SPECTRAL INDEPENDENCE

ALIREZA S GOLSEFIDY, KEIVAN MALLAHI-KARAI, AND AMIR MOHAMMADI

Abstract. We prove the spectral gap property for random walks on the product of two non-locally isomorphic analytic real or p-adic compact groups with simple Lie algebras, under the necessary condition that the marginals posses a spectral gap. Furthermore, we give additional control on the spectral gap depending on certain specific properties of the given groups and marginals; in particular, we prove some new cases of the super-approximation conjecture.

One ingredient of the proof is a local Ulam stability result which is introduced and proved in this paper. This result characterizes partially defined almost homomorphisms between two analytic compact groups with simple Lie algebras.

Contents

1. Introduction

Let G be a compact group, and let μ be a Borel symmetric probability measure on G. An ℓ-step random walk with respect to μ is

$$X^{(\ell)} := X_1 \cdots X_\ell$$

where $X_1, X_2, \ldots$ is a sequence of independent random variables with probability law μ. If the group generated by the support of μ is dense in G, then for every continuous function f,

$$\lim_{\ell\to\infty} \mathbb{E}[f(X^{(\ell)})] = \int_G f(x)\, dm_G(x),$$

A.S.G. acknowledges support by the NSF grants 1602137, 1902090, 2302519.
K.MK. acknowledges support by the DFG grant DFG MA 6503/1-1.
A.M. acknowledges support by the NSF, grants 1764246 and 2055122.

where m_G is the probability Haar measure of G. The rate of convergence is governed by the operator norm $\lambda(\mu)$ of the convolution operator

$$T_\mu : L^2_0(G, m_G) \to L^2_0(G, m_G), \qquad (T_\mu(f))(x) := \int_G f(g^{-1}x) d\mu(g),$$

where $L^2_0(G, m_G)$ is the orthogonal complement of constant functions. We say μ has the *spectral gap property* if $\lambda(\mu) < 1$. Likewise, a symmetric random variable X with values in G is said to have the spectral gap property if its probability law has spectral gap.

In this work, we will investigate the spectral gap property of a pair (X, Y) of random variables on $G_1 \times G_2$ where G_i's are compact groups. It is clear that if (X, Y) has the spectral gap property, then both X and Y must have this property. Motivated by this, we will say G_1 and G_2 are *spectrally independent* if this necessary condition is also sufficient. The main result of this paper shows that G_1 and G_2 are spectrally independent for a wide class of compact groups.

Theorem A. *For $i = 1, 2$, let F_i be the field of real $\mathbb{R}$ or p-adic $\mathbb{Q}_p$ numbers. Let $\mathbb{G}_i$ be an F_i-almost simple group, and let G_i be an open compact subgroup of $\mathbb{G}_i(F_i)$. Assume that G_1 and G_2 are not locally isomorphic. Then G_1 and G_2 are spectrally independent.*

Note that if G_1 and G_2 are spectrally independent, then they should be *algebraically independent*, that means that they have no common nontrivial topological quotients. Indeed, if $\phi_i : G_i \to H$ is a common nontrivial quotient, then $K = \{(g_1, g_2) \in G_1 \times G_2 : \phi_1(g_1) = \phi_2(g_2)\}$ carries a measure with marginals m_{G_1} and m_{G_2}, however, K is a proper closed subgroup of $G_1 \times G_2$. It is also worth noting that in Appendix C, we provide examples of algebraically independent groups which are not spectrally independent.

One of the key ingredients in the proof of Theorem A is Theorem 3, which is a local stability theorem developed in this paper and is of independent interest. Roughly speaking, this stability result states that a *partial approximate homomorphism*, see Definition 1, between G_1 and G_2 as in Theorem A is close to an isogeny between $\mathbb{G}_1$ and $\mathbb{G}_2$. Stability theorems in this vein have a long history, indeed Grove, Karcher, Ruh [14, Theorem 4.3] studied approximate homomorphisms between compact Lie groups; later Kazhdan [16] studied approximate homomorphisms from amenable groups to unitary operators on a Hilbert space. One of the main novelties of our result is that we only require the approximate homomorphism to be defined on a neighborhood of the identity rather than the entire group; this generalization requires new techniques as the averaging techniques, used in aforementioned results, can no longer be applied. We refer the reader to Theorem 3 for the precise statement, here we only state two special cases of that theorem.

In the following, 1_ρ denotes the ball of radius ρ in G with respect to the operator norm. Also, for any prime p and any positive integer k, we let $\mathrm{SL}_n(\mathbb{Z}_p)[k]$ denote the kernel of the reduction mod p^k from $\mathrm{SL}_n(\mathbb{Z}_p)$ to $\mathrm{SL}_n(\mathbb{Z}/p^k\mathbb{Z})$.

Theorem B. *(1)* (Real Case) *Let $G_i = \mathrm{SU}(n_i)$ for $n_i \geq 2$. Then there exist $m_0 > 1$ and $0 < c < 1$, depending only on n_1 and n_2, such that for every $m > m_0 m'$, the following holds. For $\rho < 1/2$, suppose $f : 1^{(1)}_\rho \to G_2$ is a map which satisfies*

$$f(g^{-1})f(g) \in 1^{(2)}_{\rho^m} \qquad \textit{and} \qquad f(g_1g_2)(f(g_1)f(g_2))^{-1} \in 1^{(2)}_{\rho^m}$$

whenever these expressions are defined. Assume further that the ρ^m neighborhood of $\mathrm{Im}(f)$ contains the $\rho^{m'}$ neighborhood of the identity in G_2. Then $n_1 = n_2 = n$ and there is an isomorphism $\Psi : \mathrm{SU}(n) \to \mathrm{SU}(n)$ so that

$$f(g)\Psi(g)^{-1} \in 1_{\rho^{cm}} \qquad \textit{for all } g \in 1^{(1)}_{\rho^m}.$$

(2) (p-adic Case) *Let $n_1, n_2 \geq 2$ be two integers. There exist $m_0 > 1$ and $0 < c < 1$, depending only on n_1 and n_2, so that for every $m > m_0 m'$ the following holds. Let $k \in \mathbb{Z}^+$. Suppose*

$$f : \mathrm{SL}_{n_1}(\mathbb{Z}_p)[k]/\,\mathrm{SL}_{n_1}(\mathbb{Z}_p)[km] \to \mathrm{SL}_{n_2}(\mathbb{Z}_p)/\,\mathrm{SL}_{n_2}(\mathbb{Z}_p)[km]$$

is a group homomorphism such that

$$\mathrm{SL}_{n_2}(\mathbb{Z}_p)[km']/\,\mathrm{SL}_{n_2}(\mathbb{Z}_p)[km] \subseteq \mathrm{Im}(f).$$

Then $n_1 = n_2 = n$ and there is an isomorphism $\Psi : \mathrm{SL}_n(\mathbb{Q}_p) \to \mathrm{SL}_n(\mathbb{Q}_p)$ so that

$$f\big(g\,\mathrm{SL}_n(\mathbb{Z}_p)[km]\big) \equiv \Psi(g) \pmod{p^{ckm}} \qquad \textit{for all } g \in \mathrm{SL}_n(\mathbb{Z}_p)[k].$$

In the next section, we will present more precise statements of results of this work, including Theorems A and B. We will also state a *global* version, Theorem 2, of Theorem A where we obtain spectral gap which is uniform across all places.

2. Statement of results

Recall that for a symmetric Borel probability measure μ on a compact group G, the contraction factor $\lambda(\mu)$ of μ is the operator norm of the convolution operator

$$T_\mu : L^2_0(G, m_G) \to L^2_0(G, m_G), \qquad (T_\mu(f))(x) := \int_G f(g^{-1}x)d\mu(g),$$

where $L^2_0(G, m_G)$ is the orthogonal complement of constant functions. Given a symmetric random variable X with values in G and probability law μ, we define

$$\mathcal{L}(X; G) := -\log(\lambda(\mu)).$$

When G is clear from the context, we often denote $\mathcal{L}(X; G)$ by $\mathcal{L}(X)$.

Theorem 1. *For $i = 1, 2$, let F_i be the field of real $\mathbb{R}$ or p-adic $\mathbb{Q}_p$ numbers. Let $\mathbb{G}_i$ be an F_i-almost simple group and let G_i be an open compact subgroup of $\mathbb{G}_i(F_i)$. Suppose G_1 and G_2 are not locally isomorphic and $X = (X_1, X_2)$ is a $G_1 \times G_2$-valued random variable with a Borel probability law. Then the marginal bounds $\min(\mathcal{L}(X_1), \mathcal{L}(X_2)) \geq c_0 > 0$ imply*

$$\mathcal{L}(X) \gg_{c_0, G_1, G_2} 1.$$

In particular, G_1 and G_2 are spectrally independent.

Recall that the *super-approximation conjecture* states that for a finite symmetric subset Ω of $\mathrm{GL}_n(\mathbb{Z}[1/q_0])$, the connected component of the Zariski closure of the group $\Gamma = \langle \Omega \rangle$ is perfect if and only if the family of Cayley graphs $\{\mathrm{Cay}(\pi_m(\Gamma), \pi_m(\Omega)\}$ forms a family of *expanders* as m ranges over the set of positive integers with $\gcd(q_0, m) = 1$. Our next theorem proves a special case of the super-approximation conjecture.

Expander graphs are roughly highly connected sparse graphs that have arbitrarily large number of vertices. The first explicit construction of expander graphs was done by Margulis [21] using Kazhdan's property (T). Later, the most optimal expanders, known as Ramanujan graphs, were constructed by Lubotzky, Phillips, and Sarnak [18], and independently by Margulis [22]. The reader can see more on expanders and their connections with the spectral gap property in [19].

In the past two decades, there has been a lot of progress on the super-approximation conjecture starting with the seminal work of Bourgain and Gamburd [2]. We refer the reader to the following articles to see the more recent results on this conjecture [3, 27, 4, 11, 24, 25, 15].

It is worth pointing out that in [15], the super-approximation conjecture is proved if Ω consists of *integral* matrices and the Zariski closure of Γ is $\mathbb{Q}$-almost simple. By refining the tools developed for the proof of Theorem 1, we provide an affirmative answer to the super-approximation conjecture with no integrality assumption in the case where the Zariski closure is absolutely almost simple and moduli have at most two distinct prime factors.

Theorem 2. *Let $\mathbb{G}$ be an absolutely almost simple $\mathbb{Q}$-algebraic group. Let $\Omega \subseteq \mathbb{G}(\mathbb{Q})$ be a finite symmetric subset such that the group $\Gamma = \langle \Omega \rangle$ is Zariski dense in $\mathbb{G}$. Denote by V_Γ the set of all places ν of $\mathbb{Q}$ such that Γ is a bounded subset of $\mathbb{G}(\mathbb{Q}_\nu)$. For distinct $\nu_1, \nu_2 \in V_\Gamma$, let Γ_{ν_1,ν_2} denote the closure of Γ in $\mathbb{G}(\mathbb{Q}_{\nu_1}) \times \mathbb{G}(\mathbb{Q}_{\nu_2})$. Then*

$$\inf_{\nu_1 \neq \nu_2 \in V_\Gamma} \mathcal{L}(X; \Gamma_{\nu_1,\nu_2}) > 0,$$

where X is a random variable with the uniform distribution on Ω.

As it was explained in the introduction, one of the main ingredients in the proof of Theorem 1 is a classification of *partial almost homomorphisms* between two compact analytic groups. To formulate our result, we need a few definitions and notation.

For a metric compact group G and $g \in G$, we let g_ρ denote the ρ-neighborhood of g.

Definition 1. Suppose G_1 and G_2 are two compact groups equipped with bi-invariant metrics compatible with their topology. Let $\delta > 0$ and S be a symmetric subset of G_1 containing the identity element $1^{(1)}$. A function $f : S \to G_2$ is called an *S-partial, δ-approximate homomorphism* if the following three properties are satisfied

(1) $f(1^{(1)}) = 1^{(2)}$,
(2) $f(g^{-1}) \in (f(g)^{-1})_\delta$ for every $g \in S$, and
(3) for all $g_1, g_2 \in S$ with $g_1 g_2 \in S$, we have $f(g_1 g_2) \in \big(f(g_1) f(g_2)\big)_\delta$.

We will refer to a $1^{(1)}_\rho$-partial δ-approximate homomorphism simply as *ρ-partial, δ-approximate homomorphism.*

Let $\mathbb{G} \subseteq (\mathrm{SL}_N)_\mathbb{Q}$ be a Zariski connected, absolutely almost simple $\mathbb{Q}$-algebraic subgroup. Let $\Sigma_\mathbb{G}$ be the set of places v of $\mathbb{Q}$ where $\mathbb{G}(\mathbb{Q}_\nu)$ contains a compact open subgroup; note that $\Sigma_\mathbb{G}$ equals the set of all places if $\mathbb{G}(\mathbb{R})$ is compact, and equals the set of all finite places otherwise. We say a family

$$\{G_\nu : \nu \in \Sigma_\mathbb{G}\}$$

of compact groups is a *coherent family* attached to $\mathbb{G}$ if the following holds: $G_\nu \subseteq \mathbb{G}(\mathbb{Q}_\nu)$ is a compact open and $G_\nu = \mathbb{G}(\mathbb{Q}_\nu) \cap \mathrm{SL}_N(\mathbb{Z}_\nu)$ for all but finitely many places ν.

In the sequel, we let $p_\nu = \nu$ if ν is non-Archimedean and $p_\nu = 2$ if $\nu = \infty$.

Theorem 3. *Let ν_1 and ν_2 be two (possibly equal) places of $\mathbb{Q}$, and let $F_i := \mathbb{Q}_{\nu_i}$. Let $\mathbb{G}_i \subseteq (\mathrm{SL}_{n_i})_{F_i}$ be a Zariski connected, F_i-almost simple subgroup, and let $G_i \subseteq \mathbb{G}_i(F_i)$ be a compact open subgroup.*

If $F_i = \mathbb{R}$, we assume that $\mathbb{G}_i$ is given by an $\mathbb{R}$-embedding in $(\mathrm{SL}_{n_i})_\mathbb{R}$ such that $\mathbb{G}_i(\mathbb{R}) \subseteq \mathrm{SO}_{n_i}(\mathbb{R})$. If $F_i = \mathbb{Q}_p$, then we assume that $G_i \subseteq \mathrm{SL}_{n_i}(\mathbb{Z}_p)$. In both cases, we consider the metric induced by the operator norm.

Then there is a positive number $c = c(\dim \mathbb{G}_1, \dim \mathbb{G}_2)$ such that for every $m \gg_{G_1,G_2} m'$ and positive number $\rho \ll_{G_1,G_2} 1$ the following holds:

If $f : 1^{(1)}_\rho \to G_2$ is a ρ-partial, ρ^m-approximate homomorphism which satisfies

$$1^{(2)}_{\rho^{m'}} \subseteq (\mathrm{Im}(f))_{\rho^m} \qquad \text{(Large image)},$$

then $F_1 = F_2 = F$, $\mathrm{Lie}(\mathbb{G}_1)(F) \simeq \mathrm{Lie}(\mathbb{G}_2)(F)$, and f is near an isogeny *in the following sense: There is an F-central isogeny $\Psi : \widetilde{\mathbb{G}}_1 \to \mathbb{G}_2$ where $\widetilde{\mathbb{G}}_1$ is the simply-connected cover of $\mathbb{G}_1$ so that*

$$f(g) \in \Psi(\widetilde{g})_{\rho^{cm}} \quad \textit{for every } g \in 1^{(1)}_\rho,$$

where $\widetilde{g}$ is the unique lift of g under the covering map from $\widetilde{\mathbb{G}}_1$ to $\mathbb{G}_1$ which belongs to the image of 4ρ-neighborhood of 0 under the exponential map.

Moreover, if $\mathcal{F}=\{G_\nu\}$ is a coherent family attached to a $\mathbb{Q}$-group $\mathbb{G}$, then there exist $\kappa:=\kappa_{\mathbb{G}}$ and $m_{\mathcal{F}},\rho_{\mathcal{F}}>0$ so that the following holds. For all distinct ν_1,ν_2 and $\rho\le\rho_{\mathcal{F}}\cdot\min\{p_{\nu_1}^{-\kappa},p_{\nu_2}^{-\kappa}\}$, there is no ρ-partial, $\rho^{m_{\mathcal{F}}}$-approximate homomorphism $f:1_\rho^{(1)}\to G_{\nu_2}$ which satisfies $1_\rho^{(2)}\subseteq(\mathrm{Im}(f))_{\rho^{m_{\mathcal{F}}}}$.

2.1. **Outline of the arguments.** We start with outlining the proof of Theorem 1. The proof will be carried out in several steps and relies heavily on the results of [20].

Step 1. The first step of the proof is to reduce proof of the spectral gap property of the random walk by μ, the probability law of X, to the study of functions which live at small scales. This is done using the notion of locally random group and the Littlewood-Paley theory for these groups, which was developed in [20]. In particular, we showed in [20, Theorem 2.10, Theorem 9.3] that if for every $\eta\le\eta_0\ll_{G_1,G_2}1$ and every function f that lives at scale η, we have the polynomial contraction

$$\|\mu^{(\ell)}*f\|_2\le\eta^c\|f\|_2\qquad\text{for some integer }\ell\le C\log(1/\eta),\tag{1}$$

then for every function $g\in L^2(G_1\times G_2)$ which is orthogonal to an *exceptional* finite dimensional subspace $\mathcal{H}_0$, we have $\|\mu*g\|_2\le 2^{-c/C}\|g\|_2$ — spectral gap for the space orthogonal to the exceptional subspace. We refer the reader to Definition 3 for the definition of functions living at scale η.

Step 2. To obtain (1), we discretize the groups G_1 and G_2 at scale $\eta^{O(1)}$ and use the spectral gap of the marginals of μ to find a coupling σ of the Haar measures m_{G_1} and m_{G_2} which is *close* to $\mu^{(\ell)}$ at scale $\eta^{O(1)}$. The aforementioned coupling is constructed using the transportation problem, see §3.7. This reduces the proof of (1) to showing that for all f that live at scale $\eta^{O(1)}$, we have

$$\|\sigma^{(n_0)}*f\|_2\le\eta^c\|f\|_2\qquad\text{for some }n_0\text{ depending only on }\dim G_1\text{ and }\dim G_2.\tag{2}$$

Step 3. In this step, we use the mixing inequality [20, Theorem 2.6] and the multiscale Bourgain-Gamburd proved in [20, Theorem 2.12] to show that the failure of (2) yields a partial approximate homomorphism between G_1 and G_2, in the sense of Definition 1.

Step 4. This step relies on Theorem 3, which is of independent interest. Indeed by loc. cit. we conclude that such partial approximate homomorphisms exists only if G_1 and G_2 are locally isomorphic. This concludes the proof of Theorem 1.

Let us now briefly outline the proof of Theorems 3. The proof relies on the Baker-Campbell-Hausdorff formula and bounded generation properties of simple Lie groups at small scales, see Proposition 39. Our argument also relies on the effective version of Nullstellensatz in the form of Łojaswicz inequality (real case) and the work of Greenberg (p-adic case). The details occupy §4 in the paper.

We end this outline by discussing the proof of Theorem 2. As alluded to before, the proof relies on Theorem 1. Indeed, our proof of Theorem 1 yields estimates on the implied constants in terms of group theoretic properties of the groups G_1 and G_2 in loc. cit. In the case considered in Theorem 2, these estimates depend only on the set Ω — this deduction relies on strong approximation theorem. This uniformity reduces the proof to the analysis of the exceptional representation $\mathcal{H}_0$ appearing in Step 1. In [20, Theorem 9.3] we showed that this representation has dimension bounded by $(p_{\nu_1}p_{\nu_2})^{O(1)}$. We use this fact and results in [24, 25] to adapt the above general outline and study functions in $\mathcal{H}_0$, thus completing the proof of Theorem 2.

3. Preliminaries and notation

In this section, we will set some notation needed for the paper and recall a number of basic facts that we will refer to in the sequel.

3.1. **Analysis on compact groups.** Let G be a compact (separable) topological group. As it is well known, G can be equipped with a bi-invariant metric that induces the topology of G. All measures on G in this paper will be assumed to be finite measures; thus they are automatically Radon measures. We let m_G denote the unique bi-invariant probability Haar measure on G. For any Borel subset $A \subseteq G$, the Haar measure of A is denoted by $m_G(A)$ or $|A|$. The cardinality of a finite set A will be denoted by $\#A$.

For a Borel measurable function $f : G \to \mathbb{C}$, the integral of f with respect to the Haar measure is denoted, interchangeably, by $\int_G f$ or $\int_G f(y)\,\mathrm{d}y$. We denote by $L^p(G)$ the space $L^p(G, m_G)$. For $f \in L^p(G)$, we write

$$\|f\|_p = \Big(\int_G |f(x)|^p \,\mathrm{d}x \Big)^{1/p}.$$

We also denote by $C(G)$ the Banach space of complex-valued continuous functions $f : G \to \mathbb{C}$, equipped with the supremum norm. For $f, g \in L^1(G)$ the convolution $f * g$ is defined by

$$(f * g)(x) = \int_G f(y)g(y^{-1}x)\,\mathrm{d}y. \tag{3}$$

It is a fact that $(L^1(G), +, *)$ is a Banach algebra and if $f \in L^1(G)$ is a class function, then f is in the center of this Banach algebra.

For Borel measures μ and ν on G, the convolution $\mu * \nu$ is the unique Borel measure on G such that for all $f \in C(G)$,

$$\int_G f \,\mathrm{d}(\mu * \nu) = \int_G \int_G f(xy)\,\mathrm{d}\mu(x)\,\mathrm{d}\nu(y).$$

For a Borel measure μ on G and $f \in L^1(G)$, the convolution $\mu * f$ is defined by

$$(\mu * f)(x) = \int_G f(y^{-1}x)\,\mathrm{d}\mu(y). \tag{4}$$

The following special cases of Young's inequality will be freely used in this paper: for $f, g \in L^2(G)$ and probability measure μ,

$$\|f * g\|_2 \le \|f\|_1\,\|g\|_2, \quad \|f * g\|_\infty \le \|f\|_2\,\|g\|_2, \quad \|\mu * f\|_2 \le \|f\|_2. \tag{5}$$

Let G be a compact Hausdorff second countable topological group. When $\mathscr{H}$ is a Hilbert space and $T : \mathscr{H} \to \mathscr{H}$ is a bounded linear operator, we define the operator norm of T by

$$\|T\|_{\mathrm{op}} := \sup_{v \in \mathscr{H} \setminus \{0\}} \frac{\|Tv\|}{\|v\|}.$$

When $\mathscr{H}$ is finite-dimensional, the Hilbert-Schmidt norm of T is defined by

$$\|T\|_{\mathrm{HS}} := (\mathrm{Tr}(TT^*))^{1/2},$$

where T^* denotes the conjugate transpose of the operator T. Note that when S and T are linear operators on a finite-dimensional Hilbert space $\mathscr{H}$, the following inequality holds

$$\|TS\|_{\mathrm{HS}} \le \|T\|_{\mathrm{op}}\|S\|_{\mathrm{HS}}.$$

3.2. **The Peter-Weyl theorem.** The set of equivalence classes of irreducible unitary representations of G is called the unitary dual of G and is denoted by $\widehat{G}$.

The group G acts on $L^2(G)$ via $(g \cdot f)(x) = f(g^{-1}x)$, preserving the L^2-norm. Hence, it defines a unitary representation of G on $L^2(G)$, the regular representation of G.

Let us enumerate a number of well known facts about unitary representations of G. It is well known that every $\pi \in \widehat{G}$ is of finite dimension, and that every unitary representation of G can be decomposed as an orthogonal direct sum of $\pi \in \widehat{G}$. A function $f \in L^2(G)$ is called G-finite if there exists a finite-dimensional G-invariant subspace of $L^2(G)$ containing f. It is clear that G-finite functions form a subspace of $L^2(G)$. We will denote this subspace by $\mathcal{E}(G)$. It follows from the classical theorem of Peter-Weyl that $\mathcal{E}(G) \subseteq C(G)$ and that $\mathcal{E}(G)$ is dense in $L^2(G)$.

For $\pi \in \widehat{G}$ and $f \in L^1(G)$, the Fourier coefficient $\widehat{f}(\pi)$ is defined by

$$\widehat{f}(\pi) = \int_G f(g)\pi(g)^* \,\mathrm{d}\mu(g).$$

One can show that for $f, g \in L^1(G)$ and $\pi \in \widehat{G}$, we have

$$\widehat{f * g}(\pi) = \widehat{g}(\pi)\widehat{f}(\pi).$$

Parseval's theorem states that for all $f \in L^2(G)$ the following identity holds:

$$\|f\|_2^2 = \sum_{\pi \in \widehat{G}} \dim \pi \, \|\widehat{f}(\pi)\|_{\mathrm{HS}}^2.$$

3.3. **Spectral Gap.** Let μ be a Borel probability measure on G and $(X_i)_{i\geq 1}$ be a sequence of independent G-valued random variables with probability law μ. An ℓ-step random walk on G with respect to μ is given by the random variable

$$X^{(\ell)} := X_1 \cdots X_\ell.$$

Let us note that the law of $X^{(\ell)}$ is given by the ℓ-fold convolution $\mu^{(\ell)}$ of μ. Assume that μ is symmetric and that the subgroup generated by the support of μ is dense in G. Consider the averaging operator

$$T_\mu : L_0^2(G) \to L_0^2(G), \quad T_\mu(f) = \mu * f,$$

where $L_0^2(G) := \{g \in L^2(G) \mid \int f = 0\}$.

Definition 2. We say that μ has *the spectral gap property* if

$$\lambda(\mu; G) := \|T_\mu\|_{\mathrm{op}} < 1.$$

More generally, for a subrepresentation $(\pi, \mathcal{H}_\pi)$ of $L_0^2(G)$, we let

$$\lambda(\mu; \mathcal{H}_\pi) := \|T_\mu|_{\mathcal{H}_\pi}\|_{\mathrm{op}} \quad \text{and} \quad \mathcal{L}(\mu; \mathcal{H}_\pi) := -\log \lambda(\mu; \mathcal{H}_\pi).$$

3.4. **Metric and Rényi entropy.** In this subsection, we will collect a number of definitions from additive combinatorics and [20] that will be needed later. Let G be as above, and let d denote a bi-invariant metric on G. The ball of radius $\eta > 0$ centered at $x \in G$ is denoted by x_η. The η-neighborhood of a set A, denoted by A_η, is the union of all x_η with $x \in A$.

A subset $A \subseteq G$ is said to be η-separated if the distance between every two points in A is at least η. An η-cover for A is a collection of balls of radius η with centers in A whose union covers A. Recall that the minimum size of an η-cover of A (which is finite by compactness of G) is denoted by $\mathcal{N}_\eta(A)$. The value

$$h(A; \eta) := \log \mathcal{N}_\eta(A)$$

is called the metric entropy of A at scale η.

The characteristic function of a set A is denoted by $\mathbb{1}_A$. For $\eta > 0$, we write $P_\eta = \frac{\mathbb{1}_{1_\eta}}{|1_\eta|}$. Note that P_η belongs to the center of the Banach algebra $L^1(G)$. For $f \in L^1(G)$ (μ a probability measure on G, respectively) we write f_η (μ_η, respectively) instead of $f * P_\eta$ ($\mu * P_\eta$, respectively).

The Rényi entropy of a G-valued Borel random variable X at scale $\eta > 0$ is defined by

$$H_2(X; \eta) := \log(1/|1_\eta|) - \log \|\mu_\eta\|_2^2,$$

where μ is the probability law of X. As $H_2(X; \eta)$ depends only on the law μ of X, we will sometimes write $H_2(\mu; \eta)$ instead of $H_2(X; \eta)$.

Let us also recall [20, Definition 8.7]:

Definition 3. We say $f \in L^2(G)$ *lives at scale* η *(with parameter* $0 < a < 1$*)* if

- (Averaging to zero) $\|f_{\eta^{1/a}}\|_2 \le \eta^{1/(2a)}\|f\|_2$.
- (Almost invariant) $\|f_{\eta^{a^2}} - f\|_2 \le \eta^{a/2}\|f\|_2$.

3.5. **Local randomness and the dimension condition.** In this subsection, we will recall the definitions of local randomness and the dimension condition from [20].

Definition 4. Let G be a compact group equipped with a compatible metric d. We say (G, d) satisfies a dimension condition $\mathrm{DC}(C_1, d_0)$ if there exist $C_1 \ge 1$ and $d_0 > 0$ such that for all $\eta \in (0, 1)$ the following bounds hold.

$$\frac{1}{C_1}\eta^{d_0} \le |1_\eta| \le C_1\eta^{d_0}. \tag{DC}$$

If the investigation involves two groups G_1 and G_2, we will distinguish their corresponding constants by an additional subscript, e.g., C_{11} and C_{12}.

Definition 5. Suppose G is a compact group and d is a compatible bi-invariant metric on G. For parameters $C_0 \ge 1$ and $L \ge 1$, we say (G, d) is L*-locally random* with coefficient C_0 if for every irreducible unitary representation π of G and all $x, y \in G$ the following inequality holds:

$$\|\pi(x) - \pi(y)\|_{\mathrm{op}} \le C_0(\dim \pi)^L d(x, y). \tag{6}$$

We say a compact group G is *locally random* if (G, d) is L-locally random with coefficient C_0 for some bi-invariant metric d on G, and some values of L and C_0.

3.6. **Standard metrics on analytic Lie groups.** We will be primarily interested in compact analytic real and p-adic Lie groups with simple Lie algebras. Let G be such a group, a bi-invariant metric d on G will be said to be *standard* if $\mathrm{diam}(G) \le 1$ and the following properties are satisfied:

(1) If G is a compact real Lie group, let d_0 be the metric induced from the Killing form. We assume there is $c_d \ge 1$ so that

$$c_d^{-1} d_0(g, 1) \le d(g, 1) \le c_d d_0(g, 1) \quad \text{for all } g \in G$$

For instance, if $G \subseteq \mathrm{SO}_n(\mathbb{R})$, the metric induced by the operator norm satisfies the above property.

(2) In the p-adic case, we assume $G \subseteq \mathrm{SL}_n(\mathbb{Z}_p)$ and take d to be the metric induced from the operator norm on $\mathrm{SL}_n(\mathbb{Z}_p)$. Note that in this case:
   (a) 1_η is a subgroup for every $0 < \eta \le 1$, and
   (b) the condition DC(C, d_0) holds for some positive constants C and $d_0 = \dim G$.

Lemma 4. *Both of the following hold.*

(1) Let G be a real and p-adic Lie groups with simple Lie algebra, then G is L-locally random with coefficient C_0.

(2) Let $\mathbb{G}$ be an absolutely almost simple simply connected $\mathbb{Q}$-algebraic group. Let $\Gamma \subseteq \mathbb{G}(\mathbb{Q})$ be a finitely generated Zariski dense subgroup in $\mathbb{G}$. Denote by V_Γ the set of all places ν of $\mathbb{Q}$ such that Γ is a bounded subset of $\mathbb{G}(\mathbb{Q}_\nu)$. For distinct $\nu_1, \nu_2 \in V_\Gamma$, let Γ_{ν_1,ν_2} denote the closure of Γ in $\mathbb{G}(\mathbb{Q}_{\nu_1}) \times \mathbb{G}(\mathbb{Q}_{\nu_2})$, and assume that $\Gamma_{\nu_1,\nu_2} = \Gamma_{\nu_1} \times \Gamma_{\nu_2}$. Then Γ_{ν_1,ν_2} is L-locally random with coefficient C_0 where L and C_0 depend only on Γ.

Proof. Part (1) is proved in [20, Section 5].

We now turn to the proof of part (2). In view of [20, Lemma 5.2] and the fact that $\Gamma_{\nu_1,\nu_2} = \Gamma_{\nu_1} \times \Gamma_{\nu_2}$. It suffices to prove that Γ_ν is L'-locally random with coefficient C_0' where L' and C_0' depend only on Γ for all $\nu \in V_\Gamma$. To see this, let $\hat{\Gamma}$ denote the closure of Γ in $\prod_{\nu \in V_{\Gamma,\mathrm{f}}} \mathbb{G}(\mathbb{Z}_\nu)$, where

$V_{\Gamma,\mathrm{f}} = V_\Gamma \setminus \{\infty\}$. Then by [25, Proposition 19], for all but finitely many representations ρ of $\hat{\Gamma}$, we have

$$\ell(\rho) < \dim(\rho)^A$$

where A depends only on Γ and $\ell(\rho)$ denotes the smallest integer k so that ρ factors through k-th congruence quotient of $\hat{\Gamma}$.

Let now $\bar{C}$ denote the maximum of levels of the finitely many exceptional representations (in the above sense); then Γ_ν is $(\bar{C}, A)$-metric quasi random in sense of [20, Definition 5.7]. This and [20, Proposition 5.9] finish the proof. □

3.7. **The transportation problem and coupling of measures.** Recall that a coupling of probability measures μ_1 and μ_2 defined on probability spaces $(\Omega_1, \mathscr{B}_1, \mu_1)$ and $(\Omega_2, \mathscr{B}_2, \mu_2)$ is a probability measure μ on the product probability space $(\Omega_1 \times \Omega_2, \mathscr{B}_1 \otimes \mathscr{B}_2)$ such that

$$\mathrm{pr}_i(\mu) = \mu_i$$

holds for $i = 1, 2$. The set of all couplings of the probability measures μ_1 and μ_2 is denoted by $\mathscr{C}(\mu_1, \mu_2)$. A simple example of coupling of two measures μ_1 and μ_2 is the product measure $\mu_1 \otimes \mu_2$. There are, however, many more examples. For instance, when Ω_1 and Ω_2 are finite sets, respectively of cardinality n_1 and n_2 elements, then $\mathscr{C}(\mu_1, \mu_2)$ is a convex set of dimension $(n_1 - 1)(n_2 - 1) + 1$. When the probability spaces Ω_1, Ω_2 are finite, the question of determining the couplings has been of interest in operation research. In the special case that Ω_1 and Ω_2 have the same cardinality, couplings of μ_1 and μ_2 correspond to doubly stochastic matrices. It is a well-known theorem that every doubly stochastic matrix can be expressed as a convex combination of permutation matrices. We will use a less well-known generalization of this result, which is established in [17].

Proposition 5. *Let Y_1 and Y_2 be two finite sets with $|Y_i| = N_i > 1$, $i = 1, 2$. Let $\tilde{\mu}$ be a probability measure on $Y_1 \times Y_2$ which satisfies the following: there exists some $A > 2$ so that*

$$\left|\pi_i \tilde{\mu}(y) - \tfrac{1}{N_i}\right| \le \tfrac{1}{(N_1 N_2)^A} \qquad \text{for all } y \in Y_i \text{ and } i = 1, 2, \tag{7}$$

where π_i denotes the projection onto the i-th coordinate. Then there exists a coupling $\tilde{\nu}$ of the uniform measures on Y_i so that

$$|\tilde{\mu}(y_1, y_2) - \tilde{\nu}(y_1, y_2)| \le \tfrac{1}{(N_1 N_2)^{A-1}}.$$

We begin by fixing some notation. Let $T(Y_1, Y_2)$ denote the set of spanning trees of the complete bipartite graph K_{Y_1,Y_2}. Given probability measures σ_i on Y_i, for $i = 1, 2$, and a spanning tree $\tau \in T(Y_1, Y_2)$, define

$$M^\tau_{\sigma_1,\sigma_2} : Y_1 \times Y_2 \to \mathbb{R} \qquad \text{by} \qquad M^\tau_{\sigma_1,\sigma_2}(y_1, y_2) := \sigma_1(Y_1') - \sigma_2(Y_2')$$

where $Y_i' \subset Y_i$ and $Y_1' \cup Y_2'$ is the connected component of $\tau \setminus \overline{y_1 y_2}$ that contains y_1; as usual, $\overline{y_1 y_2}$ denotes the edge connecting y_1 to y_2. Put

$$\mathscr{T}(\sigma_1, \sigma_2) := \{\tau \in T(Y_1, Y_2) : M^\tau_{\sigma_1,\sigma_2} \ge 0\}. \tag{8}$$

The proof of Proposition 5 is based on the following.

Theorem 6 ([17]). *Let $\sigma \in \mathscr{C}(\sigma_1, \sigma_2)$. Then σ belongs to the convex hull of*

$$\{M^\tau_{\sigma_1,\sigma_2} : \tau \in \mathscr{T}(\sigma_1, \sigma_2)\}.$$

Proof of Proposition 5. Let us denote the uniform measure on Y_i by m_{Y_i}, and write $\tilde{\mu}_i = \pi_i(\tilde{\mu})$ for $i = 1, 2$. We first show the following:

$$\mathscr{T}(\tilde{\mu}_1, \tilde{\mu}_2) \subset \mathscr{T}(m_{Y_1}, m_{Y_2}). \tag{9}$$

To see (9), let $\tau \in \mathscr{T}(\tilde\mu_1, \tilde\mu_2)$. Then

$$\begin{aligned} M^{\tau}_{m_{Y_1}, m_{Y_2}}(y_1, y_2) &= m_{Y_1}(Y_1') - m_{Y_2}(Y_2') \\ &= M^{\tau}_{\tilde\mu_1, \tilde\mu_2}(y_1, y_2) + \big(m_{Y_1}(Y_1') - \tilde\mu_1(Y_1')\big) - \big(m_{Y_2}(Y_2') - \tilde\mu_2(Y_2')\big) \end{aligned}$$

Since $M^{\tau}_{\tilde\mu_1, \tilde\mu_2}(y_1, y_2) \geq 0$, we conclude from the above and (7) that

$$M^{\tau}_{m_{Y_1}, m_{Y_2}}(y_1, y_2) \geq -\tfrac{|Y_1'|}{(N_1N_2)^A} - \tfrac{|Y_2'|}{(N_1N_2)^A} \geq -\tfrac{1}{(N_1N_2)^{A-1}}\big(\tfrac{1}{N_1} + \tfrac{1}{N_2}\big) > -\tfrac{1}{N_1N_2}, \tag{10}$$

where in the last inequality we used $A > 2$.

On the other hand, if $M^{\tau}_{m_{Y_1}, m_{Y_2}}(y_1, y_2) < 0$, then

$$M^{\tau}_{m_{Y_1}, m_{Y_2}}(y_1, y_2) = \tfrac{|Y_1'|}{N_1} - \tfrac{|Y_2'|}{N_2} \leq -\tfrac{1}{N_1N_2}.$$

This and (10), imply that $M^{\tau}_{m_{Y_1}, m_{Y_2}}(y_1, y_2) \geq 0$ as was claimed in (9).

