Random Walks in Compact Groups

Péter Pál Varjú[1]

Received: December 17, 2012

Communicated by Ursula Hamenstädt

Abstract. Let $X_1, X_2, \dots$ be independent identically distributed random elements of a compact group G. We discuss the speed of convergence of the law of the product $X_l \cdots X_1$ to the Haar measure. We give poly-log estimates for certain finite groups and for compact semi-simple Lie groups. We improve earlier results of Solovay, Kitaev, Gamburd, Shahshahani and Dinai.

2010 Mathematics Subject Classification: 60B15, 22E30, 05E15
Keywords and Phrases: random walk, spectral gap, diameter, poly-log, Solovay-Kitaev, compact group, Cayley graph

1 Introduction

Let G be a group and $S \subset G$ a finite set. We study the distribution of the product of l random elements of S. In particular, we are interested in how fast this distribution becomes uniform as l grows.
We discuss the problem in two different but very related settings: profinite groups, and compact Lie groups.

1.1 Two ways to measure uniformity

We begin by describing the details in the first setting. Let G be a finite group and $S \subset G$ be a finite generating set. For simplicity, we assume that $1 \in S$. The unit element of any group is denoted by 1 in this paper. Write

$$S^l = \{g_1 \cdots g_l : g_1, \dots, g_l \in S\}$$

for the l-fold product set of S.
The diameter of G with respect to S is defined by

$$\mathrm{diam}(G, S) = \min\{l : S^l = G\}.$$

[1] I gratefully acknowledge the support of the Simons Foundation and the European Research Council (Advanced Research Grant 267259)

The diameter is the minimal length of a product that can express any element of the group. Hence it is a (very weak) quantity to measure uniformity.
We quantify uniformity in a stronger sense, too. To this end, we introduce the notion of random walks. Denote by μ_S the normalized counting measure on the set S, and let $X_1, X_2, \ldots$ be a sequence of independent random elements of S with law μ_S. The (simple) random walk is the sequence of random elements $Y_0, Y_1, \ldots$ of G such that $Y_0 = 1$ almost surely, and $Y_{l+1} = X_{l+1} Y_l$ for all $l \geq 0$. Denote by $\mu_S^{*(l)}$ the l-fold convolution of the measure μ_S with itself. Convolution of two measures (or functions) μ, ν is defined by the usual formula

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

Observe that $\mu_S^{*(l)}$ is the law of Y_l.
We want to understand, how large l is needed to be taken such that $\mu_S^{*(l)}$ is "very close" to the uniform distribution. We make this precise with the following construction. Consider the space $L^2(G)$ which is simply the vector space of complex valued functions on G endowed with the standard scalar product. The group G acts on this by

$$\mathrm{Reg}_G(g) f(h) = f(g^{-1}h) \quad \text{for} \quad f \in L^2(G).$$

This is a unitary representation called the regular representation.
To a measure μ, we associate an operator (linear transformation) on $L^2(G)$:

$$\mathrm{Reg}_G(\mu) = \sum_{g \in G} \mu(g) \cdot \mathrm{Reg}_G(g).$$

This is an analogue of the Fourier transform of classical harmonic analysis. In particular, it has the following property:

$$\mathrm{Reg}_G(\mu * \nu) = \mathrm{Reg}_G(\mu) \cdot \mathrm{Reg}_G(\nu),$$

which is easy to verify from the definitions. We can recover μ by the formula

$$\mu = \mathrm{Reg}_G(\mu)\delta_1, \tag{1}$$

where δ_1 is the Dirac measure supported at 1. (Recall that G is finite, hence we can embed the space of probability measures into L^2.)
The operator $\mathrm{Reg}_G(\mu_S)$ is of norm 1 and it acts trivially on the one dimensional space of constant functions. We denote by Reg_G° the restriction of Reg_G to the 1 codimensional space orthogonal to constants. We define by

$$\mathrm{gap}(G, S) := 1 - \|\mathrm{Reg}_G^\circ(\mu_S)\|$$

the spectral gap of the random walk on G generated by S. Later, we will consider a slightly more general situation and replace μ_S by an arbitrary probability measure μ. Then we write

$$\mathrm{gap}(G, \mu) := 1 - \|\mathrm{Reg}_G^\circ(\mu)\|.$$

It is clear that

$$\left\| \mathrm{Reg}_G\left(\mu_S^{*(l)}\right) - \mathrm{Reg}_G(\mu_G) \right\| < e^{-l\cdot \mathrm{gap}(G,S)}. \tag{2}$$

This shows that if we take say $l = 10 \log |G| / \mathrm{gap}(G, S)$, then $\mu_S^{*(l)}$ is very close to the uniform distribution (see (1)). In particular, the support of $\mu_S^{*(l)}$ is the whole group. Thus spectral gap is a stronger quantity to measure uniformity than diameter. Somewhat surprisingly we can obtain a bound in the other direction, as well.

Lemma 1. *Let G be an arbitrary finite group, and $1 \in S \subset G$. Suppose that S is symmetric, i.e. $g \in S$ if and only if $g^{-1} \in S$. We have:*

$$(\mathrm{diam}(G, S) - 1)/ \log |G| \leq \mathrm{gap}(G, S)^{-1} \leq |S| \mathrm{diam}(G, S)^2.$$

This lemma is well-known. The second (and more difficult) inequality of it can be found for example in [13, Corollary 1 in Section 3]. The other estimate follows directly from (2). The assumption on symmetricity is not an essential one. Later, at the beginning of Section 2.1 we show how to reduce the problem to the symmetric case.

We stop for a moment to connect our terminology to the computer science and combinatorics literature. If S is symmetric, then the matrix of the operator $\mathrm{Reg}_G(\mu_S)$ is proportional to the adjacency matrix of the Cayley graph of G with respect to the generating set S. In that case $\mathrm{gap}(G, S)$ is proportional to the spectral gap of the Cayley graph. If $\mathrm{gap}(G, S) \geq c > 0$ for a family of groups and generators than the corresponding family of Cayley graphs are called expanders.

However, in this paper we are looking for weaker bounds of the form $\mathrm{gap}(G, S) \geq \log^{-A} |G|$ which, in light of the above Lemma, is equivalent to $\mathrm{diam}(G, S) \leq \log^{A'} |G|$ as long as say $|S| < 10$ and one does not care about the value of A and A'. We call such bounds poly-logarithmic.

1.2 Prior results

It was conjectured by Babai and Seress [1, Conjecture 1.7] that the family of non-Abelian finite simple groups have poly-logarithmic diameter, i.e. there is a constant A such that $\mathrm{diam}(G, S) \leq \log^A |G|$ holds for any non-Abelian finite simple group G and generating set $S \subset G$. This has been verified for the family $SL_2(\mathbf{F}_p)$ by Helfgott [22, Main Theorem] and for finite simple groups of Lie type of bounded rank by Breuillard, Green and Tao [11, Theorem 7.1] and Pyber and Szabó [30, Theorem 2] independently. The conjecture is still open for other families of finite simple groups. The best known bound to date on the diameter of alternating groups is due to Helfgott and Seress [24] and it is slightly weaker than poly-logarithmic.

However, the first results on poly-logarithmic diameter were obtained for non-simple groups. Fix a prime p and a symmetric set $S \subset \mathrm{SL}_2(\mathbf{Z})$ such that (the projection of) S generates $\mathrm{SL}_2(\mathbf{Z}/p^2\mathbf{Z})$. Gamburd and Shahshahani [20, Theorem 2.1] proved the poly-log diameter estimate

$$\mathrm{diam}(\mathrm{SL}_2(\mathbf{Z}/p^n\mathbf{Z}), S) \le C \log^A |\mathrm{SL}_2(\mathbf{Z}/p^n\mathbf{Z})|,$$

where C depends only on S and A is an absolute constant. Dinai [15, Theorem 1.2] observed that the result holds with C depending only on p and he also improved the parameter A. Thus the family $\mathrm{SL}_2(\mathbf{Z}/p^n\mathbf{Z})$ enjoys a uniform poly-log diameter bound with respect to arbitrary generators. Using the result of Helfgott [22], the constant C can be made absolute. In [16, Theorem 1.1], Dinai extended the result to the quotients of other Chevalley groups over local rings.

The result of this paper is also about non-simple finite groups. In fact, it is part of our assumptions that all simple quotients of the groups we study has poly-logarithmic diameter. Our result has a huge overlap with [20], [15] and [16]. We will remark on this later.

1.3 The role of quasirandomness

Our approach is based on representation theory, but we will use only very basic facts of the theory. We explain the key idea of the paper, which allows us to estimate the spectral gap based on lower bounds for the dimension of nontrivial representations of the group. Let μ be a probability measure on G. Write χ_G and χ_G° respectively for the characters of Reg_G and Reg_G°. Furthermore, we write

$$\chi(\mu) = \sum_{g\in G} \mu(g) \cdot \chi(g),$$

where χ is the character of a representation of G. Then $\chi_G(\mu)$ is the trace of the operator $\mathrm{Reg}_G(\mu)$.

For the moment, assume that μ is symmetric that is $\mu(g^{-1}) = \mu(g)$ for all $g \in G$. Then $\mathrm{Reg}_G(\mu)$ is selfadjoint. We can decompose $L^2(G)$ as the orthogonal sum of irreducible subrepresentations of Reg_G. As is well known, if π is an irreducible representation of G, then exactly $\dim \pi$ isomorphic copies of π appear in the decomposition of Reg_G. If λ is an eigenvalue of $\mathrm{Reg}_G(\mu)$ then there is a corresponding eigenvector in one of the irreducible representations of G. Moreover, there is one in each isomorphic copy. This leads us to the following inequality which is fundamental to us:

$$\dim(\pi) \cdot \lambda^2 \le \chi_G(\mu * \mu), \tag{3}$$

where π is an irreducible representation of G such that there is an eigenvector corresponding to λ in π. (Note that all eigenvalues of $\mathrm{Reg}_G(\mu * \mu)$ are non-negative.)

This idea goes back to Sarnak and Xue [32] and it is also one of the major steps in the work of Bourgain and Gamburd [2] on estimating spectral gaps and also in several papers [4], [5], [7], [34], [9], [31] which follow it. Gowers [21] also exploited the idea, and introduced the term quasirandom for groups that does not have low dimensional non-trivial representations. He proved several properties of such groups, in particular that they do not have large product-free subsets. Nikolov and Pyber [29, Corollary 1] pointed out that Gowers's result implies that any element of a quasirandom group can be expressed as the product of three elements of a sufficiently large subset.

In what follows, G is a profinite group and Ω denotes the family of finite index normal subgroups of it. An interesting example to have in mind is $G = \mathrm{SL}_d(\widehat{\mathbf{Z}})$, where $\widehat{\mathbf{Z}}$ is the profinite (congruence) completion of the integers. (The reader unfamiliar with the notion of profinite groups may assume that G is finite without loss of generality.) Inspired by Gowers's terminology, we make the following definition.

Definition 2. *We say that a profinite group G is (c,α)-quasirandom if for every irreducible unitary representation π of G, we have*

$$\dim \pi \geq c[G : \mathrm{Ker}(\pi)]^{\alpha}.$$

1.4 Results about profinite groups

Let G be a profinite group and Γ a finite index normal subgroup. Our plan is to prove the estimate

$$\mathrm{gap}(G/\Gamma, S) \geq c \log^{-A}[G : \Gamma] \tag{4}$$

with constants c, A independent of Γ and S. We prove this statement by induction as follows: We find a larger subgroup $\Gamma \mathrel{\underset{\neq}{\lhd}} \Gamma' \in \Omega$ and assume that the above spectral gap estimate holds for Γ'. We use this to bound the trace of the operator $\mathrm{Reg}_{G/\Gamma'}(\mu_S^{*(l)})$ for a suitable integer l. This in turn gives an estimate for the trace of $\mathrm{Reg}_{G/\Gamma}(\mu_S^{*(l)})$, and by (3) we can estimate $\mathrm{gap}(G/\Gamma, S)$ and continue by induction.

In order that the induction step works, we need to ensure that $[\Gamma' : \Gamma]$ is "not very large" compared to $[G : \Gamma]$. Of course, we cannot always find a suitable Γ', in particular when G/Γ is simple.

The statement of the following theorem is very technical, so we first explain it informally: We suppose that G is a quasirandom profinite group and partition its finite index normal subgroups into two sets: $\Omega_1 \dot{\cup} \Omega_2$. We assume that for $\Gamma \in \Omega_1$ we can find a larger subgroup $\Gamma' \in \Omega$ so that $[\Gamma' : \Gamma]$ is "not very large". We assume further that for $\Gamma \in \Omega_2$, the quotient G/Γ has poly-log spectral gap (i.e. (4) is satisfied). Then we can conclude a poly-log estimate (4) for all $\Gamma \in \Omega$.

Theorem 3. *Let c_1, α, β and A be positive numbers satisfying $\beta < \alpha$ and*

$$A \geq (\alpha - \beta)^{-1} - 1.$$

Then there is a positive number C_1 depending only on c_1, α, β and A such that the following holds.

Let G be a (c_1, α)-quasirandom profinite group. Let $\Omega_1 \dot\cup \Omega_2 = \Omega$ be a partition of the family of finite index normal subgroups of G.
Suppose that for all $\Gamma \in \Omega_1$, we have $[G : \Gamma] > C_1$ and there is $\Gamma' \in \Omega$ with $\Gamma \underset{\neq}{\trianglelefteq} \Gamma'$ such that

$$[\Gamma' : \Gamma] \leq [G : \Gamma']^\beta. \tag{5}$$

Let μ be a Borel probability measure on G and suppose that

$$\operatorname{gap}(G/\Gamma, \mu) \geq B \cdot \log^{-A}[G : \Gamma] \tag{6}$$

for all $\Gamma \in \Omega_2$.
Then for all $\Gamma \in \Omega$

$$\operatorname{gap}(G/\Gamma, \mu) \geq C_1^{-1} B \cdot \log^{-A}[G : \Gamma].$$

To verify assumption (6), we need to use known cases of the Babai–Seress conjecture mentioned above. Fortunately, this is available for many cases in the papers [22, Main Theorem], [23, Corollary 1.1], [11, Theorem 7.1], [30, Theorem 2]. In particular we can deduce the following.

