

GROWTH IN LINEAR GROUPS

SEAN EBERHARD, BRENDAN MURPHY, LÁSZLÓ PYBER, AND ENDRE SZABÓ

ABSTRACT. We prove a conjecture of Helfgott on the structure of sets of bounded tripling in bounded rank, which states the following. Let A be a finite symmetric subset of $\mathrm{GL}_n(\mathbf{F})$ for any field \mathbf{F} such that $|A^3| \leq K|A|$. Then there are subgroups $H \trianglelefteq \Gamma \trianglelefteq \langle A \rangle$ such that A is covered by $K^{O_n(1)}$ cosets of Γ , Γ/H is nilpotent of step at most $n-1$, and H is contained in $A^{O_n(1)}$. This theorem includes the Product Theorem for finite simple groups of bounded rank as a special case. As an application of our methods we also show that the diameter of sufficiently quasirandom finite linear groups is poly-logarithmic.

CONTENTS

1. Introduction	1
2. Notation	7
3. Examples	8
4. Toolbox	10
5. Finitization	13
6. Trigonalization	15
7. Main proof part 1: general to soluble	18
8. Main proof part 2: soluble to nilpotent	24
9. Diameter of quasirandom groups	33
References	37

1. INTRODUCTION

1.1. Statement of results. In this paper we characterize sets of bounded tripling in $\mathrm{GL}_n(\mathbf{F})$, where n is bounded and \mathbf{F} is an arbitrary field. Here a finite set $A \subseteq \mathrm{GL}_n(\mathbf{F})$ is said to be *K-tripling* if $|A^3| \leq K|A|$. This notion is largely the same as that of a finite *K-approximate group*, which

SE has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 803711) and from the Royal Society. BM was funded by The Leverhulme Trust through Leverhulme grant RPG 2017-371. LP was supported by the National Research, Development and Innovation Office (NKFIH) Grant K115799, ESz was supported by the NKFIH Grants K115799 and K120697. The project leading to this application has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 741420).

is a symmetric set A containing 1 such that A^2 is covered by at most K translates of A . Prototypical examples include subgroups and progressions $\{g^n : |n| \leq N\}$, as well as certain nilpotent generalizations of progressions called nilprogressions, such as the Heisenberg nilprogression

$$A = \left\{ \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} : |x|, |y| \leq N, |z| \leq N^2 \right\} \subseteq \mathrm{GL}_3(\mathbf{R}).$$

In broad qualitative terms, the most general approximate group is an extension of a subgroup by a nilprogression: this is the content of the celebrated structure theorem for approximate groups proved by Breuillard, Green, and Tao [BGT2].

Theorem 1.1 (Breuillard, Green, Tao [BGT2]). *Let A be a K -approximate subgroup of any group G . Then there are subgroups $H \trianglelefteq \Gamma \leq G$ with the following properties:*

- (1) A is covered by $O_K(1)$ cosets of Γ ,
- (2) Γ/H is nilpotent of rank and step at most $O_K(1)$,
- (3) H is contained in A^4 .

Breuillard, Green, and Tao described the statement of Theorem 1.1 as the *Helfgott–Lindenstrauss conjecture*. However, there is some inconsistency in the usage of this term. Theorem 1.1 corresponds closely to the essentially qualitative conjecture made by Lindenstrauss (Lindenstrauss, personal communication). On the other hand the conjecture of Helfgott [H1, T1] predicted *polynomial bounds*, at least for groups in *bounded rank*. This was later formulated more precisely in [GH] as well as [H2].

Unfortunately the proof of Theorem 1.1 depends on nonstandard analysis as well as the solution to Hilbert’s fifth problem, and there seems to be no way to deduce an explicit bound for the number of cosets of Γ required to cover A (the bounds on the rank and step of Γ/H are explicit and benign: see [BGT2, Remark 1.9]). However, it is now known [BT, E] (see also Section 3) that the number of cosets required is not polynomial in general in high rank.

In this paper we prove Helfgott’s original conjecture.¹ That is, we give a version of the Breuillard–Green–Tao theorem with polynomial bounds in bounded rank, over arbitrary fields.

Theorem 1.2. *Let \mathbf{F} be a field of characteristic $p \geq 0$ and let $A \subseteq \mathrm{GL}_n(\mathbf{F})$ be finite, nonempty, symmetric, and K -tripling, where $K \geq 2$. Let $G = \langle A \rangle$. Then there are subgroups $H \trianglelefteq \Gamma \leq G$ such that*

- (1) A is covered by $K^{O_n(1)}$ cosets of Γ ,
- (2) Γ/H is nilpotent of step at most $n - 1$,

¹This version of our paper differs significantly from the first version that appeared on the arXiv. In this version we prove the normality of Γ in G , which requires significant generalization of our earlier methods. We had to extend results about subgroups to results about sections. See Section 1.3 for more details.

(3) H is contained in $A^{O_n(1)}$.

Moreover, if $p > 0$ then H has a perfect soluble-by-Lie $^*(p)$ normal subgroup P contained in a translate of A^3 such that H/P is a p -group, and if $p = 0$ then H is trivial.

Here we write $\text{Lie}(p)$ for the class of finite simple groups of Lie type of characteristic p and we say G is in $\text{Lie}^*(p)$ if it is a finite direct product of members of $\text{Lie}(p)$. For convenience if $p = 0$ then $\text{Lie}(p)$ is defined to be empty and $\text{Lie}^*(p)$ consists of just the trivial group.

Remark 1.3.

- (a) Without loss of generality $H = \gamma_n(\Gamma)$, the n th term of the lower central series of Γ . Similarly P must be the perfect core of Γ (the last term of the derived series). Thus we may assume H and P are characteristic in Γ and normal in G .
- (b) The rank of Γ/H cannot be controlled without sacrificing normality of Γ in G . In fact in Section 3 we give an example in which Γ cannot be chosen to be finitely generated. However if Γ is not required to be normal then we may assume $\Gamma = \langle A^6 \cap \Gamma \rangle$ (see Lemma 4.4), and by applying the result of Tointon [T2] we may choose $H \subseteq A^{K^{O_n(1)}}$ so that $P \leq \gamma_n(\Gamma) \leq H \trianglelefteq \Gamma$, Γ/H has rank $K^{O_n(1)}$, and A is still covered by $K^{O_n(1)}$ cosets of Γ .

Roughly half the proof of Theorem 1.2 consists of establishing the existence of P . This preliminary “soluble version” was previously announced in [PS2].

Theorem 1.4. *Let hypotheses be as Theorem 1.2. Then there are subgroups $P \trianglelefteq \Gamma \trianglelefteq \langle A \rangle$ such that*

- (1) A is covered by $K^{O_n(1)}$ cosets of Γ ,
- (2) Γ/P is soluble of derived length $O(\log n)$,
- (3) P is perfect, soluble-by-Lie $^*(p)$, and contained in a translate of A^3 .

These results generalize and depend on the Product Theorem for finite simple groups, obtained independently by Breuillard, Green, and Tao [BGT1] and Pyber and Szabó [PS3].

Theorem 1.5 ([PS3, Theorem 2], see also [BGT1, Corollary 2.4]). *Let L be a finite simple group of Lie type of rank r and A a generating set of L . Then either*

- (1) $|A^3| > |A|^{1+\varepsilon}$, where $\varepsilon = \varepsilon(r)$ depends only on r , or
- (2) $A^3 = L$.

A key new ingredient in this paper is something we call the “affine conjugating trick”. See the outline below for a description. As a further application of this trick, we prove a diameter bound for quasirandom groups.

Theorem 1.6. *For each positive integer n there are positive numbers $K = K(n)$ and $c = c(n)$ with the following property. Let \mathbf{F} be a field and $G \leq \mathrm{GL}_n(\mathbf{F})$ be a K -quasirandom finite subgroup. Then the Cayley graph of G with respect to any generating set has diameter at most $(\log |G|)^c$.*

For p -generated perfect subgroups of $\mathrm{SL}_n(\mathbf{F}_p)$ this was proved in [PS3]. Like the proof of Theorem 1.2 in the special case of subsets of $\mathrm{GL}_n(\mathbf{F}_p)$ ([GH] building on [PS1, PS3]), the proof of this special case depends in an essential way on the Nori correspondence between p -generated subgroups of $\mathrm{SL}_n(\mathbf{F}_p)$ and certain closed subgroups of $\mathrm{SL}_n(\mathbf{F}_p)$. The fact that the affine conjugating trick can be used to replace the Nori correspondence in the proof of these two related results, and used in the proof of their extension to arbitrary fields, clearly shows the power of this trick.

1.2. Relation to previous literature. In characteristic zero, Theorem 1.2 was first established by Breuillard, Green, and Tao [BGT1]. We include this case mainly for the sake of having a uniform argument for all fields, and because it is hardly any extra work, but it is not the important contribution of this paper. The fact that the arbitrary field case reduces to the case of finite fields is notable, and validates the opinion of Gill and Helfgott (see [GH, Section 1.3]) that finite fields contain the heart of the matter.

The prime finite field case of Theorem 1.2 was proved by Gill and Helfgott [GH], conditional on the prime finite field case of Theorem 1.4, a result of the third and fourth authors which has not previously appeared in published form. It appears in an unpublished part of an earlier version [PS1] of [PS3] as Corollary 105, but the published version [PS3] contains only a preliminary result, Lemma 73, on perfect p -generated subgroups of $\mathrm{SL}_n(\mathbf{F}_p)$ (Theorem 85 in [PS1]). The publication of Corollary 105 has been deferred until now, in anticipation of the full version above applying to all fields uniformly.

The results proved in [PS1, PS3] depend in an essential way on the Nori correspondence between p -generated subgroups of $\mathrm{SL}_n(\mathbf{F}_p)$ and certain closed subgroups of $\mathrm{SL}_n(\overline{\mathbf{F}_p})$. Since the Nori theory does not extend to subgroups of $\mathrm{SL}_n(\mathbf{F}_q)$, q a prime power, we had to devise an entirely different argument to prove Theorem 1.4. We use arguments based on quasirandomness and a new “affine conjugating trick”, to be described below.

Similarly, the method of [GH] depends on the theory of algebraic groups over \mathbf{F}_p , and several critical features of this theory break down for fields of prime-power order, such as boundedness for chains of unipotent subgroups. Accordingly, while parts of our method are inspired by [GH] (in particular their pivoting theorem is a key tool) our deduction of Theorem 1.2 from Theorem 1.4 uses many entirely new tools, such as a growth lemma for images of bilinear maps and a more general descent argument. In fact, the subgroup Γ arises in a slightly different way in our method: we do not take

a full root kernel but we use a tricky pigeonholing argument to identify an appropriate normal subgroup.

1.3. Outline of the paper. See Section 2 for notation. In Section 3 we give a few examples that illustrate some of the subtleties of Theorem 1.2 and Theorem 1.4.

Section 4 recalls some tools that may be familiar to experts: basic group theory and arithmetic combinatorics (or nonabelian additive combinatorics), quasirandomness, and the action of p' -groups on p -groups.

In Section 5 we give a pair of short arguments that immediately reduce Theorems 1.2 and 1.4 to the case of finite fields. The key tool here is the well-known theorem of Mal'cev asserting a strong form of residual finiteness for finitely generated linear groups. In the rest of the paper we take \mathbf{F} to be finite.

In Section 6 we prove some results about trigonalization of soluble subgroups and sections of $\mathrm{GL}_n(\mathbf{F})$. A subgroup is called trigonalizable if it is conjugate to a group of upper-triangular matrices over some extension field. A section is called trigonalizable if it is the image of a trigonalizable subgroup. Another well-known theorem of Mal'cev asserts that soluble subgroups of $\mathrm{GL}_n(\mathbf{F})$ are virtually trigonalizable. We prove a variant for soluble sections that moreover provides crucial control over the “Weyl group” of the section (Theorem 6.4). This control is essential for obtaining normality of Γ in Theorem 1.2.

The rest of the paper contains the body of the proof of Theorems 1.2 and 1.4. Section 7 covers the proof of Theorem 1.4, while Section 8 covers the deduction of Theorem 1.2.

The starting points of Section 7 (the general-to-soluble reduction) are the Product Theorem and Weisfeiler’s structure theorem (Theorem 7.2) for finite linear groups. Using these tools in combination with quasirandomness arguments and a few other facts about finite simple groups, we establish the following initial structure theorem, which can be viewed as an “upside-down” version of Theorem 1.4.

Theorem 1.7. *Let \mathbf{F} be a finite field of characteristic $p > 0$. Let $A \subseteq \mathrm{GL}_n(\mathbf{F})$ be a symmetric subset such that $|A^3| \leq K|A|$, where $K \geq 2$. Then there is a normal subgroup $\Gamma \trianglelefteq \langle A \rangle$ such that*

- (1) *A is covered by $K^{O_n(1)}$ cosets of Γ ,*
- (2) *Γ is soluble-by- $\mathrm{Lie}^*(p)$,*
- (3) *$\Gamma/\mathrm{Sol}(\Gamma)$ is covered by A^6 .*

The main business of Section 7 is to turn this structure “right side up”. To do this we must show that the perfect core P of Γ is covered by $A^{O_n(1)}$ (quasirandomness upgrades this to A^3 automatically: see Lemma 4.9). An argument based on the weak Ore conjecture shows that $A^{O_n(1)}$ covers $P/\mathrm{Sol}(P)$, and it follows from Weisfeiler’s theorem that $N = [P, \mathrm{Sol}(P)]$ is a p -subgroup such that $[\mathrm{Sol}(P) : N] \leq O_n(1)$ (Corollary 7.4), so the key is to show that

$A^{O_n(1)}$ covers N . For this we use the following “affine conjugating trick” (Lemma 7.11).

Lemma 1.8 (affine conjugating trick). *Let $\Gamma = V \rtimes G$ be the semidirect product of an abelian group V and a d -generated K^{21} -quasirandom finite group G . Let $A \subseteq V$ be a symmetric G -invariant set generating V . If $|A^3| \leq K|A|$ then $A^{7d} \supseteq [V, G]$.*

Finally, we show that $A^{O_n(1)}$ covers N and hence P using a rather tricky argument based on Lemma 1.8 (see Proposition 7.12). This completes the proof of Theorem 1.4.

Two key tools are used in Section 8, which proves Theorem 1.2 starting from Theorem 1.4. The first is a powerful pivoting argument of Gill and Helfgott that can be seen as an extremely general version of a sum-product theorem: see Proposition 8.1. The second is a new growth lemma for images of a bilinear map: see Proposition 8.7. Additionally we make heavy use of the results on trigonalizable sections established in Section 6.

The deduction of Theorem 1.2 is divided into cases of increasing generality. To begin with we consider the case in which G has a trigonalizable section $\Sigma = \Gamma/N$ and A is a subset of Γ whose image in Σ has “no small roots”. Here a *root* is a homomorphism χ from Σ to some abelian p' -group defined by the conjugation action of Σ on a Σ -composition factor of $O_p(\Sigma)$: these are analogous to roots in the theory of Lie algebras. We say A has *no small roots* if $\chi(A)$ is either trivial or larger than K^C for an appropriate constant C , for every root χ of Σ . In this favorable case we can show that $\gamma_n(\Sigma)$ is covered by $A^{O_n(1)}$. This argument is a somewhat involved induction on the nilpotency class of $\gamma_n(\Sigma)$, and uses a more general form of the idea of “descent” from [GH]. This argument is given in Section 8.3.

If, more generally, A is not necessarily a subset of Γ but still acts trivially on $\Sigma/O_p(\Sigma)$, then we use pigeonholing argument to shrink Σ until the image of $A^2 \cap \Gamma$ in Σ has no small roots, and then we apply the previous case to $A^4 \cap \Gamma$. By comparing $\Sigma_1 = \langle A^2 \cap \Gamma \rangle$ with $\Sigma_2 = \langle A^4 \cap \Gamma \rangle$ and using the no-small-roots property, we prove that $\gamma_n(\Sigma_1) = \gamma_n(\Sigma_2)$. This crucial conclusion allows us to identify an appropriate normal subgroup $\Delta \trianglelefteq \langle A \rangle$ such that $P \leq \Delta \leq \Gamma$. This subgroup Δ fills the role of Γ in Theorem 1.2. See Section 8.4.

Finally, in general, the trigonalization theorem from Section 6 allows us to replace the soluble section provided by Theorem 1.4 with a trigonalizable section Σ , and moreover guarantees that $\langle A \rangle$ has a bounded-index subgroup G_0 that acts trivially on $\Sigma/O_p(\Sigma)$. We complete the proof by applying the previous case to $A^{3m} \cap G_0$, where $m = [G : G_0]$, which is a generating set for G_0 by Lemma 4.6.

The proof of Theorem 1.6 is given in Section 9. In this section we recycle some of the same ideas. In particular the affine conjugating trick plays a key role. We also rely on an important result of Steinberg on representations of finite simple groups of Lie type.

1.4. Acknowledgments. BM would like to thank Harald Helfgott for introducing him to this problem, as well as Vlad Finkelstein, Jonathan Pakianathan, Lam Pham, Misha Rudnev, Matthew Tointon, and James Wheeler for helpful conversations.

