J. Eur. Math. Soc. 14, 273–305

DOI 10.4171/JEMS/302

Péter P. Varjú

Expansion in $SL_d(\mathcal{O}_K/I)$, I square-free

Received January 21, 2010 and in revised form October 16, 2010

Abstract. Let S be a fixed symmetric finite subset of $SL_d(\mathcal{O}_K)$ that generates a Zariski dense subgroup of $SL_d(\mathcal{O}_K)$ when we consider it as an algebraic group over $\mathbb{Q}$ by restriction of scalars. We prove that the Cayley graphs of $SL_d(\mathcal{O}_K/I)$ with respect to the projections of S is an expander family if I ranges over square-free ideals of $\mathcal{O}_K$ if $d = 2$ and K is an arbitrary number field, or if $d = 3$ and $K = \mathbb{Q}$.

1. Introduction

Let $\mathcal{G}$ be a graph, and for a set of vertices $X \subset V(\mathcal{G})$, denote by ∂X the set of edges that connect a vertex in X to one in $V(\mathcal{G})\backslash X$. Define

$$c(\mathcal{G}) = \min_{X \subset V(\mathcal{G}),\, |X| \le |V(\mathcal{G})|/2} \frac{|\partial X|}{|X|},$$

where $|X|$ denotes the cardinality of the set X. A family of graphs is called a *family of expanders* if $c(\mathcal{G})$ is bounded away from zero for graphs $\mathcal{G}$ that belong to the family. Expanders have a wide range of applications in computer science (see e.g. Hoory, Linial and Widgerson [23] for a recent survey of expanders and their applications) and recently they found remarkable applications in pure mathematics as well (see Bourgain, Gamburd and Sarnak [9] and Long, Lubotzky and Reid [25]).

Let G be a group and let $S \subset G$ be a symmetric (i.e. closed for taking inverses) set of generators. The *Cayley graph* $\mathcal{G}(G, S)$ of G with respect to the generating set S is defined to be the graph whose vertex set is G, and in which two vertices $x, y \in G$ are connected exactly if $y \in Sx$. Let K be a number field and denote by $\mathcal{O}_K$ its ring of integers. Let $I \subset \mathcal{O}_K$ be an ideal, and denote by π_I the projection $\mathcal{O}_K \to \mathcal{O}_K/I$. In this paper we study the problem whether the graphs $\mathcal{G}(SL_d(\mathcal{O}_K/I), \pi_I(S))$ form an expander family, where $S \subset SL_d(\mathcal{O}_K)$ is a fixed symmetric set of matrices and I runs through certain ideals of $\mathcal{O}_K$. This problem was addressed by Bourgain and Gamburd in a series of papers [5]–[7], and by them jointly with Sarnak in [9]. It is solved for $K = \mathbb{Q}$ in the following cases: in [5] for $d = 2$ when and $I = (p)$ runs through primes, in [9] for $d = 2$

P. P. Varjú: Department of Mathematics, Princeton University, Princeton, NJ 08544, USA, and Analysis and Stochastics Research Group of the Hungarian Academy of Sciences, University of Szeged, Szeged, Hungary; e-mail: pvarju@princeton.edu

and $I = (q)$, q square-free, and in [6] and [7] when $I = (p^n)$, p^n a prime power. (When $d \geq 3$, the prime p has to be kept fixed.) A necessary and sufficient condition in each case for the Cayley graphs to be expanders is that S generates a Zariski dense subgroup $\Gamma < SL_d(\mathbb{C})$. In [9] the expander property is used for $K = \mathbb{Q}(\sqrt{-1})$ for sieving in the context of integral Apollonian packings; this is our main motivation for extending the problem to general number fields.

The starting point for our study is the work of Helfgott [21], [22]. He studies the following problem: Let $\mathcal{F}$ be a family of finite fields and let $d \geq 2$ be an integer. Is there a constant $\delta > 0$ such that for any generating set $A \subset SL_d(F)$ with $F \in \mathcal{F}$ we have

$$|A.A.A| \geq |A| \min(|A|, |SL_d(F)|/|A|)^{\delta}? \tag{1}$$

Here and everywhere in what follows, we use the notation $A.B = \{gh \mid g \in A, h \in B\}$ if A and B are subsets of a multiplicative group. Helfgott answers this question in the affirmative when $\mathcal{F}$ is the family of prime fields and $d = 2$ [21] or $d = 3$ [22]. In Section 4.1 we show that [21] (i.e. the proof for the case $d = 2$) easily extends to the case of arbitrary finite fields.

Let r be the degree of the number field K, and denote by $\sigma_1, \ldots, \sigma_r$ the embeddings of K into $\mathbb{C}$. Denote by $\widehat{\sigma} = \sigma_1 \oplus \cdots \oplus \sigma_r$ the obvious map $K \to \mathbb{C}^r$. This gives rise to an embedding (which will also be denoted by $\widehat{\sigma}$) of $SL_d(\mathcal{O}_K)$ into the direct product $SL_d(\mathbb{C})^r$. Our main result is

Theorem 1. *Let $S \subset SL_d(\mathcal{O}_K)$ be symmetric and assume that it generates a subgroup $\Gamma < SL_d(\mathcal{O}_K)$ such that $\widehat{\sigma}(\Gamma) \subset SL_d(\mathbb{C})^r$ is Zariski dense. Assume further that (1) holds for some constant $\delta > 0$ if F ranges over the fields $\mathcal{O}_K/P$, where $P \subset \mathcal{O}_K$ is a prime ideal. Then there is an ideal $J \subset \mathcal{O}_K$ such that $\mathcal{G}(SL_d(\mathcal{O}_K/I), \pi_I(S))$ is a family of expanders where $I \subset \mathcal{O}_K$ ranges over square-free ideals prime to J.*

It is part of the claim that $\pi_I(S)$ generates $SL_d(\mathcal{O}_K/I)$ if I is prime to J. In fact, J can be taken to be the product of prime ideals P for which $\pi_P(S)$ does not generate $SL_d(\mathcal{O}_K/P)$; this fact will be proven together with the theorem. We remark that the condition on Zariski density is necessary, otherwise $\pi_{(q)}(S)$ would not generate $SL_d(\mathcal{O}_K/(q))$ for any rational integer q. Note that by the above remarks on Helfgott's work, the theorem is unconditional for $d = 2$ and arbitrary K, and for $d = 3$ and $K = \mathbb{Q}$.

We introduce some notation that will be used throughout the paper. We use Vinogradov's notation $x \ll y$ as a shorthand for $|x| < Cy$ with some constant C. Let G be a discrete group. The unit element of any multiplicatively written group is denoted by 1. For given subsets A and B of G, we denote their product-set by

$$A.B = \{gh \mid g \in A,\ h \in B\},$$

while the k-fold iterated product-set of A is denoted by $\prod_k A$. We write $\widetilde{A}$ for the set of inverses of all elements of A. We say that A is *symmetric* if $A = \widetilde{A}$. The number of elements of a set A is denoted by $|A|$. The index of a subgroup H of G is denoted by $[G : H]$ and we write $H_1 \lesssim_L H_2$ if $[H_1 : H_1 \cap H_2] \leq L$ for some subgroups $H_1, H_2 < G$.

Occasionally (especially when a ring structure is present) we write groups additively, then we write

$$A + B = \{g + h \mid g \in A,\ h \in B\}$$

for the sum-set of A and B, $\sum_k A$ for the k-fold iterated sum-set of A, and 0 for the zero element.

If μ and ν are complex valued functions on G, we define their convolution by

$$(\mu * \nu)(g) = \sum_{h \in G} \mu(gh^{-1})\nu(h),$$

and we define $\widetilde{\mu}$ by the formula

$$\widetilde{\mu}(g) = \mu(g^{-1}).$$

We write $\mu^{(k)}$ for the k-fold convolution of μ with itself. As measures and functions are essentially the same on discrete sets, we use these notions interchangeably; we will also use the notation

$$\mu(A) = \sum_{g \in A} \mu(g).$$

A *probability measure* is a nonnegative measure with total mass 1. Finally, the *normalized counting measure* on a finite set A is the probability measure

$$\chi_A(B) = |A \cap B|/|A|.$$

We use the same approach to prove Theorem 1 as in [5]–[9], which goes back to [31]; we outline this here only, the details will be given in Section 5. Let $\mathcal{G}$ be an m-*regular* graph, i.e. each vertex is of degree m. It is easy to see that the largest eigenvalue of the adjacency matrix of $\mathcal{G}$ is m, and it is a simple eigenvalue if and only if the graph is connected. Denote by $\lambda_2(\mathcal{G})$ the second largest eigenvalue of the adjacency matrix. It was proven by Dodziuk [15], Alon and Milman [3] and Alon [2] that a family of graphs is an expander family if and only if $m - \lambda_2(\mathcal{G})$ is bounded away from zero (see also [23, Theorem 2.4]). For a Cayley graph $\mathcal{G}(G, S)$, the adjacency matrix is a constant multiple of convolution with χ_S from the left considered as an operator. Then the multiplicities of the nontrivial eigenvalues are at least the minimum dimension of a nontrivial representation of G. In the case of SL_d good bounds are known, hence it is enough to estimate the trace of the operator. More precisely, with the notation of Theorem 1, we need to show that for any $\varepsilon > 0$ there is a constant $C = C(\varepsilon, S)$ such that

$$\|\pi_I[\chi_S^{(C \log N(I))}]\|_2 < |SL_d(\mathcal{O}_K/I)|^{-1/2+\varepsilon}, \tag{2}$$

where $N(I)$ is the norm of the ideal. In fact, (2) means that the random walk on the graph $\mathcal{G}(SL_d(\mathcal{O}_K/I), \pi_I(S))$ is close to equidistribution after $C \log N(I)$ steps.

The proof of (2) has two parts; the first is

Theorem 2. *Let $S \subset SL_d(\mathcal{O}_K)$ be symmetric, and denote by Γ the subgroup it generates. Assume that $\widehat{\sigma}(\Gamma)$ is Zariski dense in $SL_d(\mathbb{C})^r$. Then there is a constant δ depending only on S, and there is a symmetric set $S' \subset \Gamma$ such that the following holds. For any square-free ideal I, for any proper subgroup $H < SL_d(\mathcal{O}_K/I)$ and for any even integer $l \geq \log N(I)$, we have*

$$\pi_I[\chi_{S'}^{(l)}](H) \ll [SL_d(\mathcal{O}_K/I) : H]^{-\delta}.$$

If we know that $g \in \prod_{c \log N(I)} S$, where c is a small constant depending on S, then $\pi_I(g)$ determines g uniquely. In Section 2, using Nori's [27] results we give a geometric description of the elements of $\prod_{c \log N(I)} S$ whose projection modulo I belongs to H; this will be a certain subgroup of $SL_d(\mathcal{O}_K)$. Then we will prove that the probability for the random walk on $\mathcal{G}(\Gamma, S)$ to be in this subgroup decays exponentially in the number of steps we take. (Actually, first we need to replace S by another set $S' \subset \Gamma$.) The proof of this is based on a ping-pong argument.

The second part of the proof begins with the following observation. If we apply Theorem 2 for $H = \{1\}$, then we already get

$$\|\pi_I[\chi_{S'}^{(\log N(I))}]\| \ll \|SL_d(\mathcal{O}_K/I)\|^{-\delta/2}. \tag{3}$$

Now working on the quotient $SL_d(\mathcal{O}_K/I)$, we can improve on (3), if we take the convolution of $\pi_I[\chi_{S'}^{(\log N(I))}]$ with itself. More precisely we prove in Section 3 the following

Theorem 3. *Let G be a group satisfying the assumptions* (A0)–(A5) *listed in Section 3. Then for any $\varepsilon > 0$, there is some $\delta > 0$ depending only on ε and the constants appearing in assumptions* (A0)–(A5) *such that the following holds. If μ and ν are probability measures on G such that*

$$\|\mu\|_2 > |G|^{-1/2+\varepsilon} \quad \textit{and} \quad \mu(gH) < [G : H]^{-\varepsilon}$$

for any $g \in G$ and for any proper subgroup $H < G$, then

$$\|\mu * \nu\|_2 < \|\mu\|_2^{1/2+\delta}\|\nu\|_2^{1/2}.$$

Assumptions (A0)–(A5) are too technical so we do not list them here in the introduction. Among other things, we assume that G is the direct product of quasi-simple groups that also satisfy the conclusion of the theorem. To prove the latter for the groups $SL_d(\mathcal{O}_K/P)$ we need (1), and this is the reason why we have Theorem 1 only in the cases when (1) is available. The quasi-simplicity of the factors is a severe restriction, for example it excludes factors of the form $SL_d(\mathcal{O}_K/P^k)$, where P is a prime ideal. Therefore a new idea is needed to prove Theorem 1 for general ideals.

A similar result for $G = SL_2(\mathbb{Z}/q\mathbb{Z})$, q square-free (under a stronger hypothesis on μ) is given by Bourgain, Gamburd and Sarnak [9, Proposition 4.3]. They use an argument similar to Helfgott's [21] to reduce it to a so-called sum-product theorem for the ring $\mathbb{Z}/q\mathbb{Z}$. Then they prove the latter by reducing it to the case of $\mathbb{Z}/p\mathbb{Z}$, p prime. The difference in our approach is that we use Helfgott's theorem as a black box, and extend it to the case of square-free modulus in a way that very much resembles the proof given in [9, Section 5] for the sum-product theorem.

2. Escape of mass from subgroups

We prove Theorem 2 in this section. First we note that we may assume that I is a principal ideal generated by a square-free rational integer q. Indeed, there is always a square-free rational integer $q \in I$ such that $q \leq N(I)$. Let $\widehat{H}$ be the preimage of H under the projection $SL_d(\mathcal{O}_K/(q)) \to SL_d(\mathcal{O}_K/I)$. Then we have $\log N((q)) \geq \log N(I) \geq \log N((q))/r$ and $[SL_d(\mathcal{O}_K/I) : H] = [SL_d(\mathcal{O}_K/(q)) : \widehat{H}]$. Hence the claim of the theorem for I and H follows from the claim for (q) and $\widehat{H}$. In what follows we assume that $I = (q)$ and write $\pi_q = \pi_{(q)}$. Let $q = p_1 \cdots p_n$ be the prime factorization of q and assume without loss of generality that none of the p_i ramify in K.

For $g \in SL_d(\mathbb{C})$ denote by $\|g\|$ the operator norm of g with respect to the l^2 norm on $\mathbb{C}^d$. If $\|g\| < \sqrt{q}/2$ for some $g \in SL_d(\mathcal{O}_K)$, then clearly $\|g'\| > \sqrt{q}/2$ for any other $g' \in SL_d(\mathcal{O}_K)$ with $\pi_q(g) = \pi_q(g')$ since $\|g\| \geq \sqrt{q}$ if $\pi_q(g) = 0$ and $g \neq 0$. Hence elements of small norm are uniquely determined by their projections modulo q. The first step towards the proof of Theorem 2 is to study when the projection of an element of small norm belongs to H, i.e. we study the set

$$\mathcal{L}_\delta(H) := \{h \in SL_d(\mathcal{O}_K) \mid \pi_q(h) \in H,\ \|\widehat{\sigma}(h)\| < [SL_d(\mathcal{O}_K/(q)) : H]^\delta\}$$

for $\delta > 0$ and $H < SL_d(\mathcal{O}_K/(q))$.

