Semisimple random walks on the torus

Nicolas de Saxcé, Weikun He

SEMISIMPLE RANDOM WALKS ON THE TORUS

WEIKUN HE AND NICOLAS DE SAXCÉ

Abstract. We study linear random walks on the torus and show a quantitative equidistribution statement, under the assumption that the Zariski closure of the acting group is semisimple.

1. Introduction

Let $d \geq 2$ be an integer and $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ the torus of dimension d. We study a random walk $(x_n)_{n\geq 0}$ on $\mathbb{T}^d$ given by

$$\forall n \geq 0, \quad x_n = g_n \dots g_1 x_0$$

where $(g_n)_{n\geq 1}$ is a sequence of independent identically distributed random variables with law μ on $\mathrm{GL}_d(\mathbb{Z})$. Let Γ denote the group generated by the support of μ, and G be the Zariski closure of Γ in $\mathrm{GL}_d(\mathbb{R})$. In their breakthrough paper [10], Bourgain, Furman, Lindenstrauss and Mozes showed that if G acts strongly irreducibly and proximally on $\mathbb{R}^d$, then the random walk $(x_n)_{n\geq 0}$ equidistributes in law to the Haar probability measure $\mathrm{m}_{\mathbb{T}^d}$ as soon as x_0 is irrational, i.e.

$$\forall x_0 \notin \mathbb{Q}^d/\mathbb{Z}^d, \quad \mu^{*n} * \delta_{x_0} \rightharpoonup^*_{n\to+\infty} \mathrm{m}_{\mathbb{T}^d}\,.$$

Moreover, this result is quantitative : the rate of convergence is given in function of how far away x_0 is from rational points of small denominator. Following their strategy, we showed recently in [22] that their theorem was still valid without the proximality assumption, as long as the action of G is strongly irreducible. On the other hand, the theory developed by Benoist and Quint in their series of articles [2, 4, 5, 3] made it clear that when studying random walks on homogeneous spaces, it was most natural to only assume that the acting algebraic group G was semisimple. Indeed, under this assumption they obtained a full classification of stationary measures [4, Theorem 1.1] from which they deduced in [5] very general equidistribution results. It was therefore desirable to obtain quantitative convergence results similar to those of [10] or [22] in this more general setting. This is the goal of the present article.

Of course, without the irreducibility assumption, there can exist some proper closed Γ-invariant subsets of $\mathbb{T}^d$, and the random walk may not equidistribute to the Haar measure, even if the starting point x_0 is irrational. So in order to state our main result, we need to set up some notation. Let G° denote the identity component of G for the Zariski topology. The subalgebra of $\mathcal{M}_d(\mathbb{R})$ generated by G° is denoted by E. Since G is semisimple, we may write $\mathbb{R}^d = V_0 \oplus V_1 \oplus \dots \oplus V_r$ where for $i = 0, \dots, r$, V_i is a maximal sum of simple G-modules having the same top Lyapunov exponent for the action of μ. We assume in addition that

[W.H.] The order does not matter.

$$\lambda_1(\mu, V_1) > \dots > \lambda_1(\mu, V_r) > \lambda_1(\mu, V_0) = 0.$$

2010 *Mathematics Subject Classification.* Primary 37A17, 11B75; Secondary 37A45, 11L07, 20G30.

Key words and phrases. Random walk, Toral automorphism, Equidistribution, Sum-product.

We define a quasi-norm on $\mathbb{R}^d$ by

$$|v| = \max_{0\leq i\leq r} \|v_i\|^{\frac{1}{\lambda_1(\mu,V_i)}}$$

where $v = v_0 + \ldots + v_r$ is the decomposition of v according to the direct sum $\mathbb{R}^d = \oplus_{i=1}^r V_i$. By convention, we set $\frac{1}{0} = +\infty$ and

$$\|v_0\|^{+\infty} = \begin{cases} 0 & \text{if } \|v_0\| \leq 1, \\ +\infty & \text{otherwise.} \end{cases}$$

This quasi-norm induces a quasi-distance on $\mathbb{R}^d$ given by $\tilde{d}(x,y) = |x-y|$, which projects to a quasi-distance on $\mathbb{T}^d$, still denoted by $\tilde{d}$. Our goal is the following theorem.

Theorem 1.1 (Quantitative equidistribution of semisimple linear random walk on $\mathbb{T}^d$)**.** *Let μ be a probability measure on $\mathrm{GL}_d(\mathbb{Z})$ having a finite exponential moment. Denote by $G \subset \mathrm{GL}_d(\mathbb{R})$ the algebraic group generated by μ, by G° its identity component, and let E be the subalgebra of $\mathcal{M}_d(\mathbb{R})$ generated by G°. If the algebraic group G is semisimple, then for every $\lambda \in (0,1)$, there exist $C = C(\mu,\lambda) \geq 0$ such that the following holds.*

Given $x_0 \in \mathbb{T}^d$, assume that for some $t \in (0, \frac{1}{2})$, $a_0 \in \mathbb{Z}^d \setminus \{0\}$, and $n \geq C \log \frac{\|a_0\|}{t}$,

$$|\widehat{(\mu^{*n} * \delta_{x_0})}(a_0)| \geq t.$$

Then, there exists $\gamma \in G$ such that, denoting $W_0 = (a_0 \gamma E)^\perp$ and V_0 the sum of all compact factors of G in $\mathbb{R}^d$,

$$\tilde{d}\big(x_0 - \frac{p}{q} - v, W_0\big) \leq e^{-n\lambda}$$

for some $v \in V_0$, $p \in \mathbb{Z}^d$ and $q \in \mathbb{Z} \setminus \{0\}$ such that $\max(\|v\|, |q|) \leq \left(\frac{\|a_0\|}{t}\right)^C$.

Here, of course, $\frac{p}{q}$, v and W_0 are identified with their projection to the torus $\mathbb{T}^d$.

If the sequence $(\mu^{*n} * \delta_{x_0})$ does not converge to the Haar measure $\mathrm{m}_{\mathbb{T}^d}$ in the weak-$*$ topology, then, by Weyl's equidistribution criterion, there are $a_0 \in \mathbb{Z}^d \setminus \{0\}$ and $t > 0$ such that $|\widehat{(\mu^{*n} * \delta_{x_0})}(a_0)| \geq t$ holds for an unbounded sequence of $n \in \mathbb{N}$. Letting n goes to infinity along this sequence, we deduce the following qualitative statement from the quantitative statement.

Corollary 1.2 (Qualitative statement)**.** *Let μ be a probability measure on $\mathrm{GL}_d(\mathbb{Z})$ having a finite exponential moment. Denote by $G \subset \mathrm{GL}_d(\mathbb{R})$ the algebraic group generated by μ. Assume that G is semisimple. Then for any point $x_0 \in \mathbb{T}^d$, either*

$$\mu^{*n} * \delta_{x_0} \rightharpoonup^* \mathrm{m}_{\mathbb{T}^d},$$

or

$$x_0 \in \mathbb{Q}^d + V_0 + W_0 \mod \mathbb{Z}^d,$$

where V_0 denotes the sum of all compact factors of G in $\mathbb{R}^d$ and W_0 is a rational subspace of $\mathbb{R}^d$ invariant under the action of the identity component G° of G.

As a consequence, we recovers the classification of orbit closures due to Guivarc'h and Starkov [20] and Muchnik [30].

Corollary 1.3 (Classification of orbit closures)**.** *Let $\Gamma \subset \mathrm{GL}_d(\mathbb{Z})$ be a subgroup whose Zariski closure G is a semisimple algebraic group. Let $x_0 \in \mathbb{T}^d$. Then the*

orbit closure $\overline{\Gamma x}$ *is either the whole* $\mathbb{T}^d$ *or contained in a* Γ*-invariant closed subset of the form*

$$\frac{1}{q}\mathbb{Z}^d + \mathrm{B}_{V_0}(0,R) + \bigcup_{\gamma\in G/G^\circ} \gamma W_0 \mod \mathbb{Z}^d$$

where q *is a nonzero integer,* $\mathrm{B}_{V_0}(0,R)$ *is a ball in* V_0*, the sum of all compact factors of* G *in* $\mathbb{R}^d$ *and* W_0 *is invariant under the action of the identity component* G° *of* G*.*

The qualitative statement could also be reformulated more simply as follows.

Corollary 1.4 (Equidistribution)**.** *Let* μ *be a probability measure on* $\mathrm{GL}_d(\mathbb{Z})$ *having a finite exponential moment. Denote by* $\Gamma \subset \mathrm{GL}_d(\mathbb{Z})$ *the subgroup generated by* μ*. Assume that the Zariski closure of* Γ *is semisimple. Then for any* $x_0 \in \mathbb{T}^d$*, either* $\mu^{*n} * \delta_{x_0} \rightharpoonup^* \mathrm{m}_{\mathbb{T}^d}$ *or* x_0 *is contained in a proper* Γ*-invariant closed subset.*

A particularly simple case of the above results is when the group Γ acts strongly irreducibly on $\mathbb{Q}^d$, that is, when Γ preserves no nontrivial finite union of proper subspaces of $\mathbb{Q}^d$. Then, for any $a_0 \in \mathbb{Z}^d \setminus \{0\}$ and any $\gamma \in G$, one must have $a_0\gamma E = (\mathbb{R}^d)^*$, so we obtain a simpler equidistribution statement.

Corollary 1.5 (Equidistribution of $\mathbb{Q}^d$-irreducible random walks)**.** *Assume that* G *is semisimple and acts strongly irreducibly on* $\mathbb{Q}^d$*. Then for every* $\lambda \in (0,1)$*, there exist* $C = C(\mu,\lambda) \geq 0$ *such that the following holds.*

Given $x_0 \in \mathbb{T}^d$*, assume that for some* $t \in (0, \frac{1}{2})$*,* $a_0 \in \mathbb{Z}^d$*, and* $n \geq C \log \frac{\|a_0\|}{t}$*,*

$$|\widehat{(\mu^{*n} * \delta_{x_0})}(a_0)| \geq t.$$

Then there exists $v \in V_0$*,* $p \in \mathbb{Z}^d$ *and* $q \in \mathbb{Z}\setminus\{0\}$ *such that* $\max(\|v\|, |q|) \leq \left(\frac{\|a_0\|}{t}\right)^C$ *and*

$$\tilde{d}\big(x_0 - \frac{p}{q} - v, 0\big) \leq e^{-n\lambda}.$$

In particular, if x_0 *does not lie on a rational translate of the* V_0 *leaf in* $\mathbb{T}^d$*, then* $\mu^{*n} * \delta_{x_0}$ *converges to* $\mathrm{m}_{\mathbb{T}^d}$*.*

It was observed by Benoist and Quint [4, Corollary 1.4] that if G is semisimple without compact factors and acts irreducibly on $\mathbb{Q}^d$, then $\mathrm{m}_{\mathbb{T}^d}$ is the only atom-free μ-stationary probability measure on $\mathbb{T}^d$. By the results of [5], this implies that the Cesàro averages $\frac{1}{n}\sum_{k=0}^{n-1} \mu^{*k} * \delta_{x_0}$ converge to $\mathrm{m}_{\mathbb{T}^d}$. The above corollary immediately shows that convergence also holds without the averaging process. When μ is a symmetric probability measure on $\mathrm{SL}_d(\mathbb{Z})$, a general result of Timothée Bénard [15, Theorem 1] implies this qualitative statement, but without the symmetry assumption the result seems to be new.

Corollary 1.6. *Assume that* G *is semisimple without compact factors and acts strongly irreducibly on* $\mathbb{Q}^d$*. Then, for every* x_0 *irrational in* $\mathbb{T}^d$*, the sequence of measures* $(\mu^{*n} * \delta_{x_0})_{n\geq 0}$ *converges in law to* $\mathrm{m}_{\mathbb{T}^d}$*.*

Our main motivation to carry out the rather technical proof presented here, is its application to the spectral gap property for subgroups of algebraic groups, modulo arbitrary integers. Indeed, as we shall explain in a forthcoming paper, Theorem 1.1 can be used to answer a question of Salehi Golsefidy and Varjú [31, Question 2].

1.1. **Outline of the proof.** The paper is entirely devoted to the proof of Theorem 1.1, for which we use the strategy introduced in [10], and more precisely the variant used in [22] to avoid the proximality assumption. Section 2 deals with discretized algebraic combinatorics in semisimple algebras: we prove some Fourier

decay estimate for multiplicative convolutions of measures satisfying natural non-concentration conditions, Theorem 2.1, generalizing results of Bourgain [9] for the real line. The main input for our proof is a sum-product theorem for representations of real Lie groups [23, Theorem 1.1], which easily implies the discretized sum-product theorem in semisimple algebras; then we use some L^2-flattening lemma similar to the one use by Bourgain and Gamburd in their work on the spectral gap property.

After that, in order to apply the combinatorial results of the previous section to the random walk, we need to check that the measure μ^{*n} appropriately rescaled is not concentrated near proper affine subspaces of E, nor near singular elements; this is done in Section 3. Just as in [22], the argument ultimately relies on the spectral gap property modulo primes obtained by Salehi Golsefidy and Varjú [31]. However, because the rescaling automorphism is no longer a homothety, the proof involves a detailed analysis of the behavior of the random walk with respect to a quasi-norm on the algebra E. To help the reader understand the main ideas of the proof without having to go through all the technical details, we start with the simpler case where E is simple; even in that case, the argument is different and simpler than the one presented in [22], where similar estimates are needed.

In Section 4, we prove Theorem 4.2, an important Fourier decay estimate for the law of the random walk. This simply follows from a combination of the two previous sections when the group G is connected, but becomes more complicated without this assumption. We follow the argument used in [25, Appendix B], with minor modifications.

Section 5 makes the link between the random walk on G and the random walk on $\mathbb{T}^d$. The Fourier decay obtained in the previous section shows that if $\mu^{*n} * \delta_{x_0}$ has one large Fourier coefficient, then reducing slightly the value of n, the measure $\mu^{*n} * \delta_{x_0}$ has many large Fourier coefficients. Using a quantitative version of Wiener's lemma, one infers a first "granulation statement": $\mu^{*n} * \delta_{x_0}$ is concentrated near a finite set of well-separated points in $\mathbb{T}^d$.

To conclude the proof of Theorem 1.1, we run backwards the random walk, starting from the granulation estimate mentioned above. The argument uses in particular the diophantine properties of the random walk, and the exponential unstability of closed invariant subsets, obtained using a drift function, as in Eskin-Margulis [17] or Benoist-Quint [4]. This is the content of Section 6.

1.2. **Concluding remarks.** *Affine random walks.* After some first results of J.-B. Boyer [14], it was explained in [24] how to obtain quantitative equidistribution of affine random walks on the torus, under the assumption that the action on $\mathbb{R}^d$ is strongly irreducible. The arguments in that paper could be adapted to our setting.

More general homogeneous spaces. Benoist and Quint [2, 4, 5] have obtained equidistribution results that are valid in the much more general setting of homogeneous spaces of Lie groups. One drawback is that their convergence theorems are not quantitative, and only concern the Cesàro averages $\frac{1}{n}\sum_{k=0}^{n-1} \mu^{*n} * \delta_{x_0}$. So it would be interesting to see whether the techniques used in the present paper can be used to study random walks on homogeneous spaces like $\mathrm{SL}_d(\mathbb{R})/\mathrm{SL}_d(\mathbb{Z})$, or other homogeneous spaces.

On this subject, the first author has recently obtained, in collaboration with Lakrec and Lindenstrauss, some partial results for affine random walks on nilmanifolds [25]; these spaces may be seen as the simplest generalization of tori, but the analysis already becomes much more intricate.

In a slightly different direction, W. Kim [26] studied effective equidistribution of expanding translates in the space of affine lattices. Also in a different direction,

Lindenstrauss and Mohammadi [29] studied effective density in some homogeneous spaces. Although these equidistribution results do not deal with random walks, some of the techniques used there are similar enough to ours to be mentioned here, especially those inspired by [10].

1.3. **Acknowledgement.** It is a pleasure to thank Yves Benoist and Elon Lindenstrauss for useful discussions, in particular on the existence of satellite measures in the presence of compact factors.

1.4. **Notation.** Here is a list of notation we use.

- $f \ll g$, $g \gg f$, $f = O(g)$, there exists a constant $C > 0$ such that $f \leq Cg$.
- $f \asymp g$ if $f \ll g$ and $g \ll f$.
- $\mathrm{B}(x, r)$, the ball of center x and radius r.
- $\mathrm{B}_V(\,\cdot\,,\,\cdot\,)$, ball in the ambient space V.
- V^*, the space of linear forms on a linear space V.
- $\lambda_1(\mu, V)$, the top Lyapunov exponent associated to the random walk on a Euclidean space V defined by a probability measure μ supported on a group acting linearly on V.
- $\mu * \nu$, multiplicative convolution.
- $\mu^{*k} = \mu * \cdots * \mu$, multiplicative convolution power.
- $\mu \boxplus \nu$, additive convolution.
- $\mu^{\boxplus k} = \mu \boxplus \cdots \boxplus \mu$, additive convolution power.
- $\mu \boxminus \nu$, the image measure of $\mu \otimes \nu$ under the map $(x, y) \mapsto x - y$.
- $\mathbb{1}_A(x) = 1$ if $x \in A$, $\mathbb{1}_A(x) = 0$ otherwise.
- $\#A$, cardinality of a finite set A.
- $|A|$, Lebesgue measure for subsets A of an Euclidean space or a torus.
- $|\,\cdot\,|^{\sim}$, a quasi-norm
- $\tilde{d}(\,\cdot\,,\,\cdot\,)$, a quasi-distance, usually associated to a quasi-norm.
- $\tilde{\mathrm{B}}(\,\cdot\,,\,\cdot\,)$, ball with respect to $\tilde{d}$.
- $\mathbb{P}[\,\cdot\,]$ and $\mathbb{P}[\,\cdot \mid \cdot\,]$, probability and conditional probability.
- $f_*\mu$, image measure of μ under the map f.
- $\mathcal{M}_d(\mathbb{R})$, the space of $d \times d$ real matrices.
- $\mathcal{P}(X)$, the space of Borel probability measure on a topological space.
- $\langle\,\cdot\,,\,\cdot\,\rangle$, according to the context, the natural pairing $V^* \times V \to \mathbb{R}$ or the natural pairing $\mathbb{Z}^d \times \mathbb{T}^d \to \mathbb{T}$.

2. Sum-product, L^2-flattening and Fourier decay

In this section, we study multiplicative convolutions of measures on a semisimple associative algebra E. Our goal is to derive Theorem 2.1 below, which shows that under some natural non-concentration assumptions, such multiplicative convolutions admit a polynomial Fourier decay. This generalizes results of Bourgain [9] for $E = \mathbb{R}$, of Li [28] for $E = \mathbb{R} \oplus \cdots \oplus \mathbb{R}$, and of [22] for a simple algebra E.

Let E be a normed real algebra of finite dimension. The determinant $\det_E(a)$ of an element $a \in E$ is simply defined as the determinant of the multiplication map $E \to E$, $x \mapsto ax$. Given $\rho > 0$, we let

$$S_E(\rho) = \{\, x \in E \,|\, |\det_E(x)| \leq \rho \,\}.$$

If $W \subset E$ is any subset, we let $W^{(\rho)}$ denote the ρ-neighborhood of W, defined by

$$W^{(\rho)} = \{\, x \in E \,|\, \exists w \in W : \, \|x - w\| < \rho \,\}.$$

The following definition summarizes the non-concentration conditions we shall need in order to prove some Fourier decay for multiplicative convolutions.

Definition (Non-concentration conditions)**.** Let $\varepsilon > 0$, $\kappa > 0$, $\tau > 0$ be parameters. We say a measure η on E satisfies $\mathrm{NC}_0(\varepsilon, \kappa, \tau)$ at scale $\delta > 0$ if

(i) $\operatorname{supp} \eta \subset \mathrm{B}(0, \delta^{-\varepsilon})$;
(ii) for every $x \in E$, $\eta(x + S_E(\delta^\varepsilon)) \leq \delta^\tau$;
(iii) for every $\rho \in [\delta, 1]$ and every proper affine subspace $W \subset E$, $\eta(W^{(\rho)}) \leq \delta^{-\varepsilon}\rho^\kappa$.

We say that a measure η on E satisfies $\mathrm{NC}(\varepsilon, \kappa, \tau)$ at scale $\delta > 0$ if it can be written as a sum of measures

$$\eta = \eta_0 + \eta_1 \quad \text{with} \quad \begin{cases} \eta_0 \text{ satisfying } \mathrm{NC}_0(\varepsilon, \kappa, \tau) \\ \eta_1(E) \leq \delta^\tau. \end{cases}$$

Given a finite measure μ on E, its Fourier transform $\hat{\mu}$ is the function on the dual space E^* given by the expression

$$\forall \xi \in E^*, \quad \hat{\mu}(\xi) = \int_E e^{2i\pi\langle \xi, x\rangle} \mathrm{d}\mu(x).$$

If ν is another finite measure on E, the multiplicative convolution $\mu * \nu$ is defined as the image measure of $\mu \otimes \nu$ on $E \times E$ under the map $(x, y) \mapsto xy$. It should not be confused with the additive convolution $\mu \boxplus \nu$, image of $\mu \otimes \nu$ under the map $(x, y) \mapsto x + y$.

Theorem 2.1 (Fourier decay of multiplicative convolutions)**.** *Let E be a normed finite-dimensional semisimple algebra over $\mathbb{R}$. Given $\kappa > 0$, there exists $s = s(E, \kappa) \in \mathbb{N}$ and $\varepsilon = \varepsilon(E, \kappa) > 0$ such that for any parameter $\tau \in (0, \varepsilon\kappa)$ the following holds for any scale $\delta > 0$ sufficiently small.*

If $\eta_1, \dots, \eta_s$ are probability measures on E satisfying $\mathrm{NC}(\varepsilon, \kappa, \tau)$ at scale δ, then for all $\xi \in E^*$ with $\delta^{-1+\varepsilon} \leq \|\xi\| \leq \delta^{-1-\varepsilon}$,

$$|(\eta_1 * \cdots * \eta_s)^\wedge(\xi)| \leq \delta^{\varepsilon\tau}.$$

For $E = \mathbb{R}$, this is due to Bourgain [9, Lemma 8.43]. For algebras of the form $E = \mathbb{R} \oplus \cdots \oplus \mathbb{R}$, this is due to Li [28, Theorem 1.1]. We shall first prove this theorem when all η_i are equal, i.e. $\eta_1 = \cdots = \eta_s = \eta$ and then deduce the general statement from this particular case following the argument in [24, Proof of Theorem B.3]. Alternatively, one could adapt the first part of the proof to handle directly the general case, but this would make notation cumbersome.

The proof we give for Theorem 2.1 follows a strategy originating in the work of Bourgain, Glibichuk and Konyagin [13] on exponential sums in finite fields: one deduces the bound on the exponential sum from a combinatorial "sum-product" statement, using an L^2-flattening statement. In our case, the combinatorial input is a discretized sum-product theorem in semisimple algebras, which follows from a general sum-product statement for representations of real Lie groups obtained in [23, Theorem 2.3].

2.1. **Sum-product in semisimple algebras.** Sum-product estimates go back to the work of Erdős and Szemerédi [16] who showed that there exists some positive constant ε such that for any subset A of integers,

$$|A + A| + |AA| \geq |A|^{1+\varepsilon}$$

where $A + A$ and AA denote respectively the sum-set and the product-set of A, defined by $A + A = \{a + b \ ; \ a, b \in A\}$ and $AA = \{ab \ ; \ a, b \in A\}$. In the following, we consider a normed semisimple algebra E of finite dimension over $\mathbb{R}$, and our goal is to prove a similar statement for subsets $A \subset E$, with the cardinality replaced by the covering number $\mathcal{N}(A, \delta)$ of A at small scale $\delta > 0$. Recall that by definition, $\mathcal{N}(A, \delta)$ is the minimal cardinality of a cover of A by balls of radius δ in E. In

order to ensure that the covering number of A at scale δ grows under addition or multiplication, one of course has to assume that A is not essentially equal to a ball in some subalgebra of E. We make a stronger assumption and require that A is not concentrated near any proper affine subspace of E.

Definition (Affine non-concentration)**.** Let V be a Euclidean space, and $\varepsilon, \kappa > 0$ two parameters. We say a subset $A \subset V$ satisfies $\mathrm{ANC}(\varepsilon, \kappa)$ at scale δ if

(i) $A \subset \mathrm{B}(0, \delta^{-\varepsilon})$ and
(ii) for every $\rho \geq \delta$ and every proper affine subspace $W \subset V$, $\mathcal{N}(A \cap W^{(\rho)}, \delta) \leq \delta^{-\varepsilon} \rho^{\kappa} \mathcal{N}(A, \delta)$.

