

NILPROGRESSIONS AND GROUPS WITH MODERATE GROWTH

EMMANUEL BREUILLARD AND MATTHEW C. H. TOINTON

ABSTRACT. We show that doubling at some large scale in a Cayley graph implies uniform doubling at all subsequent scales. The proof is based on the structure theorem for approximate subgroups proved by Green, Tao and the first author. We also give a number of applications to the geometry and spectrum of finite Cayley graphs. For example, we show that a finite group has moderate growth in the sense of Diaconis and Saloff-Coste if and only if its diameter is larger than a fixed power of the cardinality of the group. We call such groups almost flat and show that they have a subgroup of bounded index admitting a cyclic quotient of comparable diameter. We also give bounds on the Cheeger constant, first eigenvalue of the Laplacian, and mixing time. This can be seen as a finite-group version of Gromov's theorem on groups with polynomial growth. It also improves on a result of Lackenby regarding property (τ) in towers of coverings. Another consequence is a universal upper bound on the diameter of all finite simple groups, independent of the CFSG.

CONTENTS

1. Introduction	1
2. Approximate groups	7
3. Nilprogressions	11
4. A Gromov-type theorem for finite groups	19
5. Cheeger constant and spectral gap	28
6. Moderate growth and mixing times	31
References	36

1. INTRODUCTION

Let G be a group generated by a finite, symmetric subset S . Here, and throughout this paper, by writing that S is *symmetric* we mean that if s belongs to S then so does s^{-1} , and that S contains the identity. When speaking of the growth of G with respect to S , we refer to the behaviour of the sequence of cardinalities

$$|S|, |S^2|, |S^3|, \dots,$$

where we denote by S^n the n -fold product set $\{s_1 \cdot \dots \cdot s_n; s_i \in S\}$. This is the ball of radius n in the Cayley graph of G relative to S . In the event that G is finite, we also consider the *diameter* $\text{diam}_S(G)$ of G with respect to S , which is defined to be the minimum n such that $S^n = G$.

According to Gromov's polynomial growth theorem [23], if the sequence $\{|S^n|\}$ is bounded above by a polynomial function of n , then G has a finite-index nilpotent subgroup. In this paper we are concerned

Key words and phrases. polynomial growth, volume doubling, diameter bounds, finite groups, approximate groups, moderate growth, mixing time.

with some refinements of this theorem, mostly in the context of finite groups. In particular, we study the relations between the diameter and the cardinality of a finite group on the one hand, and various interesting invariants such as the Cheeger constant, the first non-zero eigenvalue of the Laplace operator, and the mixing time of the associated Cayley graph on the other.

Our two main results are Theorems 1.1 and 1.3 below. The first concerns the doubling property at one given scale in an arbitrary Cayley graph.

Theorem 1.1 (Doubling at some scale implies doubling at all larger scales). *For every $K \geq 1$ there exist $n_0 = n_0(K) \in \mathbb{N}$ and $\theta(K) \geq 1$, such that if S is a finite symmetric set inside some group, and if there exists $n \geq n_0$ for which*

$$|S^{2n+1}| \leq K|S^n|, \tag{1.1}$$

then for every $m \geq n$ and every $c \in \mathbb{N}$ we have $|S^{cm}| \leq \theta(K)^c |S^m|$.

We stress that the constants $n_0(K), \theta(K)$ depend only on K , and not on the group or on S . Most of our results in this paper are uniform in the generating set S and do not assume that S is bounded. Unfortunately the method does not allow for an effective computation of the constants $n_0(K)$ and $\theta(K)$ in terms of K .

The doubling property played an important role in Gromov's original proof of his polynomial-growth theorem. Indeed, the first step of his proof consisted of observing that the polynomial growth condition on $|S^n|$ implies that a doubling condition such as (1.1) holds for infinitely many n and some uniform K . Green, Tao and the first author established in [10] a structure theorem for doubling sets (or equivalently approximate subgroups; see Section 2) in arbitrary groups, allowing them to extend Gromov's theorem in a number of ways. The proof of Theorem 1.1 makes crucial use of this structure theorem via an analysis of the behaviour of large powers of nilprogressions (see Section 3, below, for the definition of nilprogressions).

Remark. Tao [47] has independently used a somewhat similar argument to classify fairly explicitly the possible behaviours of $|S^m|$ as $m \rightarrow \infty$ under the hypothesis (1.1), in particular recovering Theorem 1.1.

The motivation for Theorem 1.1 came from work of Benjamini, Finucane and Tessera [4], in which the authors identify the scaling limit of a sequence of vertex-transitive graphs of large diameter (i.e. such that (1.2) below holds). Using Theorem 1.1 they show¹ that for a given $\varepsilon > 0$, the family of all such Cayley graphs (rescaled so that edges have length $1/\text{diam}_S(G)$) is relatively compact for the Gromov-Hausdorff topology on the space of compact metric spaces. Indeed a doubling condition such as that of Theorem 1.1 is exactly what is needed to apply Gromov's relative compactness criterion [24]. Furthermore, they identify the possible limits as the flat tori $\mathbb{R}^d/\mathbb{Z}^d$, $d = d(\varepsilon)$, endowed with a translation invariant Finsler metric.

This prompts the following definition.

Definition 1.2. Let $\varepsilon > 0$. A Cayley graph of a finite group G with symmetric generating set S is called ε -almost flat if

$$\text{diam}_S(G) \geq \left(\frac{|G|}{|S|} \right)^\varepsilon. \tag{1.2}$$

¹Their original argument required only the main result of [10], not Theorem 1.1, but needed to assume a bound on $|S|$.

See [4] and the survey [7] for more details on the result of Benjamini, Finucane and Tessera. We will give further applications of Theorem 1.1 below.

Remark. Abelian groups and more generally nilpotent groups with a bound on the nilpotency class and number of generators are ε -almost flat for some positive ε depending only on these bounds. See Proposition 4.7, below. We also show in Proposition 4.15 that the notion of an almost flat group is independent of the choice of generating set, provided the size of the generating set remains bounded.

Our second main result concerns the algebraic structure of almost flat finite groups, and runs as follows.

Theorem 1.3 (Almost flat groups have large virtually abelian quotients). *Let $\varepsilon > 0$ and $\beta > \frac{1}{2}$. Every ε -almost flat Cayley graph of a finite group G with generating set S has a normal subgroup H contained in $S^{O_{\varepsilon,\beta}(\gamma^\beta)}$ such that G/H has an abelian subgroup whose rank and index are bounded in terms of ε and β only.*

The implied constants depend only on ε and β , and not on the group G or the size of the generating set S . We refer the reader to Theorem 4.1 below for two variants of Theorem 1.3. In particular, we show there that G must have a subgroup of bounded index with a cyclic quotient of comparable diameter.

We note that Gromov had already proved a finite version of his polynomial growth theorem (see at the end of [23]). The above can be seen as a refinement, in which a weaker assumption is made (a diameter condition instead of a volume growth condition) and a somewhat stronger conclusion is derived.

It is legitimate to ask whether the conclusion of Theorem 1.3 (or those of Theorem 4.1 below) continues to hold under a weaker assumption on the diameter. The answer is no. In fact, given any function $f : \mathbb{N} \rightarrow \mathbb{R}_+$ going to zero at infinity, there is a sequence of pairwise non-isomorphic finite groups G_n with uniformly bounded generating sets with respect to which their diameters are at least $|G_n|^{f(|G_n|)}$, while G_n has no proper subgroup of index at most n and no non-trivial abelian quotient. This example also provides a counter-example to a non-commutative version of the *polynomial Freiman-Ruzsa* conjecture. See Subsection 4.1 below.

We now present some applications of Theorems 1.1 and 1.3.

1.1. Diameter of finite simple groups. The following is an immediate consequence of Theorem 1.3.

Corollary 1.4 (Diameter bound for finite simple groups). *Let $\varepsilon > 0$. Then there is a constant C_ε depending only on ε such that every non-abelian finite simple group G with symmetric generating set S satisfies*

$$\text{diam}_S(G) \leq \max \left\{ \left(\frac{|G|}{|S|} \right)^\varepsilon, C_\varepsilon \right\}.$$

Proof. Write $\gamma = \text{diam}_S(G)$. Since G is simple, if $\gamma > (|G|/|S|)^\varepsilon$ then Theorem 1.3 implies either that $\gamma \leq O_\varepsilon(\gamma^{3/4})$, in which case $\gamma \leq O_\varepsilon(1)$, or that G has a normal abelian subgroup of index at most $O_\varepsilon(1)$. This normal subgroup is either all of G , in which case G is abelian, or it is trivial, in which case $|G| \leq O_\varepsilon(1)$, and so certainly $\gamma \leq O_\varepsilon(1)$. \square

This bound should be compared with the conjecture of Babai (see [2]) according to which there ought to exist universal constants $C, D > 0$ such that $\text{diam}_S(G) \leq D(\log |G|)^C$ for all finite simple groups G and all S . Such a strong bound is currently only known for finite simple groups of bounded rank (see [9, 27, 42]). In fact for these groups it is even conjectured that one can take $C = 1$ with a multiplicative

constant D that is allowed to depend on the rank [6, Conjecture 4.5]. For alternating groups A_n , weaker bounds that are still much better than that of Corollary 1.4, are known (see [2, 28]).

Corollary 1.4 is the first non-trivial bound towards Babai's conjecture valid for all finite simple groups as far as we can tell. However, even if it is not straightforward, it is likely that a better bound also valid for all finite simple groups is within reach by current methods using the classification of finite simple groups and the above mentioned results. Let us note however that a remarkable feature of the proof of Corollary 1.4 is that it does not use the classification of finite simple groups.

1.2. Cheeger constant and spectral gap. It is well known that the diameter of a Cayley graph is related to its Cheeger constant h and to the first eigenvalue λ_1 of the Laplace operator in such a way that

$$\frac{1}{8(\text{diam}_S(G))^2} \leq \frac{1}{2}h^2 \leq \lambda_1 \leq 2|S|h \leq \frac{8|S|^2 \log |G|}{\text{diam}_S(G)} \quad (1.3)$$

See Section 5 for a definition of these quantities and a proof of (1.3).

Corollary 1.5 (Groups with very small Cheeger constant or spectral gap). *Let G be a finite group with finite symmetric generating set S . Let $\varepsilon > 0$. Assume that the Cayley graph of G with respect to S satisfies one of the inequalities*

$$\lambda_1 \leq \min \left\{ 2^{-9}, \left(\frac{|S|}{|G|} \right)^\varepsilon \right\}, \quad h \leq \min \left\{ 2^{-2}, \left(\frac{|S|}{|G|} \right)^\varepsilon \right\}.$$

Then G is $\varepsilon/3$ -almost flat. In particular, the conclusion of Theorem 1.3 holds.

This is useful in a variety of contexts, for example in the work of Ellenberg, Hall and Kowalski [18] in arithmetic geometry, where they used a weak lower bound on λ_1 for certain perfect groups due to Pyber and Szabó [42]. The bound from Corollary 1.5 would work there equally well. Note that, conversely, it is easy to verify that if a finite group G has subgroup of bounded index with a cyclic quotient of comparable diameter then its λ_1 is very small, indeed of order $O_{|S|}(1)/(\text{diam}_S(G))^2$.

Corollary 1.6 (Cheeger constant and spectral gap for almost flat groups). *Given $\varepsilon > 0$ there are $C_1, C_2 > 0$ such that if G is a finite group with ε -almost flat Cayley graph, then its first eigenvalue λ_1 and Cheeger constant h satisfy*

$$\frac{1}{8(\text{diam}_S(G))^2} \leq \frac{1}{2}h^2 \leq \lambda_1 \leq C_1|S|h^2 \leq \frac{C_2|S|^2}{(\text{diam}_S(G))^2}. \quad (1.4)$$

In fact, there is $C_3 > 0$ such that

$$\lambda_1 \leq \frac{C_3|S|}{(\text{diam}_S(G))^2}. \quad (1.5)$$

In other words the Buser inequality holds in its strong form on almost flat groups, which is to say that λ_1 is comparable to h^2 . This follows from a simple argument using the doubling property at a scale comparable to the diameter. See Lemma 5.2. It is known by work of Klartag et al. [30] that for abelian G we may take $C_1 = 16$, which is essentially sharp.

1.3. Property (τ) and towers of coverings. A finitely generated group Γ is said to have property (τ) with respect to a sequence of finite-index normal subgroups $\{\Gamma_n\}_n$ if the Cheeger constant $h(\Gamma/\Gamma_n)$ of the Cayley graph of the finite group Γ/Γ_n with respect to a fixed generating set of Γ is bounded away from zero, independently of n . If Γ has property (τ) with respect to the family of all its finite-index normal subgroups then one simply says that Γ has property (τ) . This property is independent of the choice of generating set and stable under passing to a finite-index subgroup. Property (τ) was introduced and studied by Lubotzky; see [5, 37, 38] for background.

It is clear that a finitely generated group that admits a finite-index subgroup with infinite abelianisation cannot have property (τ) , as it will have large virtually cyclic quotients. The converse is not true, because of the existence of amenable finitely generated infinite torsion groups, such as the Grigorchuk group (see e.g. [38]). However, it is not known whether or not the converse holds in the category of finitely presented groups.

In connection with Thurston's virtual first Betti number conjecture regarding the topology of 3-manifolds, Lackenby [31, 32, 33] studied various extra conditions of an algebraic or topological nature on a finitely presented group without property (τ) that force the existence of a finite-index subgroup with infinite abelianisation. See also [34]. In particular, he showed that if the Cheeger constant $h(\Gamma/\Gamma_n)$ decays faster than the square root of $|\Gamma/\Gamma_n|$, then Γ has a finite-index subgroup with infinite abelianisation [33, Theorem 1.1]. As a consequence of Theorem 1.3, we can remove the finite presentability assumption and strengthen Lackenby's result as follows.

Corollary 1.7 (Strong failure of (τ) implies positive virtual Betti number – group form). *Let Γ be a finitely generated group, and $\{\Gamma_n\}_n$ an infinite family of finite-index normal subgroups such that $h(\Gamma/\Gamma_n)[\Gamma : \Gamma_n]^\varepsilon$ is bounded independently of n , for some fixed $\varepsilon > 0$. Then Γ has a finite-index subgroup with a surjective homomorphism onto \mathbb{Z} .*

Using the well-known transfer principle between spectral properties of finite covers of Riemannian manifolds and those of the Cayley graphs of the associated groups of deck transformations (see [12], [39], [5, Appendix]), one has the following geometric consequence of Corollary 1.7.

Corollary 1.8 (Strong failure of (τ) implies positive virtual Betti number – geometric form). *Let M be a compact Riemannian manifold and M_n a sequence of finite Galois covers of M of degree d_n . Assume that there is $\varepsilon > 0$ such that $h(M_n)(d_n)^\varepsilon$ is a bounded sequence, where $h(M_n)$ denotes the Cheeger constant of M_n . Then M has a finite cover with positive first Betti number.*

Note that we could have stated these corollaries with λ_1 in place of h , because of the Cheeger-Buser inequalities (1.3). We take this opportunity to raise the following question.

Question: Can the assumption that the covers are Galois in Corollary 1.8, or the assumption that Γ_n is normal in Corollary 1.7, be removed? Suppose even that the (not necessarily normal) finite-index subgroups Γ_n are such that the diameter of the Schreier graph of Γ/Γ_n is at least $[\Gamma, \Gamma_n]^\varepsilon$, for some fixed $\varepsilon > 0$. Does this imply that Γ virtually maps onto \mathbb{Z} ?

1.4. Groups with moderate growth and mixing times. Given positive constants A and d , Diaconis and Saloff-Coste [15] say that the finite group G has (A, d) -moderate growth with respect to a finite generating set S if, writing $\gamma := \text{diam}_S(G)$ for the diameter of G with respect to S , we have

$$|S^n| \geq \frac{1}{A} \left(\frac{n}{\gamma}\right)^d |G| \tag{1.6}$$

whenever $1 \leq n \leq \gamma$. It is clear from the definition, taking $n = 1$, that (G, S) is $\frac{1}{2d}$ -almost flat, provided $|G|/|S|$ is large enough (in fact this holds as soon as the diameter $\text{diam}_S(G)$ is larger than some constant depending only on A and d). In particular, by Theorem 1.3 such a group has a large virtually cyclic quotient. Using Theorem 1.1 we can prove that, conversely, every ε -almost flat group is of moderate growth.

Corollary 1.9 (Groups with moderate growth versus almost flat groups). *A group has moderate growth if and only if it is almost flat. More precisely, given $A, d > 0$ there is $C(A, d) > 0$ such that every finite group with (A, d) -moderate growth and diameter at least $C(A, d)$ is $\frac{1}{2d}$ -almost flat. Conversely, given $\varepsilon > 0$ there are $A, d > 0$ depending only on ε , such that every ε -almost flat group has (A, d) -moderate growth.*

The observation is that the almost flat condition implies the volume doubling condition at some relatively small scale (at most $(\text{diam}_S(G))^{o(1)}$; see Lemma 4.9, below), and hence at all subsequent scales by Theorem 1.1. Thus almost flat groups are doubling at all scales larger than $(\text{diam}_S(G))^{o(1)}$. This is enough to guarantee moderate growth.

Groups with moderate growth provide examples of Cayley graphs for which the mixing time of the simple random walk behaves quadratically in the diameter. This is the main result proved by Diaconis and Saloff-Coste in [15], which now translates as follows. We write μ_G for the uniform probability measure on G , μ_S for the uniform probability measure on S , and $\mu_S^{(n)}$ for the n -th convolution power of μ_S .

Corollary 1.10 (Mixing time). *The relaxation time and the mixing times of almost flat Cayley graphs are comparable to $(\text{diam}_S(G))^2$. More precisely, given $\varepsilon > 0$, there are positive constants $B, C, D > 0$, depending on ε only, such that whenever G is a finite group whose Cayley graph (G, S) is ε -almost flat and satisfies $\gamma := \text{diam}_S(G) \geq D$, for all $n \geq 1$ and all $p \in [1, \infty]$ we have*

$$\frac{1}{2} e^{-C \frac{n}{\gamma^2}} \|\mu_G\|_p \leq \|\mu_S^{(n)} - \mu_G\|_p \leq B e^{-\frac{n}{8|S|\gamma^2}} \|\mu_G\|_p,$$

where $\|\cdot\|_p$ denotes the L^p -norm on G . Indeed, the upper bound holds even without the assumption that $\gamma \geq D$.

We stress that the bounds for C and B obtained by Diaconis and Saloff-Coste are explicit when expressed in terms of the moderate growth parameters A and d . In the above reformulation in terms of ε -almost flatness, we lose the explicit dependence, because the moderate growth parameters A and d given by the second part of Corollary 1.9 depend on ε in a *non-explicit* manner (due to our reliance on the non-explicit Theorem 2.3 below).

The mixing times (for each L^p norm) are at least half the diameter, but for a general Cayley graph of bounded degree, they are always in $O(\gamma^3)$, where γ is the diameter, as easily follows from the lower bound on λ_1 from (1.3). This bound is attained for example for the uniform mixing time on the lamplighter group with cyclic base (e.g. see [40]). At the opposite extreme, if the Cayley graph is an expander graph (uniform lower bound on λ_1), then the mixing times are comparable to the diameter and a cut-off phenomenon is expected (and proved in some cases, see for that matter [36]). For almost flat Cayley graphs, no cut-off phenomenon arises.

We conclude with the following refinement of the Diaconis–Saloff-Coste result showing that doubling at one small enough scale is sufficient to guarantee quadratic mixing time.

Theorem 1.11 (Quadratic mixing under a doubling condition). *Given $K \geq 1$, there is $n_1(K)$ such that the following holds. Consider the Cayley graph \mathcal{G} of a finite group G with generating set S and diameter*

γ . Assume that $|S^{2n+1}| \leq K|S^n|$ for some $n \geq n_1(K)$ with $n \leq \gamma^{2/3}$. Then the mixing times of the simple random walk on \mathcal{G} are comparable to γ^2 up to multiplicative constants depending on K and $|S|$ only.