Recall now that $\tilde\mu \in \mathscr{C}(\tilde\mu_1, \tilde\mu_2)$, by Theorem 6, thus, there exists $\{c_\tau \in [0,1] : \tau \in \mathscr{T}(\tilde\mu_1, \tilde\mu_2)\}$, with $\sum c_\tau = 1$ so that

$$\tilde\mu = \sum_{\tau \in \mathscr{T}(\tilde\mu_1, \tilde\mu_2)} c_\tau M^{\tau}_{\tilde\mu_1, \tilde\mu_2}.$$

Define $\tilde\nu := \sum_{\tau \in \mathscr{T}(\tilde\mu_1, \tilde\mu_2)} c_\tau M^{\tau}_{m_{Y_1}, m_{Y_2}}$; we will show that the proposition holds with $\tilde\nu$. In view of (9), $\tilde\nu \in \mathscr{C}(m_{Y_1}, m_{Y_2})$. Moreover, since $\sum c_\tau = 1$, we have

$$\begin{aligned} |\tilde\mu(y_1, y_2) - \tilde\nu(y_1, y_2)| &\leq \max_{\tau \in \mathscr{T}(\tilde\mu_1, \tilde\mu_2)} |M^{\tau}_{m_{Y_1}, m_{Y_2}}(y_1, y_2) - M^{\tau}_{\tilde\mu_1, \tilde\mu_2}(y_1, y_2)| \\ &\leq |m_{Y_1}(Y_1') - \tilde\mu_1(Y_1')| + |m_{Y_2}(Y_2') - \tilde\mu_1(Y_2')| \\ &\leq \tfrac{1}{(N_1N_2)^{A-1}}\big(\tfrac{1}{N_1} + \tfrac{1}{N_2}\big) \leq \tfrac{1}{(N_1N_2)^{A-1}}. \end{aligned}$$

The proof is complete. □

4. Approximate homomorphisms

The main goal of this section is to prove Theorem 3, in which we investigate *partial approximate homomorphisms* between open compact subgroups of almost simple analytic groups, see Definition 1.

A G_1-partial δ-approximate homomorphism is simply called a δ-approximate homomorphism. Approximate homomorphisms have been studied extensively. In [16], Kazhdan used cohomological methods to show that approximate homomorphisms from an amenable group to $\mathrm{U}_n(\mathbb{C})$ are close to group homomorphisms. His argument is based on defining an averaging operator on the space of cocycles and proving that this operator is a *contraction*. These arguments do not appear to work when the domain of f is not a group or when the target is a p-adic group. In [14], authors used a similar approach to prove that approximate homomorphisms between Lie groups are close to homomorphisms.

In [9, 8], Farah, using more combinatorial techniques, proved a similar result for approximate homomorphisms from finite groups of product form.

Our arguments here are different from these works. We rely on local analysis and passing to an infinitesimal setting, we also appeal to effective Nullstellensatz and the Łojasiewicz inequality.

It will be more convenient to treat the cases where $F_2 = \mathbb{Q}_p$ and $F_2 = \mathbb{R}$, separately.

4.1. **Target is p-adic analytic.** In order to simplify the notation in this section, we will write d_1 for d_{01} and d_2 for d_{02}, also, we will write C_1 for C_{11} and C_2 for C_{12}. Without loss of generality we will assume $C_i \geq 2$; note also that $d_i \geq 3$.

In this section, we assume that $F_2 = \mathbb{Q}_p$. Note that in this case, $1^{(2)}_{\rho^m}$ is a normal subgroup of G_2. Therefore,

$$\bar{f} : 1^{(1)}_{\rho} \to G_2/1^{(2)}_{\rho^m}, \quad \bar{f}(g_1) := [f(g_1)] = f(g_1)_{\rho^m}$$

has the following properties: for every $g_1, g_1' \in 1^{(1)}_\rho$ if $g_1 g_1' \in 1^{(1)}_\rho$, then

$$\bar{f}(g_1 g_1') = \bar{f}(g_1)\bar{f}(g_1') \quad \text{and} \quad \bar{f}(g_1^{-1}) = \bar{f}(g_1)^{-1}.$$

Lemma 7. *Suppose that $F_1 = \mathbb{R}$, $F_2 = \mathbb{Q}_p$, and $\rho < 1/3$. Then there is no ρ-partial ρ^m almost homomorphism from $1^{(1)}_\rho$ into G_2 so that $1^{(2)}_{\rho^{m'}} \subset \big(f(1^{(1)}_\rho)\big)_{\rho^m}$ so long as $m \gg_{d_1,d_2} m'$.*

Proof. Assume contrary to the claim that $f : 1^{(1)}_\rho \to G_2$ is a ρ-partial, ρ^m-almost homomorphism which satisfies $1^{(2)}_{\rho^{m'}} \subset \big(f(1^{(1)}_\rho)\big)_{\rho^m}$. Define $\bar{f}$ as above.

Since $F_1 = \mathbb{R}$, $\log : 1^{(1)}_\rho \to \mathfrak{g}_1$ is a convergent series, and $\|\log g_1\| \leq 2\|g_1 - I\|$. We also note that

$$\|\exp(x) - I\| \leq 3\|x\| \tag{11}$$

for every $x \in \mathfrak{g}_1$ with $\|x\| < 1$.

Let $r := |G_2/1^{(2)}_{\rho^m}|$, and for $g_1 \in 1^{(1)}_{\rho/6}$, let $x := \frac{\log g_1}{r}$. Then $\|x\| \leq \rho/(3r)$ and for every integer $0 \leq j \leq r$ we have $\exp(jx) = \exp(x)^j \in 1^{(1)}_\rho$. Hence

$$\bar{f}(g_1) = \bar{f}(\exp(x))^r = [1^{(2)}]. \tag{12}$$

As $\bar{f}$ preserves multiplication, (12) and (DC) imply that

$$|\operatorname{Im} \bar{f}| \leq e^{h(1^{(1)}_\rho; \rho/6)} \leq C_1^2 6^{d_{01}}$$

where $h(1^{(1)}_\rho; \rho/6)$ denotes the metric entropy of $1^{(1)}_\rho$ at scale $\rho/6$.

By our assumption, $1^{(2)}_{\rho^{m'}}/1^{(2)}_{\rho^m} \subseteq \operatorname{Im} \bar{f}$, which implies that $C_2 \rho^{(m'-m)d_{02}} \leq C_1^2 6^{d_{01}}$. This is a contradiction for $m \gg_{d_1,d_2} m'$ as C_1 depends only on d_1. ☐

In view of Lemma 7, we will assume $F_1 = \mathbb{Q}_q$ in the remaining parts of §4.1. In particular, $1^{(1)}_\rho$ is a pro-q group, and $1^{(2)}_\rho/1^{(2)}_{\rho^m}$ is a finite p-group which is in the image of the group homomorphism $\bar{f}$. This implies that $p = q$. In this case, all the balls centered at the identity $1^{(i)}$ are congruence subgroups of G_i. The ball of radius p^{-k} centered at the identity $1^{(i)}$ is

$$G_{i,k} := \{g \in G_i |\, g \equiv I \pmod{p^k}\}. \tag{13}$$

The following theorem applied with $\varphi = \bar{f}$ implies Theorem 3 in this case.

Theorem 8. *Let $\mathbb{G}_1$ and $\mathbb{G}_2$ be almost $\mathbb{Q}_p$-simple groups. Suppose $G_i \subseteq \mathrm{GL}_{n_i}(\mathbb{Z}_p)$ are open compact subgroups of $\mathbb{G}_i(\mathbb{Q}_p)$ for $i = 1, 2$. Then there is $0 < c_3 \leq 1$ and a positive integer m_0 both depending only on $\dim \mathbb{G}_1$ and $\dim \mathbb{G}_2$ such that the following holds:*

If $k_0 \gg_{\mathbb{G}_1,\mathbb{G}_2} 1$ and $m \geq m_0 m'$, and

$$\varphi : G_{1,k_0} \to G_2/G_{2,k_0 m}$$

is a group homomorphism, satisfying that $G_{2,k_0 m'}/G_{2,k_0 m} \subseteq \mathrm{Im}(\varphi)$, then there is a group homomorphism $\pi : G_{1,k_0} \to G_2$ so that

$$\varphi \equiv \pi \pmod{p^{\lfloor c_3 k_0 m \rfloor}}$$

and $\mathrm{d}\pi$ *induces an isomorphism between* $\mathrm{Lie}(\mathbb{G}_1)(\mathbb{Q}_p)$ *and* $\mathrm{Lie}(\mathbb{G}_2)(\mathbb{Q}_p)$. *Moreover if* $\mathbb{G} = \widetilde{\mathbb{G}} \otimes_{\mathbb{Q}} \mathbb{Q}_p$ *where* $\widetilde{\mathbb{G}}$ *is an absolutely almost simple* $\mathbb{Q}$*-group, then the constant* k_0 *depends only on* $\widetilde{\mathbb{G}}$.

Proof. Let $m_0 \geq 1$ be a constant which will be explicated at the end of the argument.

Let $\mathfrak{g}_i := \mathfrak{gl}_{n_i}(\mathbb{Z}_p) \cap \mathrm{Lie}(\mathbb{G}_i)(\mathbb{Q}_p)$ be the Lie $\mathbb{Z}_p$-algebra of G_i. We start by recalling properties of finite logarithmic functions from [24, Lemma 34]. We can and will assume that k_0 is large enough such that for every integer $n \geq k_0$,

$$\exp : p^n\mathfrak{g}_i \to G_{i,n} \quad \text{and} \quad \log : G_{i,n} \to p^n\mathfrak{g}_i$$

are well-defined and inverse of each other; in particular, $\|\exp(x) - I\| = \|x\|$ and $\|\log g\| = \|g - I\|$. For every integers $l \geq k_0$ and $n \in [l, 2l - k_0]$, the function

$$\Psi^{p^n}_{p^l} : G_{i,l}/G_{i,n} \to \mathfrak{g}_i/p^{n-l}\mathfrak{g}_i, \quad \Psi^{p^n}_{p^l}(gG_{i,n}) := \frac{g - 1^{(i)}}{p^l} + p^{n-l}\mathfrak{g}_i$$

is a well-defined bijective G_i-equivariant function, where G_i acts by conjugation on $G_{i,l}/G_{i,n}$ and on $\mathfrak{g}_i/p^{n-l}\mathfrak{g}_i$ via the adjoint representation. Notice that we further have

$$\Psi^{p^n}_{p^l}(gG_{i,n}) = \frac{\log g}{p^l} + p^{n-l}\mathfrak{g}_i.$$

If $l, l' \geq k_0$, $n \in [l, 2l - k_0]$, and $n' \in [l', 2l' - k_0]$, then

$$\Psi^{p^{n''}}_{p^{l+l'}}([g_i, g_i']G_{i,n''}) = [\Psi^{p^n}_{p^l}(gG_{i,n}), \Psi^{p^{n'}}_{p^{l'}}(gG_{i,n'})] + p^{n''-l-l'}\mathfrak{g}_i \tag{14}$$

where $n'' := \min(n + l', n' + l)$.

By [25, Lemma 39], the Frattini subgroup $\Phi(G_{i,n})$ of $G_{i,n}$ is $G_{i,n+1}$ for every integer $n \geq k_0$. Hence for every positive integer $l \leq k_0(m - 1)$, we have

$$G_{2,k_0m'+l}/G_{2,k_0m} \subseteq \varphi(G_{1,k_0+l}) \subseteq G_{2,1+l}/G_{2,k_0m}. \tag{15}$$

In view of (15), φ induces a group homomorphism

$$\varphi_{l,n} : G_{1,l}/G_{1,n} \to G_{2,l-k_0+1}/G_{2,n-k_0+1}, \quad \varphi_{l,n}(g_1G_{1,n}) := \varphi(g_1) \pmod{G_{2,n-k_0+1}},$$

where $k_0 < l \leq n \leq k_0(m - 1)$.

Using the finite logarithmic maps, for integers $k_0 < l \leq n \leq 2l - 2k_0 + 1 \leq k_0(m - 1)$, there is an additive group homomorphism $\theta_{l,n}$ such that the following is a commuting diagram:

$$\begin{array}{ccc} G_{1,l}/G_{1,n} & \xrightarrow{\varphi_{l,n}} & G_{2,l-k_0+1}/G_{2,n-k_0+1} \\ \big\downarrow \Psi^{p^n}_{p^l} & & \big\downarrow \Psi^{p^{n-k_0+1}}_{p^{l-k_0+1}} \\ \mathfrak{g}_1/p^{n-l}\mathfrak{g}_1 & \xrightarrow{\theta_{l,n}} & \mathfrak{g}_2/p^{n-l}\mathfrak{g}_2. \end{array} \tag{16}$$

By (14) and (16), we deduce that

$$\theta_{2l-k_0+1,l+n-k_0+1} : \mathfrak{g}_1/p^{n-l}\mathfrak{g}_1 \to \mathfrak{g}_2/p^{n-l}\mathfrak{g}_2$$

is a Lie ring homomorphism for integers $k_0 < l \leq n \leq 2l - 2k_0 + 1$ and $l + n - k_0 + 1 \leq k_0(m - 1)$. We get a Lie ring homomorphism $\theta_m := \theta_{l_m,n_m}$ such that the following is a commuting diagram

$$\begin{array}{ccc} G_{1,l_m}/G_{1,n_m} & \xrightarrow{\varphi_{l_m,n_m}} & G_{2,l_m-k_0+1}/G_{2,n_m-k_0+1} \\ \big\downarrow \Psi^{p^{n_m}}_{p^{l_m}} & & \big\downarrow \Psi^{p^{n_m-k_0+1}}_{p^{l_m-k_0+1}} \\ \mathfrak{g}_1/p^{n_m-l_m}\mathfrak{g}_1 & \xrightarrow{\theta_m} & \mathfrak{g}_2/p^{n_m-l_m}\mathfrak{g}_2, \end{array} \tag{17}$$

where
$$l_m := 2\left\lceil \frac{k_0}{3}m + \frac{2k_0-2}{3}\right\rceil - k_0 + 1, \quad \text{and} \quad n_m := k_0(m-1) - 2.$$
Notice that $n_m - l_m \ge \frac{k_0}{3}(m-4) - \frac{8}{3}$.

Let $\{e_1^{(i)}, \ldots, e_{d_i}^{(i)}\}$ be a $\mathbb{Z}_p$-basis of $\mathfrak{g}_i$, and suppose $c_{jks}^{(i)} \in \mathbb{Z}_p$ are the corresponding structural constants. That is,
$$[e_j^{(i)}, e_k^{(i)}] = \sum_{s=1}^{d_i} c_{jks}^{(i)} e_s^{(i)}.$$
Then a $\mathbb{Z}_p$-linear map $T : \mathfrak{g}_1 \to \mathfrak{g}_2, T(e_j^{(1)}) = \sum_{s=1}^{d_2} x_{js} e_s^{(2)}$ is a Lie ring homomorphism if and only if $T\big([e_j^{(1)}, e_k^{(1)}]\big) = \big[T(e_j^{(1)}), T(e_k^{(1)})\big]$ for every $1 \le j,k \le d_1$. This is equivalent to having the following equations:

(18)
$$\sum_{1\le i_1,i_2\le d_2} c_{i_1 i_2 r}^{(2)} x_{j i_1} x_{k i_2} = \sum_{i=1}^{d_2} c_{jki}^{(1)} x_{ri}$$

for every integers $1 \le j,k \le d_1$ and $1 \le r \le d_2$.

Let V be the $\mathbb{Z}_p$-affine scheme given by equations in (18). By the main theorem of [13], there are positive integers $C_4(V)$ and $c_4(V)$ such that for every point in $\overline{\mathbf{x}} \in V(\mathbb{Z}_p/p^n\mathbb{Z}_p)$ there is a point $\mathbf{x} \in V(\mathbb{Z}_p)$ such that
$$\mathbf{x} \equiv \overline{\mathbf{x}} \pmod{p^{\frac{n-c_4(V)}{C_4(V)}}}.$$
A close examination of the argument in [13] yields the following: $C_4(V)$ depends only on the number of variables and the degree of the defining equations; hence $C_4(V)$ only depends on $\dim \mathbb{G}_1$ and $\dim \mathbb{G}_2$. The constant $c_4(V)$ could depend on the defining equations of V viz. the complexity of rational numbers $c_{jks}^{(i)}$. In particular, if $\mathbb{G} = \widetilde{\mathbb{G}} \otimes_{\mathbb{Q}} \mathbb{Q}_p$ where $\widetilde{\mathbb{G}}$ is an absolutely almost simple $\mathbb{Q}$-group, then the constant $c_4(V)$ depends only on $\widetilde{\mathbb{G}}$.

Note that the map θ_m in (17) is a point in $V(\mathbb{Z}_p/p^{n_m-l_m}\mathbb{Z}_p)$. Applying the above with $\overline{\mathbf{x}} = \theta_m$, there exists a Lie ring homomorphism $\theta : \mathfrak{g}_1 \to \mathfrak{g}_2$ such that

(19)
$$\theta \equiv \theta_m \pmod{p^{\lfloor c_3' k_0 m\rfloor}},$$

where $c_3' := (8C_4(V))^{-1}$ so long as k_0 is large enough so that
$$n_m - l_m \ge \tfrac{k_0}{3}(m-4) - \tfrac{8}{3} \ge \tfrac{k_0}{4}m \ge 2c_4(V).$$
In view of the above discussion thus, if $\mathbb{G} = \widetilde{\mathbb{G}} \otimes_{\mathbb{Q}} \mathbb{Q}_p$ where $\widetilde{\mathbb{G}}$ is an absolutely almost simple $\mathbb{Q}$-group, then the constant k_0 depends only on $\widetilde{\mathbb{G}}$ — recall that $m \ge m_0 m' \ge 1$.

Note that for $k_0 \ge 2$, $p^{k_0}\mathfrak{g}_i$ is a powerful Lie ring. Hence by the Baker-Campbell-Hausdorff formula (see [7, Chapter 9.4]), $\pi : G_{1,k_0} \to G_{2,k_0}, \pi(g_1) := \exp(\theta(\log(g_1)))$ is a group homomorphism. Moreover, by the definition the following is a commuting diagram

(20)
$$\begin{array}{ccc} G_{1,k_0} & \xrightarrow{\pi} & G_{2,k_0} \\ \downarrow{\scriptstyle \log} & & \downarrow{\scriptstyle \log} \\ p^{k_0}\mathfrak{g}_1 & \xrightarrow{\theta} & p^{k_0}\mathfrak{g}_2. \end{array}$$

Using the fact that log is G_i-equivariant function, where G_i acts on G_{i,k_0} by conjugation and on $p^{k_0}\mathfrak{g}_i$ via the adjoint action, by (17), (19), and (20), we deduce that for every $\overline{x} \in \mathfrak{g}_2/p^{\lfloor c_3' k_0 m\rfloor}\mathfrak{g}_2$ which is in the image of

(21)
$$\Psi_{p^{l_m-k_0+1}}^{p^{l_m-k_0+1+\lfloor c_3' k_0 m\rfloor}} \circ \varphi$$

and every $g \in G_{1,k_0}$, the following holds:

$$\mathrm{Ad}(\varphi(g))(\overline{x}) = \mathrm{Ad}(\pi(g))(\overline{x}). \tag{22}$$

Note also that, applying (15) with $\ell = \ell_m$, we conclude $p^{k_0 m'}\mathfrak{g}_2/p^{\lfloor c_3' k_0 m\rfloor}\mathfrak{g}_2$ is contained in the image of the map in (21). Therefore, by (22),

$$\mathrm{Ad}(\varphi(g)) \equiv \mathrm{Ad}(\pi(g)) \pmod{p^{\lfloor c_3' k_0 m\rfloor - k_0 m'}}$$

where $\mathrm{Ad}(\varphi(g))$ and $\mathrm{Ad}(\pi(g))$ are written in matrix form with respect to a $\mathbb{Z}_p$-basis of $\mathfrak{g}_2$. Since $\mathbb{G}_2$ is an almost simple $\mathbb{Q}_p$-group, choosing k_0 large enough we deduce that

$$\varphi(g) \equiv \pi(g) \pmod{p^{\lfloor c_3' k_0 m\rfloor - k_0 m' - k_0}}.$$

The claim follows with $c_3 = c_3'/2$ so long as $m \geq m_0 m' \geq 8m'/c_3'$. □

4.2. **Target is a Lie group.** In this section, we assume that $F_2 = \mathbb{R}$.

Let $p_0 := 2$ if $F_1 = \mathbb{R}$, and $p_0 := p$ if $F_1 = \mathbb{Q}_p$. We will also use the condition (DC); recall that in order to simplify the notation in this section we will write d_1 for d_{01} and d_2 for d_{02}, also, we will write C_1 for C_{11} and C_2 for C_{12}. Without loss of generality we will assume $C_i \geq 2$; note also that $d_i \geq 3$.

Lemma 9. *In the setting of Theorem 3, suppose $F_2 = \mathbb{R}$. If $m \geq d_1$ and $\log_{p_0}(1/\rho) \gg_{d_i} 1$. Then there is $g_1 \in 1_\rho^{(1)}$ such that $d(g_1, 1^{(1)}) \geq \rho^2$ and $d(f(g_1), 1^{(2)}) \leq \rho^{d_1/(4d_2)}$.*

Proof. Let $\{h_1, \ldots, h_l\}$ be a set of maximal ρ^2-separated points in $1_\rho^{(1)}$. Hence $l \geq \rho^{-d_1/2}$. The group G_2 can be covered by $\leq C_2\rho'^{-d_2}$-many balls of radius $\rho' := C_2\rho^{d_1/(2d_2)}$ (recall that $C_2 \geq 2$). Note that

$$C_2\rho'^{-d_2} = C_2^{1-d_2}\rho^{-d_1/2} < \rho^{-d_1/2} \leq l.$$

Hence there are $i \neq j$ such that $d(f(h_i), f(h_j)) < \rho'$. Let $g_1 := h_i h_j^{-1}$. Then $d(g_1, 1^{(1)}) \geq \rho^2$, and

$$\begin{aligned} d(f(g_1), 1^{(2)}) \leq& d(f(h_i)f(h_j^{-1}), 1^{(2)}) + \rho^m \\ \leq& d(f(h_i)f(h_j)^{-1}, 1^{(2)}) + 2\rho^m \\ \leq& \rho' + 2\rho^m = C_2\rho^{d_1/(2d_2)} + 2\rho^m \\ \leq& \rho^{d_1/(4d_2)}. \end{aligned}$$

The claim follows. □

In the next lemma we will use the constants C and c appearing in Proposition 39. We note that these two constants depend only on $\widetilde{\mathbb{G}}$ if $\mathbb{G} = \widetilde{\mathbb{G}} \otimes_{\mathbb{Q}} \mathbb{Q}_p$, see Proposition 39.

Lemma 10. *Let C and c be as in Proposition 39. In the setting of Theorem 3, suppose $m \gg_{d_i} 1$ and assume $\rho \leq c^2$ and $\log_{p_0}(1/\rho) \gg_{d_i} 1$. Then*

$$f(1^{(1)}_{\rho^{7C}}) \subseteq 1^{(2)}_{\rho^{d_1/(8d_2)}}.$$

Proof. By Lemma 9, there is $h_0 \in 1_\rho^{(1)}$ such that $d(h_0, 1^{(1)}) \geq \rho^2$ and $d(f(h_0), 1^{(2)}) \leq \rho^{d_1/(4d_2)}$. Then by Proposition 39, applied with G_1, we obtain that the following holds:

$$\left\{(g_{1,\rho}[h_0, a_1]g_{1,\rho}^{-1}) \cdots (g_{d_1^2,\rho}[h_0, a_{d_1^2}]g_{d_1^2,\rho}^{-1}) \mid a_i \in 1_{c\|h_0 - I\|\rho^C}\right\} \supseteq 1_{c^2\|h_0-I\|^2\rho^{2C}}, \tag{23}$$

for some $g_{i,\rho} \in 1_\rho^{(1)}$ where c is as in Proposition 39.

Recall that $\rho \leq c^2$; thus

$$c^2\|h_0 - I\|^2\rho^{2C} \geq c^2\rho^{4+2C} \geq \rho^{5+2C} \geq \rho^{7C}. \tag{24}$$

Since f is ρ^m-almost homomorphism,

$$d\bigg(f\Big((g_{1,\rho}[h_0,a_1]g_{1,\rho}^{-1})\cdots(g_{d_1^2,\rho}[h_0,a_{d_1^2}]g_{d_1^2,\rho}^{-1})\Big),1^{(2)}\bigg)\leq \tag{25}$$

$$d\bigg(\Big(f(g_{1,\rho})[f(h_0),f(a_1)]f(g_{1,\rho})^{-1}\Big)\cdots\Big(f(g_{d_1^2,\rho})[f(h_0),f(a_{d_1^2})]f(g_{d_2{}^2,\rho})^{-1}\Big),1^{(2)}\bigg)+9d_1^2\rho^m.$$

Since $d(hh',1^{(2)})\leq d(h,1^{(2)})+d(h',1^{(2)})$ and $d(hh'h^{-1},1^{(2)})=d(h',1^{(2)})$ for every $h,h'\in G_2$, by (25) we obtain the following upper bound:

$$d\bigg(f\Big((g_{1,\rho}[h_0,a_1]g_{1,\rho}^{-1})\cdots(g_{d_1^2,\rho}[h_0,a_{d_1^2}]g_{d_1^2,\rho}^{-1})\Big),1^{(2)}\bigg)\leq d_1^2\max_i d([f(h_0),f(a_i)],1^{(2)})+9d_1^2\rho^m. \tag{26}$$

Note also that $d([h,h'],1^{(2)})\leq d(h,1^{(2)})+d(h'h^{-1}h'^{-1},1^{(2)})\leq 2d(h,1^{(2)})$ for all $h,h'\in G_2$. Therefore, using (26), we deduce the following

$$d\bigg(f\Big((g_{1,\rho}[h_0,a_1]g_{1,\rho}^{-1})\cdots(g_{d_1^2,\rho}[h_0,a_{d_1^2}]g_{d_1^2,\rho}^{-1})\Big),1^{(2)}\bigg)\leq 2d_1^2\rho^{d_1/(4d_2)}+9d_1^2\rho^m\leq\rho^{d_1/(8d_2)}. \tag{27}$$

By (23), (24), and (27), the claim follows. □

Lemma 11. *Let the notation be as above. If* $m\geq\frac{(7C-1)d_1}{d_2}+m'+1$ *and* $\log_{p_0}(1/\rho)\gg_{d_i}1$*, then there exists some* $g_1\in 1^{(1)}_{\rho^{7C}}$ *so that*

$$d(f(g_1),1^{(2)})\geq\rho^{\frac{(7C-1)d_1}{d_2}+m'+1}.$$

Proof. Let us write $b=\frac{(7C-1)d_1}{d_2}+m'+1$. We prove the lemma by contradiction. Thus, assume that $f(1^{(1)}_{\rho^{7C}})\subset 1^{(2)}_{\rho^b}$. Fix a set $\{a_1,\ldots,a_k\}$ of coset representatives for $1^{(1)}_\rho/1^{(1)}_{\rho^{7C}}$ where $k\leq C_1^2\rho^{(1-7C)d_1}$.

Recall also our assumption that

$$1^{(2)}_{\rho^{m'}}\subset(\mathrm{Im}(f))_{\rho^m}=f(1^{(1)}_\rho)_{\rho^m}.$$

Consequently, for every $h\in 1^{(2)}_{\rho^{m'}}$, there exists some $1\leq i\leq k$ and some $g'\in 1^{(1)}_{\rho^{7C}}$ so that $d(h,f(a_ig'))\leq\rho^m$. Since f is ρ^m-almost homomorphism, we have $d(f(a_ig'),f(a_i)f(g'))\leq\rho^m$. Therefore,

$$d(h,f(a_i))\leq d(h,f(a_ig'))+d(f(a_ig'),f(a_i)f(g'))+d(f(a_i)f(g'),f(a_i))\leq 3\rho^b$$

we used $d(f(g'),1^{(2)})\leq\rho^b$ and $m\geq b$. We thus conclude that

$$1^{(2)}_{\rho^{m'}}\subset\{f(a_1),\ldots,f(a_k)\}_{3\rho^b}. \tag{28}$$

In view of (DC), one gets

$$\Big|\{f(a_1),\ldots,f(a_k)\}_{3\rho^b}\Big|\leq C_1^2\rho^{(1-7C)d_1}\cdot C_2(3\rho^b)^{d_2}=3^{d_2}C_1^2C_2\rho^{(1-7C)d_1+bd_2}\leq 3^{d_2}C_1^2C_2\rho^{(m'+1)d_2}.$$

This contradicts (28), if $\rho^{d_2}<3^{-d_2}C_1^{-2}C_2^{-2}$. □

Corollary 12. *In the setting of Theorem 3, we cannot have* $F_1=\mathbb{Q}_p$ *and* $F_2=\mathbb{R}$ *if* $m\gg_{d_1,d_2,C}m'$ *and* $\log_p(1/\rho)\gg_{d_1,d_2}1$.

Proof. Suppose to the contrary that there is such an approximate homomorphism. By Lemma 10 and Lemma 11, there exists $g_1\in 1^{(1)}_{\rho^{7C}}$ such that

$$\rho^{b_1}\leq d(f(g_1),1^{(2)})\leq\rho^{b_2}. \tag{29}$$

where $b_1=\frac{(7C-1)d_1}{d_2}+m'+1$ and $b_2=\frac{d_1}{8d_2}$.

Since $F_2 = \mathbb{R}$, assuming $\rho^{b_2} \le 1/10$, for all $0 < k \le 3\rho^{b_2-b_1}$ we have

$$d(f(g_1)^k, 1^{(2)}) \ge \frac{1}{2} k\rho^{b_1};$$

therefore, for some $0 < k \le 3\rho^{b_2-b_1}$, we have

$$d(f(g_1)^k, 1^{(2)}) > 2\rho^{b_2}. \tag{30}$$

Let m be large enough so that $3\rho^{b_2-b_1}\rho^m < \rho^{b_2}$. Since $F_1 = \mathbb{Q}_p$, we have $g_1^k \in 1^{(1)}_{\rho^{7C}}$, and

$$\begin{aligned} d(f(g_1^k), 1^{(2)}) &\ge d(f(g_1)^k, 1^{(2)}) - (k-1)\rho^m \\ &\ge 2\rho^{b_2} - 3\rho^{b_2-b_1}\rho^m > \rho^{b_2} \end{aligned} \tag{31}$$

where we used (30) and $3\rho^{b_2-b_1}\rho^m < \rho^{b_2}$.

However, since $g_1^k \in 1^{(1)}_{\rho^{7C}}$, Lemma 10 implies that $d(f(g_1^k), 1^{(2)}) < \rho^{b_2}$. This contradicts (31) and finishes the proof. □

The case where $F_1 = F_2 = \mathbb{R}$. There are certain similarities between this case and the case where $F_1 = F_2 = \mathbb{Q}_p$. However, since in this case there is no reduction mod p map, the argument is more involved.

The following lemmas can be viewed as the Archimedean analogue of (15). We start with finding a *large* subset, see also Lemma 11.

Lemma 13. *In the setting of Theorem 3, suppose $F_1 = F_2 = \mathbb{R}$. Then there is $0 < c'' := c''(G_2) \le 1$ such that for every positive integer k the following holds*

$$1^{(2)}_{c''^{k-1}\rho^{km'}} \subseteq f(1^{(1)}_{2^{k-1}\rho^k})_{6\rho^m}.$$

Proof. We proceed by the induction on k. The base of induction $k = 1$ is part of the assumption. By Proposition 65, we have

$$1^{(2)}_{c''^k\rho^{m'+km'}} \subseteq [1^{(2)}_{\rho^{m'}}, 1^{(2)}_{c''^{k-1}\rho^{km'}}]. \tag{32}$$

By the hypothesis and the induction hypothesis, for every $h' \in 1^{(2)}_{\rho^{m'}}$ and $g' \in 1^{(2)}_{c''^{k-1}\rho^{km'}}$, there are $h \in 1^{(1)}_{\rho}$ and $g \in 1^{(1)}_{2^{k-1}\rho^k}$ such that

$$f(h) \in h'_{\rho^m}, \quad \text{and} \quad f(g) \in g'_{6\rho^m}. \tag{33}$$

By (33), as part of the Solovay-Kitaev theorem (see the following claim) for $0 < \rho \ll 1$,

$$[f(h), f(g)] \in [h', g']_{O(\rho^{2m})} \subseteq [h', g']_{\rho^{2m-1}}. \tag{34}$$

Since f is a ρ^m-approximate homomorphism,

$$f([h, g]) \in [f(h), f(g)]_{5\rho^m}. \tag{35}$$

By (35) and (34), we obtain that the following holds

$$[h', g'] \in f([h, g])_{5\rho^m+\rho^{2m-1}} \subseteq f([h, g])_{6\rho^m}. \tag{36}$$

By (32) and (36), we deduce that

$$1^{(2)}_{c''^k\rho^{(k+1)m'}} \subseteq f([1^{(1)}_\rho, 1^{(1)}_{2^{k-1}\rho^k}])_{6\rho^m}. \tag{37}$$

Claim. Suppose $g, h \in \mathrm{SU}(n)$, $\|g - I\| \le r$ and $\|h - I\| \le r'$. Then $\|[g, h] - I\| \le 2rr'$.

Proof of Claim. Suppose $x, y \in \mathrm{M}_n(\mathbb{C})$ such that $\|x\| = \|y\| \leq 1$ and $g = I + rx$ and $h = I + r'y$. Then

$$\begin{aligned}\|[g,h]-I\| =& \|ghg^{-1}h^{-1}-I\| = \|(gh-hg)g^{-1}h^{-1}\| \\ =& \|(I+rx)(I+r'y)-(I+r'y)(I+rx)\| \\ =& rr'\|xy-yx\| \leq 2rr'.\end{aligned}$$

The claim follows.

By the above claim, $1^{(1)}_{2^k\rho^{k+1}} \supseteq [1^{(1)}_\rho, 1^{(1)}_{2^{k-1}\rho^k}]$. This and (37) imply that $1^{(2)}_{c''^k\rho^{(k+1)m'}} \subseteq f(1^{(1)}_{2^k\rho^{k+1}})_{6\rho^m}$, which finishes the proof. □

In order to study the image of the restriction of f to a small ball, we will be using the n-th roots of elements of a compact group. In the next lemma, we recall some basic properties of taking the n-th roots.

For $g \in 1_{1/3}$ in a compact Lie group and positive integer n, we let

$$g^{1/n} := \exp\left(\frac{1}{n}\log g\right);$$

recall from the discussion leading to (11), that exp and log are well-defined on the considered neighborhoods.

Lemma 14. *In the above setting and $n \in \mathbb{N}$, the following statements hold.*

(1) $1_{\eta/n} \subseteq 1_\eta^{1/n} \subseteq 1_{6\eta/n}$ *for every* $\eta < 1/3$.
(2) $(g^{1/n})_{\eta/n} \subseteq (g_\eta)^{1/n} \subseteq (g^{1/n})_{18\eta/n}$ *for every* $0 < \eta < 1/3$ *and* $\|g - I\| \ll 1$ *where the implied constants are universal.*

Proof. For $g \in 1_{\eta/n}$, $g^n \in 1_\eta$, and so $g \in 1_\eta^{1/n}$.

For $g \in 1_\eta^{1/n}$, we have that $g^n \in 1_\eta$, and so $\|\log g^n\| \leq 2\eta$. Hence $\|\frac{1}{n}\log g^n\| \leq 2\eta/n$, which implies

$$\|g - I\| = \|\exp\left(\frac{1}{n}\log g^n\right) - I\| \leq 3\|\frac{1}{n}\log g^n\| \leq 6\eta/n.$$

The first set of inclusions follows.

To show the second claim, we use the Baker-Campbell-Hausdorff formula. We give an extended discussion on this around (52). For now, we just mention that for x, y in a ball of radius $1/6$ in the Lie algebra $\mathfrak{g}$,

$$x\#y := \log(\exp(x)\exp(y)) \in \mathfrak{g},$$

and

$$\|x\#y - x - y\| \leq \bar{C}\|x\|\|y\|, \tag{38}$$

where $\bar{C}$ is a fixed universal constant (see 56).

Suppose $h \in (g_\eta)^{1/n}$. Then $\log h = \frac{1}{n}(x\#y)$ where $x := \log g$ and $y \in 0_{2\eta}$. Let $z := \log(hg^{-1/n})$. Then, we have

$$\frac{1}{n}(x\#y) = \log((hg^{-1/n})g^{1/n}) = z\#(x/n). \tag{39}$$

By (38) and (39), we deduce that

$$\|z\| - \frac{1}{n}(C\|x\|+1)\|y\| \leq \|\frac{1}{n}(x\#y) - \frac{x}{n} - z\| \leq \frac{\bar{C}}{n}\|z\|\|x\|. \tag{40}$$

Hence if $\|x\| \le 1/(2\bar{C})$,

$$\|z\| \le \left(\frac{1+\bar{C}\|x\|}{n-\bar{C}\|x\|}\right)\|y\| \le \frac{6}{2n-1}\eta,$$

which implies that $hg^{-1/n} \in 1_{6\eta/n}$.

On the other hand, for every $h \in 1_{\eta/n}$, we have

$$(g^{1/n}h)^n = \underbrace{g^{1/n}h\cdots g^{1/n}h}_{n \text{ times}} = (g^{1/n}hg^{-1/n})(g^{2/n}hg^{-2/n})\cdot(g^{n/n}hg^{-n/n})g \in g_\eta;$$

and so $(g^{1/n})_{\eta/n} \subseteq (g_\eta)^{1/n}$. □

The next lemma extends the result of Lemma 10 to smaller balls. Essentially, we show that $f(1_r^{(1)}) \subseteq 1_a^{(2)}$ for small values of r with the property that a/r is bounded by $\rho^{-O_{G_i}(1)}$.

Lemma 15. *In the setting of Theorem 3, suppose $F_2 = \mathbb{R}$. If $m \gg_{d_{0i}} 1$ and $C := C(G_1)$ is as in Proposition 39, see Lemma 10, then for every $4\rho^{7C+m} \le r \le \rho^{7C}/3$ the following holds*

$$f(1_r^{(1)}) \subseteq 1^{(2)}_{s_\rho(r)}$$

where $s_\rho(r) = 6(2\rho^{d_1/(8d_2)}\rho^{-7C}r + \rho^m)$.

Proof. Let $C = C(G_1)$ be as in Lemma 10; let $r_0 := \rho^{7C}$ and $a_0 := \rho^{d_1/(8d_2)}$. Then

$$f(1^{(1)}_{r_0}) \subseteq 1^{(2)}_{a_0}. \tag{41}$$

For every positive integer k and $g \in 1^{(1)}_{r_0/k}$, $g^k \in 1^{(1)}_{r_0}$. Hence using (41) and the fact that f is ρ^m-approximate homomorphism, we deduce

$$f(g)^k \in f(g^k)_{k\rho^m} \subseteq 1^{(2)}_{a_0+k\rho^m}. \tag{42}$$

Assuming that $a_0 + k\rho^m < 1/3$, by (42) and part (1) of Lemma 14, we obtain the following

$$f(g) \in (1^{(2)}_{a_0+k\rho^m})^{1/k} \subseteq 1^{(2)}_{6(\frac{a_0}{k}+\rho^m)}. \tag{43}$$

For every $4\rho^m \le \varepsilon < 1/3$, let k be an integer such that $(k+1)^{-1} < \varepsilon \le k^{-1}$. Then $a_0 + k\rho^m < 1/3$, and by (43), we have

$$f(1^{(1)}_{r_0\varepsilon}) \subseteq f(1^{(1)}_{r_0/k}) \subseteq 1^{(2)}_{6(\frac{a_0}{k}+\rho^m)} \subseteq 1^{(2)}_{6(2a_0\varepsilon+\rho^m)}. \tag{44}$$

Therefore, for every $4\rho^{7C+m} \le r < \rho^{7C}/3$, we get

$$f(1^{(1)}_r) \subseteq 1^{(2)}_{6(2a_0\rho^{-7C}r+\rho^m)}$$

which finishes the proof. □

Lemma 16. *In the setting of Theorem 3, suppose $F_1 = F_2 = \mathbb{R}$. Then there is $0 < \hat{c} := \hat{c}(G_2) \le 1$ such that for every positive integer $7C \le k \le m/3$ the following holds*

$$1^{(2)}_{\hat{c}7^{-k}\rho^{k+7Cm'}} \subseteq f\big(1^{(1)}_{\rho^{k-1}}\big)_{18k\rho^m}. \tag{45}$$

so long as $0 < \rho < 1$ is small enough depending on G_1.

