An Improved Lower Bound for Arithmetic Regularity

Kaave Hosseini∗ Shachar Lovett† Guy Moshkovitz‡ Asaf Shapira§

Abstract

The arithmetic regularity lemma due to Green [GAFA 2005] is an analogue of the famous Szemerédi regularity lemma in graph theory. It shows that for any abelian group $G$ and any bounded function $f : G \rightarrow [0, 1]$, there exists a subgroup $H \leq G$ of bounded index such that, when restricted to most cosets of $H$, the function $f$ is pseudorandom in the sense that all its nontrivial Fourier coefficients are small. Quantitatively, if one wishes to obtain that for $1 - \epsilon$ fraction of the cosets, the nontrivial Fourier coefficients are bounded by $\epsilon$, then Green shows that $|G/H|$ is bounded by a tower of twos of height $1/\epsilon^3$. He also gives an example showing that a tower of height $\Omega(\log 1/\epsilon)$ is necessary. Here, we give an improved example, showing that a tower of height $\Omega(1/\epsilon)$ is necessary.

1 Introduction

As an analogue of Szemerédi’s regularity lemma in graph theory [4], Green [2] proposed an arithmetic regularity lemma for abelian groups. Given an abelian group $G$ and a bounded function $f : G \rightarrow [0, 1]$, Green showed that one can find a subgroup $H \leq G$ of bounded index, such that when restricted to most cosets of $H$, the function $f$ is pseudorandom in the sense that all of its nontrivial Fourier coefficients are small. Quantitatively, the index of $H$ in $G$ is bounded by a tower of twos of height polynomial in the error parameter. The goal of this note is to provide an example showing that these bounds are essentially tight. This strengthens a previous example due to Green [2] which shows that a tower of height logarithmic in the error parameter is necessary; and makes the lower bounds in the arithmetic case analogous to those obtained in the graph case [1].

We restrict our attention in this paper to the group $G = \mathbb{Z}_2^n$, and note that our construction can be generalized to groups of bounded torsion in an obvious way. We first make some
basic definitions. Let $A$ be an affine subspace (that is, a translation of a vector subspace) of $\mathbb{Z}_2^n$ and let $f : A \rightarrow [0, 1]$ be a function. The Fourier coefficient of $f$ associated with $\eta \in \mathbb{Z}_2^n$ is

$$\hat{f}(\eta) = \frac{1}{|A|} \sum_{x \in A} f(x)(-1)^{(x, \eta)} = \mathbb{E}_{x \in A} [f(x)(-1)^{(x, \eta)}].$$

Any subspace $H \leq \mathbb{Z}_2^n$ naturally determines a partition of $\mathbb{Z}_2^n$ into affine subspaces $\mathbb{Z}_2^n/H = \{H + g : g \in \mathbb{Z}_2^n\}$. The number $|\mathbb{Z}_2^n/H| = 2^{n - \dim H}$ of translations is called the index of $H$.

1.1 Arithmetic regularity and the main result

For an affine subspace $A = H + g$ of $\mathbb{Z}_2^n$, where $H \leq \mathbb{Z}_2^n$ and $g \in \mathbb{Z}_2^n$, we say that a function $f : A \rightarrow [0, 1]$ is $\epsilon$-regular if all its nontrivial Fourier coefficients are bounded by $\epsilon$, that is,

$$\max_{\eta \notin H^\perp} |\hat{f}(\eta)| \leq \epsilon.$$ 

Note that a trivial Fourier coefficient $\eta \in H^\perp$ satisfies $|\hat{f}(\eta)| = |\mathbb{E}_{x \in A} f(x)|$. Henceforth, for any $f : \mathbb{Z}_2^n \rightarrow [0, 1]$ we denote by $f|_A : A \rightarrow [0, 1]$ the restriction of $f$ to $A$.

**Definition 1.1 (\(\epsilon\)-regular subspace).** Let $f : \mathbb{Z}_2^n \rightarrow [0, 1]$. A subspace $H \leq \mathbb{Z}_2^n$ is $\epsilon$-regular for $f$ if $f|_A$ is $\epsilon$-regular for at least $(1 - \epsilon)|\mathbb{Z}_2^n/H|$ translations $A$ of $H$.

Green [2] proved that any bounded function has an $\epsilon$-regular subspace $H$ of bounded index, that is, whose index depends only on $\epsilon$ (equivalently, $H$ has bounded codimension). In the following, twr($h$) is a tower of twos of height $h$; formally, twr($h$) := $2^{\text{twr}(h-1)}$ for a positive integer $h$, and twr(0) = 1.