We provide an example showing that the exponent $\frac{2}{3}$ is sharp; see Example 6.5.

The paper is organised as follows. In Section 2, we discuss approximate groups and recall the main result of [10], which describes the structure of approximate subgroups of arbitrary groups. We also prove Theorem 1.1 modulo a result on nilprogressions proved in Section 3. Section 3 is devoted to nilprogressions. We give a summary of several competing definitions and explain the relationship between them. Then we prove that their powers have uniform doubling, thus completing the proof of Theorem 1.1. Section 4 is devoted to Theorem 1.3 and several extra structural results on almost flat groups. We also describe there an example pertaining to a non-commutative version of the Freiman-Ruzsa conjecture and showing that super-polynomial bounds are required in [10]. In Section 5 we discuss bounds on the Cheeger constant and λ_1 for almost flat Cayley graphs and prove Corollaries 1.7 and 1.8 regarding property (τ) . Finally Section 6 is devoted to random walks and mixing times. We establish Corollary 1.10 and Theorem 1.11.

Acknowledgements. The authors are grateful to David Simmons, Romain Tessera and Tianyi Zheng for helpful conversations. We are also grateful to Persi Diaconis and Terry Tao for their comments on the paper. Both authors are supported by ERC grant GA617129 ‘GeTeMo’. M.T. is on leave from a Junior Research Fellowship at Homerton College, Cambridge, where some of this work was carried out.

2. APPROXIMATE GROUPS

In recent years there has been much effort to describe the structure of certain finite subsets of groups called *approximate subgroups*. The reader may consult the surveys [11, 20, 44] for details on the background to this field and its applications. The definition of an approximate group is originally due to Tao [45], and runs as follows.

Definition 2.1 (Approximate group). Let G be a group and let $K \geq 1$. A finite subset $A \subset G$ is said to be a K -*approximate subgroup* of G , or just a K -*approximate group*, if it is symmetric and if there is a subset X with $|X| \leq K$ such that $A^2 \subset XA$.

Here, and throughout the paper, for subsets X and Y of a group we denote by XY the set $\{xy : x \in X, y \in Y\}$ and by X^{-1} the set $\{x^{-1} : x \in X\}$.

It may not be immediately obvious that approximate groups should have anything to do with growth of groups. However, the following simple argument gives a reason to expect that the two should be related, at least on some level.

Lemma 2.2. *Let G be a group, and let $S \subset G$ be a finite symmetric set. Suppose that*

$$|S^{5n}| \leq K|S^n|, \tag{2.1}$$

for some integer n , then S^{2n} is a K -approximate subgroup of G .

Proof. See Ruzsa’s covering lemma [48]. Upon replacing S by S^n we may assume that $n = 1$. Let X be a subset of S^4 that is maximal with respect to the condition that the sets xS with $x \in X$ are disjoint, and note that this condition implies that $|XS| = |X||S|$. Given (2.1), and since $XS \subset S^5$, this implies in particular that $|X| \leq K$. It is therefore sufficient to show that $S^4 \subset XS^2$. This holds, because if $z \in S^4$ then, by the maximality of X , there are $x \in X$ and $s_1, s_2 \in S$ such that $zs_1 = xs_2$. The symmetry of S then implies that $z = xs_2s_1^{-1} \in XS^2$, as claimed. \square

Recently, Green, Tao and the first author [10] proved the following structure theorem for approximate subgroups of arbitrary groups.

Theorem 2.3 (Structure of approximate groups [10, Theorem 1.6]). *Let A be a K -approximate subgroup of a group G . Then there is a finite subgroup $H \subset A^4$ and a subgroup Γ containing H as a normal subgroup such that Γ/H is nilpotent of rank and class at most $O_K(1)$, and such that A is contained in fewer than $O_K(1)$ left-cosets of Γ .*

As usual $O_K(1)$ denotes a constant that depends on K only. Note that since the group Γ is finite-by-nilpotent it is also virtually nilpotent. As remarked in [10], this result easily implies Gromov's original theorem on groups with polynomial growth. More precisely one has the following somewhat stronger statement.

Corollary 2.4 (One-scale Gromov's theorem). *For every $K \geq 1$ there exists $n_0 = n_0(K) \in \mathbb{N}$ such that if G is a group generated by a finite symmetric generating set S such that $|S^{5n}| \leq K|S^n|$ for some $n \geq n_0$ then G is virtually nilpotent.*

Remark 2.5. Using Lemma 2.10 below, one sees that the hypothesis $|S^{5n}| \leq K|S^n|$ can be weakened into $|S^{2n+1}| \leq K|S^n|$. It is an interesting problem to determine whether one can push this further and assume only that $|S^{\alpha n}| \leq K|S^n|$ for some fixed $\alpha > 1$.

Remark 2.6. If we were to allow the uniform doubling constant in the conclusion of Theorem 1.1 to depend on the generating set S then that theorem would become a straightforward application of Corollary 2.4, as it is well known that large balls in Cayley graphs of finitely generated virtually nilpotent groups are doubling (see [3], for example). However, the independence from S is key to our applications.

We recall the easily seen fact that if a group has polynomial growth, which is to say that $|S^n| = O(n^D)$ for some $D > 0$, then there are infinitely many n for which $|S^{5n}| < 6^D|S^n|$, and the hypotheses of Corollary 2.4 are satisfied at infinitely many scales. By contrast, the corollary asserts that only one large scale is necessary. For the reader's convenience, and also for further use below, we now recall how to derive this corollary from Theorem 2.3, as was done in [10, Corollary 11.2].

Theorem 2.3 shows that an approximate group lies in finitely many cosets of a nilpotent group. However, Corollary 2.4 requires that the *entire* group lie in finitely many cosets of a nilpotent group. The key to overcoming this apparent difficulty is to observe that if the n -ball in a Cayley graph is contained in at most n cosets of a subgroup, then this subgroup has index at most n , as follows.

Lemma 2.7. *Let G be a group generated by a finite symmetric set S containing the identity, and let Γ be a subgroup of G . Suppose that for some $k \in \mathbb{N}$ we have S^k contained in at most k left-cosets of Γ . Then $G = S^k\Gamma$, and in particular $[G : \Gamma] \leq k$.*

Proof. Since $S\Gamma \subset S^2\Gamma \subset S^3\Gamma \subset \dots$, the number of left-cosets of Γ having non-empty intersection with $S^r\Gamma$ is non-decreasing in r . Thus, if S^k is contained in at most k left-cosets of Γ , then $S^{r+1}\Gamma = S^r\Gamma$ for some $r \leq k$. It follows by induction that $G = S^r\Gamma$, and hence that $G = S^k\Gamma$. \square

Proof of Corollary 2.4. By Lemma 2.2, S^{2n} is a K -approximate subgroup of G . By Theorem 2.3, there exists $n_0 = n_0(K)$ such that S^{2n} is contained in at most n_0 cosets of a virtually nilpotent subgroup $\Gamma \leq G$. If $n \geq n_0$ then Lemma 2.7 therefore implies that Γ is of finite index in G , and hence that G is virtually nilpotent. \square

We prove Theorem 1.1 by first reducing it to the nilpotent case, which we state as follows.

Theorem 2.8 (Nilpotent case). *Let $K \geq 1$ and $s \geq 1$. Suppose that S is K -approximate subgroup in some s -step nilpotent group G . Then there is constant $C = C(K, s) \geq 1$ such that for every $m \geq 1$, S^m is a C -approximate subgroup.*

In the remainder of this section, we explain how to perform such a reduction using Theorem 2.3. The proof of the nilpotent case will occupy the next section and rely on a study of nilprogressions.

Remark. Contrary to the general case of Theorem 1.1, where no effective estimate is known on the size of $\theta(K)$, an effective bound in the nilpotent case can be easily computed by following our argument.

The reduction of Theorem 1.1 to Theorem 2.8 is based on the following proposition, which makes use of the structure theorem for approximate groups, Theorem 2.3.

Proposition 2.9. *For every $K \geq 1$ there exists $n_0 = n_0(K)$ such that if G is a group generated by a finite symmetric generating set S such that*

$$|S^{5n}| \leq K|S^n| \tag{2.2}$$

for some $n \geq n_0$, then there are subgroups $H \leq \Gamma \leq G$ such that $H \subset S^{8n}$ is normal in G , such that Γ/H is nilpotent of rank and class at most $O_K(1)$, and such that Γ has index at most n_0 in G . Moreover, $B := S^{8n} \cap \Gamma$ is an $O_K(1)$ -approximate subgroup containing H , and there is a subset $X \subset S^n$ of size at most n_0 such that for every $b \in \mathbb{N}$ we have

$$S^{bn} \subset XB^{b-1}. \tag{2.3}$$

Proof. By Lemma 2.2, S^{2n} is a K -approximate subgroup of G . We may therefore apply Theorem 2.3 to S^{2n} and conclude that there is a natural number $n_1 = n_1(K)$, a subgroup $H_0 \subset S^{8n}$, and a subgroup $\Gamma_0 \leq G$ containing H_0 as a normal subgroup, such that Γ_0/H_0 is a nilpotent group of rank and step at most $O_K(1)$, and such that S^{2n} is contained in at most n_1 cosets of Γ_0 . If $n \geq n_1$ then Lemma 2.7 implies that Γ_0 has index at most n_1 in G . In particular, Γ_0 has a normal subgroup Γ of index at most $n_0 := n_1!$.

Since Γ is normal in G , so is $C^s(\Gamma)$, the s -th term in its descending central series, where s is the nilpotency class of Γ_0/H_0 . Setting $H := C^s(\Gamma)$, we see that Γ/H is s -step nilpotent and that $H \subset H_0 \subset S^{8n}$ is normal in G . It is well known that if F is an approximate subgroup then so is the intersection of F^2 with any subgroup (see [49, Lemma 2.10], for example). This implies in particular that $B := S^{8n} \cap \Gamma$ is an $O_K(1)$ -approximate subgroup.

It remains to exhibit a set $X \subset S^n$ satisfying (2.3). If $n \geq n_0$ then S^n intersects every coset of Γ in G , and so S^n contains a complete set of coset representatives of Γ ; we claim that it suffices to take X to be this set. Indeed, by definition this set satisfies $S^{2n} \subset X\Gamma$, and hence $S^{2n} \subset X(S^{8n} \cap \Gamma) = XB$. In particular $S^{2n} \subset S^nB$, and so $S^{bn} \subset S^{2n}B^{b-2}$ for each $b \in \mathbb{N}$ by induction. The result follows. \square

The following combinatorial lemma allows us to pass from a small doubling assumption to the more comfortable $|S^{5n}| \leq K|S^n|$.

Lemma 2.10. *For every $K \geq 2$, there is $C(K) \geq 2$ such that if S is a finite symmetric subset (containing 1) in a group G , and if $n \geq 1$ is an integer such that*

$$|S^{2n+1}| \leq K|S^n|,$$

then $|S^{5n}| \leq C(K)|S^n|$.

Proof. We apply Petridis's lemma [41] (see also [43]), according to which if A and B are finite subsets of a group G such that $|BA| \leq K|A|$, then there is a subset $A_0 \subset A$ that $|BA_0X| \leq K|A_0X|$ for every finite subset X in G . Applying this lemma with $B = S^{n+1}$ and $A = S^n$ gives a subset A_0 of S^n with the property that

$$|S^{n+1}A_0X| \leq K|A_0X|$$

for all $X \subset G$. In particular, this implies that $|S^{n+1}| \leq |S^{n+1}A_0| \leq K|A_0|$, and hence that $|S^{2n+1}| \leq K|S^n| \leq K^2|A_0|$. This also shows that $|A_0SA_0| \leq K|A_0|$, since $A_0S \subset S^{n+1}$. Finally, taking $X = A_0$ and noting that $A_0 \subset S^n \subset S^{n+1}$, we see that $|A_0^3| \leq K|A_0^2| \leq K^2|A_0|$.

Now set $A_1 := S \cup A_0$. Note that A_1^3 is contained in the union of the subsets $A_0^3, S^3, A_0SA_0, S^2A_0, A_0S^2, SA_0S, A_0^2S, SA_0^2$, of which all but A_0^3 are contained in S^{2n+1} . Since both A_0^3 and S^{2n+1} are of size at most $K^2|A_0|$, this implies that $|A_1^3| \leq 8K^2|A_1|$, and so A_1 has small tripling. It is well known that small tripling implies small k -pling for every integer k ; more precisely, [45, Lemma 3.4], for example, implies that $|(A_1 \cup A_1^{-1})^k| \leq O_{K,k}(1)|A_1|$.

We claim that there exists $k = O_K(1)$ such that S^n is contained in $(A_1 \cup A_1^{-1})^k$. This will finish the proof of the lemma, since it implies that $S^{5n} \subset (A_1 \cup A_1^{-1})^{5k}$, and hence that $|S^{5n}| \leq O_{K,k}(1)|A_1| \leq O_K(1)|S^n|$.

We build recursively sequences $x_i \in S^n$ and $k_i \in \mathbb{Z}$ for $i \geq 0$. Take x_0 to be the identity, and k_0 to be zero. Then, assuming x_0, \dots, x_i and k_1, \dots, k_i have been chosen, let k_{i+1} be the smallest integer such that there is $y \in S^{k_{i+1}}$ with the property that yA_1 is disjoint from all x_jA_1 for $j = 0, \dots, i$, and let x_{i+1} be such an element y .

Let m be the largest integer such that $k_m \leq n$ but $k_{m+1} > n$. Since the x_iA_1 are all disjoint and belong to $S^nA_1 \subset S^{2n}$, there can be at most $|S^{2n}|/|A_1| \leq K^2$ of them, and so $m \leq K^2$.

On the other hand, we claim that $x_i \in (A_1 \cup A_1^{-1})^{3i}$. We verify this by induction. Assuming this holds for i , we write $x_{i+1} = su$, where $u \in S^{k_{i+1}-1}$ and $s \in S$. By construction, uA_1 must intersect some x_jA_1 non-trivially for some $j \leq i$. This means that $u \in x_jA_1A_1^{-1} \subset (A_1 \cup A_1^{-1})^{3i+2}$. However, $S \subset A_1$, and hence $x_{i+1} \in (A_1 \cup A_1^{-1})^{3i+3}$, as desired.

If $x \in S^n$ is arbitrary then, by the definition of x_m , there must be some $j \leq m$ such that xA_1 intersects x_jA_1 non-trivially. This means that $x \in x_jA_1A_1^{-1}$, and hence that $x \in (A_1 \cup A_1^{-1})^{3m+2}$. We conclude that $S^n \subset (A_1 \cup A_1^{-1})^{3m+2}$. The claim, and hence the lemma, now follows, since $m \leq K^2$. \square

Lemma 2.11. *Let $\pi : G \rightarrow G'$ be a homomorphism with kernel N , and suppose that A is a symmetric subset of G such that $N \subset A$ and such that $\pi(A)$ is a K -approximate group. Then A^2 is a K^3 -approximate group.*

Proof. To say that $\pi(A)$ is a K -approximate group is to say that there exists $X \subset G$ with $|X| \leq K$ such that $A^2 \subset XAN$. Iterating this, we conclude that $A^4 \subset X^3AN \subset X^3A^2$. \square

We can now complete the reduction of Theorem 1.1 to the nilpotent case.

Proof of Theorem 1.1 modulo Theorem 2.8. Applying Lemma 2.10, we may assume that $|S^{5n}| \leq K|S^n|$. Proposition 2.9 therefore gives a finite-index subgroup Γ of G and a normal subgroup H contained in S^{8n} such that Γ/H is nilpotent with bounded rank and nilpotency class, such that $S^{bn} \subset XB^b$ for every $b \geq 1$, where $|X| \leq n_0(K)$ and $B := S^{8n} \cap \Gamma$ is an $O_K(1)$ -approximate subgroup of Γ containing H . Applying Theorem 2.8 to the nilpotent group Γ/H , and pulling back to Γ using Lemma 2.11, we conclude that for all $b \in \mathbb{N}$ the set B^{2b} is an $O_K(1)$ -approximate group. In particular, $S^{32bn} \subset XB^{32b} \subset XY_bB^{2b}$, for some subset Y_b of size at most $O_K(1)$. Since $B^{2b} \subset S^{16bn}$, we conclude that S^{16bn} is an $O_K(1)$ -approximate subgroup, uniformly for all $b \in \mathbb{N}$. Write X_{16bn} for the set XY_b witnessing this.

Given $m \geq 16n$, let m' be the smallest integer multiple of $32n$ that is at least m , and note that $m \geq m'/2$, and hence that $|S^{m'}| \leq O_K(1)|S^{m'/2}| \leq O_K(1)|S^m|$. We also have

$$S^{cm} \subset S^{cm'} \subset X_{m'}^{c-1} S^{m'},$$

from which the theorem then follows easily for the value of m in question. For $m < 16n$, on the other hand, the theorem follows from Lemma 2.2. \square

3. NILPROGRESSIONS

Our goal in this section is to prove Theorem 2.8. This will conclude the proof of Theorem 1.1. The proof relies on the structure theorem for nilpotent approximate groups due to the second author [49], which we recall below as Theorem 3.9, as well as an analysis of nilprogressions and their powers, which forms the basis of Proposition 3.10, below.

Nilprogressions are to nilpotent groups what arithmetic progressions are to the infinite cyclic group \mathbb{Z} . There are several natural ways to define this concept. Although they do not lead to exactly the same notions, they are roughly equivalent on some level, as we will see. We start with the following definition.

Definition 3.1 (Nilprogression, see [10]). Let G be a group, let $x_1, \dots, x_r \in G$ and let $L = (L_1, \dots, L_r)$ be a vector of non-negative integers. If the x_1, \dots, x_r generate an s -step nilpotent subgroup of G then the set of all products in the x_i and their inverses, in which each x_i and its inverse appear at most L_i times between them, is said to be a *nilprogression* of rank r and step s , and is denoted $P^*(x_1, \dots, x_r; L)$.

For the purposes of this paper we need to introduce a new variant of a nilprogression, which we call a *nilcomplete progression*. This is closely related to yet another variant, introduced by Green and the first author in [8], where it is called a *nilpotent progression*. We show below precisely how they may be thought of as ‘roughly equivalent’; see (3.2) in particular. For further background on this rough equivalence the reader may consult the work [49] of the second author.

In order to define nilcomplete progressions, we need to establish some terminology. The following definition is reproduced from [49], and follows a set up in [8, §1] that was in turn based on [26, §11.1].

Definition 3.2 (Commutators and weights). We define (*formal*) *commutators* in the letters x_1, \dots, x_r recursively by defining each x_i to be a formal commutator and, given two formal commutators α, α' in the x_j , defining $[\alpha, \alpha']$ also to be a formal commutator.

To each commutator α we assign a *weight vector* $\chi(\alpha) \in \mathbb{N}_0^r$, defined recursively by setting $\chi(x_i) := e_i$ and, given two formal commutators α, α' in the x_j , defining $\chi([\alpha, \alpha']) = \chi(\alpha) + \chi(\alpha')$. We define the *total weight* $|\chi(\alpha)|$ of a commutator α to be $\|\chi(\alpha)\|_1$. Given a weight vector $\chi \in \mathbb{N}_0^r$ and a vector $L = (L_1, \dots, L_r)$ of positive integers we write L^χ to denote the quantity

$$L^\chi := L_1^{\chi_1} \dots L_r^{\chi_r}.$$

Noting that this results in at most finitely many commutators of any given weight vector, we assign a fixed total ordering \prec to the commutators, chosen arbitrarily subject to the conditions that $x_1 \prec \dots \prec x_r$, that commutators of the same weight vector are consecutive, and that commutators of lower total weight come before commutators of higher total weight.

