Effective Multiple Equidistribution of Translated Measures

Citation for the original published paper (version of record):
Björklund, M., Gorodnik, A. (2023). Effective Multiple Equidistribution of Translated Measures.
International Mathematics Research Notices, 2023(1): 210-242.
http://dx.doi.org/10.1093/imrn/rnab274

M. Björklund and A. Gorodnik (2023) "Effective Multiple Equidistribution of Translated Measures," International Mathematics Research Notices, Vol. 2023, No. 1, pp. 210–242
Advance Access Publication October 1, 2021
https://doi.org/10.1093/imrn/rnab274

Effective Multiple Equidistribution of Translated Measures

Michael Björklund[1,*] and Alexander Gorodnik[2]

[1]Mathematical Sciences, Chalmers University of Technology and the University of Gothenburg, SE-412 96 Göteborg, Sweden and [2]Institut für Mathematik, Universität Zürich, Winterthurerstrasse, 190 8057 Zürich, Switzerland

*Correspondence to be sent to: e-mail: micbjo@chalmers.se

We study the joint distributions of translated measures supported on periodic orbits that are expanded by subgroups of diagonal matrices and generalize (special cases) of previous results of Kleinbock–Margulis, Dabbs–Kelly–Li, and Shi. More specifically, we establish quantitative estimates on higher-order correlations for measures with low regularities and derive error terms that only depend on the distances between translations.

1 Introduction

Let us consider a jointly continuous action of a topological group G on a topological space X. Given a finite Borel measure σ on X, one is often interested in the asymptotic behaviour of integrals of the form

$$\sigma(\varphi \circ g) := \int_X \varphi(gx)\,\mathrm{d}\sigma(x), \quad \text{with } \varphi \in C_b(X),$$

as $g \to \infty$ in G, as well as in the asymptotic behaviour of integrals of the form

$$\sigma\left(\varphi_1 \circ g_1 \cdots \varphi_r \circ g_r\right) := \int_X \varphi_1(g_1 x) \cdots \varphi_r(g_r x)\,\mathrm{d}\sigma(x), \quad \text{with } \varphi_1, \ldots, \varphi_r \in C_b(X),$$

Received May 12, 2021; Revised August 25, 2021; Accepted September 2, 2021
Communicated by Prof. Barak Weiss

as $g_i \to \infty$ and $g_i g_j^{-1} \to \infty$ for $i \neq j$ in G. The latter integrals are often called *r-correlations*. In this paper, we will be interested in quantitative estimates on r-correlations for actions on homogeneous spaces.

Basic models for the type of results that we are after can be found in [9, 11], which deals with the equidistribution of closed horocycles on the modular surface. More generally, one can consider the action of a partially hyperbolic one-parameter subgroup (g_t) in $\mathrm{SL}_d(\mathbb{R})$ on the homogeneous space $X := \mathrm{SL}_d(\mathbb{R})/\mathrm{SL}_d(\mathbb{Z})$, equipped with the (unique) $\mathrm{SL}_d(\mathbb{R})$-invariant probability invariant measure μ. Kleinbock and Margulis [6] showed that in this case there exists $\delta > 0$ such that for every compactly supported smooth probability measure σ on an unstable leaf of (g_t) and for every $\varphi \in C_c^\infty(X)$, there is a constant $C(\sigma, \varphi)$, such that

$$\left|\sigma(\varphi \circ g_t) - \mu(\varphi)\right| \leq C(\sigma, \varphi)\, e^{-\delta t} \tag{1.1}$$

for all $t \geq 0$. In fact, the result of [6] applies more generally to partially hyperbolic flows on homogeneous spaces of semisimple Lie groups. Subsequently, Dolgopyat [5] developed an inductive argument that allows one to deduce from the estimates of the form (1.1) quantitative estimates on higher-order correlations. In the setting described above, this argument tells us that there exists $\delta' > 0$ such that for every $t_1, \ldots, t_r \geq 0$

$$\left|\sigma\left(\varphi_1 \circ g_{t_1} \cdots \varphi_r \circ g_{t_r}\right) - \mu(\varphi_1) \cdots \mu(\varphi_r)\right| \leq C(\sigma, \varphi_1, \ldots, \varphi_r)\, e^{-\delta' D_r(t_1, \ldots, t_r)}$$

for all $\varphi_1, \ldots, \varphi_r \in C_c^\infty(X)$, where

$$D_r(t_1, \ldots, t_r) := \min(t_1, t_2 - t_1, \ldots, t_r - t_{r-1}).$$

The idea behind this argument in [6] goes back to Margulis' thesis [8] and uses that the flow (g_t) is non-expanding transversally to the unstable leaves, so that one can "thicken" the measure σ to reduce the original problem to mixing estimates for the flow (g_t) with respect to the volume measure μ.

The problem becomes significantly harder when the measure σ is supported on a proper submanifold of the unstable leaf. A particular instance of this problem was investigated by Kleinbock and Margulis in [7]. They consider the case when σ is a smooth measure, compactly supported on an orbit of the subgroup

$$U_{m,n} := \left\{ \begin{pmatrix} I_m & u \\ 0 & I_n \end{pmatrix} : \; u \in \mathrm{Mat}_{m \times n}(\mathbb{R}) \right\}, \tag{1.2}$$

which is translated by the multi-parameter flow

$$g_{\mathbf{t}} := \mathrm{diag}\left(e^{t_1}, \ldots, e^{t_m}, e^{-t_{m+1}}, \ldots, e^{-t_{m+n}}\right)$$

with $\sum_{i=1}^{m} t_i = \sum_{j=m+1}^{m+n} t_j$. In this case, the orbits of $U_{m,n}$ are unstable manifolds of $g_{\mathbf{t}}$ only when $t_1 = \cdots = t_m \to +\infty$ and $t_{m+1} = \cdots = t_{m+n} \to +\infty$, but not in general. The main result in [7] states that there exists $\delta > 0$, which is independent of σ, such that

$$\left|\sigma(\varphi \circ g_{\mathbf{t}}) - \mu(\varphi)\right| \le C(\sigma, \varphi)\, e^{-\delta\lfloor \mathbf{t} \rfloor} \tag{1.3}$$

for all $\varphi \in C_c^\infty(X)$, where

$$\lfloor \mathbf{t} \rfloor := \min(t_1, \ldots, t_{m+n}).$$

This result was generalized to homogeneous spaces of semisimple groups by Dabbs, Kelly, and Li [4] and by Shi [10]. Quantitative estimates on higher order correlations were also established in [10]: for every integer $r \ge 1$, there exists $\delta' = \delta'(r) > 0$ such that

$$\left|\sigma\left(\varphi_1 \circ g_{\mathbf{t}_1} \cdots \varphi_r \circ g_{\mathbf{t}_r}\right) - \mu(\varphi_1) \cdots \mu(\varphi_r)\right| \le C(\sigma, \varphi_1, \ldots, \varphi_r)\, e^{-\delta' D_r(\mathbf{t}_1, \ldots, \mathbf{t}_r)} \tag{1.4}$$

for all $\varphi_1, \ldots, \varphi_r \in C_c^\infty(X)$, where

$$D_r(\mathbf{t}_1, \ldots, \mathbf{t}_r) := \min\left(\lfloor \mathbf{t}_1 \rfloor, \lfloor \mathbf{t}_2 - \mathbf{t}_1 \rfloor, \ldots, \lfloor \mathbf{t}_r - \mathbf{t}_{r-1} \rfloor\right).$$

We stress that this estimate only provides non-trivial information when the vectors $\mathbf{t}_1, \ldots, \mathbf{t}_r$ are *completely ordered* with respect to the function $\lfloor \cdot \rfloor$ and *all* gaps with respect to this order go to infinity. This condition is too restrictive for some of the applications to counting problems in multiplicative Diophantine approximation that we have in mind [2].

In this paper, we generalize (a special case of) this result in two ways. Firstly, we show that one can reduce the regularity assumptions on the measure σ. Secondly, we establish a favourable estimate, which only depend on $\lfloor \mathbf{t}_1 \rfloor, \ldots, \lfloor \mathbf{t}_1 \rfloor$ and the pairwise Euclidean distances $\|\mathbf{t}_i - \mathbf{t}_j\|$ for $i \ne j$.

To formulate our first main result, we need some notation. Let Y be a compact orbit of $U_{m,n}$ in X. Then Y can be considered as a mn-dimensional torus, and we denote by m_Y the probability invariant measure on Y. We also write $\widehat{Y}$ for the character group of Y. Given a Borel measure σ on Y, we write $\hat{\sigma}(\chi) := \int_Y \chi \, d\sigma$, $\chi \in \widehat{Y}$, for the corresponding Fourier coefficients. We say that σ is a *Wiener measure* if

$$\|\sigma\|_W := \sum_{\chi \in \widehat{Y}} |\hat{\sigma}(\chi)| < \infty.$$

Note that every Wiener measure on Y is absolutely continuous with respect to m_Y, with a continuous (but possibly nowhere differentiable) density. We denote by S_d the family of norms introduced in Section 3.2 below. In particular, one can take them to be the C^k-norm on $C_c^\infty(X)$ for a fixed sufficiently large k.

Theorem 1.1. For every $r \geq 1$, there exist $d_r \in \mathbb{N}$ and $C_r, \delta_r > 0$ such that for every Wiener probability measure σ supported on a compact orbit of $U_{m,n}$ in X, $\mathbf{t}_1, \ldots, \mathbf{t}_r \in \mathbb{R}_+^{m+n}$, and $\varphi_1, \ldots, \varphi_r \in C_c^\infty(X)$,

$$\left|\sigma\left(\varphi_1 \circ g_{\mathbf{t}_1} \cdots \varphi_r \circ g_{\mathbf{t}_r}\right) - \mu(\varphi_1) \cdots \mu(\varphi_r)\right| \leq C_r \|\sigma\|_W S_{d_r}(\varphi_1) \cdots S_{d_r}(\varphi_r)\, e^{-\delta_r \Delta_r(\mathbf{t}_1,\ldots,\mathbf{t}_r)},$$

where

$$\Delta_r(\mathbf{t}_1, \ldots, \mathbf{t}_r) := \min\left(\lfloor \mathbf{t}_i \rfloor, \|\mathbf{t}_i - \mathbf{t}_j\| : 1 \leq i \neq j \leq r\right).$$

Remark 1.2. We stress that our proof techniques heavily use the assumption that the $U_{m,n}$-orbit is *compact*.

Let us now formulate a more general version of this result. Let G be a connected semisimple Lie group without compact factors and P a parabolic subgroup of G such that the projection of P to every simple factor of G is proper. Let U be the unipotent radical of P and A a maximal connected Ad-diagonalizable subgroup of P. We write $\mathfrak{a}$ and $\mathfrak{u}$ for the corresponding Lie algebras and consider the adjoint action of $\mathfrak{a}$ on $\mathfrak{u}$. Fix a norm on $\mathfrak{a}$ and let d denote the corresponding invariant distance function on A. Let $\Phi(\mathfrak{u})$ denote the set of roots (characters of $\mathfrak{a}$) arising in this action. Let $\mathfrak{a}_U^{*,+}$ be the positive cone in the dual $\mathfrak{a}^*$ spanned by the characters in $\Phi(\mathfrak{u})$. We use the usual identification $\mathfrak{a}^* \to \mathfrak{a}$ defined by the Killing form: $\alpha \mapsto s_\alpha$ given by $\langle s_\alpha, a\rangle = \alpha(a)$ for $a \in \mathfrak{a}$, and denote by $\mathfrak{a}_U^+$ the corresponding cone in $\mathfrak{a}$ given by this identification. The cone $A_U^+ := \exp(\mathfrak{a}_U^+)$ was introduced in [10] and is called the expanding cone for U. For $a \in A_U^+$, we denote by $\lfloor a \rfloor_U$ the distance from a to the boundary of the cone A_U^+. Let Γ be an irreducible lattice in G and $X := G/\Gamma$ equipped with the invariant probability measure μ.

