

# HIGHER-ORDER CORRELATIONS FOR GROUP ACTIONS

ALEXANDER GORODNIK

ABSTRACT. This survey paper discusses behaviour of higher-order correlations for one-parameter dynamical systems and more generally for dynamical systems arising from group actions. In particular, we present a self-contained proof of quantitative bounds for higher-order correlations of actions of simple Lie groups. We also outline several applications of our analysis of correlations that include asymptotic formulas for counting lattice points, existence of approximate configurations in lattice subgroups, and validity of the Central Limit Theorem for multi-parameter group actions.

## CONTENTS

1. Introduction	1
2. Decay of matrix coefficients	5
3. Application: counting lattice points	9
4. Quantitative estimates on matrix coefficients	12
5. Bounds on higher-order correlations	22
6. Application: existence of configurations	28
7. Application: Central Limit Theorem	31
References	38

## 1. INTRODUCTION

In this survey, we will be interested in gaining an insight into asymptotic properties of chaotic group actions. There are several quite distinct points of view on how this problem may be studied. Our approach here is based on the analysis of higher-order correlations which characterise random-like behaviour of observables computed along orbits. For instance, let us consider a measure-preserving transformation  $T : X \rightarrow X$  of a probability space  $(X, \mu)$ . Then for functions  $\phi_1, \dots, \phi_r \in L^\infty(X)$ , the *correlations of order  $r$*  are defined as

$$\int_X \phi_1(T^{n_1}x) \cdots \phi_r(T^{n_r}x) d\mu(x), \quad n_1, \dots, n_r \in \mathbb{N}. \quad (1.1)$$

The transformation  $T$  is called *mixing of order  $r$*  if for all  $\phi_1, \dots, \phi_r \in L^\infty(X)$ ,

$$\int_X \phi_1(T^{n_1}x) \cdots \phi_r(T^{n_r}x) d\mu(x) \longrightarrow \left( \int_X \phi_1 d\mu \right) \cdots \left( \int_X \phi_r d\mu \right) \quad (1.2)$$

as  $|n_i - n_j| \rightarrow \infty$  for all  $i \neq j$ . The multiple mixing property (i.e., mixing of all orders), in particular, implies that the family of functions  $\{\phi \circ T^n\}$  is quasi-independent asymptotically. The study of this property was initiated by Rokhlin [79] who showed that ergodic endomorphisms of compact abelian groups are mixing of all orders. In this work, Rokhlin also raised the question, which still remains

open, whether mixing of order two implies mixing of all orders for general measure-preserving transformations. Kalikow [50] established this for rank-one transformations, Ryzhikov [80] — for transformations of finite rank, and Host [44] — for transformations with singular spectrum.

The multiple mixing property has been extensively studied for flows on homogeneous spaces. Ornstein and Weiss [75] established that the geodesic flow on compact hyperbolic surfaces is Bernoulli, which implies that it is mixing of all orders. Dani [19, 20] proved that a quite general partially hyperbolic one-parameter homogeneous flow satisfies the Kolmogorov property so that, in particular, it is mixing of all orders. Sinai [85] conjectured that the horocycle flow is also mixing of all orders. This conjecture was proved by Marcus [64]. In fact, Marcus' work established mixing of all order for general flows on homogeneous spaces of semisimple groups. Ultimately Starkov [86], building on the work of Mozes [72] and the theory of unipotent flows, proved mixing of all orders for general mixing one-parameter flows on finite-volume homogeneous spaces.

Quantitative estimates on the higher-order correlations (1.1) are also of great importance. Using Fourier-analytic techniques, Lind [63] proved exponential convergence of correlations of order two for ergodic toral automorphisms, and Pène [77] proved this for correlations of all orders. Dolgopyat [25] established a general result about exponential convergence of correlations for partially hyperbolic dynamical systems under the assumption of quantitative equidistribution of translates of unstable manifolds. Gorodnik and Spatzier [37] showed exponential convergence of correlations of all orders for ergodic automorphisms of nilmanifolds.

More generally, we consider a measure-preserving action of a locally compact group  $G$  on a probability space space  $(X, \mu)$ . For functions  $\phi_1, \dots, \phi_r \in L^\infty(X)$ , we define the *correlations of order  $r$*  as

$$\int_X \phi_1(g_1^{-1}x) \cdots \phi_r(g_r^{-1}x) d\mu(x), \quad g_1, \dots, g_r \in G. \quad (1.3)$$

We assume that the group  $G$  is equipped with a (proper) metric  $d$ . We say that the action is *mixing of order  $r$*  if for functions  $\phi_1, \dots, \phi_r \in L^\infty(X)$ ,

$$\int_X \phi_1(g_1^{-1}x) \cdots \phi_r(g_r^{-1}x) d\mu(x) \longrightarrow \left( \int_X \phi_1 d\mu \right) \cdots \left( \int_X \phi_r d\mu \right) \quad (1.4)$$

as  $d(g_i, g_j) \rightarrow \infty$  for all  $i \neq j$ . At present, the available results about higher-order mixing for multi-parameter actions are limited to several particular classes of dynamical systems. del Junco and Yassawi [21, 94] proved that for finite rank actions of countable abelian groups satisfying additional technical conditions, mixing of order two implies mixing of all orders. It was discovered by Ledrappier [55] that mixing of order two does not imply mixing of order three in general for  $\mathbb{Z}^2$ -actions. The example constructed in [55] is an action by automorphisms on a (disconnected) compact abelian group. On the other hand, Schmidt and Ward [81] established that  $\mathbb{Z}^k$ -actions by automorphisms on compact connected abelian groups that are mixing of order two are always mixing of all orders. We refer to the monograph of Schmidt [82] for extensive study of mixing properties for higher-rank abelian actions by automorphisms of compact abelian groups. It turns out that this problem is intimately connected to deep number-theoretic questions that involve analysing solutions of  $S$ -unit equations. Gorodnik and Spatzier [39] proved that mixing  $\mathbb{Z}^k$ -actions by automorphisms on nilmanifolds are mixing of all orders.

Using Diophantine estimates on logarithms of algebraic numbers, the work [39] also established quantitative estimates for correlations of order up to three. The problem of producing explicit quantitative bounds on general higher-order correlations in this setting is still open, even for  $\mathbb{Z}^k$ -actions by toral automorphisms.

In this paper, we provide a self-contained accessible treatment of the analysis of higher-order correlations for measure-preserving actions of a (noncompact) simple connected Lie group  $G$  with finite centre (a less advanced reader may think about the groups like  $\mathrm{SL}_d(\mathbb{R})$  or  $\mathrm{Sp}_{2n}(\mathbb{R})$ ). This topic has long history going back at least to the works of Harish-Chandra (see, for instance, [42]). Indeed, the correlations of order two can be interpreted as matrix coefficients of the corresponding unitary representation on the space  $L^2(X)$ , and quantitative estimates on matrix coefficients have been established in the framework of the Representation Theory (see §2 and §4 for an extensive discussion). This in particular leads to the surprising corollary that every ergodic action of  $G$  is always mixing. Moreover, Mozes [72] proved that every mixing action of  $G$  is always mixing of all order, and Konstantoulas [53] and Björklund, Einsiedler, Gorodnik [7] established quantitative estimates on higher-order correlations. The main goal of these notes is to outline a proof of the following result. Let  $L$  be a connected Lie group and  $\Gamma$  a discrete subgroup of  $L$  with finite covolume. We denote by  $(X, \mu)$  the space  $L/\Gamma$  equipped with the invariant probability measure  $\mu$ . Let  $G$  be a (noncompact) simple connected higher-rank Lie group with finite center equipped with a left-invariant Riemannian metric  $d$ . We consider the measure-preserving action of  $G$  on  $(X, \mu)$  given by a smooth representation  $G \rightarrow L$ . In this setting, we establish quantitative estimates on higher-order correlations:

**Theorem.** *Assuming that the action of  $G$  on  $(X, \mu)$  is ergodic, there exists  $\delta_r > 0$  (depending only on  $G$  and  $r$ ) such that for all elements  $g_1, \dots, g_r \in G$  and all functions  $\phi_1, \dots, \phi_r \in C_c^\infty(X)$ ,*

$$\int_X \phi_1(g_1^{-1}x) \cdots \phi_r(g_r^{-1}x) d\mu(x) = \left( \int_X \phi_1 d\mu \right) \cdots \left( \int_X \phi_r d\mu \right) + O_{\phi_1, \dots, \phi_r, r} \left( e^{-\delta_r D(g_1, \dots, g_r)} \right), \quad (1.5)$$

where

$$D(g_1, \dots, g_r) = \min_{i \neq j} d(g_i, g_j).$$

As we shall explain below, a version of this theorem also holds for rank-one groups  $G$  provided that the action of  $G$  on  $L^2(X)$  satisfies the necessary condition of having the spectral gap. In this case, the exponent  $\delta_r$  also depends on the action.

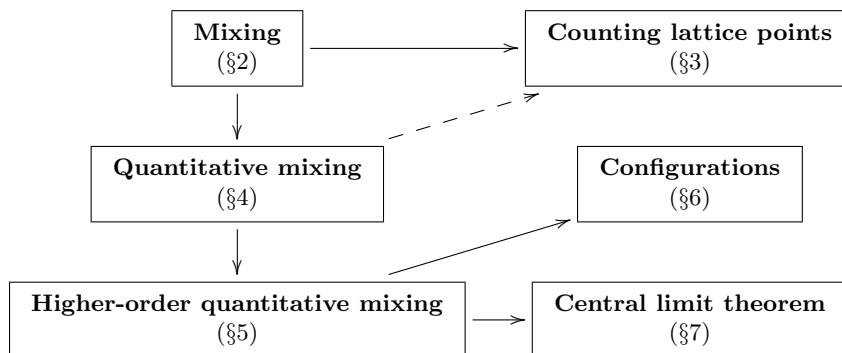
It turns out that analysis of correlations has several far-reaching applications, and here we outline how to use this approach

- to establish an asymptotic formula for the number of lattice points,
- to show existence of approximate configurations in lattice subgroups,
- to prove the Central Limit Theorem for group actions.

Other interesting applications of quantitative bounds on correlations, which we do not discuss here, involve the Kazhdan property (T) [48, §V.4.1], the cohomological equation [51, 37, 39], the global rigidity of actions [29], and analysis of the distribution of arithmetic counting functions [9, 10].

This paper is based on a series of lectures given at the Tata Institute of Fundamental Research which involved participants with quite diverse backgrounds ranging from starting PhD students to senior researchers. When I was choosing the material, I was aiming to make it accessible, but at the same time to give a reasonably detailed exposition of the developed methods as well as to survey current state of the art in the field. This inevitably required some compromises. In particular, we assumed very little knowledge of the Theory of Lie groups, and some of the arguments are carried out only in the case when  $G = \mathrm{SL}_d(\mathbb{R})$ . I hope that a less prepared reader should be able to follow this paper by thinking that a “connected simple Lie group with finite center” is  $\mathrm{SL}_d(\mathbb{R})$ , and advanced readers might be able to infer from our exposition how to deal with the general case. Besides giving a self-contained proof of the bound (1.5), we also state a number of more advanced results without proofs, which are indicated by the symbol (\*).

**Organisation of the paper.** We are not aware of any direct way for proving the main bound (1.5), and our arguments proceeds in several distinct steps. First, in §2 and §4, we study the behaviour of correlations of order two using representation-theoretic techniques. In §2 we show that the correlations of order two decay at infinity (see Theorem 2.3), and in §4 we establish quantitative bounds on the correlations of order two (see Theorem 4.10). Then the main bound (1.5) is established using an elaborate inductive argument in §5 (see Theorem 5.1). We also discuss several application of the established bounds for correlations: in §3 we derive an asymptotic formula for the number of lattice points, in §6 we establish existence of approximate configurations, and in §7 we prove the Central Limit Theorem for group actions.



**Acknowledgement.** This survey paper has grown out of the lecture series given by the author at the Tata Institute of Fundamental Research in Spring 2017. I would like to express my deepest gratitude to the Tata Institute for the hospitality and to the organisers of this programme – Shrikrishna Dani and Anish Ghosh – for all their hard work on setting up this event and making it run smoothly. I also would like to thank the referee for carefully reading the draft and for his thoughtful comments.

## 2. DECAY OF MATRIX COEFFICIENTS

Let  $G$  be a (noncompact) connected simple Lie group with finite center (e.g.,  $G = \mathrm{SL}_d(\mathbb{R})$ ). We consider a measure-preserving action  $G$  on a standard probability space  $(X, \mu)$ . The goal of this section is to show a surprising result that ergodicity of any such action implies that it is mixing:

**Theorem 2.1.** *Let  $G$  be a (noncompact) connected simple Lie group with finite centre and  $G \times X \rightarrow X$  a measurable measure-preserving action on a standard probability space  $(X, \mu)$ . We assume that the action of  $G$  on  $(X, \mu)$  is ergodic (that is, the space  $L^2(X)$  has no nonconstant  $G$ -invariant functions). Then for all  $\phi, \psi \in L^2(X)$ ,*

$$\int_X \phi(g^{-1}x)\psi(x) d\mu(x) \longrightarrow \left( \int_X \phi d\mu \right) \left( \int_X \psi d\mu \right)$$

as  $g \rightarrow \infty$  in  $G$ .

We observe that a measure-preserving action as above defines a unitary representation  $\pi$  of  $G$  on the space  $\mathcal{H} = L^2(X)$  given by

$$\pi(g)\phi(x) = \phi(g^{-1}x) \quad \text{for } g \in G \text{ and } x \in X. \quad (2.1)$$

One can also check (see, for instance, [5, A.6]) that this representation is strongly continuous (that is, the map  $g \mapsto \pi(g)\phi$ ,  $g \in G$ , is continuous).

**Convention.** Throughout these notes, we always implicitly assume that representations are strongly continuous and Hilbert spaces are separable.

Theorem 2.1 can be formulated more abstractly in terms of asymptotic vanishing of matrix coefficients of unitary representations.

**Theorem 2.2.** *Let  $G$  be a (noncompact) connected simple Lie group with finite centre and  $\pi$  a unitary representation of  $G$  on a Hilbert space  $\mathcal{H}$ . Then for all  $v, w \in \mathcal{H}$ ,*

$$\langle \pi(g)v, w \rangle \rightarrow \langle P_G v, P_G w \rangle \quad \text{as } g \rightarrow \infty \text{ in } G,$$

where  $P_G$  denotes the orthogonal projection on the subspace of the  $G$ -invariant vectors.

The study of matrix coefficients for unitary representations of semisimple Lie groups has a long history. In particular, this subject played important role in the research programme of Harish-Chandra. We refer to the monographs [92, 93, 34, 52] for expositions of this theory. Explicit quantitative bounds on the matrix coefficient, which in particular imply Theorem 2.2, were derived in the works of Borel and Wallach [11], Cowling [17], and Casselman and Milicić [13]. This initial approach to the study of asymptotic properties of matrix coefficients used elaborate analytic arguments that involved representing them as solutions of certain systems of PDE's. Subsequently, Howe and Moore [47] developed a different approach to prove Theorem 2.2 that used the Mautner phenomenon (cf. Theorem 2.4 below) and an inductive argument that derived vanishing of matrix coefficients on the whole group from vanishing along a sufficiently rich collection of subgroups. We present a version of this method here. Other treatments of Theorems 2.1 and 2.2 can be also found in the monographs [98, 48, 6].

It is worthwhile to mention that the Howe–Moore argument [47] is not restricted just to semisimple groups, and it gives the following general result. Given an irreducible unitary representation  $\pi$  of a group  $G$ , we denote by

$$R_\pi = \{g \in G : \pi(g) \in \mathbb{C}^\times \text{id}\}$$

its projective kernel. Since  $\pi$  is unitary, it is clear that the matrix coefficients  $|\langle \pi(g)v, w \rangle|$  are constant on cosets of  $R_\pi$ . One of the main results of [47] is asymptotic vanishing of matrix coefficients along  $G/R_\pi$ :

**Theorem\* 2.3.** *Let  $G$  be a connected real algebraic group and  $\pi$  an irreducible representation of  $G$  on a Hilbert space  $\mathcal{H}$ . Then for any  $v, w \in \mathcal{H}$ ,*

$$\langle \pi(g)v, w \rangle \rightarrow 0 \quad \text{as } g \rightarrow \infty \text{ in } G/R_\pi.$$

Now we start the proof of Theorem 2.2. First, we note that because of the decomposition

$$\mathcal{H} = \mathcal{H}_G \oplus \mathcal{H}_G^\perp,$$

where  $\mathcal{H}_G$  denotes the subspace of  $G$ -invariant vectors, it is sufficient to prove that for all vectors  $v, w \in \mathcal{H}_G^\perp$ ,

$$\langle \pi(g)v, w \rangle \rightarrow 0 \quad \text{as } g \rightarrow \infty,$$

and without loss of generality, we may assume that  $\mathcal{H}$  contains no nonzero  $G$ -invariant vectors.

The proof will proceed by contradiction. Suppose that, in contrary,

$$\langle \pi(g^{(n)})v, w \rangle \not\rightarrow 0 \tag{2.2}$$

for some sequence  $g^{(n)} \rightarrow \infty$  in  $G$ . We divide the proof into four steps.

Step 1: Cartan decomposition. We shall use the Cartan decomposition for  $G$ :

$$G = KA^+K,$$

where  $K$  is a maximal compact subgroup of  $G$ , and  $A^+$  is a positive Weyl chamber of a Cartan subgroup of  $G$ . For instance, when  $G = \text{SL}_d(\mathbb{R})$ , this decomposition holds with

$$K = \text{SO}(d) \quad \text{and} \quad A^+ = \{\text{diag}(a_1, \dots, a_d) : a_1 \geq a_2 \geq \dots \geq a_d > 0\}.$$

The Cartan decomposition will be applied to go from the assumption (2.2) to the corresponding non-decay statement with  $g^{(n)}$  in the Cartan subgroup  $A$ .

