(m-1)\right\}\right) & \textit{otherwise.} \end{cases}$$

Proof of Lemma 7. By Remark 6, it suffices to only prove the second part of the lemma that concerns $\mathcal{P}'_{k,\epsilon}(m, r, t)$. The proof is by an induction on the two parameters m and r.

Base of induction $m = 1$**:** If $m = 1$, then $\|g\|_{\mathbb{A}} \leq \frac{1-2\epsilon}{4}$, which implies $\|g - \mathbb{E}[g]\|_\infty \leq \frac{1-2\epsilon}{4}$. Combining with $\|\mathbf{1}_A - g\|_\infty \leq \epsilon$, we have

$$\|\mathbf{1}_A - \mathbb{E}[g]\|_\infty \leq \frac{1+2\epsilon}{4} < \frac{1}{2}.$$

Since $\mathbb{E}[g]$ is a constant and $|A| \leq \frac{|G|}{2}$, we have $A = \emptyset$. Hence, A has coset complexity 1, which is at most $\ell_{m,k}(m, r, t)$ in both cases of $r \leq k$ and $r = k+1$.

The case $r = 1, m > 1$**:** In this case, by (iv), we have $A \subseteq \mathcal{X}$, and by (iii), we have

$$\|g|_{W_i + a_i}\|_{\mathbb{A}(W_i)} \leq \left(\frac{1-2\epsilon}{4}\right)(m-1),$$

for every $W_i + a_i \subseteq \mathcal{X}$. Hence, for every $i \in [t]$, we can apply the induction hypothesis to $A|_{W_i+a_i} + a_i \subseteq W_i$ and conclude that $A|_{W_i+a_i}$ has coset complexity at most $\ell_{k,\epsilon}(m-1)$. Taking the union over all $W_i + a_i$ shows that the coset complexity of A is at most $t \times \ell_{k,\epsilon}(m-1)$. By the inequality $xy \leq 4^{\max(x,y)}$, which is valid for all positive x, y, we have

$$t \times \ell_{k,\epsilon}(m-1) \leq \mathrm{Tower}_4\left(1 + \log_4^* \max\left\{t, \ell_{k,\epsilon}(m-1)\right\}\right) = \ell_{k,\epsilon}(m, 1, t),$$

as desired.

The case $r > 1, m > 1$: Consider Definition 2 (iv). Since there are at most 2^r choices for S and two choices (a) and (b), one of the following must hold:

- **Case I:** There exists an odd $d \in [3, r]$ such that

$$\Pr_{x_1,\ldots,x_d \in A\setminus\mathcal{X}}[x_1 + \ldots + x_d \in A \setminus \mathcal{X}] \geq \frac{1}{2^{r+1}}.$$

- **Case II:** There exists a $d \in [r]$ such that

$$\Pr_{x_1,\ldots,x_d \in A\setminus\mathcal{X}}[x_1 + \ldots + x_d \in \mathcal{X}] \geq \frac{1}{2^{r+1}}.$$

Claim 1. *In Case I, there exists a coset $V + a$ such that*

$$|V + a| \geq 2^{-k^K}|A \setminus \mathcal{X}| \qquad \text{and} \qquad |(A \setminus \mathcal{X}) \cap (V + a)| \geq 2^{-k^K}|V|, \tag{12}$$

where $K = O(1)$ is a universal constant.

Proof. Consider a fixation of $x_4, \ldots, x_d$ that maximizes the probability. We conclude that with $c = x_4 + \ldots + x_d$, we have

$$\Pr_{x_1,x_2,x_3 \in A\setminus\mathcal{X}}[x_1 + x_2 + x_3 \in (A \setminus \mathcal{X}) + c] \geq \frac{1}{2^{r+1}},$$

which translates to

$$\mathop{\mathbb{E}}_{x_1,x_2,x_3 \in G}[\mathbf{1}_{A\setminus\mathcal{X}}(x_1)\mathbf{1}_{A\setminus\mathcal{X}}(x_2)\mathbf{1}_{A\setminus\mathcal{X}}(x_3)\mathbf{1}_{A\setminus\mathcal{X}}(x_1 + x_2 + x_3 + c)] \geq \frac{1}{2^{r+1}}\left(\frac{|A \setminus \mathcal{X}|}{|G|}\right)^3.$$

