Product Growth and Mixing in Finite Groups

László Babai* Nikolay Nikolov † László Pyber‡

Abstract

We prove the following inequality on the convolution of distributions over a finite group G:

$$(0.1) \quad \| X * Y - U \| \leq \sqrt{n/m} \| X - U \| \| Y - U \|,$$

where X, Y are probability distributions over G, the $*$ denotes convolution, U the uniform distribution over G, and $\| . \|$ the ℓ_2-norm; n is the order of G, and m denotes the minimum dimension of nontrivial real representations of G. This inequality can be viewed as a new expansion property of a large class of groups, including all Lie-type simple groups of bounded rank, all of which satisfy $m > cn^\beta$ (where $c > 0$ is an absolute constant and $\beta > 0$ depends on the rank bound only). Best among them are the groups $G = \mathrm{SL}_2(q)$ (2×2 matrices with determinant 1 over $\mathbb{F}_q$) where $m \sim n^{1/3}/2$.

We derive applications of the convolution inequality (0.1) to a variety of areas, ranging from stochastic processes to additive combinatorics to group theory.

An immediate consequence is a product growth inequality for subsets of G: if $A, B \subseteq G$ then $|AB| > n/(1 + \Delta)$ where $\Delta = n^2/(m|A||B|)$. On the one hand, this corollary strengthens a recent result of Gowers which served as the inspiration to the present work; on the other hand, it gives a strong (and best possible) affirmative answer to a problem regarding the product growth of subsets of $\mathrm{SL}_2(q)$ recently posed by Venkatesh and Green at a conference in the newly flourishing area of "additive combinatorics."

Another corollary to the main inequality shows that for groups with large m, mixing in the strongest sense (ℓ_∞-norm) occurs more rapidly than expected; we prove that if X, Y, Z are distributions over G then

$$(0.2) \quad \| X * Y * Z - U \|_\infty < \sqrt{\frac{n}{m}} \| X \| \| Y \| \| Z \|.$$

This generalizes a result of Gowers.

By easy induction, our main inequality generalizes to the convolution of multiple terms and thereby results in rapid mixing estimates for time-inhomogeneous Cayley walks on G. It also gives estimates for the size of the product of several subsets, resulting in diameter estimates for Cayley graphs and tying in with the broad subject of "bounded generation" in group theory.

An illustration of the connection to diameters: for $G = \mathrm{SL}_2(q)$ it follows that if $A \subseteq G$ and $|A| \geq n^{2/3+\epsilon}$ then $A^t = G$ where $t = O(1/\epsilon)$; we also show that the elements of G are represented nearly uniformly as words of length t over A.

The connection to "bounded generation" is illustrated by one of the main applications of our results: *every finite simple group of Lie type of characteristic p is the product of 5 Sylow p-subgroups.* – Results of this type are among the ingredients of the recent breakthrough result that all finite simple groups have bounded degree expander Cayley graphs [KLN]; our results improve and greatly simplify these ingredients.

The results and techniques used in this paper were inspired by a link between quasirandomness and group representation theory recently found by Gowers [Go].

1 Introduction

1.1 Expansion of finite groups. Expansion properties of graphs have been of great interest to the theory of computing for decades, and every new technique introduced in the area promises new applications. Many of the best expanders are linked to groups and specifically to the group $\mathrm{SL}_2(q)$ of 2×2 matrices with determinant 1 over the field $\mathbb{F}_q$ (q a prime power). In particular, the Ramanujan graphs constructed by Lubotzky, Phillips, Sarnak [LuPS] and by Margulis [Ma] are Cayley graphs of this group. A belief in a certain degree of expansion of finite simple groups is expressed in the conjecture [BaSe] that all Cayley graphs of such groups have polylogarithmic diameter (and therefore even for the worst generators, the random walks converge rapidly). In a recent breakthrough, Helfgott [He] confirmed this conjecture for the groups $\mathrm{SL}_2(p)$ (p prime), by proving a powerful product growth estimate

*University of Chicago. Email: laci@cs.uchicago.edu. Part of this work was done at the Rényi Mathematical Institute, Budapest, and at the Centre Interfacultaire Bernoulli, Lausanne.

†Imperial College London. Email: n.nikolov@imperial.ac.uk.

‡Rényi Mathematical Institute, Budapest. Email: pyber@renyi.hu.

for sets of generators of $\mathrm{SL}_2(p)$ (see Theorem 4.2).

1.2 Time-inhomogeneous Cayley walks. In the present paper we give yet another type of expansion result for a large class of groups. Expansion has been known to be related to rapid mixing of random walks; here we consider time-inhomogeneous Cayley walks on the finite group G. In a single step, a Cayley walk moves according to a probability distribution on G; the transition matrix is the G-circulant generated by the distribution (see Section 5.2 for the definitions). In a time-inhomogeneous Cayley walk we permit a different distribution at each time step. We consider Cayley walks where each step is "moderately random," as measured by an upper bound on the ℓ_2-norm of the corresponding distribution. This "initial randomness" requirement will be defined in terms of the parameter m, the minimum dimension of nontrivial real representations of G (the larger the dimension, the less initial randomness is required).

Our central result estimates the ℓ_2-distance to the uniform distribution of a 2-step time-inhomogeneous Cayley walk. In other words, we estimate how far the G-convolution of two distributions over G is from uniform. By easy induction, the result extends to time-inhomogeneous Cayley walks of any length. We note that our Cayley walks are not necessarily symmetric, i. e., the measure assigned to $g \in G$ and to g^{-1} may be different.

From the convolution bound we infer an unexpectedly strong bound on mixing time in the ℓ_∞ norm. In particular for $G = \mathrm{SL}_2(q)$ (of order $n = q(q^2-1)$) we show that if each distribution involved has ℓ_2-norm $\leq \epsilon n^{-4/9}$ then in three steps, the distribution is within $\sqrt{2}\epsilon^3/n$ of uniform in ℓ_∞ norm. More generally, if each probability distribution involved has ℓ_2-norm $< n^{-1/3-\epsilon}$ then in $\approx 5/(3\epsilon)$ steps the distribution will be nearly uniform in ℓ_∞ norm.

1.3 Applications. This type of estimate reduces the task of estimating the mixing time of Cayley walks on $G = \mathrm{SL}_2(q)$ to analyzing an initial phase where the walk spreads to an $n^{2/3+\epsilon}$ portion of the space. It combines well with results obtained from sum-product estimates [BoKT] which work up to n^{1-c} [He]. Indeed, such estimates can replace the (easier) second phase of Helfgott's analysis of the worst case diameter and mixing rate of $\mathrm{SL}_2(p)$ [He].