We write

$$g^{(n)} = k^{(n)}a^{(n)}\ell^{(n)} \quad \text{with } k^{(n)}, \ell^{(n)} \in K \text{ and } a^{(n)} \in A^+.$$

Since  $K$  is compact, it follows that  $a^{(n)} \rightarrow \infty$  in  $A^+$ . Passing to a subsequence, we may arrange that the sequences  $k^{(n)}$  and  $\ell^{(n)}$  converge in  $K$  so that, in particular,

$$\pi(\ell^{(n)})v \rightarrow v' \quad \text{and} \quad \pi(k^{(n)})^*w \rightarrow w'$$

for some vectors  $v', w' \in \mathcal{H}$ . We observe that

$$\begin{aligned} \langle \pi(g^{(n)})v, w \rangle - \langle \pi(a^{(n)})v', w' \rangle &= \langle \pi(a^{(n)})\pi(\ell^{(n)})v, \pi(k^{(n)})^*w \rangle - \langle \pi(a^{(n)})v', w' \rangle \\ &= \langle \pi(a^{(n)})(\pi(\ell^{(n)})v - v'), \pi(k^{(n)})^*w \rangle \\ &\quad + \langle \pi(a^{(n)})v', \pi(k^{(n)})^*w - w' \rangle. \end{aligned}$$

Using that the representation  $\pi$  is unitary, we deduce that

$$\begin{aligned} \left| \left\langle \pi(a^{(n)})\pi(\ell^{(n)})v - v', \pi(k^{(n)})^*w \right\rangle \right| &\leq \left\| \pi(a^{(n)})\pi(\ell^{(n)})v - v' \right\| \left\| \pi(k^{(n)})^*w \right\| \\ &= \left\| \pi(\ell^{(n)})v - v' \right\| \|w\| \rightarrow 0. \end{aligned}$$

Similarly, one can show that

$$\left\langle \pi(a^{(n)})v', \pi(k^{(n)})^*w - w' \right\rangle \rightarrow 0.$$

Hence, we conclude that

$$\left\langle \pi(g^{(n)})v, w \right\rangle = \left\langle \pi(a^{(n)})v', w' \right\rangle + o(1),$$

and it follows from (2.2) that

$$\left\langle \pi(a^{(n)})v', w' \right\rangle \not\rightarrow 0. \quad (2.3)$$

*Step 2: weak convergence.* We use the notion of ‘weak convergence’. We recall that a sequence of vectors  $x^{(n)}$  in a Hilbert space converges weakly to a vector  $x$  if  $\langle x^{(n)}, y \rangle \rightarrow \langle x, y \rangle$  for all  $y \in \mathcal{H}$ . We use the notation:  $x^{(n)} \xrightarrow{w} x$ . It is known that every bounded sequence has a weakly convergent subsequence. In particular, it follows that, after passing to a subsequence, we may arrange that

$$\pi(a^{(n)})v' \xrightarrow{w} v''$$

for some vector  $v'' \in \mathcal{H}$ . Then, in particular,

$$\left\langle \pi(a^{(n)})v', w' \right\rangle \rightarrow \langle v'', w' \rangle \neq 0$$

by (2.3).

*Step 3: the case when  $G = \mathrm{SL}_2(\mathbb{R})$ .* From the previous step, we know that

$$\left\langle \pi(a^{(n)})v', w' \right\rangle \rightarrow \langle v'', w' \rangle \neq 0 \quad \text{for } a^{(n)} = \begin{pmatrix} t_n & 0 \\ 0 & t_n^{-1} \end{pmatrix} \text{ with } t_n \rightarrow \infty.$$

We claim that the vector  $v''$  is invariant under the subgroup

$$U = \left\{ u(s) = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} : s \in \mathbb{R} \right\}.$$

This property will be deduced from the identity

$$(a^{(n)})^{-1}u(s)a^{(n)} = u(s/t_n^2) \rightarrow e.$$

One can easily check that

$$\pi(u(s))\pi(a^{(n)})v' \xrightarrow{w} \pi(u(s))v'',$$

so that

$$\pi(u(s))v'' = \mathrm{w}\text{-}\lim_{n \rightarrow \infty} \pi(u(s))\pi(a^{(n)})v' = \mathrm{w}\text{-}\lim_{n \rightarrow \infty} \pi(a^{(n)})\pi(u(s/t_n^2))v'.$$

Since

$$\|\pi(a^{(n)})\pi(u(s/t_n^2))v' - \pi(a^{(n)})v'\| = \|\pi(u(s/t_n^2))v' - v'\| \rightarrow 0,$$

it follows that

$$\mathrm{w}\text{-}\lim_{n \rightarrow \infty} \pi(a^{(n)})\pi(u(s/t_n^2))v' = \mathrm{w}\text{-}\lim_{n \rightarrow \infty} \pi(a^{(n)})v' = v''.$$

This proves that indeed the vector  $v''$  is invariant under  $U$ .

Next, we show that the vector  $v''$  is  $G$ -invariant. We consider the function

$$F(g) = \langle \pi(g)v'', v'' \rangle \quad \text{with } g \in G.$$

Since  $v''$  is  $U$ -invariant, the function  $F$  is bi-invariant under  $U$ . We observe that the map  $gU \mapsto ge_1$  defines an isomorphism between the homogeneous spaces  $G/U$  and  $\mathbb{R}^2 \setminus \{0\}$ . Hence, we may consider  $F$  as a function  $\mathbb{R}^2 \setminus \{0\}$ . Since the  $U$ -orbits in  $\mathbb{R}^2$  are the lines  $y = c$  with  $c \neq 0$  and the points  $(x, 0)$ , we conclude that  $F$  is constant on each line  $y = c$  with  $c \neq 0$ . By continuity, it follows that  $F$  is also constant on the line  $y = 0$ . For  $a(t) = \begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix}$ ,

$$\langle \pi(a(t))v'', v'' \rangle = F(a(t)e_1) = F(te_1) = F(e_1) = \|v''\|^2.$$

Since this gives the equality in the Cauchy–Schwarz inequality, the vectors  $\pi(a(t))v''$  and  $v''$  must be colinear, and we deduce that  $\pi(a(t))v'' = v''$ , so that the vector  $v''$  is also invariant under the subgroup  $A = \{a(t)\}$ . Hence, the function  $F$  is also constant on  $AU$ -orbits in  $\mathbb{R}^2$ . Since the half-spaces  $\{y > 0\}$  and  $\{y < 0\}$  are single  $AU$ -orbits, it follows from the continuity of  $F$ , that this function is identically constant, that is,

$$F(g) = \langle \pi(g)v'', v'' \rangle = \|v''\|^2 \quad \text{for all } g \in G.$$

This gives the equality in the Cauchy–Schwarz inequality, and as before we deduce that  $\pi(g)v'' = v''$  for all  $g \in G$ . However, we have assumed that there is nonzero  $G$ -invariant vectors. This gives a contradiction, and completes the proof of the theorem in the case  $G = \mathrm{SL}_2(\mathbb{R})$ .

We note that the above argument, in fact, implies the following ‘‘Mautner property’’ of unitary representations of  $\mathrm{SL}_2(\mathbb{R})$ : every  $U$ -invariant vector is always  $\mathrm{SL}_2(\mathbb{R})$ -invariant. More generally, one says that a closed subgroup  $H$  of topological group  $G$  has *Mautner property* if for every unitary representation of  $G$ ,  $H$ -invariant vectors are also invariant under  $G$ . Subgroups satisfying this property have appeared in a work of Segal and von Neumann [83], and Mautner [68] used this phenomenon to study ergodicity of the geodesic flow on locally symmetric spaces. The following general version of the Mautner property was established by Moore [69]:

**Theorem\* 2.4.** *Let  $G$  be a (noncompact) simple connected Lie group with finite center. Then every noncompact closed subgroup of  $G$  has the Mautner property.*

Subsequently, more general versions of this result were proved by Moore [70], Wang [89, 90], and Bader, Furman, Gorodnik, Weiss [2].

*Step 4: inductive argument.* Our next task is to develop an inductive argument which allows to deduce asymptotic vanishing of matrix coefficients using vanishing along smaller subgroup. We give a complete proof when  $G = \mathrm{SL}_d(\mathbb{R})$ , but similar ideas can be also extended to general semisimple Lie groups using their structure theory.

For  $a \in A^+$ , the set of simple roots is given by

$$\alpha_i(a) = a_i/a_{i+1} \quad \text{for } i = 1, \dots, d-1.$$

The functions  $\alpha_i$ ,  $i = 1, \dots, d-1$ , provide a coordinate system on  $A^+$ . Given a sequence  $a^{(n)} \in A^+$  such  $a^{(n)} \rightarrow \infty$ , we have  $\max_i \alpha_i(a^{(n)}) \rightarrow \infty$ . After passing to

a subsequence, we may assume that  $\alpha_i(a^{(n)}) \rightarrow \infty$  for some  $i$ . We introduce the subgroup

$$U_i = \left\{ \begin{pmatrix} I_i & u \\ 0 & I_{d-i} \end{pmatrix} : u \in M_{i,d-i}(\mathbb{R}) \right\}.$$

For  $a \in A^+$ ,

$$a^{-1} \begin{pmatrix} I_i & (u_{lk}) \\ 0 & I_{d-i} \end{pmatrix} a = \begin{pmatrix} I_i & \left(\frac{a_k}{a_l} u_{lk}\right) \\ 0 & I_{d-i} \end{pmatrix}.$$

Since for  $l \leq i < k$ ,

$$\frac{a_l}{a_k} = \frac{a_l}{a_{l+1}} \cdots \frac{a_{k-1}}{a_k} \geq \frac{a_i}{a_{i+1}} = \alpha_i(a),$$

it follows that  $a_i^{(n)}/a_k^{(n)} \rightarrow \infty$ . Hence, for  $g \in U_i$ ,

$$(a^{(n)})^{-1} g a^{(n)} \rightarrow e.$$

Using this property, we may argue exactly as in Step 3 to conclude that the vector  $v''$  is  $U_i$ -invariant. For  $1 \leq l \leq i$  and  $i+1 \leq k \leq d$ , we denote by  $U_{lk}$  the corresponding one-parameter unipotent subgroup of  $U_i$ . We observe that  $U_{lk}$  can be embedded in an obvious way as a subgroup of the group  $G_{lk} \simeq \mathrm{SL}_2(\mathbb{R})$  contained in  $G$ . Since the vector  $v''$  is invariant under  $U_{lk}$ , it follows from Step 3 that it is also invariant under  $G_{lk}$  when  $1 \leq l \leq i$  and  $i+1 \leq k \leq d$ . Finally, one can check<sup>1</sup> that these groups  $G_{lk}$  generate  $G = \mathrm{SL}_d(\mathbb{R})$ , so that the vector  $v''$  is  $G$ -invariant. This gives a contradiction and completes the proof of the theorem.

### 3. APPLICATION: COUNTING LATTICE POINTS

Given a lattice  $\Lambda$  in the Euclidean space  $\mathbb{R}^d$  and a Euclidean ball  $B$  in  $\mathbb{R}^d$ , one can show using a simple geometric argument that

$$|\Lambda \cap B| \sim \frac{\mathrm{vol}(B)}{\mathrm{vol}(\mathbb{R}^d/\Lambda)} \quad \text{as } \mathrm{vol}(B) \rightarrow \infty.$$

This result also holds for more general families of domains satisfying some regularity assumptions. The analogous lattice counting problem for the hyperbolic space is more difficult because of the exponential volume growth of the hyperbolic balls, and proving an asymptotic formula even without an error term requires analytic tools.

The hyperbolic lattice point counting problem was studied by Delsarte [22], Huber [49], and Patterson [76]. These works used spectral expansion of the counting functions in terms of the eigenfunctions of the Laplace-Beltrami operator. Margulis in his PhD thesis [65, 66] discovered that the lattice point counting problem on manifolds of variable negative curvature can be solved using solely the mixing property. Bartels [3] proved an asymptotic formula for the number of lattice points in connected semisimple Lie groups using a version of Theorem 2.1. Subsequently, this approach was generalised to counting lattice orbit points on affine symmetric varieties by Duke, Rudnick, Sarnak [26] and Eskin, McMullen [28]. We refer to the survey [1] for a comprehensive discussion of the lattice point counting problems.

In this section, we consider a more general counting problem for lattice points in locally compact groups. In particular, we prove the following result:

<sup>1</sup>The reader might find easier to verify first that the Lie algebras of  $G_{lk}$  generate the whole space of traceless matrices, which implies the original claim.

**Theorem 3.1.** *Let  $G$  be a (noncompact) connected simple matrix Lie group with finite centre, and the sets*

$$B_t = \{g \in G : \|g\| \leq t\} \quad (3.1)$$

*are defined by a norm on the space of matrices. Then for any lattice subgroup  $\Gamma$  of  $G$ ,*

$$|\Gamma \cap B_t| \sim \frac{\text{vol}(B_t)}{\text{vol}(G/\Gamma)} \quad \text{as } t \rightarrow \infty.$$

More generally, we establish an asymptotic counting formula in a setting of locally compact groups satisfying a certain mixing assumption. Let  $G$  be a locally compact second countable group and  $\Gamma$  a lattice subgroup in  $G$ . We fix a Haar measure  $m$  on  $G$  which also induced the measure  $\mu$  on the factor space  $X = G/\Gamma$  by

$$\int_{G/\Gamma} \left( \sum_{\gamma \in \Gamma} \psi(g\gamma) \right) d\mu(g\Gamma) = \int_G \psi dm \quad \text{for } \psi \in C_c(G/\Gamma).$$

We normalise the measure  $m$  so that  $\mu(X) = 1$ . Then we obtain a continuous measure-preserving action of  $G$  on the probability space  $(X, \mu)$ .

We say that a family of bounded measurable sets  $B_t$  in  $G$  is *well-rounded* (cf. [28, Section 6]) if for every  $\delta > 1$ , there exists a symmetric neighbourhood  $\mathcal{O}$  of identity in  $G$  such that

$$\delta^{-1} m \left( \bigcup_{g_1, g_2 \in \mathcal{O}} g_1 B_t g_2 \right) \leq m(B_t) \leq \delta m \left( \bigcap_{g_1, g_2 \in \mathcal{O}} g_1 B_t g_2 \right) \quad (3.2)$$

for all  $t$ .

**Theorem 3.2.** *Let  $G$  be a locally compact second countable group, and let  $B_t$  be a family of well-rounded compact sets in  $G$  such that  $m(B_t) \rightarrow \infty$ . Let  $\Gamma$  be a lattice subgroup in  $G$  such that the action of  $G$  on the space  $G/\Gamma$  is mixing. Then*

$$|\Gamma \cap B_t| \sim m(B_t) \quad \text{as } t \rightarrow \infty.$$

It follows from a Fubini-type argument that the action of  $G$  on  $X = G/\Gamma$  is ergodic (i.e., every almost everywhere invariant function is constant almost everywhere). Hence, when  $G$  is a connected simple Lie group with finite centre, it follows from Theorem 2.1 that the actions of  $G$  on  $X = G/\Gamma$  is mixing of order two. One can also check that the regularity condition (3.2) is satisfied for the norm balls (3.1) (see [26],[28]). Hence, Theorem 3.1 follows from Theorem 3.2.

We start the proof of Theorem 3.2 by realising the counting function as a function on the homogeneous space  $X = G/\Gamma$ . We set

$$F_t(g_1, g_2) = \sum_{\gamma \in \Gamma} \chi_{B_t}(g_1 \gamma g_2^{-1}). \quad (3.3)$$

In the first part of the argument, we do not impose any regularity assumptions on the compact domains  $B_t$  and just assume that  $m(B_t) \rightarrow \infty$ . Since

$$F_t(g_1 \gamma_1, g_2 \gamma_2) = F_t(g_1, g_2) \quad \text{for all } g_1, g_2 \in G \text{ and } \gamma_1, \gamma_2 \in \Gamma,$$

$F_t$  defines a function on  $G/\Gamma \times G/\Gamma$ . We note that  $F_t(e, e) = |\Gamma \cap B_t|$ , so that it remains to investigate the asymptotic behaviour of  $F_t$  at the identity coset.

The crucial connection between the original counting problem and estimating correlations is provided by the following computation. For a real-valued test-function  $\phi \in C_c(G/\Gamma)$ , we obtain that

$$\begin{aligned} \langle F_t, \phi \otimes \phi \rangle &= \int_{G/\Gamma \times G/\Gamma} F_t(g_1, g_2) \phi(g_1) \phi(g_2) d\mu(g_1\Gamma) d\mu(g_2\Gamma) \\ &= \int_{G/\Gamma \times G/\Gamma} \left( \sum_{\gamma \in \Gamma} \chi_{B_t}(g_1 \gamma g_2^{-1}) \right) \phi(g_1) \phi(g_2) d\mu(g_1\Gamma) d\mu(g_2\Gamma) \\ &= \int_{G/\Gamma \times G/\Gamma} \left( \sum_{\gamma \in \Gamma} \chi_{B_t}(g_1 (g_2 \gamma)^{-1}) \right) \phi(g_1) \phi(g_2) d\mu(g_1\Gamma) d\mu(g_2\Gamma) \\ &= \int_{G/\Gamma \times G} \chi_{B_t}(g_1 g_2^{-1}) \phi(g_1) \phi(g_2) d\mu(g_1\Gamma) dm(g_2). \end{aligned}$$

We denote by  $\pi$  the unitary representation of  $G$  on  $L^2(G/\Gamma)$  defined as in (2.1). Using a change of variables  $b = g_1 g_2^{-1}$ , we deduce that

$$\begin{aligned} \langle F_t, \phi \otimes \phi \rangle &= \int_{G/\Gamma \times G} \chi_{B_t}(b) \phi(g_1) \phi(b^{-1} g_1) d\mu(g_1\Gamma) dm(b) \\ &= \int_{B_t} \langle \pi(b) \phi, \phi \rangle dm(b). \end{aligned}$$

According to our assumption,

$$\langle \pi(b) \phi, \phi \rangle \longrightarrow \left( \int_{G/\Gamma} \phi d\mu \right)^2 \quad \text{as } b \rightarrow \infty.$$

Hence, since  $m(B_t) \rightarrow \infty$ , it follows that

$$\frac{\langle F_t, \phi \otimes \phi \rangle}{m(B_t)} = \frac{1}{m(B_t)} \int_{B_t} \langle \pi(b) \phi, \phi \rangle dm(b) \longrightarrow \left( \int_{G/\Gamma} \phi d\mu \right)^2 \quad (3.4)$$

as  $t \rightarrow \infty$ . We note that (3.4) holds for any functions  $F_t$  defined in terms of compact subsets  $B_t$  such that  $m(B_t) \rightarrow \infty$ .

Our next task is to upgrade the weak convergence of functions  $F_t$  established in (3.4) to the pointwise convergence. For this step, we use the regularity assumption (3.2) on the domains  $B_t$ . We take any  $\delta > 1$  and choose the neighbourhood  $\mathcal{O}$  of identity in  $G$  as in (3.2). We set

$$B_t^+ = \bigcup_{g_1, g_2 \in \mathcal{O}} g_1 B_t g_2 \quad \text{and} \quad B_t^- = \bigcap_{g_1, g_2 \in \mathcal{O}} g_1 B_t g_2,$$

and consider the corresponding functions  $F_t^+$  and  $F_t^-$  defined as in (3.3). It follows from (3.2) that

$$\delta^{-1} m(B_t^+) \leq m(B_t) \leq \delta m(B_t^-). \quad (3.5)$$

In particular,  $m(B_t^\pm) \rightarrow \infty$ .