Substituting the Fourier transform of $\mathbf{1}_{A\setminus\mathcal{X}}$, we obtain

$$\frac{1}{2^{r+1}}\left(\frac{|A \setminus \mathcal{X}|}{|G|}\right)^3 \leq \sum_{a\in G} |\widehat{\mathbf{1}_{A\setminus\mathcal{X}}}(a)|^4 \chi_a(c) \leq \sum_{a\in G} |\widehat{\mathbf{1}_{A\setminus\mathcal{X}}}(a)|^4.$$

Hence, by Equation (2), the additive energy of $A \setminus \mathcal{X}$ is large:

$$\mathrm{E}(A \setminus \mathcal{X}) \geq \frac{|A \setminus \mathcal{X}|^3}{2^{r+1}}. \tag{13}$$

We apply Lemma 5 with $\epsilon = 2^{-r-1} \geq 2^{-k-2}$ to conclude the existence of a coset $V + a$ with

$$|V + a| \geq 2^{-k^K}|A \setminus \mathcal{X}| \qquad \text{and} \qquad |(A \setminus \mathcal{X}) \cap (V + a)| \geq 2^{-k^K}|V|,$$

where $K = O(1)$ is a universal constant. □

We would like to obtain a similar statement for Case II. Unfortunately, this will require an application of the induction hypothesis, and it is the cause of the tower-type bound in our final result.

Claim 2. *In Case II, there exists a coset $V + a$ such that*

$$|V| \geq \frac{2^{-\ell_{k,\epsilon}(m-1)-t}}{t2^{r+1}}|A \setminus \mathcal{X}| \qquad \textit{and} \qquad |(A \setminus \mathcal{X}) \cap (V + a)| = |V|. \tag{14}$$

Proof. Considering the structure of $\mathcal{X}$, there exists an $i \in [t]$ such that

$$\Pr_{x_1,\ldots,x_d \in A\setminus\mathcal{X}}[x_1 + \ldots + x_d \in W_i + a_i] \geq \frac{1}{t2^{r+1}}.$$

Hence, there exists at least one choice of $x_2, \ldots, x_d \in A \setminus \mathcal{X}$ such that

$$\Pr_{x_1 \in A\setminus\mathcal{X}}[x_1 \in W_i + a + x_2 + \ldots + x_d] \geq \frac{1}{t2^{r+1}}.$$

Consequently, there exists $c \in G$ such that

$$\Pr_{x \in A\setminus\mathcal{X}}[x \in W_i + c] \geq \frac{1}{t2^{r+1}},$$

or equivalently

$$|(A \setminus \mathcal{X}) \cap (W_i + c)| \geq \frac{|A \setminus \mathcal{X}|}{t2^{r+1}}. \tag{15}$$

This by itself does not provide much information about $A \setminus \mathcal{X}$ as $W_i + c$ could be much larger than $A \setminus \mathcal{X}$. However, we have made progress by the decrease in $\|\cdot\|_{\mathbb{A}}$: by (iii), we have

$$\|g|_{W_i+c}\|_{\mathbb{A}(W_i)} \leq \left(\frac{1-2\epsilon}{4}\right)(m-1),$$

and thus we can apply the induction hypothesis to the restriction of A to $W_i + c$ to describe its full structure. More precisely, the coset complexity of $A|_{W_i+c}$ is at most $\ell_{k,\epsilon}(m-1)$. Since the coset complexity of $\mathcal{X}$ is at most t, it follows that the coset complexity of $(A \setminus \mathcal{X}) \cap (W_i + c)$ is at most $\ell_{k,\epsilon}(m-1) + t$. By applying Lemma 4, we find a coset $V + a \subseteq W_i + c$ such that

$$|V| \geq 2^{-\ell_{k,\epsilon}(m-1)-t}|(A \setminus \mathcal{X}) \cap (W_i + c)| \quad \text{and} \quad V + a \subseteq A \setminus \mathcal{X}.$$

Combining with Equation (15), we have

$$|V| \geq \frac{2^{-\ell_{k,\epsilon}(m-1)-t}}{t2^{r+1}}|A \setminus \mathcal{X}| \qquad \text{and} \qquad |(A \setminus \mathcal{X}) \cap (V + a)| = |V|.$$

