Abstract. Green developed an arithmetic regularity lemma to prove a strengthening of Roth’s theorem on arithmetic progressions in dense sets. It states that for every $\epsilon > 0$ there is some $N_0(\epsilon)$ such that for every $N \geq N_0(\epsilon)$ and $A \subset [N]$ with $|A| = \alpha N$, there is some nonzero $d$ such that $A$ contains at least $(\alpha^3 - \epsilon) N$ three-term arithmetic progressions with common difference $d$.

We prove that the minimum $N_0(\epsilon)$ in Green’s theorem is an exponential tower of $2$s of height on the order of $\log(1/\epsilon)$. Both the lower and upper bounds are new. It shows that the tower-type bounds that arise from the use of a regularity lemma in this application are quantitatively necessary.

1. Introduction

A celebrated theorem of Roth [25] states that for each $\alpha > 0$ there is a least positive integer $N(\alpha)$ such that if $N \geq N(\alpha)$ and $A \subset [N] := \{1, \ldots, N\}$ with $|A| \geq \alpha N$, then $A$ contains a three-term arithmetic progression. Over the past six decades, there has been great efforts made by many researchers toward understanding the growth of this function, and despite the introduction of important tools in these efforts, the growth of $N(\alpha)$ is still not well understood. The upper bound was improved by Heath-Brown [21], Szemerédi [30], Bourgain [7, 8], Sanders [28, 27], and most recently Bloom [5] (see also [6]). The lower bound of Behrend [2] was recently improved a bit by Elkin [10] (see also Green and Wolf [20] for a shorter proof). The best known bounds are of the form $\alpha^{-\Omega(\log(\alpha^{-1}))} \leq N(\alpha) \leq 2^{O(\alpha^{-1}(\log \alpha^{-1})^4)}$.

Szemerédi [29] extended Roth’s theorem to show that any dense set of integers contains arbitrarily long arithmetic progressions. Szemerédi’s proof developed an early version of Szemerédi’s regularity lemma [31], which gives a rough structural result for large graphs and is arguably the most powerful tool developed in graph theory. It roughly says that any graph can be equitably partitioned into a bounded number of parts so that between almost all pairs of parts, the graph behaves randomly-like. Szemerédi’s proof of the regularity lemma gives an upper bound on the number of parts which is tower-type in an approximation parameter, which gives a seemingly poor bound for the various applications of the regularity lemma. For over two decades there was some hope that a substantially better bound might hold leading to better bounds in the many applications. This hope was shattered by Gowers [16], who proved that the bound on the number of parts in the regularity lemma must grow as a tower-type function. Further results improving on some aspects of the lower bound were obtained in [9, 11, 24].

It has been a major program over the last few decades to find new proofs of the various applications of Szemerédi’s regularity lemma and its variants that avoid using the regularity lemma and obtain much better quantitative bounds. This program, popularized by Szemerédi and others, has been quite successful, leading to the development of powerful new methods, such as in Gowers’ new proof of Szemerédi’s theorem [17] which introduced higher order Fourier analysis [17], and in the resolution of many open problems in extremal combinatorics using the powerful probabilistic technique known as dependent random choice (see the survey [15]). However, until now it was unclear if one could...
growth of the minimum regularity lemma application like Theorem 1.1 are necessar y. Our main theorem determines the bound on the size of the partition. Green [18] proved a tower- type lower bound for the arithmetic regularity lemma and consequently shows that the bound of \(\alpha^3\) one gets by considering a random set of density \(\alpha\).

Green [18] developed an arithmetic analogue of Szemerédi’s regularity lemma and used it to prove the following theorem, which answered a question of Bergelson, Host and Kra [3]. It shows that while the total number of three-term arithmetic progressions can be much smaller than the random bound, there is a nonzero \(d\) for which the number of three-term arithmetic progressions with common difference \(d\) is at least roughly the random bound.

**Theorem 1.1** (Green’s popular progression difference theorem [18]). For each \(\epsilon > 0\) there is an integer \(N_0\) such that for any \(N \geq N_0\) and subset \(A \subset [N]\) with \(|A| = \alpha N\), there is a nonzero \(d\) such that \(A\) contains at least \((\alpha^3 - \epsilon)N\) three-term arithmetic progressions with common difference \(d\).

Similar to the graph setting, the proof of the arithmetic regularity lemma gives a tower-type upper bound on the size of the partition. Green [18] proved a tower-type lower bound for the arithmetic regularity lemma in the setting of vector spaces over \(\mathbb{F}_2\), and Hosseini, Lovett, Moshkovitz, and Shapira [22] later improved the tower height to \(\Omega(\epsilon^{-1})\). Green’s proof of Theorem 1.1 uses the arithmetic regularity lemma and consequently shows that \(N_0\) can be taken to be an exponential tower of twos of height \(\epsilon^{-O(1)}\). It was unknown if the tower-type bounds that come from any regularity lemma application like Theorem 1.1 are necessary. Our main theorem determines the growth of the minimum \(N_0\) for which Theorem 1.1 holds, showing that it is an exponential tower of twos of height \(\Theta(\log(1/\epsilon))\). Let \(\text{tower}(m)\) denote an exponential tower of twos of height \(m\).

**Theorem 1.2.** Let \(N_0(\epsilon)\) denote the smallest choice of \(N_0\) for which Theorem 1.1 holds. There exist absolute constants \(c, C > 0\) such that for all \(0 < \epsilon < 1/2\),

\[
\text{tower}(c \log(1/\epsilon)) \leq N_0(\epsilon) \leq \text{tower}(C \log(1/\epsilon)).
\]

This result, and the considerably easier analogous result in vector spaces over a fixed finite field [12, 13], are the first examples of regularity lemma applications that require the tower-type growth.

For investigations of popular differences for other patterns, including recent results on higher dimensional patterns, see [3, 4, 14, 19, 23, 26].

1.1. **Detailed statement of results.** Theorem 1.2 comes in two parts, an upper bound and a lower bound. The upper bound is as follows.

**Theorem 1.3** (Upper bound for intervals). There exists a constant \(C > 0\) such that the following is true. Let \(\epsilon > 0\) and \(N \geq \text{tower}(C \log(1/\epsilon))\). For every \(A \subset [N]\), setting \(\alpha = |A|/N\), there exists some positive integer \(d\) such that \(x, x + d, x + 2d \in A\) for at least \((\alpha^3 - \epsilon)N\) many integers \(x\).

The main part of the proof is an analogous result in abelian groups of odd order.

**Theorem 1.4** (Upper bound for abelian groups). There exists a constant \(C > 0\) such that the following is true. Let \(\epsilon > 0\) and let \(G\) be a finite abelian group of odd order with \(|G| \geq \text{tower}(C \log(1/\epsilon))\). For every \(A \subset G\), setting \(\alpha = |A|/|G|\), there exists some \(d \in G \setminus \{0\}\) such that \(x, x + d, x + 2d \in A\) for at least \((\alpha^3 - \epsilon)|G|\) many values of \(x \in G\).

For the lower bound for intervals, we prove the following result, which is somewhat stronger than that lower bound result claimed in Theorem 1.2.
Theorem 1.5 (Lower bound for intervals). There exist constants \( c, \alpha_0 > 0 \) such that for every \( 0 < \alpha \leq \alpha_0, 0 < \epsilon \leq \alpha^{12} \) and \( N \leq \text{tower}(c \log(1/\epsilon)) \), there exists \( A \subset [N] \) with \( |A| \geq \alpha N \) such that for every positive integer \( d \leq N/2 \), one has \( x, x + d, x + 2d \in A \) for at most \((\alpha^3 - \epsilon)(N - 2d)\) many integers \( x \).

Organization. In the next section, we introduce some helpful notation and preliminaries including some basic facts from discrete Fourier analysis. In Section 3, we give an overview of the proof strategies for our results. In Section 4, we prove Theorems 1.3 and 1.4 giving the upper bound results. In Section 5, we give some auxiliary results for the probabilistic lower bound construction. In Section 6, we give a lower bound construction for groups which are the product of prime cyclic groups with fast growing order. In Section 7, we then use this construction as an important ingredient to obtain the lower bound construction in intervals.

We often omit floor and ceiling signs when they are not crucial for clarity of presentation.

2. Notations and Preliminaries

Averaging and expectation. We use \( \mathbb{E} \) to denote the averaging operator: given a function \( f \) on a finite set \( S \), denote

\[
\mathbb{E} f = \mathbb{E}_{x \in S} [f(x)] := \frac{1}{|S|} \sum_{x \in S} f(x).
\]

We may write \( \mathbb{E}_x \) instead of \( \mathbb{E}_{x \in S} \) if the domain of \( x \) is clear from context (usually over a group).

The \( L^p \) norms are defined in the usual way:

\[
\|f\|_p := (\mathbb{E}[|f|^p])^{1/p}.
\]

As our lower bound construction is probabilistic, we will also need to consider expectations of random variables, for which we use the usual notation \( \mathbb{E} \) for expectation (note the difference in font compared to the averaging operator \( \mathbb{E} \)).

Fourier transform and convolutions. Given a finite abelian group \( G \), let \( \widehat{G} \) denote its dual group, whose elements are characters of \( G \), i.e., homomorphisms \( \chi : G \to \mathbb{T} := \{ z \in \mathbb{C} : |z| = 1 \} \). The Fourier transform of \( f : G \to \mathbb{C} \) is a function \( \hat{f} : \widehat{G} \to \mathbb{C} \) defined by

\[
\hat{f}(\chi) := \mathbb{E}[\chi f] = \mathbb{E}_{z \in \mathbb{T}} [f(z)\overline{\chi(z)}].
\]

We write \( \chi^{1/2} \) to denote the character given by \( x \mapsto \chi(x/2) \) (we will always work with odd abelian groups so that \( x/2 \) makes sense).

It is often convenient to explicitly identify the dual group \( \widehat{G} \) with \( G \) (they are isomorphic for finite abelian groups). For example, for \( f : \mathbb{Z}_N \to \mathbb{C} \) and \( r \in \mathbb{Z}_N \), we identify \( r \) with the character \( \chi_r(x) = e(xr/N) \) where we use the standard notation for the complex exponential

\[
e(t) := \exp(2\pi it), \quad t \in \mathbb{R}.
\]

Thus,

\[
\hat{f}(r) := \mathbb{E}_x [f(x)e(xr/N)].
\]

Likewise, for \( f : \mathbb{F}_p^n \to \mathbb{C} \) and \( r \in \mathbb{F}_p^n \), we identify \( r \) with the character \( \chi_r(x) = e(x \cdot r/p) \), where \( x \cdot r = x_1r_1 + \cdots + x_nr_n \in \mathbb{F}_p \) is the dot product in \( \mathbb{F}_p^n \). Thus,

\[
\hat{f}(r) := \mathbb{E}_x [f(x)e(x \cdot r/p)].
\]

Given two functions \( f \) and \( g \) on \( G \), their convolution \( f \ast g \) is defined by

\[
(f \ast g)(x) := \mathbb{E}_{y \in G} [f(y)g(x - y)].
\]
We recall several useful properties of the Fourier transform:

\[ f(x) = \sum_{\chi \in \hat{G}} [\hat{f}(\chi) \chi(x)], \]  
\[ \widehat{f \ast g} = \hat{f} \cdot \hat{g}, \]  
\[ E_{x \in G}[f(x)g(x)] = \sum_{\chi \in \hat{G}} \hat{f}(\chi) \overline{g(\chi)}, \]  
\[ E_{x \in G}[|f(x)|^2] = \sum_{\chi \in \hat{G}} |\hat{f}(\chi)|^2. \]

[Fourier inversion formula]
[Convolution identity]
[Plancherel’s identity]
[Parseval’s identity]

The Fourier transform is also fundamentally related to the count of 3-APs (or the count of solutions to linear equations in general), as evident in the following key identity already used in the proof of Roth’s theorem [25]. It can be easily shown by substituting the Fourier coefficients and expanding.

\[ E_{x,d \in G}[f(x)f(x + d)f(x + 2d)] = \sum_{\chi \in \hat{G}} \hat{f}(\chi)^2 \overline{\hat{f}(\chi^-2)}. \]  

(Densities) For an abelian group \( G \) with odd order and a function \( f : G \to [0, 1] \), we define the density of 3-APs of \( f \) as

\[ E_{x,d \in G}[f(x)f(x + d)f(x + 2d)]. \]

We define the density of 3-APs with common difference \( d \) of \( f \) as

\[ E_{x \in G}[f(x)f(x + d)f(x + 2d)]. \]

For a subset \( A \) of \( G \), when we say “density of 3-APs” of \( A \), we mean that of its indicator function \( 1_A \), and likewise for “density of 3-APs with common difference \( d \)” of \( A \).

Over the interval \([N]\), we have two possible notions for the density of 3-APs with common difference \( d \) of a function \( f : [N] \to [0, 1] \). One can define the density of 3-APs with common difference \( d \) of \( f \) as

\[ \frac{\sum_{x \in [N-2d]} [f(x)f(x + d)f(x + 2d)]}{N}, \]

as used in Theorem 1.3. This defines the density of 3-APs with common difference \( d \) as the average weight of the 3-APs \((x, x + d, x + 2d)\) for \( x \in [N] \), setting the value of \( f \) outside \([N]\) to 0. The other possible definition of the density of 3-APs with common difference \( d \) of \( f \) is

\[ E_{x \in [N-2d]}[f(x)f(x + d)f(x + 2d)] = \frac{\sum_{x \in [N-2d]} [f(x)f(x + d)f(x + 2d)]}{N - 2d}, \]

as used in Theorem 1.5. In this case, we take the average only over 3-APs supported in \([N]\). It is easy to see that the density from the second definition is always at least the density from the first definition. In particular, the upper bound using the first definition implies the upper bound using the second definition. Similarly, the lower bound using the second definition implies the lower bound using the first definition. Thus, we give the stronger result in each case.

Constants. We use \( c > 0 \) and \( C > 0 \) to denote small and large absolute constants, though their values may differ at every instance. One could imagine attaching a unique subscript to each appearance of \( c \) and \( C \).
3. Overview of strategy

In this section we sketch the proof ideas of our main theorems, starting with the upper bound (Theorems 1.4 and 1.3) and then followed by the lower bound (Theorem 1.5).

In both cases, we prove the “functional” versions of the theorems. That is, instead of working with subsets $A \subset G$, we work with functions $f: G \to [0,1]$, which can also be viewed as subsets with weighted elements. A subset $A \subset G$ can be represented by its indicator function $1_A$. Conversely, given a function $f: G \to [0,1]$, we can produce from it a random subset $A \subset G$ obtained by putting each element $x \in G$ independently into $A$ with probability $f(x)$. The resulting $A$ has similar statistical properties compared to $f$ due to concentration. Working with functions affords us greater flexibility, which is convenient for both parts of the proofs.

3.1. Upper bound. Theorems 1.3 and 1.4 follow from the functional forms below by setting $f = 1_A$.

**Theorem 3.1** (Upper bound for intervals, functional version). There exists a constant $C > 0$ such that the following holds. Let $\epsilon > 0$ and $N \geq \text{tower}(C \log(1/\epsilon))$. Let $f: [N] \to [0,1]$ with $\mathbb{E}f = \alpha$. Then there exists $d \in G \setminus \{0\}$ such that

$$\mathbb{E}_{x \in [N]}[f(x)f(x+d)f(x+2d)] \geq \alpha^3 - \epsilon.$$

**Theorem 3.2** (Upper bound for abelian groups, functional version). There exists a constant $C > 0$ such that the following holds. Let $\epsilon > 0$ and let $G$ be a finite abelian group of odd order with $|G| \geq \text{tower}(C \log(1/\epsilon))$. Let $f: G \to [0,1]$ with $\mathbb{E}f = \alpha$. Then there exists $d \in G \setminus \{0\}$ such that

$$\mathbb{E}_{x \in G}[f(x)f(x+d)f(x+2d)] \geq \alpha^3 - \epsilon.$$

Green [18] proved the above theorems with a slightly worse bound of $\text{tower}(C(1/\epsilon)^C)$ instead of $\text{tower}(C \log(1/\epsilon))$. Let us first give a quick sketch of Green’s approach. It is easier to first explain it for the finite field vector space setting $G = \mathbb{F}_p^n$ with $p$ fixed.

