LOCALLY RANDOM GROUPS

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

Contents

Abstract. In this work, we will introduce and study the notion of *local randomness* for compact metric groups. We prove a mixing inequality as well as a product result for locally random groups under an additional *dimension condition* on the volume of small balls, and provide several examples of such groups. In particular, this leads to new examples of groups satisfying such a mixing inequality. In the same context, we will develop a Littlewood-Paley decomposition and explore its connection to the existence of spectral gap for random walks. Moreover, under the dimension condition alone, we will prove a multi-scale entropy gain result à la Bourgain-Gamburd and Tao.

1. Introduction

The aim of this work is to introduce and study the notion of *local randomness* for the class of compact metric groups. As the name suggests, this notion aims at capturing a certain form of randomness exhibited by these groups. Before proceeding to the precise definition of this notion, let us make a few general remarks on the terminology and motivations behind the definition.

The notion of randomness is often understood as the lack of low-complexity structure. One approach towards defining randomness is *statistical randomness*. Roughly speaking, statistical randomness requires the putative random (sometimes called pseudo-random) object to pass certain randomness tests, which are passed by truly random objects. Quasi-random graphs, introduced by Chung, Graham, and Wilson [11] are examples of this kind. For instance, in such a graph the

K.M-K was partially supported by the DFG grant DI506/14-1.
A.M was partially supported by the NSF.
A.SG was partially supported by the NSF.

number of edges connecting subsets A, B of vertices is close to $\delta|A|\ |B|$, mimicking the typical behavior of Erdös-Rényi random graphs with density δ.

An alternative approach towards defining randomness is based on the non-existence of low-complexity models. In taking up such an approach, one needs to clarify what a model means and how its complexity is measured. Quasi-random groups, as named by Gowers, provide examples for this approach. Recall that a finite group G is said to be K-quasi-random when it admits no non-trivial unitary representations of degree less than K. If one views a unitary representation of a finite group as a *model* and its degree as its complexity, then qausi-random groups are precisely groups without low-complexity models.

One of the main results of Gowers's work, intertwining these two approaches, is that Cayley graphs of quasi-random groups with respect to *large* generating sets yield quasi-random graphs in the sense of Chung, Graham, and Wilson. This is based on a mixing inequality established in [16], and generalized in [1]. Let us remark that, prior to [16], the quasi-randomness had been implicitly exploited by Sarnak-Xue [20], Gamburd [13], and Bourgain-Gamburd [8].

In the present work we will define the notion of local randomness for a compact group G equipped with a compatible bi-invariant metric d by means of an inequality of the form

$$\|\pi(x)-\pi(y)\|_{\mathrm{op}} \leq C_0(\dim \pi)^L d(x,y) \tag{1.1}$$

where C_0 and L are parameters and π varies over unitary representation of G; see Definition 2.1 for the precise definition. The relation between this inequality and the non-existence of low-complexity models for G can be understood as follows. Consider an η-discretization of G, that is, a maximal set of points in G that are pairwise η-apart. From (1.1) it follows that for a unitary representation π of G (a model) to map these points to matrices that are pairwise at distance $\eta^{1-\epsilon}$, $\dim \pi$ needs to be polynomially large in η^{-1}. Thus, a group satisfying (1.1) fails to have a low complexity discretized model.

As the above definition indicates, local randomness of a compact group depends on the choice of a compatible metric. In Corollary 5.6, we show that semisimple Lie groups are exactly those groups that are locally random with respect to *all* compatible metrics. In contrast to this, compact groups that are locally random with respect to *some* compatible metric are exactly those with finitely many non-equivalent irreducible representations of a given degree, see Theorem 2.3.

One of the main properties of locally random groups is the mixing inequality proved in Theorem 2.6. This can be seen as an instance of statistical randomness and a multi-scale analogue of the mixing inequality alluded to above. This inequality is much more fruitful in the presence of a *dimension condition*, see (DC). In particular, it will enable us to prove a *product result*, Theorem 2.8, for subsets with large metric entropy, a result that can be best understood as a multi-scale version of Gowers's product theorem.

In order to study the behavior of random walks on locally random groups, we adapt the Littlewood-Paley theory [7, 10] to this context. As an application, we will show that the study of spectral gap for random walks on G can be reduced to that of *functions living at small scale*; see Theorem 2.10 and Theorem 9.3.

Notable examples of groups to which our results apply include finite products of perfect real and p-adic analytic compact Lie groups. In the special case of profinite groups, local randomness is intimately connected to the notion of quasi-randomness introduced and studied in [24]; see Proposition 5.9 for precise statements. It is also worth mentioning that inequality (1.1) has been implicitly used in [12] to establish the existence of a dimension gap for Borelean subgroups of compact Lie groups.

Our last theorem, Theorem 2.12, is an entropy gaining result in the spirit of a major ingredient of the Bourgain-Gamburd expansion machine. Roughly speaking, this theorem asserts that when X and Y are independent G-valued random variables, the Rényi entropy of XY at scale η is larger

than the average of the Rényi entropies of X and Y at scale η by a definite amount, unless algebraic obstructions exist. This can be viewed as a weighted version of Tao's result [22] and a common extension of [7, 14, 5, 10].

In a forthcoming work, we shall use Theorems 2.12 and 9.3 in proving the *spectral independence* of open compact subgroups of two non-locally isomorphic analytic simple Lie groups over local fields of characteristic zero.

This paper is structured as follows. In Section 2, we will review some basic definitions, set some notation and state the main results of the paper. In Section 3, we gather a number of basic tools, ranging from abstract harmonic analysis to notions related to metric spaces. Sections 4 and 5 feature prominent examples and fundamental properties of locally random groups. In Section 6, we will prove a number of mixing properties for locally random groups, which will be employed in Section 7 to show the product theorem. In Section 8, we will discuss in detail a Littlewood-Paley decomposition of locally random groups. The connection to the spectral gap, stated in Theorem 2.10, is established in Section 9. Finally, in Section 10, we will prove Theorem 2.12.

Acknowledgement. The authors would like to thank Péter Varjú for helpful comments. K.M-K and A.M also thank Bernoulli Center in Lausanne for its hospitality. K.M-K would like to thank the Department of Mathematics at UCSD where part of the research was conducted.

2. Basic definitions and statement of results

In this section, we will state the main results of the paper. Let us begin by defining the notion of local randomness.

Definition 2.1. Suppose G is a compact group and d is a compatible bi-invariant metric on G. For parameters $C_0 \geq 1$ and $L \geq 1$ 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:

$$\|\pi(x) - \pi(y)\|_{\mathrm{op}} \leq C_0(\dim \pi)^L d(x, y). \tag{2.1}$$

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.

Remark 2.2.
(1) It is a standard fact that every second countable compact group can be equipped with a compatible bi-invariant metric.
(2) One can easily check that (2.1) only depends on the unitary isomorphism class of π.
(3) In the rest of the paper, we will drop d from the notation and use the phrase *G is L-locally random with coefficient C_0*. Often, the implicit metric d is a standard metric on G.

Our first theorem gives a characterization of locally random groups in terms of their unitary dual.

Theorem 2.3 (Characterization)**.** *Suppose G is a compact second countable group. Then G is locally random if and only if G has only finitely many non-isomorphic irreducible representations of a given degree.*

In [3] it is proved that a finitely generated profinite group has only finitely many irreducible representations of a given degree if and only if G has the FAb property, that is, every open subgroup of G has finite abelianization. In view of Theorem 2.3, the group $\prod_{n\geq 1} \mathrm{SU}(2)$ is not a locally random group, but $\prod_{n\geq 2} \mathrm{SU}(n)$ is a locally random group.

For $\eta > 0$ and $x \in G$, denote the open ball of radius η centered at x by x_η. The L^1-normalized indicator function of the ball 1_η is denoted by $P_\eta := \frac{\mathbb{1}_{1_\eta}}{|1_\eta|}$, where $|\cdot|$ denotes the Haar measure. For $f \in L^1(G)$ and a probability measure μ on G, we write $f_\eta = f * P_\eta$ and $\mu_\eta = \mu * P_\eta$, see (3.1) and (3.2) for the definition of convolution.

Definition 2.4. 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 $C_1 \geq 1$ and $d_0 > 0$ such that for all $\eta \in (0, 1)$ the following bounds hold.

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

Remark 2.5. (1) Measures satisfying this condition is also known as Ahlfors (or Ahlfors-David) d_0-regular measures.
(2) Whenever d is clear from the context, we suppress d from the notation and simply write that G satisfies a dimension condition $\mathrm{DC}(C_1, d_0)$.

Our second theorem shows that local randomness is particularly effective in the presence of a dimension condition.

Theorem 2.6 (Scaled mixing inequality)**.** *Suppose G is an L-locally random group with coefficient C_0. Then for every $f, g \in L^2(G)$ we have*

$$\|f * g\|_2^2 \leq 2\|f_\eta * g_\eta\|_2^2 + \eta^{1/(2L)}\|f\|_2^2\|g\|_2^2$$

so long as $C_0\sqrt{\eta} \leq 0.1$.

Similar statements for finite groups, simple Lie groups and perfect Lie groups have been established thanks to work of many authors, see e.g. [16, 1, 6, 10, 4].

Definition 2.7. Suppose X is a metric space and $A \subseteq X$. For $\eta \in (0, 1)$, $\mathcal{N}_\eta(A)$ denotes the least number of open balls of radius η with centers in A required to cover A. The metric entropy of A at scale η is defined by

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

Theorem 2.8 (Product theorem for locally random groups)**.** *Suppose G is an L-locally random group with coefficient C_0. Suppose G satisfies the dimension condition $\mathrm{DC}(C_1, d_0)$. Then for every $\varepsilon > 0$ and every $\delta \ll_{L,d_0} \varepsilon$ the following holds: for all $\eta > 0$ and $A, B \subseteq G$ satisfying*

$$\frac{h(A; \eta) + h(B; \eta)}{2} \geq (1 - \delta)h(G; \eta)$$

and $\eta^\varepsilon \ll_{L,C_0,C_1,d_0} 1$, we have

$$A_\eta B_\eta B_\eta^{-1} A_\eta^{-1} \supseteq 1_{\eta^\varepsilon}.$$

Definition 2.9. Suppose G is a compact group and μ is a symmetric Borel probability measure. Denote by T_μ the convolution operator on $L^2(G)$ mapping f to $\mu * f$. For a subrepresentation $(\pi, \mathcal{H}_\pi)$ of $L_0^2(G)$, we let

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

Given a G-valued random variable X, we define the *Rényi entropy of X at scale η* by

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

where μ is the distribution (or the law) of X.

Theorem 2.10. *Suppose G is an L-locally random group with coefficient C_0. Suppose G satisfies $\mathrm{DC}(C_1, d_0)$. Then there exist $\eta_0 > 0$ small enough depending on the parameters and a subrepresentation $\mathcal{H}_0$* (exceptional subspace) *of $L^2(G)$ such that the following statements hold.*

(1) (dimension bound) $\dim \mathcal{H}_0 \leq 2C_0\eta_0^{-d_0}$.

(2) (spectral gap) *Let μ be a symmetric Borel probability measure whose support generates a dense subgroup of G. Let $a > \max(4Ld_0, 4L+2)$, and for $i \geq 1$ set $\eta_i := \eta_0^{a^i}$. If for constant $C_2 > 0$ and for every positive integer i there exists an integer $l_i \leq C_2 h(G;\eta_i)$ such that*

$$(\textit{Large entropy at scale } \eta_i) \qquad H_2(\mu^{(l_i)};\eta_i) \geq \Big(1 - \frac{1}{20Ld_0a^3}\Big)h(G;\eta_i),$$

then

$$\mathcal{L}(\mu; L^2(G) \ominus \mathcal{H}_0) \geq \frac{1}{40C_2Ld_0a^3}.$$

In particular, $\mathcal{L}(\mu; L_0^2(G)) > 0$.

Finally, we prove a multi-scale entropy gain result which is in the spirit of [9, Lemma 2.1] by Bourgain and Gamburd, and is a weighted version of [22, Theorem 6.10] by Tao. More details on the background of this result will be mentioned in Section 10. Before we state this result, we recall the definition of an approximate subgroup.

Definition 2.11. A subset X of a group G is called a *K-approximate subgroup* if X is a symmetric subset, that is, $X = X^{-1}$, and there exists subset $T \subseteq X \cdot X$ such that $\#T \leq K$ and $X \cdot X \subseteq T \cdot X$.

Theorem 2.12. *Suppose G is a compact group which satisfies the dimension condition at scale η, that is,*

$$C^{-1}\eta^{d_0} \leq |1_{a\eta}| \leq C\eta^{d_0}$$

holds for all $a \in [C'^{-1}, C']$, where $C > 1, C' \gg 1, d_0 > 0$ are fixed numbers. Suppose X and Y are independent Borel G-valued random variables. If

$$H_2(XY;\eta) \leq \log K + \frac{H_2(X;\eta) + H_2(Y;\eta)}{2}$$

for some positive number $K \geq (C2^{d_0})^{O(1)}$, then there are $H \subseteq G$ and $x, y \in G$ satisfying the following properties:

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

$$\mathbb{P}(XZ \in (xH)_\eta) \geq K^{-O(1)} \text{ and } \mathbb{P}(YZ \in (Hy)_\eta) \geq K^{-O(1)}.$$

Moreover,

$$|\{h \in H_\eta | \, \mathbb{P}(X \in (xh)_{3\eta}) \geq \widehat{C}K^{-10}2^{-H_2(X;\eta)}\}| \geq K^{-O(1)}|H_\eta|,$$

where $\widehat{C}$ is a constant of the form $(C2^{d_0})^{O(1)}$.

3. Preliminaries and notation

The purpose of this section is to provide the necessary definitions and fix the notation for the rest of the paper. For reader's convenience, these have been organized in two subsections.

Let G be a compact Hausdorff second countable topological group. It is well known that G can be equipped with a bi-invariant metric that induces the topology of G. Moreover, there exists a unique bi-invariant probability measure defined on the Borel σ-algebra of G, called the Haar measure. For a Borel measurable subset A of G, the Haar measure of A is denoted by $m_G(A)$ or $|A|$. 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 (of equivalence classes) of complex-valued functions f on G satisfying $\int_G |f(x)|^p\,\mathrm{d}x < \infty$. For $f \in L^p(G)$, we write

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

We will 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

$$(f * g)(x) = \int_G f(y)g(y^{-1}x)\,\mathrm{d}y. \tag{3.1}$$

It is a fact that $(L^1(G), +, *)$ is a unital Banach algebra and if $f \in L^1(G)$ is a class function, then f is in the center of this Banach algebra. Note also that $L^2(G)$ is naturally equipped with the inner product defined by $\langle f, g\rangle = \int_G f\overline{g}$ is a Hilbert space.

When $\mathscr{H}$ is a Hilbert space and $T : \mathscr{H} \to \mathscr{H}$ is a bounded linear operator, we will 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

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

For a Hilbert space $\mathscr{H}$, we write $\mathrm{U}(\mathscr{H})$ for the group of unitary operators of $\mathscr{H}$. A homomorphism $\pi : G \to \mathrm{U}(\mathscr{H})$ is continuous if the map

$$G \times \mathscr{H} \to \mathscr{H}, \qquad (g, v) \mapsto g \cdot v$$

is continuous. A unitary representation of G (or sometimes called a G-representation) is a pair $(\mathscr{H}, \pi)$ consisting of a Hilbert space $\mathscr{H}$ and a continuous homomorphism $\pi : G \to \mathrm{U}(\mathscr{H})$. A closed subspace $\mathscr{H}' \subseteq \mathscr{H}$ is called G-invariant (or simply invariant when G is clear from the context) if for every $g \in G$ and every $v \in \mathscr{H}'$, one has $g \cdot v \in \mathscr{H}'$. A representation $(\mathscr{H}, \pi)$ is called irreducible when $\dim \mathscr{H} \ge 1$ and the only invariant subspaces are $\{0\}$ and $\mathscr{H}$ itself. The set of equivalence classes of irreducible unitary representations of G is called the unitary dual of G and is denoted by $\widehat{G}$. If $\mathscr{H}'$ is an invariant subspace of $\mathscr{H}$, we sometimes denote by $\mathscr{H} \ominus \mathscr{H}'$ the orthogonal complement of $\mathscr{H}'$ in H, which is itself an invariant subspace of $\mathscr{H}$. The set of vectors $v \in \mathscr{H}$ satisfying $\pi(g)v = v$ for all $g \in G$ is clearly a closed invariant subspace of $\mathscr{H}$ and is denoted by $\mathscr{H}^G$.

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 $L^2(G)$, which is called the regular representation of G.

Suppose μ and ν are Borel measures on G and $f \in L^1(G)$. The convolution $\mu * f$ is defined by

$$(\mu * f)(x) = \int_G f(y^{-1}x)\,\mathrm{d}\mu(y). \tag{3.2}$$

Similarly, the convolution $\mu * \nu$ is the probability measure on G is defined through its action on continuous functions via

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

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

$$\|f * g\|_2 \leq \|f\|_1 \, \|g\|_2, \quad \|f * g\|_\infty \leq \|f\|_2 \, \|g\|_2, \quad \|\mu * f\|_2 \leq \|f\|_2. \tag{3.3}$$

Let us enumerate a number of well-known facts about unitary representations of G. First, it is 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:

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

Finally, we will remark that G is abelian if and only if every $\pi \in \widehat{G}$ is one-dimensional. In this case, the above discussion reduces to the classical Fourier analysis on abelian groups.

In this subsection, we will collect a number of definitions from additive combinatorics that will be needed later. Let G be as above, and recall that d denotes a bi-invariant metric on G. The ball of radius $\eta > 0$ centered as $x \in G$ 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 probability measure on G, respectively) we write f_η (μ_η, respectively) instead of $f * P_\eta$ ($\mu * P_\eta$, respectively). The cardinality of a finite set A is denoted by $\#A$. 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 distribution measure of X. As $H_2(X; \eta)$ depends only on the distribution measure μ of X, we will sometimes write $H_2(\mu; \eta)$ instead of $H_2(X; \eta)$.

We will use Vinogradov's notation $A \ll_{c_1, c_2} B$ to denote that $A \leq CB$, where $C = C(c_1, c_2)$ is a positive function of c_1, c_2. We write $A \ll B$ to denote that $A \leq CB$, for some absolute constant $C > 0$. We similarly define $\gg_{c_1, c_2}$ and $\gg$ for the reverse relations.

4. Local randomness and representations with bounded dimension

The main goal of this section is to prove Theorem 2.3. Along the way some basic properties of locally random groups will also be proved.

Suppose $f : \mathbb{Z}^+ \to \mathbb{R}^+$ is a strictly increasing function, and define

$$d_{G,f}(x,y) := \sup_{\pi \in \widehat{G}} \frac{\|\pi(x) - \pi(y)\|_{\mathrm{op}}}{f(\dim \pi)}. \tag{4.1}$$

Note that $\frac{\|\pi(x)-\pi(y)\|_{\mathrm{op}}}{f(\dim \pi)}$ depends only on the (unitary) isomorphism class of π. In the sequel we often assume $\pi : G \to U(n)$ for some $n \in \mathbb{N}$. Moreover, we remark that if π is a finite dimensional unitary representation of G with the orthogonal decomposition $\pi = \oplus_{i\in I}\pi_i$ into irreducible representations, then

$$\frac{\|\pi(x) - \pi(y)\|_{\mathrm{op}}}{f(\dim \pi)} \leq \max_{i\in I} \frac{\|\pi_i(x) - \pi_i(y)\|_{\mathrm{op}}}{f(\dim \pi_i)} \leq d_{G,f}(x,y). \tag{4.2}$$

Lemma 4.1. *Suppose G is a compact group and $f : \mathbb{Z}^+ \to \mathbb{R}^+$ is a strictly increasing function. Let $d_{G,f}$ be defined as in (4.1); then $d_{G,f}$ is a well-defined bounded, bi-invariant metric on G.*

Proof. Since $\|\pi(x)\|_{\mathrm{op}} = 1$ for all $\pi \in \widehat{G}$ and all $x \in G$, we have that $d_{G,L}(x,y) \leq 2/f(1)$ for any $x, y \in G$ —we also used the fact that f is increasing. As $\pi(z)$ is a unitary matrix for any $z \in G$,

$$\|\pi(x) - \pi(y)\|_{\mathrm{op}} = \|\pi(zx) - \pi(zy)\|_{\mathrm{op}} = \|\pi(xz) - \pi(yz)\|_{\mathrm{op}}.$$

This implies $d_{G,f}$ is bi-invariant. Clearly $d_{G,f}$ satisfies the triangle inequality. By the Peter-Weyl theorem, if $x \neq y$, then there is $\pi \in \widehat{G}$ such that $\pi(x) \neq \pi(y)$. Hence, if $x \neq y$, then $d_{G,f}(x,y) \neq 0$, from which the claim follows. □

Next we want to explore the conditions under which the metric $d_{G,f}$ gives us the same topology as the original topology of G. Indeed it suffices to study neighborhoods of the identity.