□

Let $\epsilon_1 := 2^{-\ell_{k,\epsilon}(m-1)-t-\log(t)-k^K} \leq \min(2^{-k^K}, \frac{2^{-\ell_{k,\epsilon}(m-1)-t}}{t2^{r+1}})$ and $\epsilon_2 := 2^{-k^K}$ so that by Equation (12) and Equation (14), in both Case I and Case II, there exists a coset $V + a$ with

$$|V| \geq \epsilon_1|A \setminus \mathcal{X}| \qquad \text{and} \qquad |(A \setminus \mathcal{X}) \cap (V + a)| \geq \epsilon_2|V|.$$

Now we are in a position to apply Corollary 1 to $A \setminus \mathcal{X}$. For $\delta := \min(\frac{1-2\epsilon}{8}, \epsilon_2)$, by applying Corollary 1, we find a subgroup W such that

- $\mathbb{E}[\mathbf{1}_{A\setminus\mathcal{X}}|W+c] \le \delta$ or $\mathbb{E}[\mathbf{1}_{A\setminus\mathcal{X}}|W+c] \ge 1-\delta$ for every $c \in G$.
- Since $\|g\|_A \le \|g\|_{\mathbb{A}} + 1 \le 2m$, the set

$$\mathcal{F}_W = \{c \in W^\perp \ : \ \mathbb{E}[\mathbf{1}_{A\setminus\mathcal{X}}|W+c] \ge 1-\delta\}$$

  satisfies $1 \le |\mathcal{F}_W| \le 2^{\frac{20m^2}{(1-2\epsilon)^2\delta}}/\epsilon_1$. Furthermore, since $|A\setminus\mathcal{X}| \le |A| \le |G|/2$, we have $\mathcal{F}_W \neq W^\perp$.
- For every $c \in G$, we have

$$\|g|_{W+c}\|_{\mathbb{A}(W)} \le \|g\|_{\mathbb{A}} - \frac{1-2\epsilon-2\delta}{2} \le \|g\|_{\mathbb{A}} - \frac{1-2\epsilon}{4}. \tag{16}$$

Fix an arbitrary $c_0 \in \mathcal{F}_W$, and let $\gamma := \frac{2^{-2k}}{t}$. We will focus on $y \in W + c_0$. Recall that $\mathcal{X} = \bigcup_{i=1}^t (W_i + a_i)$. For $i \in [t]$, define

$$E_i := \left\{ a \in W_i^\perp : \Pr_{y \in W+c_0}[y \in W_i + a_i + a] \ge \gamma \right\}.$$

Since the sets $W_i + a_i + a$ are all disjoint (for different $a \in W_i^\perp$), we have $|E_i| \le 1/\gamma$.

Now we are ready to set up for the inductive step that will decrease r. Define

$$\mathcal{X}' = \mathcal{X} \cup (W + \mathcal{F}_W) \cup (W + \mathcal{F}_W + c_0) \cup \bigcup_{i=1}^t (W_i + E_i).$$

Note that

$$\mathcal{X}' = \bigcup_{i=1}^{t'} (W_i' + a_i'),$$

where $W_i' \in \{W_1, \ldots, W_t\} \cup \{W\}$ for all $i \in [t']$, and

$$\begin{aligned}
t' &\le t + 2|\mathcal{F}_W| + \sum_{i=1}^t |E_i| \\
&\le t\left(1 + \frac{1}{\gamma}\right) + \frac{2 \times 2^{\frac{20m^2}{(1-2\epsilon)^2\delta}}}{\epsilon_1} \\
&\le t + t^2 2^{2k} + 2^{1 + \frac{20m^2}{(1-2\epsilon)^2} \times \frac{8}{\epsilon_2(1-2\epsilon)} + \ell_{k,\epsilon}(m-1) + t + \log(t) + k^K} \\
&\le t + t^2 2^{2k} + 2^{1 + \frac{160m^2 2^{k^K}}{(1-2\epsilon)^3} + \ell_{k,\epsilon}(m-1) + t + \log(t) + k^K} \\
&\le 2^{2\max\{t, \ell_{k,\epsilon}(m-1)\}},
\end{aligned} \tag{17}$$

where we assumed that the $O(1)$ term in the definition of $\ell_{k,\epsilon}$ is chosen so that $\ell_{k,\epsilon}(m-1)$ significantly dominates all the terms that do not involve t.