Proof. We will use the following two facts. First,

$$f(g^n) \in (f(g)^n)_{n\rho^m} \tag{46}$$

if g and n are so that $f(g^i)$ is defined for all $1 \le i \le n$. Second, the following consequence of the first part of Lemma 14:

$$1^{(1)}_{\rho^{k-1}} \subseteq (1^{(1)}_{6\rho^{k-1}/n})^n \subseteq (1^{(1)}_{\rho^k})^n. \tag{47}$$

Let C be as in Lemma 15. Then, by Lemma 15, $\big(f(1^{(1)}_{\rho^{k-1}})_{18k\rho^m}\big)^{1/n}$ is defined for all $7C \leq k \leq m/3$. Thus, we will assume throughout that $7C \leq k \leq m/3$.

Suppose that $1^{(2)}_{\rho_k} \subseteq f(1^{(1)}_{\rho^{k-1}})_{18k\rho^m}$ for some $\rho_k < 1/3$. Then, for $n \geq \max\{6/\rho, 18\}$, we have

$$\begin{aligned}
(1^{(2)})_{\rho_k/n} &\subseteq (1^{(2)}_{\rho_k})^{1/n} && \text{by part 1 of Lemma 14}\\
&\subseteq \big(f(1^{(1)}_{\rho^{k-1}})_{18k\rho^m}\big)^{1/n} && \\
&\subseteq (f(1^{(1)}_{\rho^{k-1}})^{1/n})_{18\times 18k\rho^m/n} && \text{by part 2 of Lemma 14}\\
&\subseteq (f((1^{(1)}_{\rho^k})^n)^{1/n})_{18\times 18k\rho^m/n} && \text{by (47)}\\
&\subseteq (((f(1^{(1)}_{\rho^k})^n)_{n\rho^m})^{1/n})_{18\times 18k\rho^m/n} && \text{by (46)}\\
&\subseteq f(1^{(1)}_{\rho^k})_{18\rho^m+18\times 18k\rho^m/n} && \text{by part 2 of Lemma 14}\\
&\subseteq f(1^{(1)}_{\rho^k})_{18(k+1)\rho^m} && \text{since } n \geq 18.
\end{aligned}$$

Hence, if we assume $\rho < 1/3$ and put $n = \lceil 6/\rho \rceil$, then applying the above, repeatedly, we conclude that, $1^{(2)}_{\rho_{k_0}} \subseteq f(1^{(1)}_{\rho^{k_0-1}})_{18k_0\rho^m}$ for some $k_0 \in \mathbb{N}$ implies

$$1^{(2)}_{\rho_{k_0}(\rho/7)^{k-k_0}} \subseteq f(1^{(1)}_{\rho^{k-1}})_{18k\rho^m} \tag{48}$$

for all $k \geq k_0$.

Applying Lemma 13 with $7C$, we deduce that (48) holds for $k = k_0 = 7C$, $\rho \leq 2^{-7C}$, and $\rho_{k_0} := c''^{7C-1}\rho^{7Cm'}$. Therefore for all $7C \leq k \leq m/3$, we have

$$1^{(2)}_{c''^{7C-1}7^{-k+7C}\rho^{k+7Cm'-7C}} \subseteq f\big(1^{(1)}_{\rho^{k-1}}\big)_{18k\rho^m},$$

and the claim follows. □

For every positive integer k, let

$$\theta_k : 0_{\rho/2} \to \mathfrak{g}_2, \quad \theta_k(x) := \frac{\log(f(\exp(\rho^k x)))}{\rho^k}.$$

By the definition, we have

$$\exp(\rho^k\theta_k(x)) = f(\exp(\rho^k x)) \tag{49}$$

for every $x \in 0_{\rho/2}$.

Lemma 15 implies that θ_k is *approximately* Lipschitz.

Lemma 17. *In the above setting, for $x \in 0_{\rho/2}$, $\rho \ll 1$, and $7C < k < m$, we have*

$$\|\theta_k(x)\| \ll \rho^{-7C}\|x\| + \rho^{m-k}. \tag{50}$$

In particular, if $\rho \ll 1$ and $7C < k < m$, then

$$\|\theta_k(x)\| < \rho^{-7C} \tag{51}$$

for all $x \in 0_{\rho/2}$.

Proof. Since $\exp(\rho^k x) \in 1^{(1)}_{\rho^k\|x\|}$, by Lemma 15 we have

$$\exp(\rho^k\theta_k(x)) = f(\exp(\rho^k x)) \in 1^{(2)}_{b(\rho^{k-7C}\|x\|+\rho^m)}.$$

Hence $\|\rho^k\theta_k(x)\| \ll \rho^{k-7C}\|x\| + \rho^m$, as we claimed in (50).

The claim in (51) follows from (50). □

To understand further properties of θ_k, we start by recalling some of the consequences of the Baker-Campbell-Hausdorff-Dynkin and the Zassenhaus formulas.

For $x, y \in \mathfrak{g}_i$ with $\|x\|, \|y\| < 1/2$, put

$$x\#y := \log(\exp(x)\exp(y)).$$

By the Baker-Campbell-Hausdorff-Dynkin formula, we have

$$x\#y = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \sum_{m_i,n_i\geq 0, m_i+n_i>0} \frac{1}{(\sum_{i=1}^{k}(m_i+n_i))\prod_{i=1}^{k}(m_i!n_i!)} Z_{\mathbf{m},\mathbf{n}}(x,y), \tag{52}$$

where $\mathbf{m} := (m_1, \ldots, m_k)$, $\mathbf{n} := (n_1, \ldots, n_k)$, and

$$Z_{\mathbf{m},\mathbf{n}}(x,y) := [\underbrace{x,\ldots,x}_{m_1}, \underbrace{y,\ldots,y}_{n_1}, \ldots, \underbrace{x,\ldots,x}_{m_k}, \underbrace{y,\ldots,y}_{n_k}]$$

is a long commutator. Let us observe that for every $\mathbf{m}$, $\mathbf{n}$, x and y we have

$$\|Z_{\mathbf{m},\mathbf{n}}(x,y)\| \leq 2^{\|\mathbf{m}\|_1+\|\mathbf{n}\|_1-1}\|x\|^{\|\mathbf{m}\|_1}\|y\|^{\|\mathbf{n}\|_1}. \tag{53}$$

Suppose $x, y \in \mathfrak{g}_i$ and $\|v\| \leq \eta\|x\|$ for some $0 < \eta < 1$. Using the multi-linearity of long commutators and (53), we deduce that

$$\|Z_{\mathbf{m},\mathbf{n}}(x+v,y) - Z_{\mathbf{m},\mathbf{n}}(x,y)\| \leq 2^{\|\mathbf{m}\|_1+\|\mathbf{n}\|_1+1}\eta\|x\|^{\|\mathbf{m}\|_1}\|y\|^{\|\mathbf{n}\|_1}. \tag{54}$$

We show this by induction on $\|\mathbf{m}\|_1 + \|\mathbf{n}\|_1$. Here is the induction step:

$$\begin{aligned}\|Z_{\mathbf{m},\mathbf{n}}(x+v,y) - Z_{\mathbf{m},\mathbf{n}}(x,y)\| &= \|\operatorname{ad}(x+v)(Z_{\mathbf{m}-\mathbf{e}_1,\mathbf{n}}(x+v,y)) - \operatorname{ad}(x)(Z_{\mathbf{m}-\mathbf{e}_1,\mathbf{n}}(x,y))\| \\ &\leq 2\eta\|x\|\|Z_{\mathbf{m}-\mathbf{e}_1,\mathbf{n}}(x+v,y)\| + 2\|x\|\|Z_{\mathbf{m}-\mathbf{e}_1,\mathbf{n}}(x+v,y) - Z_{\mathbf{m}-\mathbf{e}_1,\mathbf{n}}(x,y)\| \\ &\leq 2^{\|\mathbf{m}\|_1+\|\mathbf{n}\|_1}\eta\|x\|^{\|\mathbf{m}\|_1}\|y\|^{\|\mathbf{n}\|_1} + 2^{\|\mathbf{m}\|_1+\|\mathbf{n}\|_1}\eta\|x\|^{\|\mathbf{m}\|_1}\|y\|^{\|\mathbf{n}\|_1} \\ &\leq 2^{\|\mathbf{m}\|_1+\|\mathbf{n}\|_1+1}\eta\|x\|^{\|\mathbf{m}\|_1}\|y\|^{\|\mathbf{n}\|_1}.\end{aligned}$$

By (54) and (52), we obtain the following perturbation estimate: for every $x, y, v \in \mathfrak{g}_i$ with $\|v\| \leq \eta\|x\|$, $\|x\|, \|y\| \ll 1$, and $0 < \eta < 1$,

$$\begin{aligned}\|(x+v)\#y - x\#y\| &\leq \Big(\sum_{k=1}^{\infty}\frac{1}{k}\sum_{m_i,n_i\geq 0, m_i+n_i>0}\frac{2^{\|\mathbf{m}\|_1+\|\mathbf{n}\|_1+1}\|x\|^{\|\mathbf{m}\|_1}\|y\|^{\|\mathbf{n}\|_1}.}{(\|\mathbf{m}\|_1+\|\mathbf{n}\|_1)\prod_{i=1}^{k}(m_i!n_i!)}\Big)\eta \\ &\leq 2\Big(\sum_{k=1}^{\infty}\frac{(e^{2\|x\|+2\|y\|}-1)^k}{k}\Big)\eta = 2\log((2-e^{2\|x\|+2\|y\|})^{-1})\eta. \end{aligned} \tag{55}$$

Next we note that by (52) and an argument similar to (55), for every $0 < \eta < 1$ and $\|x\|, \|y\| < 1$ the following holds

$$(\eta x)\#(\eta y) = \eta(x+y) + \eta^2 z, \quad \text{for some } z := z(x,y) \in \mathfrak{g}_i \text{ with } \|z\| \ll \|x\|\|y\|. \tag{56}$$

By (55) and (56), for $\|x\| \leq \|y\| < 1$ and $0 < \eta \ll 1$, we obtain the following upper bound,

$$\begin{aligned}\|((\eta x + \eta y)\#(-\eta x))\#(-\eta y)\| &= \|(\eta y + \eta^2 z(\eta(x+y), -\eta x))\#(-\eta y)\| \\ &= \|(\eta y + \eta^2 z)\#(-\eta y) - (\eta y)\#(-\eta y)\| \\ &\ll \log(2 - e^{C\eta\|y\|})^{-1}\eta \ll \eta^2\|y\|,\end{aligned} \tag{57}$$

where C is a universal constant and the last inequality holds as $\lim_{s\to 0^+}\ln(2-e^s)^{-1}/s = 1$. By (57), we deduce that for $\|x\|, \|y\| < 1$ and $0 < \eta \ll 1$, there is $z' := z'(x,y) \in \mathfrak{g}_i$ such that $\|z'\| \ll \max\{\|x\|, \|y\|\}$ and

$$\exp(\eta(x+y)) = \exp(\eta x)\exp(\eta y)\exp(\eta^2 z'). \tag{58}$$

Lemma 18. *In the above setting for $x, y \in 0_{\rho/4}$, $\rho \ll 1$, and $C \ll k < m$, we have*

$$\|\theta_k(x+y) - (\theta_k(x) + \theta_k(y))\| \ll \rho^{k-14C}.$$

Proof. We start with the following computation of $\theta_k(x+y)$. By the definition of θ_k, (49), we have

$$\begin{aligned}
\exp(\rho^k\theta_k(x+y)) =& f(\exp(\rho^k(x+y))) = f(\exp(\rho^k x)\exp(\rho^k y)\exp(\rho^{2k}z')) && \text{(by (58))}\\
=& f(\exp(\rho^k x))f(\exp(\rho^k y))f(\exp(\rho^{2k}z'))\exp(w) && \text{(for some } \|w\| \ll \rho^m)\\
=& \exp(\rho^k\theta_k(x))\exp(\rho^k\theta_k(y))\exp(\rho^{2k}\theta_{2k}(z'))\exp(w)\\
=& \exp\Big((\rho^k\theta_k(x))\#(\rho^k\theta_k(y))\#(\rho^{2k}\theta_{2k}(z'))\#w\Big) \qquad (59)
\end{aligned}$$

By (59), we obtain the following

$$\rho^k\theta_k(x+y) = (\rho^k\theta_k(x))\#(\rho^k\theta_k(y))\#(\rho^{2k}\theta_{2k}(z'))\#w. \qquad (60)$$

By (51) and (56), we deduce the following

$$\begin{aligned}
(\rho^k\theta_k(x))\#(\rho^k\theta_k(y)) =& (\rho^{k-7C}(\rho^{7C}\theta_k(x)))\#(\rho^{k-7C}(\rho^{7C}\theta_k(y)))\\
=& \rho^k(\theta_k(x)+\theta_k(y)) + \rho^{2k-14C}z, \qquad (61)
\end{aligned}$$

for some $z \in \mathfrak{g}_2$ with $\|z\| \ll \rho^{14C}\|\theta_k(x)\|\|\theta_k(y)\|$. By (61) and (56), we obtain that the following holds

$$\begin{aligned}
(\rho^k\theta_k(x))\#(\rho^k\theta_k(y))\#(\rho^{2k}\theta_{2k}(z')) =& (\rho^k(\theta_k(x)+\theta_k(y)) + \rho^{2k-14C}z)\#(\rho^{2k}\theta_{2k}(z'))\\
=& (\rho^k(\theta_k(x)+\theta_k(y)) + \rho^{2k-14C}z) + (\rho^{2k}\theta_{2k}(z')) + \rho^{2k-14C}\overline{z}\\
=& \rho^k(\theta_k(x)+\theta_k(y)) + \rho^{2k-14C}(z + \overline{z} + \rho^{14C}\theta_{2k}(z')). \qquad (62)
\end{aligned}$$

By (62), (56), and (60), we deduce the following

$$\|\theta_k(x+y) - (\theta_k(x)+\theta_k(y))\| \ll \rho^{k-14C},$$

and the claim follows. □

Using Lemma 18 and Lemma 17, we can show that θ_k almost preserves scaler multiplication.

Lemma 19. *In the above setting, for $x \in 0_{\rho/2}$, $\rho \ll 1$, $-1 \le t \le 1$, and $C \ll k < m/2$, we have*

$$\|\theta_k(tx) - t\theta_k(x)\| \ll \rho^{\frac{k}{2}-14C}.$$

Proof. There is a rational number r/s such that $|t - \frac{r}{s}| \le \rho^{k/2}$ and $|r|, |s| \le \rho^{-k/2}$. By Lemma 18, we have $\|s\theta_k(\frac{x}{s}) - \theta_k(x)\| \ll s\rho^{k-14C}$, which implies

$$\left\|\theta_k\left(\frac{x}{s}\right) - \frac{1}{s}\theta_k(x)\right\| \ll \rho^{k-14C}. \qquad (63)$$

Similarly, we have

$$\left\|\theta_k\left(\frac{r}{s}x\right) - r\theta_k\left(\frac{x}{s}\right)\right\| \ll r\rho^{k-14C}. \qquad (64)$$

Combining (63) and (64), we conclude that

$$\left\|\theta_k\left(\frac{r}{s}x\right) - \frac{r}{s}\theta_k(x)\right\| \ll r\rho^{k-14C} \ll \rho^{\frac{k}{2}-14C}. \qquad (65)$$

Now Lemma 18 and Lemma 17, imply that

$$\begin{aligned}\Big\|\theta_k(tx)-\theta_k\Big(\frac{r}{s}x\Big)\Big\| \ll&\Big\|\theta_k\Big(tx-\frac{r}{s}x\Big)\Big\|+\rho^{k-14C}\\ \ll&\rho^{-7C}\Big\|tx-\frac{r}{s}x\Big\|+\rho^{k-14C}\\ \ll&\rho^{\frac{k}{2}-14C}. \end{aligned} \tag{66}$$

Note also that $\|(t-\frac{r}{s})\theta_k(x)\| \ll \rho^{-7C}|t-\frac{r}{s}| \le \rho^{\frac{k}{2}-7C}$. Therefore, by (65) and (66), we have

$$\|\theta_k(tx)-t\theta_k(x)\| \ll \rho^{\frac{k}{2}-14C},$$

as it was claimed. □

Corollary 20. *In the above setting, suppose $\rho \ll 1$ and $C \ll k < m/2$. Then there is a linear function $\widetilde{\theta}_k : \mathfrak{g}_1 \to \mathfrak{g}_2$ such that for every $x \in 0_{\rho/2}$ we have*

$$\|\theta_k(x)-\widetilde{\theta}_k(x)\| \ll \rho^{\frac{k}{2}-14C}.$$

Proof. Suppose $\{e_1,\ldots,e_{d_1}\}$ is an orthonormal basis of $\mathfrak{g}_1$. Let $\widetilde{\theta}_k : \mathfrak{g}_1 \to \mathfrak{g}_2$ be the linear map defined by $\widetilde{\theta}_k(e_i) := (\rho/2)^{-1}\theta_k((\rho/2)e_i)$ for every i.

For every $x := \sum_{i=1}^{d_1} t_ie_i \in 0_{\rho/2}$, we have

$$\begin{aligned}\|\theta_k(x)-\widetilde{\theta}_k(x)\| =&\|\theta_k(\sum_{i=1}^{d_1} t_ie_i)-\widetilde{\theta}_k(\sum_{i=1}^{d_1} t_ie_i)\|\\ \ll&\rho^{k-14C}+\Big\|\sum_{i=1}^{d_1}\theta_k(t_ie_i)-\frac{2t_i}{\rho}\theta_k((\rho/2)e_i)\Big\| && \text{(by Lemma 18)}\\ \ll&\rho^{k-14C}+\sum_{i=1}^{d_1}\Big\|\theta_k\Big(\frac{2t_i}{\rho}(\rho/2)e_i\Big)-\frac{2t_i}{\rho}\theta_k((\rho/2)e_i)\Big\|\\ \ll&\rho^{\frac{k}{2}-14C} && \text{(by Lemma 19)},\end{aligned}$$

as we claimed in the corollary. □

Our next task is to show that θ_{2k} almost preserves Lie algebra commutators. This will be done in two steps: first we show that $\theta_{2k}([x,y])$ is close to $[\theta_k(x),\theta_k(y)]$, Lemma 21, then we show that θ_{2k} and θ_k are close to each other, Lemma 22.

Let us begin with the following consequence of the Baker-Campbell-Hausdorff-Dynkin formula, see (52). For every $0<\eta\ll 1$ and $\|x\|,\|y\|<1$, we have

$$(\eta x)\#(\eta y)=\eta(x+y)+\frac{\eta^2}{2}[x,y]+\frac{\eta^3}{12}\Big([x,[x,y]]-[y,[x,y]]\Big)+\eta^4z_4, \tag{67}$$

for some $z_4 := z_4(x,y) \in \mathfrak{g}_i$ with $\|z_4\| \ll \max\{\|x\|^3\|y\|,\|x\|^2\|y\|^2,\|x\|\|y\|^3\}$. By (67), we obtain the following

$$\begin{aligned}(\eta x)\#(\eta y)\#(-\eta x)\#(-\eta y) =&\Big(\eta(x+y)+\frac{\eta^2}{2}[x,y]+\frac{\eta^3}{12}\Big([x,[x,y]]-[y,[x,y]]\Big)+\eta^4z_4\Big)\\ &\#\Big(-\eta(x+y)+\frac{\eta^2}{2}[x,y]-\frac{\eta^3}{12}\Big([x,[x,y]]-[y,[x,y]]\Big)+\eta^4z_4'\Big)\\ =&\eta^2[x,y]+\eta^3z_3',\end{aligned} \tag{68}$$

for some $z_3' := z_3'(x,y) \in \mathfrak{g}_i$ with $\|z_3'\| \ll \max\{\|x\|^2\|y\|,\|x\|\|y\|^2\}$.

Lemma 21. *In the above setting for $x, y \in 0_{\rho/16}$, $\rho \ll 1$, and $C \ll k < m$, we have*

$$\|\theta_{2k}([x,y]) - [\theta_k(x), \theta_k(y)]\| \ll \rho^{2k-14C}.$$

The implied constants are absolute.

Proof. Again using the definition of θ_k, (49), we have

$$\begin{aligned}
\exp(\rho^{2k}\theta_{2k}([x,y])) =& f(\exp(\rho^{2k}[x,y])) && \text{(by (49))}\\
=& f\Big(\exp\Big(\log([\exp(\rho^k x), \exp(\rho^k y)]) - \rho^{3k} z_3'\Big)\Big) && \text{(by (68))}\\
=& f([\exp(\rho^k x), \exp(\rho^k y)]u') && \text{(for some } u' \in 1^{(1)}_{O(\rho^{3k})})\\
=& [f(\exp(\rho^k x)), f(\exp(\rho^k y))]f(u')w' && \text{(for some } w' \in 1^{(2)}_{O(\rho^{m})})\\
=& [\exp(\rho^k\theta_k(x)), \exp(\rho^k\theta_k(y))]f(u')w' && \text{(by (49))}\\
=& \exp(\rho^{2k}[\theta_k(x), \theta_k(y)] + \rho^{3k-21C} z_3'')f(u')w' && \text{(by (51) and (68))} \qquad (69)
\end{aligned}$$

Moreover, by Lemma 15, we have

$$f(u')w' \in 1^{(2)}_{O(\rho^{3k-7C}+\rho^m)}. \qquad (70)$$

By (69) and (70), we obtain the following

$$\theta_{2k}([x,y]) = [\theta_k(x), \theta_k(y)] + \rho^{k-21C} z'', \qquad (71)$$

for some $z'' \in \mathfrak{g}_2$ with $\|z''\| \ll 1$. The claim follows. □

We now show that $\theta_k(x)$ and $\theta_{2k}(x)$ are close to each other.

Lemma 22. *In the above setting for $x \in 0_{\rho/2}$, $\rho \ll 1$, and $C \ll k < m/3$, we have*

$$\|\theta_{2k}(x) - \theta_k(x)\| \ll \rho^{k-7C}.$$

Proof. Let $\ell := \lfloor \rho^{-k} \rfloor$. By (49), we $\exp(\rho^{2k}\theta_{2k}(x)) = f(\exp(\rho^{2k}x))$, and so

$$\begin{aligned}
\exp(\ell\rho^{2k}\theta_{2k}(x)) =& \big(f(\exp(\rho^{2k}x))\big)^{\ell}\\
=& f(\exp(\ell\rho^{2k}x))u && \text{(for some } u \in 1^{(2)}_{\ell\rho^m}). \qquad (72)
\end{aligned}$$

Now note that $\|\ell\rho^{2k}x - \rho^k x\| \le \rho^{2k}$, therefore,

$$\exp(\ell\rho^{2k}x) = \exp(\rho^k x)u' \qquad (73)$$

for some $u' \in 1^{(1)}_{O(\rho^{2k})}$. By (72) and (73), we deduce the following

$$\begin{aligned}
\exp(\ell\rho^{2k}\theta_{2k}(x)) =& f(\exp(\rho^k x))f(u')w'' && \text{(for some } w'' \in 1^{(2)}_{(\ell+1)\rho^m}).\\
=& \exp(\rho^k\theta_k(x))\bar{w} && \text{(for some } \bar{w} \in 1^{(2)}_{O(\rho^{2k-7C}+\rho^{m-k})}), \qquad (74)
\end{aligned}$$

where in the last equality we used Lemma 15 and the definition of θ_k in (49).

In view of (74), we have

$$\|\ell\rho^{2k}\theta_{2k}(x) - \rho^k\theta_k(x)\| \ll \rho^{2k-7C} + \rho^{m-k}. \qquad (75)$$

Recall now that $\ell := \lfloor \rho^{-k} \rfloor$ and $\|\theta_k(x)\| < \rho^{-7C}$, see (51). Therefore,

$$\|\ell\rho^{2k}\theta_{2k}(x) - \rho^k\theta_{2k}(x)\| \le \rho^{2k-7C}.$$

This and (75) imply that $\|\rho^k\theta_{2k}(x) - \rho^k\theta_k(x)\| \ll \rho^{2k-7C} + \rho^{m-k}$. In consequence, we deduce

$$\|\theta_{2k}(x) - \theta_k(x)\| \ll \rho^{k-7C} + \rho^{m-2k} \ll \rho^{k-7C}$$

where we used $k < m/3$.

The proof is complete. □

Corollary 23. *In the above setting, for $x, y \in 0_{\rho/16}$, $\rho \ll 1$, and $C \ll k < m/3$, we have*

$$\|\theta_{2k}([x,y]) - [\theta_{2k}(x), \theta_{2k}(y)]\| \ll \rho^{k-14C}.$$

Proof. In view of Lemma 22, $\theta_{2k}(\bullet) = \theta_k(\bullet) + z_\bullet$ where $\|z_\bullet\| \ll \rho^{k-7C}$ for $\bullet = x, y$.
The claim thus follows using Lemma 21 and (51). □

Corollary 24. *In the above setting, suppose $\rho \ll 1$, $C \ll k < m/3$, $x, y \in 0_{\rho/16}$ and $\widetilde{\theta}_{2k} : \mathfrak{g}_1 \to \mathfrak{g}_2$ is the linear map as in Corollary 20. Then*

$$\|\widetilde{\theta}_{2k}(x) - \theta_{2k}(x)\| \ll \rho^{k-14C} \quad \textit{and} \quad \|\widetilde{\theta}_{2k}([x,y]) - [\widetilde{\theta}_{2k}(x), \widetilde{\theta}_{2k}(y)]\| \ll \rho^{k-21C}.$$

Proof. The first claim is proved in Corollary 20.
The second claim follows from the first claim, Corollary 23, and (51). □

Similar to the p-adic case, we consider the set of Lie algebra homomorphisms from $\mathfrak{g}_1$ to $\mathfrak{g}_2$ which can be viewed as an affine variety as follows: Let $\{e_1^{(i)}, \ldots, e_{d_{0i}}^{(i)}\}$ be an orthonormal basis of $\mathfrak{g}_i$, and suppose $c_{jks}^{(i)} \in \mathbb{R}$ are the corresponding structural constants. That is,

$$[e_j^{(i)}, e_k^{(i)}] = \sum_{s=1}^{d_{0i}} c_{jks}^{(i)} e_s^{(i)}.$$

Then a linear map $T : \mathfrak{g}_1 \to \mathfrak{g}_2, T(e_j^{(1)}) = \sum_{s=1}^{d_2} x_{js} e_s^{(2)}$ is a Lie ring homomorphism if and only if $T([e_j^{(1)}, e_k^{(1)}]) = [T(e_j^{(1)}), T(e_k^{(1)})]$ for every $1 \le j, k \le d_1$. This is equivalent to having the following equations on $\mathrm{Mat}_{d_1 \times d_2}(\mathbb{C})$:

$$f_{jkr}(\mathbf{x}) := \sum_{1 \le i_1, i_2 \le d_2} c_{i_1 i_2 r}^{(2)} x_{j i_1} x_{k i_2} - \sum_{i=1}^{d_2} c_{jki}^{(1)} x_{ri} = 0 \tag{76}$$

for every integers $1 \le j, k \le d_1$ and $1 \le r \le d_2$.

Let V be the real affine variety given by equations in (76). Note that $V(\mathbb{R})$ is non-empty as it contains the zero vector. Suppose

$$\widetilde{\theta}_{2k}(e_j^{(1)}) = \sum_{s=1}^{d_2} a_{js} e_s^{(2)}.$$

Put $\mathbf{a} = (a_{js})$. In view of Corollary 24, for every $1 \le j \le d_1$, we have

$$\left\|\widetilde{\theta}_{2k}\left(\frac{\rho}{16} e_j^{(1)}\right)\right\| \ll \left\|\theta_{2k}\left(\frac{\rho}{16} e_j^{(1)}\right)\right\| + \rho^{k-14C} \ll \rho^{-7C},$$

where we used (51) in the last inequality. Therefore,

$$\|\mathbf{a}\| \ll_{d_i} \rho^{-7C}. \tag{77}$$

Moreover, by Corollary 24, we have

$$|f_{jkr}(\mathbf{a})| \ll_{d_i} \rho^{k-21C}. \tag{78}$$

Lemma 25. *In the above setting, for $m' \ll_{d_i} k < m/3$ and $0 < \rho \ll_{G_1, G_2} 1$, there is a Lie algebra isomorphism $\widehat{\theta} : \mathfrak{g}_1 \to \mathfrak{g}_2$ with the following properties:*

(1) $\|\widehat{\theta} - \widetilde{\theta}_{2k}\|_{\mathrm{op}} \le \rho^{k/2}$.
(2) *For every $x \in \mathfrak{g}_1$ with $\|x\| < \rho/2$, $\|\widehat{\theta}(x) - \theta_{2k}(x)\| \le \rho^{k/2}$.*

Proof. Let V be the variety which was defined above. Note that $V(\mathbb{R}) \neq \emptyset$, indeed $0 \in V(\mathbb{R})$. By [26, Theorem 7], which is a quantitative version of Łojasiewicz inequality, there are positive numbers $\overline{C} := \overline{C}(V)$ and $D := D(d_1, d_2)$ such that for every $\mathbf{x} \in \mathbb{R}^{d_1 d_2}$ with $\mathrm{dist}(\mathbf{x}, V(\mathbb{R})) \leq 1$ we have

$$\mathrm{dist}(\mathbf{x}, V(\mathbb{R})) \leq \overline{C} \max_{j,k,r}\{|f_{j,k,r}(\mathbf{x})|\} \cdot (1 + \|\mathbf{x}\|)^D. \tag{79}$$

Using (78), (77), and (79), there is $\widehat{\mathbf{a}} \in V(\mathbb{R})$ such that

$$\|\widehat{\mathbf{a}} - \mathbf{a}\| \leq \overline{C} \rho^{k-21C} \rho^{-8CD} \leq \rho^{3k/4} \tag{80}$$

so long as $\max\{1, 21C + 8CD\} \leq k/8$ and $\overline{C}\rho \leq 1$.

Since $\widehat{\mathbf{a}} \in V(\mathbb{R})$, it induces a Lie algebra homomorphism $\widehat{\theta} : \mathfrak{g}_1(\mathbb{R}) \to \mathfrak{g}_2(\mathbb{R})$; moreover, (80) implies the following upper bound estimate:

$$\|\widehat{\theta} - \widetilde{\theta}_{2k}\|_{\mathrm{op}} \leq \rho^{k/2}. \tag{81}$$

In view of (81), for every $x \in 0_{\rho/2}$, we have $\|\widehat{\theta}(x) - \widetilde{\theta}_{2k}(x)\| \leq \rho^{k/2}\|x\| \leq \rho^{1+\frac{k}{2}}$. Hence, using Corollary 24, we deduce the following

$$\|\widehat{\theta}(x) - \theta_{2k}(x)\| \ll \rho^{k-14C} + \rho^{1+\frac{k}{2}} \leq \rho^{k/2}, \tag{82}$$

so long as $k \geq 28C + 1$ and ρ is small enough.

We now combine the facts that image of θ_{2k} is *large*, see Lemma 16, and that $\widehat{\theta}$ is linear with (82) to show that $\widehat{\theta}$ is surjective. More precisely, we will show that for every $0 < \varepsilon_0 \leq 0.01$ and $\rho^{\varepsilon_0} \ll 1$, we have

$$(\widehat{\theta}(\mathfrak{g}_1))_{\rho^{k/3}} \supseteq 0_{\rho^{2\varepsilon_0 k + 7Cm' + 6}}.$$

To that end, let us first recall from (11) that for every $g \in 1^{(i)}_{1/3}$ and every $x \in \mathfrak{g}_i$ with $\|x\| \leq 1/3$ we have

$$b'^{-1}\|g - 1^{(i)}\| \leq \|\log g\| \leq b'\|g - 1^{(i)}\|, \quad \text{and} \quad b'^{-1}\|x\| \leq \|\exp x - 1^{(i)}\| \leq b'\|x\|, \tag{83}$$

where $b' = 3$. Increasing b', if necessary, we further assume that $\log(g_r) \subseteq (\log g)_{b'r}$ for every $g \in 1^{(i)}_{1/6}$ and $r < 1/3$. Recall also the parameter $0 < \hat{c} \leq 1$ from Lemma 16. Fix some $0 < \varepsilon_0 \leq 0.01$. Choose ρ small enough so that

$$\rho^{\varepsilon_0} \leq \min\{0.1, b'^{-1}, \hat{c}\}. \tag{84}$$

Let $\ell = 2k + 3$; then $b'\rho^{\ell-1} \leq \frac{1}{2}\rho^{2k+1}$. Thus, Lemma 16 implies

$$1^{(2)}_{\hat{c}7^{-\ell}\rho^{\ell+7Cm'}} \subseteq f\big(\exp(0_{b'\rho^{\ell-1}})\big)_{18\ell\rho^m}. \tag{85}$$

Combining (83) and (85) implies that

$$0_{b'^{-1}\hat{c}7^{-\ell}\rho^{\ell+7Cm'}} \subseteq \log\Big(f\big(\exp(0_{b'\rho^{\ell-1}})\big)_{18\ell\rho^m}\Big) \subseteq \Big(\log f\big(\exp(0_{\frac{1}{2}\rho^{2k+1}})\big)\Big)_{18b'\ell\rho^m}. \tag{86}$$

We also recall from (49) that $\exp(\rho^{2k}\theta_{2k}(x)) = f(\exp(\rho^{2k}x))$ for every $x \in 0_{\rho/2}$. Altogether, we conclude that

$$0_{b'^{-1}\hat{c}7^{-\ell}\rho^{\ell+7Cm'}\rho^{-2k}} \subseteq (\mathrm{Im}(\theta_{2k}))_{18b'\ell\rho^{m-2k}}. \tag{87}$$

We now use (87) to complete the proof of (3). Recall from (84) that $\rho^{\varepsilon_0} \leq \min\{0.1, b'^{-1}, \hat{c}\}$, and also recall that $\ell = 2k + 3$. Hence

$$b'^{-1}\hat{c}7^{-\ell}\rho^{\ell+7Cm'}\rho^{-2k} = b'^{-1}\hat{c}7^{-\ell}\rho^{7Cm'+3} > \rho^{2\varepsilon_0 k + 7Cm' + 6}, \tag{88}$$

where we also used $3\varepsilon_0 < 1$.

Combining (82), (87), and (88), we conclude that

$$(\widehat{\theta}(\mathfrak{g}_1))_{\rho^{k/3}} \supseteq 0_{\rho^{2\varepsilon_0 k+7Cm'+6}}. \tag{89}$$

Since $\varepsilon_0 \le 0.01$, (89) implies that $\dim \widehat{\theta}(\mathfrak{g}_1) \ge \dim \mathfrak{g}_2$ so long as $k \gg m'$. This establishes part (3) and also shows that $\widehat{\theta}$ is surjective.

Furthermore, since $\mathfrak{g}_1$ is a simple Lie algebra and $\widehat{\theta}$ is not the zero morphism, $\widehat{\theta}$ is injective. Altogether, we conclude that $\widehat{\theta}$ is an isomorphism and the proof is complete. □

Proof of Theorem 3. In view of Lemma 7, Theorem 8, and Corollary 12, we may assume $F_1 = F_2 = \mathbb{R}$.

Let $\widehat{\theta}$ be as in Lemma 25. By [23, Theorem 10], there is a group homomorphism $\Psi : \widetilde{G}_1 \to G_2$ where $\widetilde{G}_1$ is the simply-connected cover of G_1 such that the following is a commuting diagram

$$\begin{array}{ccc} \widetilde{G}_1 & \xrightarrow{\Psi} & G_2 \\ \big\downarrow \log & & \big\downarrow \log \\ \mathfrak{g}_1 & \xrightarrow{\widehat{\theta}} & \mathfrak{g}_2. \end{array} \tag{90}$$

Let $\iota : \widetilde{G}_1 \to G_1$ be the covering map. Since the kernel of ι is a finite central subgroup, ι induces a homeomorphism from $\widetilde{1}^{(1)}_{O_{G_1}(1)}$ to $1^{(1)}_{O_{G_1}(1)}$. Hence we will view Ψ as a function on $1^{(1)}_{O_{G_1}(1)}$ as well.

Note that by (90), for every $x \in \mathfrak{g}_1$ and $g \in \widetilde{G}_1$, we have

$$\widehat{\theta}(\mathrm{Ad}(g)(x)) = \mathrm{Ad}(\Psi(g))(\widehat{\theta}(x)). \tag{91}$$

We will show that the theorem holds with this Ψ. In view of the definition of θ_{2k}, see (49), for every $x \in 0_{\rho/6}$ and $g \in 1^{(1)}_{\rho/6}$, we have

$$\begin{aligned} \exp\big(\rho^{2k}\,\mathrm{Ad}(f(g))(\theta_{2k}(x))\big) &= f(g)\exp(\rho^{2k}\theta_{2k}(x))f(g)^{-1} \\ &= f(g)f(\exp(\rho^{2k}x))f(g)^{-1} \\ &= f\big(g\exp(\rho^{2k}x)g^{-1}\big)u \qquad\qquad (\text{where } u \in 1^{(2)}_{3\rho^m}) \\ &= f\big(\exp(\rho^{2k}\,\mathrm{Ad}(g)(x))\big)u \\ &= \exp\big(\rho^{2k}\theta_{2k}(\mathrm{Ad}(g)(x))\big)u. \end{aligned} \tag{92}$$

By (92), we deduce that for every $x \in 0_{\rho/6}$ and $g \in 1^{(1)}_{\rho/6}$ the following holds

$$\|\,\mathrm{Ad}(f(g))(\theta_{2k}(x)) - \theta_{2k}(\mathrm{Ad}(g)(x))\| \ll \rho^{m-2k} \le \rho^{k/3}. \tag{93}$$

Moreover, by part (2) of Lemma 25, we have

$$\begin{aligned} &\|\,\mathrm{Ad}(f(g))(\theta_{2k}(x)) - \mathrm{Ad}(f(g))(\widehat{\theta}(x))\| \le \rho^{k/2} \quad \text{and} \\ &\|\theta_{2k}(\mathrm{Ad}(g)(x)) - \widehat{\theta}(\mathrm{Ad}(g)(x))\| \le \rho^{k/2}. \end{aligned} \tag{94}$$

Altogether, (91), (94), and (93), imply

$$\begin{aligned} \|\,\mathrm{Ad}(\Psi(g))(\widehat{\theta}(x)) - \mathrm{Ad}(f(g))(\widehat{\theta}(x))\| &\le \|\widehat{\theta}(\mathrm{Ad}(g)(x)) - \mathrm{Ad}(f(g))(\theta_{2k}(x))\| + \rho^{k/3} \\ &\le \|\theta_{2k}(\mathrm{Ad}(g)(x)) - \mathrm{Ad}(f(g))(\theta_{2k}(x))\| + 2\rho^{k/3} \\ &\le 3\rho^{k/3}. \end{aligned} \tag{95}$$

Now using (95), we deduce that $\|\operatorname{Ad}(\Psi(g)) - \operatorname{Ad}(f(g))\|_{\mathrm{op}} \ll \rho^{k/4}$ for every $g \in 1^{(1)}_{\rho/6}$. Finally, using the fact that Ad induces a homeomorphism on $O_{G_2}(1)$-neighborhood of $1^{(2)}$, we get

$$\|\Psi(g) - f(g)\|_{\mathrm{op}} \ll \rho^{k/4}.$$

This establishes the theorem for $F_1 = F_2 = \mathbb{R}$, and completes the proof. □

5. Discretization and couplings

The objective of this section is to show that, under mild conditions on the groups G_1 and G_2, one may reduce the question of spectral independence of G_1 and G_2 to the case of measures on $G_1 \times G_2$ whose marginals are Haar measures m_1 and m_2.