Green begins by establishing a regularity lemma. Given $f: G \to [0,1]$, one finds a subspace $H$ of $G = \mathbb{F}_p^n$ of codimension at most $\text{tower}(C(1/\epsilon)^C)$ such that inside almost all translates of $H$, $f$ behaves “pseudorandomly” in the sense of having small Fourier coefficients (other than the principal “zeroth” Fourier coefficient that records the density). This subspace $H$ is obtained iteratively, similar to the standard energy-increment proofs of regularity lemmas: starting with $H_0 = G$, at each step one checks if $H_i$ has the desired properties, and if not, then one finds a bounded-codimensional subspace $H_{i+1}$ of $H_i$ witnessing the non-uniformity. Each step increases the “energy”, or mean-squared density, by at least $\epsilon^{O(1)}$. As the energy can never exceed 1, the process terminates after at most $(1/\epsilon)^{O(1)}$ steps.

Once we have the above bounded-codimensional subspace $H$, let $g$ be the function obtained from $f$ by averaging $f$ inside each translate of $H$. In other words, consider the convolution $g = f * \beta_H$, where $\beta_H$ denotes the averaging measure on $H$ (normalized so that $\mathbb{E}\beta_H = 1$). The regularity property of $H$, namely $f - g$ having small Fourier coefficients when restricted to most $H$-cosets, is enough to deduce that $f$ and $g$ have similar densities of 3-APs with common differences lying in $H$, i.e.,

$$\mathbb{E}_{x \in G, d \in H}[f(x)f(x+d)f(x+2d)] \approx \mathbb{E}_{x \in G, d \in H}[g(x)g(x+d)g(x+2d)].$$

On the other hand, $g$ is constant along $H$-cosets, so that $g(x) = g(x+d) = g(x+2d)$ for all $d \in H$. Thus the final expression is $\mathbb{E}[g^3] \geq (\mathbb{E}g)^3$ by convexity, and we have $\mathbb{E}g \approx \mathbb{E}f$. Putting everything together, we have

$$\mathbb{E}_{x \in G, d \in H}[f(x)f(x+d)f(x+2d)] \geq (\mathbb{E}f)^3 - (\epsilon/2).$$

If $G$ is large enough, so that $H$ is large enough, then the above inequality implies that there is some nonzero common difference $d \in H$ so that

$$\mathbb{E}_{x \in G}[f(x)f(x+d)f(x+2d)] \geq (\mathbb{E}f)^3 - \epsilon,$$
thereby showing that $d$ is a popular common difference.

Let us now sketch how to improve the above bound to tower($C \log(1/\epsilon)$) in the finite field setting, which had been worked out in [12]. Instead of finding an $H$ that regularizes the function $f$, we simply seek to satisfy the inequality (2). One then shows that if (2) is violated, then we can find a bounded-codimensional subspace of $H$, via an application of the weak regularity lemma at a “local” level, so that the corresponding mean cubed density of $f$ (after averaging along the subspace) nearly doubles at each step (instead of merely increasing by $\epsilon^{-O(1)}$), so that the iteration process must end after $O(\log(1/\epsilon))$ steps (instead of $(1/\epsilon)^{O(1)}$ steps). Once we obtain an $H$ satisfying (2), the rest of the argument is essentially identical.

For general abelian groups $G$, unlike in the field field setting $\mathbb{F}_p^n$, the group might not have enough subgroups to run the above arguments. Instead, one uses Bohr sets, which play an analogous role to subgroups. Bohr sets are defined in Section 4.1, and their manipulations are much more delicate compared to subspaces, particularly as they do not have as nice closure properties. Green proved his arithmetic regularity lemma for general abelian groups using Bohr sets as basic structural objects. The strategy remains largely similar in spirit to the finite field setting, though more challenging at a technical level (e.g., the group cannot be partitioned into Bohr sets, unlike with subgroups). For instance, we obtain $g$ from $f$ by setting $g(x)$ to be a certain “smooth average” of $f$ around a certain carefully chosen Bohr neighborhood of $x$. While the values $g(x)$, $g(x + d)$, and $g(x + 2d)$ are no longer necessarily identical, they are hopefully approximately the same if $d$ lies inside some Bohr neighborhood of 0. The rest of Green’s argument is similar to the finite field vector space case.

To obtain the corresponding result for intervals, one considers embedding $[N]$ in $\mathbb{Z}_N$ and only consider Bohr sets whose elements $d$ are all small in magnitude.

In order to improve the bound from tower($(1/\epsilon)^{O(1)}$) to tower($O(\log(1/\epsilon))$) for general groups and for intervals, we carefully execute a combination of the above ideas. New ideas are required to adapt the mean cube density increment argument from [12] to Bohr sets due to complications that do not arise in the finite field setting. The proof is carried out in full detail in Section 4.

3.2. Lower bound. In this section, we give a brief overview of the proof of Theorem 1.5. We will deduce Theorem 1.5 from its functional analogue given below, where we replace the subset $A$ by a function $f : [N] \to [0,1]$ with density $\alpha$ so that for any nonzero $d$, the density of 3-APs with common difference $d$ of $f$ is at most $\alpha^3(1-\epsilon)$.

Theorem 3.3 (Lower bound for intervals, functional version). There are positive absolute constants $c, \alpha_0$ such that the following holds. If $0 \leq \alpha \leq \alpha_0$, $0 \leq \epsilon \leq \alpha^7$, and $N \leq \text{tower}(c \log(1/\epsilon))$, then there is a function $f : [N] \to [0,1]$ with $E[f] = \alpha$ such that for any integer $0 < d < N/2$,

$$E_{x \in [N-2d]} [f(x) f(x+d) f(x+2d)] \leq \alpha^3(1-\epsilon).$$

We remark that we have replaced $\epsilon$ by $\epsilon \alpha^3$, which is more convenient to work with in the lower bound construction. Since we treat $\alpha$ as a constant throughout, this has no effect on the behavior of the asymptotic bound we get. The proof of Theorem 1.5 assuming Theorem 3.3 follows from a standard sampling argument, which we defer to Appendix A. There, we also show that it suffices to prove Theorem 1.5 and Theorem 3.3 for $N \geq \epsilon^{-15}$.

In the following subsections, we sketch the construction of the function $f$ in Theorem 3.3. This construction utilizes a construction over cyclic groups which can be factored into a product of groups with appropriate growth in size. The construction in this case is inspired by the recursive construction presented in [12] over finite field vector spaces. However, several important new ideas are needed. The construction of $f$ for such product groups is sketched in Subsection 3.2.1. Using this construction, we can construct $f$ over intervals, using ideas sketched in Subsection 3.2.2, thus proving Theorem 3.3.
We remark that in this section we only give proof sketches without complete detail. The details of each construction and full proofs are presented in Sections 6 and 7.

3.2.1. Product groups. We first give the construction for groups which can be written as a product of appropriately growing cyclic groups of prime order.

**Theorem 3.4** (Lower bound for product of growing prime cyclic groups). Let $0 < \alpha \leq 1/4$, $0 < \varepsilon \leq 20^{-9}$, and $G = \mathbb{Z}_n$ where $n$ is a positive integer such that there exist distinct primes $m_1, \ldots, m_s$ with $s \leq \log_{150}(\frac{\varepsilon^{-1/4}n^6}{8})$ satisfying

- $n = \prod_{j=1}^s m_j$,
- $\varepsilon^{-1/3}/2 < n_1 \leq \varepsilon^{-1/3}$, and
- for $i \geq 2$, $n_{i-1}^6 < n_i < \exp(2^{-1} \cdot 64^{-2} \cdot 150^{i-1} \varepsilon^{1/4} n_{i-1})/2$ where $n_i = \prod_{j=1}^i m_j$.

Then, there exists a function $f : G \to [0,1]$ with $E[f] = \alpha$ such that for any $d \in G \setminus \{0\}$,

$$E_x[f(x)f(x+d)f(x+2d)] \leq \alpha^3(1-\varepsilon).$$

Furthermore, $E_x[f(x)^3] \leq 3\alpha^3/2$ and there exists $\alpha' \in [\alpha, \alpha(1+\varepsilon^{1/4})]$ such that $f(x) = \alpha'$ for at least a $3/4$ fraction of $x \in G$.

Here, the product structure of $G$ and the bound on the growth of $m_i$ allow us to conduct an iterative construction. The basic framework of the construction builds on the construction in [12], which took place in the setting of $\mathbb{F}_p^n$.

We build functions $f_1, f_2, f_3, \ldots$ in this order. Here the domain of $f_i$ is $Q_i = \prod_{j=1}^i \mathbb{Z}_{m_j}$. We will maintain that $E_{f_i} = \alpha$, and that for every $d \in Q_i \setminus \{0\}$, we have

$$E_{x \in Q_i}[f_i(x)f_i(x+d)f_i(x+2d)] \leq (1-\varepsilon)(E_{f_i})^3.$$

The function $f_s$ thus gives the desired construction.

Suppose we have already constructed $f_{i-1}$. Here is how we construct $f_i : Q_i \to [0,1]$:

(i) Design some family $\mathcal{F}$ of functions $\mathbb{Z}_{m_i} \to [0,1]$.
(ii) Choose some $M_i \subset Q_{i-1}$ (with some properties).
(iii) For each $x \in M_i$, fill in values of $f_i$ on the coset $x + \mathbb{Z}_{m_i}$ using a random function chosen from $\mathcal{F}$.
(iv) For each $x \not\in M_i$, set all values of $f_i$ on the coset $x + \mathbb{Z}_{m_i}$ to be $f_{i-1}(x)$.

We refer to this process as random modification (the phrase “perturbation” was used in [12]). We will show that with positive probability, the random function $f_s$ satisfies the desired properties.

The main difference from the finite vector space case is the choice of the family $\mathcal{F}$. In [12], we first choose $g : \mathbb{F}_p \to [0,1]$ to be a multiple of the indicator of an interval of length $2p/3$. We then choose the family $\mathcal{F}$ to be functions of the form $g(x \cdot v)$ for some nonzero $v \in \mathbb{F}_p^m$.

Here, instead, we choose a nice model function $g : \mathbb{Z}_{m_i} \to [0,1]$ satisfying certain properties to be discussed later. We choose $\mathcal{F}$ to consist of functions $g_{a,b} : \mathbb{Z}_{m_i} \to [0,1]$ defined by $g_{a,b}(x) = g(ax+b)$, indexed by $a, b \in \mathbb{Z}_{m_i}$ with $a \neq 0$.

We denote elements of $Q_i$ by $x = (x_1, x_2, \ldots, x_i)$ where $x_j \in \mathbb{Z}_{m_j}$. We write $x_{[i-1]} = (x_1, \ldots, x_{i-1}) \in Q_{i-1}$. We can prove that for any $x \in Q_i$ with $x_{[i-1]} \in M_i$, and $d \in Q_i \setminus \{0\}$ such that $d_{[i-1]} = 0$,

$$E_{a,b}[g_{a,b}(x)g_{a,b}(x+d)g_{a,b}(x+2d)] = E_{x,z \in \mathbb{Z}_{m_i}, x \neq 0}[g(y)g(y+z)g(y+2z)] \leq (1-c)(E_{g})^3,$$

where $a,b$ vary uniformly over all elements of $\mathbb{Z}_{m_i}$ with $a \neq 0$, and $y, z$ vary uniformly over all elements of $\mathbb{Z}_{m_i}$ with $z \neq 0$. The final inequality is due to a property of the model function $g$. It then follows via concentration that with positive probability,

$$E_{x \in Q_i}[f_i(x)f_i(x+d)f_i(x+2d)] \leq (1-\varepsilon)(E_{f_i})^3,$$

for every $d \in Q_i \setminus \{0\}$ with $d_{[i-1]} = 0$. 

We are left with the task of bounding the density of 3-APs with common difference \( d \) where \( d_{[i-1]} \neq 0 \). Over \( \mathbb{P}_n^i \), this is easy, as we can show, using the structure of vector spaces, that under a mild condition, if \( d_{[i-1]} \neq 0 \), then
\[
E_{x \in Q_i} [f_i(x) f_i(x+d) f_i(x+2d)] = E_{x \in Q_{i-1}} [f_{i-1}(x) f_{i-1}(x+d_{[i-1]}) f_{i-1}(x+2d_{[i-1]})] \leq \alpha^3 (1 - \epsilon).
\]
Such equality does not hold in our current setting. Even though we do not need exact equality, obtaining uniform control over all \( d \in Q_i \) with \( d_{[i-1]} \neq 0 \) using standard concentration inequalities does not work because \( n_1 \) is small compared to \( \epsilon^{-1} \). However, we can indeed guarantee with high probability the equality
\[
E_{x \in Q_i} [f_i(x) f_i(x+d) f_i(x+2d)] = E_{x \in Q_{i-1}} [f_{i-1}(x) f_{i-1}(x+d_{[i-1]}) f_{i-1}(x+2d_{[i-1]})],
\]
assuming that the model function \( g \) over \( \mathbb{Z}_{m_i} \) satisfies some additional nice properties, which we refer to as smoothness.

Roughly speaking, \( g : \mathbb{Z}_{m_i} \to [0,1] \) is smooth if, for random \( a_1, \ldots, a_h \in \mathbb{Z}_{m_i} \setminus \{0\} \), with high probability, one has
\[
E_{g} \left[ \prod_{j=1}^{h} g(a_j y + b_j) \right] = E[g]^h \text{ for all } b_1, \ldots, b_h \in \mathbb{Z}_{m_i}.
\]

To see how this smoothness property helps, assume that we are given \( x' \in Q_{i-1} \) and \( d' \in Q_{i-1} \setminus \{0\} \) such that
\[
\begin{align*}
f_i(x) &= g(a_1 x + b_1) \quad \text{for all } x \text{ such that } x_{[i-1]} = x', \\
f_i(x) &= g(a_2 x + b_2) \quad \text{for all } x \text{ such that } x_{[i-1]} = x' + d', \\
f_i(x) &= g(a_3 x + b_3) \quad \text{for all } x \text{ such that } x_{[i-1]} = x' + 2d'.
\end{align*}
\]
Then, for all \( d \) with \( d_{[i-1]} = d' \),
\[
E_{x \in Q_i : x_{[i-1]} = x'} [f_i(x) f_i(x+d) f_i(x+2d)] = E_{g \in \mathbb{Z}_{m_i}} [g(a_1 y + c_1) g(a_2 y + c_2) g(a_3 y + c_3)],
\]
where \( c_1, c_2, c_3 \) depend on \( x' \) and \( d \). Moreover, if \( g \) is smooth, then with high probability over random \( a_1, a_2, a_3 \), one has
\[
E_{g \in \mathbb{Z}_{m_i}} [g(a_1 y + c_1) g(a_2 y + c_2) g(a_3 y + c_3)] = E[g]^3, \tag{3}
\]
for all \( c_1, c_2, c_3 \in \mathbb{Z}_{m_i} \). In such case, we obtain that for all \( d \) with \( d_{[i-1]} = d' \),
\[
E_{x \in Q_i : x_{[i-1]} = x'} [f_i(x) f_i(x+d) f_i(x+2d)] = E[g]^3.
\]
Thus, as long as the parameters \( a_1, a_2, a_3 \) corresponding to each \( x' \in Q_{i-1} \) and \( d' \in Q_{i-1} \setminus \{0\} \) satisfy (3), we obtain the desired equality
\[
E_{x \in Q_i} [f_i(x) f_i(x+d) f_i(x+2d)] = E_{x \in Q_{i-1}} [f_{i-1}(x) f_{i-1}(x+d_{[i-1]}) f_{i-1}(x+2d_{[i-1]})].
\]
We guarantee this property by applying the union bound over possible values of \( d' \) and \( x' \) in \( Q_{i-1} \). Note that it is crucial here that the smoothness property allows us to avoid the union bound over \( d' \), which takes \( m_i \approx \exp(O(Q_{i-1})) \) possible values.

The model function \( g \) over \( \mathbb{Z}_{m_i} \) is constructed in Section 5. The essential idea behind the construction is that \( g \) should be supported on only a few Fourier characters. The details of the construction over product groups are included in Section 6.
3.2.2. Intervals. Using Theorem 3.4, we can prove Theorem 3.3, giving the desired lower bound over intervals. The construction of the function \( f \) in Theorem 3.3 over intervals consists of three steps.

In the first step, we construct a function \( f_1 \) with density \( \alpha \) which is 0 in the interval \([N' + 1, N]\) for \( N' \) slightly smaller than \( N \). This sets the density of 3-APs with common difference \( d \) close to \( N/2 \) to 0.

In the second step, we let \( f_2 \) be the function obtained from the following procedure applied to \( f_1 \):

(i) Partition \([N']\) to \( N'/q \) intervals \( I_1, I_2, \ldots, I_{N'/q} \) of length \( q \) where \( q \) can be written as a product of prime numbers as required in Theorem 3.4.

(ii) Using Theorem 3.4, construct a function \( g : \mathbb{Z}_q \to [0, 1] \) which satisfies \( \mathbb{E}[g] = \alpha \) and \( \mathbb{E}_{x \in \mathbb{Z}_q}[g(x)g(x + d)g(x + 2d)] \leq \alpha^3(1 - \epsilon) \) for any \( d \in \mathbb{Z}_q \setminus \{0\} \).