Claim 3. *The pair A and g satisfies $\mathcal{P}'_{k,\epsilon}(m, r-1, t')$ as witnessed by $\mathcal{X}'$.*

Proof. Conditions (i) and (ii) of Definition 2 are trivially satisfied as A and g are not altered, and $\mathcal{X}'$ is a union of t' cosets. Condition (iii) is satisfied because either $W_i' \in \{W_1, \ldots, W_t\}$, or $W_i' = W$, and in the latter case (iii) is satisfied by Equation (16).

It remains to verify (iv). Consider $x_1, \ldots, x_{r-1} \in A \backslash \mathcal{X}' \subseteq A \backslash \mathcal{X}$. For every nonempty $S \subseteq [r-1]$, let $a_S = \sum_{i \in S} x_i$. If neither (a) nor (b) hold, then

$$a_S \notin A \setminus \mathcal{X}' \qquad \text{for every } S \subseteq [r-1] \text{ where } |S| > 1 \text{ is odd;} \tag{18}$$

$$a_S \notin \mathcal{X}' \qquad \text{for every nonempty } S \subseteq [r-1]. \tag{19}$$

We will establish the existence of an $x_r \in (A \setminus \mathcal{X}) \cap (W + c_0)$ such that $x_1, \ldots, x_{r-1}, x_r$ violate both (a) and (b) for A and $\mathcal{X}$, and thus contradict our initial assumption. Pick $y \in W$ uniformly at random, and set $x_r = y + c_0$. We have the following:

- Since $c_0 \in \mathcal{F}_W$, we have $\Pr[x_r \notin A \setminus \mathcal{X}] = \Pr_y[y + c_0 \notin A \setminus \mathcal{X}] \leq \delta$.
- For every nonempty even-size $S \subseteq [r-1]$, since $a_S \notin W + \mathcal{F}_W + c_0 \subseteq \mathcal{X}'$, we have $a_S + c_0 \notin W + \mathcal{F}_W$, and thus by the definition of $\mathcal{F}_W$,

$$\Pr_y[y + c_0 + a_S \in A \setminus \mathcal{X}] \leq \delta.$$

- For every nonempty S, since $a_S \notin W_i + E_i \subseteq \mathcal{X}'$, by applying the union bound over the cosets in $\mathcal{X}$, we have

$$\Pr_y[y + c_0 + a_S \in \mathcal{X}] \leq t \max_{i \in [t]} \Pr_y[y + c_0 + a_S \in W_i + a_i] = t \max_{i \in [t]} \Pr_y[y + c_0 \in W_i + a_i + a_S] \leq t\gamma.$$

We apply the union bound to the above statements. Since $\delta + 2^{r-1}\delta + 2^{r-1}t\gamma < 2^r 2^{-k^K} + 2^{r-1}2^{-2k} < 1$, with positive probability there exists a $y \in W$ such that for $x_r = y + c_0$,

- $x_r \in A \setminus \mathcal{X}$.
- For every nonempty even-size $S \subseteq [r-1]$, we have $x_r + \sum_{i \in S} x_i \notin A \setminus \mathcal{X}$.
- For every nonempty $S \subseteq [r-1]$, $x_r + \sum_{i \in S} x_i \notin \mathcal{X}$.

These together with (18) and (19) show that the sequence $x_1, \ldots, x_r$ violates (a) and (b) for A and $\mathcal{X}$, which is a contradiction. □

To finish the proof of Lemma 7, we can apply the induction hypothesis to A and $\mathcal{X}'$. By (17) the coset complexity of A at most

$$\begin{aligned} \ell_{k,\epsilon}(m, r-1, t') &\leq \ell_{k,\epsilon}(m, r-1, 4^{\max(t, \ell_{k,\epsilon}(m-1))}) \\ &= \text{Tower}_4\left((r-1) + 1 + \log_4^* \max\left(t, \ell_{k,\epsilon}(m-1)\right)\right) = \ell_{k,\epsilon}(m, r, t). \end{aligned}$$

□

Finally, we finish the proof of Theorem 2.