We begin with the following definition.

Definition 6. Let (G, d) be a compact metric group, and let $0 < \delta < 1$. We say (G, d) is δ-discretizable if the there exists a partition $\{X_i\}$ of G satisfying the following two properties:

(1) X_i is a Borel set for all i, and $|X_i| = |X_j|$ for all i and j.
(2) $\operatorname{diam}(X_i) \leq \delta$ and X_i contains a ball of radius δ^2 for all i.

We refer to a partition $\{X_i\}$ satisfying (1) and (2) above as a δ-discretization of (G, d).

Note that in this section X_i's denote a partition for G unlike in the rest of the paper where generally X_1, X_2 and X denote random variables.

As we have done so throughout the paper, we often drop d from the notation and simply write G is δ-discretizable.

An important class of examples is provided by the following proposition.

Proposition 26. *Suppose G is a compact analytic (real or p-adic) Lie group, equipped with a standard bi-invariant metric, see §3.6. Then G is δ-discretizable for all $0 < \delta \leq \delta_0$ where δ_0 is 1 in the p-adic case and depends only on the dimension of G in the real case.*

Proof. In the real case, the claim follows from [10, Theorem 2]. Suppose G is a compact p-adic analytic group, recall from §3.6 that 1_δ is a subgroup for all $\delta > 0$. Let X_i's be the cosets of the subgroup 1_δ. Then X_i's form a partition of G that satisfy the desired conditions. □

Let G be a δ-discretizable group, and let $\{X_i\}$ be a δ-discretization of G. Then $|X_i| > 0$ for all i. If we further assume that $\frac{1}{C_1}\eta^{d_0} \leq |1_\eta| \leq C_1 \eta^{d_0}$ for $\eta = \delta, \delta^2$, then

$$C_1^{-1}\delta^{2d_0} \leq |X_i| \leq C_1 \delta^{d_0}, \qquad \text{for all } i. \tag{96}$$

The following is the main result of this section.

Theorem 27. *Let G_1 and G_2 be two compact groups. Suppose there are constants C_0, C_1, L, d_{01}, d_{02}, and $\rho \leq \frac{1}{100C_0(5C_1)^L}$ so that the following properties are satisfied.*

- *G_i is L-locally random with coefficient C_0 for $i = 1, 2$, see (6).*
- *For all $\eta = \rho^j$, $j \in \mathbb{N}$, the group G_i satisfies*

$$\frac{1}{C_1}\eta^{d_{0i}} \leq |1_\eta| \leq C_1 \eta^{d_{0i}}, \qquad \text{for } i = 1, 2. \tag{97}$$

- *For $i = 1, 2$, G_i is δ-discretizable for all $\delta = \rho^j$ with sufficiently large $j \in \mathbb{N}$.*

Let μ be a symmetric Borel probability measure on $G_1 \times G_2$ satisfying

$$\max\{\lambda(\pi_1\mu; G_1), \lambda(\pi_2\mu; G_1)\} =: \lambda < 1 \tag{98}$$

where π_i denotes the projection onto the i-th factor for $i = 1, 2$.

Then, there exists a symmetric coupling ν^ρ of m_1 and m_2 so that the following holds. Let $C > 0$ and $f \in L^2(G_1 \times G_2, m_1 \times m_2)$ satisfy that $\|P_\rho * f - f\|_2 \le \rho^C \|f\|_2$. Then

$$\left\|\mu^{(\ell)} * f - \nu^\rho * f\right\|_2 \le 6\rho^C \|f\|_2$$

so long as $\ell \gg \log_\lambda(\rho/C_1)$, see (101) for the dependence of the implied constant.

The proof will occupy the rest of this section and will be completed in several steps. Let us begin with the following lemma.

Lemma 28. *Let H be a compact group; assume that for some $0 < \eta < 1$ and constants C_1 and d_0, we have*

$$C_1^{-1}\eta^{d_0} \le |1_\eta| \le C_1 \eta^{d_0}. \tag{99}$$

Let σ be a symmetric Borel probability measure on H and assume that $\lambda(\sigma; H) < 1$. Then

$$|\sigma_\eta(X) - |X|| \le \lambda(\sigma; H)\Big(|X|/|1_\eta|\Big)^{1/2} \le \lambda(\sigma; H) C_1^{1/2}\eta^{-d_0/2}|X|^{1/2}$$

Proof. First note that the second estimate in the above upper bound is a direct consequence of (99) and the first inequality.

We now show the first inequality. Recall that $\sigma_\eta(X) = \sigma * P_\eta(X) = \langle T_\sigma(P_\eta), \mathbb{1}_X\rangle$. Similarly, we have $|X| = \langle T_\sigma(\mathbb{1}_H), \mathbb{1}_X\rangle$ (where we also used the invariance of the constant function). Thus,

$$\begin{aligned} |\sigma_\eta(X) - |X|| &= \langle T_\sigma(P_\eta - \mathbb{1}_H), \mathbb{1}_X\rangle \\ &\le \|T_\sigma(P_\eta - \mathbb{1}_H)\|_2 \|\mathbb{1}_X\|_2 \\ &\le \lambda(\sigma; H)\|P_\eta - \mathbb{1}_H\|_2 |X|^{1/2} \end{aligned} \tag{100}$$

Therefore, we need to compute $\|P_\eta - \mathbb{1}_H\|_2$. By the definition we have

$$\begin{aligned} \|P_\eta - \mathbb{1}_H\|_2^2 &= \int |P_\eta - 1|^2 dh \\ &= \int_{1_\eta} |\tfrac{1}{|1_\eta|} - 1|^2 dh + \int_{H\setminus 1_\eta} dh \\ &= |1_\eta| \tfrac{(1-|1_\eta|)^2}{|1_\eta|^2} + 1 - |1_\eta| = \tfrac{1-|1_\eta|}{|1_\eta|}. \end{aligned}$$

This and (100) imply that

$$|\sigma_\eta(X) - |X|| \le \lambda(\sigma; H)\Big(\tfrac{1-|1_\eta|}{|1_\eta|}\Big)^{1/2}|X|^{1/2}$$

and complete the proof. □

We now begin the proof of the theorem.

Proof of Theorem 27. As was mentioned before, the proof will be completed in some steps.

Let $\hat{C}$ be an integer $\ge \max\{L(C + d_{0i}) + C + 1, \frac{1}{d_{0i}} + 1\}$, and let $\delta = \rho^{\hat{C}}$. Let

$$\ell \ge -(2\hat{C}^2 + 0.5)\log_\lambda C_1 + \max\{d_{01}, d_{02}\}(0.5 + 4\hat{C}^2)\log_\lambda \rho. \tag{101}$$

Then for $i = 1, 2$, we have

$$\lambda(\pi_i\mu^{(\ell)}; G_i) \le \lambda^\ell \le C_1^{-0.5-2\hat{C}^2}\rho^{(0.5+4\hat{C}^2)d_{0i}}. \tag{102}$$

Apply Lemma 28, with $\eta = \rho$, $H = G_i$ and $\sigma = \pi_i\mu^{(\ell)}$ for $i = 1, 2$. Then (102) and (97) imply

$$\begin{aligned} \left|\pi_i\mu_\rho^{(\ell)}(Y_i) - |Y_i|\right| &\le C_1^{-0.5-2\hat{C}^2} \cdot \rho^{(0.5+4\hat{C}^2)d_{0i}} \cdot C_1^{1/2} \cdot \rho^{-d_{0i}/2}|Y_i|^{1/2} \\ &\le C_1^{-2\hat{C}^2}\rho^{4\hat{C}^2 d_{0i}}|Y_i|^{1/2} \end{aligned} \tag{103}$$

for any Borel subset $Y_i \subset G_i$.

The definition of ν^ρ. Recall that by our assumption G_1 and G_2 are δ-discretizable. For $i = 1, 2$, let $\{X^i_j : 1 \le j \le N_i\}$ be a δ-discretization of G_i. In view of (97) (with $\eta = \delta, \delta^2$) and (96), we have

$$C_1^{-1}\delta^{2d_{0i}} \le 1/N_i \le C_1\delta^{d_{0i}} \qquad \text{for } i = 1, 2. \tag{104}$$

Let $Z_i = \{1, \dots, N_i\}$ for $i = 1, 2$, and define $\tilde{\mu}$ on $Z_1 \times Z_2$ by

$$\tilde{\mu}(j, k) = \mu^{(\ell)}_\rho(X^1_j \times X^2_k).$$

Then for all $1 \le j \le N_1$, we have

$$\begin{aligned}\left|\pi_1\tilde{\mu}(j) - \tfrac{1}{N_1}\right| &= \left|\mu^{(\ell)}_\rho(X^1_j \times G_2) - |X^1_j \times G_2|\right| \\ &= \left|\pi_1\mu^{(\ell)}_\rho(X^1_j) - |X^1_j|\right| \le C_1^{-2\hat{C}^2}\rho^{4\hat{C}^2 d_{0i}}\left|X^1_j\right|^{1/2} \\ &\le (1/N_1N_2)^{\hat{C}} \end{aligned} \tag{105}$$

where the second to the last inequality follows from (103) and the last inequality follows from (104).

Similarly, for all $1 \le k \le N_2$, we have $|\pi_2\tilde{\mu}(k) - \frac{1}{N_2}| \le (1/N_1N_2)^{\hat{C}}$.

Altogether, the conditions of Proposition 5 are satisfied for $\tilde{\mu}$ with $A = \hat{C}$. Therefore, by that proposition, there exist $\{c_{j,k} \in [0, 1] : 1 \le j \le N_1, 1 \le k \le N_2\}$ so that all the following hold.

(1) For every $1 \le j \le N_1$, we have $\sum_{k=1}^{N_2} c_{j,k} = \frac{1}{N_1}$.
(2) For every $1 \le k \le N_2$, we have $\sum_{j=1}^{N_1} c_{j,k} = \frac{1}{N_2}$.
(3) For all $1 \le j \le N_1$ and all $1 \le k \le N_2$ we have

$$|\mu^{(\ell)}_\rho(X^1_j \times X^2_k) - c_{j,k}| \le (1/N_1N_2)^{\hat{C}-1}. \tag{106}$$

Let ν be the probability measure on $G_1 \times G_2$ defined using the density

$$N_1N_2\sum_{j,k} c_{j,k}\mathbb{1}_{X^1_j\times X^2_k};$$

note that ν depends on ρ. Abusing the notation, we also refer to the density of ν by ν.

Bulk of the proof is to show that ν satisfies the claim in the theorem, (possibly) except for being symmetric; the proof will then be completed by symmetrizing ν.

Sublemma. *The measure ν is a coupling of m_1 and m_2.*

Proof of the Sublemma. Since ν is absolutely continuous with respect to $m_1 \times m_2$, with density $N_1N_2\sum_{j,k} c_{j,k}\mathbb{1}_{X^1_j\times X^2_k}$, it suffices to show that for $i = 1, 2$, we have

$$\int_{G_i} N_1N_2\sum_{j,k} c_{j,k}\mathbb{1}_{X^1_j\times X^2_k}(g_1, g_2)dm_i = 1.$$

We prove this claim for $i = 1$, the other case is proved similarly. Recall that $|X^1_j| = 1/N_1$ and that $\sum_{j=1}^{N_1} c_{j,k} = \frac{1}{N_2}$ for all k. Thus we have

$$\begin{aligned}\int_{G_1} N_1N_2\sum_{j,k} c_{j,k}\mathbb{1}_{X^1_j\times X^2_k}(g_1, g_2)dm_1(g_1) &= N_1N_2\sum_{j,k} c_{j,k}|X^1_j|\mathbb{1}_{X^2_k}(g_2) \\ = N_2\sum_{j,k} c_{j,k}\mathbb{1}_{X^2_k}(g_2) = N_2\sum_k \mathbb{1}_{X^2_k}(g_2)\sum_j c_{j,k} &= \sum_k \mathbb{1}_{X^2_k}(g_2) = 1\end{aligned}$$

where in the last equality we used the fact that $\{X^2_k\}$ is a partition of G_2. □

Recall that $f \in L^2(G_1 \times G_2, m_1 \times m_2)$ and satisfies $\|f - f_\rho\|_2 \le \rho^C \|f\|_2$ where $f_\rho = P_\rho * f$.

Sublemma. *We have*

$$\left\|\mu^{(\ell)} * f - \nu * f\right\|_2 \le \left\|\mu_\rho^{(\ell)} * f_\rho - \nu * f_\rho\right\|_2 + 3\rho^C \|f\|_2. \tag{107}$$

Proof of the Sublemma. Indeed, by Young's inequality, we have

$$\begin{aligned}
\left\|\mu^{(\ell)} * f - \mu^{(\ell)} * f_\rho\right\|_2 &\le \|f - f_\rho\|_2 \le \rho^C \|f\|_2,\\
\left\|\mu^{(\ell)} * f_\rho - \mu_\rho^{(\ell)} * f_\rho\right\| &= \left\|\mu^{(\ell)} * P_\rho * (f - f_\rho)\right\|_2 \le \|f - f_\rho\|_2 \le \rho^C \|f\|_2,\\
\left\|\nu * f_\rho - \nu * f\right\|_2 &\le \|f - f_\rho\|_2 \le \rho^C \|f\|_2.
\end{aligned}$$

Now (107) follows from these estimates and the triangle inequality. □

In view of (107), thus we need to bound $\left\|\mu_\rho^{(\ell)} * f_\rho - \nu * f_\rho\right\|_2$. This will be done using the Parseval's theorem. Let us begin with the following which is a consequence of the fact that G_1 and G_2 are locally random groups, together with the fact that $\operatorname{diam}(X_j^i) \le \delta = \rho^{\hat{C}}$.

Sublemma. *Let σ be a Borel probability measure on $G = G_1 \times G_2$. Let $\varphi \in \widehat{G}$, and for all $1 \le j \le N_1$ and $1 \le k \le N_2$, let $g_{j,k} \in X_j^1 \times X_k^2$. Then*

$$\left\|\hat{\sigma}(\varphi) - \sum_{j,k} \sigma(X_j^1 \times X_k^2)\varphi(g_{j,k})\right\|_{\mathrm{op}} \le 2C_0(\dim \varphi)^L \delta.$$

Proof of the Sublemma. Since $\{X_j^1 \times X_k^2\}$ is a partition of G with Borel sets, we have

$$\begin{aligned}
\hat{\sigma}(\varphi) &= \int \varphi(g) d\sigma(g) = \sum_{j,k} \int_{X_j^1 \times X_k^2} \varphi(g) d\sigma(g)\\
&= \sum_{j,k} \Big(\int_{X_j^1 \times X_k^2} \varphi(g) - \varphi(g_{j,k}) d\sigma(g) + \sigma(X_j^1 \times X_k^2)\varphi(g_{j,k}) \Big).
\end{aligned}$$

Recall that G_1 and G_2 are L-locally random with coefficient C_0, thus, G is L-locally random with coefficient $2C_0$, [20, Lemma 5.2]. In consequence, for all $g \in X_j^1 \times X_k^2$, we have

$$\|\varphi(g) - \varphi(g_{j,k})\|_{\mathrm{op}} \le 2C_0 \dim(\varphi)^L d(g, g_{j,k}) \le 2C_0 \dim(\varphi)^L \delta,$$

where we used $\operatorname{diam}(X_j^1 \times X_k^2) \le \delta$.

Altogether, we conclude that

$$\left\|\hat{\sigma}(\varphi) - \sum_{j,k} \sigma(X_j^1 \times X_k^2)\varphi(g_{j,k})\right\|_{\mathrm{op}} \le \sum_{j,k} \int_{X_j^1 \times X_k^2} 2C_0 \dim(\varphi)^L \delta d\sigma(g) = 2C_0 \dim(\varphi)^L \delta$$

where we used the fact that $\{X_j^1 \times X_k^2\}$ is a Borel partition of G and $\sigma(G) = 1$. □

Applying the above with $\sigma = \nu$ and μ_ρ^ℓ, we conclude the following

$$\left\|\hat{\nu}(\varphi) - \sum_{j,k} c_{j,k}\varphi(g_{j,k})\right\|_{\mathrm{op}} \le 2C_0(\dim \varphi)^L \delta, \quad \text{and} \tag{109a}$$

$$\left\|\widehat{\mu_\rho^{(\ell)}}(\varphi) - \sum_{j,k} \mu_\rho^{(\ell)}(X_j^1 \times X_k^2)\varphi(g_{j,k})\right\|_{\mathrm{op}} \le 2C_0(\dim \varphi)^L \delta \tag{109b}$$

for all $\varphi \in \widehat{G}$, where we also used the fact that $\nu(X_j^1 \times X_k^2) = c_{j,k}$.

We now combine (109a), (109b), and (106) with Parseval's theorem to deduce the following

Sublemma. *We have*

(110) $$\big\|\mu_\rho^{(\ell)} * f_\rho - \nu * f_\rho\big\|_2 \le 3\rho^C \|f\|_2.$$

Proof of the Sublemma. The argument is similar to arguments in [20, §6], and as was mentioned before, is based on Parseval's theorem:

$$\big\|\mu_\rho^{(\ell)} * f_\rho - \nu * f_\rho\big\|_2^2 = \sum_{\varphi\in\widehat{G}} \dim\varphi \big\|\big(\widehat{\mu_\rho^{(\ell)}}(\varphi) - \hat{\nu}(\varphi)\big)\hat{f}_\rho(\varphi)\big\|_{\mathrm{HS}}^2$$

where $G = G_1 \times G_2$.

Let us write $d_0 = d_{01} + d_{02}$, see (97). Let $D = \lceil 4C_1^2\rho^{-2C-d_0}\rceil$, this choice will be justified later in the proof. We separate the above sum into $\sum_{\dim\varphi\le D}$ and $\sum_{\dim\varphi\ge D}$. Using the notation in [20, §6], the first sum will be denoted by $L(\mu_\rho^{(\ell)} * f_\rho - \nu * f_\rho; D)$ and second sum by $H(\mu_\rho^{(\ell)} * f_\rho - \nu * f_\rho; D)$.

First note that in view of [20, Lemma 6.1], we have

$$\begin{aligned} H\big((\mu_\rho^{(\ell)} - \nu) * f_\rho; D\big) = H\big((\mu_\rho^{(\ell)} - \nu) * f * P_\rho; D\big) &\le \frac{1}{D} H\big((\mu_\rho^{(\ell)} - \nu) * f; D\big) H(P_\rho; D) \\ &\le \frac{1}{D}\|\mu_\rho^{(\ell)} - \nu) * f\|_2^2 \|P_\rho\|_2^2 \le \frac{4}{D|1_\rho|}\|f\|_2^2 \\ &\le \frac{4C_1}{D\rho^{d_0}}\|f\|_2^2 \qquad (111) \end{aligned}$$

where we used Young's in equality in the second line and (97) with $\eta = \rho$ in the last line.

We now investigate

$$L(\mu_\rho^{(\ell)} * f_\rho - \nu * f_\rho; D) = \sum_{\varphi\in\widehat{G}, \dim\varphi\le D} \dim\varphi \big\|\big(\widehat{\mu_\rho^{(\ell)}}(\varphi) - \hat{\nu}(\varphi)\big)\hat{f}_\rho(\varphi)\big\|_{\mathrm{HS}}^2.$$

First note that $\|(\widehat{\mu_\rho^{(\ell)}}(\varphi) - \hat{\nu}(\varphi))\hat{f}_\rho(\varphi)\|_{\mathrm{HS}}^2 \le \|\widehat{\mu_\rho^{(\ell)}}(\varphi) - \hat{\nu}(\varphi)\|_{\mathrm{op}}^2 \|\hat{f}_\rho(\varphi)\|_{\mathrm{HS}}^2$. Moreover, by the triangle inequality, we have

$$\begin{aligned} \big\|\widehat{\mu_\rho^{(\ell)}}(\varphi) - \hat{\nu}(\varphi)\big\|_{\mathrm{op}} \le \big\|\widehat{\mu_\rho^{(\ell)}}(\varphi) - \sum_{j,k}\mu_\rho^\ell(X_j^1 \times X_k^2)\varphi(g_{j,k})\big\|_{\mathrm{op}} + \big\|\widehat{\nu}(\varphi) - \sum_{j,k} c_{j,k}\varphi(g_{j,k})\big\|_{\mathrm{op}} \\ + \big\|\sum_{j,k}\big(\mu_\rho^\ell(X_j^1\times X_k^2) - c_{j,k}\big)\varphi(g_{j,k})\big\|_{\mathrm{op}}. \end{aligned}$$

Hence, using (109a), (109b), and (106), we have

$$\big\|\widehat{\mu_\rho^{(\ell)}}(\varphi) - \hat{\nu}(\varphi)\big\|_{\mathrm{op}} \le 4C_0(\dim\varphi)^L\delta + N_1N_2(1/N_1N_2)^{\hat{C}-1} \le 4C_0(\dim\varphi)^L\delta + (C_1^2\delta^{d_0})^{\hat{C}-2}$$

where $d_0 = d_{01} + d_{02}$ and we used (104) for the last inequality. From this we conclude that

$$\begin{aligned} L(\mu_\rho^{(\ell)} * f_\rho - \nu * f_\rho; D) &\le \Big(4C_0D^L\delta + (C_1^2\delta^{d_0})^{\hat{C}-2}\Big)^2 \sum_{\varphi\in\widehat{G},\dim\varphi\le D} \dim\varphi \|\hat{f}_\rho(\varphi)\|_{\mathrm{HS}}^2 \\ &\le \Big(4C_0D^L\delta + (C_1^2\delta^{d_0})^{\hat{C}-2}\Big)^2 \|f\|_2^2. \qquad (112) \end{aligned}$$

Recall that $D = \lceil 4C_1\rho^{-2C-d_0}\rceil$, then $\frac{4C_1}{D\rho^{d_0}} \le \rho^{2C}$. Since $\hat{C} \ge L(C+d_0)+C+1$, and $\rho \le \frac{1}{4C_0(5C_1)^L}$, we get

$$4C_0D^L\delta \le 4C_0(4C_1\rho^{-C-d_0}+1)^L\rho^{\hat{C}} \le \rho^C;$$

moreover, since $d_0(\hat{C}-1)\geq 1$ we have $(C_1^2\delta^{d_0})^{\hat{C}-2}\leq \rho^C$. Thus, (111) and (112) imply that

$$\left\|\mu_\rho^{(\ell)} * f_\rho - \nu * f_\rho\right\|_2 \leq 3\rho^C$$

as we claimed. □

We are now in the position to complete the proof of Theorem 27. First note that in view of (107) and (110) we have

$$\left\|\mu^{(\ell)} * f - \nu * f\right\|_2 \leq 6\rho^C\|f\|_2$$

As was mentioned before, ν need not be symmetric. Define $\nu^\rho = \frac{\check{\nu}+\nu}{2}$ to be the symmetrization of ν where $\check{h}(g) = h(g^{-1})$ for any $h \in L^2(G, m_1 \times m_2)$.

Recall that $\|\check{h}\|_2 = \|h\|_2$ for all $h \in L^2(G, m_1 \times m_2)$. Since μ is symmetric, we have

$$\left\|\mu^{(\ell)} * f - \check{\nu} * f\right\|_2 = \left\|(\mu^{(\ell)} * \check{f} - \nu * \check{f})\check{}\,\right\|_2 = \left\|\mu^{(\ell)} * \check{f} - \nu * \check{f}\right\|_2 \leq 6\rho^C\|\check{f}\|_2 = 6\rho^C\|f\|_2.$$

Altogether, and using the facts that $\rho \leq 0.01$, we get that

$$\left\|\mu^{(\ell)} * f - \nu^\rho * f\right\|_2 \leq 6\rho^C\|f\|_2 \leq \rho^{C-0.5}\|f\|_2.$$

The proof is complete. □

6. Contraction of couplings at small scales

The main goals of this section are to prove Proposition 29 and Proposition 30, which are crucial ingredients for the proof of Theorem 1.

In this section we will be working with groups G_1 and G_2 which are L-locally random with coefficients C_{01} and C_{02}, respectively, and satisfy $\mathrm{DC}(d_{0i}, C_{1i})$. Throughout this section, $\mathsf{D}_\bullet$ denotes a constant of the form

$$\left(2C_{01}C_{02}C_{11}C_{12}\right)^{O_{d_{01},d_{02},L}(1)}; \tag{113}$$

this means the exponent in the definition of $\mathsf{D}_\bullet$ does not depend on other parameters introduced in the various statements throughout this section.

Proposition 29. *Suppose F_1 and F_2 are two local fields of characteristic zero, $\mathbb{G}_i$ is an almost F_i-simple group, and $\mathrm{Lie}(\mathbb{G}_1)(F_1)$ and $\mathrm{Lie}(\mathbb{G}_2)(F_2)$ are not isomorphic. For $i = 1, 2$, let $G_i \subseteq \mathbb{G}_i(F_i)$ be a compact open subgroup. For every $\bar{\delta} > 0$, there exists $\eta_0 \geq \mathsf{D}_1{}^{-1/\bar{\delta}}$ where D_1 is a constant as in (113), and a positive integer $m := m(\bar{\delta})$ such that for every $0 < \eta \leq \eta_0$ and every coupling μ of the probability Haar measures m_{G_1} and m_{G_2}, we have*

$$H_2(\mu^{(2^m)}; \eta) \geq \left(d_{01} + d_{02} - \bar{\delta}\right)\log(1/\eta).$$

The proof of this proposition will occupy the rest of this section. We will then use this proposition to prove Proposition 30 below. Before stating Proposition 30, we recall Definition 3: A function $f \in L^2(G)$ is said to live at scale η (with parameter $0 < a < 1$) if

- (Averaging to zero) $\|f_{\eta^{1/a}}\|_2 \leq \eta^{1/(2a)}\|f\|_2$.
- (Almost invariant) $\|f_{\eta^{a^2}} - f\|_2 \leq \eta^{a/2}\|f\|_2$.

Proposition 30. *In the setting of Proposition 29, there exist a positive integer m_0 and a positive number c, depending only on L, d_{01}, and d_{02} such that for every $0 < \eta \leq \mathsf{D}_2{}^{-1}$ where D_2 is a constant as in (113), every coupling μ of the probability Haar measures m_{G_1} and m_{G_2}, and every function $f \in L^2(G_1 \times G_2)$ which lives at scale η with a parameter $a \geq 4L(d_{01} + d_{02})$, we have*

$$\|\mu^{(2^{m_0})} * f\|_2 \leq \eta^c\|f\|_2.$$

Proposition 30, whose proof is based on Proposition 29, is a crucial ingredient in the proof of Theorem 1.

6.1. **Contraction, Rényi entropy, and approximate subgroups.** In this section, using the mixing inequality as in [20, Theorem 2.6] and the multi-scale version of a result of Bourgain and Gamburd (see [20, Theorem 2.12]), we justify why in the proofs of the aforementioned propositions one needs to study certain type of approximate subgroups of $G_1 \times G_2$.

We start by finding a lower bound for the Rényi entropy of every coupling of the Haar measures m_{G_1} and m_{G_2}.

Lemma 31. *Let G_1 and G_2 be two compact groups and μ be a coupling of the Haar measures m_{G_1} and m_{G_2}. Then for every $0<\eta<1$, we have*

$$H_2(\mu;\eta) \geq \max(\log(1/|1_\eta^{(1)}|), \log(1/|1_\eta^{(2)}|)).$$

Proof. By [20, Lemma 8.2], we have $\mu_\eta(x) = \mu(x_\eta)/|1_\eta|$ for every $x \in G_1 \times G_2$. Therefore

$$\mu_\eta(x,x') = \frac{\mu(x_\eta \times x'_\eta)}{|1_\eta^{(1)}||1_\eta^{(2)}|} \leq \frac{\mu(G_1 \times x'_\eta)}{|1_\eta^{(1)}||1_\eta^{(2)}|} = \frac{|x'_\eta|}{|1_\eta^{(1)}||1_\eta^{(2)}|} = \frac{1}{|1_\eta^{(1)}|}.$$

By symmetry, we have

$$\|\mu_\eta\|_\infty \leq \min\Big(\frac{1}{|1_\eta^{(1)}|}, \frac{1}{|1_\eta^{(2)}|}\Big). \tag{114}$$

Since μ_η is a probability measure, we deduce from (114) that

$$\begin{aligned} H_2(\mu;\eta) &= \log(1/|1_\eta|) - \log\|\mu_\eta\|_2^2 \\ &\geq \log(1/|1_\eta^{(1)}|) + \log(1/|1_\eta^{(2)}|) - \log\|\mu_\eta\|_\infty \\ &\geq \max(\log(1/|1_\eta^{(1)}|), \log(1/|1_\eta^{(2)}|)), \end{aligned}$$

as we claimed. □

Lemma 32. *Suppose G is an L-locally random group with coefficient C_0 which satisfies the dimension condition $\mathrm{DC}(d_0, C_1)$. Let $0<\eta<(10C_0+C_1)^{-8La}$, where a is a positive number. Let $0<\rho\leq\eta$ and suppose that ν is a probability measure on G such that*

$$H_2(\nu;\rho) \geq \Big(d_0 - \frac{1}{8La\log_\eta \rho}\Big)\log(1/\rho). \tag{115}$$

Then for every function $f \in L^2(G)$ that lives at scale η, we have

$$\|\nu_\rho * f\|_2 \leq \eta^{1/(8La)}\|f\|_2.$$

Proof. Notice that $C_0\eta^{1/(4a)} \leq 0.1$, hence by [20, Theorem 2.6], we have

$$\begin{aligned} \|\nu_\rho * f\|_2^2 &\leq 2\|(\nu_\rho)_{\eta^{1/a}} * f_{\eta^{1/a}}\|_2^2 + \eta^{1/(2La)}\|\nu_\rho\|_2^2\,\|f\|_2^2 \\ &\leq 2\eta^{1/(2a)}\|f\|_2 + \eta^{1/(2La)}\|\nu_\rho\|_2^2\,\|f\|_2^2 && (f \text{ lives at scale } \eta). \end{aligned} \tag{116}$$

On the hand, by (115), we obtain

$$\begin{aligned} \log\|\nu_\rho\|_2^2 &\leq \log(1/|1_\rho|) - d_0\log(1/\rho) + \frac{1}{8La\log_\eta\rho}\log(1/\rho) \\ &\leq \log C_1 + \frac{1}{8La}\log(1/\eta) && \text{(because of DC)} \\ &\leq \frac{1}{4La}\log(1/\eta) && (\text{as } \eta < (10C_0+C_1)^{-8La}) \end{aligned} \tag{117}$$

By (116) and (117), we deduce

$$\|\nu_\rho * f\|_2^2 \leq (2\eta^{1/(2a)} + \eta^{1/(4La)})\|f\|_2^2 \leq \eta^{1/(8La)}\|f\|_2^2,$$

and the claim follows. □

Our general strategy for the proofs of the main propositions is as follows: starting with the initial entropy provided by Lemma 31, if we can show that each time after doubling the number of steps in the random walk we can gain $\gamma_0 \log(1/\eta)$ additional Rényi entropy at scale η, then in $O(1)$-steps we reach to the desired lower bound for the Rényi entropy that is given in Lemma 32.

The following lemma, which follows from [20, Theorem 2.12], is an important tool in carrying out the above strategy. Roughly speaking, it states that the failure to gain Rényi entropy can happen only because of algebraic obstructions.

Lemma 33. *Suppose G satisfies $\mathrm{DC}(d_0, C_1)$ and X, X' are independent and identically distributed random variables with values in G. Then for every positive number γ_0, either*

$$H_2(XX';\eta) \geq H_2(X;\eta) + \gamma_0 \log(1/\eta) \tag{118}$$

or there are $H \subseteq G$ and $x, y \in G$ such that

1. *(Approximate structure) H is $R(1/\eta)^{R\gamma_0}$-approximate subgroup.*
2. *(Metric entropy) $|h(H;\eta) - H_2(X;\eta)| \leq R\gamma_0 \log(1/\eta)$.*
3. *(Almost equidistribution) Let Z be a random variable with the uniform distribution over $1_{3\eta}$ independent of X. Then*

$$\mathbb{P}(XZ \in (xH)_\eta) \geq \eta^{R\gamma_0} \text{ and } \mathbb{P}(XZ \in (Hy)_\eta) \geq \eta^{R\gamma_0}.$$

Moreover,

$$|\{h \in H_\eta | \, \mathbb{P}(X \in (xh)_{3\eta}) \geq (C_1 2^{d_0})^{-R} \eta^{10\gamma_0} 2^{-H_2(X;\mu)}\}| \geq \eta^{R\gamma_0} |H_\eta|$$

where R is a universal fixed number.

Proof. This is an immediate corollary of [20, Theorem 2.12]. □

6.2. **Approximate subgroups and approximate homomorphisms.** In view of Lemma 33, we focus on the understanding of almost subgroups H of $G_1 \times G_2$ which satisfy properties given in Lemma 33. Indeed, we will interpret this approximate structure, as a local approximate group homomorphism with *large* image from a *large* ball in G_1 to G_2. Then we apply Theorem 3 to complete the proof.

Let us begin with an application of a product result proved in [20, Theorem 2.8].

Lemma 34. *Suppose G_1 and G_2 are L-locally random with coefficients C_{01} and C_{02}, respectively. Suppose G_i satisfies $\mathrm{DC}(d_{0i}, C_{1i})$ for $i = 1, 2$. Then for every $0 < \varepsilon < 1$, there is $\gamma := \gamma(\varepsilon, L, d_{01}, d_{02}) \ll_{L, d_{01} d_{02}} \varepsilon$ such that for every η that satisfies $\eta^\varepsilon \leq \mathsf{D}_3{}^{-1}$, where D_3 is a constant as in (113), the following holds. Suppose $X := (X_1, X_2)$ is a random variable with values in $G := G_1 \times G_2$ such that X_i is uniformly distributed in G_i. Let Z be a random variable independent of X and with uniform distribution over $1_{3\eta} \subseteq G_1 \times G_2$ with respect to the maximum metric. Suppose $H \subseteq G$, $x \in G$, and $\mathbb{P}(XZ \in (xH)_\eta) \geq \eta^{R\gamma}$ where R is a universal fixed number. Then*

$$\mathrm{pr}_i(H_\eta H_\eta H_\eta^{-1} H_\eta^{-1}) \supseteq 1^{(i)}_{\eta^\varepsilon}$$

for $i = 1, 2$ where $\mathrm{pr}_i : G \to G_i$ is the projection to the i-th component.

Proof. Let γ be a constant which will be determined in the proof. Notice that $X_i Z_i$ is uniformly distributed in G_i where $Z_i := \mathrm{pr}_i(Z)$. Therefore,

$$|\,\mathrm{pr}_i(H)_\eta| = |\,\mathrm{pr}_i((xH)_\eta)| = \mathbb{P}(X_i Z_i \in \mathrm{pr}_i((xH)_\eta)) \geq \mathbb{P}(XZ \in (xH)_\eta) \geq \eta^{R\gamma}. \tag{119}$$

From (119), we deduce that

$$
\begin{aligned}
h(\mathrm{pr}_i(H);\eta) \geq& \log(1/|1^{(i)}_\eta|) - R\gamma\log(1/\eta) - 2\log C_{1i}\\
\geq& d_{0i}\Big(1-\frac{R\gamma}{d_{0i}}\Big)\log(1/\eta) - 3\log C_{1i}\\
\geq& \Big(1-\frac{R\gamma}{d_{0i}}\Big)h(G_i;\eta) - 4\log C_{1i}\\
\geq& \Big(1-\frac{2R\gamma}{d_{0i}}\Big)h(G_i;\eta), \qquad (120)
\end{aligned}
$$

where the last inequality holds as long as

$$\eta^{R\gamma} \leq (C_{11}C_{12})^{-5} \qquad (121)$$

By [20, Theorem 2.8], there is $\delta := \delta(\varepsilon, L, d_{0i}) \ll_{L,d_{0i}} \varepsilon$ such that the following holds: if $h(A;\eta) \geq (1-\delta)h(G_i;\eta)$ and

$$\eta^{\varepsilon} \leq (2C_{01}C_{02}C_{11}C_{12})^{-\overline{R}} \qquad (122)$$

(where $\overline{R}$ depends polynomially on L and d_{0i}, $i=1,2$) , then $A_\eta A_\eta A_\eta^{-1}A_\eta^{-1} \supseteq 1^{(i)}_{\eta^{\varepsilon}}$. We claim

$$\gamma = \frac{1}{2R}\min(\delta(\varepsilon, L, d_{01}), \delta(\varepsilon, L, d_{02}))$$

satisfies the claim in the lemma so long as η is small enough.

Indeed, the above definition implies

$$\gamma \ll_{L,d_{0i}} \varepsilon. \qquad (123)$$

Moreover, in view of (123), there exists R' depending on L, d_{0i} such that if $\eta^{\varepsilon} \leq (2C_{01}C_{02}C_{11}C_{12})^{-R'}$, then (121) and (122) both hold. Hence, as it was discussed above, the lemma follows by (120) and [20, Theorem 2.8]. □

For a symmetric subset H of $G_1 \times G_2$ containing $(1^{(1)}, 1^{(2)})$ and $\eta > 0$, set

$$\alpha_1(H;\eta) := \inf\{\alpha \in [0,1] |\ \exists g_1 \in G_1, d(g_1, 1^{(1)}) \geq \eta^{\alpha},\ (g_1, 1^{(2)}) \in \textstyle\prod_3 H_\eta\}, \qquad (124)$$

and similarly

$$\alpha_2(H;\eta) := \inf\{\alpha \in [0,1] |\ \exists g_2 \in G_2, d(g_2, 1^{(2)}) \geq \eta^{\alpha},\ (1^{(1)}, g_2) \in \textstyle\prod_3 H_\eta\};$$

recall that d always denotes our fixed bi-invariant metric on the underlying compact group.

Lemma 35. *Suppose G_1 and G_2 are two compact groups, η is a positive number, and $H \subseteq G_1 \times G_2$ is symmetric containing the identity. Then there is $f : \mathrm{pr}_1(H_\eta) \to G_2$ which is a $\mathrm{pr}_1(H_\eta)$-partial, $\eta^{\alpha_2(H;\eta)}$-approximate homomorphism; that means*

(1) $f(1^{(1)}) = 1^{(2)}$,
(2) if $g_1, g_1' \in \mathrm{pr}_1(H_\eta)$ and $g_1g_1' \in \mathrm{pr}_1(H_\eta)$, we have $d(f(g_1)f(g_1'), f(g_1g_1')) \leq \eta^{\alpha_2(H;\eta)}$, and
(3) for every $g_1 \in \mathrm{pr}_1(H_\eta)$, $d(f(g_1^{-1}), f(g_1)^{-1}) \leq \eta^{\alpha_2(H;\eta)}$.

Furthermore $\mathrm{pr}_2(H_\eta) \subseteq (\mathrm{Im}\, f)_{\eta^{\alpha_2(H;\eta)}}$.

Proof. For every $g_1 \in \mathrm{pr}_1(H_\eta)$, choose $f(g_1) \in G_2$ such that $(g_1, f(g_1)) \in H_\eta$. As $(1^{(1)}, 1^{(2)}) \in H$, we can and will set $f(1^{(1)}) = 1^{(2)}$. For every $g_1, g_1' \in \mathrm{pr}_1(H_\eta)$ with $g_1g_1' \in \mathrm{pr}_1(H_\eta)$,

$$(1^{(1)}, f(g_1)f(g_1')f(g_1g_1')^{-1}) \in \textstyle\prod_3 H_\eta \quad \text{and} \quad (1^{(1)}, f(g_1)f(g_1^{-1})) \in \prod_2 H_\eta \subseteq \prod_3 H_\eta.$$

Hence $d(f(g_1)f(g_1'), f(g_1g_1')) \leq \eta^{\alpha_2(H;\eta)}$ and $d(f(g_1^{-1}), f(g_1)^{-1}) \leq \eta^{\alpha_2(H;\eta)}$.