The main result of [10] is a generalization of the estimate (1.4) to compactly supported smooth measures on an orbit of U in X. Let us additionally assume that U is abelian and Y is a compact orbit of U in X. Given a Borel measure σ on U, we define the Wiener norm as above. In this setting, we establish the following general version of Theorem 1.1:

Theorem 1.3. For every $r \geq 1$, there exist $d_r \in \mathbb{N}$ and $C_r, \delta_r > 0$ such that for every Wiener probability measure σ supported on Y, $t_1, \ldots, t_r \in A_U^+$, and $\varphi_1, \ldots, \varphi_r \in C_c^\infty(X)$,

$$\left|\sigma\left(\varphi_1 \circ t_1 \cdots \varphi_r \circ t_r\right) - \mu(\varphi_1) \cdots \mu(\varphi_r)\right| \leq C_r \|\sigma\|_W S_{d_r}(\varphi_1) \cdots S_{d_r}(\varphi_r)\, e^{-\delta_r \Delta_r(t_1,\ldots,t_r)},$$

where

$$\Delta_r(t_1,\ldots,t_r) := \min\left(\lfloor t_i\rfloor_U,\ d(t_i,t_j) :\ 1\le i\ne j\le r\right).$$

As we clarify below, our method does not rely on particular properties of the horospherical subgroup U and allows us to produce quantitative estimates on $\sigma(\varphi\circ t_1\cdots\varphi\circ t_r)$ once estimates on $\sigma(\varphi\circ t)$ are available (see Theorem 2.1).

2 General Result

Let G be a real Lie group, U a closed connected subgroup of G, and T a closed connected abelian subgroup of G that normalizes U such that the adjoint action of T on the Lie algebra of U is proper and diagonalizable. Let X be a standard Borel space equipped with measurable action of G. Let μ be a G-invariant probability measure on X, and let ν be a U-invariant and U-ergodic measure on X. We assume that the measure ν has discrete spectrum. This means that there is an orthonormal basis of $L^2(\nu)$ consisting of U-eigenfunctions and allows us to introduce a Wiener norm $\|\cdot\|_{W(\nu)}$ on $L^\infty(\nu)$ (see Section 3.1 below).

We shall further assume that certain equidistribution properties hold for ν, T and a sub-algebra $\mathcal{A}$ of bounded functions on X, which is equipped with a family of norms S_d (see Section 3.2 below). For example, $\mathcal{A}$ could be $\mathbb{C}+C_c^\infty(X)$ and S_d could be the C^k-norms for a fixed sufficiently large k.

Finally, we make the following two assumptions regarding quantitative equidistribution on X in terms of a fixed suitable norm S_{d_o}:

- There exist $T_+\subset T$, $D_o\ge 1$, $\delta_o\in(0,1)$, and a function $\rho: T_+\to[1,\infty)$ such that

$$\left|\nu(\varphi\circ t)-\mu(\varphi)\right|\le D_o\,\rho(t)^{-\delta_o}\,S_{d_o}(\varphi), \tag{EQ1}$$

  for all $\varphi\in\mathcal{A}$ and $t\in T_+$.
- There exist $C\ge 1$, and $0<c<1/2$ such that

$$\left|\mu(\varphi_1\circ\exp(w)\cdot\varphi_2)-\mu(\varphi_1)\mu(\varphi_2)\right|\le C\max\left(1,\|w\|\right)^{-c}S_{d_o}(\varphi_1)\,S_{d_o}(\varphi_2) \tag{EQ2}$$

  for all $\varphi_1,\varphi_2\in\mathcal{A}$ and $w\in\mathrm{Lie}(U)$.

We fix an invariant metric d on T. Given $r\ge 2$ and $\underline{t}=(t_1,\ldots,t_r)\in T_+^r$, we define

$$\rho_r(\underline{t}) := \min\left(\rho(t_1),\ldots,\rho(t_r)\right)\quad\text{and}\quad m_r(\underline{t}) := \min_{i\ne j}\exp(d(t_i,t_j)), \tag{2.1}$$

and for $r \geq 1$, we set

$$\Delta_r(\underline{t}) := \begin{cases} \rho(t_1) & \text{if } r = 1 \\ \min\left(\rho_r(\underline{t}), m_r(\underline{t})\right) & \text{if } r \geq 2. \end{cases} \tag{2.2}$$

Later on in our argument, we shall slightly modify the definition of m_r (see (3.8) below).

Our main theorem provides quantitative estimates on r-correlations in terms of Δ_r:

Theorem 2.1. For every $r \geq 1$, there exist $d_r \in \mathbb{N}$ and $D_r, \delta_r > 0$ such that

$$\left|\nu\Big(\varphi_o \prod_{i=1}^{r} \varphi_i \circ t_i\Big) - \nu(\varphi_o) \prod_{i=1}^{r} \mu(\varphi_i)\right| \leq D_r \Delta_r(\underline{t})^{-\delta_r} \|\varphi_o\|_{W(\nu)} \prod_{i=1}^{r} S_{d_r}(\varphi_i),$$

for all $\underline{t} = (t_1, \ldots, t_r) \in T_+^r$, $\varphi_o \in W(\nu)$, and $\varphi_1, \ldots, \varphi_r \in \mathcal{A}$.

We note that Theorem 1.3 follows immediately from Theorem 2.1. Indeed, every Wiener measure on a torus has a continuous density, and the assumptions (EQ1) and (EQ2) can been verified in this setting. In particular, (EQ1) was established in [11], and (EQ2) is the well-known exponential mixing estimate (see, for instance, [6, Corollary 2.4.4]).

In Section 8 below, we also work out the parameters d_r, D_r, δ_r in Theorem 2.1 explicitly.

3 Notations

3.1 The Wiener algebra

We recall that the measure ν is assumed to have discrete spectrum with respect to U, that is, $L^2(\nu)$ has an orthonormal basis consisting of U-eigenfunctions: there exists a set Ξ of unitary characters on $\mathrm{Lie}(U)$ and an orthonormal basis $\{\psi_\xi \,:\, \xi \in \Xi\}$ of $L^2(\nu)$ such that

$$\psi_\xi \circ \exp(w) = \xi(w)\psi_\xi, \quad \text{for all } \xi \in \Xi \text{ and } w \in \mathrm{Lie}(U), \tag{3.1}$$

where the identity is interpreted in the $L^2(\nu)$-sense. Without loss of generality, $\psi_1 = 1$. Furthermore, for every $\xi \in \Xi$,

$$|\psi_\xi| = 1, \quad \nu\text{-almost everywhere}, \tag{3.2}$$

by the U-ergodicity of ν.

We denote by $B(X)$ the $*$-algebra of bounded complex-valued measurable functions on X under pointwise multiplication. The supremum norm on $B(X)$ is denoted by $\|\cdot\|_\infty$. For $\varphi \in B(X)$, we set

$$\|\varphi\|_{W(\nu)} := \sum_{\xi \in \Xi} |\nu(\varphi \cdot \overline{\psi}_\xi)|, \tag{3.3}$$

where $\{\psi_\xi : \xi \in \Xi\}$ is a fixed orthonormal basis of $L^2(\nu)$.

We define the *Wiener algebra*

$$W(\nu) := \left\{\varphi \in B(X) \;:\; \|\varphi\|_{W(\nu)} < \infty\right\}.$$

For $\varphi \in W(\nu)$,

$$\varphi = \sum_{\xi \in \Xi} \nu(\varphi \cdot \overline{\psi}_\xi)\psi_\xi, \tag{3.4}$$

with convergence in the $L^\infty(\nu)$-sense, In particular,

$$\|\varphi\|_{L^\infty(\nu)} \le \|\varphi\|_{W(\nu)}, \quad \text{for all } \varphi \in W(\nu). \tag{3.5}$$

3.2 A family of norms on the space X

Let $\mathcal{A}$ be a G-invariant sub-$*$-algebra of the algebra $B(X)$. We assume that $\mathcal{A}$ is equipped with an increasing family of norms

$$S_1 \le S_2 \le \cdots \le S_d \le \cdots$$

satisfying the following assumptions for fixed d_o and for all d:

- For all $\varphi \in \mathcal{A}$,

$$\|\varphi\|_\infty \le S_{d_o}(\varphi). \tag{S1}$$

- There exist $A \ge 1$ and $a > 0$ such that

$$\left\|\varphi \circ \exp(w) - \varphi\right\|_\infty \le A\,\|w\|^a\, S_{d_o}(\varphi) \quad \text{for all } \varphi \in \mathcal{A} \text{ and } w \in \mathrm{Lie}(U). \tag{S2}$$

- There exist $B_d \ge 1$ and $b_d > \max(1/2, a/4)$ such that

$$S_d\big(\varphi \circ \exp(w)\big) \le B_d\, \max\left(1, \|w\|\right)^{b_d} S_d(\varphi) \quad \text{for all } \varphi \in \mathcal{A} \text{ and } w \in \mathrm{Lie}(U). \tag{S3}$$

- There exists $M_d \ge 1$ such that

$$S_d(\varphi_1\varphi_2) \le M_d\, S_{d+d_o}(\varphi_1) S_{d+d_o}(\varphi_2), \quad \text{for all } \varphi_1, \varphi_2 \in \mathcal{A}. \tag{S4}$$

Since $\mathcal{A}$ is a G-invariant subalgebra, for $\varphi_1, \dots, \varphi_r \in \mathcal{A}$ and $g_1, \dots, g_r \in G$, we also have that the product $\prod_{i=1}^r \varphi_i \circ g_i \in \mathcal{A}$. We further assume that the norms S_d are convex in the following strong sense:

- For every $\varphi_1, \ldots, \varphi_r \in \mathcal{A}$, $g_1, \ldots, g_r \in G$, and compactly supported complex finite measures ω on G^r,

$$\phi := \int_{G^r} \left(\prod_{i=1}^{r} \varphi_i \circ g_i \right) \mathrm{d}\omega(g_1, \ldots, g_r) \in \mathcal{A}, \tag{S5}$$

and

$$S_d(\phi) \leq \int_{G^r} S_d \left(\prod_{i=1}^{r} \varphi_i \circ g_i \right) \mathrm{d}\omega(g_1, \ldots, g_r). \tag{S6}$$

We now provide several examples of norms for which our set-up applies when X is a finite-volume homogeneous space of a Lie group G and $\mathcal{A} = \mathbb{C} + C_c^\infty(X)$:

(i) The simplest example to which our framework applies is a fixed C^k-norm with $k \geq 1$:

$$S(\phi) := \max \|\phi\|_{C^k}.$$

In this case, there is no dependence on the index d. We note that here and in the other examples, the properties (S5)–(S6) can be verified using the dominated convergence theorem.

(ii) For some applications, one is required to approximate unbounded functions on X. Then the following refined norms are useful. Let $\|\cdot\|_{Lip}$ denote the Lipschitz norm $\mathcal{A}$ with respect to an invariant Riemannian metric and let $\|\cdot\|_{L_k^p}$ denote the L^p-Sobolev norm of order k. One can take the family of norms

$$S_d(\phi) := \max \left(\|\phi\|_{C^0}, \|\phi\|_{Lip}, \|\phi\|_{L_k^{2^d}} \right). \tag{3.6}$$

In this case, property (S3) holds with fixed B_d, b_d depending only k, and property (S4) with M_d depending only on k and with $d_o = 1$ follows from the Cauchy–Schwarz inequality.

(iii) The third example is the Sobolev norms used in [1] (see [1, Subsection 2.2]). In this case, d denotes the degree of the Sobolev norm, property (S3) holds with $B_d = L_1^d$ and $b_d = \ell d$ for some $L_1, \ell \geq 1$, and property (S4) holds with $M_d = L_2^d$ for some $L_2 \geq 1$.

Let us further assume that $X = G/\Gamma$, where G is connected semisimple Lie group without compact factors and Γ is an irreducible lattice in G. Then property (EQ2) is the well-known exponential mixing estimate (see, for instance, [6, Corollary 2.4.4]), where the bound involves the L^2_k-Sobolev norms of the test functions. Property (EQ2) has been verified in [3, 7, 10] with respect to the norm S_d defined in (3.6) (see, for instance, [3, Th. 2.2]).