Essentially, we want to show that if E is a semisimple algebra, then for every $\kappa > 0$, there exists $\varepsilon > 0$ such that for any set $A \subset \mathrm{B}_E(0,1)$ satisfying $\mathrm{ANC}(\varepsilon, \kappa)$ and $\delta^{-\kappa} \leq \mathcal{N}(A, \delta) \leq \delta^{-\dim E + \kappa}$, one has $\mathcal{N}(A+A, \delta) + \mathcal{N}(AAA, \delta) \geq \delta^{-\varepsilon} \mathcal{N}(A, \delta)$. We shall prove a slightly more technical growth statement, involving the tensor algebra $E \otimes E^{\mathrm{op}}$, where E^{op} denotes the algebra with the same linear structure as E but with multiplication $(a, b) \mapsto ba$. Note that the algebra $E \otimes E^{\mathrm{op}}$ acts naturally on E by

$$\forall a, x \in E,\ \forall b \in E^{\mathrm{op}}, \quad (a \otimes b)x = axb.$$

Theorem 2.2 (Sum-product in semisimple algebras)**.** *Let E be a finite-dimensional real semisimple algebra. Given $\kappa > 0$, there exists $\varepsilon = \varepsilon(E, \kappa)$ such that the following holds for all $\delta > 0$ sufficiently small.*

(i) Let A be a subset of E satisfying $\mathrm{ANC}(\varepsilon, \kappa)$ *and*
(ii) $\delta^{-\kappa} \leq \mathcal{N}(A, \delta) \leq \delta^{-\dim E + \kappa}$.
(iii) Let $B \subset E \otimes E^{\mathrm{op}}$ be a subset satisfying $\mathrm{ANC}(\varepsilon, \kappa)$.

Then there exists $f \in B$ such that

$$\mathcal{N}(A + A, \delta) + \mathcal{N}(A + fA, \delta) \geq \delta^{-\varepsilon} \mathcal{N}(A, \delta).$$

The theorem above is almost equivalent to the fact that one can obtain from A a small ball in E using a bounded number of sums and products. This is the content of the proposition below, which we obtain as a simple application of [23, Theorem 2.3]. For a subset A in an algebra E and $s \in \mathbb{N}^*$, we let $\langle A \rangle_s$ denote the set of elements in E that can be obtained as sums of at most s products of at most s elements of A or $-A$.

Proposition 2.3 (Bounded generation in semisimple algebras)**.** *Let E be a finite-dimensional real semisimple algebra. Given $\kappa > 0$ and ε_0, there exists $\varepsilon = \varepsilon(E, \kappa, \varepsilon_0) > 0$ and $s = s(E, \kappa, \varepsilon_0) \geq 1$ such that the following holds for all $\delta > 0$ sufficiently small. If $A \subset \mathrm{B}(0, \delta^{-\varepsilon})$ satisfies* $\mathrm{ANC}(\varepsilon, \kappa)$ *at scale δ in E, then*

$$\mathrm{B}(0, \delta^{\varepsilon_0}) \subset \langle A \rangle_s + \mathrm{B}(0, \delta).$$

Proof. We consider the action of the multiplicative group $G = E^{\times}$ on $V = E$. By semisimplicity, we may decompose E into a sum of non-trivial irreducible representations $E = \oplus_i V_i$. Let $\pi_i \colon G \to \mathrm{GL}(V_i)$ denote the representation of G on V_i. By [23, Theorem 2.3], there is a neighbourhood U of the identity in G and constants $\varepsilon = \varepsilon(E, \kappa, \varepsilon_0) > 0$ and $s = s(E, \kappa, \varepsilon) \geq 1$ such that the following holds for all $\delta > 0$ sufficiently small. Let A_0 be a subset of U and A_1 a subset of $\mathrm{B}_V(0,1)$. Assume

(i) for all $i = 1, \ldots, k$, for all $\rho \geq \delta$, $\mathcal{N}(\pi_i(A_0), \rho) \geq \delta^{\varepsilon} \rho^{-\kappa}$,
(ii) for any linear subspace $W \subset V$ which is not G-invariant, there is $a \in A_0$ such that $d(a, \mathrm{Stab}_G(W)^{\circ}) \geq \delta^{\varepsilon}$,
(iii) for any proper G-invariant linear subspace $W \subset V$, there is $a \in A_1$ such that $d(a, W) \geq \delta^{\varepsilon}$.

Then

$$\mathrm{B}_V(0,\delta^{\varepsilon_0}) \subset \langle A_0, A_1\rangle_s + \mathrm{B}(0,\delta).$$

Here, $\langle A_0, A_1\rangle_s$ denotes the set of elements in V that can be obtained as sums of at most s products of at most s elements of A_0 and elements of $A_1 \cup (-A_1)$.

Our set A is not necessarily contained in the neighborhood U, but we may cover A by translates of U in E, and then, by the pigeonhole principle, there is $a \in A$ such that $A_0 = (A-a)\cap U$ satisfies

$$\mathcal{N}(A_0,\delta) \gg_U \delta^{O(\varepsilon)}\mathcal{N}(A,\delta).$$

This set A_0 satisfies $\mathrm{ANC}(O(\varepsilon),\kappa)$. This non-concentration condition applied to the suspaces $\oplus_{j\neq i}V_j$ shows that the first condition above is verified. Moreover, if $W \subset E$ is not G-invariant, then the algebra generated by $\mathrm{Stab}_G(W)$ is a proper subalgebra of E; so the second condition is also satisfied. To conclude, take $A_1 = A$, which certainly satisfies the third condition. □

In short, Theorem 2.2 will follow from Proposition 2.3 applied to the set B in the tensor algebra $E \otimes E^{\mathrm{op}}$, and from the Plünnecke-Ruzsa inequality.

Proof of Theorem 2.2. For $K \geq 1$, define

$$R_\delta(A,K) = \big\{\, f \in E\otimes E^{\mathrm{op}} \,\big|\, \mathcal{N}(A+fA,\delta) \leq K\mathcal{N}(A,\delta)\,\big\}.$$

As explained in [21, Lemma 30], it follows from the Plünnecke-Ruzsa inequality that if $\mathcal{N}(A+A,\delta) \leq K\mathcal{N}(A,\delta)$, then the set $R_\delta(A,K)$ is almost stable under addition and multiplication. More precisely, for $s \in \mathbb{N}$,

$$\langle R_\delta(A,K)\rangle_s + \mathrm{B}_{E\otimes E^{\mathrm{op}}}(0,\delta) \subset R_\delta(A,K^{O_s(1)}).$$

Now assume for a contradiction that $B\cup\{1\} \subset R_\delta(A,\delta^{-\varepsilon})$. Since E is a semisimple algebra, $E\otimes E^{\mathrm{op}}$ is also one. Thus, by Proposition 2.3 applied to the set $B \subset E\otimes E^{\mathrm{op}}$, for any $\varepsilon_0 > 0$, there is $s = s(E,\kappa,\varepsilon_0) \geq 1$ such that

$$\mathrm{B}_{E\otimes E^{\mathrm{op}}}(0,\delta^{\varepsilon_0}) \subset \langle B\rangle_s + \mathrm{B}(0,\delta)$$

and therefore,

$$\mathrm{B}_{E\otimes E^{\mathrm{op}}}(0,\delta^{\varepsilon_0}) \subset \langle B\rangle_s + \mathrm{B}(0,\delta) \subset R_\delta(A,\delta^{-O_s(\varepsilon)}).$$

In particular, $\delta^{\varepsilon_0} \in R_\delta(A,\delta^{-O_s(\varepsilon)})$. This certainly implies $\delta^{-\varepsilon_0} \in R_\delta(A,\delta^{-O_s(\varepsilon_0+\varepsilon)})$ and then

$$\mathrm{B}_{E\otimes E^{\mathrm{op}}}(0,1) \subset R_\delta(A,\delta^{-O_s(\varepsilon_0+\varepsilon)}).$$

If ε_0 and ε are chosen small enough, this contradicts Lemma 2.4 below. □

We are left to show the next lemma, stating that if A has $\mathrm{ANC}(\varepsilon,\kappa)$ at scale δ, then $\mathrm{B}_{E\otimes E^{\mathrm{op}}}(0,1)$ is not contained in $R_\delta(A,\delta^{-\varepsilon})$.

Lemma 2.4. *Let E be a finite-dimensional real semisimple algebra with $E = E_1 \oplus \cdots \oplus E_r$ being its decomposition into minimal two-sided ideals. We write $\pi_j \colon E \to E_j$ for the corresponding projections.*

Given $\kappa > 0$, there exists $\varepsilon = \varepsilon(E,\kappa) > 0$ such that the following holds. Let $A \subset \mathrm{B}(0,\delta^{-\varepsilon})$ be a subset of E. Assume

(i) $\mathcal{N}(A,\delta) \leq \delta^{-\dim E+\kappa}$

(ii) for each $j = 1,\dots,r$, $\max_{x\in E_j} \mathcal{N}(A\cap\pi_j^{-1}(\mathrm{B}_{E_j}(x,\rho)),\delta) \leq \rho^\kappa\mathcal{N}(A,\delta)$

then there exists $f \in \mathrm{B}_{E\otimes E^{\mathrm{op}}}(0,1)$ such that

$$\mathcal{N}(A+fA,\delta) > \delta^{-\varepsilon}\mathcal{N}(A,\delta).$$

Proof. The image of $E\otimes E^{\mathrm{op}}$ in $\mathrm{End}(E)$ is equal to the image of $\bigoplus_{j=1}^{r} E_j\otimes E_j^{\mathrm{op}}$. Let f_j, $j=1,\dots,r$ be a family of jointly independent random elements of $\mathrm{B}_{E_j\otimes E_j^{\mathrm{op}}}(0,1)$ distributed according to the Lebesgue measure on $E_j\otimes E_j^{\mathrm{op}}$, and set

$$f = f_1 + \cdots + f_r$$

regarded as a random element of $\mathrm{End}(E)$. In the following argument, probabilities and expectations are taken with respect to these random variables. For each j, since the algebra E_j is simple, the action of $E_j\otimes E_j^{\mathrm{op}}$ on E_j is irreducible. Hence $E_j\otimes E_j^{\mathrm{op}}(y) = E_j$ for any non-zero $y\in E_j$ and consequently,

$$\forall \delta>0,\ \forall x,y\in E_j,\quad \mathbb{P}\big[\,\|f_j(y)-x\|\le\delta\,\big] \ll \delta^{\dim E_j}\|y\|^{-\dim E_j}. \tag{2.1}$$

Consider the map

$$\begin{array}{rccc}\varphi\colon & A\times A & \to & E\\ & (x,y) & \mapsto & x+fy\end{array}$$

The energy of the map φ at scale $\delta>0$ is defined as

$$\mathcal{E}_\delta(\varphi, A\times A) = \mathcal{N}\big(\{\,(a,a',b,b')\in A\times A\times A\times A \mid \|\varphi(a)-\varphi(a')\|\le\delta\,\},\delta\big).$$

By the Cauchy-Schwarz inequality — see also [21, Lemma 12(i)],

$$\mathcal{N}(\varphi(A\times A),\delta) = \mathcal{N}(A+fA,\delta) \ge \frac{\mathcal{N}(A,\delta)^4}{\mathcal{E}_\delta(\varphi,A\times A)}.$$

Taking expectations and applying Jensen's inequality, we find

$$\mathbb{E}\big[\mathcal{N}(A+fA,\delta)\big] \ge \frac{\mathcal{N}(A,\delta)^4}{\mathbb{E}\big[\mathcal{E}_\delta(\varphi,A\times A)\big]} \tag{2.2}$$

so it suffices to bound $\mathbb{E}\big[\mathcal{E}_\delta(\varphi,A\times A)\big]$ from above.

For that, let $\tilde{A}$ be a maximal δ-separated subset of A. By [21, Lemma 12(ii)],

$$\mathbb{E}\big[\mathcal{E}_\delta(\varphi,A\times A)\big] \le \sum_{x,y,x',y'\in\tilde{A}} \mathbb{P}\big[\,f(y'-y)\in \mathrm{B}(x-x',5\delta)\,\big].$$

Let $\rho = \delta^{\frac{\kappa}{\dim E+\kappa}}$. We split the sum into two parts according to whether

$$\forall j=1,\dots,r,\quad \|\pi_j(y'-y)\|\ge\rho.$$

If this is the case, then (2.1) implies

$$\mathbb{P}\big[\,f(y'-y)\in\mathrm{B}(x-x',5\delta)\,\big] \ll \delta^{\dim E}\rho^{-\dim E}.$$

Otherwise, there is $j\in\{1,\dots,r\}$ such that $\pi_j(y')\in\mathrm{B}(\pi_j(y),\rho)$. For fixed y the number of such y' in $\tilde{A}$ is

$$\#\big(\tilde{A}\cap\pi_j^{-1}(\mathrm{B}(\pi_j(y),\rho))\big) \ll \mathcal{N}(A\cap\pi_j^{-1}(\mathrm{B}(\pi_j(y),\rho)),\delta) \le \rho^\kappa\mathcal{N}(A,\delta).$$

Moreover for fixed y,y' and x, we have

$$\sum_{x'\in\tilde{A}} \mathbb{P}\big[\,f(y'-y)\in\mathrm{B}(x-x',5\delta)\,\big] \ll 1$$

because the balls $\mathrm{B}(x',5\delta)$ have overlap multiplicity at most $O(1)$. Putting these considerations together, we obtain

$$\begin{aligned}\mathbb{E}\big[\mathcal{E}_\delta(\varphi,A\times A)\big] &\ll \delta^{\dim E}\rho^{-\dim E}\mathcal{N}(A,\delta)^4 + \rho^\kappa\mathcal{N}(A,\delta)^3\\ &\le \big(\delta^\kappa\rho^{-\dim E}+\rho^\kappa\big)\mathcal{N}(A,\delta)^3\\ &= \delta^{\frac{\kappa^2}{\dim E+\kappa}}\mathcal{N}(A,\delta)^3\end{aligned}$$

Combined with (2.2), this finishes the proof of the lemma. □

2.2. **L^2-flattening.** Our goal is now to translate the sum-product theorem obtained above in terms of measures on the semisimple algebra E. The result we obtain is an L^2-flattening lemma for additive and multiplicative convolutions of measures on E. Statements of this form already appear implicitly in the work of Bourgain [8, 9] on the Erdős-Volkmann ring conjecture, and were later much popularized by their application to the spectral gap problem by Bourgain and Gamburd [12, 11]. They are usually derived from the analogous combinatorial growth statement, via a decomposition of the measures into dyadic level sets, combined with an application of the Balog-Szemerédi-Gowers lemma.

Before we can state our result, we give a non-concentration condition for measures on E, analogous to the one given for subsets in the previous paragraph.

Definition (Affine non-concentration for measures)**.** Let V be a Euclidean space, and $\varepsilon, \kappa > 0$ two parameters. We say that a measure η on V satisfies $\mathrm{ANC}(\varepsilon, \kappa)$ at scale δ if

(i) $\operatorname{supp}\eta \subset \mathrm{B}(0, \delta^{-\varepsilon})$;
(ii) for every $\rho \geq \delta$ and every proper affine subspace $W \subset V$, $\eta(W^{(\rho)}) \leq \delta^{-\varepsilon}\rho^{\kappa}$.

In this paper, measures are often studied at some fixed small positive scale δ. For that reason, it is convenient to define the *regularized measure* η_δ of a measure η on E at scale δ by

$$\eta_\delta = \eta \boxplus P_\delta$$

where $P_\delta = \frac{\mathbb{1}_{\mathrm{B}(0,\delta)}}{|\mathrm{B}(0,\delta)|}$ is the normalized indicator function of the ball of radius δ centered at 0. The measure η_δ will be identified with its density with respect to the Lebesgue measure on E, and we write

$$\|\eta\|_{2,\delta} = \|\eta_\delta\|_2.$$

Proposition 2.5 (L^2-flattening)**.** *Let E be finite-dimensional semisimple algebra over $\mathbb{R}$. Given $\kappa > 0$, there exists $\varepsilon = \varepsilon(E, \kappa)$ such that the following holds for all $\delta > 0$ sufficiently small. Let η be a probability measure on E satisfying*

(i) η is supported on $E \setminus S_E(\delta^\varepsilon)$;
(ii) η satisfies $\mathrm{ANC}(\varepsilon, \kappa)$ at scale δ on E;
(iii) $\delta^{-\kappa+\varepsilon} \leq \|\eta\|_{2,\delta}^2 \leq \delta^{-\dim E+\kappa-\varepsilon}$.

Then,

$$\|\eta * \eta * \eta \boxminus \eta * \eta * \eta\|_{2,\delta} \leq \delta^{\varepsilon}\|\eta\|_{2,\delta}.$$

We wish to deduce this proposition from Theorem 2.2. A first useful observation is that the non-concentration condition for measures is closely related to non-concentration for subsets.

Lemma 2.6. *Given an Euclidean space V, and parameters $\epsilon > 0$ and $\kappa > 0$, the following holds for all $\delta > 0$ sufficiently small.*

(i) If $A \subset V$ has $\mathrm{ANC}(\varepsilon, \kappa)$ at scale δ, then there is a measure supported on A which has $\mathrm{ANC}(2\varepsilon, \kappa)$ at scale δ.
(ii) Let η be a probability measure on V satisfying $\mathrm{ANC}(\varepsilon, \kappa)$ at scale δ. If $A \subset V$ is a subset such that $\eta(A) \geq \delta^\varepsilon$ then there is a subset $A' \subset A$ which satisfies $\mathrm{ANC}(6\varepsilon, \kappa)$ at scale δ.

Proof. For the first item, let $\tilde{A}$ be a maximal δ-separated subset of A. The normalised counting measure on $\tilde{A}$ satisfies the desired property. The second item is slightly more subtle. Since the normalised restriction of η to A satisfies $\mathrm{ANC}(2\varepsilon, \kappa)$, we may assume without loss of generality that $A = \operatorname{supp}\mu$. Let d denote the dimension of V. Set $A_{0,0} = A$ and for integers $0 < i \ll \log\frac{1}{\delta} + 1$,

$$A_{i,0} = \Big\{ a \in A \,\Big|\, 2^{i-1} < \frac{\eta(\mathrm{B}(a, 2\delta))}{|\mathrm{B}(0,\delta)|} \leq 2^i \Big\}.$$

Then for all $i \geq 0$, set $A_i = A_{i,0}^{(\delta)}$. It follows from this construction that

$$\eta_\delta \ll \sum_i 2^i \, \mathbb{1}_{A_i} \ll \Big(\log \frac{1}{\delta}\Big)\eta_{3\delta} + \mathbb{1}_{A_0} \, .$$

Necessarily, there is some $i \geq 0$ such that $2^i|A_i| \geq \delta^\varepsilon$. It is then easy to check that $A' = A_{i,0}$ satisfies $\mathrm{ANC}(3\varepsilon, \kappa)$ at scale δ. □

The next lemma is similar in spirit to the previous one. Roughly speaking, given measures η on V and μ on $\mathrm{GL}(V)$ such that the convolution $\mu * \eta \boxminus \mu * \eta$ has large L^2-norm at scale δ, we construct related subsets $A \subset V$ and $B \subset \mathrm{GL}(V)$ such that $A - BA$ is not much larger than A. This is the central part of the proof of Proposition 2.5; it relies on the Balog-Szemerédi-Gowers lemma.

Lemma 2.7. *Let V be a Euclidean space and μ a probability measure on $\mathrm{GL}(V)$ such that*

$$\forall g \in \operatorname{supp}\mu, \quad \|g\| + \|g^{-1}\| \leq \delta^{-\varepsilon}.$$

Let η be a probability measure on $\mathrm{B}_V(0, \delta^{-\varepsilon})$ such that

$$\|\mu * \eta \boxminus \mu * \eta\|_{2,\delta} > \delta^\varepsilon \|\eta\|_{2,\delta}.$$

Then there exist a subset $A \subset \mathrm{B}_V(0, \delta^{-O(\varepsilon)})$ and an element $g_1 \in \operatorname{supp}\mu$ such that

$$\delta^{-\dim V + O(\varepsilon)} \|\eta\|_{2,\delta}^{-2} \leq \mathcal{N}(A, \delta) \leq \delta^{-\dim V - O(\varepsilon)} \|\eta\|_{2,\delta}^{-2}$$

and

$$\mu\big(\{\, g \in \mathrm{GL}(V) \,|\, \mathcal{N}(A - g g_1^{-1} A, \delta) \leq \delta^{-O(\varepsilon)} \mathcal{N}(A, \delta) \,\}\big) \geq \delta^{O(\varepsilon)}.$$

If moreover η satisfies $\mathrm{ANC}(\varepsilon, \kappa)$ in V at scale δ for some $\kappa > 0$ then A satisfies $\mathrm{ANC}(O(\varepsilon), \kappa)$.

Proof. We use the following rough comparison notation : for positive quantities f and g, we write $f \lesssim g$ for $f \leq \delta^{-O(\varepsilon)} g$ and $f \sim g$ for $f \lesssim g$ and $g \lesssim f$.

We have

$$\|\mu * \eta_\delta \boxminus \mu * \eta_\delta\|_2 \gtrsim \|\mu * \eta \boxminus \mu * \eta\|_{2,\delta} \gtrsim \|\eta_\delta\|_2.$$

As in the proof of Lemma 2.6, we can approximate μ_δ using dyadic level sets : there are δ-discretized sets[1] $(A_i)_{i \geq 0}$ in $\mathrm{B}_V(0, \delta^{-\varepsilon})$ such that A_i is empty for $i \gg \log \frac{1}{\delta}$ and

$$\eta_\delta \ll \sum_{i \geq 0} 2^i \, \mathbb{1}_{A_i} \lesssim \eta_{3\delta} + \mathbb{1}_{A_0} \, . \tag{2.3}$$

By the pigeonhole principle, there are $i, j \geq 0$ such that

$$\begin{aligned}
\|\eta_\delta\|_2 &\lesssim \|\mu * \eta_\delta \boxminus \mu * \eta_\delta\|_2 \\
&\lesssim 2^{i+j} \|\mu * \mathbb{1}_{A_i} \boxminus \mu * \mathbb{1}_{A_j}\|_2 \\
&\lesssim 2^{i+j} \int_{\mathrm{GL}(V) \times \mathrm{GL}(V)} \|\mathbb{1}_{gA_i} \boxminus \mathbb{1}_{g'A_j}\|_2 \mathrm{d}(\mu \otimes \mu)(g, g').
\end{aligned}$$

In the last inequality, we used $g * \mathbb{1}_{A_i} = |\det g|^{-1} \, \mathbb{1}_{gA_i}$ and $|\det g| \sim 1$ for all $g \in \operatorname{supp}\mu$. By the right inequality in (2.3), we have

$$2^i |A_i| \lesssim 1 \quad \text{and} \quad 2^j |A_j|^{\frac{1}{2}} \lesssim \|\eta_\delta\|_2$$

and similarly

$$2^j |A_j| \lesssim 1 \quad \text{and} \quad 2^i |A_i|^{\frac{1}{2}} \lesssim \|\eta_\delta\|_2.$$

By Young's inequality and the estimate on $\det g$, we have for all $g, g' \in \operatorname{supp}\mu$,

$$2^{i+j} \|\mathbb{1}_{gA_i} \boxminus \mathbb{1}_{g'A_j}\|_2 \lesssim \|\eta_\delta\|_2.$$

[1] A δ-discretized set is a union of balls of radius δ.