2. NOTATION

The commutator $[x, y]$ is defined as $x^{-1}y^{-1}xy$. Iterated commutators are defined by $[x, \dots, y, z] = [[x, \dots, y], z]$. If H, K are subgroups then $[H, K] = \langle [h, k] : h \in H, k \in K \rangle$, and similarly for iterated commutators. The subgroup $[H, K]$ is a normal subgroup of $\langle H, K \rangle$ ([A, (8.5.6)]). If H and K are groups and H acts on K then we may always consider H and K as subgroups of their semidirect product $K \rtimes H$. The commutator $[H, K] \leq K$ and centralizers $C_H(K)$ and $C_K(H)$ are defined accordingly. As usual $G' = [G, G]$. The derived series of G is denoted $(G^{(n)})_{n \geq 0}$, and $G^{(\omega)} = \bigcap_{n \geq 0} G^{(n)}$. If the derived series of G terminates in finitely many steps (e.g., if G is finite), then $G^{(\omega)}$ is called the *perfect core* of G . The lower central series is denoted $(\gamma_n(G))_{n \geq 1}$, and $\gamma_\omega(G) = \bigcap_{n \geq 1} \gamma_n(G)$.

A finite group is a *p-group* if its order is a power of p ; it is a *p'-group* if its order is prime to p . The largest normal p -subgroup of a finite group G is denoted $O_p(G)$ and called the *p-core*. The largest normal soluble subgroup is denoted $\text{Sol}(G)$ and called the *soluble radical*.

If G is a group, $\text{Frat}(G)$ denotes the Frattini subgroup. If G is finite, $\text{Frat}(G)$ is nilpotent, by the Frattini argument. If G is a p -group, the Burnside basis theorem asserts that $\text{Frat}(G)$ is the smallest normal subgroup such that $G/\text{Frat}(G)$ is elementary abelian (see [A, (23.2)]).

All of the subgroups just defined are characteristic subgroups. In general if H is a characteristic subgroup of G we write $H \trianglelefteq G$.

A *section* of a group G is a quotient $\Sigma = H/N$ where $N \trianglelefteq H \leq G$. The *normalizer* of Σ is $N_G(\Sigma) = N_G(H) \cap N_G(N)$. Note that the normalizer of Σ acts naturally on Σ . The *centralizer* of Σ is $C_G(\Sigma) = \{g \in N_G(\Sigma) : g \text{ acts trivially on } \Sigma\}$. If $A \subseteq G$, the *trace* of A in Σ , denoted $\text{tr}(A, \Sigma)$, is the projection of $A \cap H$ to Σ . We say A *covers* Σ if $\text{tr}(A, \Sigma) = \Sigma$, or equivalently if $H \subseteq AN$.

Throughout \mathbf{F} denotes a field, usually finite. The (Borel) subgroup of elements represented by an upper-triangular matrix in the standard basis is denoted $B_n(\mathbf{F})$. The (toral) subgroup of diagonal matrices is denoted $T_n(\mathbf{F})$, while the subgroup of upper unitriangular matrices is denoted $U_n(\mathbf{F})$. Note that $B_n(\mathbf{F}) = U_n(\mathbf{F})T_n(\mathbf{F}) \cong U_n(\mathbf{F}) \rtimes T_n(\mathbf{F})$. There is a projection map $\pi : B_n(\mathbf{F}) \rightarrow T_n(\mathbf{F})$ such that $\ker \pi = U_n(\mathbf{F})$. Various other groups of the form $G = U \rtimes T$ will arise and usually we denote by π the projection map $G \rightarrow T$ such that $\ker \pi = U$. Maps such as π are denoted interchangeably on the left or exponentially, so $\pi(A) = A^\pi$.

3. EXAMPLES

Well-known basic examples such as subgroups and nilprogressions demonstrate the necessity of the basic elements in the main theorem Theorem 1.2. See for example [B2, Section 1.6]. In this section we give some examples which demonstrate the necessity of some subtler aspects of Theorem 1.2.

3.1. Failure of polynomiality in high rank. Unlike Theorem 1.2, the bound in [BGT2] for the number of cosets of Γ required to cover A cannot be polynomial in K , even if we only require Γ to be finite-by-soluble. The following example appeared in the unpublished manuscript [E], and was based on a similar construction in [BT, Section 4.1].

Let $G = \mathbf{Z}^n \rtimes S_n$. and let $A = A_r = [-r, r]^n S_n$ for any $r > n!$. It is easy to see that A is a 2^n -approximate group. Suppose $\Gamma \leq G$ is a subgroup such that $m = |A\Gamma/\Gamma| < n!$. Since $A_r = A_1^r$, it follows that $|A_1^{s+1}\Gamma/\Gamma| = |A_1^s\Gamma/\Gamma|$ for some $s < r$, which implies that $G = A\Gamma$ since A generates G . Hence Γ has index m in G . In particular

- (1) $\Gamma \cap S_n$ has index at most m in S_n ,
- (2) $\Gamma \cap \mathbf{Z}^n$ contains $m\mathbf{Z}^n$.

Note that Γ has no nontrivial finite normal subgroup. Indeed if $H \trianglelefteq \Gamma$ is finite then $H \cap \mathbf{Z}^n = 1$ and $[m\mathbf{Z}^n, H] \leq \mathbf{Z}^n \cap H = 1$, so H acts trivially on \mathbf{Z}^n , so $H \leq \mathbf{Z}^n$, so $H = 1$.

If Γ is soluble then so is $\Gamma \cap S_n$, which implies $m \geq n!/24^{(n-1)/3}$ by a result of Dixon [D]. If Γ is in fact nilpotent then $\Gamma \cap S_n$ must be trivial, so $m \geq n!$.

Note that A is a 2^n -approximate group, and G is isomorphic to a subgroup of $\mathrm{GL}_{n+1}(\mathbf{Z})$. This shows that the number of cosets required in Theorem 1.1 must be at least $K^{c \log \log K}$, and that in Theorem 1.2 or Theorem 1.4 must be at least $K^{c \log n}$.

3.2. Failure of normality in high rank. Normality of Γ in G , as stated in Theorem 1.2 but not in Theorem 1.1, is a special feature of bounded rank. The following example is due to Pyber. It was previously mentioned in [B2, Example 1.17] and [H2, p. 53]).

Let n be odd and let $G = S_n$ be the symmetric group of degree n . Let $m = \lfloor n/4 \rfloor$ and let Γ_0 be the elementary abelian subgroup

$$\Gamma_0 = \langle (1, 2), (3, 4), \dots, (2m-1, 2m) \rangle \cong C_2^m.$$

Let σ be the n -cycle $x \mapsto x+2$. Let $A = \Gamma_0 \cup \{\sigma^{\pm 1}\}$. Then A is a 10-approximate group which generates G , and therefore subject to Theorem 1.1. However, there are only three normal subgroups of G , and none of them is suitable.

Many variants of this construction are possible. For example we can take Γ_0 to be a direct product of copies of A_5 .

3.3. Normality vs finite generation. Even in bounded rank we cannot guarantee normality simultaneously with finite generation of Γ (so in particular the rank of Γ/H in Theorem 1.2 cannot be controlled). This example also illustrates some of the challenges of non-prime fields.

Let $\mathbf{F} = \mathbf{F}_p(t)$, where $p = 2$ (say) and t is transcendental over \mathbf{F}_p . Let $V = \mathbf{F}_p[t, t^{-1}]$ and let

$$G = \left\{ \begin{pmatrix} t^n & v \\ 0 & 1 \end{pmatrix} : n \in \mathbf{Z}, v \in V \right\} \cong \mathbf{F}_p^{\mathbf{Z}} \rtimes \mathbf{Z}.$$

Note that $G \leq \mathrm{GL}_2(\mathbf{F})$. Let $W \leq V$ be the linear subspace spanned by t^e for $-d \leq e \leq d$ and let

$$A = \begin{pmatrix} 1 & W \\ 0 & 1 \end{pmatrix} \cup \left\{ \begin{pmatrix} t & 0 \\ 0 & 1 \end{pmatrix}^{\pm 1} \right\}.$$

Then A is a $(2p + 5)$ -approximate group which generates G , but G has no finitely generated normal finite-by-nilpotent subgroup apart from the trivial group.

3.4. Lack of other structure in high rank. We can give diverse examples of large generating subsets of small growth in high-rank simple groups. The main issue is to ensure that these subsets actually generate our group, which is a crucial condition in the Product Theorem for simple groups of bounded rank. We use the following trick inspired by an idea of Bannai [B1]. Let $A = S \cup C$, where S is an arbitrary K -tripling subset of some group G and C is a conjugacy class of G . Then $A^3 \subseteq S^3 \cup S^2C \cup SC^2 \cup C^3$, so A is K' -tripling for $K' = K + K|C| + 2|C|^2$. If $|C|$ is not much larger than K , this shows that the growth of A is similar to that of S . When G is a finite simple group then certainly C generates G , and this is also often the case when G is only almost simple.

For example, let $m \leq n$ and let Γ be a copy of S_m in S_n (or similarly a copy of $\mathrm{SL}_m(\mathbf{F}_2)$ in $\mathrm{SL}_n(\mathbf{F}_2)$) and let C be a small conjugacy class, say the set of transpositions (or transvections in $\mathrm{SL}_n(\mathbf{F}_2)$). Taking $m = n/2$ or even $m = n - 100$, we get a huge generating set $A = \Gamma \cup C$ of relatively small growth.

For another example let Γ be a cyclic subgroup of $G = S_n$ of order $N \approx \exp(c\sqrt{n \log n})$, let S be an interval of length \sqrt{N} in Γ , and let C_0 be a set of $n - 1$ transpositions that generate S_n . Then $A = S \cup C_0$ is a $O(n^6)$ -tripling generating set of size $\exp(c\sqrt{n \log n})$, but A is far from containing a subgroup and not dense in any conjugacy class.

Let us note that, by a result of Guralnick and Saxl [GS], if G is almost simple then we can choose a small subset $C_0 \subseteq C$ which generates $\langle C \rangle$, so $A = S \cup C_0$ generates $\langle S, C \rangle$. If S is a K -tripling subset of G then A is a K' -tripling subset for $K' = K + K|C| + |C|^3$, but for a suitable choice of S the generating set A is far from being dense in any conjugacy class.

These examples show in particular that the Product Theorem fails completely for finite simple groups of unbounded rank.

4. TOOLBOX

4.1. Basic arithmetic combinatorics. We review some basic theory of arithmetic combinatorics and in particular sets of small tripling. Most of what we need is in [H2, Sections 3 and 4]. Throughout we will work exclusively with sets of small tripling, so we will not review any results or terms related strictly to approximate groups.

For A, B subsets of a group G , we write AB for the product set

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

Define A^k for $k \geq 0$ to be the set of all products $a_1 \cdots a_k$ with $a_1, \dots, a_k \in A$. We also define $A^{-1} = \{a^{-1} : a \in A\}$ and $A^{-k} = (A^{-1})^k$. We write $A^{\pm k}$ for $(A \cup \{1\} \cup A^{-1})^k$. We call A *symmetric* if $A = A^{-1}$.

Lemma 4.1 (tripling lemma, see [H2, (3.3)]). *Let $A \subseteq G$ be finite, nonempty, and symmetric. For $k \geq 3$,*

$$|A^k|/|A| \leq (|A^3|/|A|)^{k-2}.$$

Lemma 4.2 (orbit–stabilizer for sets, see [H2, Lemma 4.1]). *Let $A, B \subseteq G$ be finite sets and $H \leq G$ a (not necessarily normal) subgroup. Let $\pi : G \rightarrow G/H$ be the quotient map.*

- (1) $|A^\pi||B \cap H| \leq |AB|.$
- (2) $|A| \leq |A^\pi||A^{-1}A \cap H|.$

Note that if $A = B$ is a subgroup then the lemma states $|A| = |A^\pi||A \cap H|$, which is the orbit–stabilizer theorem for the action of A on G/H . The lemma implies that there is in general still some relation between the “orbit” A^π and the “stabilizer” $A \cap H$.

Lemma 4.3 (growth in subgroups, quotients, and sections). *Let $A \subseteq G$ be nonempty, finite, symmetric, and K -tripling.*

- (1) *For $H \leq G$,*

$$\frac{|A^k \cap H|}{|A^2 \cap H|} \leq \frac{|A^{k+1}|}{|A|} \leq K^{k-1}.$$

- (2) *For $H \leq G$ and $\pi : G \rightarrow G/H$ the quotient map,*

$$\frac{|(A^k)^\pi|}{|A^\pi|} \leq \frac{|A^{k+2}|}{|A|} \leq K^k.$$

- (3) *For $\Sigma = H/N$ a section of G ,*

$$\frac{|\mathrm{tr}(A^k, \Sigma)|}{|\mathrm{tr}(A^2, \Sigma)|} \leq \frac{|A^{k+5}|}{|A|} \leq K^{k+3}.$$

Proof. These all follow from Lemmas 4.1 and 4.2. See [H2, Section 4] for the first two. The third is similar. Suppose $N \trianglelefteq H \leq G$ and let $\pi : H \rightarrow H/N$ be the quotient map. Applying Lemma 4.2(1) to $(A^k \cap H, A^4 \cap H)$ we obtain

$$|(A^k \cap H)^\pi| |A^4 \cap N| \leq |(A^k \cap H)(A^4 \cap H)| \leq |A^{k+4} \cap H|,$$

and applying Lemma 4.2(2) to $A^2 \cap H$ we obtain

$$|A^2 \cap H| \leq |(A^2 \cap H)^\pi| |(A^{-2} \cap H)(A^2 \cap H) \cap N| \leq |(A^2 \cap H)^\pi| |A^4 \cap N|.$$

Multiplying these inequalities we obtain

$$\frac{|(A^k \cap H)^\pi|}{|(A^2 \cap H)^\pi|} \leq \frac{|A^{k+4} \cap H|}{|A^2 \cap H|}.$$

Now applying (1) gives (3). \square

Lemma 4.4 (covering lemma, essentially [H2, Lemma 4.2]). *Suppose $A \subseteq G$ is finite and covered by k left cosets of $H \leq G$. Then A is covered by k left translates of $A^{-1}A \cap H$.*

Proof. Suppose $A \subseteq \bigcup_{i=1}^k x_i H$. We may assume $x_1, \dots, x_k \in A$. Then $A \cap x_i H = x_i(x_i^{-1}A \cap H) \subseteq x_i(A^{-1}A \cap H)$. Hence $A \subseteq \bigcup_{i=1}^k x_i(A^{-1}A \cap H)$. \square

The follow basic and self-evident rule will come up several times.

Lemma 4.5 (modular law). *Let $A, B \subseteq G$ and assume $B \subseteq \Gamma \leq G$. Then*

$$\Gamma \cap AB = (\Gamma \cap A)B.$$

More generally, for $A, B, C \subseteq G$,

$$C \cap AB \subseteq (CB^{-1} \cap A)B \subseteq CB^{-1}B \cap AB.$$

We need the following version of Schreier's lemma (see [HS, Lemma 3.8]).

Lemma 4.6 (Schreier). *Let G be a group and $A \subseteq G$ and $H \leq G$. Suppose $AH = G$. Then $\langle A \rangle \cap H = \langle A^{\pm 3} \cap H \rangle$.*

This lemma is best known in the context of finitely generated groups, as it implies that a finite-index subgroup of a finitely generated group is finitely generated. The elements of $A^{\pm 3} \cap H$ are sometimes called Schreier generators for H .

4.2. Quasirandomness. Next we recall some facts about quasirandomness. If G is a finite group let $\deg_{\mathbb{C}}(G)$ be the minimum degree of a nontrivial complex representation of G , conventionally ∞ if G is trivial. A group is colloquially called *quasirandom* if $\deg_{\mathbb{C}}(G)$ is large. Simple groups of Lie type are examples.

Theorem 4.7 (Landazuri–Seitz [LS]). *If L is a simple group of Lie type of characteristic p and rank ℓ then $\deg_{\mathbb{C}}(L) \geq |L|^{c/\ell}$ for a constant $c > 0$.*

The relevance of quasirandomness is indicated by the following well-known result, essentially due to Gowers (see also [G, NP]).

Proposition 4.8 ([BNP, Corollary 2.6]). *Let G be a finite group and $m = \deg_{\mathbf{C}}(G)$. If $A, B, C \subseteq G$ are sets such that*

$$|A||B||C| \geq |G|^3/m,$$

then

$$ABC = G.$$

In particular if $|A| \geq |G|/m^{1/3}$ then $A^3 = G$.

Lemma 4.9. *Let A be a symmetric subset of a finite group G such that A^k contains a coset of $N \trianglelefteq G$. If $|A^3| \leq K|A|$ and $\deg_{\mathbf{C}}(N) \geq K^{3k}$ then A^3 contains a coset of N . In particular if $N = G$ then $A^3 = G$.*

Proof. We may assume $k \geq 3$ and $K > 1$. Since A^k contains a coset of N , $|A^{k+1}| \geq |N||A^\pi|$, where $\pi : G \rightarrow G/N$ is the projection. Hence there is some $x \in G$ such that

$$|A \cap xN| \geq |A|/|A^\pi| \geq (|A|/|A^{k+1}|)|N| \geq K^{-k+1}|N|$$

by Lemma 4.1. Let $B = x^{-1}A \cap N$, so $|B| \geq K^{-k+1}|N|$. Then

$$A^3 \supseteq xBxBxB = x^3B^{x^2}B^xB$$

and, by Proposition 4.8, $B^{x^2}B^xB = N$. □

Lemma 4.10. *Let G be a finite perfect group. Then*

$$\deg_{\mathbf{C}}(G) \geq c \deg_{\mathbf{C}}(G/\text{Sol}(G))^{1/2}.$$

Proof. Consider a nontrivial complex representation G^π of G of degree d . Since G is perfect, G^π is not soluble. By [NP, Corollary 2.5], $G^\pi/\text{Sol}(G^\pi)$ embeds into $\text{Sym}(d_1)$ for some $d_1 \leq Cd^2$. Since $\text{Sol}(G)^\pi \leq \text{Sol}(G^\pi)$, $G/\text{Sol}(G)$ has a nontrivial permutation representation of degree d_1 . Hence

$$\deg_{\mathbf{C}}(G/\text{Sol}(G)) \leq Cd^2. \quad \square$$

4.3. Coprime action.

Theorem 4.11 (Schur–Zassenhaus, see [A, (18.1)]). *Let G be a finite group. Assume $U \trianglelefteq G$ and $\gcd(|U|, |G/U|) = 1$. Assume U or G/U is soluble. Then there is a complement T to U in G and all complements are conjugate.*

For the rest of this subsection assume U is a finite p -group and T is a finite p' -group acting on U . Let $G = U \rtimes T$. The following lemma generalizes a familiar property of vector spaces.