For every $g_2 \in \mathrm{pr}_2(H)_\eta$, there is $g_1 \in \mathrm{pr}_1(H_\eta)$ such that $(g_1, g_2) \in H_\eta$. Hence $(1^{(1)}, g_2f(g_1)^{-1})$ is in $\prod_2 H_\eta$. Therefore $d(g_2, f(g_1)) \leq \eta^{\alpha_2(H;\eta)}$. This completes the proof. □

By Lemma 35, we get a good approximate homomorphism if $\alpha_i(H;\eta)$ is large for some subset H of $G_1\times G_2$. To get such a bound, inspired by Proposition 33, we consider an $\eta^{R\gamma}$-approximate subgroup of $G_1\times G_2$ with an upper bound for its metric entropy, and study $\alpha_i(\prod_k H;\eta)$ for a fixed positive integer k that will be determined later and depends on the dimensions of G_i's.

Let us recall the dimension condition of G_i's. For every positive number η we have

$$C_{1i}^{-1}\eta^{d_{0i}} \le |1_\eta^{(i)}| \le C_{1i}\eta^{d_{0i}} \tag{125}$$

where C_{1i} and d_{0i} are positive numbers.

We also recall the following two facts from [20, §7]. By [20, Lemma 7.1], for every non-empty subset A of $G=G_1\times G_2$, we have

$$\left|h(A;\eta)-\log\left(\frac{|A_\eta|}{|1_\eta|}\right)\right| \le \log(\mathsf{D}_4), \tag{126}$$

moreover, the same is true for the subsets of G_i's. By [20, Corollary 7.2], for every non-empty subset A of $G=G_1\times G_2$ and positive numbers η and a, we have

$$|h(A;\eta)-h(A;a\eta)| \le \log(\mathsf{D}_5). \tag{127}$$

where D_4 and D_5 are constants as in (113).

The following is an upgraded version of Lemma 35, and will be used the sequel.

Lemma 36. *Let G_1 and G_2 be two compact groups, $R,\gamma,\eta>0$, and let $H\subseteq G_1\times G_2$ be an $\eta^{-R\gamma}$-approximate subgroup. Assume further that for $i=1,2$ we have*

$$1^{(i)}_{\eta^{C_2\gamma}} \subseteq \mathrm{pr}_i(\textstyle\prod_4 H_\eta). \tag{128}$$

Then there exists an $\eta^{C_2\gamma}$-partial η^{α_2}-approximate homomorphism $f:1^{(1)}_{\eta^{C_2\gamma}}\to G_2$ satisfying that

$$1^{(2)}_{\eta^{C_2'\gamma}} \subseteq (\mathrm{Im}\, f)_{10\eta^{\alpha_2}}$$

where $\alpha_2:=\alpha_2(\prod_8 H;\eta)$ is as in (124) and C_2, C_2' depend only on L, d_{01}, and d_{02}.

Proof. Apply Lemma 35 with $\prod_8 H_\eta$ (instead of H), and let f be thus obtained. We may assume, without loss of generality, that $(g,f(g))\in H':=\prod_4 H_\eta$ for all $g\in \mathrm{pr}_1(H')$. Then since $\alpha_2(H'\cdot H';\eta)\le\alpha_2(H';\eta)$, Lemma 35, applied with H', implies that

$$\mathrm{pr}_2(H') \subseteq \big(f\big(\mathrm{pr}_1(H'_\eta)\big)\big)_{\eta^{\alpha_2}}. \tag{129}$$

In view of (128), we may restrict f to $1^{(1)}_{\eta^{C_2\gamma}}$ and obtain a $\eta^{C_2\gamma}$-partial η^{α_2}-approximate homomorphism $f:1^{(1)}_{\eta^{C_2\gamma}}\to G_2$. We now show that f also satisfies the last claim in the lemma.

To see this claim, let $C\ge 1$ be a constant which will be explicated later. Let us put

$$E_1=\big(f\big(1^{(1)}_{0.1\eta^{C_2\gamma}}\big)\big)_{\eta^{\alpha_2}} \quad\text{and}\quad E_2=\mathrm{pr}_2(H').$$

Assume first that

$$h(E_1;\eta^{\alpha_2}) \ge (1-C\gamma/\alpha_2)h(G_2;\eta^{\alpha_2}). \tag{130}$$

We want to apply [20, Theorem 2.8] with η^{α_2} and under the assumption (130). By [20, Theorem 2.8], there exists $C'=O_{L,d_{02}}(C)$ so that $\varepsilon=C'\gamma/\alpha_2$ and $\delta=C\gamma/\alpha_2$ satisfy the conditions in that theorem. Thus (130) and [20, Theorem 2.8] imply that

$$(E_1)_{\eta^{\alpha_2}}(E_1)_{\eta^{\alpha_2}}(E_1)^{-1}_{\eta^{\alpha_2}}(E_1)^{-1}_{\eta^{\alpha_2}} \supseteq 1^{(2)}_{\eta^{C'\gamma}}.$$

This and the fact that $\prod_4 1^{(1)}_{0.1\eta^{C_2\gamma}}\subseteq 1^{(1)}_{\eta^{C_2\gamma}}$ imply that

$$1^{(2)}_{\eta^{C'\gamma}} \subset \big(f(1^{(1)}_{\eta^{C_2\gamma}})\big)_{10\eta^{\alpha_2}}. \tag{131}$$

Hence, we assume that (130) fails. Since $1^{(2)}_{\eta^{C_2\gamma}} \subseteq E_2$, see (128), we have

$$h(E_2;\eta^{\alpha_2}) \geq (1 - C_2\gamma/\alpha_2) h(G_2;\eta^{\alpha_2}). \tag{132}$$

We now cover $\mathrm{pr}_1(H'_\eta)$ with $\leq (20)^{d_1} C_{11}^2 \eta^{-C_2\gamma d_1}$ many sets of the form $g1^{(1)}_{0.1\eta^{C_2\gamma}}$ where $g \in \mathrm{pr}_1(H'_\eta)$. Since $g1^{(1)}_{0.1\eta^{C_2\gamma}} \subseteq \mathrm{pr}_1(H'_\eta \cdot H'_\eta)$, we have that $f(gx)$ is defined for all $x \in 1^{(1)}_{0.1\eta^{C_2\gamma}}$. Recall from (129) that $E_2 \subseteq \big(f(\mathrm{pr}_1(H'_\eta))\big)_{\eta^{\alpha_2}}$. Therefore,

$$h(E_2;2\eta^{\alpha_2}) \leq \max_{g\in \mathrm{pr}_1(H'_\eta)} \big\{h\big(f(g1^{(1)}_{0.1\eta^{C_2\gamma}});2\eta^{\alpha_2}\big)\big\} + C_2\gamma d_1 \log(1/\eta) + \log\big((20)^{d_1} C_{11}^2\big). \tag{133}$$

Note on the other hands that

$$d(f(gx), f(g)f(x)) \leq \eta^{\alpha_2}, \qquad \text{for every } x \in 1^{(1)}_{0.1\eta^{C_2\gamma}}.$$

Thus the failure of (130) implies that

$$h\big(f(g1^{(1)}_{0.1\eta^{C_2\gamma}});2\eta^{\alpha_2}\big) \leq (1 - C\gamma/\alpha_2) h(G_2;\eta^{\alpha_2}), \quad \text{for all } g \in \mathrm{pr}_1(H'_\eta).$$

This and (133) imply

$$h(E_2;\eta^{\alpha_2}) \leq (\alpha_2 - C\gamma) d_2 \log(1/\eta) + C_2\gamma d_1 \log(1/\eta) + \log(\mathsf{D}_6); \tag{134}$$

where we used

$$h(E_2;\eta^{\alpha_2}) = h(E_2;2\eta^{\alpha_2}) + \log(\mathsf{D}_7) \quad \text{and} \quad h(G_2;\eta^{\alpha_2}) = \alpha_2 d_2 \log(1/\eta) + \log(\mathsf{D}_8),$$

for constants D_6, D_7, and D_8 as in (113). Thus, (134) contradicts (132) so long as η is small enough to account for the additive constants and $C \geq 3C_2 \max\{d_1/d_2, 1\}$. This and (131) finish the proof. □

The following two lemmas concern k fold product of approximate subgroups.

Lemma 37. *Suppose G_1 and G_2 are two compact groups, $R, \gamma, \eta > 0$, and $H \subseteq G_1 \times G_2$ is an $\eta^{-R\gamma}$-approximate subgroup. For a positive integer k, let*

$$H^{(1)}_k := \mathrm{pr}_1\Big((G_1 \times \{1^{(2)}\}) \cap \textstyle\prod_k H\Big).$$

Then

$$h(H^{(1)}_k;\eta) + h(\mathrm{pr}_2(\textstyle\prod_k H);\eta) \leq h(H;\eta) + 2kR\gamma \log(1/\eta) + \log(\mathsf{D}_9),$$

where D_9 multiplicatively depends on D_4 and D_5.

Proof. Notice that $|(\prod_{2k} H)_{2\eta}| \geq |(H^{(1)}_k)_\eta| |\mathrm{pr}_2(\prod_k H)_\eta|$. Hence

$$\log\Big(\frac{|(\prod_{2k} H)_{2\eta}|}{|1_\eta|}\Big) \geq \log\Big(\frac{|(H^{(1)}_k)_\eta|}{|1^{(1)}_\eta|}\Big) + \log\Big(\frac{|\mathrm{pr}_2(\prod_k H)_\eta|}{|1^{(2)}_\eta|}\Big).$$

Therefore, by (126) and (127), we obtain

$$h(\textstyle\prod_{2k} H;\eta) + \log(\mathsf{D}_{10}) \geq h(H^{(1)}_k;\eta) + h(\mathrm{pr}_2(\prod_k H);\eta), \tag{135}$$

where D_{10} multiplicatively depends on D_4 and D_5.

Since H is an $\eta^{-R\gamma}$-approximate subgroup, there is a symmetric set A of cardinality at most $\eta^{-R\gamma}$ such that $HH \subseteq HA$. Therefore, $\prod_{2k} H$ is a subset of $H\prod_{2k} A$, which implies that $|(\prod_{2k} H)_\eta| \leq \eta^{-2kR\gamma}|H_\eta|$. Hence, there exists D_{11}, so that

$$h(\textstyle\prod_{2k} H;\eta) \leq h(H;\eta) + 2kR\gamma \log(1/\eta) + \log(\mathsf{D}_{11}) \tag{136}$$

By (135) and (136), the claim follows. □

Lemma 38. *Suppose G_1 and G_2 are L-locally random with coefficients C_{01} and C_{02}, respectively. Suppose G_i satisfies the $\mathrm{DC}(d_{0i}, C_{1i})$. Suppose $k \geq 4$ is an integer and $\bar{\delta} > 0$. Let η and $\gamma \ll_{d_{0i},L,k} \bar{\delta}$ be positive numbers, and $\eta^{\gamma} \leq \mathsf{D}_{12}{}^{-1}$ where D_{12} is a constant as in (113). Suppose X and X' are independent and identically distributed random variables with values in G satisfying the following properties:*

- $H_2(XX';\eta) < H_2(X;\eta) + \gamma \log(1/\eta)$, *and*
- $H_2(X;\eta) < (d_{01} + d_{02} - \bar{\delta}) \log(1/\eta)$.

Then there is an $\eta^{-R\gamma}$-approximate subgroup H of $G_1 \times G_2$, where R is the universal constant given in Lemma 33, satisfying both of the following properties

$$1^{(i)}_{\eta^{C_2\gamma}} \subseteq \mathrm{pr}_i(\textstyle\prod_k H_\eta), \quad \textit{and} \quad h(H_k^{(1)};\eta) \leq \left(1 - \frac{\bar{\delta}}{2d_{01}}\right) h(G_1;\eta)$$

where C_2 depends only on L, d_{01}, and d_{02} and $H_k^{(1)}$ is defined as in Lemma 37.

Proof. We let $0 < \eta < (10C_{01} + 10C_{02} + C_{11} + C_{12})^{-1/\bar{\delta}}$ be a small constant which will be determined later. By Lemma 33 and Lemma 34, there is an $\eta^{-R\gamma}$-approximate subgroup H such that

$$1^{(i)}_{\eta^{C_2\gamma}} \subseteq \mathrm{pr}_i(\textstyle\prod_4 H_\eta) \subseteq \mathrm{pr}_i(\prod_k H_\eta),$$

where C_2 depends only on d_{0i}'s and L, and in the second containment we used $k \geq 4$.

To see the second claim, we have

$$\begin{aligned} h(\mathrm{pr}_2(\textstyle\prod_k H);\eta) \geq & h(\mathrm{pr}_2(\textstyle\prod_4 H);4\eta) - \log(\mathsf{D}_5) \geq \log\left(\frac{|\mathrm{pr}_2(\prod_4 H_\eta)|}{|1^{(2)}_\eta|}\right) - \log(\mathsf{D}_5\mathsf{D}_4) \\ \geq & C_2 d_{02}\gamma \log(\eta) + d_{02}\log(1/\eta) - \log(\mathsf{D}_{13}), \end{aligned} \tag{137}$$

where D_{13} is a constant as in (113). By Lemma 37 and (137), we obtain that the following holds

$$\begin{aligned} h(H_k^{(1)};\eta) - C_2 d_{02}\gamma\log(1/\eta) + d_{02}\log(1/\eta) \leq & h(H;\eta) + 2kR\gamma\log(1/\eta) + \log(\mathsf{D}_9\mathsf{D}_{13}) \\ \leq & H_2(X;\eta) + (2k+1)R\gamma\log(1/\eta) + \log(\mathsf{D}_9\mathsf{D}_{13}) \\ \leq & \left(d_{01} + d_{02} - \bar{\delta}\right)\log(1/\eta) + (2k+1)R\gamma\log(1/\eta) \\ & + \log(\mathsf{D}_9\mathsf{D}_{13}), \end{aligned} \tag{138}$$

where the second inequality follows from Lemma 33. By (138), we obtain

$$h(H_k^{(1)};\eta) \leq \left(d_{01} - \bar{\delta}\right)\log(1/\eta) + ((2k+1)R + C_2 d_{02})\gamma\log(1/\eta) + \log(\mathsf{D}_9\mathsf{D}_{13}). \tag{139}$$

Therefore, we can choose D_{12} so that for $\eta^\gamma \leq \mathsf{D}_{12}{}^{-1}$ and $\gamma \ll_{k,d_{0i},L} \bar{\delta}$, we have

$$h(H_k^{(1)};\eta) \leq \left(1 - \frac{\bar{\delta}}{2d_{01}}\right) h(G_1;\eta);$$

as we claimed. □

6.3. **Proof of Propositions 29 and 30 modulo a bounded generation result.** In this section we use the following bounded generation result, which is of independent interest, to complete the proofs of Propositions 29 and 30.

Proposition 39. *Suppose F is either $\mathbb{R}$ or $\mathbb{Q}_p$, $\mathbb{G} \subseteq (\mathrm{GL}_{n_0})_F$ is a connected F-almost simple subgroup, and G is a compact open subgroup of $\mathbb{G}(F)$. When $F = \mathbb{R}$, we assume that $G \subseteq O_{n_0}(\mathbb{R})$, and when $F = \mathbb{Q}_p$, we assume that $G \subseteq \mathrm{GL}_{n_0}(\mathbb{Z}_p)$. In either case, we take the metric on G that is induced by the operator norm on $\mathrm{M}_{n_0}(F)$. Let $p_0 = 2$ when $F = \mathbb{R}$ and $p_0 = p$ when $F = \mathbb{Q}_p$. Then* *for every $0<\rho\le p_0^{-2}$, there are $g_1,\ldots,g_{d^2}\in 1_\rho$ where $d:=\dim\mathbb{G}$ and positive numbers $C:=C(G)$ and $c:=c(G)$ such that the following holds. For $h\in 1_{1/4}$ and every $0<r\le c\|h-I\|\rho^C$, we have*

$$\{(g_1[h,a_1]g_1^{-1})\cdots(g_{d^2}[h,a_{d^2}]g_{d^2}^{-1})|\ a_i\in 1_r\}\supseteq 1_{cr\rho^C\|h-I\|},$$

where $[h,a_i]=ha_ih^{-1}a_i^{-1}$. Moreover, if $\mathbb{G}=\widetilde{\mathbb{G}}\otimes_{\mathbb{Q}}\mathbb{Q}_p$ where $\widetilde{\mathbb{G}}$ is an absolutely almost simple $\mathbb{Q}$-group, then the constants C and c depend only on $\widetilde{\mathbb{G}}$.

We postpone the proof of Proposition 39 to §6.4. This proposition is used in the proof of the following corollary, which in turn will play a crucial role in the proofs of Propositions 29 and 30.

Corollary 40. *Let G_1, G_2, X, and X' be as in Lemma 38. Let $k=50d_{01}^2$ and $\bar\delta>0$. Suppose η and γ are positive numbers such that $\gamma\ll_{C(G_i),d_{0i},L}\bar\delta$, where $C(G_i)$'s are given in Proposition 39, and $\eta^\gamma\le \mathsf{D}_{14}{}^{-1}$ where D_{14} is a constant as in (113). Let H be the approximate subgroup given in Lemma 38, applied with $k=50d_{01}^2$ and these $\bar\delta$, η, and γ. Let $\alpha_i:=\alpha_i(\prod_8 H;\eta)$ be as in (124). Then*

$$\alpha_i\ge\frac{\bar\delta}{10d_{0i}},\tag{140}$$

and there is $\eta^{O_{d_{0i},L}(\bar\delta)}$-approximate homomorphism

$$f:1^{(1)}_{\eta^{C_2\gamma}}\to G_2$$

such that $1^{(2)}_{\eta^{C_2'\gamma}}\subseteq(\operatorname{Im}f)_{\eta^{O_{d_{0i},L}(\bar\delta)}}$, where C_2 and C_2' depend only on L, d_{01} and d_{02}.

Proof. We first prove (140); in view of the symmetry, we show this for $i=1$. Recall that By the definition of α_1, see (124), there exists $h\in G_1$ such that $\|h-I\|\ge\eta^{2\alpha_1}$ and $(h,1)\in\prod_{24}H$. Since $1^{(1)}_{\eta^{C_2\gamma}}\subseteq\operatorname{pr}_1(\prod_4 H_\eta)$, by Proposition 39 (applied with $\rho:=\eta^{C_2\gamma}$), we deduce that

$$1^{(1)}_{c\eta^{2CC_2\gamma+4\alpha_1}}\subseteq(H^{(1)}_{50d_{01}^2})_{50d_{01}^2\eta},$$

where $C:=C(G_1)$ is as in Proposition 39 and c is a multiple of $c(G_1)$ given in the same statement. Hence, we obtain

$$\begin{aligned}h(H^{(1)}_{50d_{01}^2};\eta)\ge&\log\left(\frac{|1^{(1)}_{c\eta^{2CC_2\gamma+4\alpha_1}}|}{|1^{(1)}_\eta|}\right)-\log(\mathsf{D}_4)\\ \ge&d_{01}\Big(1-(2CC_2\gamma+4\alpha_1)\Big)\log(1/\eta)-\log(\mathsf{D}_5\mathsf{D}_4)\\ \ge&\Big(1-(3CC_2\gamma+4\alpha_1)\Big)h(G_1;\eta);\end{aligned}\tag{141}$$

notice that we can drop $\log(\mathsf{D}_5\mathsf{D}_4)$ as $\eta^\gamma\le\mathsf{D}_{14}{}^{-1}$ can be chosen small enough. By Lemma 38, applied with $k=50d_{01}^2$, and (141), we obtain the following inequality:

$$3CC_2\gamma+4\alpha_1\ge\frac{\bar\delta}{2d_{01}}.$$

Therefore, for $\gamma\ll_{C,d_{01},L,a}\bar\delta$, we obtain that

$$\alpha_1\ge\frac{\bar\delta}{10d_{01}}.$$

Recalling $1^{(1)}_{\eta^{C_2\gamma}}\subseteq\operatorname{pr}_i(\prod_4 H_\eta)$ and (140) for α_2, the remaining assertions follow from Lemma 36. $\square$

Proof of Proposition 29. Let $G = G_1 \times G_2$, then G is L-locally random with coefficient $C_0 := C_{01} + C_{02}$, see [20, Lemma 5.2]. It also satisfies $\mathrm{DC}(d_0, C_1)$ where $d_0 := d_{01} + d_{02}$ and $C_1 := C_{11}C_{12}$. Suppose

$$0 < \eta < (10C_0 + C_1)^{-1/\bar{\delta}}.$$

Let X be a random variable whose probability law is μ. Let X_i be a 2^i-random walk with respect to μ.

Claim. *For sufficiently small γ (which will be determined later and will depend linearly on $\bar{\delta}$) and for every non-negative integer i at least one of the following holds:*

$$H_2(X_{i+1}; \eta) \geq H_2(X_i; \eta) + \gamma \log(1/\eta), \tag{142a}$$

$$H_2(X_i; \eta) \geq (d_0 - \bar{\delta}) \log(1/\eta). \tag{142b}$$

Proof of the Claim. Let us assume that (142a) and (142b) fail for some i. Then by Corollary 40, there exists $\hat{c}$ depending only on $C(G_i)$ (see Proposition 39), d_{0i}, and L so that if $0 < \gamma < \hat{c}\bar{\delta}$ and $\eta^\gamma \leq \mathsf{D}_{14}{}^{-1}$, where D_{14} is a constant as in (113), then there is an η^β-approximate homomorphism

$$f : 1^{(1)}_{\eta^{C_2\gamma}} \to G_2$$

such that $1^{(2)}_{\eta^{C_2'\gamma}} \subseteq (\operatorname{Im} f)_{\eta^\beta}$, where C_2 and C_2' depend only on L, d_{01} and d_{02}, and $\beta = O_{d_{0i},L}(\bar{\delta})$.

Let m be as in Theorem 3 applied with G_1 and G_2. For small enough γ, we have $\beta/(C_2\gamma) > m$. Since G_1 and G_2 are not locally isomorphic, existence of f contradicts Theorem 3 applied with G_1, G_2 and $\rho = \eta^{C_2\gamma}$ so long as η^γ is small enough. The claim follows. □

Returning to the proof of the proposition, first note that (142a) can hold at most $i_{\max} := \lceil d_0/\gamma \rceil$-many times. Therefore, there exists some $i_0 \leq i_{\max}$ so that (142b) holds. The proof is complete. □

Proof of Proposition 30. Fix some integer $a \geq 4Ld_0$, and let $\bar{\delta} = \frac{1}{8La \log_\eta \eta^{a^2}} = \frac{1}{8La^3}$. Let $0 < \eta < \mathsf{D}_1{}^{-1/\bar{\delta}}$, where D_1 is as in Proposition 29.

Recall that $f \in L^2(G_1 \times G_2)$ lives at scale η if both of the following properties are satisfied

$$\|f_{\eta^{a^2}} - f\|_2 \leq \eta^{a/2}\|f\|_2 \quad \text{and} \quad \|f_{\eta^{1/a}}\|_2 \leq \eta^{1/(2a)}\|f\|_2.$$

Apply Proposition 29 with η^{a^2} and $\bar{\delta}$. In view of Proposition 29, conditions of Lemma 32 are satisfied for G, $\nu = \mu^{(2^m)}$ and $\rho = \eta^{a^2}$. Hence we have

$$\|\mu_\rho^{(2^m)} * f\|_2 \leq \eta^{1/(8La)}\|f\|_2.$$

Now since $\|f_\rho - f\|_2 \leq \eta^{a/2}\|f\|_2$, we conclude that

$$\|\mu^{(2^m)} * f\|_2 \leq \|\mu^{(2^m)} * f_\rho\|_2 + \eta^{a/2}\|f\|_2 = \|\mu_\rho^{(2^m)} * f\|_2 + \eta^{a/2}\|f\|_2 \leq \eta^{1/(16La)}\|f\|_2$$

This implies the proposition with $c = 1/(16La)$ and $m_0 = m$. □

6.4. **Proof of Proposition 39.** In this section we prove Proposition 39. The proof is carried out in several steps, and among other things it relies on certain quantitative inverse function theorems that are proved in the appendix. We start with a lemma which is analogous to [25, Lemma 40] for real numbers.

Lemma 41. *In the setting of Proposition 39, suppose $\{\mathrm{Ad}(g_1), \ldots, \mathrm{Ad}(g_m)\}$ is a basis of the $\mathbb{R}$-subalgebra $\mathbb{R}[\mathrm{Ad}(G)]$ of $\mathrm{End}(\mathfrak{g})$, where $\mathfrak{g} := \mathfrak{g}(\mathbb{R})$ is the Lie algebra of G. Then there is a positive number r_0 depending on g_i's such that for every unit element x of $\mathfrak{g}$,*

$$M(x) := \Big\{ \sum_{i=1}^m c_i \,\mathrm{Ad}(g_i)(x) |\; c_i \in [-1, 1] \Big\} \supseteq 0_{r_0},$$

where 0_{r_0} is the ball of radius r_0 centered at 0 in $\mathfrak{g}$.

Proof. Since $\mathbb{G}$ is $\mathbb{R}$-simple, $\mathfrak{g}$ is a simple G-module. Hence for every unit element x of $\mathfrak{g}$, $M(x)$ contains a neighborhood of 0. Let $r(x)$ be the largest positive number such that $0_{r(x)} \subseteq M(x)$. Suppose to the contrary that there is a sequence $\{x_i\}_{i=1}^{\infty}$ of unit elements of $\mathfrak{g}$ such that $\lim_{i\to\infty} r(x_i) = 0$. By the compactness of the sphere of radius 1 in $\mathfrak{g}$, after passing to a subsequence we can assume that $\{x_i\}_{i=1}^{\infty}$ converges to x, a unit element of $\mathfrak{g}$. For every $y \in 0_{r(x)}$, there are $c_i \in [-1,1]$ such that $\sum_{i=1}^m \mathrm{Ad}(g_i)(x) = y$. For every $\varepsilon > 0$, if $n \gg_\varepsilon 1$, then $\|\mathrm{Ad}(g_i)(x_n) - \mathrm{Ad}(g_i)(x)\| \le \varepsilon$. Therefore,

$$\|y - \sum_{i=1}^m \mathrm{Ad}(g_i)(x_n)\| \le \sum_{i=1}^m |c_i| \|\mathrm{Ad}(g_i)(x) - \mathrm{Ad}(g_i)(x_n)\| \le m\varepsilon.$$

Notice that $M(x_n)$ is a convex set which intersects every $m\varepsilon$-neighborhood of points of $0_{r(x)}$. Therefore for $n \gg_{m,r(x)} 1$, we have $r(x_n) \ge r(x)/2$. This contradicts $\lim_{i\to\infty} r(x_i) = 0$, and the claim follows. □

To formulate the next lemma, we start by recalling that for $g \in \mathrm{O}_{n_0}(\mathbb{R})$ or $g \in \mathrm{GL}_{n_0}(\mathbb{Z}_p)$, if $\|g - I\| < 1$ (if $p = 2$, we assume $\|g - I\| < 1/2$)), then for every $|t| \le 1$ we can define

$$g^t := \exp(t \log(g)).$$

Clearly $t \mapsto g^t$ is an analytic function, and one can see that

$$|t| \ll_g \|g^t - I\| \ll_g |t|. \tag{143}$$

Lemma 42. *In the setting of Proposition 39, suppose $\{\mathrm{Ad}(g_1), \ldots, \mathrm{Ad}(g_m)\}$ is a basis of the $\mathbb{R}$-subalgebra $\mathbb{R}[\mathrm{Ad}(G)]$ of $\mathrm{End}(\mathfrak{g})$, where $\mathfrak{g} := \mathfrak{g}(\mathbb{R})$ is the Lie algebra of G. Suppose $\|g_i - I\| < 1$ for every i. Then there is a positive integer $C := C(g_1, \ldots, g_m)$ and a positive number $c := c(g_1, \ldots, g_m)$ such that for every $0 < t \le c$ we have*

$$\widetilde{M}_t := \Big\{ \sum_{i=1}^m c_i \,\mathrm{Ad}(g_i^t) |\, c_i \in [-1,1] \Big\} \supseteq 0_{t^C},$$

where 0_{t^C} is the ball of radius t^C centered at 0 in $\mathbb{R}[\mathrm{Ad}(G)]$.

Proof. We view $\mathrm{Ad}(g_i^t)$'s as $d^2 \times 1$ column vectors, where $d := \dim G$, and let $A(t)$ be the $d^2 \times m$ matrix that have $\mathrm{Ad}(g_i^t)$ in its i-th column. Consider $f(t) := \det(A(t)^T A(t))$ where $A(t)^T$ is the transpose of $A(t)$. Then f is an analytic function, $f(1) \ne 0$, and $f(0) = 0$. Since f is an analytic function and non-zero, $f^{(\overline{C})}(0) \ne 0$ for some positive integer $\overline{C}$. As f is an analytic function, for $0 < t \le \bar{c}$ we have $\frac{f^{(\overline{C})}(0)}{2\overline{C}!} t^{\overline{C}} \le f(t)$. Since $\|A(t)\|_{\mathrm{op}} = \sqrt{d}$, $\|(A(t)^T A(t))^{-1}\|_{\mathrm{op}} \ll_{\{g_i\}} t^{-\overline{C}}$. Therefore, for every y in $\mathbb{R}[\mathrm{Ad}(G)]$, we have

$$\Big\| \Big(A(t)^T A(t)\Big)^{-1} (A(t)^T) y \Big\| \ll t^{-\overline{C}} \|y\| \quad \text{and} \quad A(t)\Big((A(t)^T A(t))^{-1} (A(t)^T) y)\Big) = y.$$

This implies the claim with $C = \overline{C}/2$ if we assume $0 < t \le c$ and $c \le \bar{c}$ is sufficiently small to account for the implied multiplicative constant above. □

Lemma 43. *In the setting of Proposition 39, for every $0 < \rho \le p_0^{-2}$ (where $p_0 = 2$ when $F = \mathbb{R}$ and $p_0 = p$ when $F = \mathbb{Q}_p$) there are $g_{1,\rho}, \ldots, g_{d^2,\rho} \in 1_\rho$ where $d := \dim \mathbb{G}$ and positive number $C := C(G)$ such that for every non-zero element $x \in \mathfrak{g}(F)$ we have*

$$\Big\{ \sum_{i=1}^{d^2} c_i \,\mathrm{Ad}(g_{i,\rho})(x) |\, c_i \in F, |c_i| \le 1 \Big\} \supseteq 0_{\rho^C \|x\|},$$

where $0_{\rho^C\|x\|}$ is the ball of radius $\rho^C\|x\|$ centered at 0 in $\mathfrak{g}$. Moreover, if $F=\mathbb{Q}_p$ and $\mathbb{G}=\widetilde{\mathbb{G}}\otimes_{\mathbb{Q}}\mathbb{Q}_p$ where $\widetilde{\mathbb{G}}$ is an absolutely almost simple $\mathbb{Q}$-group, then the constant C depends only on $\widetilde{\mathbb{G}}$.

Proof. We start with the p-adic case. Let $l:=\lceil\log_p(1/\rho)\rceil$; then $G[p^l]:=1_\rho$ is the kernel of the residue map modulo p^l. Choose $\{g_{1,\rho},\ldots,g_{d^2,\rho}\}\subseteq G[p^l]$ such that the $\mathbb{Z}_p$-linear span of $\mathrm{Ad}(g_{i,\rho})$'s is the $\mathbb{Z}_p$-algebra $\mathbb{Z}_p[\mathrm{Ad}(G[p^l])]$. Then by [25, Proposition 44'], there is a positive number $\overline{C}$ which depends on $\mathbb{G}$ (and depends only on $\widetilde{\mathbb{G}}$ if $\mathbb{G}=\widetilde{\mathbb{G}}\otimes_{\mathbb{Q}}\mathbb{Q}_p$) such that

$$\sum_{i=1}^{d^2}\mathbb{Z}_p\,\mathrm{Ad}(g_{i,\rho})\supseteq p^{\overline{C}l}\mathbb{Z}_p[\mathrm{Ad}(G[1])] \tag{144}$$

where $G[1]$ is the ball of radius 1 in $\mathbb{G}(\mathbb{Q}_p)$. By (144), for every $x\in\mathfrak{g}$ we have

$$\sum_{i=1}^{d^2}\mathbb{Z}_p\,\mathrm{Ad}(g_{i,\rho})(x)\supseteq p^{\overline{C}l}\mathbb{Z}_p[\mathrm{Ad}(G[1])]x. \tag{145}$$

On the other hand, $\mathfrak{g}(\mathbb{Q}_p)$ is a simple $\mathbb{Q}_p[\mathrm{Ad}(G[1])]$-module. Hence by [25, Lemma 40], there is a positive number $\overline{C}'$ (depending only on $\widetilde{\mathbb{G}}$ if $\mathbb{G}=\widetilde{\mathbb{G}}\otimes_{\mathbb{Q}}\mathbb{Q}_p$) such that

$$\mathbb{Z}_p[\mathrm{Ad}(G[1])]x\supseteq p^{\overline{C}'}\|x\|^{-1}(\mathfrak{g}\cap\mathrm{M}_{n_0}(\mathbb{Z}_p)). \tag{146}$$

By (145) and (146), we deduce that

$$\Big\{\sum_{i=1}^{d^2}c_i\,\mathrm{Ad}(g_{i,\rho})(x)|\; c_i\in F, |c_i|\le 1\Big\}\supseteq 0_{p^{-\overline{C}\,\overline{C}'l}\|x\|},$$

and the p-adic case follows.

Suppose $\{g_1,\ldots,g_m\}$ is as in Lemma 42. Notice that

$$\Big\{\sum_{i=1}^{m}c_i\,\mathrm{Ad}(g_i)|\; c_i\in[-1,1]\Big\}\subseteq 0_m,$$

where 0 is the zero of $\mathbb{R}[\mathrm{Ad}(G)]$. Now by Lemma 42, we have

$$\Big\{\sum_{i=1}^{m}c_i\,\mathrm{Ad}(g_i^t)|\; c_i\in[-1,1]\Big\}\supseteq 0_{t^C}\supseteq\frac{t^C}{m}\Big\{\sum_{i=1}^{m}c_i\,\mathrm{Ad}(g_i)|\; c_i\in[-1,1]\Big\},$$

for some positive numbers C. Combining this and Lemma 41, we have

$$\Big\{\sum_{i=1}^{m}c_i\,\mathrm{Ad}(g_i^t)(x)|\; c_i\in[-1,1]\Big\}\supseteq 0_{r_0t^C\|x\|/m},$$

where r_0 is the constant appearing in Lemma 41. The real case follows. □

Proof of Proposition 39. Note that

$$\mathrm{ad}:\mathrm{Lie}(\mathbb{G})\to\mathrm{Lie}(\mathbb{G})$$

is an F-isomorphism. This implies that, for every $x\in\mathrm{Lie}(\mathbb{G})(F)$,

$$\|x\|_{\mathrm{op}}\ll\|\,\mathrm{ad}(x)\|_{\mathrm{op}}\ll\|x\|_{\mathrm{op}}$$

where the implied constant depends only on $\mathbb{G}(F)$.

Moreover, if $\mathbb{G}=\widetilde{\mathbb{G}}\otimes_{\mathbb{Q}}\mathbb{Q}_p$, then ad is induced from a $\mathbb{Q}$-isomorphism of $\mathrm{Lie}(\widetilde{\mathbb{G}})$. Therefore, in this case the implied constants equal 1 if p is large enough depending only on $\widetilde{\mathbb{G}}$ — indeed, note that $(\mathrm{ad})^{-1}$ is a $\mathbb{Q}$-isomorphism, and so for p large enough (depending only on $\widetilde{\mathbb{G}}$) ad is an isometry.

We also notice that $\log : 1_{p_0^{-2}} \to \mathfrak{g}$ is a bi-Lipschitz map where $p_0 = 2$ if $F = \mathbb{R}$ and $p_0 = p$ if $F = \mathbb{Q}_p$. Indeed if $F = \mathbb{Q}_p$, then log is an isometry between $1_{p^{-2}}$ and $0_{p^{-2}}$. Thus

$$\|\operatorname{Ad}(h) - I\|_{\mathrm{op}} \gg \|\log(\operatorname{Ad}(h))\|_{\mathrm{op}} \gg \|\operatorname{ad}(\log h)\|_{\mathrm{op}} \gg \|\log h\|_{\mathrm{op}} \gg \|h - I\|_{\mathrm{op}},$$

for all $h \in 1_{p_0^{-2}}$. If $\mathbb{G} = \widetilde{\mathbb{G}} \otimes_{\mathbb{Q}} \mathbb{Q}_p$ and p is large enough, $\gg$ may be replaced with $=$ in the above.

We conclude that

$$\|(\operatorname{Ad}(h) - I)(x)\| \geq c'\|h - I\| \tag{147}$$

for a unit vector $x \in \mathfrak{g}$ and positive number $c' = c'(G)$ which depends only on $\widetilde{\mathbb{G}}$ if $\mathbb{G} = \widetilde{\mathbb{G}} \otimes_{\mathbb{Q}} \mathbb{Q}_p$.

Recall that $0 < \rho \leq p_0^{-2}$. Let $I := \{a \in F |\, |a| < 1/4\}$, and let $\{g_{1,\rho}, \ldots, g_{d^2,\rho}\}$ be given by Lemma 43. Set

$$\Phi : I \times \cdots \times I \to G,\ \Phi(t_1, \ldots, t_{d^2}) := g_{1,\rho}[h, \exp(t_1 x)]g_{1,\rho}^{-1} \cdots g_{d^2,\rho}[h, \exp(t_{d^2} x)]g_{d^2,\rho}^{-1}.$$

Then $\mathrm{d}\Phi(0) : F^{d^2} \to \mathfrak{g}$ is given by

$$\mathrm{d}\Phi(0)(c_1, \ldots, c_{d^2}) = \sum_{i=1}^{d^2} c_i \operatorname{Ad}(g_{i,\rho})(\operatorname{Ad}(h) - I)x.$$

Hence by Lemma 43 and the choice of x, as in (147), we obtain that

$$\sigma(\mathrm{d}\Phi(0)) \geq \frac{c'}{d}\|h - I\|\rho^C, \tag{148}$$

where σ of a matrix A is given by $\sup\{r \in [0, \infty)|\, 0_r \subseteq A0_1\}$ and C is a positive integer which depends on G. Writing the Taylor expansion of the exponential function, we obtain that in the p-adic case all the coefficients are p-adic integers and in the real case, for every $1 \leq j, j' \leq d^2$ and $\mathbf{x} \in I \times \cdots \times I$, $\|\partial_{j,j'}\Phi(\mathbf{x})\| \leq d^{O(1)}$. Therefore, by Theorem 56 and Theorem 57, for every $0 < r \leq c''\|h - I\|\rho^C$ (where $c'' = c'd^{-O(1)}$), we have

$$1_{\frac{c'}{4d}r\|h-I\|\rho^C} \subseteq \Phi(0_r).$$

This implies the claim with $c = \min\{\frac{c'}{4d}, c''\}$. □

7. Proof of Theorem 1

In this section, we will complete the proof of Theorem 1. We begin by recalling [20, Theorem 9.3], which will be used both in this section and §9 below.