Finally, for each commutator α we define the (*formal*) *inverse commutator* α^{-1} . We extend \prec to a partial ordering of commutators and their formal inverses, defining $\alpha^{\pm 1} \prec \beta^{\pm 1}$ when $\alpha \prec \beta$.

Certain commutators in the x_1, \dots, x_r are called *basic commutators*. These are defined precisely in each of [8, 26, 49]; for the purposes of this paper it will suffice to note that if G is the free s -step nilpotent

group on generators x_1, \dots, x_r , and $c_1, \dots, c_{t'}$ is the ordered (with respect to \prec) list of basic commutators of weight at most s in the x_i , then the elements

$$c_1^{l_1} \cdots c_{t'}^{l_{t'}}$$

are all distinct as the l_i range over \mathbb{Z} and, moreover, every element of G can be expressed in this form.

Definition 3.3 (Nilpotent progression, see [8]). Let G be a group, let x_1, \dots, x_r be elements of G and let $L = (L_1, \dots, L_r)$ be a vector of non-negative integers. If x_1, \dots, x_r generate an s -step nilpotent subgroup of G then, writing $c_1, \dots, c_{t'}$ for the ordered (with respect to \prec) list of basic commutators of weight at most s in the x_i , the set

$$P(x_1, \dots, x_r; L) := \{c_1^{l_1} \cdots c_{t'}^{l_{t'}} : |l_i| \leq L^{\chi(c_i)}\}$$

is said to be a *nilpotent progression* of rank r and step s . If the elements $c_1^{l_1} \cdots c_{t'}^{l_{t'}}$ are all distinct as the l_i range over $[-L^{\chi(c_i)}, L^{\chi(c_i)}]$ then the nilpotent progression $P(x_1, \dots, x_r; L)$ is said to be *proper*.

Remark 3.4. The nilpotent progression $P(x_1, \dots, x_r; L)$ is proper, for example, in the free s -step nilpotent group generated by x_1, \dots, x_r .

Remark 3.5. The cardinality of a nilpotent progression $P(x_1, \dots, x_r; L)$ is easily seen to be at most $\prod_{i=1}^t (2L^{\chi(c_i)} + 1)$, with equality if and only if $P(x_1, \dots, x_r; L)$ is proper. In the case that $P(x_1, \dots, x_r; L)$ is proper, this implies in particular that $|P(x_1, \dots, x_r; ML)| \leq M^{O_{r,s}(1)} |P(x_1, \dots, x_r; L)|$ for every $M \in \mathbb{N}$.

We now move on to the definition of a nilcomplete progression. The reason for introducing nilcomplete progressions is that they behave particularly well on taking powers or on scaling the side lengths L_i (see Proposition 3.10 below). The definition is similar to that of a nilpotent progression, except that we use the list of *all* commutators in the x_i instead of just the basic commutators, and that when taking commutators we allow lower-weight components to be replaced by their inverses (so, for example, we allow commutators such as $[x_1, x_2^{-1}]$ and $[x_3, [x_1^{-1}, x_2]^{-1}]$). We express this more formally as follows.

Definition 3.6 (Generalised commutators). We define the *generalised commutators* in the x_i in exactly the same way as the commutators, except that if α_1 and α_2 are two generalised commutators in the x_i then so is $[\alpha_1^{\epsilon_1}, \alpha_2^{\epsilon_2}]$ for every $\epsilon_1 = \pm 1$ and $\epsilon_2 = \pm 1$.

We extend the definition of weight vectors from commutators to generalised commutators by requiring that the weight vectors of $[\alpha_1^{\epsilon_1}, \alpha_2^{\epsilon_2}]$ are equal for all choices of ϵ_1 and ϵ_2 . We extend the order \prec arbitrarily so that the generalised commutators $[\alpha_1^{\epsilon_1}, \alpha_2^{\epsilon_2}]$ with different choices of ϵ_1 and ϵ_2 are consecutive.

Throughout this section we write ξ_1, \dots, ξ_t for the ordered (with respect to \prec) list of generalised commutators of total weight at most s in x_1, \dots, x_r . Note that t is bounded in terms of r and s only.

Definition 3.7 (Nilcomplete progression). Let G be a group, let x_1, \dots, x_r be elements of G , and let $L = (L_1, \dots, L_r)$ be a vector of non-negative integers. If x_1, \dots, x_r generate an s -step nilpotent subgroup of G then, writing ξ_1, \dots, ξ_t for the ordered (with respect to \prec) list of generalised commutators of weight at most s in the x_i , the set

$$\overline{P}(x_1, \dots, x_r; L) = \{\xi_1^{l_1} \xi_2^{l_2} \cdots \xi_t^{l_t} : |l_i| \leq L^{\chi(\xi_i)}\}$$

is said to be a *nilcomplete progression* of rank r and step s .

We recall a technical notion that was useful in [49].

Definition 3.8 (Ordered progression). Let x_1, \dots, x_r be elements of a group and let $L = (L_1, \dots, L_r)$ be a vector of non-negative integers. Then the set

$$P_{\text{ord}}(x_1, \dots, x_r; L) := \{x_1^{l_1} \cdots x_r^{l_r} : |l_i| \leq L_i\}$$

is said to be an *ordered progression* of rank r .

It is immediate from the definitions that

$$\overline{P}(x_1, \dots, x_r; L) = P_{\text{ord}}(\xi_1, \dots, \xi_t; L^{\chi(\xi_1)}, \dots, L^{\chi(\xi_t)}). \quad (3.1)$$

It was shown in [49] that nilprogressions, nilpotent progressions and nilcomplete progressions are ‘essentially’ equivalent, and that nilcomplete progressions control nilpotent approximate groups. In particular, we have

$$P_{\text{ord}}(x_1, \dots, x_r; L) \subset P^*(x_1, \dots, x_r; L) \subset P(x_1, \dots, x_r; L) \subset \overline{P}(x_1, \dots, x_r; L) \subset P_{\text{ord}}(x_1, \dots, x_r; L)^{r^{O_s(1)}}. \quad (3.2)$$

Here the first and third inclusions are trivial, while the second follows from the commutator collecting process, as described in [49]. The last inclusion is proven in [49, Proposition C.1] for P in place of \overline{P} , but precisely the same argument works and yields (3.2).

Next we recall the main result of [49].

Theorem 3.9 (Structure of nilpotent approximate groups, [49]). *Let G be an s -step nilpotent group, and suppose that A is a K -approximate subgroup of G . Then there is a subgroup H normalised by A , a natural number $r \leq K^{O_s(1)}$, elements $x_1, \dots, x_r \in G$, and a vector $L = (L_1, \dots, L_r)$ of non-negative integers such that*

$$A \subset HP^*(x_1, \dots, x_r; L) \subset HP(x_1, \dots, x_r; L) \subset H\overline{P}(x_1, \dots, x_r; L) \subset A^{K^{O_s(1)}}.$$

Again, this statement was proven in [49] only for P , and not for \overline{P} , but precisely the same proof yields the result for \overline{P} as well.

We are now ready to state the main technical result of this section, which is a description of the properties of nilcomplete progressions and their powers.

Proposition 3.10 (Powers of nilcomplete progressions). *Let G be an s -step nilpotent group, let x_1, \dots, x_r be elements of G , and let $L = (L_1, \dots, L_r)$ be a vector of positive integers. Then the following conditions hold.*

- (1) $\overline{P}(x_1, \dots, x_r; nL) \subset \overline{P}(x_1, \dots, x_r; L)^{O_{r,s}(n)}$.
- (2) $\overline{P}(x_1, \dots, x_r; L)^n \subset \overline{P}(x_1, \dots, x_r; nL)$.
- (3) For every $M \in \mathbb{N}$ there is some set $X \subset G$ of cardinality at most $O_{r,s,M}(1)$ such that

$$\overline{P}(x_1, \dots, x_r; ML) \subset \overline{P}(x_1, \dots, x_r; L)X.$$

Remark 3.11. An inspection of the arguments of [49] reveals that the exponent $r^{O_s(1)}$ in (3.2) could be written more precisely as $e^{O(s^4)}r^{O(s)}$, or simply $O_s(r^{O(s)})$. An inspection of the arguments below then shows that the constant implied by $O_{r,s}(n)$ in Proposition 3.10 (1) could be written as $e^{O(s^4)}r^{O(s)}n$, or more succinctly as $O_s(r^{O(s)}n)$. Similarly, the $O_{r,s,M}(1)$ in Proposition 3.10 (3) could be written as $r^{O(s^3)}r^{O(s)}M^{sr^{O(s)}}$, or more succinctly as $O_{r,s}(M^{sr^{O(s)}})$. We leave the details to the interested reader.

It is convenient at this point to introduce some further notation. Given a list y_1, \dots, y_k of commutators in the x_i we denote by $L^{\chi(y)}$ the k -dimensional vector

$$L^{\chi(y)} := (L^{\chi(y_1)}, \dots, L^{\chi(y_k)}).$$

Thus, for example, (3.1) can be rewritten more succinctly as

$$\overline{P}(x_1, \dots, x_r; L) = P_{\text{ord}}(\xi_1, \dots, \xi_t; L^{\chi(\xi)}). \quad (3.3)$$

In the event that $L = (L_1, \dots, L_r)$ is a vector of real numbers, as opposed to integers, we define $\lfloor L \rfloor$ and $\lceil L \rceil$ coordinate-wise; thus, for example, $\lfloor L \rfloor = (\lfloor L_1 \rfloor, \dots, \lfloor L_r \rfloor)$. When L' is another vector of integers or real numbers, we write $L < L'$ or $L \leq L'$ to indicate that the appropriate inequality holds coordinate-wise.

Finally we introduce one further definition.

Definition 3.12 (Nilcomplete set of generalised commutators). A set Y of generalised commutators in a list of group elements x_1, \dots, x_r is said to be *nilcomplete* if whenever $\alpha_1, \alpha_2 \in Y$ and $\epsilon_1 = \pm 1$ and $\epsilon_2 = \pm 1$ are such that $[\alpha_1^{\epsilon_1}, \alpha_2^{\epsilon_2}] \neq 1$, the generalised commutator $[\alpha_1^{\epsilon_1}, \alpha_2^{\epsilon_2}]$ also belongs to Y .

We now pass to the proof of Proposition 3.10.

Proof of Proposition 3.10 (1). We set $B(L) := \bigcup_1^r \{x_i^{\ell_i} : |\ell_i| \leq L_i\}$, and use the shorthand $\overline{P}(L) := \overline{P}(x_1, \dots, x_r; L)$. By definition of \overline{P} (see (3.3)) we have $\overline{P}(nL) \subset P(\xi_1; (nL)^{\chi(\xi_1)}) \cdots P(\xi_t; (nL)^{\chi(\xi_t)})$. However, [49, Lemma C.2] implies that

$$P(\xi_i; (nL)^{\chi(\xi_i)}) \subset B(nL)^{O_s(1)}.$$

Noting again that t is bounded in terms of r and s only, we see that

$$\overline{P}(nL) \subset B(nL)^{O_{r,s}(1)} \subset B(L)^{O_{r,s}(n)}, \quad (3.4)$$

and the result then follows from the fact that $B(L) \subset \overline{P}(L)$. \square

Proof of Proposition 3.10 (2). We in fact show that

$$P_{\text{ord}}(\xi_1, \dots, \xi_t; L^{\chi(\xi)})^n \subset P_{\text{ord}}(\xi_1, \dots, \xi_t; (nL)^{\chi(\xi)});$$

in light of (3.3) this is sufficient to prove Proposition 3.10 (2).

First, note the trivial identity

$$vu = uv[v, u]. \quad (3.5)$$

Since the ξ_i are generalised commutators, it is very straightforward to use this identity repeatedly to see that any string σ in the x_i is expressible in the form

$$\xi_1^{l_1} \cdots \xi_t^{l_t} \quad (3.6)$$

for *some* $l_1, \dots, l_t \in \mathbb{Z}$. For example, the string $x_2 x_1^{-1}$ is equal to $x_1^{-1} x_2 [x_2, x_1^{-1}]$. This is essentially a simplified version of the *collecting process* of [26, §11.1]. The challenge is to show that if σ is of the form $p^{(1)} \cdots p^{(n)}$ with each $p^{(i)}$ belonging to $P_{\text{ord}}(\xi_1, \dots, \xi_t; L^{\chi(\xi)})$ then we may conclude additionally that $|l_i| \leq (nL)^{\chi(\xi_i)}$.

Let $P^{(1)}, \dots, P^{(n)}$ be distinct copies of $P_{\text{ord}}(\xi_1, \dots, \xi_t; L^{\chi(\xi)})$, so that $P_{\text{ord}}(\xi_1, \dots, \xi_t; L^{\chi(\xi)})^n$ may be written as $P^{(1)} \cdots P^{(n)}$. We bound from above the number of instances of each generalised commutator ξ_i that can occur when for each j we have an element $p^{(j)}$ of $P^{(j)}$ viewed as a string in the ξ_i in the natural way, and the simplified collecting process is applied to the string $p^{(1)} \cdots p^{(n)}$.

In order to do this, we give a different label to each occurrence of each generalised commutator ξ_i in the initial string $p^{(1)} \dots p^{(n)}$. Specifically, given the string $p^{(j)} = \xi_1^{l_1^{(j)}} \dots \xi_t^{l_t^{(j)}}$ in $P^{(j)}$, for each $i = 1, \dots, t$ and for $i' = 1, \dots, l_i$ let $\xi_{i,i'}^{(j)}$ be distinct copies of the element ξ_i . Thus we have

$$\xi_1^{l_1^{(j)}} \dots \xi_t^{l_t^{(j)}} = \left(\xi_{1,1}^{(j)} \dots \xi_{1,l_1^{(j)}}^{(j)} \right) \dots \left(\xi_{t,1}^{(j)} \dots \xi_{t,l_t^{(j)}}^{(j)} \right)$$

and, in particular,

$$p^{(1)} \dots p^{(n)} = \left(\left(\xi_{1,1}^{(1)} \dots \xi_{1,l_1^{(1)}}^{(1)} \right) \dots \left(\xi_{t,1}^{(1)} \dots \xi_{t,l_t^{(1)}}^{(1)} \right) \right) \dots \left(\left(\xi_{1,1}^{(n)} \dots \xi_{1,l_1^{(n)}}^{(n)} \right) \dots \left(\xi_{t,1}^{(n)} \dots \xi_{t,l_t^{(n)}}^{(n)} \right) \right). \quad (3.7)$$

Let us apply (3.5) repeatedly to this string so that when we once again identify each occurrence of the same generalised commutator it is of the form (3.6). Specifically, beginning with the left-most occurrence of ξ_1 , use identity (3.5) to move each instance of ξ_1 to the left until it is to the left of every instance of every commutator that is later with respect to the order \prec ; then perform the same process with the instances of ξ_2 ; and so on up to the instances of ξ_t (which by this time, of course, will be at the extreme right of the expression anyway).

Once this is done, the resultant string of the form (3.6) will consist of commutators whose arguments are elements of the form $\xi_{i,i'}^{(j)}$. Some of these will simply be the original elements $\xi_{i,i'}^{(j)}$ from the string (3.7); let us call such elements *prime* instances of the commutator ξ_i . However, suppose the commutator ξ_i is equal to $[\xi_v, \xi_u]$, with $u < v$. Then there may also be instances of ξ_i in (3.6) that resulted from interchanging instances of ξ_u with instances of ξ_v ; let us call such instances of ξ_i *composite* instances. Note that a composite instance of the commutator $[\xi_v, \xi_u]$ could have arisen from interchanging a composite instance of ξ_u with a composite instance of ξ_v .

Write $p_i^{(j)}$ for the number of prime instances of ξ_i originating from the progression $P^{(j)}$, and c_i for the total number of composite instances of ξ_i . Our aim is to prove that $p_i^{(1)} + \dots + p_i^{(n)} + c_i \leq (nL)^{\chi(\xi_i)}$. Since

$$p_i^{(j)} \leq L^{\chi(\xi_i)} \quad (3.8)$$

for each j by definition, this amounts to showing that

$$c_i \leq (nL)^{\chi(\xi_i)} - nL^{\chi(\xi_i)}. \quad (3.9)$$

This is trivial for $i = 1, \dots, r$, since for each value of i less than r we have $c_i = 0$, and so by induction we may prove (3.9) for an arbitrary fixed value w of i , under the assumptions that $w > r$ and that (3.9) holds for all smaller values of i . Since $w > r$ we have $\xi_w = [\xi_v, \xi_u]$ for some u and v satisfying $u < v < w$.

Let us first bound from above the number of composite instances of ξ_w in which the ξ_u component is prime. Indeed, consider first the number of composite instances of ξ_w in which the ξ_u component is a prime instance originating from $P^{(j)}$. Prime instances of commutators originating in the same progression $P^{(j)}$ are in the desired order from the beginning, and so will never swap under the simplified collecting process, and so if ζ is a composite instance of ξ_w in which the ξ_u component is a prime instance originating from $P^{(j)}$, then its ξ_v component cannot be a prime instance originating from $P^{(j)}$.

The number of instances of ξ_v that are not prime instances coming from $P^{(j)}$ is at most $(nL)^{\chi(\xi_v)} - L^{\chi(\xi_v)}$ by (3.8) and the induction hypothesis. On the other hand, (3.8) implies that the number of prime instances of ξ_u originating from $P^{(j)}$ is at most $L^{\chi(\xi_u)}$. The total number of composite instances of ξ_w in which the ξ_u component is a prime instance originating from $P^{(j)}$ is therefore at most

$$(n^{\chi(\xi_v)} - 1)L^{\chi(\xi_u)},$$

and so the total number of composite instances of ξ_w in which the ξ_u component is a prime instance originating from *any* progression is at most

$$(n^{|\chi(\xi_v)|+1} - n)L^{\chi(\xi_w)}. \quad (3.10)$$

Let us now bound the number of composite instances of ξ_w in which the ξ_u component is composite. By induction, the number of composite instances of ξ_u is at most $(n^{|\chi(\xi_u)|} - n)L^{\chi(\xi_u)}$, whilst by induction and (3.8) the total number of instances of ξ_v is at most $n^{|\chi(\xi_v)|}L^{\chi(\xi_v)}$. The number of composite instances of ξ_w in which the ξ_u component is composite is therefore at most

$$(n^{|\chi(\xi_w)|} - n^{|\chi(\xi_v)|+1})L^{\chi(\xi_w)}. \quad (3.11)$$

The total number of composite instances of ξ_w is therefore at most the sum of the quantities (3.10) and (3.11), which is $(n^{|\chi(\xi_w)|} - n)L^{\chi(\xi_w)}$. This verifies (3.9) in the case $i = w$, and so ends the proof. \square

In light of (3.3), Proposition 3.10 (3) follows from the following more general statement.

Proposition 3.13. *Let y_1, \dots, y_k be a nilcomplete set of generalised commutators in the x_i , appearing in order with respect to \prec . Then for every $M \in \mathbb{N}$ there is some set $X \subset G$ of size bounded in terms of k, s, M such that*

$$P_{\text{ord}}(y_1, \dots, y_k; (ML)^{\chi(y)}) \subset P_{\text{ord}}(y_1, \dots, y_k; L^{\chi(y)}X).$$

The first containment of (3.2) expresses the trivial fact that an ordered progression is contained in inside the corresponding nilprogression. The following useful lemma shows that the reverse containment is also approximately true if the generators form a nilcomplete set.