Corollary 4. *Let $S \subset \mathrm{SL}_d(\widehat{\mathbf{Z}})$ be a set generating a dense subgroup. Then for all integers $q > 1$, we have*

$$\operatorname{gap}(\mathrm{SL}_d(\mathbf{Z}/q\mathbf{Z}), S) \geq c \log^{-A}(q),$$

where $A > 0$ is a number depending on d and c is a number depending on d and $|S|$.

We stress here that c depends only on the cardinality of S. Note that this dependence is necessary, owing to the following example: It may happen that all but one element of S projects into a proper subgroup of $\mathrm{SL}_d(\mathbf{Z}/q\mathbf{Z})$ in which case the characteristic function of the subgroup is an almost invariant function if $|S|$ is large.
In the case $S \subset \mathrm{SL}_d(\mathbf{Z})$ one can improve the poly-log bound to a uniform one. More precisely, Bourgain and Varjú [9, Theorem 1] proved $\operatorname{gap}(\mathrm{SL}_d(\mathbf{Z}/q\mathbf{Z}), S) \geq c$ with a constant $c > 0$ independent of q but dependent on S. This strongly resonates with the sate of affairs for compact semi-simple Lie groups to be discussed below.
Finally, we compare our result to the papers of Gamburd and Shahshahani [20, Theorem 2.1] and Dinai [15, Theorem 1.2], [16, Theorem 1.1]. These papers give results similar to Corollary 4, and even more, the proofs have some common features with our approach. However, they obtain the diameter bound directly and they use properties of commutators instead of representation theory. A flaw of our approach that it is not constructive, i.e. we can not give an efficient algorithm to express an element of G/Γ as a product of elements of S. On the contrary, [20, Theorem 2.1], [15, Theorem 1.2], [16, Theorem 1.1] give such algorithms. The advantage of our paper that it seems to apply in more general

situations, e.g. [20, Theorem 2.1], [15, Theorem 1.2], [16, Theorem 1.1] are restricted to the case when q is the power of a prime.

A comment on the exponents: If one considers the group $G = \mathrm{SL}_2(\mathbf{Z}_p)$, where p is a fixed prime, then the conditions of Theorem 3 can be satisfied for any $A > 2$. This via Lemma 1 exactly recovers the diameter bound for $\mathrm{SL}_2(\mathbf{Z}/p^k\mathbf{Z})$ in the paper [15, Theorem 1.2], but our bound for the spectral gap is better than what could be obtained from [15]. If one considers SL_d with d large than our bounds deteriorate compared to [15]. However, it seems possible that a more careful version of our argument could give better bounds, but this requires a more precise understanding of the representations. In Section 2.2 we include some remarks about what this would require. These ideas are worked out in the setting of compact Lie groups.

1.5 Results about compact Lie groups

We turn to the second setting of our paper. Let G be a semi-simple compact Lie group endowed with the bi-invariant Riemannian metric. We denote by $\mathrm{dist}(g, h)$ the distance of two elements $g, h \in G$. Let $\varepsilon > 0$ be a number and $S \subset G$ be a finite subset which generates a dense subgroup. Again, for simplicity, we assume $1 \in S$. We define the diameter of G at scale ε with respect to S by

$$\mathrm{diam}_\varepsilon(G, S) = \min\{l : \text{ for every } g \in G \text{ there is } h \in S^l \text{ such that } \mathrm{dist}(g, h) < \varepsilon\}.$$

We also introduce the relevant spectral gap notion. This requires some basic facts about the representation theory of compact Lie groups. We follow the notation in [12], but the results we need can be found in many of the textbooks on the subject, as well.

Let T be a maximal torus in G and denote by $\mathrm{L}T$ its tangent space at 1. Then T can be identified with $\mathrm{L}T/I$ via the exponential map, where I is a lattice in $\mathrm{L}T$. Denote by $\mathrm{L}T^*$ the dual of $\mathrm{L}T$ and by $I^* \subset \mathrm{L}T^*$ the lattice dual to I. We denote by $R \subset I^*$ the set of roots, and by R_+ (a choice of) the positive roots. We fix an inner product $\langle \cdot, \cdot \rangle$ on $\mathrm{L}T$ which is invariant under the Weyl group. Denote by

$$K = \{u \in \mathrm{L}T : \langle u, v \rangle > 0 \text{ for every } v \in R_+\}$$

the positive Weyl chamber and by $\overline{K}$ its closure.

It is well known (see [12, Chapter IV (1.7)]) that the irreducible representations of G can be parametrized by the elements of $I^* \cap \overline{K}$. For $v \in I^* \cap \overline{K}$, we denote by π_v the unitary representation of G with highest weight v.

For a finite Borel measure μ on G, we write

$$\pi_v(\mu) = \int \pi_v(g)\, d\mu(g),$$

which is an operator (linear transformation) on the representation space of π_v.

Let $r > 1$ be a number, and $S \subset G$ a finite set which contains 1 and generates a dense subgroup. Define the spectral gap at scale $1/r$ with respect to S by

$$\mathrm{gap}_r(G,S) := 1 - \max_{0<|v|\le r} \|\pi_v(\mu_S)\|.$$

As in the finite case, the notions of spectral gap and diameter are closely related:

Lemma 5. *Let G be a compact connected semi-simple Lie group, and let $1 \in S \subset G$ be finite. Then there is a constant $C > 0$ depending only on G such that for any $\varepsilon > 0$*

$$\mathrm{diam}_\varepsilon(G,S) \le \frac{C\log(\varepsilon^{-1})}{\mathrm{gap}_{C\varepsilon^{-C}}(G,S)}$$

and for any $r \ge 1$

$$\mathrm{gap}_r(G,S) \ge \frac{1}{|S|\mathrm{diam}_{(Cr)^{-C}}(G,S)^2}.$$

This lemma is also well-known. We give a proof in Section 4 for completeness. In Section 3, we develop an analogue for compact Lie groups of the ideas explained in the previous section. A replacement for (3) was given by Gamburd, Jacobson and Sarnak [19] and it appeared in various forms in [3] and [6], as well. However, these are based on direct calculation with characters rather than on multiplicities of eigenvalues. A more direct analogue of (3) and also of the results of Gowers [21] and Nikolov and Pyber [29] was developed very recently by Saxcé [33].
Our main result in the setting of compact Lie groups is the following:

Theorem 6. *For every semi-simple compact connected Lie group G, there are numbers c, r_0 and A such that the following holds. Let μ be an arbitrary probability measure on G. Then*

$$\mathrm{gap}_r(G,\mu) \ge c \cdot \mathrm{gap}_{r_0}(G,\mu) \cdot \log^{-A} r.$$

For simple groups, the value of A can be found in Table 1. For semi-simple groups A is the maximum of the corresponding values over all simple quotients of G. In particular, $A \le 2$ for all groups.

A_n	B_n	C_n	D_n	E_6	E_7	E_8	F_4	G_2
$1+\frac{2}{n+1}$	$1+\frac{1}{n}$	$1+\frac{1}{n}$	$1+\frac{1}{n-1}$	$\frac{7}{6}$	$\frac{10}{9}$	$\frac{16}{15}$	$\frac{7}{6}$	$\frac{4}{3}$

Table 1: The value of $A(G)$ in terms of the Dynkin diagram

A poly-logarithmic estimate for $\mathrm{diam}_\varepsilon(\mathrm{SU}(2),S)$ was found by Solovay and Kitaev independently. Nice expositions are in the paper of Dawson and Nielsen [14, Theorem 1] and in the book [26, Chapter 8.3], where they obtain the bound $\mathrm{diam}_\varepsilon(\mathrm{SU}(2),S) < \log^3 \varepsilon^{-1}$. Our theorem provides the same bound for the diameter of SU(2). On the other hand, our spectral gap bound in Theorem 6 beats anything that could be obtained from a diameter bound via Lemma 5. Dolgopyat [17, Theorems A.2 and A.3] gave an estimate for the spectral gap which is weaker than poly-logarithmic but his argument would give a poly-log estimate without significant changes. His proof consists of a version of the Solovay-Kitaev argument (that he discovered independently) and a variant of Lemma 5. The connection between the present paper and [17] was pointed out to me by Breuillard, see also his survey [10].

We note that Bourgain and Gamburd [3, Corollary 1.1], [6, Theorem 1] showed that when $\mu = \mu_S$ for some finite set $1 \in S \subset \mathrm{SU}(d)$ and the entries of the elements of S are algebraic numbers, then

$$\mathrm{gap}_r(G, S) \geq c$$

for some constant c depending on G and S. Their argument is likely to carry over to arbitrary semi-simple compact Lie groups, however the assumption on algebraicity is essential for the proof. It is a very interesting open problem whether this assumption can be removed. Moreover, we raise the following question: Is it true that there are numbers c, r_0 depending only on G such that

$$\mathrm{gap}_r(G, \mu) \geq c \cdot \mathrm{gap}_{r_0}(G, \mu)$$

for all probability measures μ on G?

Finally, we state a technical result which almost immediately follow from Theorem 6. Its purpose is that this is the version used in the paper [35] to study random walks in the group of Euclidean isometries.

For a measure ν on G, we define the measure $\widetilde{\nu}$ by

$$\int f(x)\, d\widetilde{\nu}(x) = \int f(x^{-1})\, d\nu(x)$$

for all continuous functions f. We write m_G for the Haar measure on G.

Corollary 7. *Let G be a compact Lie group with semi-simple connected component. Let μ be a probability measure on G such that $\mathrm{supp}\,(\widetilde{\mu} * \mu)$ generates a dense subgroup in G. Then there is a constant $c > 0$ depending only on μ such that the following holds. Let $\varphi \in \mathrm{Lip}(G)$ be a function such that $\|\varphi\|_2 = 1$ and $\int \varphi\, dm_G = 0$. Then*

$$\left\| \int \varphi(h^{-1}g)\, d\mu(h) \right\|_2 < 1 - c \log^{-A}(\|\varphi\|_{\mathrm{Lip}} + 2),$$

where A depends on G and is the same as in Theorem 6.

The rest of the paper is organized as follows. In Section 2 we give the proof of Theorem 3 and Corollary 4. The proof of Theorem 6 is given in Section 3. Sections 2 and 3 are independent, but there is a strong analogy between the two arguments. Finally we prove Corollary 7 and Lemma 5 in Section 4.

Throughout the paper we use the letters c and C to denote numbers which may depend on several other quantities and their value may change at each occurrence. We follow the convention that c tends to denote numbers that we consider "small" and C denotes those that we consider "large".

1.6 Motivation

In recent years there was a lot of progress on the Babai-Seress conjecture mentioned above, although it has not been settled yet. In addition, poly-log spectral gap and diameter is known to hold for some families of non-simple finite groups as well. What the scope of this phenomenon is, is an interesting question. Our result on finite groups is a (very modest) step towards understanding this. In Section 2.2 we include some remarks on how to exploit our approach for non-quasirandom groups.

Our main motivation for Theorem 6 is the application in the paper [35]. Although it seems easy to extend the Solovay-Kitaev approach to prove similar poly-log type bounds, we believe that our method gives better exponents, at least for spectral gaps.

There are many recent applications of spectral gaps. Many of these require stronger bounds than what we obtain in this paper, e.g. the results in [2], [4], [5], [7], [34], [9], [31] mentioned above. However, for some applications, the poly-log type bounds are enough, at least to obtain the same qualitative result. Prominent examples are the work of Ellenberg, Hall and Kowalski [18], the Group Large Sieve developed by Lubotzky and Meiri [28], the study of curvatures in Apollonian Circle Packings by Bourgain and Kontorovich [8] and the study of random walks on Euclidean isometries by Varjú [35]. However, our results are relevant only for the last two of the above papers, because [18] and [28] requires spectral gaps only for products of two simple groups. In addition, in the case of [8] a better uniform spectral gap is available.

Acknowledgments. I thank Jean Bourgain for discussions related to this project, in particular, for explaining to me the relation between the diameter and the spectral gap. I thank Nicolas de Saxcé for careful reading of my manuscript and for his suggestions which greatly improved the presentation of the paper. I thank Emmanuel Breuillard for calling my attention on the work of Dolgopyat [17].

2 Profinite groups

2.1 Proof of Theorem 3

Recall that G is a profinite group and Ω is the family of finite index normal subgroups of it. Assume that the hypothesis of the theorem is satisfied. We note that $\alpha \leq 1/2$, since $(\dim \pi)^2 \leq |G/\Gamma|$ for any irreducible representation π of the group G/Γ. Consequently $A \geq 1$.

We first explain how to reduce to the case when μ is symmetric. Define the probability measure $\widetilde{\mu}$ by

$$\int f(g)\, d\widetilde{\mu}(g) = \int f(g^{-1})\, d\mu(g).$$

(If G is finite, this is can be expressed as $\widetilde{\mu}(g) = \mu(g^{-1})$.) Clearly $\widetilde{\mu} * \mu$ is symmetric and

$$\mathrm{Reg}^{\circ}_{G/\Gamma}(\widetilde{\mu} * \mu) = \mathrm{Reg}^{\circ}_{G/\Gamma}(\mu)^{*} \cdot \mathrm{Reg}^{\circ}_{G/\Gamma}(\mu).$$

Hence

$$\|\mathrm{Reg}^{\circ}_{G/\Gamma}(\widetilde{\mu} * \mu)\| = \|\mathrm{Reg}^{\circ}_{G/\Gamma}(\mu)\|^{2}.$$

This yields

$$\mathrm{gap}(G/\Gamma, \widetilde{\mu} * \mu) \geq \mathrm{gap}(G/\Gamma, \mu) \geq \mathrm{gap}(G/\Gamma, \widetilde{\mu} * \mu)/2.$$

This shows that the spectral gap of $\widetilde{\mu} * \mu$ is roughly proportional to that of μ, hence it suffices to prove the theorem for the first one. From now on, we assume that μ is symmetric, hence the operator $\mathrm{Reg}_{G/\Gamma}(\mu)$ is selfadjoint and has an eigenbasis with real eigenvalues.