Lemma 4.12 ([A, (24.4–6)]). *$U = [U, T]C_U(T)$ and $[U, T] = [U, T, T]$. If U is abelian then $U = [U, T] \times C_U(T)$, and in particular $[U, T] = U$ if and only if $C_U(T) = 1$.*

In general the intersection $[U, T] \cap C_U(T)$ is nontrivial. For example, let

$$G = \left\{ \begin{pmatrix} 1 & x & z \\ 0 & a & y \\ 0 & 0 & 1 \end{pmatrix} : a \in \mathbf{F}_p^\times, x, y, z \in \mathbf{F}_p \right\} = TU,$$

where $T = G \cap T_3(\mathbf{F}_p)$ and $U = U_3(\mathbf{F}_p)$. Then $[U, T] = U$ and $C_U(T) = Z(U) \cong C_p$. However, the intersection $[U, T] \cap C_U(T)$ is always contained in the commutator subgroup of $[U, T]$.

Lemma 4.13. *Let $H = [U, T]$. Then $C_{H/H'}(T) = 1$.*

Proof. By Lemma 4.12, $H = [H, T]$. It follows that $\overline{H} = [\overline{H}, T]$ where $\overline{H} = H/H'$. On the other hand, by Lemma 4.12, $\overline{H} = [\overline{H}, T] \times C_{\overline{H}}(T)$. Hence $C_{\overline{H}}(T) = 1$. \square

Lemma 4.14. *Assume $\gamma_n(U) = \gamma_n(T) = 1$. Then $\gamma_\omega(G) = \gamma_n(G) = [U, T]$.*

Proof. By Lemma 4.12, $[U, T] = [U, T, T]$. It follows that $[U, T] \leq \gamma_\omega(G)$. On the other hand $G/[U, T] \cong U/[U, T] \times T$ is nilpotent of class at most n , so $[U, T] \geq \gamma_n(G)$. \square

5. FINITIZATION

The following well-known theorem of Mal'cev asserts that finitely generated linear groups are residually finite. More strongly, it asserts that we can distinguish the elements of any finite subset using a finite residue field. We will use it to reduce Theorem 1.2 and Theorem 1.4 to the finite field case.

Theorem 5.1 (Mal'cev, see [W1, Theorem 4.2]). *Let \mathbf{F} be a field, $G \leq \mathrm{GL}_n(\mathbf{F})$ a finitely generated subgroup, and $S \subseteq G$ a finite subset. Then there is a finite field \mathbf{K} and a homomorphism $\pi : G \rightarrow \mathrm{GL}_n(\mathbf{K})$ such that π is injective on S . If $\mathrm{char} \mathbf{F} > 0$ then $\mathrm{char} \mathbf{K} = \mathrm{char} \mathbf{F}$.*

Proposition 5.2. *If Theorem 1.2 holds for finite fields then it holds in general.*

Proof. Let \mathbf{F} be an arbitrary field and let $A \subseteq \mathrm{GL}_n(\mathbf{F})$ be finite, nonempty, symmetric, and K -tripling. Let $G = \langle A \rangle$. By Theorem 5.1, for any finite set $S \subseteq \mathrm{GL}_n(\mathbf{F})$ there is a finite field \mathbf{K} and a homomorphism $\pi : \langle A \cup S \rangle \rightarrow \mathrm{GL}_n(\mathbf{K})$ that is injective on S . Let $p = \mathrm{char} \mathbf{K}$. Note that $\langle \pi(A) \rangle = \pi(G)$. By Lemma 4.3(2), $\pi(A)$ is K^3 -tripling. Assuming Theorem 1.2 holds in the finite field case, there is $m \leq O_n(1)$ and $\Gamma_0 \trianglelefteq \pi(G)$ such that

- (i) $\pi(A)$ is covered by K^m cosets of Γ_0 ,
- (ii) $H_0 = \gamma_n(\Gamma_0)$ is contained in $\pi(A)^m$,
- (iii) H_0 has a perfect soluble-by-Lie $^*(p)$ normal subgroup P_0 contained in a translate of $\pi(A)^3$ and H_0/P_0 is a p -group.

Let $\Gamma = \pi^{-1}(\Gamma_0) \cap G$. Then $\Gamma \trianglelefteq G$, $\pi(\Gamma) = \Gamma_0$, and $\pi(\gamma_n(\Gamma)) = \gamma_n(\Gamma_0)$. We thus obtain the following properties:

- (1) A is covered by K^m cosets of Γ ,

- (2) $H_0 = \pi(\gamma_n(\Gamma))$ is contained in $\pi(A^m)$,
- (3) H_0 has a perfect soluble-by-Lie $^*(p)$ normal subgroup P_0 contained in a translate of $\pi(A)^3$ and H_0/P_0 is a p -group.

Moreover, by Lemma 4.4, A is covered by K^m translates of $A^2 \cap \Gamma$, and $P_0 \subseteq \pi(A)^6$, so we retain properties (1)–(3) if we replace Γ with the subgroup $\langle (A^6 \cap \Gamma)^G \rangle$ normally generated by $A^6 \cap \Gamma$. Thus we may append the following property:

- (4) $\Gamma = \langle (A^6 \cap \Gamma)^G \rangle$.

At the moment, Γ depends on S , but we can eliminate this dependence as follows. Call Γ *good for S* if there is a finite field \mathbf{K} (with $\text{char } K = \text{char } F$ if $\text{char } F > 0$) and some homomorphism $\pi : \langle A \cup S \rangle \rightarrow \text{GL}_n(\mathbf{K})$ injective on S such that (1)–(4) hold for Γ . By the argument above, for every finite set $S \subseteq \text{GL}_n(\mathbf{F})$ there is some subgroup $\Gamma \trianglelefteq G$ that is good for S . Moreover, if $S_1 \subseteq S_2$ and Γ is good for S_2 then Γ is good for S_1 . Since A^6 is finite, property (4) implies that there are only finitely many possibilities for Γ , say $\Gamma_1, \dots, \Gamma_N$. Suppose, for each i , Γ_i is not good for some finite set S_i . Then none of $\Gamma_1, \dots, \Gamma_N$ is good for the finite set $S = S_1 \cup \dots \cup S_N$, which is a contradiction. Thus we may assume that Γ is independent of S .

Now let $H = \gamma_n(\Gamma)$. If $x \in H$ then by applying (2) above with $S = A^m \cup \{x\}$ we obtain $x \in A^m$. Thus $H \subseteq A^m$. Moreover, since we may assume π is injective on A^m , π restricts to an isomorphism of $H = \gamma_n(\Gamma)$ with $H_0 = \pi(\gamma_n(\Gamma))$. Hence H has a perfect soluble-by-Lie $^*(p)$ normal subgroup $P \cong P_0$ such that H/P is a p -group, where $p = \text{char } \mathbf{K} > 0$. Also, $x_0 P_0 \subseteq \pi(A)^3$ for some $x_0 \in \pi(G)$. In particular $x_0 \in \pi(A)^3 = \pi(A^3)$, so $x_0 = \pi(x)$ for some $x \in A^3$ and $\pi(xP) \subseteq \pi(A^3)$. Since we may assume π is injective on $A^3 P$, this implies $xP \subseteq A^3$. Thus P is contained in a translate of A^3 .

If $\text{char } \mathbf{F} > 0$ then $p = \text{char } \mathbf{F}$ and we are done. If $\text{char } \mathbf{F} = 0$ then by choosing S to include many transvections $1 + xe_{12}$ ($x \in \mathbf{Z}$) we can force p to be larger than $|H|$, which implies that H is trivial. \square

We can reduce Theorem 1.4 to the finite field case using almost the same argument.

Proposition 5.3. *If Theorem 1.4 holds for finite fields then it holds in general.*

Proof. We argue exactly as in the proof of Proposition 5.2. We obtain $\Gamma \trianglelefteq G$ and $m = O(\log n)$ such that, for any finite subset $S \subseteq \text{GL}_n(\mathbf{F})$, there is a finite field \mathbf{K} (with $\text{char } K = \text{char } F$ if $\text{char } F > 0$) and a homomorphism $\pi : \langle A \cup S \rangle \rightarrow \text{GL}_n(\mathbf{K})$ injective on S such that

- (1) A is covered by $K^{O_n(1)}$ cosets of Γ
- (2) $P_0 = \pi(\Gamma)^{(m)}$ is perfect, soluble-by-Lie $^*(p)$, where $p = \text{char } K$, and contained in a translate of $\pi(A)^3$.

Let $P = \Gamma^{(m)}$ and note $\pi(P) = P_0$. By (2), there is some $x_0 \in \pi(G)$ such that $x_0 P_0 \subseteq \pi(A)^3$. In particular $x_0 \in \pi(A)^3 = \pi(A^3)$, so $x_0 = \pi(x)$ for

some $x \in A^3$, and $\pi(xP) \subseteq \pi(A^3)$. A priori x depends on S , but since $x \in A^3$ and A^3 is finite we can eliminate this dependence as in the previous proof. Now if $y \in P$ then by taking $S = A^3 \cup \{xy\}$ it follows that $xy \in A^3$. Thus $xP \subseteq A^3$. In particular P is finite, and taking $S \supseteq P$ gives $P \cong \pi(P) = P_0$. Thus P is perfect, soluble-by-Lie $^*(p)$, and contained in a translate of A^3 . If $\text{char } F > 0$ then $p = \text{char } F$ and we are done, and otherwise we may choose S to force p to be larger than $|P|$, which shows that P must be trivial. \square

6. TRIGONALIZATION

In this section and the next we extensively use arguments from the theory of linear groups, particularly soluble linear groups. A nice general reference on linear groups is [W1].

Call a subgroup of $\text{GL}_n(\mathbf{F})$ *trigonal* if it is contained in $B_n(\mathbf{F})$, and *trigonalizable* if it is conjugate to a subgroup of $B_n(\mathbf{E})$ for some extension \mathbf{E} of \mathbf{F} . Virtually soluble is equivalent to virtually trigonalizable, by another well-known theorem of Mal'cev.

Theorem 6.1 (Mal'cev, see [W1, Theorem 3.6]). *Every soluble subgroup G of $\text{GL}_n(\mathbf{F})$ has a trigonalizable normal subgroup G_0 of index $O_n(1)$.*

It follows immediately from Theorem 6.1 that any soluble subgroup of $\text{GL}_n(\mathbf{F})$ has derived length $O_n(1)$, a result originally proved by Zassenhaus. This bound was sharpened by Newman to $O(\log n)$: see [W1, Chapter 3]. The following lemma generalizes this bound to virtually soluble subgroups.

Lemma 6.2. *Let Γ be a finite subgroup of $\text{GL}_n(\mathbf{F})$.*

- (1) *For any $m \geq 0$ there is a soluble subgroup $S \leq \Gamma$ such that $\Gamma = \Gamma^{(m)}S$.*
- (2) *There is $m = O(\log n)$ such that $\Gamma^{(m)}$ is perfect.*

Proof. (1) Let $P = \Gamma^{(m)}$. Let S be a minimal subgroup of Γ such that $\Gamma = PS$. By minimality of S , if $M < S$ then $\Gamma > PM = P(P \cap S)M$, so $(P \cap S)M < S$. Hence $P \cap S \leq \text{Frat}(S)$. This implies that $P \cap S$ is nilpotent. On the other hand $S/(P \cap S) \cong PS/P = \Gamma/P$ is soluble. Hence S is soluble.

(2) For soluble Γ this was proved by Newman, as mentioned. In general let m be large enough for the soluble case and let Γ be arbitrary. Let $P = \Gamma^{(m)}$. By (1), there is soluble $S \leq \Gamma$ such that $\Gamma = P'S$. Then $P = \Gamma^{(m)} \leq P'S^{(m)} = P'$, so P is perfect. \square

We say a group G is *p-by-abelian* if $G/O_p(G)$ is abelian; if G is finite then by Schur–Zassenhaus (Theorem 4.11) this is equivalent to a semidirect decomposition $G = O_p(G)T$ for some abelian p' -group $T \leq G$.

Lemma 6.3. *Let \mathbf{F} be a field of characteristic $p > 0$. A finite subgroup $G \leq \text{GL}_n(\mathbf{F})$ is trigonalizable if and only if it is p-by-abelian. In this case $G/O_p(G)$ is the direct product of at most n cyclic groups and $\gamma_n(O_p(G)) = 1$.*

Proof. We may assume \mathbf{F} is algebraically closed. If G is trigonalizable then without loss of generality $G \leq B_n(\mathbf{F})$, so $O_p(G) = G \cap U_n(\mathbf{F})$ and $G/O_p(G) \cong GU_n(\mathbf{F})/U_n(\mathbf{F})$ is isomorphic to a subgroup of $T_n(\mathbf{F}) \cong (\mathbf{F}^\times)^n$, hence a direct product of at most n cyclic groups.

Conversely suppose G is p -by-abelian. Let $U = O_p(G)$. Let V be a G -composition factor of \mathbf{F}^n . Then V is an irreducible $\mathbf{F}G$ -module, so by Clifford's theorem it is a completely reducible $\mathbf{F}U$ -module, but the only irreducible $\mathbf{F}U$ -module is the trivial one-dimensional module, so U acts trivially on V . Hence V is an irreducible $\mathbf{F}(G/U)$ -module, which implies that $\dim V = 1$ since G/U is abelian and \mathbf{F} is algebraically closed. \square

For $N \trianglelefteq \Gamma \leq \mathrm{GL}_n(\mathbf{F})$, we say $\Sigma = \Gamma/N$ is a *trigonalizable section* if $\Gamma = BN$ for some trigonalizable subgroup $B \leq \mathrm{GL}_n(\mathbf{F})$. By the previous lemma, such a subgroup B must be p -by-abelian, so Σ must be p -by-abelian. The converse does not hold. For example, if Γ is a nonabelian nilpotent p' -group contained in $\mathrm{GL}_n(\mathbf{F}_p)$ then the section $\Sigma = \Gamma/\Gamma'$ is abelian but not covered by a $(p$ -by-)abelian subgroup of Γ .

The following key proposition establishes a variant of Mal'cev's theorem for soluble sections. Moreover we have control over the “Weyl group” of the trigonalizable subsection.

Theorem 6.4. *Let \mathbf{F} be a finite field of characteristic p . Let $\Sigma = \Gamma/N$ be a soluble section of $L = \mathrm{GL}_n(\mathbf{F})$. Then there is a trigonalizable subsection $\Sigma_0 \trianglelefteq \Sigma$ of index $O_n(1)$ containing $O_p(\Sigma)$. Moreover, if $R = N_L(\Sigma) = N_L(\Gamma) \cap N_L(N)$ is the section normalizer then $[R : C_R(\Sigma_0/O_p(\Sigma))] \leq n!$.*

Proof. We will establish the following properties in order (for $\Sigma_0 \trianglelefteq \Sigma$ of index $O_n(1)$ to be determined):

- (1) Σ_0 is p -by-abelian,
- (2) $\Sigma_0 = \pi(B)$ for some p -by-abelian subgroup $B \leq \Gamma$,
- (3) $B = O_p(B)T$ for some abelian p' -group T ,
- (4) if $g \in R$ then $B^g = B^\delta$ and $T^g = T^\delta$ for some $\delta \in \pi^{-1}(O_p(\Sigma))$,
- (5) $[R : C_R(\Sigma_0/O_p(\Sigma))] \leq n!$.

By Lemma 6.2(1), there is a soluble subgroup $S \leq \Gamma$ such that $\pi(S) = \Sigma$. By Theorem 6.1, there is $S_0 \trianglelefteq S$ of index $O_n(1)$ which is trigonalizable. By Lemma 6.3, $S_0 = O_p(S_0)T_0$ for some abelian p' -group $T_0 \leq S_0$ isomorphic to the direct product of at most n cyclic groups. Then $\pi(S_0) = \pi(O_p(S_0))\pi(T_0)$ is a p -by-abelian normal subgroup of Σ of index $O_n(1)$. Since $\pi(S_0)$ is p -by-abelian, its p -core $O_p(\pi(S_0))$ is the unique Sylow p -subgroup of $\pi(S_0)$, so it is normal in Σ , so it is contained in $O_p(\Sigma)$. Hence the product $\Sigma_0 = O_p(\Sigma)\pi(T_0)$ is a p -by-abelian subgroup of index $O_n(1)$. In particular, since T_0 is n -generated, $\Sigma/O_p(\Sigma)$ is $O_n(1)$ -generated, so there are only $O_n(1)$ subgroups of Σ containing $O_p(\Sigma)$ of index $[\Sigma : \Sigma_0]$. Their intersection Σ_1 is characteristic and p -by-abelian. This proves (1) with Σ_1 in the role of Σ_0 .