Lemma 3.14. *Let y_1, \dots, y_k be a nilcomplete set of generalised commutators in the x_i , appearing in order with respect to \prec . Then*

$$P^*(y_1, \dots, y_k; L^{\chi(y)}) \subset P_{\text{ord}}(y_1, \dots, y_k; (O_{k,s}(1)L)^{\chi(y)}).$$

Proof. Since the y_i are in order with respect to \prec , for each j the sequence y_j, \dots, y_k is still nilcomplete, and so by induction on k it suffices to prove that

$$P^*(y_1, \dots, y_k; L^{\chi(y)}) \subset P(y_1; L^{\chi(y_1)})P^*(y_2, \dots, y_k; (O_{k,s}(1)L)^{\chi(y_2)}, \dots, (O_{k,s}(1)L)^{\chi(y_k)}).$$

By (3.2) it is in fact sufficient to prove that

$$P(y_1, \dots, y_k; L^{\chi(y)}) \subset P(y_1; L^{\chi(y_1)})P^*(y_2, \dots, y_k; (O_{k,s}(1)L)^{\chi(y_2)}, \dots, (O_{k,s}(1)L)^{\chi(y_k)}).$$

However, the fact that the set of y_i is nilcomplete implies that each basic commutator in the y_i is itself equal to some y_j , and so this inclusion follows from the fact that there are at most $O_{k,s}(1)$ basic commutators in the y_i of total weight at most s . \square

Lemma 3.15. *Let z, y_1, \dots, y_k be a nilcomplete set of generalised commutators in the x_i , appearing in order with respect to \prec . Let $m \in \mathbb{Z}$, and suppose that*

$$|m| \leq L^{\chi(z)}. \quad (3.12)$$

Then

$$z^m P^*(y_1, \dots, y_k; L^{\chi(y)}) \subset P^*(y_1, \dots, y_k; (O_{k,s}(1)L)^{\chi(y)})z^m.$$

Proof. Let p be an arbitrary element of $P^*(y_1, \dots, y_k; L^{\chi(y)})$, viewed in the natural way as a string in the y_i and their inverses. Repeated application of the identity $uv = [u^{-1}, v^{-1}]vu$ implies that there is some formal string p' obtained by inserting into p , at various places, commutators of the form

$$[z^{-m}, y_j^{-1}] \quad (3.13)$$

or

$$[z^{-m}, y_j], \quad (3.14)$$

such that

$$z^m p = p' z^m$$

as group elements. By [49, Proposition B.3], for each j there are natural numbers t_j and $n_{j,1}, \dots, n_{j,t_j}$ with

$$t_j \leq O_s(1), \quad (3.15)$$

and generalised commutators $\zeta_{j,1}, \dots, \zeta_{j,t_j}$ and $\hat{\zeta}_{j,1}, \dots, \hat{\zeta}_{j,t_j}$ in z and y_j , such that

$$[z^{-m}, y_j^{-1}] = \zeta_{j,1}^{n_{j,1}} \dots \zeta_{j,t_j}^{n_{j,t_j}}$$

and

$$[z^{-m}, y_j] = \hat{\zeta}_{j,1}^{n_{j,1}} \dots \hat{\zeta}_{j,t_j}^{n_{j,t_j}}.$$

Moreover, writing $\omega_{j,i}$ for the z -weight of the generalised commutator $\zeta_{j,i}$ considered as a commutator in z and y_j , the bound (3.12) implies that the integers $n_{j,i}$ arising from [49, Proposition B.3] satisfy

$$|n_{j,i}| \leq L^{\omega_{j,i}\chi(z)}. \quad (3.16)$$

We may therefore replace each instance of (3.13) in p' with a string of the form $\zeta_{j,1}^{n_{j,1}} \dots \zeta_{j,t_j}^{n_{j,t_j}}$, and each instance of (3.14) with a string of the form $\hat{\zeta}_{j,1}^{n_{j,1}} \dots \hat{\zeta}_{j,t_j}^{n_{j,t_j}}$. Write p'' for the string produced by making each of these replacements, so that p' and p'' are equal as group elements.

Now, each instance of (3.13) or (3.14) appearing in p' arises from an element of the form y_j or y_j^{-1} in the string p . There are at most $L^{\chi(y_j)}$ such elements in the string p by definition, and so for each j there can be at most $L^{\chi(y_j)}$ instances of (3.13) and (3.14) between them. In particular, there are at most $L^{\chi(y_j)}|n_{j,i}|$ instances of the strings $\zeta_{j,i}$ and $\hat{\zeta}_{j,i}$ in the string p'' . The inequality (3.16) therefore implies that the total number of instances of the strings $\zeta_{j,i}$ and $\hat{\zeta}_{j,i}$ in the string p'' is at most $L^{\chi(\zeta_{j,i})}$. Here, $\chi(\zeta_{j,i})$ refers to the weight of $\zeta_{j,i}$ as a generalised commutator in the x_i .

Since the list z, y_1, \dots, y_k is nilcomplete, each $\zeta_{j,i}$ is equal to some $y_{l_{j,i}}$, and each $\hat{\zeta}_{j,i}$ is equal to some $y_{\hat{l}_{j,i}}$. Let \hat{p} be the string of y_l formed from p'' by replacing each instance of $\zeta_{j,i}$ in p'' by $y_{l_{j,i}}$, and each instance of $\hat{\zeta}_{j,i}$ by $y_{\hat{l}_{j,i}}$. It follows from (3.15) that the total number of $\zeta_{j,i}$ and $\hat{\zeta}_{j,i}$ is at most $O_{k,s}(1)$, and so the preceding paragraph and the definition of p imply that the string \hat{p} contains at most $(O_{k,s}(1)L)^{\chi(y_l)}$ instances of y_l and its inverse. In particular, $\hat{p} \in P^*(y_1, \dots, y_k; (O_{k,s}(M)L)^{\chi(y)})$, and so the lemma follows from the fact that \hat{p} and p' are equal as group elements. \square

We can now finally prove Proposition 3.13.

Proof of Proposition 3.13. We prove that there is a set X of size bounded in terms of s and M such that

$$P_{\text{ord}}(y_1, \dots, y_k; (ML)^{\chi(y)}) \subset P_{\text{ord}}(y_1, \dots, y_k; L^{\chi(y_1)}, (O_{k,s}(M)L)^{\chi(y_2)}, \dots, (O_{k,s}(M)L)^{\chi(y_k)})X;$$

by induction on k this suffices to prove the proposition. More precisely, we prove that this holds with $X = P(y_1^{L^{\chi(y_1)}}; M^{|\chi(y_1)|})$, which has cardinality at most $2M^s + 1$. To prove that it holds with this choice of X , note simply that

$$\begin{aligned} & P_{\text{ord}}(y_1, \dots, y_k; (ML)^{\chi(y)}) \\ & \subset P(y_1, L^{\chi(y_1)})P^*(y_2, \dots, y_k; (O_{k,s}(M)L)^{\chi(y_2)}, \dots, (O_{k,s}(M)L)^{\chi(y_k)})P(y_1^{L^{\chi(y_1)}}; M^{|\chi(y_1)|}) \end{aligned}$$

by Lemma 3.15, and that

$$\begin{aligned} & P^*(y_2, \dots, y_k; (O_{k,s}(M)L)^{\chi(y_2)}, \dots, (O_{k,s}(M)L)^{\chi(y_k)}) \\ & \subset P_{\text{ord}}(y_2, \dots, y_k; (O_{k,s}(M)L)^{\chi(y_2)}, \dots, (O_{k,s}(M)L)^{\chi(y_k)}) \end{aligned}$$

by Lemma 3.14. This ends the proof of Proposition 3.13, and hence of Proposition 3.10 (3). \square

Having proved Proposition 3.10, we can now easily deduce Theorem 2.8.

Proof of Theorem 2.8. First we claim that it is enough to find positive integers N_0 and C depending only on K and s such that $S^{N_0 n}$ is a C -approximate subgroup for all $n \geq 1$. Indeed, since S is a K -approximate subgroup, there is some $Y \subset G$, the size of which is bounded in terms of K and s , such that $S^{2N_0} \subset SY$. Given any integer $m \geq N_0$, there exists $n \geq 1$ such that $nN_0 \leq m < (n+1)N_0$, and so $S^{2m} \subset S^{2(n+1)N_0-1} = S^{2nN_0-1}S^{2N_0} \subset S^{2nN_0}Y \subset S^{nN_0}XY$ for some subset X of size at most C . For $m \leq N_0$, meanwhile, we have $S^{2m} \subset S^{2N_0} \subset S^m Y$. This proves the claim.

Now, without loss of generality, we may assume that S generates G . By Theorem 3.9 there is a finite subgroup $H \leq G$ normalised by S , and hence normal in G ; integers $r, M \leq K^{O_s(1)}$; and a nilcomplete progression $\overline{P}(L) := \overline{P}(x_1, \dots, x_r; L)$ for some elements x_1, \dots, x_r in G such that

$$S \subset H\overline{P}(L) \subset S^M.$$

Increasing M if necessary, Proposition 3.10 (1) implies that $\overline{P}(nL) \subset \overline{P}(L)^{Mn}$ for all $n \geq 1$. Let $N_0 := M^2$. Since H is normal in G , we have $(H\overline{P}(L))^n = H\overline{P}(L)^n$ for every $n \geq 1$. Applying Proposition 3.10 (2) and (3), we therefore see that for each $n \geq 1$ there is a subset X_n of size bounded in terms of s and K such that

$$\begin{aligned} S^{2N_0 n} & \subset (H\overline{P}(L))^{2N_0 n} = H\overline{P}(L)^{2N_0 n} \subset H\overline{P}(2N_0 nL) \\ & \subset \overline{P}(nL)X_n \subset H(\overline{P}(L))^{Mn}X_n \subset S^{M^2 n}X_n = S^{N_0 n}X_n, \end{aligned}$$

and so the theorem is proved. \square

This also completes the proof of Theorem 1.1.

3.1. Doubling of the various types of progression. We have made heavy use of the principle that sets of small doubling are controlled by nilprogressions. In this subsection we take the opportunity to explain how Proposition 3.10 can be used to shed light on the extent to which the converse holds.

First, we note that nilcomplete progressions always enjoy small doubling. Indeed, an immediate consequence of Proposition 3.10 is that for every nilcomplete progression \overline{P} of rank r and step s there is a set X of cardinality $O_{r,s}(1)$ such that $\overline{P}^2 \subset \overline{P}X$.

For nilprogressions the situation is not quite so straightforward; one may check, for example, that in the free s -step nilpotent group generated by x_1 and x_2 the ratio $|P^*(x_1, x_2; L, 1)^3|/|P^*(x_1, x_2; L, 1)|$ tends to infinity as $L \rightarrow \infty$. Nonetheless, it is shown in Tao's book [46, Proposition 2.2.3] that, provided the

lengths L_1, \dots, L_r are large enough in terms of the rank r and step s only, a nilprogression is an $O_{r,s}(1)$ -approximate group. We recover this fact, and obtain a similar statement for nilpotent progressions, as follows.

Corollary 3.16. *Fix $r, s \geq 1$. Then there is a constant $\lambda = \lambda_{r,s}$ such that whenever x_1, \dots, x_r are elements in an s -step nilpotent group and $L = (L_1, \dots, L_r)$ is a vector of integers that are at least λ , there exists a set X of cardinality at most $O_{r,s}(1)$ such that the nilprogression $P^* = P^*(x_1, \dots, x_r; L)$ and the nilpotent progression $P = P(x_1, \dots, x_r; L)$ satisfy*

$$(P^*)^2 \subset P^2 \subset P^*X \subset PX.$$

In particular, since it is symmetric the nilprogression P^* is an $O_{r,s}(1)$ -approximate group.

Proof. Note that (3.4) implies the existence of a constant $C = C_{r,s}$ such that

$$\overline{P}(x_1, \dots, x_r; L') \subset P^*(x_1, \dots, x_r; CL') \quad (3.17)$$

for every $L' = (L'_1, \dots, L'_r)$. We claim that taking $\lambda = 2C$ satisfies the requirements of the corollary; thus we may assume that $L_1, \dots, L_r \geq 2C$. This implies in particular that

$$\lceil L_i/2C \rceil \leq L_i/C \quad (3.18)$$

for each i . Denote by $\lceil L/2C \rceil$ the vector $(\lceil L_1/2C \rceil, \dots, \lceil L_r/2C \rceil)$, and let X be the set obtained by applying Proposition 3.10 (3) to $\overline{P}(x_1, \dots, x_r; \lceil L/2C \rceil)$ in the case $M = 4C$, so that

$$\overline{P}(x_1, \dots, x_r; 4C\lceil L/2C \rceil) \subset \overline{P}(x_1, \dots, x_r; \lceil L/2C \rceil)X. \quad (3.19)$$

Now note that

$$\begin{aligned} P^*(x_1, \dots, x_r; L)^2 &\subset P(x_1, \dots, x_r; L)^2 && \text{by (3.2)} \\ &\subset \overline{P}(x_1, \dots, x_r; L)^2 && \text{by (3.2)} \\ &\subset \overline{P}(x_1, \dots, x_r; 2L) && \text{by Proposition 3.10} \\ &\subset \overline{P}(x_1, \dots, x_r; 4C\lceil L/2C \rceil) \\ &\subset \overline{P}(x_1, \dots, x_r; \lceil L/2C \rceil)X && \text{by (3.19)} \\ &\subset P^*(x_1, \dots, x_r; L)X && \text{by (3.17) and (3.18)} \\ &\subset P(x_1, \dots, x_r; L)X && \text{by (3.2)}. \end{aligned}$$

□

One may show by a similar argument, making use of Remark 3.5, that a *proper* s -step nilpotent progression P of rank r satisfies $|P^k| \leq O_{k,r,s}(1)|P|$, irrespective of its side lengths. We leave the details to the interested reader.

4. A GROMOV-TYPE THEOREM FOR FINITE GROUPS

In this section we prove Theorem 1.3 about the algebraic structure of almost flat Cayley graphs, together with two variants of that statement, which we gather as follows.

Theorem 4.1 (Structure of almost flat finite groups). *Let $\varepsilon, \delta > 0$. Let G be a finite group generated by a symmetric subset S whose Cayley graph has diameter $\gamma := \text{diam}_S(G)$ and is ε -almost flat, which is to say that*

$$\gamma \geq \left(\frac{|G|}{|S|} \right)^\varepsilon.$$

Then there is a constant $D_{\varepsilon, \delta}$ depending only on ε, δ such that if $\gamma \geq D_{\varepsilon, \delta}$ then

- (1) *G has a normal subgroup H contained in $S^{[\gamma^\delta]}$ such that G/H has a nilpotent subgroup of index, rank and nilpotency class bounded in terms of ε and δ only;*
- (2) *G has a normal subgroup H contained in $S^{[\gamma^{1/2+\delta}]}$ such that G/H has an abelian subgroup of index and rank bounded in terms of ε and δ only;*
- (3) *G has a normal subgroup G_0 of index at most C_ε that admits a cyclic quotient with diameter at least $c_\varepsilon \gamma$, where $C_\varepsilon, c_\varepsilon$ are positive constants depending on ε only.*

In (3), the diameter of the cyclic quotient is understood with respect to a system of generators of G_0 derived from S by the usual Reidemeister–Schreier process, which we summarise in the following lemma.

Lemma 4.2. *Let G be a group generated by a finite symmetric set S , and let G_0 be a subgroup of G of index d . Then $G_0 \cap S^{2d-1}$ contains a symmetric generating set S_0 for G_0 satisfying $\frac{1}{d}|S| \leq |S_0| \leq d|S|$ and $\frac{1}{2d}(\text{diam}_S(G) - d) \leq \text{diam}_{S_0}(G_0) \leq \text{diam}_S(G)$.*

Proof. Let T be a set of right-coset representatives for G_0 in G , and for $x \in G$ write $\tau(x)$ for the unique element of T such that $x \in G_0\tau(x)$. It is shown in [26, Lemma 7.2.2] that the set $S_0 = \{ts\tau(ts)^{-1} : s \in S, t \in T\}$ is symmetric and generates G_0 with diameter at most $\text{diam}_S(G)$. We may assume that $d > 1$, in which case S intersects at least two cosets of G_0 . The argument of Lemma 2.7 therefore shows that we may take $T \subset S^{d-1}$, and so $S_0 \subset G_0 \cap S^{2d-1}$ and $\text{diam}_S(G) \leq 2d \text{diam}_{S_0}(G_0) + d$, as desired. It is trivial that $|S_0| \leq d|S|$, and the fact that $|S_0| \geq \frac{1}{d}|S|$ follows because there must be a coset of G_0 containing at least $\frac{1}{d}|S|$ elements of S . \square

Remark 4.3. Lemma 4.2 implies that if the index of G_0 is bounded and $\text{diam}_S(G)$ is sufficiently large then (G, S) is almost flat if and only if (G_0, S_0) is almost flat.

Remark 4.4. Theorem 4.1 (3) can be seen as a finite analogue of the key step in Gromov’s proof of his polynomial growth theorem [23] in which he establishes that every finitely generated group with polynomial growth has a finite index subgroup, which maps onto \mathbb{Z} .

Remark 4.5. Lee and Makarychev [35] show that if the Cayley graph of a finite group is uniformly doubling, then a similar conclusion to that of Theorem 4.1 (3) holds. Their proof is very different and closer to Kleiner’s approach to Gromov’s theorem via harmonic functions, and requires the doubling assumption to hold at all scales. By contrast, our argument requires doubling only at one large scale (see Prop. 4.24).

Remark 4.6. The proof of Theorem 4.1 (2) shows, more generally, that for any integer $s \geq 1$ the group G has a normal subgroup H contained in $S^{[\gamma^{\frac{1}{s+1}+\delta}]}$ such that G/H has an s -step nilpotent subgroup of index and rank bounded in terms of ε and δ only.

We also prove the following result, which may be thought of as a converse to Theorem 4.1.

Proposition 4.7 (Nilpotent groups are almost flat). *Let G be a finite s -step nilpotent group generated by a symmetric subset S . There is a positive constant $\varepsilon = \varepsilon(s, |S|)$ depending only on s and $|S|$ such that $\text{diam}_S(G) \geq |G|^\varepsilon$.*

Remark 4.8. The example of \mathbb{F}_2^n with the standard generating set shows that the ε in Proposition 4.7 has to depend on $|S|$.

We start the proof of Theorem 4.1 by translating the almost flat assumption into a doubling property.

Lemma 4.9 (Almost flat groups are doubling at some scale). *Let $\varepsilon, \delta > 0$ and set $K = 5^{\frac{2}{\varepsilon\delta}}$. Consider an ε -almost flat Cayley graph of a finite group G with symmetric generating set S . Let $\gamma := \text{diam}_S(G)$ be its diameter. Then there is an integer n satisfying $\gamma^{\delta/2} \leq n \leq \gamma^\delta$ such that*

$$|S^{5^n}| \leq K|S^n|.$$

Proof. Let k_0 be the largest k such that $5^k \gamma^{\delta/2} \leq 5\gamma^\delta$, and note therefore that $5^{k_0} \geq \gamma^{\delta/2}$. The almost flat assumption therefore implies that

$$|S^{5^{k_0} \gamma^{\delta/2}}| \leq |G| \leq |G||S^{\gamma^{\delta/2}}|/|S| \leq \gamma^{1/\varepsilon} |S^{\gamma^{\delta/2}}| = \left(\gamma^{\delta/2}\right)^{2/\delta\varepsilon} |S^{\gamma^{\delta/2}}| \leq 5^{2k_0/\delta\varepsilon} |S^{\gamma^{\delta/2}}| = K^{k_0} |S^{\gamma^{\delta/2}}|,$$

from which the conclusion follows easily. \square

Proof of Theorem 4.1 (1). Let $K = 5^{\frac{4}{\varepsilon\delta}}$ and let $n_0 = n_0(K)$ be the integer given by Proposition 2.9, and take D large enough to force $\gamma^{\delta/4} > n_0$. Lemma 4.9 then implies that there exists n with $n_0 \leq n \leq \gamma^{\delta/2}$ such that $|S^{5^n}| \leq K|S^n|$. The conclusion then follows from Proposition 2.9. \square

We now pass to the proof of part (2) of Theorem 4.1. It relies on some structural results about finite nilpotent groups, and in particular about commutators and nilprogressions; we refer the reader to Section 3 for background on these and definitions. We will also use the following standard fact, which follows from [49, Proposition B.2], for example.