By Weil restriction of scalars, we consider $SL_d(K)$ as the $\mathbb{Q}$-points of an algebraic group. To fix notation, we describe this process in detail. Let $e_1, \ldots, e_r$ be an integral basis of $\mathcal{O}_K$. Multiplication by an element $a \in K$ is an endomorphism of the $\mathbb{Q}$-vector space K. This gives rise to an embedding $\alpha : K \to Mat_r(\mathbb{Q})$ onto a subalgebra of $Mat_r(\mathbb{Q})$ which is defined by linear equations over $\mathbb{Q}$. Thus there is an algebraic subgroup $\mathbb{G}$ of SL_{dr} defined over $\mathbb{Q}$ such that $SL_d(K)$ is isomorphic to $\mathbb{G}(\mathbb{Q})$ as an abstract group; we denote this isomorphism by α as well. Moreover, we have $\alpha(SL_d(\mathcal{O}_K)) = \mathbb{G}(\mathbb{Q}) \cap SL_{dr}(\mathbb{Z})$. To shorten notation, we write $\mathbb{G}(\mathbb{Z}) = \mathbb{G}(\mathbb{Q}) \cap SL_{dr}(\mathbb{Z})$. The image of $e_1, \ldots, e_r$ under π_q is a basis of the $\mathbb{Z}/q\mathbb{Z}$-module $\mathcal{O}_K/(q)$, hence α induces an isomorphism from $SL_d(\mathcal{O}_K/(q))$ to $\mathbb{G}(\mathbb{Z}/q\mathbb{Z})$. Denote by $\mathfrak{g}$ the Lie algebra of $\mathbb{G}$. Then $\mathfrak{g}(\mathbb{Q})$ is a subspace of $Mat_{dr}(\mathbb{Q})$ defined by (linear) polynomials $\varphi_1, \ldots, \varphi_{d^2r^2-r(d^2-1)} \in \mathbb{Z}[x]$. If p is a prime which does not ramify in K, then we can write $(p) = P_1 \cdots P_k$ with different prime ideals P_i. Then $\mathbb{G}(\mathbb{Z}/p\mathbb{Z})$ is isomorphic to $SL_d(\mathcal{O}_K/P_1) \times \cdots \times SL_d(\mathcal{O}_K/P_k)$.

Proposition 4. *There are constants C and δ depending only on K such that the following holds. For any subgroup $H < SL_d(\mathcal{O}_K/(q))$, there are $u, v \in \mathfrak{g}(\mathbb{C})$ and a subgroup $H^\sharp < H$ with $[H : H^\sharp] < C^n$ such that if $h \in \mathcal{L}_\delta(H^\sharp)$ then*

$$\mathrm{Tr}(\alpha(h)u\alpha(h)^{-1}v) = 0,$$

but there is some $g_0 \in \mathbb{G}(\mathbb{Q})$ such that $\mathrm{Tr}(g_0ug_0^{-1}v) = 1$.

In what follows we often write $\mathbb{F}_{p^m}$ for the finite field of order p^m. Recall that n is the number of prime factors of q and $q = p_1 \cdots p_n$. Then $\mathbb{G}(\mathbb{Z}/q\mathbb{Z}) = \mathbb{G}(\mathbb{F}_{p_1}) \times \cdots \times \mathbb{G}(\mathbb{F}_{p_n})$. For $q_1 \mid q$, denote by

$$\pi_{q_1} : \mathbb{G}(\mathbb{Z}/q\mathbb{Z}) \to \mathop{\times}_{p \mid q_1} \mathbb{G}(\mathbb{F}_p)$$

the projection to the product of direct factors corresponding to the prime factors of q_1. Fix a proper subgroup $H < \mathbb{G}(\mathbb{Z}/q\mathbb{Z})$ and denote by q_1 the product of all primes $p \mid q$ for which $\pi_p(H) = \mathbb{G}(\mathbb{F}_p)$. In the course of the proof we will replace q by q/q_1 and H by $\pi_{q/q_1}(H)$. We need to show that $[\mathbb{G}(\mathbb{Z}/(q/q_1)\mathbb{Z}) : \pi_{q/q_1}(H)]$ is not much smaller than $[\mathbb{G}(\mathbb{Z}/q\mathbb{Z}) : H]$. For this we first give

Lemma 5. *Let p_1 and p_2 be two different primes and assume that $N \lhd H < SL_d(\mathbb{F}_{p_2^{m_1}})$ are such that H/N is isomorphic to $PSL_d(\mathbb{F}_{p_1^{m_2}})$ with some integers m_1, m_2. Then*

$$p_1 \mid \prod_{i=2}^{d}(p_2^{im_1} - 1);$$

in particular, for a fixed p_2 the product of all primes which can arise as p_1 is at most $p_2^{d^2 m_1}$.

Proof. Since $PSL_d(\mathbb{F}_{p_1^{m_2}})$ has an element of order p_1 and since the order of $SL_d(\mathbb{F}_{p_2^{m_1}})$ is $p_2^{m_1 d(d-1)/2} \prod_{i=2}^{d}(p_2^{im_1} - 1)$, the assertion is clear. □

Lemma 6. *Let H be a subgroup of $G = \mathbb{G}(\mathbb{Z}/q\mathbb{Z})$ and denote by q_1 the product of primes $p \mid q$ with $\pi_p(H) = \mathbb{G}(\mathbb{F}_p)$ and set $q_2 = q/q_1$. There is a subgroup $H_2 < \mathbb{G}(\mathbb{Z}/q_2\mathbb{Z})$ of the form $\times_{p \mid q_2} H_p$, where each H_p is a proper subgroup of $\mathbb{G}(\mathbb{F}_p)$, such that $\pi_{q_2}(H) < H_2$ and*

$$[\mathbb{G}(\mathbb{Z}/q_2\mathbb{Z}) : H_2] > [G : H]^c$$

with a constant c depending only on d and r.

Proof. If for some $p \mid q_1$, $\mathbb{G}(\mathbb{F}_p)$ is a direct factor of H then

$$[\mathbb{G}(\mathbb{Z}/(q/p)\mathbb{Z}) : \pi_{q/p}(H)] = [G : H],$$

hence we can assume without loss of generality that there is no such prime. We show that for each $p_1 \mid q_1$, there is some $p_2 \mid q_2$ such that the conditions of the previous lemma are satisfied. This will yield a bound on q_1. Set $q' = q/p_1$. By Goursat's Lemma, there is a nontrivial group N and surjective homomorphisms

$$\varphi : \pi_p(H) = \mathbb{G}(\mathbb{F}_{p_1}) \to N, \qquad \psi : \pi_{q'}(H) \to N.$$

For each factor $p \mid q'$, ψ gives rise to a surjective homomorphism

$$\psi_p : \pi_p(H) \to N_p = N/\{\psi(h) \mid h \in \pi_{q'}(H),\ \pi_p(h) = 1\}$$

in the obvious way. Since the intersection of all the subgroups $\{\psi(h) \mid h \in \pi_{q'}(H)$, $\pi_p(h) = 1\}$ is trivial, there is a prime p_2 for which N_{p_2} is nontrivial. As $\mathbb{G}(\mathbb{F}_{p_1})$ and $\mathbb{G}(\mathbb{F}_{p_2})$ have no nontrivial common factors, $p_2 \mid q_2$. It is clear that p_1 and p_2 satisfy the conditions of Lemma 5, whence $q_1 < q_2^{rd^2}$.

For each $p \mid q_2$ let H_p be a proper subgroup of $\mathbb{G}(\mathbb{F}_p)$ containing $\pi_p(H)$. Since $\mathbb{G}(\mathbb{F}_p)$ is generated by its subgroups isomorphic to $SL_2(\mathbb{F}_p)$, there must be at least one such subgroup which is not contained in H_p. Any proper subgroup of $SL_2(\mathbb{F}_p)$ is of index at least $p+1$, hence $[\mathbb{G}(\mathbb{F}_p) : H_p] > p$. This shows that for $H_2 = \times_{p|q_2} H_p$, we have

$$[\mathbb{G}(\mathbb{Z}/q_2\mathbb{Z}) : H_2] > q_2 > q^{1/(d^2r+1)} > [G : H]^c. \qquad \square$$

The proof of Proposition 4 is based on the description of subgroups of $GL_d(\mathbb{F}_p)$ given by Nori [27] that we recall now. Let H be a subgroup of $GL_d(\mathbb{F}_p)$ and denote by H^+ the subgroup of H generated by its elements of order p. [27, Theorem B] states that if p is larger than a constant depending only on d, then there is a connected algebraic subgroup $\widetilde{H}$ of GL_d defined over $\mathbb{F}_p$ such that $H^+ = \widetilde{H}(\mathbb{F}_p)^+$. Denote by $\mathfrak{h}$ the Lie algebra of $\widetilde{H}$, and define exp and log by

$$\exp(z) = \sum_{i=0}^{p-1} \frac{z^i}{i!} \quad \text{and} \quad \log(z) = -\sum_{i=1}^{p-1} \frac{(1-z)^i}{i}$$

for $z \in Mat_d(\mathbb{F}_p)$. Then for p large enough, exp and log set up a one-to-one correspondence between elements of order p of H^+ and nilpotent elements of $\mathfrak{h}(\mathbb{F}_p)$ by [27, Theorem A]. Moreover $\mathfrak{h}(\mathbb{F}_p)$ is spanned by its nilpotent elements. To understand subgroups not generated by elements of order p, we will use [27, Theorem C] which asserts that if $p \geq d$, then there is a commutative subgroup $F < H$ such that FH^+ is a normal subgroup of H and its index $[H : FH^+]$ is bounded in terms of d.

Proof of Proposition 4. We follow the argument in [7, Proposition 4.1]. Recall that H is a subgroup of $SL_d(\mathcal{O}_K/(q))$. Apply Lemma 6 to $\alpha(H)$ to get a modulus $q_2 \mid q$ and a subgroup $H_2 < \mathbb{G}(\mathbb{Z}/q_2\mathbb{Z})$. Suppose that the conclusion holds for $\alpha^{-1}(H_2)$ and for an $H_2^\sharp < SL_d(\mathcal{O}_K/(q_2))$ with $[H_2 : \alpha(H_2^\sharp)] < C^n$. Set

$$H^\sharp = \{h \in H \mid \pi_{q_2}(h) \in H_2^\sharp\},$$

and observe that $[H : H^\sharp] < C^n$ and $\mathcal{L}_\delta(H^\sharp) \subset \mathcal{L}_{\delta/c}(H_2^\sharp)$ with the constant c from Lemma 6. Therefore, if the conclusion holds for $\alpha^{-1}(H_2)$ and $H_2^\sharp$, it also holds for H and $H^\sharp$. We assume in what follows that $\alpha(H) = H_{p_1} \times \cdots \times H_{p_n}$, where $q = p_1 \cdots p_n$ is the prime factorization of q and H_{p_i} is a proper subgroup of $\mathbb{G}(\mathbb{F}_{p_i})$. For each direct factor H_{p_i}, let $H_{p_i}^\sharp < H_{p_i}$ be such that $H_{p_i}^\sharp / H_{p_i}^+$ is commutative and $[H_{p_i} : H_{p_i}^\sharp] < C$ with a constant C depending on r and d; such a subgroup exists by [27, Theorem C]. Define $H^\sharp = \alpha^{-1}(H_{p_1}^\sharp \times \cdots \times H_{p_n}^\sharp)$.

For each $g \in \mathbb{G}(\mathbb{Z})$ define the polynomial $\eta_g \in \mathbb{Z}[X, Y]$ with $X = (X_{l,k})_{1\leq l,k\leq dr}$ and $Y = (Y_{l,k})_{1\leq l,k\leq dr}$ by

$$\eta_g(X, Y) = \mathrm{Tr}(gXg^{-1}Y).$$

Let A be a fixed set of generators of $\mathbb{G}(\mathbb{Z})$ and fix an element $g_0 \in A$. Consider the system of equations

$$\begin{aligned} \varphi_i(X) &= 0, \quad 1 \leq i \leq r^2d^2 - r(d^2-1), \\ \varphi_i(Y) &= 0, \quad 1 \leq i \leq r^2d^2 - r(d^2-1), \\ \eta_{\alpha(h)}(X, Y) &= 0 \quad\ \text{for } h \in \mathcal{L}_\delta(H^\sharp), \\ \eta_{g_0}(X, Y) &= 1, \end{aligned} \tag{4}$$

where δ is a small constant depending on d and r to be chosen later. Recall that φ_i are the polynomials defining the Lie algebra $\mathfrak{g}$. The assertion follows once we show that (4) has a solution $X = u$, $Y = v \in Mat_{rd}(\mathbb{C})$ for an appropriate choice of g_0.

First we show that for each $p = p_i$, there is at least one $g_0 \in A$ such that (4) has a solution in $Mat_{dr}(\mathbb{F}_p)$. We apply the results of [27] for $H = H_p$, in particular let $\widetilde{H}$ and $\mathfrak{h}$ be the same as in the discussion preceding the proof. Conjugation by an element $g \in \mathbb{G}(\mathbb{F}_p)$ permutes elements of order p of H_p^+ if and only if it permutes nilpotent elements of $\mathfrak{h}(\mathbb{F}_p)$. Hence $\mathfrak{h}(\mathbb{F}_p)$ is invariant under g in the adjoint representation exactly if g is in the normalizer of H_p^+. First we consider the case when H_p^+ is not a normal subgroup of $\mathbb{G}(\mathbb{F}_p)$. Then there is at least one element $\pi_p(g_0) \in \pi_p(A)$ whose adjoint action does not leave $\mathfrak{h}(\mathbb{F}_p)$ invariant. Let $u \in \mathfrak{h}(\mathbb{F}_p)$ be such that $\pi_p(g_0)u\pi_p(g_0)^{-1} \notin \mathfrak{h}(\mathbb{F}_p)$ and let $v \in \mathfrak{g}(\mathbb{F}_p)$ be orthogonal to $\mathfrak{h}(\mathbb{F}_p)$ with respect to the nondegenerate bilinear form $\langle x, y\rangle = \mathrm{Tr}(xy)$ and such that $\mathrm{Tr}(\pi_p(g_0)u\pi_p(g_0^{-1})v) = 1$. This settles the claim. Now consider the case when $H_p^+ \lhd \mathbb{G}(\mathbb{F}_p)$. If $(p) = P_1 \cdots P_k$ is the factorization of (p) over K, then $\mathbb{G}(\mathbb{F}_p)$ is isomorphic to $SL_d(\mathcal{O}_K/P_1) \times \cdots \times SL_d(\mathcal{O}_K/P_k)$, and H_p^+ must be the direct product of some of these factors. Consider a direct factor $SL_d(\mathcal{O}_K/P_i)$ which does not appear in H_p^+ and denote by N the projection of $H_p^\sharp$ to this factor. There is a Lie subalgebra $\mathfrak{g}_i(\mathbb{F}_p) \subset \mathfrak{g}(\mathbb{F}_p)$ which is isomorphic to $\mathfrak{sl}_d(\mathcal{O}_K/P_i)$, invariant and irreducible in the adjoint representation of $\mathbb{G}(\mathbb{F}_p)$, and the adjoint action of an element $g \in \mathbb{G}(\mathbb{F}_p)$ on $\mathfrak{g}_i(\mathbb{F}_p)$ is determined by its projection to the factor $SL_d(\mathcal{O}_K/P_i)$. If N is nontrivial denote by V the intersection of the $\mathcal{O}_K/P_i$-linear span of N in $Mat_d(\mathcal{O}_K/P_i)$ and the Lie algebra $\mathfrak{g}_i(\mathbb{F}_p)$. If N is trivial, let V be any proper subspace of $\mathfrak{g}_i(\mathbb{F}_p)$. Then V is again invariant under $H_p^\sharp$ in the adjoint representation but not under $\mathbb{G}(\mathbb{F}_p)$ and we can establish the claim the same way as above.