(iii) For each \( j = 1, 2, \ldots, N'/q \), identify each interval \( I_j \) with \( \mathbb{Z}_q \) and place a copy of \( g \) on each of them.

For any \( d \) with \( 0 < d < N/2 \) and \( q \nmid d \), one can show that the density of 3-APs with common difference \( d \) of \( f_2 \) is at most \( \alpha^3(1 - \epsilon) \). However, for \( d \) divisible \( q \), the density of 3-APs with common difference \( d \) of \( f_2 \) is larger than \( \alpha^3 \).

Note that the function \( f_2 \) constructed in the second step is constant on each mod \( q \) residue class in \([N']\). In the third step, we construct the function \( f_3 \) as follows:

(i) Construct a subset \( X \) of \( \mathbb{Z}_{N'/q} \) with much fewer 3-APs compared to the random bound using a variant of the Behrend construction.

(ii) Let \( P_t = \{ x \in [N'] : x \equiv t \pmod{q} \} \). With some appropriate \( T \subseteq \mathbb{Z}_q \), for each \( t \in T \), take a random linear transformation of \( X \) inside set \( \mathbb{Z}_{N'/q} \), and set \( f_3 \) on \( P_t \) to be the indicator function of this randomly transformed \( X \).

(iii) On \([N'] \setminus \bigcup_{t \in T} P_t \), set \( f_3 \) to be equal to \( f_2 \).

The function \( f_3 \) has the property that in expectation, for a nonzero \( d \) divisible by \( q \), the density of 3-APs with common difference \( d \) of \( f_3 \) is at most \( \alpha^3(1 - \epsilon) \).

We let \( f = f_3 \). Using concentration inequalities, we can show that with positive probability (over the randomness in the third step), for any \( d \in [N/2] \),

\[
\mathbb{E}_{x \in [N-2d]}[f(x)f(x + d)f(x + 2d)] \leq \alpha^3(1 - \epsilon),
\]

proving Theorem 3.3. The details of this construction are contained in Section 7.

4. Upper Bound

In this section, we prove Theorem 1.4 and Theorem 1.3, showing the existence of a popular difference for 3-APs when \( |G| \geq \text{tower}(C \log(1/\epsilon)) \) or \( N \geq \text{tower}(C \log(1/\epsilon)) \). Here \( G \) always denotes a finite abelian group of odd order. For \( x \in G \), we write \( x/2 \) to mean the inverse of the isomorphism \( x \mapsto 2x \). In Section 4.1, we give some preliminaries on Bohr sets, which is an important tool to make Fourier analysis work over general abelian groups. In Section 4.2, we give the complete proofs of Theorem 1.4 and Theorem 1.3.

4.1. Bohr sets. Denote the distance from \( x \in \mathbb{R} \) to the nearest integer by \( \|x\|_{\mathbb{R}/\mathbb{Z}} := \min_{n \in \mathbb{Z}} |x - n| \). Let \( \arg(z) \) denote the argument of \( z \in \mathbb{C} \), so that \( \arg(e^{it}) \in [0, 2\pi] \) and \( e^{it} = e^{i\arg(e^{it})} \).

**Definition 4.1.** Let \( G \) be an abelian group of odd order. For a subset \( S \subseteq \hat{G} \) and a parameter \( \rho \in [0, 1] \), define the **Bohr set** \( B(S, \rho) = \{ x \in G : \| \arg(\chi(x))/(2\pi) \|_{\mathbb{R}/\mathbb{Z}} \leq \rho \forall \chi \in S \} \). We call \( S \) the **frequency set** of the Bohr set \( B(S, \rho) \) and \( \rho \) the **radius**. The **codimension** of the Bohr set is \( |S| \).

We often drop \( S \) and \( \rho \) from the notation and denote the Bohr set by \( B \) if it is clear from context. Given a Bohr set \( B \), we write \( S(B) \) to denote the frequency set of \( B \). For a real number \( \nu \geq 1 \), we denote by \( (B)_\nu \) the Bohr set with the same frequency set and scaled radius \( B(S, \nu \rho) \).
Define the normalized indicator function of Bohr sets by
\[ \beta_B(x) = \frac{|G|B(x)}{|B|}, \]
where we chose the normalization so that \( E_x[\beta_B(x)] = 1 \). Define
\[ \phi_B(x) = \beta_B \ast \beta_B(x). \]
Then we also have \( E_x[\phi_B(x)] = 1 \). The functions \( \beta_B \) can be thought of as the density with respect to the uniform distribution on \( G \) of the uniform distribution on \( B \), and \( \phi_B \) is the density of a smoothed version of the uniform distribution on \( B \). In general, a function \( \tau : G \rightarrow [0, \infty) \) with \( E\tau = 1 \) can be thought of as the density of a distribution with respect to the uniform distribution on \( G \).

**Conventions.** For simplicity of notation, we often omit the subscript \( B \) and use consistent subscripts throughout. For example, \( \phi = \phi_B, \beta = \beta_B, \phi_1 = \phi_{B_1}, \beta_1 = \beta_{B_1} \).

As introduced by Bourgain [7], it is often useful to work with regular Bohr sets, those for which a small change to the radius does not significantly change the size of the Bohr set.

**Definition 4.2.** A Bohr set \( B = B(S, \rho) \) of codimension \( d \) is regular if for all \( \delta \leq 1/(80d) \),
\[ |(B)_{1+\delta} \setminus (B)_{1-\delta}| \leq 160d|B|. \]

In the next proposition, we state some basic properties of Bohr sets, whose proofs can be found in [32, Section 4.4]. Denote by \( 2 \cdot X = \{2x, x \in X\} \) the dilation of \( X \) by a factor 2. Recall that for a character \( \chi \), we denote by \( \chi^{1/2} \) the character given by \( x \mapsto \chi(x/2) \).

**Proposition 4.3.** The following properties hold.

(i) \( |B(S, \rho)| \geq |G|\rho^{|S|} \).

(ii) \( B(S, \rho) + B(S, \rho') \subseteq B(S, \rho + \rho') \).

(iii) For every Bohr set \( B = B(S, \rho) \), there exists \( \nu \in [1/2, 1] \) such that \( (B)_{\nu} \) is regular.

(iv) For every Bohr set \( B = B(S, \rho) \), \( 2 \cdot B \) is also a Bohr set with frequency set \( \{\chi^{1/2} : \chi \in S\} \) and radius \( \rho \). Furthermore \( 2 \cdot B \subseteq (B)_2 \).

(v) For every Bohr set \( B = B(S, \rho) \), \( (2 \cdot B)_\nu = 2 \cdot (B)_\nu \). Hence, if \( (B)_\nu \) is regular then \( (2 \cdot B)_\nu \) is also regular.

**Definition 4.4.** Let \( \phi \) be function on \( G \) with \( E\phi = 1 \). Define \( f_\phi(x) = (f \ast \phi)(x) = E_y[f(x-y)\phi(y)] \).

The next estimate shows that regular Bohr sets are essentially invariant under convolutions with a distribution whose support is contained in a Bohr set with the same frequency set and smaller radius. This is analogous to the additive closure property of subgroups.

**Proposition 4.5.** Let \( B \) be a regular Bohr set of codimension \( d \), \( \beta = \beta_B \) and \( \phi = \phi_B \). Let \( \nu \leq 1/(80d) \). Let \( \tau \) be a function supported in \( B_\nu \) with \( E\tau = 1 \). Then
\[ E_x [|(\beta \ast \tau)(x) - \beta(x)|] \leq 160\nu d, \tag{4} \]
and
\[ E_x [|(\phi \ast \tau)(x) - \phi(x)|] \leq 160\nu d. \tag{5} \]
Furthermore, for any \( f : G \rightarrow [0, 1] \), letting \( \kappa \) be either \( \beta \) or \( \phi \),
\[ E_x [|(f_\tau \ast \kappa)(x) - f_\kappa(x)|] \leq 160\nu d. \tag{6} \]

**Proof.** We have
\[ (\beta \ast \tau)(x) = E_y[\beta(x-y)\tau(y)] \in [0, |G|/|B|]. \]
The support of $\beta \ast \tau$ is a subset of $\text{supp}(\beta) + \text{supp}(\tau) \subseteq B + (B)_\nu \subseteq (B)_{1+\nu}$. Thus, if $x \notin (B)_{1+\nu}$, then $(\beta \ast \tau)(x) = 0 = \beta(x)$. Furthermore, if $x \in (B)_{1-\nu}$, then for all $y \in \text{supp}(\tau)$, $x - y \in B$, so $(\beta \ast \tau)(x) = |G|/|B| = \beta(x)$. Hence,

$$E_x[|\langle \beta \ast \tau \rangle(x) - \beta(x)|] \leq \frac{1}{|G|} \sum_{x \in (B)_{1+\nu} \setminus (B)_{1-\nu}} \frac{|G|}{|B|} = \frac{|(B)_{1+\nu} \setminus (B)_{1-\nu}|}{|B|} \leq 160\nu d,$$

giving (4).

For (5), note that

$$E_x[|\langle \phi \ast \tau \rangle(x) - \phi(x)|] = E_x[|\langle \beta \ast \beta \ast \tau \rangle(x) - (\beta \ast \beta)(x)|] = E_x[|E_y[\beta(y)\langle \beta \ast \tau \rangle(x - y) - \beta(x - y)]|] \leq E_y[|E_y[\beta(y)E_x[|\langle \beta \ast \tau \rangle(x - y) - \beta(x - y)]|] \leq 160\nu dE_y[\beta(y)] = 160\nu d.$$

For (6), we have

$$E_x[|f_\tau \ast \kappa - f_\kappa|] = E_x[|E_y[f(y)(\kappa \ast \tau)(x - y) - f(y)\kappa(x - y)]|] \leq E_{x,y}[|((\kappa \ast \tau)(x - y) - \kappa(x - y))|] \leq 160\nu d.$$

The next lemma says that if $B_2 \subseteq (B_1)_{\nu/2}$ and $\tau = \beta_{B_2}$ or $\tau = \phi_{B_2}$, then the $k$-th moment of $f_\tau$ is at least the $k$-th moment of $f_{\phi_{B_1}}$ up to a small error term.

**Lemma 4.6.** Let $f : G \to [0,1]$. Let $B_1, B_2$ be regular Bohr sets such that $B_1$ has codimension $d_1$. Let $\phi_1 = \phi_{B_1}$, $\beta_2 = \beta_{B_2}$ and $\phi_2 = \phi_{B_2}$. Let $k \geq 1$ be an integer and $\nu \leq 1/(80d_1)$. Then following statements hold.

If $B_2 \subseteq (B_1)_{\nu/2}$, then

$$E_x[f_{\phi_2}(x)^k] \geq E_x[f_{\phi_1}(x)^k] - 160\nu d_1 k. \tag{7}$$

If $B_2 \subseteq (B_1)_{\nu}$, then

$$E_x[f_{\beta_2}(x)^k] \geq E_x[f_{\phi_1}(x)^k] - 160\nu d_1 k. \tag{8}$$

**Proof.** By (6) of Proposition 4.5, applied with $B = B_1$, $\tau = \phi_2$ (noting that $\text{supp}(\phi_2) \subseteq B_2 + B_2 \subseteq (B_1)_{\nu}$) and $\kappa = \phi_1$,

$$E_x[|f_{\phi_1}(x) - E_d[f_{\phi_2}(x + d)\phi_1(d)]|] \leq 160\nu d_1.$$

Thus,

$$E_x\left[|f_{\phi_1}(x)^k - E_d[f_{\phi_2}(x + d)\phi_1(d)]^k|\right] \leq 160\nu d_1 k.$$

By Jensen’s inequality applied to the convex function $t \mapsto t^k$, we obtain

$$E_x[f_{\phi_1}(x)^k] \leq E_x[|E_d[f_{\phi_2}(x + d)\phi_1(d)]|^k] + 160\nu d_1 k \leq E_x[d^k f_{\phi_2}(x + d)\phi_1(d)] + 160\nu d_1 k = E_x[f_{\phi_2}(x)^k] + 160\nu d_1 k.$$

The proof of (8) follows similarly. □
4.2. Proofs of Theorem 1.4 and Theorem 1.3. In the following, we prove two results that are used in the proof of Theorem 1.4, the counting lemma (Lemma 4.7) and the mean-cube density increment (Lemma 4.9).

For a function $\phi$ on $G$ with $E\phi = 1$ and a function $f : G \to [0, 1]$, we denote

$$\Lambda_\phi(f) = E_{x,d}[f(x)f(x+d)f(x+2d)\phi(d)].$$

**Lemma 4.7** (Counting lemma). Let $B_1 = B(S_1, \rho_1)$ and $B_2 = B(S_2, \rho_2)$ be two Bohr sets. Let $\phi_1 = \phi_{B_1}$ and $\phi_2 = \phi_{B_2}$. Then

$$\Lambda_{\phi_1}(f_{\phi_2}) \geq \Lambda_{\phi_1}(f) - 3 \sup_{\chi \in G} \left| \hat{f}(\chi) - \hat{f}_{\phi_2}(\chi) \right| E[f(x)^2] \left( \frac{|G|}{|B_1|} \right)^{1/2}.$$ 

**Proof.** By expanding in the Fourier basis,

$$\Lambda_{\phi_1}(f) = \sum_{\chi_1 \chi_2 \chi_3 = 1} \hat{f}(\chi_1)\hat{f}(\chi_2)\hat{f}(\chi_3)\hat{\phi}_1(\chi_2^{-1}\chi_3^{-2}).$$

By similar expansion for $\Lambda_{\phi_1}(f_{\phi_2})$, we can write $\Lambda_{\phi_1}(f) - \Lambda_{\phi_1}(f_{\phi_2})$ as

$$\sum_{\chi_1 \chi_2 \chi_3 = 1} \left( \hat{f}(\chi_1) - \hat{f}_{\phi_2}(\chi_1) \right)\hat{f}(\chi_2)\hat{f}(\chi_3)\hat{\phi}_1(\chi_2^{-1}\chi_3^{-2})$$

$$+ \sum_{\chi_1 \chi_2 \chi_3 = 1} \hat{f}_{\phi_2}(\chi_1)(\hat{f}(\chi_2) - \hat{f}_{\phi_2}(\chi_2))\hat{f}(\chi_3)\hat{\phi}_1(\chi_2^{-1}\chi_3^{-2})$$

$$+ \sum_{\chi_1 \chi_2 \chi_3 = 1} \hat{f}_{\phi_2}(\chi_1)\hat{f}_{\phi_2}(\chi_2)(\hat{f}(\chi_3) - \hat{f}_{\phi_2}(\chi_3))\hat{\phi}_1(\chi_2^{-1}\chi_3^{-2}).$$

Note that

$$\left| \sum_{\chi_2 \chi_3 = \chi^{-1}} \hat{f}(\chi_2)\hat{f}(\chi_3) \right| \leq \left( \sum_{\chi_2} |\hat{f}(\chi_2)|^2 \right) \left( \sum_{\chi_3} |\hat{f}(\chi_3)|^2 \right) \quad \text{(Cauchy-Schwarz)}$$

$$= \sum_{\chi} |\hat{f}(\chi)|^2$$

$$= E[f(x)^2], \quad \text{(Parseval)} \quad (9)$$

and

$$\sum_{\phi_i} |\hat{\phi}_1| \leq \left( \sum_{\chi} |\hat{\phi}_1|^2 \right)^{1/2} \quad \text{(Cauchy-Schwarz)}$$

$$= (E_x[\phi_1(x)^2])^{1/2}. \quad \text{(Parseval)} \quad (10)$$
We can now bound the first term as
\[
\left| \sum_{x_1x_2x_3=1} (\tilde{f}(x_1) - \phi_1(x_1)) \tilde{f}(x_2) \tilde{f}(x_3) \phi_1(x_2^{-1}x_3^{-2}) \right|
\]
\[
\leq \sup_x |\tilde{f}(x) - \phi_2(x)| \cdot \sum_x |\phi_1(x)| \left| \sum_{x_2x_3=1} \tilde{f}(x_2) \tilde{f}(x_3) \right|
\]
\[
\leq \sup_x |\tilde{f}(x) - \phi_2(x)| \cdot \sum_x |\phi_1(x)| \mathbb{E}[f(x)^2]
\]
(by (9))
\[
\leq \sup_x |\tilde{f}(x) - \phi_2(x)| \mathbb{E}[f(x)^2] \left( \frac{|G|}{|B|} \right)^{1/2},
\]
where in the last inequality, we used the fact that \( \sup_x \phi_1(x) \leq |G|/|B_1| \) and \( \mathbb{E}\phi_1 = 1 \).