3.3 Norms on the group T

According to our assumptions on T and $\mathfrak{u} := \mathrm{Lie}(U)$, there exists a subset $\Phi \subset \mathrm{Hom}(T, \mathbb{R}^*)$ such that

$$\mathfrak{u} = \bigoplus_{\alpha \in \Phi} \mathfrak{u}_\alpha, \quad \text{with } \mathfrak{u}_\alpha := \left\{ w \in \mathfrak{u} \,:\, \mathrm{Ad}(t)w = \alpha(t)w, \ \text{ for all } t \in T \right\}.$$

We choose a basis of each subspace $\mathfrak{u}_\alpha$. This gives a basis of $\mathfrak{u}$. For $w \in \mathfrak{u}$, we denote by $\|w\|$ the ℓ^∞-norm with respect to this basis. Given $t \in T$, define

$$\begin{aligned} \|t\|_* &:= \max\left\{ \max\left(\|\mathrm{Ad}(t)w\|, \|\mathrm{Ad}(t)^{-1}w\| \right) \,:\, \|w\| = 1 \right\} \\ &= \max\left\{ \max\left(|\alpha(t)|, |\alpha(t)|^{-1} \right) \,:\, \alpha \in \Phi \right\}. \end{aligned} \tag{3.7}$$

Note that $\|t^{-1}\|_* = \|t\|_*$ and $\|t\|_* \geq 1$ for all $t \in T$. Since the action of T on $\mathfrak{u}$ is proper, $\log \|\cdot\|_*$ defines a norm on $\mathrm{Lie}(T)$. Then the invariant metric $d(t_1, t_2)$ on T is comparable up to a constant to $\log \|t_1 t_2^{-1}\|_*$, so that it is sufficient to establish the estimates in terms of $\|\cdot\|_*$. From now on, we redefine m_r from (2.1) as

$$m_r(\underline{t}) := \min_{i \neq j} \|t_i t_j^{-1}\|_*, \quad \text{for } \underline{t} = (t_1, \ldots, t_r) \in T_+^r. \tag{3.8}$$

4 Proof of the Main Theorems

We fix $r \geq 1$, $\varphi_o \in W(\nu)$, $\varphi_1, \ldots, \varphi_r \in \mathcal{A}$ and $\underline{t} = (t_1, \ldots, t_r) \in T_+^r$. We wish to estimate the expression

$$\left| \nu\Big(\varphi_o \prod_{i=1}^r \varphi_i \circ t_i \Big) - \nu(\varphi_o) \prod_{i=1}^r \mu(\varphi_i) \right|.$$

4.1 Step I: expanding the function φ_o

By (3.4) we can write

$$\varphi_o = \sum_{\xi \in \Xi} \nu(\varphi_o \cdot \overline{\psi}_\xi) \psi_\xi,$$

where the series converges uniformly, and with the convention that $\psi_1 = 1$. Then

$$\nu\Big(\varphi_o \prod_{i=1}^{r} \varphi_i \circ t_i\Big) - \nu(\varphi_o) \prod_{i=1}^{r} \mu(\varphi_i) = \nu(\varphi_o)\Big(\nu\Big(\prod_{i=1}^{r} \varphi_i \circ t_i\Big) - \prod_{i=1}^{r} \mu(\varphi_i)\Big) + \sum_{\xi \neq 1} \nu(\varphi_o \cdot \overline{\psi}_\xi)\, \nu_\xi\Big(\prod_{i=1}^{r} \varphi_i \circ t_i\Big), \tag{4.1}$$

where we used the notation

$$\nu_\xi(\eta) := \nu(\psi_\xi \cdot \eta), \quad \text{for } \eta \in B(X). \tag{4.2}$$

Let us now define

$$D_{d,r}(\underline{t}; 1) := \sup \left\{ \Big| \nu\Big(\prod_{i=1}^{r} \varphi_i \circ t_i\Big) - \prod_{i=1}^{r} \mu(\varphi_i) \Big| \;:\; \varphi_1, \ldots, \varphi_r \in \mathcal{A},\;\; S_d(\varphi_1), \ldots, S_d(\varphi_r) \le 1 \right\}, \tag{4.3}$$

and, for $\xi \neq 1$,

$$D_{d,r}(\underline{t}; \xi) := \sup \left\{ \Big| \nu_\xi\Big(\prod_{i=1}^{r} \varphi_i \circ t_i\Big) \Big| \;:\; \varphi_1, \ldots, \varphi_r \in \mathcal{A},\;\; S_d(\varphi_1), \ldots, S_d(\varphi_r) \le 1 \right\}. \tag{4.4}$$

Finally, we set

$$E_{d,r}(\underline{t}) := \sup_{\xi \in \Xi} D_{d,r}(\underline{t}; \xi). \tag{4.5}$$

Then it follows from (4.1) that

$$\Big| \nu\Big(\varphi_o \prod_{i=1}^{r} \varphi_i \circ t_i\Big) - \nu(\varphi_o) \prod_{i=1}^{r} \mu(\varphi_i) \Big| \le E_{d,r}(\underline{t})\, \|\varphi_o\|_{W(\nu)} \prod_{i=1}^{r} S_d(\varphi_i). \tag{4.6}$$

Our goal from now on will be to estimate the quantity $E_{d,r}(\underline{t})$. This will be established through an elaborate induction scheme, so it will be convenient to define

$$\widetilde{E}_{d,r-1}(\underline{t}) := \max \Big\{ E_{d,p}(t_{i_1}, \ldots, t_{i_p}) \;:\; 1 \le p < r,\;\; \{i_1, \ldots, i_p\} \subset \{1, \ldots, r\} \Big\}. \tag{4.7}$$

4.2 Step II: an upper bound on E_1 (base of induction)

In this section, we prove the Theorem 2.1 when $r = 1$. The assumption (EQ1) asserts that

$$D_{d_o,1}(t;1) \le D_o\, \rho(t)^{-\delta_o}, \quad \text{for all } t \in T_+. \tag{4.8}$$

We aim to estimate $E_{d_1,1}(t)$ for some suitably chosen $d_1 > d_o$. In view of (4.5), it suffices to bound

$$E'_{d_1,1}(t) := \sup_{\xi \neq 1} D_{d_1,1}(t;\xi).$$

To do this we shall exploit the following U-equivariance of the complex measures ν_ξ, in combination with the equidistribution assumption (EQ2).

Lemma 4.1. For all $\xi \in \Xi$ and $w \in \mathfrak{u}$,

$$\exp(w)_* \nu_\xi = \xi(-w)\nu_\xi.$$

Proof. Indeed, it follows from (3.1) that for every $\varphi \in B(X)$,

$$\begin{aligned}\exp(w)_*\nu_\xi(\varphi) &= \nu_\xi(\varphi\circ\exp(w)) = \nu(\psi_\xi\cdot\varphi\circ\exp(w)) = \nu(\psi_\xi\circ\exp(-w)\cdot\psi)\\ &= \xi(-w)\nu(\psi_\xi\cdot\varphi) = \xi(-w)\nu_\xi(\varphi).\end{aligned}$$

■

Let $\varphi \in \mathcal{A}$, $w \in \mathfrak{u}\setminus\{0\}$, $t \in T_+$ and $\xi \neq 1$. Using Lemma 4.1, we conclude that for every $s \in \mathbb{R}$,

$$\begin{aligned}\nu_\xi(\varphi\circ t) &= \xi(sw)\nu_\xi(\varphi\circ t\circ\exp(sw)) = \xi(sw)\nu_\xi(\varphi\circ\exp(s\mathrm{Ad}(t)w)\circ t)\\ &= \xi(sw)\nu_\xi\big(\varphi\circ\exp(s\mathrm{Ad}(t)w)\circ t - \mu(\varphi)\big),\end{aligned}$$

where we used that $\nu_\xi(1) = \nu(\psi_\xi) = 0$ when $\xi \neq 1$. For $L > 0$, we set

$$\phi_L := \frac{1}{L}\int_0^L \xi(sw)\Big(\varphi\circ\exp(s\mathrm{Ad}(t)w) - \mu(\varphi)\Big)\, ds. \tag{4.9}$$

Then

$$\nu_\xi\big(\varphi\circ t\big) = \nu_\xi(\phi_L\circ t). \tag{4.10}$$

Using that $|\psi_\xi| = 1$ ν-almost everywhere, we deduce that

$$\begin{aligned}|\nu_\xi(\phi_L \circ t)| &\le \nu(|\phi_L| \circ t) \le \nu(|\phi_L|^2 \circ t)^{1/2} \\ &\le \mu(|\phi_L|^2)^{1/2} + \left|\nu(|\phi_L|^2 \circ t) - \mu(|\phi_L|^2)\right|^{1/2},\end{aligned} \tag{4.11}$$

where we used the inequality $\alpha^{1/2} \le \beta^{1/2} + |\alpha - \beta|^{1/2}$ with $\alpha, \beta \ge 0$.

The required bounds on the two terms in (4.11) is provided by the following two lemmas:

Lemma 4.2. Let $c_L := \mu(\varphi) \cdot \frac{1}{L} \int_0^L \xi(sw)\, ds$.

(*i*) $\eta_L := |\phi_L|^2 - |c_L|^2 \in \mathcal{A}$.

(*ii*) If $L\|\mathrm{Ad}(t)w\| \ge 1$, then

$$S_{d_o}(\eta_L) \le B'_{d_0}\left(L\|\mathrm{Ad}(t)w\|\right)^{2b_{2d_o}} S_{2d_o}(\varphi)^2,$$

where $B'_{d_0} := M_{d_o} B^2_{2d_o} + 2B_{d_o}$.

Lemma 4.3. If $L\|\mathrm{Ad}(t)w\| \ge 1$, then

$$\mu\left(|\phi_L|^2\right) \le 14C\left(L\|\mathrm{Ad}(t)w\|\right)^{-c} S_{d_o}(\varphi)^2,$$

where C and c are the positive constants which are defined in (EQ2).

We postpone the proofs of the lemmas until Section 5 and continue with the estimate (4.11). We note that since both ν and μ are probability measures and c_L is a constant, we have

$$\left|\nu(|\phi_L|^2 \circ t) - \mu(|\phi_L|^2)\right| = \left|\nu(\eta_L \circ t) - \mu(\eta_L)\right|.$$

Hence, using (EQ1) and Lemma 4.2, we deduce that

$$\begin{aligned}\left|\nu(|\phi_L|^2 \circ t) - \mu(|\phi_L|^2)\right| &\le D_o\, \rho(t)^{-\delta_o} S_{d_o}(\eta_L) \\ &\le D_o B'_{d_o}\left(L\|\mathrm{Ad}(t)w\|\right)^{2b_{2d_o}} \rho(t)^{-\delta_o} S_{2d_o}(\varphi)^2,\end{aligned}$$

for all $L > 0$ such that $L\|\mathrm{Ad}(t)w\| \geq 1$. Combining this estimate with Lemma 4.3, we deduce from (4.11) that

$$\left|\nu_\xi(\varphi \circ t)\right| = \left|\nu_\xi(\phi_L \circ t)\right| \leq \left(4\sqrt{C}\big(L\|\mathrm{Ad}(t)w\|\big)^{-c/2} + \sqrt{D_o B'_{d_o}}\big(L\|\mathrm{Ad}(t)w\|\big)^{b_{2d_o}} \rho(t)^{-\delta_o/2}\right) S_{2d_o}(\varphi),$$

for all $L > 0$ such that $L\|\mathrm{Ad}(t)w\| \geq 1$.

It remains to find a suitable $L > 0$ to ensure that the right-hand side in the last inequality decays like an inverse power of $\rho(t)$. If we choose $L > 0$ so that

$$L\|\mathrm{Ad}(t)w\| = \rho(t)^{\delta_o/(c+2b_{2d_o})}, \tag{4.12}$$

then, since $\rho(t) \geq 1$, we see that $L\|\mathrm{Ad}(t)w\| \geq 1$, and thus the previous bounds are available. We conclude that

$$\left|\nu_\xi(\varphi \circ t)\right| \leq D'_1\, \rho(t)^{-\delta_1}\, S_{2d_o}(\varphi),$$

where

$$D'_1 := 5 \max\left(\sqrt{C}, \sqrt{D_o B'_{d_o}}\right) \quad \text{and} \quad \delta_1 := \frac{c\delta_o}{2(c + 2b_{2d_o})} < \delta_o < 1.$$

Since this upper bound is uniform over all $\xi \neq 1$, and the constants are independent of $t \in T_+$, we obtain that

$$E'_{2d_o,1}(t) \leq D'_1\, \rho(t)^{-\delta_1}, \quad \text{for all } t \in T_+.$$

Combining this bound with (4.8), we finally deduce that

$$E_{2d_o,1}(t) \leq D_1\, \rho(t)^{-\delta_1}, \quad \text{for all } t \in T_+, \tag{4.13}$$

where $D_1 := \max(D_o, D'_1)$.

4.3 Step III: choosing a suitable one-parameter subgroup

While the estimates in Step II involve averaging along an arbitrarily chosen one-parameter subgroup of U, we will have to choose this subgroup more thoughtfully to handle higher order correlations. We will carry out this task now.