We define

$$l_\Gamma = 2\lfloor C_\Gamma \log^{A+1}[G:\Gamma]\rfloor,$$

where

$$C_\Gamma := C_0(10 - \log^{-1/10}([G:\Gamma])),$$

and

$$C_0 = \max_{\Gamma\in\Omega_2} \frac{1}{\mathrm{gap}(G/\Gamma,\mu)\cdot \log^{A}[G:\Gamma]}.$$

(The role of the subtracted term in the definition of C_Γ is simply to cancel lower order terms later.)

Recall that we denote by $\chi_{G/\Gamma}$ the character of $\mathrm{Reg}_{G/\Gamma}$. Our goal is to prove

$$\chi_{G/\Gamma}\left(\mu^{*(l_\Gamma)}\right) \leq M, \tag{7}$$

where $M \geq 2$ is a suitably large number depending on α and β. Once we proved this, the claim of the theorem will be concluded easily.

The proof is by induction with respect to the partial order $\lhd$ on $\Gamma \in \Omega$. Suppose that (7) holds for all $\Gamma' \in \Omega$ with $\Gamma \underset{\neq}{\lhd} \Gamma'$.

We distinguish two cases. First we suppose that $\Gamma \in \Omega_2$. Note that $\chi_{G/\Gamma}(\mu^{*(l_\Gamma)})$ is the trace of the operator $\mathrm{Reg}_{G/\Gamma}(\mu^{*(l_\Gamma)})$, hence it is the sum of its eigenvalues. The non-trivial eigenvalues are bounded by $e^{-l_\Gamma\cdot \mathrm{gap}(G/\Gamma,\mu)}$, hence

$$\begin{aligned}\chi_{G/\Gamma}\left(\mu^{*(l_\Gamma)}\right) &\leq 1 + [G:\Gamma]\cdot e^{-l_\Gamma\cdot\mathrm{gap}(G/\Gamma,\mu)}\\ &\leq 1 + [G:\Gamma]\cdot e^{-C_0 \log^{A+1}[G:\Gamma]\cdot \mathrm{gap}(G/\Gamma,\mu)} \leq 1+1\end{aligned}$$

using the definitions of l_Γ, C_Γ and C_0. Hence (7) follows.
Now we suppose that $\Gamma \notin \Omega_2$, hence $\Gamma \in \Omega_1$, in particular $[G:\Gamma] \geq C_1$, where C_1 can be taken as large as we need depending on α, β, c_1, A. Let Γ' be such that $\Gamma \lhd \Gamma' \in \Omega$ and $[\Gamma':\Gamma]$ is minimal but larger than 1. Then by assumption, $[\Gamma':\Gamma] \leq [G:\Gamma']^\beta$.
Since $\chi_{G/\Gamma}$ is pointwise majorized by the function $[\Gamma':\Gamma]\cdot\chi_{G/\Gamma'}$, and $\mu^{*(l'_\Gamma)}$ is a positive measure, we have

$$\chi_{G/\Gamma}\left(\mu^{*(l_{\Gamma'})}\right) \leq [\Gamma':\Gamma]\cdot\chi_{G/\Gamma}\left(\mu^{*(l_{\Gamma'})}\right) \leq M[\Gamma':\Gamma]. \tag{8}$$

We applied (7) for Γ' in the second inequality.
Denote by $\lambda_0 = 1, \lambda_1, \ldots \lambda_k$ the eigenvalues of the operator $\mathrm{Reg}_{G/\Gamma}(\mu)$ each listed as many times as its multiplicity. Then

$$\begin{aligned}\chi_{G/\Gamma}\left(\mu^{*(l_\Gamma)}\right) &= 1+\sum_{i=1}^k \lambda_i^{l_\Gamma} \leq 1+\left(\sum_{i=1}^k \lambda_i^{l_{\Gamma'}}\right)\cdot\max\{\lambda_i\}^{l_\Gamma - l_{\Gamma'}}\\ &\leq 1+M[\Gamma':\Gamma]\max\{\lambda_i\}^{l_\Gamma-l_{\Gamma'}}. \end{aligned}\tag{9}$$

We applied (8) in the last line. Also note, that all terms are positive because l_Γ is even by construction.
Our next goal is to obtain a sufficient bound on the λ_i. This can be deduced from (8) and the assumption about the dimension of faithful representations. The representation $\mathrm{Reg}(G/\Gamma)$ can be decomposed as the orthogonal sum of irreducible subrepresentations. Each irreducible representation occur with multiplicity equal to its dimension.
Consider now an eigenvalue λ_i. There is a corresponding eigenvector which is contained in an irreducible subrepresentation of $\mathrm{Reg}(G/\Gamma)$. Denote this representation by π. Write $\Gamma'' = \mathrm{Ker}(\pi)$.
First we consider the case that $\Gamma'' = \Gamma$. It follows that $\lambda_i^{l_{\Gamma'}}$ is an eigenvalue of $\mathrm{Reg}_{G/\Gamma}(\mu^{*(l_{\Gamma'})})$ with multiplicity at least

$$\dim\pi \geq c_1[G:\Gamma]^\alpha,$$

(by the assumption on quasirandomness). Hence by (8) we can conclude that

$$\lambda_i^{l_{\Gamma'}} \leq \frac{M[\Gamma':\Gamma]}{c_1[G:\Gamma]^\alpha}.$$

Next, we consider the case when $\Gamma'' \neq \Gamma$. By the assumption on the minimality of $[\Gamma':\Gamma]$, we have $[G:\Gamma''] \leq [G:\Gamma']$. We apply (7) for Γ'' along with the bound for the multiplicity of eigenvalues and get

$$\lambda_i^{l_{\Gamma''}} \leq \frac{M}{c_1[G:\Gamma'']^\alpha}.$$

An easy calculation shows that the bound we obtain for $|\lambda_i|$ is worsening when $[G:\Gamma'']$ grows.

Thus in both cases we obtain:

$$\lambda_i^{l_{\Gamma'}} \le \frac{M[\Gamma':\Gamma]}{c_1[G:\Gamma']^\alpha}. \tag{10}$$

We plug this into (9) and we want to conclude $\chi_{G/\Gamma}(\mu^{*(l_\Gamma)}) \le M$. To this end, we need

$$M[\Gamma':\Gamma]\left(\frac{M[\Gamma':\Gamma]}{c_1[G:\Gamma']^\alpha}\right)^{(l_\Gamma-l_{\Gamma'})/l_{\Gamma'}} \le M-1$$

that is

$$\frac{M}{M-1}[\Gamma':\Gamma] \le \left(\frac{c_1[G:\Gamma']^\alpha}{M[\Gamma':\Gamma]}\right)^{(l_\Gamma-l_{\Gamma'})/l_{\Gamma'}}. \tag{11}$$

For simplicity, we write $X=\log[G:\Gamma']$ and $Y=\log[\Gamma':\Gamma]$. Then by the definition of l_Γ:

$$\begin{aligned}
\frac{l_\Gamma - l_{\Gamma'}}{l_{\Gamma'}} &= \frac{2\lfloor C_\Gamma (X+Y)^{A+1}\rfloor - 2\lfloor C_{\Gamma'}X^{A+1}\rfloor}{2\lfloor C_{\Gamma'}X^{A+1}\rfloor}\\
&\ge \frac{C_\Gamma(X+Y)^{A+1} - C_{\Gamma'}X^{A+1} - 1}{C_{\Gamma'}X^{A+1}}\\
&\ge \frac{C_{\Gamma'}(X+Y)^{A+1} - C_{\Gamma'}X^{A+1}}{C_{\Gamma'}X^{A+1}} + \frac{C_\Gamma - C_{\Gamma'}}{C_{\Gamma'}} - \frac{1}{C_{\Gamma'}X^{A+1}}\\
&\ge (A+1)\frac{Y}{X} + c_2\frac{Y}{X^{11/10}},
\end{aligned} \tag{12}$$

where c_2 is an absolute constant if X is larger than an absolute constant. (Using the definition of C_Γ, we evaluate $(C_\Gamma - C_{\Gamma'})/C_{\Gamma'}$ and get the second term of (12). Notice that this is of larger order of magnitude than $1/X^{A+1}$.)
Then the logarithm of the right hand side of (11) is at least

$$\begin{aligned}
&\left((A+1)\frac{Y}{X} + c_2\frac{Y}{X^{11/10}}\right)(\alpha X - \log(M/c_1) - Y)\\
&\qquad\ge \alpha(A+1)Y - \log(M/c_1)(A+1)\frac{Y}{X} - (A+1)\frac{Y^2}{X} + c_3\frac{Y}{X^{1/10}}\\
&\qquad\ge \alpha(A+1)Y - (A+1)\frac{Y^2}{X} + c_3\frac{Y}{2X^{1/10}}
\end{aligned}$$

if X is sufficiently large depending on α,β,c_1,A. Here c_3 is a sufficiently small constant depending on α and β satisfying $c_3X \le c_2(\alpha X - \log(M/c_1) - Y)$. (Recall that $Y \le \beta X$ by assumption.)
To get (11), we need

$$\log(M/(M-1)) + Y \le \alpha(A+1)Y - (A+1)\frac{Y^2}{X} + c_3\frac{Y}{2X^{1/10}}$$

that is

$$A+1 \ge \frac{1+\log(M/(M-1))/Y - c_4X^{-1/10}}{\alpha - Y/X + c_4X^{-1/10}}, \tag{13}$$

where c_4 is yet another number depending on α, β, c_1, A.
We consider two cases. If $Y \le \sqrt{X}$ (and X is sufficiently large) then $Y/X \le c_4 X^{-1/10}$. Clearly $Y \ge \log 2$, hence (13) holds if

$$A + 1 \ge \frac{1 + \log(M/(M-1))/\log 2}{\alpha}.$$

On the other hand if $Y \ge \sqrt{X}$ (and X is sufficiently large) then $\log(M/(M-1))/Y \le c_4 X^{-1/10}$. Recall that $Y/X \le \beta$ by assumption, hence (13) holds if $A + 1 \ge (\alpha - \beta)^{-1}$.
By choosing M sufficiently large depending on α, β, we can ensure that

$$\frac{1 + \log(M/(M-1))/\log 2}{\alpha} \le \frac{1}{\alpha - \beta}.$$

Thus (13) holds in either case if $A + 1 \ge (\alpha - \beta)^{-1}$, which was assumed in the theorem. This completes the induction to prove (7).
Let λ be an eigenvalue of $\mathrm{Reg}^{\circ}_{G/\Gamma}(\mu)$. We want to show that

$$|\lambda| < 1 - C_1^{-1} B \cdot \log^{-A}[G : \Gamma].$$

Denote by π an irreducible representation of G/Γ that contains an eigenvector corresponding to λ. Without loss of generality, we can assume that $\Gamma = \mathrm{Ker}(\pi)$.
From quasirandomness and (7) we get

$$\lambda^{l_\Gamma} \le \frac{M}{c_1 [G : \Gamma]^\alpha}.$$

Thus

$$\log|\lambda| \le \frac{\log M/c_1 - \alpha \log[G : \Gamma]}{l_\Gamma}.$$

If we compare this with the definition of l_Γ we can conclude the theorem.

2.2 Some remarks about weakening the hypothesis on quasirandomness

In this section we present some ideas that lead to a refined version of (9) and (10). Using this, one could obtain a version of Theorem 3 with a weaker hypothesis instead of quasirandomness. Namely one would require that the quotient groups does not have "many" irreducible representations with "small" dimension. This weaker form of quasirandomness would be very technical hence we do not state a theorem. However, (based on analogy with compact Lie groups) it seems possible that these ideas lead to better bounds for the groups $\mathrm{SL}_d(\mathbf{Z}_p)$ than Theorem 3.

Proposition 8. *Let G be a finite group and N a normal subgroup. Let $\rho_1, \ldots, \rho_n$ be all irreducible representations of N up to isomorphism. Denote by*

$a(\rho_j)$ the number of G-conjugates of ρ_j. Denote by $d(\rho_j)$ the smallest possible dimension of a representation of G whose restriction to N contains ρ_j.
Let μ be a symmetric probability measure on G and suppose that

$$\chi_{G/N}\left(\mu^{*(2l)}\right) \le M \tag{14}$$

for some numbers l and $M > 1$.
Then for all integers $l' > l$ we have:

$$\chi_G\left(\mu^{*(2l')}\right) \le \sum_{j=1}^{n}\left[\dim(\rho_j)^2 M \cdot \left(\frac{a(\rho_j)\dim(\rho_j)^2 M}{d(\rho_j)}\right)^{(l'-l)/l}\right]. \tag{15}$$

Observe that $\sum \dim(\rho_j)^2 = |N|$ and $a(\rho_j)\dim(\rho_j)^2 \le |N|$, hence we obtain (9) combined with (10), if we estimate $d(\rho_j)$ using quasirandomness. If $d(\rho_j)$ is small only for a very few j, Proposition 8 significantly improves the argument given in the previous section.

Proof. We recall some facts from representation theory. Let π be an irreducible representation of G, and denote by χ_π its character. Denote by $\pi|_N$ the restriction of π to N. By Clifford's theorem, the irreducible components of $\pi|_N$ is a G-conjugacy class of representations and each appears with the same multiplicity. We denote by $a(\pi)$ the number of different irreducible components of $\pi|_N$, by $b(\pi)$ their common multiplicity and by $c(\pi)$ their common dimension. Thus $\dim(\pi) = a(\pi)\cdot b(\pi)\cdot c(\pi)$.
Denote one of the irreducible components of $\pi|_N$ by ρ and its character by χ_ρ. Let $\rho^{g_1}, \cdots, \rho^{g_{a(\pi)}}$ be all G-conjugates of ρ. In what follows, the function

$$\varphi = \sum_{i=1}^{a(\pi)} \chi_{\rho^{g_i}}$$

will play an important role. We also extend it to G by setting it 0 in the complement of N, i.e. we write $\widetilde{\varphi}(g) = \varphi(g)$ for $g \in N$ and $\widetilde{\varphi}(g) = 0$ otherwise. We use two inner products, one on $L^2(G)$ and one on $L^2(N)$ defined by:

$$\langle f_1, f_2\rangle_G = \frac{1}{|G|}\sum_{g\in G} f_1(g)f_2(g), \quad \langle f_1, f_2\rangle_N = \frac{1}{|N|}\sum_{g\in N} f_1(g)f_2(g).$$

These are the inner products with respect to which the irreducible characters of the corresponding groups are orthonormal.
We write:

$$\langle \chi_\pi, \widetilde{\varphi}\rangle_G = \frac{|N|}{|G|}\langle \chi_{\pi|_N}, \varphi\rangle_N = \frac{|N|a(\pi)b(\pi)}{|G|}.$$

A similar calculation shows that the inner product of $\widetilde{\varphi}$ with an irreducible character of G is always non-negative.