The two remaining terms are bounded similarly.

We next state and prove the mean-cube density increment lemma. We make use of the following classical inequality in the proof of the lemma:

**Theorem 4.8** (Schur’s inequality). For real numbers \( a, b, c \geq 0 \), one has
\[
a^3 + b^3 + c^3 + 3abc \geq a^2b + b^2a + a^2c + c^2a + b^2c + c^2b. \quad (11)
\]

**Lemma 4.9** (Mean-cube density increment). Let \( f : G \to [0, 1] \). Let \( B_1 = B(S_1, \rho_1) \) and \( B_2 = B(S_2, \rho_2) \) be two regular Bohr sets with codimension \( d_1, d_2 \) respectively. Let \( \phi_1 = \phi_{B_1} \) and \( \phi_2 = \phi_{B_2} \). Let \( \nu \leq 1/(100d_1) \). Assume that \( B_2 \subseteq (B_1)_{\nu^2/8} \cap (2 \cdot B_1)_{\nu^2/8} \). For every regular Bohr set \( B = (B_1)_{\delta \nu/2} \) with \( \delta \in [1/2, 1] \) (Proposition 4.3 guarantees the existence of such \( \delta \) that makes \( B \) regular), we have, setting \( \phi = \beta \cdot \phi_B \),
\[
\Lambda_{\phi}(f_{\phi_2}) \geq 2\mathbb{E}[f_{\phi_1}(x)^3] - \mathbb{E}[f_{\phi_2}(x)^3] - 1920\nu d_1.
\]

**Proof.** Let \( \beta = \beta_B \) and \( \phi = \beta_B \). We denote by \( \tilde{\beta} \) the normalized measure associated with the Bohr set \( 2 \cdot B \), and denote \( \phi = \beta \cdot \tilde{\beta} \).

Applying Schur’s inequality (Theorem 4.8) with \( a = f_{\phi_2}(x) \), \( b = f_{\phi_2}(x + d) \), and \( c = f_{\phi_2}(x + 2d) \) for each \( x \) and \( d \), and using linearity of expectation, we have
\[
\Lambda_{\phi}(f_{\phi_2}) = \mathbb{E}_{x,d}[f_{\phi_2}(x)f_{\phi_2}(x + d)f_{\phi_2}(x + 2d)\phi(d)]
\]
\[
\geq \frac{4\mathbb{E}_{x,d}[f_{\phi_2}(x)^2f_{\phi_2}(x + d)\phi(d)] + 2\mathbb{E}_{x,d}[f_{\phi_2}(x)^2f_{\phi_2}(x + 2d)\phi(d)]}{3} - \mathbb{E}_{x}[f_{\phi_2}(x)^3].
\]

By Property (v) in Proposition 4.3, since \( B \) is a regular Bohr set, \( 2 \cdot B \) is also a regular Bohr set. The codimensions of \( 2 \cdot B \) and \( B \) are \( d_1 \). We have that \( B_2 \subseteq (B_1)_{\nu^2/8} \cap (2 \cdot B_1)_{\nu^2/8} \subseteq (B)_{\nu/2} \cap (2 \cdot B)_{\nu/2} \). By (6) in Proposition 4.5, applied with the Bohr set \( B, \kappa = \phi, \tau = \phi_2 \),
\[
\mathbb{E}_x[|\mathbb{E}_d[f_{\phi_2}(x + d)\phi(d)] - f_{\phi}(x)|] \leq 160\nu d_1.
\]

Hence,
\[
\mathbb{E}_{x,d}[f_{\phi_2}(x)^2f_{\phi_2}(x + d)\phi(d)] \geq \mathbb{E}_x[f_{\phi_2}(x)^2f_{\phi}(x)] - 160\nu d_1.
\]

Similarly,
\[
\mathbb{E}_{x,d}[f_{\phi_2}(x)^2f_{\phi_2}(x + 2d)\phi(d)] \geq \mathbb{E}_x[f_{\phi_2}(x)^2f_{\phi}(x)] - 160\nu d_1.
\]
We have

\[ E_x[f_{\phi_2}(x)^2|f_{\phi}(x)|] = E_{x,y}[f_{\phi_2}(x)^2\beta(y)f_{\beta}(x-y)] \]
\[ = E_{x,y}[f_{\phi_2}(x)^2\beta(y)f_{\beta}(x+y)] \] (using \( \beta(-y) = \beta(y) \))
\[ = E_x[f_{\beta}(x)E_y[f_{\phi_2}(x-y)^2\beta(y)]] \]
\[ \geq E_x[f_{\beta}(x)E_y[f_{\phi_2}(x-y)\beta(y)]^2]. \] (Cauchy-Schwarz)

By (6) in Proposition 4.5, applied with the Bohr set \( B, \kappa = \beta, \tau = \phi_2, \)
\[ E_x [|E_y[f_{\phi_2}(x-y)\beta(y)] - f_{\beta}(x)|] \leq 160\nu d_1. \]

Thus,
\[ E_x[f_{\beta}(x)E_y[f_{\phi_2}(x-y)\beta(y)]^2] \geq E_x[f_{\beta}(x)^3] - 320\nu d_1. \]

Hence,
\[ E_x[f_{\phi_2}(x)^2f_{\phi_2}(x+d)\phi(d)] \geq E_x[f_{\beta}(x)^3] - 480\nu d_1. \]

Similarly, noting that \( B_2 \subseteq (2 \cdot B)_{\nu/2}, \) we have
\[ E_{x,d}[f_{\phi_2}(x)^2f_{\phi_2}(x+2d)\phi(d)] = E_{x,d}[f_{\phi_2}(x)^2f_{\phi_3}(x+d)\tilde{\phi}(d)] \geq E_x[f_{\tilde{\beta}}(x)^3] - 480\nu d_1. \]

Since \( B \subseteq (B_1)_{\delta_2}, \) by (8) of Lemma 4.6 applied to the Bohr sets \( B_1 \) and \( B, \)
\[ E_x[f_{\beta}(x)^3] \geq E_x[f_{\phi_1}(x)^3] - 480\nu d_1. \]

Thus,
\[ E_x[f_{\phi_2}(x)^2f_{\phi_2}(x+d)\phi(d)] \geq E_x[f_{\phi_1}(x)^3] - 480\nu d_1. \]

Similarly, since \( 2 \cdot B \subseteq (B_2)_{\nu} \subseteq (B_1)_{\delta_2}, \) by (8) of Lemma 4.6 applied with the Bohr sets \( B_1 \) and \( 2 \cdot B, \)
\[ E_x[f_{\beta}(x)^3] \geq E_x[f_{\phi_1}(x)^3] - 480\nu d_1. \]

Thus,
\[ E_x[f_{\phi_2}(x)^2f_{\phi_2}(x+2d)\phi(d)] \geq E_x[f_{\phi_1}(x)^3] - 960\nu d_1. \]

Combining, we get
\[ \Lambda_{\phi}(f_{\phi_2}) \geq 2E[f_{\phi_1}^3] - E[f_{\phi_2}^3] - 1920\nu d_1. \]

**Lemma 4.10.** Let \( \alpha, \epsilon > 0. \) Let \( a_1, a_2, \ldots \) be a sequence of positive real numbers such that \( \alpha^3 \leq a_i \leq 1 \) for all \( i. \) Then for some \( i \leq 2 \log_2(2/\epsilon), \) \( 2a_i - a_{i+1} \geq \alpha^3 - \epsilon/2. \)

**Proof.** Assume for the sake of contradiction that for all \( i \leq 2 \log_2(2/\epsilon), \)
\[ a_{i+1} \geq 2a_i - \alpha^3 + \epsilon/2. \]

Then \( a_2 \geq \alpha^3 + \epsilon/2 \) since \( a_1 \geq \alpha^3. \) For \( 2 \leq i \leq 2 \log_2(2/\epsilon), \) \( a_{i+1} - \alpha^3 \geq 2(a_i - \alpha^3), \) so \( a_{i+1} \geq \alpha^3 + 2\epsilon/2. \) Since \( a_{i+1} \leq 1 \) for all \( i, \) we arrive at a contradiction since \( 2^{2\log_2(2/\epsilon)}\epsilon/2 > 1. \)

**Proof of Theorem 1.4.** We define inductively parameters \( \rho_i \) such that \( \rho_1 = \epsilon^{10}, \) and for \( i \geq 2, \)
\[ \rho_i = \exp(-\rho_{i-1}^{-5}). \]

Let \( S_1 = \{ \chi \in \hat{G} : |\hat{f}(\chi)| \geq \rho_1/2 \}, \) and for \( i \geq 2, S_i = \{ \chi \in \hat{G} : |\hat{f}(\chi)| \geq \rho_1/2 \} \cup \{ \chi^{1/2} : \chi \in S_{i-1} \}. \)

By Parseval’s identity, \( |\{ \chi \in \hat{G} : |\hat{f}(\chi)| \geq \rho_1/2 \}| \leq 4\rho_i^{-2}, \) so \( |S_i| \leq \sum_{j=1}^{i} 4\rho_j^{-2} < 5\rho_i^{-2}. \) Let \( B_i = B(S_i, \rho_i/(4\pi)). \) We first note that for \( \chi \in S_i, \)
\[ |1 - \tilde{\beta}_i(\chi)| = |1 - E_x[\beta_i(x)\chi(x)]| \leq E_x[|\beta_i(x)|] |\chi(x) - 1| \leq \rho_i/2, \]
since \( ||\arg(\chi(x))/(2\pi)||_{\mathbb{R}/2} \leq \rho_i/(4\pi) \) for all \( x \) such that \( \beta_i(x) \neq 0. \) Hence,
\[ |1 - \tilde{\phi}_i(\chi)| = |1 - \tilde{\beta}_i(\chi)^2| \leq \rho_i. \]
Thus, for $\chi \in S_i$, 
\[
|\hat{f}(\chi)| \leq \rho_i,
\]
and for $\chi \notin S_i$, $|\hat{f}(\chi)| \leq \rho_i/2$ so 
\[
|\hat{f}(\chi)| \leq \rho_i.
\]

Observe that $E[f_{\phi_i}^3] \geq \alpha^3$ by convexity, and $E[f_{\phi_i}^3] \leq 1$ for all $i$. By Lemma 4.10, there exists $i \leq 2 \log_2(2/\epsilon)$ such that 
\[
2E[f_{\phi_i}(x)^3] - E[f_{\phi_{i+1}}(x)^3] \geq \alpha^3 - \epsilon/2.
\]

Fix such an $i$. We have $B_{i+1} \subseteq (B_i)_{\delta_\epsilon^2/8} \cap (2 \cdot B_i)_{\delta_\epsilon^2/8}$ since for any $\chi \in S_i$ and $x \in B_{i+1}$, 
\[
\|\text{arg}(\chi(x))/(2\pi)\|_{\mathbb{R}/\mathbb{Z}} \leq \rho_{i+1} \leq \nu_i^2 \rho_i/8,
\]
and furthermore $\chi^{1/2} \in S_{i+1}$ so 
\[
\|\text{arg}(\chi^{1/2}(x))/(2\pi)\|_{\mathbb{R}/\mathbb{Z}} \leq \rho_{i+1} \leq \nu_i^2 \rho_i/8.
\]

By Lemma 4.9, there exists a regular Bohr set $B = (B_i)_{\delta_\epsilon/2}$ for $\delta \in [1/2, 1]$ such that 
\[
\Lambda_{\phi}(f_{\phi_{i+1}}) \geq 2E[f_{\phi_i}(x)^3] - E[f_{\phi_{i+1}}(x)^3] - 1920 \nu_i |S_i| \geq \alpha^3 - \epsilon/2 - 1920 \nu_i |S_i|.
\]

By Lemma 4.7, 
\[
\Lambda_{\phi}(f_{\phi_{i+1}}) \leq \Lambda_{\phi}(f) + \sup_{\chi} |\hat{f}(\chi) - \hat{f}_{\phi_{i+1}}(\chi)| E[f(x)^2] \left(\left|\frac{G}{B}\right| \right)^{1/2} \leq \Lambda_{\phi}(f) + \rho_{i+1}(4\nu_i^{-1} \rho_i^{-1})^5 \rho_i^{-2}.
\]

By our choice of $\rho_i, \nu_i$, we have 
\[
\rho_{i+1}(4\nu_i^{-1} \rho_i^{-1})^5 \rho_i^{-2} \leq \exp(5\rho_i^{-2} \cdot \log(10^6 \epsilon^{-1} \rho_i^{-3})) \exp(-\rho_i^{-5}) < \epsilon/8,
\]
noting that $\rho_i \leq \rho_1 \leq \epsilon^{10}$. Hence, 
\[
\Lambda_{\phi}(f) \geq \alpha^3 - \epsilon/2 - 1920 \nu_i d_i - \rho_{i+1}(4\nu_i^{-1} \rho_i^{-1})^5 \rho_i^{-2} > \alpha^3 - \epsilon/2 - \epsilon/4 - \epsilon/8 = \alpha^3 - 7\epsilon/8.
\]

Observe that there exists an absolute constant $C' > 0$ so that $\rho_i \geq 1/\text{tower}(C' i)$. Furthermore, the codimension of $B_i$ is bounded above by $5\rho_i^{-2}$ and the radius of $B_i$ is $\rho_i/(4\pi)$. Hence, we obtain a Bohr set $B$ with size at least $|G|/\text{tower}(10C' \log(1/\epsilon))$ such that $\Lambda_{\phi}(f) \geq \alpha^3 - 7\epsilon/8$. Hence, for a sufficiently large constant $C > 0$, assuming that $|G| \geq \text{tower}(C \log(1/\epsilon))$, we have 
\[
E_{x,d}[f(x)f(x + d)f(x + 2d)\phi(d)\mathbb{I}(d \neq 0)] \geq \Lambda_{\phi}(f) - \frac{1}{|B|} \geq \alpha^3 - \epsilon.
\]
Thus, there exists $d \neq 0$ such that 
\[
E_{x}[f(x)f(x + d)f(x + 2d)] \geq \alpha^3 - \epsilon.
\]

\[
\square
\]

Proof of Theorem 1.3. We can assume without loss of generality that $N$ is odd by possibly increasing $N$ by 1. Let $G = \mathbb{Z}/N$. We repeat the proof of Theorem 1.4 with the inclusion of the character $\chi_0(x) = e^{2\pi ix/N}$ in the sets $S_i$. We then obtain a Bohr set $B$ whose frequency set contains $\chi_0$ such that $B$ has size at least $|G|/\text{tower}(C' \log(1/\epsilon))$ and $\Lambda_{\phi}(f) \geq \alpha^3 - 7\epsilon/8$. Assuming that $N \geq \text{tower}(C \log(1/\epsilon))$ for sufficiently large $C$, following the last step in the proof of Theorem 1.4, we obtain a positive integer $d < N/2$ such that $d \in \text{supp}(\phi)$ when viewed an element in $\mathbb{Z}/N$ and 
\[
E_{x}[f(x)f(x + d)f(x + 2d)] \geq \alpha^3 - 15\epsilon/16.
\]
Since $d \in \text{supp}(\phi) \subseteq B + B$ and $\chi_0$ is in the frequency set defining $B$, $\|\arg(\chi_0(d))/(2\pi)\|_{R/Z} \leq 2\rho \leq 2\epsilon^{10}$. Thus, as a positive integer less than $N/2$, we have $d < 2\epsilon^{10}N$. Thus, restricting to $x \in [N - 2d]$ in the above expectation, we have
\[
\sum_{x \in [N-2d]} f(x)f(x+d)f(x+2d) \geq N(\alpha^3 - 15\epsilon/16) - 2d \geq N(\alpha^3 - \epsilon).
\]
\[
\square
\]

5. LOWER BOUND CONSTRUCTION: PREPARATIONS

We assume throughout this section that $N$ is an odd prime number. As a building block in our construction, we will make use of a function $g$ which has relatively low 3-AP density (considerably smaller than the random bound given the density of $g$), but behaves random-like in the following way. If $a_1, a_2, \ldots, a_h$ are chosen independently and uniformly at random from the nonzero elements of $\mathbb{Z}_N$, then with high probability, for all $b_1, b_2, \ldots, b_h \in \mathbb{Z}_N$,
\[
E_x \left[ \prod_{j=1}^h g(a_jx + b_j) \right] = E[g]^h.
\]
(12)

In the following, we identify $\hat{\mathbb{Z}}_N$ with $\mathbb{Z}_N$, so that we write
\[
\hat{g}(r) = E_{x \in \mathbb{Z}_N} \left[ g(x)e\left(\frac{rx}{N}\right) \right].
\]

**Lemma 5.1.** Suppose $g : \mathbb{Z}_N \rightarrow [0, 1]$ and $a_1, a_2, \ldots, a_h \in \mathbb{Z}_N \setminus \{0\}$ satisfy the following properties:

(i) The support of $\hat{g}$ has size at most $\ell$.