We take a nonnegative function  $\tilde{\phi} \in C_c(G)$  such that

$$\text{supp}(\tilde{\phi}) \subset \mathcal{O} \quad \text{and} \quad \int_G \tilde{\phi} dm = 1,$$

and define a function  $\phi \in C_c(G/\Gamma)$  as  $\phi(g) = \sum_{\gamma \in \Gamma} \tilde{\phi}(g\gamma)$ . Then

$$\begin{aligned} \langle F_t^+, \phi \otimes \phi \rangle &= \int_{G/\Gamma \times G/\Gamma} F_t^+(g_1, g_2) \phi(g_1) \phi(g_2) d\mu(g_1\Gamma) d\mu(g_2\Gamma) \\ &= \int_{G/\Gamma \times G/\Gamma} F_t^+(g_1, g_2) \left( \sum_{\gamma_1, \gamma_2 \in \Gamma} \tilde{\phi}(g_1\gamma_1) \tilde{\phi}(g_2\gamma_2) \right) d\mu(g_1\Gamma) d\mu(g_2\Gamma) \\ &= \int_{G/\Gamma \times G/\Gamma} \left( \sum_{\gamma_1, \gamma_2 \in \Gamma} F_t^+(g_1\gamma_1, g_2\gamma_2) \tilde{\phi}(g_1\gamma_1) \tilde{\phi}(g_2\gamma_2) \right) d\mu(g_1\Gamma) d\mu(g_2\Gamma) \\ &= \int_{G \times G} F_t^+(g_1, g_2) \tilde{\phi}(g_1) \tilde{\phi}(g_2) dm(g_1) dm(g_2). \end{aligned}$$

We observe that when  $g_1, g_2 \in \mathcal{O}$ ,

$$F_t^+(g_1, g_2) = \sum_{\gamma \in \Gamma} \chi_{g_1^{-1}B_t^+g_2}(\gamma) \geq \sum_{\gamma \in \Gamma} \chi_{B_t}(\gamma) = |\Gamma \cap B_t|,$$

so that since  $\text{supp}(\tilde{\phi}) \subset \mathcal{O}$ , we obtain that

$$\langle F_t^+, \phi \otimes \phi \rangle \geq |\Gamma \cap B_t| \left( \int_G \tilde{\phi} dm \right)^2 \geq |\Gamma \cap B_t|.$$

Hence, it follows from (3.4) and (3.5) that

$$\limsup_{t \rightarrow \infty} \frac{|\Gamma \cap B_t|}{m(B_t)} \leq \delta \limsup_{t \rightarrow \infty} \frac{\langle F_t^+, \phi \otimes \phi \rangle}{m(B_t^+)} = \delta$$

for all  $\delta > 1$ . A similar argument applied to the function  $F_t^-$  gives

$$\liminf_{t \rightarrow \infty} \frac{|\Gamma \cap B_t|}{m(B_t)} \geq \delta^{-1} \liminf_{t \rightarrow \infty} \frac{\langle F_t^-, \phi \otimes \phi \rangle}{m(B_t^-)} = \delta^{-1}$$

for all  $\delta > 1$ . This implies Theorem 3.2.

It is worthwhile to mention that Gorodnik and Nevo [35, 36] showed the asymptotic formula for counting lattice points can be deduced solely from an ergodic theorem for averages along the sets  $B_t$  on the space  $X = G/\Gamma$ . ‘Ergodic theorem’ is a much more prolific phenomenon than ‘mixing property’.

#### 4. QUANTITATIVE ESTIMATES ON MATRIX COEFFICIENTS

The goal of this section is to establish quantitative estimates on matrix coefficients for unitary representations  $\pi$  of higher-rank simple groups  $G$  (for instance, for  $G = \text{SL}_d(\mathbb{R})$  with  $d \geq 3$ ). It is quite remarkable that this quantitative bound for higher-rank groups holds uniformly for all representations without invariant vectors.

A qualitative bound on matrix coefficients may only hold on a proper subset of vectors, and to state such a bound, we introduce a notion of  $K$ -finite vectors. Let  $K$  be a maximal compact subgroup of  $G$ . By the Peter–Weyl Theorem, a unitary representation  $\pi|_K$  splits as a sum of finite-dimensional irreducible representations. A vector  $v$  is called  $K$ -finite if the span of  $\pi(K)v$  is of finite dimension. We set

$$d_K(v) = \dim \langle \pi(K)v \rangle.$$

The space of  $K$ -finite vectors is dense in the representation space.

With this notation, we prove:

**Theorem 4.1.** *Let  $G$  be a (noncompact) connected simple higher-rank matrix Lie group with finite centre and  $K$  a maximal compact subgroup of  $G$ . Then there exist  $c, \delta > 0$  such that for any unitary representation  $\pi$  of  $G$  on a Hilbert space  $\mathcal{H}$  without nonzero  $G$ -invariant vectors, the following estimate holds: for all elements  $g \in G$  and all  $K$ -finite vectors  $v, w \in \mathcal{H}$ ,*

$$|\langle \pi(g)v, w \rangle| \leq c d_K(v)^{1/2} d_K(w)^{1/2} \|v\| \|w\| \|g\|^{-\delta}.$$

As we already remarked in Section 2, asymptotic properties of matrix coefficients for semisimple Lie groups has been studied extensively starting with foundational works of Harish-Chandra (see, for instance, [42]). Explicit quantitative bounds on matrix coefficients have been obtained, in particular, in the works of Borel and Wallach [11], Cowling [17], Howe [46], Casselman and Milicić [13], Cowling, Haagerup, and Howe [18], Li [61], Li and Zhu [62], and Oh [73, 74]. Here we follow the elegant elementary approach of Howe and Tan [48] to prove Theorem 4.1.

We start our investigation by analysing the unitary representations of the semidirect product

$$L = \mathrm{SL}_2(\mathbb{R}) \ltimes \mathbb{R}^2.$$

We shall use the following notation:

$$\begin{aligned} S &= \mathrm{SL}_2(\mathbb{R}), \\ K_S &= \mathrm{SO}(2) = \left\{ k(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} : \theta \in [0, 2\pi) \right\}, \\ A_S &= \left\{ a(t) = \begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix} : t > 0 \right\}, \\ U_S &= \left\{ u(s) = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} : s \in \mathbb{R} \right\}. \end{aligned} \quad (4.1)$$

**Proposition 4.2.** *Let  $\pi$  be a unitary representation of  $L$  on the Hilbert space  $\mathcal{H}$  that does not have any nonzero  $\mathbb{R}^2$ -invariant vectors. Then for all vectors  $v, w$  belonging to a  $K_S$ -invariant dense subspace of  $\mathcal{H}$ ,*

$$|\langle \pi(a(t))v, w \rangle| \leq c(v, w) t^{-1} \quad \text{when } t \geq 1. \quad (4.2)$$

*Proof.* We consider the restricted representation  $\pi|_{\mathbb{R}^2}$  which can be decomposed with respect irreducible one-dimensional unitary representations of  $\mathbb{R}^2$  — the unitary characters of  $\mathbb{R}^2$ :

$$r \mapsto \chi_z(r) = e^{i\langle z, r \rangle}, \quad z \in \mathbb{R}^2.$$

Namely, there exists a Borel projection-valued measure  $P$  on  $\mathbb{R}^2$  such that

$$\pi(r) = \int_{\mathbb{R}^2} \chi_z(r) dP_z \quad \text{for } r \in \mathbb{R}^2.$$

We shall use that the measure  $P$  satisfies an equivariance property with respect to the action of  $S$ : since

$$\pi(g)\pi(r)\pi(g)^{-1} = \pi(g(r)) \quad \text{for } g \in S \text{ and } r \in \mathbb{R}^2,$$

and

$$\pi(g(r)) = \int_{\mathbb{R}^2} \chi_z(g(r)) dP_z = \int_{\mathbb{R}^2} \chi_{g^t(z)}(r) dP_z,$$

it follows that

$$\pi(g)P_B\pi(g)^{-1} = P_{(g^t)^{-1}B} \quad \text{for Borel } B \subset \mathbb{R}^2 \text{ and } g \in S. \quad (4.3)$$

For  $s > 1$ , we set

$$\Omega_s = \{r \in \mathbb{R}^2 : s^{-1} \leq \|r\| \leq s\},$$

and consider the closed subspace  $\mathcal{H}_s = \text{Im}(P_{\Omega_s})$ . Since the set  $\Omega_s$  is invariant under  $K_S$ , it follows from (4.3) that the subspace  $\mathcal{H}_s$  is  $K_S$ -invariant. Using that the projection-valued measure  $P$  is strongly continuous, we deduce that for all  $v \in \mathcal{H}$ ,

$$P_{\Omega_s}v \rightarrow P_{\mathbb{R}^2 \setminus \{0\}}v = v - P_0v \quad \text{as } s \rightarrow \infty.$$

Moreover, since we assumed that there is no nonzero  $\mathbb{R}^2$ -invariant vectors,  $P_0 = 0$ . This shows that  $\cup_{s>1}\mathcal{H}_s$  is dense in  $\mathcal{H}$ . Using that the span of  $K_S$ -eigenvectors is dense in  $\mathcal{H}_s$ , it remains to show that (4.2) holds for all  $K_S$ -eigenvectors in  $\mathcal{H}_s$  with  $s > 1$ . We recall that  $P_{B_1}P_{B_2} = P_{B_1 \cap B_2}$  for Borel  $B_1, B_2 \subset \mathbb{R}^2$ . This allows us to compute that for  $v, w \in \mathcal{H}_s$ ,

$$\begin{aligned} \langle \pi(a(t))v, w \rangle &= \langle \pi(a(t))P_{\Omega_s}v, P_{\Omega_s}w \rangle = \langle P_{a(t)^{-1}\Omega_s}\pi(a(t))v, P_{\Omega_s}w \rangle \\ &= \langle P_{a(t)^{-1}\Omega_s}\pi(a(t))v, P_{a(t)^{-1}\Omega_s}P_{\Omega_s}w \rangle \\ &= \langle P_{a(t)^{-1}\Omega_s}\pi(a(t))v, P_{a(t)^{-1}\Omega_s \cap \Omega_s}w \rangle \\ &= \langle P_{a(t)^{-1}\Omega_s \cap \Omega_s}\pi(a(t))v, P_{a(t)^{-1}\Omega_s \cap \Omega_s}w \rangle \\ &= \langle \pi(a(t))P_{\Omega_s \cap a(t)\Omega_s}v, P_{a(t)^{-1}\Omega_s \cap \Omega_s}w \rangle. \end{aligned}$$

Hence, by the Cauchy–Schwarz inequality,

$$|\langle \pi(a(t))v, w \rangle| \leq \|P_{\Omega_s \cap a(t)\Omega_s}v\| \|P_{a(t)^{-1}\Omega_s \cap \Omega_s}w\|. \quad (4.4)$$

We observe that the region  $\Omega_s \cap a(t)\Omega_s$  is contained in two sectors of angle

$$\alpha \leq 2 \sin^{-1}(s^2/t)$$

(see Figure 1). We take  $\theta_m = 2\pi/m$  such that  $\theta_{m+1} \leq \alpha < \theta_m$ , and consider the partition

$$\mathbb{R}^2 \setminus \{0\} = \bigsqcup_{i=1}^m S_i \quad (4.5)$$

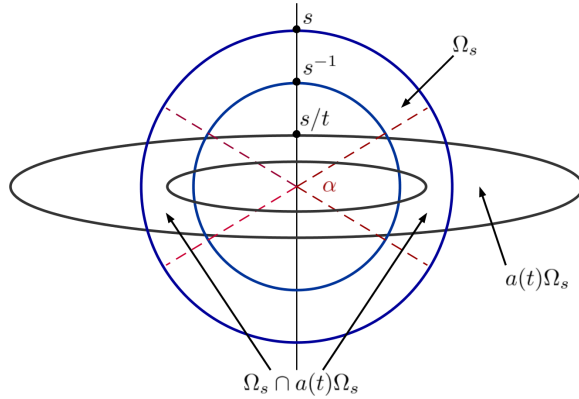


FIGURE 1. Estimating the angle  $\alpha$ .

into sectors such that

$$\Omega_s \cap a(t)\Omega_s \subset S_1. \quad (4.6)$$

We note that  $k_{\theta_m}(S_i) = S_{(i+1) \bmod m}$ . Now we suppose that  $v$  is an eigenvector of  $K_S$ , that is,  $\pi(k_{\theta})v = e^{i\lambda\theta}v$  for some  $\lambda \in \mathbb{R}$ . Then using (4.3), we deduce that

$$\pi(k_{\theta_m})P_{S_i}v = P_{k_{\theta_m}S_i}\pi(k_{\theta_m})v = e^{i\lambda\theta_m}P_{S_{(i+1) \bmod m}}v.$$

Hence, it follows that

$$\|P_{S_i}v\| = \|P_{S_{(i+1) \bmod m}}v\|.$$

For (4.5), we obtain the orthogonal decomposition

$$v = P_{\mathbb{R}^2 \setminus \{0\}}v = \sum_{i=1}^m P_{S_i}v,$$

so that

$$\|v\|^2 = \sum_{i=1}^m \|P_{S_i}v\|^2 \quad \text{and} \quad \|P_{S_i}v\| = m^{-1/2}\|v\|.$$

It follows from the inclusion (4.6) that

$$\|P_{\Omega_s \cap a(t)\Omega_s}v\| \leq \|P_{S_1}v\| \ll (\sin^{-1}(s^2/t))^{1/2}\|v\| \ll_s t^{-1/2}\|v\|.$$

A similar argument also gives that when  $w$  is an eigenvector of  $K_S$ ,

$$\|P_{a(t)^{-1}\Omega_s \cap \Omega_s}w\| \ll_s t^{-1/2}\|w\|.$$

Finally, we conclude from (4.4) that

$$|\langle \pi(a(t))v, w \rangle| \ll_s t^{-1}\|v\|\|w\|,$$

which proves the proposition.  $\square$

The above argument has been generalised by Konstantoulas [53] to give bounds on correlations of higher orders.

Proposition 4.2 gives the best possible bound in terms of the parameter  $t$ , but the drawback is that the dependence on the vectors  $v, w$  is not explicit. Our goal will be to derive a more explicit estimate for matrix coefficients. We observe that Proposition 4.2 implies an integrability estimate on the functions  $g \mapsto \langle \pi(g)v, w \rangle$ ,  $g \in G$ . This will eventually allow us to reduce our study to the case of the regular representation. We recall that the invariant measure on  $S = \mathrm{SL}_2(\mathbb{R})$  with respect to the Cartan decomposition  $S = K_S A_S K_S$  is given by

$$\int_S f(g) dg = \int_{[0, 2\pi) \times [1, \infty) \times [0, 2\pi)} f(k(\theta_1)a(t)k(\theta_2))(t^2 - t^{-2}) d\theta_1 \frac{dt}{t} d\theta_2$$

for  $f \in C_c(G)$ . Hence, from Proposition 4.2, we deduce that

**Corollary 4.3.** *With the notation as in Proposition 4.2, for vectors  $v, w$  belonging to a dense  $K_S$ -invariant subspace of  $\mathcal{H}$ , the functions  $g \mapsto \langle \pi(g)v, w \rangle$ ,  $g \in S$ , are  $L^{2+\epsilon}$ -integrable for all  $\epsilon > 0$ .*

We say that a unitary representation  $\pi$  of a group  $G$  is  $L^p$ -integrable if the matrix coefficients  $g \mapsto \langle \pi(g)v, w \rangle$ ,  $g \in G$ , belong to  $L^p(G)$  for vectors  $v, w$  from a dense subset.

The following result concerns representations of a general locally compact group  $G$ . We define the *regular representation*  $\lambda_G$  on  $L^2(G)$  by

$$\lambda_G(g)\phi(x) = \phi(xg) \quad \text{for } \phi \in L^2(G).$$

**Proposition 4.4.** *Let  $\rho$  be a unitary representation of  $G$  on a Hilbert space  $\mathcal{H}$ . We assume that the functions  $g \mapsto \langle \rho(g)v, w \rangle$  belong to  $L^2(G)$  for vectors  $v, w$  belonging to a dense subspace of  $\mathcal{H}$ . Then there exists an isometric embedding*

$$\mathcal{I} : \mathcal{H} \rightarrow \bigoplus_{n \geq 1} L^2(G)$$

such that for  $g \in G$ , we have

$$\begin{array}{ccc} \mathcal{H} & \xrightarrow{\mathcal{I}} & \bigoplus_{n \geq 1} L^2(G) \\ \rho(g) \downarrow & & \downarrow \bigoplus_{n \geq 1} \lambda_G(g) \\ \mathcal{H} & \xrightarrow{\mathcal{I}} & \bigoplus_{n \geq 1} L^2(G) \end{array}$$

*Proof.* Let  $\mathcal{H}_0$  be a countable orthonormal dense subset of  $\mathcal{H}$  such that the functions  $g \mapsto \langle v, \rho(g)w \rangle$  belong to  $L^2(G)$  for all  $v, w \in \mathcal{H}_0$ . For  $v, w \in \mathcal{H}$ , we set

$$f_{v,w}(x) = \langle v, \rho(x)w \rangle, \quad x \in G,$$

and define the map

$$\mathcal{I} : \langle \mathcal{H}_0 \rangle \rightarrow \bigoplus_{n \geq 1} L^2(G) : w \mapsto (f_{v,w} : v \in \mathcal{H}_0).$$

Since

$$\lambda_G(g)f_{v,w} = f_{v,\rho(g)w} \quad \text{for } g \in G \text{ and } v, w \in \mathcal{H},$$

we conclude that

$$\mathcal{I} \circ \rho(g) = \left( \bigoplus_{n \geq 1} \lambda_G(g) \right) \circ \mathcal{I} \quad \text{for } g \in G.$$

Moreover, since  $\mathcal{H}_0$  forms an orthonormal basis of  $\mathcal{H}$ ,

$$\sum_{v \in \mathcal{H}_0} \|f_{v,w}\|_2^2 = \|w\|^2 \quad \text{for } w \in \mathcal{H}_0,$$

it follows that

$$\|\mathcal{I}w\| = \|w\| \quad \text{for } w \in \mathcal{H}_0.$$

Hence, one can check that  $\mathcal{I}$  extends to an isometric embedding, as required.  $\square$

Although Proposition 4.4 can not be directly applied to the representation  $\pi$  appearing in Proposition 4.2, we deduce the following corollary about its tensor square  $\pi \otimes \pi$ .

**Corollary 4.5.** *Let  $\pi$  be a unitary representation of  $L$  as in Proposition 4.2. Then the representation  $(\pi \otimes \pi)|_S$  embeds as a subrepresentation of  $\bigoplus_{n \geq 1} \lambda_S$ .*

*Proof.* According to Corollary 4.3, the functions  $g \mapsto \langle \pi(g)v, w \rangle$ ,  $g \in S$ , are  $L^p$ -integrable for all  $p > 2$  when  $v, w$  belongs to a suitable orthonormal basis  $\mathcal{H}_0$  of  $\mathcal{H}$ . Then it follows from the Cauchy–Schwarz inequality that for  $v_1, v_2, w_1, w_2 \in \mathcal{H}_0$ , the functions

$$\langle (\pi \otimes \pi)(g)(v_1 \otimes v_2), w_1 \otimes w_2 \rangle = \langle \pi(g)v_1, w_1 \rangle \langle \pi(g)v_2, w_2 \rangle$$

belong to  $L^2(S)$ . Hence, the claim is implied by Proposition 4.4.  $\square$

The above result ultimately reduces our original problem regarding representations  $\pi|_S$  to the study of matrix coefficients for the regular representation  $\lambda_S$ . It turns out that the matrix coefficients of the latter representation can be estimated in terms of an explicit function that we now introduce. In fact, this is true for general connected semisimple Lie groups  $G$ . We recall that in this setting, there is the *Iwasawa decomposition*

$$G = UAK,$$

where  $K$  is a maximal compact subgroups of  $G$ ,  $A$  is a Cartan subgroup, and  $U$  is the subgroup generated by positive root subgroups. The invariant measure with respect to the Iwasawa decomposition is given by

$$\int_G f(g) dg = \int_{U \times A \times K} f(uak) \Delta(a) du da dk \quad \text{for } f \in C_c(G), \quad (4.7)$$

where  $\Delta$  denotes the modular function of the group  $UA$ , and  $du$ ,  $da$ , and  $dk$  denote the invariant measures on the corresponding factors. For example, for the group  $S = \mathrm{SL}_2(\mathbb{R})$ , using the notation (4.1), we have the the Iwasawa decomposition  $S = U_S A_S K_S$ , and the modular function is given by  $\Delta(a(t)) = t^{-2}$ . The product map  $U \times A \times K \rightarrow G$  defines a diffeomorphism, and for  $g \in G$ , we denote by  $u(g) \in U$ ,  $a(g) \in A$ , and  $k(g) \in K$  the unique elements such that

$$g = u(g)a(g)k(g).$$

The *Harish-Chandra function* is defined as

$$\Xi(g) = \int_K \Delta(a(kg))^{-1/2} dk \quad \text{for } g \in G.$$

It is easy to check that the function  $\Xi$  is bi- $K$ -invariant.