Lemma 4.10. *Let α be a commutator form of weight k in the sense of [49, Definition 3.2]. Then the value of $\alpha(y_1, \dots, y_k) \bmod G_{k+1}$ is multilinear in the variables y_1, \dots, y_k .*

This means in particular that the value of $\alpha(y_1, \dots, y_k) \bmod G_{k+1}$ depends only on the values of $y_i \bmod [G, G]$.

Lemma 4.11 (Commutators in finite nilpotent groups are close to the identity). *Let G be a finite s -step nilpotent group generated by the nilprogression $P = P^*(x_1, \dots, x_r; L)$. Then the commutator subgroup $[G, G]$ is contained in $P^{O_{r,s}(\gamma^{1/2})}$, where $\gamma := \text{diam}_P(G)$.*

Proof. By induction on s , it is enough to prove that G_s , the last non-trivial term in the lower central series of G , is contained in $P^{O_{r,s}(\gamma^{1/2})}$. In fact, since G_s is an abelian group generated by the basic commutators of weight s in the x_i , of which there are at most $O_{r,s}(1)$, it suffices to prove that every power of each such commutator is contained in $P^{O_{r,s}(\gamma^{1/2})}$. However, Lemma 4.10 means that if $\alpha(x_{j_1}, \dots, x_{j_s})$ is a commutator of weight s in the x_i then $\alpha(x_{j_1}, \dots, x_{j_s})^n = \alpha(x_{j_1}^n, x_{j_2}, \dots, x_{j_s})$. We have $x_{j_1}^n \in P^\gamma$ by definition, and so there are ℓ_1, \dots, ℓ_r such that $|\ell_i| \leq L_i \gamma$ and such that $x_i^n = x_1^{\ell_1} \cdots x_r^{\ell_r} \pmod{[G, G]}$, and hence that $\alpha(x_{j_1}^n, x_{j_2}, \dots, x_{j_s}) = \prod_{i=1}^r \alpha(x_i, x_{j_2}, \dots, x_{j_s})^{\ell_i}$, again by Lemma 4.10. It will therefore suffice to show that each $\alpha(x_i, x_{j_2}, \dots, x_{j_s})^{\ell_i}$ with $|\ell_i| \leq L_i \gamma$ belongs to $P^{O_{r,s}(\gamma^{1/2})}$. However, $\alpha(x_i, x_{j_2}, \dots, x_{j_s})^{\ell_i} \in \overline{P}(x_1, \dots, x_r; [\gamma^{1/s}]L)$ by definition, and this last set is contained in $P^{O_{r,s}(\gamma^{1/s})}$ by the proof of Proposition 3.10 (1), since in the notation of that proof we have $B(L) \subset P^*(L)$. The lemma is therefore proved. \square

Remark 4.12. Essentially the same proof shows that the same result holds if P is a nilpotent progression or nilcomplete progression.

Proof of Theorem 4.1 (2). By part (1) applied to $\delta = \frac{1}{3}$, we may pass to a quotient and assume that G has a normal nilpotent subgroup Γ whose index, class and rank are bounded in terms of ε only. Indeed, the ratio $|G|/|S|$ does not increase when passing to a group quotient, and the diameter of the quotient will be at least $\gamma - \gamma^{1/3}$. Lemma 4.2 then gives a generating set $S_0 \subset S^{O_\varepsilon(1)}$ for Γ such that $\text{diam}_S(\Gamma)$ is comparable to γ and such that (Γ, S_0) is ε' -almost flat for some positive ε' depending on ε only. Arguing as in the proof of Theorem 4.1 (1), and using Lemma 2.2, we see that there exists $n \leq O_{\varepsilon, \delta}(\gamma^\delta)$ such that S_0^n is an $O_{\varepsilon, \delta}(1)$ -approximate subgroup. By Theorem 3.9, therefore, and possibly after passing to a further quotient, there exists a nilprogression P of bounded rank such that $S_0 \subset P \subset S_0^{O_{\varepsilon, \delta}(\gamma^\delta)}$, and hence that $\text{diam}_P(\Gamma) \geq \Omega_{\varepsilon, \delta}(\gamma^{1-\delta})$. Lemma 4.11 therefore implies and we conclude that $[\Gamma, \Gamma] \subset P^{O_\varepsilon(\gamma^{1/2})}$, and hence $[\Gamma, \Gamma] \subset S^{O_{\varepsilon, \delta}(\gamma^{1/2+\delta})}$. The subgroup $H := [\Gamma, \Gamma]$ is characteristic in Γ , and hence normal in G , and the quotient G/H has the desired properties. \square

Lemma 4.13. *Let G be a finite nilpotent group of rank r and step s . Then*

$$|G/[G, G]| \geq |G|^{\Omega_{r, s}(1)}.$$

Proof. We denote by $G = G_0 = G_1 \geq G_2 \geq \dots \geq G_s \geq G_{s+1} = \{1\}$ the lower central series of G . Observe that

$$|G_k| = \prod_{i=k}^s |G_i/G_{i+1}|, \quad (4.1)$$

that G_k/G_{k+1} is abelian and generated by the basic commutators in the x_i of total weight k (see Section 3 or [26]). Let x_1, \dots, x_r be a generating set for G . Note that the group $G/[G, G]$ is generated by the elements $x_i[G, G]$. Write d for the order of the element $x_i[G, G]$ of maximum order in $G/[G, G]$, and note that $|G/[G, G]| \geq d$. It follows from Lemma 4.10 that the order in the group G_k/G_{k+1} of a commutator of total weight k featuring generators x_{i_1}, \dots, x_{i_k} is at most d . Therefore, writing $b(r, k)$ for the number of basic commutators of total weight k in r generators we have

$$|G_{k-1}/G_k| \leq d^{b(r, k)}.$$

In particular, writing $c(r, s)$ for the number of basic commutators of total weight at most s in r generators, (4.1) implies that $|[G, G]| \leq d^{c(r, s)}$. The result follows. \square

Proof of Proposition 4.7. Let $\gamma := \text{diam}_S(G)$. Suppose first that G is abelian. Then $|S^n| \leq n^{|S|}$, and hence $\gamma \geq |G|^{1/|S|}$. If G is nilpotent of class s , then Lemma 4.13 shows that $|G| \leq |G/[G, G]|^{C_{|S|, s}} \leq \gamma^{|S|C_{|S|, s}}$ as desired. \square

We now pass to the proof of the part (3) of Theorem 4.1. For this we require the following fact about flat tori.

Lemma 4.14 (Euclidean tori have large one-dimensional quotients). *Let Δ be a discrete subgroup of rank n in \mathbb{R}^n and consider the torus $T := \mathbb{R}^n/\Delta$, endowed with the standard Euclidean metric. Then there is a one-dimensional quotient torus with diameter at least $\frac{\pi}{2nn!}$ times the diameter of T .*

Proof. Let v_1, \dots, v_n be a basis of successive minima of Δ , with $\|v_1\| \leq \dots \leq \|v_n\|$. Let H be the \mathbb{R} -span of v_1, \dots, v_{n-1} . Let T_0 be the quotient torus $(\mathbb{R}^n/H)/(\Delta \cap H)$. The diameter of T_0 is the supremum over all $x \in \mathbb{R}^n$ of the distance from x to $H + \Delta$. This is half the distance from v_n to H . On the other hand

the diameter $\text{diam}(T)$ of T is the supremum $\sup_{x \in \mathbb{R}^n} \inf_{\delta \in \Delta} \|x - \delta\|$, which is at most $\frac{1}{2} \sum_1^n \|v_i\| \leq \frac{n}{2} \|v_n\|$. On the other hand, it follows from Minkowski's second theorem that if V_n denotes the volume of the Euclidean n -ball,

$$2^n/n! \leq \frac{V_n \|v_1\| \dots \|v_n\|}{\|v_1 \wedge \dots \wedge v_n\|} \leq 2^n$$

and similarly for $v_1 \wedge \dots \wedge v_{n-1}$. Since $\|v_1 \wedge \dots \wedge v_{n-1}\| d(v_n, H) = \|v_1 \wedge \dots \wedge v_n\|$, we conclude that $\|v_n\| \leq 2(n-1)! d(v_n, H) V_{n-1}/V_n$. On the other hand, it is well known that $V_{n-1}/V_n \leq n/\pi$, and so the result follows. \square

Proof of Theorem 4.1 (3). Pick $\delta = 1/6$. By part (2) of the theorem there is a normal subgroup H contained in $S^{O_\varepsilon(\gamma^{2/3})}$ such that G/H has an abelian subgroup of rank and index bounded in terms of ε only. Applying Lemma 4.2, we may pass to a subgroup of bounded index and assume that G/H is abelian of bounded rank. Lemma 4.9 implies that, provided γ is larger than some constant depending on ε only, there is some $n_1 \leq \gamma^{1/3}$ such that S^{2n_1} is an $O_\varepsilon(1)$ -approximate subgroup. Its image \overline{S}^{2n_1} in G/H is therefore also an $O_\varepsilon(1)$ -approximate group, and so there is a subgroup $H_0 \in \overline{S}^{8n_1}$ and a generalised arithmetic progression $P = P(v_1, \dots, v_r; L)$ of bounded rank such that $S^{2n_1} \subset H_0 P \subset S^{O_\varepsilon(n_1)}$ [22]. On replacing H by the pullback of H_0 to G , we may in fact assume that H_0 is trivial. Note that if $\gamma = \text{diam}_S(G)$ is larger than a constant depending on ε only,

$$2n_1 \text{diam}_P(G/H) \leq \text{diam}_{\overline{S}}(G/H) \leq O_\varepsilon(n_1 \text{diam}_P(G/H)).$$

The same comparison holds for the diameter of every group quotient. The generators v_1, \dots, v_r allow us to write G/H as \mathbb{Z}^r/Δ for some lattice Δ ; thus, we write $x \in G/H$ in coordinate form as $x = (x_1, \dots, x_r)$. Since G/H is abelian, $P^n = P(v_1, \dots, v_r; nL)$ for every n . Let $\|\cdot\|$ be the norm on \mathbb{R}^r defined by $\|x\| := \max_1^r |x_i/L_i|$, and define $\ell(x) := \max_1^r \lceil |x_i/L_i| \rceil$. Note that $\|x\| \leq \ell(x) \leq \|x\| + 1$ and that

$$\text{diam}_P(G/H) = \sup_{x \in \mathbb{Z}^r} \inf_{\delta \in \Delta} \ell(x - \delta).$$

We may compare this diameter with the diameter of the Euclidean torus $T := \mathbb{R}^r/\Delta_L$, where Δ_L is the lattice of all vectors of the form $(x_1/L_1, \dots, x_r/L_r)$ with $(x_1, \dots, x_r) \in \Delta$. Indeed,

$$\text{diam}_P(G/H) \leq 1 + \sup_{x \in \mathbb{Z}^r} \inf_{\delta \in \Delta} \|x - \delta\| \leq 2 + \sup_{x \in \mathbb{R}^r} \inf_{\delta \in \Delta} \|x - \delta\|,$$

where we have used that $\inf_{y \in \mathbb{Z}^r} \|x - y\| \leq 1$ for all $x \in \mathbb{R}^r$, because $L_i \geq 1$. Hence

$$\text{diam}_P(G/H) \leq 2 + \sup_{y \in \mathbb{R}^r} \inf_{\delta \in \Delta_L} \|y - \delta\|_\infty \leq 2 + \sup_{y \in \mathbb{R}^r} \inf_{\delta \in \Delta_L} \|y - \delta\|_2 = 2 + \text{diam}(\mathbb{R}^r/\Delta_L),$$

where $\|\cdot\|_2$ is the standard Euclidean norm on \mathbb{R}^r . Lemma 4.14 implies that there is a hyperplane \mathcal{H}_L in \mathbb{R}^r intersecting Δ_L in a lattice such that the diameter of the one-dimensional torus quotient $\mathbb{R}^r/(\mathcal{H}_L + \Delta_L)$ satisfies

$$\text{diam}(\mathbb{R}^r/\Delta_L) \leq \frac{2r r!}{\pi} \text{diam}(\mathbb{R}^r/(\mathcal{H}_L + \Delta_L)).$$

Moreover,

$$\text{diam}(\mathbb{R}^r/(\mathcal{H}_L + \Delta_L)) = \sup_{y \in \mathbb{R}^r} \inf_{\delta \in \mathcal{H}_L + \Delta_L} \|y - \delta\|_2 \leq \sqrt{r} \sup_{y \in \mathbb{R}^r} \inf_{\delta \in \mathcal{H}_L + \Delta_L} \|y - \delta\|_\infty = \sqrt{r} \sup_{x \in \mathbb{R}^r} \inf_{\delta \in \mathcal{H} + \Delta} \|x - \delta\|,$$

where $\mathcal{H} = \{(L_1 x_1, \dots, L_r x_r); (x_1, \dots, x_r) \in \mathcal{H}_L\}$. Writing $\mathcal{H}_\mathbb{Z} := \mathcal{H} \cap \mathbb{Z}^r$, we then have

$$\sup_{x \in \mathbb{R}^r} \inf_{\delta \in \mathcal{H} + \Delta} \|x - \delta\| \leq 1 + \sup_{x \in \mathbb{Z}^r} \inf_{\delta \in \mathcal{H}_\mathbb{Z} + \Delta} \|x - \delta\| \leq 1 + \sup_{x \in \mathbb{Z}^r} \inf_{\delta \in \mathcal{H}_\mathbb{Z} + \Delta} \ell(x - \delta).$$

However, $\sup_{x \in \mathbb{Z}^r} \inf_{\delta \in \mathcal{H}_{\mathbb{Z}} + \Delta} \ell(x - \delta)$ is precisely the diameter of the cyclic quotient $Q := \mathbb{Z}^r / (\mathcal{H}_{\mathbb{Z}} + \Delta)$ with respect to the image of P , and so, combining all of the above, we obtain

$$\text{diam}_{\overline{S}}(G/H) \leq O_{\varepsilon}(n_1 \text{diam}_P(G/H)) \leq O_{\varepsilon}(n_1 \text{diam}_{\overline{P}}(Q)) \leq O_{\varepsilon}(\text{diam}_{\overline{S}}(Q)).$$

□

This ends the proof of Theorem 4.1.

We now record the following simple consequence of Part (1) of Theorem 4.1.

Corollary 4.15 (Almost flatness is a group property). *Let $\varepsilon > 0$ and suppose that G is a finite group with a symmetric generating set S with respect to which G is ε -almost flat. Suppose that E is some other symmetric generating set for G . Then there is $\varepsilon' = \varepsilon'(\varepsilon, |E|) > 0$ and $c_{\varepsilon} > 0$ depending on ε only such that $\text{diam}_E(G) \geq c_{\varepsilon} |G|^{\varepsilon'}$.*

Proof. By Theorem 4.1 G has a quotient Q of comparable diameter containing a nilpotent subgroup of bounded rank, index and nilpotency class, and so the conclusion follows from Proposition 4.7 and Lemma 4.2. □

Remark 4.16. We caution that, in spite of Corollary 4.15, for every $\varepsilon > 0$ there exists $C = C_{\varepsilon}$ such that every finite group G satisfying $|G| \geq C$ has at least one Cayley graph that is ε -almost flat and at least one Cayley graph that is not ε -almost flat. Indeed, taking $S = G \setminus \{x, x^{-1}\}$ for some $x \in G$ gives an almost flat Cayley graph. On the other hand, as is well known and easy to prove, every finite group G of size at most 2^n has a generating set of size at most n and diameter at most n .

4.1. Explicit bounds, the polynomial Freiman-Ruzsa conjecture and an example. The only source of ineffectiveness in our arguments comes from Theorem 2.3, the structural result for approximate subgroups of arbitrary groups proved in [10]. Any explicit bound in terms of K on the number of translates of Γ needed to cover the K -approximate subgroup A would yield to an explicit bound on $n_0(K)$ in Theorem 1.1 and on the index of the abelian subgroup in Theorem 1.3. In this subsection we present an example to show that

- (1) Theorem 1.3 is sharp in the sense that the power ε cannot be allowed to tend to zero without losing the conclusion that the index of the abelian subgroup remains bounded; and
- (2) the bound on the number of translates required in Theorem 2.3, and hence the quantity $n_0(K)$ in Theorem 1.1, is worse than polynomial in K .

We give more precise versions of these statements below as Facts 4.17 and 4.18, as well as giving more detail on the context in which they should be viewed. However, first we establish some of the notation of our example.

Let n be an even integer ≥ 8 , and let $p = p(n)$ be a prime such that $p > n$. We define groups $L_n \geq G'_n \geq G_n$ as follows. First, take

$$L_n = \text{Sym}(n) \rtimes \mathbb{F}_p^n,$$

with the semidirect product defined by the natural action of $\text{Sym}(n)$ on the basis vectors, and write $\pi : L_n \rightarrow \text{Sym}(n)$ for the natural projection. Writing V for the subspace of \mathbb{F}_p^n consisting of those vectors whose coordinates sum to zero, define the subgroup G'_n of L_n by

$$G'_n = \text{Sym}(n) \rtimes V.$$

Finally, define the subgroup G_n of index 2 in G'_n by

$$G_n = \text{Alt}(n) \times V.$$

Let \tilde{S} be the generating set of L_n consisting of the identity and the three elements $(c; 0)$, $(\tau; 0)$, $(1; 1, 0, \dots, 0)$ and their inverses. Here c is the long cycle $(12 \dots n)$ and τ is the transposition (12) . It is clear that L_n is the direct product of G'_n with the central subgroup $Z \cong \mathbb{Z}/p\mathbb{Z}$ made of elements of the form $(1, v)$ in which all coordinates of v are equal, and so we may set S' to be the image of \tilde{S} under projection to G'_n . Finally, let S be the generating set for G_n obtained from S' using Lemma 4.2, and note that $|S| \leq 8$.

The reason for making these definitions is that the groups they describe satisfy the following facts.

Fact 4.17. *For every function $f : \mathbb{N} \rightarrow (0, +\infty)$ going to zero at infinity there exist choices of $p(n)$ such that $\text{diam}_S(G_n) \geq |G_n|^{f(|G_n|)}$ but such that no subgroup of G_n of index $< n/2$ admits a non-trivial abelian quotient.*

Fact 4.18. *For every function $f : \mathbb{N} \rightarrow \mathbb{N}$, there are a choice of $p(n)$, an increasing sequence of integers K_n , and K_n -approximate subgroups A_n in L_n such that, for large enough n , the approximate subgroup A_n is not contained in fewer than $K_n^{\frac{1}{200} \log^{(4)} K_n}$ cosets of any subgroup $\Gamma \leq L_n$ such that Γ has a normal subgroup H contained in $A_n^{f(K_n)}$ with Γ/H solvable.*

Here we have denoted by $\log^{(4)}$ the 4th iteration of log in base e .

Fact 4.17 establishes statement (1) above, while Fact 4.18 corresponds to statement (2). Fact 4.18 should be compared with the classical case of approximate subgroups of abelian groups, where it is expected that only $O(K^{O(1)})$ translates of some coset-progression are needed to cover the approximate subgroup (this is the so-called *polynomial Freiman-Ruzsa type conjecture*, see [21] and references therein). In the abelian case bounds of the form $e^{(\log K)^{O(1)}}$ have been established by Sanders [44].

Remark. After a first draft of this paper circulated, Eberhard [17] adapted the construction of Fact 4.18 to show that for arbitrarily large K there exists a group G with a finite K -approximate subgroup that is not covered by $K^{\log \log K}$ cosets of any subgroup Γ having a finite normal subgroup H with Γ/H solvable. This nicely improves the bound in our Fact 4.18. Note that the example in [17] generates an infinite group, while we work here in the setting of finite groups.