A number of corollaries regarding the product growth of subsets of G follow. We show that subsets of size just above $n^{2/3}$ have rapid product growth, i. e., for two such subsets $A, B \subseteq \mathrm{SL}_2(q)$, the set AB is much larger than $\max\{|A|, |B|\}$. We also infer that a product of $t \approx 1/(3\epsilon)$ subsets of G of size $\geq n^{2/3+\epsilon}$ is the entire group and the elements of the group are nearly uniformly represented as products of length t from these sets.

We obtain a similar result, with $n^{2/3}$ replaced by n^{1-c}, for all other classes of (quasi)simple groups of Lie type of bounded rank. (A *quasisimple group* G is a simple group with possibly a small center tucked underneath; G has to be *perfect*, i. e., $G = G'$ where G' is the commutator subgroup. An example is $\mathrm{SL}_2(q)$ which has a center of order 2 if q is odd; and the quotient by the center is the simple group $\mathrm{PSL}_2(q)$.)

Our results have already found a number of applications in the area of "bounded generation." Typical results in this area state that most finite simple groups are the product of a bounded number of subsets of a certain type such as abelian subgroups, Sylow subgroups, certain conjugacy classes, commutators and more complex words [LP, Shv, Shm, NS, Ni, NP, APPS]. Bounded generation results for *finite* groups are motivated by a multitude of application to major problems in the theory of *infinite* groups (Serre's conjecture on the topology of profinite groups via "word width" of finite groups [NS], model theoretic characterization of simple pseudofinite groups via commutator width [Wil], profinite completions of arithmetic groups via Sylow width [LP, HruP], expansion/Kazhdan constants of certain infinite matrix groups via product decomposition into elementary subgroups [Shm], etc.)

Because of the technical nature of the consequences of our work to bounded generation, we do not go into the details here but mention one of the main applications which will appear in the journal version of this paper.

Application 1.1. *Every finite simple group of Lie type defined over a field of characteristic p can be written as a product of at most 5 of its Sylow p-subgroups.*

It was previously known that a bounded number of Sylow p-subgroups suffice; the previous bound was 25 [NP]. The proof was based on detailed structural arguments. For the most part we can argue by the size of certain subsets, ignoring the structure and thus greatly simplifying the proofs and at the same time obtaining considerably better, nearly optimal bounds.

This result has an important connection to a recent major result, the proof that all finite simple groups have expander Cayley graphs (Kassabov, Lubotzky, Nikolov [KLN]). Bounded Sylow-width is an ingredi-

ent; this ingredient has now been greatly improved and simplified.

1.4 Methods. Our main convolution estimate (Theorem 2.1), of which all other result proved in this paper are corollaries, is proved by invoking the representation theory of finite groups to estimate the second eigenvalue that is the key to expansion and mixing, in the spirit of a recent paper by Gowers [Gow2] (which incidentally was motivated by giving a strong negative answer to a question stated in [BaSo]).

Specifically, the key parameter of the group G that plays a role in this connection is the minimum degree m of nontrivial representations of G, i. e., the minimum dimension of complex vector spaces on which the group can act nontrivially. For instance, the order of the group $G = \mathrm{SL}_2(q)$ is $n = q(q^2-1)$ and the minimum degree of its nontrivial representations is $m = (q-1)/2 \sim n^{1/3}/2$.

The connection between the parameter m and eigenvalue estimation has previously been used by several authors in a similar fashion [SarX, BoG]. Recently Bourgain and Gamburd [BoG] used it to obtain uniform and almost sure expansion bounds for $\mathrm{SL}_2(q)$. In the sweet 2003 monograph by Davidoff, Sarnak, and Valette, this method is used to give an elementary proof of a strong but not optimal estimate of the second eigenvalue of the LPS/Margulis Ramanujan graphs [DSV, pp. 124–128].

Because of the magnitude of the parameter m for the groups $\mathrm{SL}_2(q)$, these groups were the primary targets in each case mentioned.

Gowers made the relation between the magnitude of m and quasirandomness properties of the group explicit, an insight that has guided our work although the quasirandomness philosophy [Th, CGW] (cf. [AS, Gow1]) is no longer present and is subsumed by the linear algebra machinery of mixing.

1.5 Outline of the paper. In Section 2.1 we describe the mixing results, starting with the Main Convolution Bound (Theorem 2.1), the main result of this paper. In Section 2.2 we state the applications to "arithmetic combinatorics" (product growth of sets, Gowers' original context). Section 2.3 describes the resulting tight solution to Venkatesh's problem on the product growth of subsets of $\mathrm{SL}_2(q)$. In Section 2.4 we take a look at the key parameter m (minimum degree of irreducible representations) for finite simple groups. We review the main result of Gowers' paper [Gow2] and its connection to our work in Section 2.5. In Section 3 we derive all the corollaries from the Main Convolution Bound (Theorem 2.1). Applications to diameter is discussed in Section 4. Section 5 is devoted to proving Theorem 2.1.

2 The main results

Throughout this paper we use the following notation: G is a finite group of order n and m denotes the minimum dimension of its nontrivial real representations. (A "trivial representation" represents every group element by the identity transformation of the representation space.) $\|\,.\,\|$ denotes the ℓ_2-norm, $*$ denotes convolution, and U the uniform distribution over G.

To put the results in a more concrete perspective, it may be helpful to keep the most important special case of $G = \mathrm{SL}_2(q)$ in mind; in this case, $m \sim n^{1/3}/2$.

2.1 Convolutions. This paper has one new theorem and countless corollaries. We start with the theorem.

Theorem 2.1. (Main Convolution Bound) *Let X and Y be two probability distributions over G. Then*

$$\| X * Y - U \| \le \sqrt{n/m} \| X - U \| \| Y - U \|, \tag{2.3}$$

A generalization to multiple terms is immediate.

Corollary 2.2. (Multi-term Convolution Bound) *Let $X_1, \ldots, X_t$ $(t \ge 1)$ be probability distributions over G. Then*

$$\left\| \prod_{i=1}^{t}{}^{*} X_i - U \right\| \le \left(\frac{n}{m}\right)^{(t-1)/2} \prod_{i=1}^{t} \| X_i - U \| \tag{2.4}$$

where $\prod^*$ denotes convolution.