Write $\Sigma_1 = \Gamma_1/N$ and $O_p(\Sigma) = \Delta/N$. Let P be a Sylow p -subgroup of Δ . Then $\pi(P) = O_p(\Sigma)$, so $\Delta = PN$, and by the Frattini argument

$$\Gamma_1 = N_{\Gamma_1}(P)\Delta = N_{\Gamma_1}(P)PN = N_{\Gamma_1}(P)N.$$

Hence $\pi(N_{\Gamma_1}(P)) = \Sigma_1$. Since Γ_1/Δ is a p' -group, $N_{\Gamma_1}(P)/P$ is a p' -group, so $N_{\Gamma_1}(P) = PH$ for some p' -group H by Schur–Zassenhaus. Since H is a p' -group, Jordan’s theorem (see [W1, Theorems 9.2 and 9.3]) implies that H has an abelian subgroup H_0 of index $O_n(1)$. Being an abelian p' -subgroup of $\mathrm{GL}_n(\mathbf{F})$, H_0 is n -generated (e.g., by Lemma 6.3), and $[H : H_0]$ is $O_n(1)$, so H is $O_n(1)$ -generated and has only $O_n(1)$ subgroups of index $[H : H_0]$. Thus, replacing H_0 with the intersection of these subgroups, we may assume $H_0 \trianglelefteq H$. The image $\pi(PH_0)$ also has index $O_n(1)$ in $\pi(PH) = \Sigma_1$. Let $\Sigma_2 = \Gamma_2/N$ be the intersection of all subgroups of Σ_1 containing $O_p(\Sigma)$ of index $[\Sigma_1 : \pi(PH_0)]$. Then $O_p(\Sigma) \leq \Sigma_2 \trianglelefteq \Sigma_1$, $[\Sigma_1 : \Sigma_2] \leq O_n(1)$, and $\Sigma_2 \leq \pi(PH_0)$. Let $T = H_0 \cap \Gamma_2$ and $B = PT = PH_0 \cap \Gamma_2$. Then $\pi(B) = \Sigma_2$, so (1), (2), and (3) hold with Σ_2 in the role of Σ_0 .

The normalizer $R = N_L(\Sigma)$ preserves Γ , N , and Σ (by definition), the characteristic subgroups $O_p(\Sigma) \trianglelefteq \Sigma_2 \trianglelefteq \Sigma_1 \trianglelefteq \Sigma$, and also their preimages $\Delta \trianglelefteq \Gamma_2 \trianglelefteq \Gamma_1 \trianglelefteq \Gamma$. Additionally $N_R(P)$ preserves P and $N_{\Gamma_1}(P) = PH$ as well as $PH_0 \trianglelefteq PH$ and $PT = PH_0 \cap \Gamma_2$. By Sylow’s theorem Δ is transitive on its Sylow p -subgroups, so $R = N_R(P)\Delta$. By Schur–Zassenhaus, P is transitive on its complements in PT , so $N_R(P) = N_{N_R(P)}(T)P$. Thus

$$R = N_R(P)\Delta = N_{N_R(P)}(T)P\Delta = N_{N_R(P)}(T)\Delta.$$

This proves (4).

Finally, let $W = R/C_R(\Sigma_2/O_p(\Sigma))$. Then W acts faithfully on $\Sigma_2/O_p(\Sigma) = \Gamma_2/\Delta$. By (4), $R = N_R(T)\Delta$. On the other hand since $\Sigma_2 = O_p(\Sigma)\pi(T)$ we have $C_R(T)\Delta \leq C_R(\Sigma_2/O_p(\Sigma))$. Hence W is a section of $N_R(T)\Delta/C_R(T)\Delta$, which is isomorphic to a section of $N_L(T)/C_L(T)$: indeed, there is a natural surjective map $N_R(T)/C_R(T) \rightarrow N_R(T)\Delta/C_R(T)\Delta$ as well as a natural injective map $N_R(T)/C_R(T) \rightarrow N_L(T)/C_L(T)$. Hence it suffices to prove $[N_L(T) : C_L(T)] \leq n!$.

Since T is an abelian p' -group, it is diagonalizable over $\overline{\mathbf{F}}$ (see, e.g., [W1, Corollaries 1.3 and 1.6]). Let V_1, \dots, V_k be the eigenspaces. Then $k \leq n$ and $C_L(T)$ consists of those elements of L which map each V_i into itself. Elements of $N_L(T)$ must permute V_1, \dots, V_k . Thus $N_L(T)/C_L(T)$ permutes V_1, \dots, V_k faithfully, so $[N_L(T) : C_L(T)] \leq k! \leq n!$. This proves (5). \square

Although the bound $[R : C_R(\Sigma_0/O_p(\Sigma))] \leq n!$ is all we need in this paper, the following more general result is included for independent interest. If Σ is a section of $L = \mathrm{GL}_n(\mathbf{F})$ then the *Weyl group* of Σ is $W(\Sigma) = N_L(\Sigma)/C_L(\Sigma)$.

Corollary 6.5. *If Σ is an abelian p' -section of $\mathrm{GL}_n(\mathbf{F})$ then the Weyl group $W(\Sigma)$ has order $O_n(1)$.*

Proof. Let $W = W(\Sigma)$. By Theorem 6.4, there is a trigonalizable subgroup $\Sigma_0 \trianglelefteq \Sigma$ of index $O_n(1)$ such that $[W : C_W(\Sigma_0)] \leq n!$. Since $W/C_W(\Sigma/\Sigma_0)$

is isomorphic to a subgroup of $\text{Aut}(\Sigma/\Sigma_0)$, it has order $O_n(1)$. Hence $W_0 = C_W(\Sigma_0) \cap C_W(\Sigma/\Sigma_0)$ has index $O_n(1)$ in W . But W_0 can be identified with a subgroup of $\text{Hom}(\Sigma/\Sigma_0, \Sigma_0)$. Since Σ_0 is a direct product of at most n cyclic groups, it follows that W_0 has order $O_n(1)$. \square

7. MAIN PROOF PART 1: GENERAL TO SOLUBLE

7.1. Structure of finite linear groups. In this section we need a great deal of information about finite groups of Lie type. A good reference is [KL, Chapter 5].

Recall from the introduction that $\text{Lie}(p)$ is the class of finite simple groups of Lie type of characteristic p and $\text{Lie}^*(p)$ is the class of finite direct products of members of $\text{Lie}(p)$. We will write $\text{Lie}^n(p)$ for the class of direct products of at most n simple groups of Lie type of characteristic p and of rank at most n . A *quasisimple group* (of Lie type of characteristic p) is a perfect central extension of some finite simple group (of Lie type of characteristic p).

Lemma 7.1. *Suppose $\Sigma = \Gamma/N$ is a $\text{Lie}^*(p)$ section of $\text{GL}_n(\mathbf{F})$, where \mathbf{F} is a finite field of characteristic p . Then Σ is $\text{Lie}^n(p)$. More precisely, Σ has at most $n/2$ simple factors and each has Lie rank at most $n - 1$.*

Proof. The bound on the number of factors is proved in [LP3, Corollary 3.3], while the bound on the Lie ranks follows from [FT] and [KL, Proposition 5.2.12]. Indeed, let S be one of the $\text{Lie}(p)$ factors of Σ and let $\gamma : \Gamma \rightarrow S$ be the natural projection homomorphism. Let H be a minimal subgroup of Γ such that $\gamma(H) = S$ and let $\lambda : H \rightarrow \text{PGL}_m(\mathbf{F})$ be a nontrivial projective representation of H of minimal degree. Then $m \leq n$. By the main result of [FT] (see also [FT, Proposition 4.1]), λ factorizes through S . Thus S embeds in $\text{PGL}_m(\mathbf{F})$, and so [KL, Proposition 5.2.12(i)] implies that S has Lie rank at most $m - 1$. \square

We will need the following deep theorem on the structure of finite linear groups. It was proved by Weisfeiler [W2] using the classification of finite simple groups, and later by Larsen and Pink [LP1] without the classification.

Theorem 7.2 (Weisfeiler [W2, Theorem 1], Larsen–Pink [LP1, Theorem 0.2]). *Let \mathbf{F} be a field of characteristic $p > 0$ and let G be a finite subgroup of $\text{GL}_n(\mathbf{F})$. Then G has a normal subgroup Γ of index $O_n(1)$ containing $O_p(G)$ such that $\Gamma/O_p(G)$ is a central extension of a member of $\text{Lie}^*(p)$.*

Remark 7.3. Unfortunately this statement is not quite explicit in either [W2, Theorem 1] or [LP1, Theorem 0.2].

According to [W2, Theorem 1], $G/O_p(G)$ has a subgroup $\Gamma = TL$ of index $O_n(1)$ such that T is an abelian p' -group and L is a central extension of a group of $\text{Lie}^*(p)$ type. However, it is explicit in the proof on p. 5279 (though not in the statement of the theorem) that T is the centre of Γ , so Γ itself is a central extension of a $\text{Lie}^*(p)$ group.

Similarly, according to [LP1, Theorem 0.2], G has normal subgroups $\Gamma_3 \leq \Gamma_2 \leq \Gamma_1 \leq G$ such that Γ_1 has index $O_n(1)$, $\Gamma_1/\Gamma_2 \in \text{Lie}^*(p)$, Γ_2/Γ_3 is an abelian p' -group, and Γ_3 is a p -group. We may assume $\Gamma_3 = O_p(G)$ by replacing Γ_i with $\Gamma_i O_p(G)$ for $i = 1, 2, 3$. Now see the definition of Γ_2 on p. 1156 and the last line of the proof of Theorem 0.2 for the fact that $\Gamma_2/\Gamma_3 = Z(\Gamma_1/\Gamma_3)$. (See also [LP2, Corollary 3.1].)

A central extension of a $\text{Lie}^*(p)$ group is the same as a central product of an abelian group and a set of quasisimple groups of Lie type of characteristic p : see [A, (31.1)].

Corollary 7.4. *Let \mathbf{F} be a finite field of characteristic $p > 0$ and let P be a subgroup of $\text{GL}_n(\mathbf{F})$ with $\deg_{\mathbf{C}}(P)$ sufficiently large in terms of n . Then*

- (a) P is soluble-by- $\text{Lie}^*(p)$,
- (b) $N = [P, \text{Sol}(P)]$ is a p -subgroup such that $[P, N] = N$,
- (c) $|\text{Sol}(P)/N| \leq (2n + 1)^n$.

Proof. Apply Theorem 7.2 to P . Since P has no proper normal subgroups of index less than $\deg_{\mathbf{C}}(P)$ and $\deg_{\mathbf{C}}(P)$ is sufficiently large, it follows that $P/O_p(P)$ is a central extension of some $\Gamma \in \text{Lie}^*(p)$. Hence (a) holds, and $N = [P, \text{Sol}(P)] \leq O_p(P)$, so N is a p -group. For X any normal subgroup of a perfect group P we have $[X, P, P] = [X, P]$ by the three subgroup lemma (see [A, (8.9)]), so (b) holds.

By Lemma 7.1, P is soluble-by- $\text{Lie}^n(p)$. Suppose $P/\text{Sol}(P) = L_1 \times \cdots \times L_t$, where $t \leq n$, $L_1, \dots, L_t \in \text{Lie}(p)$, and the Lie rank of L_i is at most n for each i . Since P/N is a perfect central extension of $P/\text{Sol}(P)$,

$$|\text{Sol}(P)/N| \leq |M(P/\text{Sol}(P))| = \prod_{i=1}^t |M(L_i)|,$$

where $M(L)$ denotes the Schur multiplier of L (see [A, Section 33 and Exercise 11.2]). By [KL, Theorem 5.1.4], $|M(L_i)| \leq 2n + 1$ provided that $|L_i|$ is larger than some constant, which we may assume since $\deg_{\mathbf{C}}(P)$ is sufficiently large. Hence (c) holds. \square

Remark 7.5. In fact $N = O_p(P)$. This follows from the fact that, with finitely many exceptions, if Γ is a finite simple group of Lie type of characteristic p then $|M(\Gamma)|$ is prime to p (see [KL, Theorem 5.1.4]).

7.2. Reduction to soluble-by-Lie*. The goal of this section is to prove the upside-down version of Theorem 1.4, Theorem 1.7. The following statement is slightly stronger: it asserts that we may additionally assume the $\text{Lie}^*(p)$ part of Γ is highly quasirandom.

Theorem 7.6. *Let \mathbf{F} be a finite field of characteristic $p > 0$. Let $A \subseteq \text{GL}_n(\mathbf{F})$ be a symmetric subset such that $|A^3| \leq K|A|$, where $K \geq 2$. Let $d \geq 1$. Then there is a normal subgroup $\Gamma \trianglelefteq \langle A \rangle$ such that*

- (1) A is covered by $K^{O_n(d)}$ cosets of Γ ,
- (2) Γ is soluble-by- $\text{Lie}^n(p)$

- (3) $\Gamma/\text{Sol}(\Gamma)$ is covered by A^6 , and
 (4) $\deg_{\mathbf{C}}(\Gamma/\text{Sol}(\Gamma)) \geq K^d$.

Remark 7.7. Theorem 7.6 is also true when \mathbf{F} is infinite, but we do not need this generalization. It is even true in characteristic zero, where $\text{Lie}^*(0)$ is interpreted as trivial; in this case the theorem simply states that A is covered by $K^{O_n(1)}$ cosets of a soluble normal subgroup (this was first proved in [BGT1]).

This result is a relatively direct consequence of Theorem 7.2 and Theorem 1.5 (the Product Theorem). Indeed, already by Theorem 7.2, $G = \langle A \rangle$ is covered by $O_n(1)$ cosets of a soluble-by- $\text{Lie}^*(p)$ normal subgroup $\Gamma \trianglelefteq G$. To get such a subgroup Γ for which A^6 covers $\Gamma/\text{Sol}(\Gamma)$ we need one auxiliary result on $\text{Lie}^*(p)$ groups.

Lemma 7.8. *Let $\Gamma \in \text{Lie}^n(p)$, and let A be a symmetric generating set of Γ that projects onto all simple quotients of Γ . Then $A^{O_n(1)} = \Gamma$.*

Proof. By induction it suffices to prove that if A generates a group of the form $H \times L$ with $L \in \text{Lie}(p)$ and A projects onto both H and L then $A^{O(\ell)} = H \times L$, where ℓ is the Lie rank of L . If A is a subgroup we are done, since A is a generating set by hypothesis. Otherwise let $x \in A^2 \setminus A$ and let $y \in A$ such that x and y have the same projection in H . Then $a = xy^{-1}$ is a nontrivial element of $A^3 \cap L$. By [LL, Theorem 1] (and recalling that untwisted Lie rank is at most twice the Lie rank), every element of L is the product of m conjugates of a for some $m \leq O(\ell)$. Since A projects onto L , it follows that A^{5m} contains L . Since A projects onto H , $A^{5m+1} = H \times L$. \square

Now we can prove Theorem 7.6.

Proof of Theorem 7.6. By Theorem 7.2 and Lemma 7.1, $G = \langle A \rangle$ has a soluble-by- $\text{Lie}^n(p)$ normal subgroup Γ of index $m \leq O_n(1)$. We claim that G has a normal subgroup Δ such that $\text{Sol}(\Gamma) \leq \Delta \leq \Gamma$ and $\Delta/\text{Sol}(\Gamma)$ is covered by A^6 and A is covered by $K^{O_n(1)}$ cosets of Δ .

By Lemma 4.6, Γ is generated by $B = A^{3m} \cap \Gamma$. By Lemma 4.3, B has tripling $K^{O_n(1)}$.

Write \mathcal{L} for the set of simple factors of $\Gamma/\text{Sol}(\Gamma)$. For $L \in \mathcal{L}$ denote by $\pi_L : \Gamma \rightarrow L$ the natural projection. Since $\Gamma \trianglelefteq G$, G permutes \mathcal{L} . We claim that $\mathcal{L} = \mathcal{L}_1 \cup \mathcal{L}_2$ for G -invariant sets $\mathcal{L}_1, \mathcal{L}_2$ such that

$$\begin{aligned} \pi_L(B)^3 = L \text{ and } \deg_{\mathbf{C}}(L) &\geq K^d & (L \in \mathcal{L}_1) \\ |\pi_L(B)| &\leq K^{O_n(d)} & (L \in \mathcal{L}_2). \end{aligned}$$

Let $\mathcal{L}_1 \subseteq \mathcal{L}$ be the largest G -invariant set such $\pi_L(B)^3 = L$ and $\deg_{\mathbf{C}}(L) \geq K^d$ for every $L \in \mathcal{L}_1$ and let $\mathcal{L}_2 = \mathcal{L} \setminus \mathcal{L}_1$. It suffices to show $|\pi_L(B)| \leq K^{O_n(d)}$ for all $L \in \mathcal{L}_2$. Suppose $L_1 \in \mathcal{L}$ satisfies $\pi_{L_1}(B)^3 \neq L_1$. By Lemma 4.3, $\pi_{L_1}(B) \subseteq L_1$ has tripling $K^{O_n(1)}$, so Theorem 1.5 implies that $|\pi_{L_1}(B)| \leq K^{O_n(1)}$. Similarly, if $\deg_{\mathbf{C}}(L_1) < K^d$ then $|L_1| \leq \deg_{\mathbf{C}}(L_1)^{O(n)} \leq K^{O_n(d)}$.

by Theorem 4.7, so certainly $|\pi_{L_1}(B)| \leq K^{O_n(d)}$. Now for any $L \in \mathcal{L}_1$ and $a \in A$ we have, by Lemma 4.3(3),

$$\begin{aligned} |\pi_{L^a}(B)| &= |\pi_L(B^{a^{-1}})| \leq |\pi_L(A^{3m+2} \cap \Gamma)| \leq K^{O_n(1)} |\pi_L(A^2 \cap \Gamma)| \\ &\leq K^{O_n(1)} |\pi_L(B)|. \end{aligned}$$

Since $|\mathcal{L}| \leq n$ and G is generated by A , it follows that $|\pi_L(B)| \leq K^{O_n(d)}$ for every L in the G -orbit of L_1 . Since \mathcal{L}_2 is the union of such G -orbits, the claim holds.