Thus, by the pigeonhole principle again, there exists $g_0 \in \operatorname{supp}\mu$ and a set $B_0 \subset \operatorname{supp}\mu$ such that $\mu(B_0) \gtrsim 1$ and for all $g \in B_0$,

$$\|\eta_\delta\|_2 \gtrsim 2^{i+j}\|\mathbb{1}_{g_0A_i} \boxminus \mathbb{1}_{gA_j}\|_2 \gtrsim \|\eta_\delta\|_2.$$

By the above estimates, this implies

$$\begin{aligned}\|\mathbb{1}_{g_0A_i} \boxminus \mathbb{1}_{gA_j}\|_2^2 &\gtrsim 2^{-2i-2j}\|\eta_\delta\|_2^2\\ &\gtrsim 2^{-i-j}|A_i|^{\frac12}|A_j|^{\frac12}\\ &\gtrsim |A_i|^{\frac32}|A_j|^{\frac32}\\ &\sim |g_0A_i|^{\frac32}|gA_j|^{\frac32}.\end{aligned}$$

By the Balog-Szemerédi-Gowers lemma [32, Theorem 6.10], for each $g \in B_0$ there are δ-discretized subsets $A_g \subset A_i$ and $A'_g \subset A_j$ such that

$$|A_g| \sim |A_i|,\ |A'_g| \sim |A_j|, \text{ and } \quad \mathcal{N}(g_0A_g - gA'_g, \delta) \lesssim \mathcal{N}(g_0A_g, \delta)^{\frac12}\mathcal{N}(gA'_g, \delta)^{\frac12}.$$

By a Cauchy-Schwarz and pigeonhole argument there are $g_1 \in B_0$ and $B_1 \subset B_0$ such that $\mu(B_1) \gtrsim 1$ and

$$\forall g \in B_1, \quad |A_{g_1} \cap A_g| \sim |A_{g_1}| \sim |A_g| \text{ and } |A'_{g_1} \cap A'_g| \sim |A'_{g_1}| \sim |A'_g|.$$

For subsets $A, A' \in V$, write $A \approx A'$ if $\mathcal{N}(A - A', \delta) \lesssim \mathcal{N}(A, \delta)^{\frac12}\mathcal{N}(A', \delta)^{\frac12}$. By Ruzsa's triangle inequality (see [32, Lemma 3.2 and Section 6], if $A \approx A'$ and $A' \approx A''$, then $A \approx A''$, and moreover, if $A \approx A'$ for some sets A and A', then $A \approx A$ and $A' \approx A'$.

The above shows that for every $g \in B_0$, $g_0A_g \approx gA'_g$. This implies $g_0A_g \approx g_0A_g$, and since g is $\delta^{-\varepsilon}$-Lipschitz, $A_g \approx A_g$. Therefore, for $g_1 \in B_0$ and $g \in B_1$ as above, we find $A_{g_1} \approx A_{g_1} \cap A_g \approx A_g$. Similarly, $A'_{g_1} \approx A'_g$,. Finally

$$g_0A_{g_1} \approx g_0A_g \approx gA'_g \approx gA'_{g_1} \approx gg_1^{-1}g_0A_{g_1},$$

showing that $A = g_0A_{g_1}$ has all the desired properties.

For the "moreover" part, note that $\eta(A_i) \gtrsim 1$. By the proof of Lemma 2.6(ii), A_i satisfies $\mathrm{ANC}(O(\varepsilon), \kappa)$ and hence so does A_{g_1} and A. □

To prove Proposition 2.5 we use the above lemma for the action of $E \otimes E^{\mathrm{op}}$ on E, and the apply the sum-product theorem in E.

Proof of Proposition 2.5. Let μ be the image measure of $\eta \otimes \eta$ in $\mathrm{GL}(E)$, so that $\mu * \eta \boxminus \mu * \eta = \eta * \eta * \eta \boxminus \eta * \eta * \eta$. Assume for a contradiction that

$$\|\mu * \eta \boxminus \mu * \eta\|_{2,\delta} > \delta^{\varepsilon}\|\eta\|_{2,\delta}.$$

Then by Lemma 2.7 there is a subset $A \subset E$ satisfying the $\mathrm{ANC}(O(\varepsilon), \kappa)$ at scale δ and an element $g_1 \in \operatorname{supp}\mu$ such that

$$\delta^{-\kappa+O(\varepsilon)} \le \mathcal{N}(A, \delta) \le \delta^{-\dim E+\kappa-O(\varepsilon)}.$$

and

$$\mu\big(\{\, g \in \mathrm{GL}(V) \,\big|\, \mathcal{N}(A - gg_1^{-1}A, \delta) \le \delta^{-O(\varepsilon)}\mathcal{N}(A, \delta)\,\}\big) \ge \delta^{O(\varepsilon)}.$$

By definition of μ, we may write $g_1 = a_1 \otimes b_1$, and the above inequality becomes

$$(\eta * a_1^{-1}) \otimes (b_1^{-1} * \eta)\big(\{\, (a,b) \in E \times E \,\big|\, \mathcal{N}(A - aAb, \delta) \le \delta^{-O(\varepsilon)}\mathcal{N}(A, \delta)\,\}\big) \ge \delta^{O(\varepsilon)}.$$

Since $a_1, b_1 \notin S_E(\delta^{\varepsilon})$, the measures $\eta * a_1^{-1}$ and $b_1^{-1} * \eta$ satisfy $\mathrm{ANC}(O(\varepsilon), \kappa)$ at scale δ. Moreover, Lemma 2.8 below shows that $(\eta * a_1^{-1})\dot{\otimes}(b_1^{-1} * \eta)$ satisfies $\mathrm{ANC}(O(\varepsilon), \frac{\kappa}{2})$. In view of Lemma 2.6(ii), this contradicts Theorem 2.2, provided ε is chosen small enough. □

Lemma 2.8. *Let V_1 and V_2 be finite-dimensional linear spaces. For each $i = 1, 2$, let η_i be a measure on V_i and denote by $\eta_1 \dot{\otimes} \eta_2$ the image measure of $\eta_1 \otimes \eta_2$ by the natural bilinear map $V_1 \times V_2 \to V_1 \otimes V_2$.*

Given two parameters $\varepsilon, \kappa > 0$, the following holds for $\delta > 0$ sufficiently small. If η_1 and η_2 both satisfy $\mathrm{ANC}(\varepsilon, \kappa)$ at scale δ, then $\eta_1 \dot{\otimes} \eta_2$ satisfies $\mathrm{ANC}(2\varepsilon, \frac{\kappa}{2})$ in $V_1 \otimes V_2$ at scale δ.

Proof. Let v_1 and v_2 be independent random variables taking values respectively in V_1 and V_2 and distributed according to η_1 and η_2. To establish $\mathrm{ANC}(2\varepsilon, \frac{\kappa}{2})$ for $\eta_1 \dot{\otimes} \eta_2$, it is enough to show that for any linear form $\varphi \in (V_1 \otimes V_2)^*$ with $\|\varphi\| = 1$, any $t \in \mathbb{R}$ and any $\rho \geq \delta$, we have

$$\mathbb{P}\big[\, |\varphi(v_1 \otimes v_2) - t| < \rho \,\big] \ll \delta^{-\varepsilon} \rho^{\kappa/2}. \tag{2.4}$$

Note that $(V_1 \otimes V_2)^* = V_1^* \otimes V_2^*$. Hence, letting $(\psi_1, \dots, \psi_d)$ be a orthonormal basis of V_1^*, we can write $\varphi \in (V_1 \otimes V_2)^*$ as

$$\varphi = \psi_1 \otimes \varphi_1 + \cdots + \psi_d \otimes \varphi_d$$

where $\varphi_1, \dots, \varphi_d \in V_2^*$ are uniquely determined. Moreover,

$$1 = \|\varphi\|^2 = \|\varphi_1\|^2 + \cdots + \|\varphi_d\|^2. \tag{2.5}$$

On the one hand, when v_2 is fixed, the map

$$v_1 \mapsto \varphi(v_1 \otimes v_2) = \sum_{i=1}^{d} \psi_i(v_1) \varphi_i(v_2)$$

is the linear form $\sum_{i=1}^d \varphi_i(v_2)\psi_i \in V_1^*$, which has norm $\sum_{i=1}^d |\varphi_i(v_2)|^2$. Thus, by independence of v_1 and v_2 and property $\mathrm{ANC}(\varepsilon, \kappa)$ for η_1, we can estimate the conditional probability

$$\mathbb{P}\Big[\, |\varphi(v_1 \otimes v_2) - t| < \rho \,\Big|\, \sum_{i=1}^{d} |\varphi_i(v_2)|^2 \geq \rho^{1/2} \,\Big] \leq \delta^{-\varepsilon} \rho^{\kappa/2}. \tag{2.6}$$

On the other hand, by property $\mathrm{ANC}(\varepsilon, \kappa)$ for η_2, for each $i = 1, \dots, d$,

$$\mathbb{P}\big[\, |\varphi_i(v_2)| \leq \rho^{1/2} \|\varphi_i\| \,\big] \leq \delta^{-\varepsilon} \rho^{\kappa/2}.$$

Hence, on account of (2.5),

$$\mathbb{P}\Big[\sum_{i=1}^{d} |\varphi_i(v_2)|^2 < \rho^{1/2} \,\Big] \leq d\delta^{-\varepsilon} \rho^{\kappa/2}. \tag{2.7}$$

Inequalities (2.6) and (2.7) together imply (2.4) and finish the proof of the lemma. □

2.3. **Fourier decay.** To prove Theorem 2.1 we apply the L^2-flattening Proposition 2.5 repeatedly. The measures we obtain are images of tensor powers η^k under polynomial maps $E^k \to E$, and we need to compare their Fourier decay to that of simple multiplicative convolutions of η. This is the content of the next lemma. This technique will also be useful to weaken slightly the assumptions of Theorem 2.1, see Corollary 2.11 below.

Lemma 2.9. *Let E be any real associative algebra, and let η be a measure on E with $\eta(E) \leq 1$. Let $\mu = \eta * \eta * \eta \boxminus \eta * \eta * \eta$ then for any integer $m \geq 1$,*

$$\forall \xi \in E^*, \quad |\widehat{\eta^{*3m}}(\xi)|^{2^m} \leq |\widehat{\mu^{*m}}(\xi)|.$$

Proof. By [25, Lemma B.6], if η, η', η'' be probability measures on E, then the Fourier transform of $\eta * (\eta' \boxminus \eta') * \eta''$ takes non-negative real values and moreover,

$$\forall \xi \in E^*, \quad |(\eta * \eta' * \eta'')^\wedge(\xi)|^2 \leq \big(\eta * (\eta' \boxminus \eta') * \eta''\big)^\wedge(\xi).$$

By a simple renormalising argument we see that the same holds η, η', η'' are measures with total mass $\eta(E), \eta'(E), \eta''(E) \leq 1$. Using this inequality m times with measure $\eta' = \eta^{*3}$, so that $\mu = \eta' \boxminus \eta'$, we get

$$|\widehat{\mu^{*m}}(\xi)| \geq |\widehat{\mu^{*(m-1)}} * \eta^{*3}(\xi)|^2 \geq \cdots \geq |\widehat{\eta^{*3m}}(\xi)|^{2^m}.$$

□

We shall also need a lemma on Fourier decay for multiplicative convolutions of measures with small L^2-norm. In the case where $E = \mathbb{R}$, such bounds originate in the work of Falconer [18] on projection theorems, and appear explicitly in Bourgain [9, Theorem 7]. The result below is taken from [22, Lemma 2.9].

Lemma 2.10. *Let E be a finite-dimensional real associative algebra with unit. The following holds for any parameters $\kappa > 0$ and $\varepsilon > 0$ and any scale $\delta > 0$ small enough. Let η and ν be probability measures on E. Assume*

(i) $\|\eta\|_{2,\delta}^2 \leq \delta^{-\kappa}$,
(ii) $\operatorname{supp}\eta \subset \mathrm{B}(0, \delta^{-\varepsilon})$ *and* $\operatorname{supp}\nu \subset \mathrm{B}(0, \delta^{-\varepsilon})$,
(iii) for every proper affine subspace $W \subset E$, $\nu(W^{(\delta)}) \leq \delta^{2\kappa}$.

Then for $\xi \in E^*$ with $\delta^{-1+\varepsilon} \leq \|\xi\| \leq \delta^{-1-\varepsilon}$,

$$|\widehat{\eta * \nu}(\xi)| \leq \delta^{\frac{\kappa}{\dim E + 3} - O(\varepsilon)}.$$

We can finally derive Theorem 2.1.

Proof of Theorem 2.1. First case: $\eta_1 = \cdots = \eta_s = \eta$.
For a measure η satisfying $\mathrm{NC}(\varepsilon, \kappa, \tau)$ at scale δ, we write $\mathrm{ess}(\eta)$ to denote the essential part of η, defined as a measure on E satisfying

(i) $\mathrm{ess}(\eta) \leq \eta$ and $\mathrm{ess}(\eta)(E) \geq \eta(E) - 3\delta^\tau$,
(ii) $\mathrm{ess}(\eta)$ is supported on $\mathrm{B}(0, \delta^{-\varepsilon}) \setminus S_E(\delta^\varepsilon)$,
(iii) $\mathrm{ess}(\eta)$ satisfies $\mathrm{ANC}(\varepsilon, \kappa)$ at scale δ.

If η and η' satisfy $\mathrm{NC}(\varepsilon, \kappa, \tau)$ at scale δ, then $\eta \boxplus \eta'$, $\eta \boxminus \eta'$ and $\eta * \eta'$ all satisfy $\mathrm{NC}(O(\varepsilon), \kappa, \frac{\tau}{2})$ at scale δ. We may therefore define inductively $\eta_0 = \mathrm{ess}(\eta)$, and for $k \geq 0$,

$$\eta_{k+1} = \mathrm{ess}\big(\eta_k^{*3} \boxminus \eta_k^{*3}\big),$$

to get, for each $k \geq 0$,

(i) $\eta_k(E) \geq 1 - \delta^{\frac{\tau}{O_k(1)}}$,
(ii) η_k is supported on $\mathrm{B}(0, \delta^{-O_k(\varepsilon)}) \setminus S_E(\delta^{O_k(\varepsilon)})$,
(iii) η_k satisfies $\mathrm{ANC}(O_k(\varepsilon), \kappa)$ at scale δ.

Note that $\mathrm{ANC}(O_k(\varepsilon), \kappa)$ implies

$$\|\eta_k\|_{2,\delta}^2 \leq \delta^{-\dim E + \kappa - O_k(\varepsilon)}.$$

By Proposition 2.5, there exists $\varepsilon_1 = \varepsilon_1(E, \kappa)$ such that, provided $\varepsilon > 0$ is small enough, we have for each $0 \leq k \leq \left\lceil \frac{\dim E}{\varepsilon_1} \right\rceil$, either $\|\eta_k\|_{2,\delta}^2 \leq \delta^{-\kappa}$ or

$$\|\eta_{k+1}\|_{2,\delta}^2 \leq \delta^{\varepsilon_1} \|\eta_k\|_{2,\delta}^2.$$

Hence there exists $s \leq \left\lceil \frac{\dim E}{\varepsilon_1} \right\rceil$ such that

$$\|\eta_s\|_{2,\delta}^2 \leq \delta^{-\kappa}.$$

By Lemma 2.10, for $\xi \in E^*$ with $\delta^{-1+\varepsilon} \leq \|\xi\| \leq \delta^{-1-\varepsilon}$,

$$\left|\widehat{\eta_s^{*2}}(\xi)\right| \leq \delta^{\frac{\kappa}{O(1)} - O(\varepsilon)} \leq \delta^{\tau}.$$

Now, a first application of Lemma 2.9 to $\mu = \eta_{s-1}^{*3} \boxminus \eta_{s-1}^{*3}$ with $m = 2$ yields

$$\widehat{\eta_s^{*2}}(\xi) + \delta^{\frac{\tau}{O_s(1)}} = \widehat{\mu^{*2}}(\xi) \geq |\widehat{\eta_{s-1}^{*2\cdot 3}}(\xi)|^{2^2}.$$

A second application of the same lemma to $\mu = \eta_{s-2}^{*3} \boxminus \eta_{s-2}^{*3}$ with $m = 2 \cdot 3$ gives

$$\widehat{\eta_{s-1}^{*2\cdot 3}}(\xi) + \delta^{\frac{\tau}{O_s(1)}} = \widehat{\mu_1^{*2\cdot 3}}(\xi) \geq |\widehat{\eta_{s-2}^{*2\cdot 3^2}}(\xi)|^{2^{2\cdot 3}}$$

and repeating this process s times, we finally obtain

$$\left|\widehat{\eta_0^{*2\cdot 3^s}}(\xi)\right|^{O_s(1)} \leq \left|\widehat{\eta_s^{*2}}(\xi)\right| + \delta^{\frac{\tau}{O_s(1)}} \leq \delta^{\frac{\kappa}{O(1)} - O(\varepsilon)} \leq \delta^{\frac{\tau}{O_s(1)}}.$$

This allows to conclude the proof of the theorem in this first case:

$$\left|\widehat{\eta^{*2\cdot 3^s}}(\xi)\right| \leq \left|\widehat{\eta_0^{*2\cdot 3^s}}(\xi)\right| + O_s(\delta^{\tau}) \leq \delta^{\varepsilon\tau}.$$

General case

To deduce the general case from the previous one, we follow [25, Proof of Theorem B.3]. In short, one applies the previous case to the measures

$$\eta_\lambda = \lambda_1(\eta_1 \boxminus \eta_1) + \cdots + \lambda_s(\eta_s \boxminus \eta_s),$$

where $\lambda = (\lambda_1, \ldots, \lambda_s)$ in $\mathbb{R}_+^s$ is such that $\lambda_1 + \cdots + \lambda_s \leq 1$. The Fourier decay for $\eta_1 * \cdots * \eta_s$ can be deduced from that of $\eta_\lambda * \cdots * \eta_\lambda$ for every λ using the fact that Fourier coefficients of $\eta_\lambda * \cdots * \eta_\lambda$ can be written as polynomials in λ whose coefficients are essentially Fourier coefficients of $\eta_1 * \cdots * \eta_s$. The reader is referred to [25] for details. □

We conclude this section by showing that the conclusion of Theorem 2.1 still holds if the non-concentration assumption is only satisfied for some additive convolution of the measures η_i, $i = 1, \ldots, s$. This will be useful when we study Fourier decay of random walks on linear groups.

Corollary 2.11. *Let E be a normed finite-dimensional semisimple algebra over $\mathbb{R}$. Given $D \in \mathbb{N}^*$ and $\kappa > 0$, there exists $s = s(E, \kappa) \in \mathbb{N}$ and $\varepsilon = \varepsilon(E, \kappa, D) > 0$ such that for any parameter $\tau \in (0, \varepsilon\kappa)$ the following holds for any scale $\delta > 0$ sufficiently small.*

If η_i, $i = 1, \ldots, s$ are probability measures on E such that each $\eta_i^{\boxplus D}$ satisfies $\mathrm{NC}(\varepsilon, \kappa, \tau)$ at scale δ, then for all $\xi \in E^*$ with $\delta^{-1+\varepsilon} \leq \|\xi\| \leq \delta^{-1-\varepsilon}$,

$$|(\eta_1 * \cdots * \eta_s)^\wedge(\xi)| \leq \delta^{\varepsilon\tau}.$$

Proof. Let $\xi \in E^*$ with $\delta^{-1+\varepsilon} \leq \|\xi\| \leq \delta^{-1-\varepsilon}$. Since all the measures $\eta_i^{\boxplus D} \boxminus \eta_i^{\boxplus D}$, $i = 1, \ldots, s$ satisfy $\mathrm{NC}(\varepsilon, \kappa, \tau)$ at scale δ, Theorem 2.1 shows that

$$\left|\left((\eta_1^{\boxplus D} \boxminus \eta_1^{\boxplus D}) * \cdots * (\eta_s^{\boxplus D} \boxminus \eta_s^{\boxplus D})\right)^\wedge(\xi)\right| \leq \delta^{\varepsilon\tau}.$$

Applying [25, Lemma B.6] repeatedly s times, we see that

$$\begin{aligned}
&\left|\left((\eta_1^{\boxplus D} \boxminus \eta_1^{\boxplus D}) * \cdots * (\eta_s^{\boxplus D} \boxminus \eta_s^{\boxplus D})\right)^\wedge(\xi)\right| \\
&\qquad \geq \left|\left((\eta_1^{\boxplus D} \boxminus \eta_1^{\boxplus D}) * \cdots * (\eta_{s-1}^{\boxplus D} \boxminus \eta_{s-1}^{\boxplus D}) * \eta_s\right)^\wedge(\xi)\right|^{2D} \\
&\qquad \geq \cdots \\
&\qquad \geq |(\eta_1 * \cdots * \eta_s)^\wedge(\xi)|^{(2D)^s}
\end{aligned}$$

so that

$$|(\eta_1 * \cdots * \eta_s)^\wedge(\xi)| \leq \delta^{\frac{\varepsilon\tau}{(2D)^s}}.$$

□

3. Non-concentration for random walks on semisimple groups

In this section we consider a probability measure μ on $\mathrm{SL}_d(\mathbb{Z})$, and we prove some non-concentration property for the law of the associated random walk, viewed as a measure on the algebra generated by μ. Let Γ be the group generated by the support of μ and G be the Zariski closure of Γ in $\mathrm{SL}_d(\mathbb{R})$. We assume that G is semisimple and Zariski connected.

Let E denote the $\mathbb{R}$-linear span of G in $\mathcal{M}_d(\mathbb{R})$, which is also the subalgebra generated by G in $\mathcal{M}_d(\mathbb{R})$. Since G is semisimple, one may decompose E into a direct sum of irreducible G-modules. This gives a decomposition of E into minimal left ideals, so that by the fundamental theorem of semisimple rings [33, §117], E is a semisimple algebra. Let

$$E = E_1 \oplus \cdots \oplus E_r \tag{3.1}$$

be the decomposition of E into simple factors, i.e. into minimal two-sided ideals. For $j = 1, \dots, r$, let $\pi_j \colon E \to E_j$ denote the corresponding projections. Consider the top Lyapunov exponent associated to μ on each of the factors E_j, defined by

$$\lambda_1(\mu, E_j) = \lim_{n \to +\infty} \frac{1}{n} \int \log \|\pi_j(g)\| \mathrm{d}\mu^{*n}(g).$$

In order to study the law at time n of the random walk, we shall use the rescaling automorphism $\varphi_n \colon E \to E$ defined by

$$\varphi_n(g) = \sum_{j=1}^{r} e^{-n\lambda_1(\mu, E_j)} \pi_j(g). \tag{3.2}$$

Recall that by Furstenberg's theorem [19] on the positivity of the Lyapunov exponent, one has $\lambda_1(\mu, E_j) \geq 0$ with equality if and only if $\pi_j(G)$ is compact. After reordering the factors, we may assume that $\lambda_1(\mu, E_j) > 0$ if and only if $j \leq s$ for some integer $s \leq r$. Let $E' = E_1 \oplus \cdots \oplus E_s$ and $\pi' \colon E \to E'$ the corresponding projection. Finally, for $n \geq 1$ we define

$$\mu_n = (\pi' \circ \varphi_n)_*(\mu^{*n}).$$

The goal of this section is as follows.

Proposition 3.1 (Non-concentration)**.** *Let μ be a probability measure on $\mathrm{SL}_d(\mathbb{Z})$ having a finite exponential moment. Let G denote the algebraic group generated by μ. Assume that G is semisimple and Zariski connected, and denote by $E \subset \mathcal{M}_d(\mathbb{R})$ the algebra generated by G. Writing $D = \dim E$, there exists $\kappa = \kappa(\mu) > 0$ such that for any $\varepsilon > 0$ there exists $\tau > 0$ such that $\mu_n^{\boxplus D}$ satisfies $\mathrm{NC}(\varepsilon, \kappa, \tau)$ at scale e^{-n} in E' for all n sufficiently large.*

The reader can easily convince themselves that μ_n does not satisfy $\mathrm{NC}(\varepsilon, \kappa, \tau)$, especially the nonconcentration condition near singular matrices. Hence taking an additive convolution power is necessary.

3.1. Non-concentration near affine subspaces.

In this subsection we show that if μ is a probability measure on $\mathrm{SL}_d(\mathbb{Z})$ generating a connected semisimple algebraic group G, the law at time n of the random walk associated to μ is not concentrated near proper affine subspaces of the algebra generated by μ.

We introduce a quasi-norm adapted to the random walk on the algebra E generated by μ. Given an element g in E, we write $g = \sum_{i=1}^{r} g_i$ according to the direct sum decomposition (4.1) and set

$$|g|^{\sim} = \max_{1 \leq i \leq s} \|g_i\|^{\frac{1}{\lambda_1(\mu, E_i)}}.$$

Note that $|g| = 0$ if and only if g lies in the sum

$$E_0 := E_{s+1} \oplus \cdots \oplus E_r$$

of all compact factors. We denote by $\tilde{d}$ the quasi-distance on E given by $\tilde{d}(x,y) = |x-y|^{\sim}$. For instance, if W is any affine subspace of E, we write

$$\tilde{d}(g,W) = \inf_{w \in W} |g-w|^{\sim}.$$

Our goal is the following proposition.