Near uniformity in the strongest sense, i. e., in ℓ_∞ norm, sets in much sooner than predicted by Corollary 2.2.

Corollary 2.3. (Max-Norm Uniformity) *If $X_1, \ldots, X_t$ $(t \ge 2)$ are distributions over G then*

$$\left\| \prod_{i=1}^{t}{}^{*} X_i - U \right\|_\infty < \left(\frac{n}{m}\right)^{t/2-1} \prod_{i=1}^{t} \| X_i \|. \tag{2.5}$$

2.2 Product Growth. In this section we consider results on product growth of subsets of G. Recall that for $A, B \subseteq G$ we define the set $AB \subseteq G$ as $AB = \{ab : a \in A, b \in B\}$.

COROLLARY 2.4. (MULTISTEP PRODUCT GROWTH) *Let $A_1, \dots, A_t$ $(t \geq 1)$ be nonempty subsets of G. Then*

$$\left| \prod_{i=1}^{t} A_i \right| > \frac{n}{1+\Delta}, \tag{2.6}$$

where

$$\Delta = \frac{n^t}{m^{t-1} \prod_{i=1}^{t} |A_i|}. \tag{2.7}$$

We especially highlight the case $t = 2$ because (a) it solves a published open problem; and (b) it directly improves over Gowers' result (see Section 2.5).

COROLLARY 2.5. (2-STEP PRODUCT GROWTH) *Let A, B be nonempty subsets of G. Then*

$$|AB| > \frac{n}{1+\frac{n^2}{m|A||B|}} \geq \min\left\{\frac{n}{2}, \frac{m|A||B|}{2n}\right\}. \tag{2.8}$$

Note that the second inequality can be rephrased as follows:

If $|B| \geq \alpha n/m$ then

$$|AB| > \min\left\{n/2, (\alpha/2)|A|\right\}. \tag{2.9}$$

Next we state a particularly important corollary which is the basis of all our applications to the theory of "bounded generation" as well as to diameter bounds.

COROLLARY 2.6. (PRODUCT DECOMPOSITION 1) *Let $A_1, \dots, A_t$ $(t \geq 2)$ be nonempty subsets of G. Assume*

$$\prod_{i=1}^{t} |A_i| \geq \frac{n^t}{m^{t-2}}. \tag{2.10}$$

Then each of the following holds:

(a) $A_t \cap \prod_{i=1}^{t-1} A_i \neq \emptyset$.

(b) $\prod_{i=1}^{t} A_i = G$.

In fact, as we shall see from the proof in Section 3, this result already follows from the inequality

$$\left| \prod_{i=1}^{t} A_i \right| > n(1-\Delta) \tag{2.11}$$

in Corollary 2.4, a considerable weakening of inequality (2.6). If we use the full strength of inequality (2.6), we obtain the following weaker sufficient condition for the conclusions of Corollary 2.6.

COROLLARY 2.7. (PRODUCT DECOMPOSITION 2) *Let $A_1, \dots, A_t$ $(t \geq 2)$ be nonempty subsets of G. Assume*

$$\prod_{i=1}^{t} |A_i| \geq \frac{n^{t-1}(n - |A_t|)}{m^{t-2}}. \tag{2.12}$$

Then the conclusions of Corollary 2.6 follow.

If the condition of Corollary 2.6 is *amply satisfied* then the elements of G are nearly uniformly represented as products $\prod_{i=1}^{t} a_i$ where $a_i \in A_i$. This is formalized in the next statement.

COROLLARY 2.8. (UNIFORM DECOMPOSITION) *Let $A_1, \dots, A_t$ $(t \geq 2)$ be nonempty subsets of G. For $g \in G$, let $N_g = N_g(A_1, \dots, A_t)$ denote the number of solutions to the equation $g = x_1 \dots x_t$ where $x_i \in A_i$ and let $N = N(A_1, \dots, A_t)$ denote the number of solutions to the equation $x_1 \dots x_{t-1} = x_t$ where $x_i \in A_i$. If $\alpha > 0$ and $\prod_{i=1}^{t} |A_i| \geq \frac{\alpha n^t}{m^{t-2}}$ then for all $g \in G$,*

$$\left| N_g - \frac{\prod_{i=1}^{t} |A_i|}{n} \right| < \frac{1}{\sqrt{\alpha}} \frac{\prod_{i=1}^{t} |A_i|}{n}, \tag{2.13}$$

and the same inequality holds for N in place of N_g.

The $t = 3$ cases of Corollaries 2.6 and 2.8 are equivalent to results of Gowers (see Section 2.5).

2.3 Product growth in SL_2. The following question, posed by A. Venkatesh, was communicated by Ben Green at the 2004 AIM Workshop on Additive Combinatorics, an emerging field best defined by the beautiful 2006 monograph by Tao and Vu [TaV][1].

PROBLEM 2.9. [CL, Problem 3.4]. *Let $X \subset \mathrm{SL}_2(p)$ satisfy $|A| \sim p^{5/2}$; does this imply $|A^2| > p^{5/2+\delta}$?*

We confirm this in a strong form. The following is an immediate consequence of Corollary 2.5 and the fact that for $G = \mathrm{SL}_2(q)$ we have $m = (q-1)/2$.

COROLLARY 2.10. *If $A \subseteq \mathrm{SL}_2(q)$ and $A \sim q^{5/2}$ then $|A^2| \gtrsim p^3/3 \sim n/3$ (A^2 takes up at least about one third the group). Here the asymptotic notation refers to the limit $q \to \infty$.*

In fact, the desired kind of expansion already starts at $|A| \gtrsim p^{2+\delta}$.

[1] The field has been renamed "arithmetic combinatorics" in a Fall 2007 IAS semester.

COROLLARY 2.11. *Assume the subset* $A \subset \mathrm{SL}_2(q)$ *satisfies* $|A| \sim q^{2+\delta}$ *where* $0 < \delta < 1/2$. *Then* $|A^2| \gtrsim q^{2+2\delta}$. *(Here* δ *is fixed while* $q \to \infty$.*)*

REMARK 2.12. This result is best possible in the sense that $|A| \sim q^2$ would not imply any growth. Indeed we could then choose A to be a subgroup of order $q(q-1)$ (upper triangular matrices). Therefore in the range $|A| \le q(q-1)$, Gowers-type arguments cannot seem to replace Helfgott's result about triple products, $|AAA| \ge |A|^{1+\delta}$, where A generates $\mathrm{SL}_2(q)$.