For a particular $g_0 \in A$, denote by q_{g_0} the product of the primes $p \mid q$ for which (4) has a solution over $\mathbb{F}_p$. As there are only a finite number (and bounded in terms of K) of possibilities for g_0, there is an appropriate choice such that $q_{g_0} > q^c$. Here and everywhere below, c is a constant depending only on K which need not be the same at different occurrences. Now assume to the contrary that the system (4) has no solution over $\mathbb{C}$. We can clearly replace the family of polynomials $\eta_\alpha(h)$, $h \in \mathcal{L}_\delta(H^\sharp)$, by a linearly independent subset of at most $M \le r^4 d^4$ elements that we denote by $\eta_1, \ldots, \eta_M$. Note that the coefficients of all the polynomials in (4) are bounded by $[G : H]^{c\delta} < q^{c'\delta}$. Using the effective Bézout identities proved by Berenstein and Yger [4, Theorem 5.1] we obtain polynomials

$$\begin{aligned}
\psi_1(X,Y), \ldots, \psi_M(X,Y) &\in \mathbb{Z}[X,Y],\\
\psi_1'(X,Y), \ldots, \psi'_{r^2d^2-r(d^2-1)}(X,Y) &\in \mathbb{Z}[X,Y],\\
\psi_1''(X,Y), \ldots, \psi''_{r^2d^2-r(d^2-1)}(X,Y) &\in \mathbb{Z}[X,Y],\\
\psi'''(X,Y) &\in \mathbb{Z}[X,Y]
\end{aligned}$$

and a positive integer $0 < D < q^{c\delta}$ such that

$$\begin{aligned}
D = \sum_{i=1}^{M} \eta_i(X,Y)\psi_i(X,Y) &+ \sum_{i=1}^{r^2d^2-r(d^2-1)} \varphi_i(X)\psi_i'(X,Y)\\
&+ \sum_{i=1}^{r^2d^2-r(d^2-1)} \varphi_i(Y)\psi_i''(X,Y) + (\eta_{g_0}(X,Y) - 1)\psi'''(X,Y).
\end{aligned}$$

Substituting the solution of (4) over $\mathbb{F}_p$ for all $p \mid q_{g_0}$, we see that $q_{g_0} \mid D$, a contradiction if δ is small enough. □

Corollary 7. *There are constants δ and C depending only on K, and for each $H < SL_d(\mathcal{O}_K/(q))$ there is an $H^\sharp < H$ with $[H : H^\sharp] < C^n$ such that at least one of the following holds:*

(i) *There is an embedding $\sigma : K \to \mathbb{C}$ and a proper subspace $V \subset \mathfrak{sl}_d(\mathbb{C})$ such that if $h \in \mathcal{L}_\delta(H^\sharp)$, then*

$$\sigma(h)V\sigma(h^{-1}) = V. \tag{5}$$

(ii) *There are two embeddings $\sigma_1, \sigma_2 : K \to \mathbb{C}$ and an invertible linear transformation $T : \mathfrak{sl}_d(\mathbb{C}) \to \mathfrak{sl}_d(\mathbb{C})$ such that*

$$T(\sigma_1(h)v\sigma_1(h^{-1})) = \sigma_2(h)T(v)\sigma_2(h^{-1}) \tag{6}$$

for any $h \in \mathcal{L}_\delta(H^\sharp)$ and $v \in \mathfrak{sl}_d(\mathbb{C})$.

Proof. Choose δ to be $1/r(d^2-1)$ times the δ in Proposition 4. Then there are $u, v \in \mathfrak{g}(\mathbb{C})$ and $g_0 \in \mathbb{G}(\mathbb{Q})$ such that $\mathrm{Tr}(\alpha(h)u\alpha(h^{-1})v) = 0$ for

$$h \in \textstyle\prod_{r(d^2-1)} \mathcal{L}_\delta(H^\sharp) \subset \mathcal{L}_{\delta r(d^2-1)}(H^\sharp),$$

while $\mathrm{Tr}(g_0ug_0^{-1}v) = 1$. Let U_l be the linear span of $\{\alpha(g)u\alpha(g^{-1}) \mid g \in \prod_l \mathcal{L}_\delta(H^\sharp)\}$ in $\mathfrak{g}(\mathbb{C})$. Comparing dimensions, we see that for some $l \leq r(d^2-1)$ we have $U_l = U_{l+1}$, and then it is invariant under $\alpha(\mathcal{L}_\delta(H^\sharp))$ in the adjoint representation. Write $U = U_l$. Then for any $x \in U$, we have $\mathrm{Tr}(xv) = 0$, hence $g_0ug_0^{-1} \notin U$, and U is not invariant under the full group $\mathbb{G}(\mathbb{C})$ in the adjoint representation.

Consider the embedding $\alpha : K \to Mat_r(\mathbb{Q})$. Let $a \in K$ be a generator of K over $\mathbb{Q}$. Note that the minimal polynomial of a over $\mathbb{Q}$ is the same as the minimal polynomial of $\alpha(a)$ in $Mat_r(\mathbb{Q})$. This polynomial has r different roots $\sigma_1(a), \dots, \sigma_r(a)$ in $\mathbb{C}$, hence there is a basis over $\mathbb{C}$ in which $\alpha(a)$ is diagonal. Any element $b \in K$ can be expressed as the value at a of a polynomial with rational coefficients. Thus in that basis the matrix of b is $\mathrm{diag}(\sigma_1(b), \dots, \sigma_l(b))$. Therefore there is an appropriate basis in which any $g \in \mathbb{G}(\mathbb{C})$ is a block diagonal matrix with $\sigma_1(g), \dots, \sigma_r(g)$ along the diagonal. This gives rise to an isomorphism $\beta : \mathbb{G}(\mathbb{C}) \to SL_d(\mathbb{C})^r$ such that $\sigma = \beta \circ \alpha$. Moreover β also induces an isomorphism between the Lie algebras $\mathfrak{g}(\mathbb{C})$ and $\mathfrak{sl}_d(\mathbb{C})^r$; denote by W the image of U.

Assume that W is a subspace of minimal dimension which is invariant under $\widehat{\sigma}[\mathcal{L}_\delta(H^\sharp)]$ in the adjoint representation, but not under the whole group $SL_d(\mathbb{C})^r$. Denote by $\mathfrak{g}_1(\mathbb{C}), \dots, \mathfrak{g}_r(\mathbb{C})$ the r copies of $\mathfrak{sl}_d(\mathbb{C})$ in $\mathfrak{sl}_d(\mathbb{C})^r$ and denote by π_i the projection to $\mathfrak{g}_i(\mathbb{C})$. For $1 \leq i \leq r$, the spaces $\pi_i(W)$ and $W \cap \mathfrak{g}_i(\mathbb{C})$ are invariant under $\sigma_i[\mathcal{L}_\delta(H^\sharp)]$ in the adjoint representation, hence (i) holds if the dimension of any of the above spaces is strictly between 0 and d^2-1. Suppose that this is not the case. Since W is not the direct sum of some $\mathfrak{g}_i(\mathbb{C})$, we may assume that say $W \cap \mathfrak{g}_1(\mathbb{C}) = \{0\}$ and $\pi_1(W) = \mathfrak{g}_1(\mathbb{C})$. By the minimality of the dimension of W, $\mathrm{Ker}(\pi_1) \cap W$ must be the direct sum of some $\mathfrak{g}_i(\mathbb{C})$. Since $\dim W > \dim \mathrm{Ker}(\pi_1) \cap W$, we can assume that say $\pi_2(\mathrm{Ker}(\pi_1) \cap W) = \{0\}$ and $\pi_2(W) = \mathfrak{g}_2(\mathbb{C})$. Then $T = \pi_2 \circ \pi_1^{-1}$ is well-defined and satisfies (ii). □

Recall that we are given a symmetric $S \subset SL_d(\mathcal{O}_K)$ which generates the subgroup Γ. We will choose an appropriate $S' \subset \Gamma$ and study the random walk on $\mathcal{G}(\langle S' \rangle, S')$, where $\langle S' \rangle$ is the subgroup generated by S'. In particular, we prove an exponential decay for the probability that after k steps we are in the subgroup of $SL_d(\mathcal{O}_K)$ whose elements satisfy (5) for some fixed V or in the one whose elements satisfy (6) for some fixed T.

Proposition 8. *Assume that $\widehat{\sigma}(\Gamma)$ is Zariski dense in $SL_d(\mathbb{C})^r$. Let V be a proper subspace of $\mathfrak{sl}_d(\mathbb{C})$, let $\sigma : K \to \mathbb{C}$ be an embedding, and denote by H_V the subgroup of elements $h \in SL_d(\mathcal{O}_K)$ for which (5) holds. Then*

$$\chi_S^{(k)}(H_V) \ll c^k$$

with some constant $c < 1$ depending only on S.

Proposition 9. *Assume that $\widehat{\sigma}(\Gamma)$ is Zariski dense in $SL_d(\mathbb{C})^r$. Then there is a symmetric set $S' \subset \Gamma$ and a constant $c < 1$ depending only on S such that the following holds. Let σ_1, σ_2 be two different embeddings of K into $\mathbb{C}$ and let T be an invertible linear transformation on $\mathfrak{sl}_d(\mathbb{C})$. Denote by H_T the subgroup of elements $h \in SL_d(\mathcal{O}_K)$ for which (6) holds. Then*

$$\chi_{S'}^{(k)}(H_T) \ll c^k.$$

Proposition 8 can be proved as outlined in [6, Section 9]; we omit the details. A weaker form analogous to Proposition 9, which is sufficient for our purposes, can be proved by the same method as we prove Proposition 9 below.

Let $A \subset \Gamma$ be a subset that freely generates a subgroup. By abuse of notation, by a *word* w over $A \cup \widetilde{A}$, we mean a finite sequence $g_1 \cdots g_k$, where $g_1, \ldots, g_k \in A \cup \widetilde{A}$. Recall that $\widetilde{A}$ is the set of inverses of all elements of A. We will refer to the elements of $A \cup \widetilde{A}$ as *letters*. We say that w is *reduced* if $g_i g_{i+1} \neq 1$ for any $1 \leq i < k$. There is a natural bijection between the set of reduced words and the group $\langle A \rangle$ generated by $A \subset \Gamma$. For the sake of clarity we write $w_1.w_2$ for concatenation of the sequences w_1 and w_2 and $w_1 w_2$ for the product in Γ, i.e. for concatenation followed by all possible reductions. Denote by B_l the set of reduced words of length l. Note that $|B_l| = 2m(2m-1)^{l-1}$ for $l \geq 1$.

Lemma 10. *Let notation be as above, and suppose that $H < \langle A \rangle$ is a subgroup such that for any $h \in \langle A \rangle$, there is a letter $g_0 \in A \cup \widetilde{A}$ such that $w \notin hHh^{-1}$ whenever w is a reduced word starting with g_0. Then*

$$|B_l \cap H| \leq (2m-1)^{l/2+1}(2m-2)^{l/2-1}.$$

We remark that the condition for $h = 1$ can be interpreted as follows. We can remove one edge incident to 1 from the Schreier graph of $H \backslash G$ such that we get two connected components and one of these is a tree.

Proof. Let w_0 be the longest word (possibly the empty word 1) such that w_0 is a prefix of all non-unit elements of H. Let w_1 be a reduced word of length at most $\lceil l/2 \rceil - 1$. We want to bound the number of letters $g' \in A \cup \widetilde{A}$ that can be the next letter in a reduced word of length l which belongs to H. We will show that if $|w_1| > |w_0|$ then there are at most $2l-2$ such letters. If $|w_1| = |w_0|$, we will see that there are at most $2l-1$ choices for g', this being trivial if $w_0 \neq 1$. If $|w_1| < |w_0|$ then we always have exactly one choice. Thus if we pick the letters of $w \in S_l \cap H$ one by one, then at the first $\lceil l/2 \rceil$ steps we have at most $2l-2$ choices with possibly one exception, when we might have $2l-1$; this gives the claim.

Now assume that $|w_0| < |w_1| \leq \lceil l/2 \rceil - 1$, but if $w_0 = 1$, we allow $w_1 = 1$. Using the assumption for $h = w_1^{-1}$, we get a letter g_0 such that if $g_0.w_2$ is a reduced word (i.e. the first letter of w_2 is not g_0^{-1}), then $g_0.w_2 \notin w_1^{-1}Hw_1$. We show that the last letter of w_1 is not g_0^{-1}. If w_1 is not the empty word, it is longer than w_0, hence there is a word $u \in H$ such that w_1 is not a prefix of u. Now if g_0^{-1} were the last letter of w_1, we would have $w_1^{-1}uw_1 \in w_1^{-1}Hw_1$ which begins with g_0, a contradiction.

Obviously we cannot continue w_1 with the inverse of its last letter to get a reduced word. We show that we cannot continue it with g_0 either to get one in $B_l \cap H$. Assume to the contrary that for some w_2, $w_1.g_0.w_2$ is a reduced word in $B_l \cap H$. Then $g_0w_2w_1 \in w_1^{-1}Hw_1$ and the length of w_1 is less than the length of w_2, hence $g_0w_2w_1$ starts with g_0, a contradiction. □

Let V be a vector space over $\mathbb{C}$, and denote by $\mathbb{P}(V)$ the corresponding projective space. For a vector $v \in V$ (resp. a subspace $W \subset V$) denote by $\bar{v}$ (resp. $\overline{W}$) its projection to $\mathbb{P}(V)$. Any invertible linear transformation T of V acts naturally on $\mathbb{P}(V)$; this action will be denoted by the same letter. We say that T is *proximal* if V is spanned by an eigenvector z_T and an invariant subspace V_T of T, and the eigenvalue corresponding to z_T is strictly larger than any other eigenvalue of T. In short, T is proximal if it has a unique simple eigenvalue of maximal modulus. It is clear that whenever z_T and V_T exist, V_T is unique and z_T is unique up to a constant multiple. Define the distance on $\mathbb{P}(V)$ by

$$d(\bar{x}, \bar{y}) = \frac{\|x \wedge y\|}{\|x\| \, \|y\|},$$

where $\|\cdot\|$ is the norm coming from the standard Hermitian form. We recall from Tits [35] a simple criterion for a transformation T to be proximal. Let $Q \subset \mathbb{P}(V)$ be compact and assume that $T(Q)$ is contained in the interior of Q. Assume further that $d(T(x), T(y)) < d(x, y)$ for $x, y \in Q$. Then T is proximal and $\bar{z}_T \in Q$ (see [35, Lemma 3.8(ii)]).

Let notation be as in Proposition 9. For $i \in \{1, 2\}$, denote by ρ_i the representation of $SL_d(\mathcal{O}_K)$ on $\mathfrak{sl}_d(\mathbb{C})$ defined by

$$\rho_i(h)v = \sigma_i(h)v\sigma_i(h^{-1}) \quad \text{for } v \in \mathfrak{sl}_d(\mathbb{C}) \text{ and } h \in SL_d(\mathcal{O}_K).$$

We study the action of $SL_d(\mathcal{O}_K)$ on the space $\mathbb{P}(\mathfrak{sl}_d(C)) \times \mathbb{P}(\mathfrak{sl}_d(C))$ via $\rho_1 \oplus \rho_2$. If T is an invertible linear transformation of $\mathfrak{sl}_d(\mathbb{C})$ and $h \in H_T$ is such that $\rho_1(h)$ and $\rho_2(h)$ are both proximal, then clearly

$$T(\bar{z}_{\rho_1(h)}) = \bar{z}_{\rho_2(h)}. \tag{7}$$

Our aim is to find a subset $A \subset \Gamma$ such that A freely generates a subgroup of $SL_d(\mathcal{O}_K)$ and for any linear transformation T of $\mathfrak{sl}_d(\mathbb{C})$, there is a letter $g_0 \in A \cup \widetilde{A}$ such that (7) fails when $h = w$ is a reduced word starting with g_0. Then Proposition 9 will follow easily from Lemma 10.

