IMPROVED BOUNDS FOR SZEMERÉDI'S THEOREM

JAMES LENG, ASHWIN SAH, AND MEHTAAB SAWHNEY

ABSTRACT. Let $r_k(N)$ denote the size of the largest subset of $[N] = \{1, \ldots, N\}$ with no k-term arithmetic progression. We show that for $k \geq 5$, there exists $c_k > 0$ such that

$$r_k(N) \ll N \exp(-(\log \log N)^{c_k}).$$

Our proof is a consequence of recent quasipolynomial bounds on the inverse theorem for the Gowers U^k-norm as well as the density increment strategy of Heath–Brown and Szemerédi as reformulated by Green and Tao.

1. INTRODUCTION

Let $[N] = \{1, \ldots, N\}$ and $r_k(N)$ denote the size of the largest $S \subseteq [N]$ such that S has no k-term arithmetic progressions. The first nontrivial upper bound on $r_3(N)$ came from work of Roth [20] which proved

$$r_3(N) \ll N(\log \log N)^{-1}.$$

A long series of works improved this bound, including works of Heath-Brown [14], Szemerédi [26], Bourgain [4,5], Sanders [21,22], Bloom [1], and Bloom and Sisask [2]. In breakthrough work, Kelley and Meka [15] very recently proved

$$r_3(N) \ll N \exp(-c(\log N)^{1/12});$$

the constant $1/12$ was refined to $1/9$ in work of Bloom and Sisask [3].

For higher k, a long-standing conjecture of Erdős and Turán stated that $r_k(N) = o(N)$. In seminal works, Szemerédi [24, 25] first established the estimate $r_4(N) = o(N)$ and then established his eponymous theorem that

$$r_k(N) = o(N).$$

Due to uses of van der Waerden theorem and the regularity lemma (which was introduced in this work), Szemerédi's density saving was exceedingly small. In breakthrough work, Gowers [6, 7] introduced higher order Fourier analysis and proved the first "reasonable" bounds for Szemerédi's theorem:

$$r_k(N) < N(\log \log N)^{-2^{-2^{k+9}}}.$$

The only improvement to this result for $k \geq 4$ was work of Green and Tao [9,13] which ultimately established that

$$r_4(N) \ll N(\log N)^{-c},$$

and recent work of the authors [17] which proved

$$r_5(N) \ll N \exp(-(\log \log N)^c).$$

Our main result is an extension of this bound for all $k \geq 5$.

Theorem 1.1. *Fix $k \geq 5$. There is $c_k \in (0, 1)$ such that*

$$r_k(N) \ll N \exp(-(\log \log N)^{c_k}).$$

1.1. **Proof outline and techniques.**

1.1.1. *Local and global inverse theorems.* The primary input to our result will be the main result of recent work of the authors [18], i.e., quasipolynomial bounds on the inverse theorem for the Gowers U^{k+1}-norm. Given an inverse theorem, the deduction of Szemerédi's theorem via a standard density increment strategy is essentially folklore and was recorded in work of Green and Tao [11] (although, prior to [18] the resulting bounds would be far from matching those of Gowers [7]). However, if one naively follows this script using [18], one obtains a bound of $N \exp(-(\log\log\log N)^{-\Omega_k(1)})$ which is weaker than the work of Gowers [7]. Furthermore, Gowers's argument makes use of a "local" inverse theorem that in fact gives a slightly stronger correlation compared to the bound given for the "global" inverse theorem in [18] (namely, polynomial versus quasipolynomial). Thus, this global nature of [18] must be exploited. Additionally, use of global inverse theorems necessitates understanding of nilsequences and polynomial sequences on nilpotent Lie groups, as opposed to merely polynomials as in the work of Gowers [7].

1.1.2. *Schmidt-type decomposition problems.* This is done via the improved density increment strategy of Heath–Brown [14] and Szemerédi [26] which involves extracting a set of functions to correlate with instead of simply one and using this to give a multiplicative density increment. Such a strategy was given a robust formulation in work of Green and Tao [9] on four-term progressions; in particular, their reformulation avoided the explicit Fourier-analytic formulas used in [14, 26] and thus is applicable to the higher order setting. The strategy here runs smoothly given the inverse theorem, modulo resolving a certain Schmidt-type problem for nilsequences. In particular, given a polynomial sequence $g(n)$ with $g(0) = \mathrm{id}_G$ on a nilmanifold G/Γ of degree k with complexity M and dimension d, one needs to prove that

$$\min_{1 \le n \le N} d_{G/\Gamma}(\mathrm{id}_G, g(n)\Gamma) \ll M^{O_k(d^{O_k(1)})} N^{-1/d^{O_k(1)}}.$$

In particular, the polynomial dependence on dimension within the exponent is key.

We in fact require a certain slightly stronger result (decomposing $[N]$ into long arithmetic progressions P such that the diameters of the sets $\{g(n)\Gamma \colon n \in P\}$ are small), which is the heart of the matter for this work. When the underlying nilpotent group G is abelian, this is easily deduced from a result of Schmidt [23] (see Lemma 2.3, or [9, Section 6] in the quadratic case).

For general degree 2 nilmanifolds such a problem was implicitly solved in work of Green and Tao [9] and for degree 3 nilmanifolds it was essentially solved in recent work of the authors [17]. More precisely, [17] essentially proves that given a list of bracket expressions $(a_i n \lfloor b_i n \rfloor \lfloor c_i n \rfloor)_{1 \le i \le d}$ that

$$\min_{1 \le n \le N} \|a_i n \lfloor b_i n \rfloor \lfloor c_i n \rfloor\|_{\mathbb{R}/\mathbb{Z}} \le N^{-1/d^{O(1)}}$$

and via an explicit computation with fundamental domains on degree 3 nilmanifolds one may reduce to such a situation. The proof given in [17] relies on the fact that 3 is sufficiently small and in particular that it is possible to reduce to a situation in which there are no "nested integer part operations" as one attempts to solve the "bracket Schmidt" problem in one go.

1.1.3. *Iterative Schmidt refinement.* The key observation required for our work, at least at a heuristic level, is a procedure for solving such "bracket Schmidt" problems even when there are nested brackets. As a simple example, consider bracket expressions $(a_i n \lfloor b_i n \lfloor c_i n \rfloor \rfloor)_{1 \le i \le d}$. We will solve the Schmidt problem via iteratively "reducing" the number of brackets from the inside-out (at the cost of passing to subprogressions). In particular, using Dirichlet's theorem, one can break $[N]$ into arithmetic progressions P each of length $N^{1/d^{O(1)}}$ such that when restricted to each arithmetic progression, every function $\lfloor c_i n \rfloor$ is a linear function (i.e., it is a *locally linear*

function on each P). Since the only locally linear functions on a progression agree with genuinely linear functions, we can replace $\lfloor c_i n \rfloor$ by $d_{i,P} n + e_{i,P}$ and reduce to considering the bracket expression $a_i n \lfloor b_i n (d_{i,P} n + e_{i,P}) \rfloor$ when restricted to P. One can then iterate this argument on the "inner quadratics" $b_i n (d_{i,P} n + e_{i,P})$ (essentially using abelian Schmidt as discussed above for degree 2 in this case). We may find a decomposition into long arithmetic progressions Q such that $\lfloor b_i n (d_{i,P} n + e_{i,P}) \rfloor$ is locally quadratic (and hence agrees with a global quadratic) on each Q. Thus, restricted to any such Q, our original functions $a_i n \lfloor b_i n \lfloor c_i n \rfloor \rfloor$ agrees with a genuine cubic. Finally, we can decompose these progressions Q into ones where the cubics are approximately constant mod 1 (using abelian Schmidt for degree 3). While in theory this approach can be made to work for all such bracket Schmidt problems, this however necessitates working with bracket functions and rather quickly becomes messy to handle.