Let $r \geq 2$ and $\underline{t} = (t_1, \ldots, t_r) \in T_+^r$. We define

$$M_r(\underline{t}) := \max_{i,j} \|t_i t_j^{-1}\|_*, \tag{4.14}$$

where $\|\cdot\|_*$ is given by (3.7). By the definition of $\|\cdot\|_*$, we can find indices $i, j = 1, \ldots, r$ and $\alpha \in \Phi$ such that

$$M_r(\underline{t}) = \|t_i t_j^{-1}\|_* = |\alpha(t_i t_j^{-1})| = \left\|\mathrm{Ad}(t_i t_j^{-1}) e_{\alpha,1}\right\|,$$

where $e_{\alpha,1}$ belongs to the (fixed) basis of $\mathfrak{u}_\alpha$. We set $i_1 = i$ and pick indices $i_2, \ldots, i_r$, all distinct, such that

$$\left\|\mathrm{Ad}(t_{i_1} t_j^{-1}) e_{\alpha,1}\right\| \geq \left\|\mathrm{Ad}(t_{i_2} t_j^{-1}) e_{\alpha,1}\right\| \geq \cdots \geq \left\|\mathrm{Ad}(t_{i_r} t_j^{-1}) e_{\alpha,1}\right\|. \tag{4.15}$$

In particular, there exists an index $l = 1, \ldots, r$ such that $i_l = j$, and thus

$$1 = \left\|\mathrm{Ad}(t_{i_l} t_j^{-1}) e_{\alpha,1}\right\| \geq \left\|\mathrm{Ad}(t_{i_r} t_j^{-1}) e_{\alpha,1}\right\|. \tag{4.16}$$

Since the expression $E_{d,r}(\underline{t})$ that we wish to estimate is invariant under permutations of the elements in $\underline{t}$, we may henceforth without loss of generality adopt the following convention:

The indices are relabelled so that $i_k = k$ for $k = 1, \ldots, r$.

Assuming this convention, we set

$$w := \mathrm{Ad}(t_l^{-1}) e_{\alpha,1} \quad \text{and} \quad w^{(i)} := \mathrm{Ad}(t_i) w, \quad \text{for } i = 1, \ldots, r. \tag{4.17}$$

We note that

$$\|w^{(1)}\| = M_r(\underline{t}) \quad \text{and} \quad \|w^{(r)}\| \leq 1,$$

and it follows from (4.15) and (4.16) that

$$\|w^{(1)}\| \geq \cdots \geq \|w^{(r)}\| \quad \text{and} \quad \|w^{(r)}\| \leq M_r(\underline{t})^{-1} \|w^{(1)}\|. \tag{4.18}$$

We stress that the r-tuple $w^{(1)}, \ldots, w^{(r)}$ is uniquely determined up to permutations of indices by $\underline{t}$.

4.4 Step IV: a recursive estimate of E_r in terms of $\widetilde{E}_{r-1}$ (inductive step)

Let $r \geq 2$, $\underline{t} = (t_1, \dots, t_r) \in T_+^r$, and $\varphi_1, \dots, \varphi_r \in \mathcal{A}$. For $\xi \in \Xi$, we need to estimate

$$\nu_\xi\Big(\prod_{i=1}^{r} \varphi_i \circ t_i\Big) = \nu\Big(\psi_\xi \prod_{i=1}^{r} \varphi_i \circ t_i\Big).$$

We exploit the invariance of the measure ν under U and consider the one-parameter subgroup $\exp(sw)$, where w is determined by the tuple $\underline{t}$ and is defined in (4.17). Then

$$\nu_\xi\Big(\prod_{i=1}^{r} \varphi_i \circ t_i\Big) = \nu\Big(\psi_\xi \circ \exp(sw) \prod_{i=1}^{r} \varphi_i \circ t_i \circ \exp(sw)\Big) = \nu_\xi\Big(\xi(sw) \prod_{i=1}^{r} \varphi_i \circ \exp(sw^{(i)}) \circ t_i\Big),$$

where $w^{(i)} := \mathrm{Ad}(t_i)w$. Therefore, for any $L > 0$,

$$\nu_\xi\Big(\prod_{i=1}^{r} \varphi_i \circ t_i\Big) = \nu_\xi\Big(\frac{1}{L}\int_0^L \xi(sw) \prod_{i=1}^{r} \varphi_i \circ \exp(sw^{(i)}) \circ t_i \, \mathrm{d}s\Big). \tag{4.19}$$

Our argument uses induction on the number of factors r. For an index $1 \leq p < r$, we set

$$I_{1,\xi} := \nu_\xi\Big(\prod_{i=1}^{r} \varphi_i \circ t_i\Big) - \nu_\xi\left(\Big(\frac{1}{L}\int_0^L \xi(sw) \prod_{i=1}^{p} \varphi_i \circ \exp\left(sw^{(i)}\right) \circ t_i \, \mathrm{d}s\Big) \prod_{i=p+1}^{r} \varphi_i \circ t_i\right), \tag{4.20}$$

$$I_{2,\xi} := \nu_\xi\left(\Big(\frac{1}{L}\int_0^L \xi(sw)\Big(\prod_{i=1}^{p} \varphi_i \circ \exp\left(sw^{(i)}\right) \circ t_i - \prod_{i=1}^{p} \mu(\varphi_i)\Big) \mathrm{d}s\Big) \prod_{i=p+1}^{r} \varphi_i \circ t_i\right), \tag{4.21}$$

$$I_{3,\xi} := \begin{cases} \prod_{i=1}^{p} \mu(\varphi_i)\,\left(\nu\Big(\prod_{i=p+1}^{r} \varphi_i \circ t_i\Big) - \prod_{i=p+1}^{r} \mu(\varphi_i)\right) & \text{if } \xi = 1 \\ \prod_{i=1}^{p} \mu(\varphi_i)\, \nu_\xi\Big(\prod_{i=p+1}^{r} \varphi_i \circ t_i\Big) & \text{if } \xi \neq 1 \end{cases}. \tag{4.22}$$

Then

$$I_{1,\xi} + I_{2,\xi} + \Big(\frac{1}{L}\int_0^L \xi(sw)\,\mathrm{d}s\Big) I_{3,\xi} = \begin{cases} \nu\Big(\prod_{i=1}^{r} \varphi_i \circ t_i\Big) - \prod_{i=1}^{r} \mu(\varphi_i) & \text{if } \xi = 1 \\ \nu_\xi\Big(\prod_{i=1}^{r} \varphi_i \circ t_i\Big) & \text{if } \xi \neq 1 \end{cases}.$$

Since $|\xi| = 1$, we see that when $S_d(\phi_1), \dots, S_d(\phi_r) \leq 1$,

$$E_{d,r}(\underline{t}) \leq \sup_{\xi \in \Xi} |I_{1,\xi}| + \sup_{\xi \in \Xi} |I_{2,\xi}| + \sup_{\xi \in \Xi} |I_{3,\xi}|. \tag{4.23}$$

We thus wish to prove that each term in (4.23) is small, provided that $L > 0$ and the index $p \in [1, r)$ are chosen appropriately. An important ingredient towards achieving this is the following technical proposition, whose proof we postpone until Section 6.

Proposition 4.4. For every $L > 0$ and $1 \leq p < r$ such that $L\|w^{(p)}\| \geq 1$,

(I) $\sup_{\xi\in\Xi} |I_{1,\xi}| \leq rA\big(L\|w^{(p+1)}\|\big)^{a} \prod_{i=1}^{r} S_{d_o}(\varphi_i)$.

(II) $\sup_{\xi\in\Xi} |I_{2,\xi}| \leq P_1\Big(\big(P_d L\|w^{(1)}\|\big)^{rb_{d+d_o}} D_{d,p}(t_1,\dots,t_p;1)^{1/2} + \sqrt{r}\big(L\|w^{(p)}\|\big)^{-c/2}\Big)$ $\prod_{i=1}^{r} S_{d+d_o}(\varphi_i)$, where $P_1 := \sqrt{14C}$ and $P_d := \big(M_d B_{d+d_o}^2 + 2B_d^2\big)^{1/2b_{d+d_o}}$.

(III) $\sup_{\xi\in\Xi} |I_{3,\xi}| \leq E_{d,r-p}(t_{p+1},\dots,t_r) \prod_{i=1}^{r} S_d(\varphi_i)$ when $d \geq d_o$.

The key point of Proposition 4.4 is that it will eventually allow us to establish an upper bound $E_{d,r}(\underline{t})$ in terms of $\widetilde{E}_{d,r-1}(\underline{t})$. Using induction, we can then use our bound on $E_{d,1}$ (and thus on $\widetilde{E}_{d,1}$) from the previous steps to provide an upper bound on $E_{d,r}$.

4.5 Step V: minimize the bound with respect to p and L

Here we specify the parameters p and L for which estimates from Step IV will be applied. The following version of the pigeonhole principle will be useful (see also Subsection 2.6 in [1] for a similar application).

Lemma 4.5. Fix an integer $r \geq 2$ and a real number $\theta \in (0, 1)$. Then, for every r-tuple $(\beta_1, \dots, \beta_r)$ of non-negative real numbers, which satisfies

$$\beta_r \leq \dots \leq \beta_1 \quad \text{and} \quad \beta_r \leq \beta_1\,\theta,$$

there exist $1 \leq p \leq r-1$ and $0 \leq q \leq r-2$ such that

$$\beta_{p+1} < \beta_1\,\theta^{(q+1)/r} < \beta_1\,\theta^{q/r} \leq \beta_p.$$

Proof. Let $\gamma_q = \beta_1\,\theta^{q/r}$ for $0 \leq q \leq r$. Since $\theta \in (0, 1)$ and $\beta_r \leq \beta_1\theta$, we have

$$\beta_r \leq \beta_1\,\theta = \gamma_r < \dots < \gamma_1 < \gamma_0 = \beta_1,$$

and thus the r distinct (and linearly ordered) points $\gamma_0, \ldots, \gamma_{r-1}$ belong to the interval $(\beta_r, \beta_1]$. Let us partition this interval into $r-1$ half-open (possibly empty) intervals as

$$(\beta_r, \beta_1] = \bigsqcup_{p=1}^{r-1} (\beta_{p+1}, \beta_p].$$

Then, by the pigeonhole principle, there are two consecutive points that belong to same partition interval, that is, there exist $1 \le p \le r-1$ and $0 \le q \le r-2$ such that

$$\beta_{p+1} < \gamma_{q+1} < \gamma_q \le \beta_p,$$

which finishes the proof. ■

Let $\underline{t} \in T_+^r$. Throughout the rest of the argument, we assume that

$$M_r(\underline{t}) = \max_{i,j} \|t_i t_j^{-1}\|_* > 1,$$

or equivalently, that $t_i \neq t_j$ for at least one pair (i,j) of distinct indices. We shall apply Lemma 4.5 above to

$$\theta \in [M_r(\underline{t})^{-1}, 1) \quad \text{and} \quad \beta_i = \|w^{(i)}\|, \quad i = 1, \ldots, r,$$

where the elements $w^{(i)}$ are defined in (4.17). Note that (4.18) implies that the conditions in the lemma are satisfied. We conclude that there are indices $1 \le p < r-1$ and $0 \le q \le r-2$ such that

$$\|w^{(p+1)}\| < \|w^{(1)}\|\, \theta^{(q+1)/r} < \|w^{(1)}\|\, \theta^{q/r} \le \|w^{(p)}\|,$$

and we set

$$L := \|w^{(1)}\|^{-1} \theta^{-(q+1/2)/r}.$$

With this choice of L, we obtain

$$L\|w^{(1)}\| = \theta^{-(q+1/2)/r} \le \theta^{-1}, \tag{4.24}$$

$$L\|w^{(p)}\| = \|w^{(1)}\|^{-1}\, \theta^{-(q+1/2)/r} \|w^{(p)}\| \ge \theta^{-\frac{(q+1/2)}{r}}\, \theta^{\frac{q}{r}} = \theta^{-1/2r} > 1, \tag{4.25}$$

$$L\|w^{(p+1)}\| = \|w^{(1)}\|^{-1}\, \theta^{-(q+1/2)/r} \|w^{(p+1)}\| < \theta^{-(q+1/2)/r} \theta^{(q+1)/r} = \theta^{1/2r}. \tag{4.26}$$

We stress that these bounds are independent of the indices p and q.