Lemma 4.2. *In the above setting, $d_{G,f}$ induces the original topology of G if and only if*

$$\lim_{x\to 1} d_{G,f}(x,1) = 0.$$

Proof. In order to distinguish the two topologies on G, we let G_f denote the topological space whose point set is G and whose topology is generated by the metric $d_{G,f}$.

If G and G_f coincide, then $\lim_{x\to 1} d_{G,f}(x,1) = 0$.

Conversely, let $I_G : G \to G_f$ be the identity map. Since $d_{G,f}$ is bi-invariant, $\lim_{x\to 1} d_{G,f}(x,1) = 0$ implies $\lim_{x\to y} d_{G,f}(x,y) = 0$ for all $y \in G$. Hence I_G is continuous. Since G is compact and I_G is a continuous bijection, it is a homeomorphism; this finishes the argument. □

The following is a generalization of Theorem 2.3.

Theorem 4.3. *Suppose G is a compact group. The following statements are equivalent.*

1. *For any strictly increasing function $f : \mathbb{Z}^+ \to \mathbb{R}^+$, the metric $d_{G,f}$ induces the original topology of G.*
2. *For some strictly increasing function $f : \mathbb{Z}^+ \to \mathbb{R}^+$, the metric $d_{G,f}$ induces the original topology of G.*
3. *For any positive integer n, $\{\pi \in \widehat{G} |\ \dim \pi \leq n\}$ is finite.*

We start by proving that the second condition implies the *FAb condition*.

Definition 4.4. A compact group G has the *FAb property* if $H^{\mathrm{ab}} := H/\overline{[H,H]}$ is finite for any open subgroup H of G.

Lemma 4.5. *Suppose G is a compact group and for some strictly increasing function $f : \mathbb{Z}^+ \to \mathbb{R}^+$, the metric $d_{G,f}$ induces the original topology of G. Then G has the FAb property.*

Proof. We first show that $N/\overline{[N,N]}$ is finite for any *normal* open subgroup N of G.

Let $\xi \in \widehat{N}$, $\dim \xi = 1$, and $n \in \mathbb{Z}^+$, then $x \mapsto \xi(x)^n$ also defines a character of N; we denote this charatcter by $\xi_n \in \widehat{N}$.

Let $\pi_n := \mathrm{ind}_N^G(\xi_n)$. We can identify the space $\mathcal{H}_{\pi_n}$ of the representation π_n with

$$\mathbb{C}[G] \otimes_{\mathbb{C}[N]} \mathcal{H}_{\xi_n} = \bigoplus_{j=1}^{m} \mathbb{C}(g_j \otimes 1)$$

where $\mathbb{C}[G]$ and $\mathbb{C}[N]$ are the corresponding group rings and $\{g_j\}_{j=1}^m$ is a set of coset representatives of N with $g_1 = 1$; note that the inner product is induced from $\langle g_i \otimes 1, g_j \otimes 1\rangle = \delta_{ij}$, and for any $y \in N$ we have

$$\pi_n(y)(g_j \otimes 1) = g_j \otimes \xi_n(g_j^{-1} y g_j) = \xi(g_j^{-1} y g_j)^n (g_j \otimes 1).$$

Therefore, $\|\pi_n(y) - I\|_{\mathrm{op}} = \max_j |\xi(g_j^{-1} y g_j)^n - 1| \geq |\xi(y)^n - 1|$. By Lemma 4.2, we get that for any $\varepsilon > 0$ there exists $\eta > 0$ with the following property: for any $\xi \in \widehat{N}$ with dimension 1 and any $n \in \mathbb{Z}^+$ we have

$$|\xi(y)^n - 1| \leq f([G:N])\varepsilon \qquad \text{for all } y \in N \cap 1_\eta \tag{4.3}$$

see (4.2).

Note that if $\zeta \in \mathbb{S}^1 \setminus \{1\}$ is a norm 1 complex number that is not 1, then there is a positive integer n such that $|\zeta^n - 1| \geq \sqrt{3}$. Hence, (4.3) implies that if $\varepsilon < \sqrt{3}/f([G:N])$, then $\xi(x) = 1$ for $x \in N \cap 1_\eta$. Therefore, there is $\eta > 0$ such that $1_\eta \subseteq \ker \xi$ for all $\xi \in \widehat{N}$ that has dimension 1; thus, $\overline{[N,N]} = \bigcap_{\xi \in \widehat{N}, \dim \xi = 1} \ker \xi$ is an open subgroup of G. In particular, $N/\overline{[N,N]}$ is finite.

Suppose now that H is an arbitrary open subgroup of G; then G acts on the finite set G/H by the left multiplication. The kernel N of this action is an open normal subgroup of G. Since $\overline{[N,N]} \subseteq \overline{[H,H]}$ and $\overline{[N,N]}$ is an open subgroup, the claim follows. □

Lemma 4.6. *Suppose G is a compact group and $d_{G,f}$ induces the original topology of G for some strictly increasing function $f : \mathbb{Z}^+ \to \mathbb{R}^+$. Then G has only finitely many open subgroups of index at most n for any positive integer n.*

Proof. Suppose to the contrary that G has infinitely many open subgroups $\{H_i\}_{i=1}^\infty$ of index at most n. Let π_i be the representation of G on $L^2(G/H_i)$. Since $\pi_i(x)$ is a permutation for any $x \in G$, we have $\|\pi_i(x) - I\|_{\mathrm{op}} \geq 1$ for $x \notin \ker \pi_i$. And so for any $x \notin N := \bigcap_{i=1}^\infty \ker \pi_i$ we have $d_{G,f}(x, 1) \geq 1/f(n)$. Therefore by Lemma 4.2, we have that N is an open subgroup of G.

Since $\ker \pi_i \subseteq H_i$, the sequence $\{H_i/N\}_{i=1}^\infty$ consists of distinct subgroups of a finite group G/N, which is a contradiction. □

Next we prove a lemma on compact Lie groups which is essentially due to Platonov [18], and its proof has some similarities with the proof of the Schur-Zassenhaus theorem in finite group theory. We will provide a proof for the convenience of reader.

Lemma 4.7. *Let G be a closed subgroup of the unitary group $\mathrm{U}_n(\mathbb{C})$. Then there exists a finite subgroup $F \leq G$ such that $G = FG^\circ$.*

Proof. Let T be a maximal torus in G°. For every $g \in G$, the torus $T^g := gTg^{-1}$ is also a maximal torus in G°, and hence, by conjugacy of maximal tori, there exists $g_0 \in G^\circ$ such that $T^g = T^{g_0}$, or, equivalently, $g_0^{-1} g \in N_G(T)$. This establishes that $G = N_G(T)G^\circ$. In view of the fact that $\mathrm{Aut}(T) \simeq \mathrm{GL}_d(\mathbb{Z})$ is discrete, we have $N_G(T)^\circ \subseteq C_G(T)$. Moreover, as T is a maximal torus,

we have $C_G(T)^\circ = T$, and hence $N_G(T)^\circ = T$. From the compactness of $N_G(T)$, we obtain that $[N_G(T) : N_G(T)^\circ] < \infty$, and hence we have the following exact sequence:

$$1 \to T \to N_G(T) \to \overline{F} \to 1,$$

where $\overline{F}$ is a finite group. Since T is abelian, the conjugation action of $N_G(T)$ on T induces an action of $\overline{F}$ on T. For any $f \in \overline{F}$ and $t \in T$, we denote the action of f on t by t^f. That means if $s : \overline{F} \to N_G(T)$ is a section for the projection map from $N_G(T)$ to $\overline{F}$, then $t^f = s(f)ts(f)^{-1}$ for any $f \in \overline{F}$ and $t \in T$. For the section s, let $c(f_1, f_2) = s(f_1f_2)^{-1}s(f_1)s(f_2)$ for $f_1, f_2 \in \overline{F}$. Note that $c(\overline{F} \times \overline{F}) \subseteq T$, and we have

$$s(f_1f_2)c(f_1, f_2) = s(f_1)s(f_2).$$

From here one can verify the following 2-cocycle relation:

$$c(f_1, f_2f_3)c(f_2, f_3) = c(f_1f_2, f_3)^{f_3^{-1}}c(f_1, f_2). \tag{4.4}$$

Let $\alpha : \overline{F} \to T$ be defined in such a way that $\alpha(f)^{\#\overline{F}} = \prod_{f' \in \overline{F}} c(f', f)$. From (4.4) and the definition of α we have

$$\begin{aligned} \alpha(f_1f_2)^{\#\overline{F}} &= \prod_{f' \in \overline{F}} c(f', f_1f_2) = \prod_{f' \in \overline{F}} \Big(c(f'f_1, f_2)^{f_2^{-1}}c(f', f_1)c(f_1, f_2)^{-1}\Big) \\ &= \Big(\alpha(f_2)^{f_2^{-1}}\alpha(f_1)c(f_1, f_2)\Big)^{\#\overline{F}}. \end{aligned} \tag{4.5}$$

Let $T_{\overline{F}}$ be the subgroup of T consisting of all elements of order dividing $\#\overline{F}$; then from (4.5), it follows that there exists a map $\zeta : \overline{F} \times \overline{F} \to T_{\overline{F}}$ such that for all $f_1, f_2 \in \overline{F}$ we have

$$\alpha(f_1f_2) = \alpha(f_1)^{f_2^{-1}}\alpha(f_2)c(f_1, f_2)\zeta(f_1, f_2).$$

Now consider the modified section $\widetilde{s} : \overline{F} \to T$ defined by $\widetilde{s}(f) = s(f)\alpha(f)$. A simple computation shows that

$$\widetilde{s}(f_1f_2) = \widetilde{s}(f_1)\widetilde{s}(f_2)\zeta(f_1, f_2).$$

It follows that $F := \widetilde{s}(\overline{F})T_{\overline{F}}$ is a finite subset which is closed under multiplication, and hence a subgroup of G. Clearly, we have $N_G(T) = TF$. As $T \subseteq G^\circ$, $G = N_G(T)G^\circ$, and $N_G(T) = TF$, the claim follows. □

Lemma 4.8. *Let G_0 be a compact connected simple Lie group of adjoint type. Suppose G is a closed subgroup of $\prod_{i=1}^n G_0$ so that $\mathrm{pr}_i(G) = G_0$ for all i, where pr_i denotes the projection to the i-th component; and assume that for all $i \neq j$ and all $\theta \in \mathrm{Aut}(G_0)$ there exists $x \in G$ such that $\theta \circ \mathrm{pr}_i(x) \neq \mathrm{pr}_j(x)$. Then $G = \prod_{i=1}^n G_0$.*

Proof. We proceed by induction on n. The base of induction is clear. For any i, let $j_i : G_0 \to \prod_{i=1}^n G_0$ be the natural injection to the i-th component and $N_i := G \cap j_i(G_0)$.

Suppose contrary to the claim that $N_i \neq j_i(G_0)$ for some i. Without loss of generality we can and will assume that $N_1 \neq j_1(G_0)$. This and our assumption on G imply that N_1 is a proper normal subgroup of $j_1(G_0)$, hence, it is trivial.

We conclude that the projection $\mathrm{pr}_{[2..n]} : G \to \prod_{i=2}^n G_0$ to the components $2, \dots, n$ is injective. Let $H_1 := \mathrm{pr}_{[2..n]}(G) \subseteq \prod_{i=2}^n G_0$. Clearly H_1 satisfies the same properties as G, hence by the inductive hypothesis, we have $H_1 = \prod_{i=2}^n G_0$.

Let $\xi : \prod_{i=2}^n G_0 \to G$ be the inverse of the isomorphism $\mathrm{pr}_{[2..n]} : G \to H_1$. Let

$$\phi = \mathrm{pr}_1 \circ \xi : \prod_{i=2}^n G_0 \to G_0.$$

Then G can be identified with the graph of ϕ, and ϕ is onto. Since $\ker(\phi)$ is a normal subgroup of $\prod_{i=2}^n G_0$, it follows that

$$[\ker(\phi), j_i(G_0)] = j_i([\mathrm{pr}_i(\ker(\phi)), G_0]) \subseteq \ker(\phi).$$

As $G_0 = [G_0, G_0]$, we have $\ker(\phi) = \prod_{i\in I} G_0$ for some $I \subseteq [2..n]$.

This shows that $\prod_{i\in[2..n]\setminus I} G_0 \simeq G_0$, that is, $I = [2..n] \setminus \{i_1\}$ for some i_1, and

$$\phi|_{j_{i_1}(G_0)} : j_1(G_0) \to G_0$$

is an isomorphism. Hence, there is $\theta \in \mathrm{Aut}(G_0)$ such that $\theta(\pi_1(x)) = \mathrm{pr}_{i_1}(x)$ for any $x \in G$, which contradicts our assumption. □

Lemma 4.9. *Let G be a compact group and assume that $d_{G,f}$ induces the original topology of G for some strictly increasing function $f : \mathbb{Z}^+ \to \mathbb{R}^+$. Then G has only finitely many pairwise non-isomorphic irreducible representations of dimension n for any positive integer n.*

Proof. Suppose contrary to the claim that there are infinitely many pairwise non-isomorphic n-dimensional irreducible representations $\{\pi_i\}_{i=1}^\infty \subseteq \widehat{G}$.

Let $G_i := \pi_i(G) \subseteq \mathrm{U}_n(\mathbb{C})$. By Lemma 4.7, for every i there exists a finite subgroup $F_i \subseteq G_i$ such that $G_i = F_i G_i^\circ$. By Jordan's theorem, F_i contains a normal abelian subgroup A_i of bounded index, depending on n. Let $N_i = \pi_i^{-1}(A_i G_i^\circ)$. Then N_i is a normal open subgroup of G with $[G : N_i] \ll_n 1$. By Lemma 4.6, passing to a subsequence, we can and will assume that N_i is a fixed subgroup N for all $i \geq 1$. In particular, it follows that N surjects onto all $A_i/(A_i \cap G_i^\circ)$ for $i \geq 1$. By Lemma 4.5, N has a finite abelianization. Hence we must have $\sup_i [A_i : A_i \cap G_i^\circ] < \infty$, which implies that

$$\sup_i [G : \pi_i^{-1}(G_i^\circ)] \leq [G : N] \sup_i [A_i : A_i \cap G_i^\circ] < \infty. \tag{4.6}$$

As $H_i := \pi_i^{-1}(G_i^\circ)$ is an open subgroup for all i, by the assumption and (4.6), after passing to a subsequence, we may and will assume that H_i is the same subgroup H for all $i \geq 1$.

Altogether, we have proved that there is an open subgroup H of G such that $\pi_i(H)$ are connected subgroups of $\mathrm{U}_n(\mathbb{C})$ for all i. Since H has a finite abelianization, $\pi_i(H)$ are semisimple connected subgroups of $\mathrm{U}_n(\mathbb{C})$. There are only finitely many such subgroups, up to isomorphism. Hence, after passing to factors of $\pi_i(H)$ and a subsequence, we can and will assume that there is a compact connected simple Lie group of adjoint type G_0 such that $\pi_i(H) \simeq G_0$ for any i.

The order of the group $\mathrm{Aut}(G_0)/\mathrm{Inn}(G_0)$ is bounded by a function of n and π_i's are pairwise not G-conjugate; thus, after passing to a subsequence we can and will assume that for any $i \neq j$ and $\theta \in \mathrm{Aut}(G_0)$, there is $x \in H$ such that $\theta(\pi_i(x)) \neq \pi_j(x)$.

By Lemma 4.8, for any positive integer m,

$$\pi_{[1..m]} : H \to \prod_{i=1}^m G_0, \qquad \pi_{[1..m]}(x) := (\pi_i(x))_{i=1}^m \tag{4.7}$$

is an onto group homomorphism. Now let us consider the group homomorphism

$$\pi : H \to \prod_{i=1}^\infty G_0, \qquad \pi(x) := (\pi_i(x))_{i=1}^\infty.$$

For $x_0 \in G_0 \setminus \{1\}$, by (4.7), we get a sequence $\{g_m\}_{m=1}^\infty$ of elements of G such that

$$\pi(g_m) \in \prod_{i=1}^m \{1\} \times \{x_0\} \times \prod_{i=m+2}^\infty G_0; \tag{4.8}$$

and so $\lim_{m\to\infty}\pi(g_m)=1$. Since $\ker\pi$ is a compact subgroup, we can choose g_m's in a way that $\lim_{m\to\infty} g_m = 1$ and (4.8) holds. Let ϕ be a non-trivial irreducible representation of G_0, and let $\phi_i := \phi\circ \mathrm{pr}_i\circ\pi \in \widehat{G}$, where pr_i is the projection to the i-th component. Then by (4.8) we have

$$\phi_{m+1}(g_m) = \phi(x_0);$$

and so

$$d_{G,f}(g_m,1) \geq \frac{\|\phi_{m+1}(g_m)-I\|_{\mathrm{op}}}{f(\dim\phi_{m+1})} = \frac{\|\phi(x_0)-I\|_{\mathrm{op}}}{f(\dim\phi)} > 0.$$

On the other hand, by Lemma 4.2 and $\lim_{m\to\infty} g_m = 1$, we have

$$\lim_{m\to\infty} d_{G,f}(g_m,1) = 0,$$

which is a contradiction. □

Proof of Theorem 4.3. Clearly (1) implies (2). Lemma 4.9 proves that (2) implies (3). Next we want to show that (3) implies (1). Suppose $f:\mathbb{Z}^+\to\mathbb{R}^+$ is a strictly increasing function. By Lemma 4.2 we need to show that $\lim_{x\to 1} d_{G,f}(x,1)=0$. For a given $\varepsilon>0$, there are only finitely many representations $\{\pi_1,\dots,\pi_n\}\subset\widehat{G}$ such that $f(\dim\pi_i)<2/\varepsilon$. Hence, for all $\pi\in\widehat{G}\setminus\{\pi_1,\dots,\pi_n\}$ and all $x\in G$ we have

$$\frac{\|\pi(x)-I\|_{\mathrm{op}}}{f(\dim\pi)} \leq \frac{2}{f(\dim\pi)} \leq \varepsilon.$$

Since π_i's are continuous and G is compact, π_i's are uniformly continuous. And so there is $\eta>0$ such that for all $x\in 1_\eta$ we have

$$\|\pi_i(x)-I\|_{\mathrm{op}} \leq \varepsilon f(1)$$

for all $i\in[1..n]$. Altogether we get that for all $x\in 1_\eta$ and all $\pi\in\widehat{G}$ we have

$$\frac{\|\pi(x)-I\|_{\mathrm{op}}}{f(\dim\pi)} \leq \varepsilon,$$

which implies that $d_{G,f}(x,1)\leq\varepsilon$ for all $x\in 1_\eta$; and the claim follows. □

Proof of Theorem 2.3. Suppose G is locally random; that means G has a metric such that for all $x\in G$ and $\pi\in\widehat{G}$ we have

$$\|\pi(x)-I\|_{\mathrm{op}} \leq C_0(\dim\pi)^L d(x,1).$$

Let $f:\mathbb{Z}^+\to\mathbb{R}^+, f(n):=C_0n^L$; then f is strictly increasing and $\lim_{x\to1} d_{G,f}(x,1)=0$. Hence, for every $n\geq 1$, it follows from Theorem 4.3, that there are only finitely many elements of $\widehat{G}$ of dimension at most n.

Conversely, suppose that for all integers $n\geq 1$, there are only finitely many elements of $\widehat{G}$ of dimension at most n. Set $f:\mathbb{Z}^+\to\mathbb{R}^+, f(n):=n$. By Theorem 4.3, $d_{G,f}$ induces the original topology of G, and with respect to this metric for all $x,y\in G$ and $\pi\in\widehat{G}$ we have

$$\|\pi(x)-\pi(y)\|_{\mathrm{op}} \leq (\dim\pi)d_{G,f}(x,y);$$

therefore, G is locally random. □

5. Local randomness, dimension condition, and important examples.

As we pointed out earlier, local randomness is particularly powerful when in addition the chosen metric has a *dimension condition*, (DC). Furthermore, several important examples, e.g., analytic compact groups, come equipped with a natural metric and we would like to know whether G is locally random with respect to this natural metric.

In this section we address this question. In particular, we show that compact simple Lie groups (with respect to their natural metric) are locally random; we also provide a connection between quasi-randomness and local randomness for profinite groups.