**Theorem 1** (Arithmetic regularity lemma in $\mathbb{Z}_2^n$, Theorem 2.1 in [2]). For every $0 < \epsilon < \frac{1}{2}$ there is $M(\epsilon)$ such that every function $f : \mathbb{Z}_2^n \rightarrow [0, 1]$ has an $\epsilon$-regular subspace of index at most $M(\epsilon)$. Moreover, $M(\epsilon) \leq \text{twr}([1/\epsilon^3])$.

A lower bound on $M(\epsilon)$ of about twr($\log_2(1/\epsilon)$) was given in the same paper [2], following the lines of Gowers’ lower bound on the order of $\epsilon$-regular partitions of graphs [1]. While Green’s lower bound implies that $M(\epsilon)$ indeed has a tower-type growth, it is still quite far from the upper bound in Theorem 1.

Our main result here nearly closes the gap between the lower and upper bounds on $M(\epsilon)$, showing that $M(\epsilon)$ is a tower of twos of height at least linear in $1/\epsilon$. Our construction follows the same initial setup as in [2], but will diverge from that point on. Our proof is inspired by the recent simplified lower bound proof for the graph regularity lemma in [3] by a subset of the authors.

**Theorem 2.** For every $\epsilon > 0$ it holds that $M(\epsilon) \geq \text{twr}([1/16\epsilon])$. 

2
1.2 A variant of Theorem 2 for binary functions

One can also deduce from Theorem 2 a similar bound for $\epsilon$-regular sets, that is, for binary functions $f : \mathbb{Z}_2^n \to \{0, 1\}$. For this, all we need is the following easy probabilistic argument.

**Claim 1.2.** Let $\tau > 0$ and $f : \mathbb{Z}_2^n \to [0, 1]$. There exists a binary function $S : \mathbb{Z}_2^n \to \{0, 1\}$ satisfying, for every affine subspace $A$ of $\mathbb{Z}_2^n$ of size $|A| \geq 4n^2/\tau^2$ and any vector $\eta \in \mathbb{Z}_2^n$, that

$$|\widehat{S}|_A(\eta) - \widehat{f}|_A(\eta)| \leq \tau.$$

**Proof.** Choose $S : \mathbb{Z}_2^n \to \{0, 1\}$ randomly by setting $S(x) = 1$ with probability $f(x)$, independently for each $x \in \mathbb{Z}_2^n$. Let $A, \eta$ be as in the statement. The random variable

$$\widehat{S}|_A(\eta) = \frac{1}{|A|} \sum_{x \in A} S(x)(-1)^{\langle x, \eta \rangle}$$

is an average of $|A|$ mutually independent random variables taking values in $[-1, 1]$, and its expectation is $\widehat{f}|_A(\eta)$. By Hoeffding’s bound, the probability that $|\widehat{S}|_A(\eta) - \widehat{f}|_A(\eta)| > \tau$ is smaller than

$$2 \exp(-\tau^2 |A| / 2) \leq 2^{-2n^2+1}.$$

The number of affine subspaces over $\mathbb{Z}_2^n$ can be trivially bounded by $2^{n^2}$, the number of sequences of $n$ vectors in $\mathbb{Z}_2^n$. Hence, the number of pairs $(A, \eta)$ is bounded by $2^{n^2+n}$. The claim follows by the union bound. \hfill $\square$

Applying Claim 1.2 with $\tau = \epsilon/2$ (say) implies that if $f : \mathbb{Z}_2^n \to [0, 1]$ has no $\epsilon$-regular subspace of index smaller than $\text{twr}([1/16\epsilon])$ then, provided $n$ is sufficiently large in terms of $\epsilon$, there is $S : \mathbb{Z}_2^n \to \{0, 1\}$ that has no $\epsilon/2$-regular subspace of index smaller than $\text{twr}([1/16\epsilon])$.

2 Proof of Theorem 2

2.1 The Construction

To construct a function witnessing the lower bound in Theorem 2 we will use pseudo-random spanning sets.

**Claim 2.1.** Let $V$ be a vector space over $\mathbb{Z}_2$ of dimension $d$. Then there is a set of $8d$ nonzero vectors in $V$ such that any subset of $\frac{3}{4}$ of them spans $V$.

**Proof.** Choose random vectors $v_1, \ldots, v_{8d} \in V \setminus \{0\}$ independently and uniformly. Let $U$ be a subspace of $V$ of dimension $d - 1$. The probability that a given $v_i$ lies in $U$ is at most $\frac{1}{2}$. By Chernoff’s bound, the probability that more than $6d$ of our vectors $v_i$ lie in $U$ is smaller than $\exp(-2(2d)^2/8d) = \exp(-d)$. By the union bound, the probability that there exists a
subspace $U$ of dimension $d - 1$ for which the above holds is at most $2^d \exp(-d) < 1$. This completes the proof. \hfill \Box