This procedure can be adapted to work with polynomial sequences on nilmanifolds directly due to an unpublished observation of Green and Tao. This is the approach we take in the present work. The crucial point is that given a polynomial sequence $g(n)$ with respect to a group G given a filtration $G_0 = G_1 \geqslant G_2 \geqslant \cdots \geqslant G_k \geqslant \mathrm{id}_G$, the polynomial sequence $g(n) \bmod G_2$ is a standard polynomial. Thus one can apply Schmidt to a standard polynomial and therefore (after passing to long subprogressions) one may factor $g(n) \bmod G_2 = \varepsilon(n) \cdot \gamma(n)$ where ε is smooth and γ lies in the lattice $\Gamma \bmod G_2$. One may then lift ε, γ from $G \bmod G_2$ to $\widetilde{\varepsilon}, \widetilde{\gamma}$ on G and analyze the polynomial sequence $\widetilde{\varepsilon}^{-1} g \widetilde{\gamma}^{-1}$ which now lives in the group G_2. One can iterate this procedure and inductively reduce G_2 to G_3 and so on, which allows us to solve the Schmidt problem for our nilmanifold. We remark that this procedure is an induction on the length of the filtration whereas the (closely related) approach taken in [11] is phrased as an induction on dimension. This difference is crucial for getting bounds in which the exponent depends polynomially on dimension.

1.2. **Organization and notation.** All definitions regarding nilsequences and associated complexity will be exactly as in [18, Section 3,4]. We refer the reader to that paper for all such definitions; we will only require degree filtrations in this paper.

We use standard asymptotic notation. Given functions $f = f(n)$ and $g = g(n)$, we write $f = O(g)$, $f \ll g$, $g = \Omega(f)$, or $g \gg f$ to mean that there is a constant C such that $|f(n)| \leq Cg(n)$ for sufficiently large n. We write $f \asymp g$ or $f = \Theta(g)$ to mean that $f \ll g$ and $g \ll f$, and write $f = o(g)$ or $g = \omega(f)$ to mean $f(n)/g(n) \to 0$ as $n \to \infty$. Subscripts on asymptotic notation indicate dependence of the bounds on those parameters. We will use the notation $[x] = \{1, 2 \ldots, \lfloor x \rfloor\}$. In this paper $x = \lfloor x \rfloor + \{x\}$ where $\{x\} \in [0, 1)$ and $\lfloor x \rfloor \in \mathbb{Z}$; we remark this is different than in [18]. We write $\|x\|_{\mathbb{R}/\mathbb{Z}} = \mathrm{dist}(x, \mathbb{Z})$ for $x \in \mathbb{R}$. Furthermore throughout this paper we abusively write $\log$ for $\max(\log(\cdot), e^e)$; this is to avoid trivial issues with small numbers.

Finally, in terms of organization, in Section 2 we solve the Schmidt problem for nilsequences and in Section 3 we prove Theorem 1.1.

2. Schmidt's problem for nilsequences

In this section, we prove that given a list of nilsequences on $[N]$, one can decompose $[N]$ into a controlled set of arithmetic progressions such that the nilsequences are almost constant on these sequences.

Lemma 2.1. *Consider nilmanifolds G_i/Γ_i for $1 \leq i \leq T$, each given a degree k filtration, having complexity bounded by M, dimension bounded by d, and for each $1 \leq i \leq T$ let $g_i(n)$ be a polynomial sequence with respect to the specified degree k-filtration on G_i.*

We may decompose $[N]$ into disjoint arithmetic progressions $\mathcal{P}_1, \ldots, \mathcal{P}_L$ such that following conditions hold:

- $N/L \geq N^{\Omega_k(1/(Td)^{O_k(1)})};$
- *We have*

$$\max_{\substack{1 \leq i \leq T \\ 1 \leq j \leq L}} \max_{n,n' \in \mathcal{P}_j} d_{G_i/\Gamma_i}(g_i(n)\Gamma_i, g_i(n')\Gamma_i) \leq M^{O_k(d^{O_k(1)})} \cdot N^{-\Omega_k(1/(Td)^{O_k(1)})}.$$

The key ingredient in this proof is a result of Schmidt [23] regarding finding small fractional parts of polynomials. We will need a version of this result with explicit quantification; this is explicitly stated in work of the authors [17, Proposition 3.7] although the argument is essentially verbatim from a paper of Green and Tao [9, Appendix A] generalized from quadratics to all degrees.

Proposition 2.2. *Fix an integer $k \geq 1$. There exist $c_k > 0$ such that the following holds. Let $\alpha_1, \ldots, \alpha_d$ be real numbers. Then*

$$\min_{1 \leq n \leq N} \max_{1 \leq i \leq d} \|\alpha_i n^k\|_{\mathbb{R}/\mathbb{Z}} \ll_k dN^{-c_k/d^2}.$$

As stated this result is for pure monomial phases and only provides a single point with small fractional part. This statement however can be "upgraded" via a straightforward iterative argument which is implicit in say [9, Proposition 6.4] (where the quadratic case is handled).

Lemma 2.3. *Fix an integer $k \geq 0$. Consider polynomials $Q_1, \ldots, Q_d$ of degree k. Then there exist disjoint arithmetic progressions $\mathcal{P}_1, \ldots, \mathcal{P}_L$ such that following conditions hold:*

- $N/L \geq N^{\Omega_k(1/d^{O_k(1)})}$
- *We have*

$$\max_{\substack{1 \leq i \leq T \\ 1 \leq j \leq L}} \max_{n,n' \in \mathcal{P}_j} \|Q_j(n) - Q_j(n')\|_{\mathbb{R}/\mathbb{Z}} \leq 2 \cdot N^{-\Omega_k(1/d^{O_k(1)})}.$$

Proof. We proceed by induction on k. The case $k = 0$ is trivial as $Q_j(\cdot)$ are constant. Furthermore we may assume that $N \geq \exp(d^{\Omega_k(1)})$ else we may break $[N]$ into singleton arithmetic progressions.

Let $Q_j(n) = \sum_{\ell=0}^{k} \alpha_{j,\ell} n^\ell$. Applying Proposition 2.2, there exists $D \leq N^{1/2}$ such that

$$\max_{1 \leq j \leq d} \|\alpha_{j,k} D^k\|_{\mathbb{R}/\mathbb{Z}} \ll_k dN^{-c_k/(2d^2)} =: \tau.$$