(ii) For all $r_1, r_2, \ldots, r_h \in \mathbb{Z}_N$ such that $\sum_{j=1}^h r_ja_j = 0$ and $(r_1, r_2, \ldots, r_h) \neq (0, 0, \ldots, 0)$, there is some $j \in [h]$ such that $r_j$ is not contained in the support of $\hat{g}$.

Then for all $b_1, b_2, \ldots, b_h \in \mathbb{Z}_N$,
\[
E_x \left[ \prod_{j=1}^h g(a_jx + b_j) \right] = E[g]^h.
\]

Furthermore, if $a_1, a_2, \ldots, a_h$ are chosen from $\mathbb{Z}_N \setminus \{0\}$ uniformly and independently at random, then Property (ii) is satisfied with probability at least $1 - \ell^h/(N - 1)$.

**Proof.** By the Fourier inversion formula,
\[
E_x \left[ \prod_{j=1}^h g(a_jx + b_j) \right] = E_x \left[ \prod_{j=1}^h \left( \sum_{r_j \in \mathbb{Z}_N} \hat{g}(r_j)e\left(\frac{r_ja_jx + r_jb_j}{N}\right) \right) \right] = \sum_{r_1, r_2, \ldots, r_h \in \mathbb{Z}_N} \prod_{j=1}^h e\left(\frac{r_jb_j}{N}\right) \hat{g}(r_j) \cdot E_x \left[ e\left(\frac{\sum_{j=1}^h r_ja_jx}{N}\right) \right] = \sum_{r_1, r_2, \ldots, r_h \in \mathbb{Z}_N} \sum_{\sum_{j=1}^h r_ja_j = 0} \prod_{j=1}^h e\left(\frac{r_jb_j}{N}\right) \hat{g}(r_j).
Note that $\prod_{j=1}^{h}g(0) = E[g]^h$. Consider $(r_1, r_2, \ldots, r_h) \neq (0, \ldots, 0)$ where $\sum_{j=1}^{h}r_ja_j = 0$. Property (ii) guarantees that $\hat{g}(r_j) = 0$ for some $j \in [h]$, so $\prod_{j=1}^{h}\hat{g}(r_j) = 0$. Hence, if $a_1, a_2, \ldots, a_h$ satisfy Property (ii), then

$$E_x\left[\prod_{j=1}^{h}g(a_jx + b_j)\right] = \hat{g}(0)^h = E[g]^h.$$  

Next, we show that if $a_1, a_2, \ldots, a_h$ are chosen uniformly and independently at random from $\mathbb{Z}_N \setminus \{0\}$, then Property (ii) is satisfied with probability at least $1 - \frac{\ell^h}{(N-1)}$. Indeed, consider a fixed $(r_1, r_2, \ldots, r_h) \neq (0, 0, \ldots, 0)$ such that $r_j$ is in the support of $\hat{g}$ for each $j \in [h]$. There exists $i \in [h]$ such that $r_i \neq 0$. For each fixed choice of $a_j$ for $j \in [h] \setminus \{i\}$, there is a unique choice of $a_i$ such that $\sum_{j=1}^{h}r_ja_j = 0$. Hence, the probability that $\sum_{j=1}^{h}r_ja_j = 0$ is at most $1/(N-1)$. By the union bound over the choice of $r_j$ in the support of $\hat{g}$, we obtain that Property (ii) is violated with probability at most $\frac{\ell^h}{(N-1)}$.  

Next, for each $\alpha \leq 1/2$, we construct a function $g_\alpha$ with mean $\alpha$ and prove that $g_\alpha$ has the desired properties in Lemma 5.1. We recall that the 3-AP density of a function $g$ is denoted by $\Lambda(g) = E_{x,d}[g(x)g(x+d)g(x+2d)]$.

**Lemma 5.2.** For $\alpha \leq \frac{1}{3}$, define a function $g_\alpha : \mathbb{Z}_N \to [0, 1]$ by

$$g_\alpha(x) = \alpha - \frac{\alpha \cos(2\pi x/N)}{2} - \frac{\alpha \cos(4\pi x/N)}{2}.$$  

Then $g_\alpha$ satisfies the following properties.

(i) $E[g_\alpha] = \alpha \leq \frac{1}{3}$ and $g_\alpha(x) \in [0, 2\alpha]$ for all $x \in \mathbb{Z}_N$.

(ii) $\Lambda(g_\alpha) = E_{x,d}[g_\alpha(x)g_\alpha(x+d)g_\alpha(x+2d)] = (1 - \frac{\alpha}{2})\alpha^3$, and $E_x[g_\alpha(x)^3] \leq \frac{3}{4}\alpha^3$.

(iii) For $h$ a positive integer, if we choose $a_1, a_2, \ldots, a_h$ uniformly and independently at random from $\mathbb{Z}_N \setminus \{0\}$, then with probability at least $1 - \frac{5h^2}{(N-1)}$, for all choices of $b_1, b_2, \ldots, b_h \in \mathbb{Z}_N$,

$$E_x\left[\prod_{j=1}^{h}g_\alpha(a_jx + b_j)\right] = E[g_\alpha]^h.$$  

(iv) $E_{x \neq y}[g_\alpha(x)g_\alpha(y)] \leq \alpha^2$.

(v) $E_x[g_\alpha(x)^2] = \frac{5}{4}\alpha^2$.

**Proof.** We first observe that

$$\widehat{g_\alpha}(0) = \alpha, \quad \text{and} \quad \widehat{g_\alpha}(1) = \widehat{g_\alpha}(N-1) = \widehat{g_\alpha}(2) = \widehat{g_\alpha}(N-2) = \alpha/4,$$

and for all $r \notin \{0, 1, 2, N-1, N-2\}$,

$$\widehat{g_\alpha}(r) = 0.$$  

Furthermore, it is clear from the definition of $g_\alpha$ that for all $x \in \mathbb{Z}_N$, $g_\alpha(x) \in [0, 2\alpha]$. Moreover, $E[g_\alpha] = \widehat{g_\alpha}(0) = \alpha$. This proves Property (i).

From (1),

$$\Lambda(g_\alpha) = \frac{1}{N} \sum_{r} \widehat{g_\alpha}(r)^2 \widehat{g_\alpha}(-2r) = \alpha^3 - 2\left(\frac{\alpha}{4}\right)^3 = \left(1 - \frac{1}{32}\right)\alpha^3,$$

and

$$E_x[g_\alpha^2] = \frac{5}{4}\alpha^2.$$
and
\[
E_x[g_\alpha(x)^3] = E_x \left[ \left( \sum_r \hat{g}_\alpha(r)e \left( \frac{r x}{N} \right) \right)^3 \right]
\]
\[
= \sum_{r_1, r_2, r_3 \in \mathbb{Z}_N} \hat{g}_\alpha(r_1)\hat{g}_\alpha(r_2)\hat{g}_\alpha(r_3) E_x \left[ e \left( \frac{r_1 x + r_2 x + r_3 x}{N} \right) \right]
\]
\[
= \sum_{\substack{r_1, r_2, r_3 \in \mathbb{Z}_N: \\
r_1 + r_2 + r_3 = 0}} \hat{g}_\alpha(r_1)\hat{g}_\alpha(r_2)\hat{g}_\alpha(r_3)
\]
\[
= \alpha^3 + 6\alpha^2 \frac{\alpha^2}{16} - 6\frac{\alpha^3}{64} < \frac{3}{2} \alpha^3.
\]
This proves Property (ii).

Property (iii) follows directly from Lemma 5.1 applied to the function \( g_\alpha \) and \( \ell = 5 \).

To prove Property (iv), notice that
\[
\left| \sum_x g_\alpha(x) \right|^2 \geq \frac{1}{N}(\sum_x g_\alpha(x))^2 = \alpha^2 N
\]
so
\[
E_x[g_\alpha(x)g_\alpha(y)] \leq \frac{\alpha^2 N^2 - \alpha^2 N}{N(N-1)} = \alpha^2.
\]
Finally, Property (v) follows from Parseval’s identity,
\[
E_x[g_\alpha(x)^2] = \sum_r |\hat{g}_\alpha(r)|^2 = \frac{5\alpha^2}{4}.
\]

\[\square\]

6. Lower bound construction for product groups

In this section, we prove Theorem 3.4. For convenience, we recall the theorem statement here.

**Theorem.** Let \( 0 < \alpha \leq 1/4 \), \( 0 < \epsilon \leq 20^{-9} \), and \( G = \mathbb{Z}_n \) where \( n \) is a positive integer such that there exist distinct primes \( m_1, \ldots, m_s \) with \( s \leq \log_{150} \left( \frac{\epsilon^{-1/4} \alpha^6}{8} \right) \) satisfying

- \( n = \prod_{j=1}^i m_j \),
- \( \epsilon^{-1/3}/2 < m_1 \leq \epsilon^{-1/3} \), and
- \( \epsilon^{-1} \) for \( i \geq 2 \), \( n_{i-1}^{6} < m_i \leq \exp(2^{-1} \cdot 64^{-2} \cdot 150^{i-1} \cdot \epsilon^{1/4} n_{i-1})/2 \) where \( n_i = \prod_{j=1}^i m_j \).

Then, there exists a function \( f : G \to [0,1] \) with \( E[f] = \alpha \) such that for any \( d \in G \setminus \{0\} \),
\[
E_x[f(x)f(x+d)f(x+2d)] \leq \alpha^2(1 - \epsilon).
\]
Furthermore, \( E_x[f(x)^3] \leq 3\alpha^3/2 \) and there exists \( \alpha' \in [\alpha, \alpha(1 + \epsilon^{1/4})] \) such that \( f(x) = \alpha' \) for at least a 3/4 fraction of \( x \in G \).

We first make a few notation conventions. Note that if \( n = \prod_{i=1}^s m_i \) for distinct primes \( m_i \), then
\[
G = \mathbb{Z}_n \cong \prod_{i=1}^s \mathbb{Z}_{m_i}.
\]

Each element of \( G \) can be represented by an \( s \)-tuple \((x_1, x_2, \ldots, x_s)\) where \( x_i \in \mathbb{Z}_{m_i} \). Let \( Q_i = \prod_{j=1}^i \mathbb{Z}_{m_i} \). We can think of \( Q_i \) as a quotient of \( G \) by the subgroup \( \mathcal{H}_i = \{ x \in G : x_j = 0 \text{ for all } j \leq i \} \). We identify \( \mathbb{Z}_{m_i} \) as the subgroup of \( Q_i \) consisting of elements with \( x_j = 0 \) for \( j < i \), and we identify the quotient \( Q_i/\mathbb{Z}_{m_i} \) with \( Q_{i-1} \). We hence use elements of \( Q_{i-1} \) to index \( \mathbb{Z}_{m_i} \)-cosets in \( Q_i \).

For an element \( x \in G \) or \( x \in Q_i \) with \( j \geq i \), we denote \( x[j] = (x_1, \ldots, x_i) \). For \( j < i \), we say that an element \( x \) of \( Q_i \) is a lift of an element \( y \) in \( Q_j \) if \( x[j] = y \). In the following discussion, when the level \( i \) is clear from context, if not specified otherwise, the 3-APs would refer to 3-APs in \( Q_i \).
6.1. The Construction. Let \( s = \lceil \log_{150} \left( \frac{e^{1/4}a^6}{8} \right) \rceil \). In each level \( i \), for \( i \in [s] \), we construct a function \( f_i : Q_i \rightarrow [0, 1] \). Finally, we let \( f = f_s : G \rightarrow [0, 1] \).

We introduce parameters \( \mu_1 = e^{1/4} \) and \( \mu_i = 150^{i-1}e^{1/4} \) for \( i \geq 2 \), where \( \alpha' = \alpha(1 + \frac{1}{m_{i-1}}) \).

In the first level, define \( f_1 : Q_1 \rightarrow [0, 1] \) by \( f_1(0) = 0 \) and \( f_1(x) = \alpha \left( 1 + \frac{1}{m_{i-1}} \right) \) for each \( x \in Q_1 \setminus \{0\} \).

For \( i \geq 2 \), we extend \( f_{i-1} : Q_{i-1} \rightarrow [0, 1] \) to a function \( \tilde{f}_{i-1} : Q_i \rightarrow [0, 1] \) by setting \( \tilde{f}_{i-1}(x) = f_{i-1}(x_{[i-1]}) \) for each \( x \in Q_i \). Let \( M_{i-1} \) be any set of \( \mu_{i-1} \) elements of \( Q_{i-1} \) so that \( f_{i-1}(x) = \alpha' \) for any \( x \in M_{i-1} \). In level \( i \), we define \( f_i \) to be a random function as follows.

For each \( x \in M_{i-1} \), we choose \( a_x \in \mathbb{Z}_{m_i} \setminus \{0\} \) and \( b_x \in \mathbb{Z}_{m_i} \) uniformly and independently at random. For each \( y \in Q_i \) such that \( y_{[i-1]} = x \), we define

\[
f_i(y) = g_{\alpha'}(a_x y_i + b_x) = \alpha' - \underbrace{\alpha' \cos(2\pi (a_x y_i + b_x)/m_i)}_{
2} - \underbrace{\alpha' \cos(4\pi (a_x y_i + b_x)/m_i)}_{
2},
\]

where \( g_{\alpha'} \) is the function with density \( \alpha' \) and with low 3-AP density defined earlier in Lemma 5.2. Otherwise, for \( x \notin M_{i-1} \) and \( y \in Q_i \) such that \( y_{[i-1]} = x \), we define

\[
f_i(y) = f_{i-1}(x).
\]

We refer to this as the random modification in level \( i \). This defines (random) \( f_i : Q_i \rightarrow [0, 1] \).

Finally, we let \( f = f_s : G \rightarrow [0, 1] \). We will show that with positive probability, for each level \( i \), we can pick \( f_i \) such that the function \( f \) has the desired properties claimed in Theorem 3.4.

6.2. Proof of Theorem 3.4. We first claim that the construction is feasible with the above choice of parameters. Note that \( \mu_1 \geq 1/m_1 \), so \( f_1(x) = \alpha' \) for all but a \( \mu_1 \) fraction of elements \( x \in Q_1 \).

For \( i \geq 2 \), observe that if \( f_i(y) \neq \alpha' \), then we must have \( y_{[i]} = 0 \) or \( y_{[j]} \in M_j \) for some \( j < i \). Thus, the fraction of \( y \in Q_i \) for which \( f_i(y) \neq \alpha' \) is at most \( \sum_{j=1}^{i} \mu_j \). Since

\[
\sum_{j=1}^{s} \mu_j < 2 \mu_s = 2 \cdot 150^{s-1}e^{1/4} \alpha'^{-6} < 1/4
\]

as \( s - 1 \leq \log_{150} \left( \frac{e^{1/4}a^6}{8} \right) < \log_{150} \left( \frac{e^{1/4}a^6}{8} \right) \), it is possible to choose \( M_i \) for each \( i \leq s - 1 \) such that \( f_i(x) = \alpha' \) for any \( x \in M_i \).

We next prove that the function \( f_i \) has density \( \alpha \) and \( f_i \) maps \( Q_i \) to \([0, 1]\). This is true for \( i = 1 \). Assume that \( f_{i-1} \) has density \( \alpha \) and takes values in \([0, 1]\), we show that \( f_i \) also has these properties. For \( x \in Q_i \) such that \( x_{[i-1]} \notin M_{i-1} \), \( f_i(x) = f_{i-1}(x_{[i-1]}) \in [0, 1] \). If \( x_{[i-1]} \in M_{i-1} \) then \( f_{i-1}(x_{[i-1]}) = \alpha' \leq 1/2 \). Hence, if \( x_{[i-1]} \in M_{i-1} \) then \( f_i(x) = g_{\alpha'}(ax_i + b) \) for some \( a \in \mathbb{Z}_{m_i} \setminus \{0\}, b \in \mathbb{Z}_{m_i} \). Since \( g_{\alpha'} \) also has density \( \alpha' \) and takes values in \([0, 1]\), we have \( f_i(x) = g_{\alpha'}(ax_i + b) \in [0, 1] \) and the density of \( f_i \) over the \( \mathbb{Z}_{m_i} \)-coset \( x_{[i-1]} \) is \( f_{i-1}(x_{[i-1]}) = \alpha' \). Hence, the density of \( f_i \) is the same as the density of \( f_{i-1} \), and \( f_i \) takes values in \([0, 1]\). By induction, the density of \( f_i \) is \( \alpha \) and the values of \( f_i \) are in \([0, 1]\) for all \( i \in [s] \).