We now describe a function $f : \mathbb{Z}_2^n \to [0, 1]$ which, as we will later prove, has no $\epsilon$-regular subspace of small index. Henceforth set $s = \lceil 1/16\epsilon \rceil$. Furthermore, let $d_i$ be the following sequence of integers of tower-type growth:

$$d_{i+1} = \begin{cases} 2D_i & \text{if } i = 1, 2, 3 \\ 2D_{i-3} & \text{if } i > 3 \end{cases} \quad \text{where } D_i = \sum_{j=1}^{i} d_j \text{ and } D_0 = 0 .$$

Note that the first values of $d_i$ for $i \geq 1$ are $1, 2, 8, 2^8, 2^{264}$, etc., and it is not hard to see that $d_i \geq \text{twr}(i - 1)$ for every $i \geq 1$. Set $n = D_s \geq \text{twr}(s - 1))$. For $x \in \mathbb{Z}_2^n$, partition its coordinates into $s$ blocks of sizes $d_1, \ldots, d_s$, and identify $x = (x^1, \ldots, x^s) \in \mathbb{Z}_2^{d_1 + \cdots + d_s} = \mathbb{Z}_2^n$.

Let $1 \leq i \leq s$. Bijectively associate with each $v \in \mathbb{Z}_2^{D_{i-1}} = \mathbb{Z}_2^{d_1 + \cdots + d_{i-1}}$ a nonzero vector $\xi_i(v) \in \mathbb{Z}_2^{d_i}$ such that the set of vectors $\{\xi_i(v) : v \in \mathbb{Z}_2^{D_{i-1}}\}$ has the property that any subset of $\frac{3}{4}$ of its elements spans $\mathbb{Z}_2^{d_i}$. The existence of such a set, which is a subset of size $2^{D_{i-1}}$ in a vector space of dimension $d_i$, follows from Claim 2.1 when $i > 3$, since then $2^{D_{i-1}} = 8d_i$. When $i \leq 3$ the existence of such a set is trivial since $\lceil (3/4)i \rceil = i$, hence any basis would do (and we take $2^{D_{i-1}} = d_i$). With a slight abuse of notation, if $x \in \mathbb{Z}_2^n$ we write $\xi_i(x)$ for $\xi_i((x^1, \ldots, x^{i-1}))$.

We define our function $f : \mathbb{Z}_2^n \to [0, 1]$ as

$$f(x) = \frac{|\{1 \leq i \leq s : \langle x^i, \xi_i(x) \rangle = 0\}|}{s} .$$

The following is our main technical lemma, from which Theorem 2 immediately follows.

**Lemma 2.2.** The only $\epsilon$-regular subspace for $f$ is the zero subspace $\{0\}$.

*Proof of Theorem 2.* The index of $\{0\}$ is $|\mathbb{Z}_2^n/\{0\}| = 2^n \geq \text{twr}(s) = \text{twr}(\lceil 1/16\epsilon \rceil))$. \hfill \Box

### 2.2 Proof of Lemma 2.2

Let $H$ be an $\epsilon$-regular subspace for $f$, and assume towards contradiction that $H \neq \{0\}$. Let $1 \leq i \leq s$ be minimal such that there exists $v \in H$ for which $v^i \neq 0$. For any $g \in \mathbb{Z}_2^n$, let

$$\gamma_g = (0, \ldots, 0, \xi_i(g), 0, \ldots, 0) \in \mathbb{Z}_2^n$$

where only the $i$-th component is nonzero. We will show that for more than an $\epsilon$-fraction of the translations $H + g$ of $H$ it holds that $\gamma_g \notin H^\perp$ and

$$\frac{|\langle f |_{H+g}(\gamma_g) \rangle|}{1} > \epsilon .$$

This will imply that $H$ is not $\epsilon$-regular for $f$, thus completing the proof.

First, we argue that $\gamma_g \notin H^\perp$ for a noticeable fraction of $g \in \mathbb{Z}_2^n$. We henceforth let $B = \{g \in \mathbb{Z}_2^n : \gamma_g \in H^\perp\}$ be the set of ”bad” elements.
Claim 2.3. \( |B| \leq \frac{3}{4} |Z_2^n| \).

Proof. If \( g \in B \) then \( \langle \xi_i(g), v_i \rangle = 0 \). Hence, \( \{ \xi_i(g) : g \in B \} \) does not span \( Z_2^{d_i} \). By the construction of \( \xi_i \), this means that \( \{ (g^1, \ldots, g^{i-1}) : g \in B \} \) accounts to at most \( \frac{3}{4} \) of the elements in \( Z_2^{D_i-1} \), and hence \( |B| \leq \frac{3}{4} |Z_2^n| \). \( \square \)

Next, we argue that typically \( \hat{f}_{H+g}(\gamma_g) \) is large. Let \( W \leq Z_2^n \) be the subspace spanned by the last \( s - i \) blocks, that is, \( W = \{ w \in Z_2^n : w^1 = \ldots = w^i = 0 \} \). Note that for any \( g \in Z_2^n, w \in W \) we have \( \gamma_{g+w} = \gamma_g \). In particular, \( g + w \in B \) if and only if \( g \in B \).