We break $[N]$ into arithmetic progressions of common difference D and with lengths between $2^{-1}\tau^{-1/(2k)}$ and $\tau^{-1/(2k)}$. Label these progressions $\mathcal{R}_1, \ldots, \mathcal{R}_{L'}$ with starting points s_i for $1 \leq i \leq L'$. We have

$$Q_j(Dn + s_i) = \alpha_{j,k} D^k n^k + Q_{j,i}(n)$$

for appropriately defined polynomials $Q_{j,i}(n)$ of degree at most $k - 1$. Note that for $n, n' \in [\tau^{-1/(2k)}]$, we have

$$\|Q(Dn + s_i) - Q(Dn' + s_i)\|_{\mathbb{R}/\mathbb{Z}} = \|\alpha_{j,k}D^k(n^k - (n')^k) + Q_{j,i}(n) - Q_{j,i}(n')\|_{\mathbb{R}/\mathbb{Z}}$$

$$\leq 2\tau^{-1/2} \cdot \|\alpha_{j,k}D^k\|_{\mathbb{R}/\mathbb{Z}} + \|Q_{j,i}(n) - Q_{j,i}(n')\|_{\mathbb{R}/\mathbb{Z}}$$

$$\leq 2\tau^{1/2} + \|Q_{j,i}(n) - Q_{j,i}(n')\|_{\mathbb{R}/\mathbb{Z}}.$$

4

The result now follows by induction applied to each $Q_{j,i}(n)$ for $1 \leq i \leq L'$ on the interval $[\tau^{-1/(2k)}]$ and using these decompositions to split the $\mathcal{R}_i$ into our final decomposition. Letting $N' = \tau^{-1/(2k)}$, the number of arithmetic progressions resulting is bounded by

$$(2N/N') \cdot (N')^{1-c_1/d^{c_2}} \leq N^{1-\Omega_k(1/d^{O_k(1)})},$$

where c_1, c_2 are the implicit constants for the inductive hypothesis $k-1$. The result follows. $\square$

We next require the following lemma controlling coefficients of polynomials which live in a restricted range mod 1. It will be convenient to recall the smoothness norm of a polynomial $P(n) = \sum_{i=0}^{k} \alpha_i \binom{n}{i}$ which is defined as

$$\|P\|_{C^{\infty}[N]} := \max_{1 \leq i \leq k} N^i \|\alpha_i\|_{\mathbb{R}/\mathbb{Z}}.$$

Lemma 2.4. *Fix an integer $k \geq 1$. There exists $c_k > 0$ such that if $\varepsilon \in (0, c_k)$ then the following holds. Consider a polynomial $P(n) = \sum_{i=0}^{k} \alpha_i \binom{n}{i}$. Suppose that for $n, n' \in [N]$, we have $\|P(n) - P(n')\|_{\mathbb{R}/\mathbb{Z}} \leq \varepsilon$. Then*

$$\|P\|_{C^{\infty}[N]} \ll_k \varepsilon.$$

Proof. Note that

$$\left| \sum_{n=1}^{N} e(P(n)) \right| \geq N/2.$$

By a quantitative version of Weyl which may be found in Green and Tao [12, Proposition 4.3], there exists $q \in \mathbb{N}$ with $q \ll_k 1$ such that

$$\|qP\|_{C^{\infty}[N]} \ll_k 1.$$

Let $1 \leq t \leq \lfloor N/(2k) \rfloor$ be an integer and note that

$$\alpha_k \cdot t^k = \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} \cdot P(t \cdot i + 1).$$

Via the triangle inequality, we therefore have

$$\|\alpha_k \cdot t^k\|_{\mathbb{R}/\mathbb{Z}} \leq 2^{k-1} \varepsilon.$$

Take t to be a prime between $\lfloor N/(4k) \rfloor$ and $\lfloor N/(2k) \rfloor$. Combining this with the estimate $\|q\alpha_k\|_{\mathbb{R}/\mathbb{Z}} \ll_k N^{-k}$ implies that $\|\alpha_k\|_{\mathbb{R}/\mathbb{Z}} \ll_k \varepsilon \cdot N^{-k}$. The result then follows by induction on k and applying the result for the degree $(k-1)$ polynomial $P'(n) = \sum_{i=0}^{k-1} \alpha_i \binom{n}{i}$. $\square$

With this we are in position to deduce the result for nilsequences along the lines sketched in Section 1.1.3.

Proof of Lemma 2.1. Consider the degree k filtration of the group G_i, $G_{i,0} = G_{i,1} \geqslant G_{i,2} \geqslant \cdots \geqslant G_{i,k} \geqslant \mathrm{id}_{G_i}$. We say the group G_i has a degree k filtration of type t if $G_{i,t} = G_i$ (i.e., the first $(t+1)$ groups in the filtration match). We prove the result by backwards induction on t assuming that all groups G_i have degree k filtrations of type t; note that the result is trivial when $t = k+1$ and we aim to prove the claim when $t = 1$. So, consider the case where the filtration has type t for some $1 \leq t \leq k$ and suppose that we already know cases of larger type.

Let $\mathcal{X}_i = \{X_{1,i}, \ldots, X_{\dim(G_i),i}\}$ denote the Mal'cev basis for G_i. By the classification of polynomial sequences (see [10, Lemma 6.7]), we have

$$g_i(n) = \exp\left(\sum_{j=1}^{\dim(G_i)} P_{i,j}(n) \cdot X_{j,i} \right)$$

5

where if $X_{j,i} \in (\mathcal{X}_i \cap \log(G_{i,\ell})) \setminus (\mathcal{X}_i \cap \log(G_{i,\ell+1}))$ then the degree of polynomial $P_{i,j}(n)$ is bounded by ℓ.

We consider the polynomials $P_{i,j}(n)$ for $1 \le i \le T$ and $1 \le j \le \dim(G_i) - \dim(G_{i,t+1})$. The degrees of $P_{i,j}(n)$ are all at most $t \le k$ and the total number of polynomials number consideration is bounded by $T \cdot d$. By applying Lemma 2.3, there exists a decomposition of $[N]$ into arithmetic progressions $\mathcal{P}_1, \ldots, \mathcal{P}_L$ such that:

- $N/L \ge N^{\Omega_k(1/(dT)^{O_k(1)})}$
- We have

$$\max_{\substack{1 \le i \le T \\ 1 \le j \le \dim(G_i) - \dim(G_{i,t+1})}} \max_{1 \le s \le L} \max_{n,n' \in \mathcal{P}_s} \|P_{i,j}(n) - P_{i,j}(n')\|_{\mathbb{R}/\mathbb{Z}} \le 2 \cdot N^{-\Omega_k(1/(dT)^{O_k(1)})}.$$

We break the progressions $\mathcal{P}_s$ into two classes: the first class ($s \in \mathcal{S}$) if the progression has length bounded by $\sqrt{N/L}$ and the second class ($s \in \mathcal{L}$) otherwise. For progressions which are short, we break each such progression into singletons; after this there are at most $L + \sqrt{N/L} \cdot L \le 2\sqrt{NL}$ progressions which is qualitatively identical to before. For each $s \in \mathcal{L}$, we write $\mathcal{P}_s = \{a_s n + b_s\}_{n \in [|\mathcal{P}_s|]}$ where $|\mathcal{P}_s|$ denotes the length of the progression.

Using the second condition above and applying Lemma 2.4, we see that for each long progression $\mathcal{P}_s$, we have for all i,j that

$$P_{i,j}(a_s n + b_s) = P_{i,j,s,\text{small}}(n) + P_{i,j,s,\text{int}}(n)$$

where:

- $\deg(P_{i,j,s,\text{int}}), \deg(P_{i,j,s,\text{small}}) \le \deg(P_{i,j})$
- $P_{i,j,s,\text{int}}$ maps $\mathbb{Z} \to \mathbb{Z}$
- If $P_{i,j,s,\text{small}}(n) = \sum_{r=0}^{t} \alpha_{i,j,s,\text{small},r} \binom{n}{r}$ then

$$|\alpha_{i,j,s,\text{small},r}| \le N^{-r} \cdot N^{-\Omega_k(1/(dT)^{O_k(1)})}$$

for $1 \le r \le t$ and $|\alpha_{i,j,s,\text{small},0}| \le 1$.

We have implicitly used $|\mathcal{P}_s| \ge \sqrt{N/L} \ge N^{\Omega_k(1/(dT)^{O_k(1)})}$ for $s \in \mathcal{L}$ here.