Let Γ_1 be the subgroup of Γ corresponding to the product of the factors in \mathcal{L}_1 . Since the projection of B to any factor of Γ/Γ_1 (i.e., any $L \in \mathcal{L}_2$) has size $K^{O_n(d)}$, B is covered by $K^{O_n(d)}$ cosets of Γ_1 . By Lemma 4.4, A is also covered by $K^{O_n(d)}$ cosets of Γ_1 . Since B^3 projects onto each factor of $\Gamma_1/\text{Sol}(\Gamma)$, B^k projects onto $\Gamma_1/\text{Sol}(\Gamma)$ by Lemma 7.8 for some $k \leq O_n(1)$. In particular A^{3mk} covers $\Gamma_1/\text{Sol}(\Gamma)$. Applying Lemma 4.3(3) with $\Sigma = \Gamma_1/\text{Sol}(\Gamma)$, it follows that

$$|\text{tr}(A^2, \Sigma)| \geq |\text{tr}(A^{3mk}, \Sigma)|/K^{3mk+3}.$$

Since $\deg_{\mathbb{C}}(\Sigma) \geq K^d$, Proposition 4.8 implies that $\text{tr}(A^2, \Sigma)^3 = \Sigma$, provided that $d \geq 3(3mk + 3)$, which we may assume. Thus A^6 covers $\Gamma_1/\text{Sol}(\Gamma)$. This completes the proof of the theorem with Γ_1 in the role of Γ . \square

7.3. Covering the perfect core. In this section we start with the soluble-by-Lie $^*(p)$ group Γ furnished by Theorem 1.7. By Lemma 6.2, Γ has derived length $O(\log n)$. Let $P = \Gamma^{(\omega)} = \Gamma^{(O(\log n))}$ be the perfect core of Γ . Our goal is to prove that $A^{O_n(1)}$ covers P . It is easy to show that $A^{O_n(1)}$ covers $P/\text{Sol}(P)$, by iterating the following lemma. The main work of this section consists of showing that $A^{O_n(1)}$ also covers $\text{Sol}(P)$.

Lemma 7.9. *Let Γ be soluble-by-Lie $^*(p)$. Let $A \subseteq \Gamma$ be a symmetric set covering $\Gamma/\text{Sol}(\Gamma)$. Then $A^{O(1)}$ covers $\Gamma'/\text{Sol}(\Gamma')$.*

Proof. Let $\bar{\Gamma} = \Gamma/\text{Sol}(\Gamma)$. Since A covers $\Gamma/\text{Sol}(\Gamma)$, the projection of $A^4 \cap \Gamma'$ to $\bar{\Gamma}$ contains the set of commutators C . By the weak Ore conjecture, $\bar{\Gamma} = C^{O(1)}$ (see [W3, Proposition 2.4] or [S1] or [NP, Theorem 3]). Hence $(A^4 \cap \Gamma')^{O(1)}$ covers $\Gamma/\text{Sol}(\Gamma)$. Since $\Gamma' \cap \text{Sol}(\Gamma) = \text{Sol}(\Gamma')$, this implies that $(A^4 \cap \Gamma')^{O(1)}$ covers $\Gamma'/\text{Sol}(\Gamma')$. \square

We need the following elementary result of Rhemtulla: see [R, Lemma 2].

Lemma 7.10 (Rhemtulla). *If $\Gamma = V \rtimes G$ where V is abelian and $G = \langle x_1, \dots, x_d \rangle$ then*

$$[V, G] = \{[v_1, x_1] \cdots [v_d, x_d] : v_1, \dots, v_d \in V\}.$$

Lemma 7.11 (affine conjugating trick, Lemma 1.8 restated). *Let $\Gamma = V \rtimes G$ be the semidirect product of an abelian group V and a d -generated finite group G with $\deg_{\mathbb{C}}(G) \geq K^{21}$. Let $A \subseteq V$ be a symmetric G -invariant set generating V . If $|A^3| \leq K|A|$ then $A^{7d} \supseteq [V, G]$.*

Proof. Let $B = AG$. Since $B^n = A^n G$ for all $n \neq 0$, $|B^3| \leq K|B|$ and B is a symmetric generating set for Γ .

We claim that B^6 contains all conjugates of G . It suffices to prove if G_0 is a conjugate of G contained in B^6 and $b \in B$ then $G_1 = G_0^b \subseteq B^6$. Certainly $G_1 \subseteq B^8$, so by Lemma 4.3(1),

$$|G_1| = |B^8 \cap G_1| \leq K^7 |B^2 \cap G_1|.$$

Hence $(B^2 \cap G_1)^3 = G_1$ by Proposition 4.8, so $B^6 \supseteq G_1$, as required.

Hence if $g \in G$ and $v \in V$ the element $[v, g] = (g^{-1})^v g$ is contained in $B^6 B \cap V = A^7$. Let x_1, \dots, x_d be generators of G . Then

$$A^{7d} \supseteq \{[v_1, x_1] \cdots [v_d, x_d] : v_1, \dots, v_d \in V\}.$$

Finally, $\{[v_1, x_1] \cdots [v_d, x_d] : v_1, \dots, v_d \in V\} = [V, G]$ by Lemma 7.10. \square

Proposition 7.12. *For each $d \geq 0$ there is a constant $m = m(d)$ such that the following holds. Let $N \leq G$ be finite normal subgroups of a group Γ such that*

- (i) $[G, N] = N$,
- (ii) N is nilpotent,
- (iii) G/N is d -generated, and
- (iv) $\deg_{\mathbf{C}}(G) \geq K^m$,

where $K \geq 2$. Let A be a finite symmetric K -tripling set generating Γ and covering G/N .

- (1) If $G = \Gamma$ then $A^3 = G$.
- (2) In general $(A^2 \cap G)^3 = G$.

Proof. Call A hereditarily K -tripling if $|(A^\pi)^3| \leq K|A^\pi|$ for every quotient $\pi : \Gamma \rightarrow \Gamma/\ker \pi$ of Γ . It follows from Lemma 4.3(2) that A is hereditarily K^3 -tripling. Hence it suffices to prove the proposition for hereditarily K -tripling sets. This observation allows us to use induction.

We begin with (1). We argue by induction on $|N|$. If $N = 1$ then the claim is trivial because A covers G/N by hypothesis, so assume $N \neq 1$. If V is a nontrivial normal subgroup of G contained in N then the hypotheses hold for G/V , so by induction $A^3 V = G$.

Suppose we can find a nontrivial commutator $[a, b] \in Z(G) \cap N$. Let

$$V = \langle [a, b] \rangle = \{[a^k, b] : k \in \mathbf{Z}\}.$$

Then V is normal in G and contained in N so by induction $G = A^3 V$. In particular $G = A^3 Z(G)$, so all commutators are contained in A^{12} . In particular $V \subseteq A^{12}$, so $G = A^{15}$. Hence $A^3 = G$ by Lemma 4.9, assuming $m \geq 45$.

Hence assume no nontrivial commutator is in $Z(G) \cap N$. Since $[G, N] = N$, N is not contained in $Z(G)$. Let $V \leq N$ be a minimal normal subgroup of G not contained in $Z(G)$. If $[V, G] < V$ then $[V, G] \leq Z(G) \cap N$ by minimality of V , but this contradicts the assumption that no nontrivial commutator is in $Z(G) \cap N$. Since N is nilpotent, $[V, N] < V$, for otherwise $\gamma_i(N) \geq V$ for

all i . Hence $[V, N] \leq Z(G) \cap N$ by minimality of V again, a contradiction unless $[V, N] = 1$. Hence $[V, G] = V$ and $V \leq Z(N)$.

In particular V is an abelian group (since $V \leq Z(N)$) and the conjugation action of G on V factors through G/N (since $[V, N] = 1$), so we may identify V with a $\mathbf{Z}(G/N)$ -module satisfying $[V, G/N] = V$.

Next we note that A^7 contains an element $x \in V \setminus Z(G)$. If $V \cap Z(G) \neq 1$ then by induction $A^3(V \cap Z(G)) = G$, so A^3 must contain an element of $V \setminus Z(G)$. If $V \cap Z(G) = 1$ then, since $A^3V = G$ and A generates G , A^4 contains two elements of some coset of V and A^3 also intersects this coset, so A^7 contains a nontrivial element of V , so we are done.

By minimality of V , x generates V as a $\mathbf{Z}(G/N)$ -module. Hence $A^7 \cap V$ also generates V . Since A covers G/N , taking the union of all A -conjugates of $A^7 \cap V$ produces a symmetric G/N -invariant generating set B of V contained in A^9 , and

$$\frac{|B^3|}{|B|} \leq \frac{|A^{27} \cap V|}{|A^7 \cap V|} \leq \frac{|A^{27} \cap V|}{|A^2 \cap V|} \leq K^{26}$$

by Lemma 4.3. Hence by Lemma 7.11, $B^{7d} = [V, G/N] = V$. Hence $A^{63d} \supseteq V$ and $G = A^3V = A^{63d+3}$. Again Lemma 4.9 implies $G = A^3$ provided $m > 3(63d + 3)$.

Next we prove (2). By Lemma 4.3(1), $A^2 \cap G$ is K^5 -tripling. Let $H = \langle A^2 \cap G \rangle$. By (1) it suffices to prove $H = G$.

We again argue by induction on $|N|$. If $N = 1$ then $H = G$ because A covers $G/N = G$, so assume $N \neq 1$. If V is a nontrivial A -invariant subgroup of N then the hypotheses hold for Γ/V , so by induction $HV = G$.

Suppose $H < G$. Let $V = Z(G) \cap N$. If $V \neq 1$ then $HV = G$, so $H \trianglelefteq G$ and G/H is abelian, in contradiction to $\deg_{\mathbf{C}}(G) > 1$. Hence $Z(G) \cap N = 1$.

Since N is nilpotent, $Z(N) \neq 1$. Let V be a minimal A -invariant subgroup of $Z(N)$ and $U = V \cap H$. Since $V \trianglelefteq \Gamma$, $U \trianglelefteq H$. Since $G = HV$, $U \trianglelefteq G$.

By the modular law, $N = N \cap HV = (N \cap H)V$. Hence $N = [G, N]$ implies

$$(N \cap H)V = [HV, (N \cap H)V] \leq [H, N \cap H]V.$$

Moreover, $[H, N \cap H] \leq [HV, (N \cap H)V]$ and $V \leq (N \cap H)V$, hence

$$(N \cap H)V = [H, N \cap H]V.$$

Intersecting with H and applying the modular law again,

$$N \cap H = [H, N \cap H]U. \tag{7.1}$$

We may identify V with a simple $\mathbf{Z}(\Gamma/N)$ -module. Since $G \trianglelefteq \Gamma$, V is a semisimple $\mathbf{Z}(G/N)$ -module by Clifford's theorem. Since $U \trianglelefteq G$, U is a $\mathbf{Z}(G/N)$ -submodule and hence also semisimple. If S is any simple submodule of U then $[G/N, S]$ is a submodule of S , and it cannot be trivial since $Z(G) \cap N = 1$, so $S = [G/N, S]$. Hence also $U = [G/N, U]$. Since H covers G/N , $U = [H, U] \leq [H, N \cap H]$.

Hence, from (7.1),

$$N \cap H = [H, N \cap H] \leq [H, H].$$

Since $H/(N \cap H) \cong G/N$ and G is perfect (since $\deg_{\mathbf{C}}(G) > 1$), H is perfect. By Lemma 4.10 it follows that $\deg_{\mathbf{C}}(H) \geq K^{cm}$ (recall that $K \geq 2$). Hence by (1) applied to $A^2 \cap G$ and H it follows that $H = (A^2 \cap G)^3 \subseteq A^6$.

Since $\deg_{\mathbf{C}}(H) \geq K^{cm}$, every proper subgroup of H has index at least K^{cm} . If $a \in A$, $H^a \subseteq A^8$. On the other hand $A^2 \cap H^a = A^2 \cap G \cap H^a \subseteq H \cap H^a$. Hence

$$[H : H \cap H^a] = \frac{|H^a|}{|H \cap H^a|} \leq \frac{|A^8 \cap H^a|}{|A^2 \cap H^a|} \leq K^7$$

by Lemma 4.3. It follows that $H = H^a$. Hence $H \trianglelefteq G$. Since $G = HZ(N)$, G/H is abelian and perfect, hence trivial. \square

We are now ready to prove Theorem 1.4.

Proof of Theorem 1.4. By Proposition 5.3 we may assume \mathbf{F} is a finite field. By Theorem 7.6, there is $\Gamma \trianglelefteq \langle A \rangle$ such that A is covered by $K^{O_n(1)}$ cosets of Γ , $\Gamma/\text{Sol}(\Gamma)$ is covered by A^6 , and $\Gamma/\text{Sol}(\Gamma) \in \text{Lie}^n(p)$, and $\deg_{\mathbf{C}}(\Gamma/\text{Sol}(\Gamma)) > K^d$, where $d = d(n)$ is sufficiently large for the rest of the argument.

By Lemma 6.2, Γ has derived length $O(\log n)$. Let $P = \Gamma^{(O(\log n))}$ be the perfect core. By iterating Lemma 7.9, $A^{O_n(1)}$ covers $P/\text{Sol}(P)$. Note that $P/\text{Sol}(P) \cong \Gamma/\text{Sol}(\Gamma)$ since $\Gamma/\text{Sol}(\Gamma)$ is perfect.

By Lemma 4.10, $\deg_{\mathbf{C}}(P) \geq cK^{d/2}$. By Corollary 7.4, there is a normal p -subgroup $N \trianglelefteq P$ contained in $\text{Sol}(P)$ such that $[P, N] = N$ and $|\text{Sol}(P)/N| \leq (2n+1)^n$. Since $A^{O_n(1)}$ covers $P/\text{Sol}(P)$, it projects to a subset of P/N of size at least $|P/\text{Sol}(P)| \geq (2n+1)^{-n}|P/N|$. Assuming $\deg_{\mathbf{C}}(P)$ is larger than $(2n+1)^{3n}$, this implies that $A^{O_n(1)}$ covers P/N by Proposition 4.8. By the fact that every finite simple group is 2-generated, $P/\text{Sol}(P)$ is $2n$ -generated, so P/N is $O_n(1)$ -generated. Applying Proposition 7.12 to the section P/N , it follows that $A^{O_n(1)}$ covers P . Finally, Lemma 4.9 implies that A^3 contains a coset of P , and the proof is complete. \square

8. MAIN PROOF PART 2: SOLUBLE TO NILPOTENT

8.1. Pivoting. The workhorse of the rest of the proof is a pivoting argument due to Gill and Helfgott related to sum-product theory. Let T be an abelian group acting on a nontrivial group U by automorphisms. We use multiplicative notation in T and U and exponential notation $(t, u) \mapsto u^t$ for the action of T on U . For $W \subseteq U$ and $X \subseteq T$ we write W^X for $\{w^x : w \in W, x \in X\}$. Let $F = F(T, U) \subseteq T$ be the set of $t \in T$ having a fixed point in $U \setminus \{1\}$.

Proposition 8.1 ([GH, Proposition 2.11]). *Let $X \subseteq T$ and $W \subseteq U$. Then either*

$$|(W^{X^{\pm 2}})^{\pm 6}| \geq \frac{1}{2} \frac{|W||X|}{|X^{-1}X \cap F|} \quad (8.1)$$

or

$$(W^X)^{\pm 8} = \langle W^{\langle X \rangle} \rangle. \quad (8.2)$$

Remark 8.2. The factor of $1/2$ in (8.1) does not appear in [GH, Proposition 2.11], but [GH, (2.5)] appears to be unjustified (the argument given shows only $|Y|^2|W|^2 \geq |\langle\langle X \rangle\rangle(\langle W \rangle)|$). Omitting this inequality, the rest of the proof is only impacted by a factor of 2, arguing as in [H1, Proposition 3.1].

The statement [GH, Proposition 2.11] also assumes that T acts faithfully on U , but the proof does not use this hypothesis.

The idea of the rest of this section is to apply Proposition 8.1 in appropriate trigonalizable sections of $\mathrm{GL}_n(\mathbf{F})$, or equivalently quotients of trigonalizable subgroups. The material is somewhat technical and the reader is encouraged to keep in mind the case of trigonalizable subgroups as a representative case. However, to prove our main theorem, the more general case of trigonalizable sections seems to be necessary.

Let $\Sigma = \Gamma/N$ be a trigonalizable section of $\mathrm{GL}_n(\mathbf{F})$, where \mathbf{F} is a finite field of characteristic p . By definition this means that $\Gamma = BN$ for some trigonalizable subgroup B . Lemma 6.3 implies that B and hence Σ is p -by-abelian. By Schur–Zassenhaus we therefore have a semidirect decomposition $\Sigma = UT$ where $U = O_p(\Sigma)$ and T is some abelian p' -group. Note that Σ acts naturally on U by conjugation. Let V be a Σ -composition factor of U and let $K = C_\Sigma(V)$. Then U acts trivially on V (as in the proof of Lemma 6.3), so $U \leq K$ and $\Sigma/K \cong T/T \cap K$ is an abelian p' -group. We call K a *root kernel* and the corresponding homomorphism $\chi : \Sigma \rightarrow \Sigma/K$ a *root*. The set of nontrivial roots of Σ is denoted $\Phi^*(\Sigma)$.

Lemma 8.3.

- (1) If K is a root kernel then Σ/K is isomorphic to a subgroup of \mathbf{F}^\times . Thus we may assume roots take values in \mathbf{F}^\times .
- (2) If V is a Σ -composition factor of U and $\chi : \Sigma \rightarrow \mathbf{F}^\times$ is the corresponding root then $V \cong \mathbf{F}_p(\chi(T))$, the subfield of \mathbf{F} generated by the image of χ with the action $v^g = \chi(g)v$.
- (3) $|\Phi^*(\Sigma)| < n^2$.