In the case when  $S = \mathrm{SL}_2(\mathbb{R})$ , the Harish-Chandra function can be explicitly computed as

$$\Xi(a(t)) = \frac{1}{2\pi} \int_0^{2\pi} (t^{-2} \cos^2 \theta + t^2 \sin^2 \theta)^{-1/2} d\theta.$$

Moreover, one can check for all  $\epsilon > 0$

$$\Xi(a(t)) \ll_{\epsilon} t^{-1+\epsilon} \quad \text{when } t \geq 1. \quad (4.8)$$

Surprisingly, it turns out that matrix coefficients of general  $K$ -eigenfunctions in  $L^2(G)$  can be explicitly estimated in terms of the Harish-Chandra function:

**Proposition 4.6.** *For all  $K$ -eigenfunctions  $\phi, \psi \in L^2(G)$ ,*

$$|\langle \lambda_G(g)\phi, \psi \rangle| \leq \Xi(g) \|\phi\|_2 \|\psi\|_2.$$

The following argument is a version of Herz's majoration principle [43].

*Proof of Proposition 4.6.* Replacing  $\phi$  and  $\psi$  by  $|\phi|$  and  $|\psi|$ , we may assume without loss of generality that  $\phi, \psi \geq 0$ , and the functions  $\phi$  and  $\psi$  are  $K$ -invariant. Then because of the Cartan decomposition  $G = KAK$ , it is sufficient to prove the estimate when  $g = a \in A$ . Using the decomposition of the invariant measure on  $G$  given by (4.7), we obtain that

$$\langle \lambda_G(a)\phi, \psi \rangle = \int_G \phi(ga)\psi(g) dg = \int_{U \times A \times K} \phi(ubka)\psi(ubk)\Delta(b) dudbdk.$$

Then by the Cauchy–Schwarz inequality,

$$\langle \lambda_G(a)\phi, \psi \rangle \leq \int_K \left( \int_{U \times A} \phi^2(ubka)\Delta(b)dudb \right)^{1/2} \left( \int_{U \times A} \psi^2(ubk)\Delta(b)dudb \right)^{1/2} dk.$$

Using that  $\psi$  is  $K$ -invariant, we obtain that

$$\int_{U \times A} \psi^2(ubk)\Delta(b)dudb = \int_{U \times A \times K} \psi^2(ubk)\Delta(b)dudbdk = \|\psi\|_2^2.$$

To estimate the other term, we write

$$ubka = ub \cdot u(ka)a(ka)k(ka) = ubu(ka)b^{-1} \cdot ba(ka) \cdot k(ka).$$

Since  $\phi$  is  $K$ -invariant, using the invariance of the integrals, we deduce that

$$\begin{aligned} & \int_K \left( \int_{U \times A} \phi^2(ubka)\Delta(b)dudb \right)^{1/2} dk \\ &= \int_K \left( \int_{U \times A} \phi^2(ubu(ka)b^{-1} \cdot ba(ka))\Delta(b)dudb \right)^{1/2} dk \\ &= \int_K \left( \int_{U \times A} \phi^2(u \cdot b)\Delta(ba(ka)^{-1})dudb \right)^{1/2} dk \\ &= \left( \int_K \Delta(a(ka))^{-1/2} dk \right) \left( \int_{U \times A} \phi^2(ub)\Delta(b)dudb \right)^{1/2}. \end{aligned}$$

Finally, because of the  $K$ -invariance of  $\phi$ ,

$$\int_{U \times A} \phi^2(ub)\Delta(b)dudb = \int_{U \times A \times K} \phi^2(ubk)\Delta(b)dudbdk = \|\phi\|_2^2$$

so that

$$\int_K \left( \int_{U \times A} \phi^2(ubka)\Delta(b)dudb \right)^{1/2} dk = \Xi(a)\|\phi\|_2.$$

This implies the required estimate.  $\square$

Using Proposition 4.6, we deduce our main result about representations of the group  $L = \mathrm{SL}_2(\mathbb{R}) \ltimes \mathbb{R}^2$ :

**Theorem 4.7.** *Let  $\pi$  be a unitary representation of  $L = \mathrm{SL}_2(\mathbb{R}) \ltimes \mathbb{R}^2$  on a Hilbert space  $\mathcal{H}$  such that there is no nonzero  $\mathbb{R}^2$ -invariant vectors. Then for all elements  $g \in \mathrm{SL}_2(\mathbb{R})$  and all  $\mathrm{SO}(2)$ -finite vectors  $v, w \in \mathcal{H}$ ,*

$$|\langle \pi(g)v, w \rangle| \leq d_{\mathrm{SO}(2)}(v)^{1/2} d_{\mathrm{SO}(2)}(w)^{1/2} \|v\| \|w\| \Xi(g)^{1/2}.$$

More general results giving quantitative bounds for representations of semidirect products have been established by Oh [74] and Wang [91].

In relation to Proposition 4.6, we mention that Cowling, Haagerup, and Howe [18] discovered that the bound in Proposition 4.6 holds more generally for any representation which is  $L^{2+\epsilon}$ -integrable for all  $\epsilon > 0$ :

**Theorem\* 4.8.** *Let  $G$  be a semisimple real algebraic group and  $\pi$  a unitary representation of  $G$  on a Hilbert space  $\mathcal{H}$  which is  $L^{2+\epsilon}$ -integrable for all  $\epsilon > 0$ . Then for all elements  $g \in G$  and all  $K$ -finite vectors  $v, w \in \mathcal{H}$ ,*

$$|\langle \pi(g)v, w \rangle| \leq d_K(v)^{1/2} d_K(w)^{1/2} \|v\| \|w\| \Xi(g).$$

In view of Corollary 4.3, Theorem 4.8 applies to the setting of Theorem 4.7 and implies a bound which is essentially optimal in terms of the decay rate along  $G$ .

**Theorem\* 4.9.** *Let  $\pi$  be a unitary representation of  $L = \mathrm{SL}_2(\mathbb{R}) \ltimes \mathbb{R}^2$  on a Hilbert space  $\mathcal{H}$  such that there is no nonzero  $\mathbb{R}^2$ -invariant vectors. Then for all elements  $g \in \mathrm{SL}_2(\mathbb{R})$  and all  $\mathrm{SO}(2)$ -finite vectors  $v, w \in \mathcal{H}$ ,*

$$|\langle \pi(g)v, w \rangle| \leq d_{\mathrm{SO}(2)}(v)^{1/2} d_{\mathrm{SO}(2)}(w)^{1/2} \|v\| \|w\| \Xi(g).$$

Here we only prove the weaker bound given by Theorem 4.7:

*Proof of Theorem 4.7.* First, we consider the case when  $v$  and  $w$  are eigenvectors of  $K_S = \mathrm{SO}(2)$ . We recall that by Corollary 4.5, the representation  $(\pi \otimes \pi)|_S$ , where  $S = \mathrm{SL}_2(\mathbb{R})$ , embeds as a subrepresentation of  $\bigotimes_{n \geq 1} \lambda_S$ . Hence, it follows from Proposition 4.6 that for  $g \in S$ ,

$$|\langle \pi(g)v, w \rangle| = |\langle (\pi \otimes \pi)(g)(v \otimes v), w \otimes w \rangle|^{1/2} \leq \|v\| \|w\| \Xi(g)^{1/2}.$$

In general, we write  $v$  and  $w$  as  $v = \sum_{i=1}^n v_i$  and  $w = \sum_{j=1}^m w_j$ , where  $v_i$ 's and  $w_j$ 's are orthogonal  $K_S$ -eigenvectors. Then for  $g \in S$ ,

$$\begin{aligned} |\langle \pi(g)v, w \rangle| &\leq \sum_{i=1}^n \sum_{j=1}^m |\langle \pi(g)v_i, w_j \rangle| \leq \left( \sum_{i=1}^n \|v_i\| \right) \left( \sum_{j=1}^m \|w_j\| \right) \Xi(g)^{1/2} \\ &\leq n^{1/2} \left( \sum_{i=1}^n \|v_i\|^2 \right)^{1/2} m^{1/2} \left( \sum_{j=1}^m \|w_j\|^2 \right)^{1/2} \Xi(g)^{1/2} \\ &\leq d_{K_S}(v)^{1/2} d_{K_S}(w)^{1/2} \|v\| \|w\| \Xi(g)^{1/2}. \end{aligned}$$

This proves the theorem.  $\square$

Now we can derive uniform bounds for matrix coefficients of higher-rank simple Lie groups and prove Theorem 4.1:

*Proof of Theorem 4.1.* We give a proof of the theorem for

$$G = \mathrm{SL}_d(\mathbb{R}) \supset K = \mathrm{SO}(d).$$

Because of the Cartan decomposition

$$G = KAK \quad \text{where } A = \{(a_1, \dots, a_d) : a_1, \dots, a_d > 0, a_1 \cdots a_d = 1\},$$

it is sufficient to prove this estimate when  $g = a \in A$ .

We consider the subgroup  $L = S \ltimes \mathbb{R}^2$ , where  $S = \mathrm{SL}_2(\mathbb{R})$ , embedded into the top left corner of  $G$ . It follows from Theorem 2.2 that for all  $v, w \in \mathcal{H}$ ,

$$\langle \pi(g)v, w \rangle \rightarrow 0 \quad \text{as } g \rightarrow \infty \text{ in } G.$$

In particular, it follows that there is no nonzero  $\mathbb{R}^2$ -invariant vectors. Hence, Theorem 4.7 can be applied to the representation  $\pi|_L$ . We write  $a \in A$  as  $a = a' a''$  with

$$\begin{aligned} a' &= \mathrm{diag} \left( (a_1/a_2)^{1/2}, (a_2/a_1)^{1/2}, 1, \dots, 1 \right), \\ a'' &= \mathrm{diag} \left( (a_1 a_2)^{1/2}, (a_1 a_2)^{1/2}, a_3, \dots, a_d \right). \end{aligned}$$

We note that  $a' \in A_S \subset L$  and  $a''$  commutes with  $S$ . In particular, it commutes with  $K_S = \mathrm{SO}(2)$ . This implies that the vector  $\pi(a'')v$  is  $K_S$ -finite, and

$$d_{K_S}(\pi(a'')v) \leq d_{K_S}(v) \leq d_K(v).$$

Hence, we deduce from Theorem 4.7 and (4.8) that

$$\begin{aligned} |\langle \pi(a)v, w \rangle| &= |\langle \pi(a')\pi(a'')v, w \rangle| \leq d_{K_S}(\pi(a'')v)^{1/2} d_{K_S}(w)^{1/2} \|v\| \|w\| \Xi(a')^{1/2} \\ &\ll_{\epsilon} d_K(v)^{1/2} d_K(w)^{1/2} \|v\| \|w\| \left( \frac{a_1}{a_2} \right)^{-1/4+\epsilon} \end{aligned}$$

for all  $\epsilon > 0$ . The same argument can be applied to other embeddings of  $\mathrm{SL}_2(\mathbb{R}) \times \mathbb{R}^2$  into  $\mathrm{SL}_d(\mathbb{R})$ . This gives the bound

$$|\langle \pi(a)v, w \rangle| \ll_{\epsilon} d_K(v)^{1/2} d_K(w)^{1/2} \|v\| \|w\| \left( \max_{i \neq j} \frac{a_i}{a_j} \right)^{-1/4+\epsilon}$$

for all  $\epsilon > 0$ , and proves the theorem.  $\square$

It is useful for applications to have the estimate as Theorem 4.1 in terms of Hölder norms or Sobolev norms of smooth vectors, as in the works of Moore [71], Ratner [78], and Katok and Spatzier [51]. Given a unitary representation  $\pi$  of  $G$  on a Hilbert space  $\mathcal{H}$ , one can also define an action of the Lie algebra  $\mathrm{Lie}(G)$  on a dense subspace  $\mathcal{V}$  of  $\mathcal{H}$  that satisfies

$$\pi(\mathcal{X})v = \frac{d}{dt} \pi(\exp(t\mathcal{X}))v|_{t=0} \quad \text{for } \mathcal{X} \in \mathrm{Lie}(G) \text{ and } v \in \mathcal{V}.$$

We fix an (ordered) basis  $(\mathcal{X}_1, \dots, \mathcal{X}_n)$  of the Lie algebra  $\mathrm{Lie}(G)$ . Then the Sobolev norm of order  $\ell$  is defined as

$$S_{\ell}(v)^2 = \sum_{(i_1, \dots, i_{\ell})} \|\pi(\mathcal{X}_{i_1}) \dots \pi(\mathcal{X}_{i_{\ell}})v\|^2 \quad (4.9)$$

With this notation, we prove:

**Theorem 4.10.** *Let  $G$  be a (noncompact) connected simple higher-rank matrix Lie group with finite centre and  $K$  a maximal compact subgroup of  $G$ . Then there exist  $c, \delta, \ell > 0$  such that for any unitary representation  $\pi$  of  $G$  on a Hilbert space  $\mathcal{H}$  without nonzero  $G$ -invariant vectors,*

$$|\langle \pi(g)v, w \rangle| \leq c S_{\ell}(v) S_{\ell}(w) \|g\|^{-\delta} \quad \text{for all } g \in G \text{ and all } v, w \in \mathcal{V}.$$

*Proof.* The proof will require several more advanced facts about representations of semisimple groups. We indicate how to complete the proof using these facts when  $\pi$  is irreducible. Then the general case will follow by using the integral decomposition (see, for instance, [30, Section 7.4]).

We decompose  $\mathcal{H}$  as a direct sum of irreducible representations of a maximal compact subgroup of  $K$ . This gives the decomposition

$$\mathcal{H} = \bigoplus_{\sigma \in \hat{K}} \mathcal{H}_{\sigma}, \quad (4.10)$$

where  $\hat{K}$  denotes the unitary dual of  $K$ , and  $\mathcal{H}_{\sigma}$  is the direct sum of the irreducible components isomorphic to  $\sigma$ . There is the Casimir operator  $\mathcal{C}$  of  $K$  which is the second order differential operator commuting with the action of  $K$ . It leaves each

of the subspaces  $\mathcal{H}_\sigma$  invariant. Moreover, one can deduce from the Schur Lemma that

$$\mathcal{C}|_{H_\sigma} = \lambda_\sigma \text{id}_{\mathcal{H}_\sigma}$$

for some  $\lambda_\sigma > 0$ . The eigenvalues  $\lambda_\sigma$  and the dimensions  $\dim(\sigma)$  are computed explicitly in the Representation Theory of compact groups, and one can verify that

$$\dim(\sigma) \leq \lambda_\sigma^{c_1} \quad \text{and} \quad \sum_{\sigma \in \hat{K}} \dim(\sigma)^{-c_2} < \infty$$

for some  $c_1, c_2 > 0$ . We shall also use a result of Harish-Chandra regarding admissibility of irreducible unitary representation (see, for instance, [92, Section 4.4.2]) that gives the bound

$$\dim(H_\sigma) \leq \dim(\sigma)^2.$$

Now utilising these estimates, we proceed with the proof of the theorem. We decompose the vectors with respect to the decomposition (4.10) and deduce from Theorem 4.1 that

$$\begin{aligned} |\langle \pi(g)v, w \rangle| &\leq \sum_{\sigma, \tau \in \hat{K}} |\langle \pi(g)v_\sigma, w_\tau \rangle| \\ &\ll \left( \sum_{\sigma \in \hat{K}} \dim(H_\sigma)^{1/2} \|v_\sigma\| \right) \left( \sum_{\tau \in \hat{K}} \dim(H_\tau)^{1/2} \|w_\tau\| \right) \|g\|^{-\delta}. \end{aligned}$$

Since the above sums can be estimates as

$$\begin{aligned} \sum_{\sigma \in \hat{K}} \dim(H_\sigma)^{1/2} \|v_\sigma\| &\leq \sum_{\sigma \in \hat{K}} \dim(\sigma) \lambda_\sigma^{-s} \|\pi(\mathcal{C})^s v_\sigma\| \\ &\leq \sum_{\sigma \in \hat{K}} \dim(\sigma)^{1-s/c_1} \|\pi(\mathcal{C})^s v_\sigma\| \\ &\leq \left( \sum_{\sigma \in \hat{K}} \dim(\sigma)^{-2(s/c_1-1)} \right)^{1/2} \left( \sum_{\sigma \in \hat{K}} \|\pi(\mathcal{C})^s v_\sigma\|^2 \right)^{1/2} \\ &\ll \|\pi(\mathcal{C})^s v\| \end{aligned}$$

for sufficiently large  $s$ , this implies the theorem.  $\square$

We note that if the assumption that  $G$  is of higher rank is removed, the statements of Theorems 4.1 and 4.10 are not true. Although we know from Theorem 2.2 that

$$\langle \pi(g)v, w \rangle \rightarrow 0 \quad \text{as } g \rightarrow \infty \text{ in } G,$$

there are unitary representations without invariant vectors whose matrix coefficients do not possess explicit estimates. For example, in the case of  $\text{SL}_2(\mathbb{R})$ , the complementary series representations provide examples with arbitrary slow decay rate (see, for instance, [48, Section V.3.1]). Nonetheless, it is known that for every nontrivial irreducible representation  $\pi$ , there exists  $c(\pi), \delta(\pi) > 0$  such that

$$|\langle \pi(g)v, w \rangle| \leq c(\pi) S_\ell(v) S_\ell(w) \|g\|^{-\delta(\pi)}.$$

Moreover, this bound also holds for any unitary representation  $\pi$  which is isolated from the trivial representation in the sense of the Fell topology.

Theorem 4.10 can be applied to finite-volume homogeneous spaces  $X$  of  $G$ . Indeed, a simple Fubini-type argument implies that the corresponding unitary representation of  $G$  on  $L^2(X)$  has no nonconstant invariant vectors so that the bound of Theorem 4.10 can be applied to all function in  $L_0^2(X)$ , which denotes the subspace of functions with zero integral. Even when  $G$  has rank one, one can show that the unitary representation of  $G$  on  $L_0^2(X)$  is isolated from the trivial representation (see, for instance, [4, Lemma 3]) so that the quantitative bounds on correlations hold in this case as well.