Before we prove Facts 4.17 and 4.18, let us isolate a few helpful statements about the groups L_n , G'_n and G_n . We start by bounding their diameters, which we denote from now on by

$$\gamma = \text{diam}_{\tilde{S}}(L_n), \quad \gamma' = \text{diam}_{S'}(G'_n), \quad \gamma_0 = \text{diam}_S(G_n).$$

Lemma 4.19. *We have*

- (1) $\frac{1}{2}(p-1) \leq \gamma \leq np + O(n^2)$;
- (2) $\frac{1}{2}(p^{1-1/n} - 1) \leq \gamma' \leq np + O(n^2)$;
- (3) $\frac{1}{10}p^{1-1/n} \leq \gamma_0 \leq np + O(n^2)$.

Proof. Item (3) follows from item (2) by Lemma 4.2. The upper bound of (2) trivially follows from the upper bound of (1). To prove the lower bound of (1) note that the elements of \tilde{S}^R are all in $\text{Sym}(n) \times [-R, R]^n$, and hence that $(2\gamma + 1)^n \geq p^n$. To prove the lower bound of (2), note that the projections of these elements to G'_n lie in $\text{Sym}(n) \times B_n$ for some subset B_n of V with size at most $(2R + 1)^n$, and hence

that $(2\gamma + 1)^n \geq p^{n-1}$. For the upper bound of (1), note that $O(n^2)$ steps are enough to cover $Sym(n)$ (see [37], for example), while p steps are enough to obtain the line $(1; \mathbb{F}_p, 0, \dots, 0)$ and thus (conjugating by powers of c) we obtain each coordinate line in $p + n$ steps and thus all of \mathbb{F}_p^n in $n(p + n)$ steps. Hence the diameter of L_n with respect to \tilde{S} is at most $np + O(n^2)$. \square

Lemma 4.20. *The only subgroups of G'_n of index $< n$ are G'_n itself and G_n .*

Proof. It is well known that the only proper subgroup of $Sym(n)$ with index $< n$ is $Alt(n)$, so if Γ is a subgroup of G'_n with index $< n$ then $\pi(\Gamma)$ contains $Alt(n)$. Since $p > n$, the projection π is not injective in restriction to Γ . However, $\ker \pi \cap \Gamma$ is invariant under the action of $Alt(n)$. It is well known that $Sym(n)$ and $Alt(n)$ act irreducibly on V , and so this implies that $\ker \pi \cap \Gamma = V$, and hence that $Alt(n) \times V \subset \Gamma$. \square

Proof of Fact 4.17. Since $|G_n| \leq n!p^{n-1} \leq (np)^{n-1} \leq p^{2(n-1)}$, Lemma 4.19 implies that $\gamma_0 \gg |G_n|^{\frac{1}{2n}}$. In light of Lemma 4.20, therefore, Fact 4.17 is satisfied upon setting $p = p(n)$ sufficiently large. \square

We now pass to the proof of Fact 4.18. We use the following lemma.

Lemma 4.21. *Let $\beta \in (0, 1)$ and suppose that $(\log \log n)^{-\beta} \leq d < 1$. Write $t = d^{3/d}$ and $b = 9/t^2$. Then, provided $n \geq n(\beta)$, every subgroup $H \leq Sym(n)$ of order at least $d^n n!$ has a subgroup H_0 with $[H : H_0] \leq b$ that stabilises a subset $\Omega_0 \subset \{1, \dots, n\}$ of size at least $\frac{d}{3b}n$ on which it acts as $Sym(\Omega_0)$ or $Alt(\Omega_0)$.*

We prove Lemma 4.21 shortly. First, let us use it to prove Fact 4.18.

Proof of Fact 4.18. By Lemma 4.19 we have $p^{1-\frac{1}{n}} \leq 2\gamma + 1 \leq O(np)$. Setting $\varepsilon = \frac{1}{4n}$ and $\delta = \frac{1}{2}$, and taking $K_n = 5^{\frac{2}{\varepsilon\delta}} = 5^{16n}$, Lemmas 4.9 and 2.2 therefore give a scale $k \in [\gamma^{1/4}, \gamma^{1/2}]$ such that $A_n := \tilde{S}^{2k}$ is a K_n -approximate subgroup.

Assume that there are subgroups $H \leq \Gamma$ in L_n , such that H is normal in Γ , contained in $\tilde{S}^{2kf(K_n)}$, Γ/H is solvable, and \tilde{S}^{2k} is contained in fewer than $K_n^{\alpha \log \log \log \log K_n}$ cosets of Γ for some fixed $\alpha > 0$ to be determined shortly. We may choose p large enough that $K_n^{\alpha \log \log \log \log K_n} < \gamma^{1/4} < 2k$, which by Lemma 2.7 implies that Γ has index at most $K_n^{\alpha \log \log \log \log K_n}$ in L_n . Now set $B = 5^{16}$, so that $K_n = B^n$, and set $\alpha := \frac{1}{3 \log B}$. Note that $\alpha \geq \frac{1}{200}$.

If n is large enough, we then see that the index $K_n^{\alpha \log \log \log \log K_n}$ is bounded above by d^{-n} , where $d = (\log \log n)^{-\frac{1}{2}}$. We may thus apply Lemma 4.21 to the subgroup $\pi(\Gamma) \leq Sym(n)$, because its index is at most d^{-n} with $d = (\log \log n)^{-\frac{1}{2}}$. We get a subgroup $\Gamma_0 \leq \Gamma$ of index at most b such that Γ_0 fixes a subset Ω_0 of $\{1, \dots, n\}$ of size at least $\frac{d}{3b}n$, where it acts as $Sym(\Omega_0)$ or $Alt(\Omega_0)$. Note that $b \leq \log n$, and that $\frac{d}{3b}n$ is at least $\frac{n}{(\log n)^2}$, when n is large. Let $H_0 = H \cap \Gamma_0$, and π_{Ω_0} the natural projection from $Sym(\Omega_0) \times \mathbb{F}_p^{\Omega_0}$ to $Sym(\Omega_0)$. There is a natural restriction homomorphism from Γ_0 to $Sym(\Omega_0) \times \mathbb{F}_p^{\Omega_0}$, which consists in forgetting the entries outside of Ω_0 . Let $\bar{\Gamma}_0, \bar{H}_0$ be the images of Γ_0, H_0 under this map.

Now observe that if n is large enough $\bar{\Gamma}_0$ must contain $Alt(\Omega_0) \times V_{\Omega_0}$, where V_{Ω_0} is the subspace of $\mathbb{F}_p^{\Omega_0}$ of vectors whose coordinates sum to 0. Indeed, to see this it is enough to show that $|\bar{\Gamma}_0| > p|\Omega_0|!$, because then π_{Ω_0} will not be injective on Γ_0 and there will be an element in the kernel whose coordinates are not all equal, so that the span of its orbit under $Alt(\Omega_0)$ will contain V_{Ω_0} . On the other hand if $|\bar{\Gamma}_0| \leq p|\Omega_0|!$, then $|\Gamma_0| \leq p|\Omega_0|!p^{n-|\Omega_0|}(n-|\Omega_0|)! \leq n!p^n p^{1-|\Omega_0|}$, so that $p^n n! = |L_n| \leq b|\Gamma_0| |L_n : \Gamma| \leq bn!p^n p^{1-|\Omega_0|} K_n^{\alpha \log^{(4)} n}$,

which implies that $p^{|\Omega_0|-1} \leq e^{n \log n}$ if n is large enough. But $|\Omega_0| \geq n/(\log n)^2$ and $p \geq e^{n^2}$, so this is impossible unless n is bounded.

The group $\pi_{\Omega_0}(\overline{H_0})$ is a normal subgroup of $Sym(\Omega_0)$. Hence it is either trivial, or must contain $Alt(\Omega_0)$. If it contains $Alt(\Omega_0)$, then, since $\overline{H_0}$ is normal in $\overline{\Gamma_0}$, and $\overline{\Gamma_0}$ contains V_{Ω_0} , $\overline{H_0}$ must contain $[V_{\Omega_0}, Alt(\Omega_0)] = V_{\Omega_0}$, hence $\overline{H_0}$ contains $Alt(\Omega_0) \times V_{\Omega_0}$. However we observed in the proof of Lemma 4.19, the vector coordinates in \tilde{S}^R are contained in the interval $[-R, R]$. Since H_0 lies in $\tilde{S}^{2kf(K_n)}$ we must have $p \leq 2kf(K_n) = O(f(K_n)\sqrt{np})$, which we may rule out by choosing p large enough in terms of n . It follows that $\pi_{\Omega_0}(\overline{H_0})$ is trivial. But this contradicts the fact $\pi_{\Omega_0}(\overline{\Gamma_0})/\pi_{\Omega_0}(\overline{H_0})$ is solvable (being a homomorphic image of Γ_0/H_0), because $Alt(\Omega_0)$ is not solvable if n is large enough. This contradiction ends the proof of Fact 4.18. \square

The proof of Lemma 4.21 requires the following two preliminary steps.

Lemma 4.22. *There is $n_0 \in \mathbb{N}$ such that if $d \in [0, 1]$, $n \geq n_0$, and $H \leq Sym(n)$ is a subgroup of order at least $d^n n!$, then H has an orbit of size at least $\frac{d}{3}n$.*

Proof. Let n_1, \dots, n_s be the size of the orbits of H . We have $H \leq Sym(n_1) \times \dots \times Sym(n_s)$, and hence $|H| \leq n_1! \dots n_s! \leq n_1^{n_1} \dots n_s^{n_s}$. If all the n_i are at most $\frac{d}{3}n$, then we get $|H| \leq d^n (\frac{n}{3})^n$, and hence $n! \leq (\frac{n}{3})^n$, which implies that n is bounded. \square

Lemma 4.23. *There is $m_0 \in \mathbb{N}$ such that if $t \in [0, 1]$, $m \geq m_0$, and $H \leq Sym(m)$ is a transitive subgroup of order at least $t^m m!$, then the maximal number of imprimitivity blocks of H is at most $(3/t)^2$.*

Proof. Let a be the maximal possible number of blocks of imprimitivity. If H is not primitive, then a is a proper divisor of m and thus $a \leq \frac{m}{2}$. Moreover H is contained in the wreath product of $Sym(a)$ with $Sym(\frac{m}{a})$, so that $|H| \leq a!((\frac{m}{a})!)^a$. Hence

$$t^m m! \leq a^a \left(\frac{m}{a}\right)^m \leq a^{m/2} \left(\frac{m}{3}\right)^m \left(\frac{3}{a}\right)^m.$$

But $(\frac{m}{3})^m \leq m!$ as soon as m is large enough. So this yields $t^m \leq (\frac{9}{a})^{m/2}$ and the result follows. \square

Proof of Lemma 4.21. By Lemma 4.22, H has an orbit Ω of size $m \geq \frac{d}{3}n$. The action of H on Ω may not be primitive. Let Ω_0 be a minimal block of imprimitivity, and H_0 the stabiliser of Ω_0 in H . Let K be the restriction of H to Ω and K_0 the restriction of H_0 to Ω_0 . Note that $[Sym(\Omega) : K] \leq [Sym(n) : H] \leq d^{-n} \leq t^{-m}$. We may thus apply Lemma 4.23 to K and get that the maximal number a of blocks of imprimitivity is most $b = 9t^{-2}$. In particular $[H : H_0] \leq b$. Similarly $[Sym(\Omega_0) : K_0] \leq [Sym(n) : H_0] \leq bd^{-n} \leq bt^{-m} \leq t^{-2m}$, and so $|K_0| \geq (\frac{m}{a})! t^{2m}$.

On the other hand, since the action of K_0 on Ω_0 is primitive, Bochert's theorem [16, Theorem 3.3B] implies that unless K_0 acts as $Sym(\Omega_0)$ or $Alt(\Omega_0)$ on Ω_0 it satisfies $|K_0| \leq \frac{(m/a)!}{((m/a+1)/2)!}$. Combining these two inequalities yields $(m/4a)^{m/4a} \leq t^{-2m}$, thus $m \leq 4at^{-8a}$, and hence $n \leq \frac{12a}{d}t^{-8a}$. This implies that n is bounded in terms of β if $d \geq 1/(\log \log n)^\beta$. \square

4.2. Doubling close to the diameter and a question. We end this section by asking a question. Lemma 4.9 shows that the almost flat condition implies doubling at some controlled scale. Moreover, doubling at some scale is enough to apply Theorem 2.3, yielding the following.

Proposition 4.24. *Consider a finite group G with symmetric generating set S . Let $K \geq 1$. There is an integer $n_1 = n_1(K)$ depending only on K such that if n is such that $n \geq n_1$ and*

$$|S^{2n+1}| \leq K|S^n|,$$

then S^{8n} contains a normal subgroup H of G such that G/H has a nilpotent subgroup whose index, rank and nilpotency class are bounded in terms of K only.

Proof. By Lemma 2.10, $|S^{5n}| \leq C(K)|S^n|$ for some constant $C(K)$ depending on K only. Now apply Proposition 2.9. \square

Note that the quotient G/H is non-trivial whenever $n < \frac{1}{8}\gamma$, where γ is the diameter of G with respect to S .

Question: Assume that $|S^{2n+1}| \leq K|S^n|$ holds for some $n \in [\frac{1}{8}\gamma, (1 - \varepsilon)\gamma]$, say. Is it true that G has a normal subgroup $H \neq G$ such that G/H has a nilpotent subgroup whose index, rank and nilpotency class are bounded in terms of K, ε only?

5. CHEEGER CONSTANT AND SPECTRAL GAP

In this section we prove Corollaries 1.5, 1.6, 1.7 and 1.8. We first recall some definitions.

Let \mathcal{G} be a finite connected k -regular graph. Given a set $A \subset V(\mathcal{G})$, denote by ∂A the *edge boundary* of A , defined as the set of edges between A and $V(\mathcal{G}) \setminus A$. The *Cheeger constant* $h(\mathcal{G})$ of a graph \mathcal{G} is defined by

$$h(\mathcal{G}) = \inf \left\{ \frac{|\partial A|}{|A|} : A \subset V(\mathcal{G}), |A| \leq \frac{1}{2}|\mathcal{G}| \right\}.$$

Recall that \mathcal{G} is called an ε -*expander* if $h(\mathcal{G}) \geq \varepsilon$; see the survey [29] for detailed background. Expansion may be considered in terms of the spectrum of the Laplacian operator on the graph. Specifically, the Laplacian Δ on our connected k -regular graph \mathcal{G} is an operator on the finite-dimensional vector space $\ell^2(\mathcal{G})$, and is defined by

$$\Delta f(x) = kf(x) - \sum_{y \sim x} f(y),$$

where $y \sim x$ means that y and x are connected by an edge in the graph. The Laplacian is self-adjoint on $\ell^2(\mathcal{G})$, and so its eigenvalues

$$\lambda_0(\mathcal{G}) = 0 < \lambda_1(\mathcal{G}) \leq \dots \leq \lambda_{|\mathcal{G}|-1}(\mathcal{G})$$

are real. The link between this and expansion as defined above is given by the *discrete Cheeger–Buser inequality*, which states that

$$\frac{1}{2}h(\mathcal{G})^2 \leq \lambda_1(\mathcal{G}) \leq 2k \cdot h(\mathcal{G}). \tag{5.1}$$

See the book [37] for a proof of these inequalities and the references therein for further historical background. Furthermore, we record the following well-known additional inequalities, which together with (5.1) imply (1.3).

Lemma 5.1 (Diameter versus expansion). *For every finite k -regular graph \mathcal{G} we have*

$$h(\mathcal{G}) \leq \frac{4k \log |\mathcal{G}|}{\text{diam}(\mathcal{G})}.$$

If \mathcal{G} is vertex transitive then

$$\frac{1}{2 \text{diam}(\mathcal{G})} \leq h(\mathcal{G}).$$

Lemma 5.1 is both standard (see e.g. [14] for a variant) and straightforward, but for completeness we present a short proof. Just as the proof of the Cheeger-Buser inequalities, this is based on the variational characterisation of λ_1 and h , which gives

$$\lambda_1(\mathcal{G}) = \inf \left\{ \frac{\|\nabla f\|_2^2}{\|f\|_2^2} : \int f = 0 \right\}; \quad (5.2)$$

$$h(\mathcal{G}) = \inf \left\{ \frac{\|\nabla f\|_1}{\inf_{m \in \mathbb{R}} \|f - m\|_1} \right\}. \quad (5.3)$$

Here the infimums are taken over all real valued functions on the set of vertices of the graph. The norm $\|f\|_2$ is the ℓ^2 norm, while $\|f\|_1$ denotes the ℓ^1 norm, defined as usual by $\|f\|_1 = \sum_{x \in V(\mathcal{G})} |f(x)|$. Finally, ∇f is the non-negative function defined on the set of edges by $\nabla f(e) = |f(e^+) - f(e^-)|$, where e^+ and e^- are the vertices of the edge e . The identity (5.2) may be found in [37]. Since we do not have a readily available reference for (5.3), we provide a proof below. First let us use it to prove Lemma 5.1.

Proof of Lemma 5.1. Abbreviate $h = h(\mathcal{G})$, the Cheeger constant of \mathcal{G} . Given $x \in \mathcal{G}$ and $r > 0$, write $B_r(x)$ for the closed ball of radius r centred at x for the graph metric on \mathcal{G} , and write $S_r(x) = B_r(x) \setminus B_{r-1}(x)$ for the sphere of radius r centred at x .

By definition of h , if $|B_r(x)| \leq \frac{1}{2}|\mathcal{G}|$ then we have $|\partial B_r(x)| \geq h|B_r(x)|$. Since each vertex of $S_{r+1}(x)$ is incident with at most k edges in $\partial B_r(x)$ we have $|S_{r+1}(x)| \geq k^{-1}|\partial B_r(x)|$, which implies that

$$|B_{r+1}(x)| \geq \left(1 + \frac{h}{k}\right) |B_r(x)|.$$

Noting that $\frac{h}{k} \leq 1$, and hence that $1 + \frac{h}{k} \geq e^{\frac{h}{2k}}$, it follows immediately that $|B_r(x)| \geq \min \left\{ \frac{1}{2}|\mathcal{G}|, e^{rh/2k} \right\}$. In particular, if $r \frac{h}{k} \geq 2 \log |\mathcal{G}|$ then the balls $B_r(x)$ and $B_r(y)$ have size at least $\frac{1}{2}|\mathcal{G}|$, and hence non-trivial intersection for every choice of x and y . This proves the first inequality of the lemma.