REMARK 2.13. Corollary 2.11 is in a somewhat surprising contrast with the following result of Kedlaya [Ke] which was also part of Gowers' motivation: $\mathrm{SL}_2(q)$ *has a subset* A *of size* $\ge cq^{5/2}$ *such that* $AA \cap A = \emptyset$ *(where* $c = 1/\sqrt{31}$*).* The question arises whether the exponent $5/2$ is best possible in this result. According to Gowers, $|A| < \sqrt[3]{2}q^{8/3}$ (Theorem 2.14, part (b)).

2.4 Minimum degree of irreducible representations for finite simple groups. The usefulness of our bounds depends heavily on the value of the parameter m. Let m' denote the minimum degree of the nontrivial complex representations of G. Trivially, $m \ge m'$ (in fact either $m = m'$ or $m = 2m'$), so our bounds remain valid with m' in the place of m.

We mentioned the 1896 result of Frobenius [Fr] that for $\mathrm{SL}_2(q)$ we have $m' = (q-1)/2 \sim n^{1/3}/2$. A proof can be found e. g. in [DSV, p. 102].

Landazuri and Seitz [LanS] obtained strong lower bounds for the value of m' for all finite (quasi)simple groups of Lie type; with some slight errors corrected, the results can be found in [KL, p. 188]. Comparing these with the table of the orders of the finite simple groups [KL, p. 170], we can infer lower bounds of the form $m' > cn^\beta$ where c is a positive absolute constant and $\beta > 0$ only depends on to which of the families of Lie type (quasi)simple groups G belongs. For the special linear groups and the unitary groups, $\beta(\mathrm{SL}_r) = \beta(U_r) = 1/(r+1)$; for the symplectic groups, $\beta(Sp_{2r}) = (r-1)/(2r^2+r)$ (the bound is $(2r-1)/(2r^2+r)$ in characteristic 2); for the even dimensional orthogonal groups $\beta(P\Omega^+_{2r}) = (2r-3)/(2r^2-r)$, and for the odd dimensional orthogonal groups $\beta(P\Omega_{2r+1}) = (2r-2)/(2r^2+r)$. This takes care of the six series of "classical simple groups." The remaining 10 families of "exceptional simple groups" all have bounded rank and satisfy the absolute bound $\beta \ge 29/248 \approx 0.1169$.

The remaining finite simple groups are either sporadic or alternating. The "sporadic" finite simple groups, being finite in number, have no asymptotic significance. The alternating groups A_r are notable exceptions on our list: A_r has order $n = r!/2$, yet the corresponding value of m' is $r-1$, so in this case, $m' \sim \log n / \log\log n$. This may be one explanation why the alternating groups do not expand so rapidly.

2.5 Comparison with Gowers' results. The main result of Gowers' paper, [Gow2, Theorem 3.3], can be stated as follows.

THEOREM 2.14. **(Gowers)** *Let* $A, B, C \subseteq G$ *and let* $N = N(A,B,C)$ *denote the number of solutions to the equation* $xy = z$ *where* $x \in A$, $y \in B$, $z \in C$.

(a) If $\alpha > 0$ *and* $|A||B||C| \ge \alpha n^3/m$ *then*

$$\left| N - \frac{|A||B||C|}{n} \right| < \frac{1}{\sqrt{\alpha}} \frac{|A||B||C|}{n}. \tag{2.14}$$

(b) In particular, if $|A||B||C| \ge n^3/m$ *then* $N > 0$, *i. e., a solution exists.*

Statement (b) is equivalent to saying that under the condition $|A||B||C| \ge n^3/m$, we have $|AB| > n - |C|$. In other words, statement (b) is equivalent to the bound

$$|AB| > n - \frac{n^3}{m|A||B|} \tag{2.15}$$

for any pair of nonempty subsets A, B of G. This can be written as

$$|AB| > n(1-\Delta), \tag{2.16}$$

where $\Delta = \dfrac{n^2}{m|A||B|}$. In Corollary 2.5, we state that in fact,

$$|AB| > \frac{n}{1+\Delta}, \tag{2.17}$$

a uniformly stronger bound which gives valuable information even when Δ is large (the sets A, B are relatively small) whereas inequality (2.16) becomes vacuous when $\Delta \ge 1$. To get a sense of the difference, assume $|A| = |B| = k$. Then the condition $\Delta < 1$ means $k > n/\sqrt{m}$. In contrast, we obtain fast product growth as soon as k is large compared to n/m. Indeed, if $k = \alpha n/m$ then inequality (2.17) implies $|AB| \gtrsim \alpha k$ for $k = o(n/\sqrt{m})$ and $|AB| > \min\{n/2, \alpha k/2\}$ always. This difference is critical to our solution to Venkatesh's product growth problem in $\mathrm{SL}_2(p)$ (Problem 2.9). The size of the subsets of $\mathrm{SL}_2(p)$ for which the the question was asked corresponds precisely to the threshold at which the Gowers bound ceases to be useful.

Statement (a) of Gowers' theorem (Theorem 2.14) is identical with the special case $t = 3$ of Corollary 2.8.

It is possible to generalize Gowers' proof directly to obtain Corollary 2.8; this indeed was how we first proved this result. This, however, did not yield the important corollary about 2-term products (Corollary 2.5). So we discarded our initial approach and were led, via a deeper analysis of Gowers' ideas, to our central result, Theorem 2.1, from which we were able derive all the intended results as simple corollaries (including the mentioned generalization of Gowers' uniformity result (part (a) of Theorem 2.14)).

All this told, we are happy to acknowledge the guidance we received from reading some of the fine details of Gowers' paper.

3 Deriving the corollaries

In this section we derive the stated corollaries from the Main Convolution Bound, Theorem 2.1.

The proof of Corollary 2.2 is straightforward induction.

For the proof of Corollary 2.3, we need the following two observations.