We begin with investigating local randomness of quotients and products. Indeed, Theorem 2.3 implies that

(1) if G is locally random and N is a closed normal subgroup, then G/N is locally random;
(2) if G_1 and G_2 are locally random, then $G_1 \times G_2$ is locally random.

These statements, however, do not provide information regarding the metrics (or the involved parameters) with respect to which these groups are locally random. The following two lemmas prove the above statements with some control on the involved metrics.

Lemma 5.1. *Suppose G is L-locally random with coefficient C_0, and let N be a closed normal subgroup of G. Then G/N equipped with the natural quotient metric is L-locally random with coefficient C_0.*

Proof. Let us recall that given a bi-invariant metric d on G, the natural quotient metric on G/N is $d(xN, yN) := \inf_{h,h' \in N} d(xh, yh')$.

For $\overline{\pi} \in \widehat{G/N}$, let $\pi(x) := \overline{\pi}(xN)$; then $\pi \in \widehat{G}$. For $x, y \in G$, and every $\varepsilon > 0$, there exist $h, h' \in N$ such that

$$d(xh, xh') < d(xN, yN) + \varepsilon.$$

From this we conclude

$$\begin{aligned} \|\overline{\pi}(xN) - \overline{\pi}(yN)\|_{\mathrm{op}} = \|\pi(xh) - \pi(yh')\|_{\mathrm{op}} &\leq C_0(\dim \pi)^L d(xh, yh') \\ &\leq C_0(\dim \overline{\pi})^L (d(xN, yN) + \varepsilon). \end{aligned}$$

The claim follows as ε is an arbitrary positive number. □

Lemma 5.2. *Suppose G_i is an L_i-locally random group with coefficient C_i for $i = 1, 2$. Then $G_1 \times G_2$ is an $\max\{L_1, L_2\}$-locally random group with coefficient $C_1 + C_2$ with respect to the maximum metric.*

Proof. We know that any $\pi \in \widehat{G_1 \times G_2}$ is of the form $\pi_1 \otimes \pi_2$ for some $\pi_i \in \widehat{G}_i$. It is also well-known that for any two matrices a and b we have $\|a \otimes b\|_{\mathrm{op}} = \|a\|_{\mathrm{op}} \|b\|_{\mathrm{op}}$. Let $L := \max\{L_1, L_2\}$ and $C_0 := C_1 + C_2$. Then for any $(g_1, g_2) \in G_1 \times G_2$ we have

$$\begin{aligned} \|\pi(g_1, g_2) - I\|_{\mathrm{op}} =& \|\pi_1(g_1) \otimes \pi_2(g_2) - I \otimes I\|_{\mathrm{op}} \\ \leq& \|\pi_1(g_1) \otimes \pi_2(g_2) - I \otimes \pi_2(g_2)\|_{\mathrm{op}} + \|I \otimes \pi_2(g_2) - I \otimes I\|_{\mathrm{op}} \\ =& \|(\pi_1(g_1) - I) \otimes \pi_2(g_2)\|_{\mathrm{op}} + \|I \otimes (\pi_2(g_2) - I)\|_{\mathrm{op}} \\ =& \|\pi_1(g_1) - I\|_{\mathrm{op}} + \|\pi_2(g_2) - I\|_{\mathrm{op}} \\ \leq& C_1(\dim \pi_1)^L d_1(g_1, 1) + C_2(\dim \pi_2)^L d_2(g_2, 1) \\ \leq& C_0(\dim \pi)^L d((g_1, g_2), (1, 1)), \end{aligned}$$

from which the claim follows. □

The following is essentially proved for the standard metric d_0 in [12, Lemme 3.1, 3.2].

Proposition 5.3. *Suppose G is a compact semisimple Lie group with a compatible bi-invariant metric d. Then (G, d) is 1-locally random with coefficient $C_0 := C_0(G, d)$.*

We start with the following lemma.

Lemma 5.4. *Let d be a bi-invariant metric on a connected compact semisimple Lie group with the standard metric d_0. Then for all $g \in G$ we have*

$$d(g, 1) \gg_{d_0,d} d_0(g, 1)$$

Proof. Let us recall the construction of the standard metric d_0. It is well known that the Killing from is a negative definite bilinear form on the Lie algebra $\mathfrak{g}$, therefore,

$$\langle X, Y\rangle := -\operatorname{Tr}(\operatorname{ad}(X)\operatorname{ad}(Y))$$

defines a bi-invariant inner product on $\mathfrak{g}$ and hence a bi-invariant metric on G, which induces d_0. Fix a maximal torus T of G with the Lie algebra $\mathfrak{t}$. There exists $\eta_0' := \eta_0'(G)$ such that for every $X \in \mathfrak{t}$ with $\|X\| \le \eta_0'$ we have

$$\|X\| \ll d_0(\exp_T(X), 1) \ll \|X\|.$$

Since every element of G is in a conjugate of T, the map from $G \times T$ to G sending (g, t) to $g^{-1}tg$ is open. This implies that there exists $\eta_0 = \eta_0(G, d)$ such that $\bigcup_{g\in G} g^{-1}\exp_T\big(\{X \in \mathfrak{t} : \|X\| \le \eta_0'\}\big)g$ contains a ball of radius η_0 with respect to the metric d. Let

$$K_0 = \min\{d(\exp_T(X), 1) : X \in \mathfrak{t}, \|X\| \in [\eta_0'/2, \eta_0']\}.$$

For $X \in \mathfrak{t}$ with $\|X\| \le \eta_0'/4$, let n be the least positive integer such that $\|nX\| \ge \eta_0'/2$. By the triangle inequality we have

$$d(\exp_T(X), 1) \ge \frac{1}{n} d(\exp_T(nX), 1) \ge \frac{K_0}{n} \ge \frac{K_0\eta_0'\|X\|}{4}.$$

For every $g \in G$ with $d(g, 1) < \eta_0$, find $X \in \mathfrak{t}$ of norm less than η_0' such that g is conjugate to $\exp_T(X)$. It follows that

$$d(g, 1) = d(\exp_T(X), 1) \ge \frac{K_0\eta_0'\|X\|}{4} \gg_{d_0,d} d_0(\exp_T(X), 1) = d_0(g, 1).$$

This establishes the inequality for all g in a neighborhood of 1. On the complement of this set, $d(g, 1)$ is bounded from below, from which the claim follows. □

Proof of Proposition 5.3. Without loss of generality, we can assume that G is connected. In view of Lemma 5.4, it suffices to prove the claim for the natural metric d_0. Let Φ be the set of roots with respect to T and let Φ^+ be a set of positive roots. Let π be an irreducible unitary representation of G. Let $W_\pi := \{\lambda_1, \dots, \lambda_n\}$ be the set of weights of π, and let λ denote the highest weight of π with respect to Φ^+. We have

$$\mathcal{H}_\pi = \bigoplus_{j=1}^{n} \ker(\pi(\exp_T(X)) - e^{i\lambda_j(X)} I).$$

where $\exp_T$ denotes the restriction of the exponential map $\exp_G$ to $\mathfrak{t}$.

As before, let $\eta_0' := \eta_0'(G)$ be such that for any $X \in \mathfrak{t}$ with $\|X\| \le \eta_0'$ we have

$$\|X\| \ll d_0(\exp_T(X), 1) \ll \|X\|,$$

and choose $\eta_0 = \eta_0(G)$ such that $1_{\eta_0} \subset \bigcup_{g\in G} g^{-1}\exp_T\big(\{X \in \mathfrak{t} : \|X\| \le \eta_0'\}\big)g$.

Let $g \in 1_{\eta_0}$. Then,

$$\begin{aligned}\|\pi_\lambda(g) - I\|_{\mathrm{op}} &= \|\pi_\lambda(\exp_T(X)) - I\|_{\mathrm{op}} = \max_{\lambda_j \in W_{\pi_\lambda}} |e^{i\lambda_j(X)} - 1| \\ &\leq \max_{\lambda_j \in W_{\pi_\lambda}} |\lambda_j(X)| \ll \|\lambda\|\|X\| \ll \|\lambda\| d_0(\exp_T(X), 1) = \|\lambda\| d_0(g, 1).\end{aligned} \tag{5.1}$$

On the other hand, by Weyl's formula

$$\dim \pi = \prod_{\alpha \in \Phi^+} \frac{\langle \lambda + \rho, \alpha \rangle}{\langle \rho, \alpha \rangle}$$

where ρ is the half of the sum of the positive roots. For every $\alpha \in \Phi^+$, we have $\frac{\langle \lambda+\rho,\alpha\rangle}{\langle \rho,\alpha\rangle} \geq 1$. Moreover, since elements of Φ^+ contains a basis for the dual space of $\mathfrak{t}$, it follows that there exists $\alpha \in \Phi^+$ for which $\langle \lambda, \alpha \rangle \gg_G \|\lambda\|$. This implies that

$$\dim \pi \gg_G \|\lambda\|. \tag{5.2}$$

By (5.1) and (5.2) we get

$$\|\pi_\lambda(g) - I\|_{\mathrm{op}} \leq C_0'(G)(\dim \pi) d_0(g, 1),$$

for some $C_0'(G)$ and any $g \in 1_{\eta_0}$. Therefore, G is 1-locally random with coefficient $C_0 := \frac{2C_0'(G)}{\eta_0}$. □

Lemma 5.5. *Let G be a compact metrizable group such that for every compatible metric d, the pair (G, d) is L-locally random with coefficient C for some $C, L > 0$. Then G is a Lie group.*

Proof. The claim is clear when G is finite. Henceforth, we will assume that G is an infinite compact metrizable group. Let $(\pi_i)_{i\geq 1}$ be an enumeration of all elements of $\widehat{G}$ ordered such that the sequence $(\dim \pi_i)_{i\geq 1}$ is non-decreasing. For $m \geq 1$, let

$$\rho_m : G \to \prod_{1 \leq n \leq m} \pi_n(G)$$

denote the direct sum $\pi_1 \oplus \cdots \oplus \pi_m$. Equip G with the bi-invariant metric defined by

$$d(g, 1) = \sum_{n=1}^{\infty} e^{-nD_n} \|\pi_n(g) - I\|_{\mathrm{op}},$$

where $D_n = \deg \pi_n$. It is not hard to see that $d(g, 1) \to 0$ if $g \to 1$ in G. By virtue of Lemma 4.2, this metric is compatible with the topology of G. If ρ_m is injective for some $m \geq 1$, then it follows that G is homeomorphic to a closed subgroup of the Lie group $\rho_m(G)$, and hence is itself a compact Lie group. Suppose this fails for all $m \geq 1$ and pick a sequence $g_m \in \ker \rho_m \setminus \{1\}$. For each $g \in G \setminus \{1\}$ write $j(g)$ for the least index j such that $g \notin \ker \rho_j$. It follows from the choice of g_m that

$$d(g_m, 1) \leq 2 \sum_{i=m+1}^{\infty} e^{-iD_i} \leq 4e^{-j_m D_{j_m}}$$

where $j_m = j(g_m) \geq m + 1$. Let $\lambda \neq 1$ be an eigenvalue of $\rho_{j_m}(g_m)$. There exists an integer k such that $|\lambda^k - 1| \geq \sqrt{3}$. This implies that after replacing g_m by an appropriate power (if necessary) we can assume that $\|\rho_{j_m}(g_m) - I\|_{\mathrm{op}} \geq \sqrt{3}$. Hence for every choice of $L, C > 0$, all sufficiently large $m \geq 1$ we have

$$C(\dim \rho_{j_m})^L d(g_m, 1) \leq 4CD_{j_m}^L e^{-j_m D_{j_m}} < \sqrt{3} \leq \|\rho_{j_m}(g_m) - I\|_{\mathrm{op}}.$$

This shows that (G, d) is not L-locally random with coefficient C. □

Corollary 5.6. *Let G be a metrizable compact group. Then G is a (possibly disconnected) semisimple Lie group if and only if for every compatible metric d, we have (G,d) is $L(d)$-locally random with coefficient $C(d)$.*

Proof. Assuming that G is a semisimple Lie group, the claim follows from Proposition 5.3. Conversely, by Lemma 5.5, G is a Lie group. Now, it follows from Lemma 4.5 that the connected component of the identity in G has a finite abelianization, implying that G is semisimple. □

We now turn to the case of profinite groups. Following Varjú [24], a profinite group G will be called (c,α)-*quasi-random* if for all $\pi \in \widehat{G}$ we have

$$\dim \pi \geq c\,(\#\pi(G))^{\alpha}.$$

This is a natural extension of Gowers's notion of quasi-randomness to profinite setting.

Our next objective in this section is to relate this notion, which does not depend on the metric structure of G, to local randomness. Indeed, if G is a finitely generated (c,α)-quasi-random group, then it has only finitely many irreducible unitary representations of a given dimension. Therefore, by Theorem 2.3, we deduce that such a group is locally random. We will investigate this relationship in more details.

The following discussion is inspired by the p-adic setting. Suppose G is equipped with a bi-invariant metric, and define the *level* of $\pi \in \widehat{G}$ as:

$$\ell(\pi) := \inf\{\eta^{-1} |\ 1_\eta \not\subseteq \ker \pi\},$$

so for all $\varepsilon > 0$ we have $1_{(\ell(\pi)+\varepsilon)^{-1}} \subseteq \ker \pi$. If 1_η is a normal subgroup for every $\eta > 0$, then $\pi(G)$ is a factor of $G/1_{(\ell(\pi)+\varepsilon)^{-1}}$. Hence

$$\#\pi(G) \leq |1_{(\ell(\pi)+\varepsilon)^{-1}}|^{-1}. \tag{5.3}$$

If, in addition, G satisfies (DC), then we conclude from (5.3) that $\#\pi(G) \leq C_1(\ell(\pi)+\varepsilon)^{d_0}$ for all $\varepsilon > 0$. Therefore,

$$\#\pi(G) \leq C_1 \ell(\pi)^{d_0}. \tag{5.4}$$

In view of the above inequality, we define a *metric quasi-randomness* for profinite groups.

Definition 5.7. A compact group G with a given bi-invariant metric is said to be (C,A)-*metric quasi-random* if the following two conditions are satisfied:

(1) For all $\eta > 0$, 1_η is a subgroup of G.
(2) For all $\pi \in \widehat{G}$, we have $\ell(\pi) \leq C(\dim \pi)^A$.

Hence by (5.4) we get the following:

Lemma 5.8. *Suppose G is an (C,A)-metric quasi-random group and $|1_\eta| \leq C_1\eta^{d_0}$ for all $\eta > 0$ where C_1 and d_0 are positive constants. Then G is $((C_1C^{d_0})^{-1}, 1/(Ad_0))$-quasi-random.*

Next we prove that L-local randomness (with some parameters) and metric quasi-randomness are equivalent when balls centered at 1 are subgroups.

Proposition 5.9. *Suppose G is a compact group with a bi-invariant metric. Suppose $G = 1_1$, and 1_η is a subgroup of G for all $\eta \in (0,1]$. Then G is L-locally random with coefficient C_0 if and only if G is (C,L)-metric quasi-random, where $C = C_0$ in one direction, and $C_0 = 2C$ in the other direction.*

Proof. Suppose G is locally random, and let $\pi \in \widehat{G}$ be non-trivial. For $x \in 1_\eta$ we have

$$\|\pi(x) - I\|_{\mathrm{op}} \leq C_0(\dim \pi)^L \eta.$$

In particular, if $\eta < C_0^{-1}(\dim \pi)^{-L}$ and $x \in 1_\eta$, then for any $n \in \mathbb{Z}$, $\|\pi(x)^n - I\|_{\mathrm{op}} < 1$. This implies $\log(\pi(x)^n)$ is well-defined for all integer n —recall that $\pi(x) \in U_{\dim \pi}(\mathbb{C})$. Furthermore, $\log(\pi(x)^n) = n \log(\pi(x))$. Since G is profinite, $\pi(G)$ is a finite group, and hence, $\pi(x)$ is torsion for any $x \in G$. Therefore, for some positive integer n we have $0 = \log(\pi(x)^n) = n \log(\pi(x))$, which implies that $\pi(x) = I$. That is:

$$1_\eta \subseteq \ker \pi \quad \text{if } \eta < C_0^{-1}(\dim \pi)^{-L}. \tag{5.5}$$

By (5.5) we have

$$\ell(\pi) \leq C_0(\dim \pi)^L,$$

which implies that G is (C_0, L)-metric quasi-random.

To see the other implication, note that for all $\pi \in \widehat{G}$ and any $x \in G$, $\pi(x) \neq I$ implies that $d(x, 1) \geq 1/\ell(\pi)$. Therefore,

$$\|\pi(x) - I\|_{\mathrm{op}} \leq 2 \leq 2\ell(\pi)d(x, 1) \leq 2C(\dim \pi)^L d(x, 1),$$

which implies that G is L-locally random with coefficient $2C$. □

In [19, Lemma 20] using Howe's Kirillov theory, it is proved that an open compact subgroup G of a p-adic analytic group with a perfect Lie algebra is (C, A)-metric quasi-random for some positive numbers C and A depending on G. Thus, by Proposition 5.9 we obtain an important family of locally random groups.

Proposition 5.10. *Suppose G is a compact open subgroup of a p-adic analytic group with a perfect Lie algebra. Then, for some positive number L and C_0, G is L-locally random with coefficient C_0.*

6. Mixing inequality for locally random groups

In this section, we will prove Theorem 2.6 and derive a number of its corollaries.

6.1. **High and low frequencies and the proof of Theorem 2.6.** The proof of Theorem 2.6 involves splitting the terms in Parseval's theorem for $\|f\|^2$ into the sum of contributions from *low frequency* and *high frequency* terms. By low (resp. high) frequency terms, we mean terms coming from irreducible representations of small (resp. large) degree. The low frequency terms can be bounded by the local randomness assumption whereas high frequency terms are dealt with using a trivial bound. For $f \in L^2(G)$ and a threshold parameter D, write

$$L(f; D) := \sum_{\pi \in \widehat{G}, \dim \pi \leq D} \dim \pi \, \|\widehat{f}(\pi)\|_{\mathrm{HS}}^2 \tag{6.1}$$

for the low-frequency terms and

$$H(f; D) := \sum_{\pi \in \widehat{G}, \dim \pi > D} \dim \pi \, \|\widehat{f}(\pi)\|_{\mathrm{HS}}^2 \tag{6.2}$$

for the high frequency terms. By Parseval's theorem, $\|f\|_2^2 = L(f; D) + H(f; D)$ holds.

Lemma 6.1. *In the above setting, we have*

$$L(f * g; D) \leq L(f; D)L(g; D) \quad \text{and} \quad H(f * g; D) < \frac{1}{D} H(f; D)H(g; D).$$

Proof. We have $\|\widehat{f*g}(\pi)\|_{\mathrm{HS}} = \|\widehat{g}(\pi)\widehat{f}(\pi)\|_{\mathrm{HS}} \leq \|\widehat{g}(\pi)\|_{\mathrm{HS}}\|\widehat{f}(\pi)\|_{\mathrm{HS}}$, and

$$\begin{aligned} L(f*g;D) =& \sum_{\pi\in\widehat{G},\dim\pi\leq D} \dim\pi\|\widehat{f*g}(\pi)\|_{\mathrm{HS}}^2 \\ \leq& \left(\sum_{\pi\in\widehat{G},\dim\pi\leq D} \dim\pi\|\widehat{f}(\pi)\|_{\mathrm{HS}}^2\right)\left(\sum_{\pi\in\widehat{G},\dim\pi\leq D} \dim\pi\|\widehat{g}(\pi)\|_{\mathrm{HS}}^2\right) \\ \leq& L(f;D)L(g;D). \end{aligned}$$

Similarly, we have the inequality

$$\begin{aligned} H(f*g;D) =& \sum_{\pi\in\widehat{G},\dim\pi> D} \dim\pi\|\widehat{f*g}(\pi)\|_{\mathrm{HS}}^2 \\ <& \frac{1}{D}\left(\sum_{\pi\in\widehat{G},\dim\pi> D} \dim\pi\|\widehat{f}(\pi)\|_{\mathrm{HS}}^2\right)\left(\sum_{\pi\in\widehat{G},\dim\pi> D} \dim\pi\|\widehat{g}(\pi)\|_{\mathrm{HS}}^2\right) \\ \leq& \frac{1}{D}H(f;D)H(g;D), \end{aligned}$$

as we claimed. □

Lemma 6.2 (Fourier terms in low frequencies)**.** *Suppose G is an L-locally random with coefficient C_0. Then for all $\eta > 0$ and $\pi \in \widehat{G}$ we have*

$$\|\widehat{P}_\eta(\pi) - I\|_{\mathrm{op}} \leq C_0(\dim\pi)^L\,\eta.$$

Proof. For all $x \in 1_\eta$, we have $\|\pi(x) - I\|_{\mathrm{op}} \leq C_0(\dim\pi)^L d(1,x) \leq C_0(\dim\pi)^L\,\eta$. Therefore,

$$\|\widehat{P}_\eta(\pi) - I\|_{\mathrm{op}} = \left\|\int P_\eta(x)(\pi(x)^* - I)\,\mathrm{d}x\right\|_{\mathrm{op}} \leq C_0(\dim\pi)^L\,\eta.$$

□

Lemma 6.3. *Suppose G is an L-locally random group with coefficient C_0. Let $\eta > 0$ and $D \geq 1$ be two parameters satisfying $C_0D^L\eta < 1$. Then*

$$L(f;D) \leq (1 - C_0D^L\eta)^{-2}L(f_\eta;D) \leq (1 - C_0D^L\eta)^{-2}\|f_\eta\|_2^2$$

where $f_\eta = P_\eta * f$.