Claim 2.4. Fix \( g \in Z_2^n \) such that \( \gamma_g \notin H^\perp \). Then
\[
\mathbb{E}_{w \in W} \left[ \hat{f}_{H+g+w}(\gamma_g) \right] = \frac{1}{2s}.
\]

Proof. Write \( f(x) = \frac{1}{s} \sum_{j=1}^{s} B_j(x) \) where \( B_j(x) : Z_2^n \to \{0, 1\} \) is the characteristic function for the set of vectors \( x \) satisfying \( \langle x^j, \xi_j(x) \rangle = 0 \). Hence, for any affine subspace \( A \) in \( Z_2^n \),
\[
\hat{f}_{|A}(\gamma_g) = \frac{1}{s} \sum_{j=1}^{s} B_j|_{A}(\gamma_g). \tag{1}
\]

Set \( A = H + g + w \) for an arbitrary \( w \in W \). We next analyze the Fourier coefficient \( \hat{B}_j|_{A}(\gamma_g) \) for each \( j \leq i \), and note that in these cases we have \( \xi_j(x) = \xi_j(g) \) for any \( x \in A \). First, if \( j < i \) then for every \( x \in A \) we have \( x^j = g^j \), which implies that \( B_j|_{A} \) is constant. Since a nontrivial Fourier coefficient of a constant function equals 0, we have
\[
\hat{B}_j|_{A}(\gamma_g) = 0, \quad \forall j < i. \tag{2}
\]

Next, for \( j = i \), write \( B_i|_{A}(x) = \frac{1}{2}((-1)^{\langle x^i, \xi_i(x) \rangle} + 1) \). Since \( \langle x, \gamma_g \rangle = \langle x^i, \xi_i(x) \rangle \), we have
\[
\hat{B}_i|_{A}(\gamma_g) = \mathbb{E}_{x \in A} \left[ \frac{1}{2}((-1)^{\langle x^i, \xi_i(x) \rangle} + 1) \cdot (-1)^{\langle x^i, \xi_i(x) \rangle} \right] = \mathbb{E}_{x \in A} [B_i(x)] = \frac{1}{2}. \tag{3}
\]

Finally, for \( j > i \) we average over all \( w \in W \). Let \( H + W \) be the subspace spanned by \( H, W \). Writing \( B_j(x) = \frac{1}{2}((-1)^{\langle x^j, \xi_j(x) \rangle} + 1) \), the average Fourier coefficient is
\[
\mathbb{E}_{w \in W} \mathbb{E}_{x \in H+g+w} \left[ B_j(x)(-1)^{\langle x^j, \xi_j(x) \rangle} \right] = \frac{1}{2} \mathbb{E}_{x \in H+W+g} \left[ (-1)^{\langle x^i, \xi_i(g) \rangle + \langle x^i, \xi_i(x) \rangle} \right].
\]

Note that for every fixing of \( x^1, \ldots, x^{j-1} \), we have that \( x^j \) is uniformly distributed in \( Z_2^{d_j} \) (due to \( W \)), and that \( (-1)^{\langle x^i, \xi_i(g) \rangle} \) is constant. Since \( \xi_i(x) \neq 0 \), we conclude that
\[
\mathbb{E}_{w \in W} \left[ B_j|_{H+g+w}(\gamma_g) \right] = 0, \quad \forall j > i. \tag{4}
\]

The proof now follows by substituting (2), (3) and (4) into (1). \( \square \)
Since $f|_{H+g+w}(\gamma_g) \leq 1$, we infer (via a simple averaging argument) the following corollary.

**Corollary 2.5.** Fix $g \in \mathbb{Z}_2^n$ such that $\gamma_g \notin H^\perp$. Then for more than $1/4s$ fraction of the elements $w \in W$,

$$f|_{H+g+w}(\gamma_g) > \frac{1}{4s}.$$  

We are now ready to conclude the proof of Lemma 2.2. Partition $\mathbb{Z}_2^n$ into translations of $W$, and recall that $\gamma_g$ depends just on the translation $g+W$. By Claim 2.3, for at least $\frac{1}{4}$ of the translations, $\gamma_g \notin H^\perp$. By Corollary 2.5, in each such translation, more than $1/4s$-fraction of the elements $g+w$ satisfy $f|_{H+g+w}(\gamma_g) > 1/4s$. Since $1/16s \geq \epsilon$, the subspace $H$ cannot be $\epsilon$-regular for $f$.

**References**