Proposition 3.2 (Affine non-concentration on E)**.** *Let μ and G be as in There exists $\kappa = \kappa(\mu) > 0$ such that for every $n \geq 0$ and $\rho \geq e^{-n}$, for every affine hyperplane $W \subset E$ such that $W - W \supset E_0$,*

$$\mu^{*n}\big(\{\, g \in G \,|\, \tilde{d}(g,W) < \rho \min_{j \in J_W} |\pi_j(g)|^{\sim} \,\}\big) \ll \rho^{\kappa}$$

where $J_W = \{\, 1 \leq j \leq r \,|\, V_j \not\subset W - W \,\}$.

Remark. In general, it is not possible to replace the minimum $\min_{j \in J_W} |\pi_j(g)|^{\sim}$ by $|g|^{\sim}$. This can be seen for example by taking $G = G_1 \times G_1$ and $\mu = \mu_1 \otimes \mu_1$; in other words, the random walk is the direct product of two independent copies of a random walk on G_1. By the central limit theorem for random matrix products, the probability to obtain at time n an element $g = (g_1, g_2)$ such that $\|g_1\| \leq e^{-\sqrt{n}} \|g_2\|$ has a positive limit c. Therefore, for large n,

$$\mu^{*n}\big(\{\, g = (g_1,g_2) \in G \,|\, \|g_1\| < e^{-\sqrt{n}} \|g\| \,\}\big) \geq \frac{c}{2}.$$

and taking $W = \{0\} \times \mathrm{Span}_{\mathbb{R}}(G_1)$, we find

$$\mu^{*n}\big(\{\, g \in G \,|\, \tilde{d}(g,W) < e^{-\sqrt{n}} |g|^{\sim} \,\}\big) \geq \frac{c}{2}.$$

3.1.1. *The case of a simple algebra.* For clarity, we first explain the proof of Proposition 3.2 when the algebra E generated by G is simple. In that case, the quasi-norm is a norm on E and $\min_{j \in J_W} |\pi_j(g)|^{\sim} = \|g\|$. The key result in the proof is following proposition, which we shall later apply to the irreducible action of $G \times G$ on E.

Proposition 3.3. *Let μ be a probability measure on $\mathrm{SL}_d(\mathbb{Z})$ with a finite exponential moment. Assume that the algebraic group G generated by μ is Zariski connected and acts irreducibly on $V = \mathbb{R}^d$. There exists $\kappa = \kappa(\mu)$ such that for every $v \in V$, and any affine hyperplane $W \subset V$,*

$$\mu^{*n}\big(\{\, g \in G \,|\, d(gv, W) \leq \rho \|gv\| \,\}\big) \ll \rho^{\kappa}.$$

Proof. First step: escape from affine subvarieties.
There exists $c > 0$ such that for every affine map f on E that is not identically zero on G,

$$\mu^{*n}\big(\{\, g \in G \,|\, f(g) = 0 \,\}\big) \ll e^{-cn}.$$

Indeed, by [2, Lemme 8.5], the group G is semisimple. So the desired inequality follows from the spectral gap property modulo prime integers [31], combined with the Lang-Weil estimates on the number of points on algebraic varieties in finite fields. We refer the reader to [22, Proposition 3.7] for a detailed argument.
Second step: a small neighborhood via a Diophantine property.
There exist $C, c > 0$ such that for every non-zero polynomial map f of degree at most 1 on G,

$$\mu^{*n}\big(\{\, g \in G \,|\, |f(g)| \leq e^{-Cn} \|f\| \,\}\big) \ll e^{-cn}.$$

By the large deviation principle (see Theorem 3.11), there exists $c > 0$ such that for all n large enough, $\mu^{*n}(\{g \in G \mid \|g\| > e^{2\lambda_1(\mu,V)}\}) \leq e^{-cn}$. Therefore, to prove

the desired inequality, it suffices to show that for $C \geq 0$ large enough, the subset $A_n \subset \mathcal{M}_d(\mathbb{Z})$ defined as

$$A_n = \big\{ g \in \Gamma \,\big|\, |f(g)| \leq e^{-Cn}\|f\| \text{ and } \|g\| \leq e^{2\lambda_1(\mu,V)} \big\}$$

is included in $G \cap \ker \psi$ for some affine map $\psi : E \to \mathbb{R}$ not identically zero on G.

Suppose for a contradiction that this is not the case. Letting $k = \dim \mathbb{R}_{\leq 1}[G]$ be the dimension of the space of polynomial maps on G of degree at most 1, we may choose $g_1, \dots, g_k$ in A_n such that the linear map

$$\begin{array}{rccc} L\colon & \mathbb{R}_{\leq 1}[G] & \to & \mathbb{R}^k \\ & \psi & \mapsto & (\psi(g_1), \dots, \psi(g_k)) \end{array}$$

is bijective. Since it has integer coefficients and norm at most $e^{C_0 n}$, we get $\|L^{-1}\| \leq e^{C_1 n}$. In particular, $\|f\| \leq e^{C_1 n}\|Lf\| = e^{C_1 n} \max_{1\leq i\leq k} |f(g_i)| \leq e^{C_1 n} e^{-Cn}\|f\|$, which is the desired contradiction.

Third step: distance to proper subspaces.

There exist $C, c > 0$ such that for every $v \in V$ and every affine hyperplane $W \subset V$,

$$\mu^{*n}(\{g \in G \mid d(gv, W) \leq e^{-Cn}\|v\|\}) \ll e^{-cn}.$$

Then, let $\varphi_W : V \to \mathbb{R}$ be an affine map such that $\ker \varphi_W = W$, and consider the affine map on G given by

$$f_{v,W}(g) = \frac{\varphi_W(gv)}{\|\varphi_W\|}.$$

Note that $|f_{v,W}(g)| \asymp d(gv, W)$. Let $B = \mathrm{B}_G(1, 1)$ denote the unit ball centered at the identity in G. Note that $\|v\| \asymp \sup_{g\in B} |f_{v,W}(g)|$ within constants independent of v and W. Indeed, otherwise, we may find v_n and W_n such that $\sup_{g\in B} d(gv_n, W_n) \to 0$. Extracting subsequences if necessary, we may assume that $v_n \to v$ and $W_n \to W$; then for every $g \in B$, $d(gv, W) = \lim d(gv_n, W_n) = 0$. This implies that $G \cdot v \subset W$ and contradicts the assumption that G acts irreducibly on V. The desired inequality therefore follows from the previous step.

Fourth step: scaling.

First observe that increasing C slightly, we can assume that for every $v \in V$ and every affine hyperplane $W \subset V$,

$$\mu^{*n}\big(\{\, g \in G \,|\, d(gv, W) \leq e^{-Cn}\|gv\| \,\}\big) \ll e^{-cn}.$$

Indeed, by the large deviation estimate,

$$\mu^{*n}\big(\{\, g \in G \,|\, \|gv\| \leq e^{2\lambda_1 n}\|v\| \,\}\big) \geq 1 - e^{-cn}$$

where $\lambda_1 = \lambda_1(\mu, V)$ is the top Lypunov exponent of μ.

Let $\kappa = \frac{c}{C}$, where $C, c > 0$ are the constants obtained above. Choose $m \in \mathbb{N}^*$ such that $\rho = e^{-Cm}$ and write

$$\begin{aligned} &\mu^{*n}\big(\{\, g \in G \,|\, d(gv, W) \leq \rho\|gv\| \,\}\big) \\ &\qquad = \int \mu^{*m}\big(\{\, g \in G \,|\, d(gg_1v, W) \leq e^{-Cm}\|gg_1v\| \,\}\big) \mathrm{d}\mu^{*(n-m)}(g_1) \\ &\qquad \leq e^{-cm} = \rho^{\kappa}. \end{aligned}$$

□

Proof of Proposition 3.2, case where E is simple. For $x \in E$, let $L_x : E \to E$ and $R_x : E \to E$ denote the left and right multiplication by x, respectively. Given a probability measure μ on G, we define a probability measure $\bar{\mu}$ on $\mathrm{GL}(E)$ by

$$\bar{\mu} = \frac{1}{2} L_*\mu + \frac{1}{2} R_*\mu.$$

The group generated by $\bar{\mu}$ is isomorphic to $G \times G$ and acts irreducibly on E. Moreover, in an appropriate basis, the elements of $\operatorname{supp} \bar{\mu}$ have integer coefficients, so we may apply Proposition 3.3 to $\bar{\mu}$, with vector $v = 1_E$, the unit of E. Note that if $\bar{g}$ is a random element distributed according to $\bar{\mu}^{*n}$, then $\bar{g} \cdot 1_E$ has law μ^{*n}, and therefore we find, uniformly over all affine hyperplanens $W \subset E$,

$$\mu^{*n}\big(\{\, g \in G \,|\, d(g, W) \leq \rho \|g\| \,\}\big) \ll \rho^{\kappa},$$

which is exactly the content of Proposition 3.2 in the case where E is simple. □

3.1.2. *General case.* The proof of Proposition 3.2 in the general case follows the same strategy as in the simple case, but the argument becomes slightly more technical, because the norm on E is replaced by a quasi-norm, and E contains proper ideals.

To state the appropriate generalization of Proposition 3.3, we consider a probability measure μ on $\mathrm{SL}_d(\mathbb{Z})$ with some finite exponential moment, and let G be the algebraic group generated by μ. We assume that G is Zariski connected and that the space $V = \mathbb{R}^d$ can be decomposed into a sum of irreducible representations of G:

$$V = V_1 \oplus \cdots \oplus V_r.$$

We denote by $\pi_j \colon V \to V_j$, $j = 1, \dots, r$ the corresponding projections. To define a quasi-norm on V, we fix $\alpha = (\alpha_1, \dots, \alpha_s)$ an s-tuple of positive real numbers, where s is some fixed integer $1 \leq s \leq r$, and set

$$|v|^{\sim} = |v|^{\sim}_{\alpha} = \max_{1 \leq j \leq s} \|\pi_j(v)\|^{\alpha_j}.$$

For example, $|v|^{\sim} = 0$ if and only if $v \in V_0 := V_{s+1} \oplus \cdots \oplus V_r$. The quasi-distance associated to $|\cdot|^{\sim}$ on V is given by

$$\tilde{d}(v, w) = \tilde{d}_{\alpha}(v, w) = |v - w|^{\sim}_{\alpha}.$$

It satisfies a weak form of the triangle inequality:

$$\forall u, v, w \in V, \quad \tilde{d}(u, w) \ll_{\alpha} \tilde{d}(u, v) + \tilde{d}(v, w).$$

Given a subset $W \subset V$, and $v \in V$, we define the distance from v to W by

$$\tilde{d}(v, W) = \inf_{w \in W} \tilde{d}(v, w).$$

Remark. In all our applications, we shall take s so that $V_0 = V_{s+1} \oplus \cdots \oplus V_r$ is the sum of all compact factors and

$$\alpha_j = \frac{1}{\lambda_1(\mu, V_j)} \quad \text{for } j = 1, \dots, s$$

to obtain a quasi-norm adapted to the random walk associated to μ, same as the one defined in the introduction. However, the proof works in the more general setting of any choice of s and α.

In the remainder of this subsection, s and α are fixed and the implied constants in all Landau and Vinogradov notations may depend on d and α.

Proposition 3.4. *Assume that G is Zariski connected and that the linear span of G in $\mathrm{End}(V)$ contains π_j for $j = 1, \dots, s$. Then there exists $\kappa = \kappa(\mu, \alpha) > 0$ such that for any $v \in V$ and any affine hyperplane $W \subset V$ with $V_0 \subset W - W$,*

$$\forall n \geq 0,\ \forall \rho \geq e^{-n}, \quad \mu^{*n}\big(\big\{\, g \in G \,\big|\, \tilde{d}(gv, W) < \rho \min_{j \in J_W} |\pi_j(gv)|^{\sim} \,\big\}\big) \ll \rho^{\kappa}$$

where $J_W = \{\, 1 \leq j \leq r \,|\, V_j \not\subset W - W \,\}$.

Remark. The requirement that the linear span of G contain π_j, $j = 1, \dots, s$ is here to exclude examples such as $V = V_1 \oplus V_1$, with G acting irreducibly on V_1. Indeed, in that case the diagonal subspace $W = \{(v_1, v_1)\ ;\ v_1 \in V_1\}$ is stable under G, so the proposition cannot hold.

In the proof, we will use Lemma 3.5 and Lemma 3.6, whose proofs will be given right after.

Proof. First and second step: spectral gap and Diophantine property.
Arguing exactly as in the proof of Proposition 3.3 we obtain that there exist $C, c > 0$ such that for every polynomial map f of degree at most 1 on G,

$$\mu^{*n}\big(\{\, g \in G \,|\, |f(g)| \le e^{-Cn}\|f\| \,\}\big) \ll e^{-cn}.$$

Third step: distance to proper subspaces.
There exist $C_1, c > 0$ such that for every $v \in V$ and every affine hyperplane W such that $W - W \supset V_0$,

$$\mu^{*n}\big(\{\, g \in G \,\big|\, \tilde{d}(gv, W) \le e^{-C_1 n} \min_{j \in J_W} |\pi_j(v)|^{\sim} \,\}\big) \ll e^{-cn}.$$

To prove this, Lemma 3.5 below shows that it is enough to show that if B is some large ball in G, then

$$\mu^{*n}\big(\{\, g \in G \,\big|\, \tilde{d}(gv, W) \le e^{-C_1 n} \sup_{h \in B} \tilde{d}(hv, W) \,\}\big) \ll e^{-cn}.$$

Now let $\varphi_W \colon V \to \mathbb{R}$ be an affine map such that $\ker \varphi_W = W$, and denote by ℓ_W the linear part of φ_W; by Lemma 3.6 below, the distance to W for the quasi-norm is given by

$$\forall v \in V, \quad \tilde{d}(v, W) \asymp \min_{i : \ell_W(u_i) \neq 0} \left| \frac{\varphi_W(v)}{\ell_W(u_i)} \right|^{\alpha_{j(i)}},$$

where $(u_i)_{1 \le i \le d}$ is an orthonormal basis compatible with the quasi-norm and $u_i \in V_{j(i)}$ for $i = 1, \dots, d$. Therefore, if g satisfies $\tilde{d}(gv, W) \le e^{-C_1 n} \sup_{h \in B} \tilde{d}(hv, W)$, there must exist i such that

$$\left| \frac{\varphi_W(gv)}{\ell_W(u_i)} \right|^{\alpha_{j(i)}} \ll_\alpha e^{-C_1 n} \sup_{h \in B} \left| \frac{\varphi_W(hv)}{\ell_W(u_i)} \right|^{\alpha_{j(i)}}$$

whence

$$|\varphi_W(gv)| \ll_\alpha e^{-\frac{C_1 n}{\alpha_{j(i)}}} \sup_{h \in B} |\varphi_W(hv)|.$$

If $\frac{C_1}{\alpha_{j(i)}} > C$, the previous step applied to the affine map $f \colon g \mapsto \varphi_W(gv)$ shows that the μ^{*n}-measure of such points is bounded above by e^{-cn}, so the desired statement is proved.
Fourth step: scaling
Note that there are $C_2 = C_2(\mu, \alpha) > 1$ and $c = c(\mu) > 0$ such that for any vector $v \in V$ and any affine hyperplane $W \subset V$ with $V_0 \subset W - W$,

$$\forall n \ge 0, \quad \mu^{*n}\big(\{\, g \in G \,\big|\, \tilde{d}(gv, W) \le e^{-C_2 n} \min_{j \in J_W} |\pi_j(gv)|^{\sim} \,\}\big) \ll e^{-cn}. \tag{3.3}$$

This readily follows from the previous step, and from the fact that, by the exponential moment assumption, there are $C_3 = C_3(\mu) > 1$ and $c = c(\mu) > 0$ such that

$$\mu^{*n}\big(\{\, g \in G \,|\, \|g\| \ge e^{C_3 n} \,\}\big) \ll e^{-cn}.$$

Noting that for any $j = 1, \dots, s$, $|\pi_j(gv)|^{\sim} \le \|g\|^{\alpha_j} |\pi_j(v)|^{\sim}$, we obtain (3.3) by taking $C_2 = C_1 + (\max_{1 \le j \le s} \alpha_j) C_3$.

Finally, given $e^{-n} \leq \rho \leq 1$, set $m = \left\lfloor \frac{-\log \rho}{C_2} \right\rfloor$. Writing $\mu^{*n} = \mu^{*m} * \mu^{*(n-m)}$ and using the fact that (3.3) holds uniformly in v, we find

$$\begin{aligned}
&\mu^{*n}\big(\{\, g \in G \,|\, \tilde{d}(gv, W) < \rho \min_{j \in J_W} |\pi_j(gv)|^{\sim} \,\}\big) \\
&\leq \int_G \mu^{*m}\big(\{\, g \in G \,|\, \tilde{d}(ghv, W) \leq e^{-C_2 m} \min_{j \in J_W} |\pi_j(ghv)|^{\sim} \,\}\big) \mathrm{d}\mu^{*(n-m)}(h) \\
&\ll e^{-cm} \\
&\ll \rho^{c/C_2}.
\end{aligned}$$

This finishes the proof of Proposition 3.4. □

We are left to show the two technical lemmas on quasi-norms and distances to hyperplanes that we used in the proof.

Lemma 3.5. *Assume that the linear span of G in $\mathrm{End}(V)$ contains π_j, for $j = 1, \dots, s$. Then there exists a ball $B \subset G$ such that for any affine hyperplane $W \subset V$ with $V_0 \subset W - W$, and any $v \in V$,*

$$\min_{j \in J_W} |\pi_j(v)|^{\sim} \ll \sup_{g \in B} \tilde{d}(gv, W).$$

Proof. In the particular case $r = 1$ (irreducible case), one may assume that the quasi-norm is equal to the euclidean norm. So the desired inequality with $B = \mathrm{B}_G(1, 1)$ has already been proved in the third step of the proof of Proposition 3.3.

For the general case, first observe that by working in the quotient space V/V_0, we may assume that $V_0 = \{0\}$, that is, $V = V_1 \oplus \cdots \oplus V_s$. Then assume for a contradiction that for arbitrarily large R and arbitrarily small $c > 0$, there exists $v \in V$ and $W \subset V$ such that

$$\forall g \in \mathrm{B}_G(1, R), \quad \tilde{d}(gv, W) \leq c \min_{j \in J_W} |\pi_j(v)|^{\sim}. \tag{3.4}$$

Applying several times the irreducible case in each V_j, $j = 1, \dots, s$ to the vector $\pi_j(v)$, we obtain vectors

$$v_{j,1}, \dots, v_{j,\dim V_j} \in \mathrm{B}_G(1, 1)\pi_j(v) - \pi_j(v) \tag{3.5}$$

forming a basis of V_j such that

$$\big\|v_{j,1} \wedge \cdots \wedge v_{j,\dim V_j}\big\| \gg \|\pi_j(v)\|^{\dim V_j}.$$

Concatenate these bases to get a basis $(u_1, \dots, u_d)$ of V, which has the property that

$$\|u_1 \wedge \cdots \wedge u_d\| \gg \prod_{j=1}^{s} \|\pi_j(v)\|^{\dim V_j}. \tag{3.6}$$

Let $W_0 = W - W$ denote the direction of W. By assumption, for $j = 1, \dots, s$, there exist constants $\beta_{j,k} \in \mathbb{R}$ and elements g_k in G such that

$$\pi_j = \sum_k \beta_{j,k} g_k.$$

Set R large enough so that for all k, $g_k \, \mathrm{B}_G(1, 1) \subset \mathrm{B}_G(1, R)$. Taking (3.4) and (3.5) into account, this implies that for each $i = 1, \dots, d$,

$$\tilde{d}(u_i, W_0) \leq 2c \sum_k |\beta_{j,k}| \min_{j \in J_W} |\pi_j(v)|^{\sim}.$$

For each $i = 1, \dots, d$, let $w_i \in W_0$ be such that

$$\tilde{d}(u_i, w_i) \ll c \min_{j \in J_W} |\pi_j(v)|^{\sim},$$

where the involved constant depends on the numbers $\beta_{j,k}$. Using the assumption that $\bigoplus_{j\notin J_W} V_j \subset W_0$, after adjusting w_i, we can moreover ensure that

$$w_i - u_i \in \bigoplus_{j\in J_W} V_j.$$

We can bound

$$\big\|w_1\wedge\cdots\wedge w_d - u_1\wedge\cdots\wedge u_d\big\| \ll c^{\min_j \frac{1}{\alpha_j}} \prod_{j=1}^{s} \|\pi_j(v)\|^{\dim V_j}. \tag{3.7}$$

Indeed, developing the first wedge product using $w_i = u_i + (w_i - u_i)$ and then decomposing each vector along $V_1\oplus\cdots\oplus V_s$ and further developing the sum, we can express $w_1\wedge\cdots\wedge w_d - u_1\wedge\cdots\wedge u_d$ as a sum of wedge products of d vectors of the following types

(i) (first type) $\pi_j(w_i - u_i)$ with $j\in J_W$, or
(ii) (second type) $\pi_j(u_i)$ with $1\le j\le s$.

In each wedge product, the first type appears at least once and the product is zero unless π_j appears exactly $\dim V_j$ times. We can bound vectors of the first type by

$$\|\pi_j(w_i-u_i)\| \le \tilde{d}(w_i,u_i)^{\frac{1}{\alpha_j}} \ll c^{\frac{1}{\alpha_j}}\|\pi_j(v)\|$$

and vectors of the second type by

$$\|\pi_j(u_i)\| \ll \|\pi_j(v)\|.$$

This proves (3.7).

We then choose c be to small enough so that (3.7) combined with (3.6) implies $w_1\wedge\cdots\wedge w_d \neq 0$ contradicting the condition that W_0 is a proper linear subspace of V. □

The second lemma is an elementary computation using the definition of the quasi-norm. It is instructive to convince oneself with a picture that the lemma holds when the quasi-norm is simply the euclidean norm on $\mathbb{R}^d$.

Lemma 3.6. *Let $(u_i)_{1\le i\le d}$ be a union of orthonormal bases of each of the V_j, $j=1,\dots,r$. For $i=1,\dots,d$, denote by $j(i)$ the unique integer such that $u_i\in V_{j(i)}$.*

Let $v\in V$ and let $W\subset V$ be an affine hyperplane with with $V_0\subset W-W$. Let $\varphi_W\colon V\to\mathbb{R}$ be an affine map such that

$$W = \{\, v\in V \mid \varphi_W(v) = 0\,\}.$$

Let $\ell_W\colon V\to\mathbb{R}$ denote the linear part of φ_W. We have for any $v\in V$,

$$\tilde{d}(v,W) \asymp \min_{i:\ell_W(u_i)\neq 0} \left|\frac{\varphi_W(v)}{\ell_W(u_i)}\right|^{\alpha_{j(i)}}.$$

Proof. Note that $\ell_W(u_i)\neq 0$ implies that $j(i)\in J_W$ and $J_W\subset\{1,\dots,s\}$ because $V_0\subset W-W$. It follows that $\alpha_{j(i)}$ is defined and positive.

For any $i\in\{1,\dots,d\}$ with $\ell_W(u_i)\neq 0$, we have $v - \frac{\varphi_W(v)}{\ell_W(u_i)}u_i\in W$. Hence

$$\tilde{d}(v,W) \le \left|\frac{\varphi_W(v)}{\ell_W(u_i)}\right|^{\alpha_{j(i)}}.$$

Let $u\in V$ be such that $v-u\in W$. Write $u=\sum_{i=1}^d x_i u_i$. Then

$$\varphi_W(v) = \varphi_W(v-u) + \ell_W(u) = \sum_{i=1}^{d} x_i \ell_W(u_i).$$

It follows that there exists i with $\ell_W(u_i) \neq 0$ and such that

$$|x_i| \geq \frac{1}{d} \left| \frac{\varphi_W(v)}{\ell_W(u_i)} \right|.$$

This allows to conclude since

$$\tilde{d}(v, v-u) = |u|^{\sim} \geq \|\pi_{j(i)}(u)\|^{\alpha_{j(i)}} \geq |x_i|^{\alpha_{j(i)}}.$$

□

To conclude, we explain how to obtain Proposition 3.2 from Proposition 3.4. The argument is essentially the same as the one used in the particular case where E is simple.