OBSERVATION 3.1. *If X is a probability distribution over a set of n elements then*

$$\| X - U \|^2 = \| X \|^2 - \frac{1}{n}. \tag{3.18}$$

OBSERVATION 3.2. *Let W and Z be probability distributions over the finite group G. Then $\| W*Z - U \|_\infty \le \| W - U \| \| Z \|$.*

Proof. Let $h \in G$ and let us choose $w, z \in G$ independently according to the distributions W and Z, respectively. Now, by Cauchy-Schwarz,

$$|P(wz = h) - (1/n)| = \left| \sum_{g \in G} (W(g) - (1/n)) Z(g^{-1}h) \right| \le \| W - U \| \| Z \|. \qquad \blacksquare$$

Proof of Corollary 2.3. Let $W = \prod_{i=1}^{*\,t-1} X_i$ and $Z = X_t$. By Observation 3.2, we have $\| W*Z - U \|_\infty \le \| W - U \| \| Z \|$. Now plug in the bound of Corollary 2.2 for $\| W - U \|$ (changing t to $t-1$). $\blacksquare$

REMARK 3.3. Note that in fact we proved the slightly stronger statement

$$\left\| \prod_{i=1}^{t} {}^* X_i - U \right\|_\infty < \left(\frac{n}{m} \right)^{t/2-1} \| X_t \| \prod_{i=1}^{t-1} \| X_i - U \|. \tag{3.19}$$

For the proof of Corollary 2.4, we need the following observation.

OBSERVATION 3.4. *Let A be a nonempty subset of the finite set Ω and let X be a probability distribution over Ω, concentrated on A. Then $\| X \| \ge 1/\sqrt{|A|}$, and equality holds if and only if X is uniform over A.*

Proof of Corollary 2.4. Let X_i be the uniform probability measure on A_i viewed as a measures on G. So $\| X_i \| = 1/\sqrt{|A_i|}$. Let $\prod_{i=1}^{t} A_i = C$ and $Z = \prod_{i=1}^{*\,t} X_i$; so Z is concentrated on C. By Observation 3.4, we have $|C| \ge \| Z \|^{-2}$. Therefore, by Corollary 2.2 and Observation 3.1 we obtain

$$\begin{aligned} \frac{1}{|C|} \le \| Z \|^2 &= \frac{1}{n} + \| Z - U \|^2 \\ &< \frac{1}{n} + \left(\frac{n}{m} \right)^{t-1} \prod_{i=1}^{t} \| X_i \|^2 \\ &= \frac{1}{n} + \frac{n^{t-1}}{m^{t-1} \prod_{i=1}^{t} |A_i|}. \end{aligned}$$

By rearranging, we obtain the inequality stated in Corollary 2.4. $\blacksquare$

Proof of Corollary 2.5. The first inequality is a special case of Cor. 2.4. The second inequality is obvious. $\blacksquare$

Proof of Corollary 2.6. We show that inequality (2.10) implies that

$$|A_t| + \left| \prod_{i=1}^{t-1} A_i \right| > n. \tag{3.20}$$

This will clearly imply statement (a). To see that it also implies (b), note that for $g \in G$, the statement $g \in \prod_{i=1}^{t} A_i$ is equivalent to saying that $gA_t^{-1} \cap \prod_{i=1}^{t-1} A_i \neq \emptyset$, and this follows from part (a) given that $|gA_t^{-1}| = |A_t|$.

Set $P = \left| \prod_{i=1}^{t-1} A_i \right|$ and $Q = \prod_{i=1}^{t-1} |A_i|$. Set $\Delta = \dfrac{n^{t-1}}{m^{t-2} Q}$. We need to show that $P > n - |A_t|$. By Corollary 2.4 we have

$$P > \frac{n}{1 + \Delta} > n(1 - \Delta) = n - n\Delta, \tag{3.21}$$

so we only need to show $n\Delta \le |A_t|$. Substituting the values of Δ and Q we see that this inequality is exactly the same as the assumption, inequality (2.10). $\blacksquare$

Proof of Corollary 2.7. We use the notation of the proof of Corollary 2.6. Like in that proof, we need to show that $P > n - |A_t|$. By Corollary 2.4, we

have $P > n/(1+\Delta)$, so it suffices to show that $n \geq (1+\Delta)(n-|A_t|)$, or equivalently $(1+1/\Delta)|A_t| \geq n$, i. e., $|A_t| \geq \Delta(n-|A_t|)$. Substituting the values of Δ and Q we see that this inequality is exactly the same as the assumption, inequality (2.12). ■

Proof of Corollary 2.8. We derive this result from Corollary 2.3. As in the proof of Corollary 2.4, let X_i be the uniform distribution over A_i, viewed as a distribution over G. Let $Q = \prod_{i=1}^{t} |A_i|$; so $\prod_{i=1}^{t} \| X_i \| = Q^{-1/2}$. Let $W = \prod_{i=1}^{*\,t} X_i$. Then, for $g \in G$, we have $W(g) = N_g/Q$. Therefore inequality (2.5) asserts that

$$\left| \frac{N_g}{Q} - \frac{1}{n} \right| < \left(\frac{n}{m}\right)^{t/2-1} \frac{1}{\sqrt{Q}}. \tag{3.22}$$

Multiplying by Q and using the definition of α, inequality (2.13) follows. ■

4 Applications

4.1 Diameter of Cayley graphs. In this section we comment on the relation of Helfgott's recent polylogarithmic bound on the diameter of $\mathrm{SL}_2(p)$ and Gowers' result, deduce a general upper bound on the diameter from our main results, and as an application indicate a new worst-case polylogarithmic upper bound on the diameter of a certain class of groups.

For a group G and a set S of generators, the *Cayley graph* $\Gamma(G,S)$ is defined as the directed graph on vertex set G where there is an edge from $g \in G$ to $gs \in G$ for all $s \in S$. This graph is *undirected* if $S = S^{-1}$.

For a group G and a set S of generators, let $\mathrm{diam}(G,S)$ denote the diameter of the (undirected) Cayley graph $\Gamma(G, S \cup S^{-1})$.

The following conjecture appears in [BaSe] (1992).

CONJECTURE 4.1. *For every finite simple group G and every set S of generators,* $\mathrm{diam}(G,S)$ *is polylogarithmic (in terms of $n = |G|$).*

In a recent breakthrough, Helfgott [He] confirmed this conjecture for the groups $(\mathrm{P})\mathrm{SL}_2(p)$. This remains the only class of groups for which the conjecture has been confirmed.

Helfgott's proof proceeds in two phases. The first (more difficult) phase is based on the following remarkable lemma, derived using recent estimates on the sum-product growth of finite fields of prime order [BoKT]. We continue to use the notation $n = |G|$.

THEOREM 4.2. **(Helfgott)** *To every $\epsilon > 0$ there exists $\delta > 0$ such that if S is a set of generators of $G = (\mathrm{P})\mathrm{SL}_2(p)$ and $|S| \leq n^{1-\epsilon}$ then $|S^3| \geq |S|^{1+\delta}$.*

This result implies that for every set S of generators, $|S^\ell| \geq n^{1-\epsilon}$ where ℓ is polylogarithmic (in terms of n): $\ell < (\ln n)^{\ln 3/\delta}$.