Let us now utilize these bounds in combination with Proposition 4.4. From (4.26) and Proposition 4.4 (I), we see that

$$\sup_{\xi\in\Xi}|I_{1,\xi}| \le rA\theta^{a/2r}\prod_{i=1}^{r} S_{d_o}(\varphi_i).$$

Using (4.24), (4.25), and the trivial bound

$$D_{d,p}(t_1,\ldots,t_p;1) \le \widetilde{E}_{d,r-1}(\underline{t}),$$

we deduce from Proposition 4.4(II) that

$$\sup_{\xi\in\Xi}|I_{2,\xi}| \le P_1\left((P_d\theta^{-1})^{rb_{d+d_o}}\,\widetilde{E}_{d,r-1}(\underline{t})^{1/2} + \sqrt{r}\,\theta^{c/4r}\right)\prod_{i=1}^{r} S_{d+d_o}(\varphi_i).$$

Finally, since $E_{d,r-p}(t_{p+1},\ldots,t_r) \le \widetilde{E}_{d,r-1}(\underline{t}) \le 2$, we conclude from Proposition 4.4(III) that

$$\sup_{\xi\in\Xi}|I_{3,\xi}| \le \widetilde{E}_{d,r-1}(\underline{t})\prod_{i=1}^{r} S_d(\varphi_i) \le \sqrt{2}\,\widetilde{E}_{d,r-1}(\underline{t})^{1/2}\prod_{i=1}^{r} S_d(\varphi_i)$$

Finally, using (4.23), we conclude that when $S_{d+d_o}(\phi_1),\ldots,S_{d+d_o}(\phi_r)\le 1$,

$$\begin{aligned} E_{d+d_o,r}(\underline{t}) &\le \sup_{\xi\in\Xi}|I_{1,\xi}| + \sup_{\xi\in\Xi}|I_{2,\xi}| + \sup_{\xi\in\Xi}|I_{3,\xi}| \\ &\le 2P_1(P_d\theta^{-1})^{rb_{d+d_o}}\widetilde{E}_{d,r-1}(\underline{t})^{1/2} + r\,Q\,\theta^{c_1/r}, \end{aligned} \tag{4.27}$$

for all $\underline{t}\in T_+^r$, where $Q := 2\max(A,P_1)$ and $c_1 := \min(a/2, c/4)$. We recall that this bound holds under the standing assumptions that $M_r(\underline{t}) > 1$ and $\theta\in[M_r(\underline{t})^{-1},1)$.

4.6 Proof of Theorem 2.1

Given $r\ge 2$ and $\underline{t} = (t_1,\ldots,t_r)\in T_+^r$, we use the quantities $\rho_r(\underline{t})$, $m_r(\underline{t})$, and $\Delta_r(\underline{t})$ introduced in (2.1), (2.2), and (3.8). Throughout the computation, we assume that

$$\Delta_r(\underline{t}) > 1. \tag{4.28}$$

We note that for any exhaustion

$$\{t_{i_1}\}\subset\{t_{i_1},t_{i_2}\}\subset\ldots\subset\{t_{i_1},\ldots,t_{i_r}\},$$

where $t_{i_1},\ldots,t_{i_r}$ are distinct entries in the vector $\underline{t}$, we have

$$\Delta_1(t_{i_1}) \ge \Delta_2(t_{i_1},t_{i_2}) \ge \ldots \ge \Delta_r(\underline{t}) > 1. \tag{4.29}$$

In (4.13), we proved

$$E_{d_1,1}(t_i) \le D_1\, \rho(t_i)^{-\delta_1}, \quad \text{for all } i = 1, \dots, r,$$

where $d_1 := 2d_o$, and $D_1 \ge 1$ and $\delta_1 \in (0,1)$ are explicit constants. Hence, by (4.29),

$$\widetilde{E}_{d_1,1}(t_{i_1}, t_{i_2}) \le D_1\, \Delta_r(\underline{t})^{-\delta_1}, \tag{4.30}$$

for all indices i_1, i_2, where $\widetilde{E}_{d,r}$ is defined by (4.7).

We introduce the following inductive assumption:

Ind$r$: *There exist* $d_{r-1} \in \mathbb{N}$, $D_{r-1} \ge 1$ *and* $\delta_{r-1} \in (0,1)$ *such that*

$$\widetilde{E}_{d_{r-1},r-1}(t_1, \dots, t_r) \le D_{r-1}\, \Delta_r(\underline{t})^{-\delta_{r-1}}. \tag{4.31}$$

We note that (4.30) implies that the base of induction Ind_2 holds. In what follows, we shall use our recursive bound (4.27) (for suitable θ) to show that for every integer $r \ge 2$

$$\mathrm{Ind}r \implies \mathrm{Ind}r+1,$$

and provide explicit estimates for constants d_r, D_r, δ_r.

We verify the inductive step (4.31) under the assumption $\Delta_r(\underline{t}) > 1$. Let

$$\theta := \Delta_r(\underline{t})^{-\varepsilon_r} < 1$$

with a parameter $\varepsilon_r \in (0,1)$, which will be specified later. We have

$$M_r(\underline{t})^{-1} \le M_r(\underline{t})^{-\varepsilon_r} \le m_r(\underline{t})^{-\varepsilon_r} \le \Delta_r(\underline{t})^{-\varepsilon_r} = \theta,$$

and thus $\theta \in [M_r(\underline{t})^{-1}, 1)$, so the inequality (4.27) can be applied. Combining (4.31) with (4.27), we conclude that

$$\begin{aligned} E_{d_{r-1}+d_o,r}(\underline{t}) &\le 2P_1 (P_{d_{r-1}} \Delta_r(\underline{t})^{\varepsilon_r})^{rb_{d_{r-1}+d_o}} \widetilde{E}_{d_{r-1},r-1}(\underline{t})^{1/2} + r\,Q \Delta_r(\underline{t})^{-\varepsilon_r c_1/r} \\ &\le 2P_1 (P_{d_{r-1}} \Delta_r(\underline{t})^{\varepsilon_r})^{rb_{d_{r-1}+d_o}} D_{r-1}^{1/2} \Delta_r(\underline{t})^{-\delta_{r-1}/2} + r\,Q \Delta_r(\underline{t})^{-\varepsilon_r c_1/r} \\ &\le 2P_1 P_{d_{r-1}}^{rb_{d_{r-1}+d_o}} D_{r-1}^{1/2}\, \Delta_r(\underline{t})^{-(\delta_{r-1} - 2\varepsilon_r r b_{d_{r-1}+d_o})/2} + r\,Q \Delta_r(\underline{t})^{-\varepsilon_r c_1/r}. \end{aligned}$$

Let us choose

$$d_r := d_{r-1} + d_o \quad \text{and} \quad \varepsilon_r := \frac{\delta_{r-1}}{\frac{2c_1}{r} + 2rb_{d_{r-1}+d_o}}, \tag{4.32}$$

so that the two exponents in the last expression match. Since $b_{d_{r-1}+d_o} > 1/2$ and $\delta_{r-1} \in (0,1)$, we have $\varepsilon_r \in (0,1)$. Then we obtain

$$E_{d_r,r}(\underline{t}) \leq D_r \Delta_r(\underline{t})^{-\delta_r},$$

where

$$D_r := 2P_1 P_{d_{r-1}}^{rb_{d_{r-1}+d_o}} D_{r-1}^{1/2} + rQ \quad \text{and} \quad \delta_r := \frac{c_1 \delta_{r-1}}{r(\frac{2c_1}{r} + 2rb_{d_{r-1}+d_o})} < \delta_{r-1}, \tag{4.33}$$

provided that $\Delta_r(\underline{t}) > 1$. Finally, we note that $D_r \geq 2$ and $E_r \leq 2$ on T_+^r, so the last inequality holds trivially if $\Delta_r(\underline{t}) = 1$. This completes the proof of Theorem 2.1.

5 Proof of Lemmas 4.2 and 4.3

Proof of Lemma 4.2. To prove (i), we note that upon expanding the square, we have

$$\begin{aligned} |\phi_L|^2 &= \frac{1}{L^2} \int_0^L \int_0^L \xi((s_1 - s_2)w) \big(\varphi \circ \exp(s_1 \mathrm{Ad}(t)w) - \mu(\varphi)\big) \big(\overline{\varphi \circ \exp(s_2 \mathrm{Ad}(t)w)} - \overline{\mu(\varphi)}\big)\, \mathrm{d}s_1 \mathrm{d}s_2 \\ &= R_{L,1} + R_{L,2} + |c_L|^2, \end{aligned}$$

where

$$R_{L,1} := \frac{1}{L^2} \int_0^L \int_0^L \xi((s_1 - s_2)w) \cdot \varphi \circ \exp(s_1 \mathrm{Ad}(t)w) \cdot \overline{\varphi \circ \exp(s_2 \mathrm{Ad}(t)w)}\, \mathrm{d}s_1 \mathrm{d}s_2,$$

and

$$R_{L,2} := -2\,\mathrm{Re}\left(\overline{c_L} \cdot \frac{1}{L} \int_0^L \xi(sw) \cdot \varphi \circ \exp(s\ \mathrm{Ad}(t)w)\, \mathrm{d}s\right).$$

It readily follows from our assumption (S5) that $R_{L,1}, R_{L,2} \in \mathcal{A}$, so that

$$\eta_L = |\phi_L|^2 - |c_L|^2 = R_{L,1} + R_{L,2} \in \mathcal{A}.$$

This gives (i).

To prove (ii), we first note that our assumptions (S6) and (S4) imply that

$$\begin{aligned} S_{d_o}(R_{L,1}) &\leq \frac{1}{L^2} \int_0^L \int_0^L S_{d_o}\big(\varphi \circ \exp(s_1\ \mathrm{Ad}(t)w) \cdot \overline{\varphi \circ \exp(s_2\ \mathrm{Ad}(t)w)}\big)\, \mathrm{d}s_1 \mathrm{d}s_2 \\ &\leq \frac{M_{d_o}}{L^2} \int_0^L \int_0^L S_{2d_o}\big(\varphi \circ \exp(s_1\ \mathrm{Ad}(t)w)\big)\, S_{2d_o}\big(\overline{\varphi \circ \exp(s_2\ \mathrm{Ad}(t)w)}\big)\, \mathrm{d}s_1 \mathrm{d}s_2 \\ &= M_{d_o} \left(\frac{1}{L} \int_0^L S_{2d_o}\big(\varphi \circ \exp(s\ \mathrm{Ad}(t)w)\big)\, \mathrm{d}s\right)^2. \end{aligned}$$

Similarly, by (S6),

$$S_{d_o}(R_{L,2}) \leq 2\|\varphi\|_\infty \Big(\frac{1}{L}\int_0^L S_{d_o}\big(\varphi \circ \exp(s\,\mathrm{Ad}(t)w)\big)\,\mathrm{d}s\Big).$$

By (S3), we have

$$S_{2d_o}\big(\varphi \circ \exp(s\,\mathrm{Ad}(t)w)\big) \leq B_{2d_o} \max\big(1, (s\|\,\mathrm{Ad}(t)w\|)^{b_{2d_o}}\big)\, S_{2d_o}(\varphi), \tag{5.1}$$

for all $s \geq 0$.

Let us now assume that $L\|\,\mathrm{Ad}(t)w\| \geq 1$. Then,

$$\max(1, (s\|\,\mathrm{Ad}(t)w\|)^{b_{2d_o}}) \leq \big(L\|\,\mathrm{Ad}(t)w\|\big)^{b_{2d_o}} \quad \text{for all } s \in [0, L],$$

so that we conclude from (5.1) that

$$\frac{1}{L}\int_0^L S_{2d_o}\big(\varphi \circ \exp(s\,\mathrm{Ad}(t)w)\big)\,\mathrm{d}s \leq B_{2d_o}(L\|\,\mathrm{Ad}(t)w\|)^{b_{2d_o}}\, S_{2d_o}(\varphi).$$

Hence,

$$S_{d_o}(R_{L,1}) \leq M_{d_o} B_{2d_o}^2\, (L\|\,\mathrm{Ad}(t)w\|)^{2b_{2d_o}}\, S_{2d_o}(\varphi)^2.$$

Furthermore, by our assumption (S1), $\|\varphi\|_\infty \leq S_{d_o}(\varphi)$, and thus

$$S_{d_o}(R_{L,2}) \leq 2B_{d_o}\, (L\|\,\mathrm{Ad}(t)w\|)^{b_{d_o}}\, S_{d_o}(\varphi)^2.$$

We conclude that

$$\begin{aligned} S_{d_o}(\eta_L) &\leq S_{d_o}(R_{L,1}) + S_{d_o}(R_{L,2}) \leq \Big(M_{d_o} B_{2d_o}^2 (L\|\,\mathrm{Ad}(t)w\|)^{2b_{2d_o}} + 2B_{d_o}\, (L\|\,\mathrm{Ad}(t)w\|)^{b_{d_o}}\Big) S_{2d_o}(\varphi)^2 \\ &\leq (M_{d_o} B_{2d_o}^2 + 2B_{d_o})\, (L\|\,\mathrm{Ad}(t)w\|)^{2b_{d_o}}\; S_{2d_o}(\varphi)^2, \end{aligned}$$

where we have used in the last inequality that $L\|\,\mathrm{Ad}(t)w\| \geq 1$. This proves (ii). ■