Proof. The second inequality is clear because of $L(f_\eta;D) \leq \|f_\eta\|_2^2$.

We now show the first inequality. Note that

$$L(f_\eta;D) = \sum_{\pi\in\widehat{G},\dim\pi\leq D} \dim\pi\,\|\widehat{P}_\eta(\pi)\widehat{f}(\pi)\|_{\mathrm{HS}}^2. \tag{6.3}$$

We have

$$\begin{aligned} \|\widehat{P}_\eta(\pi)\widehat{f}(\pi)\|_{\mathrm{HS}} =& \|\widehat{f}(\pi) - (I - \widehat{P}_\eta(\pi))\widehat{f}(\pi)\|_{\mathrm{HS}} \\ \geq& (1 - C_0(\dim\pi)^L\,\eta)\|\widehat{f}(\pi)\|_{\mathrm{HS}} && \text{(By Lemma 6.2)} \\ \geq& (1 - C_0D^L\,\eta)\|\widehat{f}(\pi)\|_{\mathrm{HS}} \end{aligned}$$

This estimate and (6.3) imply that

$$L(f_\eta;D) \geq (1 - C_0D^L\eta)^2L(f;D) \tag{6.4}$$

which finishes the proof. □

Proof of Theorem 2.6. Let η be as in the statement of the theorem, and let $D = (\sqrt{\eta})^{-1/L}$.

By Lemmas 6.1 and 6.3, we have

$$\begin{aligned}\|f*g\|_2^2 =& L(f*g;D) + H(f*g;D)\\ \le& L(f;D)L(g;D) + \frac{1}{D}H(f;D)H(g;D)\\ \le& (1-C_0D^L\eta)^{-4}\|f_\eta\|_2^2\|g_\eta\|_2^2 + \frac{1}{D}\|f\|_2^2\|g\|_2^2.\end{aligned}$$

Note that $(1-C_0D^L\eta)^{-4} = (1-C_0\sqrt{\eta})^{-4} \le 0.9^{-4} \le 2$. The claim follows from here. □

6.2. **An almost orthogonality and further mixing inequalities.** The inequality in Theorem 2.6 is non-trivial only when $\|f_\eta\|_2$ and $\|g_\eta\|_2$ are small. In this section, we show that $(f-f_\eta)_{\eta'}$ is small when η is polynomially smaller than η'. Thus applying the mixing inequality of Theorem 2.6 to $(f-f_\eta)_{\eta'}$ and g, we get a meaningful mixing. We will then use this to prove a product theorem. To get a better understanding of the discussion, consider the case when 1_η is a subgroup of G. Then $f \mapsto f_\eta$ is the orthogonal projection onto the space of 1_η-invariant functions in $L^2(G)$ and $(f-f_\eta)_\eta = 0$; hence, one may let $\eta' = \eta$.

Results in this section require only a dimension condition at a given scale. This is implied by (DC), but is more general.

Let us recall that any class function in $L^1(G)$ is in the center of the Banach algebra $(L^1(G), +, *)$; therefore, P_η is in the center of $L^1(G)$ for any η.

Lemma 6.4. *Suppose G is an L-locally random group with coefficient C_0. For every $C_1 > 0$ and every $\eta \ll_{C_0,C_1,L} 1$ we have the following. Suppose $\eta' \ge \eta^{1/(4Ld_0)}$ satisfies $|1_{\eta'}| \ge \frac{1}{C_1}\eta'^{d_0}$. Then for every $f \in L^2(G)$ we have*

$$\|(f-f_\eta)_{\eta'}\|_2 \le \eta^{1/(8L)}\|f\|_2.$$

Proof. Let D be a threshold parameter which will be set later. Then

$$\begin{aligned}L((f-f_\eta)_{\eta'};D) =& \sum_{\pi\in\widehat{G},\dim\pi\le D} \dim\pi\|\widehat{f}(\pi)\widehat{P}_{\eta'}(\pi)(I-\widehat{P}_\eta(\pi))\|_{\mathrm{HS}}^2\\ \le& \sum_{\pi\in\widehat{G},\dim\pi\le D} \dim\pi\|I-\widehat{P}_\eta(\pi)\|_{\mathrm{op}}^2\|\widehat{P}_{\eta'}(\pi)\|_{\mathrm{op}}^2\|\widehat{f}(\pi)\|_{\mathrm{HS}}^2\\ \le& (C_0D^L\eta)^2 L(f;D) \qquad \text{(By Lemma 6.2).}\end{aligned}$$

We used $\|AB\|_{\mathrm{HS}} \le \|A\|_{\mathrm{op}}\|B\|_{\mathrm{HS}}$ for matrices A and B for the first inequality, and $\|\widehat{P}_{\eta'}(\pi)\|_{\mathrm{op}} \le 1$ in the final inequality. For the high frequencies we have

$$\begin{aligned}H((f-f_\eta)_{\eta'};D) =& \sum_{\pi\in\widehat{G},\dim\pi> D} \dim\pi\|\widehat{f}(\pi)\widehat{P}_{\eta'}(\pi)(I-\widehat{P}_\eta(\pi))\|_{\mathrm{HS}}^2\\ \le& \sum_{\pi\in\widehat{G},\dim\pi> D} \dim\pi\|I-\widehat{P}_\eta(\pi)\|_{\mathrm{op}}^2\|\widehat{P}_{\eta'}(\pi)\|_{\mathrm{op}}^2\|\widehat{f}(\pi)\|_{\mathrm{HS}}^2\\ \le& \frac{4}{D}H(P_{\eta'};D)H(f;D) \le \frac{4}{D}\|P_{\eta'}\|_2^2H(f;D) = \frac{4}{D\,|1_{\eta'}|}H(f;D),\end{aligned}$$

where we used the trivial bound $\|I-\widehat{P}_\eta(\pi)\|_{\mathrm{op}} \le 2$. Combining these two estimates, we conclude

$$\|(f-f_\eta)_{\eta'}\|_2^2 \le \left((C_0D^L\eta)^2 + \frac{4}{D\,|1_{\eta'}|}\right)\|f\|_2^2.$$

Setting $D = \eta^{-1/(2L)}$ we get the desired inequality. □

In the rest of this section we will prove a number of mixing inequalities.

Lemma 6.5. *Suppose G is an L-locally random group with coefficient C_0. For every integer $m \geq 2$, $C_1 > 0$, and $\eta \ll_{C_0,C_1,L} 1$ we have the following. Suppose $\eta' \geq \eta^{1/(4Ld_0)}$ satisfies $C_0\sqrt{\eta'} < 0.1$ and $|1_{\eta'}| \geq \frac{1}{C_1}\eta'^{d_0}$. Then for all $f_1, \dots, f_m \in L^2(G)$ we have*

$$\|(f_1 - (f_1)_\eta) * f_2 * \cdots * f_m\|_2 \leq \sqrt{3}^m \eta'^{1/(4L)} \prod_{i=1}^m \|f_i\|_2.$$

Proof. We proceed by induction on m. Let us start with the base case $m = 2$. By Theorem 2.6, we have that $\|(f_1 - (f_1)_\eta) * f_2\|_2^2$ is bounded from above by

$$2\|(f_1 - (f_1)_\eta)_{\eta'}\|_2^2 \|(f_2)_{\eta'}\|_2^2 + \eta'^{1/(2L)} \|f_1 - (f_1)_\eta\|_2^2 \|f_2\|_2^2. \tag{6.5}$$

By Lemma 6.4 we have

$$\|(f_1 - (f_1)_\eta)_{\eta'}\|_2 \leq \eta^{1/(8L)} \|f_1\|_2. \tag{6.6}$$

Since $\|f * g\|_2 \leq \|f\|_1 \|g\|_2$, we have

$$\|f_1 - (f_1)_\eta\|_2 \leq \|1 - P_\eta\|_1 \|f_1\|_2 \leq 2\|f_1\|_2 \quad \text{and} \quad \|(f_2)_{\eta'}\|_2 \leq \|f_2\|_2. \tag{6.7}$$

By (6.5), (6.6), and (6.7), we get that $\|(f_1 - (f_1)_\eta) * f_2\|_2^2$ is bounded from above by

$$2(\eta^{1/(8L)} \|f_1\|_2)^2 \|f_2\|_2^2 + 4\eta'^{1/(2L)} \|f_1\|_2^2 \|f_2\|_2^2.$$

Therefore

$$\|(f_1 - (f_1)_\eta) * f_2\|_2 \leq \sqrt{6}\eta'^{1/(4L)} \|f_1\|_2 \|f_2\|_2.$$

This concludes the proof for $m = 2$. Now, suppose that the inequality holds for some value of m, and set

$$F_m := (f_1 - (f_1)_\eta) * f_2 * \cdots * f_m.$$

By Theorem 2.6, we have that $\|F_m * f_{m+1}\|_2^2$ is at most

$$2\|(F_m)_{\eta'}\|_2^2 \|(f_{m+1})_{\eta'}\|_2^2 + \eta'^{1/(2L)} \|F_m\|_2^2 \|f_{m+1}\|_2^2.$$

Since $\|(F_m)_{\eta'}\|_2 \leq \|F_m\|_2$, by the induction hypothesis we have

$$\|(F_m)_{\eta'}\|_2^2 \leq \|F_m\|_2^2 \leq 3^m \eta'^{1/(2L)} \prod_{i=1}^m \|f_i\|_2^2.$$

Hence by the above inequalities we get $\|F_m * f_{m+1}\|_2^2$ is at most

$$3^m(2 + \eta'^{1/(2L)})\eta'^{1/(2L)} \prod_{i=1}^{m+1} \|f_i\|_2^2 \leq 3^{m+1} \eta'^{1/(2L)} \prod_{i=1}^{m+1} \|f_i\|_2^2;$$

and the claim follows. □

Proposition 6.6. *Suppose G is an L-locally random group with coefficient C_0. For every integer $m \geq 2$, $C_1 > 0$, and $\eta \ll_{C_0,C_1,L} 1$ we have the following. Suppose $\eta' \geq \eta^{1/(4Ld_0)}$ satisfies $C_0\sqrt{\eta'} < 0.1$ and $|1_{\eta'}| \geq \frac{1}{C_1}\eta'^{d_0}$. Suppose $f_1, \dots, f_m, f_{m+1} \in L^2(G)$. Then*

$$\|f_1 * \cdots * f_{m+1} - f_1 * \cdots * f_{m+1} * P_\eta\|_\infty \leq \sqrt{3}^m \eta'^{1/(4L)} \prod_{i=1}^{m+1} \|f_i\|_2.$$

Proof. Recall that $\|f * g\|_\infty \leq \|f\|_2 \|g\|_2$, see (3.3). Therefore, from the fact that P_η is in the center of $(L^1(G), +, *)$ we obtain

$$\|f_1 * \cdots * f_{m+1} - f_1 * \cdots * f_{m+1} * P_\eta\|_\infty \leq \|(f_1 - (f_1)_\eta) * \cdots * f_m\|_2 \|f_{m+1}\|_2.$$

The claim thus follows from Lemma 6.5. □

Corollary 6.7. *Suppose G is an L-locally random group with coefficient C_0. For every integer $m \geq 2$, $C_1 > 0$, and $\eta \ll_{C_0,C_1,L} 1$ we have the following. Suppose $\eta' \geq \eta^{1/(4Ld_0)}$ satisfies $C_0\sqrt{\eta'} < 0.1$ and $|1_{\eta'}| \geq \frac{1}{C_1}\eta'^{d_0}$. Suppose $f_1, \dots, f_m, f_{m+1} \in L^2(G)$. Then*

$$\|f_1 * \cdots * f_{m+1} - (f_1)_\eta * \cdots * (f_{m+1})_\eta\|_\infty \leq m\sqrt{3}^m \eta'^{1/(4L)} \prod_{i=1}^{m+1} \|f_i\|_2.$$

Proof. Let $F_1 := f_1 * \cdots * f_{m+1}$ and $F_{k+1} := (f_1)_\eta * \cdots * (f_k)_\eta * f_{k+1} * \cdots * f_{m+1}$ for any $1 \leq k \leq m$. By Proposition 6.6 and the fact that P_η is in the center of $(L^1(G), +, *)$ for any k we have

$$\begin{aligned}\|F_k - F_{k+1}\|_\infty \leq& \sqrt{3}^m \eta'^{1/(4L)} \prod_{i=1}^{k-1} \|(f_i)_\eta\|_2 \prod_{i=k}^{m+1} \|f_i\|_2 \\ \leq& \sqrt{3}^m \eta'^{1/(4L)} \prod_{i=1}^{m+1} \|f_i\|_2.\end{aligned}$$

Therefore, $\|F_1 - F_{m+1}\|_\infty \leq m\sqrt{3}^m \eta'^{1/(4L)} \prod_{i=1}^{m+1} \|f_i\|_2$, and the claim follows. □

7. A product result for large subsets.

The main goal of this section is to prove Theorem 2.8. We start by recalling a number of definitions and setting some notation. Suppose X is a metric space and A is a non-empty subset of X. Recall that for $\eta \in (0,1)$, x_η denotes the ball of radius η centered at x, and similarly A_η denotes the union of all x_η with $x \in A$. We write $\mathcal{N}_\eta(A)$ for the least number of open balls of radius η with centers in A that cover A. The metric entropy of A at scale η is defined by $h(A;\eta) := \log \mathcal{N}_\eta(A)$. A maximal η-separated subset $\mathcal{C}$ of A has the property that every distinct $x, x' \in \mathcal{C}$ are at least η apart and its η-neighborhood covers A.

The metric space we will be working with is a metrizable compact group G equipped with bi-invariant metric denoted by $d(\cdot,\cdot)$. We will assume further that the pair (G,d) enjoys the dimension condition $\mathrm{DC}(C,d_0)$ defined in (DC).

Lemma 7.1 (Uniformly comparable quantities)**.** *Fix a subset $A \subseteq X$ and $\eta > 0$, and let $A^* \subseteq A$ be a maximal η-separated subset of A, and write $\overline{A} = (A^*)_\eta$. Then A^* is finite, $\overline{A}$ is open, and $A^* \subseteq A \subseteq \overline{A}$. Moreover, the ratio of any two quantities among*

$$|\overline{A}|/|1_\eta|, \quad |A_\eta|/|1_\eta|, \quad \mathcal{N}_\eta(A), \quad \#A^*$$

is bounded above by $\Omega = 2^{d_0}C^2$.

Proof. Write $N = \mathcal{N}_\eta(A)$ and denote by $\{(x_i)_\eta\}_{i=1}^N$ a minimal η-cover of A with centers in A. For each $x \in A_\eta$ there exists some $1 \leq i \leq N$ such that $x \in (x_i)_{2\eta}$, implying that $A_\eta \subseteq \bigcup_{i=1}^N (x_i)_{2\eta}$. Therefore

$$|A_\eta| \leq N|1_{2\eta}| \leq 2^{d_0}C^2 N|1_\eta|. \tag{7.1}$$

where the last inequality follows from an application of (DC).

Since A^* is a maximal η-separated subset of A, the open balls $\{x_\eta\}_{x\in A^*}$ form an η-cover of A with centers in A, and hence

$$\mathcal{N}_\eta(A) \leq \#A^*. \tag{7.2}$$

Finally, since A^* is η-separated, each two balls in the family $\{x_{\eta/2} : x \in A^*\}$ are pairwise disjoint, yielding

$$|A^*_{\eta/2}| = (\#A^*)\ |1_{\eta/2}|.$$

This implies that

$$\#A^* \leq \frac{|A^*_\eta|}{|1_{\eta/2}|} \leq 2^{d_0} C^2 \frac{|A_\eta|}{|1_\eta|}. \tag{7.3}$$

This completes the proof. □

Remark 7.2. From now on, whenever two positive quantities X and Y are within a multiplicative factor of the form $\Omega^{O(1)}$ of one another, we will write $X \approx Y$. Similarly, we write $X \preccurlyeq Y$ to state that X/Y is bounded from above by an expression of the form $\Omega^{O(1)}$, where the implied constants are not of importance. Using this notation we can now write

$$\mathcal{N}_\eta(A) \approx \#A^* \approx \frac{|A_\eta|}{|1_\eta|}.$$

Remark 7.3. The proof of Lemma 7.1 only uses the dimension condition for $\eta, 2\eta$ and $\eta/2$. We will use this fact later.

Corollary 7.4. *Suppose G is a compact group that satisfies (DC). Then for every fixed constant $c \geq 1$ and every non-empty subset A of G and every $0 < \eta < 1$ we have*

$$|A_{c\eta}| \approx |A_\eta|$$

Proof. Since $|A_{c\eta}| \geq |A_\eta|$, we will need to prove the reverse inequality. Denote by $A^*(\eta)$ and $A^*(c\eta)$, respectively, maximal η-separated and $c\eta$-separated subsets of A. By Lemma 7.1 we have that $\#A^*(c\eta) \approx \frac{|A_{c\eta}|}{|1_{c\eta}|}$. Clearly we have $\#A^*(c\eta) \leq \#A^*(\eta)$, implying

$$\frac{|A_{c\eta}|}{|1_{c\eta}|} \preccurlyeq \frac{|A_\eta|}{|1_\eta|}.$$

Hence

$$|A_{c\eta}| \preccurlyeq \frac{|1_{c\eta}|}{|1_\eta|}|A_\eta| \approx |A_\eta|;$$

and the claim follows. □

For a Borel measurable set $A \subseteq G$ with $|A| > 0$ and $\eta > 0$, define

$$\chi_{A,\eta} = \left(\frac{1}{|A|}\mathbb{1}_A\right) * 1_\eta.$$

Some basic properties of $\chi_{A,\eta}$ are summarized in the next lemma:

Lemma 7.5. *Let G be as above and $0 < \eta < 1$.*

(1) *For a measurable subset of positive measure $A \subseteq G$, we have*

$$\chi_{A,\eta}(x) = \frac{|A \cap x_\eta|}{|A||1_\eta|}.$$

(2) *$\chi_{A,\eta}$ is supported on η-neighborhood of A and has L^∞ norm at most $1/|A|$.*

(3) *For $A \subseteq B$ of positive measure*

$$\chi_{A,\eta}(x) \leq \frac{|B|}{|A|}\chi_{B,\eta}(x).$$

(4) *If $d(x,y) < \rho < 1$, then*

$$\chi_{A,\eta}(x) \preccurlyeq \left(\frac{\eta+\rho}{\eta}\right)^{d_0} \chi_{A,\eta+\rho}(y).$$

Proof. Since 1_η is a symmetric subset, we have

$$\chi_{A,\eta}(x) = \frac{1}{|A||1_\eta|}\int_G \mathbb{1}_{A\cap x_\eta}(y)\,\mathrm{d}y,$$

from which part (a) follows. Part (b) follows immediately from part (a). To show part (c), observe that $y_{\eta+\rho} \supseteq x_\eta$. It thus follows from the dimension condition that

$$\chi_{A,\eta}(x) \le \frac{|A\cap y_{\eta+\rho}|}{|A||1_\eta|} = \frac{|1_{\eta+\rho}|}{|1_\eta|}\chi_{A,\eta+\rho}(y) \preccurlyeq \left(\frac{\eta+\rho}{\eta}\right)^{d_0}\chi_{A,\eta+\rho}(y). \tag{7.4}$$

□

The next lemma, which is a version of Markov's inequality, establishes another quantity that is comparable to the ones in Lemma 7.1.