Proof of Proposition 3.2, general case. For $x \in E$, let $L_x : E \to E$ and $R_x : E \to E$ denote the left and right multiplication by x, respectively. Then, define

$$L\colon \begin{array}{ccc} E & \to & \operatorname{End} E \\ x & \mapsto & L_x \end{array} \quad \text{and} \quad R\colon \begin{array}{ccc} E & \to & \operatorname{End} E \\ x & \mapsto & R_x \end{array}$$

Given a probability measure μ on G, we define a probability measure $\bar{\mu}$ on $\mathrm{GL}(E)$ by

$$\bar{\mu} = \frac{1}{2} L_* \mu + \frac{1}{2} R_* \mu.$$

The group $\bar{G}$ generated by $\bar{\mu}$ is isomorphic to $G \times G$ and the decomposition of E into irreducible $\bar{G}$-submodules is simply the decomposition into simple ideals $E = \oplus_j E_j$. By definition, the algebra generated by G contains the unit 1_{E_j} of E_j for each j. It follows that the linear span of $\bar{G}$ contains all projections $\pi_j \colon E \to E_j$. Moreover, in an appropriate basis, the elements of $\operatorname{supp} \bar{\mu}$ have integer coefficients, so we may apply Proposition 3.4 to $\bar{\mu}$, with vector $v = 1_E$. Note that if $\bar{g}$ is a random element distributed according to $\bar{\mu}^{*n}$, then $\bar{g} \cdot 1_E$ has law μ^{*n}, and therefore we obtain $\kappa > 0$ such that uniformly over all affine hyperplanes $W \subset E$ with $W - W \supset E_0$,

$$\forall n \geq 0,\ \forall \rho \geq e^{-n}, \quad \mu^{*n}\big(\big\{\, g \in G \,\big|\, \tilde{d}(g, W) \leq \rho \min_{j \in J_W} |\pi_j(g)|^{\sim} \,\big\}\big) \ll \rho^{\kappa}.$$

□

3.2. **Non-concentration at singular matrices.** As in the previous paragraph, μ denotes a probability measure on $\mathrm{SL}_d(\mathbb{Z})$. We assume that the algebraic group G generated by μ is semisimple and connected, and let E be the algebra generated by G in $\mathcal{M}_d(\mathbb{R})$. Recall that for $x \in E$, we defined $\det_E(x)$ to be the determinant of the map $E \to E$, $y \mapsto xy$. Note that $\det_E$ is a homogeneous polynomial function on E of degree equal to $\dim E$. Recall also that

$$\mu_n = (\pi' \circ \varphi_n)_*(\mu^{*n}),$$

where $\pi' \colon E \to E'$ is the projection to the direct sum $E' = E_1 \oplus \cdots \oplus E_s$ of all simple ideals with non-zero Lyapunov exponent, and $\varphi_n \colon E \to E$ is the renormalisation map defined in (4.2).

Lemma 3.7. *Given $\omega > 0$ there exists $c = c(\mu, \omega) > 0$ such that the following holds.*

$$\forall n \geq 0,\ \forall y \in E', \quad \mu_n^{\boxplus D}\big(\big\{\, x \in E' \,\big|\, |\det_{E'}(x - y)| \leq e^{-\omega n} \,\big\}\big) \ll e^{-cn}.$$

Note that for all x in E', $\det_{E'}(x) = \prod_{j=1}^{s} \det_{E_j}(\pi_j(x))$ and hence for every $n \geq 0$, and every $x \in E$,

$$\det_{E'}(\pi' \circ \varphi_n(x)) = \prod_{j=1}^{s} e^{-(\dim E_j)\lambda_1(\mu, E_j)n} \det_{E_j}(\pi_j(x)).$$

This immediately reduces the proof of Lemma 3.7 to the following.

Lemma 3.8. *Given $\omega > 0$ there exists $c = c(\mu, \omega) > 0$ such that the following holds for every $j = 1, \dots, s$, all $n \geq 0$ and all $y \in E_j$,*

$$(\mu^{*n})^{\boxplus D}\big(\{\, x \in E \,|\, |\det_{E_j}(\pi_j(x) - y)| \leq e^{(\dim E_j)\lambda_1(\mu, E_j)n - \omega n}\,\}\big) \ll e^{-cn}.$$

The idea is to apply [22, Proposition 3.2], where the case where E is simple were treated. However, upon projecting to a simple factor, the random walk might no longer be defined with integer coefficients: simple factors of E are only defined over a number field. So we could not apply [22, Proposition 3.2] as it is stated. Nevertheless, we can remark that, in the proof of [22, Proposition 3.2], [22, Lemma 3.13] holds more generally for the projected random walk from E to each E_j and then the rest of the proof of [22, Proposition 3.2] for a projected random walk is identical.

Here is the detailed proof. We need two ingredients from [22]. For a probability measure μ on a semisimple Lie group G and a finite dimensional linear representation (ρ, V) of G over $\mathbb{R}$, recall that

$$\lambda_1(\mu, V) = \lim_{n \to +\infty} \frac{1}{n} \int_G \log\|\rho(g)\| \mathrm{d}\mu^{*n}(g)$$

denotes the top Lyapunov exponent associated to the random walk induced on V. By semisimplicity V is a sum of irreducible sub-representations. The sum of irreducible sub-representations of same top Lyapunov exponent is a sum of isotypical components. For $\lambda \in \mathbb{R}$, we will denote by $p_\lambda \colon V \to V$ the G-equivariant projection onto

$$\sum_{\substack{V' \subset V, \text{ irreducible} \\ \lambda_1(\mu, V') \geq \lambda}} V'.$$

The following is [22, Proposition 3.17].

Proposition 3.9. *Let μ be a probability measure on $\mathrm{SL}_d(\mathbb{Z})$ having a finite exponential moment. Let G denote the Zariski closure of the subgroup generated by $\mathrm{supp}(\mu)$ in $\mathrm{SL}_d(\mathbb{R})$. Assume that G is semisimple and Zariski connected. Given $D \geq 1$, $\lambda \geq 0$, and $\omega > 0$, there is $c = c(\mu, D, \lambda, \omega) > 0$ such that the following holds for every $f \in \mathbb{R}[G]_{\leq D}$.*

$$\forall n \geq 0, \quad \mu^{*n}\big(\{\, g \in G \,|\, |f(g)| \leq e^{(\lambda - \omega)n} \|p_\lambda(f)\|\,\}\big) \ll e^{-cn}.$$

Here $p_\lambda \colon \mathbb{R}[G]_{\leq D} \to \mathbb{R}[G]_{\leq D}$ is defined as above.

The following is [22, Lemma 3.18]. For $k \geq 1$ and a measure μ on G, $\mu^{\otimes k} = \mu \otimes \cdots \otimes \mu$ denotes the product measure on $G^k = G \times \cdots \times G$.

Lemma 3.10. *Let V be a Euclidean space. Let μ be a Borel probability measure on $\mathrm{SL}(V)$ having a finite exponential moment. Let G denote the Zariski closure of the subgroup generated by $\mathrm{supp}(\mu)$ in $\mathrm{SL}_d(\mathbb{R})$. Assume that G is Zariski connected, is not compact and acts irreducibly on V.*

Let E denote the $\mathbb{R}$-span of G in $\mathrm{End}(V)$. Let $f \in \mathbb{R}[E]_{\leq D}$ and let $k \geq \dim E$ be an integer. Define $F \in \mathbb{R}[G^k]_{\leq D}$ to be the polynomial function

$$\forall (x_1, \dots, x_k) \in G^k, \quad F(x_1, \dots, x_k) = f(x_1 + \cdots + x_k).$$

Then we have

$$\|f_D\|_{\mathbb{R}[E]_{\leq D}} \ll_{\mu, D, k} \|p_{D\lambda_1(\mu, V)}(F)\|_{\mathbb{R}[G^k]_{\leq D}},$$

where f_D is the degree D homogeneous part of f and $p_{D\lambda_1(\mu,V)} \colon \mathbb{R}[G^k]_{\leq D} \to \mathbb{R}[G^k]_{\leq D}$ denotes the projection to the sum of irreducible G^k-subrepresentations $M \subset \mathbb{R}[G^k]_{\leq D}$ with $\lambda_1(\mu^{\otimes k}, M) \geq D\lambda_1(\mu, V)$.

Proof of Lemma 3.8. Fix $j = 1, \dots, s$. Remember that E_j is a simple algebra over $\mathbb{R}$. Using Wedderburn's structure theorem, we can find a irreducible faithful linear representation $E_j \to \mathrm{End}(V_j)$. It is easy to see that $\lambda_1(\mu, E_j) = \lambda_1\big(\pi_{j*}\mu, V_j\big)$. The Zariski closure of the subgroup generated by $\mathrm{supp}(\pi_{j*}\mu)$ is precisely $\pi_j(G)$. It spans E_j, is Zariski connected, acts irreducibly on V_j and is not compact. Thus, we can apply Lemma 3.10 to $\pi_{j*}\mu$.

Let $y \in E_j$ and consider the polynomial function $f \in \mathbb{R}[E_j]$, $f(x) = \det_{E_j}(x-y)$. The degree of f is $D_j = \dim E_j$ and its degree D_j homogeneous part is $\det_{E_j}$. Recall $D = \dim E$. Consider $F \in \mathbb{R}[\pi_j(G)^D]_{\leq D_j}$ defined as

$$\forall x_1, \dots, x_D \in \pi_j(G), \quad F(x_1, \dots, x_D) = f(x_1 + \cdots + x_D).$$

By Lemma 3.10,

$$\big\|p_{D_j\lambda_1(\mu,E_j)}(F)\big\|_{\mathbb{R}[\pi_j(G)^D]_{\leq D_j}} \gg \|\det_{E_j}\|_{\mathbb{R}[E_j]_{\leq D_j}} \gg_E 1.$$

The linear map $\Theta_j \colon \mathbb{R}[\pi_j(G)^D] \to \mathbb{R}[G^D]$ obtained by precomposing $(\pi_j, \dots, \pi_j)$ is injective and sends irreducible $\pi_j(G)^D$-subrepresentations to irreducible G^D-subrepresentations. Moreover, for any irreducible $\pi_j(G)^D$-subrepresentation $M \subset \mathbb{R}[\pi_j(G)^D]$, we have

$$\lambda_1\big((\pi_{j*}\mu)^{\otimes D}, M\big) = \lambda_1\big(\mu^{\otimes D}, \Theta_j(M)\big).$$

It follows that

$$\big\|p_{D_j\lambda_1(\mu,E_j)}(F \circ (\pi_j, \dots, \pi_j))\big\|_{\mathbb{R}[G^D]_{\leq D_j}} \gg_E 1.$$

Then we obtain Lemma 3.8 by applying Proposition 3.9 to the measure $\mu^{\otimes D}$ and the polynomial function $F \circ (\pi_j, \dots, \pi_j) \in \mathbb{R}[G^D]_{\leq D_j}$. □

3.3. **Proof of Proposition 3.1.** In order to obtain the required non-concentration properties for the measure μ_n, we shall use the basic large deviation estimates for matrix products that have already been used in the proof of Proposition 3.4. The statement below is taken from Boyer [14, Theorem A.5], which generalizes previous results of Le Page [27] and Bougerol [7, Theorem V.6.2].

Theorem 3.11 (Large deviation estimates)**.** *Let μ be a Borel probability measure on $\mathrm{GL}_d(\mathbb{R})$ having a finite exponential moment. For any $\omega > 0$, there is $c = c(\mu, \omega) > 0$, such that the following holds.*

(i) For all $n \geq 1$,

$$\mu^{*n}\Big(\Big\{\, g \in \Gamma \;\Big|\; \Big|\frac{1}{n}\log\|g\| - \lambda_1(\mu, \mathbb{R}^d)\Big| \geq \omega \,\Big\}\Big) \ll_\omega e^{-cn}.$$

(ii) Assume further that the group generated by $\mathrm{supp}(\mu)$ acts irreducibly on $\mathbb{R}^d$. For all $n \geq 1$ and all $v \in \mathbb{R}^d \setminus \{0\}$,

$$\mu^{*n}\Big(\Big\{\, g \in \Gamma \;\Big|\; \Big|\frac{1}{n}\log\frac{\|gv\|}{\|v\|} - \lambda_1(\mu, \mathbb{R}^d)\Big| \geq \omega \,\Big\}\Big) \ll_\omega e^{-cn}.$$

To prove Proposition 3.1, we shall only need the first item; the second item will be used later in Section 6.

Proof of Proposition 3.1. Note that condition $\mathrm{NC}(\varepsilon, \kappa, \tau)$ was defined for algebras endowed with a norm, and not with a quasi-norm. However, for some constants $\alpha, \beta > 0$, we have, for every $v \in E'$,

$$\begin{cases} \|v\| \leq |v|^{\sim\,\alpha} & \text{if } \|v\| \geq 1 \\ \|v\| \leq |v|^{\sim\,\beta} & \text{if } \|v\| < 1. \end{cases}$$

So if some measure satisfies condition $\mathrm{NC}(\varepsilon, \kappa, \tau)$ for the quasi-norm $|\cdot|^\sim$ on E', then it satisfies $\mathrm{NC}(\alpha\varepsilon, \frac{\kappa}{\beta}, \tau)$ for the usual norm $\|\cdot\|$ on E'. It is therefore sufficient

to check the non-concentration properties of μ_n with respect to the quasi-distance $\tilde{d}$.

For that, let $\varepsilon > 0$ be some small parameter. By Theorem 3.11(i) applied to each $\pi_{j*}\mu$, there exists $\tau = \tau(\mu, \varepsilon) > 0$ such that

$$\mu_n\big(\{\, g \in E' \,|\, \forall j = 1, \dots, s,\ |\pi_j(g)|^\sim \geq e^{-\varepsilon n} \,\}\big) \geq 1 - e^{-\tau n}.$$

Let ν_0 be the restriction of μ_n to such g and write

$$\mu_n = \nu_0 + \nu_1$$

so that $\nu_1(E) \leq e^{-\tau n}$. By Proposition 3.2, there exists $\kappa = \kappa(\mu) > 0$ such that for any affine hyperplane $W \subset E$ with $E_0 \subset W - W$,

$$\forall \rho \geq e^{-n}, \quad \mu^{*n}\big(\{\, g \in G \,|\, \tilde{d}(g, W) < \rho \min_{j \in J_W} |\pi_j(g)|^\sim \,\}\big) \ll \rho^\kappa.$$

By definition of φ_n and of the quasi-norm $|\cdot|$ on E, we have $|\varphi_n(g)| = e^{-n}|g|$ for every $g \in G$ and therefore, for every affine hyperplane $W \subset E'$,

$$\forall \rho \geq e^{-n}, \quad \mu_n\big(\{\, g \in E' \,|\, \tilde{d}(g, W) < \rho \min_{j \in J_W} |\pi_j(g)|^\sim \,\}\big) \ll \rho^\kappa.$$

By definition of ν_0, this implies

$$\forall \rho \geq e^{-n}, \quad \nu\big(\{\, g \in E' \,|\, \tilde{d}(g, W) < \rho e^{-\varepsilon n} \,\}\big) \ll \rho^\kappa \tag{3.8}$$

and this inequality is still valid for any convolution $\nu \boxplus \eta$. On the other hand, Lemma 3.7 shows that

$$\forall y \in E', \quad \mu_n^{\boxplus D}\big(\{\, x \in E' \,|\, |\det_{E'}(x - y)| \leq e^{-\omega n} \,\}\big) \ll e^{-\tau n}. \tag{3.9}$$

Let η_0 be the restriction of $\nu \boxplus \mu_n^{\boxplus(D-1)}$ to $\mathrm{B}_{E'}(0, e^{2\varepsilon n})$, and write

$$\mu_n^{\boxplus D} = \eta_0 + \eta_1.$$

By Theorem 3.11(i), we have $\eta_1(E') \leq e^{\tau n}$ for some $\tau = \tau(\varepsilon) > 0$, and by equations (3.8) and (3.9), the measure η_0 satisfies $\mathrm{NC}_0(2\varepsilon, \frac{\kappa}{2}, \tau)$. □

4. Fourier spectrum of the random walk

Let μ be a probability measure on $\mathrm{GL}_d(\mathbb{Z})$. Denote by $\Gamma \subset \mathrm{GL}_d(\mathbb{Z})$ the subgroup generated by $\mathrm{supp}(\mu)$ and $G \subset \mathrm{GL}_d(\mathbb{R})$ the Zariski closure of Γ in $\mathrm{GL}_d(\mathbb{R})$. Under the assumption that G is semisimple, we want to show some Fourier decay property for the measure μ^{*n} on the algebra E generated by G.

4.1. Connected case.

The result we need about Fourier decay for random walks is particularly transparent and easy to prove when the algebraic group G generated by μ is Zariski connected. So we first explain this particular case. Recall that $\varphi_n : E \to E$ is the rescaling automorphism given by (4.2), that $\pi' : E \to E'$ denotes the projection to the direct sum of all non-compact factors in E, and that for any integer $n \geq 1$, we let

$$\mu_n = (\pi' \circ \varphi_n)_*(\mu^{*n})$$

be the image of μ^{*n} after rescaling and projection to E'. We will denote by E'^* the space of linear forms on E' over the real numbers.

Theorem 4.1 (Fourier decay for random walks in E')**.** *Assume that G is semisimple and Zariski connected. Then there exists $\alpha_0 = \alpha_0(\mu) > 0$ such that for every $\alpha_1 \in (0, \alpha_0)$, there exists $c_0 = c_0(\mu, \alpha_1) > 0$ such that for all n sufficiently large, for all $\xi \in E'^*$ with*

$$e^{\alpha_1 n} \leq \|\xi\| \leq e^{\alpha_0 n}$$

the following estimate on the Fourier transform of μ^{*n} holds:

$$|\widehat{\mu_n}(\xi)| \leq e^{-c_0 n}.$$

We let E' act on E'^* on the right by

$$\forall \xi \in E'^*,\ \forall x, y \in E', \quad (\xi \cdot x)(y) = \xi(xy).$$

Moreover, we let E act on E'^* via π'.

Proof. Let $D = \dim E'$, $\varepsilon = \varepsilon(E', \kappa, D)$ and $s = s(E', \kappa)$ be the quantities given by Corollary 2.11. By Proposition 3.1, given $\alpha_1 \in (0,1)$, there exists $\kappa > 0$ such that for any $\varepsilon > 0$, there exists $\tau > 0$ such that $\mu_n^{\boxplus D}$ satisfies $\mathrm{NC}(\frac{\alpha_1 \varepsilon}{2}, \kappa, \tau)$ at scale e^{-n} in E' for all n sufficiently large. This formally implies that $\mu_n^{\boxplus D}$ satisfies $\mathrm{NC}(\varepsilon, \kappa, \tau)$ at all scales $\delta \in [e^{-n}, e^{-\frac{\alpha_1 n}{2}}]$.

Without loss of generality, we may of course assume that $\tau \in (0, \varepsilon\kappa)$. Let $\xi \in E'^*$ be such that $e^{\frac{\alpha_1 n}{2}} \leq \|\xi\| \leq e^n$. Taking $\delta = \|\xi\|^{-1}$, we have $\delta \in [e^{-n}, e^{-\frac{\alpha_1 n}{2}}]$ so $\mu_n^{\boxplus D}$ satisfies $\mathrm{NC}(\varepsilon, \kappa, \tau)$ at scale δ. Therefore, Corollary 2.11 shows that $\mu_{sn} = \mu_n * \cdots * \mu_n$ satisfies

$$|\widehat{\mu_{sn}}(\xi)| \leq e^{-\varepsilon\tau n}.$$

This shows the desired property if $n \in s\mathbb{Z}$. In general, take $\alpha_0 = \frac{1}{4s}$. For n large and $\xi \in E'^*$ such that $e^{\alpha_1 n} \leq \|\xi\| \leq e^{\alpha_0 n}$, write $n = sm + r$, with $0 \leq r < s$ and

$$\widehat{\mu_n}(\xi) = \int_G \widehat{\mu_{sm}}(\xi \cdot x) \mathrm{d}\mu_r(x).$$

Then, observe from the exponential moment assumption that outside of a set of μ_r-measure at most e^{-cn}, one has $e^{-\frac{\alpha_1 n}{2}} \|\xi\| \leq \|\xi \cdot x\| \leq e^{\frac{n}{2s}} \|\xi\|$ and so

$$e^{\frac{\alpha_1 m}{2}} \leq e^{\frac{\alpha_1 n}{2}} \leq \|\xi \cdot x\| \leq e^{\frac{n}{2s}} e^{\frac{n}{4s}} \leq e^m.$$

For such $\xi \cdot x$, we may bound

$$|\widehat{\mu_{sm}}(\xi \cdot x)| \leq e^{-\varepsilon\tau m} \leq e^{-\frac{\varepsilon\tau n}{2s}}$$

whence

$$|\widehat{\mu_n}(\xi)| \leq e^{-\frac{\varepsilon\tau n}{2s}} + e^{-cn} \leq e^{-c_0 n}$$

with $c_0 = \min(\frac{c}{2}, \frac{\varepsilon\tau}{4s})$. □

4.2. **Disconnected case.** As before, μ denotes a probability measure on $\mathrm{GL}_d(\mathbb{Z})$, and G the algebraic group generated by μ. We still assume that G is semisimple but no longer that it is Zariski connected. The identity component G° is then a finite index subgroup in G. We now write $\bar{E}$ for the subalgebra generated by G in $\mathcal{M}_d(\mathbb{R})$. As before, we decompose

$$\bar{E} = \bar{E}_1 \oplus \cdots \oplus \bar{E}_r \tag{4.1}$$

into simple ideals and define the rescaling automorphism $\varphi_n \colon \bar{E} \to \bar{E}$ by

$$\varphi_n(g) = \sum_{j=1}^{r} e^{-n\lambda_1(\mu, \bar{E}_j)} \pi_j(g) \tag{4.2}$$

where $\lambda_1(\mu, \bar{E}_j)$ denotes the top Lyapunov exponent associated to μ on each of the factors $\bar{E}_j$. Also, we assume that $\lambda_1(\mu, \bar{E}_j) = 0$ if and only if $j > s$ and denote by $\pi' \colon \bar{E} \to \bar{E}' = \bar{E}_1 \oplus \cdots \oplus \bar{E}_s$ the projection to the non-compact factors.

Example. When G is not Zariski connected, we shall write $\bar{E}$ for the algebra generated by G, and let E denote the algebra generated by the identity component G° of G. Let $a_0 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $a_1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ and $w = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Then define by blocks $A_0 = \begin{pmatrix} w & 0 \\ 0 & a_0 \end{pmatrix}$ and $A_1 = \begin{pmatrix} w & 0 \\ 0 & a_1 \end{pmatrix}$ in $\mathrm{SL}_4(\mathbb{Z})$ and set

$$\mu = \frac{1}{4}(\delta_{A_0} + \delta_{A_1} + \delta_{A_0^{-1}} + \delta_{A_1^{-1}}).$$

One has $G \simeq (\mathbb{Z}/4\mathbb{Z}) \times \mathrm{SL}_2(\mathbb{R})$ and $\bar{E} \simeq \mathbb{C} \times M_2(\mathbb{R})$. On the other hand, the algebra generated by G° is $E \simeq \mathbb{R} \times M_2(\mathbb{R})$ if one identifies $\mathbb{C} \simeq \{\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \ ; \ a, b \in \mathbb{R}\}$ and $\mathbb{R} \simeq \mathbb{R}1$. The law μ^{*n} of the random walk at time n is supported by E if n is even, and by $A_0 E \simeq i\mathbb{R} \times M_2(\mathbb{R})$ if n is odd. It is always concentrated on a proper subspace of $\bar{E}$.

To overcome this issue, we shall use the algebra $E \subset \bar{E}$ generated by the identity component G° in G. The group G° has finite index in G and we let

$$F = G/G^\circ.$$

With a slight abuse of notation, we identify F with a set of representatives in G and write G as a disjoint union

$$G = \bigsqcup_{\gamma \in F} \gamma G^\circ.$$

Any measure ν on G can then be decomposed uniquely in the form

$$\nu = \sum_{\gamma \in F} \gamma_* \nu_\gamma$$

where each ν_γ is a measure on E. Finally, we let $E' = \pi'(E)$, and for $n \geq 1$ and $\gamma \in F$,

$$\mu_{n,\gamma} = (\pi' \circ \varphi_n)_* [(\mu^{*n})_\gamma].$$

Fourier decay for (integer coefficient) random walks on non-connected semisimple groups can be stated as follows.