We say that $A \subset SL_d(\mathcal{O}_K)$ is *generic* if for any $g \in A \cup \widetilde{A}$, $\rho_1(g)$ and $\rho_2(g)$ are both proximal, and the following hold:

(i) for every $g_1, g_2 \in A \cup \widetilde{A}$ with $g_1 g_2 \neq 1$ and $i \in \{1,2\}$, we have $z_{\rho_i(g_1)} \notin V_{\rho_i(g_2)}$,
(ii) for any proper subspace V of $\mathfrak{sl}_d(\mathbb{C})$ of dimension k and $i \in \{1,2\}$, we have

$$|\{g \in A \cup \widetilde{A} \mid z_{\rho_i(g)} \in V\}| \leq k+1,$$

(iii) for any linear transformation T on $\mathfrak{sl}_d(\mathbb{C})$, we have

$$|\{g \in A \cup \widetilde{A} \mid T(\bar{z}_{\rho_1(g)}) = \bar{z}_{\rho_2(g)}\}| \leq d^2+1.$$

Note that $\mathfrak{sl}_d(\mathbb{C})$ is of dimension d^2-1. Actually the above definition would be more natural if we replaced the right hand sides of the inequalities in (ii) and (iii) by k and d^2 respectively, however doing so would make the next proof slightly more complicated. We prove the existence of generic sets in

Lemma 11. *Assume that $\widehat{\sigma}(\Gamma)$ is Zariski dense in $SL_d(\mathbb{C})^r$. Then for m positive integer, there is a generic set $A_m \subset \Gamma$ of cardinality m.*

Proof. Goldsheid and Margulis [18] prove (see also Sections 3.12–3.14 in Abels, Margulis and Soifert [1]) that if a real algebraic subgroup of $GL_d(\mathbb{R})$ is strongly irreducible (i.e. does not leave a finite union of proper subspaces invariant) and contains a proximal element, then a Zariski dense subgroup of it also contains a proximal element. If σ_1 is a real embedding, then it follows from the Zariski density of $\sigma_1(\Gamma)$ in $SL_d(\mathbb{R})$ that there is an element $g_0 \in \Gamma$ such that $\sigma_1(g_0)$ is proximal. If σ_1 is complex, then let $\bar{\sigma}_1$ denote its complex conjugate. Since $(\sigma_1 \oplus \bar{\sigma}_1)(\Gamma)$ is Zariski dense in $SL_d(\mathbb{C}) \times SL_d(\mathbb{C})$, we find that $\sigma_1(\Gamma)$ is Zariski dense in $SL_d(\mathbb{C})$ over the reals as well, i.e. considered as a subgroup of $SL_{2d}(\mathbb{R})$. Consider $\mathbb{C}^d$ as a real vector space, and take the wedge product $\mathbb{C}^d \wedge \mathbb{C}^d$. Denote by U the subspace spanned by the images of complex lines in $\mathbb{C}^d$; this is also the subspace fixed by the linear transformation induced by multiplication by i on $\mathbb{C}^d$. It is clear that $SL_d(\mathbb{C})$ (as a real group) acts on U strongly irreducibly and proximally in the natural way, hence there is an element $g_0 \in \Gamma$ such that $\sigma_1(g_0)$ is proximal on U. This implies in turn that $\sigma_1(g_0)$ is proximal on $\mathbb{C}^d$ now considered as a complex vector space.

So far we saw that we can find proximal elements in $\sigma_1(\Gamma)$ (σ_1 real or complex) and this follows for σ_2 in a similar fashion. Now we show that there is an element which is proximal for both σ_1 and σ_2 and this is also true for its inverse. Denote by σ_i' (for $i \in \{1,2\}$) the representation of Γ which assigns the transpose inverse of the matrix assigned by σ_i. Applying [1, Lemma 5.15] for the representation $\sigma_1 \oplus \sigma_1' \oplus \sigma_2 \oplus \sigma_2'$, we get an element $g_0 \in \Gamma$ such that $\sigma_1(g_0)$, $\sigma_1(g_0^{-1})$, $\sigma_2(g_0)$ and $\sigma_2(g_0^{-1})$ are proximal simultaneously. This implies in turn that $\rho_1(g_0)$, $\rho_1(g_0^{-1})$, $\rho_2(g_0)$ and $\rho_2(g_0^{-1})$ are also proximal.

We can set $A_1 = \{g_0\}$ and get the claim for $m = 1$. Proceeding by induction, assume that we can construct A_m for some $m \geq 1$. We try to find an element $h \in \Gamma$ such that

$A_{m+1} := A_m \cup \{hg_0h^{-1}\}$ is generic. Clearly $\bar{z}_{\rho_1(hg_0h^{-1})} = \rho_1(h)\bar{z}_{\rho_1(g_0)}$. One condition h needs to satisfy is that neither $\rho_1(h)z_{\rho_1(g_0)}$ nor $\rho_1(h)z_{\rho_1(g_0^{-1})}$ should belong to those proper subspaces V of $\mathfrak{sl}_d(\mathbb{C})$ for which

$$|\{g \in A_m \cup \widetilde{A}_m \mid z_{\rho_1(g)} \in V\}| \geq \dim V.$$

There are a finite number of such subspaces, hence this is a Zariski open condition on $\sigma_1(h)$. It can be seen in a similar fashion that A_{m+1} is generic if $(\sigma_1(h), \sigma_2(h))$ belongs to a certain Zariski dense open subset of $SL_d(\mathbb{C}) \times SL_d(\mathbb{C})$, and the lemma follows by induction. □

We remark that it is easy to see from the proof that A_m can be chosen in such a way that it is generic with respect to any pair of embeddings σ_1 and σ_2.

Lemma 12. *Let $A \subset \Gamma$ be a generic set of cardinality at least $(d^2+2)/2$. Then for each $g \in A \cup \widetilde{A}$ and $i \in \{1,2\}$, there is a neighborhood $U_g^{(i)} \subset \mathbb{P}(\mathfrak{sl}_d(\mathbb{C}))$ of $\bar{z}_{\rho_i(g)}$ with the following property. For any invertible linear transformation T on $\mathfrak{sl}_d(\mathbb{C})$ there is a $g \in A \cup \widetilde{A}$ such that $T(U_g^{(1)}) \cap U_g^{(2)} = \emptyset$.*

First we recall [13, Proposition 2.1]. Let $T_1, T_2, \ldots$ be a sequence of invertible linear transformations on $\mathfrak{sl}_d(\mathbb{C})$. There is a not necessarily invertible linear transformation $T \neq 0$ and a subsequence of $T_1, T_2, \ldots$ that considered as maps on $\mathbb{P}(\mathfrak{sl}_d(\mathbb{C}))$ converge uniformly to T on compact subsets of $\mathbb{P}(\mathfrak{sl}_d(\mathbb{C})) \setminus \overline{\mathrm{Ker}(T)}$.

Proof of Lemma 12. Assume to the contrary that the claim is false. Then there is a sequence $\{T_k\}$ of linear transformations such that for any choice of the neighborhoods $U_g^{(i)}$ ($i \in \{1,2\}$ and $g \in A \cup \widetilde{A}$), we have $T_k(U_g^{(1)}) \cap U_g^{(2)} \neq \emptyset$ for k large enough. By the aforementioned result, we may assume that $\{T_k\}$ converges uniformly to a linear transformation T on compact subsets of $\mathbb{P}(\mathfrak{sl}_d(\mathbb{C})) \setminus \overline{\mathrm{Ker}(T)}$. This implies that if $z_{\rho_1(g)} \notin \mathrm{Ker}(T)$, then $T(\bar{z}_{\rho_1(g)}) = \bar{z}_{\rho_2(g)}$. When T is invertible, this violates (iii) in the definition of generic sets. If T is not invertible, we get a contradiction with (ii) of that definition, either for $V = \mathrm{Ker}(T)$ or for $V = \mathrm{Im}(T)$, and the lemma follows. □

Lemma 13. *Let $A \subset \Gamma$ be generic, and for each $g \in A \cup \widetilde{A}$ and $i \in \{1,2\}$ let $U_g^{(i)} \subset \mathbb{P}(\mathfrak{sl}_d(\mathbb{C}))$ be a sufficiently small neighborhood of $\bar{z}_{\rho_i(g)}$. Then there is a positive integer M such that $\{g^M \mid g \in A\}$ freely generates a subgroup of Γ and if $h = g_1^M \cdots g_k^M$ is a reduced word, then $\rho_1(h)$ and $\rho_2(h)$ are proximal with $\bar{z}_{\rho_i(h)} \in U_{g_1}^{(i)}$.*

Proof. To simplify the notation we omit those subscripts and superscripts that indicate which of the representations ρ_1 or ρ_2 the object in question is related to. If U_g are sufficiently small, then there are compact sets $Q_g \subset \mathbb{P}(\mathfrak{sl}_d(\mathbb{C})) \setminus \overline{V_{\rho(g)}}$ for $g \in A \cup \widetilde{A}$ and an integer M such that

$$d(\rho(g^M)\bar{x}, \rho(g^M)\bar{y}) < d(\bar{x}, \bar{y}) \quad \text{for } x, y \in Q_g$$

and

$$U_{g'} \subset Q_g \quad \text{if } gg' \neq 1.$$

Here we have used property (i) of generic sets. If M is large enough we clearly also have $\rho(g^M)Q_g \subset U_g$. By induction, we see that if $h = g_1^M \cdots g_k^M$ is a reduced word then $\rho(h)Q_{g_k} \subset U_{g_1}$, and $d(\rho(h)\bar{x}, \rho(h)\bar{y}) < d(\bar{x}, \bar{y})$ for $\bar{x}, \bar{y} \in Q_{g_k}$. If $g_1 g_k \neq 1$, then $U_{g_1} \subset Q_{g_k}$ and the claim follows for h by the aforementioned lemma of Tits [35, Lemma 3.8(ii)]. If $g_1 g_k = 1$, then write $h = g_1^M h' g_1^{-M}$. If h' is proximal with $\bar{z}_{\rho(h')} \in U_{g_2}$, then h is also proximal with $\bar{z}_{\rho(h)} = \rho(g_1)\bar{z}_{\rho(h')}$, and the claim follows by induction. Now $\{g^M \mid g \in A\}$ generates freely a group since the identity is not proximal. □

Proof of Proposition 9. Let A be a generic set of cardinality $m \geq (d^2+2)/2$, and set $S' = \{g^M \mid g \in A \cup \widetilde{A}\}$, where M is as in Lemma 13. For $g \in S'$ and $i \in \{1, 2\}$ let $U_g^{(i)}$ be a neighborhood of $\bar{z}_{\rho_i(g)}$ which is sufficiently small for Lemmata 12 and 13. Then there is an element $g_0 \in S'$ such that $T(U_{g_0}^{(1)}) \cap U_{g_0}^{(2)} = \emptyset$. For $h \in H_T$ we clearly have $T\bar{z}_{\rho_1(h)} = \bar{z}_{\rho_2(h)}$, so if h is a reduced word of the form $g_1 \cdots g_k$ with $g_i \in S'$, then $g_1 \neq g_0$ by Lemma 13. If $h \in SL_d(\mathcal{O}_K)$, a similar result holds for $hH_Th^{-1} = H_{\rho_2(h)T\rho_1(h^{-1})}$. Therefore by Lemma 10, we have

$$|B_l \cap H_T| \leq (2m-1)^{l/2+1}(2m-2)^{l/2-1},$$

where B_l is the set of reduced words of length l over the alphabet S'.

Set $P_k(l) = \chi_{S'}^{(2k)}(w)$, where $w \in B_l$. Since $|B_l| = 2m(2m-1)^{l-1}$ for $l \geq 1$,

$$1 = P_k(0) + \sum_{l \geq 1} 2m(2m-1)^{l-1} P_k(l). \tag{8}$$

By a result of Kesten [24, Theorem 3.], we have

$$\limsup_{k \to \infty} (P_k(0))^{1/k} = (2m-1)/m^2.$$

From general properties of Markov chains (see [36, Lemma 1.9]) it follows that

$$P_k(0) \leq \left(\frac{2m-1}{m^2}\right)^k.$$

Since $\chi_{S'}^{(2k)}$ is symmetric, we have $P_k(0) = \sum_g \chi_{S'}^{(k)}(g)^2$, hence $P_k(l) \leq P_k(0)$ for all l by the Cauchy–Schwarz inequality. Now we can write

$$\begin{aligned}
\chi_{S'}^{(2k)}(H_T) &= \sum_l |B_l \cap H_T| P_k(l) \leq \sum_l (2m-1)^{l/2+1}(2m-2)^{l/2-1} P_k(l) \\
&\leq \sum_{l \leq k/10} (2m-1)^{l/2+1}(2m-2)^{l/2-1} \left(\frac{2m-1}{m^2}\right)^k \\
&\quad + \left(\frac{2m-1}{2m}\right)^{k/20} \sum_{l \geq k/10} 2m(2m-1)^{l-1} P_k(l) \\
&< \left(\frac{2m-1}{2m}\right)^{k/2} + \left(\frac{2m-1}{2m}\right)^{k/20},
\end{aligned}$$

which was to be proven. The inequality between the third and fourth lines follows from (8). □

Proof of Theorem 2. Let S' be as in Proposition 9 and let C and δ be as in Corollary 7. As we remarked after Lemma 11, we can choose S' in such a way that it works for any pair of embeddings σ_1 and σ_2. There is a constant c depending on the set S' such that $\log \|\widehat{\sigma}(g)\| \leq cl$ for $g \in \prod_l S'$. Then for $l = \delta \log[SL_d(\mathcal{O}_k/(q)) : H^\sharp]/c$, we have

$$\pi_q[\chi_{S'}^{(l)}](H^\sharp) = \chi_{S'}^{(l)}(\mathcal{L}_\delta(H^\sharp)).$$

Combining Corollary 7 with either Proposition 8 or Proposition 9 we get

$$\chi_{S'}^{(l)}(\mathcal{L}_\delta(H^\sharp)) \ll [SL_d(\mathcal{O}_k/(q)) : H^\sharp]^{-\delta c'}$$

with some $c' > 0$. If l is even, then by the symmetry of S',

$$(\pi_q[\chi_{S'}^{(l/2)}](gH^\sharp))^2 \leq \pi_q[\chi_{S'}^{(l)}](H^\sharp)$$

for any coset $gH^\sharp$, and from $[H : H^\sharp] < C^n$ we then have

$$\pi_q[\chi_{S'}^{(l/2)}](H) \leq C^n(\pi_q[\chi_{S'}^{(l)}](H^\sharp))^{1/2}.$$

If $l_1 \leq l_2$, then clearly

$$\pi_q[\chi_{S'}^{(l_2)}](H) \leq \max_g \pi_q[\chi_{S'}^{(l_1)}](gH).$$

Now it is straightforward to get the theorem by putting together the above inequalities. □

3. A product theorem

Recall that $H_1 \lesssim_L H_2$ is shorthand for $[H_1 : H_1 \cap H_2] \leq L$. We denote by $Z(G)$ the center of the group G, by $\mathcal{C}(g)$ the centralizer of the element $g \in G$ and by $\mathcal{N}_G(H)$ the normalizer of the subgroup $H < G$. In this section, K is not a number field; it usually stands for a large positive real number. We begin by listing the assumptions already mentioned in Theorem 3. When we say that something depends on the constants appearing in the assumptions (A1)–(A5) we mean L and the function $\delta(\varepsilon)$ for which (A4) holds.

- (A0) $G = G_1 \times \cdots \times G_n$ is a direct product, and the collection of the factors satisfy (A1)–(A5) for some sufficiently large constant L.
- (A1) There are at most L isomorphic copies of the same group in the collection.
- (A2) Each G_i is quasi-simple and $|Z(G_i)| < L$.
- (A3) Any nontrivial representation of G_i is of dimension at least $|G_i|^{1/L}$.

(A4) For any $\varepsilon > 0$, there is a $\delta > 0$ such that the following holds. If μ and ν are probability measures on G_i with

$$\|\mu\|_2 > |G_i|^{-1/2+\varepsilon} \quad \text{and} \quad \mu(gH) < |G_i|^{-\varepsilon}$$

for any $g \in G_i$ and for any proper $H < G_i$, then

$$\|\mu * \nu\|_2 \ll \|\mu\|_2^{1/2+\delta} \|\nu\|_2^{1/2}. \tag{9}$$

(A5) For some $m < L$, there are classes $\mathcal{H}_0, \mathcal{H}_1, \ldots, \mathcal{H}_m$ of subgroups of G_i having the following properties.

(i) $\mathcal{H}_0 = \{Z(G)\}$.
(ii) Each $\mathcal{H}_j$ is a set of proper subgroups of G_i and this collection is closed under conjugation by elements of G_i.
(iii) For each proper $H < G_i$ there is an $H^\sharp \in \mathcal{H}_j$ for some j with $H \lesssim_L H^\sharp$.
(iv) For every pair of subgroups $H_1, H_2 \in \mathcal{H}_j$, $H_1 \neq H_2$, there is some $j' < j$ and $H^\sharp \in \mathcal{H}_{j'}$ for which $H_1 \cap H_2 \lesssim_L H^\sharp$.