**Theorem\* 4.11.** *Let  $G$  be a (noncompact) connected simple matrix Lie group with finite centre and  $(X, \mu)$  is a probability homogeneous space of  $G$ . Then there exist  $\delta, \ell > 0$  such that for all elements  $g \in G$  and all functions  $\phi, \psi \in C_c^\infty(X)$ ,*

$$\int_X \phi(g^{-1}x)\psi(x) d\mu(x) = \left( \int_X \phi d\mu \right) \left( \int_X \psi d\mu \right) + O(S_\ell(\phi)S_\ell(\psi) \|g\|^{-\delta}).$$

Maucourant [67] used estimates on the correlations from Theorem 4.11 to prove a version of Theorem 3.2 that gives an asymptotic formula for the number of lattice points with an error term.

In conclusion, we note that in Theorems 4.1, 4.10 and 4.11, the matrix norm  $\|\cdot\|$  can be estimated in terms of a (left) invariant Riemannian metric  $d$  on  $G$  as

$$e^{c_1 d(g,e)} \leq \|g\| \leq e^{c_2 d(g,e)} \quad \text{for all } g \in G.$$

In particular, in Theorem 4.11, this gives the error term

$$O\left(S_\ell(\phi)S_\ell(\psi) e^{-\delta' d(g,e)}\right)$$

for some  $\delta' > 0$ .

## 5. BOUNDS ON HIGHER-ORDER CORRELATIONS

Building on the results from Section 4, we intend to establish quantitative estimates on correlations of arbitrary order. We follow the argument of Björklund, Einsiedler, Gorodnik [7]. Throughout this section,  $G$  denotes a (noncompact) connected simple matrix Lie group with finite center. We consider a measure preserving action of  $G$  on a standard probability space  $(X, \mu)$ . To simplify notation, we set

$$(g \cdot \phi)(x) = \phi(g^{-1}x) \quad \text{for } g \in G \text{ and } \phi \in L^\infty(X).$$

Our goal is to estimate the correlations

$$\mu((g_1 \cdot \phi_1) \cdots (g_r \cdot \phi_k)) = \int_X \phi_1(g_1^{-1}x) \cdots \phi_r(g_k^{-1}x) d\mu(x)$$

for suitable functions  $\phi_1, \dots, \phi_r$  on  $X$ . We shall assume that we know how to estimate correlations of order two. Namely, we assume that there exist a subalgebra  $\mathcal{A}$  of  $L^\infty(X)$  and  $\delta > 0$  such that for all functions  $\phi_1, \phi_2 \in \mathcal{A}$  and all  $g \in G$ ,

$$\mu((g \cdot \phi_1) \phi_2) = \mu(\phi_1)\mu(\phi_2) + O(S_\ell(\phi_1)S_\ell(\phi_2) \|g\|^{-\delta}), \quad (5.1)$$

where  $S_\ell$  denotes a norm on the algebra  $\mathcal{A}$ . The precise definition of the family of norms

$$S_1 \leq S_2 \leq \cdots \leq S_\ell \leq \cdots$$

will not be important for our arguments. We shall only use that these norms satisfy the following properties:

(N<sub>1</sub>) there exists  $\ell_1$  such that

$$\|\phi\|_{L^\infty} \ll S_{\ell_1}(\phi),$$

(N<sub>2</sub>) there exists  $\ell_2$  such that

$$\|g \cdot \phi - \phi\|_{L^\infty} \ll \|g - e\| S_{\ell_2}(\phi) \quad \text{for all } g \in G,$$

(N<sub>3</sub>) for all  $\ell$ , there exists  $\sigma_\ell > 0$  such that

$$S_\ell(g \cdot \phi) \ll_\ell \|g\|^{\sigma_\ell} S_\ell(\phi) \quad \text{for all } g \in G,$$

(N<sub>4</sub>) for every  $\ell$ , there exists  $\ell'$  such that

$$S_\ell(\phi_1 \phi_2) \ll_\ell S_{\ell'}(\phi_1) S_{\ell'}(\phi_2).$$

For instance, when  $X = L/\Gamma$  where  $L$  is a connected Lie group and  $\Gamma$  a discrete cocompact subgroup, it follows from a version of the Sobolev embedding theorem that the Sobolev norms defined in (4.9) satisfy these properties. More generally, when  $\Gamma$  is discrete subgroup of finite covolume, one can also introduce a family of norms majorating the usual Sobolev norms satisfying these properties (see, for instance, [27, §3.7]). In particular, it follows from Section 4 that the bound (5.1) holds in this setting.

The main result of this section is the following:

**Theorem 5.1.** *For every  $r \geq 2$ , there exist  $\delta_r, \ell_r > 0$  such that for all elements  $g_1, \dots, g_r \in G$  and all functions  $\phi_1, \dots, \phi_r \in \mathcal{A}$ ,*

$$\begin{aligned} \mu((g_1 \cdot \phi_1) \cdots (g_r \cdot \phi_r)) &= \mu(\phi_1) \cdots \mu(\phi_r) \\ &\quad + O_r(S_{\ell_r}(\phi_1) \cdots S_{\ell_r}(\phi_r) N(g_1, \dots, g_r)^{-\delta_r}), \end{aligned}$$

where

$$N(g_1, \dots, g_r) = \min_{i \neq j} \|g_i^{-1} g_j\|.$$

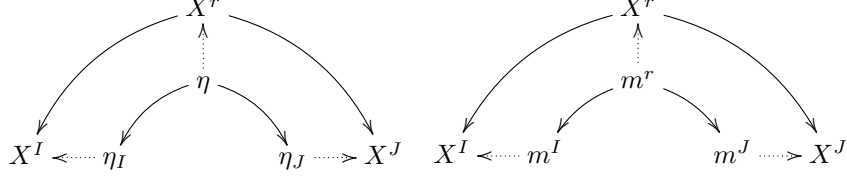
We first explain the strategy of the proof of Theorem 5.1. It will be convenient to consider the correlation of order  $r$  as a measure on the product space  $X^r$ : we introduce a measure  $\eta = \eta_{g_1, \dots, g_r}$  on  $X^r$  defined by

$$\eta(\phi) = \int_X \phi(g_1^{-1}x, \dots, g_r^{-1}x) d\mu(x) \quad \text{for } \phi \in L^\infty(X^r).$$

Theorem 5.1 amounts to showing that the measure  $\eta$  is “approximately” equal to the product measure  $\mu^r$  on  $X^r$ . Our argument will ultimately proceed by induction on the number of factors  $r$ . For now, we keep  $r$  fixed and introduce a nontrivial partition  $\{1, \dots, r\} = I \sqcup J$  which defines the projection maps  $X^r \rightarrow X^I$  and  $X^r \rightarrow X^J$ . We obtain the measures  $\eta_I$  and  $\eta_J$  on  $X^I$  and  $X^J$  respectively that are the projections of the measure  $\eta$ . Formally, these measure are defined as

$$\begin{aligned} \eta_I(\phi) &= \eta(\phi \otimes 1) \quad \text{for } \phi \in L^\infty(X^I), \\ \eta_J(\psi) &= \eta(1 \otimes \psi) \quad \text{for } \psi \in L^\infty(X^J). \end{aligned}$$

Ultimately, our proof will involve comparing the diagrams:



We may assume by induction that

$$\eta_I \approx \mu^I \quad \text{and} \quad \eta_J \approx \mu^J$$

and need to show that

$$\eta \approx \mu^r = \mu^I \otimes \mu^J. \quad (5.2)$$

To establish this estimate, we use that the measure  $\eta$  is invariant under the subgroup

$$D = \{(g_1^{-1}hg_1, \dots, g_r^{-1}hg_r) : h \in G\}.$$

We take a one-parameter subgroup

$$h(t) = (\exp(tZ_1), \dots, \exp(tZ_r))$$

in  $D$  which also can be written as

$$h(t) = (h_I(t), h_J(t))$$

for one-parameter subgroups acting on  $X^I$  and  $X^J$ . We note that the measure  $\eta_I$  is  $h_I(t)$ -invariant, and the measure  $\eta_J$  is  $h_J(t)$ -invariant. We consider the averaging operator

$$P_T : L^\infty(X^I) \rightarrow L^\infty(X^I) : \phi \mapsto \frac{1}{T} \int_0^T \phi(h_I(t)x) dt \quad (5.3)$$

that preserves the measure  $\eta_I$ . Given functions  $\phi_1, \dots, \phi_r \in L^\infty(X)$ , we write

$$\phi_I = \otimes_{i \in I} \phi_i \quad \text{and} \quad \phi_J = \otimes_{i \in J} \phi_i.$$

To simplify notation, we write  $S_\ell(\phi_I) = \prod_{i \in I} S_\ell(\phi_i)$  below. We establish (5.2) using the following key estimate

$$\begin{aligned} |\eta(\phi_I \otimes \phi_J) - \mu^r(\phi_1 \otimes \dots \otimes \phi_r)| &= |\eta(\phi_I \otimes \phi_J) - \mu^I(\phi_I)\mu^J(\phi_J)| & (5.4) \\ &\leq |\eta(\phi_I \otimes \phi_J) - \eta(P_T\phi_I \otimes \phi_J)| & \text{(I)} \\ &\quad + |\eta(P_T\phi_I \otimes \phi_J) - \eta_I(\phi_I)\eta_J(\phi_J)| & \text{(II)} \\ &\quad + |\eta_I(\phi_I)\eta_J(\phi_J) - \mu_I(\phi_I)\mu_J(\phi_J)|. & \text{(III)} \end{aligned}$$

These inequalities provide an approximate decoupling of  $\eta(\phi_I \otimes \phi_J)$  into the product  $\eta_I(\phi_I)\eta_J(\phi_J)$  at the cost of introducing the block correlations  $\eta(P_T\phi_I \otimes \phi_J)$  that we will control inductively. We will estimate the terms (I), (II), and (III) separately for a carefully chosen partition  $\{I, J\}$  and a carefully chosen one-parameter subgroup  $h(t)$ . To simplify notation, we shall assume that for all  $i = 1, \dots, r$ ,

$$S_{\ell'}(\phi_i) \leq 1$$

for a fixed sufficiently large  $\ell'$ . In particular, it also follows from properties (N<sub>1</sub>) and (N<sub>2</sub>) of the norms that for all  $i = 1, \dots, r$  and  $g \in G$ ,

$$\|\phi_i\|_{L^\infty} \ll 1 \quad \text{and} \quad \|g \cdot \phi_i - \phi_i\|_{L^\infty} \ll \|g - e\|. \quad (5.5)$$

It will be convenient to replace the matrix norm by a different norm defined in terms of the adjoint representation  $\text{Ad} : G \mapsto \text{GL}(\text{Lie}(G))$  with

$$\text{Ad}(g) : X \mapsto gXg^{-1} \quad \text{for } X \in \text{Lie}(G).$$

We fix a norm on the Lie algebra  $\text{Lie}(G)$  and set

$$\|g\| = \max \{ \|\text{Ad}(g)Z\| : \|Z\| = 1 \}.$$

One can arrange that for any  $g \in G$ ,

$$\|g\| \geq 1 \quad \text{and} \quad \|g\| = \|\text{Ad}(g)Z\| \quad \text{for some nilpotent } Z \text{ with } \|Z\| = 1.$$

This might require to modify the original choice of the norm (see [7, Section 2.1]), but we omit these details here.

Now we describe the choice of the one-parameter  $h(t)$  subgroup that we use. Let

$$Q = \max_{i \neq j} \|g_i^{-1}g_j\| \quad \text{and} \quad q = \min_{i \neq j} \|g_i^{-1}g_j\| \geq 1$$

We take  $i_1 \neq i_s$  such that

$$Q = \|g_{i_1}^{-1}g_{i_s}\| = \|\text{Ad}(g_{i_1}^{-1}g_{i_s})Z\|$$

for some nilpotent  $Z$  with  $\|Z\| = 1$ . Then

$$\|\text{Ad}(g_{i_1}^{-1}g_{i_s})Z\| \geq \|\text{Ad}(g_i^{-1}g_j)Z\| \quad \text{for all } i \neq j.$$

For a suitable choice of indices, we obtain that

$$\|\text{Ad}(g_{i_1}^{-1}g_{i_s})Z\| \geq \|\text{Ad}(g_{i_2}^{-1}g_{i_s})Z\| \geq \cdots \geq \|\text{Ad}(g_{i_r}^{-1}g_{i_s})Z\|.$$

In fact, after relabelling, we may assume that

$$\|\text{Ad}(g_1^{-1}g_s)Z\| \geq \|\text{Ad}(g_2^{-1}g_s)Z\| \geq \cdots \geq \|\text{Ad}(g_r^{-1}g_s)Z\|.$$

We note that

$$\|\text{Ad}(g_r^{-1}g_s)Z\| \leq \|\text{Ad}(g_s^{-1}g_s)Z\| = 1.$$

We set

$$Z_j = \frac{\text{Ad}(g_j^{-1}g_s)Z}{\|\text{Ad}(g_1^{-1}g_s)Z\|} \quad \text{and} \quad w_j = \|Z_j\|.$$

Then

$$1 = w_1 \geq w_2 \geq \cdots \geq w_r \quad \text{and} \quad w_r \leq Q^{-1} \leq q^{-1} \quad (5.6)$$

We take

$$I = \{1, \dots, p\} \quad \text{and} \quad J = \{p+1, \dots, r\},$$

where the index  $p$  will be specified later.

We observe that with these choices, the one-parameter subgroups  $h_I(t)$  and  $h_J(t)$  satisfy the following properties with some exponents  $a, b > 0$ ,

- (a)  $\|h_J(t) \cdot \phi_J - \phi_J\|_{L^\infty} \ll w_{p+1}|t|$ ,
- (b)  $S_\ell(h_I(t) \cdot \phi_I) \ll \max(1, |t|)^a$ ,
- (c)  $|\mu^I((h_I(t) \cdot \phi_I)\phi_I) - \mu^I(\phi_I)^2| \ll \max(1, w_p|t|)^{-b}$ .

Indeed, (a) can be deduced from the property (N<sub>2</sub>) of the Sobolev norms, (b) — from the property (N<sub>3</sub>), and (c) — from the bound (5.1) on correlations of order two.

Now we proceed to estimate (5.4). Our argument proceeds by induction on  $r$ , and we suppose that we have established existence of  $E = E(g_1, \dots, g_r)$  such that for all proper subsets  $L$  of  $\{1, \dots, r\}$  and functions  $\psi_1, \dots, \psi_r \in \mathcal{A}$ ,

$$|\eta_L(\psi_L) - \mu^L(\psi_L)| \leq E S_\ell(\psi_L). \quad (5.7)$$

We estimate each of the terms (I), (II), (III) appearing in (5.4). We note that it follows immediately from the assumption (5.7) and (5.5) that

$$|\eta_I(\phi_I)\eta_J(\phi_J) - \mu^I(\phi_I)\mu^J(\phi_J)| \ll E. \quad (5.8)$$

This provides an estimate for the term (III).

To estimate the term (I), we observe that

$$\eta(P_T\phi_I \otimes \phi_J) = \eta\left(\frac{1}{T} \int_0^T (h_I(t) \cdot \phi_I) \otimes \phi_J dt\right).$$

Using that the measure  $\eta$  is invariant under  $h(t) = (h_I(t), h_J(t))$ , we obtain that

$$\eta(\phi_I \otimes \phi_J) = \eta\left(\frac{1}{T} \int_0^T (h_I(t) \cdot \phi_I) \otimes (h_J(t) \cdot \phi_J) dt\right).$$

Hence, the term (I) can be estimated as

$$\begin{aligned} & |\eta(\phi_I \otimes \phi_J) - \eta(P_T\phi_I \otimes \phi_J)| \\ & \leq \eta\left(\frac{1}{T} \int_0^T |(h_I(t) \cdot \phi_I) \otimes (h_J(t) \cdot \phi_J) - (h_I(t) \cdot \phi_I) \otimes \phi_J| dt\right) \\ & \leq \frac{1}{T} \int_0^T \|(h_I(t) \cdot \phi_I) \otimes (h_J(t) \cdot \phi_J - \phi_J)\|_{L^\infty} dt \\ & \leq \|\phi_I\|_{L^\infty} \cdot \max_{0 \leq t \leq T} \|h_J(t) \cdot \phi_J - \phi_J\|_{L^\infty} \ll w_{p+1}T, \end{aligned} \quad (5.9)$$

where we used (5.5) and (a).

To estimate the term (II), we use that

$$\eta_I(\phi_I)\eta_J(\phi_J) = \eta_I(\phi_I)\eta(1 \otimes \phi_J) = \eta(\eta_I(\phi_I) \otimes \phi_J).$$

We first show that the term (II) can be estimated in terms of the quantity

$$D_T(\eta_I) = \eta_I(|P_T\phi_I - \eta_I(\phi_I)|^2)^{1/2}.$$

Indeed, we obtain that

$$\begin{aligned} |\eta(P_T\phi_I \otimes \phi_J) - \eta_I(\phi_I)\eta_J(\phi_J)| &= |\eta((P_T\phi_I - \eta_I(\phi_I)) \otimes \phi_J)| \\ &\leq \eta(|P_T\phi_I - \eta_I(\phi_I)| \otimes |\phi_J|) \\ &\leq \eta(|P_T\phi_I - \eta_I(\phi_I)|) \|\phi_J\|_{L^\infty} \\ &\leq D_T(\eta_I) \end{aligned}$$

by (5.5) and the Cauchy–Schwarz inequality. To deal with  $D_T(\eta_I)$ , we use that it can be approximated by

$$D_T(\mu^I) = \mu^I(|P_T\phi_I - \mu^I(\phi_I)|^2)^{1/2}.$$

The corresponding estimate is given by:

**Lemma 5.2.**  $|D_T(\eta_I) - D_T(\mu^I)| \ll T^{a/2}E^{1/2}$ .

*Proof.* Using the inequality  $|x - y| \leq \sqrt{|x^2 - y^2|}$  with  $x, y \geq 0$ , we obtain that

$$|D_T(\eta_I) - D_T(\mu^I)| \leq \sqrt{|D_T(\eta_I)^2 - D_T(\mu^I)^2|}.$$

Expanding the averaging operator (5.3) and changing the order of integration, we deduce that

$$\begin{aligned} D_T(\eta_I)^2 &= \int_{X^I} |P_T \phi_I - \eta_I(\phi_I)|^2 d\eta_I \\ &= \frac{1}{T^2} \int_0^T \int_0^T (\eta_I((h_I(s-t) \cdot \phi_I)\phi_I) - \eta_I(\phi_I)^2) dsdt. \end{aligned}$$

Similarly,

$$D_T(\mu^I)^2 = \frac{1}{T^2} \int_0^T \int_0^T (\mu^I((h_I(s-t) \cdot \phi_I)\phi_I) - \mu^I(\phi_I)^2) dsdt.$$

Hence,

$$\begin{aligned} &|D_T(\eta_I)^2 - D_T(\mu^I)^2| \\ &\leq \frac{1}{T^2} \int_0^T \int_0^T \left( |\eta_I((h_I(s-t) \cdot \phi_I)\phi_I) - \mu^I((h_I(s-t) \cdot \phi_I)\phi_I)| \right. \\ &\quad \left. + |\eta_I(\phi_I)^2 - \mu^I(\phi_I)^2| \right) dsdt. \end{aligned}$$