Proof. By hypothesis $\Gamma = BN$ for some trigonalizable subgroup $B \leq \mathrm{GL}_n(\mathbf{F})$. By replacing \mathbf{F} with an extension and Σ with a conjugate we may assume $B \leq B_n(\mathbf{F})$. Then B acts on $U_n(\mathbf{F})$, $O_p(B) \leq U_n(\mathbf{F})$, and $U \cong O_p(B)/O_p(B) \cap N$. By Jordan–Hölder, the Σ -composition factors (equivalently, B -composition factors) of U appear among the B -composition factors of $U_n(\mathbf{F})$.

Consider the following B -invariant series for $U_n(\mathbf{F})$. Let U_d be the subgroup of all $g \in U_n(\mathbf{F})$ such that $g_{ij} = 0$ for $0 < j - i < d$. Then U_d/U_{d+1} is a direct sum of copies V_{ij} of \mathbf{F} , for $j - i = d$, where B acts on V_{ij} according

to $v^g = \chi_{ij}(g)v$, where $\chi_{ij}(g) = g_{ii}/g_{jj}$. Note that V_{ij} is a direct sum of isomorphic copies of the irreducible B -module $\mathbf{F}_p(\chi_{ij}(B))$, and $C_B(v) = \ker \chi_{ij}$ for every nonzero $v \in V_{ij}$.

Thus if V is a Σ -composition factor of U then $V \cong \mathbf{F}_p(\chi_{ij}(B))$ for some i, j . If K is the corresponding root kernel then $K = \ker \chi_{ij}N/N$ (and $B \cap N \leq \ker \chi_{ij}$, since $B \cap N$ must act trivially). Since there are fewer than n^2 possibility for (i, j) , this proves the three claims. \square

Lemma 8.4. *Let V be a T -invariant section of $U = O_p(\Sigma)$ such that $C_V(T) = 1$. Then*

$$F(T, V) \subseteq \bigcup_{\chi \in \Phi^*(\Sigma)} \ker \chi.$$

Proof. Suppose $t \in F(T, V)$, so there is some nontrivial $v \in V$ such that $v^t = v$. By replacing V with $\langle v^T \rangle$ we may assume $V = \langle v^T \rangle$. Since V is nontrivial, $V \neq \text{Frat}(V)$, so $v^T \not\subseteq \text{Frat}(V)$, which implies that $v \notin \text{Frat}(V)$. By replacing V with $V/\text{Frat}(V)$ we may therefore assume $\text{Frat}(V)$ is trivial (by Lemma 4.12, $V = [V, T]$, so the quotient has the same property, so the condition $C_V(T) = 1$ is preserved). Hence V is elementary abelian and we may identify it with an $\mathbf{F}_p T$ -module. By Maschke's theorem, V is completely reducible. By projecting to one of the irreducible components we may assume V is irreducible. Note then $[V, U] = 1$, so V is a Σ -composition factor of U . Now $C_V(t)$ is a nontrivial submodule of V , so $C_V(t) = V$, so t is contained in the root kernel $K = C_\Sigma(V)$, and $K \neq \Sigma$ because $C_V(T) = 1$. Hence t is contained in a nontrivial root kernel. \square

Proposition 8.5. *Let V be a Σ -invariant section of $U = O_p(\Sigma)$ such that $[V, U] = C_V(T) = 1$. Let $A \subseteq \Sigma$ be a symmetric K -tripling subset such that $\langle A^\pi \rangle = T$, where $\pi : \Sigma \rightarrow T$ is the natural projection. Then either*

- (a) *there is some $\chi \in \Phi^*(\Sigma)$ such that $1 < |\chi(A)| \leq 2n^2 K^{O(1)}$, or*
- (b) *$A^{O(1)}$ covers $\langle (A^2 \cap V)^T \rangle$.*

Proof. Assume that $|\chi(A)| \geq R$ for every $\chi \in \Phi^*(\Sigma)$. Then by Lemma 4.2 and Lemma 4.3(2),

$$|(A^\pi)^2 \cap \ker \chi| \leq R^{-1} |(A^\pi)^3| \leq R^{-1} K^3 |A^\pi|.$$

Hence by the previous two lemmas

$$|(A^\pi)^2 \cap F(T, V)| \leq n^2 K^3 R^{-1} |A^\pi|.$$

Since $[V, U] = 1$, the action of Σ on V factors through $\pi : \Sigma \rightarrow T$. Apply Proposition 8.1 with $W = A^2 \cap V$ and $X = A^\pi$. If (8.1) holds then

$$|((A^2 \cap V)^{A^2})^6| \geq \frac{1}{2} \frac{|A^2 \cap V| |A^\pi|}{|(A^\pi)^2 \cap F(T, V)|} \geq \frac{1}{2} n^{-2} K^{-3} R |A^2 \cap V|,$$

while

$$|((A^2 \cap V)^{A^2})^6| \leq |A^{36} \cap V| \leq K^{35} |A^2 \cap V|$$

by Lemma 4.3(1). This implies (a). On the other hand if (8.2) holds then

$$\langle (A^2 \cap V)^T \rangle = ((A^2 \cap V)^A)^8 \subseteq A^{32},$$

which implies (b). \square

8.2. Growth of bilinear images. In this section we consider abelian groups only, so we use additive notation.

We briefly recall the definition of the Fourier transform on a finite abelian group G . The dual group \widehat{G} is the group of all homomorphisms $\chi : G \rightarrow S^1$. We endow G with the counting measure and \widehat{G} with the uniform measure. The Fourier transform of a function $f : G \rightarrow \mathbf{C}$ is then defined by

$$\widehat{f}(\chi) = \sum_{x \in G} f(x) \chi(-x) \quad (\chi \in \widehat{G}),$$

and the Fourier inversion formula is

$$f(x) = \frac{1}{|G|} \sum_{\chi \in \widehat{G}} \widehat{f}(\chi) \chi(x) \quad (x \in G).$$

Parseval's identity is

$$\sum_G |f|^2 = \frac{1}{|G|} \sum_{\widehat{G}} |\widehat{f}|^2.$$

The convolution of two functions $f_1, f_2 : G \rightarrow \mathbf{C}$ is defined by

$$f_1 * f_2(x) = \sum_{y \in G} f_1(y) f_2(x - y).$$

With this definition we have the rule

$$\widehat{f_1 * f_2} = \widehat{f_1} \widehat{f_2}.$$

In the following lemma, by a *probability measure* on G we simply mean a function $\mu : G \rightarrow [0, 1]$ such that $\sum_{x \in G} \mu(x) = 1$.

Lemma 8.6. *Let G be a finite abelian group, and let μ be a probability measure on G such that, for all $\chi \in \widehat{G}$, either $|\widehat{\mu}(\chi)| \leq 1/R$ or $\widehat{\mu}(\chi) = 1$. Let S be the support of μ , and let $A \subseteq G$. Then*

$$|A + S| > \frac{1}{2} \min(R^2 |A|, |\langle S \rangle|).$$

Moreover, if $S \subseteq A \subseteq \langle S \rangle$ and $|2A| \leq \frac{1}{2} R^2 |A|$, then $4A = \langle S \rangle$.

Proof. The convolution $1_A * \mu$ is supported on $A + S$, so, by Cauchy-Schwarz and Parseval's identity,

$$|A|^2 = \left(\sum_G 1_A * \mu \right)^2 \leq |A + S| \sum_G |1_A * \mu|^2 = |A + S| \frac{1}{|G|} \sum_{\widehat{G}} |\widehat{1_A}|^2 |\widehat{\mu}|^2.$$

Let $H = \langle S \rangle$, and note $\widehat{\mu}(\chi) = 1$ if and only if $\chi \in H^\perp = \{\psi \in \widehat{G} : \psi(H) = 1\}$. Hence, using Parseval's identity again,

$$\begin{aligned} \frac{1}{|G|} \sum_{\widehat{G}} |\widehat{1}_A|^2 |\widehat{\mu}|^2 &\leq \frac{|H^\perp|}{|G|} |A|^2 + R^{-2} \frac{1}{|G|} \sum_{\widehat{G} \setminus H^\perp} |\widehat{1}_A|^2 \\ &< \frac{|A|^2}{|H|} + R^{-2} |A|. \end{aligned}$$

Rearranging,

$$|A + S| > \frac{|A|}{|A|/|H| + R^{-2}} \geq \frac{1}{2} \min(|H|, R^2 |A|).$$

For the last statement, if $|A + S| \leq \frac{1}{2} R^2 |A|$ then $|A + S| > \frac{1}{2} |H|$, so $4A \supseteq 2(A + S) = H$ by a standard exercise. \square

Proposition 8.7. *Let U, V, Z be finite abelian groups, let T be a group acting on U and V , and let $\beta : U \times V \rightarrow Z$ be a bilinear map which is T -invariant in the sense that*

$$\beta(u^t, v^t) = \beta(u, v) \quad (u \in U, v \in V, t \in T).$$

Suppose that the T -simple composition factors of U have size at least R . Write $\beta(U, V)$ for the image of $U \times V$ and let $W = \langle \beta(U, V) \rangle$. Then for any $A \subseteq Z$ we have

$$|A + \beta(U, V)| > \frac{1}{2} \min(R^2 |A|, |W|).$$

Moreover, if $\beta(U, V) \subseteq A \subseteq W$ and $|2A| \leq \frac{1}{2} R^2 |A|$ then $4A = W$.

Proof. Let μ be the pushforward of the uniform measure on $U \times V$, so that

$$\mu(z) = \frac{\#\{(u, v) \in U \times V : \beta(u, v) = z\}}{|U||V|}.$$

Then for $\chi \in \widehat{Z}$,

$$\widehat{\mu}(\chi) = \sum_{z \in Z} \mu(z) \chi(z) = \frac{1}{|U||V|} \sum_{u \in U, v \in V} \chi(\beta(u, v)) = [U : U_\chi]^{-1},$$

where U_χ is the subgroup of all $u \in U$ such that $\chi(\beta(u, V)) = 1$. Since U_χ is T -invariant, either $U_\chi = U$ or $[U : U_\chi] \geq R$, so the lemma applies. \square

8.3. The no-small-roots case. Suppose $\Sigma = \Gamma/N$ is a trigonalizable section of $\mathrm{GL}_n(\mathbf{F})$. Recall from Section 8.1 that $\Phi^*(\Sigma)$ is the set of nontrivial roots of Σ . Call $\chi \in \Phi^*(\Sigma)$ an R -small root for $A \subseteq \Sigma$ if $1 < |\chi(A)| \leq R$. In this section we establish the trigonal no-small-roots case of Theorem 1.2.

Proposition 8.8. *There is a constant C_n such that the following holds. Let \mathbf{F} be a finite field. Let $\Sigma = \Gamma/N$ be a trigonalizable section of $\mathrm{GL}_n(\mathbf{F})$. Let $A \subseteq \Sigma$ be a nonempty symmetric set containing 1 such that*

- (i) $|A^3| \leq K|A|$,
- (ii) A has no K^{C_n} -small roots,

(iii) A generates Σ .

Then $\gamma_n(\Sigma) \subseteq A^{C_n}$.

By Lemma 6.3, Σ is p -by-abelian and we have a decomposition $\Sigma = UT$, where T is an abelian p' -group and $U = O_p(\Sigma)$, and moreover $\gamma_n(U) = 1$. Let $\pi : \Sigma \rightarrow T$ be the natural projection. Let $H = [U, T]$. By Lemmas 4.12 and 4.14, $U = HC_U(T)$, $H = [H, T]$, and $H = \gamma_\omega(\Sigma) = \gamma_n(\Sigma)$. We will prove that $H \subseteq A^{O_n(1)}$.

Several times in the proof we will replace A by a small power. This operation is justified by Lemma 4.1.

8.3.1. *Central case.* First we will prove that if Z is a normal subgroup of Σ contained in H such that $[Z, U] = C_Z(T) = 1$ and A covers H/Z then $A^{O_n(1)}$ covers Z .

We use the idea of “descent” from [GH]. Let $\Sigma_1 = ZC_U(T)T$. Let $A_1 = A^3 \cap \Sigma_1$. Then we claim

- (1) A_1 generates Σ_1 ,
- (2) A_1 has no K^{C_n} -small roots.

Since A covers H/Z we have $\Sigma = A\Sigma_1$, so Lemma 4.6 implies (1). Let $A_H = A \cap H$. Then A_H also covers H/Z , so $\Sigma = A_H\Sigma_1$ for the same reason. In particular $A \subseteq A_H\Sigma_1$, so $A \subseteq A_H(A_H A \cap \Sigma_1)$. Applying π , $A^\pi \subseteq (A_H A \cap \Sigma_1)^\pi \subseteq A_1^\pi$. Hence (2) holds.

By Lemmas 4.12 and 4.14, $Z = [Z, T] \times C_Z(T) = [Z, T]$ and $Z = \gamma_n(\Sigma_1)$. The n -fold commutators $[a_1, \dots, a_n]$ with $a_1, \dots, a_n \in A_1$ all lie in $C = A_1^{2^{n+1}} \cap Z$, and they normally generate $Z = \gamma_n(\Sigma_1)$ since A_1 generates Σ_1 . Since Z is central in U , $C^{\Sigma_1} = C^T$, so $\langle C^T \rangle = Z$. Now we can apply Proposition 8.5 and we find that $Z \subseteq A_1^{O_n(1)} \subseteq A^{O_n(1)}$, as claimed.

8.3.2. *Abelian case.* Now suppose V is an abelian normal subgroup of Σ contained in H such that $C_V(T) = 1$, and assume that A covers H/V . We claim that $A^{O_n(1)}$ covers H . Suppose $Z \leq V$ and $Z \trianglelefteq \Sigma$ and A^m covers H/Z . Then $C_Z(T) \leq C_V(T) = 1$, so by the central case $A^{O_n(m)}$ covers $Z/[Z, U]$, so $A^{O_n(m)}$ covers $H/[Z, U]$. Iterating this $n - 1$ times starting with $Z = V$ and using

$$[V, \underbrace{U, \dots, U}_{n-1}] \leq \gamma_n(U) = 1,$$

it follows that $A^{O_n(1)}$ covers H .

8.3.3. *General case.* Finally consider H itself. Let $k \geq 1$ be minimal such that $\gamma_{k+1}(H) = 1$. Then $V = \gamma_k(H)$ is a characteristic central subgroup of H . By induction on k we may assume A^m covers H/V for some $m \leq O_n(1)$. Replacing A with A^m , we may assume A covers H/V . We claim that $V \subseteq A^{O_n(1)}$.

Let $W = [V, T]$. Since W is centralized by H and normalized by $C_U(T)$ and T , we have $W \trianglelefteq \Sigma$. By Lemma 4.12, $[W, T] = W$ and $C_W(T) = 1$. Hence $A^{O_n(1)}$ covers W by the abelian case.

Thus we may assume $[V, T] = 1$. Since $H = [H, T]$, either H is trivial or $k \geq 2$. If $k \geq 2$ the commutator induces a well-defined map

$$\beta : H/\gamma_2(H) \times \gamma_{k-1}(H)/\gamma_k(H) \rightarrow V$$

which is bilinear (see [A, (8.5.4)]) and whose image generates V . Moreover, β is T -invariant in the sense of Proposition 8.7, because $[x^t, y^t] = [x, y]^t = [x, y]$ for $x \in H$ and $y \in \gamma_{k-1}(H)$. Since $H \subseteq A\gamma_k(H)$, the image of β is contained in the set $B = A^4 \cap V$, so B generates V . On the other hand, B has tripling at most K^{11} by Lemma 4.3(1). By Lemma 4.13, $C_{H/\gamma_2(H)}(T) = 1$. By Lemma 8.3, any T -simple composition factor of $H/\gamma_2(H)$ is isomorphic to a T -module of the form $\mathbf{F}_p(\chi(T))$ for some root $\chi : T \rightarrow \mathbf{F}$. Since A has no K^{C_n} -small-roots, $H/\gamma_2(H)$ has no T -simple composition factor of size less than K^{C_n} . Hence, by Proposition 8.7, $B^4 = V$, so $V \subseteq A^{16}$, as claimed.

This completes the induction on k , and we have proved that $H \subseteq A^{O_n(1)}$. The proof of Proposition 8.8 is then complete provided that C_n is at least as large as this implicit constant.

8.4. Groups with a good section. The hypotheses (i)–(iv) of the following proposition are guaranteed by Theorem 1.4 and Theorem 6.4. The hypothesis (v) on the other hand is mildly restrictive, and will be removed in the next (and final) section.

Proposition 8.9. *Let \mathbf{F} be a finite field. Let $A \subseteq \mathrm{GL}_n(\mathbf{F})$ be a finite, nonempty, symmetric, K -tripling set, and let $G = \langle A \rangle$. Assume there is a section $\Sigma = \Gamma/P$ where $P \trianglelefteq \Gamma \trianglelefteq G$ such that*

- (i) $|A\Gamma/\Gamma| \leq K^{O_n(1)}$,
- (ii) Γ/P is soluble,
- (iii) $P \subseteq A^6$,
- (iv) $\Sigma = \mathrm{tr}(B, \Sigma)$ for some trigonalizable subgroup $B \leq \Gamma$,
- (v) G acts trivially on $\Sigma/O_p(\Sigma)$.

Then there is a normal subgroup $\Delta \trianglelefteq G$ such that $P \leq \Delta \leq \Gamma$ and

- (1) $|A\Delta/\Delta| \leq K^{O_n(1)}$, and
- (2) $\gamma_n(\Delta) \subseteq A^{O_n(1)}$.