Theorem 44. *Suppose G is L-locally random with coefficients C_0 which satisfies the dimension condition $\mathrm{DC}(C_1, d_0)$, see (DC). Let μ be a symmetric Borel probability measure on G, and the group generated by the support of μ is dense in G. Suppose that there exist $C_3 > 0$, $b > 0$, and $0 < \eta_0 < 1$ such that for every $\eta \leq \eta_0$ and every function $g \in L^2(G)$ which lives at scale η there exists $l \leq C_3 \log(1/\eta)$ such that*

$$\|\mu^{(l)} * g\|_2 \leq \eta^b \|g\|_2.$$

Then there is a subrepresentation $\mathcal{H}_0$ of $L^2(G)$ with $\dim \mathcal{H}_0 \leq 2C_0\eta_0^{-d_0}$ such that

$$\mathcal{L}(\mu; L^2(G) \ominus \mathcal{H}_0) \geq \frac{b}{C_3}.$$

In particular, $\mathcal{L}(\mu; G) > 0$.

In addition to Theorem 44, the proof also relies on results in [20, §5], as well as Theorem 27 and Proposition 30 in this paper.

Proof of Theorem 1. The proof will be completed in some steps. We will write $F_i = \mathbb{Q}_{\nu_i}$. Recall that G_i is a compact open subgroup of $\mathbb{G}_i(\mathbb{Q}_{\nu_i})$ where for $i = 1, 2$, $\mathbb{G}_i$ is $\mathbb{Q}_{\nu_i}$-simple group. Recall also our notation $p_\nu = \nu$ if ν is non-Archimedean and $p_\infty = 2$.

Let $X = (X_1, X_2)$ be as in the statement, and let μ be the probability law of X. Since $\min\{\mathcal{L}(X_1), \mathcal{L}(X_2)\} \geq c_0 > 0$, we have

$$\max\{\lambda(\pi_{1*}\mu; G_1), \lambda(\pi_{2*}\mu; G_2)\} =: \lambda < 1 \tag{149}$$

where π_i denotes the projection onto the i-th factor.

Reduction to functions living at certain scale. There exist L, C_0, C_1 and d_{0i} for $i = 1, 2$ so that if we let $\rho_0 \ll 1$, then all of the following properties are satisfied:

(G-1) G_i is L-locally random with coefficient C_0 for $i = 1, 2$.

(G-2) For all $\hat\eta = \rho_0^j$, $j \in \mathbb{N}$, the group G_i satisfies
$$\frac{1}{C_1}\hat\eta^{d_{0i}} \leq |1_{\hat\eta}^{(i)}| \leq C_1\hat\eta^{d_{0i}}, \qquad \text{for } i = 1, 2.$$

(G-3) For $i = 1, 2$, G_i is δ-discretizable for all $\delta = \rho_0^j$, and all $j \in \mathbb{N}$ in the p-adic case and sufficiently large $j \in \mathbb{N}$ in the real case.

See [20, §5] for the first statement, the second assertion holds in view of the choice of ρ_0, and the third property is satisfied by Proposition 26. We note that L and d_{01}, d_{02} depend only on $\dim \mathbb{G}_i$, C_0 depends on G_1 and G_2, and C_1 depends on $\max\{p_{\nu_1}, p_{\nu_2}\}$.

Let m_i denote the probability Haar measure on G_i for $i = 1, 2$, and let $m = m_1 \times m_2$ denote the probability Haar measure on G.

From (G-1) and (G-2), we conclude that $G = G_1 \times G_2$ is L-locally random with coefficient $2C_0$, see [20, Lemma 5.2]. The group G also satisfies $\mathrm{DC}(d_0, C_1)$ where $d_0 = d_{01} + d_{02}$ and $C_1 = C_{11}C_{12}$.

Fix some integer $a \geq 4Ld_0$. Let $\eta = \rho_0^j$ for some $j \geq j_0$; the parameter j_0 will be explicated later. Recall that $f \in L^2(G, m)$ lives at scale η if both of the following properties are satisfied

$$\|f_{\eta^{a^2}} - f\|_2 \leq \eta^{a/2}\|f\|_2, \quad \text{and} \quad \|f_{\eta^{1/a}}\|_2 \leq \eta^{1/(2a)}\|f\|_2. \tag{150}$$

We claim

Claim 45. *There exists $\ell \leq \bar{C}\log(1/\eta)$ where $\bar{C}$ depends on d_{0i}, L so that*

$$\|\mu^{(\ell)} * f\|_2 \leq \eta^{c/2}\|f\|_2 \tag{151}$$

where b depends only on L and d_{01}, d_{02}

Note that in view of Theorem 44, (151) finishes the proof. Thus it remains to prove the claim.

Proof of the Claim. The proof relies on Theorem 27 and Proposition 30 as we now explicate.

Reduction to couplings of Haar measures. Let $\eta = \rho_0^j$ for some $j \geq j_0$. Properties (G-1), (G-2), (G-3), together with (149) imply that Theorem 27 is applicable with G_1, G_2, μ, and $\rho = \eta^{a^2}$. In view of that theorem thus there exists a symmetric coupling $\sigma = \sigma^\rho$ of m_1 and m_2 so that the following holds. Let $f \in L^2(G, m)$ satisfy that $\|f_\rho - f\|_2 \leq \rho^{1/(4a)}\|f\|_2$; then

$$\left\|\mu^{(\ell_1)} * f - \sigma * f\right\|_2 \leq 6\rho^{1/(4a)}\|f\|_2, \tag{152}$$

so long as $\ell_1 \gg_{a, d_{0i}, L} \log_\lambda(\rho/C_1)$, see (101).

Conclusion of the proof. Let us write $\sigma = \sigma^\rho$. Let j_0 be large enough so that $\eta \le \eta_0$ where η_0 is as in Proposition 30, in particular, $\eta_0 = \max\{p_{\nu_1}, p_{\nu_2}\}^{-O_{\dim G}(1)}$. Then by Proposition 30, for every f which lives at scale η we have

$$\|\sigma^{(2^{n_0})} * f\|_2 \le \eta^c \|f\|_2 \tag{153}$$

where c and n_0 depend only on L and d_{01}, d_{02}. Without loss of generality, we assume $c < 1/(4a)$.

Let $0 < n_1 \le 2^{n_0}$ be the smallest integer so that

$$\|\sigma^{(n_1)} * f\|_2 \le \rho^{1/4a}\|f\|_2$$

if such exists, otherwise let $n_1 = 2^{n_0}$. Now for all $0 \le i < n_1$, we have $\rho^{-1/4a}\|\sigma^{(i)} * f\|_2 \ge \|f\|_2$; hence, using (5), we have

$$\|\sigma^{(i)} * f_\rho - \sigma^{(i)} * f\|_2 \le \|f_\rho - f\|_2 \le \rho^{1/2a}\|f\|_2 \le \rho^{1/4a}\|\sigma^i * f\|_2,$$

where we also used $\sigma^{(i)} * f_\rho = (\sigma^{(i)} * f)_\rho$ and the first estimate in (150). Therefore by (152), we deduce that

$$\|\mu^{(\ell_1)} * \sigma^{(i)} * f - \sigma^{(i+1)} * f\|_2 \le 6\rho^{1/(4a)}\|f\|_2,$$

for every $i < n_1$. Using the triangle inequality and (5), we get

$$\begin{aligned}\|\mu^{(\ell_1 n_1)} * f - \sigma^{(n_1)} * f\|_2 \le& \sum_{i=0}^{n_1-1} \|\mu^{(\ell_1(n_1-i))} * \sigma^{(i)} * f - \mu^{(\ell_1(n_1-i-1))} * \sigma^{(i+1)} * f\|_2 \\ =& \sum_{i=0}^{n_1-1} \|\mu^{(\ell_1(n_1-i-1))} * (\mu^{(\ell_1)} * \sigma^{(i)} * f - \sigma^{(i+1)} * f)\|_2 \\ \le& 6n_1\rho^{1/(4a)}\|f\|_2.\end{aligned}$$

By (153) and the choice of n_1, we have that $\|\sigma^{n_1} * f\|_2 \le \eta^c\|f\|_2$. Hence,

$$\|\mu^{(\ell_1 n_1)} * f\|_2 \le \eta^{c/2}\|f\|_2.$$

This implies that (151) holds with $b = c/2$ and $\ell = \ell_1 n_1 \le \bar{C}\log(1/\eta)$ where $\bar{C}$ depends on d_{0i}, L, and a. This completes the proof of the claim, and hence the proof of the theorem. □

8. Spectral independence and exceptional representations

The objective of this section is to prove Proposition 46, which will be used in the proof of Theorem 2. To obtain Theorem 2, we apply Theorem 1 to the group Γ_{ν_1,ν_2} where $\nu_i \in V_\Gamma$; see the notation in Theorem 2. Then using Claim 45 in the proof of Theorem 1 and Theorem 44, we reduce the analysis to the study of *exceptional representations* of Γ_{ν_1,ν_2} whose dimension is $\ll_\Gamma (\max\{p_{\nu_1}, p_{\nu_2}\})^{O(1)}$. The proof of Proposition 46 relies on the results in [24, 25] and Proposition 39 in this paper.

We begin by recalling some of the notation from Theorem 2. Let $\mathbb{G}$ be an absolutely almost simple, connected, simply connected $\mathbb{Q}$-algebraic group. As it was done before, we fix a $\mathbb{Q}$-embedding $\mathbb{G} \subseteq (\mathrm{SL}_N)_\mathbb{Q}$, for some N. Let $\Omega \subseteq \mathbb{G}(\mathbb{Q})$ be a finite symmetric subset such that $\Gamma = \langle\Omega\rangle$ is Zariski dense in $\mathbb{G}$. Denote by V_Γ the set of all places ν of $\mathbb{Q}$ such that Γ is a bounded subset of $\mathbb{G}(\mathbb{Q}_\nu)$, and let $V_{\mathrm{f},\Gamma} \subseteq V_\Gamma$ denote the subset of finite places.

For distinct $\nu_1, \nu_2 \in V_\Gamma$, let Γ_{ν_1,ν_2} denote the closure of Γ in $\mathbb{G}(\mathbb{Q}_{\nu_1}) \times \mathbb{G}(\mathbb{Q}_{\nu_2})$. The following are our standing assumption in this section:

$$\begin{aligned}&\Gamma_\nu \subseteq \mathrm{SL}_N(\mathbb{Z}_{p_\nu}) &&\text{for all } \nu \in V_{\mathrm{f},\Gamma}, \text{ and}\\ &\Gamma_{\nu_1,\nu_2} = \Gamma_{\nu_1} \times \Gamma_{\nu_2}, &&\text{for all } \nu_1, \nu_2 \in V_\Gamma,\end{aligned}$$

where Γ_ν denotes the closure of Γ in $\mathbb{G}(\mathbb{Q}_\nu)$.

Let $W_\Gamma \subset V_{\mathrm{f},\Gamma}$ denote the set of places where $\mathbb{G}(\mathbb{Q}_\nu)\cap \mathrm{SL}_N(\mathbb{Z}_{p_\nu})$ is a hyperspecial subgroup of $\mathbb{G}(\mathbb{Q}_{p_\nu})$ and

$$\Gamma_\nu = \mathbb{G}(\mathbb{Q}_{p_\nu})\cap \mathrm{SL}_N(\mathbb{Z}_{p_\nu}).$$

For every $n\geq 0$, let $\pi_{\nu,n} : \mathrm{SL}_N(\mathbb{Z}_{p_\nu}) \to \mathrm{SL}_N(\mathbb{Z}/p_\nu^n\mathbb{Z})$. For every $\nu\in V_{\mathrm{f},\Gamma}$, let

$$\Gamma_{\nu,n} = \Gamma_\nu \cap \ker\big(\pi_{\nu,n}\big)$$

denote the n-th congruence subgroup of Γ_ν.

If $\infty\in V_\Gamma$, we put $\Gamma_{\infty,n}=\{1\}$ for all $n\in\mathbb{N}$ — recall also our convention $p_\infty=2$.

Proposition 46. *For all $D,d>0$, there exists $\varrho=\varrho(\Omega,d,D)>0$ so that the following holds. For $i=1,2$, let $\nu_i\in V_\Gamma$ and $n_i\in\mathbb{N}$. Put $G_i=\Gamma_{\nu_i}/\Gamma_{\nu_i,n_i}$ and assume*

$$\text{the number of connected components of } G_1\times G_2 \leq D\max\{p_{\nu_1},p_{\nu_2}\}^d. \tag{154}$$

Then $\mathcal{L}(X;G_1\times G_2)>\varrho$, where X is a random variable with the uniform distribution on Ω.

The following theorem is a special case of [25, Theorem 24], applied with $\mathcal{C}=p_\nu^n$ for $\nu\in V_{\mathrm{f},\Gamma}$, and will play a key role in the proof of Proposition 46.

Theorem 47. *There exist positive numbers $\bar{E}=\bar{E}(\mathbb{G})$, $R_1=R_1(\Omega)$, and $\varepsilon_0=\varepsilon_0(\Omega)$, such that for every $0<\varepsilon\leq\varepsilon_0(\Omega)$, $0<\delta\ll_{\Omega,\varepsilon}1$ and $R_2\gg_{\Omega,\varepsilon}1$ the following holds. Let $n\in\mathbb{N}$ and assume $n\varepsilon^{\bar{E}}\log(p_\nu)\geq\log(R_1)$. Suppose for a finite symmetric subset $A\subseteq\Gamma$ and some $\ell\geq(n\log p_\nu)/\delta$, we have*

$$\mathcal{P}_\Omega^{(\ell)}(A)\geq p_\nu^{-\delta n},$$

then there exists a non-negative integer $n'\leq\varepsilon n$ so that

$$\pi_{\nu,n}(\Gamma_{\nu,n'})\subseteq \textstyle\prod_{R_2}\pi_{\nu,n}(A).$$

Let us now return to the proof of Proposition 46. Note that increasing D and d makes the assumption weaker; therefore throughout the argument, we may assume $D,d>100$.

8.1. **Both places are finite.** We will first consider the case

$$\nu_1,\nu_2\in V_\Gamma \qquad \text{are finite places.} \tag{155}$$

Let us write $G=G_1\times G_2$, and put $M=\max\{|G_1|,|G_2|\}$. As before, m_{G_i} and m_G denote the uniform measures on G_i and G, respectively. To ensure consistency with most of the existing literature, in this case where G is finite, we deviate from the notation in the rest of the paper and define the convolution of $f_1,f_2\in\ell^2(G)$ using the counting measure. That is:

$$f_1 * f_2(x)=\sum_{y\in G} f_1(xy^{-1})f_2(y).$$

Reduction to deep levels verses large primes. Since almost every ν belongs to W_Γ, in the proof of Proposition 46 we may assume that $\{\nu_1,\nu_2\}\cap W_\Gamma\neq\emptyset$. Thus, for the rest of the proof we will assume that $\max\{p_{\nu_1},p_{\nu_2}\}=p_{\nu_1}\geq D$ and that $\nu_1\in W_\Gamma$. Note that

$$p_{\nu_1}^{n_1}\leq|G_1\times G_2|\leq Dp_{\nu_1}^d\leq p_{\nu_1}^{d+1}. \tag{156}$$

In consequence, we will assume for the rest of the argument that

$$n_1\leq d+1. \tag{157}$$

Recall from Lemma 4 that Γ_{ν_1,ν_2} is L-locally random with coefficient C_0 depending only on Γ. This in particular implies that for every $\rho\in\hat{G}$, we have

$$\dim(\rho)\geq c_\Gamma[G:\ker\rho]^{\alpha_\Gamma}, \qquad \text{for some } c_\Gamma,\alpha_\Gamma>0. \tag{158}$$

We will assume, for the rest of the argument, that

$$\max\{n_1, n_2\} > 8d^2(d+1)E_2/(\varepsilon_0(\Omega)\alpha_\Gamma), \tag{159}$$

where d is as in Proposition 46, $\varepsilon_0(\Omega)$ is as in Theorem 47, and $E_2 = 2C/c$ with C and c as in Proposition 39, see (180) below — as it will be explicated later, if (159) fails, Proposition 46 follows from [25, Theorem 1].

Equidistribution of marginals. Let $X = (X_1, X_2)$ be as in the statement of Proposition 46, and let μ be the probability law of X. Since $\langle\Omega\rangle = \Gamma$ is Zariski dense in $\mathbb{G}$, which is absolutely almost simple, we have $\min\{\mathcal{L}(X_1), \mathcal{L}(X_2)\} \geq c_0 > 0$ where $c_0 = c_0(\Omega)$, see [24, Theorem 1] and references therein. Thus

$$\max\{\lambda(\pi_{1*}\mu; G_1), \lambda(\pi_{2*}\mu; G_2)\} \leq \lambda < 1 \tag{160}$$

where π_i denotes the projection onto the i-th factor and $\lambda := \lambda(\Omega)$. Thus, there exists $E_1 \geq 1$ so that for all $\ell_1 \geq E_1 \log M$, we have

$$\|\pi_{i*}\mu^{(\ell_1)} - m_{G_i}\|_2 \leq |G|^{-1}; \tag{161}$$

this can be seen, e.g., by applying Lemma 28 with $\eta = 1/M$ and choosing E_1 so that $\lambda^{\ell_1/2}M^2 \leq 1$ for all $\ell_1 \geq E_1 \log M$.

Let us fix one such ℓ_1, and put $\sigma := \mu^{(\ell_1)}$. The goal now is to show the following

Lemma 48. *With the above notation, there exists $m_0 = m_0(\Omega) \in \mathbb{N}$ so that*

$$\|\sigma^{(2^{m_0})} - m_G\|_2 \ll_\Omega |G|^{-1}.$$

The argument is similar to the one we used in the proof of Propositions 29 and 30; we now turn to the details.

Proof. The proof will be completed in some steps. For every integer $m \geq 0$, let $\sigma_m = \sigma^{(2^m)}$.

Quasi randomness and its consequences. Recall from (158) that

$$\dim(\rho) \geq c_\Gamma [G : \ker\rho]^{\alpha_\Gamma} \qquad \text{for any } \rho \in \hat{G}. \tag{162}$$

Let ρ be an irreducible component of $\ell^2(G)$. If $[G : \ker\rho] < p_{\nu_1}$, then $G_1 \subseteq \ker\rho$. That is: ρ may be identified as a representation of $G/G_1 \simeq G_2$ and the claim in the lemma follows from (161). Therefore, in the proof of the lemma we may restrict to representations ρ so that

$$[G : \ker\rho] \geq p_{\nu_1} \geq |G|^{1/(d+1)}, \tag{163}$$

where in the last inequality we used (156). In view of (162), thus, it suffices to consider

$$\dim\rho \geq c_\Gamma |G|^{\alpha_\Gamma/(d+1)}. \tag{164}$$

Now if $f \in (\ell^2(G))_\rho$ (the ρ-isotypic submodule of $\ell^2(G)$) for some ρ satisfying (164), then

$$\|\chi * f\|_2 \ll_\Gamma |G|^{-\alpha_\Gamma/(2d+2)}|G|^{1/2}\|\chi\|_2\|f\|_2, \quad \text{for any } \chi \in \ell^2(G), \tag{165}$$

see [20, Lemma 6.1] and references therein, note also that in this proof convolution operator is defined using the counting measure.

Thus in order to prove the lemma, it suffices to show that

$$\|\sigma_{m_1}\|_2^2 \leq |G|^{-1+\beta} \tag{166}$$

for some $0 < \beta < \alpha := \alpha_\Gamma/(2d+2)$ and some $m_1 = m_1(\Omega)$.

Initial entropy and its consequences. In view of (161),

$$\sigma_m((g_1,g_2)) \le \min\{\pi_{1*}\sigma(g_1), \pi_{2*}\sigma(g_2)\} \le 2M^{-1}$$

for all $m \ge 0$ and all $(g_1,g_2)\in G$. Since σ_m is a probability measure, we conclude that

$$\|\sigma_m\|_2^2 \le 2M^{-1} \qquad \text{for all } m\ge 0; \tag{167}$$

see also Lemma 31 with $\eta = 1/|G|$.

In view of (167), (166) holds unless

$$|G_i| \ge DM^{\alpha/4} \qquad \text{for } i=1,2, \tag{168}$$

where D is as in the statement of proposition. Therefore, for the rest of the argument we will assume that (168) holds.

Let d be as in the statement of the proposition. For every $0<\varepsilon<1$, put

$$\mathrm{Exc}(\varepsilon) := \{(\nu,n)\in V_{\mathrm{f},\Gamma}\times\mathbb{N} : p_\nu^n \le R_1^{1/\varepsilon^{\bar E}}\},$$

where $\bar E$ and R_1 are as in Theorem 47.

Let

$$\varepsilon = \varepsilon_0(\Omega)\alpha/(8dE_2). \tag{169}$$

Note that in the proof of Proposition 46, we are allowed to assume

$$(\alpha\log M)/4 > (\log R_1)/\varepsilon^{\bar E}. \tag{170}$$

Since $\mathrm{Exc}(\varepsilon)$ is a finite set, (170) and (168) imply that we may assume $(\nu_1,n_1),(\nu_2,n_2)\notin \mathrm{Exc}(\varepsilon)$ for the rest of the proof.

Let δ and R_2 be as in Theorem 47 applied with ε, and let m_2 be so that $2^{m_2}E_1 > 1/\delta$. Then

$$2^{m_2}\ell_1 \ge 2^{m_2}E_1\log M \ge (\log M)/\delta. \tag{171}$$

ℓ^2-Flattening lemma. Let $K=M^\kappa$ for some $\kappa>0$ which will be explicated momentarily. Assume now that for some $m\ge m_2$, we have

$$\|\sigma_m * \sigma_m\|_2 > K^{-1}\|\sigma_m\|_2. \tag{172}$$

Then, there exists $H\subseteq G$ which is symmetric and contains 1, so that all of the following hold:

(H-1) $K^{-E}\|\sigma_m\|_2^{-2} \le |H| \le K^E\|\sigma_m\|_2^{-2}$,

(H-2) $|H\cdot H\cdot H| \le K^E|H|$, and

(H-3) $\sigma_m*\sigma_m(H) \ge K^{-E}$,

where E is absolute. See, e.g., [20, Theorem 2.12] and references therein.

We will apply the above with $\kappa = \alpha_\Gamma(\min\{\varepsilon,\delta\})/(10R_2E\dim\mathbb{G})$. This in particular implies

$$K^{-E} = M^{-E\kappa} \ge M^{-\delta\alpha/4} \ge p_{\nu_i}^{-\delta n_i}, \tag{173}$$

where in the last inequality we used (168) and $|G_i| = p_{\nu_i}^{n_i}$.

Bounded generation and the structure of H. Let ε be as in (169), and let $A \subseteq \Gamma$ be a finite symmetric lift of H. Then by (H-3) and (173), we have

$$\mathcal{P}_\Omega^{(2^{m+1}\ell_1)}(A) \geq K^{-E} \geq p_{\nu_i}^{-\delta n_i} \quad \text{for } i = 1, 2. \tag{174}$$

Apply Theorem 47 with ε and A. Since $\pi_{\nu_i,n_i}(A) = \pi_i(H)$, (174) and Theorem 47 imply that for $i = 1, 2$, there exists $0 \leq n_i' \leq n_i\varepsilon$ so that if we put $H_1 = \prod_{R_2} H$, then

$$\pi_{\nu_i,n_i}(\Gamma_{\nu_i,n_i'}) \subseteq \pi_i(H_1), \qquad \text{for } i = 1, 2. \tag{175}$$

Recall from (157) that $n_1 \leq d+1$. Hence, $\varepsilon n_1 \leq \varepsilon(d+1) < 1$, which implies $n_1' = 0$. We also record for later use that $n_1 \leq d+1$ and (159) imply

$$\varepsilon n_2 > 2. \tag{176}$$

A homomorphism from G_1 into G_2. In view of (175) and since $n_1' = 0$, we have

$$G_1 = \pi_1(H_1) \qquad \text{and} \qquad \pi_{\nu_2,n_2}(\Gamma_{\nu_2,n_2'}) \subseteq \pi_2(H_1), \tag{177}$$

where $0 \leq n_2' \leq \varepsilon n_2$.

For every $g_2 \in G_2$, put $\ell(g_2) := \max\{\ell \geq 0 : g_2 \in \ker(\pi_{\nu_2,\ell})\}$. Define a function $s : G_1 \to G_2$ by

$$(g_1, s(g_1)) \in H_1 \quad \text{and} \quad \ell(s(g_1)) = \min\{\ell(g_2) : (g_1, g_2) \in H_1\}.$$

Define ℓ_s by $\mathrm{Im}(s(G_1)) \subseteq \pi_{\nu_2,n_2}(\Gamma_{\nu_2,\ell_s})$ but $\mathrm{Im}(s(G_1)) \not\subseteq \pi_{\nu_2,n_2}(\Gamma_{\nu_2,\ell_s+1})$. In view of the definition of s and (177), we have

$$\ell_s \leq n_2' \leq \varepsilon n_2. \tag{178}$$

Let us also define $k_s = \min\{k_s', k_s''\}$ where

$$k_s' = \min\big\{\ell\big(s(g_1)s(g_1^{-1})\big) : g_1 \in G_1\big\}, \quad \text{and}$$
$$k_s'' = \min\big\{\ell\big(s(g_1g_1')s(g_1')^{-1}s(g_1)^{-1}\big) : g_1, g_1' \in G_1\big\}.$$

Then s induces a homomorphism $f : G_1 \to G_2/\pi_{\nu_2,n}(\Gamma_{\nu_2,k_s})$. Recall now that $p_{\nu_1} > p_{\nu_2}$, $n_1 \leq d+1$, and Γ_{ν_1} is a hyperspecial subgroup. Thus $G_1/\Gamma_{\nu_1,1}$ does not arise as a composition factor of any subgroup of $G_2/\pi_{\nu_2,n}(\Gamma_{\nu_2,k_s})$. We conclude that f is trivial, and hence

$$k_s \leq \ell_s.$$

Large fibers for H_1. Note that

$$\big(1, s(g_1)s(g_1^{-1})\big) \in H_1 \cdot H_1 \quad \text{and} \quad \big(1, s(g_1g_1')s(g_1')^{-1}s(g_1)^{-1}\big) \in H_1 \cdot H_1 \cdot H_1$$

for all $g_1, g_1' \in G_1$. This and $k_s \leq \ell_s \leq \varepsilon n_2$, see (178), imply the following:

$$\text{There exists } h \in G_2 \text{ with } \ell(h) \leq \varepsilon n_2 \text{ so that } (1, h) \in H_1 \cdot H_1 \cdot H_1 \tag{179}$$

Apply Proposition 39 with the group Γ_{ν_2}, h as in (179), and $\rho = p_{\nu_2}^{-\varepsilon n_2}$ (recall from (176) that $\varepsilon n_2 > 2$). Then by Proposition 39 combined with (179) and (177) (for ν_2), there exists E_2 (depending on $\mathbb{G}$) so that if we put $t = E_2\varepsilon n_2$, then

$$\{(1, g_2) : g_2 \in \pi_{\nu_2,n_2}(\Gamma_{\nu_2,t})\} \subset \prod_{7\dim\mathbb{G}} H_1. \tag{180}$$

The conclusion of the proof. Recall now from (177) that $G_1 = \pi_1(H_1)$. This and (180) imply

$$\Big|\prod\nolimits_{R_2(1+7\dim\mathbb{G})} H\Big| \geq |G_1||\pi_{\nu_2,n_2}(\Gamma_{\nu_2,t})| = |G_1||G_2|^{1-E_2\varepsilon}, \tag{181}$$

where we also used $H_1 = \prod_{R_2} H$.

In view of (H-1) and (H-2), we have

$$K^{E+R_2E(1+7\dim\mathbb{G})}\|\sigma_m\|_2^{-2} \geq K^{R_2E(1+7\dim\mathbb{G})}|H| \geq \Big|\prod\nolimits_{R_2(1+7\dim\mathbb{G})} H\Big|$$

Combining this with (181), we get

$$K^{E+R_2E(1+7\dim\mathbb{G})}\|\sigma_m\|_2^{-2} \geq |G_1||G_2|^{1-E_2\varepsilon}.$$

Recall that $K = \max\{|G_1|, |G_2|\}^\kappa$ where $\kappa = (\alpha\min\{\varepsilon,\delta\})/(10R_2E\dim\mathbb{G})$. In consequence,

$$\|\sigma_m\|_2^{-2} \geq |G_1||G_2|^{1-E_2\varepsilon}K^{-8R_2E\dim\mathbb{G}} \geq |G|^{1-\frac{\alpha}{4}},$$

where we also used the choice of ε, see (169). In particular, (166) holds for σ_m. Altogether,

$$\text{for all } m \geq m_2 \text{ either } \|\sigma_m * \sigma_m\|_2 \leq K^{-1}\|\sigma_m\|_2 \text{ or (166) holds for } \sigma_m \tag{182}$$

see (172).

Since $K = M^\kappa$ and $\|\sigma_m\|_2 \ll M^{-1/2}$ for all m, see (167), we conclude from (182) that there exists $m_2 \leq m_1 \ll_\Omega 1$ so that (166) holds. The proof is complete. □

8.2. **Infinite and finite place.** We now investigate Proposition 46 in the case

$$\nu_1, \nu_2 \in V_\Gamma, \quad \nu_2 = \infty, \quad \text{and} \quad \nu_1 \text{ is a finite place.} \tag{183}$$

Preliminary reductions. Since almost every ν belongs to W_Γ, in the proof of Proposition 46 we may assume that $\nu_1 \in W_\Gamma$.

Note that G_1 is a finite group, we thus equip G with the metric d induced by the admissible metric on G_2 and the discrete metric on G_1. To be more explicit,

$$1_\eta = \{1^{(1)}\} \times 1_\eta^{(2)} \qquad \text{for all } 0 < \eta < 1.$$

We let m_{G_i} denote the probability Haar measures on G_i for $i = 1, 2$, and let m_G be the product measure. Then $|1_\eta| = |1_\eta^{(2)}/|G_1||$ for all $0 < \eta < 1$. In particular, we have

$$\frac{1}{|G_1|C_1}\eta^{d_0} \leq |1_\eta| \leq \frac{C_1}{|G_1|}\eta^{d_0}. \tag{184}$$

where $d_0 = \dim\mathbb{G}$ and $C_1 \geq 2$ depends only on d_0. This in particular implies that for any $\theta > 0$ and all $2^{-\ell-1} < \eta \leq 2^{-\ell} \leq |G_1|^{-\theta}$, we have

$$\frac{1}{2C_1}\eta^{d_0(\ell,\theta)} \leq |1_\eta| \leq 2C_1\eta^{d_0(\ell,\theta)} \tag{185}$$

where $d_0(\ell,\theta) = d_0 + \frac{1}{\ell}\log_2|G_1|$; consequently, $d_0 \leq d_0(\ell,\theta) \leq d_0 + \theta^{-1}$.

We will use the following notation:

$$P_\eta^{(2)} = \frac{1}{|1_\eta^{(2)}|}1_\eta^{(2)} \quad \text{and} \quad P_\eta = \frac{1}{|1_\eta|}1_\eta.$$

Throughout the argument, and whenever necessary, we will assume

$$0 < \eta \leq (C_1d_0)^{-O_\Gamma(1)} \qquad \text{and} \qquad p_{\nu_1} \geq (C_1d_0)^{O_\Gamma(1)}.$$

As it was done in §8.1, let $X = (X_1, X_2)$ be as in the statement of Proposition 46, and let μ be the probability law of X. Since $\langle\Omega\rangle = \Gamma$ is Zariski dense in $\mathbb{G}$, which is absolutely almost simple, we have $\min\{\mathcal{L}(X_1), \mathcal{L}(X_2)\} \geq c_0 > 0$ where $c_0 = c_0(\Omega)$, see [1, 24] and references therein. Thus

$$\max\{\lambda(\pi_{1*}\mu; G_1), \lambda(\pi_{2*}\mu; G_2)\} \leq \lambda < 1 \tag{186}$$

where π_i denotes the projection onto the i-th factor and $\lambda := \lambda(\Omega)$. Therefore, there exists $E_1 \geq 1$ so that both of the following hold

$$\begin{aligned} &\left|\pi_{1*}\mu^{(\ell_1)}(f_1)\right| \leq |G_1|^{-1}\|f_1\|_2 && \text{for all } \ell_1 \geq E_1 \log|G_1| \\ &\left|\pi_{2*}\mu^{(\ell_1)}(f_{2,\eta})\right| \leq \eta\|f_2\|_2 && \text{for all } 0 < \eta < 1 \text{ and all } \ell_1 \geq E_1|\log\eta| \end{aligned} \tag{187}$$

where $f_1 \in L^2_0(G_1)$, $f_2 \in L^2_0(G_2)$, and $f_{2,\eta} = f_2 * P^{(2)}_\eta$. This can be seen, e.g., by applying Lemma 28 with η (and in the first case $\eta = |G_1|^{-1}$) and choosing E_1 so that $\lambda^{\ell_1/2}\eta^{-2} \leq 1$ for all $\ell_1 \geq E_1|\log\eta|$.

Recall that by the Peter–Weyl theorem we have

$$L^2(G) = \bigoplus_{\rho_{1,i}\in\hat{G}_1, \rho_{2,j}\in\hat{G}_2} \dim(\rho_{1,i}\otimes\rho_{2,j})\rho_{1,i}\otimes\rho_{2,j}. \tag{188}$$

We will use the following fact in the sequel. Let $V_2 = \oplus_{\rho_{2,j}\neq 1}\dim(1\otimes\rho_{2,j})1\otimes\rho_{2,j}$. Then, V_2 is the natural embedding of $L^2_0(G_2)$ in $L^2_0(G)$, thus (186) implies that

$$\mathcal{L}(\mu; V_2) \geq c_2(\Omega) > 0. \tag{189}$$

Finally, note that (154) and the fact that G_2 is connected imply that

$$p_{\nu_1}^{n_1} \leq |G_1| \leq Dp_{\nu_1}^d \leq p_{\nu_1}^{d+1}.$$

In consequence, we will assume for the rest of the argument that

$$n_1 \leq d + 1. \tag{190}$$

In view of (190) and part (1) in Lemma 4, we have: for any nontrivial representation $\rho \in \hat{G}_1$

$$\dim(\rho) \geq c'_\Gamma\,|G_1|^{\alpha'_\Gamma}, \qquad \text{for some } 0 < c'_\Gamma, \alpha'_\Gamma \leq 1. \tag{191}$$

We begin with the following.

Lemma 49. *There exists some $\alpha = \alpha(C_0, L, d_0) > 0$ so that the following holds. Assume there exists E'_1 so that for all $0 < \eta \leq |G_1|^{-1}$ there exists some $\ell \leq E'_1|\log\eta|$ so that*

$$\|(\mu^{(\ell)})_\eta\|_2 \leq \eta^{-\alpha},$$

then Proposition 46 holds for $G = G_1 \times G_2$.

Proof. This is essentially proved in [20, §8], we explicate some of the details for the convenience of the reader. First note that by Lemma 4, there exists (C_0, L) depending only on Γ so that $G = G_1 \times G_2$ is L-locally random with coefficient C_0. Moreover, G satisfies the dimension condition (184). Therefore, [20, Theorem 2.10] implies that there exists D, depending only on d_0, and α', depending only on L and C_0, so that the following holds: there is an exceptional representation $\mathcal{H}_0$ of $G_1 \times G_2$ with $\dim\mathcal{H}_0 \leq 2C_0C_1^D|G_1|^D$ so that, if for all $\eta \leq |G_1|^{-D}$ there exists $\ell \leq E'|\log\eta|$ satisfying

$$\|(\mu^{(\ell)})_\eta\|_2 \leq \eta^{-\alpha'},$$

then we have

$$\mathcal{L}(\mu; L^2(G) \ominus \mathcal{H}_0) \geq c, \tag{192}$$

where $c = O_{E',L,C_0}(\alpha') > 0$.

Thus, we may restrict our attention to

$$V := \oplus(\dim(\rho_{1,i}\otimes\rho_{2,j}))\rho_{1,i}\otimes\rho_{2,j} \supseteq \mathcal{H}_0$$

where the sum is over representations with $\dim(\rho_{1,i}\otimes\rho_{2,j}) \le 2C_0C_1^D|G_1|^{-D}$.

By [20, Lemma 8.8], there exists $D' \ge 1$, depending on C_0, L, D, so that if we put $\delta = |G_1|^{-D'}$,

$$\|P_\delta * f - f\|_2 \le \delta^{1/2}\|f\|_2 \tag{193}$$

for any $f \in V$.

Moreover, in view of (189), we may replace V by $V' \subseteq V$ where in the sum above we further assume that $\rho_{1,i} \ne 1$. Then in view of (191), for all components of V', we have

$$\dim(\rho_{1,i}) \ge c'_\Gamma\,|G_1|^{\alpha'_\Gamma}.$$

Recall again that there exists (C_0, L) depending only on Γ so that $G_1 \times G_2$ is L-locally random with coefficient C_0, see Lemma 4. Thus by [20, Lemma 6.1], for any $\chi \in L^2(G)$ and any $f \in V'$,

$$\|\chi * f\|_2 \ll_\Gamma |G_1|^{-\alpha'_\Gamma/2}\|\chi\|_2\|f\|_2. \tag{194}$$

We will show the claim in this case holds if $\alpha \le \alpha'_\Gamma/(4D')$. Recall that $\delta = |G_1|^{-D'}$, and let $\ell \le E'_1|\log\delta|$ be so that the condition of the lemma holds for this choice of δ. Then applying (194) with $\chi = (\mu^{(\ell)})_\delta$ and $f \in V'$, we have

$$\begin{aligned}\|\mu^{(\ell)} * f\| &\le \delta^{1/2}\|f\|_2 + \|\mu^{(\ell)} * P_\delta * f\|_2\\ &\ll \delta^{1/2}\|f\|_2 + |G_1|^{-\alpha'_\Gamma/2}\delta^{-\alpha}\|f\|_2 \le 2\delta^{\alpha'_\Gamma/(4D')}\|f\|_2,\end{aligned} \tag{195}$$

where we used (193) in the first inequality and (194) in the second.

Now, (195) implies that

$$E'_1|\log\delta|\mathcal{L}(\mu; V') \ge \alpha'_\Gamma|\log\delta|/(8D').$$

This, together with (192) and (189) finish the proof. □

In view of Lemma 49, thus, Proposition 46 in this case will follow from the following.

Lemma 50. *There exists E'_1 so that for all $0 < \eta \le |G_1|^{-1}$, there is some $\ell \le E'_1|\log\eta|$ with*

$$\|(\mu^{(\ell)})_\eta\|_2 \le \eta^{-\alpha},$$

where α is as in Lemma 49.

Proof. The proof of the lemma, which will be completed in several steps, is similar to the proof of Lemma 48. Let E_1 be as in (187) and let $\ell_1 \ge E_1|\log\eta|$. We will show that the claim holds with some $E'_1 = 2^m E_1$ for some m. Let $\sigma = \mu^{(\ell_1)}$, and for every nonnegative integer m, let $\sigma_m = \sigma^{(2^m)}$.

Let $\bar{E}$, R_1 and $\varepsilon_0(\Omega)$ as in Theorem 47, and $E_2 = C/c$ with C and c as in Proposition 39. Throughout the argument, we will assume $p_{\nu_1} \ge R_1^{1/\varepsilon^{\bar{E}}}$, where

$$\varepsilon = \varepsilon_0(\Omega)\alpha/(8dE_2). \tag{196}$$

We will also assume that η^ε is smaller than various constants which depend only on d_0, as needed.