We remark that by considering the induced representation, (A3) implies that for any proper subgroup $H < G_i$ we have

$$[G_i : H] > |G_i|^{1/L}. \tag{10}$$

One may think about (A5) that there is a notion of dimension for the subgroups of G_i.

In the next section we show that Theorem 3 is a simple corollary of the following seemingly weaker result.

Proposition 14. *Let G be a group satisfying* (A0)–(A5). *For any $\varepsilon > 0$ there is a $\delta > 0$ depending only on ε and on the constants in the assumptions such that the following holds. If $S \subset G$ is symmetric such that*

$$|S| < |G|^{1-\varepsilon} \quad \textit{and} \quad \chi_S(gH) < [G : H]^{-\varepsilon}|G|^{\delta}$$

for any $g \in G$ and any proper $H < G$, then $|\prod_3 S| \gg |S|^{1+\delta}$.

3.1. Proof of Theorem 3 using Proposition 14

We make use of the following result which appeared first implicitly in the proof of Proposition 2 in Bourgain and Gamburd [5].

Lemma 15 (Bourgain and Gamburd). *Let μ and ν be two probability measures on an arbitrary group G and let $K > 2$ be a number. If*

$$\|\mu * \nu\|_2 > \frac{\|\mu\|_2^{1/2} \|\nu\|_2^{1/2}}{K}$$

then there is a symmetric set $S \subset G$ with

$$\frac{1}{K^R \|\mu\|_2^2} \ll |S| \ll \frac{K^R}{\|\mu\|_2^2}, \qquad |\textstyle\prod_3 S| \ll K^R |S|, \qquad \min_{g \in S} (\widetilde{\mu} * \mu)(g) \gg \frac{1}{K^R |S|},$$

where R and the implied constants are absolute.

Proof. We include the proof only for the sake of completeness; the argument is essentially the same as in the proof of [5, Proposition 2].

First we note that by Young's inequality $\|\mu * \nu\|_2 \leq \|\mu\|_2$ and hence $\|\nu\|_2 < K^2\|\mu\|_2$ and similarly $\|\mu\|_2 < K^2\|\nu\|_2$. Let λ be a nonnegative measure with $\|\lambda\| \leq 1$ and $\|\lambda\|_2^2 < c$. Observe that if $\lambda(g) \geq K'c$ for some K' for every $g \in \operatorname{supp}\lambda$, then $\|\lambda\|_2^2 \geq K'c\|\lambda\|_1$, hence $\|\lambda\|_1 < 1/K'$. Similarly, if $\lambda(g) \leq c/K'$ for all g, then $\|\lambda\|_2^2 < c/K'$. Now define the sets

$$A_i = \{g \in G \mid 2^{i-1}\|\mu\|_2^2 < \mu(g) \leq 2^i\|\mu\|_2^2\},$$
$$B_i = \{g \in G \mid 2^{i-1}\|\nu\|_2^2 < \nu(g) \leq 2^i\|\nu\|_2^2\}$$

for $|i| < 10\log K$. By Young's inequality,

$$\|\mu * \nu\|_2 \leq \sum_{|i|,|j| \leq 10\log K} 2^{i+j}\|\mu\|_2^2\|\nu\|_2^2|A_i|\,|B_j|\,\|\chi_{A_i} * \chi_{B_j}\|_2 + K^{-5}(\|\mu\|_2 + \|\nu\|_2),$$

hence there must be a pair of indices i, j such that

$$2^{i+j}\|\mu\|_2^2\|\nu\|_2^2\,|A_i|\,|B_j|\,\|\chi_{A_i} * \chi_{B_j}\|_2 \gg \frac{\|\mu\|_2^{1/2}\|\nu\|_2^{1/2}}{K\log^2 K}. \tag{11}$$

By construction, for $g \in A_i$ we have

$$2^i\|\mu\|_2^2 \gg \mu(g) \gg 2^i\|\mu\|_2^2,$$

and by (11) and Young's inequality, $1 \geq \mu(A_i) \gg 1/K^R$. Here and throughout, R denotes an absolute constant which need not be the same at different occurrences. These together give

$$\frac{K^R}{\|\mu\|_2^2} \gg |A_i| \gg \frac{1}{K^R\|\mu\|_2^2}.$$

We may get the analogous inequalities

$$\frac{K^R}{\|\mu\|_2^2} \gg |B_j| \gg \frac{1}{K^R\|\mu\|_2^2}$$

in a similar way and using the relations between $\|\mu\|_2$ and $\|\nu\|_2$. Applying our inequalities to (11), we get

$$\|\chi_{A_i} * \chi_{B_j}\|_2^2 \gg \frac{1}{K^R|A_i|^{1/2}|B_j|^{1/2}}.$$

We invoke the noncommutative version of the Balog–Szemerédi–Gowers theorem proven by Tao [32, Theorem 5.2] (note that we use a different normalization). This gives subsets $A \subset A_i$ and $B \subset B_i$ with $|A| \gg |A_i|/K^R$ and $|A.B| \ll K^R|A|^{1/2}|B|^{1/2}$. Ruzsa's triangle inequality [32, Lemma 3.2] for the sets A and $\widetilde{B}$ gives $|A.\widetilde{A}| \ll K^R|A|$. Using [32, Proposition 4.5] with $n = 3$, we get a symmetric set S with $|S| > |A|/K^R$ and

$$|\textstyle\prod_3 S| \ll K^R|A| \ll K^{R'}|S|.$$

In the proof of Proposition 4.5 of [32] the set S is defined by

$$\{g \in G \mid |A \cap (A.\{g\})| > |A|/C\}$$

with $C = 2|A.\widetilde{A}|/|A|$. For $g \in S$, we have

$$(\widetilde{\mu} * \mu)(g) \geq 2^{2i-2}\|\mu\|_2^4|A \cap (A.\{g\})| \gg \frac{1}{K^R|S|}.$$

The expression in the middle is bounded below by $\|\mu\|_2^2/K^R$ also, which gives the required upper bound for $|S|$, since $\|\widetilde{\mu} * \mu\|_1 = 1$. □

Proof of Theorem 3. Assume that the conclusion of the theorem fails, i.e. there is an ε such that for any δ there are probability measures μ and ν with

$$\|\mu\|_2 > |G|^{-1/2+\varepsilon} \quad \text{and} \quad \mu(gH) < [G : H]^{-\varepsilon}$$

for any $g \in G$ and for any proper $H < G$, and yet

$$\|\mu * \nu\|_2 \geq \|\mu\|_2^{1/2+\delta}\|\nu\|_2^{1/2}.$$

Take $K = \|\mu\|_2^{-\delta}$ in Lemma 15. Note that by the third property of the set S we have

$$\chi_S(gH) \ll K^R\widetilde{\mu} * \mu(gH) \leq K^R \max_{h \in G} \mu(hH) \ll |G|^{R\delta/2}[G : H]^{-\varepsilon}.$$

Now $|\prod_3 S| \ll K^R|S|$ contradicts Proposition 14 if δ is small enough. □

3.2. *Proof of Proposition 14*

Throughout Sections 3.2–3.4, we assume that $G = G_1 \times \cdots \times G_n$ satisfies (A0)–(A5) with some L. Moreover ε and S are as in Proposition 14, and we fix a sufficiently small δ. By sufficiently small, we mean that we are free to use inequalities $\delta < \delta'$, where δ' is any function of ε and the constants in (A1)–(A5). We use $c, \delta', \delta'', Q, Q'$, etc. to denote positive constants that may depend only on ε and the constants in (A1)–(A5). These need not be the same at different occurrences. We will also use inequalities of the form

$$Q \log |G_i| < |G_i|^{\delta\delta'}. \tag{12}$$

Let N be the product of those factors G_i for which such an inequality fails. Since the same group appears at most L times among the G_i, the size of N is bounded. Replace G by G/N. For any $H < G/N$, we have $[G/N : H] = [G : HN]$ and if $\bar{S}$ denotes the projection of S in G/N, then we have $|\prod_3 S| \geq |\prod_3 \bar{S}|$ and $|S| \leq |\bar{S}|\,|N|$. Hence the proposition for the group G/N implies the proposition for G with a larger implied constant. Thus we can use (12) without loss of generality.

In a similar fashion we may replace each G_i by $G_i/Z(G_i)$, hence from now on, we assume that all the G_i are simple. This may introduce a factor of size at most L^n which is $\ll |G|^\delta$ for any $\delta > 0$.

We follow the argument of Bourgain, Gamburd and Sarnak [9, Section 5]. First we introduce some notation. Denote by π_i for $1 \leq i \leq n$ the projection from G to G_i. Set $G_{\leq i} = \times_{j \leq i} G_i$ and denote by $\pi_{\leq i}$ the projection from G to $G_{\leq i}$. To the set S, we associate a tree of $n+1$ levels. Level 0 consists of a single vertex, while for $i > 0$ the vertices of level i are the elements of the set $\pi_{\leq i}(S)$, and a vertex g on level $i-1$ is connected to those vertices on level i which are of the form (g,h) with some $h \in G_i$. By removing some vertices, we can get a regular tree, that is, a tree which has vertices of equal degree on each level. More precisely, using [9, Lemma 5.2] we obtain a subset $A \subset S$ and a sequence $\{D_i\}_{1\leq i\leq n}$ of positive integers with $D_i \geq |G_i|^\delta$ or $D_i = 1$ such that for any $g \in \pi_{\leq i-1}A$, we have

$$|\{h \in G_i \mid (g,h) \in \pi_{\leq i}(A)\}| = D_i,$$

and

$$|A| > \left[\prod_{i=1}^{n}(|G_i|^\delta \log |G_i|)\right]^{-1}|S| > |G|^{-2\delta}|S|. \tag{13}$$

The second inequality in (13) is of type (12).

We briefly outline the proof. Consider the set $\prod_k A$ for some integer k and the tree associated to it in the way described above. If $g \in \pi_{\leq i-1}(\prod_k A)$ is a vertex on level $i-1$ and $g = g_1 \cdots g_k$ with $g_l \in \pi_{\leq i-1}(A)$, then (g,h) is connected to g for every h in the product-set

$$\{h_1 \mid (g_1,h_1) \in \pi_{\leq i}(A)\}\cdots\{h_k \mid (g_k,h_k) \in \pi_{\leq i}(A)\}.$$

Let I_s be the set of indices $1 \leq i \leq n$ for which $D_i < |G_i|^{1-1/3L}$ (i.e. indices corresponding to small degrees); for such an index, there is hope that we can apply (A4) for G_i and deduce that the above product-set is of size $D_i^{1+\delta'}$ for some $\delta' > 0$. We make this speculation precise in Section 3.3. Set $I_l = \{1,\dots,n\}\setminus I_s$ (indices corresponding to large degrees), $G_s = \times_{i\in I_s} G_i$ and $G_l = \times_{i\in I_l} G_i$, and denote by π_s and π_l the projections from $G = G_s \times G_l$ to G_s and G_l respectively. We infer from a result of Gowers [19] that $\pi_l(S.S.S) = G_l$. In Subsection 3.4, we prove using a result of Farah [16] on approximate homomorphisms that $\pi_l^{-1}(1) \cap \prod_9 S$ contains an element g whose centralizer $\mathcal{C}(g)$ is of large index. Then S will contain elements from at least $[G:\mathcal{C}(g)]^\varepsilon|G|^{-\delta}$ cosets of $\mathcal{C}(g)$, hence there are many $h \in \prod_{11} S$ with $\pi_l(h) = 1$, and $\prod_{12} S$ is much larger than G_l.

Finally, we mention that there is a useful result of Helfgott [21, Lemma 2.2] that allows us to bound $|S.S.S|$ in terms of larger iterated product-sets. He proves that if S is a symmetric subset of an arbitrary group G and $k \geq 3$ is an integer, then

$$\frac{|\prod_k S|}{|S|} \leq \left(\frac{|S.S.S|}{|S|}\right)^{k-2}. \tag{14}$$

3.3. *The case of many small degrees*

In this section we prove

Proposition 16. *There are positive constants δ' and Q, depending only on ε and the constants in the assumptions, such that*

$$|\textstyle\prod_{2^{m+1}} S| > |S|\,|G|^{-Q\delta}\displaystyle\prod_{i\in I_s} D_i^{\delta'},$$

where m is as in (A5).

The biggest issue here is that beside its size, we have no information about a set of the form $\{b \mid (a,b) \in \pi_{\le i}(A)\}$. A large part of it might be contained in a coset of a proper subgroup and then (A4) does not apply with μ being the normalized counting measure on that set. To resolve this problem, we multiply sets of this form together with random elements of G_i. We need to construct a probability distribution supported on S whose projection to most factors G_i is well-behaved in the following sense.

Lemma 17. *There is a subset $B \subset S$, and there is a partition of the indices $1,\dots,n$ into two parts J_g and J_b, such that*

$$\prod_{i \in J_b} |G_i| \le |G|^{\delta/\delta'}, \tag{15}$$

and for any $i \in J_g$ and for any proper coset $gH \subset G_i$, we have

$$\chi_B(\{x \in G \mid \pi_i(x) \in gH\}) \le |G_i|^{-\delta'}, \tag{16}$$

where $\delta' > 0$ is a constant depending on ε and on L.

Proof. We obtain the set B by the following algorithm. First set $B = S$ and $J_g = \{1,\dots,n\}$. Then iterate the following step as long as possible. If there is an index $i \in J_g$ and a coset $gH \subset G_i$ such that (16) fails, then replace B by

$$\{x \in B \mid \pi_i(x) \in gH\}$$

and put i into J_b. It is clear that (16) holds when this process terminates. As for (15), note that

$$\chi_S(B) \ge \prod_{i \in J_b} |G_i|^{-\delta'}$$

and B is contained in a coset of a subgroup of index at least $\prod_{i \in J_b} |G_i|^{1/L}$ by (10). Together with the assumption of Proposition 14 on S, this implies

$$\prod_{i \in J_b} |G_i|^{-\delta'} < \Big(\prod_{i \in J_b} |G_i|^{1/L}\Big)^{-\varepsilon} |G|^{\delta},$$

and (15) follows easily if we set $\delta' = \varepsilon/2L$. □

Now assume that $i \in J_g$. Then, starting from arbitrary sets $A_1,\dots,A_{2^m} \subset G_i$ of the same size $|G_i|^{\delta} < D < |G_i|^{1-1/3L}$, we construct a measure λ_m for which (A4) is applicable.

Choose the elements x_j for $1 \le j \le 2^m - 1$ independently at random according to the distribution χ_B. Set $y_j = \pi_i(x_j)$. For $0 \le k \le m$ define

$$\lambda_k = \chi_{A_1} * 1_{y_1} * \chi_{A_2} * 1_{y_2} * \cdots * 1_{y_{2^k-1}} * \chi_{A_{2^k}},$$

where 1_y denotes the unit mass measure at y.