Theorem 4.2 (Fourier decay for random walks in E')**.** *Let μ, G, G° and F be as above. Then there exists $\alpha_0 = \alpha_0(\mu) > 0$ such that for every $\alpha_1 \in (0, \alpha_0)$, there exists $c_0 = c_0(\mu, \alpha_1) > 0$ such that for all n sufficiently large, all $\gamma \in F$ and $\xi \in E'^*$ with*

$$e^{\alpha_1 n} \leq \|\xi\| \leq e^{\alpha_0 n}$$

the following estimate on the Fourier transform of μ^{*n} holds:

$$|\widehat{\mu_{n,\gamma}}(\xi)| \leq e^{-c_0 n}.$$

One can ask whether the same statement holds if $\mathrm{supp}(\mu)$ is not assumed to have integer coefficients.

4.3. **Induced random walk on the identity component.** In order to prove Theorem 4.2, we shall use the induced random walk on G°, whose definition is given below. Since by definition G° is connected, this will allow us to use the results of Section 3. The drawback is that we can no longer use the simple identity $\mu_{sn} = \mu_n * \cdots * \mu_n$; so we shall have to write μ_{sn} as a weighted sum of convolutions related to the induced measure μ° on the identity component, which makes the argument more technical.

Let $(g_n)_{n \geq 1}$ be a sequence of independent random variables distributed according to μ. Consider the return times to G°,

$$\tau(1) = \inf\{\, n \geq 1 \mid g_n \cdots g_1 \in G^\circ \,\}$$

and recursively for $m \geq 2$,

$$\tau(m) = \inf\{\, n > \tau(m-1) \mid g_n \cdots g_1 \in G^\circ \,\}.$$

Those are the return times of a Markov chain on the finite space G/G°, so that for every $m \geq 1$, $\tau(m)$ is almost surely finite. In fact, by Kac's formula [6, Lemma 5.4]

$$\mathbb{E}[\tau(1)] = [G : G^\circ].$$

The random variables $(g_{\tau(m)} \cdots g_{\tau(m-1)+1})_{m\geq 0}$ are independent and identically distributed with law μ°, the law of $g_{\tau(1)} \cdots g_1$. Note that μ° is a probability measure on G° and has the following properties [1, Lemmas 4.40 and 4.42].

Lemma 4.3. *Let μ be a probability measure on a real algebraic group G and μ° the induced measure on the identity component G°. Let $T = [G : G^\circ]$. If μ admits some finite exponential moment, then:*

(i) The measure μ° has some finite exponential moment;
(ii) For every $\omega > 0$, there exists $c = c(\mu, \omega) > 0$ such that for all m sufficiently large, $\mathbb{P}\big[|\tau(m) - Tm| \geq \omega m\big] \leq e^{-cm}$.

In order to prove Theorem 4.2, we shall need to relate the random walk defined by μ and the one defined by μ°. For that, we introduce, for $m \geq 1$ and $\ell \geq 1$, the law ν_ℓ of the random variable

$$g_{\tau(m)} \cdots g_1 \quad \text{conditional to the event} \quad \tau(m) = \ell.$$

Naturally, ν_ℓ is also the law of the variable $g_\ell \cdots g_1$ conditional to $\tau(m) = \ell$. On the one hand, we may relate the measures ν_ℓ to $(\mu^\circ)^{*m}$ with the formula

$$(\mu^\circ)^{*m} = \sum_{\ell\in\mathbb{N}} p_\ell \nu_\ell. \tag{4.3}$$

where $p_\ell = \mathbb{P}[\tau(m) = \ell]$. Here, we are hiding the dependency of ν_ℓ and p_ℓ on m in order to make notation less cumbersome.

On the other hand, writing $\ell_1 + \cdots + \ell_s + k = n$ for some natural integers n, s and $\ell_1, \ldots, \ell_s$, we have

$$\mu^{*n} = \sum_{\ell_1+\cdots+\ell_s+k=n} p_{\ell_1} \cdots p_{\ell_s} \mu^{*k} * \nu_{\ell_s} * \cdots * \nu_{\ell_1} + ((\mathbb{P}[\tau(sm) > n])) \tag{4.4}$$

where the notation $((t))$ for some positive quantity t means some unspecified positive measure of total mass at most t. These two formulae will allow us to use the non-concentration properties of $(\mu^\circ)^{*m}$ to prove some Fourier decay estimate for μ^{*n}.

Before we derive Theorem 4.2, we note that the scaling automorphism φ_m° on E associated to μ° is simply given by $\varphi_n^\circ = \varphi_{mT}$, where $T = [G : G^\circ]$. This readily follows from the fact that if $\bar{E}_i$ is any simple ideal in $\bar{E}$ and V any G°-irreducible submodule of $\bar{E}_i$, then $\lambda_1(\mu^\circ, V) = T\lambda_1(\mu, \bar{E}_i)$.

Proof of Theorem 4.2. Let $\alpha_1 > 0$ be a given small number. Since the algebraic group generated by μ° is connected, Proposition 3.1 applies to the induced random walk on G°. We let $\kappa = \kappa(\mu^\circ) > 0$ be the constant given by that proposition. Let $D = \dim E$ and $s = s(E, \kappa) \geq 1$ and $\varepsilon = \varepsilon(E, \kappa, D) > 0$ be the constants given by Corollary 2.11.

Given $\alpha_1 > 0$, Proposition 3.1 shows that for all m large enough, the measure

$$(\pi' \circ \varphi_{mT})_*\big(((\mu^\circ)^{*m})^{\boxplus D} \boxminus ((\mu^\circ)^{*m})^{\boxplus D}\big)$$

satisfies $\mathrm{NC}(\frac{\alpha_1\varepsilon}{2}, \kappa, \tau)$ in E' at scale e^{-m} for some $\tau > 0$. This implies that the same measure satisfies $\mathrm{NC}(\frac{\varepsilon}{2}, \kappa, \tau)$ in E' at all scales $\delta \in [e^{-m}, e^{-\alpha m}]$. Without loss of generality, we may assume that $\tau < \kappa\varepsilon/2$ and $\tau < \varepsilon/2$.

Let $\omega = \omega(\mu, \alpha_1)$ be a constant whose value is to be determined later. Fix $n \geq 1$ large, and set $m = \big\lfloor (1 - 2\omega)\frac{n}{Ts} \big\rfloor$, where $T = [G : G^\circ]$. Everything below is true for n sufficiently large (larger than some n_0 depending on μ and α_1). The letter c denotes a small positive constant, whose value may vary from one line to the other, depending on μ and α_1 but independent of n.

By Lemma 4.3, we have

$$\mathbb{P}[\tau(sm) > n - \omega n] \leq e^{-cn}$$

and

$$\mathbb{P}[\tau(sm) < n - 3\omega n] \leq e^{-cn}.$$

Put

$$\mathcal{L} = \{\, \ell \in \mathbb{N} \mid p_\ell \geq e^{-\frac{\alpha_1 \tau}{2D} m} \,\}.$$

We can bound

$$\sum_{(\ell_1,\dots,\ell_s)\notin\mathcal{L}^s} p_{\ell_1}\cdots p_{\ell_s} \leq sne^{-\frac{\alpha_1\tau}{2D}m} \leq e^{-cn}.$$

Thus, (4.4) becomes

$$\mu^{*n} = \sum_{\substack{\ell_1,\dots,\ell_s\in\mathcal{L},\, \omega n\leq k\leq 3\omega n\\ \ell_1+\cdots+\ell_s+k=n}} p_{\ell_1}\cdots p_{\ell_s}\mu^{*k} * \nu_{\ell_s} * \cdots * \nu_{\ell_1} + ((e^{-cn})).$$

Let $\gamma \in F$. To finish the proof of the theorem, it suffices to establish an upper bound of the form e^{-cn} for the quantity

$$\begin{aligned} I_{\ell_1,\dots,\ell_s,k}(\xi) &:= \int_{\gamma G^\circ} e\big(\xi\circ\pi'\circ\varphi_n(\gamma^{-1}g)\big)\mathrm{d}\big(\mu^{*k} * \nu_{\ell_s} * \cdots * \nu_{\ell_1}\big)(g)\\ &= \iint_{g\in\gamma G^\circ} e\big(\xi\circ\pi'\big(\varphi_{n-smT}(\gamma^{-1}g)\varphi_{smT}(h)\big)\big)\mathrm{d}\mu^{*k}(g)\mathrm{d}(\nu_{\ell_s}*\cdots*\nu_{\ell_1})(h)\\ &= \int_{\gamma G^\circ} \big((\pi'\circ\varphi_{smT})_*(\nu_{\ell_s}*\cdots*\nu_{\ell_1})\big)^\wedge\big(\xi\cdot\varphi_{n-smT}(\gamma^{-1}g)\big)\mathrm{d}\mu^{*k}(g). \end{aligned}$$

uniformly for all $\ell_1,\dots,\ell_s \in \mathcal{L}$ and $\omega n \leq k \leq 3\omega n$ with $\ell_1+\cdots+\ell_s+k=n$.

First, we claim that uniformly for all $\ell \in \mathcal{L}$, the measure

$$(\pi'\circ\varphi_{mT})_*\big(\nu_\ell^{\boxplus D} \boxminus \nu_\ell^{\boxplus D}\big)$$

satisfies $\mathrm{NC}(\varepsilon,\kappa,2\tau)$ in E' at all scales $\delta \in [e^{-m}, e^{-\alpha_1 m}]$, provided that $m \geq 1$ is large enough. Indeed, developing $((\mu^\circ)^{*m})^{\boxplus D} \boxminus ((\mu^\circ)^{*m})^{\boxplus D}$ using (4.3), we see that for any $\ell \geq 1$,

$$((\mu^\circ)^{*m})^{\boxplus D} \boxminus ((\mu^\circ)^{*m})^{\boxplus D} = p_\ell^{2D}\big(\nu_\ell^{\boxplus D} \boxminus \nu_\ell^{\boxplus D}\big) + ((1)).$$

The inequality $p_\ell^{2D} \geq e^{-\alpha_1\tau m}$ for $\ell \in \mathcal{L}$ together with the fact that the left-hand side rescaled by $\pi'\circ\varphi_{mT}$ satisfies $\mathrm{NC}(\frac{\varepsilon}{2},\kappa,\tau)$ in E' at all scales $\delta \in [e^{-m}, e^{-\alpha_1 m}]$ show our claim. By Corollary 2.11, this implies, for all $\zeta \in (E')^*$ such that $e^{\alpha_1 m} \leq \|\zeta\| \leq e^m$,

$$\big|\big((\pi'\circ\varphi_{smT})_*(\nu_{\ell_s}*\cdots*\nu_{\ell_1})\big)^\wedge(\zeta)\big| \leq e^{-\frac{\alpha_1\varepsilon\tau}{(2D)^s}m} \leq e^{-cn}.$$

Note that for any $g \in \gamma G^\circ$,

$$\|\xi\|\|g^{-1}\|^{-1} \ll \|\xi\cdot\varphi_{n-smT}(\gamma^{-1}g)\| \ll \|\xi\|\|\varphi_{n-smT}\|\|g\|. \tag{4.5}$$

On the one hand, we have $0 \leq n - smT \leq 3\omega n$. Hence, there exists a constant $C = C(\mu) \geq 1$ such that

$$\|\varphi_{n-smT}\| \leq e^{C\omega n}.$$

On the other hand, using the assumption that μ has a finite exponential moment and Markov's inequality, we can find a constant $C = C(\mu) \geq 1$ such that for any $k \geq 1$, the μ^{*k}-measure of the set of $g \in \Gamma$ such that

$$\|g\| \leq e^{Ck} \quad \text{and} \quad \|g^{-1}\| \leq e^{Ck} \tag{4.6}$$

is at least $1 - e^{-k}$.

Set $\alpha_0 = \frac{1}{4Ts}$ and let $\xi \in (E')^*$ be such that $e^{\alpha_1 n} \leq \|\xi\| \leq e^{\alpha_0 n}$. Using $k \leq 3\omega n$, we have, for any $g \in \mathrm{supp}(\mu^{*k})$ satisfying (4.6),

$$e^{(\alpha_1 - 4C\omega)n} \leq \|\xi\cdot\varphi_{n-smT}(\gamma^{-1}g)\| \leq e^{(\alpha_0+5C\omega)n}.$$

With the choice $\omega = \min\{\frac{\alpha_1}{8C}, \frac{1}{20CTs}\}$, we can guarantee that this implies

$$e^{\alpha_1 m} \leq e^{\alpha_1 n/2} \leq \|\xi \cdot \varphi_{n-smT}(\gamma^{-1} g)\| \leq e^m.$$

Putting everything together, we obtain

$$|I_{\ell_1,\dots,\ell_s,k}(\xi)| \leq e^{-cn} + e^{-k} \leq e^{-cn} + e^{-\omega n}.$$

for all $\ell_1, \dots, \ell_s \in \mathcal{L}$, $\omega n \leq k \leq 3\omega n$ with $\ell_1 + \cdots + \ell_s + k = n$. This concludes the proof of the theorem. □

5. From Fourier decay to granular structure

As in the previous section, μ denotes a probability measure on $\mathrm{GL}_d(\mathbb{Z})$ and we study the random walk associated to μ on $\mathbb{T}^d$, with starting distribution $\nu \in \mathcal{P}(\mathbb{T}^d)$. The law of the walk at time n is $\nu_n = \mu^{*n} * \nu$. The goal of this section is to show that if ν_n has a large Fourier coefficient, then the starting distribution ν must have some strong concentration property.

5.1. **Concentration statement for the random walk.** In order to state the main proposition of this section, we need to set up some notation. As before, G denotes the algebraic subgroup generated by μ, $E \subset \mathcal{M}_d(\mathbb{R})$ denotes the algebra generated by the identity component G° of G, F denotes the finite group G/G° and $T = \#F$.

Changing notation slightly, we now consider a decomposition of E

$$E = E_0 \oplus E_1 \oplus \cdots \oplus E_r$$

into maximal sums of minimal ideals with same Lyapunov exponent for the action of μ°. We assume that the summands are ordered so that

$$\lambda_1(\mu^\circ, E_1) > \cdots > \lambda_1(\mu^\circ, E_r) > 0 = \lambda_1(\mu^\circ, E_0).$$

Here, E_0 is eventually trivial.

The group G acts naturally on the space $V = \mathbb{R}^d$ and for $1 \leq j \leq r$, we let V_i be the sum of all simple G°-submodules $W \subset V$ such that $\lambda_1(\mu^\circ, W) = \lambda_1(\mu^\circ, E_i)$. Equivalently, V_i is also the sum of all simple G-submodules $W \subset V$ such that $\lambda_1(\mu, W) = \frac{1}{T}\lambda_1(\mu^\circ, E_i)$. We have

$$V = V_0 \oplus V_1 \oplus \cdots \oplus V_r.$$

Let $\pi_i \colon V \to V_i$ denote the corresponding projection. Define a quasi-norm on V by

$$|v|^\sim = \max_{0 \leq i \leq s} \|\pi_i(v)\|^{\frac{1}{\lambda_1(\mu, V_i)}}$$

where by convention

$$\|\pi_0(v)\|^{\frac{1}{0}} = \|\pi_0(v)\|^{+\infty} = \begin{cases} 0 & \text{if } \|\pi_0(v)\| \leq 1 \\ +\infty & \text{otherwise.} \end{cases}$$

This induces a quasi distance on $\mathbb{T}^d$. For $x, y \in \mathbb{T}^d$, define

$$\tilde{d}(x,y) = \begin{cases} |v - w|^\sim & \text{if there are lifts } v \in V \text{ of } x \text{ and } w \in V \text{ of } y \text{ such that } \|v - w\| \leq \frac{1}{2}, \\ 1 & \text{otherwise.} \end{cases}$$

Neighborhoods of subsets of $\mathbb{T}^d$ with respect to this quasi-distance will be denoted by $\tilde{\mathrm{N}}\mathrm{bd}(\,\cdot\,,\,\cdot\,)$. Finally, for a rational subspace $W \subset V$, we let $W \bmod \mathbb{Z}^d$ denote its projection in $\mathbb{T}^d$, which is a subtorus.

Proposition 5.1. *Let μ be a probability measure on $\mathrm{GL}_d(\mathbb{Z})$ and ν be a Borel probability measure on $\mathbb{T}^d$. If the algebraic group G generated by μ is semisimple, then there exist $C = C(\mu) \geq 0$ and $\tau > 0$ such that the following holds.*

Assume that for some $t \in (0, \frac{1}{2})$,

$$|\widehat{\mu^{*n} * \nu}(a_0)| \geq t \quad \text{for some } a_0 \in \mathbb{Z}^d \text{ and } n \geq C \log \frac{\|a_0\|}{t}.$$

Then, there exists $\gamma \in F$ such that, denoting

$$W = (a_0 \gamma E)^\perp$$

there exists a finite subset $X \subset \mathbb{T}^d$ such that

$$(X - X) \cap \tilde{\mathrm{N}}\mathrm{bd}(W \bmod \mathbb{Z}^d, e^{-(1-2\tau)n}) = \{0\}$$

and

$$\nu\big(X + \tilde{\mathrm{N}}\mathrm{bd}(W \bmod \mathbb{Z}^d, e^{-(1-\tau)n})\big) \geq t^{O(1)}.$$

The proof of Proposition 5.1 is in two steps: First, using the results of Section 4, one shows that the inequality $|\widehat{\mu^{*n} * \nu}(a_0)| \geq t$ implies that $\mu^{*n} * \nu$ has many large Fourier coefficients (reducing slightly the value of n) and then, one applies a Fourier analysis lemma originating in the work of Bourgain, Furman, Lindenstrauss and Mozes [10, Proposition 7.5]. We start with the statement and proof of a general version of that lemma adapted to our needs.

5.2. **A quantitative version of Wiener's lemma.** Wiener's lemma in harmonic analysis states that a measure ν on the torus $\mathbb{T}^d$ is atom-free if and only if its Fourier series tends to zero in density, i.e. given $t > 0$, the proportion of vectors $a \in \mathbb{Z}^d$ in a large ball $B(0, N)$ such that $|\hat{\nu}(a)| \geq t$, tends to zero as N goes to infinity. In their paper [10], Bourgain, Furman, Lindenstrauss and Mozes observed that this statement could be made quantitative: If $B(0, N)$ contains a proportion at least $s > 0$ of integer vectors satisfying $|\hat{\nu}(a)| \geq t$, then there exists a ball $B = B(x, \frac{1}{N})$ of radius $\frac{1}{N}$ in $\mathbb{T}^d$ such that $\nu(B) \gg (st)^3$, where the involved constant depends only on d. In order to later be able to use the quasi-norm adapted to a random walk, we need to generalize this statement. It turns out to be most convenient to formulate the lemma in terms of convex sets and polar pairs.

We will need to generalize slightly the notation $\mathcal{N}(\,\cdot\,, \delta)$ of covering number. Instead of covering a set by balls, we will use translates of a convex body. Given a symmetric convex body $B \subset \mathbb{R}^d$ and $A \subset V$ a bounded non-empty subset, we define the *covering number of A by B* by

$$\mathcal{N}(A, B) = \min\Big\{ N \geq 1 \,\Big|\, \exists x_1, \dots, x_N \in V,\ A \subset \bigcup_{i=1}^{N} (x_i + B) \Big\}.$$

We shall also say that A is *B-separated*, if $(A - A) \cap B = \{0\}$. Let us briefly list some useful properties of covering numbers. These may be used without explicit mention in the sequel; the elementary proofs are left to the reader.

- Let $f \colon V \to W$ be a linear map to another Euclidean space W. Then, for any set $A \subset V$ and any convex body $B \subset W$,

$$\mathcal{N}\big(f(A), f(B)\big) \leq \mathcal{N}(A, B) \tag{5.1}$$

  with equality if f is a linear isomorphism.
- Let $B' \subset V$ be another convex body, then

$$\mathcal{N}(A, B') \leq \mathcal{N}(A, B)\mathcal{N}(B, B').$$

- In A, maximal B-separated subsets have cardinality at least $N(A, B)$.
- If B is symmetric and A is $2B$-separated then $\mathcal{N}(A, B) \geq \#A$.

- Let $f\colon V \to W$ be a surjective linear map between Euclidean spaces. Let $B, C \subset V$ be convex bodies and let $A \subset C$ be a subset of C. We have

$$\frac{\mathcal{N}\big(f(A), f(B)\big)}{\mathcal{N}\big(f(C), f(B)\big)} \gg \frac{\mathcal{N}(A,B)}{\mathcal{N}(C,B)} \tag{5.2}$$

  where the implied constant in the $\gg$ notation depends only on $\dim V$.

Let $\mathbb{R}^d$ be endowed with the usual scalar product. Given a symmetric convex body $C \subset \mathbb{R}^d$, its *polar set* C^* is defined by

$$C^* = \{\, x \in \mathbb{R}^d \,|\, \forall y \in C,\ \langle x, y\rangle \leq 1 \,\}.$$

If $C \supset B(0,2)$, then $C^* \subset B(0, \frac{1}{2})$ and we naturally identify C^* with its projection to $\mathbb{T}^d$. The quantitative version of Wiener's lemma that we need is given by the proposition below.

Proposition 5.2. *Let ν be a probability measure on $\mathbb{T}^d$ and write for $t > 0$*

$$A_t = \{\, a \in \mathbb{Z}^d \,|\, |\hat{\nu}(a)| \geq t \,\}.$$

Assume that for some symmetric convex subsets $B \subset C \subset \mathbb{R}^d$ containing $B(0,1)$, we have, for some $c_0 \in \mathbb{Z}^d$ and some $s > 0$

$$\mathcal{N}(A_t \cap (c_0 + C), B) \geq s \cdot \frac{|C|}{|B|}. \tag{5.3}$$

Then there exists a B^*-separated subset $X \subset \mathbb{T}^d$ such that

$$\nu(X + C^*) \gg_d s^{3/2} t^6.$$

Proof. The implied constants in the Vinogradov notation in this proof depend only on d. We shall need two auxiliary functions; the first one corresponds to the pair of convex sets (C, C^*), the second to (B, B^*):

(1) There exists a smooth function $\psi\colon \mathbb{T}^d \to \mathbb{R}_{\geq 0}$ such that
   (a) $\int_{\mathbb{T}^d} \psi = 1$,
   (b) $\psi \ll \frac{1}{|C^*|}\, \mathbb{1}_{C^*}$,
   (c) $\widehat{\psi} \gg \mathbb{1}_{2C \cap \mathbb{Z}^d}$.

(2) There exists a smooth function $\varphi\colon \mathbb{T}^d \to \mathbb{R}_{\geq 0}$ such that
   (i) $\varphi \gg \mathbb{1}_{B^*}$,
   (ii) $\widehat{\varphi}$ is real and positive and $\widehat{\varphi} \ll \frac{1}{|B|^2}\, \mathbb{1}_B \boxminus \mathbb{1}_B \leq \frac{1}{|B|}\, \mathbb{1}_{2B}$.

One obtains ψ by taking any smooth symmetric bump function supported on $\frac{1}{16}C^*$ with integral $\int \psi = 1$. The third property follows from the fact that for every $\xi \in 2C \cap \mathbb{Z}^d$ and $x \in \frac{1}{16}C^*$, one has $\langle \xi, x\rangle \leq \frac{1}{8}$ and hence $\Re\big(e(\langle \xi, x\rangle)\big) \geq \frac{1}{2}$. The function φ can be given explicitly by the formula $\varphi(x) = \Big|\frac{1}{|B|}\sum_{a \in \frac{1}{8}B \cap \mathbb{Z}^d} e(\langle a, x\rangle)\Big|^2$ for all x in $\mathbb{T}^d$. The second item is immediate by definition of φ, and the first one follows from the fact that by Minkowski's first theorem on convex bodies, one has $\#(\frac{1}{8}B \cap \mathbb{Z}^d) \gg |B|$.