We denote by \( E_{f_i} \) the expectation over the randomness of \( f_i \) (the local modifications in level \( i \)), conditioned on a fixed choice of \( f_{i-1} \). Furthermore, all of the probability we consider will be conditioned on this fixed choice of \( f_{i-1} \), hence in level \( i \) we only consider the randomness of the random modification in level \( i \).
The random modification in level $i$ has the following key property. For any $x = (x_1, \ldots, x_i) \in Q_i$ such that $x_{[i-1]} = (x_1, \ldots, x_{i-1}) \in M_{i-1}$ and $d \in Q_i \setminus \{0\}$ such that $d_{[i-1]} = 0 \in Q_{i-1}$, we have

$$
\mathbb{E}_{f_i}[f_i(x) f_i(x + d) f_i(x + 2d)] = \mathbb{E}_{aq \in \mathbb{Z}_{m_i} \setminus \{0\}, \beta \in \mathbb{Z}_{m_i}} [g_{a\alpha'}(a x_i + b) g_{a\alpha'}(a x_i + ad_i + b) g_{a\alpha'}(a x_i + 2ad_i + b)]
\leq \Lambda(g_{a\alpha'})
= \frac{31}{32} \epsilon^3.
$$

This is since when $a$ is chosen uniformly at random from $\mathbb{Z}_{m_i} \setminus \{0\}$ and $b$ is chosen uniformly at random from $\mathbb{Z}_{m_i}$, then for any fixed $x_i$ and nonzero $d_i$, $(a x_i + b, a x_i + ad_i + b, a x_i + 2ad_i + b)$ is distributed uniformly among all 3-APs with nonzero common difference in $\mathbb{Z}_{m_i}$.

We now proceed to prove that there exists a choice of the modification in each level so that for any $d \in G \setminus \{0\}$,

$$
\mathbb{E}_{x \in Q_i} [f_i(x) f_i(x + d) f_i(x + 2d)] \leq \alpha^3 (1 - \epsilon).
$$

The main idea is to maintain by induction that for any $i \in [s]$, we can choose $f_i$ which is a random modification of $f_{i-1}$ so that for any $d \in Q_i \setminus \{0\}$,

$$
\mathbb{E}_{x \in Q_i} [f_i(x) f_i(x + d) f_i(x + 2d)] \leq \alpha^3 (1 - \epsilon).
$$

For all $d$ such that $d_{[i-1]} = 0$, the above property follows from observation (14) and concentration inequalities. If $d_{[i-1]} \neq 0 \in Q_{i-1}$, we have $\mathbb{E}_{x \in Q_{i-1}} [f_{i-1}(x) f_{i-1}(x + d_{[i-1]}) f_{i-1}(x + 2d_{[i-1]})]$ is small by the induction hypothesis. We guarantee that with large probability, $\mathbb{E}_{x \in Q_i} [f_i(x) f_i(x + d) f_i(x + 2d)] = \mathbb{E}_{x \in Q_{i-1}} [f_{i-1}(x) f_{i-1}(x + d_{[i-1]}) f_{i-1}(x + 2d_{[i-1]})]$ for all $d$ such that $d_{[i-1]} \neq 0$, so we also have $\mathbb{E}_{x \in Q_i} [f_i(x) f_i(x + d) f_i(x + 2d)]$ is small. Combining these two cases, we obtain a modification $f_i$ of $f_{i-1}$ whose density of 3-APs with common difference $d$ is small for all nonzero $d \in Q_i$.

We now give the proof of Theorem 3.4.

**Proof of Theorem 3.4.** It is easy to see that

$$
\mathbb{E}_{x \in Q_i} [f_i(x)]^3 = \alpha^3 \left(1 + \frac{m_{i-1}}{m_i}ight)^3 (m_i - 1)
< \alpha^3 \left(1 + \frac{3}{m_1}\right) \leq \alpha^3 (1 + 6 \epsilon^{1/3}) \leq \alpha^3 (1 + 2 \epsilon^{1/4})
< \alpha^3 (1 + 2 \mu_1),
$$

for $\epsilon \leq 20^{-9}$. Inductively, if

$$
\mathbb{E}_{x \in Q_{i-1}} [f_{i-1}(x)]^3 \leq \alpha^3 (1 + 2 \mu_{i-1}),
$$

then

$$
\mathbb{E}_{x \in Q_i} [f_i(x)]^3 \leq \alpha^3 (1 + 2 \mu_{i-1}) + \frac{1}{2} \mu_i \alpha^3 < \alpha^3 (1 + 2 \mu_i),
$$

where the first inequality is by (ii) in Lemma 5.2 as we apply the local modification to a $\mu_i$ fraction of the $\mathbb{Z}_{m_i}$-cosets, getting at most a $\frac{1}{2} \alpha^3$ increment in the mean cube density over each of them, and the second inequality follows from our choice of parameters $\mu_i \geq 150 \mu_{i-1}$ for all $i \geq 2$.

Let $\mathcal{P}(i)$ be the property that for all $d \in Q_i \setminus \{0\}$,

$$
\mathbb{E}_{x \in Q_i} [f_i(x) f_i(x + d) f_i(x + 2d)] \leq \alpha^3 (1 - \epsilon).
$$

We prove by induction that in level $i$, the modifications can be chosen so that $\mathcal{P}(i)$ holds.
Consider the base case $i = 1$. Recall that $\epsilon^{-1/3}/2 \leq m_1 \leq \epsilon^{-1/3}$. For any $d \in Q_1 \setminus \{0\}$,

$$E_{x \in Q_1}[f_i(x)f_i(x + d)f_i(x + 2d)] = \alpha^3 \left( 1 + \frac{1}{m_1-1} \right)^3 \frac{m_1}{3} (m_1 - 3)$$

$$= \alpha^3 \frac{m_1^2 (m_1 - 3)}{(m_1 - 1)^3} = \alpha^3 \left( 1 - \frac{3m_1 - 1}{(m_1 - 1)^2} \right)$$

$$\leq \alpha^3 \left( 1 - \frac{1}{m_1^2} \right) \leq \alpha^3 (1 - \epsilon).$$

This establishes $\mathcal{P}(1)$. Next, we continue with the inductive step. Assume that $\mathcal{P}(i-1)$ holds. We prove that we can choose the modification in level $i$ so that $\mathcal{P}(i)$ also holds. This follows from the following two claims.

**Claim 6.1.** With probability larger than $1/2$, conditioned on a fixed choice of $f_{i-1}$ satisfying $\mathcal{P}(i-1)$, for all $d \in Q_i \setminus \{0\}$ with $d_{[i-1]} = 0$,

$$E_{x \in Q_i}[f_i(x)f_i(x + d)f_i(x + 2d)] \leq \alpha^3(1 - \epsilon).$$

**Claim 6.2.** With probability larger than $1/2$, conditioned on a fixed choice of $f_{i-1}$ satisfying $\mathcal{P}(i-1)$, for all $d \in Q_i \setminus \{0\}$ with $d_{[i-1]} \neq 0$,

$$E_{x \in Q_i}[f_i(x)f_i(x + d)f_i(x + 2d)] = E_{x \in Q_{i-1}}[f_{i-1}(x)f_{i-1}(x + d')f_{i-1}(x + 2d')].$$

Combining Claims 6.1 and 6.2, by the union bound, the modification in level $i$ fails to satisfy $\mathcal{P}(i)$ with probability strictly less than 1. Thus we can choose a modification satisfying $\mathcal{P}(i)$ in level $i$. This completes the induction. Thus, there exists $f = f_s$ which satisfies $\mathcal{P}(s)$, so for any nonzero $d$ in $G$,

$$E_{x \in G}[f(x)f(x + d)f(x + 2d)] = E_{x \in G}[f_s(x)f_s(x + d)f_s(x + 2d)] \leq \alpha^3(1 - \epsilon).$$

This completes the proof of Theorem 3.4.

Now we turn to the proofs of Claims 6.1 and 6.2.

**Proof of Claim 6.1.** Let $d \in Q_i \setminus \{0\}$ such that $d_{[i-1]} = 0$. By (14), for any $x \in Q_i$ with $x_{[i-1]} \in M_{i-1}$,

$$E_{f_i}[f_i(x)f_i(x + d)f_i(x + 2d)] \leq \frac{31}{32} \alpha^3.$$ 

Hence, for $y \in M_{i-1}$,

$$E_{f_i}E_{x \in Q_i, x_{[i-1]} = y}[f_i(x)f_i(x + d)f_i(x + 2d)] \leq \frac{31}{32} \alpha^3.$$ 

Note that the random variables $E_{x \in Q_i, x_{[i-1]} = y}[f_i(x)f_i(x + d)f_i(x + 2d)]$, for $y \in M_{i-1}$, are independent (under the randomness of the modification in level $i$, conditioned on a fixed choice of $f_{i-1}$). Thus the probability that

$$E_{y \in M_{i-1}}E_{x \in Q_i, x_{[i-1]} = y}[f_i(x)f_i(x + d)f_i(x + 2d)] \geq \frac{63}{64} \alpha^3$$

is at most $\exp(-2^{1 - 64^{-2} \mu_i n_{i-1} \alpha^6})$ by Hoeffding’s inequality.

For $d \in Q_i$ such that $d_{[i-1]} = 0$, we have

$$E_{x \in Q_i}[f_i(x)f_i(x + d)f_i(x + 2d)]$$

$$= E_{y \in Q_{i-1}}[f_{i-1}(y)^3] + \frac{|M_{i-1}|}{|Q_{i-1}|} \left( E_{y \in M_{i-1}}E_{x \in Q_i, x_{[i-1]} = y}[f_i(x)f_i(x + d)f_i(x + 2d)] - E_{y \in M_{i-1}}[f_{i-1}(y)^3] \right)$$

$$\leq \alpha^3(1 + 2\mu_{i-1}) + \mu_i \cdot \left( E_{y \in M_{i-1}}E_{x \in Q_i, x_{[i-1]} = y}[f_i(x)f_i(x + d)f_i(x + 2d)] - \alpha^3 \right)$$
where the first equality follows from \( f_i(x) = f_i(x + d) = f_i(x + 2d) = f_{i-1}(y) \) if \( d_{i-1} = 0 \) and \( x_{i-1} = y \notin M_{i-1} \), and the inequality follows from (15) and \( f_{i-1}(y) = \alpha' \) for \( y \in M_{i-1} \). Thus, if

\[
E_{y \in M_{i-1}} E_{x \in Q_i, x_{i-1} = y} [f_i(x) f_i(x + d) f_i(x + 2d)] \leq \frac{63}{64} \alpha_3^3,
\]

then

\[
E_{x \in Q_i} [f_i(x) f_i(x + d) f_i(x + 2d)] \leq \alpha_3^3 (1 + 2 \mu_{i-1}) - \mu_i \alpha_3^3 / 64.
\]

Since

\[
\alpha_3^3 (1 - \epsilon) \geq \alpha_3^3 (1 - \epsilon^{1/4}) > \alpha_3^3 (1 + 2 \mu_{i-1}) - \mu_i \alpha_3^3 / 64,
\]

by the union bound, the probability that there exists \( d \in Q_i \setminus \{0\} \) with \( d_{i-1} = 0 \) and

\[
E_{x \in Q_i} [f_i(x) f_i(x + d) f_i(x + 2d)] \geq \alpha_3^3 (1 - \epsilon)
\]

is at most

\[
m_i \exp(-2^{-1} \cdot 64^{-2} \mu_{i-1} \alpha_6^6) < 1/2,
\]

where we used the upper bound on \( m_i \) in the theorem statement.

\[\square\]

**Proof of Claim 6.2.** Recall that for a \( \mathbb{Z}_{m_i} \)-coset representing by \( w \in Q_i \), \( f_i \) is either a constant function on \( w \) if \( w \notin M_{i-1} \), or otherwise \( f_i(x) = g_{i'}(a_w x_i + b_w) \) where

\[
g_{i'}(x) = \alpha' - \frac{\alpha' \cos(2\pi/m_i)}{2} - \frac{\alpha' \cos(4\pi/m_i)}{2}
\]

as defined in Lemma 5.2 and \( a_w \in \mathbb{Z}_{m_i} \setminus \{0\} \) and \( b_w \in \mathbb{Z}_{m_i} \) are chosen uniformly and independently for each \( w \in M_{i-1} \). For each 3-AP \( (w, w + d', w + 2d') \) with common difference \( d' \in Q_{i-1} \setminus \{0\} \), and for any lift \( d \) of \( d' \), we have

\[
E_{x \in Q_i, x_{i-1} = w} [f_i(x) f_i(x + d) f_i(x + 2d)] = E_{y \in \mathbb{Z}_{m_i}} [g(1)(a_1 y + b_1) g(2)(a_2 y + a_2 d_i + b_2) g(3)(a_3 y + 2 a_3 d_i + b_3)] = E_{y \in \mathbb{Z}_{m_i}} [g(1)(a_1 y + c_1) g(2)(a_2 y + c_2) g(3)(a_3 y + c_3)],
\]

where \( g_{(j)} : \mathbb{Z}_{m_i} \to [0,1] \) can be either the function \( g_{i'} \) or a constant function, \( a_j \in \mathbb{Z}_{m_i} \setminus \{0\}, b_j \in \mathbb{Z}_{m_i} \) are chosen uniformly and independently at random, and \( c_1 = b_1, c_2 = a_2 d_i + b_2, c_3 = 2 a_3 d_i + b_3 \). Note that if we fix the modification (i.e., fixing each \( a_j \) and \( b_j \)), changing \( d \) to a different lift of \( d' \) would only change \( c_j \) in equation (17), and would not change the coefficients of \( y \) in \( g(1), g(2), g(3) \) in the last line of equation (17). Let \( J \subseteq [3] \) be the set of indices such that \( g_{(j)} = g_{i'} \). By Lemma 5.1 applied to the function \( g_{i'} \) and \( h = \lvert J \rvert \leq 3 \), with probability at least \( 1 - 125/(m_i - 1) \),

\[
E_{y \in \mathbb{Z}_{m_i}} \left[ \prod_{j \in J} g_{(j)}(a_j y + u_j) \right] = \prod_{j \in J} E_{y \in \mathbb{Z}_{m_i}} [g_{(j)}(y)]
\]

for all \( u_j \in \mathbb{Z}_{m_i} \). Since \( g_{(j)} \) is a constant function for \( j \notin J \), we obtain that with probability at least \( 1 - 125/(m_i - 1) \),

\[
E_{y \in \mathbb{Z}_{m_i}} [g(1)(a_1 y + c_1) g(2)(a_2 y + c_2) g(3)(a_3 y + c_3)] = E_{y \in \mathbb{Z}_{m_i}} [g(1)(y)] E_{y \in \mathbb{Z}_{m_i}} [g(2)(y)] E_{y \in \mathbb{Z}_{m_i}} [g(3)(y)] = f_{i-1}(w) f_{i-1}(w + d') f_{i-1}(w + 2d').
\]

Thus, by the union bound, with probability at least \( 1 - 125 \mu_{i-1}^2/(m_i - 1) \), for every 3-AP \( (w, w + d', w + 2d') \) in \( Q_{i-1} \) with nonzero common difference \( d' \), and for all lifts \( d \) of \( d' \) in \( Q_i \),

\[
E_{x \in Q_i, x_{i-1} = w} [f_i(x) f_i(x + d) f_i(x + 2d)] = f_{i-1}(w) f_{i-1}(w + d') f_{i-1}(w + 2d').
\]

For \( i \geq 2, m_i \geq n_{i-1}^2, \) and \( m_i \geq m_2 \geq \epsilon^{-2}/64 > 10^6 \), so \( 125 \mu_{i-1}^2/(m_i - 1) < 1/2 \). Hence with probability larger than \( 1/2 \), for all \( d \in Q_i \) such that \( d_{i-1} = d' \neq 0 \),

\[
E_{x \in Q_i} [f_i(x) f_i(x + d) f_i(x + 2d)] = E_{x \in Q_{i-1}} [f_{i-1}(x) f_{i-1}(x + d') f_{i-1}(x + 2d')].
\]

\[\square\]
Thus, assuming that \( \mathcal{P}(i - 1) \) holds, we can choose the modification in level \( i \) so that \( \mathcal{P}(i) \) holds. By induction, we can find a function \( f_s \) which satisfies \( \mathcal{P}(s) \). Notice that \( f_s(x) = \alpha' \) for at least a \( 3/4 \) fraction of \( x \in G \) by (13), and \( \mathbb{E}_x[f_s(x)^3] \leq 3\alpha'^3/2 \) by (15) with \( i = s \). The function \( f = f_s \) then satisfies the conclusion of Theorem 3.4. \( \square \)

7. LOWER BOUND CONSTRUCTION FOR INTERVALS

In this section we prove Theorem 3.3, restated below for convenience.