Proof of Theorem 2. Theorem 2 (i) is an immediate corollary of Lemma 3. To prove Theorem 2 (ii), by assumption, A and A^c are k-affine connected and there exists $g : G \to \mathbb{R}$ such that $\|\mathbf{1}_A - g\|_\infty \leq \epsilon$ and $\|g\|_{\mathbb{A}} \leq M$. Hence, A and g satisfy $\mathcal{P}_{k,\epsilon}(m)$ for $m = \left\lceil \frac{4M}{1-2\epsilon} \right\rceil$ and by Lemma 7, the coset complexity of A is at most

$$\ell_{k,\epsilon}(m) = \mathrm{Tower}_2\left(O\left(\frac{Mk}{1-2\epsilon}\right)\right).$$

□

4 Concluding remarks and open problems

We conclude the paper with some suggestions for future research:

- Can Theorem 2 be extended to all locally compact Abelian groups, or more generally to all locally compact groups? As we discussed in Section 1.1, we believe this is possible.
- Is the tower type bound in Theorem 2 necessary or is it an artifact of our proof? Note that for Theorem 1, Sanders' bound in [San19] is only exponential in $O(M^{3+o(1)})$.
- What can be said about the structure of the sets $A \subseteq G$ that have small approximate algebra norm if we do not assume affine connectivity? The following conjecture from [HHH21] remains open.

  Conjecture 1. *If $A \subseteq G$ satisfies $\|\mathbf{1}_A\|_{A,\epsilon} \leq M$, then there is a coset $V + a \subseteq G$ of codimension at most $\ell = O_{M,\epsilon}(1)$ such that $V + a \subseteq A$ or $V + a \subseteq A^c$.*

  Note that by Theorem 2 and Lemma 4, such a coset $V + a$ exists if we further assume that A and A^c are $O(1)$-affine connected.
- Theorem 2 belongs to a more general program that aims to characterize the functions that have complexity $O(1)$ in various natural communication and query models. As we discussed earlier, the approximate algebra norm, randomized parity decision tree complexity, and randomized communication complexity of the XOR-lift are exponentially equivalent [STV17, KLMY21, HHH21]. Therefore, Theorem 2 and Conjecture 1 are steps towards achieving such a characterization for the randomized parity decision tree model and the randomized communication complexity of the XOR-lifts.

  Another possible application of Theorem 2 in this program is a potential characterization of the XOR-lifts with communication complexity $O(1)$ in the unbounded-error model of Paturi and Simon [PS86]. We conjecture that those are precisely the XOR-lifts of the Boolean functions that have coset complexity $O(1)$. We intend to investigate this problem in future works.

References

[BS21] Nikhil Bansal and Makrand Sinha. k-Forrelation optimally separates quantum and classical query complexity. In *53rd Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2021, pages 1303–1316. ACM, 2021.

[CHHL19] Eshan Chattopadhyay, Pooya Hatami, Kaave Hosseini, and Shachar Lovett. Pseudorandom generators from polarizing random walks. *Theory of Computing*, 15(10):1–26, 2019.

[Coh60] Paul J. Cohen. On a conjecture of Littlewood and idempotent measures. *Amer. J. Math.*, 82:191–212, 1960.

[FK18] Michael A. Forbes and Zander Kelley. Pseudorandom generators for read-once branching programs, in any order. In *59th IEEE Annual Symposium on Foundations of Computer Science*, STOC 2018, pages 946–955. IEEE Computer Society, 2018.

[GS08a] Ben Green and Tom Sanders. Boolean functions with small spectral norm. *Geometric and Functional Analysis*, 18(1):144–162, 2008.

[GS08b] Ben Green and Tom Sanders. A quantitative version of the idempotent theorem in harmonic analysis. *Ann. of Math. (2)*, 168(3):1025–1054, 2008.

[GTW21] Uma Girish, Avishay Tal, and Kewen Wu. Fourier growth of parity decision trees. In *36th Computational Complexity Conference*, volume 200 of *LIPIcs*, pages 39:1–39:36, 2021.

[GW11] W. T. Gowers and J. Wolf. Linear forms and higher-degree uniformity for functions on $\mathbb{F}_p^n$. *Geom. Funct. Anal.*, 21(1):36–69, 2011.

[Hel53] Henry Helson. Note on harmonic functions. *Proc. Amer. Math. Soc.*, 4(5):686–691, 1953.

[HHH21] Lianna Hambardzumyan, Hamed Hatami, and Pooya Hatami. Dimension-free bounds and structural results in communication complexity. *Israel J. Math.*, 2021. to appear.

[HMP86] B. Host, J.-F. Méla, and F. Parreau. Analyse harmonique des mesures. *Astérisque*, (135-136):261, 1986.

[Hos86] B. Host. Le théorème des idempotents dans $B(G)$. *Bull. Soc. Math. France*, 114(2):215–223, 1986.