Since $\widetilde{\varphi}$ is a class function on G, it can be decomposed as a linear combination of irreducible characters. According to the above calculation, the coefficient of χ_π is $|N|a(\pi)b(\pi)/|G|$ and all other characters have non-negative contribution. Thus

$$\chi_\pi\left(\mu^{*(2l)}\right) \le \frac{|G|}{|N|a(\pi)b(\pi)}\widetilde{\varphi}\left(\mu^{*(2l)}\right). \tag{16}$$

(Note that $\chi_\pi(\mu^{*(2l)}) \ge 0$ being the trace of a positive operator.)
On the other hand,

$$\|\widetilde{\varphi}\|_\infty \le \sum_{i=1}^{a(\pi)} \|\chi_{\rho^{g_i}}\|_\infty = a(\pi)c(\pi).$$

Thus for every $g \in G$, we have

$$|\widetilde{\varphi}(g)| \le \frac{|N|a(\pi)c(\pi)}{|G|}\chi_{G/N}(g).$$

(Note that for $g \notin N$ both sides are 0.) Therefore

$$\widetilde{\varphi}\left(\mu^{*(2l)}\right) \le \frac{|N|a(\pi)c(\pi)}{|G|}\chi_{G/N}\left(\mu^{*(2l)}\right). \tag{17}$$

We combine (16) and (17) and get

$$\chi_\pi\left(\mu^{*(2l)}\right) \le \frac{c(\pi)}{b(\pi)}\chi_{G/N}\left(\mu^{*(2l)}\right).$$

This implies that for all eigenvalues λ of $\pi(\mu^{*(2l)})$, we have

$$|\lambda| < \frac{c(\pi)}{b(\pi)}M = \frac{a(\pi)c(\pi)^2}{\dim(\pi)}M. \tag{18}$$

(Here we also used the hypothesis (14).)
Now let $\pi_1, \pi_2, \ldots, \pi_k$ denote all the irreducible representations (up to isomorphism) of G whose restriction to N contain ρ. By a calculation very similar to the one leading to (16) we get

$$\sum_{i=1}^{k}\frac{|N|a(\pi_i)b(\pi_i)}{|G|}\chi_{\pi_i}\left(\mu^{*(2l)}\right) \le \widetilde{\varphi}\left(\mu^{*(2l)}\right).$$

Combining with (17), we get

$$\sum_{i=1}^{k}\frac{b(\pi_i)}{c(\pi_i)}\chi_{\pi_i}\left(\mu^{*(2l)}\right) \le \chi_{G/N}\left(\mu^{*(2l)}\right) \le M.$$

(We used again the hypothesis (14).) Multiplying by $a(\pi)c(\pi)^2 = a(\pi_i)c(\pi_i)^2$ (which is independent of i) we get

$$\sum_{i=1}^{k} \dim(\pi_i)\chi_{\pi_i}\left(\mu^{*(2l)}\right) \leq a(\pi)c(\pi)^2 M. \tag{19}$$

We use (18) and write for an integer $l' \geq l$:

$$\sum_{i=1}^{k} \dim(\pi_i)\chi_{\pi_i}\left(\mu^{*(2l')}\right) \leq a(\pi)c(\pi)^2 M \left(\frac{a(\pi)c(\pi)^2 M}{\dim(\pi)}\right)^{(l'-l)/l}. \tag{20}$$

We sum (20) for $\rho = \rho_1, \ldots, \rho_n$ and get (15). $\square$

2.3 Proof of Corollary 4

We first discuss quasirandomness. This was already proved for $\mathrm{SL}_d(\widehat{\mathbf{Z}})$ by Bourgain and Varjú [9]. In fact, it is easy to deduce it from the corresponding result about $\mathrm{SL}_d(\mathbf{Z}_p)$ which was obtained by Bourgain and Gamburd [4, Lemma 7.1] for $d = 2$ and by Saxcé [33, Lemme 5.1] for $d \geq 2$. For completeness, we explain this deduction.

For an integer $q > 0$ denote by

$$\Gamma_q = \{g \in \mathrm{SL}_d(\widehat{\mathbf{Z}}) : g \equiv 1 \pmod{q}\}$$

the mod q congruence subgroup of $\mathrm{SL}_d(\widehat{\mathbf{Z}})$.

Let p be a prime, $k \geq 1$ be an integer, and let π be a representation of $\mathrm{SL}_d(\widehat{\mathbf{Z}})/\Gamma_{p^k}$ which is not a representation of $\mathrm{SL}_d(\widehat{\mathbf{Z}})/\Gamma_{p^{k-1}}$. By [33, Lemme 5.1], π is of dimension at least

$$c \cdot [\mathrm{SL}_d(\widehat{\mathbf{Z}}) : \Gamma_{p^k}]^{1/(d+1)},$$

where $c > 0$ is a number depending on d. For any $\varepsilon > 0$, we can replace this bound by $|\mathrm{SL}_d(\widehat{\mathbf{Z}})/\Gamma_q|^{1/(d+1)-\varepsilon}$ if p is sufficiently large depending on ε.

Let π be an irreducible unitary representation of $\mathrm{SL}_d(\widehat{\mathbf{Z}})$. Since Γ_q, $q > 1$ form a system of neighborhoods of 1 in $\mathrm{SL}_d(\widehat{\mathbf{Z}})$, there is q such that $\mathrm{Ker}(\pi) \supset \Gamma_q$. Let q be minimal with this property. Let $q = p_1^{k_1} \cdots p_n^{k_n}$ such that p_i are primes. Then

$$\mathrm{SL}_d(\widehat{\mathbf{Z}})/\Gamma_q = \left(\mathrm{SL}_d(\widehat{\mathbf{Z}})/\Gamma_{p_1^{k_1}}\right) \times \cdots \times \left(\mathrm{SL}_d(\widehat{\mathbf{Z}})/\Gamma_{p_n^{k_n}}\right).$$

Any representation of this group is a tensor product of representations of $\mathrm{SL}_d(\widehat{\mathbf{Z}})/\Gamma_{p_i^{k_i}}$. Hence

$$\dim \pi \geq c^m [\mathrm{SL}_d(\widehat{\mathbf{Z}}) : \Gamma_{p^k}]^{1/(d+1)-\varepsilon}, \tag{21}$$

where m is the number of not large enough primes in the sense of the previous paragraph. Thus $\mathrm{SL}_d(\widehat{\mathbf{Z}})$ is $(c^m, 1/(d+1) - \varepsilon)$-quasirandom.

We refer the interested reader to the paper of Kelmer and Silberman [25, Section 4], where quasirandomness is proved with optimal parameter α for some other arithmetic groups.

We define Ω_1 and Ω_2. We fix an integer M that we will set later depending on d. Let $\Gamma \lhd \mathrm{SL}_d(\widehat{\mathbf{Z}})$ be a finite index normal subgroup. Denote by q the smallest integer such that $\Gamma_q \lhd \Gamma$. We put Γ in Ω_1, if q has at least $M+1$ prime factors (taking multiplicities into account) and q is sufficiently large. We put Γ in Ω_2 otherwise.

Let $\Gamma \in \Omega_1$ and let q be the same as above. Let $p|q$ be the smallest prime divisor of q. By simple calculation:

$$[\Gamma_{q/p} : \Gamma_q] \leq [\mathrm{SL}_d(\widehat{\mathbf{Z}}) : \Gamma_{q/p}]^{1/M}.$$

Now we define $\Gamma' := \Gamma_{q/p}\Gamma$. Clearly

$$[\Gamma' : \Gamma] \leq [\Gamma_{q/p} : \Gamma_q],$$

so we only need to estimate $[\mathrm{SL}_d(\widehat{\mathbf{Z}}) : \Gamma']$ in terms of $[\mathrm{SL}_d(\widehat{\mathbf{Z}}) : \Gamma_{q/p}]$. For our purposes the very crude bound

$$[\mathrm{SL}_d(\widehat{\mathbf{Z}}) : \Gamma'] \geq c^m [\mathrm{SL}_d(\widehat{\mathbf{Z}}) : \Gamma_{q/p}]^{1/(d+1)-\varepsilon}$$

is sufficient which follows from (21). This implies

$$[\Gamma' : \Gamma] \leq [\mathrm{SL}_d(\widehat{\mathbf{Z}}) : \Gamma']^{(d+2)/M}$$

if q is sufficiently large (and ε is sufficiently small).

We choose α and β in such a way that $\beta < \alpha < 1/(d+1)$. Then the quasirandomness is satisfied, and also (5) if we set $M \geq (d+2)/\beta$.

It is left to verify (6) for $\mu = \mu_S$. First we note that by the same argument as at the beginning of the proof of Theorem 3, we can assume that S is symmetric. We show that the groups $\mathrm{SL}_d(\widehat{\mathbf{Z}})/\Gamma_q$ for $\Gamma_q \in \Omega_2$ have poly-logarithmic diameter with respect to any generating set S. In light of Lemma 1 this implies (6).

Let now $\Gamma_q \in \Omega_2$. There are two possibilities. If q is small (e.g. $q \leq C_1$ or as in the definition of Ω_1), then we have the trivial bound

$$\mathrm{diam}(\mathrm{SL}_d(\widehat{\mathbf{Z}})/\Gamma_q, S) \leq |\mathrm{SL}_d(\widehat{\mathbf{Z}})/\Gamma_q| \leq C \log^A(\mathrm{SL}_d(\widehat{\mathbf{Z}})/\Gamma_q)$$

for some suitably large constant C.

The other situation that may happen is that q contains at most M prime factors counting multiplicities. In this case we can easily deduce the poly-log diameter bound from [11, Theorem 7.1] and [30, Theorem 2] which contain this result in the case when q is prime. This deduction is very similar to [9, Proof of Proposition 3].

Let $q_0 = 1, q_1, q_2, \ldots, q_n = q$ be a sequence of integers such that q_{i+1}/q_i is a prime number for all i. We will apply the following lemma repeatedly to prove the diameter bound we are looking for.

Lemma 9. *Fix i and write $p = q_i/q_{i-1}$. Then*

$$\operatorname{diam}(\mathrm{SL}_d(\mathbf{Z}/q_i\mathbf{Z}), S) \leq C(\operatorname{diam}(\mathrm{SL}_d(\mathbf{Z}/q_{i-1}\mathbf{Z}), S) + \operatorname{diam}(\mathrm{SL}_d(\mathbf{Z}/p\mathbf{Z}), S)),$$

where C is a number depending on d.

Proof. Let

$$D = \max\{\operatorname{diam}(\mathrm{SL}_d(\mathbf{Z}/q_{i-1}\mathbf{Z}), S), \operatorname{diam}(\mathrm{SL}_d(\mathbf{Z}/p\mathbf{Z}), S)\}.$$

Then

$$S^D \cdot \Gamma_{q_{i-1}} = \mathrm{SL}_d(\mathbf{Z}/q_i\mathbf{Z}).$$

Since S is generating, S^{D+1} must intersect some $\Gamma_{q_{i-1}}$-coset in at least two points. Thus there is

$$1 \neq g_0 \in S^{2D+1} \cap \Gamma_{q_{i-1}}.$$

If $p \nmid q_{i-1}$, and hence $\Gamma_{q_{i-1}}/\Gamma_{q_i} = \mathrm{SL}_d(\mathbf{Z}/p\mathbf{Z})$, we also want to show that g_0 can be taken non-central in $\mathrm{SL}_d(\mathbf{Z}/p\mathbf{Z})$. With the same argument as above, we can show that $S^{(j+1)D+j}$ intersects all $\Gamma_{q_{i-1}}$-cosets in at least $j+1$ points. Taking $j = |Z(SL_d(\mathbf{Z}/p\mathbf{Z}))|$, we can find a suitable g_0 in $S^{(j+1)D+j}$.

We put

$$X = \{g^{-1}g_0 g, g^{-1}g_0^{-1}g : g \in S^D\}.$$

We show that

$$X^C = \Gamma_{q_{i-1}}/\Gamma_{q_i} \tag{22}$$

for some constant C depending on d.

We have two cases. First, we suppose that $p \nmid q_{i-1}$. Then $\Gamma_{q_{i-1}}/\Gamma_{q_i} = \mathrm{SL}_d(\mathbf{Z}/p\mathbf{Z})$, and X is a non-trivial conjugacy class. In this case (22) is a result of Lev [27, Theorem 2].

Now suppose that $p|q_{i-1}$. In this case $\Gamma_{q_{i-1}}/\Gamma_{q_i}$ is isomorphic to $\mathfrak{sl}_d(\mathbf{Z}/p\mathbf{Z})$, and the conjugation action

$$h \mapsto g^{-1}hg, \quad g \in \mathrm{SL}_d(\mathbf{Z}/q_i\mathbf{Z}),\ h \in \Gamma_{q_{i-1}}/\Gamma_{q_i}$$

factors through $G/\Gamma_p = \mathrm{SL}_d(\mathbf{Z}/p\mathbf{Z})$. Now the claim (22) follows from [9, Lemma 5].

Therefore, we have $S^{C\cdot D} \supset \Gamma_{q_{i-1}}/\Gamma_{q_i}$ for some other constant C depending on d. This implies $S^{(C+1)\cdot D} = \mathrm{SL}_d(\mathbf{Z}/q_i\mathbf{Z})$ which was to be proved. ☐

Now Corollary 4 is immediate. By [11, Theorem 7.1] and [30, Theorem 2] we have

$$\operatorname{diam}(\mathrm{SL}_d(\mathbf{Z}/p\mathbf{Z}), S) \leq C \log^A(p)$$

for all primes p. We can use Lemma 9 repeatedly to show

$$\operatorname{diam}(\mathrm{SL}_d(\widehat{\mathbf{Z}})/\Gamma_q, S) \leq C \log^A(p)$$

for $\Gamma_q \in \Omega_2$, where p is the largest prime factor of q and C is a different constant depending on M. This is precisely the poly-log diameter estimate we were looking for.