Proof of Lemma 4.3. Setting $\psi := \varphi - \mu(\varphi)$, we obtain from (4.9) that

$$\mu(|\phi_L|^2) \leq \frac{1}{L^2}\int_0^L \int_0^L \big|\mu\big(\psi \circ \exp(s_1\,\mathrm{Ad}(t)w) \cdot \overline{\psi \circ \exp(s_2\,\mathrm{Ad}(t)w)}\big)\big|\,\mathrm{d}s_1 \mathrm{d}s_2.$$

Since the measure μ is U-invariant,

$$\mu\big(\psi \circ \exp(s_1 \operatorname{Ad}(t)w) \cdot \overline{\psi \circ \exp(s_2 \operatorname{Ad}(t)w)}\big) = \mu\big(\psi \circ \exp((s_1 - s_2) \operatorname{Ad}(t)w) \cdot \overline{\psi}\big)$$
$$= \mu(\varphi \circ \exp((s_1 - s_2) \operatorname{Ad}(t)w) \cdot \overline{\varphi}) - |\mu(\varphi)|^2,$$

for all $s_1, s_2 \in \mathbb{R}$. It follows from (EQ2) that

$$\left|\mu\big(\varphi \circ \exp((s_1 - s_2) \operatorname{Ad}(t)w) \cdot \overline{\varphi}\big) - |\mu(\varphi)|^2\right| \leq C \max\big(1, |s_1 - s_2| \| \operatorname{Ad}(t)w\|\big)^{-c} S_{d_o}(\varphi)^2.$$

Hence, we conclude that

$$\mu(|\phi_L|^2) \leq C\Big(\frac{1}{L^2}\int_0^L \int_0^L \max\big(1, |s_1 - s_2| \| \operatorname{Ad}(t)w\|\big)^{-c} \,\mathrm{d}s_1 \mathrm{d}s_2\Big) S_{d_o}(\varphi)^2$$
$$= C\Big(\frac{1}{R^2}\int_0^R \int_0^R \max\big(1, |s_1 - s_2|\big)^{-c} \,\mathrm{d}s_1 \mathrm{d}s_2\Big) S_{d_o}(\varphi)^2,$$

where $R := L\| \operatorname{Ad}(t)w\|$. We shall use the following integral estimate, which proof we leave to the reader: for every $R \geq 1$ and $c \in (0, 1)$,

$$\frac{1}{R^2}\int_0^R \int_0^R \max(1, |u - v|)^{-c} \,\mathrm{d}u\mathrm{d}v \leq \frac{7R^{-c}}{1 - c}.$$

Then we obtain that

$$\mu(|\phi_L|^2) \leq \frac{7C}{1 - c}\big(L\| \operatorname{Ad}(t)w\|\big)^{-c} S(\varphi)^2 \leq 14C\big(L\| \operatorname{Ad}(t)w\|\big)^{-c} S_{d_o}(\varphi)^2,$$

where we used that $c < 1/2$. This proves the lemma. ■

6 Proof of Proposition 4.4

6.1 Upper bound on $I_{1,\xi}$

We recall that $\nu_\xi\Big(\prod_{i=1}^r \varphi_i \circ t_i\Big)$ can be rewritten using an additional average along a one-parameter subgroup $\exp(sw)$ as in (4.19). Then

$$\nu_\xi\Big(\prod_{i=1}^r \varphi_i \circ t_i\Big) - \nu_\xi\Big(\Big(\frac{1}{L}\int_0^L \xi(sw) \prod_{i=1}^p \varphi_i \circ \exp\big(sw^{(i)}\big) \circ t_i \,\mathrm{d}s\Big) \prod_{i=p+1}^r \varphi_i \circ t_i\Big)$$

equals

$$\nu_\xi\Big(\frac{1}{L}\int_0^L \xi(sw)\prod_{i=1}^{p}\varphi_i\circ\exp\big(sw^{(i)}\big)\circ t_i\Big(\prod_{i=p+1}^{r}\varphi_i\circ\exp\big(sw^{(i)}\big)\circ t_i-\prod_{i=p+1}^{r}\varphi_i\circ t_i\Big)\mathrm{d}s\Big).$$

Thus, since $|\xi|=1$ and $|\psi_\xi|=1$ almost everywhere,

$$|I_{1,\xi}|\le\Big(\prod_{i=1}^{p}\|\varphi_i\|_\infty\Big)\cdot\sup_{s\in[0,L]}\Big\|\prod_{i=p+1}^{r}\varphi_i\circ\exp\big(sw^{(i)}\big)\circ t_i-\prod_{i=p+1}^{r}\varphi_i\circ t_i\Big\|_\infty. \tag{6.1}$$

Furthermore,

$$\begin{aligned}&\Big\|\prod_{i=p+1}^{r}\varphi_i\circ\exp\big(sw^{(i)}\big)\circ t_i-\prod_{i=p+1}^{r}\varphi_i\circ t_i\Big\|_\infty=\Big\|\prod_{i=p+1}^{r}\varphi_i\circ\exp\big(sw^{(i)}\big)-\prod_{i=p+1}^{r}\varphi_i\Big\|_\infty\\
&\le\sum_{k=r}^{p+1}\Big\|\Big(\prod_{i=p+1}^{k}\varphi_i\circ\exp\big(sw^{(i)}\big)\Big)\prod_{i=k+1}^{r}\varphi_i-\Big(\prod_{i=p+1}^{k-1}\varphi_i\circ\exp\big(sw^{(i)}\big)\Big)\prod_{i=k}^{r}\varphi_i\Big\|_\infty\\
&\le\sum_{k=r}^{p+1}\big\|\varphi_k\circ\exp\big(sw^{(k)}\big)-\varphi_k\big\|_\infty\prod_{p+1\le i\ne k\le r}\|\varphi_i\|_\infty.\end{aligned}$$

From (S2), we have

$$\big\|\varphi_k\circ\exp\big(sw^{(k)}\big)-\varphi_k\big\|_\infty\le A\big(s\|w^{(k)}\|\big)^a S_{d_o}(\varphi_k).$$

Therefore, using (S1) and (4.18), we conclude that

$$\Big\|\prod_{i=p+1}^{r}\varphi_i\circ\exp\big(sw^{(i)}\big)\circ t_i-\prod_{i=p+1}^{r}\varphi_i\circ t_i\Big\|_\infty\le rA\big(s\|w^{(p+1)}\|\big)^a\prod_{i=p+1}^{r}S_{d_o}(\varphi_i)$$

for all $s\ge 0$. Hence, using (S1) one more time, we conclude from (6.1) that

$$|I_{1,\xi}|\le rA\big(L\|w^{(p+1)}\|\big)^a\prod_{i=1}^{r}S_{d_o}(\varphi_i).$$

This verifies Proposition 4.4(I).

6.2 Upper bound on $I_{2,\xi}$

This proof here is similar to the argument in Section 4.2, but is more involved. Let us introduce a function $\widetilde{\psi}_L : X^p \to \mathbb{C}$ defined by

$$\widetilde{\psi}_L(\underline{x}) := \frac{1}{L}\int_0^L \xi(sw)\Big(\prod_{i=1}^{p} \varphi_i(\exp(sw^{(i)})t_i.x_i) - \prod_{i=1}^{p}\mu(\varphi_i)\Big)\,\mathrm{d}s \tag{6.2}$$

for $\underline{x} = (x_1, \ldots, x_p) \in X^p$. We denote by ψ_L the restriction of $\widetilde{\psi}_L$ to the diagonal in X^p, that is,

$$\psi_L(x) := \widetilde{\psi}_L(x, \ldots, x), \quad \text{for } x \in X.$$

Furthermore, we set $\lambda := \prod_{i=p+1}^{r} \varphi_i \circ t_i$, so that we can write

$$I_{2,\xi} = \nu_\xi(\psi_L \cdot \lambda) = \nu(\psi_\xi \cdot \psi_L \cdot \lambda).$$

Hence, by Cauchy–Schwarz inequality,

$$|I_{2,\xi}| \leq \nu(|\psi_L|^2)^{1/2}\|\lambda\|_\infty. \tag{6.3}$$

Now, using that the inequality $u^{1/2} \leq |u - v|^{1/2} + v^{1/2}$ for all $u, v \geq 0$, we see that

$$\nu(|\psi_L|^2)^{1/2} \leq \left|\nu(|\psi_L|^2) - \mu^{\otimes p}(|\widetilde{\psi}_L|^2)\right|^{1/2} + \mu^{\otimes p}(|\widetilde{\psi}_L|^2)^{1/2}, \tag{6.4}$$

where $\mu^{\otimes p}$ denotes the product measure on X^p induced from μ.

6.2.1 *Upper bound on* $\left|\nu(|\psi_L|^2) - \mu^{\otimes p}(|\widetilde{\psi}_L|^2)\right|^{1/2}$

Setting

$$\theta_L := \frac{1}{L}\int_0^L \xi(sw)\,\mathrm{d}s \quad \text{and} \quad |c_L|^2 := |\theta_L|^2 \prod_{i=1}^{p} |\mu(\varphi_i)|^2,$$

we have

$$|\widetilde{\psi}_L|^2 = \widetilde{R}_{1,L} + \widetilde{R}_{2,L} + \widetilde{R}_{3,L} + |c_L|^2,$$

where

$$\widetilde{R}_{1,L}(\underline{x}) := \frac{1}{L^2}\int_0^L\int_0^L \xi((s_1-s_2)w)\prod_{i=1}^{p}\varphi_i(\exp(s_1w^{(i)})t_i.x_i)\overline{\varphi_i(\exp(s_2w^{(i)})t_i.x_i)}\,\mathrm{d}s_1\mathrm{d}s_2, \quad (6.5)$$

$$\widetilde{R}_{2,L}(\underline{x}) := -\overline{\theta}_L\left(\frac{1}{L}\int_0^L \xi(sw)\prod_{i=1}^{p}\varphi_i(\exp(sw^{(i)})t_i.x_i)\,\mathrm{d}s\right)\prod_{i=1}^{p}\overline{\mu(\varphi_i)}, \quad (6.6)$$

$$\widetilde{R}_{3,L}(\underline{x}) := \overline{\widetilde{R}_{2,L}(\underline{x})}. \quad (6.7)$$

We denote by $R_{1,L}$, $R_{2,L}$, and $R_{3,L}$ the restrictions of $\widetilde{R}_{1,L}$, $\widetilde{R}_{2,L}$, and $\widetilde{R}_{3,L}$ respectively to the diagonal in X^p. Then

$$|\psi_L|^2 = R_{1,L} + R_{2,L} + R_{3,L} + |c_L|^2,$$

and

$$\left|\nu(|\psi_L|^2) - \mu^{\otimes p}(|\widetilde{\psi}_L|^2)\right| \le |J_1| + |J_2| + |J_3|, \quad (6.8)$$

where

$$J_k := \nu(R_{k,L}) - \mu^{\otimes p}(\widetilde{R}_{k,L}), \quad \text{for } k = 1,2,3.$$

We further note that $|J_2| = |J_3|$. The following lemma, whose proof is postponed until Section 7, provides estimates on the J_k's.

Lemma 6.1. For every $L > 0$ such that $L\|w^{(1)}\| \ge 1$,

(i) $|J_1| \le (M_d B_{d+d_o}^2)^p (L\|w^{(1)}\|)^{2pb_{d+d_o}} D_{d,p}(t_1,\ldots,t_p)\prod_{i=1}^{p} S_{d+d_o}(\varphi_i)^2$.
(ii) $|J_2| = |J_3| \le B_d^p (L\|w^{(1)}\|)^{pb_d} D_{d,p}(t_1,\ldots,t_p)\prod_{i=1}^{p} S_d(\varphi_i)^2$.

In particular, since $B_d, M_S, L\|w^{(1)}\| \ge 1$, we conclude from this lemma and (6.8) that

$$\left|\nu(|\psi_L|^2) - \mu^{\otimes p}(|\widetilde{\psi}_L|^2)\right| \le B_{d,p}(L\|w^{(1)}\|)^{2pb_{d+d_o}} D_{d,p}(t_1,\ldots,t_p)\prod_{i=1}^{p} S_{d+d_o}(\varphi_i)^2, \quad (6.9)$$

where $B_{d,p} := (M_d B_{d+d_o}^2)^p + 2B_d^p$.