The second phase takes us from a subset of size $n^{1-\epsilon}$ to all of G. This phase, admittedly the easier part of Helfgott's argument, can be entirely replaced by Gowers' result: if $|S^\ell| \gtrsim 2n^{8/9}$ then by the product-decomposition version of Theorem 2.14 (the $t = 3$ case of Corollary 2.6) it follows that $\mathrm{diam}(G,S) \leq 3\ell$.

One can generalize this argument, and go even slightly below the n/m threshold.

THEOREM 4.3. *Let G be a finite group and S a set of generators. Suppose $\alpha > 0$ and $|S^\ell| \geq \alpha n/m$ where, as before, m is the smallest dimension of nontrivial real representations of G. Then* $\mathrm{diam}(G,S) \leq t\ell$ *where*

(a) if $\alpha \geq \mathrm{e}$ then $t = \lceil 2 \log m / \log \alpha \rceil$;

(b) if $\alpha < \mathrm{e}$ then $t = \lceil 2\mathrm{e}/\alpha \rceil \cdot \lceil 2 \ln m \rceil$.

Note that part (b) breaks the n/m barrier; we achieve this by appealing to a "noncommutative Cauchy-Davenport theorem," proved by Hamidoune [Ha] and Olson [Ol], for an initial modest but unconditional growth, before a phase of rapid growth starts.

It follows from this result that to prove Conjecture 4.1, it would suffice to prove that there exists a polylogarithmic value ℓ such that $|S^\ell| \geq cn/m$ for any constant $c > 0$ (in fact c could even go to zero at an inverse polylogarithmic rate).

We note that while Helfgott's first phase does not require *inversion* of the generators, the second phase does. The application of Gowers' result eliminates this need. In fact, the proof of Theorem 4.3 does not require inversions, so the bounds stated there remain valid for $\mathrm{diam}^+(G,S)$, defined as the directed diameter of the directed Cayley graph $\Gamma(G,S)$.

It should be mentioned, however, that inversion can always be avoided at the cost of a modest increase in the diameter [Ba].

In the journal version of this paper we give an application of Theorem 4.3 where case (b) is required (we do need to go slightly below the n/m). Combining Helfgott's result with the Gowers idea, we obtain a polylogarithmic bound on the diameter of an interesting extension of $\mathrm{SL}_2(p)$ ("symplectic normalizers," of significance to the structure theory of simple groups of Lie type, cf. [KL]).

4.2 Sylow width. This application, along with its significance, in particular to expander constructions,

has been highlighted in the introduction (Application 1.1). It will constitute a major technical section of the journal version of this paper.

5 Proof of the main result

5.1 Preliminary observations. For a symmetric real matrix M, let $\lambda_1(M) \geq \lambda_2(M) \geq \dots$ denote the eigenvalues.

We call a (not necessarily square) matrix *biregular* if all of its row sums are equal and all of its colum sums are equal. Let $s_r(A)$ denote the sum of each row of a biregular matrix A; and $s_c(A) = s_r(A^T)$ denote the sum of each column. Note that the product of biregular matrices (if defined) is biregular and the quantities s_r and s_c are multiplicative.

The following propositions are easy to verify.

PROPOSITION 5.1. *Let B be a (not necessarily symmetric) nonnegative real $n \times n$ matrix with an all-positive eigenvector corresponding to eigenvalue ρ. Then for all (complex) eigenvalues λ of B we have $\rho \geq |\lambda|$.*

Let $\mathbf{1}_n \in \mathbb{R}^n$ denote the all-ones vector $\mathbf{1}_n = (1, \dots, 1)^T$.

PROPOSITION 5.2. *Let A be a nonnegative biregular $k \times n$ matrix with row sum s_r and column sum s_c. Then $\lambda_1(A^T A) = s_r s_c$ and a corresponding eigenvector is $\mathbf{1}_n$.*

PROPOSITION 5.3. *Let A be a $k \times n$ biregular matrix. If $\mathbf{u} \in \mathbb{R}^n$ and $\mathbf{u} \perp \mathbf{1}_n$ then $A\mathbf{u} \perp \mathbf{1}_k$.*

The next lemma was an ingredient in our original proof of the t-step extension of Gowers' result (Corollaries 2.6 and 2.8). We still need its first (trivial) part; the second part is stated as a hint for the alternative proof.

LEMMA 5.4. *Let $A_1, \dots, A_\ell$ be nonnegative biregular matrices such that the product $A = A_1 \dots A_\ell$ is defined. Then*

$$\lambda_1(A^T A) = \prod_{i=1}^{\ell} \lambda_1(A_i^T A_i) \tag{5.23}$$

and

$$\lambda_2(A^T A) \leq \prod_{i=1}^{\ell} \lambda_2(A_i^T A_i). \tag{5.24}$$

Proof. A is nonnegative biregular and $s_r(A) = \prod_{i=1}^{\ell} s_r(A_i)$ and $s_c(A) = \prod_{i=1}^{\ell} s_c(A_i)$. Combining these with Proposition 5.2 yields equation (5.23).

Inequality (5.24) is not needed in this paper so we omit its proof. ∎

5.2 Time-inhomogeneous Cayley walks and G-circulants. A *probability distribution* over the set Ω of $n \geq 1$ "states" is a function $X : \Omega \to \mathbb{R}$ such that $X(i) \geq 0$ for all $i \in \Omega$ and $\sum_{i \in \Omega} X(i) = 1$.

An $n \times n$ real matrix labeled by $\Omega \times \Omega$ is *stochastic* if its rows are probability distributions over Ω. Such a matrix can be viewed as the transition matrix for a random step on the state space. A *time-inhomogeneous Markov Chain* on Ω is described by a sequence $B_1, B_2, \dots$ of stochastic matrices; the transition at time i is governed by B_i. So the combined transition probabilities from time i to j are described by the product $\prod_{k=i}^{j-1} B_i$. This matrix is also stochastic.