3 Compact Lie groups

The purpose of this section is to prove Theorem 6. Recall the definitions of $T, \mathrm{L}T, \mathrm{L}T^*, I, I^*, R, R_+, K$ and π_v from Section 1. Let μ be a probability measure on G. Denote by χ_v the character of π_v. For a continuous function f and a measure ν on G, we write

$$f(\nu) = \int f(x)\, d\nu(x).$$

If f is a continuous class function on G, then by the Peter-Weyl theorem, we can decompose it as a linear combination of irreducible characters. Denote by $m_v(f)$ the coefficient of χ_v in this decomposition.
We introduce two partial orders on the space of continuous class functions on G. We write $f_1 \le f_2$, if $f_1(g) \le f_2(g)$ for every $g \in G$. We write $f_1 \sqsubseteq f_2$ if $m_v(f_1) \le m_v(f_2)$ for all $v \in \overline{K} \cap I^*$. Denote by $\preceq$ the transitive closure of the union of $\le$ and $\sqsubseteq$, i.e. we write $f_1 \preceq f_2$ if there is a sequence of class functions φ_i such that

$$f_1 \le \varphi_1 \sqsubseteq \varphi_2 \le \varphi_3 \sqsubseteq \ldots \le \varphi_n \sqsubseteq f_2.$$

These relations have a crucial property contained in the following Lemma. Recall that for a measure ν on G, we define the measure $\widetilde{\nu}$ by

$$\int f(x)\, d\widetilde{\nu}(x) = \int f(x^{-1})\, d\nu(x)$$

for all continuous functions f. We say that ν is symmetric if $\nu = \widetilde{\nu}$.

Lemma 10. *Let ν be a symmetric probability measure and let $f_1 \preceq f_2$ be two continuous class functions on G. Then*

$$f_1(\nu * \nu) \le f_2(\nu * \nu).$$

Proof. Clearly, it is enough to prove the statements for $\le$ and $\sqsubseteq$ in place of $\preceq$. For $\le$ it easily follows from the definitions and from the fact that $\nu * \nu$ is a positive measure.
Suppose $f_1 \sqsubseteq f_2$. Observe that

$$f_i(\nu * \nu) = \sum_{v \in \overline{K} \cap I^*} m_v(f_i) \chi_v(\nu * \nu).$$

Hence the claim follows from $m_v(f_1) \le m_v(f_2)$ once we prove that $\chi_v(\nu * \nu) \ge 0$ for all v. This follows from $\chi_v(\nu * \nu) = \mathrm{Tr}(\pi_v(\nu * \nu))$ and from the fact that $\pi_v(\nu * \nu) = \pi_v(\nu) \cdot \pi_v(\nu)^*$ is a positive selfadjoint operator. ☐

Now we explain the strategy of the proof. First of all, we note that by the argument at the beginning of Section 2.1 we can assume that μ is symmetric. Hence Lemma 10 applies for $\nu = \mu^{*(2l)}$ for all positive integers l.

We write for $r \geq 1$

$$\chi_r = \left(\sum_{|v| \leq r} \chi_v \right)^2 \cdot r^{-\dim \mathrm{L}T}$$

which plays the role of $\chi_{G/\Gamma}$ used in the previous section. We also write

$$l_r = 2\lfloor C_r \log^{A+1} r \rfloor,$$

where

$$C_r = C_0 \left(10 - \log^{-1/10}(r) \right)$$

and C_0 is a suitably large constant to be set later.
Our goal is to prove the inequality

$$\chi_r \left(\mu^{*(l_r)} \right) \leq E \tag{23}$$

for some constant E independent of r. This will easily imply the theorem.
We assume that (23) holds for some range $1 \leq r \leq r_1$. (This can be verified easily for $r_1 = r_0$ if C_0 is suitably large in terms of $\mathrm{gap}_{r_0}(\mu)$.) And then we show that (23) also holds for a suitable $r = r_2$. Iterating this argument, we can prove the claim for all r.
We prove the "induction step" in the following way. We find suitable functions $\varphi_1, \ldots, \varphi_n$ such that

$$\chi_{r_2} \sqsubseteq \varphi_1 + \ldots + \varphi_n.$$

Then it will be enough to estimate $\varphi_i(\mu^{*(l_{r_2})})$. We will show that $\varphi_i \preceq B_i \chi_{r_1}$, where B_i is a number depending on i, r_1 and r_2. This allows us to estimate $\varphi_i(\mu^{*(l_{r_1})})$. We will also show that $m_v(\varphi_i)$ is either "large" or 0, and this will yield an estimate on the eigenvalues of $\pi_v(\mu^{*(l_{r_1})})$ for all v which contributes to φ_i. Finally this allows us to get a refined estimate on $\varphi_i(\mu^{*(l_{r_2})})$.
To implement the above plan we need methods to estimate $m_v(f)$. In our examples f will always be the character of a representation which is obtained from other representations using tensor products. Explicit formulas are available for $m_v(f)$ in such cases, however, they do not seem very practical for our purposes. Instead, we will use elementary methods to estimate these coefficients based on double-counting dimensions. However, we still need some very basic facts about the representations π_v.
The first fact is Weyl's dimension formula [12, Chapter VI. (1.7) (iv)]:

$$\dim \pi_v = \prod_{u \in R_+} \frac{\langle u, v + \rho \rangle}{\langle u, \rho \rangle},$$

where

$$\rho = \frac{1}{2} \sum_{u \in R_+} u$$

is the half sum of the positive roots.
The second fact is the content of the following lemma which bounds the highest weight of possible irreducible constituents of $\pi_v \otimes \pi_u$. Recall that we denote the Haar measure on G by m_G.

Lemma 11. *There is a constant D depending only on G such that if*

$$\int \chi_v \chi_u \overline{\chi_w}\, dm_G \neq 0 \tag{24}$$

for some $v, u, w \in \overline{K} \cap I^*$, then $|v - w| < D|u|$.

Proof. If (24) holds then π_w is contained in $\pi_v \otimes \pi_u$. By [12, Chapter VI, Lemma (2.8)], $v + u$ dominates w that is

$$\langle v + u, t\rangle \geq \langle w, t\rangle$$

for every $t \in \overline{K}$. Similarly, if (24) holds then π_v is contained in $\pi_w \otimes \overline{\pi_u}$, hence

$$\langle w + \overline{u}, t\rangle \geq \langle v, t\rangle.$$

Here $\overline{u}$ is the highest weight of $\overline{\chi_u}$.
Now let $t_1, \ldots, t_{\dim T}$ be a basis of $\mathrm{L}T^*$ consisting of unit vectors in $\overline{K}$. By the above inequalities, we have

$$|\langle w - v, t_j\rangle| \leq |u|$$

for all $1 \leq j \leq \dim T$. The claim follows from this with the constant D equal to the length of the longest vector in the set

$$\{x \in \mathrm{L}T^* : |\langle x, t_j\rangle| \leq 1\}.$$

□

We proceed by some Lemmata which bound the multiplicities of some irreducible constituents in certain tensor products.

Lemma 12. *We have*

$$\int \left(\sum_{|v| \leq r} \chi_v\right)^2 dm_G = |\{v \in \overline{K} \cap I^* : |v| \leq r\}|.$$

In particular

$$c \leq m_{\chi_0}(\chi_r) \leq C,$$

where $c, C > 0$ are constants depending on G

Proof. Since characters form an orthonormal basis, we have

$$\int \left(\sum_{|v| \le r} \chi_v \right)^2 dm_G = \int \sum_{|v| \le r} \chi_v \overline{\chi_v}\, dm_G = |\{v \in \overline{K} \cap I^* : |v| \le r\}|.$$

For the second statement notice that

$$m_{\chi_0}(\chi_r) = \frac{1}{r^{\dim \mathrm{L}T}} \int \left(\sum_{|v| \le r} \chi_v \right)^2 \chi_0\, dm_G = \frac{|\{v \in \overline{K} \cap I^* : |v| \le r\}|}{r^{\dim \mathrm{L}T}}.$$

□

LEMMA 13. *There is a constant $C > 0$ depending on G such that the following holds. Let $u \in \overline{K} \cap I^*$, and $r \ge 1$.*

$$m_{\chi_u}(\chi_r) \le C \dim \chi_u.$$

Moreover, $m_{\chi_u}(\chi_r) = 0$ if $|u| > Cr$.

Proof. For the first part of the lemma we write

$$\begin{aligned} r^{\dim \mathrm{L}T} m_{\chi_u}(\chi_r) &= \int \left(\sum_{|v| \le r} \chi_v \right)^2 \overline{\chi_u}\, dm_G \\ &\le 2 \sum_{\substack{|v|,|w| \le r \\ \dim \chi_v \le \dim \chi_w}} \int \chi_v \chi_w \overline{\chi_u}\, dm_G \\ &= 2 \sum_{|v| \le r} \left[\sum_{\substack{|w| \le r \\ \dim \chi_w \ge \dim \chi_v}} m_{\overline{\chi_w}}(\chi_v \overline{\chi_u}) \right]. \end{aligned}$$

The sum of the multiplicities of some irreducible components of $\pi_v \otimes \overline{\pi_u}$ can not be bigger than the dimension of $\pi_v \otimes \overline{\pi_u}$ divided by the minimal dimension of the irreducible components we consider. Thus

$$r^{\dim \mathrm{L}T} m_{\chi_u}(\chi_r) \le 2 \sum_{|v| \le r} \frac{\dim(\chi_v \overline{\chi_u})}{\dim(\chi_v)} \le C r^{\dim \mathrm{L}T} \dim \chi_u.$$

The second part follows immediately from Lemma 11. □

For $u \in \overline{K} \cap I^*$ and $r \ge 1$, we write

$$\psi_{u,r} = \chi_r \cdot \sum_{w : |u - w| < 3Dr} \chi_w.$$

We show that χ_z is contained in $\psi_{u,r}$ with high multiplicity if $|u - z| \le Dr$.

Lemma 14. *Let $r \geq 1$ and $z, u \in \overline{K} \cap I^*$ with $|z-u| \leq Dr$. Then*

$$m_{\chi_z}(\psi_{u,r}) \geq c \cdot \frac{\dim(\chi_z) \cdot r^{\dim G}}{\max_{w:|u-w|<3Dr}\{\dim(\chi_w)\}},$$

where $c > 0$ is a constant depending only on G.

Proof. We can write

$$\begin{aligned} r^{\dim \mathrm{L}T} m_{\chi_z}(\psi_{u,r}) &= \sum_{w:|u-w|<3Dr} \left[\sum_{|v_1|,|v_2|\leq r} \int \chi_{v_1}\chi_{v_2}\chi_w\overline{\chi_z}\, dm_G \right] \\ &= \sum_{w:|u-w|<3Dr} m_{\chi_w} \left(\sum_{|v_1|,|v_2|\leq r} \overline{\chi_{v_1}}\,\overline{\chi_{v_2}}\chi_z \right) \\ &\geq \frac{\sum_{|v_1|,|v_2|\leq r} \dim\chi_{v_1} \cdot \dim\chi_{v_2} \cdot \dim\chi_z}{\max_{w:|u-w|<3Dr} \dim\chi_w}. \end{aligned} \tag{25}$$

In the last line we used the fact that all irreducible components of $\overline{\chi_{v_1}}\,\overline{\chi_{v_2}}\chi_z$ has highest weight w satisfying $|w-z| \leq 2Dr$ and hence $|u-w| < 3Dr$ which follows from two applications of Lemma 11. Hence all possible irreducible component appears in the range of summation, and the inequality follows by comparing dimensions.

Let K' be a closed convex cone strictly contained in the positive Weyl chamber K. For $v \in K' \cap I^*$ and $|v| \geq r/2$ it follows from Weyl's dimension formula that $\dim(\chi_v) \geq cr^{|R_+|}$ for some constant $c > 0$ depending only on G, K'. Thus the numerator in (25) is bounded below by

$$c\dim(\chi_z) \cdot r^{2\dim \mathrm{L}T + 2|R_+|}.$$

The proof is finished by noting that $\dim G = \dim \mathrm{L}T + 2|R_+|$. ☐

We continue implementing our plan described above. Recall that $r_2 > r_1 \geq 1$, are numbers that we will choose later and that we assume

$$\chi_r\left(\mu^{*(l_r)}\right) \leq E \tag{26}$$

for $r \leq r_1$ and for some number E to be chosen later, as well. Our goal is to prove (26) with $r = r_2$.

Let $u_0 = 0, u_1, \ldots u_m$ be a maximal Dr_1 separated subset of $\{v \in \overline{K} \cap I^* : |v| \leq Cr_2\}$, where C is the constant from Lemma 13. For $i = 0, \ldots, m$, let

$$M_i = \frac{C\max_{w:|u_i-w|<3Dr_1}\{\dim\chi_w\}}{cr_1^{\dim G}},$$

where c is the constant from Lemma 14 and C is as above. Write

$$\varphi_i = M_i \sum_{w:|u_i-w|<Dr_1} m_{\chi_w}(\psi_{u_i,r_1})\chi_w,$$

i.e. we removed from ψ_{u_i,r_1} those irreducible components whose multiplicities we cannot bound below. Hence $m_{\chi_w}(\varphi_i) \geq C \cdot \dim(\chi_w)$ if $|w - u_i| < Dr_1$ by Lemma 14 (applied with $r = r_1$).
It follows from Lemma 13 (applied with $r = r_2$) that

$$\chi_{r_2} \sqsubseteq \sum_{i=0}^{m} \varphi_i.$$

Moreover, we have

$$\varphi_i \sqsubseteq M_i \psi_{u_i,r_1} \leq \left[M_i \sum_{w:|u_i - w| < 3Dr_1} \dim(\chi_w) \right] \cdot \chi_{r_1},$$

because χ_{r_1} is non-negative, and

$$\sum_{w:|u_i - w| < 3Dr_1} \chi_w(g) \leq \sum_{w:|u_i - w| < 3Dr_1} \dim \chi_w$$

for every $g \in G$. Clearly

$$\sum_{w:|u_i - w| < 3Dr_1} \dim \chi_w \leq C r_1^{\dim \mathrm{LT}} \max_{w:|u_i - w| < 3Dr_1} \{\dim \chi_w\},$$

hence

$$\varphi_i \preceq C \frac{(\max_{w:|u_i - w| < 3Dr_1} \{\dim \chi_w\})^2}{r_1^{2|R_+|}} \cdot \chi_{r_1}. \tag{27}$$

Denote by N_i the number of positive roots $v \in R_+$ such that $\langle u_i, v \rangle \leq 4Dr_1|v|$. Then it follows from Weyl's dimension formula that

$$\max_{w:|u_i - w| < 3Dr_1} \dim \chi_w \leq C r_1^{N_i} \min_{w:|u_i - w| < 3Dr_1} \dim \chi_w. \tag{28}$$

After these preparations, we can give an estimate on $\varphi_i(\mu^{*(l_{r_2})})$. This is done in the next two Lemmata.