The key trick is to now "reduce" the polynomial sequence g_i to one which lives in $G_{i,t+1}$. Define

- $\varepsilon_{i,s}(n) = \exp\left(\sum_{j=1}^{\dim(G_i) - \dim(G_{i,t+1})} P_{i,j,s,\text{small}}(n) \cdot X_j\right)$
- $\gamma_{i,s}(n) = \prod_{j=1}^{\dim(G_i) - \dim(G_{i,t+1})} \exp(X_j)^{P_{i,j,s,\text{int}}(n)}$
- $g'_{i,s}(n) = \varepsilon_{i,s}(n)^{-1} \cdot g(a_s n + b_s) \cdot \gamma_{i,s}(n)^{-1}$

Note that $\varepsilon_{i,s}$, $\gamma_{i,s}$ are polynomial sequences with respect to the filtration given on G_i by the classification of polynomial sequences (see [10, Lemma 6.7]) and the fact that the set of polynomial sequences form a group. Therefore $g'_{i,s}$ is also seen to be a polynomial sequence. The crucial point, however, is that by the Baker–Campbell–Hausdorff formula, we have that $g'_{i,s}$ only takes on values in $G_{i,t+1}$. (We are using the assumption on type that $G_i = G_{0,i} = G_{t,i}$, so any commutator is in $G_{2t,i} \le G_{t+1,i}$ since $t \ge 1$.)

Therefore we may inductively apply the claim for each long progression $\mathcal{P}_s$, to the polynomials $g'_{i,s}$ on $G_{i,t+1}$ where we take the filtration on G_i intersected with $G_{i,t+1}$ (note that the filtration is still degree k). The corresponding Mal'cev basis is given by taking the last $\dim(G_{i,t+1})$ elements of $\mathcal{X}_i$. By induction therefore we may break each long $\mathcal{P}_s$ into L_s such progressions $\mathcal{P}_{s,r}$ where $L_s \le |\mathcal{P}_s|^{1-\Omega_k(1/(Td)^{O_k(1)})}$ and such that

$$\max_{\substack{s \in \mathcal{L} \\ 1 \le r \le L_s}} \max_{n,n' \in \mathcal{P}_{s,r}} d_{G_i/\Gamma_i}(g'_{i,s}(n)\Gamma_i, g'_{i,s}(n')\Gamma_i) \le M^{O_k(d^{O_k(1)})} \cdot N^{-\Omega_k(1/d^{O_k(1)})}.$$

Here we are using [16, Lemma B.9] to compare distances between G_i and $G_{i,t+1}$.

6

Furthermore note that $\gamma_{i,s}$ takes values only in Γ by the definition of a Mal'cev basis and that for $n, n' \in [|\mathcal{P}_s|]$ we have

$$d_{G_i}(\varepsilon_{i,s}(n), \mathrm{id}_{G_i}) \leq M^{O_k(d^{O_k(1)})} \text{ and } d_{G_i}(\varepsilon_{i,s}(n), \varepsilon_{i,s}(n')) \leq M^{O_k(d^{O_k(1)})} \cdot N^{-\Omega_k(1/(dT)^{O_k(1)})}.$$

This is due to our bounds on the smoothness norm of $P_{i,j,s,\mathrm{small}}$ and [16, Lemma B.3].

It therefore follows by [16, Lemma B.4] that for any s, r we have

$$
\begin{aligned}
\max_{n,n' \in \mathcal{P}_{s,r}} & d_{G_i/\Gamma_i}(g_i(a_s n + b_s)\Gamma_i, g_i(a_s n' + b_s)\Gamma_i) \\
&= \max_{n,n' \in \mathcal{P}_{s,r}} d_{G_i/\Gamma_i}(\varepsilon_{i,s}(n)g'_{i,s}(n)\Gamma_i, \varepsilon_{i,s}(n')g'_{i,s}(n')\Gamma_i) \\
&\leq \max_{n,n' \in \mathcal{P}_{s,r}} d_{G_i/\Gamma_i}(\varepsilon_{i,s}(n)g'_{i,s}(n)\Gamma_i, \varepsilon_{i,s}(n)g'_{i,s}(n')\Gamma_i) \\
&\qquad + \max_{n,n' \in \mathcal{P}_{s,r}} d_{G_i/\Gamma_i}(\varepsilon_{i,s}(n)g'_{i,s}(n')\Gamma_i, \varepsilon_{i,s}(n')g'_{i,s}(n')\Gamma_i) \\
&\leq M^{O_k(d^{O_k(1)})} \sup_{n,n' \in \mathcal{P}_{s,r}} d_{G_i/\Gamma_i}(g'_{i,s}(n)\Gamma_i, g'_{i,s}(n')\Gamma_i) + \max_{n,n' \in \mathcal{P}_{s,r}} d_{G_i}(\varepsilon_{i,s}(n), \varepsilon_{i,s}(n')) \\
&\leq M^{O_k(d^{O_k(1)})} \cdot N^{-\Omega_k(1/(dT)^{O_k(1)})}
\end{aligned}
$$

which completes the inductive step (our final decomposition is composed of all elements of the short $\mathcal{P}_s$ indexed by $s \in \mathcal{S}$ and all $\mathcal{P}_{s,r}$ arising from the long progressions indexed by $s \in \mathcal{L}$). We are done, noting that the number of inductive steps (hence the decay in parameters) is bounded in terms of k. $\qquad\square$

3. Completing the proof

We are now run the Heath–Brown [14] and Szemerédi [26] density increment strategy as reformulated by Green and Tao [9]. In the first subsection we recall a number of preliminaries for the proof and in the second subsection we prove Theorem 1.1. Our treatment at this point is quite close to that of [9] and we borrow certain elements from the density increment portion of [19] as well.

3.1. **Preliminaries for density increment.** We first recall the definition of the Gowers U^s-norm over the integers.

Definition 3.1. *Given $f \colon \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ and $s \geq 1$, we define*

$$\|f\|_{U^s(\mathbb{Z}/N\mathbb{Z})}^{2^s} = \mathbb{E}_{x,h_1,\ldots,h_s \in \mathbb{Z}/N\mathbb{Z}} \Delta_{h_1,\ldots,h_s} f(x)$$

where $\Delta_h f(x) = f(x)\overline{f(x+h)}$ is the multiplicative discrete derivative (extended to vectors h in the natural way).

Given a natural number N and a function $f \colon [N] \to \mathbb{C}$, we choose a number $\widetilde{N} \geq 2^s N$ and define $\widetilde{f} \colon \mathbb{Z}/\widetilde{N}\mathbb{Z} \to \mathbb{C}$ via $\widetilde{f}(x) = f(x)$ for $x \in [N]$ and 0 otherwise. Then

$$\|f\|_{U^s[N]} := \|\widetilde{f}\|_{U^s(\mathbb{Z}/\widetilde{N}\mathbb{Z})} / \|\mathbb{1}_{[N]}\|_{U^s(\mathbb{Z}/\widetilde{N}\mathbb{Z})}.$$

One can check that this definition does not depend on the choice of $\widetilde{N}$. This is well known to be a seminorm for $s \geq 1$ and a norm for $s \geq 2$.

As mentioned, the main input for our result will be the following improved bound for the U^s-norm inverse theorem given as [18, Theorem 1.2].