Now assume that \mathcal{G} is vertex transitive. Set $\gamma := \text{diam}(\mathcal{G})$, and write $G = \text{Aut}(\mathcal{G})$, the (finite) group of automorphisms of the graph. Pick a base vertex x_0 in the graph and connect every other vertex y to x_0 by a geodesic path, say $x_0 = y_0, y_1, \dots, y_n = y$, noting that $n \leq \gamma$. For any real-valued function f on the vertices of \mathcal{G} we have

$$\sum_{\sigma \in G} |f(\sigma(x_0)) - f(\sigma(y))| \leq \sum_1^n \sum_{\sigma \in G} |f(\sigma(y_{i-1})) - f(\sigma(y_i))| \leq \|\nabla f\|_1 \sum_1^n |G_{e_i}|,$$

where G_{e_i} is the stabiliser of the edge e_i between y_{i-1} and y_i . Clearly G_{e_i} has a subgroup of index at most 2 that fixes y_i , and hence $|G_{e_i}| \leq 2|G|/|\mathcal{G}|$. Summing over all vertices y of \mathcal{G} , it follows that

$$\inf_{m \in \mathbb{R}} \|f - m\|_1 \leq \frac{1}{|G|} \sum_{\sigma \in G} \|f - f(\sigma(x_0))\|_1 = \frac{1}{|G|} \sum_{\sigma \in G} \sum_{y \in \mathcal{G}} |f(\sigma(x_0)) - f(\sigma(y))| \leq 2\gamma \|\nabla f\|_1$$

and so the lemma follows from (5.3). \square

Proof of (5.3). To see that $h(\mathcal{G})$ dominates the right hand side, we simply pick the function $f := 1_A - 1_{\mathcal{G} \setminus A}$, where A is a subset of vertices of size at most $|\mathcal{G}|/2$ realising the infimum in the definition of $h(\mathcal{G})$. Then note that $\|\nabla f\|_1 = 2|\partial A|$ and that $\|f - m\|_1$ is minimised by taking $m = -1$, which makes its value equal to $2|A|$. Now to verify the opposite inequality, given a function f on the vertices consider the subsets of vertices $\Omega_t := \{x : f(x) \geq t\}$ and let $m := \inf\{t \in \mathbb{R}; |\Omega_t| \leq |\mathcal{G}|/2\}$. Note that if $t > m$, then $|\Omega_t| \leq |\mathcal{G}|/2$, while if $t \leq m$, then $|\mathcal{G} \setminus \Omega_t| \leq |\mathcal{G}|/2$. On the other hand, $\|f - m\|_1 = \sum_{x \in \mathcal{G}} [(f - m)^+(x) + (m - f)^+(x)]$, where $u^+ = \max\{0, u\}$ is the positive part of the real number u . Hence

$$\|f - m\|_1 = \int_{-\infty}^m |\mathcal{G} \setminus \Omega_t| dt + \int_m^{+\infty} |\Omega_t| dt \leq \frac{1}{h} \int_{-\infty}^{+\infty} |\partial \Omega_t| dt,$$

since $\partial \Omega_t = \partial(\mathcal{G} \setminus \Omega_t)$. On the other hand, observe that for every edge e we have $\nabla f(e) = \int_{-\infty}^{+\infty} 1_{\partial \Omega_t}(e) dt$, and so $\|\nabla f\|_1 = \int_{-\infty}^{+\infty} |\partial \Omega_t| dt$. The inequality follows. \square

We now pass to the proofs of Corollaries 1.5 and 1.6.

Proof of Corollary 1.5. We have $\gamma^2 \geq \frac{1}{8\lambda_1}$ by the left hand side of (1.3), and so if $\lambda_1 \leq 2^{-9}$ then $\gamma \geq 8$, and hence $\gamma^3 \geq 1/\lambda_1$. On the other hand, $\gamma \geq 1/2h$ by Lemma 5.1, so if $h \leq \frac{1}{4}$ then $\gamma \geq 2$, and hence $\gamma^2 \geq 1/h$. \square

To prove Corollary 1.6 we use the following general lemma.

Lemma 5.2. *Let \mathcal{G} be the Cayley graph of a finite group G with symmetric generating set S and diameter γ . Then*

$$\lambda_1(\mathcal{G}) \leq \frac{9|S|}{\gamma^2} \frac{|G|}{|S^{\gamma/3}|}$$

Proof. Pick $a, b \in G$ such that $d(a, b) = \gamma$, and let $f(x) := d(x, a) - d(x, b)$, where $d(x, y)$ is the graph distance in \mathcal{G} . Note that f has zero mean and that $|\nabla f(e)| \leq 2$ for every edge e . Hence $\|\nabla f\|_2^2 \leq 2|S||G|$. On the other hand $|f(x)| \geq \gamma/3$ on the balls of radius $\gamma/3$ around a and around b . Hence $\|f\|_2^2 \geq 2(\gamma/3)^2 |S^{\gamma/3}|$. The result therefore follows from (5.2). \square

Proof of Corollary 1.6. Set $K = 5^{4/\varepsilon}$, and let $n_0 = n_0(K)$ be the constant appearing in the conclusion of Theorem 1.1. Provided $\gamma \geq n_0^4$, Lemma 4.9 implies that there is n satisfying $n_0 \leq n \leq \gamma^{1/2}$ such that $|S^{5n}| \leq |S^n|$. Theorem 1.1 therefore implies that $|G| = |S^\gamma| \leq O_K(1)|S^{\gamma/3}|$. The bound (1.5) therefore follows from Lemma 5.2, and the bounds (1.4) then follow from (1.3) and (1.5). \square

Remark 5.3 (Cheeger constant versus λ_1). We have shown that almost flat groups satisfy the strong form of the Buser inequality; that is to say λ_1 is comparable to the square of the Cheeger constant. This property does not characterise almost flatness. Indeed it is shared for example by the Cayley graph of the lamplighter group based on a finite cyclic group (see [40]), which is by no means almost flat. By contrast, there are Cayley graphs of finite groups at the opposite extreme permitted by (1.3), which is to say λ_1 is comparable to the Cheeger constant itself, not its square. This happens, for example, for the lamplighter group based on $\mathbb{Z}/n^d\mathbb{Z}$ for any given $d \geq 3$. For all $d \geq 1$ the Cheeger constant is comparable to $1/\gamma$ (almost extremal sets are obtained by setting one fixed light off), but λ_1 is comparable to $1/\gamma^2$ if $d = 1$, to $1/\gamma \log \gamma$ if $d = 2$, and to $1/\gamma$ if $d \geq 3$ [40].

Remark 5.4 (Expansion and number of generators). Almost flat groups are not expanders, since their λ_1 values are comparable to $1/\gamma^2$ by Corollary 1.6. In fact, in order to construct a Cayley graph of a finite group G with a nilpotent subgroup of bounded index, rank and nilpotency class with a positive lower bound on λ_1 , one would need at least a constant times $\log |G|$ generators. This follows immediately from Lemma 4.13 and the corresponding result for abelian groups, which is proved in [38, Propositions 3.9 & 3.10]. Note that this is extremal in some sense, because according to Alon and Roichman [1, Theorem 1], given $\varepsilon > 0$, every finite group G has a generating set S of size at most $C_\varepsilon \log |G|$ with ε -expanding generating set.

We now pass to the proof of Corollaries 1.7 and 1.8.

Proof of Corollary 1.7. We fix a finite symmetric generating set for Γ and consider the Cayley graphs of the quotients Γ/Γ_n with respect to this generating set. The second inequality in Lemma 5.1 implies that $\text{diam}(\Gamma/\Gamma_n) \geq \frac{1}{2}|\Gamma/\Gamma_n|^\varepsilon$. Being finitely generated, Γ has only finitely many subgroups of index at most M for any given $M \geq 1$, and so it follows that Γ/Γ_n is $\frac{\varepsilon}{2}$ -almost flat as soon as n is large enough. It then follows from Theorem 1.3 that Γ/Γ_n has a quotient of diameter at least $\frac{1}{2} \text{diam}(\Gamma/\Gamma_n)$ that has an abelian subgroup whose index is bounded independently of n . This implies that there is a finite-index subgroup $H \leq \Gamma$ that maps onto abelian groups of unbounded order, which in particular means that $H/[H, H]$ is infinite. Since $H/[H, H]$ is also finitely generated and abelian, it admits a surjective homomorphism onto \mathbb{Z} . \square

The proof of Corollary 1.8, the manifold version of Corollary 1.7, is a simple application of the Brooks-Burger transfer principle, for which we refer the reader to the paper [39] and the appendix to the lecture notes [5], as well as the original sources [12, 13]. This states that given a compact connected Riemannian manifold M , and a finite symmetric generating set S of its fundamental group, there are positive constants c_1, c_2 such that for every finite-degree Riemannian cover $M' \rightarrow M$, we have

$$c_1 \lambda_1(M') \leq \lambda_1(\pi_1(M)/\pi_1(M'), S) \leq c_2 \lambda_1(M') \quad (5.4)$$

where the $\lambda_1(M')$ denotes the first non-zero eigenvalue of the Laplace-Beltrami operator on M' and $\lambda_1(\pi_1(M)/\pi_1(M'), S)$ that of the corresponding Schreier graph. Similarly,

$$c_1 h(M') \leq h(\pi_1(M)/\pi_1(M'), S) \leq c_2 h(M'). \quad (5.5)$$

It is worth emphasising that the constants c_1, c_2 here depend only on the base manifold M and choice of S , but not on M' .

Proof of Corollary 1.8. This is just the combination of (5.5) and Corollary 1.7. \square

6. MODERATE GROWTH AND MIXING TIMES

In this section we prove Corollary 1.9, which shows that almost flat groups and groups with moderate growth coincide. We then discuss mixing-time estimates for random walks on almost flat groups, and prove Corollary 1.10 and Theorem 1.11.

We start with a crude, yet general, bound on the diameter of a Cayley graph (which is meaningful only when S is large, since otherwise it is worse than the trivial bound $\text{diam}_S(G) \leq |G|$).

Lemma 6.1. *Let G be a finite group and S a symmetric generating set. Then*

$$\text{diam}_S(G) \leq 2(|G|/|S|)^{\frac{7}{4}}.$$

Proof. Let m be the least integer such that $|S^{2^{m+1}}| \leq \frac{3}{2}|S^{2^m}|$, noting that this implies that $|G|/|S| \geq (\frac{3}{2})^m$. Freiman's lemma [19] shows that the bound $|S^{2^{m+1}}| \leq \frac{3}{2}|S^{2^m}|$ implies that S^{2^m} is contained in a coset of a subgroup $H \leq G$ satisfying $|H| \leq \frac{3}{2}|S^{2^m}|$. Since $1 \in S$, we in fact have $S^{2^m} \subset H$, and since S generates G this implies that $H = G$, and in particular that $|S^{2^m}| \geq \frac{2}{3}|G|$. This in turn implies that $S^{2 \cdot 2^m} = G$, and hence that $\text{diam}_S(G) \leq 2 \cdot 2^m$. The result follows since $(\frac{3}{2})^{\frac{7}{4}} \geq 2$. \square

Proof of Corollary 1.9. In the first direction, suppose that (G, S) has (A, d) -moderate growth, which is to say that (1.6) holds for all $n \geq 1$. The case $n = 1$ implies that $\gamma \geq A^{-1/d}(|G|/|S|)^{1/d}$. However, $\gamma \leq 2(|G|/|S|)^2$ by Lemma 6.1, so (G, S) is $\frac{1}{2A}$ -almost flat whenever $\gamma \geq 2A^4$.

In the converse direction, suppose that (G, S) is ε -almost flat, let $K = 5^{\frac{4}{\varepsilon}}$, and let $n_0 = n_0(K)$ be the integer given by Theorem 1.1. If $\gamma < n_0^4$ then the almost flat assumption implies that for every $n \leq \gamma$ we have

$$|S^n| \geq |S| \geq \left(\frac{n}{n_0^4}\right)^{\varepsilon-1} |S| \geq \frac{1}{n_0^{4\varepsilon-1}} \left(\frac{n}{\gamma}\right)^{\varepsilon-1} |G|,$$

and so G has $(n_0^{4\varepsilon-1}, \varepsilon^{-1})$ -moderate growth. We may therefore assume that $\gamma \geq n_0^4$. Lemma 4.9 therefore implies that there is some n satisfying $n_0 \leq n \leq \gamma^{1/2}$ such that $|S^{5n}| \leq K|S^n|$, and so Theorem 1.1 implies that $|S^{5n}| \leq \theta(K)^5|S^n|$ for all $n \geq \gamma^{1/2}$. Given $n \geq \gamma^{1/2}$, therefore, writing r for the smallest integer such that $5^r n \geq \gamma$ we have $|S^{5^r n}| \leq \theta(K)^{5^r}|S^n|$ and $\theta(K)^{5^r} \leq \theta(K)^{5+5\log_5(\gamma/n)} = \theta(K)^5(\gamma/n)^{5\log_5(\theta(K))}$, and so

$$|S^n| \geq \frac{1}{A} \left(\frac{n}{\gamma}\right)^d |G| \tag{6.1}$$

holds with $A = \theta(K)^5$ and $d = 5\log_5 \theta(K)$. When $n \leq \gamma^{1/2}$, on the other hand, the almost flat assumption implies that $|S|/|G| \geq 1/\gamma^{\varepsilon-1}$, and so

$$\frac{|S^n|}{|G|} \geq \frac{|S|}{|G|} \geq \frac{1}{\gamma^{\varepsilon-1}} = \left(\frac{\gamma^{1/2}}{\gamma}\right)^{2\varepsilon-1} \geq \left(\frac{n}{\gamma}\right)^{2\varepsilon-1}.$$

Thus (6.1) holds with $d = 2\varepsilon^{-1}$ and $A = 1$. \square

The main result of the work [15] of Diaconis and Saloff-Coste shows that groups with moderate growth have mixing time quadratic in the diameter. Inspired by this, we spend the rest of this section discussing various other sufficient conditions on the Cayley graph of a finite group for its mixing time to be quadratic in the diameter.

We should emphasise that the constants in the results of Diaconis and Saloff-Coste are effective in the moderate-growth parameters of the group in question; the ineffectiveness of the constants appearing in the results of this section arise only due to our reliance on Theorem 2.3.

By definition, the L^p mixing time T_p is the first integer n such that $\|\mu_S^n - \mu_G\|_p \leq \frac{1}{10}\|\mu_G\|_p$, where μ_S is the uniform probability measure on the generating set S and μ_G the uniform probability measure on G . We view them as functions on G taking non-negative values. Of course, the choice of factor $\frac{1}{10}$ is purely a matter of convention. It is customary also to define the relaxation time T_{rel} as the inverse spectral gap, or more specifically $T_{rel} = 1/(1 - \beta_S)$, where β_S is the norm of the operator $1 - \Delta/|S|$. For simplicity, we will assume that $\beta_S = 1 - \lambda_1/|S|$, which amounts to assuming that Δ has no eigenvalue too close to $2|S|$. This is the case if $\lambda_1 \leq 2$ and $1 \in S$.

The mixing times T_p dominate the relaxation time T_{rel} and are non-decreasing in p , and in fact there are essentially only two mixing times, T_1 and T_∞ , because T_p is comparable to T_∞ if $p > 1$. The example of the lamplighter group with cyclic base shows that T_∞ really can be much larger than T_1 [25, 40]. Writing $\gamma = \text{diam}_S(G)$, we have $\frac{1}{2} \text{diam}_S(G) \leq T_p \leq O_{|S|}(\gamma^3)$, and the same lamplighter example shows that the upper bound can indeed be reached for T_∞ [25, 40]. It is not known, however, whether T_1 is always $O_{|S|}(\gamma^2)$. We summarise these facts and some others in the following proposition, which is well known and encapsulates the basic properties of mixing times on Cayley graphs.

Proposition 6.2 (Basic facts on mixing times). *Let G be a finite group with symmetric generating set S containing 1 and denote by γ the diameter of the Cayley graph. Assume that $\lambda_1 \leq 2$, and set $\beta_S := 1 - \lambda_1/|S|$. Then, for $p \in [1, \infty]$ and $n \in \mathbb{N}$, the following conditions hold.*

- (1) *The quantity $\|\mu_S^{(n)} - \mu_G\|_p$ is non-increasing in n .*
- (2) *The quantity $\|\mu_S^{(n)} - \mu_G\|_p / \|\mu_G\|_p$ is non-decreasing in p , and hence so is T_p .*
- (3) *We have $\beta_S^n \leq \|\mu_S^{(n)} - \mu_G\|_p / \|\mu_G\|_p$.*
- (4) *We have*

$$\frac{\|\mu_S^{(2n)} - \mu_G\|_p}{\|\mu_G\|_p} \leq \left(\frac{\|\mu_S^{(n)} - \mu_G\|_2}{\|\mu_G\|_2} \right)^2.$$

- (5) *We have $\|\mu_S^{(n)} - \mu_G\|_2 \leq \beta_S^n$.*
- (6) *For every $p > 1$ there exists $C(p) > 0$ such that $T_\infty \leq C(p)T_p$. For $p \geq 2$ we may take $C(p) = 2$.*
- (7) *We have $T_p \geq \frac{1}{2}\gamma$ for all p , and $T_\infty \geq \gamma$.*
- (8) *We have $T_2 \leq 8|S|\gamma^2 \log |G| \leq 8|S| \log |S|\gamma^3$.*
- (9) *We have $T_{rel} \leq \min\{T_1, 8|S|\gamma^2\}$.*

We include the straightforward proof for the reader's convenience.

Proof. Condition (1) holds with $\|\cdot\|_p$ replaced by any G -invariant norm $\|\cdot\|$, since the triangle inequality implies that $\|\mu_S * f\| \leq \|f\|$ for every $f : G \rightarrow \mathbb{R}$. To prove (2), assume that $p < p' < \infty$ and note that the function $u \mapsto u^{p'/p}$ is convex, and hence that we may apply Jensen's inequality in the form $\mu_G(f^{p'/p}) \geq (\mu_G(f))^{p'/p}$ to the function $f(x) = \mu_S^{(n)}(x) - \mu_G(x)$, from which (2) follows (the case $p' = \infty$ being a trivial calculation).

For the bound (3), let q be such that $1/p + 1/q = 1$, so that by duality we have

$$\|\mu_S^{(n)} - \mu_G\|_p = \sup_{\|f\|_q \leq 1} |\mu_S^{(n)}(f) - \mu_G(f)|. \quad (6.2)$$

Let f be an eigenfunction of the Laplacian with eigenvalue λ_1 , translated so that $f(1) = \|f\|_\infty$ and scaled so that $\|f\|_q = 1$. Note that $\mu_G(f) = 0$, and also that $\mu_S^{(n)} * f = \beta_S^n f$, and hence that $\mu_S^{(n)}(f) = \beta_S^n f(1) = \beta_S^n \|f\|_\infty$. It therefore follows from (6.2) that $\|\mu_S^{(n)} - \mu_G\|_p \geq \beta_S^n \|f\|_\infty \geq \beta_S^n \|f\|_q |G|^{-1/q} = \beta_S^n \|\mu_G\|_p$, and so (3) follows.

It is enough to prove (4) for $p = \infty$, but this follows from the Cauchy-Schwarz inequality applied to the obvious identity $(\mu_S^{(2n)} - \mu_G)(x) = (\mu_S - \mu_G)^{(2n)}(x) = \sum_y (\mu_S - \mu_G)^{(n)}(xy^{-1})(\mu_S - \mu_G)^{(n)}(y)$, which is valid for every $x \in G$, and yields $\|\mu_S^{(2n)} - \mu_G\|_\infty \leq \|\mu_S^{(n)} - \mu_G\|_2^2$, and hence (4) after observing that $\|\mu_G\|_\infty = \|\mu_G\|_2^2$.

To see (5), write P_{μ_S} for the operator of left convolution by μ_S , and note that (6.2) gives

$$\|\mu_S^{(n)} - \mu_G\|_2 = \sup_{\|f\|_2 \leq 1, \mu_G(f)=0} |P_{\mu_S}^n f(1)| \leq \|P_{\mu_S}^n\|_2 = \beta_S^n. \quad (6.3)$$

It is trivial that $T_\infty \geq \gamma$, so to prove (7) it suffices to check that $T_1 > \frac{1}{2}\gamma$. To see this, note that the ball of radius T_1 contains at least half of the elements of G , since otherwise $\|\mu_S^{T_1} - \mu_G\|_1 \geq \frac{1}{2}$, and so the ball of radius $2T_1$ is all of G .

The fact that (6) holds with $C(p) = 2$ for $p \geq 2$ follows from (2) and (4). When $p \in (1, 2)$, applying Young's inequality to the identity $(\mu_S^{(2n)} - \mu_G) = (\mu_S^{(n)} - \mu_G)^{(2)}$ implies that $\|\mu_S^{(2n)} - \mu_G\|_{\frac{p}{2-p}} \leq \|\mu_S^{(n)} - \mu_G\|_p^2$, and hence that $T_{\frac{p}{2-p}} \leq 2T_p$. Iterating this a finite number of times if necessary, we see that $T_2 \leq C'(p)T_p$ for some constant $C'(p) > 0$, and (6) follows.