The first term inside the integral is estimated using (5.7) as

$$\begin{aligned} &\ll E S_\ell((h_I(s-t) \cdot \phi_I)\phi_I) \ll E S_{\ell'}((h_I(s-t) \cdot \phi_I)S_{\ell'}(\phi_I)) \\ &\ll E \max(1, |s-t|)^a, \end{aligned}$$

where we used (N<sub>4</sub>), (b), and (5.5). The second term inside the integral is estimated using (5.7) and (5.5) as

$$= |\eta_I(\phi_I) - \mu^I(\phi_I)| \cdot |\eta_I(\phi_I) + \mu^I(\phi_I)| \leq E \cdot 2\|\phi_I\|_{L^\infty} \ll E.$$

Finally, the lemma follows from the bound

$$\frac{1}{T^2} \int_0^T \int_0^T \max(1, |s-t|)^a dsdt \ll T^a.$$

□

It follows from (c) that

$$\begin{aligned} D_T(\mu^I)^2 &= \frac{1}{T^2} \int_0^T \int_0^T (\mu^I((h_I(s-t) \cdot \phi_I)\phi_I) - \mu^I(\phi_I)^2) dsdt \\ &\ll \frac{1}{T^2} \int_0^T \int_0^T \max(1, w_p|s-t|)^{-b} dsdt \ll (w_p T)^{-b}. \end{aligned}$$

Hence, we conclude from Lemma 5.2 that the term (II) can be estimated as

$$|\eta(P_T \phi_I \otimes \phi_J) - \eta_I(\phi_I)\eta_J(\phi_J)| \ll \max(T^{a/2} E^{1/2}, (w_p T)^{-b/2}). \quad (5.10)$$

Now combining the bounds (5.9), (5.10), and (5.8), we deduce from (5.4) that for all  $T \geq 1$ ,

$$|\eta(\phi_I \otimes \phi_J) - \mu^I(\phi_I)\mu^J(\phi_J)| \ll \max(w_{p+1}T, T^{a/2} E^{1/2}, (w_p T)^{-b/2}).$$

This estimate will be used to complete the proof of Theorem 5.1 by induction on  $r$ . We suppose that (5.7) holds with  $E = q^{-\tau}$  for some  $\tau > 0$ . We have to pick the index  $p$  and the parameter  $T$  to minimise

$$\max(w_{p+1}T, T^{a/2}q^{-\tau/2}, (w_pT)^{-b/2}). \quad (5.11)$$

We seek a bound which is uniform on the parameters  $w_1, \dots, w_r$  satisfying (5.6). We take  $\theta > 0$  with  $\theta < (r-1)^{-1}$ . Then since  $w_r \leq q^{-1}$ , all the  $r$  points

$$1, q^{-\theta}, \dots, q^{-(r-1)\theta}$$

are contained in the union of  $r-1$  intervals

$$[w_r, w_{r-1}], \dots, [w_2, w_1].$$

Hence, by the Pigeonhole Principle, there exist  $p$  and  $i$  such that

$$w_{p+1} \leq q^{-(i+1)\theta} < q^{-i\theta} \leq w_p.$$

Taking  $T = q^{(i+1/2)\theta}$ , we obtain that (5.11) is estimated by  $q^{-\tau'}$  with  $\tau' > 0$ . This completes the proof of Theorem 5.1.

## 6. APPLICATION: EXISTENCE OF CONFIGURATIONS

Analysis of higher-order correlation can be used to study existence of combinatorial configurations. Perhaps, the most striking example of this is the Szemerédi theorem [87] which states that any subset of integers of positive upper density contains arbitrary long arithmetic progressions. Furstenberg [31] discovered that this problem can be modelled using dynamical systems. His approach is based on the ‘‘Furstenberg Correspondence Principle’’ which associates to a subset of positive density in  $\mathbb{Z}$  a shift-invariant measure on the space  $\{0, 1\}^{\mathbb{Z}}$ . The crux of Furstenberg’s proof [31] of the Szemerédi theorem is the following result which implies nonvanishing of higher-order correlations.

**Theorem\* 6.1.** *Let  $T : X \rightarrow X$  be a measure-preserving transformation of a probability space  $(X, \mu)$ . Then for every nonnegative  $\phi \in L^\infty(X)$  which is not zero almost everywhere,*

$$\liminf_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^{N-1} \int_X \phi(x) \phi(T^i x) \cdots \phi(T^{(r-1)i} x) d\mu(x) > 0.$$

This result was generalised by Furstenberg and Katznelson [32] to systems of commuting transformations which allowed to prove the following generalisation of the Szemerédi theorem.

**Theorem\* 6.2.** *Let  $\Omega$  be a subset of  $\mathbb{Z}^d$  of positive upper density. Then for any finite subset  $F$  of  $\mathbb{Z}^d$ , there exist  $a \in \mathbb{Z}^d$  and  $t \in \mathbb{N}$  such that*

$$a + tF \subset \Omega.$$

We recall that a set  $\Omega$  has *positive upper density* if there exists a sequence of boxes  $B_n$  with lengths of all sides going to infinity such that

$$\limsup_{n \rightarrow \infty} \frac{|\Omega \cap B_n|}{|B_n|} > 0.$$

Existence of configurations in subsets of the Euclidean space  $\mathbb{R}^d$  has been also extensively studied. The following results was proved by Furstenberg, Katznelson,

and Weiss [33] for  $d = 2$  using ergodic-theoretic techniques and by Bourgain [12] in general using Fourier analysis.

**Theorem\* 6.3.** *Let  $\Omega$  be a subset of positive density in  $\mathbb{R}^d$ , and  $F = \{0, x_1, \dots, x_{d-1}\}$  is a subset of points in  $\mathbb{R}^d$  in general position. Then there exists  $t_0$  such that for every  $t \geq t_0$ , the set  $\Omega$  contains an isometric copy of  $tF$ .*

It was shown by Bourgain [12] and Graham [41] that an analogue of this theorem fails for general configurations. Nonetheless, one may ask whether the set  $\Omega$  contains approximate configurations. This was settled by Furstenberg, Katznelson, and Weiss [33] by configurations of three points and by Ziegler [96] in general:

**Theorem\* 6.4.** *Let  $\Omega$  be a subset of positive density in  $\mathbb{R}^d$  and  $x_1, \dots, x_{r-1} \in \mathbb{R}^d$ . Then there exists  $t_0$  such that for every  $t \geq t_0$  and  $\epsilon > 0$ , one can find  $(y_0, y_1, \dots, y_{r-1}) \subset \Omega^r$  and an isometry of  $I$  of  $\mathbb{R}^d$  such that*

$$d(0, I(y_0)) < \epsilon \quad \text{and} \quad d(tx_i, I(y_i)) < \epsilon \quad \text{for } i = 1, \dots, r-1.$$

The proof of Theorem 6.4 requires more detailed analysis of the averages of correlations

$$\frac{1}{N} \sum_{i=0}^{N-1} \int_X \phi_0(x) \phi_1(T^i x) \cdots \phi_{r-1}(T^{(r-1)i} x) d\mu(x) \quad (6.1)$$

for  $\phi_0, \dots, \phi_{r-1} \in L^\infty(X)$ . While the case when  $r = 3$  can be reduced to investigating translations on compact abelian groups. The general case have presented a significant challenge that was solved in the ground-breaking works of Host and Kra [45], and Ziegler [97]. These works developed a comprehensive method that allowed to understand limits in  $L^2(X)$  of the averages (6.1). It turns out that this reduces to analysing this limit for the so-called characteristic factors which are shown to be inverse limits of dynamical systems which are translations on nilmanifolds. Thus, remarkably to investigate the general limits of the averages (6.1) it suffices to deal with these limits for nilsystems. We also mention that Leibman [56] and Ziegler [95] established existence of the limit of (6.1) for translations on nilmanifolds.

More generally, let us consider a locally compact group  $G$  equipped with a left-invariant metric. Given a “large” subset of  $G$ , we would like to show that it approximately contains an isometric copy of a given configuration  $(g_1, \dots, g_r) \in G^r$ . It is not clear what a natural notion of largeness in  $G$  is, especially when the group  $G$  is not amenable. In any case, one definitely views a subgroup  $\Gamma$  in  $G$  with finite covolume as being “large”. We will be interested in investigated how rich the set of configurations  $(\gamma_1, \dots, \gamma_r) \in \Gamma^r$  is. In particular, one may wonder whether general configurations  $(g_1, \dots, g_r) \in G$  can be approximated by isometric copies of the configurations  $(\gamma_1, \dots, \gamma_r) \in \Gamma^r$  (see Figure 2), namely, whether for every  $\epsilon > 0$ , there exists an isometry  $I : G \rightarrow G$  such that

$$d(g_i, I(\gamma_i)) < \epsilon \quad \text{for } i = 1, \dots, r. \quad (6.2)$$

It was observed by Björklund, Einsiedler, and Gorodnik [7] that the estimates on higher-order correlations (Theorem 5.1) can be used to solve this problem in an optimal way when  $G$  is a connected simple Lie group with finite center, and  $\Gamma$  is a discrete subgroup of  $G$  with finite covolume. It is clear that since  $\Gamma$  is discrete, the approximation (6.2) can not hold when the points  $g_i$  are not too

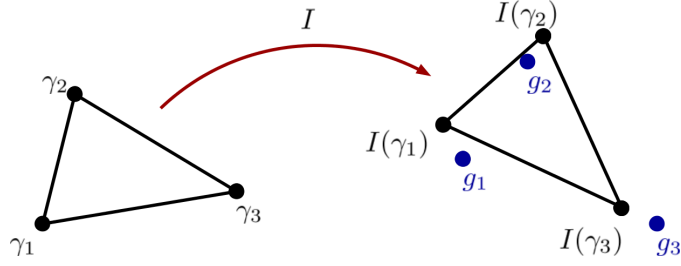


FIGURE 2. Existence of approximate configurations.

“clustered” together. To address this issue, we introduce the notion of *width*: for  $(g_1, \dots, g_r) \in G^r$ , we set

$$w(g_1, \dots, g_r) = \min_{i \neq j} d(g_i, g_j).$$

We shall show that (6.2) can be established provided that the points are sufficiently spread out in terms of  $\epsilon$ .

**Theorem 6.5.** *For every  $r \geq 2$ , there exist  $c_r, \epsilon_r > 0$  such that for all tuples  $(g_1, \dots, g_r) \in G^r$  satisfying*

$$w(g_1, \dots, g_r) \geq c_r \log(1/\epsilon) \quad \text{with } \epsilon \in (0, \epsilon_r),$$

*there exists a tuples  $(\gamma_1, \dots, \gamma_r) \in \Gamma^r$  and  $g \in G$  such that*

$$d(g_i, g \cdot \gamma_i) < \epsilon \quad \text{for } i = 1, \dots, r.$$

Let us illustrate Theorem 6.5 by an example of the orbit  $\Gamma \cdot i$  in the hyperbolic plane  $\mathbb{H}^2$  for  $\Gamma = \text{PSL}_2(\mathbb{Z})$ . For  $g \in \text{PSL}_2(\mathbb{R})$ ,

$$d(gi, i) = \cosh^{-1}(\|g\|^2/2)$$

where  $\|\cdot\|$  is the Euclidean norm. In this case, Theorem 6.5 with  $r = 2$  reduces to showing that any distance  $D > 0$  can be approximated by distances from the set

$$\Delta = \{\cosh^{-1}((a^2 + b^2 + c^2 + d^2)/2) : a, b, c, d \in \mathbb{Z}^4, ad - bc = 1\}.$$

Namely, when  $D \geq c_2 \log(1/\epsilon)$ , there exists  $\delta \in \Delta$  such that  $|D - \delta| < \epsilon$ . On the other hand, one can show that the set  $\Delta$  is not  $\epsilon$ -dense in an interval  $[a_\epsilon, \infty)$  with  $a_\epsilon = o(\log(1/\epsilon))$  as  $\epsilon \rightarrow 0^+$ .

*Proof of Theorem 6.5.* We consider the action of  $G$  on the space  $X = G/\Gamma$  equipped with the normalised invariant measure  $\mu$  and apply Theorem 5.1 to a suitably chosen family of test functions supported on  $X$ . We take nonnegative  $\tilde{\phi}_\epsilon \in C_c^\infty(G)$  such that

$$\text{supp}(\tilde{\phi}_\epsilon) \subset B_\epsilon(e), \quad \mu(\tilde{\phi}_\epsilon) = 1, \quad S_\ell(\tilde{\phi}_\epsilon) \ll \epsilon^{-\alpha},$$

for some fixed  $\alpha > 0$  depending only on  $\ell$  and  $G$ . Such a family of function can be constructed using a local coordinate system in a neighbourhood of identity in  $G$ . We set

$$\phi_\epsilon(g\Gamma) = \sum_{\gamma \in \Gamma} \tilde{\phi}_\epsilon(g\gamma), \quad g \in G,$$

which defines a function in  $C_c^\infty(G/\Gamma)$ . Then Theorem 5.1 give that

$$\mu((g_1 \cdot \phi_\epsilon) \cdots (g_r \cdot \phi_\epsilon)) = 1 + O_r\left(e^{-\delta w(g_1, \dots, g_r)} \epsilon^{-\alpha r}\right) = 1 + O_r\left(\epsilon^{c_r \delta - \alpha r}\right).$$

If we take  $c_r > \alpha r / \delta$ , then it follows from this estimate that for all sufficiently small  $\epsilon$ ,

$$\mu((g_1 \cdot \phi_\epsilon) \cdots (g_r \cdot \phi_\epsilon)) > 0.$$

Since

$$\mu((g_1 \cdot \phi_\epsilon) \cdots (g_r \cdot \phi_\epsilon)) = \int_{G/\Gamma} \left( \sum_{\gamma_1, \dots, \gamma_r \in \Gamma} \tilde{\phi}_\epsilon(g_1^{-1}g\gamma_1) \cdots \tilde{\phi}_\epsilon(g_r^{-1}g\gamma_r) \right) d\mu(g\Gamma),$$

it follows that there exist  $(\gamma_1, \dots, \gamma_r) \in \Gamma^r$  and  $g \in G$  such that

$$g_i^{-1}g\gamma_i \in \text{supp}(\tilde{\phi}_\epsilon) \subset B_\epsilon(e) \quad \text{for } i = 1, \dots, r,$$

so that

$$d(g_i, g\gamma_i) = d(g_i^{-1}g\gamma_i, e) < \epsilon \quad \text{for } i = 1, \dots, r,$$

as required.  $\square$

## 7. APPLICATION: CENTRAL LIMIT THEOREM

Suppose that the time evolution of a physical system is given by a one-parameter flow  $T_t : X \rightarrow X$  on the phase space  $X$ . Observables of this system are represented by functions  $\phi$  on  $X$  so that studying the transformation of this system as time progresses involves the analysis of the values  $\phi(T_t x)$  with  $t \geq 0$  and  $x \in X$ . Often these values fluctuate quite erratically which makes it difficult to understand them in deterministic terms. Instead, one might attempt to study their statistical properties. Formally, we consider  $\{\phi \circ T_t : t \geq 0\}$  as a family of random variables on  $X$ . For chaotic flows, these family typically exhibits quasi-independence properties, and it is natural to expect that they satisfy probabilistic limit laws known for independent random variables. For instance, we mention one of the first results in this direction which was proved by Sinai [84]:

**Theorem 7.1.** *Let  $g_t : T^1(M) \rightarrow T^1(M)$  be the geodesic flow on a compact manifold  $M$  with constant negative curvature. Then for any  $\phi \in C^{1+\alpha}(X)$  with zero integral, the family of functions*

$$F_t(x) = t^{-1/2} \int_0^t \phi(g_s x) ds$$

*converges in distribution to the Normal Law as  $t \rightarrow \infty$ ; that is, for all  $\xi \in \mathbb{R}$ ,*

$$\frac{\text{vol}(\{x \in T^1(M) : F_t(x) < \xi\})}{\text{vol}(T^1(M))} \rightarrow \text{Norm}_\sigma(\phi)(\xi) \quad \text{as } t \rightarrow \infty,$$

*where*

$$\text{Norm}_\sigma(\xi) = (\sqrt{2\pi}\sigma)^{-1} \int_{-\infty}^{\xi} e^{-s^2/(2\sigma^2)} ds$$

*denotes the Normal Distribution with variance  $\sigma$ .*

Validity of the Central Limit Theorem for one-parameter dynamical systems has been extensively studied in the last decades, and we refer to surveys [24, 23, 40, 54, 88] for an introduction to this vast area of research. However, there was very little known about the distribution of averages for more general groups actions. In this section, we present a method developed by Björklund and Gorodnik [8] for proving the Central Limit Theorem, which is based on the quantitative estimates for higher-order correlations established in the previous sections.

Let us consider an action of a group  $H$  on a standard probability space  $(X, \mu)$ . Given a function  $\phi$  on  $X$ , we consider the family of its translations

$$(h \cdot \phi)(x) = \phi(h^{-1}x) \quad \text{with } h \in H.$$

One may think about  $\{h \cdot \phi : h \in H\}$  as a collection of identically distributed random variables on the probability space  $(X, \mu)$ . When the action exhibits chaotic behaviour, it is natural to expect that these random variables are quasi-independent in a suitable sense which leads to the question whether these random variables satisfy analogues of the standard probabilistic laws such as, for instance, the Central Limit Theorem, the Law of Iterated Logarithms, etc.

Here we prove a general Central Limit theorem for group actions. From the perspective of this notes, the chaotic nature of group actions is reflected in the asymptotic behaviour of the higher-order correlations. We demonstrate that quantitative estimates on correlations imply the Central Limit Theorem. Although we do not pursue this direction here, we mention that this approach has found interesting applications in Number Theory to study the distribution of arithmetic counting functions (see [9, 10]).

Let  $H$  be a (noncompact) locally compact group  $H$  equipped with a left-invariant metric  $d$ . We consider a measure-preserving action of  $H$  on a standard probability space  $(X, \mu)$ . We assume that this action is mixing of all orders in the following quantitative sense. There exists a subalgebra  $\mathcal{A}$  of  $L^\infty(X)$  equipped with a family of norms

$$S_1 \leq S_2 \leq \dots \leq S_\ell \leq \dots$$

satisfying the following properties:

(N<sub>1</sub>) there exists  $\ell_1$  such that

$$\|\phi\|_{L^\infty} \ll S_{\ell_1}(\phi),$$

(N<sub>3</sub>) for all  $\ell$ , there exists  $\sigma_\ell > 0$  such that

$$S_\ell(h \cdot \phi) \ll_\ell e^{\sigma_\ell d(h, e)} S_\ell(\phi) \quad \text{for all } h \in H,$$

(N<sub>4</sub>) for every  $\ell$ , there exists  $\ell'$  such that

$$S_\ell(\phi_1 \phi_2) \ll_\ell S_{\ell'}(\phi_1) S_{\ell'}(\phi_2).$$

We suppose that for every  $r \geq 2$  there exist  $\delta_r, \ell_r > 0$  such that for all elements  $h_1, \dots, h_r \in H$  and all functions  $\phi_1, \dots, \phi_r \in \mathcal{A}$ ,

$$\begin{aligned} \mu((h_1 \cdot \phi_1) \cdots (h_r \cdot \phi_r)) &= \mu(\phi_1) \cdots \mu(\phi_r) \\ &+ O_r \left( S_{\ell_r}(\phi_1) \cdots S_{\ell_r}(\phi_r) e^{-\delta_r D(h_1, \dots, h_r)} \right), \end{aligned} \quad (7.1)$$

where

$$D(h_1, \dots, h_r) = \min_{i \neq j} d(h_i, h_j).$$

We shall additionally assume that the group  $H$  has *subexponential growth* which means that the balls

$$B_t = \{h \in H : d(h, e) < t\}$$

satisfy

$$\frac{\log \text{vol}(B_t)}{t} \rightarrow 0 \quad \text{as } t \rightarrow \infty. \quad (7.2)$$

Our main result is the following:

**Theorem 7.2.** *For every  $\phi \in \mathcal{A}$  with integral zero, the family of functions*

$$F_t(x) = \text{vol}(B_t)^{-1/2} \int_{B_t} \phi(h^{-1}x) dh \quad (7.3)$$

*converges in distribution as  $t \rightarrow \infty$  to the Normal Law with mean zero and the variance*

$$\sigma(\phi)^2 = \int_H \langle h \cdot \phi, \phi \rangle dh.$$

*Explicitly, this means that for every  $\xi \in \mathbb{R}$ ,*

$$\mu(\{x \in X : F_t(x) < \xi\}) \longrightarrow \text{Norm}_{\sigma(\phi)}(\xi) \quad \text{as } t \rightarrow \infty.$$

We remark that the condition that  $H$  has subexponential growth is important in Theorem 7.2. Indeed, Gorodnik and Ramirez [38] constructed examples of actions of rank-one simple Lie groups on homogeneous spaces which are exponentially mixing of all orders, but do not satisfy the Central Limit Theorem.