Initial entropy and the range of η. Arguing as in the proof of Lemma 31, with σ for ν and $0 < \eta \le |G_1|^{-1}$, the estimate in (187) implies that

$$\|\sigma_\eta\|_\infty \le |G_1|. \tag{197}$$

Now (197) implies that if $\eta \le |G_1|^{-1/\alpha}$, then

$$\|\sigma_\eta\|_2 \le \eta^{-\alpha}. \tag{198}$$

In particular, (198) implies the lemma with ℓ_1 so long as $\eta \le |G_1|^{-1/\alpha}$. For the rest of the argument, thus, we assume

$$|G_1|^{-1/\alpha} \le \eta \le |G_1|^{-1}. \tag{199}$$

We also recall that

$$\frac{1}{C_1'} P_\eta * \sigma_m \le (\sigma_\eta)^{(2^m)} \le C_1' P_{2^m\eta} * \sigma_m \tag{200}$$

where $C_1' \ge 1$ depends only on C_1 and d_0. This and (197) imply that

$$\|(\sigma_m)_\eta\| \le |G_1| \le \eta^{-1} \qquad \text{for all } m. \tag{201}$$

Let δ and $R_2 \ge 4$ be as in Theorem 47 applied with ε. Apply [20, Theorem 2.8] with ε as above (for ε in loc. cit.); replacing δ with a smaller constant, if necessary, we assume that δ also satisfies the claim in [20, Theorem 2.8] with this ε.

Let m_2 be so that $2^{m_2} E_1 > 1/\delta$. Then

$$2^{m_2}\ell_1 \ge 2^{m_2} E_1 |\log \eta| \ge |\log \eta|/\delta \ge \log |G_1|/\delta. \tag{202}$$

Flattening lemma. Recall from (185) that in the range (199), $G = G_1 \times G_2$ satisfies dimension condition with $2C_1$ and some d' satisfying $d_0 \le d' \le d_0 + 1$. Thus [20, Theorem 2.12] is applicable with $2C_1$ and d' for G and all η in this range.

Assume now that for some $m \ge m_2$, we have

$$\|(\sigma_m * \sigma_m)_\eta\|_2 > \eta^\kappa \|(\sigma_m)_\eta\|_2 \tag{203}$$

for some $\kappa > 0$, which is explicated in (204). Then, by [20, Theorem 2.12] there exists E depending on C_1 and d_0, and $H \subseteq G$ which is symmetric and contains 1, so that all of the following hold:

(H-i) $\eta^{E\kappa} \|(\sigma_m)_\eta\|_2^{-2} \le |H_\eta| \le \eta^{-E\kappa} \|(\sigma_m)_\eta\|_2^{-2}$,

(H-ii) $H \cdot H \subset T \cdot H$ where $T \subset H \cdot H$ satisfies $\#T \le \eta^{-E\kappa}$, and

(H-iii) $\sigma_m * \sigma_m(H_{3\eta}) \ge \eta^{E\kappa}$.

We will apply the above with $\kappa = (\alpha \min\{\varepsilon, \delta\})/(10R_2 E d_0)$. This in particular implies

$$\eta^{E\kappa} \ge \eta^{\delta\alpha/(2d_0)} \ge |G_1|^{-\delta}, \tag{204}$$

where in the last inequality we used (199).

Bounded generation and the structure of H. Let ε be as in (196), and let $A \subseteq \Gamma$ be a finite symmetric lift of H. Then by (H-iii) and (204), we have

$$\mathcal{P}_\Omega^{(2^{m+1}\ell_1)}(A) \ge \eta^{E\kappa} \ge |G_1|^{-\delta} \tag{205}$$

Apply Theorem 47 with ε and A. Since $\pi_{\nu_1,n_1}(A) = \pi_1(H)$, (205) and Theorem 47 imply that there exists $0 \le n_1' \le n_1\varepsilon$ so that if we put $H_1 = \prod_{R_2} H_{4\eta}$, then $\pi_{\nu_1,n_1}(\Gamma_{\nu_1,n_1'}) \subseteq \pi_1(H_1)$. Recall from (157) that $n_1 \le d+1$. Hence, $\varepsilon n_1 \le \varepsilon(d+1) < 1$, and we conclude $n_1' = 0$. Altogether,

$$\pi_1(H_1) = G_1 \tag{206}$$

Let $Y = (\pi_2(H_{3\eta}))_\eta$, then $|Y| \asymp |\pi_2(H_{3\eta})|$, see (200). Moreover, by (187), we have

$$\left|(\pi_{2*}\sigma_{m+1})(Y) - |Y|\right| \le \eta |Y|^{1/2} \le \eta$$

Using (204), this and (H-iii) imply that

$$|\pi_2(H_{3\eta})| \ge \eta^{2E\kappa} \ge \eta^{\delta/d_0},$$

we also used $|Y| \asymp |\pi_2(H_1)|$ and assume η is small.

Since $R_2 \ge 4$, the above and [20, Theorem 2.8], recall also the choice of δ, imply that

$$\pi_2(H_1) \supseteq 1^{(2)}_{\eta^\varepsilon}. \tag{207}$$

As it was done in (124), let

$$\beta_2 := \inf\{\beta \in [0,1] |\ \exists g_2 \in G_2, d(g_2, 1^{(2)}) \ge \eta^\beta,\ (1^{(1)}, g_2) \in \textstyle\prod_3 (H_1)_\eta\};$$

where d denotes our fixed bi-invariant metric on G_2.

We claim

$$\beta_2 \le 10\varepsilon. \tag{208}$$

Let us first assume (208) and complete the proof of the lemma.

Large fibers for H_1. In view of (208),

$$\text{There exists } h \in G_2 \text{ with } d(1,h) \ge \eta^{10\varepsilon} \text{ so that } (1,h) \in \textstyle\prod_3 (H_1)_\eta \tag{209}$$

Apply Proposition 39 with the group G_2, h as in (209), and $\rho = \eta^{10\varepsilon}$. Then by Proposition 39 combined with (209) and (207), there exists E_2 depending only on d_0 (indeed $E_2 = C/c$) so that

$$\big\{(1, g_2) : g_2 \in 1^{(2)}_{\eta^{E_2\varepsilon}}\big\} \subset \textstyle\prod_{7d_0}(H_1)_\eta. \tag{210}$$

The conclusion of the proof. Recall now from (206) that $G_1 = \pi_1(H_1)$. This and (210) imply that there exists some t, again depending only on d_0, so that

$$\Big|\textstyle\prod_{R_2(1+7d_0)} H_{t\eta}\Big| \ge m_{G_2}\big(1^{(2)}_{\eta^{E_2\varepsilon}}\big) \ge C_1^{-1}\eta^{E_2 d_0\varepsilon} \tag{211}$$

where we also used $H_1 = \prod_{R_2} H_{4\eta}$ and the triangle inequality.

Then using (H-i) and (H-ii), we have

$$\eta^{-E\kappa(1+R_2(1+7d_0))}\|(\sigma_m)_\eta\|_2^{-2} \ge \eta^{-E\kappa(R_2(1+7\dim\mathbb{G}))}|H_\eta| \ge c_1'\Big|\textstyle\prod_{R_2(1+7d_0)} H_{t\eta}\Big|$$

where c_1' depends only on d_0.

Combining this with (211), we get

$$\eta^{-E\kappa(1+R_2(1+7d_0))}\|(\sigma_m)_\eta\|_2^{-2} = \eta^{-\kappa(E+R_2E(1+7d_0))}\|(\sigma_m)_\eta\|_2^{-2} \ge c_1' C_1^{-1}\eta^{E_2 d_0\varepsilon}$$

Recall that $\kappa = (\alpha\min\{\varepsilon,\delta\})/(10R_2Ed_0)$. In consequence,

$$\|(\sigma_m)_\eta\|_2^{-2} \ge c_1' C_1^{-1}\eta^{E_2d_0\varepsilon}\eta^{8R_2Ed_0\kappa} \ge \eta^{2E_2d_0\varepsilon}.$$

This and the choice of ε, see (196), imply that

$$\|(\sigma_m)_\eta\|_2 \le \eta^{-E_2d_0\varepsilon} \le \eta^{-\alpha/4}. \tag{212}$$

Altogether,

$$\text{for all } m \ge m_2 \text{ either } \|\sigma_m * \sigma_m\|_2 \le \eta^{E\kappa}\|\sigma_m\|_2 \text{ or (212) holds for } \sigma_m \tag{213}$$

see (203).

Since $\|(\sigma_m)_\delta\|_2 \le \eta^{-1}$ for all m, see (201), we conclude from (213) that there exists $m_2 \le m_1 \ll_\Omega 1$ so that (212) holds. This completes the proof of the lemma assuming (208). □

Proof of (208). In view of Lemma 35 applied with H_1, and using (206) and (207), there exists

$$f : G_1 \to G_2 \qquad \text{with} \qquad f(1^{(1)}) = 1^{(2)}$$

so that all the following hold

$$d(f(g_1)f(g_1'), f(g_1g_1')) \le \eta^{\beta_2} \qquad \text{for all } g_1, g_1' \in G_1, \tag{214a}$$

$$d(f(g_1^{-1}), f(g_1)^{-1}) \le \eta^{\beta_2} \qquad \text{for all } g_1 \in G_1, \text{ and} \tag{214b}$$

$$1^{(2)}_{\eta^\varepsilon} \subseteq (\operatorname{Im} f)_{\eta^{\beta_2}} \tag{214c}$$

Assume contrary to the claim that $\beta_2 > 10\varepsilon$. We will now use a simpler version of the argument in §4.2 in the case $F_2 = \mathbb{R}$ and $F_1 = \mathbb{Q}_p$ to get a contradiction.

Using (214c) and $\beta_2 > 10\varepsilon$, there exists $g_1 \in G_1$ so that $f(g_1) \in 1^{(2)}_{\eta^\varepsilon} \setminus 1^{(2)}_{\eta^\varepsilon/2}$. Recall now from [25] that there exists some C depending only on d_0 so that

$$\langle g_1 \rangle \subseteq \prod_C \{h[g_1, h']h^{-1} : h, h' \in G_1\},$$

This, the bi-invariance of the metric, and the triangle inequality imply that

$$d(f(g_1^k), 1^{(2)}) \leq C'(\eta^\varepsilon + \eta^{\beta_2}) \leq \eta^{\varepsilon/2} \qquad \text{for all } k \tag{215}$$

where C' depends only on d_0, see the proof of Lemma 10.

Now sing (214a), (214b), and (215), for all $\eta^{-\varepsilon}/100 \leq k \leq \eta^{-\beta_2/2}$, we have

$$d(f(g_1)^k, 1^{(2)}) \leq d(f(g_1^k), 1^{(2)}) + \varepsilon^{\beta_2/2} \leq \eta^{\varepsilon/3}.$$

However, if we write $f(g_1) = \exp(z)$ where $z \in \mathfrak{g}_2$ satisfies $\eta^\varepsilon/2 \leq \|z\| \leq \eta^\varepsilon$, then there exists some k in the above range so that

$$d(f(g_1)^k, 1^{(2)}) = d(\exp(kz), 1^{(2)}) \gg 1$$

This contradiction finishes the proof. □

Proof of Proposition 46. We first assume that both ν_1 and ν_2 are finite places.

If we have

$$\max\{n_1, n_2\} \leq 4d^2(d+1)E_2/(\varepsilon_0(\Omega)\alpha_\Gamma), \tag{216}$$

the proposition follows from [25, Theorem 1], with M_0 in that theorem equal to $\frac{4d^2(d+1)E_2}{\varepsilon_0(\Omega)\alpha_\Gamma}$. If (216) fails on the other hand, the proposition follows from Lemma 48.

Altogether, the proof is complete if $\nu_1, \nu_2 \in V_{\mathrm{f},\Gamma}$.

The case $\nu_2 = \infty$ and $\nu_1 \in V_{\mathrm{f},\Gamma}$ follows from the discussion in §8.2, see in particular, Lemma 49 and Lemma 50. □

9. Proof of Theorem 2

We now begin the proof of Theorem 2. As it was mentioned before, the proof relies on Theorem 1, Theorem 44 and Proposition 46.

Proof of Theorem 2. Recall that we fixed a $\mathbb{Q}$-embedding $\mathbb{G} \subseteq (\mathrm{SL}_N)_\mathbb{Q}$; our constants are allowed to depend on this embedding.

Step 1: By the strong approximation theorem, there exists a finite index subgroup $\Lambda \subset \Gamma$ so that

$$\Lambda_{\nu_1,\nu_2} = \Lambda_{\nu_1} \times \Lambda_{\nu_2} \qquad \text{for all } \nu_1, \nu_2 \in V_\Gamma, \tag{217}$$

where Λ_ν denotes the closure of Λ in $\mathbb{G}(\mathbb{Q}_\nu)$. In view of Proposition 55, it suffices to prove the theorem for Λ. Thus, we assume (217) holds for Γ.

Step 2: There is a finite set of places E_Γ so that for all $\nu \notin E_\Gamma$, we have

$$\Gamma \subseteq \mathrm{SL}_N(\mathbb{Z}_\nu).$$

For every finite place $\nu \in E_\Gamma$, there exists $g_\nu \in \mathrm{GL}_N(\mathbb{Q}_\nu)$ such that

$$g_\nu \Gamma g_\nu^{-1} \subseteq \mathrm{SL}_N(\mathbb{Z}_\nu).$$

In view of the strong approximation theorem, there exists $g \in \mathrm{PGL}_N(\mathbb{Q})$ so that $g \in \mathrm{PGL}_N(\mathbb{Z}_\nu)$ for all $\nu \notin E_\Gamma$ and $g^{-1}g_\nu \in \mathrm{PGL}_N(\mathbb{Z}_\nu)$ for every finite place $\nu \in E_\Gamma$. Therefore,

$$g\Gamma g^{-1} \subseteq \mathrm{SL}_N(\mathbb{Z}_\nu) \quad \text{for all non-archimedean places } \nu. \tag{218}$$

In view of (218), we will assume

$$\Gamma \subseteq \mathrm{SL}_N(\mathbb{Z}_\nu) \quad \text{for all non-archimedean places } \nu. \tag{219}$$

for the rest of the argument.

Step 3: If the archimedean place belongs to V_Γ, i.e., $\mathbb{G}(\mathbb{R})$ is compact, fix some $g \in \mathrm{SL}_N(\mathbb{R})$ so that

$$g\mathbb{G}(\mathbb{R})g^{-1} \subseteq \mathrm{SO}_N(\mathbb{R}).$$

In this case, we replace $\mathbb{G}(\mathbb{R})$ by $g\mathbb{G}(\mathbb{R})g^{-1}$ and we work with the corresponding copy of Γ in the archimedean place. This means, for every non-archimedean place ν, we are replacing $\Gamma_{\infty,\nu}$ with

$$(g,1)\Gamma_{\infty,\nu}(g^{-1},1) \subseteq \mathrm{SO}_N(\mathbb{R}) \times \mathrm{SL}_N(\mathbb{Z}_\nu).$$

Note that $\mathbb{G}(\mathbb{Q}_\nu)$ is unchanged for all non-archimedean places ν.

Step 4: Let $\nu_1, \nu_2 \in V_\Gamma$. In view of Lemma 4, we have Γ_{ν_1,ν_2} is L-locally random with coefficient C_0, where the parameters L and C_0 depend only on Γ.

Let μ be the law of X. By Claim 45 in the proof of Theorem 1, applied with $G_1 := \Gamma_{\nu_1}$ and $G_2 := \Gamma_{\nu_2}$, we have

$$\|\mu^{(\ell)} * f\|_2 \le \eta^b \|f\|_2, \tag{220}$$

for all functions f which live at scale $\eta \le \eta_0$ and $\ell \le \bar{C}\log(1/\eta)$, where $\bar{C}$ and b depend only on $\dim \mathbb{G}$, L, and $\eta_0 = \max\{p_{\nu_1}, p_{\nu_2}\}^{-O_{\mathbb{G}}(1)}$.

In view of (220) and Theorem 44, applied with $G = \Gamma_{\nu_1,\nu_2}$, there exists a subrepresentation $\mathcal{H}_0 := \mathcal{H}_{\nu_1,\nu_2,0} \subset L^2_0(\Gamma_{\nu_1,\nu_2})$ with

$$\dim \mathcal{H}_0 \le \eta_0^{-c_0(\Gamma)} \eta_0^{-O_{\mathbb{G}}(1)} \quad \text{and} \tag{221a}$$

$$\mathcal{L}(X; L^2_0(\Gamma_{\nu_1,\nu_2}) \ominus \mathcal{H}_0) \ge \frac{b}{\bar{C}}. \tag{221b}$$

Step 5: We now investigate $\mathcal{L}(X;\mathcal{H}_0)$. In view of (221a) and [24, Proposition 33], the representation $\mathcal{H}_0$ of Γ_{ν_1,ν_2} factors through $G_1 \times G_2$ where

$$G_i = \Gamma_{\nu_i}/\Gamma_{\nu_i,n_i}$$

and the number of connected components of $G_1 \times G_2$ is $\le \eta_0^{-O_\Gamma(1)} \eta_0^{-O_{\mathbb{G}}(1)}$, see §8 for the notation. In particular, condition (154) of Proposition 46 is satisfied. Hence, by Proposition 46, we have

$$\mathcal{L}(X;\mathcal{H}_0) \ge \mathcal{L}(X; G_1 \times G_2) > \varrho_\Gamma > 0.$$

This and (221b) complete the proof. □

Appendix A. Passing to an open subgroup

The main goal of this section is to show how we can study the spectral gap property of an action on a compact group by passing to an open subgroup and vice versa (see Proposition 55). Along the way, we review the connection between the spectral gap property and *almost invariant* functions, and we also give the connection between $\mathcal{L}(\mathcal{P}_{\Omega_1})$ and $\mathcal{L}(\mathcal{P}_{\Omega_2})$ where Ω_1 and Ω_2 generate the same dense subgroup Γ of G.

We start by recalling the concept of *almost invariant functions* with respect to a given finite symmetric set. For a compact group G, a finite symmetric subset Ω of G, and a non-zero function $f \in L^2(G)$, we let

$$\delta_\Omega(f) := \max_{w\in\Omega} \frac{\|w \cdot f - f\|_2}{\|f\|_2}.$$

We let $\delta(\Omega) := \inf\{\delta_\Omega(f) \mid f \in L^2(G)^\circ, \|f\|_2 = 1\}$ where $L^2(G)^\circ$ is the space of functions in $L^2(G)$ that are orthogonal to the constant functions on G. A function is called *ε-almost invariant with respect to* Ω if $\delta_\Omega(f) \le \varepsilon$. It is well-known that $\mathcal{L}(\mathcal{P}_\Omega) > 0$ if and only if $\delta(\Omega) > 0$. The next lemma is a quantitative version of this statement.

Lemma 51. *Suppose G is a compact group and Ω is a finite symmetric subset of G. Then*

$$\frac{1}{|\Omega|}\delta(\Omega)^2 \ll \mathcal{L}^\bullet(\mathcal{P}_\Omega) \ll \delta(\Omega),$$

where $\mathcal{L}^\bullet(\mathcal{P}_\Omega) := \min\{\mathcal{L}(\mathcal{P}_\Omega), 1\}$ and the implied constants are fixed positive numbers.

Proof. The lemma is proved in [6, Proposition I]. See in particular, parts (4) and (6) in that proposition, and note that since $\mathcal{P}_\Omega$ is self adjoint, its spectral radius equals its operator norm. □

In the next lemma, we show how the Lyapunov exponents with respect to two generating sets are related to each other.

Lemma 52. *Suppose G is a compact group and Ω_1 and Ω_2 are two finite symmetric subsets of G such that $\langle\Omega_1\rangle = \langle\Omega_2\rangle$ is dense in G. Suppose k_0 is a positive integer such that $\Omega_1 \subseteq \prod_{k_0}\Omega_2$ and $\Omega_2 \subseteq \prod_{k_0}\Omega_1$. Then*

$$\frac{1}{k_0}\delta(\Omega_1) \le \delta(\Omega_2) \le k_0\delta(\Omega_1),$$

and

$$\frac{1}{k_0^2|\Omega_2|}\mathcal{L}^\bullet(\mathcal{P}_{\Omega_1})^2 \ll \mathcal{L}^\bullet(\mathcal{P}_{\Omega_2}) \ll k_0|\Omega_1|^{1/2}\mathcal{L}^\bullet(\mathcal{P}_{\Omega_1})^{1/2},$$

where $\mathcal{L}^\bullet(\mathcal{P}_{\Omega_i}) := \min\{\mathcal{L}(\mathcal{P}_{\Omega_i}), 1\}$ and the implied constants are fixed positive numbers.

Proof. Since $\Omega_2 \subseteq \prod_{k_0}\Omega_1$, for every function $f \in L^2(G)^\circ$ we have

$$\delta_{\Omega_2}(f) \le k_0\delta_{\Omega_1}(f).$$

Hence $\delta(\Omega_2) \le k_0\delta(\Omega_1)$. By symmetry, we have that $\delta(\Omega_1) \le k_0\delta(\Omega_2)$. This completes proof of the first set of inequalities.

By Lemma 51 and the first set of inequalities, we obtain that

$$\mathcal{L}^\bullet(\mathcal{P}_{\Omega_2}) \ll \delta(\Omega_2) \le k_0\delta(\Omega_1) \ll k_0|\Omega_1|^{1/2}\mathcal{L}^\bullet(\mathcal{P}_{\Omega_1})^{1/2}.$$

The proof follows by symmetry. □

Lemma 53. *Suppose F is a finite group. Then $\delta(F) \ge (2|F|^{-1})^{1/2}$.*

Proof. Suppose $f \in L^2(F)^\circ$. Then

$$\begin{aligned}\sum_{x\in F}\|x\cdot f - f\|_2^2 &= \sum_{x,y\in F}|f(x^{-1}y) - f(y)|^2\\ &= \sum_{x,y\in F}\left(|f(x)|^2 + |f(y)|^2 - 2\mathrm{Re}(f(x)\overline{f(y)})\right)\\ &= 2\|f\|_2^2 + 2|\sum_{x\in F}f(x)|^2 = 2\|f\|_2^2. \end{aligned} \tag{222}$$

By (222), we obtain that $\delta_F(f) \ge (2|F|^{-1})^{1/2}$, which completes the proof. □

The next lemma is a well-known result which provides us with a generating set with interesting properties for a subgroup of finite index in a finitely generated group.

Lemma 54. *Suppose G is a compact group, Γ is a finitely generated dense subgroup of G, and H is an open subgroup of G. Suppose Ω is a finite symmetric generating set of Γ which intersects all the cosets of H in G. Let $s : G/H \to \Omega$ be a section; that means $s(xH) \in xH$ for every $x \in G$. Then $\Omega_H := \overline{\Omega}_H \cup \overline{\Omega}_H^{-1}$ is a symmetric generating set of $\Gamma \cap H$ where*

$$\overline{\Omega}_H := \{s(w_1w_2H)^{-1}w_1w_2 \mid w_1, w_2 \in \Omega\}.$$

Moreover for every $x \in \Gamma \cap H$ we have

$$\frac{1}{3}\ell_\Omega(x) \le \ell_{\Omega_H}(x) \le \ell_\Omega(x)$$

where $\ell_\Omega(x)$ is the least non-negative integer k such that $x \in \prod_k \Omega$ and $\ell_{\Omega_H}(x)$ is defined in a similar way.

Proof. Since every element of Ω_H is a product of at most three elements of Ω, we have that $\ell_\Omega(x) \le 3\ell_{\Omega_H}(x)$ for every x in the group generated by Ω_H. Next by induction on $\ell_\Omega(x)$, we prove that $\ell_{\Omega_H}(x) \le \ell_\Omega(x)$ for every $x \in \Gamma \cap H$; this, in particular, implies that Ω_H generates $\Gamma \cap H$. The base case of $\ell_\Omega(x) = 0$ is clear. Suppose $\ell_\Omega(x) = n$ for some $x \in \Gamma \cap H$. Then $x = w_1 \dots w_n$ where w_i's are in Ω. Let $x' := w_1 \cdots w_{n-2} s(w_{n-1}w_nH)$ and $w' := s(w_{n-1}w_nH)^{-1}w_{n-1}w_n$. Then $w' \in \Omega_H$, $\ell_\Omega(x') \le n-1$, and $x = x'w'$. Since x and w' are in $\Gamma \cap H$, so is x'. Therefore, by the induction hypothesis, we have $\ell_{\Omega_H}(x') \le n-1$. Hence $\ell_{\Omega_H}(x) = \ell_{\Omega_H}(x'w') \le \ell_{\Omega_H}(x') + 1 \le n$. This completes the proof. □

Proposition 55. *Let G be a compact group and Γ be finitely generated dense subgroup of G. Let H be an open subgroup of G. Then $\Gamma \curvearrowright G$ has spectral gap if and only if $\Gamma \cap H \curvearrowright H$ has spectral gap. More precisely, if Γ has a generating set of size n, then Γ has a finite symmetric generating set Ω and $\Gamma \cap H$ has a finite symmetric generating set Ω' such that $|\Omega|, |\Omega'| \ll_{n,[G:H]} 1$ and*

$$\mathcal{L}^\bullet(\mathcal{P}_\Omega)^8 \ll_{n,[G:H]} \mathcal{L}^\bullet(\mathcal{P}_{\Omega'}) \ll_{n,[G:H]} \mathcal{L}^\bullet(\mathcal{P}_\Omega)^{1/8},$$

where $\mathcal{L}^\bullet(\cdot) = \min\{\mathcal{L}(\cdot), 1\}$ and $\mathcal{L}(\mathcal{P}_{\Omega'})$ is defined with respect to a random walk in the compact group H.

Proof. Since H is an open subgroup and Γ is a dense subgroup of G, every coset of H has a representative in Γ. Let $s : G/H \to \Gamma$ be a section such that $s(H) = 1$. Let Ω be a finite symmetric generating set of Γ which contains the image of s as a subset. Let Ω_H be as in Lemma 54 with respect to Ω and section s.

We start with the case that H is a normal subgroup of G.

Claim 1. *In the above setting, $\mathcal{L}^\bullet(\mathcal{P}_\Omega) \gg \frac{1}{|\Omega|[G:H]}\mathcal{L}^\bullet(\mathcal{P}_{\Omega_H})^2$ where the implied constant is a fixed positive number.*

Proof of Claim 1. If $\mathcal{L}(\mathcal{P}_{\Omega_H}) = 0$, there is nothing to prove. Thus assume that $c_H := \mathcal{L}(\mathcal{P}_{\Omega_H}) > 0$. We will estimate $\delta(\Omega)$ from below in terms of c_H, the claim will then follow from Lemma 51.

To that end, suppose $f \in L^2(G)$ and $\|f\|_2 = 1$. Let f_H denote the projection of f into the space of H invariant functions $L^2(G)^H$, then

$$\|\mathcal{P}_{\Omega_H} * f - f_H\|_2^2 \le 2^{-2c_H}\|f - f_H\|_2^2, \tag{223}$$

We include the proof of (223) for completeness. Let us first recall the definition of f_H: for every coset $\overline{x} := xH$, let $f_x := f\mathbb{1}_{\overline{x}}$ where $\mathbb{1}_{\overline{x}}$ is the characteristic function of the coset $\overline{x}$. Then

$$f = \sum_{\overline{x} \in G/H} f_{\overline{x}} \quad \text{and } f_{\overline{x}}\text{'s are pairwise orthogonal} \tag{224}$$

Similarly, since for every $\overline{x} \in G/H$ and $y \in H$, $(y \cdot f - f)\mathbb{1}_{\overline{x}} = y \cdot f_{\overline{x}} - f_{\overline{x}}$, we have

$$y \cdot f - f = \sum_{\overline{x} \in G/H} y \cdot f_{\overline{x}} - f_{\overline{x}} \tag{225}$$

For every $\overline{x} = xH \in G/H$, let $a_{\overline{x}} := \int_G f_{\overline{x}}(z)dz$. Then $f_{\overline{x}} - a_{\overline{x}}\mathbb{1}_{\overline{x}}$ is in the space $L^2(Hx)^\circ$ of functions on $Hx = xH$ that are orthogonal to the constant functions; moreover

$$f_H = \sum_{\overline{x} \in G/H} a_{\overline{x}}\mathbb{1}_{\overline{x}}.$$

In particular, we have

$$\|f\|_2^2 = \|f_H\|_2^2 + \|f - f_H\|_2^2 \quad \text{and} \tag{226a}$$

$$\|\mathcal{P}_{\Omega_H} * (f_{\overline{x}} - a_{\overline{x}}\mathbb{1}_{\overline{x}})\|_2 \le 2^{-c_H}\|f_{\overline{x}} - a_{\overline{x}}\mathbb{1}_{\overline{x}}\|_2 \tag{226b}$$

for every $\overline{x} \in G/H$. By (224), (225), and (226b), we obtain that

$$\|\mathcal{P}_{\Omega_H} * f - f_H\|_2^2 = \|\mathcal{P}_{\Omega_H} * (f - f_H)\|_2^2 = \sum_{\overline{x}\in G/H} \|\mathcal{P}_{\Omega_H} * (f_{\overline{x}} - a_{\overline{x}}\mathbb{1}_{\overline{x}})\|_2^2 \le 2^{-2c_H}\|f - f_H\|_2^2,$$

where we also used $\mathcal{P}_{\Omega_H} * f_H = f_H$ in the first equality. As it was claimed in (223).

We will now use (223) to bound $\|w \cdot f_H - f_H\|_2$, for $w \in \Omega$, in terms on c_H and $\delta_\Omega(f)$, see (230) below. Indeed, since Ω_H is a subset of the three fold product of Ω, for every $y \in \Omega_H$ we have

$$\|y \cdot f - f\|_2 \le 3\delta_\Omega(f). \tag{227}$$

The estimate (227) implies that

$$\|\mathcal{P}_{\Omega_H} * f - f\|_2 = \|\sum_{y\in\Omega_H} \mathcal{P}_{\Omega_H}(y)(y \cdot f - f)\|_2 \le 3\delta_\Omega(f). \tag{228}$$

Now (223) and (228) give $\|f - f_H\|_2 \le 2^{-c_H}\|f - f_H\|_2 + 3\delta_\Omega(f)$, which we write as

$$\|f - f_H\|_2 \le 3(1 - 2^{-c_H})^{-1}\delta_\Omega(f). \tag{229}$$

By (229) and the definition of $\delta_\Omega(f)$, we conclude that for every $w \in \Omega$ the following holds

$$\|w \cdot f_H - f_H\|_2 \le (6(1 - 2^{-c_H})^{-1} + 1)\delta_\Omega(f). \tag{230}$$

We find a lower bound for $\sup_\Omega \|w \cdot f_H - f_H\|_2$. Indeed, $L^2(G)^H$ can be isometrically identified with $L^2(G/H)$. This way we view f_H as an element of $L^2(G/H)$. Since f is orthogonal to constant functions on G, f_H is orthogonal to constant functions on G/H. Therefore, by Lemma 53, we obtain that there is $\overline{x}_0 \in G/H$ such that

$$\|\overline{x_0} \cdot f_H - f_H\|_2 \ge (2[G : H]^{-1})^{1/2}\|f_H\|_2. \tag{231}$$

Since the image of s is in Ω, (231) and (230) imply

$$(2[G : H]^{-1})^{1/2}\|f_H\|_2 \le \|\overline{x_0} \cdot f_H - f_H\|_2 \le (6(1 - 2^{-c_H})^{-1} + 1)\delta_\Omega(f). \tag{232}$$

Moreover, we have

$$\|f_H\|_2 \ge \|f\|_2 - \|f - f_H\|_2 = 1 - \|f - f_H\|_2 \ge 1 - 3(1 - 2^{-c_H})^{-1}\delta_\Omega(f), \tag{233}$$

where in the last estimate we used (229). Using (233) in (232), we conclude

$$(2[G : H]^{-1})^{1/2}\big(1 - 3(1 - 2^{-c_H})^{-1}\delta_\Omega(f)\big) \le (6(1 - 2^{-c_H})^{-1} + 1)\delta_\Omega(f).$$

This gives

$$\delta_\Omega(f) \ge (12[G : H]^{1/2})^{-1}(1 - 2^{-c_H}). \tag{234}$$

Using (234) and the fact that $1 - e^{-x} \ge \frac{1}{3}\min\{x, 1\}$ for every positive x, we conclude that

$$\delta(\Omega) \gg [G : H]^{-1/2}\mathcal{L}^\bullet(\mathcal{P}_{\Omega_H}). \tag{235}$$

By (235) and Lemma 51, we deduce that

$$\mathcal{L}(\mathcal{P}_\Omega) \gg \frac{1}{|\Omega|[G : H]}\mathcal{L}^\bullet(\mathcal{P}_{\Omega_H})^2,$$

which completes proof of Claim 1. □

Claim 2. *In the above setting* $\mathcal{L}^\bullet(\mathcal{P}_{\Omega_H}) \gg \frac{1}{|\Omega|^2}\mathcal{L}^\bullet(\mathcal{P}_\Omega)^2$ *where the implied constant is a fixed number.*

Proof of Claim 2. If $\mathcal{L}(\mathcal{P}_\Omega)=0$, there is nothing to prove. So without loss of generality, we can assume that $c_G:=\mathcal{L}(\mathcal{P}_\Omega)>0$. Let $f\in L^2(H)^\circ$ with $\|f\|_2^2=1$, and define $\widetilde{f}:G\to\mathbb{C}$, $\widetilde{f}(x):=f(s(xH)^{-1}x)$. Notice that

$$\|\widetilde{f}\|_2^2=\frac{1}{[G:H]}\sum_{xH\in G/H}\int_H|\widetilde{f}(s(xH)y)|^2dm_H(y)=\frac{1}{[G:H]}\sum_{xH\in G/H}\int_H|f(y)|^2dy=1 \tag{236}$$

and

$$\int_G\widetilde{f}(x)dx=\frac{1}{[G:H]}\sum_{xH\in G/H}\int_H\widetilde{f}(s(xH)y)dm_H(y)=\frac{1}{[G:H]}\sum_{xH\in G/H}\int_H f(y)dm_H(y)=0, \tag{237}$$

where m_H is the probability Haar measure on H. In view of (236) and (237), Lemma 51 implies that there exists $w_0\in\Omega$ such that

$$\|w_0\cdot\widetilde{f}-\widetilde{f}\|_2=\delta_\Omega(\widetilde{f})\geq\delta(\Omega)\gg\mathcal{L}^\bullet(\mathcal{P}_\Omega). \tag{238}$$

Next, we want to give an upper bound for $\delta_\Omega(\widetilde{f})$ in terms of $\delta_{\Omega_H}(f)$. Note that for every $w\in\Omega$,

$$w'(\overline{x},w):=s(\overline{x})^{-1}w\,s(w^{-1}\overline{x})\in\Omega_H, \tag{239}$$

where as before $\overline{x}:=xH\in G/H$.

For every $w\in\Omega$ and $y\in\overline{x}=xH$, we have

$$\begin{aligned}(w\cdot\widetilde{f})(y)&=\widetilde{f}(w^{-1}y)=f(s(w^{-1}\overline{x})^{-1}w^{-1}y)\\&=f(w'(\overline{x},w)^{-1}s(\overline{x})^{-1}y)=(w'(\overline{x},w)\cdot f)(s(\overline{x})^{-1}y).\end{aligned} \tag{240}$$

Let now $w\in\Omega$, then

$$\begin{aligned}\|w\cdot\widetilde{f}-\widetilde{f}\|_2^2=&\frac{1}{[G:H]}\sum_{\overline{x}\in G/H}\int_H|(w\cdot\widetilde{f})(s(\overline{x})y)-\widetilde{f}(s(\overline{x})y)|^2dm_H(y)\\=&\frac{1}{[G:H]}\sum_{\overline{x}\in G/H}\int_H|(w'(\overline{x},w)\cdot f)(y)-f(y)|^2dm_H(y)\qquad\text{by (240)}\\=&\frac{1}{[G:H]}\sum_{\overline{x}\in G/H}\|w'(\overline{x},w)\cdot f-f\|_2^2\leq\delta_{\Omega_H}(f)^2.\end{aligned} \tag{241}$$

By (238) and (241), we obtain that

$$\delta_{\Omega_H}(f)\gg\mathcal{L}^\bullet(\mathcal{P}_\Omega). \tag{242}$$

By (242) and Lemma 51, we deduce that

$$\mathcal{L}^\bullet(\mathcal{P}_\Omega)\ll|\Omega_H|^{1/2}\mathcal{L}^\bullet(\mathcal{P}_{\Omega_H})\leq|\Omega|\mathcal{L}^\bullet(\mathcal{P}_{\Omega_H}). \tag{243}$$

This completes the proof of Claim 2. □

Next we consider the case where H is an arbitrary open subgroup. Let N be the normal core of H in G, i.e., N is the largest normal subgroup of G which is a subgroup of H. Since H is an open subgroup of G and G is a compact group, N is also an open subgroup of G. Let $s:G/N\to\Gamma$ be a section such that $s(N)=1$, and Ω be a finite symmetric generating set of Γ which contains the image of s as a subset. Let Ω_N be as in Lemma 54 with respect Ω and the section s. Let

$$\Omega'_H:=\Omega_N\cup\{s(xN)^{\pm1}\mid x\in H\}.$$

Notice that the restriction of s to H/N gives us a section whose image is a subset of Ω'_H. Let Ω'_N be a generating set of $\Gamma \cap N$ which is given by Lemma 54 with respect to Ω'_H and the section $s : H/N \to \Omega'_H$. Then by Claim 1 and Claim 2, we have

$$\frac{1}{|\Omega|^2}\mathcal{L}^\bullet(\mathcal{P}_\Omega)^2 \ll \mathcal{L}^\bullet(\mathcal{P}_{\Omega_N}) \ll |\Omega|^{1/2}[G:N]^{1/2}\mathcal{L}^\bullet(\mathcal{P}_\Omega)^{1/2} \tag{244}$$

and

$$\frac{1}{|\Omega'_H|^2}\mathcal{L}^\bullet(\mathcal{P}_{\Omega'_H})^2 \ll \mathcal{L}^\bullet(\mathcal{P}_{\Omega'_N}) \ll |\Omega'_H|^{1/2}[H:N]^{1/2}\mathcal{L}^\bullet(\mathcal{P}_{\Omega'_H})^{1/2}. \tag{245}$$

Notice that since $s(N) = 1$, 1 is in Ω. Therefore for every $x \in \Omega_N$, $s(1xN)^{-1}1x \in \Omega'_N$ as $1, x \in \Omega'_H$. Notice that $s(xN) = 1$, and so $x \in \Omega'_N$ for every $x \in \Omega_N$. This means $\Omega_N \subseteq \Omega'_N$.