Theorem 3.2. *Fix $\delta \in (0, 1/2)$. Suppose that $f \colon [N] \to \mathbb{C}$ is 1-bounded and*

$$\|f\|_{U^{s+1}[N]} \geq \delta.$$

7

Then there exists a nilmanifold G/Γ of degree s, complexity at most M, and dimension at most d as well as a function F on G/Γ which is at most K-Lipschitz such that

$$|\mathbb{E}_{n\in[N]}[f(n)\overline{F(g(n)\Gamma)}]| \geq \varepsilon,$$

where we may take

$$d \leq \log(1/\delta)^{O_s(1)} \text{ and } \varepsilon^{-1}, K, M \leq \exp(\log(1/\delta)^{O_s(1)}).$$

We now define the k-fold linear operator corresponding to counting k-term arithmetic progressions. Given functions $f_i \colon [N] \to \mathbb{C}$, define

$$\Lambda_k(f_1,\ldots,f_k) = \mathbb{E}_{x,y\in\{0,\ldots,N\}} \prod_{j=1}^{k} f_j(x+(j-1)y)$$

where f_i are extended by 0 outside of $[N]$. We also write

$$\Lambda_k(f) := \Lambda_k(f,\ldots,f).$$

We have the following basic inequalities regarding the operator Λ_k. The proof is by now standard and hence is omitted (see [9, Lemma 3.2] and [8, Theorem 3.2]).

Lemma 3.3. *Consider functions $f_i \colon [N] \to \mathbb{C}$ for $1 \leq i \leq k$. Then we have*

$$\Lambda_k(f_1,\ldots,f_k) \leq \min_{1\leq i\leq k} \|f_i\|_{L^1[N]} \cdot \prod_{j\neq i}\|f_j\|_{L^\infty[N]},$$

$$\Lambda_k(f_1,\ldots,f_k) \ll_k \min_{1\leq i\leq k} \|f_i\|_{U^{k-1}[N]} \cdot \prod_{j\neq i}\|f_j\|_{L^\infty[N]}.$$

We next define factors and the factor induced by function g of resolution K.

Definition 3.4. *We define a* factor *$\mathcal{B}$ of $[N]$ to be a partition $[N] = \bigsqcup_{B\in\mathcal{B}} B$. We define $\mathcal{B}(x)$ for $x \in [N]$ to be the part of $\mathcal{B}$ that contains x. We say $\mathcal{B}'$ refines $\mathcal{B}$ if every part of $\mathcal{B}$ can be written as a disjoint union of parts of $\mathcal{B}'$. We define a* join *of a sequence of factors to be the partition (discarding empty parts)*

$$\mathcal{B}_1 \vee \cdots \vee \mathcal{B}_d := \{B_1 \cap \cdots \cap B_d \colon B_i \in \mathcal{B}_i\}.$$

Next given a function $g \colon [N] \to \mathbb{R}$ and a resolution K, we define the factor induced by g of resolution K to be

$$\mathcal{B}_{g,K} = \bigsqcup_{j\in\mathbb{Z}} \{x \in [N] \colon g(x) \in [j/K, (j+1)/K).$$

Finally, given a factor $\mathcal{B}$, we define $\Pi_{\mathcal{B}}f$ by

$$\Pi_{\mathcal{B}}f(x) = \mathbb{E}_{y\in\mathcal{B}(x)}f(y).$$

A technical annoyance is that one may potentially have a large set of points near the cutoffs when defining $\mathcal{B}_{g,K}$. We define a notion of regularity capturing when a function g avoids such issues, which is related to an idea introduced by Bourgain [4] with regards to Bohr sets.

Definition 3.5. *The factor $\mathcal{B}_{g,K}$ is C-regular if*

$$\sup_{r>0} \left(\frac{1}{2r}\frac{1}{N}\left|\{x\in[N] \colon \|K\cdot g(x)\|_{\mathbb{R}/\mathbb{Z}} \leq r\}\right|\right) \leq C.$$

It turns out to be easy to obtain "regular" factors; a useful trick (motivated by the proof of [10, Corollary 2.3]) is to consider a random shift of g and then apply the Hardy—Littlewood maximal inequality. Given a function g and resolution K, we define the maximal function

$$M_{g,K}(t) := \sup_{r>0} \frac{1}{2r}\frac{1}{N}\left|\{x\in[N] \colon \|K\cdot g(x) - t\|_{\mathbb{R}/\mathbb{Z}} \leq r\}\right|.$$

The Hardy–Littlewood maximal inequality (on the torus $\mathbb{R}/\mathbb{Z}$) implies that

$$\mathbb{E}_{t\in[0,1]}[M_{g,K}(t)] = O(1).$$

Therefore we have the following elementary fact which will prove useful.

Fact 3.6. *There exists a constant $C = C_{3.6} > 0$ such that the following holds. Given a function $g\colon [N] \to \mathbb{R}$ and a resolution K, there exists a shift $t \in [0, 1/K)$ such that $\mathcal{B}_{g-t,K}$ is C-regular.*

3.2. **Constructing factor approximation and density increment.** The key claim which we need to prove Theorem 1.1 is the following density increment lemma, phrased as a trichotomy.

Lemma 3.7. *Fix an integer $k \geq 5$ and a constant $c > 0$. Consider a function $f\colon [N] \to [0,1]$ such that $\mathbb{E}_{n\in[N]}f(n) = \delta$. There exist $c' = c'(c,k)$ and $C = C(c,k)$ such that one of the following always holds:*

- $N \leq \exp(\log(1/\delta)^C)$;
- $|\Lambda_k(f) - \Lambda_k(\delta \cdot \mathbb{1}_{[N]})| \leq c\delta^k$;
- *There exists an arithmetic progression $\mathcal{P} \subseteq [N]$ of length at least $N^{1/\exp(\log(1/\delta)^C)}$ such that*

$$\mathbb{E}_{n\in\mathcal{P}}f(n) \geq (1+c')\delta.$$

We prove Theorem 1.1 given Lemma 3.7; this is the standard density increment strategy.

Proof of Theorem 1.1 given Lemma 3.7. Suppose $A \subseteq [N]$ has no k-term arithmetic progressions. We iteratively increase the density of A; set $A = A_1$, $N = N_1$ and $\delta = \delta_1$ and we iteratively define $A_i \subseteq [N_i]$, and $\delta_i = |A_i|/N_i$.

If $N_i \leq \exp(\log(1/\delta_i)^C)$, we immediately terminate. Otherwise, note that as A_i is free of k-term arithmetic progressions, we have that

$$|\Lambda_k(\mathbb{1}_{A_i}) - \Lambda_k(\delta_i \cdot \mathbb{1}_{[N_i]})| \geq \delta_i^k \cdot |\Lambda_k(\mathbb{1}_{[N_i]})| - |A_i| \cdot N_i^{-2} \gg_k \delta_i^k$$

where we have used that $N_i \geq \exp(\log(1/\delta_i)^C) \gg \delta_i^{-k}$. Therefore, the third case in Lemma 3.7 occurs and there exists $\mathcal{P}_{i+1}$ such that

$$|A_i \cap \mathcal{P}_{i+1}|/|\mathcal{P}_{i+1}| \geq (1+c')\delta_i$$

and $|\mathcal{P}_{i+1}| \geq N_i^{1/\exp(\log(1/\delta_i)^C)}$. We now rescale the arithmetic progression $\mathcal{P}_{i+1}$ to $[|\mathcal{P}_{i+1}|] =: [N_{i+1}]$, which sends $A_i \cap \mathcal{P}_{i+1}$ to a new set A_{i+1}, and then we continue the iteration.