Let us say that a matrix is a *permutant* if every row is a permutation of the first row. Let the first row of a permutant B be X; then obviously

$$\| X \|^2 = \frac{1}{n} \operatorname{Tr}(B^T B). \tag{5.25}$$

Let $\Omega = G$ be a finite group. A matrix B, labeled by $G \times G$, is a *G-circulant* if for every $g, h \in G$ we have $B(g, h) = B(1, g^{-1}h)$. We note that (a) a G-circulant is a permutant; (b) the transpose of a G-circulant is again a G-circulant; (c) the product of G-circulants is a G-circulant; (d) a G-circulant is biregular; (e) a G-circulant is determined ("generated") by its first row. If X is the first row then $B(g, h) = X(g^{-1}h)$.

A *time-inhomogeneous Cayley walk* on G is a time-inhomogeneous Markov Chain described by a sequence of stochastic G-circulants $B_1, B_2, \dots$ generated by the probability distributions $X_1, X_2, \dots$. The *convolution* $X * Y$ of two distributions X and Y over G is defined as $(X * Y)(h) = \sum_{g \in G} X(g) Y(g^{-1}h)$.

OBSERVATION 5.5. *If A and B are the G-circulants generated by X and Y, respectively, then the G-circulant generated by $X * Y$ is AB.*

Combining this observation with equation (5.25) we obtain the following identity.

PROPOSITION 5.6. *Let X, Y be two distributions over G and let A and B be the corresponding stochastic G-circulants. Then*

$$\| X * Y \|^2 = \frac{1}{n} \operatorname{Tr}(B^T A^T A B). \tag{5.26}$$

5.3 Symmetry implies degeneracy. If A, B are two commuting $n \times n$ matrices then it is easy to see that the eigensubspaces of A are invariant under B. Therefore if A commutes with a group of matrices B_g ($g \in G$) which constitute a representation of the group G (i. e., $B_{g^{-1}h} = B_g^{-1} B_h$), then each eigensubspace

U_λ of A is invariant under the B_g and therefore it contains a G-irreducible subspace. Let m denote the minimum dimension of irreducible representations of G over the appropriate field (containing the eigenvalue λ) and let m_λ be the multiplicity of the eigenvalue λ. We conclude that unless the G-action on U_λ is trivial, we have

$$m_\lambda \geq m. \tag{5.27}$$

This simple observation was made by physicist Eugene Wigner around 1930 [Wig] who used it to great advantage in classifying the eigenvibrations of physical systems according to the irreducible representations of their group of symmetries. ("Multiplicity" is regarded "degeneracy" by physicists.)

G-circulants are linear combinations of the permutation matrices describing the right translations $R_g : h \mapsto hg$ and therefore they commute with the matrices $L_g : h \mapsto g^{-1}h$ of left translations. The matrices L_g constitute the *left regular representation* of G. We can now state the key tool that connects eigenvalue estimation to representation theory of finite groups.

LEMMA 5.7. *Let m be the minimum degree of nontrivial real representations of G. If A is a nonnegative G-circulant then*

$$\lambda_2(A^TA) \leq \frac{\mathrm{Tr}(A^TA) - \lambda_1(A^TA)}{m}. \tag{5.28}$$

Proof. Write $C = A^TA$. Then C is positive semidefinite, so $\lambda_1(C) \geq \lambda_2(C) \geq \cdots \geq \lambda_n(C) \geq 0$. Since C is stochastic, the all-ones vector $\mathbf{e}$ is an eigenvector. Moreover, it corresponds to $\lambda_1(C)$ according to Proposition 5.1 (since C is nonnegative).

Let U denote the orthogonal complement of $\mathbf{e}$ in the real euclidean space $\mathbb{R}^G$ (under the standard dot product). This subspace is then invariant both under C and under the left regular representation of G. Moreover, no nonzero vector in U is fixed under the left regular G-action (because the sum of coordinates of each vector in U is zero and G permutes the coordinates transitively). It follows by inequality (5.27) that the multiplicity of every eigenvalue of the restriction of C to U is $\geq m$; this in particular applies to the multiplicity of λ_2 on U. But the trace of C restricted to U is $\mathrm{Tr}(C) - \lambda_1(C)$, hence this quantity is $\geq m\lambda_2(C)$. ■

Note that this proof works even when $\lambda_2 = \lambda_1$; we did not use any irreducibility assumption on C.

5.4 The convolution bound. In this section we conclude the proof of our main result (Theorem 2.1). We need the following linear algebra lemma.

LEMMA 5.8. *Let A and B be nonnegative biregular matrices such that the product AB is defined. Then*

$$\mathrm{Tr}(B^TA^TAB) \leq \lambda_1(A^TA)\lambda_1(B^TB) + \lambda_2(A^TA)\left(\mathrm{Tr}(B^TB) - \lambda_1(B^TB)\right).$$

Proof. Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be an orthonormal eigenbasis for B^TB. Assuming B is $k \times n$, by Proposition 5.2 we can take $\mathbf{e}_1 = \mathbf{1}_n/\sqrt{n}$. We have

$$\| B\mathbf{e}_i \|^2 = \mathbf{e}_i^T B^T B \mathbf{e}_i = \lambda_i(B^TB). \tag{5.29}$$

Let $C = B^TA^TAB = (AB)^T(AB)$. Then

$$\mathrm{Tr}(C) = \sum_{i=1}^n \mathbf{e}_i^T C \mathbf{e}_i = \sum_{i=1}^n \| AB\mathbf{e}_i \|^2. \tag{5.30}$$

According to equation (5.23) in Lemma 5.4 we have

$$\| AB\mathbf{e}_1 \|^2 = \lambda_1(C) = \lambda_1(B^TB)\lambda_1(A^TA). \tag{5.31}$$

We need to prove

$$\sum_{i=2}^n \| AB\mathbf{e}_i \|^2 \leq \lambda_2(A^TA)\left(\mathrm{Tr}(B^TB) - \lambda_1(B^TB)\right). \tag{5.32}$$

But for $i \geq 2$, $B\mathbf{e}_i$ is perpendicular to $\mathbf{1}_k$ (by Proposition 5.3) and therefore, for $i \geq 2$,

$$\| AB\mathbf{e}_i \|^2 \leq \lambda_2(A^TA) \| B\mathbf{e}_i \|^2 \tag{5.33}$$

by the Rayleigh principle. We conclude that

$$\sum_{i=2}^n \| AB\mathbf{e}_i \|^2 \leq \lambda_2(A^TA) \sum_{i=2}^n \| B\mathbf{e}_i \|^2 = \lambda_2(A^TA)\left(\mathrm{Tr}(B^TB) - \lambda_1(B^TB)\right).$$