Notice that by Lemma 54, $\ell_\Omega(w) \leq 3$ for every $w \in \Omega'_H$. There by another application of Lemma 54, we deduce that for every $x \in \Omega'_N$, $\ell_{\Omega_N}(x) \leq 7$. This is the case as $x^{\pm 1}$ is equal to $s(w_1w_2N)^{-1}w_1w_2$ for some $w_1, w_2 \in \Omega'_H$, and so

$$\begin{aligned}\ell_{\Omega_N}(x) = \ell_{\Omega_N}(s(w_1w_2N)^{-1}w_1w_2) &\leq \ell_\Omega(s(w_1w_2N)^{-1}w_1w_2)\\ &\leq \ell_\Omega(s(w_1w_2N)^{-1}) + \ell_\Omega(w_1) + \ell_\Omega(w_2) \leq 7.\end{aligned}$$

This means that $\Omega'_N \subseteq \prod_7 \Omega_N$. Hence by Lemma 52, we have

$$\frac{1}{|\Omega'_N|}\mathcal{L}^\bullet(\mathcal{P}_{\Omega_N})^2 \ll \mathcal{L}^\bullet(\mathcal{P}_{\Omega'_N}) \ll |\Omega_N|^{1/2}\mathcal{L}^\bullet(\mathcal{P}_{\Omega_N})^{1/2} \tag{246}$$

By (244), (245), (246), and the facts that $|\Omega'_H| \ll |\Omega|^2$, $|\Omega'_N| \ll |\Omega|^4$, $|\Omega_N| \leq |\Omega|^2$, we conclude

$$|\Omega|^{-18}[G:N]^{-1}\mathcal{L}^\bullet(\mathcal{P}_\Omega)^8 \ll \mathcal{L}^\bullet(\mathcal{P}_{\Omega'_H}) \ll |\Omega|^{21/8}[G:N]^{1/8}\mathcal{L}^\bullet(\mathcal{P}_\Omega)^{1/8}. \tag{247}$$

Notice that $[G:N] \leq [G:H]!$, hence, the claim follows from (247). □

Appendix B. Quantitative inverse function theorem

In this section, we recall and state a quantitative version of the inverse function theorem. We start with setting up a few notation. Here F is either $\mathbb{R}$ or $\mathbb{Q}_p$. For $\mathbf{v} \in \mathbb{Q}_p^d$, $\|\mathbf{v}\|$ denotes the max norm, and for $\mathbf{v} \in \mathbb{R}^d$, $\|\mathbf{v}\|$ denotes the Euclidean norm. For a positive real number r and $\mathbf{v} \in F^d$, $\mathbf{v}_r$ denotes the closed ball of radius r centered at $\mathbf{v}$. For $A \in \mathrm{M}_{m,n}(F)$, we let

$$\sigma(A) := \sup\{r \in [0,\infty) |\, 0_r \subseteq A0_1\}.$$

For $r_0 \in \mathbb{R}^+$, $\mathbf{x}_0 \in F^n$, and an analytic function $\Phi : (\mathbf{x}_0)_{r_0} \to F^m$, we view $\mathrm{d}\Phi(\mathbf{x})$ as an m-by-n matrix with entries in F; the ij-entry of $\mathrm{d}\Phi(\mathbf{x})$ is $\partial_j\Phi_i(\mathbf{x})$ where $\Phi = (\Phi_1, \ldots, \Phi_m)$.

Next we state a p-adic analytic inverse function theorem which is essentially given in [25, Lemma 54].

Theorem 56 (Quantitative inverse function theorem: the p-adic case)**.** *Suppose $r_0 \leq 1$, $\mathbf{x}_0 \in \mathbb{Z}_p^n$ and $\Phi : (\mathbf{x}_0)_{r_0} \to \mathbb{Z}_p^m$ is an analytic function with the following properties.*

(1) There are $c_{\mathbf{i},j} \in \mathbb{Z}_p$ such that

$$\Phi(\mathbf{x}) = \sum_{\mathbf{i}} (c_{\mathbf{i},1}(\mathbf{x}-\mathbf{x}_0)^{\mathbf{i}}, \ldots, c_{\mathbf{i},m}(\mathbf{x}-\mathbf{x}_0)^{\mathbf{i}}),$$

where $(\mathbf{x}-\mathbf{x}_0)^{\mathbf{i}} = \prod_{j=1}^n (x_j - x_{0j})^{i_j}$ for a multi-index $\mathbf{i} = (i_1, \ldots, i_n)$ and $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{x}_0 = (x_{01}, \ldots, x_{0n})$.

(2) $\sigma(\mathrm{d}\Phi(\mathbf{x}_0)) \geq p^{-k_0}$ for some positive integer k_0.

Then for every integer $l \geq k_0 + 1$ we have

$$\Phi(\mathbf{x}_0)_{p^{-k_0-l}} \subseteq \Phi((\mathbf{x}_0)_{p^{-l}}).$$

Proof. See proofs of [25, Lemma 54, Lemma 54']. □

Now the real case of Theorem 56 will be discussed.

Theorem 57 (Quantitative inverse function theorem: the real case)**.** *Suppose* $r_0 \le 1$, $\mathbf{x}_0 \in \mathbb{R}^n$ *and* $\Phi : (\mathbf{x}_0)_{r_0} \to \mathbb{R}^m$ *is a* C^2*-function with the following properties.*

(1) For every $1 \le j, j' \le n$ *and* $\mathbf{x} \in (\mathbf{x}_0)_{r_0}$, $\|\partial_{j,j'}\Phi(\mathbf{x})\| \le \alpha$.
(2) For some positive number σ_0, $\sigma(\mathrm{d}\Phi(\mathbf{x}_0)) \ge \sigma_0$.

Then for every $0 < r < \max(\frac{\sigma_0}{2mn\sqrt{\alpha}}, r_0)$ *we have*

$$\Phi(\mathbf{x}_0)_{(\frac{\sigma_0}{4})r} \subseteq \Phi((\mathbf{x}_0)_r).$$

We start with a linear algebra lemma.

Lemma 58. *Suppose* $\mathbf{v}_1, \dots, \mathbf{v}_n, \mathbf{v}'_1, \dots, \mathbf{v}'_n \in \mathbb{R}^m$*, and let* $A := [\mathbf{v}_1 \cdots \mathbf{v}_n]$ *and* $A' := [\mathbf{v}'_1 \cdots \mathbf{v}'_n]$*. Suppose* σ_0 *is a positive number and* $0 < \varepsilon < \sigma_0/\sqrt{n}$*. Suppose* $A = K_1[D\, 0_{n-m}]K_2$ *is a singular value decomposition of A; that means* $K_1 \in O_m(\mathbb{R})$*,* $K_2 \in O_n(\mathbb{R})$*, and* $D = \mathrm{diag}(s_1, \dots, s_m)$ *for some* $s_1 \ge \cdots \ge s_m \ge 0$*. If* $\sigma(A) \ge \sigma_0$ *and* $\|\mathbf{v}_i - \mathbf{v}'_i\| < \varepsilon$ *for every i, then*

$$\sigma(A') \ge \sigma(A'L) \ge \sigma_0 - \sqrt{n}\varepsilon,$$

where $L := K_2^{-1}\begin{bmatrix} I_m \\ 0_{n-m} \end{bmatrix}$.

Proof. Since K_1 and K_2 are orthogonal,

$$\sigma(A'L) = \sigma(K_1^{-1}A'L) \quad \text{and} \quad \sigma(A) = \sigma(K_1^{-1}AK_2^{-1}) = \sigma(D) = s_m. \tag{248}$$

By the assumption, all the columns of $Y := \varepsilon^{-1}(A - A')$ have length at most 1. As K_1 and K_2 are orthogonal, all the columns of $X := K_1^{-1}YL$ have length at most $\sqrt{n}$. Notice that by the definition of X, we have $K_1^{-1}A'L = D - \varepsilon X$. By (248), we deduce that $\sigma(D - \varepsilon X) = \sigma(A'L)$, where $D := \mathrm{diag}(s_1, \dots, s_m)$. Notice that for every $\mathbf{w} \in \mathbb{R}^m$ we have

$$\begin{aligned}
\|(D - \varepsilon X)^{-1}\mathbf{w}\| =& \|D^{-1}(I - \varepsilon XD^{-1})^{-1}\mathbf{w}\| \le s_m^{-1}\|(I - \varepsilon XD^{-1})^{-1}\mathbf{w}\| \\
\le& \sigma_0^{-1}\Big(\sum_{i=0}^{\infty}(\varepsilon\sqrt{n}\sigma_0^{-1})^i\Big)\|\mathbf{w}\| = \sigma_0^{-1}\frac{1}{1 - \varepsilon\sqrt{n}\sigma_0^{-1}}\|\mathbf{w}\| \\
=& \frac{1}{\sigma_0 - \sqrt{n}\varepsilon}.
\end{aligned} \tag{249}$$

By (249), we deduce that $\sigma(D - \varepsilon X) \ge \sigma_0 - \sqrt{n}\varepsilon$. and the claim follows. □

Lemma 59. *Under the assumptions of Theorem 57, let* $\mathrm{d}\Phi(\mathbf{x}_0) = K_1[D\, 0_{m,n-m}]K_2$ *be a singular value decomposition; that means* $K_1 \in O_m(\mathbb{R})$*,* $K_2 \in O_n(\mathbb{R})$*, and* $D = \mathrm{diag}(s_1, \dots, s_m)$ *for some* $s_1 \ge \cdots \ge s_m \ge 0$*. Let* $L := K_2^{-1}\begin{bmatrix} I_m \\ 0_{n-m} \end{bmatrix}$*. If* $\|\mathbf{x} - \mathbf{x}_0\| \le \min(r_0, \frac{\sigma_0}{2n\sqrt{m\alpha}})$*, then the following holds*

$$\sigma(\mathrm{d}\Phi(\mathbf{x})L) \ge \sigma_0/2.$$

Proof. By the mean value theorem, for every i, j, and k, there is a point $\mathbf{x}_{ijk}$ on the segment connecting $\mathbf{x}_0$ to $\mathbf{x}$ such that

$$\partial_j\Phi_i(\mathbf{x}) - \partial_j\Phi_i(\mathbf{x}_0) = \sum_{k=1}^{n}\partial_{k,j}\Phi_i(\mathbf{x}_{ijk})(x_k - x_{0k}). \tag{250}$$

By (250), we deduce that for every j the following holds

$$\|\partial_j\Phi(\mathbf{x}) - \partial_j\Phi(\mathbf{x}_0)\| \leq \sqrt{mn\alpha}\|\mathbf{x}-\mathbf{x}_0\| \leq \sqrt{mn\alpha}\frac{\sigma_0}{2n\sqrt{m\alpha}} = \frac{\sigma_0}{2\sqrt{n}}. \tag{251}$$

By Lemma 58 and (251), the claim follows. □

Proof of Theorem 57. By the singular value decomposition of $\mathrm{d}\Phi(\mathbf{x}_0)$, there are $K_1 \in O_m(\mathbb{R})$ and $K_2 \in O_n(\mathbb{R})$ and $s_1 \geq \ldots \geq s_m > 0$ such that $\mathrm{d}\Phi(\mathbf{x}_0) = K_1[D\, 0_{m,n-m}]K_2$ where $D := \mathrm{diag}(s_1,\ldots,s_m)$ is the diagonal matrix with diagonal entries $s_1,\ldots,s_m$. Then $s_m = \sigma(\mathrm{d}\Phi(\mathbf{x}_0))$. Let $L:\mathbb{R}^m \to \mathbb{R}^n, L(\mathbf{v}) := K_2^{-1}\begin{pmatrix}\mathbf{v}\\ 0_{n-m}\end{pmatrix}$. Notice that $\|L(\mathbf{v})\| = \|\mathbf{v}\|$ for every $\mathbf{v}\in\mathbb{R}^m$ and L can be represented by the matrix $K_2^{-1}\begin{bmatrix} I_m \\ 0_{m,n-m}\end{bmatrix}$; this matrix is denoted by L. Furthermore for every $\mathbf{v}$ we have

$$\|\,\mathrm{d}\Phi(\mathbf{x}_0)L(\mathbf{v})\| \geq \sigma(\mathrm{d}\Phi(\mathbf{x}_0))\|\mathbf{v}\| \geq \sigma_0\|\mathbf{v}\|. \tag{252}$$

Let

$$\Psi: 0_{r_0} \to \mathbb{R}^m, \Psi(\mathbf{v}) := \Phi(L(\mathbf{v})+\mathbf{x}_0) - \Phi(\mathbf{x}_0) - \mathrm{d}\Phi(\mathbf{x}_0)L(\mathbf{v}). \tag{253}$$

Suppose $\Psi = (\Psi_1,\ldots,\Psi_m)$. By the mean value theorem, for every i, there is $\mathbf{v}_i$ on the segment connecting $\mathbf{v}$ to 0 such that

$$\Psi_i(\mathbf{v}) - \Psi_i(0) = \nabla\Psi_i(\mathbf{v}_i)\cdot\mathbf{v}. \tag{254}$$

Notice that $\nabla\Psi_i(\mathbf{w})$ is the i-th row of $\mathrm{d}\Psi(\mathbf{w})$, and by the chain rule, we have

$$\mathrm{d}\Psi(\mathbf{w}) = \Big(\mathrm{d}\Phi(L(\mathbf{w})+\mathbf{x}_0) - \mathrm{d}\Phi(\mathbf{x}_0)\Big)L. \tag{255}$$

By (251), we have

$$\|\partial_j\Phi(L(\mathbf{w})+\mathbf{x}_0) - \partial_j\Phi(\mathbf{x}_0)\| \leq \sqrt{mn\alpha}\|L(\mathbf{w})\| = \sqrt{mn\alpha}\|\mathbf{w}\|. \tag{256}$$

By (254), (255), and (256), the following holds

$$|\Psi_i(\mathbf{v}) - \Psi_i(0)| \leq (\sqrt{n}\|\mathbf{v}\|)(\sqrt{mn\alpha}\|\mathbf{v}\|) = n\sqrt{m\alpha}\|\mathbf{v}\|^2. \tag{257}$$

Inequality given in (257) implies

$$\|\Psi(\mathbf{v}) - \Psi(0)\| \leq mn\sqrt{\alpha}\|\mathbf{v}\|^2. \tag{258}$$

Since $\Psi(0)=0$, by (253), (258), and (252), we obtain

$$\begin{aligned}\|\Phi(L(\mathbf{v})+\mathbf{x}_0) - \Phi(\mathbf{x}_0)| \geq& \|\,\mathrm{d}\Phi(\mathbf{x}_0)L(\mathbf{v})\| - mn\sqrt{\alpha}\|\mathbf{v}\|^2\\ \geq&\sigma_0\|\mathbf{v}\| - mn\sqrt{\alpha}\|\mathbf{v}\|^2\\ \geq&\Big(\sigma_0 - mn\sqrt{\alpha}\|\mathbf{v}\|\Big)\|\mathbf{v}\|\end{aligned} \tag{259}$$

By (259), if $\|\mathbf{v}\| \leq \frac{\sigma_0}{2mn\sqrt{\alpha}}$, we obtain

$$\|\Phi(L(\mathbf{v})+\mathbf{x}_0) - \Phi(\mathbf{x}_0)\| \geq \frac{\sigma_0}{2}\|\mathbf{v}\|. \tag{260}$$

Let $0 < r < \frac{\sigma_0}{2mn\sqrt{\alpha}}$, and for $\mathbf{y}\in\Phi(\mathbf{x}_0)_{\sigma_0 r/4}$, consider the function

$$f_{\mathbf{y}}:\overline{0}_r \to \mathbb{R}, f(\mathbf{v}) := \|\Phi(L(\mathbf{v})+\mathbf{x}_0) - \mathbf{y}\|^2,$$

where $\overline{0}_r$ is the closed ball of radius r centered at 0. By (260), if $\|\mathbf{v}\| = r$, then

$$\begin{aligned}\sqrt{f_{\mathbf{y}}(\mathbf{v})} \geq& \|\Phi(L(\mathbf{v}) + \mathbf{x}_0) - \Phi(\mathbf{x}_0)\| - \|\Phi(\mathbf{x}_0) - \mathbf{y}\| \\ \geq& \frac{\sigma_0}{2}\|\mathbf{v}\| - \|\Phi(\mathbf{x}_0) - \mathbf{y}\| \\ >& \frac{\sigma_0 r}{2} - \frac{\sigma_0 r}{4} = \frac{\sigma_0 r}{4} > \sqrt{f_{\mathbf{y}}(0)}.\end{aligned} \tag{261}$$

By (261), the minimum of $f_{\mathbf{y}}$ occurs at at a critical point $\mathbf{v}_{\mathbf{y}} \in 0_r$ of $f_{\mathbf{y}}$. Knowing that $\nabla f_{\mathbf{y}}(\mathbf{v}_{\mathbf{y}}) = 0$, using the chain rule, we obtain that the following holds

$$(\Phi(L(\mathbf{v}_{\mathbf{y}}) + \mathbf{x}_0) - \mathbf{y})^T \, \mathrm{d}\Phi(L(\mathbf{v}_{\mathbf{y}}) + \mathbf{x}_0)L = 0,$$

where $(\Phi(L(\mathbf{v}_{\mathbf{y}})+\mathbf{x}_0)-\mathbf{y})^T$ is the row matrix form of the vector $\Phi(L(\mathbf{v}_{\mathbf{y}})+\mathbf{x}_0)-\mathbf{y}$. By Lemma 59, $\sigma(\mathrm{d}\Phi(L(\mathbf{v}_{\mathbf{y}}) + \mathbf{x}_0)L) \geq \sigma_0/2 > 0$. Therefore $\mathrm{d}\Phi(L(\mathbf{v}_{\mathbf{y}}) + \mathbf{x}_0)L$ is injective, which implies that $\mathbf{y} = \Phi(L(\mathbf{v}_{\mathbf{y}}) + \mathbf{x}_0)$. The claim follows. □

Appendix C. The case of abelian groups

In this appendix we will prove the following.

Theorem 60. *The groups $\mathbb{Z}_p$ and $\mathbb{R}/\mathbb{Z}$ are not spectrally independent.*

We start by a general criterion characterizing when a coupling of the Haar measures on two compact abelian groups has spectral gap.

Lemma 61. *Let G be a compact abelian group, and let μ be a symmetric probability measure on G. Then μ does not have spectral gap iff there exists $\gamma_j \in \widehat{G} \setminus \{1\}$ such that γ_j converges to the constant function 1 on G, μ-a.e.*

Proof. Since G is an abelian group, the spectrum of the convolution operator T_μ consists of $\widehat{\mu}(\gamma)$ for $\gamma \in \widehat{G}$. Since μ is a symmetric probability measure $\widehat{\mu}(\gamma) \in [-1, 1]$ for all $\gamma \in \widehat{G}$. Assuming μ does not have spectral gap, we obtain a sequence $\{\gamma_j\}_{j=1}^{\infty} \subseteq \widehat{G}$ such that $|\widehat{\mu}(\gamma_j)| \to 1$ and γ_j's are pairwise distinct. We will show that passing to a subsequence, if necessary, we have

$$\gamma_j^2 \to 1 \qquad \mu\text{-a.e.} \tag{262}$$

To see this, write $\gamma_j = x_j + iy_j$. Indeed, after passing to a subsequence, which we continue to denote by $\{\gamma_j\}$, we have

$$\int \gamma_j d\mu = \int x_j d\mu \to \epsilon$$

where $\epsilon = \pm 1$. Since $|\gamma_j| = 1$, we conclude that $\int |\gamma_j - \epsilon| \, d\mu \to 0$. Thus passing to a further subsequence, if necessary, we get that $\gamma_j \to \epsilon$, μ-a.e.; this implies (262).

For the converse, note that if $\gamma_j \to 1$, μ-a.e., then by Lebesgue's dominant convergence theorem, we have $\widehat{\mu}(\gamma_i) \to 1$; hence μ does not have spectral gap. □

Lemma 62. *Let (Z, ν), (X_1, μ_1), and (X_2, μ_2) be probability spaces. Assume that for $i = 1, 2$, $f_i : (Z, \nu) \to (X_i, \mu_i)$ is a measurable map so that $f_{i*}\nu = \mu_i$. Let $\delta : Z \to X_1 \times X_2$ be defined by $\delta(z) = (f_1(z), f_2(z))$. Then*

$$\mu = \delta_*(\nu) \tag{263}$$

is a coupling of μ_1 and μ_2.

Proof. For $i = 1, 2$, denote the canonical projections from $X_1 \times X_2$ onto X_i by π_i. Then since $\pi_i \circ \delta = f_i$ and $f_{i*}\nu = \mu_i$, we have

$$(\pi_i)_*\mu = (\pi_i)_* \circ \delta_*(\nu) = \mu_i \qquad \text{for } i = 1, 2,$$

as we claimed. □

For the rest of this section, put $Z = \{0, 1, \dots, p-1\}^{\mathbb{N}_0}$ and let ν denote the product measure of probability counting measure on $\{0, 1, \dots, p-1\}$. We will apply Lemma 62 with (Z, ν) and $(X_1, \mu_1) = (\mathbb{R}/\mathbb{Z}, m_{\mathbb{R}/\mathbb{Z}})$ and $(X_2, \mu_2) = (\mathbb{Z}_p, m_{\mathbb{Z}_p})$.

We will also need the following to define the marginals.

Lemma 63. *Let $f_\infty : Z \to \mathbb{R}/\mathbb{Z}$ and $f_p : Z \to \mathbb{Z}_p$ be defined by*

$$f_\infty(\{x_i\}_{i=0}^\infty) := \left(\sum_{i=0}^\infty x_i p^{-i-1}\right) + \mathbb{Z}, \qquad \textit{and} \qquad f_p(\{x_i\}_{i=0}^\infty) := \sum_{i=0}^\infty x_i p^i.$$

Then $f_{\infty,*}\nu = m_{\mathbb{R}/\mathbb{Z}}$ and $f_{p,*}\nu = m_{\mathbb{Z}_p}$.

Proof. First, we observe that f_∞ and f_p are continuous surjective functions; in particular, they are measurable functions. We also notice that f_p is injective, therefore it is a homeomorphism. As for f_∞, note that $f_\infty^{-1}(x + \mathbb{Z})$ has one element unless x is of the form $\frac{j}{p^k}$ in which case $f_\infty^{-1}(x + \mathbb{Z})$ has two elements. Therefore, there are ν-null subset $\mathcal{N}_1 \subseteq Z$ and $m_{\mathbb{R}/\mathbb{Z}}$-null subset $\mathcal{N}_2 \subseteq \mathbb{R}/\mathbb{Z}$, so that $f_\infty : Z \setminus \mathcal{N}_1 \to (\mathbb{R}/\mathbb{Z}) \setminus \mathcal{N}_2$ is a bijective measurable map whose inverse is also measurable.

To prove the lemma for f_∞, it suffices to give a generating set $\mathcal{U}$ of the Borel σ-algebra of $\mathbb{R}/\mathbb{Z}$ such that for every $U \in \mathcal{U}$, the sets $U \setminus \mathcal{N}_2$ and $f_\infty^{-1}(U \setminus \mathcal{N}_2)$ have the same measure. The proof for f_p is similar. For $i \in \mathbb{N}_0$ and $0 \le j \le p^{i+1} - 1$, let

$$U_{i,j} := \left(\frac{j}{p^{i+1}}, \frac{j+1}{p^{i+1}}\right) + \mathbb{Z}.$$

Notice that $\mathcal{U} := \{U_{i,j} \mid i \in \mathbb{N}_0, 0 \le j \le p-1\}$ generates the Borel σ-algebra on $\mathbb{R}/\mathbb{Z}$, and

$$f_\infty^{-1}(U_{i,j}) = \left(\{a_0\} \times \cdots \times \{a_i\} \times \prod_{k=i+1}^\infty \{0, \dots, p-1\}\right) \setminus \mathcal{N}_1$$

where $j = a_0 + a_1 p + \cdots + a_i p^i$ and $a_0, \dots, a_i \in [0, p-1]$. Therefore, both $U_{i,j}$ and $f_\infty^{-1}(U_{i,j})$ have measure $\frac{1}{p^{i+1}}$. Altogether, we deduce that f_∞ is measure-preserving, which means $f_{\infty,*}\nu = m_{\mathbb{R}/\mathbb{Z}}$. The f_p case is similar. □

We note that for any permutation $\sigma : \mathbb{N}_0 \to \mathbb{N}_0$, the induced map $h_\sigma : Z \to Z$ is a homeomorphism which preserves ν. Therefore, combining Lemma 62 and Lemma 63, we have

Corollary 64. *Let $\sigma : \mathbb{N}_0 \to \mathbb{N}_0$ be a permutation, and let $\delta_\sigma : Z \to \mathbb{R}/\mathbb{Z} \times \mathbb{Z}_p$ be*

$$\delta_\sigma(z) = \big(f_\infty(z), f_p(h_\sigma(z))\big),$$

and let $\mu_\sigma = \delta_{\sigma,*}\nu$. Then μ_σ is a coupling of $m_{\mathbb{R}/\mathbb{Z}}$ and $m_{\mathbb{Z}_p}$. □

Proof of Theorem 60. Let $\sigma : \mathbb{N}_0 \to \mathbb{N}_0$ be the permutation fixing 0 and 1 and inverting each block $\{2^j, \dots, 2^{j+1} - 1\}$. More precisely, all $j \ge 1$ and $0 \le i \le 2^j - 1$ we have $\sigma(i + 2^j) = 2^{j+1} - i - 1$. Let $\mu = \mu_\sigma$ be as in Corollary 64 applied with this σ. Then μ is a coupling of the probability Haar measures on $\mathbb{R}/\mathbb{Z}$ and $\mathbb{Z}_p$ supported on

$$\big\{(\textstyle\sum_{i=0}^\infty x_i p^{-(i+1)}, \sum_{i=0}^\infty y_i p^i) \in \mathbb{R}/\mathbb{Z} \times \mathbb{Z}_p | x_{i+2^j} = y_{2^{j+1}-i-1}, \text{for all } j \ge 1,\ 0 \le i \le 2^j - 1\big\}. \tag{264}$$

We will use Lemma 61 to show that μ does not have spectral gap, which will finish the proof of the Theorem as μ is a coupling of the probability Haar measures on $\mathbb{R}/\mathbb{Z}$ and $\mathbb{Z}_p$.

To apply Lemma 61, we construct a family of characters for $\mathbb{R}/\mathbb{Z}$ and $\mathbb{Z}_p$. Let

$$\mathrm{e} : \mathbb{R} \to \mathbb{C}^\times, \mathrm{e}(x) := e^{2\pi i x},$$

and recall that

$$\widehat{\mathbb{R}/\mathbb{Z}} = \{e_{\infty,n} \mid n \in \mathbb{Z}\},$$

where $e_{\infty,n}(x + \mathbb{Z}) := \mathrm{e}(nx)$ — notice that $e_{\infty,n}$ is well-defined as e is $\mathbb{Z}$-invariant.

To describe the dual of $\mathbb{Z}_p$, we recall that every element $x \in \mathbb{Q}_p$ can be written as a sum of a rational number r_x and a p-adic integer z_x. Let

$$e_p : \mathbb{Q}_p \to \mathbb{C}^\times, \quad e_p(x) := \mathrm{e}(r_x),$$

and notice that e_p does not depend on the choice of decomposition $x = r_x + z_x$ as e is $\mathbb{Z}$-invariant. Then

$$\widehat{\mathbb{Z}_p} = \big\{e_{p,r} \mid r = \tfrac{k}{p^j}, k, j \in \mathbb{Z}\big\},$$

where $e_{p,r} : \mathbb{Z}_p \to \mathbb{C}^\times, e_{p,r}(x) = e_p(rx)$.

For $j \geq 1$, define $\alpha_j := e_{\infty,p^{2^j}}$, $\beta_j := e_{p,p^{-2^{j+1}}}$, and

$$\gamma_j : \mathbb{R}/\mathbb{Z} \times \mathbb{Z}_p \to \mathbb{C}^\times, \qquad \gamma_j(x, y) := \alpha_j(x)/\beta_j(y).$$

We claim that for μ-a.e. (x, y) we have $\gamma_j(x, y) \to 1$. By (264), we have that if (x, y) is in the support of μ, then there are digits $x_i, y_i \in \{0, 1, \ldots, p-1\}$ such that $x = \sum_{i=0}^\infty x_i p^{-(i+1)}$ and $y = \sum_{i=0}^\infty y_i p^{i+1}$, where

$$y_{2^{j+1}-i-1} = x_{2^j+i} \qquad \text{for all } j \geq 1 \text{ and } 0 \leq i < 2^j. \tag{265}$$

Then for all j

$$\begin{aligned} \alpha_j(x) &= e_{\infty,p^{2^j}}(x) = \mathrm{e}\Big(\sum_{i=0}^\infty x_i p^{2^j-i-1}\Big) \\ &= \mathrm{e}\Big(\sum_{i=0}^\infty x_{i+2^j} p^{-(i+1)}\Big) = \mathrm{e}\Big(\sum_{i=0}^{2^j-1} x_{i+2^j} p^{-(i+1)}\Big) + O(p^{-2^j}). \end{aligned} \tag{266}$$

Similarly, we have

$$\begin{aligned} \beta_j(y) &= e_{p,p^{-2^{j+1}}}(y) = e_p\Big(\sum_{i=0}^\infty y_i p^{i+1-2^{j+1}}\Big) \\ &= \mathrm{e}\Big(\sum_{i=0}^{2^{j+1}-1} y_i p^{i+1-2^{j+1}}\Big) = \mathrm{e}\Big(\sum_{i=0}^{2^j-1} y_{2^{j+1}-i-1} p^{-(i+1)}\Big) + O(p^{-2^j}). \end{aligned} \tag{267}$$

Hence, by (265), (266), and (267), we obtain that

$$|\gamma_j(x, y) - 1| = |\alpha_j(x) - \beta_j(y)| = O(p^{-2^j}).$$

Therefore by Lemma 61, we deduce that μ, which is a coupling of the probability Haar measures of $\mathbb{R}/\mathbb{Z}$ and $\mathbb{Z}_p$, does not have spectral gap property. Thus, $\mathbb{R}/\mathbb{Z}$ and $\mathbb{Z}_p$ are not spectrally independent. □

Appendix D. Commutator of small neighborhoods in compact semisimple Lie groups.

In [12] it is proved that if G is a perfect connected compact group, then every element of G is a commutator element. More recently in [5], it is proved that the commutator map

$$\psi : G \times G \to G, \psi(g_1, g_2) := [g_1, g_2] = g_1 g_2 g_1^{-1} g_2^{-1}$$

is an open function if G is a compact semisimple Lie group. Here we show the following quantitative version of their result. For the case of unitary groups, this is part of the Solovay-Kitaev algorithm.

Proposition 65. *Suppose G is a compact semisimple Lie group. Suppose $G \subseteq \mathrm{SO}(n)$, and it is equipped with the metric induced by the operator norm. Let $\psi : G \times G \to G, \psi(g_1, g_2) := [g_1, g_2]$ be the commutator map. Then there is a positive number $c'' := c''(G)$ such that for every $0 < \rho_1, \rho_2 < 1$ we have $\psi(1_{\rho_1} \times 1_{\rho_2}) \supseteq 1_{c''\rho_1\rho_2}$.*

Proof. Let $\mathfrak{g}$ be the Lie algebra of G, and $\overline{\psi} : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}, \overline{\psi}(x, y) := [x, y]$. By [5, Theorem 2.1], there is a positive number c_1'' such that $\overline{\psi}(0_1 \times 0_1) \supseteq 0_{c_1''}$. Since $\overline{\psi}$ is bilinear, for every positive numbers ρ and ρ' we have

$$\overline{\psi}(0_\rho \times 0_{\rho'}) \supseteq 0_{c_1''\rho\rho'}. \tag{268}$$

Let $\phi(t) := \frac{e^t - 1}{t}$. Notice that ϕ is an analytic function, $\phi(0) = 1$, and $|\phi(t) - 1| \leq \frac{te^t}{2}$ for every $0 < t < 1$. We set

$$\xi : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}, \quad \xi(x, y) := \exp(\mathrm{ad}(y))(x) - x,$$

and we notice that

$$\begin{aligned} \xi(x, y) =& (\exp(\mathrm{ad}(y)) - \mathrm{id})(x) = \mathrm{ad}(y)(\phi(\mathrm{ad}(y))(x)) \\ =& [y, \phi(\mathrm{ad}(y))(x)] = \psi(y, \phi(\mathrm{ad}(y))(x)). \end{aligned} \tag{269}$$

For every $0 < \rho \ll_{\dim G} 1$, $0 < \rho' < 1$, and $y \in 0_\rho$, $\phi(\mathrm{ad}(y))(0_{\rho'}) \supseteq 0_{\rho'/2}$. Therefore by (268) and (269), we obtain

$$\xi(0_\rho \times 0_{\rho'}) \supseteq 0_{\frac{1}{2}c_1''\rho\rho'} \tag{270}$$

for every $0 < \rho < 1$ and $0 < \rho' \ll_{\dim G} 1$. Using the Baker-Campbell-Hausdorff formula, it is deduced in [5, Proposition 3.1] that there are analytic functions P and Q from a neighborhood $\mathcal{O}$ of $(0, 0) \in \mathfrak{g} \times \mathfrak{g}$ to G such that

$$P(0, 0) = Q(0, 0) = 1, \ \exp(x + y) = \exp(\mathrm{Ad}(P(x, y))(x)) \exp(\mathrm{Ad}(Q(x, y))(y)), \tag{271}$$

for every $(x, y) \in \mathcal{O}$. By (271), we have

$$\begin{aligned} \exp(-\xi(x, y)) =& [A, B], \quad \text{where} \\ A :=& P(x, -\exp(\mathrm{ad}(y))(x)) \exp(x) P(x, -\exp(\mathrm{ad}(y))(x))^{-1} \text{ and} \\ B :=& Q(x, -\exp(\mathrm{ad}(y))(x)) \exp(y) P(x, -\exp(\mathrm{ad}(y))(x))^{-1}. \end{aligned} \tag{272}$$

For $x \in 0_\rho$, we have $\| -\exp(\mathrm{ad}(y))(x)\| = \| \mathrm{Ad}(\exp(y))(x)\| \leq \rho$. Since P and Q are analytic and $P(0, 0) = Q(0, 0) = 1$, for $0 < \rho \ll_G 1$ and $x \in 0_\rho$ we have

$$P := P(x, -\exp(\mathrm{ad}(y))(x)) \in 1_{C''\rho} \quad \text{and} \quad Q := Q(x, -\exp(\mathrm{ad}(y))(x)) \in 1_{C''\rho}, \tag{273}$$

for some $C'' := C''(G)$. By (273), for $x \in 0_\rho$ and $y \in 0_{\rho'}$, we have

$$A \in 1_{2\rho} \quad \text{and} \quad B \in 1_{4C''\rho + 2\rho'}. \tag{274}$$

By (270), (272), and (274), we deduce that the following holds:

$$[1_{2\rho}, 1_{4C''\rho+2\rho'}] \supseteq 1_{\frac{1}{4}c_1''\rho\rho'}.$$

For every $0 < \rho \ll_G 1$ and $\rho \le \rho'' \ll 1$, we have

$$[1_\rho, 1_{\rho'}] \supseteq [1_{2\frac{\rho}{8C''}}, 1_{4C''\frac{\rho}{8C''}+2\frac{\rho'}{4}}] \supseteq 1_{\frac{c_1''}{128C''}\rho\rho'}.$$

This completes the proof. □

References

[1] Yves Benoist and Nicolas de Saxcé. A spectral gap theorem in simple Lie groups. *Invent. Math.*, 205(2):337–361, 2016.

[2] Jean Bourgain and Alex Gamburd. Uniform expansion bounds for Cayley graphs of $\mathrm{SL}_2(\mathbb{F}_p)$. *Ann. of Math. (2)*, 167(2):625–642, 2008.

[3] Jean Bourgain, Alex Gamburd, and Peter Sarnak. Affine linear sieve, expanders, and sum-product. *Invent. Math.*, 179(3):559–644, 2010.

[4] Jean Bourgain and Péter P. Varjú. Expansion in $SL_d(\mathbf{Z}/q\mathbf{Z})$, q arbitrary. *Invent. Math.*, 188(1):151–173, 2012.

[5] Alessandro D'Andrea and Andrea Maffei. Commutators of small elements in compact semisimple groups and Lie algebras. *J. Lie Theory*, 26(3):683–690, 2016.

[6] Pierre de la Harpe, A. Guyan Robertson, and Alain Valette. On the spectrum of the sum of generators for a finitely generated group. *Israel J. Math.*, 81(1-2):65–96, 1993.

[7] J.D. Dixon, M.P.F. Du Sautoy, A. Mann, and D. Segal. *Analytic Pro-P Groups*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2003.

[8] Ilijas Farah. Approximate homomorphisms. *Combinatorica*, 18(3):335–348, 1998.

[9] Ilijas Farah. Approximate homomorphisms. II. Group homomorphisms. *Combinatorica*, 20(1):47–60, 2000.

[10] Giacomo Gigante and Paul Leopardi. Diameter bounded equal measure partitions of Ahlfors regular metric measure spaces. *Discrete Comput. Geom.*, 57(2):419–430, 2017.

[11] A. Salehi Golsefidy and Péter P. Varjú. Expansion in perfect groups. *Geom. Funct. Anal.*, 22(6):1832–1891, 2012.

[12] Morikuni Gotô. A theorem on compact semi-simple groups. *J. Math. Soc. Japan*, 1:270–272, 1949.

[13] Marvin J. Greenberg. Strictly local solutions of Diophantine equations. *Pacific J. Math.*, 51:143–153, 1974.

[14] Karsten Grove, Hermann Karcher, and Ernst A. Ruh. Group actions and curvature. *Invent. Math.*, 23:31–48, 1974.

[15] Weikun He and Nicolas de Saxcé. Trou spectral dans les groupes simples, 2021.

[16] D. Kazhdan. On ε-representations. *Israel J. Math.*, 43(4):315–323, 1982.

[17] Victor Klee and Christoph Witzgall. Facets and vertices of transportation polytopes. In *Mathematics of the Decision Sciences, Part 1 (Seminar, Stanford, Calif., 1967)*, pages 257–282. Amer. Math. Soc., Providence, R.I., 1968.

[18] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.

[19] Alexander Lubotzky. *Discrete groups, expanding graphs and invariant measures*. Modern Birkhäuser Classics. Birkhäuser Verlag, Basel, 2010. With an appendix by Jonathan D. Rogawski, Reprint of the 1994 edition.

[20] Keivan Mallahi-Karai, Amir Mohammadi, and Alireza Salehi Golsefidy. Locally random groups, 2020.

[21] G. A. Margulis. Explicit constructions of expanders. *Problemy Peredači Informacii*, 9(4):71–80, 1973.

[22] G. A. Margulis. Explicit group-theoretic constructions of combinatorial schemes and their applications in the construction of expanders and concentrators. *Problemy Peredachi Informatsii*, 24(1):51–60, 1988.

[23] A. L. Onishchik and È. B. Vinberg. *Lie groups and algebraic groups*. Springer Series in Soviet Mathematics. Springer-Verlag, Berlin, 1990. Translated from the Russian and with a preface by D. A. Leites.

[24] Alireza Salehi Golsefidy. Super-approximation, I: $\mathfrak{p}$-adic semisimple case. *Int. Math. Res. Not. IMRN*, 2017(23):7190–7263, 2017.

[25] Alireza Salehi Golsefidy. Super-approximation, II: the p-adic case and the case of bounded powers of square-free integers. *J. Eur. Math. Soc. (JEMS)*, 21(7):2163–2232, 2019.

[26] Pablo Solernó. Effective łojasiewicz inequalities in semialgebraic geometry. *Appl. Algebra Engrg. Comm. Comput.*, 2(1):2–14, 1991.

[27] Péter P. Varjú. Expansion in $\mathrm{SL}_d(\mathcal{O}_K/I)$, I square-free. *J. Eur. Math. Soc. (JEMS)*, 14(1):273–305, 2012.

Mathematics Department, University of California, San Diego, CA 92093-0112, USA
Email address: golsefidy@ucsd.edu

Constructor University, Campus Ring I, 28759, Bremen, Germany
Email address: kmallahikarai@constructor.university

Mathematics Department, University of California, San Diego, CA 92093-0112, USA
Email address: ammohammadi@ucsd.edu