Lemma 18. *If $i \in J_g$, then there is a constant δ' depending only on ε and L such that the probability of the event that*

$$\lambda_k(gH) < D^{-\delta'/10^k} \tag{17}$$

holds for any proper coset $gH < G_i$ with $H \in \mathcal{H}_l$ for some $l \leq k$ is at least

$$1 - (2^k - 1)|G_i|^{-\delta'}. \tag{18}$$

Proof. Let δ' be twice the δ' of the previous lemma. For $k = 0$, the claim follows from $L/D < D^{-\delta'}$, which is an inequality of form (12). We assume that $k > 0$ and that the claim holds for $k - 1$. Set

$$\eta_{k-1} = \chi_{A_{2^{k-1}+1}} * 1_{y_{2^{k-1}+1}} * \chi_{A_{2^{k-1}+2}} * 1_{y_{2^{k-1}+2}} * \cdots * 1_{y_{2^k-1}} * \chi_{A_{2^k}}$$

and assume that $y_1, \ldots, y_{2^{k-1}-1}$ and $y_{2^{k-1}+1}, \ldots, y_{2^k-1}$ are chosen in such a way that λ_{k-1} and η_{k-1} satisfy

$$\lambda_{k-1}(gH) < D^{-\delta'/10^{k-1}} \quad \text{and} \quad \eta_{k-1}(gH) < D^{-\delta'/10^{k-1}}$$

for subgroups $H \in \mathcal{H}_{k-1}$. By the induction hypothesis, the probability of such a choice is at least $1 - (2^k - 2)|G_i|^{-\delta'}$. Now assume that $\lambda_k = \lambda_{k-1} * 1_{y_{2^{k-1}}} * \eta_{k-1}$ violates (17) for some $g \in G_i$ and $H \in \mathcal{H}_k$. To shorten the notation write $y = y_{2^{k-1}}$. We prove that y is in a set of $\pi_i(\chi_B)$ measure at most $|G_i|^{-\delta'}$, and this set will depend only on λ_{k-1} and η_{k-1}, in particular it will be independent of the choice of H and g. Let $\{h_j\}$ be a left transversal for H (i.e. a system of representatives for left H-cosets). Then it is easy to see that $\{gh_j^{-1}\}$ is a right transversal for gHg^{-1}, hence

$$\lambda_k(gH) = \sum_j \lambda_{k-1}(gHg^{-1}gh_j^{-1})\eta_{k-1}(y^{-1}h_jH).$$

We claim that for some index j, we have

$$\lambda_{k-1}(B_j) \geq D^{-\delta'/10^k}/2 \quad \text{and} \quad \eta_{k-1}(C_j) \geq D^{-\delta'/10^k}/2, \tag{19}$$

where $B_j = gHh_j^{-1}$ and $C_j = y^{-1}h_jH$. Assume to the contrary that this fails. Then

$$\begin{aligned}\sum_j \lambda_{k-1}(B_j)\eta_{k-1}(C_j) &= \sum_{j:\,\lambda_{k-1}(B_j)<D^{-\delta'/10^k}/2} \lambda_{k-1}(B_j)\eta_{k-1}(C_j) \\ &\quad + \sum_{j:\,\eta_{k-1}(C_j)<D^{-\delta'/10^k}/2} \lambda_{k-1}(B_j)\eta_{k-1}(C_j) \\ &< D^{-\delta'/10^k},\end{aligned}$$

a contradiction.

Let j be such that (19) holds. Define $H_1 = h_j H h_j^{-1}$ and $H_2 = y^{-1} H_1 y$. Notice that $\widetilde{B}_j.B_j \subset H_1$ and $C_j.\widetilde{C}_j \subset H_2$. This shows that there are subgroups $H_1, H_2 \in \mathcal{H}_k$ such that

$$(\widetilde{\lambda}_{k-1} * \lambda_{k-1})(H_1) \geq D^{-2\delta'/10^k}/4 \quad \text{and} \quad (\eta_{k-1} * \widetilde{\eta}_{k-1})(H_2) \geq D^{-2\delta'/10^k}/4 \tag{20}$$

and $H_1 = y H_2 y^{-1}$. For fixed H_1 and H_2, this restricts y to a single $\mathcal{N}(H_2)$-coset. By Lemma 17, this is a set of χ_B measure at most $|G_i|^{\delta'/2}$. The final step is to show that the number of possible pairs H_1, H_2 such that (20) holds is at most $|G_i|^{\delta'/2}$.

Suppose that we have M distinct subgroups $H_1 \in \mathcal{H}_k$ such that

$$\widetilde{\lambda}_{k-1} * \lambda_{k-1}(H_1) \geq D^{-2\delta'/10^k}/4.$$

If H_1 and H_1' are two such subgroups, then $H_1 \cap H_1' \lesssim_L H^\sharp$ for some $H^\sharp \in \mathcal{H}_{k-1}$. By the induction hypothesis, we have $\widetilde{\lambda}_{k-1} * \lambda_{k-1}(H^\sharp) \leq D^{-\delta'/10^{k-1}}$, hence $\widetilde{\lambda}_{k-1} * \lambda_{k-1}(H_1 \cap H_2) \leq L D^{-\delta'/10^{k-1}}$. By the inclusion-exclusion principle, we have

$$M D^{-2\delta'/10^k}/4 - M^2 L D^{-\delta'/10^{k-1}} \leq 1.$$

This is violated if $M = D^{\delta'/4\cdot 10^{k-1}}$, in fact we need $D^{\delta'/2\cdot 10^k} > 4(1+L)$, which is an inequality of form (12). Thus $M < D^{\delta'/4\cdot 10^{k-1}}$, and as the case of H_2 is similar, the proof is complete. $\square$

Using property (A4), we get the following simple

Corollary 19. *Assume that $|G_i|^\delta < D < |G_i|^{1-1/3L}$, and let $A' \subset G_i$ be any set of cardinality D. There is a positive number δ' depending only on ε and the constants in* (A1)–(A5) *such that for the above defined λ_m, we have*

$$\|\lambda_m * \chi_{A'}\|_2 \ll D^{-1/2-\delta'}$$

with probability at least $1/2$.

Proof. By Lemma 18 (and using (12)), with probability at least $1/2$ we have $\lambda_m(gH) < L D^{-\delta''}$ with some $\delta'' > 0$ for every proper coset gH. By (12), we have $L < D^{-\delta''/2}$. If say $\|\lambda_m\|_2 > |G_i|^{-1/2+1/12L}$, then we get

$$\|\lambda_m * \chi_{A'}\|_2 \leq \|\lambda_m\|_2^{1/2+\delta'} \|\chi_A'\|_2^{1/2}$$

by (A4) with $\mu = \lambda_m$ and $\nu = \chi_{A'}$. Otherwise the claim is trivial by Young's inequality. $\square$

In what follows, we need some basic facts about entropy. Let μ be a probability measure on G, and let $\mathcal{A}$ be a partition of G. The *entropy* of $\mathcal{A}$ is defined by

$$H_\mu(\mathcal{A}) = \sum_{A \in \mathcal{A}} -\mu(A) \log(\mu(A)),$$

with the convention $0 \cdot \log 0 = 0$. We also use the notation H_μ for the entropy of the partition consisting of one-element sets. The inequalities

$$|\mathrm{supp}\,\mu| \geq e^{H_\mu} \geq 1/\|\mu\|_2^2$$

are well-known. If $B \subset G$, we write $\mu|_B(A) = \mu(A \cap B)/\mu(B)$, and if $\mathcal{B}$ is another partition, we define the *conditional entropy* by

$$H_\mu(\mathcal{A}\,|\,\mathcal{B}) = \sum_{B \in \mathcal{B}} H_{\mu|_B}(\mathcal{A})\mu(B).$$

It is easy to see that

$$H_\mu(\mathcal{A} \vee \mathcal{B}) = H_\mu(\mathcal{A}\,|\,\mathcal{B}) + H_\mu(\mathcal{B}),$$

where $\mathcal{A} \vee \mathcal{B}$ denotes the coarsest partition that is finer than both $\mathcal{A}$ and $\mathcal{B}$. On finite sets, partitions and σ-algebras are essentially the same, hence we make no distinction.

Proof of Proposition 16. First we introduce a couple of σ-algebras (partitions) on the set $A^{\times(2^m+1)}$, the 2^m+1-fold Cartesian product of A. Let $\mathcal{A}_i$ be the coarsest σ-algebra for which the projection map

$$\pi_{\leq i} : A^{\times(2^m+1)} \to G_{\leq i}^{\times(2^m+1)}$$

is measurable. Furthermore, let $\mathcal{B}$ be the coarsest σ-algebra for which the map

$$(a_1, \ldots, a_{2^m}, a_{2^m+1}) \mapsto a_1 x_1 a_2 x_2 \cdots x_{2^m-1} a_{2^m} a_{2^m+1}$$

is measurable, where the elements $x_1, \ldots, x_{2^m-1}$ are chosen independently at random according to the distribution χ_B, hence the partition $\mathcal{B}$ is random. Denote by μ the measure $\chi_A^{\otimes(2^m+1)}$ on $A^{\times(2^m+1)}$. It follows from the definition that the entropy of the measure

$$\chi_A * 1_{x_1} * \chi_A * 1_{x_2} * \cdots * 1_{x_{2^m-1}} * \chi_A * \chi_A$$

equals $H_\mu(\mathcal{B})$. We write for the expectation of $H_\mu(\mathcal{B})$:

$$\begin{aligned}
\mathbf{E}[H_\mu(\mathcal{B})] &\geq \sum_{i=1}^{n} \mathbf{E}[H_\mu(\mathcal{B} \wedge \mathcal{A}_i\,|\,\mathcal{A}_{i-1})] \\
&\geq \sum_{i \in I_s \cap J_g} \left(\frac{\log D_i}{2} + \frac{(1+2\delta')\log D_i}{2} - \log c \right) + \sum_{i \notin I_s \cap J_g} \log D_i \\
&\geq \log|A| + \sum_{i \in I_s \cap J_g} \delta' \log D_i - n \log c.
\end{aligned}$$

The second inequality follows from Corollary 19 and c is the implied constant there; and $\mathcal{A} \wedge \mathcal{B}$ denotes the finest partition that is coarser than both $\mathcal{A}$ and $\mathcal{B}$. This implies in turn that for some choices of $x_1, \ldots, x_{2^m-1}$, we have

$$|A.x_1.A.x_2 \ldots x_{2^m-1}.A.A| \geq c^{-n}|A|\,|G|^{-\delta} \prod_{i \in I_s} D_i^{\delta'},$$

where we have also used (15). Note that we can assume $c^n < |G|^\delta$ by (12), and recall that $|A| > |S|\,|G|^{-2\delta}$ by (13), hence Proposition 16 follows. $\square$

3.4. The case of many large degrees

This section is devoted to the proof of

Proposition 20. *There is a positive constant δ', depending only on ε and L, such that*

$$|\prod_{12} S| \geq |G|^{\delta'-\delta} \prod_{i \in I_l} D_i.$$

Recall that $G_s = \times_{i \in I_s} G_i$, $G_l = \times_{i \in I_l} G_i$ and π_s and π_l are the projections to these subgroups respectively.

By (A3), any nontrivial representation of G_i is of dimension at least $|G_i|^{1/L}$. It was pointed out by Nikolov and Pyber [26, Corollary 1] that a result of Gowers [19, Theorem 3.3] implies that if $A, B, C \subset G_i$ are subsets such that $|A|\,|B|\,|C| > |G_i|^{3-1/L}$ then $A.B.C = G_i$.

Let $i_1 \leq \cdots \leq i_{n'}$ be the indices in I_l and for $1 \leq n'' \leq n'$ set $G_{\{i_1,\dots,i_{n''}\}} = G_{i_1} \times \cdots \times G_{i_{n''}}$ and denote by $\pi_{\{i_1,\dots,i_{n''}\}}$ the projection to this subgroup. We prove by induction that

$$\pi_{\{i_1,\dots,i_{n''}\}}(A.A.A) = G_{\{i_1,\dots,i_{n''}\}}.$$

For $n'' = 1$, this follows directly from [26, Corollary 1] and from $\pi_{i_1}(A) \geq D_{i_1} \geq |G_i|^{1-1/3L}$. Now assume that the claim holds for some n'' and take an arbitrary element $g \in G_{\{i_1,\dots,i_{n''+1}\}}$. By the induction hypothesis there are elements $h_1, h_2, h_3 \in A$ such that

$$\pi_{\{i_1,\dots,i_{n''}\}}(h_1 h_2 h_3) = \pi_{\{i_1,\dots,i_{n''}\}}(g).$$

Define the sets

$$B_i = \{x \in A \mid \pi_{\{i_1,\dots,i_{n''}\}}(x) = \pi_{\{i_1,\dots,i_{n''}\}}(h_i)\}$$

and note that

$$\pi_{i_{n''+1}}(B_i) \supset \pi_{i_{n''+1}}(\{x \in A \mid \pi_{\leq i_{n''+1}-1}(x) = \pi_{\leq i_{n''+1}-1}(h_i)\}),$$

hence $|\pi_{i_{n''+1}}(B_i)| \geq D_{i_{n''+1}} \geq |G_{i_{n''+1}}|^{1-1/3L}$. Now an application of [26, Corollary 1] to the sets $\pi_{i_{n''+1}}(B_i)$ gives that $g \in \pi_{\{i_1,\dots,i_{n''+1}\}}(A.A.A)$, whence the claim follows.

Define the distance of two elements $g, h \in G_s$ by

$$d(g,h) = \sum_{i \in I_s:\, \pi_i(g) \neq \pi_i(h)} \log |G_i|.$$

Lemma 21. *If $|S.S.S| \leq |G|^{1-\varepsilon+\delta}$ then there is an element $g \in \prod_9 S$ such that*

$$\pi_l(g) = 1 \quad \textit{and} \quad d(\pi_s(g), 1) > \delta' \log |G|,$$

where $\delta' > 0$ is a constant depending only on ε and L.

Following Farah [16], we say that a map $\psi : G_l \to G_s$ is a δ'-*approximate homomorphism* if

$$d(\psi(g)\psi(h), \psi(gh)) \leq \delta' \quad \text{and} \quad d(\psi(g), (\psi(g^{-1}))^{-1}) \leq \delta'$$

for all $g, h \in G_l$. Note that in [16], such a ψ is called an *approximate homomorphism of type II*. We recall a result of Farah [16, Theorem 2.1] that will be crucial in the proof. Let $\psi : G_l \to G_s$ be a δ'-approximate homomorphism. Then there is a homomorphism $\varphi : G_l \to G_s$ such that

$$d(\psi(g), \varphi(g)) \leq 24\delta' \qquad \text{for all } g \in G_l.$$

Proof of Lemma 21. Assume to the contrary that for any $g \in \prod_9 S$ with $\pi_l(g) = 1$, we have $d(\pi_s(g), 1) \leq \delta' \log |G|$. For each $g \in G_l$, pick an element $h \in S.S.S$ with $\pi_l(h) = g$ and set $\psi(g) = \pi_s(h)$. This gives rise to a map $\psi : G_l \to G_s$, which of course depends on our choices for h. It follows in turn that for any $g \in G_l$ and $h \in S.S.S$ with $\pi_l(h) = g$, we have $d(\pi_s(h), \psi(g)) < \delta' \log |G|$ and that ψ is a $\delta' \log |G|$-approximate homomorphism. By [16, Theorem 2.1], there is a homomorphism φ with $d(\psi(g), \varphi(g)) \leq 24\delta' \log |G|$ for any $g \in G_l$. The elements $g \in G$ satisfying

$$\pi_s(g) = \varphi(\pi_l(g))$$

constitute a subgroup $H < G$ of index $|G_s|$, since the cosets of H are represented by the elements g with $\pi_l(g) = 1$. For $h_1 \in S.S.S$, the coset $h_1 H$ is represented by the element g_1 with $\pi_l(g_1) = 1$ and $\pi_s(g_1) = \pi_s(h_1)\varphi(\pi_l(h_1))^{-1}$. Since

$$\begin{aligned} d(\pi_s(h_1), \varphi(\pi_l(h_1))) &\leq d(\pi_s(h_1), \psi(\pi_l(h_1))) + d(\psi(\pi_l(h_1)), \varphi(\pi_l(h_1))) \\ &< 25\delta' \log |G|, \end{aligned}$$

there is an index set $I \subset I_s$ with $\prod_{i \in I} |G_i| < |G|^{25\delta'}$ such that $\pi_i(g_1) \neq 1$ exactly if $i \in I$. If I is given, there are at most $|G|^{25\delta'}$ choices for g_1. Thus $S.S.S$ is contained in $2^n |G|^{25\delta'} < |G|^{26\delta'}$ cosets of H. This is a contradiction if

$$|G_s|^{-\varepsilon} |G|^{26\delta' + \delta} < 1.$$

Since $|G_l| \leq |S.S.S| \leq |G|^{1-\varepsilon+\delta}$, we have $|G_s| \geq |G|^{\varepsilon - \delta}$. Now, if δ is small enough (e.g. $\delta < \varepsilon^2/10$) we can get the desired contradiction by an appropriate choice of δ'. □