Lemma 15. *Fix $0 \leq i \leq m$ such that $N_i < |R_+|$. If r_1 is sufficiently large and*

$$\log^{1/3} r_1 \leq \log r_2 - \log r_1 \leq \log^{1/2} r_1,$$

then

$$\varphi_i \left(\mu^{*(l_{r_2})} \right) \leq \left(\frac{r_1}{r_2} \right)^{(A-1)(|R_+| - N_i)}.$$

Recall that the value of A is given in the statement of Theorem 6 and it also appears in the definition of l_r above.

Proof. For notational simplicity, write

$$X = \max_{w:|u_i - w| < 3Dr_1} \dim \chi_w,$$

and note that by Weyl's dimension formula, we have

$$X \le C r_1^{N_i} r_2^{|R_+| - N_i}. \tag{29}$$

By (27), the induction hypothesis and Lemma 10, we have

$$\varphi_i \left(\mu^{*(l_{r_1})} \right) \le CE \frac{X^2}{r_1^{2|R_+|}}. \tag{30}$$

Denote by $\lambda_{\max}$ the maximum of the absolute values of the eigenvalues of the operators $\pi_v(\mu)$ for $|v - u_i| \le Dr_1$, i.e. for the irreducible characters contained in φ_i. Clearly

$$\lambda_{\max}^{l_{r_1}} \le CE \frac{X^2}{r_1^{2|R_+|} m_{\chi_v}(\varphi_i)}, \tag{31}$$

where χ_v is the character of the representation, which contains $\lambda_{\max}$. Recall that by Lemma 14 and the definition of φ_i, we have

$$m_{\chi_v}(\varphi_i) \ge c \dim \chi_v.$$

If we combine this with (28) and (31), we get

$$\lambda_{\max}^{l_{r_1}} \le CE \frac{X}{r_1^{2|R_+| - N_i}} \le CE \frac{r_2^{|R_+| - N_i}}{r_1^{2|R_+| - 2N_i}}. \tag{32}$$

(For the second inequality we used (29).)
By (30), we clearly have

$$\begin{aligned} \varphi_i \left(\mu^{*(l_{r_2})} \right) &\le CE \frac{X^2}{r_1^{2|R_+|}} \cdot \lambda_{\max}^{l_{r_2} - l_{r_1}} \\ &\le CE \frac{r_2^{2|R_+| - 2N_i}}{r_1^{2|R_+| - 2N_i}} \cdot \left(CE \frac{r_2^{|R_+| - N_i}}{r_1^{2|R_+| - 2N_i}} \right)^{\frac{l_{r_2} - l_{r_1}}{l_{r_1}}}. \end{aligned} \tag{33}$$

By a computation very similar to (12), we get

$$\frac{l_{r_2} - l_{r_1}}{l_{r_1}} \ge \left(A + 1 + \frac{c}{\log^{1/10} r_1} \right) \frac{\log r_2 - \log r_1}{\log r_1}, \tag{34}$$

where c is an absolute constant.
Now we write

$$(\log r_2 - 2 \log r_1) \frac{\log r_2 - \log r_1}{\log r_1} = (\log r_1 - \log r_2) \left(1 - \frac{\log r_2 - \log r_1}{\log r_1} \right). \tag{35}$$

and

$$\left(A+1+\frac{c}{\log^{1/10} r_1}\right)\left(1-\frac{\log r_2-\log r_1}{\log r_1}\right) \geq A+1+\frac{c}{2\log^{1/10} r_1} \tag{36}$$

which follows from our assumption $\log r_2 - \log r_1 \leq \log^{1/2} r_1$.
Combining (34), (35) and (36) we get

$$\left(\frac{r_2^{|R_+|-N_i}}{r_1^{2|R_+|-2N_i}}\right)^{\frac{l_{r_2}-l_{r_1}}{l_{r_1}}} \leq \left(\frac{r_1}{r_2}\right)^{(A+1)(|R_+|-N_i)+c\cdot\log^{-1/10} r_1}.$$

If we plug this into (33), we get

$$\varphi_i\left(\mu^{*(l_{r_2})}\right) \leq \left(\frac{r_1}{r_2}\right)^{(A-1)(|R_+|-N_i)+c\cdot\log^{-1/10} r_1} (CE)^{l_{r_2}/l_{r_1}}.$$

Finally we note that by $\log r_2 - \log r_1 \geq \log^{1/3} r_1$, we have

$$\left(\frac{r_2}{r_1}\right)^{c/\log^{1/10} r_1} \geq e^{c\log^{7/30} r_1} \geq (CE)^2$$

if r_1 is sufficiently large depending on G and E. On the other hand, by the computation leading to (34), we have $l_{r_2}/l_{r_1} < 2$. This finishes the proof. □

Lemma 16. *Fix $0 \leq i \leq m$ such that $N_i = |R_+|$. If $\log r_2 - \log r_1 > \log^{1/3} r_1$ and r_1 is sufficiently large then*

$$\varphi_i(\mu^{*(l_{r_2})}) < C,$$

where C is a constant depending only on G (and not on E).

Proof. Let $v \in \overline{K} \cap I^*$, such that $m_{\chi_v}(\varphi_i) > 0$. By Weyl's formula, we have $\dim(\chi_v) \leq Cr_1^{|R_+|}$, since $N_i = |R_+|$. Let λ be an eigenvalue of $\pi_v(\mu)$. Then by (31) we get

$$\lambda^{l_{r_1}} \leq \frac{CE}{\dim \chi_v}.$$

In fact, we can get a better bound if $|v| \leq r_1$. By a similar argument and using the induction hypothesis for $r = |v| \leq r_1$, we get

$$\lambda^{l_{|v|}} \leq \frac{CE}{\dim \chi_v}.$$

Weyl's dimension formula gives $\dim \chi_v \geq c|v|$, and an easy calculation shows that

$$\lambda^{l_{r_1}} \leq \left(\frac{CE}{|v|}\right)^{l_{r_1}/l_{|v|}} \leq \frac{CE}{r_1} \tag{37}$$

if $|v|$ is sufficiently large (depending on G and E).
We suppose that r_0 is so large that $|v| < r_0$ for those v which are too small for the above argument. We set $C_0 > \mathrm{gap}_{r_0}^{-1}(G,\mu)$, hence for $r_0 \geq |v| \neq 0$ we have

$$\lambda^{l_{r_1}} \leq e^{-\mathrm{gap}_{r_0}(G,\mu) l_{r_1}} \leq e^{-\log^{A+1}(r_1)},$$

which is stronger than (37).
By (27), the induction hypothesis and Lemma 10, we have

$$\varphi_i\left(\mu^{*(l_{r_1})}\right) \leq CE,$$

hence

$$\varphi_i\left(\mu^{*(l_{r_1})}\right) \leq m_{\chi_0}(\varphi_i) + CE\left(\frac{CE}{r_1}\right)^{\frac{l_{r_2}-l_{r_1}}{l_{r_1}}}.$$

Since $\log r_2 - \log r_1 > \log^{1/3} r_1$ and

$$\frac{l_{r_2}-l_{r_1}}{l_{r_1}} > c\frac{\log r_2 - \log r_1}{\log r_1}$$

we easily get

$$\varphi_i\left(\mu^{*(l_{r_1})}\right) \leq m_{\chi_0}(\varphi_i) + 1$$

if r_1 is sufficiently large (depending on G and E) which was to be proved. □

It is left to estimate the number of u_i for which N_i takes a particular value. This is done with the help of the next lemma.
Denote by S the set of simple roots. (This is not to be confused with the generating set of the random walk, which is denoted by S in other sections.)

Lemma 17. *Let $S' \subsetneq S$ be a set of simple roots. Denote by $R' \subset R_+$ the set of all positive roots which can be expressed as a combination of elements of S'. We have*

$$\frac{|S|-|S'|}{|R_+|-|R'|} \leq A(G) - 1,$$

where the value of $A(G)$ is given in Table 1, for simple Lie groups and for non-simple ones it is defined to be $A(G) = \max\{A(H)\}$ where H runs through all simple quotients.

Proof. If the Dynkin diagram of G is not connected, we can write $R_+ = R_1 \cup \ldots \cup R_n$, where R_i is a system of positive roots in a root system with connected diagram. Clearly

$$\frac{|S|-|S'|}{|R_+|-|R'|} \leq \max_{1\leq i\leq n} \frac{|S\cap R_i| - |S'\cap R_i|}{|R_i| - |R'\cap R_i|}.$$

Hence we can assume without loss of generality that the diagram of G is connected.

By a simple calculation, one can verify that $|S|/|R_+| = A(G) - 1$ as given in Table 1. Now notice that R' is itself (the set of positive roots in) a root system and its diagram is the subgraph spanned by S' in the diagram of G. Examining Table 1 it is easy to check that

$$\frac{|S'|}{|R'|} \geq A(G) - 1 = \frac{|S|}{|R_+|}.$$

(In fact, it is enough to check that the value in the table is never smaller for connected subdiagrams.) Then

$$\frac{|S| - |S'|}{|R_+| - |R'|} \leq \frac{|S| - |S||R'|/|R_+|}{|R_+| - |R'|} = \frac{|S|}{|R_+|} = A(G) - 1.$$

□

Let $S' \subset S$ be a subset of the simple roots. We write $\Omega(S')$ for the set of indices i such that $\langle u_i, v\rangle \leq 4Dr_1$ for $v \in S$ if and only if $v \in S'$.

We estimate $|\Omega(S')|$. Since S' consist of linearly independent vectors, the elements of $\Omega(S')$ are in a Cr_1 neighborhood of a subspace of $\mathrm{L}T$ of dimension $|S| - |S'|$. Since they are Dr_1-separated, we have

$$|\Omega(S')| \leq C\left(\frac{r_2}{r_1}\right)^{|S|-|S'|}. \tag{38}$$

All positive roots are positive linear combinations of simple roots, hence $N_i \leq |R'|$ for $i \in \Omega(S')$, where R' is the set of positive roots which are combinations of the elements of S'. If $S' \neq S$, then Lemmata 15 and 17 together with (38) gives

$$\sum_{i\in\Omega(S')} \varphi_i\left(\mu^{*(l_{r_2})}\right) \leq C$$

with a constant C depending on G. If $S' = S$, the same follows from Lemma 16. Summing this up for all $S' \subset S$, we get

$$\sum_i \varphi_i\left(\mu^{*(l_{r_2})}\right) \leq C. \tag{39}$$

This completes the proof of (26) for $r = r_2$ with $E = C$, where C is the constant in (39).

We explain how to set the various parameters and how to complete the induction. We set $E = C$ with the constant C from (39) in the previous paragraph. Then we pick r_0 to be sufficiently large (depending on E and G) so that all of the above arguments are valid with $r_1 \geq r_0$.

For $r_0 \geq r \geq 1$, we have

$$\begin{aligned} \chi_r\left(\mu^{*(l_r)}\right) &\leq m_{\chi_0}(\chi_r) + e^{-l_r \cdot \mathrm{gap}_{Cr}(G,\mu)} \cdot \sum_{0<|v|\leq Cr} m_{\chi_v}(\chi_r) \\ &\leq C + Cr^{-C_0 \cdot \mathrm{gap}_{Cr}(G,\mu)} \cdot \sum_{0<|v|\leq Cr} \dim(\chi_v) \\ &\leq C + Cr^{C(\dim LT + |R_+|) - C_0 \cdot \mathrm{gap}_{Cr}(G,\mu)}. \end{aligned} \tag{40}$$

Here we first used the definition of $\mathrm{gap}_r(G,\mu)$, then the definition of l_r and Lemmata 12 and 13, finally Weyl's dimension formula.
We put $C_0 = C \cdot \mathrm{gap}^{-1}_{Cr_0}(G,\mu)$, where C is a suitable constant depending on the constant in (40) such that

$$\chi_r\left(\mu^{*(l_r)}\right) \leq E$$

for $1 \leq r \leq r_0$. (Recall that the only constraint we had for C_0 above is in the proof of Lemma 16 and it is satisfied with this choice.)
Thus we see that (26) hold for $1 \leq r \leq r_0$. Once we know that (26) holds on an interval $r \in [1,a]$, we can extend it to $r \in [1, e^{\log(a)+\log^{1/2}(a)}]$. This follows from the above argument with the choice $r_2 = r$ and any $r_1 \leq a$ which satisfies

$$\log^{1/3} r_1 \leq \log r_2 - \log r_1 \leq \log^{1/2} r_1.$$

If we apply this repeatedly, we can conclude that (26) holds for all $r \geq 1$.
Finally, we conclude the proof of the theorem. Fix $r \geq 1$ and suppose that π_{v_0} is the representation for which the maximum in the definition of $\mathrm{gap}_r(G,\mu)$ is attained. We use Lemma 14 with $z = u = v_0$ and (26) to get

$$\begin{aligned} \chi_{v_0}\left(\mu^{*(l_r)}\right) &\leq C\frac{\max_{w:|v_0-w|<3Dr}\{\dim(\chi_w)\}}{\dim(\chi_{v_0}) \cdot r^{\dim G}} \cdot \psi_{v_0,r}\left(\mu^{*(l_r)}\right) \\ &\leq CE\frac{(\max_{w:|v_0-w|<3Dr}\{\dim(\chi_w)\}) \cdot \sum_{w:|v_0-w|<3Dr}\dim(\chi_w)}{\dim(\chi_{v_0}) \cdot r^{\dim G}}. \end{aligned}$$

We evaluate dimensions using Weyl's formula and get

$$\left\|\pi_{v_0}\left(\mu^{*(l_r)}\right)\right\| \leq \chi_{v_0}\left(\mu^{*(l_r)}\right) \leq CE/\dim(\chi_{v_0}) \leq CE/r \leq r^{-1/2},$$

if $r \geq r_0$ and r_0 is sufficiently large, as we may assume. This implies

$$\mathrm{gap}_r(G,\mu) \geq (1/10)\frac{\log r}{l_r}.$$

Inspecting the definition of l_r and the above choice of C_0, we see that this is exactly what was to be proved.