In particular, Theorem 7.2 immediately implies the following results about higher-rank abelian actions on homogeneous spaces.

**Corollary 7.3.** *Let  $G$  be a (noncompact) connected simple matrix Lie group with finite centre and  $H$  a (noncompact) closed subgroup of a Cartan subgroup of  $G$ . Then a measure-preserving action of  $H$  on finite-volume homogeneous spaces  $X$  of  $G$  satisfies the Central Limit Theorem. Namely, for every  $\phi \in C_c^\infty(X)$  with zero integral, the family of functions*

$$F_t = \text{vol}(B_t)^{-1/2} \int_{B_t} (h \cdot \phi) dh$$

*converges in distribution to the Normal Law as  $t \rightarrow \infty$ .*

We note that when  $M$  is compact surface with constant negative curvature, its unit tangent bundle can be realised as

$$T^1(M) \simeq \text{PSL}_2(\mathbb{R})/\Gamma,$$

where  $\Gamma$  is a discrete cocompact subgroup of  $\text{PSL}_2(\mathbb{R})$ , and the geodesic flow is given by

$$g_t : x \mapsto \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix} x \quad \text{for } x \in \text{PSL}_2(\mathbb{R})/\Gamma.$$

Hence, our method also provides a new proof of Theorem 7.1.

It is well-known from Probability that in order to establish that a family of bounded random variables  $X_t$  converges in distribution to a normal random variable  $N$ , it is sufficient to establish convergence of all moments, that is, that for all  $r \geq 1$

$$\mathbb{E}(X_t^r) \rightarrow \mathbb{E}(N^r) \quad \text{as } t \rightarrow \infty.$$

We essentially follow this route, but it will be more convenient to work with cumulants instead of moments. Given random variables  $X_1, \dots, X_r$ , the joint *cumulant* is defined as

$$\text{cum}(X_1, \dots, X_r) = (-i)^r \frac{\partial^r}{\partial z_1 \cdots \partial z_r} \log \mathbb{E} \left[ e^{i \sum_{k=1}^r z_k X_k} \right] \Big|_{z_1 = \dots = z_r = 0}.$$

It is useful to keep in mind that the joint cumulants can be expressed in terms of joint moments and conversely (see, for instance, [59]):

$$\begin{aligned} \text{cum}(X_1, \dots, X_r) &= \sum_{P \in \mathcal{P}_r} (-1)^{|P|-1} (|P|-1)! \prod_{I \in P} \mathbb{E} \left( \prod_{i \in I} X_i \right), \\ \mathbb{E}(X_1 \cdots X_r) &= \sum_{P \in \mathcal{P}_r} \prod_{I \in P} \text{cum}(X_i : i \in I), \end{aligned}$$

where the sums are taken over the set  $\mathcal{P}_r$  consisting of all partitions of  $\{1, \dots, r\}$ . Hence, studying cumulants is essentially equivalent to studying moments. However, it turns out that cumulants have several very convenient additional vanishing properties that will be crucial for our argument:

- If there exists a nontrivial partition  $\{1, \dots, r\} = I \sqcup J$  such that  $\{X_i : i \in I\}$  and  $\{X_i : i \in J\}$  are independent, then

$$\text{cum}(X_1, \dots, X_r) = 0. \quad (7.4)$$

- if  $N$  is a normal random variable, then the cumulants of order at least three satisfy

$$\text{cum}(N, \dots, N) = 0$$

Now we adopt this probabilistic notation to our setting. For functions  $\phi_1, \dots, \phi_r \in L^\infty(X)$  and a subset  $I \subset \{1, \dots, r\}$ , we set

$$\phi_I = \prod_{i \in I} \phi_i.$$

We use the convention that  $\phi_\emptyset = 1$ . Then we define the joint *cumulant* of  $\phi_1, \dots, \phi_r$  as

$$\text{cum}_r(\phi_1, \dots, \phi_r) = \sum_{P \in \mathcal{P}_r} (-1)^{|P|-1} (|P|-1)! \prod_{I \in P} \mu(\phi_I).$$

For a function  $\phi \in L^\infty(X)$ , we also set

$$\text{cum}_r(\phi) = \text{cum}_r(\phi, \dots, \phi).$$

The following proposition, which is essentially equivalent to the more widely known Method of Moments, provides a convenient criterion for proving the Central Limit Theorem.

**Proposition 7.4.** *Let  $F_t \in L^\infty(X)$  be a family of functions such that as  $t \rightarrow \infty$ ,*

$$\mu(F_t) \rightarrow 0, \quad (7.5)$$

$$\|F_t\|_{L^2} \rightarrow \sigma, \quad (7.6)$$

$$\text{cum}_r(F_t) \rightarrow 0 \quad \text{for all } r \geq 3. \quad (7.7)$$

Then for every  $\xi \in \mathbb{R}$ ,

$$\mu(\{x \in X : F_t(x) < \xi\}) \rightarrow \text{Norm}_\sigma(\xi) \quad \text{as } t \rightarrow \infty.$$

In the setting of one-parameter dynamical systems, this criterion was developed by Leonov [57] who used it in [58] to establish the Central Limit Theorem for ergodic automorphisms of compact abelian groups. Estimates on cumulants were subsequently used by Cohen and Conze [14, 15, 16] to derive the Central Limit Theorem for higher-rank abelian actions by automorphisms of compact abelian groups. Similar results regarding higher-rank abelian actions by automorphisms

have been previously proved by Levin [60] who managed to compute the asymptotics of all moments directly using delicate analysis of solutions of the  $S$ -unit equation.

We begin the proof of Theorem 7.2. In view of Proposition 7.4, it remains to verify that the family of functions  $F_t$  defined in (7.3) satisfies (7.5), (7.6), and (7.7). The first condition is immediate, and the second is verified as follows. We observe that

$$\begin{aligned} \|F_t\|_{L^2}^2 &= \text{vol}(B_t)^{-1} \int_{B_t \times B_t} \langle h_1 \cdot \phi, h_2 \cdot \phi \rangle dh_1 dh_2 \\ &= \text{vol}(B_t)^{-1} \int_{H \times H} \chi_{B_t}(h_1) \chi_{B_t}(h_2) \langle (h_1^{-1} h_2) \cdot \phi, \phi \rangle dh_1 dh_2 \\ &= \int_H \frac{\text{vol}(B_t \cap B_t h^{-1})}{\text{vol}(B_t)} \langle h \cdot \phi, \phi \rangle dh. \end{aligned}$$

It is not hard to check using the subexponential growth property (7.2) that the balls  $B_t$  satisfy the Følner property, that is, for all  $h \in H$ ,

$$\frac{\text{vol}(B_t \cap B_t h^{-1})}{\text{vol}(B_t)} \rightarrow 1 \quad \text{as } t \rightarrow \infty.$$

Moreover, it follows from (7.1) with  $r = 2$  that the function  $h \mapsto \langle h \cdot \phi, \phi \rangle$  is in  $L^1(H)$ . Thus, using the Dominated Convergence Theorem, we deduce that

$$\|F_t\|_{L^2}^2 \rightarrow \int_H \langle h \cdot \phi, \phi \rangle dh \quad \text{as } t \rightarrow \infty.$$

This implies (7.6).

Verification of (7.7) is the most challenging part of the proof because it requires to show asymptotic vanishing of the cumulants

$$\text{cum}_r(F_t) = \text{vol}(B_t)^{-r/2} \int_{B_t^r} \text{cum}_r(h_1 \cdot \phi, \dots, h_r \cdot \phi) dh,$$

which is even more than the square-root cancellation in this integral. The first crucial input for estimating  $\text{cum}_r(F_t)$  comes from the bound on correlations (7.1). However, these bound will be only useful for certain ranges of tuples  $h = (h_1, \dots, h_r)$ .

To utilise the bound (7.1) most efficiently, we introduce a decomposition of the product  $H^r$  into a union of domains where the components  $h_i$  are either separated or clustered on suitable scales. For subsets  $I, J \subset \{1, \dots, r\}$  and  $h = (h_1, \dots, h_r) \in H^r$ , we set

$$\begin{aligned} d^I(h) &= \max\{d(h_i, h_j) : i, j \in I\} \\ d_{I,J}(h) &= \min\{d(h_i, h_j) : i \in I, j \in J\}, \end{aligned}$$

and for a partition  $Q \in \mathcal{P}_r$ , we set

$$\begin{aligned} d^Q(h) &= \max\{d^I(h) : I \in Q\}, \\ d_Q(h) &= \min\{d_{I,J}(h) : I \neq J \in Q\}. \end{aligned}$$

Using this notation, we define for  $0 \leq \alpha \leq \beta$ ,

$$\begin{aligned} \Delta_Q(\alpha, \beta) &= \{h \in H^r : d^Q(h) \leq \alpha, d_Q(h) > \beta\}, \\ \Delta(\beta) &= \{h \in H^r : d(h_i, h_j) \leq \beta \text{ for all } i, j\}. \end{aligned}$$

For  $h = (h_1, \dots, h_r) \in \Delta_Q(\alpha, \beta)$ , we think about components  $h_i$  with  $i$  in the same atom of  $Q$  as “clustered” and about  $h_i$  with  $i$  in different atoms of  $Q$  as “separated” (see Figure 3).

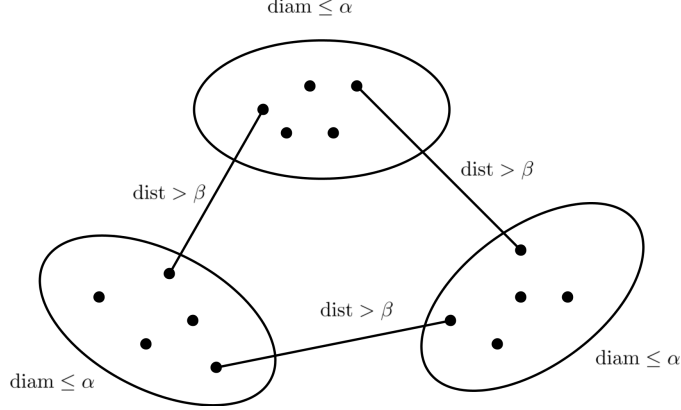


FIGURE 3. Tuples in the sets  $\Delta(\alpha, \beta)$ .

These features allow to estimate  $\text{cum}_r(h_1 \cdot \phi, \dots, h_r \cdot \phi)$  on the sets  $\Delta_Q(\alpha, \beta)$ :

**Proposition 7.5.** *There exist  $\delta_r, \sigma_r > 0$  such that for every  $0 \leq \alpha \leq \beta$ ,  $Q \in \mathcal{P}_r$  with  $|Q| \geq 2$ , and  $(h_1, \dots, h_r) \in \Delta_Q(\alpha, \beta)$ ,*

$$\text{cum}_r(h_1 \cdot \phi, \dots, h_r \cdot \phi) \ll_{r, \phi} e^{-\delta_r \beta - \sigma_r \alpha}.$$

*Proof.* The proof will exploit a certain cancellation property of cumulants. For  $Q \in \mathcal{P}_r$  and  $\phi_1, \dots, \phi_r \in L^\infty(X)$ , we define the conditional cumulant as

$$\text{cum}_r(\phi_1, \dots, \phi_r | Q) = \sum_{P \in \mathcal{P}_r} (-1)^{|P|-1} (|P| - 1)! \prod_{I \in P} \prod_{J \in Q} \mu(\phi_{I \cap J}).$$

One can show that when the partition  $Q$  is nontrivial,

$$\text{cum}_r(\phi_1, \dots, \phi_r | Q) = 0. \tag{7.8}$$

This fact is an analogue of (7.4), but it is not a probabilistic property, but rather a combinatorial cancellation feature of the cumulant sums, and we refer, for instance, to [8] for a self-contained proof of (7.8).

In order to bound  $\text{cum}_r(h_1 \cdot \phi, \dots, h_r \cdot \phi)$ , we shall show that when  $(h_1, \dots, h_r) \in \Delta_Q(\alpha, \beta)$ ,

$$\text{cum}_r(h_1 \cdot \phi, \dots, h_r \cdot \phi) \approx \text{cum}_r(h_1 \cdot \phi, \dots, h_r \cdot \phi | Q)$$

which reduces to verifying that for  $I \in P$ ,

$$\mu \left( \prod_{i \in I} h_i \cdot \phi \right) \approx \prod_{J \in Q} \mu \left( \prod_{i \in I \cap J} h_i \cdot \phi \right).$$

This is where the full strength of the estimate (7.1) on higher-order correlations comes into play. For each  $J$ , we pick  $h_j$  as one of  $h_i$ ,  $j \in I \cap J$ . Then

$$\mu \left( \prod_{i \in I} h_i \cdot \phi_i \right) = \mu \left( \prod_{J \in Q} h_J \Phi_J \right),$$

where  $\Phi_J = \prod_{i \in I \cap J} (h_J^{-1} h_i) \cdot \phi$ . Since  $(h_1, \dots, h_r) \in \Delta_Q(\alpha, \beta)$ , we have

$$\begin{aligned} d(h_J^{-1} h_i, e) &= d(h_i, h_J) \leq \alpha && \text{for } i \in J \in Q, \\ d(h_{J_1}, h_{J_2}) &> \beta && \text{for } J_1 \neq J_2 \in Q. \end{aligned}$$

Hence, it follows from (7.1) that

$$\mu \left( \prod_{J \in Q} h_J \Phi_J \right) = \prod_{J \in Q} \mu(\Phi_J) + O_r \left( \prod_{J \in Q} S_{\ell_r}(\Phi_J) e^{-\delta_r \beta} \right),$$

and by the properties (N<sub>3</sub>) and (N<sub>4</sub>) of the norms,

$$S_\ell(\Phi_J) \ll_\ell \prod_{i \in I \cap J} S_{\ell'}((h_J^{-1} h_i) \cdot \phi) \ll_{\ell', \phi} e^{r \sigma_\ell \alpha}.$$

This implies that for some  $\sigma_r > 0$ ,

$$\mu \left( \prod_{i \in I} h_i \cdot \phi_i \right) = \prod_{J \in Q} \mu \left( \prod_{i \in I \cap J} h_i \cdot \phi_i \right) + O_{r, \phi} \left( e^{-(\delta_r \beta - \sigma_r \alpha)} \right).$$

which can be used to prove the proposition.  $\square$

We shall use the following decomposition of the space of tuples  $H^r$ .

**Proposition 7.6.** *Given parameters*

$$0 = \beta_0 < \beta_1 < 3\beta_1 \leq \beta_2 < \dots < \beta_{r-1} < 3\beta_{r-1} \leq \beta_r,$$

*we have the decomposition*

$$H^r = \Delta(\beta_r) \cup \left( \bigcup_{j=0}^{r-1} \bigcup_{Q: |Q| \geq 2} \Delta_Q(3\beta_j, \beta_{j+1}) \right).$$

The proof of Proposition 7.6 uses the following lemma:

**Lemma 7.7.** *Let  $Q \in \mathcal{P}_r$  with  $|Q| \geq 2$  and  $0 \leq \alpha \leq \beta$ . Suppose that for  $h \in H^r$ ,*

$$d^Q(h) \leq \alpha \quad \text{and} \quad d_Q(h) \leq \beta.$$

*Then there exists a partition  $Q_1$  which is strictly coarser than  $Q$  such that*

$$d^{Q_1}(h) \leq 3\beta.$$

*Proof.* We observe that the sets  $\{h_i : i \in I\}$  with  $I \in Q$  have diameters at most  $\alpha$ , and the distance between at least two of these sets is bounded by  $\beta$ . We define the new partition  $Q_1$  by combining the sets whose distance at most  $\beta$  between them. This gives a strictly coarser partition. It follows from the triangle inequality that the diameters of the sets  $\{h_i : i \in J\}$  with  $J \in Q_1$  are at most  $2\alpha + \beta \leq 3\beta$ . This implies that  $d^{Q_1}(h) \leq 3\beta$ .  $\square$

*Proof of Proposition 7.6.* Let us take arbitrary  $h \in H^r$ . Suppose that  $h \notin \Delta_{Q_0}(0, \beta_1)$  for  $Q_0 = \{\{1\}, \dots, \{r\}\}$ . It is clear that  $d^{Q_0}(h) = 0$  so that also  $d_{Q_0}(h) \leq \beta_1$ . Hence, it follows from Lemma 7.7 that there exists a partition  $Q_1$  coarser than  $Q_0$  such that  $d^{Q_1}(h) \leq 3\beta_1$ . If  $d_{Q_1}(h) > \beta_2$ , then  $h \in \Delta_{Q_1}(3\beta_1, \beta_2)$  and  $h$  belongs to the union. On the other, if  $d_{Q_1}(h) \leq \beta_2$ , we apply Lemma 7.7 again to conclude that there exists a partition  $Q_2$  coarser than  $Q_1$  such that  $d^{Q_2}(h) \leq 3\beta_2$ . This argument can be continued, and we deduce that after at most  $r$  steps, we see that  $h$  belongs to the union of  $\Delta_{Q_j}(3\beta_j, \beta_{j+1})$  with  $|Q_j| \geq 2$ , or we get  $Q_i = \{\{1, \dots, r\}\}$  and  $d^{Q_i}(h) \leq 3\beta_i < \beta_r$ . In the latter case, we deduce that  $h \in \Delta(\beta_r)$ . This proves the required decomposition.  $\square$

Now we are ready to complete the proof of Theorem 7.2. As we have already remarked, it remains to prove (7.7). Using the decomposition established in Proposition 7.6, we deduce that

$$\begin{aligned} \text{cum}_r(F_t) &= \text{vol}(B_t)^{-r/2} \int_{B_t^r} \text{cum}_r(h_1 \cdot \phi, \dots, h_r \cdot \phi) dh \\ &\ll_r \text{vol}(B_t)^{-r/2} \left( \text{vol}(B_t^r \cap \Delta(\beta_r)) \|\phi\|_{L^\infty}^r \right. \\ &\quad \left. + \max_{j, Q: |Q| \geq 2} \int_{B_t^r \cap \Delta_Q(3\beta_j, \beta_{j+1})} |\text{cum}_r(h_1 \cdot \phi, \dots, h_r \cdot \phi)| dh \right). \end{aligned}$$

It follows from invariance of the volume on  $H$  that

$$\text{vol}(B_t^r \cap \Delta(\beta_r)) \leq \int_{B_t} \text{vol}(B(h, \beta_r))^{r-1} dh = \text{vol}(B_t) \text{vol}(B_{\beta_r})^{r-1}.$$

Hence, using Proposition 7.5, we conclude that

$$\text{cum}_r(F_t) \ll_{r, \phi} \text{vol}(B_t)^{1-r/2} \text{vol}(B_{\beta_r})^{r-1} + \text{vol}(B_t)^{r/2} \left( \max_j e^{-\delta_r \beta_{j+1} - 3\sigma_r \beta_j} \right).$$

For a parameter  $\theta > 0$ , we choose  $\beta_j$ 's recursively as

$$\beta_0 = 0, \beta_{j+1} = \max\{3\beta_j, \delta_r^{-1}(\theta + 3\sigma_r \beta_j)\}.$$

Then  $\beta_r \leq c_r \theta$  with some  $c_r > 0$ , and

$$\text{cum}_r(F_t) \ll_{r, \phi} \text{vol}(B_t)^{1-r/2} \text{vol}(B_{c_r \theta})^{r-1} + \text{vol}(B_t)^{r/2} e^{-\theta}.$$

We take  $\theta = r \log \text{vol}(B_t)$ . Then it follows from the subexponential growth condition (7.2) that when  $r \geq 3$ ,

$$\text{cum}_r(F_t) \rightarrow 0 \quad \text{as } t \rightarrow \infty.$$

This completes the proof of Theorem 7.2.