**Theorem.** There are positive absolute constants \( c, \alpha_0 \) such that the following holds. If \( 0 \leq \alpha \leq \alpha_0 \), \( 0 < \epsilon \leq \alpha^7 \), and \( N \leq \text{tower}(c \log(1/\epsilon)) \), then there is a function \( f : [N] \to [0,1] \) with \( \mathbb{E}[f] = \alpha \) such that for any \( 0 < d < N/2 \),

\[
\mathbb{E}_{x \in [N-d]}[f(x)f(x+d)f(x+2d)] \leq \alpha^3(1-\epsilon).
\]

By Appendix A, in order to prove Theorem 3.3, we can (and will) assume that \( N \geq \epsilon^{-15} \).

Before proving Theorem 3.3, we first need an auxiliary construction of a set with relatively low 3-AP density given its density. Recall from the introduction that \( N(\alpha) \) is the least positive integer such that if \( N \geq N(\alpha) \) and \( A \subset [N] \) with \( |A| \geq \alpha N \), then \( A \) contains a 3-AP.

**Lemma 7.1.** For \( \alpha > 0 \) sufficiently small, there is a subset \( T \subset \mathbb{Z}_n \) with \( |T| \geq \alpha n \) and with 3-AP density at most \( \max \left( \frac{1}{n}, \frac{2\alpha}{N(6\alpha)} \right) \).

**Proof.** Let \( N = N(6\alpha) - 1 \), so there is \( A \subset [N] \) with \( |A| = [6\alpha N] \) which has nontrivial 3-AP.

First assume \( n \leq 4N \). Partition \([N]\) into at most \( 2N/n + 1 \leq 6N/n \) intervals of length at most \( \lceil n/2 \rceil \). The set \( A \) contains at least \( |A|/(6N/n) \geq \alpha n \) elements in one of these intervals. Viewed as a subset of \( \mathbb{Z}_n \), we have a subset of \( \mathbb{Z}_n \) with density at least \( \alpha \) and with no nontrivial 3-AP, and hence 3-AP density at most \( 1/n \).

So we may assume \( n > 4N \). Integers \( x, y, z \) form an approximate 3-AP if \( |2z - x - y| \leq 1 \). Let \( S := \{2a : a \in A\} \), so \( S \) has no approximate 3-AP. Let \( t = \lfloor \frac{n}{6N} \rfloor \). Consider the set \( I_i := \{(i-1)t + 1, (i-1)t + 2, \ldots, (i-1)t + t\} \) of \( t \) consecutive integers. Let \( T \) be the union of the sets \( I_i \) with \( i \in S \). The set \( T \) has size \( |T| = |A|t \geq \alpha n \). Also, every element of \( T \) is a positive integer at most \((2N-1)t + t \leq n/2 \). So if \( x, y, z \in T \) are such that \( (x, y, z) \ (\text{mod } n) \) form a 3-AP in \( \mathbb{Z}_n \), then \( (x, y, z) \) is also a 3-AP of integers. Since \( S \) has no approximate 3-AP, it follows that the only 3-APs in \( T \) are those where the three terms are in the same interval \( I_i \). In each interval \( I_i \), which has size \( t \), the number of 3-APs (with any integer difference allowed) is \( t + 2 \lfloor \frac{t-1}{3} \rfloor \). There are \( |A| \) intervals \( I_i \) whose union is \( T \). The number of 3-APs in \( \mathbb{Z}_n \) is \( n^3 \). Hence, the 3-AP density of \( T \) as a subset of \( \mathbb{Z}_n \) is \((t + 2 \lfloor \frac{t-1}{3} \rfloor) \alpha^3/n^2 \leq \frac{2\alpha}{N(6\alpha)} \).

The Behrend construction [2] implies that on \( N(\alpha) \) implies that if \( \alpha > 0 \) is sufficiently small, then \( N(6\alpha) \geq 2^{\frac{1}{2}}(\log_2 1/\alpha)^2 \). Together with the previous lemma, we have the following immediate corollary.

**Lemma 7.2.** If \( \alpha > 0 \) is sufficiently small, then for any positive integer \( n \), there is a subset of \( \mathbb{Z}_n \) with density at least \( \alpha \) and 3-AP density at most \( \max \left( \frac{1}{n}, \frac{2\alpha}{N(6\alpha)} \right) \).

7.1. The construction and proof of Theorem 3.3. We next construct a function \( f : [N] \to [0,1] \) with \( \mathbb{E}[f(x)] = \alpha \) such that for any \( 0 < d < N/2 \), \( \mathbb{E}_{x \in [N-d]}[f(x)f(x+d)f(x+2d)] \leq \alpha^3(1-\epsilon) \).

The construction is done in three steps.

**Step 1:** Choose \( \beta \) so that \( \beta \leq \epsilon^2 \) and \( N' = N(1 - \beta) \) has a divisor \( q \) such that \( N^{1/5} < q < \sqrt{3\alpha^3}(1-\epsilon)^N \) \( N'/q \) is prime, and \( q \) satisfies the condition in Theorem 3.4 with parameters \( \alpha_{3.4} = \alpha \) and \( \epsilon_{3.4} = 4\epsilon \). Here, for clarity, we include the theorem index in the subscript of the parameters in the theorem. Note that if \( q \) satisfies the condition in Theorem 3.4 with parameters \( \alpha_{3.4} = \alpha \) and
$\epsilon_{3.4} = 4\epsilon$, then for any $\alpha_1 \in [\alpha, 1/4]$, $g$ also satisfies the condition in Theorem 3.4 with parameters $\alpha_{3.4} = \alpha_1$ and $\epsilon_{3.4} = 4\epsilon$. The existence of $\beta, N'$, $q$ is guaranteed by Lemma 7.4, which is deferred to the end of the section. Let

$$\alpha' := (1 - \beta)^{-1}\alpha.$$ 

**Step 2:** Since $\alpha' \in [\alpha, 1/4]$, we can apply Theorem 3.4 with $G = \mathbb{Z}_q$, $\alpha'_{4.4} = \alpha'$, $\epsilon_{4.4} = 4\epsilon$. We obtain $g : \mathbb{Z}_q \to [0, 1]$ with density $\alpha'$, mean cube density at most $\frac{2}{3}\alpha^3$, such that for each $d \in \mathbb{Z}_q \setminus \{0\}$,

$$E_x[g(x)g(x + d)g(x + 2d)] \leq \alpha^3(1 - 4\epsilon) \leq \alpha^3(1 - 3\epsilon),$$

and there exists $\alpha^* \in [\alpha', \alpha'(1 + e^{1/4})]$ such that $|\{x \in \mathbb{Z}_q : g(x) = \alpha^*\}| \geq 3q/4$. For an integer $x$, denote $\bar{x}_q = x \mod q \in \mathbb{Z}_q$. Define $f_2 : [N] \to [0, 1]$ by $f_2(x) = g(\bar{x}_q)$ for $x \in [N]$ and $f_2(x) = 0$ for $x > N'$.

**Step 3:** Let $n = \frac{N'}{q}$, which is prime. Apply Lemma 7.2 to find $X \subset \mathbb{Z}_n$ with density at least $\alpha^*$ and 3-AP density at most $\min \left\{ \frac{1}{n}, 2^{-\log_2((1/\alpha^*)^2/9)} \right\}$. We use $\xi$ to denote the characteristic function of $X$ scaled by $\alpha^*n/|X|$, so $\xi(x) = \alpha^*nX(x)/|X|$. Then $E_{x \in \mathbb{Z}_n} [\xi(x)] = \alpha^*$ and since $|X|/n \geq \alpha^*$,

$$E_{x,d \in \mathbb{Z}_n} [\xi(x)\xi(x + d)\xi(x + 2d)] \leq E_{x,d \in \mathbb{Z}_n} [X(x)X(x + d)X(x + 2d)] \leq \min \left\{ \frac{1}{n}, 2^{-\log_2((1/\alpha^*)^2/9)} \right\}.$$

For each $t \in \mathbb{Z}_q$, we define $P_t = \{x \in [N'] : \bar{x}_q = t\}$, which forms an arithmetic progression of length $n$. Let $x_1 < x_2 < \ldots < x_n$ be the elements of $P_t$ in increasing order. We define a bijection $\phi_t : P_t \to \mathbb{Z}_n$ such that $\phi_t(x_i) \equiv t \mod n$. Observe that if $(x, y, z)$ is a 3-AP in $P_t$, then $(\phi(x), \phi(y), \phi(z))$ is a 3-AP in $\mathbb{Z}_n$.

For each $t \in \mathbb{Z}_q$, we choose independently and uniformly at random $a_t \in \mathbb{Z}_n \setminus \{0\}, b_t \in \mathbb{Z}_n$, with independent choices for different $t$. Define $f_3$ by $f_3(x) = \xi(a_t\phi_t(x) + b_t)$ for $x$ such that $\bar{x}_q = t$ and $g(t) = \alpha^*$, and $f_3(x) = f_2(x)$ otherwise.

We let $f = f_3$. It is easy to see that $E_x[f_3(x)] = \alpha$. We now prove that there exists a choice of randomness (in Step 3) such that for each positive integer $d < N/2$,

$$E_{x \in [N - 2d]} [f(x)f(x + d)f(x + 2d)] \leq \alpha^3(1 - \epsilon).$$

**Proof of Theorem 3.3.** We reuse the notations from the description of the construction. For a 3-AP $(x, x + d, x + 2d)$ in $[N]$, we refer to $f_3(x)f_3(x + d)f_3(x + 2d)$ as its weight. For any common difference $d \geq \frac{N'}{2}$, all the 3-APs $(x, x + d, x + 2d)$ in $[N]$ have zero weight since $x + 2d > N'$. Hence the density of 3-APs with common difference $d$ of $f_3$ is 0. For any common difference $d \geq \frac{N' - \beta N\alpha^3(1 - \epsilon)}{2}$, let $t = N' - 2d$. Then $t \leq \beta N\alpha^3(1 - \epsilon)$. The number of 3-APs with common difference $d$ in $[N']$ is at most $t$, and hence the number of 3-APs in $[N]$ with nonzero weight is at most $t$. The number of 3-APs with common difference $d$ in $[N]$ is $N - 2d = \beta N + t$, so the density of 3-APs with common difference $d$ in $[N]$ is at most $\frac{t}{\beta N + t} < \alpha^3(1 - \epsilon)$ since $t \leq \frac{\beta N\alpha^3(1 - \epsilon)}{1 - \alpha^3(1 - \epsilon)}$. For $d$ such that $0 < d < \frac{N' - \beta N\alpha^3(1 - \epsilon)}{2}$, the number of 3-APs of common difference $d$ in $[N']$ is at least $\beta N\alpha^3(1 - \epsilon)$. Partition the 3-APs with common difference $d$ in $[N']$ into different classes according to the congruence class modulo $q$ of the 3-AP (so the class a 3-AP belongs to is determined by the congruence class modulo $q$ of the first element of the 3-AP). Since $\beta N\alpha^3(1 - \epsilon) > q^2$, all classes of 3-APs modulo $q$ with common difference $\bar{d}_q$ appear, each class with at least $q$ elements, and any two classes differ in size by at most 1. Hence,

$$|E_{x \in [N' - 2d]} [f_2(x)f_2(x + d)f_2(x + 2d)] - E_{y \in \mathbb{Z}_q} [g(y)g(y + \bar{d}_q)g(y + 2\bar{d}_q)]| \leq \frac{1}{q}.$$

By the construction, if $\bar{d}_q \neq 0$ then

$$E_{y \in \mathbb{Z}_q} [g(y)g(y + \bar{d}_q)g(y + 2\bar{d}_q)] \leq \alpha^3(1 - 3\epsilon),$$

and if $\bar{d}_q = 0$ then

$$E_{y \in \mathbb{Z}_q} [g(y)g(y + \bar{d}_q)g(y + 2\bar{d}_q)] \leq \frac{3}{2}\alpha^3.$$
Thus, for $d$ nonzero modulo $q$,
\[ E_{x \in \mathbb{N} - 2d} [f_2(x)f_2(x + d)f_2(x + 2d)] \leq \alpha^3 (1 - 3\epsilon) + \frac{1}{q} < \alpha^3 (1 - 2\epsilon), \]
and for $d$ divisible by $q$,
\[ E_{x \in \mathbb{N} - 2d} [f_2(x)f_2(x + d)f_2(x + 2d)] \leq \frac{3}{2} \alpha^3 + \frac{1}{q}. \]

(18)

In the third step, suppose $d$ is nonzero and divisible by $q$ and let $t \in \mathbb{Z}_q$ with $g(t) = \alpha^*$. For $x \in \mathbb{N} - 2d$ with $x_q = t$, one has $(f_3(x), f_3(x + d), f_3(x + 2d)) = (\xi(a_t\phi_t(x) + b_t), \xi(a_t\phi_t(x + d) + b_t), \xi(a_t\phi_t(x + 2d) + b_t))$. Recall that $a_t$ is uniformly distributed over $\mathbb{Z}_n \setminus \{0\}$ and $b_t$ is uniformly distributed over $\mathbb{Z}_n$, so $(a_t\phi_t(x) + b_t, a_t\phi_t(x + d) + b_t, a_t\phi_t(x + 2d) + b_t)$ is uniformly distributed over the 3-APs in $\mathbb{Z}_n$ with nonzero common difference. Thus,
\[ \mathbb{E}_{f_3} E_{x_q = t, x \in \mathbb{N} - 2d} [f_3(x)f_3(x + d)f_3(x + 2d)] = \lambda(\xi) \leq \Lambda(\xi) \leq \max \left( \frac{1}{n}, 2^{-\log_2(1/\alpha^*)^2/9} \right) \leq \frac{\alpha^*}{10}, \]
where $\lambda(\xi)$ is the density of 3-APs with nonzero common difference of $\xi$, $\Lambda(\xi)$ is the density of 3-APs of $\xi$, and we have used that $n \geq \sqrt{n}/2 \geq \epsilon^{-15/2}/2 \geq \alpha^{-10}$ and $\alpha \leq \alpha_0$ is sufficiently small. Thus, for each $t \in \mathbb{Z}_q$ such that $g(t) = \alpha^*$,
\[ \mathbb{E}_{f_3} E_{x_q = t, x \in \mathbb{N} - 2d} [f_3(x)f_3(x + d)f_3(x + 2d)] \leq \mathbb{E}_{x_q = t, x \in \mathbb{N} - 2d} [f_2(x)f_2(x + d)f_2(x + 2d)] - \frac{9\alpha^3}{10}. \]

(19)

For each $t \in \mathbb{Z}_q$ such that $g(t) \neq \alpha^*$,
\[ \mathbb{E}_{f_3} E_{x_q = t, x \in \mathbb{N} - 2d} [f_3(x)f_3(x + d)f_3(x + 2d)] = \mathbb{E}_{x_q = t, x \in \mathbb{N} - 2d} [f_2(x)f_2(x + d)f_2(x + 2d)]. \]

(20)

Hence,
\[ \mathbb{E}_{f_3} E_{x \in \mathbb{N} - 2d} [f_3(x)f_3(x + d)f_3(x + 2d)] \leq \frac{3\alpha^3}{2} + \frac{1}{q} - \frac{3}{4} \cdot \frac{9\alpha^3}{10} \leq \frac{5\alpha^3}{6}, \]
where in the first inequality we used (18), (19) and (20) together with the fact that $g(t) = \alpha^*$ for at least a $3/4$ fraction of $t \in \mathbb{Z}_q$, and in the second inequality we used that $q > N^{1/5} \geq \epsilon^{-3} \geq \alpha^{-21} > 120/\alpha^{*3}$. Notice that for fixed nonzero $d$ divisible by $q$, the random variables $E_{x_q = t, x \in \mathbb{N} - 2d} [f_3(x)f_3(x + d)f_3(x + 2d)]$, for $t \in \mathbb{Z}_q$, are independent. By Hoeffding’s inequality, the probability that
\[ E_{x \in \mathbb{N} - 2d} [f_3(x)f_3(x + d)f_3(x + 2d)] \geq \alpha^3 - \frac{\alpha^*}{12} \]
is at most $\exp \left( -2 \cdot (12^{-1} \alpha^*^3)^2 q \right) = \exp \left( -72^{-1} \alpha^*^6 q \right)$. Noting that $\alpha^*^3 - \alpha^*^3/12 \leq \alpha^3 (1 - \epsilon)$, by the union bound, the probability that there exists a nonzero common difference $d$ which is divisible by $q$ such that
\[ E_{x \in \mathbb{N} - 2d} [f_3(x)f_3(x + d)f_3(x + 2d)] \geq \alpha^3 (1 - \epsilon) \]
is at most $(N/q) \exp \left( -72^{-1} \alpha^*^6 q \right) < q^4 \exp \left( -72^{-1} \alpha^*^6 q \right) < 1/2$, as $N < q^5$, $q > \epsilon^{-3}$ and $\epsilon \leq \alpha^7$.