4 Some technicalities

We begin this section by proving Corollary 7. First we give a lemma which will be used for reducing the problem to the connected case:

Lemma 18. *Let G be a Lie group and let μ be a symmetric probability measure on it such that* $\operatorname{supp}\mu$ *generates a dense subgroup and* $1 \in \operatorname{supp}\mu$*. Write G° for the connected component of G and let $n = [G : G^\circ]$. Then* $\operatorname{supp}(\mu^{*(2n-1)}) \cap G^\circ$ *generates a dense subgroup of G°.*

Proof. Suppose to the contrary that for some $h \in G^\circ$ and $\varepsilon > 0$ there is no h' in the group generated by $\operatorname{supp}(\mu^{*(2n-1)}) \cap G^\circ$ such that $\operatorname{dist}(h, h') < \varepsilon$. On the other hand, by assumption, there is $g = g_1 \cdots g_l$ with $g_i \in \operatorname{supp}(\mu)$ and $\operatorname{dist}(h, g) < \varepsilon$.
We show that g is in the group generated by $\operatorname{supp}(\mu^{*(2n-1)}) \cap G^\circ$, a contradiction. If $l \leq 2n-1$, then by definition, $g \in \operatorname{supp}(\mu^{*(2n-1)}) \cap G^\circ$. Suppose $l > 2n-1$. By the pigeon hole principle, there are $i \leq j \leq n$ such that $g_i \cdots g_j \in G^\circ$. We can write $g = g' \cdot g''$, where

$$\begin{aligned} g' &= g_1 \cdots g_{i-1} g_i \cdots g_j g_{i-1}^{-1} \cdots g_1^{-1} \quad \text{and} \\ g'' &= g_1 \cdots g_{i-1} g_{j+1} \cdots g_l. \end{aligned}$$

Now clearly $g', g'' \in G^\circ$, and $g' \in \operatorname{supp}(\mu^{*(2n-1)})$, since it is a product of length at most $2n-1$. Since the length of g'' is strictly less than that of g the proof can be completed by induction. □

If G is a compact connected semi-simple Lie group, the space $L^2(G)$ can be decomposed as an orthogonal sum of finite dimensional irreducible representations. We write $\mathcal{H}_r \subset L^2(G)$ for the sum of those constituents which have highest weight v with $|v| \leq r$. In the next lemma we construct an approximate identity in $\mathcal{H}_r$.

Lemma 19. *Let G be a compact connected semi-simple Lie group. Then for each r, there is a non-negative function $f_r \in \mathcal{H}_r$ such that*

$$\int f_r \, dm_G = 1, \quad \|f_r\|_2 \leq C r^{\dim G/4} \quad \text{and} \quad \int f_r(g) \operatorname{dist}(g, 1) \, dm_G(g) \leq C/\sqrt{r},$$

where C is a constant depending on G.

Proof. We fix a maximal torus $T \subset G$. Let π be a faithful (not necessarily irreducible) finite dimensional unitary representation of G with real character χ. We can decompose the representation space as the sum of weight spaces, i.e. there is an orthonormal basis $\varphi_1, \ldots, \varphi_m$ (where $m = \dim \pi$) such that the following holds. For each φ_i, there is a weight $u_j \in I^*$ such that for the elements $g \in T$ we have $\pi(g)\varphi_j = e^{2\pi i \langle \log g, u_j \rangle} \varphi_j$. (Here $\log : T \to LT$ is a branch of the inverse of the exponential map.) Since χ is real, we have

$$\chi(g) = \sum_{j=1}^m e^{2\pi i \langle \log g, u_j \rangle} = \sum_{j=1}^m \cos(2\pi \langle \log g, u_j \rangle).$$

Since π is faithful, $\chi(g) < m$ for $g \neq 1$ and we can deduce from the above formula that

$$m - c_1 \mathrm{dist}(g,1)^2 \leq \chi(g) \leq m - c_2 \mathrm{dist}(g,1)^2, \tag{41}$$

where $c_1, c_2 > 0$ are constants depending on G.
Denote by r_0 the length of the highest weight of the irreducible components of π. We define

$$f_r(g) = c_r(\chi(g) + m)^{\lfloor r/r_0 \rfloor},$$

where c_r is a normalizing constant so that $\int f_r \, dm_G = 1$. Observe that $f_r \in \mathcal{H}_r$.
By simple calculation based on (41), we get

$$c \frac{\lfloor r/r_0 \rfloor^{\dim G/2}}{(2m)^{\lfloor r/r_0 \rfloor}} \leq c_r \leq C \frac{\lfloor r/r_0 \rfloor^{\dim G/2}}{(2m)^{\lfloor r/r_0 \rfloor}},$$

where $c, C > 0$ are numbers depending on G.
We have

$$\|f_r\|_\infty = c_r (2m)^{\lfloor r/r_0 \rfloor} \leq C r^{\dim G/2}.$$

The L^2 bound now follows from $\|f_r\|_2^2 \leq \|f_r\|_1 \|f_r\|_\infty$.
Now using again (41), we get

$$\begin{aligned}
&\int f_r(g) \mathrm{dist}(1,g) \, dm_G(g) \\
&\qquad \leq \quad c_r \int (2m - c_2 \mathrm{dist}(g,1)^2)^{\lfloor r/r_0 \rfloor} \mathrm{dist}(1,g) \, dm_G(g) \\
&\qquad \leq \quad C \lfloor r/r_0 \rfloor^{\dim G/2} \int e^{-c_3 \mathrm{dist}(g,1)^2 \lfloor r/r_0 \rfloor} \mathrm{dist}(1,g) \, dm_G(g) \\
&\qquad \leq \quad C \lfloor r/r_0 \rfloor^{\dim G/2} \int_{\mathbf{R}^{\dim G}} e^{-c_3 |x|^2 \lfloor r/r_0 \rfloor} |x| \, dx \\
&\qquad = \quad C \lfloor r/r_0 \rfloor^{-1/2} \int_{\mathbf{R}^{\dim G}} e^{-c_3 |y|^2} |y| \, dy \leq C/\sqrt{r},
\end{aligned}$$

which was to be proved. □

Proof of Corollary 7. Write $\mathrm{Reg}(g)f(h) = f(g^{-1}h)$ for $f \in L^2(G)$, which is the left regular representation of G.
Assume to the contrary that $f \in \mathrm{Lip}(G)$, $\int f = 0$, $\|f\|_2 = 1$ and yet

$$\left\| \int \mathrm{Reg}(g) f \, d\mu(g) \right\|_2 = \|\mathrm{Reg}(\mu) f\|_2 \geq 1 - c_0 \log^{-A}(\|f\|_{\mathrm{Lip}} + 2) \tag{42}$$

with a constant c_0 which will be chosen to be sufficiently small depending on μ.
By the same argument as in the beginning of Section 2, we have $\|\mathrm{Reg}(\mu)f\|_2^2 = \|\mathrm{Reg}(\widetilde{\mu} * \mu) f\|_2$. Thus (42) holds (with a different c_0) for μ replaced by $\widetilde{\mu} * \mu$, hence we can assume that μ is symmetric and $1 \in \operatorname{supp} \mu$.

Write $n = [G : G^\circ]$ as in Lemma 18. Furthermore, we define μ_1 to be the probability measure on G° that we obtain from $\mu^{*(2n-1)}$ by restricting it to G° and then normalizing it. By Lemma 18, we can apply Theorem 6 for the measure μ_1. Now we need to find a suitable test function related to f.

First, we want to rule out the possibility that f is "almost constant" on cosets of G°. It follows from (42) that there is a set $X \subset G$ with $\mu(X) > 1 - \varepsilon$ such that

$$\|f - \mathrm{Reg}(g)f\|_2 < 1/(10n) \tag{43}$$

where $\varepsilon > 0$ is as small as we wish, if we choose c_0 in (42) sufficiently small. (In fact, we could obtain a much stronger estimate from (42)). In particular, we can ensure that $X \cdot G^\circ$ generates G/G°.

Since $\int f \, dm_G = 0$,

$$\int \langle \mathrm{Reg}(g)f, f\rangle \, dm_G(g) = 0.$$

Hence there is $g \in G$ such that $\langle \mathrm{Reg}(g)f, f\rangle \leq 0$, in particular $\|\mathrm{Reg}(g)f - f\|_2 \geq \sqrt{2}$. We can write $g = h_1 \cdots h_n g_0$, where $h_i \in X$ and $g_0 \in G^\circ$. By the triangle inequality, either

$$\|\mathrm{Reg}(h_1 \cdots h_n g_0)f - \mathrm{Reg}(h_1 \cdots h_n)f\|_2 > 1 \tag{44}$$

or

$$\|\mathrm{Reg}(h_1 \cdots h_j)f - \mathrm{Reg}(h_1 \cdots h_{j-1})f\|_2 > (\sqrt{2} - 1)/n \tag{45}$$

for some $1 \leq j \leq n$. Since Reg is unitary, the second case yields $\|\mathrm{Reg}(h_j)f - f\|_2 > (\sqrt{2} - 1)/n$ which is a contradiction to (43). Thus only the first case is possible, from which we conclude $\|\mathrm{Reg}(g_0)f - f\|_2 > 1$ again by unitarity.

This shows that f can not be "almost constant" on all cosets of G°. Let $x_1, \ldots, x_n \in G$ be a system of representatives for G°-cosets. Write f_i for the restriction of f to the coset $G^\circ \cdot x_i$ considered a function on G°. More formally:

$$f_i(g) = f(gx_i) \in L^2(G^\circ).$$

We have

$$\|\mathrm{Reg}(g_0)f - f\|_2^2 = \sum_{i=1}^{n} \|\mathrm{Reg}(g_0)f_i - f_i\|_2^2,$$

hence there is $1 \leq i_0 \leq n$ such that $\|\mathrm{Reg}(g_0)f_{i_0} - f_{i_0}\|_2 \geq 1/\sqrt{n}$.

We define

$$\varphi = \frac{f_{i_0} - \int f_{i_0} \, dm_G}{\|f_{i_0} - \int f_{i_0} \, dm_G\|_2}$$

aiming to estimate $\|\mathrm{Reg}(\mu_1)\varphi\|_2$ using Theorem 6.

It follows from the above considerations that $\|f_{i_0} - \int f_{i_0} \, dm_G\|_2 \geq 1/(2\sqrt{n})$. Thus $\|\varphi\|_{\mathrm{Lip}} < 2\sqrt{n}\|f\|_{\mathrm{Lip}}$.

From Lemma 18, we know that $\operatorname{supp} \mu_1$ is not contained in a proper closed subgroup of G°. Thus we have

$$\mathrm{gap}_r(G^\circ, \mu_1) > 0$$

for all r. Then Theorem 6 implies that

$$\mathrm{gap}_r(G^\circ, \mu_1) > c \log^{-A(G)} r$$

with a constant $c > 0$ depending on μ.
To apply this spectral gap estimate, we need to approximate φ by a function in $\mathcal{H}_r$ with small r. We use Lemma 19 with $r = D(\|\varphi\|_{\mathrm{Lip}} + 2)^4$; we will chose the sufficiently large number D later. Then the Lemma gives:

$$\|f_r * \varphi - \varphi\|_\infty \leq \int f_r(g)\mathrm{dist}(g, 1)\|\varphi\|_{\mathrm{Lip}}\, dm_G(g) \leq \frac{C}{D^{1/2}(\|\varphi\|_{\mathrm{Lip}} + 2)}.$$

Clearly $f_r * \varphi \in \mathcal{H}_r$, and moreover

$$\int f_r * \varphi\, dm_G = 0.$$

Thus

$$\begin{aligned}
\|\mathrm{Reg}(\mu_1)\varphi\|_2 &\leq \|\mathrm{Reg}(\mu_1)(f_r * \varphi)\|_2 + \frac{C}{D^{1/2}(\|\varphi\|_{\mathrm{Lip}} + 2)} \\
&\leq 1 - \mathrm{gap}_r(G^\circ, \mu_1) + \frac{C}{D^{1/2}(\|\varphi\|_{\mathrm{Lip}} + 2)} \\
&\leq 1 - c \log^{-A(G)}(D^{1/2}(\|\varphi\|_{\mathrm{Lip}} + 2)) + \frac{C}{D^{1/2}(\|\varphi\|_{\mathrm{Lip}} + 2)}.
\end{aligned}$$

Now we choose D sufficiently large depending on μ so that for the quantities in the last line we have

$$c \log^{-A(G)}(D^{1/2}(\|\varphi\|_{\mathrm{Lip}} + 2)) \geq \frac{2C}{D^{1/2}(\|\varphi\|_{\mathrm{Lip}} + 2)}.$$

This in turn implies with a different constant c:

$$\|\mathrm{Reg}(\mu_1)\varphi\|_2 \leq 1 - c \log^{-A(G)}(\|\varphi\|_{\mathrm{Lip}} + 2).$$

Furthermore, by the definition of φ, we have

$$\|\mathrm{Reg}(\mu_1) f_{i_0}\|_2 \leq \|f_{i_0}\|_2 - c \log^{-A(G)}(\|f\|_{\mathrm{Lip}} + 2).$$

This yields

$$\|\mathrm{Reg}(\mu_1) f\|_2 \leq 1 - c \log^{-A(G)}(\|f\|_{\mathrm{Lip}} + 2).$$

By selfadjointness

$$\|\mathrm{Reg}(\mu) f\|_2^{2n-1} \leq \|\mathrm{Reg}(\mu^{*(2n-1)}) f\|_2 \leq 1 - c'(1 - \|\mathrm{Reg}(\mu_1) f\|_2),$$

where $c' = \mu^{*(2n-1)}(G^\circ)$ is the normalizing constant in the definition of μ_1. This is a contradiction to (42), if we choose there c_0 to be sufficiently small. $\square$

The rest of the section is devoted to the proof of Lemma 5. We start with a Lemma which provides an estimate on the Lipschitz norm of a function in $\mathcal{H}_r$.