## REFERENCES

- [1] M. Babillot, Points entiers et groupes discrets: de l'analyse aux systèmes dynamiques. Panor. Synthèses, 13, Rigidité, groupe fondamental et dynamique, 1–119, Soc. Math. France, Paris, 2002.
- [2] U. Bader, A. Furman, A. Gorodnik, B. Weiss, Rigidity of group actions on homogeneous spaces, III. *Duke Math. J.* 164 (2015), no. 1, 115–155.
- [3] H.-J. Bartels, Nichteuklidische Gitterpunktprobleme und Gleichverteilung in linearen algebraischen Gruppen. *Comment. Math. Helv.* 57 (1982), no. 1, 158–172.
- [4] B. Bekka, On Uniqueness of Invariant Means. *Proc. Amer. Math. Soc.* 126 (1998), 507–514.
- [5] B. Bekka, P. de la Harpe, A. Valette, Kazhdan's property (T). *New Mathematical Monographs*, 11. Cambridge University Press, Cambridge, 2008.
- [6] B. Bekka and M. Mayer, Ergodic theory and topological dynamics of group actions on homogeneous spaces. *London Mathematical Society Lecture Note Series*, 269. Cambridge University Press, Cambridge, 2000.
- [7] M. Björklund, M. Einsiedler, A. Gorodnik, Quantitative multiple mixing. to appear in *J. Eur. Math. Soc.*; ArXiv:1701.00945.
- [8] M. Björklund and A. Gorodnik, Central Limit Theorems for group actions which are exponentially mixing of all orders. accepted to *Journal d'Analyse Mathématiques*; ArXiv:1706.09167.
- [9] M. Björklund and A. Gorodnik, Central limit theorems in the geometry of numbers. *Electron. Res. Announc. Math. Sci.* 24 (2017), 110–122.
- [10] M. Björklund and A. Gorodnik, Central limit theorems for Diophantine approximants. ArXiv:1804.06084.

- [11] A. Borel and N. Wallach, Continuous cohomology, discrete subgroups, and representations of reductive groups. *Annals of Mathematics Studies*, 94. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1980.
- [12] J. Bourgain, A Szemerédi type theorem for sets of positive density in  $R^k$ . *Israel J. Math.* 54 (1986), 307–316.
- [13] W. Casselman and D. Milčić, Asymptotic behavior of matrix coefficients of admissible representations. *Duke Math. J.* 49 (1982), no. 4, 869–930.
- [14] G. Cohen and J.-P. Conze, CLT for Random Walks of Commuting Endomorphisms on Compact Abelian Groups. *J. Theoret. Probab.* 30 (2017), no. 1, 143–195.
- [15] G. Cohen and J.-P. Conze, Central limit theorem for commutative semigroups of toral endomorphisms. *Arxiv:1304.4556*.
- [16] G. Cohen and J.-P. Conze, Almost mixing of all orders and CLT for some  $Z^d$ -actions on subgroups of  $F_p^{Z^d}$ . *Arxiv:1609.06484*.
- [17] M. Cowling, Sur les coefficients des représentations unitaires des groupes de Lie simples. *Analyse harmonique sur les groupes de Lie (Sém., Nancy–Strasbourg 1976–1978)*, II, Springer, Berlin (1979), 132–178.
- [18] M. Cowling, U. Haagerup, R. Howe, Almost  $L^2$  matrix coefficients. *J. Reine Angew. Math.* 387 (1988), 97–110.
- [19] S. G. Dani, Kolmogorov automorphisms on homogeneous spaces. *Amer. J. Math.* 98 (1976), no. 1, 119–163.
- [20] S. G. Dani, Spectrum of an affine transformation. *Duke Math. J.* 44 (1977), no. 1, 129–155.
- [21] A. del Junco and R. Yassawi, Multiple mixing and rank one group actions. *Canad. J. Math.* 52 (2000), no. 2, 332–347.
- [22] J. Delsarte, Sur le gitter fuchsien. *C. R. Acad. Sci. Paris* 214 (1942), 147–149; *Oeuvres de Jean Delsarte*, vol. II, Editions du CNRS, Paris, 1971, pp. 829–845.
- [23] M. Denker, The central limit theorem for dynamical systems. *Banach Center Pub.* 23 (1989), 33–61.
- [24] Y. Derriennic, Some aspects of recent works on limit theorems in ergodic theory with special emphasis on the “central limit theorem”. *Discrete Contin. Dyn. Syst.* 15 (2006), no. 1, 143–158.
- [25] D. Dolgopyat, Limit theorems for partially hyperbolic systems. *Trans. Amer. Math. Soc.* 356 (2004), no. 4, 1637–1689.
- [26] W. Duke, Z. Rudnick, P. Sarnak, Density of integer points on affine homogeneous varieties. *Duke Math. J.* 71 (1993), no. 1, 143–179.
- [27] M. Einsiedler, G. Margulis, A. Venkatesh, Effective equidistribution for closed orbits of semisimple groups on homogeneous spaces. *Invent. Math.* 177 (2009), no. 1, 137–212.
- [28] A. Eskin and C. McMullen, Mixing, counting, and equidistribution in Lie groups. *Duke Math. J.* 71 (1993), no. 1, 181–209.
- [29] D. Fisher, B. Kalinin, R. Spatzier, Global rigidity of higher rank Anosov actions on tori and nilmanifolds. With an appendix by James F. Davis. *J. Amer. Math. Soc.* 26 (2013), no. 1, 167–198.
- [30] G. Folland, *A course in abstract harmonic analysis*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1995.
- [31] H. Furstenberg, Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. *J. Analyse Math.* 31 (1977), 204–256.
- [32] H. Furstenberg and Y. Katznelson, An ergodic Szemerédi theorem for commuting transformations. *J. Analyse Math.* 34 (1978), 275–291.
- [33] H. Furstenberg, Y. Katznelson and B. Weiss, Ergodic theory and configurations in sets of positive density. *Mathematics of Ramsey Theory, Algorithms and Combinatorics*, Vol. 5, Springer, Berlin, 1990, pp. 184–198.
- [34] R. Gangolli and V. S. Varadarajan, Harmonic analysis of spherical functions on real reductive groups. *Ergebnisse der Mathematik und ihrer Grenzgebiete*, 101. Springer-Verlag, Berlin, 1988.
- [35] A. Gorodnik and A. Nevo, The ergodic theory of lattice subgroups. *Annals of Mathematics Studies*, 172. Princeton University Press, Princeton, NJ, 2010.
- [36] A. Gorodnik and A. Nevo, Counting lattice points. *J. Reine Angew. Math.* 663 (2012), 127–176.

- [37] A. Gorodnik and R. Spatzier, Exponential mixing of nilmanifold automorphisms, *J. Anal. Math.* 123 (2014), 355–396.
- [38] A. Gorodnik and F. Ramirez, Limit theorems for rank-one Lie groups. *Proc. Amer. Math. Soc.* 142 (2014), no. 4, 1359–1369.
- [39] A. Gorodnik and R. Spatzier, Mixing properties of commuting nilmanifold automorphisms. *Acta Math.* 215 (2015), no. 1, 127–159.
- [40] S. Gouëzel, Limit theorems in dynamical systems using the spectral method. *Hyperbolic dynamics, fluctuations and large deviations*, 161–193, *Proc. Sympos. Pure Math.*, 89, Amer. Math. Soc., Providence, RI, 2015.
- [41] R. L. Graham, Recent trends in Euclidean Ramsey theory. *Trends in discrete mathematics. Discrete Math.* 136 (1994), no. 1–3, 119–127.
- [42] Harish-Chandra, Spherical functions on a semisimple Lie group. I. *Amer. J. Math.* 80 (1958), 241–310.
- [43] C. Herz, Sur le phénomène de Kunze-Stein. *C. R. Acad. Sci. Paris Sér. A-B* 271 (1970), A491–A493.
- [44] B. Host, Mixing of all orders and pairwise independent joinings of systems with singular spectrum. *Israel J. Math.* 76 (1991), no. 3, 289–298.
- [45] B. Host and B. Kra, Nonconventional ergodic averages and nilmanifolds. *Ann. Math.* 161 (2005), no. 1, 397–488.
- [46] R. Howe, On a notion of rank for unitary representations of the classical groups. *Harmonic analysis and group representations*, 223–331, Liguori, Naples, 1982.
- [47] R. Howe and C. Moore, Asymptotic properties of unitary representations. *J. Funct. Anal.* 32 (1979), no. 1, 72–96.
- [48] R. Howe and E.-C. Tan, *Nonabelian harmonic analysis. Applications of  $SL(2, \mathbb{R})$* . Universitext. Springer-Verlag, New York, 1992.
- [49] H. Huber, Zur analytischen Theorie hyperbolischer Raumformen und Bewegungsgruppen, I, *Math. Ann.* 138 (1959), 1–26; II *Math. Ann.*, 142 (1961), 385–398 and 143 (1961), 463–464.
- [50] S. Kalikow, Twofold mixing implies threefold mixing for rank one transformations. *Ergodic Theory Dynam. Systems* 4 (1984), no. 2, 237–259.
- [51] A. Katok and R. Spatzier, First cohomology of Anosov actions of higher rank abelian groups and applications to rigidity. *Inst. Hautes Études Sci. Publ. Math.* 79 (1994), 131–156.
- [52] A. Knapp, *Representation theory of semisimple groups. An overview based on examples. Princeton Landmarks in Mathematics.* Princeton University Press, Princeton, NJ, 2001.
- [53] I. Konstantoulas, Effective decay of multiple correlations in semidirect product actions. *J. Mod. Dyn.* 10 (2016), 81–111.
- [54] S. Le Borgne, *Martingales in hyperbolic geometry. Analytic and probabilistic approaches to dynamics in negative curvature*, 1–63, Springer INdAM Ser., 9, Springer, Cham, 2014.
- [55] F. Ledrappier, Un champ markovien peut être d’entropie nulle et mélangeant. *C. R. Acad. Sci. Paris Sér. A-B*, 287 (1978), A561–A563.
- [56] A. Leibman, Pointwise convergence of ergodic averages for polynomial sequences of translations on a nilmanifold. *Ergodic Theory Dynam. Systems* 25 (2005), no. 1, 201–213.
- [57] V.P. Leonov, The use of the characteristic functional and semi-invariants in the ergodic theory of stationary processes. *Dokl. Akad. Nauk SSSR* 133 (1960), 523–526.
- [58] V.P. Leonov, On the central limit theorem for ergodic endomorphisms of compact commutative groups. *Dokl. Akad. Nauk SSSR* 135 (1960), 258–261.
- [59] V.P. Leonov and A.N. Shiryayev, On a method of calculations of semi-invariants. *Theory of Probability and its Applications* 4 (1959), 319–329.
- [60] M. Levin, Central limit theorem for  $Z^{d+}$ -actions by toral endomorphisms. *Electron. J. Probab.* 18 (2013), no. 35, 42 pp.
- [61] J.-S. Li, The minimal decay of matrix coefficients for classical groups. *Harmonic analysis in China*, 146–169, *Math. Appl.*, 327, Kluwer Acad. Publ., Dordrecht, 1995.
- [62] J.-S. Li and C.-B. Zhu, On the decay of matrix coefficients for exceptional groups. *Math. Ann.* 305 (1996), no. 2, 249–270.
- [63] D. A. Lind, Dynamical properties of quasihyperbolic toral automorphisms. *Ergodic Theory Dynamical Systems* 2 (1982), no. 1, 49–68.
- [64] B. Marcus, The horocycle flow is mixing of all degrees. *Invent. Math.* 46 (1978), no. 3, 201–209.

- [65] G. Margulis, Certain applications of ergodic theory to the investigation of manifolds of negative curvature. *Funkcional. Anal. i Prilozen.* 3 (1969), no. 4, 89–90.
- [66] G. Margulis, On some aspects of the theory of Anosov systems. *Springer Monographs in Mathematics*. Springer-Verlag, Berlin, 2004
- [67] F. Maucourant, Homogeneous asymptotic limits of Haar measures of semisimple linear groups and their lattices. *Duke Math. J.* 136 (2007), no. 2, 357–399.
- [68] F. I. Mautner, Geodesic flows on symmetric Riemann spaces. *Ann. Math.* 65 (1957), 416–431.
- [69] C. Moore, Ergodicity of flows on homogeneous spaces. *Amer. J. Math.* 88 (1966), 154–178.
- [70] C. Moore, The Mautner phenomenon for general unitary representations. *Pacific J. Math.* 86 (1980), no. 1, 155–169.
- [71] C. Moore, Exponential decay of correlation coefficients for geodesic flows. *Group representations, ergodic theory, operator algebras, and mathematical physics (Berkeley, Calif., 1984)*, 163–181, *Math. Sci. Res. Inst. Publ.*, 6, Springer, New York, 1987.
- [72] S. Mozes, Mixing of all orders of Lie groups actions. *Invent. Math.* 107 (1992), no. 2, 235–241; erratum: *Invent. Math.* 119 (1995), no. 2, 399.
- [73] H. Oh, Tempered subgroups and representations with minimal decay of matrix coefficients. *Bull. Soc. Math. France* 126 (1998), no. 3, 355–380.
- [74] H. Oh, Uniform pointwise bounds for matrix coefficients of unitary representations and applications to Kazhdan constants. *Duke Math. J.* 113 (2002), no. 1, 133–192.
- [75] D. Ornstein and B. Weiss, Geodesic flows are Bernoullian, *Israel J. Math.* 14 (1973), 184–198.
- [76] S. J. Patterson, A lattice-point problem in hyperbolic space. *Mathematika* 22 (1975), no. 1, 81–88; erratum: *Mathematika* 23 (1976), no. 2, 227.
- [77] F. Pène, Averaging method for differential equations perturbed by dynamical systems. *ESAIM Probab. Statist.* 6 (2002), 33–88.
- [78] M. Ratner, The rate of mixing for geodesic and horocycle flows. *Ergodic Theory Dynam. Systems* 7 (1987), 267–288.
- [79] V. A. Rokhlin, On endomorphisms of compact commutative groups. *Izvestiya Akad. Nauk SSSR. Ser. Mat.* 13 (1949), 329–340.
- [80] V. V. Ryzhikov, Joinings and multiple mixing of the actions of finite rank. *Funct. Anal. Appl.* 27 (1993), no. 2, 128–140.
- [81] K. Schmidt and T. Ward, Mixing automorphisms of compact groups and a theorem of Schlickewei. *Invent. Math.* 111 (1993), no. 1, 69–76.
- [82] K. Schmidt, *Dynamical systems of algebraic origin*. *Progress in Mathematics*, 128. Birkhäuser Verlag, Basel, 1995.
- [83] I.E. Segal and J. von Neumann, A theorem on unitary representations of semisimple Lie groups. *Ann. Math.* 52 (1950), 509–517.
- [84] Ya. Sinai, The central limit theorem for geodesic flows on manifolds of constant negative curvature. *Soviet Math. Dokl.* 1 (1960), 938–987.
- [85] Ya. Sinai, Probabilistic concepts in ergodic theory. 1963 *Proc. Internat. Congr. Mathematicians (Stockholm, 1962)* pp. 540–559.
- [86] A. Starkov, Multiple mixing of homogeneous flows. *Dokl. Akad. Nauk* 333 (1993), no. 4, 442–445; translation in *Russian Acad. Sci. Dokl. Math.* 48 (1994), no. 3, 573–578.
- [87] E. Szemerédi, On sets of integers containing no  $k$  elements in arithmetic progression. *Acta Arith.* 27 (1975), 199–245.
- [88] M. Viana, *Stochastic dynamics of deterministic systems*. *Lecture Notes XXI Bras. Math. Colloq. IMPA, Rio de Janeiro, 1997*.
- [89] S.P. Wang, On the Mautner phenomenon and groups with property (T). *Amer. J. Math.* 104 (1982), no. 6, 1191–1210.
- [90] S.P. Wang, The Mautner phenomenon for  $p$ -adic Lie groups. *Math. Z.* 185 (1984), no. 3, 403–412.
- [91] Z. Wang, Uniform pointwise bounds for matrix coefficients of unitary representations on semidirect products. *J. Funct. Anal.* 267 (2014), no. 1, 15–79.
- [92] G. Warner, *Harmonic analysis on semi-simple Lie groups. I. Die Grundlehren der mathematischen Wissenschaften, Band 188*. Springer-Verlag, New York-Heidelberg, 1972.
- [93] G. Warner, *Harmonic analysis on semi-simple Lie groups. II. Die Grundlehren der mathematischen Wissenschaften, Band 189*. Springer-Verlag, New York-Heidelberg, 1972.
- [94] R. Yassawi, Multiple mixing and local rank group actions. *Ergodic Theory Dynam. Systems* 23 (2003), no. 4, 1275–1304.

- [95] T. Ziegler, A non-conventional ergodic theorem for a nilsystem. *Ergodic Theory Dynam. Systems* 25 (2005), no. 4, 1357–1370.
- [96] T. Ziegler, Nilfactors of  $R^m$ -actions and configurations in sets of positive upper density in  $R^m$ . *J. Anal. Math.* 99 (2006), 249–266.
- [97] T. Ziegler, Universal characteristic factors and Furstenberg averages. *J. Amer. Math. Soc.* 20 (2007), no. 1, 53–97.
- [98] R. Zimmer, *Ergodic theory and semisimple groups*. *Monographs in Mathematics*, 81. Birkhäuser Verlag, Basel, 1984.