Proof. First we shrink Σ (and Γ) until there are no small roots, as follows. Let $\chi_1, \dots, \chi_m \in \Phi^*(\Sigma)$ be the R -small roots for $\mathrm{tr}(A^2, \Sigma)$, for some value of R . Note that $m \leq n^2$ by Lemma 8.3. Let $\Sigma_m = \bigcap_{i=1}^m \ker \chi_i \trianglelefteq \Sigma$. Then

$$|\mathrm{tr}(A^2, \Sigma/\Sigma_m)| \leq \prod_{i=1}^m |\chi_i(\mathrm{tr}(A^2, \Sigma))| \leq R^m.$$

Now if $\chi \in \Phi^*(\Sigma)$ is an R -small root for $\mathrm{tr}(A^2, \Sigma_m)$ then, since $\mathrm{tr}(A^2, \Sigma)$ is covered by R^m cosets of Σ_m , and therefore by R^m translates of $\mathrm{tr}(A^2, \Sigma)^2 \cap$

$\Sigma_m \subseteq \text{tr}(A^4, \Sigma_m)$ (see Lemma 4.4), it follows from Lemma 4.3(3) that

$$|\chi(\text{tr}(A^2, \Sigma))| \leq R^m |\chi(\text{tr}(A^4, \Sigma_m))| \leq K^7 R^m |\chi(\text{tr}(A^2, \Sigma_m))| \leq K^7 R^{m+1},$$

so χ is a $K^7 R^{m+1}$ -small root for $\text{tr}(A^2, \Sigma)$. By the pigeonhole principle we can choose R so that $K^{C'_n} \leq R \leq K^{O_n(1)}$, where $C'_n = 15C_n + 3$ and C_n is the constant in Proposition 8.8, and such that there is no $\chi \in \Phi^*(\Sigma)$ such that

$$R < |\chi(\text{tr}(A^2, \Sigma))| \leq K^7 R^{m+1}.$$

For this value of R , it follows that $\text{tr}(A^2, \Sigma_m)$ has no R -small roots $\chi \in \Phi^*(\Sigma)$. Since $O_p(\Sigma_m) = O_p(\Sigma)$, we have $\Phi^*(\Sigma_m) \subseteq \Phi^*(\Sigma)$, so a fortiori $\text{tr}(A^2, \Sigma_m)$ has no R -small roots in $\Phi^*(\Sigma_m)$. Write $\Sigma_m = \Gamma_m/P$. Then $\Gamma_m \trianglelefteq \Gamma$, and in fact $\Gamma_m \trianglelefteq G$ since G acts trivially on $\Sigma/O_p(\Sigma)$. From Lemma 4.2(2),

$$|A\Gamma_m/\Gamma_m| \leq |A\Gamma/\Gamma| |\text{tr}(A^2, \Sigma/\Sigma_m)| \leq K^{O_n(1)}.$$

Hence we may replace Γ with Γ_m , and thus we may assume $\text{tr}(A^2, \Sigma)$ has no $K^{C'_n}$ -small roots.

Next we replace Γ with the preimage of $\langle \text{tr}(A^2, \Sigma/O_p(\Sigma)) \rangle$, ensuring that $\text{tr}(A^2, \Sigma/O_p(\Sigma))$ generates $\Sigma/O_p(\Sigma)$. This does not change the value of $|\chi(A)|$ for any root χ , nor does it change the value of $|A\Gamma/\Gamma|$, by Lemma 4.4. Again this does not compromise normality of Γ in G because G acts trivially on $\Sigma/O_p(\Sigma)$.

Let $U = O_p(\Sigma)$. For $i = 1, 2$ define

$$\begin{aligned} \Sigma_i &= \langle \text{tr}(A^{2^i}, \Sigma) \rangle, \\ U_i &= \Sigma_i \cap U, \\ H_i &= \gamma_n(\Sigma_i). \end{aligned}$$

By Lemma 6.3, Σ is p -by-abelian and $\gamma_n(U) = 1$. By Schur–Zassenhaus, $\Sigma_1 = U_1 T$ for an abelian p' -group T . Since $\text{tr}(A^2, \Sigma/U)$ generates Σ/U we also have $\Sigma = UT$ as well as $\Sigma_2 = \Sigma_2 \cap UT = U_2 T$ by the modular law. By Lemmas 4.12 and 4.14, $H_i = [U_i, T] = [H_i, T]$. Also, note that

$$\Sigma_1^a \leq \Sigma_2 \quad (a \in A).$$

Let $V = H_2/H'_2$. Since $[H_2, T] = H_2$ it follows that $[V, T] = V$ and, by Lemma 4.12, $C_V(T) = 1$. Similarly for any subgroup $L \leq V$ we have $C_L(T) \leq C_V(T) = 1$ and, by Lemma 4.12, $L = [L, T]$. Applying this to $L = \langle \text{tr}(A^2, V) \rangle$, we have

$$\text{tr}(A^2, V) \subseteq L = [L, T] \leq \text{tr}([U_1, T], V) = \text{tr}(H_1, V) = H_1 H'_2/H'_2.$$

By Proposition 8.8 applied to $\text{tr}(A^4, \Sigma)$, which has tripling at most K^{15} by Lemma 4.3 and no K^{15C_n} -small roots, we have $H_2 \subseteq \text{tr}(A^{15C_n}, H_2)$. Hence, by Lemma 4.3,

$$[H_2 : H_1 H'_2] \leq \frac{|\text{tr}(A^{15C_n}, H_2)|}{|\text{tr}(A^2, H_2)|} \leq K^{15C_n+3}.$$

By Lemma 8.3, any T -simple quotient of $H_2/(H_1H'_2)$ is isomorphic to a T -module of the form $\mathbf{F}_p(\chi(T))$ for some root $\chi : T \rightarrow \mathbf{F}$. Since $H_2 = [H_2, T]$, we have $W = [W, T]$, so χ must be nontrivial. On the other hand the above bound shows that $|W| \leq K^{15C_n+3}$. Since T has no K^{15C_n+3} -small roots, there is no such W . Thus $H_2 = H_1H'_2$. By the Burnside basis theorem it follows that $H_2 = H_1$.

Let $H = H_1 = H_2$. Hence $\Sigma_i = HC_{U_i}(T)T$ by Lemma 4.12. Since $H^a = \gamma_n(\Sigma_1^a) \leq \gamma_n(\Sigma_2) = H$ for $a \in A$, it follows that H is normalized by G . At last define

$$\Omega = \Sigma_1 C_U(T) = \Sigma_2 C_U(T) = HC_U(T)T.$$

For $a \in A$ we have $T^a \leq \Sigma_1^a \leq \Sigma_2 \leq \Omega$. Since $\Omega = TC_U(T)H$, it follows that $T^a = T^h$ for some $h \in H$ by Schur–Zassenhaus. Hence also

$$C_U(T)^a = C_U(T)^h \leq C_U(T)H \leq \Omega.$$

Thus Ω is normalized by G . By Lemma 4.14 (and recalling $\gamma_n(U) = 1$), we have $\gamma_n(\Omega) = [HC_U(T), T] = [H, T] = H$, and we saw earlier that $H = H_2$ is covered by A^{15C_n} .

Let Δ be the preimage of Ω in Γ . Recall that A is covered by $|A\Gamma/\Gamma|$ translates of $A^2 \cap \Gamma$ by Lemma 4.4. Therefore since $\text{tr}(A^2, \Sigma) \subseteq \Omega$, we have $|A\Delta/\Delta| = |A\Gamma/\Gamma|$. Now since $\gamma_n(\Omega)$ is covered by A^{15C_n} and $P \subseteq A^6$ it follows that $\gamma_n(\Delta) \subseteq A^{15C_n+6}$. This completes the proof. \square

8.5. Creating a good section. Finally, let A be as in Theorem 1.2: A is a nonempty, symmetric, K -tripling subset of $\text{GL}_n(\mathbf{F})$. By Proposition 5.2 we may assume \mathbf{F} is finite. Let $G = \langle A \rangle$. By Theorem 1.4, there is a soluble section $\Sigma = \Gamma/P$ where $P \triangleleft \Gamma \triangleleft G$ such that (i)–(iii) of Proposition 8.9 are satisfied. Moreover P is perfect, soluble-by-Lie $^*(p)$, and contained in a translate of A^3 . By Theorem 6.4 (and replacing Σ with Σ_0), we can assume (iv) holds too, and moreover we can assume that $G_0 = C_G(\Sigma/O_p(\Sigma))$ has index at most $n!$ in G .

Let $m = [G : G_0]$, so $m \leq n!$. Since A generates G , we have $G = A^m G_0$, so G_0 is generated by $A_0 = A^{3m} \cap G_0$ by Lemma 4.6. By Lemma 4.3(1), A_0 has tripling $K^{O(m)}$. By Lemma 4.3(2), $|A_0\Gamma/\Gamma| \leq K^{O(m)}|A\Gamma/\Gamma| \leq K^{O_n(1)}$. Since P acts trivially on Σ we have $P \subseteq A^6 \cap G_0 \subseteq A_0^6$. By Lemma 6.3, $\Sigma/O_p(\Sigma)$ is abelian, so $\Gamma \leq G_0$. Hence the hypotheses of Proposition 8.9 hold with $(A_0, G_0, K^{O(m)})$ in the role of (A, G, K) . Thus there is $\Delta \trianglelefteq G_0$ such that $P \leq \Delta \leq \Gamma$ and $|A_0\Delta/\Delta| \leq K^{O_n(1)}$ and $\gamma_n(\Delta) \subseteq A_0^{O_n(1)}$. Note also that $\gamma_n(\Delta)/P$ is a p -group since $\gamma_n(\Gamma)/P \cong \gamma_n(\Sigma)$ and $\Sigma/O_p(\Sigma)$ is abelian, as noted.

Since A is covered by m cosets of G_0 , it is covered by m translates of $A^2 \cap G_0 \subseteq A_0$ (Lemma 4.4), so $|A\Delta/\Delta| \leq m|A_0\Delta/\Delta| \leq K^{O_n(1)}$.

The only remaining issue is that while Δ is normal in G_0 it may not be normal in G . But since $[G : G_0] = m$, there are at most m conjugates of Δ in G , and since $G = A^m G_0$ each of them has the form Δ^a for some $a \in A^m$.

Let $\Delta_0 \trianglelefteq G$ be their intersection. If $a \in A^m$ then

$$|A\Delta^a/\Delta^a| = |A^{a^{-1}}\Delta/\Delta| \leq |A^{2m+1}\Delta/\Delta| \leq K^{2m+1}|A\Delta/\Delta|$$

by Lemma 4.3(2). It follows that $|A\Delta_0/\Delta_0| \leq (K^{2m+1}|A\Delta/\Delta|)^m \leq K^{O_n(1)}$, and obviously $\gamma_n(\Delta_0) \subseteq \gamma_n(\Delta) \subseteq A^{O_n(1)}$. Moreover $P \leq \Delta_0$ and $\gamma_n(\Delta_0)/P$ is a p -group. This finally completes the proof of Theorem 1.2.

9. DIAMETER OF QUASIRANDOM GROUPS

As promised in the introduction, here we show that one of our key new ideas in the proof of Theorem 1.2, namely the “affine conjugating trick” (Lemma 1.8) has another application to growth-type questions.

Theorem 9.1 (Theorem 1.6 restated). *For each positive integer n there are positive numbers $K = K(n)$ and $c = c(n)$ with the following property. Let \mathbf{F} be a field and $G \leq \mathrm{GL}_n(\mathbf{F})$ be a K -quasirandom finite subgroup. Then the Cayley graph of G with respect to any generating set has diameter at most $(\log |G|)^c$.*

This theorem is sharp in the following sense. In [PS3, Example 75] a perfect subgroup G of $\mathrm{SL}_5(\mathbf{F}_q)$ is constructed which has diameter $|G|^c$ for a constant $c > 0$. Slightly modifying this example, one can obtain subgroups $G \leq \mathrm{SL}_n(\mathbf{F}_q)$ with $\deg_{\mathbf{C}}(G) \geq n - 1$ and diameter $|G|^{c_n}$.

The proof of Theorem 1.6 is based on Lemma 1.8 and on a boundedness property of finite linear groups (Theorem 9.6) which is of independent interest. The proof of Theorem 9.6 in turn uses a result of McNinch concerning connected algebraic groups acting on connected unipotent groups, and on an important result of Steinberg on representations of finite simple groups of Lie type.

Definition 9.2. Let G be a group acting algebraically on a connected unipotent group U over an algebraically closed field. The action is *linearizable* if there is a G -invariant normal chain of connected closed subgroups $1 = U_0 \trianglelefteq U_1 \trianglelefteq \dots \trianglelefteq U_k = U$ such that each quotient U_j/U_{j-1} is G -equivariantly isomorphic to a vector space with a linear G -action.

Lemma 9.3. *Let \mathbf{F} be an algebraically closed field of characteristic $p > 0$, let $U_0 \leq \mathrm{GL}_n(\mathbf{F})$ be a unipotent subgroup, and let $G \leq \mathrm{GL}_n(\mathbf{F})$ be a subgroup normalizing U_0 . Then there is a G -invariant connected closed unipotent subgroup $U \leq \mathrm{GL}_n(\mathbf{F})$ containing U_0 and G has a subgroup H of index $O_n(1)$ whose action on U is linearizable.*

Proof. Let $M_n(\mathbf{F})$ denote the linear space of n -by- n matrices over \mathbf{F} . Then $\mathrm{GL}_n(\mathbf{F})$ acts linearly on $M_n(\mathbf{F})$ via conjugation. Let $N \leq M_n(\mathbf{F})$ be the linear span of $\{u - 1 : u \in U_0\}$. Then N is a G -invariant nilpotent subalgebra of $M_n(\mathbf{F})$ and $U = 1 + N$ is a G -invariant connected closed unipotent subgroup.

Let S be the normalizer of U in $\mathrm{GL}_n(\mathbf{F})$. We claim that $[S : S^\circ] = O_n(1)$. Observe that $\mathrm{Ad}(S)$ is just the intersection of $\mathrm{Ad}(\mathrm{GL}_n(\mathbf{F}))$ with the stabilizer

of N in $\mathrm{GL}(M_n(\mathbf{F}))$, which is itself the intersection of $\mathrm{GL}(M_n(\mathbf{F}))$ with a linear subspace (of codimension $d(n^2 - d)$, where $d = \dim N$). Therefore by Bezout's theorem the number of components of S is bounded by the degree of $\mathrm{Ad}(\mathrm{GL}_n(\mathbf{F})) = \mathrm{Ad}(\mathrm{SL}_n(\mathbf{F}))$, which is at most n^{n^2} (see [V]).

Let $H = G \cap S^\circ$. Since $G \leq S$ we have $[G : H] \leq [S : S^\circ] = O_n(1)$.

The result of McNinch [M] implies that the S° -action on U is linearizable. Hence the H -action on U is linearizable. \square

Definition 9.4. Let G be a finite group acting on a p -group P . Assume that $G/O_p(G)$ is a central product of quasisimple groups S_1, \dots, S_l . We say that a G -invariant normal chain $1 = P_0 \trianglelefteq P_1 \trianglelefteq \dots \trianglelefteq P_k = P$ of length k is B -bounded for some $B > 0$ if

- (a) each quotient P_j/P_{j-1} is elementary abelian,
- (b) as an $\mathbf{F}_p G$ -module P_j/P_{j-1} is isomorphic to the direct sum of isomorphic copies of some irreducible $\mathbf{F}_p G$ -module W_j ,
- (c) W_j has a tensor product decomposition $W_j \cong X_{j,1} \otimes_{\mathbf{F}_p} \dots \otimes_{\mathbf{F}_p} X_{j,l}$ where each $X_{j,i}$ is an irreducible $\mathbf{F}_p S_i$ -module,
- (d) either $X_{j,i} \cong \mathbf{F}_p$ with trivial S_i -action, or

$$X_{j,i} = (S_i x)^{\pm B} \quad \text{for all nonzero } x \in X_{j,i}.$$

Lemma 9.5. Let G be a finite group acting on a p -group P . If P has a B -bounded G -invariant normal chain of length k then each G -invariant section of P has a B -bounded G -invariant normal chain of length k .

Proof. Clear. \square

Theorem 9.6. For all $n > 0$ there is an integer $K = K(n) > 0$ with the following property. Let \mathbf{F} be a field of characteristic $p > 0$, let $G \leq \mathrm{GL}_n(\mathbf{F})$ be a finite subgroup and let $P < \mathrm{GL}_n(\mathbf{F})$ be a p -group normalized by G . If $\deg_{\mathbf{C}}(G) \geq K$ then any G -invariant section of P has an $O_n(1)$ -bounded G -invariant normal chain of length less than n^2 .

Proof. If we choose K large enough then by Weisfeiler's theorem $G/O_p(G)$ is isomorphic to a central product of at most n quasisimple groups of Lie type of characteristic p with Lie rank at most n (see Theorem 7.2 and Lemma 7.1). Let S_1, \dots, S_l denote the factors, and let S be their direct product.

Let $\overline{\mathbf{F}}$ be the algebraic closure of \mathbf{F} . Lemma 9.3 gives us a subgroup $H \leq G$ of index $O_n(1)$ and a G -invariant connected unipotent subgroup $U < \mathrm{GL}_n(\overline{\mathbf{F}})$ containing P with linearizable H -action (it is well known that p -subgroups of $\mathrm{GL}_n(\mathbf{F})$ are unipotent). By choosing K large enough we make sure that $G = H$. We shall prove that U has an $O_n(1)$ -bounded G -invariant normal chain of length less than n^2 .