Lemma 20. *Let G be a connected compact semi-simple Lie group and let $f \in \mathcal{H}_r$ with $\|f\|_2 = 1$. Then $\|f\|_{\mathrm{Lip}} \le Cr^{\dim G/2+1}$, where C is a constant depending on G.*

Proof. We fix a maximal torus T and write

$$\|f\|_{\mathrm{Lip}(T)} = \max_{t\in T, g\in G}\left\{\frac{f(tg)-f(g)}{\mathrm{dist}(1,t)}\right\}$$

for $f \in \mathrm{Lip}(G)$. Since every element of G is contained in a maximal torus, it is enough to bound this new semi-norm which a priori could be smaller.
As in the proof of Lemma 19, we decompose $\mathcal{H}_r$ as the sum of weight spaces for T. There is an orthonormal basis $\varphi_1, \ldots, \varphi_m \in \mathcal{H}_r$ (where $m = \dim \mathcal{H}_r$) such that for $g \in T$ we have

$$\pi(g)\varphi_j = e^{2\pi i\langle \log g, u_j\rangle}\varphi_j,$$

where $u_j \in I^*$ is a weight and $|u_j| \le r$. Thus $\|\varphi_j\|_{\mathrm{Lip}(T)} \le Cr$. We can write $f = \sum \alpha_j \varphi_j$ such that $\sum \alpha_j^2 = 1$. Then

$$\|f\|_{\mathrm{Lip}(T)} \le Cr \sum \alpha_j \le Cr\sqrt{m}.$$

By Weyl's dimension formula it follows that each irreducible representation in $\mathcal{H}_r$ is of dimension at most $Cr^{|R_+|}$ and each appears with multiplicity equal to its dimension. The number of irreducible components is at most $Cr^{\dim T}$. Putting these together we get $m \le Cr^{\dim G}$ which proves the Lemma. $\square$

Proof of Lemma 5. First we bound the diameter in terms of the spectral gap. Let $\varepsilon > 0$ be a number, and $g_0 \in G$. Suppose that for some integer l, there is no $g \in S^l$ such that $\mathrm{dist}(g, g_0) \le \varepsilon$.
We fix a number D that we will specify later, and write $r = D\varepsilon^{-2\dim G-2}$ and let f_r be the approximate identity constructed in Lemma 19. Then

$$\begin{aligned}\int\limits_{\mathrm{dist}(g,g_0)<\frac{\varepsilon}{2}} \mathrm{Reg}(\mu_S)^l f_r(g)\, dm_G(g) &= \int\limits_{\mathrm{dist}(g,g_0)<\frac{\varepsilon}{2}} \sum_{h\in S^l} \mu_S^{*(l)}(h) f_r(h^{-1}g)\, dm_G(g) \\ &\le \int\limits_{\mathrm{dist}(g,1)>\frac{\varepsilon}{2}} f_r(g)\, dm_G(g) \le \frac{C}{\varepsilon\sqrt{r}}. \end{aligned} \tag{46}$$

For the inequality between the first and second lines, we used that $\mathrm{dist}(h, g) \ge \varepsilon/2$, hence $\mathrm{dist}(1, h^{-1}g) \ge \varepsilon/2$.
On the other hand, we have

$$\|1 - \mathrm{Reg}(\mu_S)^l f_r\|_2 \le Cr^{\dim G/4} e^{-l\cdot \mathrm{gap}_r(G,S)}.$$

(Recall $\|f_r\|_2 \le Cr^{\dim G/4}$ from Lemma 19.) Hence

$$\begin{aligned}\int\limits_{\mathrm{dist}(g,g_0)<\frac{\varepsilon}{2}} \mathrm{Reg}(\mu_S)^l f_r(g)\, dm_G(g) &\ge \int\limits_{\mathrm{dist}(g,g_0)<\frac{\varepsilon}{2}} 1\, dm_G(g) - \frac{Cr^{\dim G/4}}{e^{l\cdot \mathrm{gap}_r(G,S)}}\\ &\ge c\varepsilon^{\dim G}. \qquad (47)\end{aligned}$$

provided

$$l > \frac{\log\left(2Cr^{\dim G/4}/(c\varepsilon^{\dim G})\right)}{\mathrm{gap}_r(G,S)} \ge C' \cdot \frac{\log(\varepsilon^{-1})}{\mathrm{gap}_r(G,S)},$$

where C' depends on G and D. Now we choose D such that

$$\frac{C}{\varepsilon\sqrt{r}} \le c\varepsilon^{\dim G},$$

where C and c are the constants form (46) and (47), respectively. This is impossible. We can conclude

$$\mathrm{diam}(G,S) \le \frac{C\log(\varepsilon^{-1})}{\mathrm{gap}_{D\varepsilon^{-2\dim G-2}}(G,S)}.$$

Now we estimate the spectral gap in terms of the diameter. This argument was communicated to me by Jean Bourgain. Let $r > 0$ be a number, and set $\varepsilon = Dr^{-\dim G/2-1}$, where $D > 0$ depends on G and will be set later. Let $f \in \mathcal{H}_r$ and assume that $\|f\|_2 = 1$ and $\int f = 0$. Then $\int \langle \mathrm{Reg}(g)f, f\rangle\, dm_G(g) = 0$, hence there is $g \in G$ such that $\langle \mathrm{Reg}(g)f, f\rangle \le 0$. Thus $\|\mathrm{Reg}(g)f - f\|_2 \ge \sqrt{2}$.
Let $l = \mathrm{diam}_\varepsilon(G,S)$ and $g_0 = g_1\cdots g_l \in S^l$ such that $\mathrm{dist}(g,g_0) \le \varepsilon$. By Lemma 20 we have

$$\begin{aligned}\|\mathrm{Reg}(g_0)f - f\|_2 &\ge \|\mathrm{Reg}(g)f - f\|_2 - \|\mathrm{Reg}(g_0)f - \mathrm{Reg}(g)f\|_2\\ &\ge \sqrt{2} - \varepsilon Cr^{\dim G/2+1} \ge 1\end{aligned}$$

if we choose D to be sufficiently small in the definition of ε.
By the triangle inequality, there is $1 \le j \le l$ such that

$$\|\mathrm{Reg}(g_j)f - f\|_2 = \|\mathrm{Reg}(g_1\cdots g_j)f - \mathrm{Reg}(g_1\cdots g_{j-1})f\|_2 \ge 1/l.$$

This implies

$$\|\mathrm{Reg}(g_j)f + f\|_2 \le 2 - 1/l^2.$$

Finally, we can conclude

$$\begin{aligned}\|\mathrm{Reg}(\mu_S)f\|_2 &\le \frac{1}{|S|}\left(\|\mathrm{Reg}(g_j)f + f\|_2 + \left\|\sum_{g\in S\setminus\{1,g_j\}} \mathrm{Reg}(g)f\right\|_2\right)\\ &\le 1 - \frac{1}{|S|\mathrm{diam}_\varepsilon(G,S)^2}\end{aligned}$$

which was to be proved. (Recall the assumption $1 \in S$.) □

REFERENCES

[1] L. Babai and Á. Seress, *On the diameter of permutation groups*, European J. Combin. 13 No. 4 (1992), 231–243. 1139

[2] J. Bourgain and A. Gamburd, *Uniform expansion bounds for Cayley graphs of $SL_2(\mathbb{F}_p)$,* Ann. of Math. 167 (2008), 625–642. 1140, 1146

[3] J. Bourgain and A. Gamburd, *On the spectral gap for finitely-generated subgroups of $SU(2)$* Invent. math. 171 (2008), no. 1, 83–121. 1144, 1145

[4] J. Bourgain and A. Gamburd, *Expansion and random walks in $SL_d(\mathbb{Z}/p^n\mathbb{Z})$:I*, J. Eur. Math. Soc. 10 (2008), 987–1011. 1140, 1146, 1153

[5] J. Bourgain and A. Gamburd, *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 No. 5. (2009), 1057–1103. 1140, 1146

[6] J. Bourgain and A. Gamburd, *A spectral gap theorem in $SU(d)$*, J. Eur. Math. Soc. (JEMS) 14 No. 5 (2012), 1455-1511. 1144, 1145

[7] J. Bourgain, A. Gamburd and P. Sarnak, *Affine linear sieve, expanders, and sum-product*, Invent. math., 179 No. 3. (2010), 559–644. 1140, 1146

[8] J. Bourgain and A. Kontorovich, *On the local–global conjecture for integral Apollonian gaskets, With an appendix by Péter P. Varjú*, Invent. math., to appear http://dx.doi.org/10.1007/s00222-013-0475-y 1146

[9] J. Bourgain and P. P. Varjú, *Expansion in $SL_d(\mathbf{Z}/q\mathbf{Z})$, q arbitrary*, Invent. math. 188, No 1. (2012) 151–173. 1140, 1142, 1146, 1153, 1154, 1155

[10] E. Breuillard, *Random walks on Lie groups*, preprint, http://www.math.u-psud.fr/~breuilla/part0gb.pdf 1145

[11] E. Breuillard, B. J. Green and T. C. Tao, *Approximate subgroups of linear groups*, Geom. Funct. Anal. 21 (2011) 774–819. 1139, 1142, 1154, 1155

[12] T. Bröcker and T. tom Dieck, *Representations of compact Lie groups*, Graduate Texts in Mathematics, 98 Springer-Verlag, New York, 1995. 1143, 1157, 1158

[13] P. Diaconis and L. Saloff-Coste, *Comparison techniques for random walk on finite groups*, Ann. Probab. 21, No. 4 (1993) 2131–2156. 1139

[14] C. M. Dawson and M. A. Nielsen, *The Solovay-Kitaev algorithm*, Quantum Inf. Comput. 6 No. 1 (2006), 81–95. 1144

[15] O. Dinai, *Poly-log diameter bounds for some families of finite groups,* Proc. Amer. Math. Soc. 134 No. 11. (2006), 3137–3142. 1140, 1142, 1143

[16] O. Dinai, *Diameters of Chevalley groups over local rings*, Arch. Math. (Basel) 99 No. 5 (2012), 417-424. 1140, 1142, 1143

[17] D. Dolgopyat, *On mixing properties of compact group extensions of hyperbolic systems*, Israel J. Math. 130 (2002), 157–205. 1145, 1146

[18] J. S. Ellenberg, C. Hall and E. Kowalski, *Expander graphs, gonality, and variation of Galois representations* Duke Math. J. 161 No. 7, (2012), 1233–1275. 1146

[19] A. Gamburd, D. Jakobson and P. Sarnak, *Spectra of elements in the group ring of* SU(2), J. Eur. Math. Soc. 1, (1999) 51–85. 1144

[20] A. Gamburd, M. Shahshahani, *Uniform diameter bounds for some families of Cayley graphs.* Int. Math. Res. Not. 71 (2004), 3813–3824. 1140, 1142, 1143

[21] W. T. Gowers, *Quasirandom Groups*, Combin. Probab. Comput. 17 (2008), 363–387. 1140, 1144

[22] H. A. Helfgott, *Growth and generation in* $SL_2(\mathbb{Z}/p\mathbb{Z})$, Ann. of Math. 167 (2008), 601–623. 1139, 1140, 1142

[23] H. A. Helfgott, *Growth in* $SL_3(\mathbb{Z}/p\mathbb{Z})$, J. Eur. Math. Soc., 13 No. 3 (2011), 761–851. 1142

[24] H. A. Helfgott, Á. Seress, *On the diameter of permutation groups*, Ann. of Math., to appear, http://arxiv.org/abs/1109.3550v4 1139

[25] D. Kelmer and L. Silberman, *A uniform spectral gap for congruence covers of a hyperbolic manifold*, Amer. J. Math. 135 No 4 (2013), 1067–1085. 1154

[26] A. Yu. Kitaev, A. H. Shen and M. N. Vyalyi, *Classical and quantum computation*, Translated from the 1999 Russian original by Lester J. Senechal. Graduate Studies in Mathematics, 47 American Mathematical Society, Providence, RI, 2002. 1144

[27] A. Lev, *The covering number of the group* $\mathrm{PSL}_n(F)$ J. Algebra 182 No. 1 (1996), 60–84. 1155

[28] A. Lubotzky and C. Meiri, *Sieve methods in group theory I: Powers in linear groups* J. Amer. Math. Soc. 25 No. 4 (2012), 1119–1148. 1146

[29] N. Nikolov and L. Pyber, *Product decompositions of quasirandom groups and a Jordan type theorem*, J. Eur. Math. Soc. 13 No. 4 (2011), 1063–1077. 1141, 1144

[30] L. Pyber, E. Szabó, *Growth in finite simple groups of Lie type of bounded rank*, preprint, http://arxiv.org/abs/1005.1858 1139, 1142, 1154, 1155

[31] A. Salehi Golsefidy and P. P. Varjú, *Expansion in perfect groups*, Geom. Funct. Anal. 22 No. 6 (2012), 1832-1891. 1140, 1146

[32] P. Sarnak and X. X. Xue, *Bounds for multiplicities of automorphic representations*, Duke Math. J., 64 No. 1, (1991), 207–227. 1140

[33] N. de Saxcé, *Trou dimensionel dans les groupes de Lie compacts semisimples via les séries de Fourier*, J. Anal. Math., to appear, http://www.ma.huji.ac.il/∼saxce/dimensionnel.pdf 1144, 1153

[34] P. P. Varjú, *Expansion in $SL_d(O_K/I)$, I square-free,* J. Eur. Math. Soc. 14 No 1. (2012) 273–305. 1140, 1146

[35] P. P. Varjú, *Random walks in Euclidean space*, preprint, http://arxiv.org/abs/1205.3399 1145, 1146

Péter P. Varjú
Centre for Mathematical Sciences
Wilberforce Road
Cambridge CB3 0WA
England
pv270@dpmms.cam.ac.uk
and
The Einstein Institute of Mathematics
Edmond J. Safra Campus
Givat Ram
The Hebrew University of Jerusalem
Jerusalem, 91904
Israel