Pick a maximal $2B$-separated subset $A' \subset A_t \cap (c_0 + C)$ such that all coefficients $\hat{\nu}(a)$, $a \in A'$ fall in the same quadrant of $\mathbb{C}$. One still has $\#A' \gg \mathcal{N}(A_t \cap (c_0 + C), B)$ and moreover

$$\Big|\sum_{a \in A'} \hat{\nu}(a)\Big| \geq \frac{t \#A'}{\sqrt{2}}.$$

By the Cauchy-Schwarz inequality,

$$\sum_{a,b \in A'} \hat{\nu}(a-b) = \int_{\mathbb{T}^d} \Big|\sum_{a \in A'} e(\langle a, x\rangle)\Big|^2 \mathrm{d}\nu(x) \geq \Big|\int_{\mathbb{T}^d} \sum_{a \in A'} e(\langle a, x\rangle) \mathrm{d}\nu(x)\Big|^2 \gg t^2 (\#A')^2.$$

Hence, there exists a translate A of A' such that $A \subset A' - A' \subset 2C$ and

$$\Big|\sum_{a\in A} \hat{\nu}(a)\Big| \gg t^2 \# A$$

and

$$\# A = \# A' \gg s \cdot \frac{|C|}{|B|}. \tag{5.4}$$

Consider the function $f\colon \mathbb{T}^d \to \mathbb{R}$ defined by

$$\forall x \in \mathbb{T}^d, \quad f(x) = \sum_{a\in A} e(\langle a, x\rangle).$$

On the one hand, using the definition of f, the properties of φ and the fact that A is $2B$-separated, one has, for any $y \in \mathbb{T}^d$,

$$\begin{aligned}
\int_{y+B^*} |f|^2 &= \int_{\mathbb{T}^d} \mathbb{1}_{B^*}(x-y)|f(x)|^2 \mathrm{d}x \\
&\ll \sum_{a_1,a_2\in A} \int \varphi(x-y) e(\langle a_1 - a_2, x\rangle) \mathrm{d}x \\
&\leq \sum_{a_1,a_2\in A} \hat{\varphi}(a_1 - a_2) \\
&\ll \frac{1}{|B|} \sum_{a_1,a_2\in A} \mathbb{1}_{2B}(a_1 - a_2) \\
&\ll \frac{\# A}{|B|}.
\end{aligned}$$

On the other hand, from the properties of ψ and of those of A,

$$\Big|\int_{\mathbb{T}^d} f \mathrm{d}(\nu \boxplus \psi)\Big| = \Big|\sum_{a\in A} \hat{\nu}(a)\hat{\psi}(a)\Big| \gg \Big|\sum_{a\in A} \hat{\nu}(a)\Big| \gg t^2 \# A.$$

Let $(y_i)_{i\in I}$ be a maximal family of $(4B^*)$-separated points in $\mathbb{T}^d$. Then the translates $(y_i + B^*)_{i\in I}$ are disjoint and have a total volume $\gg 1$. By Fubini's theorem,

$$\int \mathrm{d}x \sum_{i\in I} \int_{x+y_i+B^*} f \mathrm{d}(\nu \boxplus \psi) = \sum_{i\in I} |y_i + B^*| \int_{\mathbb{T}^d} f \mathrm{d}(\nu \boxplus \psi).$$

Hence, translation all y_i by some $x \in \mathbb{T}^d$ if necessary, we may assume that

$$\sum_{i\in I} \Big|\int_{y_i+B^*} f \mathrm{d}(\nu \boxplus \psi)\Big| \gg t^2 \# A. \tag{5.5}$$

By the Cauchy-Schwarz inequality, for each $i \in I$,

$$\begin{aligned}
\Big|\int_{y_i+B^*} f \mathrm{d}(\nu \boxplus \psi)\Big| &\leq \sqrt{\int_{y_i+B^*} |f|^2} \sqrt{\int_{y_i+B^*} (\nu \boxplus \psi)^2} \\
&\ll \sqrt{\frac{\# A}{|B|}} \sqrt{(\nu \boxplus \psi)(y_i + B^*) \max_{y_i+B^*} \nu \boxplus \psi} \\
&\ll \nu(y_i + B^* + C^*) \sqrt{\frac{h_i \# A |C|}{|B|}}
\end{aligned}$$

where $h_i = \frac{|C^*| \max_{y_i+B^*} \nu \boxplus \psi}{\nu(y_i+B^*+C^*)}$. Recalling (5.4) and (5.5), we obtain some constant $L = L(d) > 1$ depending only on d such that

$$\sum_{i\in I} \nu(y_i + B^* + C^*) h_i^{1/2} \geq \frac{s^{1/2}t^2}{L}.$$

On the other hand, since for every x in $\mathbb{T}^d$, one has $(\nu \boxplus \psi)(x) \ll \frac{\nu(x+C^*)}{|C^*|}$, increasing the value of L if necessary, we get

$$\forall i, \quad h_i \leq L.$$

Finally, $(y_i)_{i\in I}$ is $4B^*$-separated so

$$\sum_{i\in I} \nu(y_i + B^* + C^*) \leq \sum_{i\in I} \nu(y_i + 2\cdot B^*) \leq 1$$

and we may set $J = \{\, i \in I \,|\, h_i \geq \frac{st^4}{4L^2} \,\}$ to find

$$\sum_{i\in J} \nu(y_i + B^* + C^*) \geq \frac{s^{1/2}t^2}{2L}.$$

For each $i \in J$, fix $x_i \in y_i + B^*$ such that

$$(\nu \boxplus \psi)(x_i) = \max_{y_i+B^*} \nu \boxplus \psi$$

and let

$$X = \{x_i \ ; \ i \in J\}.$$

Since the family $(y_i)_{i\in I}$ is $4B^*$-separated, the set X is B^*-separated. For the second property, note that for each i in J,

$$\nu(x_i + C^*) \geq |C^*|(\nu \boxplus \psi)(x_i) = h_i \nu(y_i + B^* + C^*) \gg st^4 \nu(y_i + B^* + C^*).$$

so that

$$\nu(X + C^*) = \sum_{i\in J} \nu(x_i + C^*) \gg st^4 \sum_{i\in J} \nu(y_i + B^* + C^*) \gg s^{3/2}t^6.$$

□

5.3. **Proof of Proposition 5.1.** To prove Proposition 5.1, we follow the same pattern as in [22]. The only difference here is that we need to find the correct polar pairs (B, B^*) and (C, C^*) to apply Proposition 5.2. First, let us state two lemmas that are probably well-known.

Lemma 5.3 ([22, Lemma 4.3] Additive structure of large Fourier coefficients)**.** *Let μ be a Borel probability measure on $\mathrm{SL}_d(\mathbb{Z})$ and ν a Borel probability measure on $\mathbb{T}^d$. If*

$$|\widehat{\mu * \nu}(a_0)| \geq t_0 > 0,$$

then for any integer $k \geq 1$, the set

$$A = \left\{ g \in \mathcal{M}_d(\mathbb{Z}) \mid |\hat{\nu}(a_0 g)| \geq t_0^{2k}/2 \right\}$$

satisfies

$$\left(\mu^{\boxplus k} \boxminus \mu^{\boxplus k}\right)(A) \geq \frac{t_0^{2k}}{2}.$$

Lemma 5.4 ([22, Lemma 4.4] Regularity from Fourier decay)**.** *Given $D \geq 1$ and $\alpha > 0$ sufficiently small, there exist constants $c = c(D,\alpha) > 0$ and $C_1 = C_1(D,\alpha) > 0$ such that the following holds for all $0 < \delta < ct$. Let η be a Borel measure on $\mathbb{R}^D$, of total mass $\mu(\mathbb{R}^D) \leq 1$. Let A be a subset of $\mathbb{R}^D$. Assume*

(i) $\mathrm{supp}(\eta) \subset B(0, \delta^{-\alpha})$,
(ii) for all $\xi \in \mathbb{R}^D$ with $\delta^{-\alpha} \leq \|\xi\| \leq \delta^{-1-\alpha}$, $|\hat{\eta}(\xi)| \leq \|\xi\|^{-C_1}$,

(iii) $\eta(A) \geq t$.

Then there exists $a \in \mathbb{R}^D$ *such that*

$$\mathcal{N}(A \cap B(a, \delta^\beta), \delta) \geq c t^{D+1} \Big(\frac{\delta^\beta}{\delta}\Big)^D,$$

where $\beta = (2D+1)\alpha$.

We are ready to prove the main proposition of this section.

Proof of Proposition 5.1. We shall use Lemma 5.4 with $D = \dim E'$ and

$$\alpha = \frac{\min_{1\leq i\leq r} \lambda_1(\mu, V_i)}{3(2D+1)\max_i \lambda_1(\mu, V_i)} \quad \text{so} \quad \beta = \frac{\min_{1\leq i\leq r} \lambda_1(\mu, V_i)}{3\max_{1\leq i\leq r} \lambda_1(\mu, V_i)}.$$

Let $C_1 = C_1(D, \alpha)$ be as in the lemma. Let $\alpha_0 = \alpha_0(\mu)$, $\alpha_1 = \frac{\alpha\alpha_0}{2}$ and $c_0 = c_0(\mu, \alpha_1)$ be as in Theorem 4.2. Set also $k_0 = \lceil \frac{\alpha_0 C_1}{c_0} \rceil$ and $k = Tk_0$, where $T = \#F$.

Assume

$$|\widehat{\mu^{*n} * \nu}(a_0)| \geq t.$$

By Lemma 5.3, there is a subset $A \subset \mathcal{M}_d(\mathbb{Z})$ such that

$$\forall g \in A, \quad |\hat{\nu}(a_0 g)| \geq \frac{t^{2k}}{2}$$

and

$$\big((\mu^{*n})^{\boxplus k} \boxminus (\mu^{*n})^{\boxplus k}\big)(A) \geq \frac{t^{2k}}{2}.$$

Note that μ^{*n} is supported on $\bigcup_{\gamma\in F} \gamma E$. We can cover μ^{*n} by its restrictions $(\mu^{*n})_{|\gamma E}$ to each subspace γE. Thanks to the choice of k and the commutativity of additive convolutions, there exists $\gamma \in F$ such that

$$(\mu^{*n})_{|\gamma E}^{\boxplus k_0} \boxplus (\text{a probability measure})(A) \geq \frac{t^{2k}}{2T^{2k}}.$$

Hence for some $x_1 \in \mathcal{M}_d(\mathbb{Z})$, we have

$$(\mu^{*n})_{|\gamma E}^{\boxplus k_0}(x_1 + A) \geq \frac{t^{2k}}{2T^{2k}}.$$

Let η' be the pushforward of $(\mu^{*n})_{|\gamma E}^{\boxplus k_0}$ under the map $g \mapsto (\pi' \circ \varphi_n)(\gamma^{-1} g)$ and let

$$A' = (\pi' \circ \varphi_n)\big(E \cap \gamma^{-1}(x_1 + A)\big) \subset E'$$

so that

$$\eta'(A') \gg t^{2k}. \tag{5.6}$$

Lemma 5.4 will be used at scale $\delta = e^{-\frac{\alpha_0 n}{2}}$. By the definition of $\mu_{n,\gamma}$ in Section 4. We have $\eta' = \mu_{n,\gamma}^{\boxplus k_0}$. By Theorem 4.2, for all $\xi \in E'^*$ with $\delta^{-\alpha} = e^{\alpha_1 n} \leq \|\xi\| \leq e^{\alpha_0 n} = \delta^{-2}$, we have

$$|\widehat{\eta'}(\xi)| \leq e^{-k_0 c_0 n} \leq \|\xi\|^{-C_1}.$$

In view of the large deviation principle for μ^{*n}, we may replace η' by its restriction to $B(0, \delta^{-\alpha})$ while maintaining (5.6) and the conclusion of Theorem 4.2. Thus by Lemma 5.4 applied to η' and A', there exists $x_2 \in \mathrm{B}_{E'}(0, \delta^{-\alpha})$ such that

$$\mathcal{N}\big(A' \cap B(x_2, \delta^\beta), \delta\big) \gg t^{O(1)} \delta^{-D(1-\beta)}. \tag{5.7}$$

Now define convex bodies in E by

$$C_0 = \varphi_{-n}\big(\mathrm{B}_{E'}(0, \delta^\beta) \times \mathrm{B}_{E_0}(0, R)\big) \quad \text{and} \quad B_0 = \varphi_{-n}\big(\mathrm{B}_{E'}(0, \delta) \times \mathrm{B}_{E_0}(0, R)\big)$$

where $R = O_\mu(k)$ is a constant large enough so that $\gamma^{-1}A \subset E' \times \mathrm{B}_{E_0}(0,R)$. Then the inequality (5.7) implies that for some x_3 in E,

$$\mathcal{N}(\gamma^{-1}A \cap (C_0 + x_3), B_0) \gg t^{O(1)}\delta^{-D(1-\beta)} \asymp t^{O(1)} \cdot \frac{|C_0|}{|B_0|}.$$

Indeed, with

$$\begin{array}{rccc} f_1\colon & E & \to & E' \\ & x & \mapsto & \pi' \circ \varphi_n(\gamma^{-1}x_1 + x) \end{array}$$

one has $f_1(\gamma^{-1}A) \supset A'$, $\pi' \circ \varphi_n(B_0) = \mathrm{B}_{E'}(0,\delta)$, and taking $x_3 \in E'$ such that $\pi' \circ \varphi_n(\gamma^{-1}x_1 + x_3) = x_2$, $f_1(C_0 + x_3) = \mathrm{B}_{E'}(x_2, \delta^\beta)$. The choice of R guarantees that $f_1\big(\gamma^{-1}A \cap (C_0 + x_3)\big) = f_1(\gamma^{-1}A) \cap f_1(C_0 + x_3)$. One concludes using the inequality (5.1) on f_1.

Now let

$$C = a_0\gamma C_0 + \mathrm{B}_{\mathbb{R}^d}(0,2) \quad \text{and} \quad B = a_0\gamma B_0 + \mathrm{B}_{\mathbb{R}^d}(0,2)$$

and apply (5.2) to $f\colon x \mapsto a_0\gamma x$ to obtain, with $c_0 = a_0\gamma x_3$,

$$\frac{\mathcal{N}\big(a_0A \cap (C + c_0), B\big)}{\mathcal{N}(C,B)} \gg \frac{\mathcal{N}\big(\gamma^{-1}A \cap (C_0 + x_3), B_0\big)}{\mathcal{N}(C_0,B_0)} \gg t^{O(1)}$$

whence

$$\mathcal{N}\big(a_0A \cap (C + c_0), B\big) \geq t^{O(1)}\frac{|C|}{|B|}.$$

Since $\hat\nu(a_0g) \geq t^{O(1)}$ for every $g \in A$, Proposition 5.2 shows that there exists a B^*-separated subset $X \subset \mathbb{T}^d$ such that

$$\nu(X + C^*) \geq t^{O(1)}.$$

To conclude, it remains to describe the sets B^* and C^*. For that, first consider a decomposition of the space of linear forms on $\mathbb{R}^d$ into irreducible components under the right action of G°

$$(\mathbb{R}^d)^* = V' = V^{(1)} \oplus \cdots \oplus V^{(r)}$$

and write p_i, $i = 1, \dots, k$ for the corresponding projections. Since $a_0\gamma$ is an integer vector, and each $V^{(i)}$ is defined over a number field, by Liouville's Theorem, there exists a constant $C > 0$ such that for each i such that $p_i(a_0\gamma) \neq 0$, one has

$$\|a_0\|^{-C} \ll \|p_i(a_0\gamma)\| \ll \|a_0\|.$$

Therefore, for any $\varepsilon > 0$, we may choose $C_\varepsilon \geq 0$ such that $n \geq C_\varepsilon \log \frac{\|a_0\|}{t}$ implies, for $i = 1, \dots, k$,

$$p_i(a_0\gamma) = 0 \quad \text{or} \quad e^{-\varepsilon n} \leq \|p_i(a_0\gamma)\| \leq e^{\varepsilon n}.$$

Thus, if $p_i(a_0\gamma) \neq 0$ and $\lambda_1(\mu, V^{(i)}) \neq 0$, then

$$p_i(a_0\gamma)B_0 \subset \mathrm{B}_{V^{(i)}}(0, e^{(\lambda_1(\mu,V^{(i)})+\varepsilon)n}\delta)$$

and

$$p_i(a_0\gamma)C_0 \supset \mathrm{B}_{V^{(i)}}(0, e^{(\lambda_1(\mu,V^{(i)})-\varepsilon)n}\delta^\beta).$$

Now consider the decomposition of $V = \mathbb{R}^d$ according to Lyapunov exponents

$$V = V_0 \oplus V_1 \oplus \cdots \oplus V_r,$$

where for $i = 0, \dots, r$, V_i is the sum of all irreducible G-submodules of V with Lyapunov exponent $\lambda_1(\mu, E_i)$. Since $W = (a_0\gamma E)^\perp$ is a submodule, it can be written

$$W = W_0 \oplus \cdots \oplus W_r, \quad \text{where} \quad W_i = W \cap V_i, \quad i = 0, \dots, r.$$

An elementary computation based on the above observations shows that for some compact subset $A_0 \subset V_0$ containing $\mathrm{B}_{V_0}(0, \frac{1}{2})$, (we identify subsets of $\mathrm{B}_V(0, \frac{1}{2})$ with their projections in $\mathbb{T}^d$)

$$B^* \supset A_0 \times \prod_{1 \leq i \leq r} \{\, v_i \in V_i \,|\, d(v_i, W_i) \leq e^{-(\lambda_1(\mu, V_i)+\varepsilon)n} \delta^{-1} \text{ and } \|v_i\| \leq \frac{1}{2} \,\}$$

and

$$C^* \subset A_0 \times \prod_{1 \leq i \leq r} \{\, v_i \in V_i \,|\, d(v_i, W_i) \leq e^{-(\lambda_1(\mu, V_i)-\varepsilon)n} \delta^{-\beta} \text{ and } \|v_i\| \leq \frac{1}{2} \,\}.$$

Recalling $\delta = e^{-\frac{\alpha_0 n}{2}}$ and setting $\tau = \frac{\alpha_0}{5 \max_i \lambda_1(\mu, V_i)} > 0$, we may choose $\varepsilon > 0$ small enough so that

$$e^{-(\lambda_1(\mu, V_i)+\varepsilon)n} \delta^{-1} = e^{-(\lambda_1(\mu, V_i)+\varepsilon-\frac{\alpha_0}{2})n} \geq e^{-(1-2\tau)\lambda_1(\mu, V_i)n}$$

and

$$e^{-(\lambda_1(\mu, V_i)-\varepsilon)n} \delta^{-\beta} = e^{-(\lambda_1(\mu, V_i)-\varepsilon-\frac{\beta\alpha_0}{2})n} \leq e^{-(1-\tau)\lambda_1(\mu, V_i)n}.$$

Finally, since A_0 can be covered by a bounded number of translates of $\mathrm{B}_{V_0}(0, \frac{1}{2})$, we may assume replace A_0 by $\mathrm{B}_{V_0}(0, \frac{1}{2})$, and then $B^* \supset \tilde{\mathrm{N}}\mathrm{bd}(W \bmod \mathbb{Z}^d, e^{-(1-2\tau)n})$ while $C^* \subset \tilde{\mathrm{N}}\mathrm{bd}(W \bmod \mathbb{Z}^d, e^{-(1-\tau)n})$. □

For a technical reason, instead of on the torus, we will have to work on a union a tori $\mathbb{T}^d \times F$, where Γ acts diagonally. For a measure ν on $\mathbb{T}^d \times F$ and $a \in \mathbb{Z}^d$, we write $\widehat{\mu^{*n} * \nu}(a, 1)$ for the Fourier coefficient at frequency a of the restriction of ν to $\mathbb{T}^d \times \{1\}$ viewed as a measure on $\mathbb{T}^d$.

A more careful look at the proof will give us the following slightly more precise version of Proposition 5.1.

Corollary 5.5. *Let μ be a probability measure on $\mathrm{GL}_d(\mathbb{Z})$ Assume the algebraic group G generated by μ is semisimple and write $F = G/G^\circ$. Let ν be a Borel probability measure on $\mathbb{T}^d \times F$. Then there exist $C = C(\mu) \geq 0$ and $\tau > 0$ such that the following holds.*

Assume that for some $t \in (0, \frac{1}{2})$,

$$\big|\widehat{\mu^{*n} * \nu}(a_0, 1)\big| \geq t \quad \text{for some } a_0 \in \mathbb{Z}^d \text{ and } n \geq C \log \frac{\|a_0\|}{t}.$$

Then, there exists $\gamma \in F$ such that, denoting

$$W = (a_0 \gamma E)^\perp$$

there exists a finite subset $X \subset \mathbb{T}^d \times \{\gamma^{-1}\}$ such that

$$(X - X) \cap \tilde{\mathrm{N}}\mathrm{bd}(W \bmod \mathbb{Z}^d, e^{-(1-2\tau)n}) = \{0\}$$

and

$$\nu\big(X + \tilde{\mathrm{N}}\mathrm{bd}(W \bmod \mathbb{Z}^d, e^{-(1-\tau)n})\big) \geq t^{O(1)}.$$

Here, of course, the addition on $\mathbb{T}^d \times \{\gamma^{-1}\}$ is defined for the torus coordinate.

6. Concentration and unstability of the random walk

In this section, we finally prove the main result of the paper. We consider a probability measure μ on $\mathrm{GL}_d(\mathbb{Z})$ and the associated random walk on the torus $\mathbb{T}^d$, starting from a point $x_0 \in \mathbb{T}^d$. Letting Γ be the group generated by $\operatorname{supp} \mu$ and G the Zariski closure of Γ, we assusme that G is semisimple as an algebraic group, and we show — in a quantitative way — that if the law $\mu^{*n} * \delta_{x_0}$ of the random walk is not exponentially close to the Haar measure on $\mathbb{T}^d$, then the starting point x_0 is exponentially close to a proper closed invariant subset.

But let us first introduce a new space on which it is convenient to study the random walk, especially to overcome issues related to being Zariski disconnected. As before, G° denotes the identity component of G, and $F = G/G^\circ$. The subalgebra of $\mathcal{M}_d(\mathbb{R})$ generated by G° is denoted by E. In order to keep track of the coset modulo G°, we let

$$Y_0 = \mathbb{T}^d \times F$$

and let Γ act on Y_0 diagonally.

Let W_0 be a rational G°-invariant subspace of $V = \mathbb{R}^d$. We can define a factor of $Y_0 \to Y$ by

$$Y = \bigsqcup_{\gamma \in F} \mathbb{R}^d/(\gamma W_0 + \mathbb{Z}^d) \times \{\gamma\}.$$

The action of Γ on Y is defined in the obvious way. This way, the natural projection $Y_0 \to Y$ is Γ-equivariant.

Given $a_0 \in \mathbb{Z}^d$ for which the random walk $\mu^{*n} * \delta_{x_0}$ has large Fourier at a_0. We will set

$$W_0 = V^\circ + (a_0 E)^\perp$$

and study the random walk on Y associated to this W_0, because information given by the assumption is all contained in this factor.

In the introduction, we defined the quasi-distance adapted to the random walk on V and on $\mathbb{T}^d$. Similarly, we can define a quasi-distance on each of the tori $\mathbb{R}^d/(\mathbb{Z}^d + \gamma W_0)$ in Y. Together, theses quasi-distances define a quasi-distance on Y by the formula

$$\tilde{d}\big((x_1, \gamma_1), (x_2, \gamma_2)\big) = \begin{cases} \tilde{d}(x_1, x_2) & \text{if } \gamma_1 = \gamma_2, \\ +\infty & \text{otherwise.} \end{cases}$$

Below we will state a slightly more precise version of Theorem 1.1. Let V_0 denote the sum of all compact factors of G in $\mathbb{R}^d$, that is, the sum of irreducible subrepresentations $W \subset V$ such that $\lambda_1(\mu, W) = 0$. We fix a G-invariant Euclidean norm on V_0 and write $\mathrm{B}_{V_0}(0, R)$ for the closed ball in V_0 with radius $R > 0$ with respect to this norm. Given some parameter $Q > 0$, we note that

$$\mathrm{B}_{V_0}(0, Q) + \bigcup_{q \le Q} \frac{1}{q}\mathbb{Z}^d \subset \mathbb{R}^d$$

is Γ-invariant. As a consequence, the set

$$Z_Q = \bigsqcup_{\gamma \in F} \big(\mathrm{B}_{V_0}(0, Q) + \bigcup_{q \le Q} \frac{1}{q}\mathbb{Z}^d \bmod \gamma W_0 + \mathbb{Z}^d\big) \times \{\gamma\}$$

is a Γ-invariant closed subset of Y.

Theorem 6.1. *Assume that μ has a finite exponential moment and the algebraic group G is semisimple. Then for every $\lambda \in (0, 1)$, there exist $C = C(\mu, \lambda) \ge 0$ such that the following holds.*

Given $a_0 \in \mathbb{Z}^d$, let Y be defined as above. For any $x_0 \in \mathbb{T}^d$, if

$$\big|\widehat{(\mu^{*n} * \delta_{x_0})}(a_0)\big| \ge t \quad \text{for some } t \in (0, \frac{1}{2}) \text{ and } n \ge C \log \frac{\|a_0\|}{t}, \tag{6.1}$$

then there is $\gamma_0 \in F$ such that writing $y_0 = (x_0 \mod \gamma_0 W_0, \gamma_0) \in Y$, we have

$$\tilde{d}(y_0, Z_Q) \le e^{-\lambda n} \quad \text{for some } Q \le \left(\frac{\|a_0\|}{t}\right)^C.$$

To obtain Theorem 1.1 from this theorem, it suffices lift y_0 and Z_Q to Y_0 and then project to $\mathbb{T}^d$.