If $d$ is not divisible by $q$, each 3-AP with common difference $d$ occupies three different modulo $q$ classes, and hence the weights of the elements in the 3-AP are independent random variables. By construction, for each $x \in \mathbb{N}$, $E_{f_3} [f_3(x)] = f_2(x)$. Hence, by independence, if $(x, x + d, x + 2d) \in [\mathbb{N}]^3$,
\[ E_{f_3} [f_3(x)f_3(x + d)f_3(x + 2d)] = f_2(x)f_2(x + d)f_2(x + 2d) \]
Thus,
\[ E_{f_3} E_{x \in \mathbb{N} - 2d} [f_3(x)f_3(x + d)f_3(x + 2d)] = E_{x \in \mathbb{N} - 2d} [f_2(x)f_2(x + d)f_2(x + 2d)] \leq \alpha^3 (1 - 2\epsilon). \]
For this fixed $d$, we can partition $\mathbb{Z}_q$ into five sets $S_1, S_2, \ldots, S_5$ such that for each $i, 1 \leq i \leq 5$, the 3-APs $(t, t + d, t + 2d), t \in S_i$ are disjoint, and $|S_i| \geq q/10$. For each set $S_i$, the random variables
\[ E_{x_q=t,x \in [N'-2d]}[f_3(x)f_3(x+d)f_3(x+2d)], \text{ for } t \in S_i, \text{ are independent. By Hoeffding's inequality, the probability that} \]
\[ E_{t \in S_i}E_{x_q=t,x \in [N'-2d]}[f_3(x)f_3(x+d)f_3(x+2d)] \geq \]
\[ E_{f_3}E_{t \in S_i}E_{x_q=t,x \in [N'-2d]}[f_3(x)f_3(x+d)f_3(x+2d)] + \epsilon \alpha^3 \]
\[ \text{is at most } \exp(-2(\epsilon \alpha^3)^q/10) = \exp(-5^{1-1}\cdot \epsilon^2 \alpha^6 q). \text{ By the union bound, the probability that there exists a common difference } d \text{ not divisible by } q \text{ with} \]
\[ E_{x \in [N'-2d]}[f_3(x)f_3(x+2d)f_3(x+2d)] \geq \alpha^3(1-2\epsilon) + \epsilon \alpha^3 = \alpha^3(1-\epsilon) \]
\[ \text{is at most } 5N \exp(-5^{1-1}\cdot \epsilon^2 \alpha^6 q) < 1/2, \text{ where we used } N < q^5, q > \epsilon^{-3} \text{ and } \epsilon \leq \alpha^7. \]
\[ \text{Since } f_3(x) = 0 \text{ for all } x \notin [N'], \]
\[ E_{x \in [N'-2d]}[f_3(x)f_3(x+2d)f_3(x+2d)] \leq E_{x \in [N'-2d]}[f_3(x)f_3(x+2d)f_3(x+2d)] \]
\[ \text{Hence, with positive probability, the function } f_3 \text{ satisfies the required properties in Theorem 3.3.} \]

To finish the proof, we prove that the parameters in Step 1 of the construction described at the beginning of the subsection can be chosen. We first prove that we can approximate any large integer with one satisfying the conditions in Theorem 3.4.

**Lemma 7.3.** There exist constants \( c, \alpha_0 > 0 \) such that if \( 0 \leq \alpha \leq \alpha_0, \ 0 < \epsilon \leq \alpha^7, \) and \( r \) is an integer satisfying \( \epsilon^{-15} \leq r \leq \text{tower}(c \log(1/\epsilon)), \) then we can choose \( s \in \left[2, \left\lceil \log_{150} \left( \frac{\epsilon^{-1/4} \alpha^6}{8} \right) \right\rceil \right] \) and primes \( m_1, \ldots, m_s \) and \( q \) satisfying the following properties:

- \( \epsilon^{-1/3}/2 \leq m_1 \leq \epsilon^{-1/3}, \) and,
- for \( i \geq 2, \) \( n_i^{1 \epsilon_i-1} < m_i < \exp(2^{-1} \cdot 64^{-2} \cdot 150^{i-1 \epsilon/4} n_{i-1}) \) where \( n_i = \prod_{j=1}^i m_j, \) and,
- for \( n = m_s, \) we have \( n \in \lceil r(1 - \epsilon^2), r \rceil. \)

**Proof.** Choose a sequence of \( s \) real numbers \( \hat{m}_1, \ldots, \hat{m}_s \) satisfying the following properties: \( \hat{m}_1 \) is a prime number with \( \epsilon^{-1/3}/2 \leq \hat{m}_1 \leq \epsilon^{-1/3}, \) \( \hat{m}_i^{1 \epsilon_i-1} < \hat{m}_i < \exp(2^{-1} \cdot 64^{-2} \cdot 150^{i-1 \epsilon/4} n_{i-1}) \) for \( i \geq 2 \) with \( \hat{m}_i = \prod_{j=1}^i \hat{m}_j, \) and \( \prod_{i=1}^s \hat{m}_i = r. \) The existence of \( \hat{m}_1 \) is guaranteed by Bertrand’s postulate. Since \( \epsilon^{-15} \leq r \leq \text{tower}(c \log(1/\epsilon)) \) and \( \epsilon \leq \alpha^7, \) if we choose \( c, \alpha_0 \) sufficiently small, it is easy to see that there exists a choice of \( s \) and \( m_1, \ldots, m_s \) satisfying these properties.

For each \( i \) let \( m_i \) be the largest prime such that \( m_i \leq \hat{m}_i. \) So \( m_1 = \hat{m}_1, \) and from [1], for all \( \hat{m}_i \) large enough, \( m_i \geq \hat{m}_i - \hat{m}_i^{0.525}. \) Hence
\[ n = \prod_{i=1}^s m_i \geq \hat{m}_1 \prod_{i=2}^s (\hat{m}_i - \hat{m}_i^{0.525}) \]
\[ \geq r \prod_{i=2}^s (1 - \hat{m}_i^{-0.475}) \geq r \exp(-\sum_{i=2}^s \frac{2}{\hat{m}_i^{0.475}}) \]
\[ \geq r \exp(-\frac{4}{\hat{m}_2^{0.475}}) \geq r \exp(-\epsilon^{2.5}) \]
\[ \geq r(1 - \epsilon^{2.5}), \]
where we have applied the inequality \( \exp(-x) \geq 1 - x \geq \exp(-2x) \) for \( 0 \leq x \leq 1/2, \) \( \sum_{i=2}^s \frac{1}{m_i^{0.475}} \leq \frac{1}{m_1^{0.475}}, \) \( \sum_{i=0}^1 \frac{1}{m_i^{0.475}} \leq \frac{2}{m_2^{0.475}} \) as \( \hat{m}_i+1 \geq \hat{m}_i^{1/0.475} \hat{m}_i \geq \ldots \geq 2^{i(1-1/0.475)} \hat{m}_2 \) and \( \hat{m}_i \geq \hat{m}_2 \geq \epsilon^{-6} \) for \( i \geq 2. \) Thus we can choose \( n = \prod_{i=1}^s m_i \) and \( r(1 - \epsilon^2) \leq n \leq r. \)

Using Lemma 7.3, we prove that the parameters \( N', q \) in Step 1 of the construction can be chosen.

**Lemma 7.4.** Let \( N \geq \epsilon^{-15} \). There exists \( q, N' \) such that \( N'/q \) is prime, \( q \) satisfies the conditions in Theorem 3.4, and \( (1 - \epsilon^2)N \leq N' \leq N. \)
Proof. We choose $p$ to be a prime number in $(N^{1/5}, \sqrt{\epsilon^4 \alpha^3(1-\epsilon)N})$. Let $r = \left\lfloor N(1-\epsilon^2/4)/p \right\rfloor \geq N(1-\epsilon^2/2)/p \geq \epsilon^{-7}$. Apply Lemma 7.3 with the above choice of $r$ and with $\epsilon_{7.3} = 4\epsilon$, we find $n$ satisfying the conditions in Theorem 3.4, applied with $\alpha_{3.4} \in [\alpha, \alpha(1+\epsilon)]$ and $\epsilon_{3.4} \leq 4\epsilon$, such that $(1-\epsilon^2.5)r \leq n \leq r$. Then we let $N' = pn$ and $q = n$.

We have

$$N' = pn \geq (1-\epsilon^2.5)rp \geq (1-\epsilon^2/2)(1-\epsilon^2.5)N \geq (1-\epsilon^2)N,$$

thus $(1-\epsilon^2)N \leq N' \leq N$, finishing the proof. \qed

References

[1] R. C. Baker, G. Harman, J. Pintz, The difference between consecutive primes, II., Proc. London Math. Soc. (3) 83 (2001), 532–562.
[2] F. Behrend, On sets of integers which contain no three terms in arithmetic progression, Proc. Nat. Acad. Sci. 32 (1946), 331–332.
[3] V. Bergelson, B. Host, and B. Kra, Multiple recurrence and nilsequences, with an appendix by Imre Ruzsa, Invent. Math. 160 (2005), 261–303.
[4] A. Berger, Popular differences for corners in abelian groups, arXiv:1909.12350.
[5] T. Bloom, A quantitative improvement for Roth’s theorem on arithmetic progressions, J. Lond. Math. Soc. 93 (2016), 643–663.
[6] T. Bloom and O. Sisask, Logarithmic bounds for Roth’s theorem via almost-periodicity, Discrete Anal. 4 (2019), 20 pp.
[7] J. Bourgain, On triples in arithmetic progression, Geom. Funct. Anal. 9 (1999), 968–984.
[8] J. Bourgain, Roth’s theorem on progressions revisited, J. Anal. Math. 104 (2008), 155–192.
[9] D. Conlon and J. Fox, Bounds for graph regularity and removal lemmas, Combin. Probab. Comput. 26 (2017), 505–527.
[10] M. Elkin, An improved construction of progression-free sets, Int. Math. Res. Not. IMRN (2017), 465–588.
[11] M. Mandache, A variant of the corners theorem, arXiv:1804.03972.
In this appendix, we show how to deduce Theorem 1.5 from its functional version Theorem 3.3. We first show that Theorem 1.5 holds if it holds in the case $N \geq \epsilon^{-15}$. Assume that $N < \epsilon^{-15}$. Recall that $N(\alpha)$ denotes the least positive integer such that if $N \geq N(\alpha)$ then any $A \subset [N]$ with $|A| \geq \alpha N$ contains a nontrivial 3-AP. If $N < N(\alpha)$, there exists a subset $A$ of $[N]$ with $|A| \geq \alpha N$ and $A$ does not contain a nontrivial 3-AP. In this case, for all $d \neq 0$, 

$$E_{x \in [N-2d]} [A(x)A(x+d)A(x+2d)] = 0,$$

so the conclusion of Theorem 1.5 holds. By Behrend’s bound ([2], [10]), we have that $N(\alpha) \geq \exp((\log 1/\alpha)^2/6)$ for $\alpha > 0$ sufficiently small. For $N \geq N(\alpha) \geq \exp((\log 1/\alpha)^2/6)$, let $\epsilon_0 = N^{-1/15}$ so $N = \epsilon_0^{-15}$. We have $\epsilon_0 > \epsilon$ and $\epsilon_0 \leq \exp(-(\log(1/\epsilon)^2)/90) \leq \alpha^{12}$ (assuming $\alpha_0$ is small enough). By choosing $\alpha_0$ in Theorem 1.5 small enough, we may assume that for all $x \leq \alpha_0$, $x^{-15} < \text{tower}(c(\log(1/\epsilon_0)))$. Then $N = \epsilon_0^{-15} < \text{tower}(c(\log(1/\epsilon_0)))$ as $\epsilon_0 < \alpha \leq \alpha_0$. Thus, $\epsilon_0^{-15} \leq N \leq \text{tower}(c(\log(1/\epsilon_0)))$, so we can apply Theorem 1.5 with $\epsilon_0$ in place of $\epsilon$ to obtain the desired set $A$. The same argument also shows that we only need to prove Theorem 3.3 when $N \geq \epsilon^{-15}$.

We next discuss how to obtain a set $A$ with the properties in Theorem 1.5 from Theorem 3.3 when $N \geq \epsilon^{-15}$. This follows via a standard sampling argument which is essentially similar to Lemma 9 in [12]. However, there are some small differences to the argument which we now highlight. Given a function $f : [N] \to [0,1]$ such that the density of 3-APs with common difference $d$ of $f$ is small for all $0 < d < N/2$, we sample a set $A$ where each element $x \in [N]$ is in $A$ with probability $f(x)$ independent of each other. If the density of 3-APs with common difference $d$ in $A$ is concentrated around its expectation, which is the density of 3-APs with common difference $d$ of $f$, then it is small with high probability. However, for $d$ near $N/2$, there are very few 3-APs with common difference $d$, and we do not have sufficiently strong concentration to be able to take a union bound over all such $d$. To get around this, we define a function $f'$ which is 0 for all $x$ close to $N$, and which is equal to $f$ elsewhere, and sample the set $A$ from $f'$. This ensures that for common differences $d$ which are close to $N/2$, set $A$ contains very few 3-APs with common difference $d$.

We now carry out the details. By Theorem 3.3 applied with $\alpha$ replaced by $\alpha + 2\epsilon$ and $\epsilon$ replaced by $12\epsilon/\alpha^3 \leq \alpha^3$, we can find a function $f : [N] \to [0,1]$ such that for any $0 < d < N/2$, 

$$E_{x \in [N-2d]} [f(x)f(x+d)f(x+2d)] \leq (\alpha + 2\epsilon)^3(1 - 12\epsilon/\alpha^3).$$

Define $f' : [N] \to [0,1]$ by $f'(x) = 0$ if $x \geq N(1-\epsilon)$ and $f'(x) = f(x)$ otherwise. We define $A$ to be a random subset of $[N]$ where each $x \in [N]$ is in $A$ with probability $f'(x)$, independently of the other elements. In particular, $A$ does not contain $x$ if $x \geq N(1-\epsilon)$. Hence, if the common difference $d$ is larger than $N(1-\epsilon)/2$ then 

$$E_{x \in [N-2d]} [A(x)A(x+d)A(x+2d)] = 0.$$

If $d$ is at most $N(1-\epsilon)/2$, then $N-2d \geq \epsilon N$, so by Hoeffding’s inequality, with probability at least $1 - \exp(-\epsilon^2(N-2d)) \geq 1 - \exp(-\epsilon^3 N)$, 

$$|E_{x \in [N-2d]} [A(x)A(x+d)A(x+2d)] - E_{x \in [N-2d]} [f''(x)f''(x+d)f''(x+2d)]| \leq \epsilon.$$

Furthermore, note that 

$$E_{x \in [N-2d]} [f''(x)f''(x+d)f''(x+2d)] \leq E_{x \in [N-2d]} [f(x)f(x+d)f(x+2d)] \leq (\alpha + 2\epsilon)^3(1 - 12\epsilon/\alpha^3) \leq \alpha^3 - 2\epsilon.$$
By Hoeffding’s inequality, with probability at least $1 - \exp(-\epsilon^2 N)$, the density of $A$ is at least

$$E_x[f'(x)] - \epsilon \geq E_x[f(x)] - \frac{\epsilon N}{N} - \epsilon = \alpha.$$ 

Thus, by the union bound, with probability at least $1 - N \exp(-\epsilon^3 N) - \exp(-\epsilon^2 N)$, $A$ is a set with density at least $\alpha$ such that for all $0 < d < N/2$,

$$E_{x \in [N-2d]} [A(x)A(x+d)A(x+2d)] \leq \alpha^3 - \epsilon.$$ 

Since $N > \epsilon^{-15}$, for $\epsilon$ sufficiently small, $1 - N \exp(-\epsilon^3 N) - \exp(-\epsilon^2 N) > 0$. This gives Theorem 1.5, assuming Theorem 3.3.

Fox, Department of Mathematics, Stanford University, Stanford, CA, USA
E-mail address: jacobfox@stanford.edu

Pham, Department of Mathematics, Stanford University, Stanford, CA, USA.
E-mail address: huypham@stanford.edu

Zhao, Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, USA
E-mail address: yufeiz@mit.edu