6.2.2 *Upper bound on $\mu^{\otimes p}(|\widetilde{\psi}_L|^2)^{1/2}$*

We write the function $\widetilde{\psi}_L$ as

$$\widetilde{\psi}_L(\underline{x}) = \frac{1}{L}\int_0^L \xi(sw)\gamma_s(\underline{x})\,\mathrm{d}s,$$

with

$$\gamma_s(\underline{x}) := \prod_{i=1}^{p} \varphi_i(\exp(sw^{(i)})t_i.x_i) - \prod_{i=1}^{p} \mu(\varphi_i).$$

Let us set $\rho_j := \phi_j - \mu(\phi_j)$. Then we can write

$$\gamma_s(\underline{x}) = \sum_{j=1}^{p} \gamma_s^{(j)}(\underline{x}),$$

where

$$\gamma_s^{(j)}(\underline{x}) := \Big(\prod_{i=1}^{j-1} \mu(\varphi_i)\Big)\rho_j(\exp(sw^{(j)})t_j.x_j)\Big(\prod_{i=j+1}^{p} \varphi_i(\exp(sw^{(i)})t_i.x_i)\Big).$$

Hence,

$$\mu^{\otimes p}(|\widetilde{\psi}_L|^2) = \sum_{j,k=1}^{p} \frac{1}{L^2}\int_0^L\int_0^L \xi((s_1-s_2)w)\mu^{\otimes p}\big(\gamma_{s_1}^{(j)}\cdot\overline{\gamma_{s_2}^{(k)}}\big)\,\mathrm{d}s_1\mathrm{d}s_2.$$

Since the measure μ is G-invariant, we have $\mu(\rho_j) = 0$, so that for $j < k$,

$$\begin{aligned}\mu^{\otimes p}\big(\gamma_{s_1}^{(j)}\cdot\overline{\gamma_{s_2}^{(k)}}\big) =& \Big(\prod_{i=1}^{j-1}|\mu(\varphi_i)|^2\Big)\mu(\rho_j)\overline{\mu(\phi_j)}\Big(\prod_{i=j+1}^{k-1}|\mu(\varphi_i)|^2\Big)\mu\big(\phi_k\circ\exp(s_1w^{(k)})\cdot\overline{\rho_k\circ\exp(s_2w^{(k)})}\big)\\ &\times\Big(\prod_{i=k+1}^{p}\mu\big(\phi_i\circ\exp(s_1w^{(i)})\cdot\overline{\phi_i\circ\exp(s_2w^{(i)})}\big)\Big) = 0.\end{aligned}$$

Also,

$$\begin{aligned}\mu^{\otimes p}\big(\gamma_{s_1}^{(j)}\cdot\overline{\gamma_{s_2}^{(j)}}\big) =& \Big(\prod_{i=1}^{j-1}|\mu(\varphi_i)|^2\Big)\mu\big(\rho_j\circ\exp(s_1w^{(j)})\cdot\overline{\rho_j\circ\exp(s_2w^{(j)})}\big)\\ &\times\Big(\prod_{i=j+1}^{p}\mu\big(\phi_i\circ\exp(s_1w^{(i)})\cdot\overline{\phi_i\circ\exp(s_2w^{(i)})}\big)\Big).\end{aligned}$$

Therefore, we deduce that

$$\begin{aligned}\mu^{\otimes p}(|\widetilde{\psi}_L|^2) &\le \sum_{j=1}^{p}\frac{1}{L^2}\int_0^L\int_0^L \mu^{\otimes p}\big(|\gamma_{s_1}^{(j)}|^2\big)\,\mathrm{d}s_1\mathrm{d}s_2\\ &\le \sum_{j=1}^{p}\Big(\frac{1}{L^2}\int_0^L\int_0^L\big|\mu(\rho_j\circ\exp((s_1-s_2)w^{(j)})\cdot\overline{\rho}_j)\big|\,\mathrm{d}s_1\mathrm{d}s_2\Big)\prod_{i\neq j}\|\varphi_i\|_\infty^2.\end{aligned}$$

By arguing verbatim as in the proof of Lemma 4.3, we see that

$$\frac{1}{L^2}\int_0^L\int_0^L \left|\mu(\rho_j\circ\exp((s_1-s_2)w^{(j)})\cdot\overline{\rho}_j)\right| \mathrm{d}s_1\mathrm{d}s_2 \le 14C\big(L\|w^{(j)}\|\big)^{-c} S_{d_o}(\varphi_j)^2, \tag{6.10}$$

for all $j=1,\ldots,p$, provided that $L\|w^{(j)}\|\ge 1$.

Let us now assume that $L\|w^{(p)}\|\ge 1$. By (4.18), we then also have $L\|w^{(j)}\|\ge 1$ for all $j=1,\ldots,p$, and thus we can use the bound (6.10) for every j. We conclude, once again using (4.18), that

$$\mu^{\otimes p}(|\widetilde{\psi}_L|^2) \le 14p\,C\big(L\|w^{(p)}\|\big)^{-c}\prod_{i=1}^{p} S_{d_o}(\varphi_i)^2. \tag{6.11}$$

6.2.3 *Combining the two bounds*

Our goal is to bound $|I_{2,\xi}|$ from above, uniformly over $\xi\in\Xi$. Recall from (6.3) and (6.4) that

$$|I_{2,\xi}| \le \left(\left|\nu(|\psi_L|^2)-\mu^{\otimes p}(|\widetilde{\psi}_L|^2)\right|^{1/2} + \mu^{\otimes p}(|\widetilde{\psi}_L|^2)^{1/2}\right)\|\lambda\|_\infty.$$

By (S1), we have $\|\lambda\|_\infty \le \prod_{i=p+1}^{r} S_{d_o}(\varphi_i)$.

We recall that we have proved in (6.9) that

$$\left|\nu(|\psi_L|^2)-\mu^{\otimes p}(|\widetilde{\psi}_L|^2)\right| \le B_{d,p}\big(L\|w^{(1)}\|\big)^{2pb_{d+d_o}} D_{d,p}(t_1,\ldots,t_p)\prod_{i=1}^{p} S_{d+d_o}(\varphi_i)^2,$$

provided that $L\|w^{(1)}\|\ge 1$, and in (6.11) we have proved that

$$\mu^{\otimes p}(|\widetilde{\psi}_L|^2) \le 14p\,C\big(L\|w^{(p)}\|\big)^{-c}\prod_{i=1}^{p} S_{d_o}(\varphi_i)^2,$$

provided that $L\|w^{(p)}\|\ge 1$. Hence,

$$\begin{aligned}|I_{2,\xi}| &\le \left(B_{d,p}^{1/2}\big(L\|w^{(1)}\|\big)^{pb_{d+d_o}} D_{d,p}(t_1,\ldots,t_p)^{1/2} + (14pC)^{1/2}\big(L\|w^{(p)}\|\big)^{-c/2}\right)\prod_{i=1}^{r} S_{d+d_o}(\varphi_i)\\ &\le P_1\left(\big(P_d L\|w^{(1)}\|\big)^{rb_{d+d_o}} D_{d,p}(t_1,\ldots,t_p)^{1/2} + \sqrt{r}\big(L\|w^{(p)}\|\big)^{-c/2}\right)\prod_{i=1}^{r} S_{d+d_o}(\varphi_i),\end{aligned}$$

where

$$P_1 := \sqrt{14C} \quad \text{and} \quad P_d := \left(M_d B_{d+d_o}^2 + 2B_d^2\right)^{1/2b_{d+d_o}}.$$

This completes the proof of Proposition 4.4(II).

6.3 Upper bound on $I_{3,\xi}$

Let us first consider $\xi = 1$. It follows immediately from the definition of $D_{d,r-p}(t_{p+1}, \ldots, t_r)$ in (4.3) and (S1) that

$$|I_{3,1}| \le \Big(\prod_{i=1}^{p} \|\varphi_i\|_\infty\Big) \Big| \nu\Big(\prod_{i=p+1}^{r} \varphi_i \circ t_i\Big) - \prod_{i=p+1}^{r} \mu(\varphi_i) \Big| \le D_{d,r-p}(t_{p+1}, \ldots, t_r) \prod_{i=1}^{r} S_d(\varphi_i)$$

when $d \ge d_o$. Similarly, if $\xi \ne 1$, by the definition of $D_{d,r-p}(t_{p+1}, \ldots, t_r; \xi)$ in (4.4) and (S1),

$$|I_{3,\xi}| \le \Big(\prod_{i=1}^{p} \|\varphi_i\|_\infty\Big) \Big| \nu_\xi\Big(\prod_{i=p+1}^{r} \varphi_i \circ t_i\Big) \Big| \le D_{d,r-p}(t_{p+1}, \ldots, t_r; \xi) \prod_{i=1}^{r} S_d(\varphi_i)$$

when $d \ge d_o$. Therefore, from (4.5) we conclude that

$$\sup_{\xi \in \Xi} |I_{3,\xi}| \le E_{d,r-p}(t_{p+1} \ \ldots, t_r) \prod_{i=1}^{r} S_d(\varphi_i),$$

which verifies Proposition 4.4(III).

7 Proof of Lemma 6.1

Proof of Lemma 6.1(i). Given $\underline{s} = (s_1, s_2) \in \mathbb{R}^2$ and $x \in X$, we define

$$\phi_i(\underline{s}, x) := \varphi_i(\exp(s_1 w^{(i)})x)\overline{\varphi_i(\exp(s_2 w^{(i)})x)}, \quad \text{for } i = 1, \ldots, p.$$

Then

$$\widetilde{R}_{1,L}(\underline{x}) = \frac{1}{L^2} \int_0^L \int_0^L \xi((s_1 - s_2)w) \prod_{i=1}^{p} \phi_i(\underline{s}, t_i x_i) \, \mathrm{d}s_1 \mathrm{d}s_2,$$

and

$$|J_1| = \left|\nu(R_{1,L}) - \mu^{\otimes p}(\widetilde{R}_{1,L})\right| \leq \frac{1}{L^2}\int_0^L\int_0^L \left|\nu\Big(\prod_{i=1}^p \phi_i(\underline{s},\cdot)\circ t_i\Big) - \prod_{i=1}^p \mu\big(\phi_i(\underline{s},\cdot)\big)\right| \mathrm{d}s_1\,\mathrm{d}s_2. \quad (7.1)$$

Since $\mathcal{A}$ is a G-invariant $*$-algebra and $\varphi_i \in \mathcal{A}$, we note that $\phi_i(\underline{s},\cdot) \in \mathcal{A}$. Hence, by the definition of $D_{d,p}(t_1,\ldots,t_p)$ in (4.3),

$$\left|\nu\Big(\prod_{i=1}^p \phi_i(\underline{s},\cdot)\circ t_i\Big) - \prod_{i=1}^p \mu\big(\phi_i(\underline{s},\cdot)\big)\right| \leq D_{d,p}(t_1,\ldots,t_p)\prod_{i=1}^p S_d\big(\phi(\underline{s},\cdot)\big),$$

for every $\underline{s} \in \mathbb{R}^2$. By (S4),

$$S_d\big(\phi_i(\underline{s},\cdot)\big) \leq M_d\, S_{d+d_o}\big(\varphi_i\circ\exp(s_1 w^{(i)})\big)\, S_{d+d_o}\big(\varphi_i\circ\exp(s_2 w^{(i)})\big),$$

and by (S3),

$$S_{d+d_o}\big(\varphi_i\circ\exp(s w^{(i)})\big) \leq B_{d+d_o}\, \max\big(1, s\|w^{(i)}\|\big)^{b_{d+d_o}}\, S_{d+d_o}(\varphi_i), \quad \text{for all } s \geq 0.$$

Let us now assume that $L\|w^{(1)}\| \geq 1$. Then, by (4.18),

$$\max(1, s\|w^{(i)}\|) \leq L\|w^{(1)}\|, \quad \text{for all } i = 1,\ldots,p \text{ and } s \in [0,L].$$

Therefore, we conclude that for all $\underline{s} \in [0,L]^2$,

$$\left|\nu\Big(\prod_{i=1}^p \phi_i(\underline{s},\cdot)\circ t_i\Big) - \prod_{i=1}^p \mu\big(\phi_i(\underline{s},\cdot)\big)\right| \leq (M_d B_{d+d_o}^2)^p \big(L\|w^{(1)}\|\big)^{2pb_{d+d_o}} D_{d,p}(t_1,\ldots,t_p)\prod_{i=1}^p S_{d+d_o}(\varphi_i)^2,$$

and it now follows from (7.1) that

$$|J_1| \leq (M_d B_{d+d_o}^2)^p \big(L\|w^{(1)}\|\big)^{2pb_{d+d_o}} D_{d,p}(t_1,\ldots,t_p)\prod_{i=1}^p S_{d+d_o}(\varphi_i)^2,$$

provided that $L\|w^{(1)}\| \geq 1$. ∎

Proof of Lemma 6.1(ii). We recall that

$$\widetilde{R}_{2,L}(\underline{x}) = -\overline{\theta}_L\left(\frac{1}{L}\int_0^L \xi(sw)\prod_{i=1}^p \varphi_i(\exp(sw^{(i)})t_i x_i)\,\mathrm{d}s\right)\prod_{i=1}^p \overline{\mu(\varphi_i)}.$$