Note that at every iteration the density δ_i increases by a multiplicative factor of at least $(1+c')$, so we must terminate in at most $O_k(\log(1/\delta))$ iterations. Thus there exists an index $j \leq O_k(\log(1/\delta))$ such that

$$N^{1/\exp(O_k(\log(1/\delta)^{C+1}))} \leq N_j \leq \exp(\log(1/\delta_j)^C) \leq \exp(\log(1/\delta)^C).$$

This implies that

$$\log N \leq \exp(O_k(\log(1/\delta)^{O_k(1)}))$$

and thus

$$\delta \leq \exp(-(\log\log N)^{\Omega_k(1)}). \qquad \square$$

In order to prove Lemma 3.7, we first iterate Theorem 3.2 to obtain the following result.

Lemma 3.8. *Fix a parameter $\eta \in (0, 1/2)$ and $k \geq 5$. There exists a constant $C = C_k > 0$ such that the following statement holds. If $N \geq \exp(\log(1/\eta)^C)$ and $f\colon [N] \to \mathbb{R}$ is 1-bounded then there exist functions $h_1, \ldots, h_T\colon [N] \to \mathbb{R}$ and $d, M, K \geq 1$ such that:*

- $\mathcal{B} = \bigvee_{1 \leq i \leq T} \mathcal{B}_{h_i,K}$ *satisfies* $\|f - \Pi_{\mathcal{B}}f\|_{U^{k-1}[N]} \leq \eta$;
- $T, M, K \leq \exp(\log(1/\eta)^C)$ *and* $d \leq \log(1/\eta)^C$;

- $h_i = F_i(g_i(n)\Gamma_i)$ *is a nilsequence where $g_i(n)$ takes values in a group G_i which is given a degree $(k-2)$ filtration, G_i/Γ_i has complexity bounded by M and dimension bounded by d, and $F_i\colon G_i/\Gamma_i \to \mathbb{R}$ is M-Lipschitz;*
- $\mathcal{B}_{h_i,K}$ *is C-regular for $1 \le i \le T$.*

Proof. The proof follows via applying Theorem 3.2 repeatedly. We begin the iteration by setting $\mathcal{B}_0 = [N]$ (i.e., the trivial partition). At each stage we will construct h_{i+1} and then set $\mathcal{B}_{i+1} = \mathcal{B}_i \vee \mathcal{B}_{h_{i+1},K}$ with $K = \lceil \exp(\log(1/\eta)^{O_k(1)}) \rceil$, where the implicit constant is chosen sufficiently large.

Step 1: If $\|f - \Pi_{\mathcal{B}_i} f\|_{U^{k-1}[N]} \le \eta$, we terminate.

Step 2: If $\|f - \Pi_{\mathcal{B}_i} f\|_{U^{k-1}[N]} > \eta$, then by Theorem 3.2, there exists nilsequence $h_{i+1}(n) = F_{i+1}(g_{i+1}(n)\Gamma_{i+1})$ such that

$$\left| \mathbb{E}_{n\in[N]}[(f - \Pi_{\mathcal{B}_i} f)(n)\overline{F_{i+1}(g_{i+1}(n)\Gamma_{i+1})}] \right| \ge \exp(-\log(1/\eta)^{O_k(1)})$$

and where G_{i+1}/Γ_{i+1} has complexity bounded by $\exp(\log(1/\eta)^{O_k(1)})$ and dimension bounded by $\log(1/\eta)^{O_k(1)}$, $F_{i+1}\colon G_{i+1}/\Gamma_{i+1} \to \mathbb{C}$ is $\exp(\log(1/\eta)^{O_k(1)})$–Lipschitz, G_{i+1} has been given a degree $(k-2)$ filtration, and where $g_{i+1}(n)$ is a polynomial sequence with respect to this filtration. Taking either the real or imaginary part of F_{i+1}, we may assume that $F_{i+1}\colon G_{i+1}/\Gamma_{i+1} \to \mathbb{R}$ and thus that

$$\left| \mathbb{E}_{n\in[N]}[(f - \Pi_{\mathcal{B}_i} f)(n)F_{i+1}(g_{i+1}(n)\Gamma_{i+1})] \right| \ge \exp(-\log(1/\eta)^{O_k(1)}).$$

Note that for any $t \in [0, 1/K)$, this implies that

$$\left| \mathbb{E}_{n\in[N]} \left[(f - \Pi_{\mathcal{B}_i} f)(n)\frac{\lfloor K(F_{i+1}(g_{i+1}(n)\Gamma_{i+1}) + t)\rfloor}{K} \right] \right|$$
$$\ge \left| \mathbb{E}_{n\in[N]}[(f - \Pi_{\mathcal{B}_i} f)(n)F_{i+1}(g_{i+1}(n)\Gamma_{i+1})] \right| - 2/K$$
$$\ge \exp(-\log(1/\eta)^{O_k(1)})$$

given that the implicit constant defining K is chosen sufficiently large. Recall here $\lfloor x \rfloor$ is defined in the standard manner that $x = \lfloor x \rfloor + \{x\}$ where $\lfloor x \rfloor \in \mathbb{Z}$ and $\{x\} \in [0, 1)$. We then take $t \in [0, 1/K)$, such that $\mathcal{B}_{F_{i+1}(g_{i+1}(n)\Gamma_{i+1})+t,K}$ is C-regular; this exists for C larger than an absolute constant by Fact 3.6.

Note that

$$\frac{\lfloor K(F_{i+1}(g_{i+1}(n)\Gamma_{i+1}) + t)\rfloor}{K}$$

is measurable with respect to $\mathcal{B}_{h_{i+1},K}$ by construction and it is bounded by $\log(1/\eta)^{O_k(1)}$. Therefore since $\Pi_{\mathcal{B}_{h_{i+1},K}}$ is self-adjoint we have

$$\mathbb{E}_{n\in[N]}[|\Pi_{\mathcal{B}_{h_{i+1},K}}(f - \Pi_{\mathcal{B}_i} f)(n)|]$$
$$\ge (1 + \|F_{i+1}\|_{L^\infty(G_{i+1}/\Gamma_{i+1})})^{-1} \cdot \left| \mathbb{E}_{n\in[N]} \left[(f - \Pi_{\mathcal{B}_i} f)(n)\frac{\lfloor K(F_{i+1}(g_{i+1}(n)\Gamma_{i+1}) + t)\rfloor}{K} \right] \right|$$
$$\ge \exp(-\log(1/\eta)^{O_k(1)}).$$

Step 3: We now return back to Step 1 and keep on iterating this procedure until it terminates. This completes the proof modulo showing that the iteration terminates in a small number of steps. To show this, note that

$$\|\Pi_{\mathcal{B}_{h_{i+1},K}}(f - \Pi_{\mathcal{B}_i} f)\|_{L^1[N]} \le \|\Pi_{\mathcal{B}_{h_{i+1},K}}(f - \Pi_{\mathcal{B}_i} f)\|_{L^2[N]} = \|\Pi_{\mathcal{B}_{h_{i+1},K}}\Pi_{\mathcal{B}_{i+1}}(f - \Pi_{\mathcal{B}_i} f)\|_{L^2[N]}$$
$$\le \|\Pi_{\mathcal{B}_{i+1}}(f - \Pi_{\mathcal{B}_i} f)\|_{L^2[N]} = \|\Pi_{\mathcal{B}_{i+1}} f - \Pi_{\mathcal{B}_i} f\|_{L^2[N]}$$
$$= (\|\Pi_{\mathcal{B}_{i+1}} f\|^2_{L^2[N]} - \|\Pi_{\mathcal{B}_i} f\|^2_{L^2[N]})^{1/2}.$$

The final equality is the Pythagorean theorem with respect to projections (this follows from e.g. [19, Lemma 4.3(iv)]). We deduce

$$\|\Pi_{\mathcal{B}_{i+1}}f\|_{L^2[N]}^2 - \|\Pi_{\mathcal{B}_i}f\|_{L^2[N]}^2 \geq \exp(-\log(1/\eta)^{O_k(1)}).$$

Since for all i we have $\|\Pi_{\mathcal{B}_i}f\|_{L^2[N]} \leq \|f\|_{L^2[N]} \leq 1$, there are at most $\exp(\log(1/\eta)^{O_k(1)})$ iterations as desired. $\qquad\square$

We now complete the proof of Lemma 3.7 and therefore the proof of Theorem 1.1. The first part of the proof is finding a density increment on a factor derived from nilsequences, which is essentially identical to that of [9, Lemma 5.8]. In the second part, we apply our nilsequence Schmidt-type result Lemma 2.1 to find a long arithmetic progression with density increment.