This completes the proof. ■

Proof of Theorem 2.1. Let A and B be the stochastic G-circulants generated by the distributions X and Y, respectively. Combining Proposition 5.6 and Lemma 5.8 we obtain

$$\| X * Y \|^2 \leq \frac{1}{n}\lambda_1(A^TA)\lambda_1(B^TB) + \frac{1}{n}\lambda_2(A^TA)\left(\mathrm{Tr}(B^TB) - \lambda_1(B^TB)\right).$$

Now $\lambda_1(A^TA) = \lambda_1(B^TB) = 1$ (because both matrices are stochastic), therefore the first term on the right-hand side is $1/n$. By Observation 3.1 and Lemma 5.7

we obtain

$$\begin{aligned}\| X * Y - U \|^2 &\le \frac{1}{n}\lambda_2(A^TA)\left(\mathrm{Tr}(B^TB)-1\right)\\ &\le \frac{1}{nm}(\mathrm{Tr}(A^TA)-1)(\mathrm{Tr}(B^TB)-1)\\ &= \frac{n}{m}\| X-U\|^2\| Y-U\|^2. \qquad \blacksquare\end{aligned}$$

References

[APPS] M. Abért, P. P. Pálfy, L. Pyber, B. Szegedy: Groups of finite abelian or solvable width. *In preparation*

[AS] N. Alon, J. H. Spencer: *The Probabilistic Method.* Wiley, 1992.

[Ba] L. Babai: On the diameter of Eulerian orientations of a graph. In: *Proc. 17th Ann. Symp. on Discr. Alg. (SODA'06)*, ACM–SIAM 2006, pp. 822-831.

[BaSe] L. Babai, Á. Seress: On the diameter of permutation groups. *Europ. J. Comb.* **13** (1992), 231–243.

[BaSo] L. Babai, V. T. Sós: Sidon sets in groups and induced subgraphs of Cayley graphs. *Europ. J. Combinatorcs* **6** (1985), 101–114.

[BoG] J. Bourgain, A. Gamburd: Uniform expansion bound for Cayley graphs of $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$. *Ann. Math.* to appear.

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

[Ca] R. Carter: *Simple groups of Lie type.* Pure and Applied Mathematics, Vol. 28. John Wiley & Sons, London-New York-Sydney, 1972.

[CGW] F. R. K. Chung, R. L. Graham, R. M. Wilson: Quasi-random graphs. *Combinatorica* **9** (1989), 345–362.

[CL] Problems presented at the workshop on Recent Trends in Additive Combinatorics, collected by E. Croot and V. F. Lev. American Institute of Mathematics, Palo Alto, 2004. `http://www.aimath.org/WWW/additivecomb/additivecomb.pdf`

[DSV] G. Davidoff, P. Sarnak, A. Valette: Elementary Number Theory, Group Theory, and Ramanujan Graphs. Cambridge U. Press, 2003.

[Fr] G. Frobenius: Über Gruppencharactere. *Sitzungsber. Königl. Preuß. Akad. Wiss. Berlin* (1896), 985–1021.

[Gow1] W. T. Gowers: Quasirandomness, counting, and regularity for 3-uniform hypergraphs. *Combinatorics, Probability, and Computing* **15** (2006), 143–184.

[Gow2] W. T. Gowers: Quasirandom groups. Manuscript, 2006.

[Ha] Y. O. Hamidoune: An application of connectivity theory in graphs to factorizations of elements in groups. *Eur. J. Comb.* **2** (1981) 349–355.

[He] H. A. Helfgott: Growth and generation in $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$. *Ann. Math.* to appear.

[HruP] E. Hrushovski, A. Pillay: Definable subgroups of algebraic groups over finite fields. *J. Reine Angew. Math.* 462 (1995), 69–91.

[KLN] M. Kassabov, A. Lubotzky, N. Nikolov: Finite simple groups as expanders. *Proc. Nat. Acad. USA* **103** (2006), 6116–6119.

[Ke] K. S. Kedlaya: Large product-free subsets of finite groups. *J. Combin. Theory - A* **77** (1997), 339–343.

[LanS] V. Landazuri, G. Seitz: On the minimal degrees of projective representations of the finite Chevalley groups. *J. Algebra* **32** (1974), 418-443.

[LP] M. W. Liebeck, L. Pyber: Finite linear groups and bounded generation. *Duke Math. J.* **107** (2001), 159–171.

[LuPS] A. Lubotzky, R. Phillips, P. Sarnak: Ramanujan graphs. *Combinatorica* **8** (1988), 261–277.

[KL] P. Kleidman, M. Liebeck: *The subgroup structure of the finite classical groups.* LMS Lecture Notes 129, Cambridge Univ. Press 1990.

[Ma] G. Margulis: Explicit group-theoretical constructions of combinatorial schemes and their applications to the design of expanders and concentrators. *J. Problems Info. Transmission* **24/1** (1988), 39–46.

[Ni] N. Nikolov: A product decomposition for classical quasisimple groups. *J. Group Th.* **10** (2007), 43–53.

[NP] N. Nikolov, L. Pyber: Product decomposition of quasirandom groups and a Jordan type theorem. Manuscript. arXiv:math/0703343 (March 2007)

[NS] N. Nikolov, D. Segal: On finitely generated profinite groups II. Products in quasisimple groups. *Ann. Math.* **165** (2007), 239–273.

[Ol] J. E. Olson: On the sum of two sets in a group. *J. Number Theory* **18** (1984), 110-120.

[SarX] P. Sarnak, X. Xue: Bounds for multiplicities of automorphic representations. *Duke Math. J.* **64** (1991), 207-227.

[Shv] A. Shalev: Word maps, conjugacy classes, and a non-commutative Waring-type theorem. *Ann. Math.* to appear

[Shm] Y. Shalom: Bounded generation and Kazhdan's property (T). *Publ. Math. IHES* **90** (1999/2001) 145–168.

[TaV] T. Tao, Van H. Vu: *Additive Combinatorics.* Cambridge University Press, 2006.

[Th] A. Thomason: Pseudo-random graphs. In: *Proc. Conf. on Random Graphs, Poznań 1985.* (M. Karoński, ed.) *Annals of Discrete Math.* **33** (North Holland 1987), 307–331.

[Wig] E. P. Wigner: Über die elastischen Eigenschwingungen symmetrischer Systeme. *Nachr. der Ges. der Wiss. zu Göttingen. Math-Phys. Kl.* (1930) 133–146.

[Wil] J. S. Wilson: On simple pseudofinite groups. *J. London Math. Soc.* **51** (1995), 471–490.