Since $|\theta_L| \le 1$ and the measure μ is T-invariant, we have

$$|J_2| \le \left|\nu(R_{k,L}) - \mu^{\otimes p}(\widetilde{R}_{k,L})\right|$$
$$\le \left(\frac{1}{L}\int_0^L \left|\nu\Big(\prod_{i=1}^{p} \varphi_i \circ \exp(sw^{(i)}) \circ t_i\Big) - \prod_{i=1}^{p} \mu\big(\varphi_i \circ \exp(sw^{(i)})\big)\right| \mathrm{d}s\right) \prod_{i=1}^{p} \|\varphi_i\|_\infty.$$

By the definition of $D_{d,p}(t_1,\ldots,t_p)$ in (4.3), we see that

$$\left|\nu\Big(\prod_{i=1}^{p} \varphi_i \circ \exp(sw^{(i)}) \circ t_i\Big) - \prod_{i=1}^{p} \mu\big(\varphi_i \circ \exp(sw^{(i)})\big)\right| \le D_{d,p}(t_1,\ldots,t_p) \prod_{i=1}^{p} S_d\big(\varphi_i \circ \exp(sw^{(i)})\big),$$

for all $s \in \mathbb{R}$. By (S2),

$$S_d\big(\varphi_i \circ \exp(sw^{(i)})\big) \le B_d \max\big(1, s\|w^{(i)}\|\big)^{b_d} S_d(\varphi_i), \quad \text{for all } s \ge 0.$$

Let us now assume that $L\|w^{(1)}\| \ge 1$. Then, by (4.18),

$$\max(1, s\|w^{(i)}\|) \le L\|w^{(1)}\|, \quad \text{for all } i = 1,\ldots,p \text{ and } s \in [0,L].$$

Therefore,

$$\left|\nu\Big(\prod_{i=1}^{p} \varphi_i \circ \exp(sw^{(i)}) \circ t_i\Big) - \prod_{i=1}^{p} \mu\big(\varphi_i \circ \exp(sw^{(i)})\big)\right| \le B_d^p \big(L\|w^{(1)}\|\big)^{pb_d} D_{d,p}(t_1,\ldots,t_p) \prod_{i=1}^{p} S_d(\varphi_i),$$

for all $s \in [0,L]$. Now it readily follows that

$$|J_2| \le B_d^p \big(L\|w^{(1)}\|\big)^{pb_d} D_{d,p}(t_1,\ldots,t_p) \prod_{i=1}^{p} S_d(\varphi_i)^2,$$

provided that $L\|w^{(1)}\| \ge 1$. ■

8 Explicit Version of the Main Theorem

Let us now present a version of Theorem 2.1 with explicit parameters. Here we additionally assume that the estimates (S3) and (S4) hold for

$$B_d = L_1^d \text{ and } b_d = \ell d \text{ with } L_1, \ell \ge 1 \quad \text{and} \quad M_d = L_2^d \text{ with } L_2 \ge 1$$

respectively. Such bounds can be verified for the Sobolev norms (see Section 3.2, examples (i)–(iii)). Then we prove the following:

Theorem 8.1. There exist $H_1, H_2 > 0$ and $\lambda > 1$ such that for all $r \ge 1$, $\varphi_o \in W(\nu)$, $\varphi_1,\ldots,\varphi_r \in \mathcal{A}$ and $\underline{t} \in T_+^r$ satisfying $\Delta_r(\underline{t}) > H_2^{(r-1)!r!(r+1)!\lambda^r}$,

$$\left|\nu\Big(\varphi_o \prod_{i=1}^{r} \varphi_i \circ t_i\Big) - \nu(\varphi_o) \prod_{i=1}^{r} \mu(\varphi_i)\right| \le rH_1\, \Delta_r(\underline{t})^{-\frac{1}{(r!)^2(r+1)!\lambda^r}} \prod_{i=1}^{r} S_{(r+1)d_o}(\varphi_i),$$

Proof. We recall from the proof of Theorem 8.1 that the parameters δ_r, D_r, and δ_r are determined by the following relations:

$$d_r = d_{r-1} + d_o \quad \text{with } d_1 = 2d_o,$$

$$D_r = 2P_1 P_{d_{r-1}}^{rb_{d_{r-1}+d_o}} D_{r-1}^{1/2} + rQ, \quad \text{where } P_d := \left(M_d B_{d+d_o}^2 + 2B_d^2\right)^{1/2b_{d+d_o}},$$

$$\delta_r = \frac{c_1 \delta_{r-1}}{r(\frac{2c_1}{r} + 2rb_{d_{r-1}+d_o})}.$$

It is clear from the recursive formulas that

$$d_r = (r+1)d_o$$

and

$$\delta_r = \frac{c_1 \delta_{r-1}}{r(\frac{2c_1}{r} + 2\ell r(r+1))} \geq \frac{1}{(r!)^2 (r+1)!\, \lambda^r}$$

for an explicit $\lambda > 1$, which depends only on a, b, c, and ℓ. However, with this choice, D_r grows super-exponentially fast in r. We modify the proof of Theorem 2.1 choosing θ differently to get rid of the $P_{d_{r-1}}^{rb_{d_{r-1}}}$-factor. This will then imply that the constant grows linearly. Unfortunately, this type of argument only applies to those $\underline{t} \in T_+^r$ such that $\Delta_r(\underline{t})$ is sufficiently large depending on r.

We assume that the inductive assumption (4.31) holds and choose

$$\theta := P_{d_{r-1}} \Delta_r(\underline{t})^{-\varepsilon_r},$$

where ε_r is given by (4.32), and $P_{d_{r-1}}$ is defined in Proposition 4.4. We note that

$$P_{d_{r-1}} = \left(M_{rd_o} B_{(r+1)d_o}^2 + 2B_{rd_o}^2\right)^{1/(2(r+1)d_o)} = \left(L_2^{rd_o} L_1^{2(r+1)d_o} + 2L_1^{2rd_o}\right)^{1/(2(r+1)d_o)} \leq L_1(L_2 + 2).$$

Since $P_{d_{r-1}} \geq 1$, we have $M_r(\underline{t})^{-1} \leq \theta$, and we note that $\theta < 1$ provided that

$$\Delta_r(\underline{t}) > P_{d_{r-1}}^{1/\varepsilon_r}. \tag{8.1}$$

Hence, if $\underline{t}$ satisfies (8.1), then, just as in the previous proof, it follows from (4.27) combined with (4.31) that

$$E_{d+d_o,r}(\underline{t}) \leq D_r \Delta_r(\underline{t})^{-\delta_r},$$

where

$$D_r := 2P_1 D_{r-1}^{1/2} + rQ_r \quad \text{with } Q_r := QP_{d_{r-1}}^{c_1/r}, \tag{8.2}$$

and δ_r as above. Then

$$1/\varepsilon_r = \frac{\frac{2c_1}{r} + 2rb_{d_{r-1}+d_o}}{\delta_{r-1}} \leq \left(\frac{2c_1}{r} + 2\ell r(r+1)\right)((r-1)!)^2 r!\, \lambda^{r-1} \leq \gamma\,(r-1)!\, r!\,(r+1)!\, \lambda^r,$$

where $\gamma := \frac{2\max(c_1, 2\ell)}{\lambda}$. By induction on r, one can also show that there exists $H_1 \geq 1$ such that $D_r \leq H_1 r$ for all r. Therefore, we conclude that

$$E_r(\underline{t}) \leq rH_1 \Delta_r(\underline{t})^{-\frac{1}{(r!)^2 (r+1)!\lambda^r}},$$

for all $\underline{t} \in T^r$ satisfying

$$\Delta_r(\underline{t}) > H_2^{(r-1)!r!(r+1)!\,\lambda^r}, \quad \text{where } H_2 := L_1^{\gamma}(L_2+2)^{\gamma}.$$

This finishes the proof. ■

In the case of the examples (i)–(ii) of the norms introduced in Section 3.2, the above estimates can be simplified. Let us now additionally assume that $X = G/\Gamma$, where G is connected semisimple Lie group without compact factors and Γ an irreducible lattice in G, and $\mathcal{A} = \mathbb{C} + C_c^{\infty}(X)$.

Corollary 8.2. Let S_d denote the norms defined in either (i) or (ii) in Section 3.2. Then there exist $H_1, H_2 > 0$ and $\lambda > 1$ such that for all $r \geq 1$, $\varphi_o \in W(\nu)$, $\varphi_1, \ldots, \varphi_r \in \mathcal{A}$ and $\underline{t} \in T_+^r$ satisfying $\Delta_r(\underline{t}) > H_2^{(r-1)!r!\,\lambda^r}$,

$$\left|\nu\left(\varphi_o \prod_{i=1}^{r} \varphi_i \circ t_i\right) - \nu(\varphi_o)\prod_{i=1}^{r}\mu(\varphi_i)\right| \leq r\, H_1\, \Delta_r(\underline{t})^{-\frac{1}{(r!)^2\lambda^r}} \prod_{i=1}^{r} S_{r+1}(\varphi_i).$$

Proof. We note that in this case, properties (S3)–(S4) hold with fixed constants independent of d. Using this we obtain that

$$\delta_r = \frac{c_1 \delta_{r-1}}{r(\frac{2c_1}{r} + 2rb)} \geq \frac{1}{(r!)^2\, \lambda^r}$$

and

$$1/\varepsilon_r = \frac{\frac{2c_1}{r} + 2rb}{\delta_{r-1}} \leq \gamma\,(r-1)!\, r!\, \lambda^r.$$

This implies the estimate. ■

Funding

This work was supported by the Swedish Research Council [2019-04774 to M.B.]; and the SNF [200021–182089 to A.G.].

References

[1] Björklund, M., M. Einsiedler, and A. Gorodnik. “Quantitative multiple mixing.” *J. Eur. Math. Soc. (JEMS)* 22, no. 5 (2020): 1475–529.

[2] Björklund, M., R Fregoli, and A. Gorodnik. “A metrical theorem in multiplicative Diophantine approximation.” (2021): Preprint.

[3] Björklund, M. and A. Gorodnik. “Central limit theorems for Diophantine approximants.” *Math. Ann.* 374, no. 3–4 (2019): 1371–437.

[4] Dabbs, K., M. Kelly, and H. Li. “Effective equidistribution of translates of maximal horospherical measures in the space of lattices.” *J. Mod. Dyn.* 10 (2016): 229–54.

[5] Dolgopyat, D. “Limit theorems for partially hyperbolic systems.” *Trans. Amer. Math. Soc.* 356, no. 4 (2004): 1637–89.

[6] Kleinbock, D. Y. and G. A. Margulis. “Bounded Orbits of Nonquasiunipotent Flows on Homogeneous Spaces.” In *Sinai’s Moscow Seminar on Dynamical Systems*. 141–72. Amer. Math. Soc. Transl. Ser. 2, 171, Adv. Math. Sci., 28. Providence, RI: Amer. Math. Soc., 1996.

[7] Kleinbock, D. Y. and G. A. Margulis. “On Effective Equidistribution of Expanding Translates of Certain Orbits in the Space of Lattices.” In *Number Theory, Analysis and Geometry*. 385–96. New York: Springer, 2012.

[8] Margulis, G. *On Some Aspects of the Theory of Anosov Systems*. Springer Monogr. Math. Berlin: Springer, 2004.

[9] Sarnak, P. “Asymptotic behavior of periodic orbits of the horocycle flow and Eisenstein series.” *Comm. Pure Appl. Math.* 34, no. 6 (1981): 719–39.

[10] Shi, R. “Expanding cone and applications to homogeneous dynamics.” *Int. Math. Res. Not. IMRN* 9 (2021): 7060–95.

[11] Zagier, D. “Eisenstein Series and the Riemann Zeta Function.” In *Automorphic Forms, Representation Theory and Arithmetic (Bombay, 1979)*. 275–301. Tata Inst. Fund. Res. Studies in Math., 10. Bombay: Tata Inst. Fundamental Res., 1981.