Proof of Proposition 20. First we calculate the index of the centralizer $\mathcal{C}(g)$ of g, the element constructed in Lemma 21. An element h commutes with g if and only if $\pi_i(h) \in \mathcal{C}(\pi_i(g))$ for all indices i for which $\pi_i(g) \neq 1$. For such an i, $[G_i : \mathcal{C}(\pi_i(g))] > |G_i|^{1/L}$. Recall that we assume that all the G_i are simple, in particular their centers are trivial. Now we see that $[G : \mathcal{C}(g)] > |G|^{\delta'/L}$ with the δ' of Lemma 21. Then S contains elements from at least $|G|^{\varepsilon\delta'/L - \delta}$ cosets of $\mathcal{C}(g)$. Thus the set

$$\{sgs^{-1} \mid s \in S\} \subset \prod_{11} S$$

contains at least $|G|^{\varepsilon\delta'/L - \delta}$ different elements h with $\pi_l(h) = 1$, whence

$$|\prod_{12} S| \geq |G|^{\varepsilon\delta'/L - \delta} \prod_{i \in I_l} D_i,$$

which was to be proven. □

Proof of Proposition 14. By Propositions 16 and 20, we have

$$|\prod_{2^m+1} S| > |S||G|^{-Q\delta} \prod_{i\in I_s} D_i^{\delta'_1} \quad \text{and} \quad |\prod_{12} S| > |G|^{\delta'_2-\delta} \prod_{i\in I_l} D_i$$

with some constants δ'_1, δ'_2 and Q. Multiply the first inequality with the δ'_1th power of the second one, and use $|G| \geq |S|$ and $\prod D_i = |A| \geq |S|\,|G|^{-2\delta}$ to get

$$|\prod_{2^m+1} S|\,|\prod_{12} S|^{\delta'_1} > |S|^{1+\delta'_1+\delta'_1\delta'_2}|G|^{-Q'\delta}.$$

By the hypothesis on the set S for $H = \{1\}$, we get $|S| > |G|^{\varepsilon-\delta}$. Therefore (14) gives the claim if δ is sufficiently small. □

4. (A1)–(A5) for $G_i = SL_d(\mathbb{F}_{p^k})$

Let K be a number field and let $I \subset \mathcal{O}_K$ be a square-free ideal. Then $I = P_1 \cdots P_n$ for some prime ideals, and $G = SL_d(\mathcal{O}_K/I) = SL_d(\mathcal{O}_K/P_1) \times \cdots \times SL_d(\mathcal{O}_K/P_n)$. The last ingredient we need for the proof of Theorem 1 is that the groups $G_i = SL_d(\mathcal{O}_K/P_1)$ satisfy the assumptions (A1)–(A5). We write $\mathbb{F}_{p^k}$ for the finite field of order p^k.

(A1) is immediate, and (A2) is a classical result of Jordan. Regarding (A3), Harris and Hering [20] proved that any nontrivial representation of $SL_d(\mathbb{F}_q)$ is of dimension at least $q^{d-1}-1$ or $(q-1)/2$ when $d=2$ and q is odd. In fact for our purposes it is enough to note that any such representation restricted to an appropriate subgroup isomorphic to $SL_2(\mathbb{F}_p)$ gives rise to a nontrivial representation, which is of dimension at least $(p-1)/2$ by a classical result of Frobenius [17].

We study (A4) and (A5) in the next two sections.

4.1. Assumption (A4)

We recall some results of Helfgott. Let $G = SL_d(\mathbb{F}_p)$, and let $S \subset G$ be a set which is not contained in any proper subgroup. Suppose further that $|S| < |G|^{1-\varepsilon}$ for some $\varepsilon > 0$. Then if $d = 2$ [21, Key Proposition] or if $d = 3$ [22, Main Theorem], there is a $\delta > 0$ depending only on ε such that $|S.S.S| \gg |S|^{1+\delta}$. These results imply (A4) for $G_i = SL_d(\mathbb{F}_{p_i})$ if $d = 2$ or $d = 3$ the same way as we proved Theorem 3 using Proposition 14. We show below that the argument in [21] extends easily to groups $G = SL_2(\mathbb{F}_{p^k})$. After the circulation of an early version of this paper I have learnt that this extension of Helfgott's theorem was recently proven by Oren Dinai in his PhD thesis [14].

Let Λ be a subset of the multiplicative group $\mathbb{F}_{p^k}^*$. Denote by Λ^r the set of rth powers of the elements of Λ and set

$$w(\Lambda) = \{w(a) \mid a \in \Lambda\}, \quad \text{where} \quad w(a) = a + a^{-1}.$$

The only notable change needed to extend Helfgott's argument to the case $k > 1$ is to replace [21, Proposition 3.3] by the following

Proposition 22. *Let $\Lambda \subset \mathbb{F}_{p^k}^*$ be a set which contains 1 and is closed under taking multiplicative inverses. Let $a_1, a_2 \in \mathbb{F}_{p^k}^*$, and assume that if $w(\Lambda^2)$ is contained in a proper subfield F of $\mathbb{F}_{p^k}$, then $a_1/a_2 \notin F$. Now if $|\Lambda| < p^{(1-\delta)k}$, then*

$$|\{a_1 w(bc) + a_2 w(bc^{-1}) \mid b, c \in \textstyle\prod_4 \Lambda\}| \gg |\Lambda|^{1+\varepsilon}$$

with a constant ε depending only on δ.

The proof follows the same lines as that of [21, Proposition 3.3].

Proof. Set $\Lambda_1 = \Lambda^2.\Lambda^2$. Using the substitution $b = \bar{b}\bar{c}$ and $c = \bar{b}\bar{c}^{-1}$, we see that

$$\begin{aligned} a_1 w(\Lambda_1) + a_2 w(\Lambda_1) &= \{a_1 w(\bar{b}^2) + a_2 w(\bar{c}^2) \mid \bar{b}, \bar{c} \in \Lambda.\Lambda\} \\ &\subset \{a_1 w(bc) + a_2 w(bc^{-1}) \mid b, c \in \textstyle\prod_4 \Lambda\}. \end{aligned}$$

If $w(\Lambda^2)$ is contained in a subfield F, then $a_1/a_2 \notin F$ by assumption, and then trivially

$$|a_1 w(\Lambda_1) + a_2 w(\Lambda_1)| \geq |w(\Lambda^2)|^2 \geq \frac{1}{16}|\Lambda|^2,$$

and the claim follows.

Therefore we will assume now that $w(\Lambda^2)$ generates $\mathbb{F}_{p^k}$. Assume that

$$|(a_1/a_2) w(\Lambda_1) + w(\Lambda_1)| \leq K|\Lambda| \tag{21}$$

for some constant K. By the Ruzsa–Plünnecke inequalities [30] (see also [34, Corollary 6.9])

$$|w(\Lambda_1) + w(\Lambda_1) - w(\Lambda_1) - w(\Lambda_1)| \ll K^4 |\Lambda|.$$

Note that $w(a)w(b) = w(ab) + w(ab^{-1})$, hence

$$w(\Lambda^2).w(\Lambda^2) \subset w(\Lambda_1) + w(\Lambda_1)$$

and

$$|w(\Lambda^2).w(\Lambda^2) - w(\Lambda^2).w(\Lambda^2)| \ll K^4 |\Lambda|.$$

This would contradict the sum-product theorem if $K = |\Lambda|^\varepsilon$ with ε small enough. The most convenient reference for us is [33, Theorem 1.5] that we can apply with $A = w(\Lambda^2)$ and $a = w(1) = 2$. However the contradiction could also be deduced from the results of [11] or [10]. □

To use this proposition we need to replace [21, Corollary 4.5] by

Lemma 23. *Let $S \subset SL_2(\mathbb{F}_{p^k})$ be symmetric containing 1, and assume that it is not contained in any proper subgroup. Let F be a proper subfield of $\mathbb{F}_{p^k}$. Then there is an absolute constant R such that there is a matrix*

$$x = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \textstyle\prod_R S$$

with $abcd \neq 0$ and $ad/bc \notin F$.

Proof. In this proof the value of R may be different at different occurrences. First note that for any matrix x with entries as above, $ad + bc = 1 \in F$ and hence

$$bc = \frac{1}{ad/bc - 1},$$

so x satisfies the requirements of the lemma exactly if $bc \notin F$. If x does not satisfy this, look at x^2 and notice that the product of the off-diagonal entries is $bc(\operatorname{Tr} x)^2$, hence it remains to show that $\prod_N S$ contains an element with nonzero off-diagonals and with $(\operatorname{Tr} x)^2 \notin F$.

Note that if $\operatorname{span}(\prod_l S) = \operatorname{span}(\prod_{l+1} S)$, where $\operatorname{span}(X)$ denotes $\mathbb{F}_{p^k}$-linear span in $Mat_2(\mathbb{F}_{p^k})$, then $\operatorname{span}(\prod_m S) = \operatorname{span}(\prod_l S)$ for any $m > l$. From this we conclude that as S is not contained in a proper subgroup, $\prod_4 S$ must span $Mat_2(\mathbb{F}_{p^k})$. Let $y_1, y_2, y_3, y_4 \in \prod_4 S$ be a basis of $Mat_2(\mathbb{F}_{p^k})$ and let z_1, z_2, z_3, z_4 be the dual basis with respect to the nondegenerate form $\operatorname{Tr}(yz)$. Denote by ω an element of $\mathbb{F}_{p^k}$ which is not in F but $\omega^2 \in F$. If there is no such element, the rest of the proof is even simpler. Consider the 16 F-vector spaces

$$\omega^{\alpha_1} F \cdot z_1 + \omega^{\alpha_2} F \cdot z_2 + \omega^{\alpha_3} F \cdot z_1 + \omega^{\alpha_4} F \cdot z_1,$$

where the α_i takes the values 0 and 1 independently. Now we invoke Lemma 4.4 from Helfgott [21], which shows that there is a matrix $\bar{x} \in \prod_R S$ which is not contained in any of the above subspaces if R is large enough. By definition, there is an index i such that $(\operatorname{Tr}(y_i \bar{x}))^2 \notin F$. It may happen that one or both off-diagonal entries are zero. Using [21, Lemma 4.4] now for the representation of $SL_2(\mathbb{F}_{p^k})$ acting on $Mat_2(\mathbb{F}_{p^k})$ by conjugations, we see that $w y_i \bar{x} w^{-1}$ has no zero entries for some $w \in \prod_R S$. This proves the claim. $\square$

We remark that in the way [21, Lemma 4.4] is stated, it gives an R which depends on the dimension of $Mat(\mathbb{F}_{p^k})$ over F, however it is easily seen by a careful analysis of the proof in [21] that the dependence is only on the dimension of the subspaces we want to avoid.

Extending [21, Key Proposition] to arbitrary finite fields. The proof on pp. 616 of [21] is given for arbitrary finite fields up to the point when the set V is constructed, except that we get $|V| < p^{k(1-\delta/3)}$ not $|V| < p^{1-\delta/3}$. If $w(V)$ is contained in a proper subfield of $\mathbb{F}_{p^k}$ then denote this subfield by F, and instead of [21, Corollary 4.5] use Lemma 23 to construct the matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. In what follows simply use Proposition 22 instead of [21, Proposition 3.3]. $\square$

4.2. *Assumption (A5)*

We prove that $SL_d(\mathbb{F}_{p^k})$ satisfies (A5) with L depending on d and k. Note that we can embed $SL_d(\mathbb{F}_{p^k})$ into $GL_{kd}(\mathbb{F}_p)$ by Weil restriction. We again rely on the description of the subgroup structure of $GL_d(\mathbb{F}_p)$ given by Nori [27]. Recall that for a group $H < GL_d(\mathbb{F}_p)$, H^+ denotes the subgroup generated by elements of order p. By [27, Theorem B] there is a connected algebraic subgroup $\widetilde{H} < GL_d$ such that $\widetilde{H}(\mathbb{F}_p)^+ = H^+$. By [27, Theorem C] there is a commutative $F < H$ such that $p \nmid |F|$ and $H \lesssim_{L_1} FH^+$ with a constant L_1

depending only on d. Moreover, it follows from the proof there that if P is any p-Sylow subgroup of H^+, then F can be chosen to satisfy

$$F < \mathcal{N}_H(P), \qquad F \cap P = \emptyset \quad \text{and} \quad [\mathcal{N}_H(P) : FP] < L_1. \tag{22}$$

The choice of F is not unique, even for a fixed Sylow subgroup P, however the following is true. Let $K < \mathcal{N}_H(P)/P$ be a group whose order is prime to P. Then there is an $F < \mathcal{N}_H(P)$ with $K = FP/P$ by [29, Theorem 7.41] and all such subgroups F are conjugates of each other by elements of P (see Rotman [29, Theorem 7.42]).

Proposition 24. *Let G be a quasi-simple subgroup of $GL_d(\mathbb{F}_p)$ such that $G = G^+$. There are classes $\mathcal{H}_0, \dots, \mathcal{H}_m$ of subgroups of G such that the following hold with some constants L, m depending only on d:*

(i) $\mathcal{H}_0 = \{Z(G)\}$,
(ii) *each $\mathcal{H}_i$ is a set of proper subgroups of G, closed under conjugation by elements of G,*
(iii) *for every proper subgroup $H < G$ there is some i and a subgroup $H^\sharp \in \mathcal{H}_i$ such that $H \lesssim_L H^\sharp$,*
(iv) *for every pair of subgroups $H_1, H_2 \in \mathcal{H}_i$, $H_1 \neq H_2$, there is some $i' < i$ and $H^\sharp \in \mathcal{H}_{i'}$ such that $H_1 \cap H_2 \lesssim_L H^\sharp$.*

Proof. In each subgroup $H < G$ which is generated by elements of order p, distinguish a p-Sylow subgroup P. This can be arbitrary, but should be fixed throughout the proof. For integers i and j we define classes $\mathcal{H}_{i,j}$. A proper subgroup $H < G$ belongs to $\mathcal{H}_{i,j}$ precisely if $Z(G) < H$, $\dim \widetilde{H} = i$ and j is the least integer for which the following hold. There is a commutative subgroup $F < \mathcal{N}_H(P)$ such that

$$Z(G) < F, \qquad H = FH^+, \qquad F \cap P = \emptyset, \tag{23}$$

$$[\mathcal{N}_{H^+}(P) : (F \cap H^+)P] < L_1^{2^{d-j}}, \tag{24}$$

and there is a j-dimensional subspace V of $Mat_d(\mathbb{F}_p)$ such that

$$F = V \cap \mathcal{N}_G(P) \cap \mathcal{N}_G(H^+). \tag{25}$$

Order the nonempty classes $\mathcal{H}_{i,j}$ in such a way that $\mathcal{H}_{i,j}$ precedes $\mathcal{H}_{i',j'}$ if $i < i'$ or $i = i'$ and $j < j'$. Condition (24) may look artificial, but it plays an important role in the proof of (iv). Up to that point it can be safely ignored.

The first nonempty class is $\mathcal{H}_{0,j} = \{Z(G)\}$ for some j. Indeed, if two matrices belong to the center of G, then any linear combination of them is central, too, provided it belongs to the group. Furthermore, if $Z(G) \neq H < G$ belongs to $\mathcal{H}_{0,j'}$ for some j', then by (23), the linear span of H contains the span of $Z(G)$, so $j' > j$. Hence (i) follows.

Since conjugation is a linear transformation on $Mat_d(\mathbb{F}_p)$, (ii) is clear.