Lemma 7.6 (Density points)**.** *Let G be as above, $A \subseteq G$ and $0 < \eta < \rho < 1$. Fixing η, let A^* be a maximal η-separated subset of A, and $\overline{A} := A^*_\eta$. For a threshold parameter $0 < \tau < 1$, we let*

$$A_{\mathrm{high}} := \{x \in A^* :\ \chi_{\overline{A},3\rho}(x) > \tau\}.$$

Under the condition that $\tau \preccurlyeq 1$, we have

$$|\overline{A}| \preccurlyeq |(A_{\mathrm{high}})_\rho|.$$

Proof. Every point x in the support of $\chi_{\overline{A},\rho}$ lies at distance less than ρ from $\overline{A}$ and hence at distance less than $\eta+\rho < 2\rho$ from a point $\overline{x} \in A^*$:

$$\operatorname{supp}\chi_{\overline{A},\rho} \subseteq (A^*)_{2\rho}.$$

By part (4) of Lemma 7.5 we have

$$\chi_{\overline{A},\rho}(x) \preccurlyeq \chi_{\overline{A},3\rho}(\overline{x}). \tag{7.5}$$

Write $Z = (A_{\mathrm{high}})_{2\rho}$. If $x \in G\setminus Z$, the above $\overline{x}$ is in $A^*\setminus A_{\mathrm{high}}$, which means

$$\chi_{\overline{A},3\rho}(\overline{x}) \le \tau. \tag{7.6}$$

By (7.5) and (7.6) we deduce that for $x \in G\setminus Z$ and $\tau \preccurlyeq 1$

$$\chi_{\overline{A},\rho}(x) \le 1/2.$$

This means that the density function $\chi_{\overline{A},\rho}$ is concentrated on Z:

$$1/2 \le \int_Z \chi_{\overline{A},\rho}(x)\,\mathrm{d}x \le \frac{|(A_{\mathrm{high}})_{2\rho}|}{|\overline{A}|}; \tag{7.7}$$

where the last inequality follows from the fact that $\chi_{\overline{A},\rho}$ is bounded by $1/|\overline{A}|$. The claim now follows from Corollary 7.4. □

Proof of Theorem 2.8. As before we will choose maximal η-separated subsets $A^* \subseteq A$ and $B^* \subseteq B$, set $\overline{A} = (A^*)_\eta$ and $\overline{B} = (B^*)_\eta$. Also write $C = B^{-1}A^{-1}$, and $\overline{C} = \overline{B}^{-1}\overline{A}^{-1}$. Note that in this proof we are deviating from the notation we used earlier in that here $\overline{C}$ is *not* defined to be $(C^*)_\eta$.

By the mixing inequality given in Proposition 6.6 for $\rho := \eta^\varepsilon$ we have

$$\|\chi_{\overline{A}} * \chi_{\overline{B}} * \chi_{\overline{C}} - \chi_{\overline{A},5\rho} * \chi_{\overline{B},5\rho} * \chi_{\overline{C},5\rho}\|_\infty \le \rho^{O_{L,d_0}(1)}\|\chi_{\overline{A}}\|_2\|\chi_{\overline{B}}\|_2\|\chi_{\overline{C}}\|_2. \tag{7.8}$$

The main step of the proof is to show the following inequality:

$$\chi_{\overline{A},5\rho} * \chi_{\overline{B},5\rho} * \chi_{\overline{C},5\rho}(x) \succcurlyeq \frac{(|\overline{A}||\overline{B}|)^{3/2}}{|\overline{C}|}.$$

Let $\tau \preccurlyeq 1$ be as in Lemma 7.6. For any $y \in (A_{\text{high}})_\rho$ there is $y' \in A_{\text{high}}$ such that $d(y', y) < \rho$. By part (4) of Lemma 7.5, we have that

$$\chi_{\overline{A},5\rho}(y) \succcurlyeq \chi_{\overline{A},4\rho}(y) \succcurlyeq \chi_{\overline{A},3\rho}(y') \succcurlyeq 1. \tag{7.9}$$

Similarly for $z \in (B_{\text{high}})_\rho$ we have

$$\chi_{\overline{B},4\rho}(z) \succcurlyeq 1. \tag{7.10}$$

For $y \in (A_{\text{high}})_\rho$, $z \in (B_{\text{high}})_\rho$, and $x \in 1_\rho$, by part (4) of Lemma 7.5 we have

$$\chi_{\overline{C},4\rho}(z^{-1}y^{-1}x) \succcurlyeq \chi_{\overline{C},3\rho}(z^{-1}y^{-1}). \tag{7.11}$$

On the other hand, by part (1) of Lemma 7.5 we have

$$\chi_{\overline{C},3\rho}(z^{-1}y^{-1}) = \chi_{\overline{C}^{-1}z^{-1},3\rho}(y) = \chi_{y^{-1}\overline{C}^{-1},3\rho}(z). \tag{7.12}$$

Since $z \in (B_{\text{high}})_\rho$, there exists some $z' \in B_{\text{high}}$ so that $d(z, z') \le \rho$. Moreover, using the definition $C = B^{-1}A^{-1}$, we have that $\overline{A} \subseteq \overline{C}^{-1}z'^{-1}$. Similarly, from $d(y, y') \le \rho$, we see that $\overline{B} \subseteq y'\overline{C}^{-1}$. Hence by (7.11), and (7.12) and the estimate (7.9) we have

$$\begin{aligned}
\chi_{\overline{C},5\rho}(z^{-1}y^{-1}x) &\succcurlyeq \chi_{\overline{C},4\rho}(z'^{-1}y^{-1}x) && \text{(part (4) of Lemma 7.5)}\\
&\succcurlyeq \chi_{\overline{C},3\rho}(z'^{-1}y^{-1}) && \text{(by (7.11))}\\
&= \chi_{\overline{C}^{-1}z'^{-1},3\rho}(y) && \text{(by (7.12))}\\
&\succcurlyeq \frac{|\overline{A}|}{|\overline{C}|}\chi_{\overline{A},3\rho}(y) && \text{(part (3) of Lemma 7.5)}\\
&\succcurlyeq \frac{|\overline{A}|}{|\overline{C}|}. && \text{(by } y \in A_{\text{high}})
\end{aligned}$$

Similarly,

$$\chi_{\overline{C},5\rho}(z^{-1}y^{-1}x) \succcurlyeq \frac{|\overline{B}|}{|\overline{C}|}.$$

Combining these two inequalities gives

$$\chi_{\overline{C},5\rho}(z^{-1}y^{-1}x) \succcurlyeq \max\left\{\frac{|\overline{A}|}{|\overline{C}|}, \frac{|\overline{B}|}{|\overline{C}|}\right\} \ge \frac{|\overline{A}|^{1/2}|\overline{B}|^{1/2}}{|\overline{C}|}. \tag{7.13}$$

By (7.9), (7.10), and (7.13), Lemma 7.6, Corollary 7.4, for $x \in 1_\rho$, we get that

$$\chi_{\overline{A},5\rho} * \chi_{\overline{B},5\rho} * \chi_{\overline{C},5\rho}(x) \succcurlyeq |(A_{\text{high}})_\rho|\ |(B_{\text{high}})_\rho| \cdot \frac{|\overline{A}|^{1/2}|\overline{B}|^{1/2}}{|\overline{C}|} \succcurlyeq \frac{(|\overline{A}||\overline{B}|)^{3/2}}{|\overline{C}|}.$$

In order to show $x \in \overline{A} \cdot \overline{B} \cdot \overline{C}$, by (7.8), it suffices to prove that for δ small enough we have

$$\frac{(|\overline{A}||\overline{B}|)^{3/2}}{|\overline{C}|} > \alpha\rho^\beta(|\overline{A}||\overline{B}||\overline{C}|)^{-1/2}$$

where β is a fixed positive number that depends on L, d_0, and α is a fixed positive number that depends on L, d_0, C_0, C_1. This inequality holds if and only if $|\overline{A}||\overline{B}| > \sqrt{\alpha}\rho^{\beta/2}|\overline{C}|^{1/4}$, which, in view of $|\overline{C}| \le 1$, follows from

$$|\overline{A}||\overline{B}| > \sqrt{\alpha}\eta^{(\beta/2)\varepsilon}. \tag{7.14}$$

Now, recall the condition $\frac{h(A;\eta)+h(B;\eta)}{2} > (1-\delta)h(G;\eta)$. This implies

$$|A_\eta||B_\eta| \succcurlyeq |1_\eta|^{-2\delta} \succcurlyeq \eta^{2\delta d_0}. \tag{7.15}$$

Consequently, applying Lemma 7.1 we obtain $|\overline{A}||\overline{B}| \geq E^{-1}\eta^{2\delta d_0}$ where $E = \Omega^{O(1)}$. Finally note that if $\eta^{\varepsilon} \ll_{\alpha,\beta,d_0} 1$, then for $\delta \ll_{\beta,d_0} \varepsilon$ we have $E^{-1}\eta^{2\delta d_0} > \sqrt{\alpha}\eta^{(\beta/2)\varepsilon}$. This and (7.15) imply (7.14). The proof is complete. □

8. A Littlewood-Paley decomposition for locally random groups

In this section, we will give a decomposition of $L^2(G)$ into almost orthogonal subspaces of functions, each consisting of functions *living at a different scale*. This notion will be defined later (see Definition 8.7). We will first treat the case of profinite groups, which is somewhat simpler and sharper results can be obtained. Then, in the next subsection, we will deal with the general case of locally random groups.

8.1. The case of profinite groups.

Let G be a profinite group, equipped with a bi-invariant metric d such that balls centered at the identity element form a family of normal subgroups. Such a metric always exists. In fact, if G is presented as the inverse limit of finite groups $(G_i)_{i\geq 1}$, one can define the distance $d(g,h)$ to be 2^{-i} where i is the largest index with the property that $\pi_i(g) = \pi_i(h)$. Here $\pi_i : G \to G_i$ denotes the natural projection.

Lemma 8.1. *Suppose G is a compact group and N is a normal open subgroup of G. Let $f_N := \frac{\mathbb{1}_N}{|N|}$. Then $T_N : L^2(G) \to L^2(G), T_N(g) := f_N * g$ is the orthogonal projection onto the subspace $L^2(G)^N := \{f \in L^2(G) |\ f(gn) = f(g)$ for all $n \in N, g \in G\}$ of N-invariant functions. In addition*

$$q : L^2(G)^N \to L^2(G/N), \quad q(g)(xN) := g(x)$$

is a well-defined unitary G-module isomorphism.

Proof. The proof is a standard computation. □

Given a G-valued random variable X with distribution measure μ, let $X_\eta = XZ$, where Z is a random variable with distribution $P_\eta = \frac{\mathbb{1}_{1_\eta}}{|1_\eta|}$ independent from X.

Lemma 8.2. *Let μ_η denote the density function of X_η. Then $\mu_\eta(x) = \frac{\mu(x_\eta)}{|1_\eta|}$ for all $x \in G$.*

Proof. By definition, for all $f \in C(G)$ we have

$$\int_G f(x)\mu_\eta(x)\ \mathrm{d}x = \int_G \int_G f(xy)P_\eta(y)\ \mathrm{d}y\ \mathrm{d}\mu(x). \tag{8.1}$$

Notice that the right hand side of (8.1) is equal to

$$\int_G \int_G f(z)P_\eta(x^{-1}z)\ \mathrm{d}z\ \mathrm{d}\mu(x) = \int_G f(z) \int_G P_\eta(x^{-1}z)\ \mathrm{d}\mu(x)\ \mathrm{d}z = \int_G f(z)\frac{\mu(z_\eta)}{|1_\eta|}\ \mathrm{d}z.$$

□

Define the *Rényi entropy of X at scale η* by

$$H_2(X;\eta) := \log(1/|1_\eta|) - \log \|\mu_\eta\|_2^2 \tag{8.2}$$

where μ is the distribution of X. We also write $H_2(\mu;\eta)$ instead of $H_2(X;\eta)$. Let us observe that by Lemma 8.2 we have

$$\|\mu_\eta\|_\infty \leq 1/|1_\eta|; \text{ and so } \|\mu_\eta\|_2^2 \leq 1/|1_\eta|,$$

which implies that $H_2(X;\eta) \geq 0$.

Proposition 8.3. *Suppose G is a compact group with a given bi-invariant metric such that 1_η is a subgroup of G for all $\eta > 0$. Suppose G is an L-locally random group with coefficient C_0. Suppose G satisfies the dimension condition $DC(C_1, d_0)$. Let μ be a symmetric Borel probability measure on G whose support generates a dense subgroup of G. Fix a number $a > 1$ and $\eta_0 < 1$, and for all $i \geq 1$, let $\eta_i := \eta_0^{a^i}$ and $\mathcal{H}_i := L^2(G)^{1_{\eta_i}}$. Suppose that $C_2 > 0$ is such that for every $i \gg 1$, there exists an integer $l_i \leq C_2 h(G; \eta_i)$ such that*

$$\text{(Large entropy at scale } \eta) \qquad H_2(\mu^{(l_i)}; \eta_i) \geq \left(1 - \frac{1}{8Ld_0a}\right) h(G; \eta_i).$$

Then there exists $i_0 \geq 1$ such that

$$\mathcal{L}(\mu; L^2(G) \ominus \mathcal{H}_{i_0}) \geq \frac{1}{16C_2Ld_0a}.$$

In particular, $\mathcal{L}(\mu; G) > 0$.

Proof. For all i and $f \in \mathcal{H}_{i+1} \ominus \mathcal{H}_i$, we have $f_{\eta_i} = 0$ and $f_{\eta_{i+1}} = f$. Hence for $f \in \mathcal{H}_{i+1} \ominus \mathcal{H}_i$ and every symmetric Borel probability measure ν we have $\|\nu * f\|_2 = \|\nu_{\eta_{i+1}} * f_{\eta_{i+1}}\|_2$. Applying Theorem 2.6 for $i \gg 1$ we obtain

$$\begin{aligned}\|\nu * f\|_2^2 \leq& 2\|\nu_{\eta_i}\|_2^2 \|f_{\eta_i}\|_2^2 + \eta_i^{1/(2L)} \|\nu_{\eta_{i+1}}\|_2^2 \|f_{\eta_{i+1}}\|_2^2 \\ =& \eta_i^{1/(2L)} \|\nu_{\eta_{i+1}}\|_2^2 \|f\|_2^2 = \eta_{i+1}^{1/(2La)} \|\nu_{\eta_{i+1}}\|_2^2 \|f\|_2^2.\end{aligned}$$

This implies

$$2\mathcal{L}(\nu; \mathcal{H}_{i+1} \ominus \mathcal{H}_i) \geq H_2(\nu; \eta_{i+1}) - h(G; \eta_{i+1}) - \frac{1}{2La} \log \eta_{i+1}.$$

Since 1_η is a group, $h(G; \eta) = \log(1/|1_\eta|)$; and so by the dimension condition we have

$$|h(G; \eta) + d_0 \log \eta| \leq \log C_1.$$

Therefore by the previous inequality, for $\eta_{i+1} \ll_{C_1} 1$, we have

$$2\mathcal{L}(\nu; \mathcal{H}_{i+1} \ominus \mathcal{H}_i) \geq H_2(\nu; \eta_{i+1}) - \left(1 - \frac{1}{4Ld_0a}\right) h(G; \eta_{i+1}).$$

Applying the above inequality for $\nu := \mu^{(l_i)}$ coupled with $\mathcal{L}(\mu^{(l_i)}; \mathcal{H}_{i+1} \ominus \mathcal{H}_i) = l_i \mathcal{L}(\mu; \mathcal{H}_{i+1} \ominus \mathcal{H}_i)$ implies that for $i \gg 1$ we have

$$\mathcal{L}(\mu; \mathcal{H}_{i+1} \ominus \mathcal{H}_i) \geq \frac{1}{16C_2Ld_0a}.$$

As a result, $\mathcal{L}(\mu; L^2(G) \ominus \mathcal{H}_{i_0}) \geq \frac{1}{16C_2Ld_0a}$ for some i_0. Since $\mathcal{H}_{i_0} \ominus \mathbb{C}\mathbb{1}_G$ is a finite dimensional subspace of $L_0^2(G)$, and the support of μ generates a dense subgroup, we have $\mathcal{L}(\mu; G) > 0$. □

Now we interpret the spaces $\mathcal{H}_{i+1} \ominus \mathcal{H}_i$'s in terms of certain convolution operators. This point of view will be extended to an arbitrary locally random group.

Lemma 8.4. *Suppose G is a compact group, $G := N_1 \supseteq N_2 \supseteq \cdots$ is a sequence of normal open subgroups of G that form a basis for the neighborhoods of 1. For integers $i \geq 1$, let*

$$\Delta_i : L^2(G) \to L^2(G), \quad \Delta_i(g) := f_{N_{i+1}} * g - f_{N_i} * g,$$

and let $\Delta_0(g) := f_{N_1} * g$. Then the following statements hold.

1. *For all $g \in L^2(G)$ we have $g = \sum_{i=0}^{\infty} \Delta_i(g)$ in $L^2(G)$.*
2. *For all $i \neq j$ and $g \in L^2(G)$, we have $\Delta_i(g) \perp \Delta_j(g)$.*
3. *For all $g \in L^2(G)$ we have $\|g\|_2^2 = \sum_{i=0}^{\infty} \|\Delta_i(g)\|_2^2$*
4. *If μ is a Borel probability measure on G, then $\Delta_i(\mu * f) = \mu * \Delta_i(f)$ for all i.*

Proof. For every integer $k \geq 1$, define $\mathcal{H}_k := L^2(G)^{N_k}$. Since $G := N_1 \supseteq N_2 \supseteq \cdots$, we have $\mathbb{C}\mathbb{1}_G = \mathcal{H}_1 \subseteq \mathcal{H}_2 \subseteq \cdots$. By Lemma 8.1, we have that $f_{N_j} * g$ is the orthogonal projection of g onto $\mathcal{H}_j$ for any j. And so $\Delta_i(g) \in \mathcal{H}_{i+1} \ominus \mathcal{H}_i$ for positive integer i, and $\Delta_0(g) \in \mathcal{H}_1$. This implies (2).

Every element g of the matrix algebra $\mathcal{E}(G)$ generates a finite dimensional G-submodule M of $L^2(G)$, defining a unitary representation $\pi_M : G \to \mathcal{U}(M)$. Since G is profinite, we have that $\pi_M(G)$ is a finite group. Hence $\ker \pi_M$ is an open subgroup of G. Therefore, $N_k \subseteq \ker \pi_M$ for some k, implying $g \in \mathcal{H}_k$. It follows that $g = \sum_{i=0}^{j} \Delta_i(g)$ for all $j \geq k-1$. By the Peter-Weyl Theorem $\mathcal{E}(G)$ is dense in $L^2(G)$, from which part (1) follows. Part (3) is an immediate implication of (1) and (2).

In order to prove (4), note that if f is a class function, then

$$\mu * (f * g) = f * (\mu * g).$$

Since f_{N_j}'s are class functions, the claim follows. □

Remark 8.5. It follows from the above argument that $\mathcal{E}(G) = \bigcup_{i=1}^{\infty} L^2(G)^{N_i}$.

8.2. **The general case.** In the rest of this section we will prove a generalization of Lemma 8.4 that applies to general locally random groups. The results of this section will be crucially used in the next section to prove a generalization of Proposition 8.3. Another result of this section, Proposition 8.8, is a Fourier theoretic interpretation of the notion of *living at a given scale* (see Definition 8.7 for definition), which parallels the classical Paley-Littlewood theory.

A major difficulty in dealing with the general case is that unlike profinite groups, neighborhoods of identity are only *approximate* subgroups in general compact groups. Throughout this section, we will assume that the group G satisfies the following two properties:

(1) G is a compact group which is L-locally random with coefficient C_0.
(2) DC(C_1, d_0): for all $\eta > 0$

$$C_1^{-1}\eta^{d_0} \leq |1_\eta| \leq C_1\eta^{d_0}.$$

As in Proposition 8.3 we let η_0 be a small positive number, whose value will be specified later. Also fix

$$a \geq 4Ld_0, \qquad \text{and set} \qquad \eta_i := \eta_0^{a^i}, \text{for } i \geq 1.$$

As in Lemma 8.4, we define a family of operators $\Delta_i : L^2(G) \to L^2(G)$ by setting $\Delta_0(g) := P_{\eta_0} * g$ and for every $i \geq 1$

$$\Delta_i(g) := (P_{\eta_{i+1}} - P_{\eta_i}) * g. \tag{8.3}$$

Since P_η is invariant under conjugation, Δ_i's commute with any convolution operator (including convolution by a Borel probability measure), and for all $x, x' \in G$ we have

$$\lambda(x) \circ \rho(x') \circ \Delta_i = \Delta_i \circ \lambda(x) \circ \rho(x'),$$

where λ and ρ denote, respectively, the left and right-regular representations of G.