[W.H.] If we want γ_0 to be the neutral element, Y needs to be modified/shifted.

We proceed to the proof of Theorem 6.1. We fix the meaning of $a_0 \in \mathbb{Z}^d$, $W_0 \subset V$, $x_0 \in \mathbb{T}^d$ as in the statement. By the pigeonhole principle, (6.1) implies that there is $\gamma_0 \in F$ such that

$$\left|\widehat{\mu^{*n} * \delta_{(x_0,\gamma_0)}}(a_0, 1)\right| \geq \frac{t}{\#F}, \tag{6.2}$$

where the notation $\widehat{\mu^{*n} * \delta_{(x_0,\gamma_0)}}(a_0, 1)$ is defined in the paragraph preceding Corollary 5.5. This choice of γ_0 determines $y_0 \in Y$. We fix this y_0 for the rest of the proof.

6.1. **Bootstrapping concentration.** In order to prove Theorem 6.1, we start from the granulation estimate obtained in the previous section as Proposition 5.1. The first step is then to run backwards the random walk to increase the concentration.

Proposition 6.2 (High concentration)**.** *Assume* (6.1). *Given $\eta > 0$, there exists $n_1 \asymp_\eta \log \frac{\|a_0\|}{t}$ and $\rho > 0$ with $|\log \rho| \asymp n_1$ such that for some $y \in Y$,*

$$\mu^{*(n-n_1)} * \delta_{y_0}(\tilde{\mathrm{B}}(y, \rho)) \geq \rho^\eta.$$

Proof. Using (6.2) and Corollary 5.5 and observing that $(a_0\gamma E)^\perp = \gamma^{-1}(a_0 E)^\perp$, we obtain that for $n_0 \geq \log \frac{\|a_0\|}{t}$, there exists an $e^{-(1-2\tau)n_0}$-separated subset X_0 contained in a single torus in Y such that

$$\mu^{*(n-n_0)} * \delta_{y_0}\left(\tilde{\mathrm{N}}\mathrm{bd}(X_0, e^{-(1-\tau)n_0})\right) \geq t^{C_0}.$$

Increasing C_0 if necessary, we may also assume that $\#X_0 \leq e^{C_0 n_0}$. Fix some large $k \in \mathbb{N}$ and then $\varepsilon > 0$ such that $2k\varepsilon < 1$. Starting with

$$m_0 = \lfloor \frac{\tau n_0}{2d} \rfloor, \quad r_0 = e^{-(1-2\tau)n_0}, \quad \text{and} \quad \rho_0 = e^{-(1-\tau)n_0},$$

we apply Lemma 6.3 below k times successively. This yields integers m_i, and scales $r_i > \rho_i$, defined inductively by

$$\begin{cases} r_{i+1} = e^{-m_i(1+\varepsilon)} r_i \\ \rho_{i+1} = e^{-m_i(1-\varepsilon)} \rho_i \\ m_{i+1} = \lfloor m_i(1 - \frac{\varepsilon}{d}) \rfloor \end{cases}$$

and at each step, an r_i-separated set X_i such that $\#X_i \leq \#X_0$ and

$$\mu^{*(n-n_0-m_0-\cdots-m_i)} * \delta_{y_0}\big(\tilde{\mathrm{N}}\mathrm{bd}(X_i, \rho_i)\big) \geq \left(\frac{t^{C_0}}{2}\right)^{d^i}.$$

Set $n_1 = n_0 + m_0 + \cdots + m_k$ and $\rho = \rho_k$. One has $\#X_k \leq \#X_0$ and

$$\mu^{*(n-n_1)} * \delta_{y_0}\big(\tilde{\mathrm{N}}\mathrm{bd}(X_k, \rho)\big) \geq \left(\frac{t^{C_0}}{2}\right)^{d^k}$$

so that for some $y \in X_k$,

$$\mu^{*(n-n_1)} * \delta_{y_0}\left(\tilde{\mathrm{B}}(y, \rho)\right) \geq \frac{1}{\#X_0}\left(\frac{t^{C_0}}{2}\right)^{d^k} \geq e^{-C_0 n_0}\left(\frac{t^{C_0}}{2}\right)^{d^k}$$

Now, since $m_0 + \cdots + m_k \geq \frac{k m_0}{3}$ we may choose k large enough so that $m_0 + \cdots + m_k \geq \frac{3C_0 n_0}{\eta}$, and then $n_0 \asymp \log \frac{\|a_0\|}{t}$ large enough to ensure that

$$\rho = \rho_k \leq e^{-(1-\varepsilon)(m_0+\cdots+m_k)} \leq e^{-\frac{2C_0 n_0}{\eta}} \leq e^{-\frac{C_0 n_0}{\eta}}\left(\frac{t^{C_0}}{2}\right)^{\frac{d^k}{\eta}}.$$

The proposition follows. □

After using Corollary 5.5, we can now forget how W_0 is constructed from a_0. All what we need is that W_0 is a G°-invariant rational subspace containing V°.

We now prove the lemma that was used in the above proof. The notion of r-seperated sets in Y are with respect to the quasi-distance on Y.

Lemma 6.3. *For any $\varepsilon > 0$ there exist $c > 0$ and $m_0 \in \mathbb{N}$ depending only on μ and ε such that the following holds for any Borel probability measure ν on Y and any $m \geq m_0$.*
Let $r > 0$ and $\rho > 0$ be such that $e^{(d+1)m}\rho < r$. Set

$$r_1 = e^{-m(1+\varepsilon)}r \quad \text{and} \quad \rho_1 = e^{-m(1-\varepsilon)}\rho.$$

If X is an r-separated subset contained in a single torus in Y, then there is an r_1-separated subset $X_1 \subset Y$, contained in a single torus, with cardinality $\#X_1 \leq \#X$ and such that

$$\nu\big(\tilde{\mathrm{N}}\mathrm{bd}(X_1, \rho_1)\big) \geq (\mu^{*m} * \nu)\big(\tilde{\mathrm{N}}\mathrm{bd}(X, \rho)\big)^d - e^{-cm}.$$

Proof. In this proof, we write $X^{(\rho)}$ for $\tilde{\mathrm{N}}\mathrm{bd}(X, \rho)$. By Jensen's inequality and the definition of $\mu^{*m} * \nu$, (see [10, Lemma 7.6] for details),

$$(\mu^{*m} * \nu)(X^{(\rho)})^d \leq \sum_{g_1,\dots,g_d \in \Gamma} \mu^{*m}(g_1) \dots \mu^{*m}(g_d)\nu(g_1^{-1}X^{(\rho)} \cap \dots \cap g_d^{-1}X^{(\rho)}).$$

This implies that the set of d-tuples $(g_i)_{1\leq i\leq d}$ such that

$$\nu(g_1^{-1}X^{(\rho)} \cap \dots \cap g_d^{-1}X^{(\rho)}) \geq (\mu^{*m} * \nu)(X^{(\rho)})^d - e^{-cm} \tag{6.3}$$

has $(\mu^{*m})^{\otimes d}$-measure at least e^{-cm}. Using the fact that the large deviation estimates Theorem 3.11(i) and (ii) are valid under the only assumption that the action is irreducible, one readily checks that [22, Lemma 5.5] and its proof are also valid under this assumption. Applying this lemma in each irreducible subrepresentation of $\mathbb{R}^d$, one obtains that if c is chosen small enough, there must exist $g_1, \dots, g_d \in \Gamma$ satisfying (6.3) and moreover,

$$\forall v \in \mathbb{R}^d/W_0 \setminus \{0\}, \quad e^{(1-\varepsilon)m} \leq \max_{1\leq i\leq d} \frac{|g_i v|^\sim}{|v|^\sim} \tag{6.4}$$

and (using the large deviation estimates again and the fact that $-\lambda_{\dim W}(\mu, W) < (\dim W - 1)\lambda_1(\mu, W)$ for any G-invariant $W \subset V$)

$$\forall i \in \{1, \dots, d\}, \quad |g_i|^\sim \leq e^{(1+\varepsilon)m} \quad \text{and} \quad |g_i^{-1}|^\sim \leq e^{(d-1+\varepsilon)m}. \tag{6.5}$$

We fix such elements $g_1, \dots, g_d$ for the rest of the proof.

Since X is contained in a single torus, all the g_i's are contained in the same class in G/G°.

We claim that the set $g_1^{-1}X^{(\rho)} \cap \dots \cap g_d^{-1}X^{(\rho)}$ is included in a union of at most $\#X$ balls of radius $\rho_1 = e^{-(1-\varepsilon)m}\rho$. Indeed, from (6.5) and the fact that $e^{(d+1)m}\rho < r$, we find, for a given $x \in X$ and $i \geq 1$, that the set $g_1^{-1}\tilde{\mathrm{B}}(x, \rho)$ meets at most one component $g_i^{-1}\tilde{\mathrm{B}}(y, \rho)$, $y \in X$. Therefore, there are at most $\#X$ non-empty intersections $g_1^{-1}\tilde{\mathrm{B}}(x_1, \rho) \cap \dots \cap g_d^{-1}\tilde{\mathrm{B}}(x_d, \rho)$, for $x_1, \dots, x_d \in X$.

If x, y lie inside such an intersection, then, for each i, $\tilde{d}(g_i x, g_i y) \leq \rho$. Then (6.4) and (6.5) implies that $\tilde{d}(x, y) \leq e^{-(1-\varepsilon)m}\rho = \rho_1$. Thus, each intersection $g_1^{-1}\tilde{\mathrm{B}}(x_1, \rho) \cap \dots \cap g_d^{-1}\tilde{\mathrm{B}}(x_d, \rho)$ is included in a ball of radius ρ_1.

Finally, using (6.5), we see that these ρ_1-balls are separated by at least $r_1 = e^{-(1+\varepsilon)m}r$. Moreover they are contained in the same torus in Y because the g_i's are in the same G° coset. This finishes the proof of the proposition. □

6.2. **A diophantine property.** From the high concentration property obtained in the previous paragraph, we want to infer that $\mu^{*(n-n_1)} * \delta_{y_0}$ is concentrated near a proper Γ-invariant subset. The argument relies on a diophantine property of the random walk, coming from the fact that μ is supported on $\mathrm{GL}_d(\mathbb{Z})$.

Proposition 6.4 (Concentration near a closed invariant subset)**.** *Given $\beta > 0$, there exists $C > 0$ such that the following holds.*

Assume (6.1)*. Then there exist $n_1 \in \mathbb{N}^*$ such that $\frac{1}{C}\log\frac{\|a_0\|}{t} \leq n_1 \leq C\log\frac{\|a_0\|}{t}$ and $\rho \in [e^{-Cn_1}, e^{-\frac{n_1}{C}}]$ and $Q \leq \rho^{-\beta}$ such that*

$$\mu^{*(n-n_1)} * \delta_{y_0}\big(\tilde{\mathrm{N}}\mathrm{bd}(Z_Q, \rho)\big) \geq \rho^{\beta}.$$

This proposition is an immediate consequence of Proposition 6.2 and of a diophantine property of the random walk, given by the following lemma.

Lemma 6.5 (Diophantine property)**.** *For every $\beta > 0$, there exist constants C and $\eta > 0$ depending on μ and β such that for any $y_0, y \in Y$, if for $n \geq C|\log\rho|$, one has*

$$(\mu^{*n} * \delta_{y_0})(\tilde{\mathrm{B}}(y, \rho)) \geq \rho^{\eta}$$

then $\tilde{d}(y, Z_Q) \leq \rho^{1-\beta}$ for some $Q \leq \rho^{-\beta}$.

Proof. Consider the action of G on $V_F = \bigsqcup_{\gamma\in F} V/\gamma W_0 \times \{\gamma\}$. Since V/W_0 contains no G°-invariant vector, for every non-zero (v_1, v_2) in $V_F \times V_F$, the set $\{\, g \,|\, gv_1 = v_2 \,\}$ is a linear subvariety in G of dimension less than $\dim G$. Using the spectral gap property modulo prime numbers [31] (the same way it is used in the first step of the proof of Proposition 3.3), we obtain that for m large enough, for some c_0 independent of (v_1, v_2),

$$\mu^{\otimes m}(\{\, (g_1, \dots, g_m) \,|\, g_m \cdots g_1 v_1 = v_2 \,\}) \leq e^{-c_0 m}. \tag{6.6}$$

Fix m such that

$$e^{-\frac{c_0 m}{2}} \geq \rho^{\eta} > e^{-c_0 m}.$$

If C is large enough, the condition $n \geq C|\log\rho|$ ensures that $n \geq m$. From the assumed inequality, it follows that there exists some $y_1 \in Y$ such that

$$(\mu^{*m} * \delta_{y_1})(\tilde{\mathrm{B}}(y, \rho)) \geq \rho^{\eta},$$

which implies that the set

$$A_m = \{\, (g_1, \dots, g_m) \in (\operatorname{supp}\mu)^m \,|\, \tilde{d}(g_m \dots g_1 y_1, y) \leq \rho \,\}$$

satisfies

$$\mu^{\otimes m}(A_m) \geq \rho^{\eta} > e^{-c_0 m}.$$

Using the finite exponential moment of μ and reducing the set A_m, we can assume further that for all $(g_i) \in A_m$, $\|g_m \cdots g_1\| \leq e^{C_0 m}$ for some $C_0 = C_0(\mu)$.

Write $y_1 = (x_1, \gamma_1)$ and $y = (x, \gamma)$. Recalling (6.6) above, one infers that the linear map

$$\begin{array}{rccc} \theta\colon & V/\gamma_1 W_0 \times V/\gamma W_0 & \to & (V/\gamma W_0)^{A_m} \\ & (v_1, v_2) & \mapsto & (g_n \dots g_1 v_1 - v_2)_{(g_i)\in A_m} \end{array}$$

is injective. Moreover, in the canonical bases, its matrix has coefficients in $\mathbb{Z}$ bounded by $e^{C_0 m}$, so its inverse has coefficients in $\frac{1}{Q}\mathbb{Z}$ for some $Q \leq e^{C_1 m}$, and bounded by $e^{C_1 m}$. Therefore, any solution (v_1, v_2) in $V/\gamma_1 W_0 \times V/\gamma W_0$ to

$$\forall (g_1, \dots, g_m) \in A_m, \quad g_m \dots g_1 v_1 - v_2 \in \tilde{\mathrm{B}}(0, \rho) + \mathbb{Z}^d \bmod \gamma W_0$$

can be written, for some $w_1, w_2 \in \mathbb{Z}^d$ and u_1, u_2 in $\tilde{\mathrm{B}}(0, e^{C_1 m}\rho)$,

$$v_1 = \frac{1}{Q} w_1 + u_1 \bmod \gamma_1 W_0 \quad \text{and} \quad v_2 = \frac{1}{Q} w_2 + u_2 \bmod \gamma W_0.$$

This applies in particular to representatives of (x_1, x) in $V/\gamma_1 W_0 \times V/\gamma W_0$. It follows that

$$\tilde{d}(y, Z_{e^{C_1 m}}) \leq e^{C_1 m}\rho.$$

If $\eta > 0$ is chosen so small that $\frac{2C_1}{c_0}\eta < \beta$, one has

$$e^{C_1 m} = e^{\frac{2C_1}{c_0}\frac{c_0 m}{2}} \leq \rho^{-\frac{2C_1}{c_0}\eta} \leq \rho^{-\beta}$$

so the lemma is proved. □

6.3. **Unstability of closed invariant subsets.** To conclude the proof of Theorem 6.1, we use a variant of the argument given in [24, §3]. It is based on Foster's exponential recurrence criterion, applied to a well-chosen function associated to a closed invariant subset. This technique has been used extensively in homogeneous dynamics since the work of Eskin and Margulis [17], in particular by Benoist and Quint for their study of stationary measures [2, 4, 5, 3].

Lemma 6.6 (Margulis inequality)**.** *For every $\lambda \in (0,1)$, there exist constants $C, \alpha > 0$ depending only on μ such that the following holds. For $Q \geq 2$, define a function $\varphi_Q \colon Y \to \mathbb{R} \cup \{+\infty\}$ by*

$$\varphi_Q(y) = \begin{cases} \tilde{d}(y, Z_Q)^{-\alpha} & \text{if } \tilde{d}(y, Z_Q) > 0 \\ +\infty & \text{otherwise.} \end{cases}$$

For all $y \in Y$ and all integers $n \geq 1$,

$$\int \varphi_Q(gy)\mathrm{d}\mu^{*n}(g) \leq e^{-\lambda\alpha n}\varphi_Q(y) + Q^C.$$

The proof of such inequalities is an application of Furstenberg's law of large numbers [6, Theorem 4.28], using also the exponential moment assumption on μ. Since it is rather standard, we leave it to the reader, and turn to the proof of Theorem 6.1.

Proof of Theorem 6.1. Let $C, \alpha > 0$ be the parameters given by Lemma 6.6 applied with $\lambda' = \frac{1+\lambda}{2}$ instead of λ. Then set $\beta = \frac{\alpha}{C+2}$.

Proposition 6.4 shows that for some $n_1 \asymp_\beta \log \frac{\|a_0\|}{t}$ and some $\rho \in [e^{-C_1 n_1}, e^{-c_1 n_1}]$, there exist $Q \leq \rho^{-\beta}$ such that

$$\mu^{*(n-n_1)} * \delta_{y_0}\big(\tilde{\mathrm{N}}\mathrm{bd}(Z_Q, \rho)\big) \geq \rho^\beta.$$

Applying Lemma 6.6 yields

$$\begin{aligned} \rho^{-\alpha+\beta} &\leq \int \varphi_Q(gy_0)\mathrm{d}\mu^{*(n-n_1)}(g) \\ &\leq e^{-\lambda'\alpha(n-n_1)}\varphi_Q(y_0) + Q^C. \end{aligned}$$

Note that $Q^C \leq \rho^{-C\beta} \leq \frac{1}{2}\rho^{-\alpha+\beta}$ and therefore

$$\varphi_Q(y_0) = \tilde{d}(y_0, Z_Q)^{-\alpha} \gg e^{\lambda'\alpha(n-n_1)}\rho^{-\alpha+\beta} \gg e^{\lambda'\alpha n}e^{-C_1\alpha(1+\frac{1}{C+2})n_1}.$$

Since $n_1 \asymp \log \frac{\|a_0\|}{t}$ and $\lambda' = \lambda + \frac{1-\lambda}{2}$, taking $n \gg \frac{1}{1-\lambda}\log \frac{\|a_0\|}{t}$ yields

$$\tilde{d}(y_0, Z_Q) \leq e^{-\lambda n}$$

and the theorem is proved. □

References

[1] R. Aoun. Random subgroups of linear groups are free. *Duke Math. J.*, 160(1):117–173, 2011.

[2] Y. Benoist and J.-F. Quint. Mesures stationnaires et fermés invariants des espaces homogènes. *Ann. of Math. (2)*, 174(2):1111–1162, 2011.

[3] Y. Benoist and J.-F. Quint. Random walks on finite volume homogeneous spaces. *Invent. Math.*, 187(1):37–59, 2012.

[4] Y. Benoist and J.-F. Quint. Stationary measures and invariant subsets of homogeneous spaces (II). *J. Amer. Math. Soc.*, 26(3):659–734, 2013.

[5] Y. Benoist and J.-F. Quint. Stationary measures and invariant subsets of homogeneous spaces (III). *Ann. of Math. (2)*, 178(3):1017–1059, 2013.

[6] Y. Benoist and J.-F. Quint. *Random walks on reductive groups*, volume 62 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics*. Springer, Cham, 2016.

[7] P. Bougerol and J. Lacroix. *Products of random matrices with applications to Schrödinger operators*, volume 8 of *Progress in Probability and Statistics*. Birkhäuser Boston, Inc., Boston, MA, 1985.

[8] J. Bourgain. On the Erdős-Volkmann and Katz-Tao ring conjectures. *Geom. Funct. Anal.*, 13(2):334–365, 2003.

[9] J. Bourgain. The discretized sum-product and projection theorems. *J. Anal. Math.*, 112:193–236, 2010.

[10] J. Bourgain, A. Furman, E. Lindenstrauss, and S. Mozes. Stationary measures and equidistribution for orbits of nonabelian semigroups on the torus. *J. Amer. Math. Soc.*, 24(1):231–280, 2011.

[11] J. Bourgain and A. Gamburd. On the spectral gap for finitely-generated subgroups of SU(2). *Invent. Math.*, 171(1):83–121, 2008.

[12] J. Bourgain and A. Gamburd. Uniform expansion bounds for Cayley graphs of $\mathrm{SL}_2(\mathbb{F}_p)$. *Ann. of Math. (2)*, 167(2):625–642, 2008.

[13] J. Bourgain, A. A. Glibichuk, and S. V. Konyagin. Estimates for the number of sums and products and for exponential sums in fields of prime order. *J. Lond. Math. Soc., II. Ser.*, 73(2):380–398, 2006.

[14] J.-B. Boyer. On the affine random walk on the torus. *arXiv e-prints*, page arXiv:1702.08387, Feb 2017.

[15] T. Bénard. Equidistribution of mass for random processes on finite-volume spaces, 2021.

[16] P. Erdős and E. Szemerédi. On sums and products of integers. Studies in Pure Mathematics, Mem. of P. Turán, 213-218 (1983)., 1983.

[17] A. Eskin and G. Margulis. Recurrence properties of random walks on finite volume homogeneous manifolds. In *Random walks and geometry. Proceedings of a workshop at the Erwin Schrödinger Institute, Vienna, June 18 – July 13, 2001. In collaboration with Klaus Schmidt and Wolfgang Woess. Collected papers.*, pages 431–444. Berlin: de Gruyter, 2004.

[18] K. J. Falconer. Hausdorff dimension and the exceptional set of projections. *Mathematika*, 29:109–115, 1982.

[19] H. Furstenberg. Noncommuting random products. *Trans. Amer. Math. Soc.*, 108:377–428, 1963.

[20] Y. Guivarc'h and A. N. Starkov. Orbits of linear group actions, random walks on homogeneous spaces and toral automorphisms. *Ergodic Theory Dynam. Systems*, 24(3):767–802, 2004.

[21] W. He. Discretized sum-product estimates in matrix algebras. *J. Anal. Math.*, 139(2):637–676, 2019.

[22] W. He and N. de Saxcé. Linear random walks on the torus, 2019. preprint.

[23] W. He and N. de Saxcé. Sum-product for real Lie groups. *J. Eur. Math. Soc. (JEMS)*, 23(6):2127–2151, 2021.

[24] W. He, T. Lakrec, and E. Lindenstrauss. Affine Random Walks on the Torus. *International Mathematics Research Notices*, 01 2021. rnaa322.

[25] W. He, T. Lakrec, and E. Lindenstrauss. Equidistribution of affine random walks on some nilmanifolds, 2021. preprint.

[26] W. Kim. Effective equidistribution of expanding translates in the space of affine lattices, 2021. preprint.

[27] E. Le Page. Théorèmes limites pour les produits de matrices aléatoires. In *Probability measures on groups (Oberwolfach, 1981)*, volume 928 of *Lecture Notes in Math.*, pages 258–303. Springer, Berlin-New York, 1982.

[28] J. Li. Discretized sum-product and Fourier decay in $\mathbb{R}^n$. *J. Anal. Math.*, 143(2):763–800, 2021.

[29] E. Lindenstrauss and A. Mohammadi. Polynomial effective density in quotients of $\mathbb{H}^3$ and $\mathbb{H}^2 \times \mathbb{H}^2$, 2021.
[30] R. Muchnik. Semigroup actions on $\mathbb{T}^n$. *Geom. Dedicata*, 110:1–47, 2005.
[31] A. Salehi Golsefidy and P. P. Varjú. Expansion in perfect groups. *Geom. Funct. Anal.*, 22(6):1832–1891, 2012.
[32] T. Tao. Product set estimates for non-commutative groups. *Combinatorica*, 28(5):547–594, 2008.
[33] B. L. Van der Waerden. *Modern Algebra. Volume II. Based in part on lectures by E. Artin and E. Noether.* Frederick Ungar Publishing Co., New York, transl. from the 3rd german edition, 1950.

Email address: heweikun@kias.re.kr

CNRS – Université Paris 13, LAGA, 93430 Villetaneuse, France.
Email address: desaxce@math.univ-paris13.fr