Let $H < G$ be a proper subgroup, and replace it by $Z(G)H$ if necessary, to ensure that $Z(G) < H$. Let F be a subgroup of $\mathcal{N}_H(P)$ that satisfies (22). Without loss of generality, we can assume that $Z(G) < F$. Set

$$F^\sharp = \operatorname{span}(F) \cap \mathcal{N}_G(P) \cap \mathcal{N}_G(H^+),$$

where $\mathrm{span}(F)$ is the linear span of F in the vector space $Mat_d(\mathbb{F}_p)$. First we remark that $F^\sharp$ does not contain an element of order p, in fact its elements can be mutually diagonalized over an appropriate extension field. This implies that $F^\sharp \cap P = \emptyset$. Since $F^\sharp \subset \mathcal{N}_G(H^+)$, we can define the subgroup $H^\sharp = F^\sharp H^+$, and we have $(H^\sharp)^+ = H^+$. Since $[H : FH^+] < L_1$ and $FH^+ < H^\sharp$, for (iii) we only need to show that $H^\sharp \in \mathcal{H}_{i,j}$ for some i and j. First we remark that $H^\sharp$ is a proper subgroup, since $(H^\sharp)^+ = H^+$. (Note that G is generated by elements of order p, hence $H^+ < G^+ = G$.) Second, (23) and (25) hold with $i = \dim \widetilde{H}$ and with $F^\sharp$ and $V = \mathrm{span}(F^\sharp)$. For (24), we can write

$$[\mathcal{N}_{H^+}(P) : (F^\sharp \cap H^+)P] \le [\mathcal{N}_{H^+}(P) : (F \cap H^+)P] = [\mathcal{N}_{FH^+}(P) : FP] \le L_1.$$

Here the equality in the middle follows from the fact $\mathcal{N}_{FH^+}(P) = F\mathcal{N}_{H^+}(P)$, while the last inequality is contained in (22).

It remains to show (iv). Let H_1 and H_2 be two different groups in $\mathcal{H}_{i,j}$. If $\widetilde{H}_1 \ne \widetilde{H}_2$, then

$$\dim(\widetilde{H}_1 \cap \widetilde{H}_2) \le \dim \widetilde{H_1 \cap H_2} < i$$

and $(H_1 \cap H_2)^\sharp \in \mathcal{H}_{i',j'}$ with some $i' < i$, as we saw in the previous paragraph. Therefore we may assume $\widetilde{H}_1 = \widetilde{H}_2$ and hence $H_1^+ = H_2^+$. Let P be the distinguished p-Sylow subgroup and denote by $F_l < \mathcal{N}_{H_l}(P)$ and V_l ($l = 1, 2$) the subgroups and subspaces for which (23)–(25) hold. We show that there is an $H \in \mathcal{H}_{i,j'}$ for some $j' < j$ such that $H_1 \cap H_2 \lesssim_{L_1^{2^{d-j+1}}} H$. We have $[\mathcal{N}_{H_l}(P) : F_l P] < L_1^{2^{d-j}}$ for $l = 1, 2$, hence

$$[\mathcal{N}_{H_1 \cap H_2}(P) : F_1 P \cap F_2 P] < L_1^{2^{d-j+1}}. \tag{26}$$

Now let $F < F_1$ be such that $FP = F_1 P \cap F_2 P$, and define $H = FH^+$. Then $H_1 \cap H_2 \lesssim_{L_1^{2^{d-j+1}}} H$ by (26). By [29, Theorem 7.42] (as mentioned before the statement of the proposition), there is an element $g \in P$ such that $gFg^{-1} < F_2$. Then $F = F_1 \cap g^{-1}F_2 g$. Now (23) follows by construction. For (25), we have

$$F = V_1 \cap g^{-1}V_2 g \cap \mathcal{N}_G(P) \cap \mathcal{N}_G(H^+)$$

and $\dim(V_1 \cap g^{-1}V_2 g) < j$ since $H_1 \ne H_2$. Finally for (24), we can write

$$[\mathcal{N}_{H^+}(P) : (F \cap H^+)P] \le [\mathcal{N}_{H^+}(P) : (F_1 \cap H^+)P] \cdot [\mathcal{N}_{H^+}(P) : (g^{-1}F_2 g \cap H^+)P]. \quad \square$$

5. Proof of Theorem 1

First we note that by [23, Claim 11.19], it is enough to prove that $\mathcal{G}(SL_d(\mathcal{O}_K/I), \pi_I(S'))$ form a family of expanders with some $S' \subset \Gamma$, hence we can assume without loss of generality that Theorem 2 holds with $S = S'$. If $H < SL_d(\mathcal{O}_K/I)$ and $\pi_I(S) \subset H$, then by Theorem 2, $[SL_d(\mathcal{O}_K/I) : H] < C$ for some constant C which depends on the δ and the implied constant of that theorem. Let J be a square-free ideal for whose prime factors P, $\pi_P(S)$ does not generate $SL_d(\mathcal{O}_K/P)$. Since each proper subgroup in $SL_d(\mathcal{O}_K/P)$ is of index at least $N(P)^{\delta'}$ for some $\delta' > 0$, we get $N(J) < C^{\delta'}$. Here and below, δ' is a constant which may depend on S and which need not be the same at different occurrences. Thus there are at most a finite number of prime ideals P such that $\pi_P(S)$ is not generating, and from now on, we denote by J the product of those prime ideals.

Let I be an ideal which is prime to J and write $G = SL_d(\mathcal{O}_K/I)$, and $\overline{S} = \pi_I(S)$. Denote by $l^2(G)$ the vector space of complex valued functions on G. Consider the operator on $l^2(G)$ which is convolution with $\chi_{\overline{S}}$ from the left. Denote its matrix in the standard basis by M. It is plain that $|S|M$ is the adjacency matrix of the graph $\mathcal{G}(G, \overline{S})$. In light of the results of Dodziuk [15], Alon and Milman [3] and Alon [2] already mentioned in the introduction, we have to give an upper bound on the second largest eigenvalue of M independently of I. For $g \in G$, denote by $\alpha(g)$ left translation by g on $l^2(G)$. Then α is called the *regular representation* of G, and it is well known that $l^2(G)$ decomposes as a direct sum $V_0 \oplus V_1 \oplus \cdots \oplus V_m$ such that each $\alpha|_{V_i}$ is irreducible and the multiplicity of every irreducible representation of G in this decomposition is the same as its dimension. Therefore it remains to show that if β is a nontrivial irreducible representation of G, and λ is an eigenvalue of the operator

$$\frac{1}{|S|}\sum_{g\in\overline{S}}\beta(s),$$

then $\lambda < c < 1$ for some constant c independent of I. Replacing I by a larger ideal if necessary, we may assume that the representation is faithful. Faithful representations of G are tensor products of nontrivial representations of the direct factors, hence they are of dimension at least $|G|^{\delta'}$ as we noted at the beginning of Section 4. Hence λ is an eigenvalue of M with multiplicity at least $|G|^{\delta'}$.

Denote by $(M)_{i,j}$ the i, j entry of M and notice that for an integer k, the rows of M^k are translates of $\chi_{\overline{S}}^{(k)}$. Then

$$\mathrm{Tr}(M^{2k}) = \sum_{i,j\le|G|}(M^k)_{i,j}^2 = |G|\,\|\chi_{\overline{S}}^{(k)}\|_2^2,$$

whence

$$\lambda^{2k} \le |G|^{1-\delta'}\|\chi_{\overline{S}}^{k}\|_2^2. \tag{27}$$

If the index of a subgroup $H < SL_d(\mathcal{O}_K/I)$ is large, then we can cancel the implied constant in Theorem 2 by making δ smaller. If the index is small, then we can get a nontrivial bound $\chi_{\overline{S}}^{(k)}(H) < c < 1$, since we assumed that $\overline{S}$ generates the group. Thus if I is restricted to ideals prime to J, Theorem 2 holds with the implied constant set to 1. Now apply it for $H = \{1\}$ to get

$$\|\chi_{\overline{S}}^{(\log N(I))}\|_2^2 < |G|^{-\delta'}.$$

We saw in Section 4 that G satisfies (A0)–(A3) and (A5). It also satisfies (A4) if $d = 2$ or if $d = 3$ and $K = \mathbb{Q}$ or if we assume that (1) holds when F ranges over the fields $\mathcal{O}_K/P$, P prime. Therefore we can apply Theorem 3 repeatedly to get

$$\|\chi_{\overline{S}}^{(C\log(N(I)))}\|_2^2 < |G|^{-1+\varepsilon}$$

for arbitrary $\varepsilon > 0$ with some constant C depending on ε. If ε is less than the δ' in (27), the theorem follows.

Acknowledgments. I am grateful to my advisor Jean Bourgain for suggesting this problem and for guiding me during my research. I also had very useful discussions with Elon Lindenstrauss, Alireza Salehi Golsefidy and Peter Sarnak; I thank them for their interest and valuable remarks.

While writing this paper, I was supported by a Fulbright Science and Technology Award, grant no. 15073240.

Note added in proof. Soon after the submission of this paper, equation (1) was proven independently by Breuillard, Green and Tao [12] and Pyber and Szabó [28], even in the more general context of finite simple groups of Lie type. Those results imply that Theorem 1 holds unconditionally for arbitrary d and K.

References

[1] Abels, H., Margulis, G. A., Soifer, G. A.: Semigroups containing proximal linear maps. Israel J. Math. **91**, 1–30 (1995) Zbl 0845.22004 MR 1348303

[2] Alon, N.: Eigenvalues and expanders. Combinatorica **6**, 83–96 (1986) Zbl 0661.05053 MR 0875835

[3] Alon, N., D. Milman, V.: λ_1, isoperimetric inequalities for graphs, and superconcentrators. J. Combin. Theory Ser. B **38**, 73–88 (1985) Zbl 0549.05051 MR 0782626

[4] Berenstein, C. A., Yger, A.: Effective Bezout identities in $\mathbb{Q}[z_1, \dots, z_n]$. Acta Math. **166**, 69–120 (1991) Zbl 0724.32002 MR 1088983

[5] Bourgain, J., Gamburd, A.: Uniform expansion bounds for Cayley graphs of $SL_2(\mathbb{F}_p)$, Ann. of Math. **167**, 625–642 (2008) Zbl 1216.20042 MR 2415383

[6] Bourgain, J., Gamburd, A.: Expansion and random walks in $SL_d(\mathbb{Z}/p^n\mathbb{Z})$: I. J. Eur. Math. Soc. **10**, 987–1011 (2008) Zbl 1193.20059 MR 2443926

[7] Bourgain, J., Gamburd, A.: Expansion and random walks in $SL_d(\mathbb{Z}/p^n\mathbb{Z})$: II. With an appendix by J. Bourgain. J. Eur. Math. Soc. **11**, 1057–1103 (2009) Zbl 1193.20060 MR 2538500

[8] Bourgain, J., Gamburd, A., Sarnak, P.: Sieving and expanders. C. R. Math. Acad. Sci. Paris **343**, 155–159 (2006) Zbl 1217.11081 MR 2246331

[9] Bourgain, J., Gamburd, A., Sarnak, P.: Affine linear sieve, expanders, and sum-product. Invent. Math. **179**, 559–644 (2010) Zbl pre05675041 MR 2587341

[10] Bourgain, J., Glibichuk, A. A., Konyagin, S. V.: Estimates for the number of sums and products and for exponential sums in fields of prime order. J. London Math. Soc. **73**, 380–398 (2006) Zbl 1093.11057 MR 2225493

[11] Bourgain, J., Katz, N., Tao, T.: A sum-product estimate in finite fields, and applications. Geom. Funct. Anal. **14**, 27–57 (2004) Zbl 1145.11306 MR 2053599

[12] Breuillard, E., Green, B., Tao, T.: Approximate subgroups of linear groups. Geom. Funct. Anal. **21**, 774–819 (2011) Zbl pre05950685

[13] Cano, A., Seade, J.: On the equicontinuity region of discrete subgroups of $PU(1,n)$. J. Geom. Anal. **20**, 291–305 (2010) Zbl 1218.37059 MR 2579511

[14] Dinai, O.: Expansion properties of finite simple groups. PhD thesis, Hebrew Univ. (2009)

[15] Dodziuk, J.: Difference equations, isoperimetric inequality and transience of certain random walks. Trans. Amer. Math. Soc. **284**, 787–794 (1984) Zbl 0512.39001 MR 0743744

[16] Farah, I.: Approximate homomorphisms II: group homomorphisms. Combinatorica **20**, 47–60 (2000) Zbl 1008.28004 MR 1770537

[17] Frobenius, G.: Ueber Gruppencharaktere. Sitzungsber. Königlich Preuß. Akad. Wiss. Berlin **1896**, 985–1021 JFM 27.0092.01

[18] Goldsheid, I. Ya., Margulis, G. A.: Lyapunov exponents of a product of random matrices. Uspekhi Mat. Nauk **44**, no. 5, 13–60 (1989) (in Russian); English transl.: Russian Math. Surveys **44**, no. 5, 11–71 (1989) Zbl 0705.60012 MR 1040268

[19] Gowers, W. T.: Quasirandom groups. Combin. Probab. Comput. **17**, 363–387 (2008) Zbl 1191.20016 MR 2410393

[20] Harris, M. E., Hering, C.: On the smallest degrees of projective representations of the groups $PSL(n,q)$. Canad. J. Math. **23**, 90–102 (1971) Zbl 0221.20009 MR 0272917

[21] Helfgott, H. A., Growth and generation in $SL_2(\mathbb{Z}/p\mathbb{Z})$. Ann. of Math. **167**, 601–623 (2008) Zbl 1213.20045 MR 2415382

[22] Helfgott, H. A.: Growth in $SL_3(\mathbb{Z}/p\mathbb{Z})$. J. Eur. Math. Soc. **13**, 761–851 (2011) Zbl pre05873844 MR 2781932

[23] Hoory, S., Linial, N., Widgerson, A.: Expander graphs and their applications. Bull. Amer. Math. Soc. **43**, 439–561 (2006) Zbl 1147.68608 MR 2247919

[24] Kesten, H.: Symmetric random walks on groups. Trans. Amer. Math. Soc. **92**, 336–354 (1959) Zbl 0092.33503 MR 0109367

[25] Long, D. D., Lubotzky, A., Reid, A. W.: Heegaard genus and property τ for hyperbolic 3-manifolds. J. Topol. **1**, 152–158 (2008) Zbl 1158.57018 MR 2365655

[26] Nikolov, N., Pyber, L.: Product decompositions of quasirandom groups and a Jordan type theorem. J. Eur. Math. Soc. **13**, 1063–1077 (2011) Zbl pre05919471 MR 2800484

[27] Nori, M. V.: On subgroups of $GL_n(\mathbb{F}_p)$. Invent. Math. **88**, 257–275 (1987) Zbl 0632.20030 MR 0880952

[28] Pyber, L., Szabó, E.: Growth in finite simple groups of Lie type of bounded rank. http://arxiv.org/abs/1005.1858

[29] Rotman, J. J.: An Introduction to the Theory of Groups. 4rth ed., Grad. Texts in Math. 148, Springer, New York (1995) Zbl 0810.20001 MR 1307623

[30] Ruzsa, I. Z.: An application of graph theory to additive number theory. Sci. Ser. A Math. Sci. (N.S.) **3**, 97–109 (1989) Zbl 0743.05052 MR 2314377

[31] Sarnak, P., Xue, X. X.: Bounds for multiplicities of automorphic representations. Duke Math. J. **64**, 207–227 (1991) Zbl 0741.22010 MR 1131400

[32] Tao, T.: Product set estimates for non-commutative groups. Combinatorica **28**, 547–594 (2008) Zbl pre05494691 MR 2501249

[33] Tao, T.: The sum-product phenomenon in arbitrary rings. Contrib. Discrete Math. **4**, 59–82 (2009) Zbl pre05704242 MR 2592424

[34] Tao, T., Vu, V. H.: Additive Combinatorics. Cambridge Univ. Press, Cambridge (2006) Zbl 1127.11002 MR 2289012

[35] Tits, J.: Free subgroups in linear groups. J. Algebra **20**, 250–270 (1972) Zbl 0236.20032 MR 0286898

[36] Woess, W.: Random Walks on Infinite Graphs and Groups. Cambridge Tracts in Math. 138, Cambridge Univ. Press, Cambridge (2000) Zbl 0951.60002 MR 1743100