We showed previously that if 1_η's are subgroups, then $(\Delta_i(g))_{\eta_i} = 0$ and $(\Delta_i(g))_{\eta_{i+1}} = \Delta_i(g)$. We start by showing an approximate version of these equalities. In this section, we only establish properties of the operators Δ_i's, and postpone the discussion on their connections with spectral gap properties of T_μ to the next section.

Proposition 8.6. *In the setting of this section, if integers i, j, k satisfy $0 \leq j < i$ and $k > i+1$, then the following hold:*

- *(Averaging to zero)* $\|\Delta_i(g)_{\eta_j}\|_2 \ll_{C_0,C_1,L} \eta_0^{a^i/(4L+2)}\|g\|_2$.
- *(Almost invariant)* $\|\Delta_i(g)_{\eta_k} - \Delta_i(g)\|_2 \leq 2\eta_0^{a^k/(8L)}\|g\|_2$.

Proof. The argument for the first part is fairly similar to the one presented for Lemma 6.4. We let D be a threshold parameter whose value will be set later and estimate the corresponding low frequency and high frequency terms. By Lemma 6.2 we have

$$\|\widehat{P}_\eta(\pi) - I\|_{\mathrm{op}} \leq C_0(\dim \pi)^L \eta.$$

Combined with the trivial bound $\|\widehat{P}_\eta(\pi)\|_{\mathrm{op}} \leq 1$, this implies

$$\begin{aligned} L(\Delta_i(g)_{\eta_j}; D) &= \sum_{\pi\in\widehat{G},\dim\pi\leq D} \dim\pi \left\|\widehat{P}_{\eta_j}(\pi)\big(\widehat{P}_{\eta_{i+1}}(\pi) - \widehat{P}_{\eta_i}(\pi)\big)\widehat{g}(\pi)\right\|_{\mathrm{HS}}^2 \\ &\leq C_0^2 D^{2L}(\eta_{i+1}+\eta_i)^2 L(g;D) \leq 4C_0^2 D^{2L}\eta_i^2\|g\|_2^2. \end{aligned} \tag{8.4}$$

For the high frequency term, by Lemma 6.1 and the trivial bound $\|\widehat{P}_{\eta_{i+1}}(\pi) - \widehat{P}_{\eta_i}(\pi)\|_{\mathrm{op}} \leq 2$, we have

$$\begin{aligned} H(\Delta_i(g)_{\eta_j}; D) &= \sum_{\pi\in\widehat{G},\dim\pi> D} \dim\pi \left\|\widehat{P}_{\eta_j}(\pi)\big(\widehat{P}_{\eta_{i+1}}(\pi) - \widehat{P}_{\eta_i}(\pi)\big)\widehat{g}(\pi)\right\|_{\mathrm{HS}}^2 \\ &\leq \frac{4}{D}H(P_{\eta_j};D)H(g;D) \leq \frac{4}{D|1_{\eta_j}|}\|g\|_2^2 \leq \frac{4C_1}{D\eta_j^{d_0}}\|g\|_2^2. \end{aligned} \tag{8.5}$$

We choose D such that $4C_0^2D^{2L}\eta_i^2 = \frac{4C_1}{D\eta_j^{d_0}}$, which implies that D equals $\eta_j^{-d_0/(2L+1)}\eta_i^{-2/(2L+1)}$ up to a multiplicative factor, which is a function of the constants C_0, C_1, and L. Hence by (8.4) and (8.5) we get

$$\|\Delta_i(g)_{\eta_j}\|_2^2 \ll_{C_0,C_1,L} \eta_j^{-d_0+d_0/(2L+1)}\eta_i^{2/(2L+1)}\|g\|_2^2.$$

Notice that

$$\eta_j^{-d_0+d_0/(2L+1)}\eta_i^{2/(2L+1)} = \eta_0^{\frac{2}{2L+1}a^i - \frac{2Ld_0}{2L+1}a^j};$$

and $a^i - (2Ld_0)a^j \geq a^i(1-(2Ld_0)a^{-1}) \geq a^i/2$. Therefore,

$$\|\Delta_i(g)_{\eta_j}\|_2^2 \ll_{C_0,C_1,L} \eta_0^{a^i/(2L+1)}\|g\|_2^2;$$

and the first part follows.

For the second part we use Lemma 6.4 to obtain

$$\begin{aligned} \|\Delta_i(g)_{\eta_k} - \Delta_i(g)\|_2 =& \|\Delta_i(g_{\eta_k} - g)\|_2 \\ \leq& \|(g_{\eta_k}-g)_{\eta_{i+1}}\|_2 + \|(g_{\eta_k}-g)_{\eta_i}\|_2 \\ \leq& 2\eta_k^{1/(8L)}\|g\|_2. \end{aligned}$$

□

Definition 8.7. We say $g \in L^2(G)$ *lives at scale* η *(with parameter* a*)* if

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

From Proposition 8.6 we deduce that if $\|\Delta_i(g)\|_2/\|g\|_2 \gg 1$, then $\Delta_i(g)$ lives at scale η_i. The next proposition provides a Fourier theoretic understanding of this notion.

For every $\pi \in \widehat{G}$, let H_π denote the subspace of $L^2(G)$ spanned by the matrix coefficients of π. Given an interval $I \subset \mathbb{R}$, set

$$\mathcal{H}_I := \oplus_{\pi\in\widehat{G},\dim\pi\in I} H_\pi,$$

and denote by $\pi_I : L^2(G) \to \mathcal{H}_I$ the corresponding orthogonal projection.

Proposition 8.8. *Let* $0 < \eta < 1$ *be a parameter.*

(1) *Suppose $f \in L^2(G)$ lives at scale η. Then*

$$\|\pi_{I_\eta}(f)\|_2^2 \geq \big(1 - 8\eta^{1/(2a)}\big)\|f\|_2^2.$$

where $I_\eta = [\frac{1}{2C_0}\eta^{-1/(La)}, 2C_0\eta^{-d_0a^2}]$.

(2) *Let $I'_\eta = [C_1\eta^{-\frac{d_0+1}{a}}, C_0^{\frac{-1}{L}}\eta^{\frac{-2a^2+a}{2L}}]$. Then every $f \in \mathcal{H}_{I'_\eta}$ lives at scale η.*

Proof. Without loss of generality, assume that $\|f\|_2 = 1$. To see part (1) it suffices to show that

$$L(f; (2C_0)^{-1}\eta^{-1/(La)}) \leq 4\eta^{1/2a} \text{ and } \quad H(f; 2C_0\eta^{-d_0a^2}) \leq 4\eta^{a/2}. \tag{8.6}$$

By Lemma 6.3, for an arbitrary threshold D satisfying $C_0D^L\eta^{1/a} < 1$, we have

$$L(f; D) \leq (1 - C_0D^L\eta^{1/a})^{-2}L(f_{\eta^{1/a}}; D) \leq (1 - C_0D^L\eta^{1/a})^{-2}\eta^{1/(2a)}.$$

In the last inequality we used $\|f_{\eta^{1/a}}\|_2 \leq \eta^{1/(2a)}\|f\|_2$, which holds since f lives at scale η. Setting $D := \frac{1}{2C_0}\eta^{-1/(La)}$, the first inequality in (8.6) follows.

To show the second inequality in (8.6), we note that

$$\|f\|_2^2 - \|f_{\eta^{a^2}}\|_2^2 = (\|f\|_2 - \|f_{\eta^{a^2}}\|_2)(\|f\|_2 + \|f_{\eta^{a^2}}\|_2) \leq 2\|f - f_{\eta^{a^2}}\|_2 \leq 2\eta^{a/2}. \tag{8.7}$$

Since $\|P_{\eta^{a^2}}\|_1 = 1$, for all $\pi \in \widehat{G}$ we have $\|\widehat{P}_{\eta^{a^2}}(\pi)\|_{\mathrm{op}} \leq 1$. In consequence, Lemma 6.1 implies that for an arbitrary threshold D' we have

$$L(f; D') - L(f_{\eta^{a^2}}; D') \geq 0.$$

This and (8.7) imply that

$$H(f; D') - H(f_{\eta^{a^2}}; D') \leq 2\eta^{a/2}.$$

Altogether, we deduce

$$\begin{aligned}
H(f; D') \leq& 2\eta^{a/2} + H(f_{\eta^{a^2}}; D') \\
\leq& 2\eta^{a/2} + \frac{1}{D'}H(P_{\eta^{a^2}}; D')H(f; D') && \text{(by Lemma 6.1)} \\
\leq& 2\eta^{a/2} + \frac{1}{D'|1_{\eta^{a^2}}|}H(f; D') && \text{(by } H(P_{\eta^{a^2}}; D') \leq \|1_{\eta^{a^2}}\|_2^2\text{)} \\
\leq& 2\eta^{a/2} + \frac{C_0}{D'\eta^{d_0a^2}}H(f; D').
\end{aligned}$$

Therefore $\Big(1 - \frac{C_0}{D'\eta^{d_0a^2}}\Big)H(f; D') \leq 2\eta^{a/2}$. Setting $D' := 2C_0\eta^{-d_0a^2}$, the claim in part (1) follows.

We now turn to part (2). Let $f \in \mathcal{H}_{I'_\eta}$ be a unit vector. Note that for every π with $\dim\pi \notin I'_\eta$, $\hat{f}(\pi) = 0$. In particular, $L(f; D) = 0$ for any $D < C_1\eta^{-\frac{d_0+1}{a}}$. Therefore, by Lemma 6.1, we have

$$\begin{aligned}
\|f_{\eta^{1/a}}\|_2^2 = \|P_{\eta^{1/a}} * f\|_2^2 &\leq C_1^{-1}\eta^{(d_0+1)/a}\|P_{\eta^{1/a}}\|_2^2\|f\|_2^2 \\
&\leq C_1^{-1}\eta^{(d_0+1)/a}\frac{1}{|1_{\eta^{1/a}}|} \leq \eta^{1/a};
\end{aligned}$$

we used (DC) in the second inequality.

To verify the required bound for $\|f_{\eta^{a^2}} - f\|_2$, we use Lemma 6.2 combined with the fact that for every π with $\dim \pi \notin I'_\eta$, $\hat{f}(\pi) = 0$, and conclude that

$$\begin{aligned}\|f_{\eta^{a^2}} - f\|_2^2 &= \sum_{\dim \pi \in I'_\eta} \dim(\pi) \|(I - \hat{P}_{\eta^{a^2}}(\pi))\hat{f}(\pi)\|_{\mathrm{HS}}^2 \\ &\leq \sum_{\dim \pi \in I'_\eta} \dim(\pi) \|I - \hat{P}_{\eta^{a^2}}(\pi)\|_{\mathrm{op}}^2 \|\hat{f}(\pi)\|_{\mathrm{HS}}^2 \\ &\leq \sum_{\dim \pi \in I'_\eta} C_0^2 \dim(\pi)^{2L} \eta^{2a^2} \dim(\pi) \|\hat{f}(\pi)\|_{\mathrm{HS}}^2 \leq \eta^a.\end{aligned}$$

This completes the proof of part (2) and the lemma. □

We will now prove an almost orthogonality of the images of Δ_i's and show that their sum is dense in $L^2(G)$.

Lemma 8.9. *In the setting of this section, for non-negative integers $j < i - 1$, and $g \in L^2(G)$ we have*

$$\|\Delta_i \Delta_j\|_{\mathrm{op}} \ll_{C_0,C_1,L} \eta_i^{1/(4L+2)} \text{ and } |\langle \Delta_i(g), \Delta_j(g)\rangle| \ll_{C_0,C_1,L} \eta_i^{1/(4L+2)} \|g\|_2^2.$$

Proof. Since Δ_i is a self-adjoint operator, we have $\langle \Delta_i(g), \Delta_j(g)\rangle = \langle g, \Delta_i(\Delta_j(g))\rangle$; this implies

$$|\langle \Delta_i(g), \Delta_j(g)\rangle| \leq \|\Delta_i \Delta_j\|_{\mathrm{op}} \|g\|_2^2.$$

By the first part of Proposition 8.6 for $j > 0$ we have

$$\|\Delta_i \Delta_j(g)\|_2 = \|\Delta_i(g)_{\eta_{j+1}} - \Delta_i(g)_{\eta_j}\|_2 \ll_{C_0,C_1,L} \eta_i^{1/(4L+2)} \|g\|_2.$$

For $j = 0$ it is similar and the claims follow. □

Lemma 8.10. *In the setting of this section, $g = \sum_{i=0}^{\infty} \Delta_i(g)$ for any $g \in L^2(G)$.*

Proof. It suffices to show that for all $g \in L^2(G)$, $\|g - \sum_{i=1}^{n} \Delta_i(g)\|_2 = \|g - g_{\eta_{n+1}}\|_2$ tends to zero as $n \to \infty$. By the Peter-Weyl Theorem, for every $\varepsilon > 0$ there is $f \in C(G)$ such that $\|f - g\|_2 \leq \varepsilon$. Since G is compact, f is uniformly continuous. Let $\eta > 0$ be such that

$$d(x, y) \leq \eta \text{ implies that } |f(x) - f(y)| \leq \varepsilon.$$

For $n \gg_\varepsilon 1$, we have $\|f_{\eta_n} - f\|_\infty \leq \varepsilon$. Hence $\|f_{\eta_n} - f\|_2 \leq \varepsilon$. On the other hand, $\|f - g\|_2 \leq \varepsilon$ implies that $\|f_{\eta_n} - g_{\eta_n}\|_2 \leq \varepsilon$. Therefore for $n \gg_\varepsilon 1$ we have

$$\|g - g_{\eta_n}\|_2 \leq \|g - f\|_2 + \|f - f_{\eta_n}\|_2 + \|f_{\eta_n} - g_{\eta_n}\|_2 \leq 3\varepsilon.$$

Thus $\lim_{n\to\infty} g_{\eta_n} = g$ in L^2, from which the claim follows. □

By a similar argument as in the proof of the Cotlar-Stein Lemma (see [10, Lemma 6.3], and also [21, Chapter VII]), we will prove

Proposition 8.11. *In the setting of this section, for $\eta_0 \ll_{C_0,C_1,L} 1$, and $g \in L^2(G)$ we have*

$$\|g\|_2^2 \ll \sum_{i=0}^{\infty} \|\Delta_i(g)\|_2^2 \ll \|g\|_2^2. \tag{8.8}$$

In preparation for the proof we will need to establish some inequalities.

Lemma 8.12. *In the setting of this section, for a non-negative integer i, we have*

$$\sum_{j=0}^{\infty} \|\Delta_i \Delta_j\|_{\mathrm{op}}^{1/2} \ll_{C_0,C_1,L} 1.$$

Proof. By Lemma 8.9 and $\|\Delta_j\|_{op} \leq 2$, we get that

$$\sum_{j=0}^{\infty} \|\Delta_i \Delta_j\|_{\mathrm{op}}^{1/2} \leq 6 + O_{C_0,C_1,L}\Big(\sum_{j=1}^{\infty} \eta_0^{a^j/(4L+2)}\Big) \ll_{C_0,C_1,L} 1.$$

□

The proof of the next lemma is based on the proof of the Cotlar-Stein lemma.

Lemma 8.13. *In the above setting, for every $g \in L^2(G)$, we have*

$$\sum_{i,j} |\langle \Delta_i(g), \Delta_j(g)\rangle| \ll \|g\|_2^2.$$

Proof. For a given $g \in L^2(G)$, for every $i \neq j$, choose $u_{i,j} \in \mathbb{S}^1 \cup \{0\}$ such that $|\langle \Delta_i(g), \Delta_j(g)\rangle| = u_{i,j}\langle \Delta_i(g), \Delta_j(g)\rangle$ where $u_{i,j} = 0$ if $\langle \Delta_i(g), \Delta_j(g)\rangle = 0$. Then for every integer $N \geq 1$ we have

$$\sum_{0\leq i,j\leq N} |\langle \Delta_i(g), \Delta_j(g)\rangle| = \langle R_N(g), g\rangle,$$

where $R_N = \sum_{0\leq i,j\leq N} u_{i,j}\Delta_j\Delta_i$. Thus, it is enough to prove that for all possible choices of $u_{i,j}$ and all $N \geq 1$ we have $\|R_N\|_{\mathrm{op}} \leq \Phi$ for a fixed positive number Φ. Since Δ_i's are self-adjoint and pairwise commuting, for every positive integer k we have $\|R_N^k\|_{\mathrm{op}} = \|R_N\|_{\mathrm{op}}^k$. By the triangle inequality, we have

$$\|R_N\|_{op}^k \leq \sum_{0\leq i_l,j_l\leq N, \forall 1\leq l\leq k} \|\Delta_{i_1}\Delta_{j_1}\cdots\Delta_{i_k}\Delta_{j_k}\|_{\mathrm{op}}.$$

Since

$$\|\Delta_{i_1}\Delta_{j_1}\cdots\Delta_{i_k}\Delta_{j_k}\| \leq \min\Big(\prod_{l=1}^{k}\|\Delta_{i_l}\Delta_{j_l}\|_{\mathrm{op}}, \|\Delta_{i_1}\|_{\mathrm{op}}\|\Delta_{j_k}\|_{\mathrm{op}}\prod_{l=1}^{k-1}\|\Delta_{j_l}\Delta_{i_{l+1}}\|_{\mathrm{op}}\Big),$$

we have that

$$\|\Delta_{i_1}\Delta_{j_1}\cdots\Delta_{i_k}\Delta_{j_k}\|_{\mathrm{op}} \leq 4\Big(\prod_{l=1}^{k}\|\Delta_{i_l}\Delta_{j_l}\|_{\mathrm{op}}\prod_{l=1}^{k-1}\|\Delta_{j_l}\Delta_{i_{l+1}}\|_{\mathrm{op}}\Big)^{1/2}.$$

Altogether we get

$$\begin{aligned}\|R_N\|_{\mathrm{op}}^k &\leq 4\sum_{i_1=0}^{N}\sum_{j_1=0}^{N}\cdots\sum_{j_k=0}^{N}\Big(\prod_{l=1}^{k}\|\Delta_{i_l}\Delta_{j_l}\|_{\mathrm{op}}\prod_{l=1}^{k-1}\|\Delta_{j_l}\Delta_{i_{l+1}}\|_{\mathrm{op}}\Big)^{1/2}\\ &= 4\sum_{i_1=0}^{N}\sum_{j_1=0}^{N}\cdots\sum_{i_k=0}^{N}\Big(\prod_{l=1}^{k-1}\|\Delta_{i_l}\Delta_{j_l}\|_{\mathrm{op}}\prod_{l=1}^{k-1}\|\Delta_{j_l}\Delta_{i_{l+1}}\|_{\mathrm{op}}\Big)^{1/2}\Big(\sum_{j_k=0}^{N}\|\Delta_{i_k}\Delta_{j_k}\|_{\mathrm{op}}^{1/2}\Big).\end{aligned} \tag{8.9}$$

By repeatedly using Lemma 8.12, it follows that there is a constant $M := M(C_0, C_1, L)$ such that

$$\|R_N\|_{\mathrm{op}}^k \leq 4(N+1)M^{2k-1},$$

which implies $\|R_N\|_{\mathrm{op}} \leq 4^{1/k}(N+1)^{1/k}M^2$ for any positive integer k. The claim follows from here. □

Corollary 8.14. *In the setting of this section, for $g \in L^2(G)$ we have that*

$$\sum_{i=0}^{\infty}\|\Delta_i(g)\|_2^2 \ll \|g\|_2^2, \text{ and } \sum_{i=0}^{\infty}|\langle \Delta_i(g), \Delta_{i+1}(g)\rangle| \leq \sum_{i=0}^{\infty}\|\Delta_i(g)\|_2^2.$$

Proof. The first inequality is a weaker version of the inequality given in Lemma 8.13. Applying the Cauchy-Schwarz inequality twice, we obtain

$$\begin{aligned}\sum_{i=0}^{\infty}|\langle\Delta_i(g),\Delta_{i+1}(g)\rangle| &\le \sum_{i=0}^{\infty}\|\Delta_i(g)\|_2\|\Delta_{i+1}(g)\|_2\\ &\le\Big(\sum_{i=0}^{\infty}\|\Delta_i(g)\|_2^2\Big)^{1/2}\Big(\sum_{i=0}^{\infty}\|\Delta_{i+1}(g)\|_2^2\Big)^{1/2}\\ &\le\sum_{i=0}^{\infty}\|\Delta_i(g)\|_2^2.\end{aligned}$$

□

Proof of Proposition 8.11. By Lemma 8.10 we have $g=\sum_{i=1}^{\infty}\Delta_i(g)$. It follows that

$$\begin{aligned}\|g\|_2^2 &= \sum_{0\le i,j}\langle\Delta_i(g),\Delta_j(g)\rangle\\ &=\sum_{i=0}^{\infty}\|\Delta_i(g)\|_2^2+2\sum_{i=0}^{\infty}\langle\Delta_i(g),\Delta_{i+1}(g)\rangle+2\sum_{0\le i<j,|i-j|>1}\langle\Delta_i(g),\Delta_j(g)\rangle\\ &\le 3\sum_{i=0}^{\infty}\|\Delta_i(g)\|_2^2+2\sum_{0\le i<j,|i-j|>1}|\langle\Delta_i(g),\Delta_j(g)\rangle| \qquad (8.10)\\ &\le 3\sum_{i=0}^{\infty}\|\Delta_i(g)\|_2^2+O_{C_0,C_1,L}\Big(\sum_{0\le i<j,|i-j|>1}\eta_0^{a^j/(4L+2)})\|g\|_2^2\Big) \qquad (8.11)\\ &\le 3\sum_{i=0}^{\infty}\|\Delta_i(g)\|_2^2+O_{C_0,C_1,L}(\eta_0^{a/(4L+2)})\|g\|_2^2\\ &\le 3\sum_{i=0}^{\infty}\|\Delta_i(g)\|_2^2+(1/2)\|g\|_2^2,\end{aligned}$$

where (8.10) is deduced from Corollary 8.14 and (8.11) follows from Lemma 8.9. The reverse inequality is already proven in Corollary 8.14. □

9. Littlewood-Paley decomposition and spectral gap

The main goal of this section is to prove Theorem 2.10 which is a generalization of Proposition 8.3 for general locally random groups. At the end, we will show how the existence of spectral gap can be reduced to study of the gap for functions that live at small scales, Theorem 9.3.