Proof of Lemma 3.7. Without loss of generality, we may assume that c is smaller than an absolute constant. Furthermore we may assume that $N \geq \exp(\log(1/\delta)^{\Omega(1)})$ (where the implicit constant may depend on c, k) and $|\Lambda_k(f) - \Lambda_k(\delta \cdot \mathbb{1}_{[N]})| \geq c\delta^5$.

Step 1: Increment on a factor. By applying Lemma 3.8, there exists factor $\mathcal{B}$ (derived from nilsequences of appropriate complexity, with parameters below) such that

$$\|\Pi_{\mathcal{B}}f - f\|_{U^{k-1}[N]} \leq c^*\delta^k$$

where we choose c^* sufficiently small in terms of c. Via telescoping and the second inequality in Lemma 3.3, we have

$$|\Lambda_k(f) - \Lambda_k(\Pi_{\mathcal{B}}f)| \leq c\delta^k/2$$

as long as c^* was chosen appropriately, and therefore

$$|\Lambda_k(\Pi_{\mathcal{B}}f) - \Lambda_k(\delta \cdot \mathbb{1}_{[N]})| \geq c\delta^k/2.$$

Take $c' = \min(c, 1)/(10k)^5$. Let $g = \min(\Pi_{\mathcal{B}}f, (1 + c')\delta)$. The crucial claim is that if $\Omega' = \{n \in [N] : g(n) \neq \Pi_{\mathcal{B}}f(n)\} = \{n \in [N] : \Pi_{\mathcal{B}}f(n) > (1 + c')\delta\}$ then Ω' must have sufficiently large measure. To see this note that:

$$|\Lambda_k(\Pi_{\mathcal{B}}f) - \Lambda_k(g)| \leq k\|\Pi_{\mathcal{B}}f - g\|_{L^1[N]} \leq k\mathbb{P}_{n\in[N]}[n \in \Omega'],$$
$$|\Lambda_k(\delta\mathbb{1}_{[N]}) - \Lambda_k(g)| \leq k(1 + c')^{k-1}\delta^{k-1}\|\delta\mathbb{1}_{[N]} - g\|_{L^1[N]},$$
$$\|g - \delta\mathbb{1}_{[N]}\|_{L^1[N]} \leq \mathbb{P}_{n\in[N]}[n \in \Omega'] + \|\delta\mathbb{1}_{[N]} - \Pi_{\mathcal{B}}f\|_{L^1[N]}.$$

The first and second inequality follow from the first part of Lemma 3.3 and telescoping while the final inequality follows from the triangle inequality. We simplify the inequalities slightly; as $\mathbb{E}_{n\in[N]}[\delta\mathbb{1}_{[N]}] = \mathbb{E}_{n\in[N]}[f] = \mathbb{E}_{n\in[N]}[\Pi_{\mathcal{B}}f]$, we have

$$\|\delta\mathbb{1}_{[N]} - \Pi_{\mathcal{B}}f\|_{L^1[N]} = 2\|\max(\Pi_{\mathcal{B}}f - \delta\mathbb{1}_{[N]}, 0)\|_{L^1[N]} \leq 2c'\delta + 2\mathbb{P}_{n\in[N]}[n \in \Omega'].$$

Given this and using the upper bound on c', we deduce

$$|\Lambda_k(\Pi_{\mathcal{B}}f) - \Lambda_k(g)| \leq k\|\Pi_{\mathcal{B}}f - g\|_{L^1[N]} \leq k\mathbb{P}_{n\in[N]}[n \in \Omega'],$$
$$|\Lambda_k(\delta\mathbb{1}_{[N]}) - \Lambda_k(g)| \leq 2k\delta^{k-1}\|\delta\mathbb{1}_{[N]} - g\|_{L^1[N]},$$
$$\|g - \delta\mathbb{1}_{[N]}\|_{L^1[N]} \leq 3\mathbb{P}_{n\in[N]}[n \in \Omega'] + 2c'\delta.$$

Therefore

$$c\delta^k/2 \leq |\Lambda_k(\delta \cdot \mathbb{1}_{[N]}) - \Lambda_k(\Pi_{\mathcal{B}}f)| \leq |\Lambda_k(\delta \cdot \mathbb{1}_{[N]}) - \Lambda_k(g)| + |\Lambda_k(\Pi_{\mathcal{B}}f) - \Lambda_k(g)|$$
$$\leq k\mathbb{P}_{n\in[N]}[n \in \Omega'] + 2k\delta^{k-1}\|\delta\mathbb{1}_{[N]} - g\|_{L^1[N]} \leq 7k\mathbb{P}_{n\in[N]}[n \in \Omega'] + 4kc'\delta^k;$$

thus we have $\mathbb{P}_{n\in[N]}[n \in \Omega'] \geq c\delta^k/(20k)$.

11

Step 2: Increment on a progression. We are now in position to apply the nilsequence Schmidt-type result Lemma 2.1. Recall that we applied Lemma 3.8 to find $\mathcal{B}$, and hence we may write $\mathcal{B} = \bigvee_{1 \leq i \leq T} \mathcal{B}_{h_i,K}$ where:

- $T, M, K \leq \exp(\log(1/\delta)^C)$ and $d \leq \log(1/\delta)^C$;
- $h_i = F_i(g_i(n)\Gamma_i)$ is a nilsequence where $g_i(n)$ takes values in a group G_i which is given a degree $(k-2)$ filtration, G_i/Γ_i has complexity bounded by M and dimension bounded by d, and $F_i \colon G_i/\Gamma_i \to \mathbb{R}$ is M-Lipschitz;
- $\mathcal{B}_{h_i,K}$ is C-regular for $1 \leq i \leq T$.

Here C is a slightly larger value than the constant C_k in Lemma 3.8, depending only on k.