Since the action of G on U is linearizable, there is a G -invariant normal chain $1 = U_0 \trianglelefteq U_1 \trianglelefteq \dots \trianglelefteq U_k = U$ of connected closed subgroups such that each quotient $V_j = U_j/U_{j-1}$ is isomorphic to a vector space over $\overline{\mathbf{F}}$ with a linear G -action. We can refine this chain so that each V_j is an irreducible

$\overline{\mathbf{F}}G$ -module. The length of this chain is $k \leq \dim(U) < n^2$. We shall prove that the normal chain $U_0 \trianglelefteq \dots \trianglelefteq U_k$ in $O_n(1)$ -bounded.

Since V_j is irreducible, $O_p(G)$ acts on it trivially, so it is an irreducible $\overline{\mathbf{F}}S$ -module. Therefore it can be written as a tensor product

$$V_j \cong Y_{j,1} \otimes_{\overline{\mathbf{F}}} \dots \otimes_{\overline{\mathbf{F}}} Y_{j,l}$$

where each $Y_{j,i}$ is an irreducible $\overline{\mathbf{F}}S_i$ -module.

Let \mathbf{F}_{q_i} be the defining field of S_i . By a result of Steinberg [S2], there is an $\mathbf{F}_{q_i}S_i$ module $M_{j,i}$ such that

$$Y_{j,i} \cong M_{j,i} \otimes_{\mathbf{F}_{q_i}} \overline{\mathbf{F}},$$

and $\dim_{\mathbf{F}_{q_i}}(M_{j,i}) < n^2$. Note that in Steinberg's result not all simple groups are covered: odd-dimensional unitary groups and Ree groups of type 2G_2 are excluded. The reason for this exclusion is that one needs to lift projective representations to ordinary representations of the group he called Γ_q^1 constructed from the simply connected algebraic group. Since we now know the Schur multipliers of these groups, we can include them as well, with a small number of exceptions with bounded size. We exclude those exceptions from the S_i by making K large enough.

Each $M_{j,i}$ as an $\mathbf{F}_p S_i$ -module must be the direct sum of isomorphic copies of some irreducible $\mathbf{F}_p S_i$ -module $X_{j,i}$. Therefore V_j as an $\mathbf{F}_p S$ -module is the direct sum of isomorphic copies of $X_{j,1} \otimes_{\mathbf{F}_p} \dots \otimes_{\mathbf{F}_p} X_{j,l}$.

Finally suppose that S_i acts nontrivially on $X_{j,i}$ and $x \in X_{j,i}$. The bound $\dim_{\mathbf{F}_{q_i}}(M_{j,i}) < n^2$ implies that $|M_{j,i}| = |S_i|^{O_n(1)}$ and therefore also $|X_{j,i}| = |S_i|^{O_n(1)}$. Let $A = (S_i x)^{\pm 1} \subseteq X_{j,i}$. Then A is a symmetric S_i -invariant generating set for $X_{j,i}$. Now we apply the affine conjugating trick to $A \subseteq X_{j,i} \rtimes S_i$ with $K = |S_i|^{c/n}$ (see Theorem 4.7). Let $k \geq 0$ be minimal such that $|A^{3^{k+1}}| < K|A^{3^k}|$. Then

$$|S_i|^{ck/n} \leq K^k |A| \leq |A^{3^k}| \leq |X_{j,i}| = |S_i|^{O_n(1)},$$

which implies that $k = O_n(1)$. Since $|A^{3^k}|$ has tripling less than K , Lemma 1.8 implies that $A^{3^{k \cdot 14}} = X_{j,i}$. Thus $X_{j,i} = A^{O_n(1)}$.

This proves that the normal chain $U_0 \trianglelefteq \dots \trianglelefteq U_k$ is $O_n(1)$ -bounded. By Lemma 9.5 the G -action on each G -invariant section of P has an $O_n(1)$ -bounded G -invariant normal chain of length less than n^2 . \square

Lemma 9.7. *Let G be a finite group acting on an abelian p -group V such that $V = [V, G]$. Assume that $G/O_p(G)$ is d -generated and a central product of quasisimple groups. Suppose that V has a B -bounded G -invariant normal chain of length k . If $A \subseteq V$ is a G -invariant symmetric generating set containing 1 then $A^{O(Bd \log |V|)^k} = V$.*

Proof. Let $1 = V_0 \trianglelefteq V_1 \trianglelefteq \dots \trianglelefteq V_k = V$ be a B -bounded normal chain. We prove the statement by induction on k .

If $k = 0$ then V is trivial and the statement clearly holds.

For the induction step we assume that the statement holds for the group V/V_1 . This gives us an exponent $c = O(Bd \log |V|)^{k-1}$ such that A^c covers V/V_1 . According to Definition 9.4(d) we have two possibilities.

Suppose first that V_1 is a trivial $\mathbf{F}_p G$ -module. Let g_1, \dots, g_d be elements of G whose images generate $G/O_p(G)$. Then

$$\{[v_1, g_1] \cdots [v_d, g_d] : v_1, \dots, v_d \in V\} = [V, G] = V$$

by Lemma 7.10. Since $[V_1, G] = 1$, $[v_i, g_i]$ depends only on the image of v_i in V/V_1 . Therefore

$$\{[v_1, g_1] \cdots [v_d, g_d] : v_1, \dots, v_d \in A^c\} = V.$$

Hence $A^{2dc} = V$ and the induction step is complete in this case.

Consider now the case when V_1 is a nontrivial $\mathbf{F}_p G$ -module. Let S be one of the quasisimple factors of G/V which acts on V_1 nontrivially. By Definition 9.4(d), as an $\mathbf{F}_p S$ -module V_1 is the direct sum of isomorphic copies of an irreducible $\mathbf{F}_p S$ -module X such that

$$X = (Sx)^{\pm B} \quad \text{for all nonzero } x \in X.$$

By Lemma 4.6 (Schreier's lemma), $A_1 = A^{3c} \cap V_1$ is a symmetric generating set of V_1 . Each element $a \in A_1$ generates an $\mathbf{F}_p S$ -submodule $X_a \leq V_1$ isomorphic to X , so $X_a = (Sa)^{\pm B}$. Hence A_1^B contains X_a for every $a \in A_1$. Since V_1 is the sum of the submodules X_a , it is the sum of some $\log_p |V_1|$ of them. Therefore $V_1 = A_1^{B \log_p |V_1|}$, and so $V \subseteq A^{c+cB \log_p |V_1|} \subseteq A^{O(Bd \log |V|)^k}$. The induction step is complete in this case too. \square

Proof of Theorem 1.6 (=Theorem 9.1). If \mathbf{F} has characteristic 0 then by Jordan's theorem G has an abelian subgroup of index $O_n(1)$. By choosing $K(n)$ larger than this bound we can ensure that G is abelian and hence G is trivial. The statement holds in this case.

Hence assume that \mathbf{F} has characteristic $p > 0$. By Theorem 7.2 and Corollary 7.4, $G/O_p(G)$ is a central product of quasisimple groups of Lie type of characteristic p , $P = [G, \text{Sol}(G)]$ is a p -subgroup such that $P = [P, G]$, and $|\text{Sol}(G)/P| \leq (2n+1)^n$. By Lemma 7.1, $G/\text{Sol}(G) \in \text{Lie}^n(p)$. In particular, G/P is $O_n(1)$ -generated.

Let $A \subseteq G$ be a symmetric generating set. Since $G/\text{Sol}(G) \in \text{Lie}^n(p)$ and $\text{Sol}(G)/P$ has bounded order, G/P has poly-logarithmic diameter. Hence A^ℓ covers G/P , where $\ell = c_1(\log |G|)^{c_2}$ and $c_1 = c_1(n)$ and $c_2 = c_2(n)$.

By Schreier's lemma (Lemma 4.6), $A^{3\ell} \cap P$ generates P . Therefore $A^{5\ell}$ contains a set $B \subseteq P$ whose image in $P/[P, P]$ is a symmetric G -invariant generating set containing 1.

If K is large enough then by Theorem 9.6 the section $P/[P, P]$ has an $O_n(1)$ -bounded G -invariant normal chain of length less than n^2 . Applying Lemma 9.7 to G and $P/[P, P]$, we obtain that $X = B^{O_n(\log |P|)^{n^2}}$ maps surjectively onto $P/[P, P]$.

The nilpotency class of P is at most $n - 1$. Consider the set $X_i \subseteq X^{2^{i+1}}$ of all weight- i left-normed commutators $[x_1, \dots, x_i]$, where $x_1, \dots, x_i \in X$. The image of X_i in $\gamma_i(P)/\gamma_{i+1}(P)$ is a generating set and a union of subgroups, so $X_i^{O(\log |P|)}$ covers $\gamma_i(P)/\gamma_{i+1}(P)$. This shows that $X^{O(2^i \log |P|)}$ covers $\gamma_i(P)/\gamma_{i+1}(P)$ for each $i = 1, \dots, n - 1$, so $X^{O(2^n \log |P|)}$ covers P . Thus $A^{O_n(\log |G|)^{O_n(1)}}$ covers G . \square

REFERENCES

- [A] M. Aschbacher, *Finite group theory*, Second, Cambridge Studies in Advanced Mathematics, vol. 10, Cambridge University Press, Cambridge, 2000. MR1777008 [↑2](#), [4.11](#), [4.12](#), [7.3](#), [7.1](#), [8.3.3](#)
- [B1] E. Bannai, *Maximal subgroups of low rank of finite symmetric and alternating groups*, Journal of the Faculty of Science, the University of Tokyo. Sect. 1 A, Mathematics **18** (1972), no. 3, 475–486. [↑3.4](#)
- [B2] E. Breuillard, *A brief introduction to approximate groups*, Thin groups and super-strong approximation, 2014, pp. 23–50. MR3220883 [↑3](#), [3.2](#)
- [BGT1] E. Breuillard, B. Green, and T. Tao, *Approximate subgroups of linear groups*, Geom. Funct. Anal. **21** (2011), no. 4, 774–819. MR2827010 [↑1.1](#), [1.5](#), [1.2](#), [7.7](#)
- [BGT2] ———, *The structure of approximate groups*, Publ. Math. Inst. Hautes Études Sci. **116** (2012), 115–221. MR3090256 [↑1.1](#), [1.1](#), [1.1](#), [3.1](#)
- [BNP] L. Babai, N. Nikolov, and L. Pyber, *Product growth and mixing in finite groups*, Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms, 2008, pp. 248–257. MR2485310 [↑4.8](#)
- [BT] E. Breuillard and M. C. H. Tointon, *Nilprogressions and groups with moderate growth*, Adv. Math. **289** (2016), 1008–1055. MR3439705 [↑1.1](#), [3.1](#)
- [D] J. D. Dixon, *The Fitting subgroup of a linear solvable group*, J. Austral. Math. Soc. **7** (1967), 417–424. MR0230814 [↑3.1](#)
- [E] S. Eberhard, *A note on nonabelian Freiman-Ruzsa*, arXiv e-prints (Nov. 2015), arXiv:1511.06758, available at [1511.06758](#). [↑1.1](#), [3.1](#)
- [FT] W. Feit and J. Tits, *Projective representations of minimum degree of group extensions*, Canadian J. Math. **30** (1978), no. 5, 1092–1102. MR498824 [↑7.1](#)
- [G] W. T. Gowers, *Quasirandom groups*, Combin. Probab. Comput. **17** (2008), no. 3, 363–387. MR2410393 [↑4.2](#)
- [GH] N. Gill and H. A. Helfgott, *Growth in solvable subgroups of $GL_r(\mathbb{Z}/p\mathbb{Z})$* , Math. Ann. **360** (2014), no. 1–2, 157–208. MR3263161 [↑1.1](#), [1.1](#), [1.2](#), [1.3](#), [8.1](#), [8.2](#), [8.3.1](#)
- [GS] R. M. Guralnick and J. Saxl, *Generation of finite almost simple groups by conjugates*, Journal of Algebra **268** (2003), no. 2, 519–571. [↑3.4](#)
- [H1] H. A. Helfgott, *Growth in $SL_3(\mathbb{Z}/p\mathbb{Z})$* , J. Eur. Math. Soc. (JEMS) **13** (2011), no. 3, 761–851. MR2781932 [↑1.1](#), [8.2](#)
- [H2] H. A. Helfgott, *Growth in groups: ideas and perspectives*, Bull. Amer. Math. Soc. (N.S.) **52** (2015), no. 3, 357–413. MR3348442 [↑1.1](#), [3.2](#), [4.1](#), [4.1](#), [4.2](#), [4.1](#), [4.4](#)
- [HS] H. A. Helfgott and Á. Seress, *On the diameter of permutation groups*, Annals of mathematics (2014), 611–658. [↑4.1](#)
- [KL] P. Kleidman and M. Liebeck, *The subgroup structure of the finite classical groups*, London Mathematical Society Lecture Note Series, vol. 129, Cambridge University Press, Cambridge, 1990. MR1057341 [↑7.1](#), [7.1](#), [7.1](#), [7.5](#)
- [LL] R. Lawther and M. W. Liebeck, *On the diameter of a Cayley graph of a simple group of Lie type based on a conjugacy class*, J. Combin. Theory Ser. A **83** (1998), no. 1, 118–137. MR1629452 [↑7.2](#)

- [LP1] M. J. Larsen and R. Pink, *Finite subgroups of algebraic groups*, J. Amer. Math. Soc. **24** (2011), no. 4, 1105–1158. MR2813339 ↑7.1, 7.2, 7.3
- [LP2] M. W. Liebeck and L. Pyber, *Finite linear groups and bounded generation*, Duke Math. J. **107** (2001), no. 1, 159–171. MR1815254 ↑7.3
- [LP3] M. W. Liebeck and L. Pyber, *Upper bounds for the number of conjugacy classes of a finite group*, J. Algebra **198** (1997), no. 2, 538–562. MR1489911 ↑7.1
- [LS] V. Landazuri and G. M. Seitz, *On the minimal degrees of projective representations of the finite Chevalley groups*, J. Algebra **32** (1974), 418–443. MR360852 ↑4.7
- [M] G. J. McNinch, *Linearity for actions on vector groups*, Journal of Algebra **397** (2014), 666–688. ↑9
- [NP] N. Nikolov and L. Pyber, *Product decompositions of quasirandom groups and a Jordan type theorem*, J. Eur. Math. Soc. (JEMS) **13** (2011), no. 4, 1063–1077. MR2800484 ↑4.2, 4.2, 7.3
- [PS1] L. Pyber and E. Szabó, *Growth in finite simple groups of Lie type of bounded rank*, arXiv e-prints (May 2010), arXiv:1005.1858, available at 1005.1858. ↑1.1, 1.2
- [PS2] L. Pyber and E. Szabó, *Growth in linear groups*, Thin groups and superstrong approximation, 2014, pp. 253–268. MR3220893 ↑1.1
- [PS3] ———, *Growth in finite simple groups of Lie type*, J. Amer. Math. Soc. **29** (2016), no. 1, 95–146. MR3402696 ↑1.1, 1.5, 1.1, 1.2, 9
- [R] A. Rhemtulla, *Commutators of certain finitely generated soluble groups*, Canadian Journal of Mathematics **21** (1969), 1160–1164. ↑7.3
- [S1] A. Shalev, *Word maps, conjugacy classes, and a noncommutative Waring-type theorem*, Ann. of Math. (2) **170** (2009), no. 3, 1383–1416. MR2600876 ↑7.3
- [S2] R. Steinberg, *Representations of algebraic groups*, Nagoya Mathematical Journal **22** (1963), 33–56. ↑9
- [T1] T. Tao, *Freiman’s theorem for solvable groups*, 2009. <https://terrytao.wordpress.com/2009/06/21/freimans-theorem-for-solvable-groups/#comment-39705>. ↑1.1
- [T2] M. C. H. Tointon, *Freiman’s theorem in an arbitrary nilpotent group*, Proc. Lond. Math. Soc. (3) **109** (2014), no. 2, 318–352. MR3254927 ↑(b)
- [V] A. Vistoli, *Degree of image of a polynomial map*. <https://mathoverflow.net/q/63463> (version: 2011-04-29). ↑9
- [W1] B. A. F. Wehrfritz, *Infinite linear groups. An account of the group-theoretic properties of infinite groups of matrices*, Springer-Verlag, New York-Heidelberg, 1973. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 76. MR0335656 ↑5.1, 6, 6.1, 6, 6
- [W2] B. Weisfeiler, *Post-classification version of Jordan’s theorem on finite linear groups*, Proc. Nat. Acad. Sci. U.S.A. **81** (1984), no. 16, , Phys. Sci., 5278–5279. MR758425 ↑7.1, 7.2, 7.3
- [W3] J. S. Wilson, *On simple pseudofinite groups*, J. London Math. Soc. (2) **51** (1995), no. 3, 471–490. MR1332885 ↑7.3

SEAN EBERHARD, MATHEMATICAL SCIENCES RESEARCH CENTRE, QUEEN'S UNIVERSITY BELFAST, BELFAST BT7 1NN, UK

Email address: `s.eberhard@qub.ac.uk`

BRENDAN MURPHY, ATMOSPHERIC CHEMISTRY RESEARCH GROUP, SCHOOL OF CHEMISTRY, UNIVERSITY OF BRISTOL, BRISTOL BS8 1TS, UK

Email address: `brendan.murphy@bristol.ac.uk`

LÁSZLÓ PYBER, A. RÉNYI INSTITUTE OF MATHEMATICS, EÖTVÖS LORÁND RESEARCH NETWORK, P.O. BOX 127, H-1364 BUDAPEST, HUNGARY

Email address: `pyber@renyi.hu`

ENDRE SZABÓ, A. RÉNYI INSTITUTE OF MATHEMATICS, EÖTVÖS LORÁND RESEARCH NETWORK, P.O. BOX 127, H-1364 BUDAPEST, HUNGARY

Email address: `endre@renyi.hu`