We continue to assume that G is a compact group satisfying the following two properties:

(1) G is an L-locally random group with coefficient C_0.
(2) DC(C_1,d_0): for all $\eta>0$
$$C_1^{-1}\eta^{d_0}\le|1_\eta|\le C_1\eta^{d_0}.$$

Fix $a>\max(4Ld_0,4L+2)$, and set η_0 to be a sufficiently small positive number whose value will be determined later and $\eta_i:=\eta_0^{a^i}$. Define $(\Delta_j)_{j\ge0}$ as in (8.3). We begin with a basic property of these operators.

Lemma 9.1. *For all $j\ge0$, Δ_j is a compact operator. Moreover, for any symmetric Borel probability measure μ on G, there exists an orthonormal basis $\{e_i\}_{i=1}^{\infty}$ of $L^2(G)$ consisting of common eigenfunctions of $\{\Delta_j:j\ge0\}$ and T_μ.*

Proof. Since Δ_j is a convolution operator by a function in $L^2(G)$, it is a compact operator. Further, since 1_η is a symmetric subset, Δ_j is a self-adjoint operator.

The construction of an orthonormal basis consisting of eigenvectors for $\{\Delta_j\}$ and T_μ follows from standard arguments in view of commutativity of the family, compactness of $\{\Delta_j\}$, and the fact that $f = \sum_{j=0}^{\infty} \Delta_j(f)$ for any $f \in L^2(G)$. □

Lemma 9.2. *In the setting of this section, suppose $\{e_i\}_{i=1}^{\infty}$ is an orthonormal basis of $L^2(G)$ which consists of common eigenfunctions of Δ_j's (see Lemma 9.1). Suppose $\Delta_j(e_i) = \alpha_{ji} e_i$ for all $i \geq 1$ and $j \geq 0$. Then*

- $\|(e_i)_{\eta_{j-1}}\|_2 \ll_{C_0,C_1,L} |\alpha_{ji}|^{-1} \eta_j^{1/(4L+2)}$.
- $\|(e_i)_{\eta_{j+2}} - e_i\|_2 \leq 2|\alpha_{ji}|^{-1} \eta_{j+2}^{1/(8L)}$.

In particular, if $|\alpha_{ji}| \geq \eta_j^{1/(8L+4)}$, then e_i lives at scale η_j.

Proof. This is an immediate consequence of Proposition 8.6. □

Proof of Theorem 2.10. We will use the above notation. Let $I_j := \{i \in \mathbb{Z}^+ | \, |\alpha_{ji}| \geq \eta_j^{1/(8L+4)}\}$, $E := \mathbb{Z}^+ \setminus \bigcup_{j=1}^{\infty} I_j$, and for $i \in I_j$ we let $\mathcal{H}_{ji} := \ker(\Delta_j - \alpha_{ji} I)$.

We will show the claim holds with $\mathcal{H}_0$ the space spanned by $\{e_i : i \in E\}$. Let us first show that $\mathcal{H}_0$ is finite dimensional. By definition, for all $i \in E$ and all positive integers j, we have

$$|\alpha_{ji}| \leq \eta_j^{1/(8L+4)}.$$

On the other hand, by Lemma 8.10 we have $\sum_{j=0}^{\infty} \alpha_{ji} = 1$. Therefore

$$|1 - \alpha_{0i}| \leq \sum_{j=1}^{\infty} \eta_j^{1/(8L+4)} \leq \eta_0^{1/(8L+4)}.$$

Therefore $\alpha_{0i} > 1 - \eta_0^{1/(8L+4)}$ for any $i \in E$.

Notice that Δ_0 is a Hilbert-Schmidt operator with kernel $k(x,y) := P_{\eta_0}(xy^{-1})$. Therefore $P_{\eta_0}(xy^{-1}) = \sum_i \alpha_{0i} e_i(x)\overline{e_i(y)}$. This implies that

$$\frac{1}{|1_{\eta_0}|} = \int_G \int_G P_{\eta_0}(xy^{-1})^2 \, \mathrm{d}y \, \mathrm{d}x = \sum_i |\alpha_{0i}|^2.$$

By the above equality, we get

$$(1 - \eta_0^{1/(8L+4)})^2 \, \#E \leq \frac{1}{|1_{\eta_0}|};$$

which implies that $\dim \mathcal{H}_0 \leq \frac{2}{|1_{\eta_0}|}$.

We now investigate spectral properties of T_μ on $\mathcal{H}_{ji} = \ker(\Delta_j - \alpha_{ji} I)$. It is clear that $\mathcal{H}_{ji}$ is a finite-dimensional subrepresentation of $L^2(G)$. Since e_k's are also eigenfunctions of T_μ,

$$\mathcal{L}(\mu; \mathcal{H}_{ji}) = \min\{-\log \|\mu * e_k\|_2 : e_k \in \mathcal{H}_{ij}\}.$$

Let $\nu = \mu^{(l)}$ for some positive integer l to be specified later, and let $e_k \in \mathcal{H}_{ij}$; note that $\alpha_{jk} = \alpha_{ji}$. By the definition of $\mathcal{H}_{ij}$ and Lemma 9.2, e_k lives at scale η_j. Thus we have

$$\big| \|(e_k)_{\eta_{j+2}} * \nu\|_2 - \|(e_k * \nu)\|_2 \big| \leq \|((e_k)_{\eta_{j+2}} - e_k) * \nu\|_2 \leq \eta_j^{a/2},$$

which implies that $|\|(e_k)_{\eta_{j+2}} * \nu\|_2^2 - \|e_k * \nu\|_2^2| \leq 2\eta_j^{a/2}$. Therefore,

$$\|e_k * \nu\|_2^2 \leq 2\eta_j^{a/2} + \|(e_k)_{\eta_{j+2}} * \nu\|_2^2. \tag{9.1}$$

On the other hand, by the Mixing Inequality (see Theorem 2.6), we have

$$\begin{aligned}\|(e_k)_{\eta_{j+2}} * \nu\|_2^2 =& \|e_k * \nu_{\eta_{j+2}}\|_2^2 \\ \le& 2\|(e_k)_{\eta_j^{1/a}}\|_2^2 \|(\nu_{\eta_{j+2}})_{\eta_j^{1/a}}\|_2^2 + \eta_j^{1/(8aL)}\|\nu_{\eta_{j+2}}\|_2^2 \\ \le& (2\eta_j^{1/a} + \eta_j^{1/(8aL)})\|\nu_{\eta_{j+2}}\|_2^2 \le 3\eta_j^{1/(8aL)}\|\nu_{\eta_{j+2}}\|_2^2 \end{aligned} \tag{9.2}$$

where the second inequality follows from the fact that e_k lives as scale η_j.

By (9.1) and (9.2), for every $k \in I_j$, we have

$$\begin{aligned}-2\log\big(\|e_k * \nu\|_2\big) \ge& -\log\big(2\eta_j^{a/2} + 3\eta_j^{1/(8aL)}\|\nu_{\eta_{j+2}}\|_2^2\big) \\ \ge& -\log 5 - \log\big(\max(\eta_j^{a/2}, \eta_j^{1/(8aL)}\|\nu_{\eta_{j+2}}\|_2^2)\big).\end{aligned}$$

For $\eta_0 \ll_{L,d_0} 1$ small enough, one obtains

$$-2\log(\|e_k * \nu\|_2) \ge \min\Big(-\frac{1}{3a}\log\eta_{j+2}, -\frac{1}{9a^3L}\log\eta_{j+2} - \log\|\nu_{\eta_{j+2}}\|_2^2\Big). \tag{9.3}$$

By Lemma 7.1 and the dimension condition, we have

$$|h(G;\eta) - \log(1/|1_\eta|)| \ll_{d_0,C_1} 1, \text{ and } |\log(1/|1_\eta|) + d_0\log\eta| \ll_{d_0,C_1} 1. \tag{9.4}$$

Hence for $\eta_0 \ll_{C_0,C_1,L} 1$, by (9.3) and (9.4) we have

$$\begin{aligned}-2\log(\|e_k * \nu\|_2) \ge& \min\left(\frac{1}{4d_0a}h(G;\eta_j), \frac{1}{10Ld_0a^3}h(G;\eta_j) - \log\|\nu_{\eta_{j+2}}\|_2^2\right) \\ \ge& \min\left(\frac{1}{4d_0a}h(G;\eta_j), H_2(\nu;\eta_{j+2}) - \Big(1 - \frac{1}{10Ld_0a^3}\Big)h(G;\eta_j)\right).\end{aligned} \tag{9.5}$$

By the assumption for some $l_{j+2} \le C_2 h(G;\eta_{j+2})$, we have

$$H_2(\mu^{(l_{j+2})};\eta_{j+2}) \ge \Big(1 - \frac{1}{20Ld_0a^3}\Big)h(G;\eta_{j+2});$$

and so by applying the inequality (9.5) to $\nu = \mu^{(l_{j+2})}$ for every $i \in I_j$ we have

$$\mathcal{L}(\mu;\mathcal{H}_{ji}) \ge \min\Big(\frac{1}{8C_2d_0a}, \frac{1}{40C_2Ld_0a^3}\Big) = \frac{1}{40C_2Ld_0a^3}. \tag{9.6}$$

Altogether, (9.6) and the definition of $\mathcal{H}_0$ imply

$$\mathcal{L}(\mu; L^2(G) \ominus \mathcal{H}_0) \ge \frac{1}{40C_2Ld_0a^3},$$

as we claimed.

Since the group generated by the support of μ is dense in G and $\dim\mathcal{H}_0 < \infty$, it follows that $\mathcal{L}(\mu; L_0^2(G)) > 0$. □

The following theorem is a corollary of the proof of Theorem 2.10.

Theorem 9.3. *In the above setting, suppose μ is a symmetric Borel probability measure on G, and the group generated by the support of μ is dense in G. Suppose that there exist $C_3 > 0$, $c > 0$, and $0 < \eta_0 < 1$ such that for every $\eta \le \eta_0$ and every function $g \in L^2(G)$ which lives at scale η there exists $l \le C_3\log(1/\eta)$ such that*

$$\|\mu^{(l)} * g\|_2 \le \eta^c\|g\|_2.$$

Then there is a subrepresentation $\mathcal{H}_0$ of $L^2(G)$ with $\dim\mathcal{H}_0 \le 2C_0\eta_0^{-d_0}$ such that

$$\mathcal{L}(\mu; L^2(G) \ominus \mathcal{H}_0) \ge \frac{c}{C_3}.$$

In particular, $\mathcal{L}(\mu; G) > 0$.

Proof. Without loss of generality, assume that η_0 is sufficiently small so that Theorem 2.10 holds. As before, fix $a > \max(4Ld_0, 4L+2)$, and for $i \geq 1$, set $\eta_i := \eta_0^{a^i}$. Let $\{e_i\}_{i=1}^{\infty}$, the sets I_j's, and E be as in the proof of Theorem 2.10. Define $\mathcal{H}_0$ as in that proof as well.

For all $i \in I_j$, e_i is function which lives at scale η_j. This, together with the assumption, implies that $\|\mu^{(l_{ji})} * e_i\|_2 \leq \eta_j^c$ for some positive integer $l_{ji} \leq C_3 \log(1/\eta_j)$. Hence

$$C_3 \log(1/\eta_j)\mathcal{L}(\mu; \mathcal{H}_{ji}) \geq -c \log \eta_j,$$

where $\mathcal{H}_{ji} := \ker(\Delta_j - \alpha_{ji} I)$. In view of this, we have $\mathcal{L}(\mu; L^2(G) \ominus \mathcal{H}_0) \geq c/C_3$.

Finally, since the group generated by the support of μ is dense in G and $\mathcal{H}_0$ is finite dimensional, it follows that $\mathcal{L}(\mu; L_0^2(G)) > 0$. □

10. Gaining entropy in a multi-scale setting

The goal of this section is to prove Theorem 2.12. In their seminal work [7], Bourgain and Gamburd proved that, if X and Y are random variables taking values in a finite group G, then the Rényi entropy of XY will be substantially larger than the average of the Rényi entropies of X and Y, unless there is an algebraic obstruction, see also [23, Lemma 15]. This type of result had been proved earlier for random variables X and Y that are uniformly distributed in subsets A and B, respectively. For abelian groups, this is due to Balog and Szemerédi [2] and Gowers [15]. For general groups, this was proved by Tao [22]. In the same work, Tao also proves a multi-scale version of this result. In this section, we will prove a multi-scale version of the aforementioned result of Bourgain and Gamburd, which can be considered as a weighted version of [22]. Similar results have been proved earlier for some specific groups in [14, 5, 17, 10]. We start by recalling the definition of an approximate subgroup.

Definition 10.1. For $K \geq 1$, a subset X of a group G is called a *K-approximate subgroup* if X is symmetric, that is, $X = X^{-1}$ and there exists $T \subseteq X \cdot X$ with $\#T \leq K$, such that $X \cdot X \subseteq T \cdot X$.

Recall also that if X is a random variable taking finitely many values, then the Rényi entropy (of order 2) of X is defined by

$$H_2(X) = -\log\left(\sum_x \mathbb{P}(X = x)^2\right),$$

where, here and in what follows log refers to logarithm in base 2. It is easy to see that when X and Y take values in a group G then $H_2(XY) \geq \frac{H_2(X)+H_2(Y)}{2}$ holds.

Theorem 10.2 (Bourgain-Gamburd)**.** *Let G be a finite group and suppose X and Y are two G-valued random variables. If*

$$H_2(XY) \leq \frac{H_2(X) + H_2(Y)}{2} + \log K$$

for some positive number $K \geq 2$, then there exists $H \subseteq G$ such that:

1. *(Approximate structure) H is an $O(K^{O(1)})$-approximate subgroup.*
2. *(Controlling the order) $|\log(\#H) - H_2(X)| \ll \log K$.*
3. *(Almost equidistribution) There are elements $x, y \in G$ such that for all $h \in H$*

$$\mathbb{P}(X = xh) \geq K^{-O(1)}(\#H)^{-1}, \qquad \mathbb{P}(Y = hy) \geq K^{-O(1)}(\#H)^{-1}.$$

More generally, suppose that G is an arbitrary compact group and $A, B \subseteq G$ are two measurable subsets of positive measure. The energy of the pair (A, B) is defined by

$$E(A, B) := \|\mathbb{1}_A * \mathbb{1}_B\|_2^2. \tag{10.1}$$

When G is finite, this reduces to

$$E(A,B) = \#Q(A,B)/(\#G)^3,$$

where

$$Q(A,B) := \{(a,b,a',b') \in A\times B\times A\times B |\ ab = a'b'\}.$$

For general compact groups, the notion of *η-approximate energy* has been introduced in [22]. We will work with two different metrics on G^4: For $(g_i)_{1\le i\le 4}$ and $(g'_i)_{1\le i\le 4}$ in G^4, define

$$\begin{aligned} d^+\big((g_i)_{1\le i\le 4},(g'_i)_{1\le i\le 4}\big) &:= \sum_{1\le i\le 4} d(g_i,g'_i), \quad \text{and} \\ d\big((g_i)_{1\le i\le 4},(g'_i)_{1\le i\le 4}\big) &:= \max_{1\le i\le 4} d(g_i,g'_i). \end{aligned} \tag{10.2}$$

For non-empty $A,B\subseteq G$ and $\eta>0$, we let

$$E_\eta(A,B) := \mathcal{N}_\eta(Q_\eta(A,B)), \tag{10.3}$$

where

$$Q_\eta(A,B) := \{(a,b,a',b') \in A\times B\times A\times B |\ ab \in (a'b')_\eta\}$$

where $\mathcal{N}_\eta$ is computed with respect to d^+.

The results of this section are proved under a weaker dimension condition that we now define. We say that (G,d) satisfies *the dimension condition at scale η with parameter C'* if there exist $C>1$ and $d_0>0$ such that

$$C^{-1}\eta^{d_0} \le |1_{c\eta}| \le C\eta^{d_0}$$

holds for all $c\in[C'^{-1},C']$.

Abusing the notation, for two positive quantities X and Y we write $X\preccurlyeq Y$ if X/Y is bounded from above by an expression of the form $\Omega^{O(1)}$, where $\Omega = 2^{d_0}C^2$. If $X\preccurlyeq Y$ and $Y\preccurlyeq X$, we write $X\approx Y$.

Theorem 10.3 ([22], Theorem 6.10)**.** *Suppose G is a compact group with a fixed bi-invariant metric. Suppose $A,B\subseteq G$ are non-empty. For every $\eta>0$ and $K\succcurlyeq 1$, if G satisfies the dimension condition at scale η with parameter C' (which is a large universal constant), and the energy bound*

$$E_\eta(A,B) \gg K^{-O(1)}\mathcal{N}_\eta(A)^{3/2}\mathcal{N}_\eta(B)^{3/2} \tag{EB}$$

holds, then there is $H\subseteq G$ such that

1. *(Approximate structure) H is an $K^{O(1)}$-approximate subgroup.*
2. *(Controlling the metric entropy) $|h(H;\eta) - \frac{h(A;\eta)+h(B;\eta)}{2}| \le \log K$.*
3. *(Large intersection) There are $x,y\in G$, such that $|h(A\cap xH;\eta) - h(A;\eta)| \le \log K$ and $|h(B\cap Hy;\eta) - h(B;\eta)| \le \log K$.*

Theorem 2.12 is both a multi-scale version of Theorem 10.2 and a weighted version of Theorem 10.3.

Let X and Y be Borel random variables whose distributions are given by measures μ and ν, respectively. Let $\mu_\eta := \mu * P_\eta$ and $\nu_\eta := \nu * P_\eta$. The idea of the proof is to approximate μ_η and ν_η by step functions, and find subsets of η-neighborhoods of supports of μ and ν with large η-approximate energy. We will then apply Theorem 10.3 to finish the proof. The following lemma summarizes some of the properties of the function μ_η.

Lemma 10.4. *Suppose G is a compact group and G satisfies the dimension condition at scale η with parameter C' for some $C'\gg 1$ (larger than a universal constant). Suppose μ and ν are two Borel probability measures on G and $f\in L^2(G)$ is non-negative. Then*

(1) *For all* $y \in x_\eta$ *and* $c \in [C'^{-1}, C'-1]$, *we have* $\mu_{c\eta}(y) \preccurlyeq \mu_{(c+1)\eta}(x)$, *and* $f_{c\eta}(y) \preccurlyeq f_{(c+1)\eta}(x)$; *in particular* $\mu_\eta(y) \preccurlyeq \mu_{2\eta}(x) \preccurlyeq \mu_{3\eta}(y)$.
(2) *For any* $\eta, \eta' > 0$ *and* $y \in G$, *we have* $P_{\eta'}(y) \leq \frac{|1_{\eta+\eta'}|}{|1_{\eta'}|} P_{\eta'+\eta} * P_\eta(y)$. *(see [10, Lemma A.5])*
(3) *For* $c \in [(C'-1)^{-1}, (C'-1)]$, *we have* $\|\mu_{c\eta}\|_2 \approx \|\mu_\eta\|_2$ *and* $\|f_{c\eta}\|_2 \approx \|f_\eta\|_2$.
(4) $\|\mu_\eta * \nu_\eta\|_2 \leq \|(\mu * \nu)_\eta\|_2 \preccurlyeq \|\mu_\eta * \nu_\eta\|_2$.