We now apply Lemma 2.1 to $g_i(n)$ for $1 \leq i \leq T$. We obtain a decomposition of $[N]$ into arithmetic progressions $\mathcal{P}_1, \ldots, \mathcal{P}_L$ such that

- $N/L \geq N^{-1/\exp(\log(1/\delta)^{O_k(1)})}$;
- We have

$$\max_{\substack{1 \leq i \leq T \\ 1 \leq j \leq L}} \max_{n,n' \in \mathcal{P}_j} d_{G_i/\Gamma_i}(g_i(n)\Gamma_i, g_i(n')\Gamma_i) \leq \exp(\log(1/\delta)^{O_k(1)}) \cdot N^{-1/\exp(\log(1/\delta)^{O_k(1)})}.$$

We now consider $\mathcal{P}_j$ which intersect Ω'. Call a progression in the decomposition *crossing* if it intersects Ω' and $[N] \setminus \Omega'$ and a progression *contained* if it is fully within in Ω'. Since Ω' is measurable in terms of $\mathcal{B}$, for a progression to be crossing it must "cross a boundary" defining $\mathcal{B}_{h_i,K}$ for at least one $1 \leq i \leq T$. If a progression $\mathcal{P}_j$ crosses one of these boundaries defined by h_i then all points in map close to this boundary, since the function F_i is M-Lipschitz. In particular, by regularity of each $\mathcal{B}_{h_i,K}$, the measure (with respect to the uniform distribution on $[N]$) of improper progressions is bounded by

$$\ll_k T \cdot \exp(\log(1/\delta)^{O_k(1)}) \cdot N^{-1/\exp(\log(1/\delta)^{O_k(1)})} = \exp(\log(1/\delta)^{O_k(1)}) \cdot N^{-1/\exp(\log(1/\delta)^{O_k(1)})}.$$

Let Ω^* denote the union of all the contained progressions which have length at least $N' = cc'\delta^{k+1}/(400k) \cdot N/L$ (hence certainly $\Omega^* \subseteq \Omega'$). Let $\mathcal{I}$ be the set of all $1 \leq i \leq L$ so that $\mathcal{P}_i$ either has length at most N' or is crossing. We easily see that

$$0 \leq \mathbb{P}_{n \in [N]}[n \in \Omega'] - \mathbb{P}_{n \in [N]}[n \in \Omega^*] \leq \sum_{i \in \mathcal{I}} \mathbb{P}_{n \in [N]}[n \in \mathcal{P}_i] \leq cc'\delta^{k+1}/(200k);$$

in the final inequality we have used that $N \geq \exp(\log(1/\delta)^{\Omega(1)})$ for a sufficiently large implicit constant.

Finally, this implies that

$$\begin{aligned}
\mathbb{E}_{n \in \Omega^*}[f] = \frac{\mathbb{E}_{n \in [N]}[f \cdot \mathbb{1}_{n \in \Omega^*}]}{\mathbb{P}_{n \in [N]}[n \in \Omega^*]} &\geq \frac{\mathbb{E}_{n \in [N]}[f \cdot \mathbb{1}_{n \in \Omega'}] - cc'\delta^{k+1}/(200k)}{\mathbb{P}_{n \in [N]}[n \in \Omega']} \\
&\geq \frac{\mathbb{E}_{n \in [N]}[f \cdot \mathbb{1}_{n \in \Omega'}]}{\mathbb{P}_{n \in [N]}[n \in \Omega']} - \frac{cc'\delta^{k+1}/(200k)}{c\delta^k/(20k)} \\
&= \frac{\mathbb{E}_{n \in [N]}[\Pi_{\mathcal{B}} f \cdot \mathbb{1}_{n \in \Omega'}]}{\mathbb{P}_{n \in [N]}[n \in \Omega']} - c'\delta/10 \\
&\geq (1 + c')\delta - c'\delta/10 \geq (1 + c'/2)\delta.
\end{aligned}$$

By pigeonhole, this implies that there exists a contained arithmetic progression $\mathcal{P}_i$ having length at least $cc'\delta^{k+1}/(400k) \cdot N/L \geq N^{-1/\exp(\log(1/\delta)^{O_k(1)})}$ on which the density of f is at least $(1 + c'/2)\delta$. Adjusting the value of c', this completes the proof. $\qquad\square$

References

[1] T. F. Bloom, *A quantitative improvement for Roth's theorem on arithmetic progressions*, J. Lond. Math. Soc. (2) **93** (2016), 643–663.

[2] T. F. Bloom and O. Sisask, *Breaking the logarithmic barrier in Roth's theorem on arithmetic progressions*, arXiv:2007.03528.

[3] T. F. Bloom and O. Sisask, *An improvement to the Kelley-Meka bounds on three-term arithmetic progressions*, arXiv:2309.02353.

[4] J. Bourgain, *On triples in arithmetic progression*, Geom. Funct. Anal. **9** (1999), 968–984.

[5] J. Bourgain, *Roth's theorem on progressions revisited*, J. Anal. Math. **104** (2008), 155–192.

[6] W. T. Gowers, *A new proof of Szemerédi's theorem for arithmetic progressions of length four*, Geom. Funct. Anal. **8** (1998), 529–551.

[7] W. T. Gowers, *Arithmetic progressions in sparse sets*, Current developments in mathematics, 2000, Int. Press, Somerville, MA, 2001, pp. 149–196.

[8] W. T. Gowers, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal. **11** (2001), 465–588.

[9] B. Green and T. Tao, *New bounds for Szemerédi's theorem. II. A new bound for $r_4(N)$*, Analytic number theory, Cambridge Univ. Press, Cambridge, 2009, pp. 180–204.

[10] B. Green and T. Tao, *An arithmetic regularity lemma, an associated counting lemma, and applications*, An irregular mind, Bolyai Soc. Math. Stud., vol. 21, János Bolyai Math. Soc., Budapest, 2010, pp. 261–334.

[11] B. Green and T. Tao, *Yet another proof of Szemerédi's theorem*, An irregular mind, Bolyai Soc. Math. Stud., vol. 21, János Bolyai Math. Soc., Budapest, 2010, pp. 335–342.

[12] B. Green and T. Tao, *The quantitative behaviour of polynomial orbits on nilmanifolds*, Ann. of Math. (2) **175** (2012), 465–540.

[13] B. Green and T. Tao, *New bounds for Szemerédi's theorem, III: a polylogarithmic bound for $r_4(N)$*, Mathematika **63** (2017), 944–1040.

[14] D. R. Heath-Brown, *Integer sets containing no arithmetic progressions*, J. London Math. Soc. (2) **35** (1987), 385–394.

[15] Z. Kelley and R. Meka, *Strong bounds for 3-progressions*, arXiv:2302.05537.

[16] J. Leng, *Efficient Equidistribution of Periodic Nilsequences and Applications*, arXiv:2306.13820.

[17] J. Leng, A. Sah, and M. Sawhney, *Improved bounds for five-term arithmetic progressions*, arXiv:2312.10776.

[18] J. Leng, A. Sah, and M. Sawhney, *Quasipolynomial bounds for the inverse theorem for the Gowers $U^{s+1}[N]$-norm*, manuscript.

[19] S. Peluse and S. Prendiville, *A polylogarithmic bound in the nonlinear Roth theorem*, Int. Math. Res. Not. IMRN (2022), 5658–5684.

[20] K. F. Roth, *On certain sets of integers. II*, J. London Math. Soc. **29** (1954), 20–26.

[21] T. Sanders, *On Roth's theorem on progressions*, Ann. of Math. (2) **174** (2011), 619–636.

[22] T. Sanders, *On certain other sets of integers*, J. Anal. Math. **116** (2012), 53–82.

[23] W. M. Schmidt, *Small fractional parts of polynomials*, Regional Conference Series in Mathematics, No. 32, American Mathematical Society, Providence, RI, 1977.

[24] E. Szemerédi, *On sets of integers containing no four elements in arithmetic progression*, Number Theory (Colloq., János Bolyai Math. Soc., Debrecen, 1968), Colloq. Math. Soc. János Bolyai, vol. 2, North-Holland, Amsterdam-London, 1970, pp. 197–204.

[25] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245.

[26] E. Szemerédi, *Integer sets containing no arithmetic progressions*, Acta Math. Hungar. **56** (1990), 155–158.

Department of Mathematics, UCLA, Los Angeles, CA 90095, USA
Email address: `jamesleng@math.ucla.edu`

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
Email address: `{asah,msawhney}@mit.edu`