[KI40] Yukiyosi Kawada and Kiyosi Itô. On the probability distribution on a compact group. I. *Proc. Phys.-Math. Soc. Japan (3)*, 22:977–998, 1940.

[KLMY21] A. Knop, S. Lovett, S. McGuire, and W. Yuan. Guest column: Models of computation between decision trees and communication. *SIGACT News*, 52(2):46–70, June 2021.

[KM93] Eyal Kushilevitz and Yishay Mansour. Learning decision trees using the Fourier spectrum. *SIAM J. Comput.*, 22(6):1331–1348, 1993.

[Lef72] Marcel Lefranc. Sur certaines algèbres de fonctions sur un groupe. *C. R. Acad. Sci. Paris Sér. A-B*, 274:A1882–A1883, 1972.

[MŚ82] J.-F. Méla. Mesures ε-idempotentes de norme bornée. *Studia Math.*, 72(2):131–149, 1982.

[MRT19] Raghu Meka, Omer Reingold, and Avishay Tal. Pseudorandom generators for width-3 branching programs. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, page 13–23, 2019.

[PS86] Ramamohan Paturi and Janos Simon. Probabilistic communication complexity. *Journal of Computer and System Sciences*, 33(1):106–123, 1986.

[Riv90] Theodore J. Rivlin. *Chebyshev polynomials*. Pure and Applied Mathematics (New York). John Wiley & Sons, Inc., New York, second edition, 1990.

[RSV13] Omer Reingold, Thomas Steinke, and Salil P. Vadhan. Pseudorandomness for regular branching programs via fourier analysis. In *APPROX-RANDOM 2013*, volume 8096 of *Lecture Notes in Computer Science*, pages 655–670. Springer, 2013.

[RT19] Ran Raz and Avishay Tal. Oracle separation of BQP and PH. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, page 13–23, 2019.

[Rud59a] Walter Rudin. Idempotent measures on Abelian groups. *Pacific J. Math.*, 9:195–209, 1959.

[Rud59b] Walter Rudin. Measure algebras on abelian groups. *Bull. Amer. Math. Soc.*, 65:227–247, 1959.

[Rud90] Walter Rudin. *Fourier analysis on groups*. Wiley Classics Library. John Wiley & Sons, Inc., New York, 1990. Reprint of the 1962 original, A Wiley-Interscience Publication.

[Run07] Volker Runde. Cohen-Host type idempotent theorems for representations on Banach spaces and applications to Figà-Talamanca-Herz algebras. *J. Math. Anal. Appl.*, 329(1):736–751, 2007.

[San11] Tom Sanders. A quantitative version of the non-abelian idempotent theorem. *Geom. Funct. Anal.*, 21(1):141–221, 2011.

[San19] Tom Sanders. Boolean functions with small spectral norm, revisited. *Math. Proc. Cambridge Philos. Soc.*, 167(2):335–344, 2019.

[San20] Tom Sanders. Bounds in Cohen's idempotent theorem. *J. Fourier Anal. Appl.*, 26(2):Paper No. 25, 64, 2020.

[San21] Tom Sanders. Coset decision trees and the Fourier algebra. *J. Anal. Math.*, 144(1):227–259, 2021.

[STV17] Amir Shpilka, Avishay Tal, and Ben lee Volk. On the structure of Boolean functions with small spectral norm. *Comput. Complexity*, 26(1):229–273, 2017.

[Tal17] Avishay Tal. Tight bounds on the fourier spectrum of AC0. In *32nd Computational Complexity Conference*, volume 79 of *LIPIcs*, pages 15:1–15:31, 2017.

[Tal20] Avishay Tal. Towards optimal separations between quantum and randomized query complexities. In *61st IEEE Annual Symposium on Foundations of Computer Science*, FOCS 2020, pages 228–239, 2020.

[TV10] Terence Tao and Van H. Vu. *Additive combinatorics*, volume 105 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.

[TWXZ13] Hing Yin Tsang, Chung Hoi Wong, Ning Xie, and Shengyu Zhang. Fourier sparsity, spectral norm, and the Log-rank conjecture. In *54th IEEE Annual Symposium on Foundations of Computer Science*, FOCS 2013, pages 658–667, 2013.

[Wen54] J. G. Wendel. Haar measure and the semigroup of measures on a compact group. *Proc. Amer. Math. Soc.*, 5:923–929, 1954.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il