The first inequality of (8) follows immediately from (5) and the lower bound on λ_1 from (1.3), while the second inequality follows from the fact that $|G| = |S^\gamma| \leq |S|^\gamma$. Finally, to prove (9), note that (1.3) shows that $T_{rel} \leq 8|S|\gamma^2$, while the fact that $T_{rel} \leq T_1$ follows from (3) and the assumption that $\lambda_1 \leq 2$. \square

Proof of Corollary 1.10. The lower bound follows immediately from (3) and the bound (1.5) from Corollary 1.6. For the upper bound, we need to improve Proposition 6.2 (5) into $\|\mu_S^{(n)} - \mu_G\|_2 \leq B\beta_S^n \|\mu_G\|_2$. Getting this extra factor $\|\mu_G\|_2$ is the heart of the matter, and it was established by Diaconis and Saloff-Coste in [15, Theorem 3.2] for groups of moderate growth with an explicit dependence of the constant B in terms of the moderate growth parameters A and d . Corollary 1.9 implies that G has such parameters depending only on ε , and so the desired bound follows. (The reader should note that, while [15, Theorem 3.2] is stated for $\|\mu_S^{(n)} - \mu_G\|_1$, it also holds for $\|\mu_S^{(n)} - \mu_G\|_2 / \|\mu_G\|_2$ by Remark 2 on page 6 of [15].) \square

We now discuss and prove Theorem 1.11. Almost flat groups or groups of moderate growth are doubling at at least one scale $\leq \gamma^\delta$ for any fixed $\delta > 0$ by Lemma 4.9, and it is legitimate to ask whether this property alone is enough to give a mixing time quadratic in γ . The following reformulation of Theorem 1.11 shows that if a Cayley graph is doubling at one scale $\leq \gamma^{2/3}$ then this is indeed enough to guarantee quadratic mixing time, thus extending the validity of the Diaconis–Saloff-Coste theorem [15].

Theorem 6.3 (Doubling and quadratic mixing time). *Let G be a finite group with symmetric generating set S and denote by γ the diameter of the Cayley graph. Let $K \geq 1$ be a parameter. Then there is a constant $C_K > 0$ depending on K only such that if $|S^{2n+1}| \leq K|S^n|$ for some $n \geq C_K$ the*

- if $n \leq \gamma/3$, then T_{rel} is quadratic in the sense that $T_{rel} \geq \gamma^2/C_K$;
- if $n \leq \gamma^{2/3}$, then T_p is also quadratic for each $p \in [1, \infty]$, in the sense that $T_p = O_{K,|S|}(\gamma^2)$.

In proving Theorem 6.3 we make use of the following variant of Lemma 5.1.

Lemma 6.4. *Let \mathcal{G} be the Cayley graph of a finite group G with symmetric generating set S . Let $H \leq G$ be a normal subgroup and let γ_H be the diameter of H in the graph distance. Assume that f is a function on G such that $\sum_{x \in gH} f(x) = 0$ for every $g \in G$. Then*

$$\frac{\|\nabla f\|_2^2}{\|f\|_2^2} \geq \frac{1}{\gamma_H^2}$$

Proof. Pick $y \in H$ and connect it to the identity by a geodesic path $1 = y_0, y_1, \dots, y_n = y$, where $y_i = s_i y_{i-1}$ for some $s_i \in S$, and $n \leq \gamma_H$. Observe that for each i we have $\sum_{g \in G} |f(y_{i-1}g) - f(y_i g)|^2 \leq \|\nabla f\|_2^2$.

The Cauchy–Schwarz inequality in the form $(\sum_{i=1}^n |v_i|)^2 \leq n\|v\|_2^2$ then gives

$$\sum_{g \in G} |f(g) - f(yg)|^2 \leq \sum_{g \in G} \gamma_H \sum_1^n |f(y_{i-1}g) - f(y_i g)|^2 \leq \gamma_H^2 \|\nabla f\|_2^2. \quad (6.4)$$

On the other hand, for $g \in G$ we have $\sum_{y \in H} |f(g) - f(yg)|^2 = \|f - f(g)\|_{L^2(gH)}^2$ since $gH = Hg$, and so

$$\sum_{y \in H} |f(g) - f(yg)|^2 \geq \|f\|_{L^2(gH)}^2 \quad (6.5)$$

since f has mean zero on gH . Summing (6.4) over $y \in H$ and (6.5) over $g \in G$ shows that $|H|\|f\|_2^2 \leq \sum_{g \in G} \sum_{y \in H} |f(g) - f(yg)|^2 \leq |H|\gamma_H^2 \|\nabla f\|_2^2$, and so the lemma is proved. \square

Proof of Theorem 6.3. The first item follows directly from the combination of Theorem 1.1 and Lemma 5.2.

For the second item, we first note that if C_K is chosen large enough then Proposition 4.24 implies that there is a normal subgroup $H \leq G$ contained in $S^{8\gamma^{2/3}}$ such that G/H has a nilpotent subgroup whose index, rank and nilpotency class are bounded in terms of K only. Note that the diameter $\gamma_{G/H}$ of the quotient Cayley graph is at least $\gamma/2$ (unless γ is bounded), and hence that, by Proposition 4.7 and Lemma 4.2, this quotient Cayley graph is ε -almost flat for some ε depending only on $|S|$ and K . Corollary 1.10 therefore applies, and we conclude (denoting by $P := P_{\mu_S}$ the operator of left convolution by μ_S and recalling (6.3)) that if u is a zero mean function in $\ell^2(G/H)$, then $\|P^n u\|_2 \leq B e^{-n/C\gamma^2} \|u\|_2 \|\mu_{G/H}\|_2$ for all $n \geq 1$ and for some positive constants B, C depending only on $|S|$ and K .

Now we decompose the space $\ell^2(G)$ as a direct sum $\ell^2(G) = \ell^2(G/H) \oplus \ell^2(G/H)^\perp$, where $\ell^2(G/H)$ is this time viewed as the subspace of H -invariant functions and $\ell^2(G/H)^\perp$ its orthogonal, the subspace of functions with zero mean on each coset of H . Both spaces are G -invariant, and given a function f on G we may write f uniquely as $f = u + v$, where $u \in \ell^2(G/H)$ and $v \in \ell^2(G/H)^\perp$, and $\|f\|_2^2 = \|u\|_2^2 + \|v\|_2^2$. In particular, this implies that $\|P^n f\|_2^2 = \|P^n u\|_2^2 + \|P^n v\|_2^2$. However, the norm of P in restriction to $\ell^2(G/H)^\perp$ is given by $1 - \lambda/|S| \leq e^{-\lambda/|S|}$, where λ is the smallest eigenvalue of Δ in restriction to $\ell^2(G/H)^\perp$. Lemma 6.4 shows that $\lambda \geq 1/\gamma_H^2$, where γ_H is the diameter of H in G , and so we have $\|P^n v\|_2 \leq e^{-\frac{n}{|S|\gamma_H^2}} \|v\|_2$.

Now note that $\mu_S^{(n)} - \mu_G = P^n(f)$, where $f = 1_e - \mu_G$. Decompose f as $f = \bar{f} + (f - \bar{f})$, where \bar{f} is the H -invariant function $\bar{f}(g) := \frac{1}{|H|} \sum_{x \in gH} f(x)$, so that $f - \bar{f}$ has zero mean on H -cosets, and note therefore that

$$\|\mu_S^{(n)} - \mu_G\|_2^2 = \|P^n(\bar{f})\|_2^2 + \|P^n(f - \bar{f})\|_2^2 \leq B^2 e^{-2n/C\gamma^2} \|\bar{f}\|_2^2 \|\mu_{G/H}\|_2^2 + e^{-\frac{2n}{|S|\gamma_H^2}} \|f - \bar{f}\|_2^2$$

Finally, observe that $\bar{f} = \frac{1}{|H|} 1_H - \mu_G$, that $\|\bar{f}\|_2^2 = 1/|H| - 1/|G| \leq 1/|H| = \|\mu_G\|_2^2 / \|\mu_{G/H}\|_2^2$, and that $\|f - \bar{f}\|_2^2 = 1 - \frac{1}{|H|} \leq 1$. Hence

$$\|\mu_S^{(n)} - \mu_G\|_2^2 \leq B^2 e^{-2n/C\gamma^2} \|\mu_G\|_2^2 + e^{-\frac{2n}{|S|\gamma_H^2}}.$$

Now note that $\gamma_H^2 \log |G| = \gamma_H^2 \log(|G|/|H|) + \gamma_H^2 \log |H|$. But G/H is ε -almost flat, so $|G/H| \leq |S|\gamma^{1/\varepsilon}$, and clearly $|H| \leq |S|^{\gamma_H}$, so $\log |H| \leq \gamma_H \log |S|$. On the other hand H is contained in the ball of radius

$8\gamma^{2/3}$ so $\gamma_H \leq 8\gamma^{2/3}$. It follows that

$$\gamma_H^2 \log |G| \ll_{K,|S|} \gamma_H^2 \log \gamma + \gamma_H^3 \ll \gamma^2,$$

and therefore that $\|\mu_S^{(n)} - \mu_G\|_2 \leq \frac{1}{10}\|\mu_G\|_2$ as soon as $n \gg_{K,|S|} \gamma^2$. This gives the result for T_2 , and hence for all T_p by Proposition 6.2. \square

We finish by presenting a simple example showing that the exponent $\frac{2}{3}$ in the second item of Theorem 1.11 is sharp. We thank T. Zheng for suggesting it.

Example 6.5 (Sharpness of the $\frac{2}{3}$ in Theorem 1.11). Let L_M be the lamplighter group $L_M := \mathbb{Z}/M\mathbb{Z} \times B$, where B is the direct product of M copies of $\mathbb{Z}/2\mathbb{Z}$ that are cyclically permuted by the action of $\mathbb{Z}/M\mathbb{Z}$, and let $G := L_M \times \mathbb{Z}/N\mathbb{Z}$. Let $S := S_1 \times S_2$, where S_1 is the standard generating set of L_M (consisting of the identity, a move of one space to the left or right, or a switch of the current lamp), and S_2 is the standard generating set $\{0, \pm 1\}$ of $\mathbb{Z}/N\mathbb{Z}$. Clearly the diameter γ of G is comparable to $\max\{M, N\}$. On the other hand, the uniform mixing time $T_\infty(G)$ is at least $\max\{T_\infty(L_M), T_\infty(\mathbb{Z}/N\mathbb{Z})\}$, which is comparable to $\max\{M^3, N^2\}$ (see [25]). Let $\alpha \in (\frac{2}{3}, 1)$ and set $M := N^\alpha$. Then $\gamma \simeq N$, and $T_\infty(G)$ grows faster than $N^2 \simeq \gamma^2$. However, $M \simeq \gamma^\alpha$, and we have uniform doubling at all scales $\gtrsim M$.

REFERENCES

- [1] N. Alon and Y. Roichman. Random Cayley graphs and expanders, *Random structures algorithms* **5**(2) (1994), 271-284.
- [2] L. Babai and A. Seress. On the diameter of Cayley graphs of the symmetric group. *Journal of combinatorial theory, Series A* **49** (1988), 175–179.
- [3] H. Bass. The degree of polynomial growth of finitely generated nilpotent groups, *Proc. London Math. Soc. (3)* **25**(4) (1972), 603-614.
- [4] I. Benjamini, H. Finucane and R. Tessera. On the scaling limit of finite vertex transitive graphs with large diameter. [arXiv:1203.5624](https://arxiv.org/abs/1203.5624).
- [5] E. Breuillard. Expander graphs, property (τ) and approximate groups, in *Geometric Group Theory, IAS/Park City Mathematics Series*, Eds. M. Bestvina, M. Sageev, and K. Vogtmann, (2014) 399 pp, Volume 21.
- [6] E. Breuillard. Diophantine geometry and uniform growth of finite and infinite groups. *Proceedings of the International Congress of Mathematicians, 2014 ICM (Seoul)*.
- [7] E. Breuillard. Approximate groups and Hilbert’s fifth problem, in preparation, to appear in IMA Volume *Recent Trends in Combinatorics*, Springer-Verlag.
- [8] E. Breuillard and B. J. Green. Approximate groups. I. The torsion-free nilpotent case, *J. Inst. Math. Jussieu* **10**(1) (2011), 37-57.
- [9] E. Breuillard, B. Green and T. Tao. Approximate subgroups of linear groups, *Geom. Funct. Anal.* **21**(4) (2011), 774-819.
- [10] E. Breuillard, B. J. Green and T. C. Tao. The structure of approximate groups, *Publ. Math. IHES* **116**(1) (2012), 115-221.
- [11] E. Breuillard, B. J. Green and T. C. Tao. Small doubling in groups, *Proc. of the Erdős centenary conference* (2013). [arXiv:1301.7718](https://arxiv.org/abs/1301.7718).
- [12] R. Brooks, Combinatorial problems in spectral geometry, Curvature and topology of Riemannian manifolds (Katata, 1985), 14–32, Lecture Notes in Math., 1201, Springer, Berlin, (1986).
- [13] M. Burger, Grandes valeurs propres du laplacien et graphes Séminaire de Théorie Spectrale et Géométrie, No. 4, Année 1985–1986, 95–100, Univ. Grenoble I, Saint-Martin-dHères, (1986).
- [14] P. Diaconis and L. Saloff-Coste. Comparison techniques for random walk on finite groups, *Ann. Probab.* **21**(4) (1993), 2131-2156.
- [15] P. Diaconis and L. Saloff-Coste. Moderate growth and random walk on finite groups, *Geom. Funct. Anal.* **4**(1) (1994), 1-36.
- [16] J. Dixon and B. Mortimer. Permutation groups. *Graduate Texts in Mathematics* **163**, Springer-Verlag, New York, 1996. xii+346 p
- [17] S. Eberhard. A note on nonabelian Freiman–Ruzsa. [arXiv:1511.06758](https://arxiv.org/abs/1511.06758).

- [18] J. S. Ellenberg, C. Hall and E. Kowalski. Expander graphs, gonality and variation of Galois representations, *Duke Math. J.* **161**(7) (2012), 1233-1275.
- [19] G. A. Freiman. On finite subsets of nonabelian groups with small doubling, *Proc. Amer. Math. Soc.* **140**(9) (2012), 2997-3002.
- [20] B. J. Green. Approximate groups and their applications: work of Bourgain, Gamburd, Helfgott and Sarnak, *Current events bulletin of the AMS* (2010). [arXiv:0911.3354](https://arxiv.org/abs/0911.3354).
- [21] B. J. Green. The polynomial Freiman-Ruzsa conjecture. *Open Problems in Mathematics [Online]*, Volume 2 (2014).
- [22] B. J. Green and I. Z. Ruzsa. Freiman's theorem in an arbitrary abelian group, *J. Lond. Math. Soc.* **75**(1) (2007), 163-175.
- [23] M. Gromov. Groups of polynomial growth and expanding maps, *Publ. Math. IHES* **53** (1981), 53-73.
- [24] M. Gromov. Metric structures for Riemannian and non-Riemannian spaces. Based on the 1981 French original, *Modern Birkhuser Classics*, Birkhuser Boston, Inc., Boston, MA, (2007), xx+585 pp.
- [25] Olle Häggström and Johan Jonasson, Rates of convergence for lamplighter processes, *Stochastic Process. Appl.* **67**(2) (1997), 227-249.
- [26] M. Hall. *The theory of groups*, Amer. Math. Soc./Chelsea, Providence, RI (1999).
- [27] H. A. Helfgott. Growth and generation in $SL_2(\mathbb{Z}/p\mathbb{Z})$, *Ann. of Math. (2)* **167**(2) (2008), 601-623.
- [28] H. A. Helfgott and A. Seress. On the diameter of permutation groups, *Ann. of Math. (2)* **179**(2) (2014), 611-658.
- [29] S. Hoory, N. Linial and A. Wigderson. Expander graphs and their applications, *Bull. Amer. Math. Soc.* **43** (2006), 439-561.
- [30] B. Klartag, G. Kozma, P. Ralli and P. Tetali. Discrete curvature and abelian groups. [arXiv:1501.00516](https://arxiv.org/abs/1501.00516).
- [31] M. Lackenby. Expanders, rank and graphs of groups. *Israel J. Math.* **146** (2005), 357-370.
- [32] M. Lackenby. A characterisation of large finitely presented groups. *J. Algebra* **287**(2) (2005), 458-473.
- [33] M. Lackenby. Heegaard splittings: The virtually Haken conjecture and property (τ) , *Invent. Math.* **164** (2006), 317-359.
- [34] M. Lackenby. Finite covering spaces of 3-manifolds, *Proceedings of the ICM*, volume II, 1042-1070. Hindustan Book Agency, New Delhi (2010).
- [35] J. Lee and Y. Makarychev. Eigenvalue multiplicity and volume growth. [arXiv:0806.1745v1](https://arxiv.org/abs/0806.1745v1).
- [36] E. Lubetzky and Y. Peres. Cutoff on all Ramanujan graphs. [arXiv:1507.04725](https://arxiv.org/abs/1507.04725).
- [37] A. Lubotzky. Discrete groups, expanding graphs and invariant measures. With an appendix by Jonathan D. Rogawski. Reprint of the 1994 edition. *Modern Birkhuser Classics*. Birkhuser Verlag, Basel, 2010. iii+192 pp.
- [38] A. Lubotzky and B. Weiss. Groups and expanders, *Expanding graphs* (Princeton, NJ, 1992), *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, volume 10, 95-109. Amer. Math. Soc., Providence, RI (1993).
- [39] T. Mantuano, Discretization of Compact Riemannian Manifolds applied to the spectrum of the Laplacian, *Annals of Global Analysis and Geometry* **27** (2005) 33-46.
- [40] Y. Peres and D. Revelle, Mixing times for random walks on finite lamplighter groups. *Electron. J. Probab.* **9** (2004), no. 26, 825-845.
- [41] G. Petridis. New proofs of Plünnecke-type estimates for product sets in groups, *Combinatorica* **32**(6) (2012), 721-733.
- [42] L. Pyber and E. Szabó. Growth in finite simple groups of Lie type, *J. Amer. Math. Soc.* **29** (2016), 95-146.
- [43] I. Z. Ruzsa. An application of graph theory to additive number theory, *Scientia*, Ser. A **3** (1989), 97-109.
- [44] T. Sanders. The structure theory of set addition revisited, *Bull. Amer. Math. Soc.* **50** (2013), 93-127.
- [45] T. C. Tao. Product set estimates for non-commutative groups, *Combinatorica* **28**(5) (2008), 547-594.
- [46] T. C. Tao. Hilbert's 5th problem and related topics. *Graduate Studies in Mathematics* **153**, American Mathematical Society, Providence, RI, 2014. xiv+338
- [47] T. C. Tao. Inverse theorems for sets and measures of polynomial growth, to appear in *Q. J. Math.*. [arXiv:1507.01276](https://arxiv.org/abs/1507.01276).
- [48] T. C. Tao and V. H. Vu. Additive combinatorics, *Cambridge studies in advanced mathematics* **105**, Cambridge Univ. Press (2006).
- [49] M. C. H. Tointon. Freiman's theorem in an arbitrary nilpotent group, *Proc. London Math. Soc.* **109** (2014), 318-352.

LABORATOIRE DE MATHÉMATIQUES D'ORSAY, UNIV. PARIS-SUD, CNRS, UNIVERSITÉ PARIS-SACLAY, 91405 ORSAY, FRANCE

E-mail address: emmanuel.breuillard@math.u-psud.fr

LABORATOIRE DE MATHÉMATIQUES D'ORSAY, UNIV. PARIS-SUD, CNRS, UNIVERSITÉ PARIS-SACLAY, 91405 ORSAY, FRANCE

E-mail address: matthew.tointon@math.u-psud.fr