Proof. The sequence of inequalities

$$\mu_{c\eta}(y) = \frac{\mu(y_{c\eta})}{|1_{c\eta}|} \leq \frac{|1_{(c+1)\eta}|}{|1_{c\eta}|} \cdot \frac{\mu(x_{(c+1)\eta})}{|1_{(c+1)\eta}|} \preccurlyeq \mu_{(c+1)\eta}(x)$$

proves the first claim of part (1). The second claim of (1) is a special case. Part (2) is an easy consequence of the fact that, if $y \in 1_{\eta'}$, then for any $x \in 1_\eta$ we have $x^{-1}y \in 1_{\eta'+\eta}$.

For part (3), by symmetry we can and will assume that $c > 1$. Note that

$$\mu_\eta(y) = \frac{\mu(y_\eta)}{|1_\eta|} \leq \frac{|1_{c\eta}|}{|1_\eta|} \cdot \frac{\mu(y_{c\eta})}{|1_{c\eta}|} \preccurlyeq \mu_{c\eta}(y).$$

Hence, we have $\|\mu_\eta\|_2 \preccurlyeq \|\mu_{c\eta}\|_2$, and, in particular, $\|f_\eta\|_2 \preccurlyeq \|f_{c\eta}\|_2$. In order to prove the reverse inequality, note that by (2) we have $\mu_{c\eta} \preccurlyeq P_{(c+1)\eta} * \mu_\eta$ and $f_{c\eta} \preccurlyeq P_{(c+1)\eta} * f_\eta$. These imply that

$$\|\mu_{c\eta}\|_2 \preccurlyeq \|P_{(c+1)\eta} * \mu_\eta\|_2 \leq \|\mu_\eta\|_2 \quad \text{and} \quad \|f_{c\eta}\|_2 \preccurlyeq \|P_{(c+1)\eta} * f_\eta\|_2 \leq \|f_\eta\|_2.$$

Finally, to prove (4), first note that

$$\|\mu_\eta * \nu_\eta\|_2 = \|P_\eta * (\mu * \nu)_\eta\|_2 \leq \|(\mu * \nu)_\eta\|_2.$$

Part (2) implies that $P_\eta \preccurlyeq P_{2\eta} * P_\eta$, which, in turn, shows that

$$(\mu * \nu)_\eta \preccurlyeq \mu_{2\eta} * \nu_\eta. \tag{10.4}$$

On the other hand, using (3) and the fact that $\mu * \nu_\eta$ is a non-negative function, we have

$$\|\mu_{2\eta} * \nu_\eta\|_2 = \|(\mu * \nu_\eta)_{2\eta}\|_2 \approx \|(\mu * \nu_\eta)_\eta\|_2 = \|\mu_\eta * \nu_\eta\|_2; \tag{10.5}$$

applying (10.4) and (10.5) we obtain the desired inequality. □

From now on, we will assume that μ and ν denote the distributions of the random variables X and Y, respectively, and that the inequality

$$H_2(XY;\eta) \leq \log K + \frac{H_2(X;\eta) + H_2(Y;\eta)}{2}$$

holds. Hence we have

$$\|(\mu * \nu)_\eta\|_2 \geq K^{-1} \|\mu_\eta\|_2^{1/2} \|\nu_\eta\|_2^{1/2}.$$

By Lemma 10.4 and the above inequality we deduce that

$$\|\mu_\eta * \nu_\eta\|_2 \succcurlyeq K^{-1} \|\mu_\eta\|_2^{1/2} \|\nu_\eta\|_2^{1/2}. \tag{10.6}$$

By (3.3), we have $\|\mu_\eta * \nu_\eta\|_2 \leq \min(\|\mu_\eta\|_2, \|\nu_\eta\|_2)$, which implies

$$K^{-2} \|\mu_\eta\|_2 \preccurlyeq \|\nu_\eta\|_2 \preccurlyeq K^2 \|\mu_\eta\|_2. \tag{10.7}$$

To find the desired step function approximation of μ_η, we discretize G and then choose subsets of this discrete model according to the value of μ_η. We fix a maximal η-separating subset $\mathcal{C}$ of G.

As it was mentioned in Remark 7.3, the proof of Lemma 7.1 only uses the dimension condition for $\eta, \eta/2$ and 2η. Hence for $c \in [(C'/2)^{-1}, C'/2]$ we have

$$\mathcal{N}_{c\eta}(A) \approx \frac{|A_\eta|}{|1_\eta|}. \tag{10.8}$$

We partition $\mathcal{C}$ according to the value of $\mu_{2\eta}$ as follows:

$$\mathcal{C}(\mu;>) := \{x \in \mathcal{C} |\ \mu_{2\eta}(x) > K^{10}\|\mu_\eta\|_2^2\}, \tag{10.9}$$

$$\mathcal{C}(\mu;<) := \{x \in \mathcal{C} |\ \mu_{2\eta}(x) < K^{-10}\|\mu_\eta\|_2^2\}, \tag{10.10}$$

and

$$\mathcal{C}(\mu;\sim) := \{x \in \mathcal{C} |\ K^{-10}\|\mu_\eta\|_2^2 \le \mu_{2\eta}(x) \le K^{10}\|\mu_\eta\|_2^2\}. \tag{10.11}$$

We also define the following functions:

$$\mu_\eta^> := \mathbb{1}_{\mathcal{C}(\mu;>)_\eta} \cdot \mu_\eta,\ \ \mu_\eta^< := \mathbb{1}_{\mathcal{C}(\mu;<)_\eta} \cdot \mu_\eta, \tag{10.12}$$

and

$$\mu_\eta^\sim(x) := \begin{cases} \mu_\eta(x) & \text{if } x \notin \mathcal{C}(\mu;>)_\eta \cup \mathcal{C}(\mu;<)_\eta \\ 0 & \text{otherwise.} \end{cases}$$

And so $\mu_\eta(x) \le \mu_\eta^>(x)+\mu_\eta^<(x)+\mu_\eta^\sim(x)$, and inequality can possibly occur only in $\mathcal{C}(\mu;>)_\eta \cap \mathcal{C}(\mu;<)_\eta$. The functions $\mu_\eta^>$ and $\mu_\eta^<$ should be viewed as *tails* of μ_η and will now be shown to be negligible.

Lemma 10.5. *In the above setting,* $\|\mu_\eta^>\|_1 \preccurlyeq K^{-10}$ *and* $\|\mu_\eta^<\|_2 \preccurlyeq K^{-5}\|\mu_\eta\|_2$.

Proof. For any $y \in \mathcal{C}(\mu;>)_\eta$, there is $x \in \mathcal{C}(\mu,>)$ such that $y \in x_\eta$. Applying part (1) of Lemma 10.4 we have

$$\mu_{3\eta}(y) \succcurlyeq \mu_{2\eta}(x) > K^{10}\|\mu_\eta\|_2^2.$$

On the other hand, by part (3) of Lemma 10.4 we have $\|\mu_\eta\|_2 \approx \|\mu_{3\eta}\|_2$. Hence, we have

$$\|\mu_\eta\|_2^2 \succcurlyeq \int_{\mathcal{C}(\mu,>)_\eta} \mu_{3\eta}(y)^2\ \mathrm{d}y \succcurlyeq K^{10}\|\mu_\eta\|_2^2 \int_{\mathcal{C}(\mu,>)_\eta} \mu_\eta(y)\ \mathrm{d}y = K^{10}\|\mu_\eta\|_2^2\|\mu_\eta^>\|_1,$$

which implies the first inequality.

For any $y \in \mathcal{C}(\mu,<)_\eta$, there is $x \in \mathcal{C}(\mu,<)$ such that $y \in x_\eta$; and so by part (1) of Lemma 10.4 we have $\mu_\eta(y) \preccurlyeq \mu_{2\eta}(x) \le K^{-10}\|\mu_\eta\|_2^2$. Therefore

$$\|\mu_\eta^<\|_2^2 = \int_{\mathcal{C}(\mu,<)_\eta} \mu_\eta(y)^2\ \mathrm{d}y \preccurlyeq K^{-10}\|\mu_\eta\|_2^2 \int_{\mathcal{C}(\mu,<)_\eta} \mu_\eta(y)\ \mathrm{d}y \le K^{-10}\|\mu_\eta\|_2^2;$$

and the second inequality follows. □

Corollary 10.6. *In the above setting,* $\|\mu_\eta^\sim * \nu_\eta^\sim\|_2 \ge (2K)^{-1}\|\mu_\eta\|_2^{1/2}\|\nu_\eta\|_2^{1/2}$ *if* $K \succcurlyeq 1$.

Proof. For all $y \in G$, we have $\mu_\eta(y) \ge \mu_\eta^\sim(y)$. By Lemma 10.5, and (10.7), we have

$$\|\mu_\eta^> * \nu_\eta\|_2 \preccurlyeq K^{-10}\|\nu_\eta\|_2 \preccurlyeq K^{-9}\|\mu_\eta\|_2^{1/2}\|\nu_\eta\|_2^{1/2}, \tag{10.13}$$

$$\|\mu_\eta^< * \nu_\eta\|_2 \le \|\mu_\eta^<\|_2 \preccurlyeq K^{-5}\|\mu_\eta\|_2 \preccurlyeq K^{-4}\|\mu_\eta\|_2^{1/2}\|\nu_\eta\|_2^{1/2}, \tag{10.14}$$

$$\|\mu_\eta^\sim * \nu_\eta^>\|_2 \preccurlyeq K^{-10}\|\mu_\eta^\sim\|_2 \le K^{-10}\|\mu_\eta\|_2 \preccurlyeq K^{-9}\|\mu_\eta\|_2^{1/2}\|\nu_\eta\|_2^{1/2}, \tag{10.15}$$

$$\|\mu_\eta^\sim * \nu_\eta^<\|_2 \le \|\mu_\eta^\sim\|_1\|\mu_\eta^<\|_2 \preccurlyeq K^{-5}\|\nu_\eta\|_2 \preccurlyeq K^{-4}\|\mu_\eta\|_2^{1/2}\|\nu_\eta\|_2^{1/2}. \tag{10.16}$$

Hence by the triangle inequality and $\mu_\eta(y) \le \mu_\eta^>(y) + \mu_\eta^<(y) + \mu_\eta^\sim(y)$ we get

$$\|\mu_\eta^\sim * \nu_\eta^\sim\|_2 \ge (K^{-1} - \Omega^{O(1)}(2K^{-4} + 2K^{-9}))\|\mu_\eta\|_2^{1/2}\|\nu_\eta\|_2^{1/2}.$$

For $K \succcurlyeq 1$, the claim follows. □

We will now apply Corollary 10.6 to prove that the energy $E_{16\eta}(\mathcal{C}^\sim(\mu;\eta), \mathcal{C}^\sim(\nu;\eta))$ is *large*. Using this bound and Theorem 10.3, we deduce Theorem 2.12.

Lemma 10.7. *For non-empty sets $A, B \subseteq G$, we have*

$$E_{\eta/16}(A,B) \preccurlyeq \frac{E(A_\eta, B_\eta)}{|1_\eta|^3} \preccurlyeq E_{6\eta}(A,B).$$

Proof. By definition $E_\eta(A,B) := \mathcal{N}_\eta(Q_\eta(A,B))$ with d^+-metric on G^4. Hence by Lemma 7.1 we have

$$E_\eta(A,B) \approx \frac{|(Q_\eta(A,B))_\eta|}{|(1,1,1,1)^+_\eta|},$$

where $+$ indicates that we are using the d^+-metric. Since

$$(1,1,1,1)_{\eta/4} \subseteq (1,1,1,1)^+_\eta \subseteq (1,1,1,1)_\eta,$$

by $|1_{c\eta}| \approx |1_\eta|$ we deduce

$$E_\eta(A,B) \approx \frac{|(Q_\eta(A,B))_\eta|}{|1_\eta|^4}. \tag{10.17}$$

Based on (10.17), we will focus on $|Q_\eta(A,B)_\eta|$ and relate it to energies of thickened sets. First, we will exprees $E(A_\eta, B_\eta)$ as the measure of a subset of G^3:

$$\begin{aligned}
E(A_\eta, B_\eta) =& \|\mathbb{1}_{A_\eta} * \mathbb{1}_{B_\eta}\|_2^2 \\
=& \int_G \int_G \int_G \mathbb{1}_{A_\eta}(x)\mathbb{1}_{B_\eta}(x^{-1}y)\mathbb{1}_{A_\eta}(z)\mathbb{1}_{B_\eta}(z^{-1}y)\ \mathrm{d}x\ \mathrm{d}z\ \mathrm{d}y \\
=& |\{(x,z,y) \in A_\eta \times A_\eta \times G |\ x^{-1}y \in B_\eta, z^{-1}y \in B_\eta\}| \\
=& |\{(x,z,t) \in A_\eta \times A_\eta \times B_\eta |\ z^{-1}xt \in B_\eta\}|. \qquad (10.18)
\end{aligned}$$

Using (10.18), we can find an upper bound for $|Q_\eta(A,B)_\eta|$. We have

$$\begin{aligned}
|Q_\eta(A,B)_\eta| \leq& |\{(x_1,x_2,y_1,y_2) \in A_\eta \times A_\eta \times B_\eta \times B_\eta |\ y_2^{-1}x_2^{-1}y_1x_1 \in 1_{5\eta}\}| \\
=& |\{(x_1,x_2,y_1,h) \in A_\eta \times A_\eta \times B_\eta \times 1_{5\eta} |\ x_2^{-1}x_1y_2h^{-1} \in B_\eta\}| \\
\leq& |\{(x_1,x_2,y_1,h) \in A_\eta \times A_\eta \times B_\eta \times 1_{5\eta} |\ x_2^{-1}x_1y_2 \in B_{6\eta}\}| \\
\preccurlyeq& |1_\eta||\{(x_1,x_2,y_1) \in A_{6\eta} \times A_{6\eta} \times B_{6\eta} |\ x_2^{-1}x_1y_2 \in B_{6\eta}\}| \\
=& |1_\eta| E(A_{6\eta}, B_{6\eta}). \qquad (10.19)
\end{aligned}$$

Again using (10.18), we find a lower bound for $|Q_\eta(A,B)_\eta|$:

$$\begin{aligned}
|Q_\eta(A,B)_\eta| \geq& |\{(x_1,x_2,y_1,y_2) \in A_{\eta/8} \times A_{\eta/8} \times B_{\eta/8} \times B_{\eta/8} |\ y_2^{-1}x_2^{-1}y_1x_1 \in 1_{\eta/2}\}| \\
=& |\{(x_1,x_2,y_1,h) \in A_{\eta/8} \times A_{\eta/8} \times B_{\eta/8} \times 1_{\eta/2} |\ x_2^{-1}y_1x_1h^{-1} \in B_{\eta/8}\}| \\
\geq& |\{(x_1,x_2,y_1,h) \in A_{\eta/8} \times A_{\eta/8} \times B_{\eta/8} \times 1_{\eta/16} |\ x_2^{-1}y_1x_1h^{-1} \in B_{\eta/16}\}| \\
\succcurlyeq& |1_\eta| E(A_{\eta/16}, B_{\eta/16}). \qquad (10.20)
\end{aligned}$$

By (10.17), (10.19), and (10.20), claim follows. □

Lemma 10.8. *In the above setting, $\frac{1}{K^{O(1)}\|\mu_\eta\|_2^2} \leq |\mathcal{C}(\mu,\sim)_\eta| \leq \frac{K^{O(1)}}{\|\mu_\eta\|_2^2}$.*

Proof. For all $y \in \mathcal{C}(\mu,\sim)_\eta$, there exists $x \in \mathcal{C}(\mu,\sim)$ such that $y \in x_\eta$. Hence by part (1) of Lemma 10.4 we have

$$\mu_{3\eta}(y) \succcurlyeq \mu_{2\eta}(x) \succcurlyeq K^{-20}\|\mu_\eta\|_2^2,$$

which implies that

$$\|\mu_{3\eta}\|_2^2 \succcurlyeq K^{-20}\|\mu_\eta\|_2^4 |\mathcal{C}(\mu,\sim)_\eta|.$$

Therefore by part (3) of Lemma 10.4 we deduce that

$$|\mathcal{C}(\mu,\sim)_\eta| \preccurlyeq \frac{K^{20}}{\|\mu_\eta\|_2^2}.$$

It follows from the definition of $\mu_\eta^\sim$ that the support of $\mu_\eta^\sim$ is a subset of $\mathcal{C}(\mu,\sim)_\eta$. Hence if $\mu_\eta^\sim(y)\neq 0$, then there is $x\in\mathcal{C}(\mu,\sim)$ such that $y\in x_\eta$. So, by part (1) of Lemma 10.4 we have

$$\mu_\eta(y) \preccurlyeq \mu_{2\eta}(x) \le K^{10}\|\mu_\eta\|_2^2, \text{ which implies } \|\mu_\eta^\sim\|_\infty \preccurlyeq K^{10}\|\mu_\eta\|_2^2. \tag{10.21}$$

Therefore we get

$$\|\mu_\eta^\sim\|_2^2 \le \|\mu_\eta^\sim\|_\infty^2|\mathcal{C}(\mu,\sim)_\eta| \preccurlyeq K^{20}\|\mu_\eta\|_2^4|\mathcal{C}(\mu,\sim)_\eta|. \tag{10.22}$$

By (10.7), Corollary 10.6, and (10.22), we get

$$\begin{aligned} K^{-2}\|\mu_\eta\|_2^2 \preccurlyeq& \|\mu_\eta\|_2\|\nu_\eta\|_2 \preccurlyeq K^2\|\mu_\eta^\sim * \nu_\eta^\sim\|_2^2 \\ \le& K^2\|\mu_\eta^\sim\|_2^2 \preccurlyeq K^{22}\|\mu_\eta\|_2^4|\mathcal{C}(\mu,\sim)_\eta|; \end{aligned}$$

Therefore

$$\frac{1}{K^{24}\|\mu_\eta\|_2^2} \preccurlyeq |\mathcal{C}(\mu,\sim)_\eta|;$$

and the claim follows. □

Proposition 10.9. *In the above setting the inequality*

$$E_{16\eta}(\mathcal{C}(\mu;\sim),\mathcal{C}(\nu;\sim)) \succcurlyeq \frac{1}{K^{O(1)}}\mathcal{N}_{16\eta}(\mathcal{C}(\mu;\sim))^{3/2}\mathcal{N}_{16\eta}(\mathcal{C}(\nu;\sim))^{3/2}$$

holds, where $\mathcal{C}(\mu;\sim)$ is defined in (10.11).

Proof. By (10.21), we have

$$\mu_\eta^\sim \preccurlyeq (K^{10}\|\mu_\eta\|_2^2)\,\mathbb{1}_{\mathcal{C}(\mu,\sim)_\eta} \text{ and } \nu_\eta^\sim \preccurlyeq (K^{10}\|\nu_\eta\|_2^2)\,\mathbb{1}_{\mathcal{C}(\nu,\sim)_\eta}.$$

It follows that

$$\|\mu_\eta^\sim * \nu_\eta^\sim\|_2^2 \preccurlyeq K^{40}\|\mu_\eta\|_2^4\|\nu_\eta\|_2^4\|\mathbb{1}_{\mathcal{C}(\mu,\sim)_\eta} * \mathbb{1}_{\mathcal{C}(\nu,\sim)_\eta}\|_2^2 = K^{40}\|\mu_\eta\|_2^4\|\nu_\eta\|_2^4 E(\mathcal{C}(\mu,\sim)_\eta,\mathcal{C}(\nu,\sim)_\eta).$$

By Corollary 10.6 and the above inequality we have

$$K^{-2}\|\mu_\eta\|_2\|\nu_\eta\|_2 \preccurlyeq K^{40}\|\mu_\eta\|_2^4\|\nu_\eta\|_2^4 E(\mathcal{C}(\mu,\sim)_\eta,\mathcal{C}(\nu,\sim)_\eta). \tag{10.23}$$

By Lemma 10.8 and (10.23), we obtain

$$K^{-O(1)}|\mathcal{C}(\mu,\sim)_\eta)|^{3/2}|\mathcal{C}(\nu,\sim)_\eta)|^{3/2} \preccurlyeq E(\mathcal{C}(\mu,\sim)_\eta,\mathcal{C}(\nu,\sim)_\eta); \tag{10.24}$$

and so by Lemma 7.1 and Lemma 10.7, we deduce

$$K^{-O(1)}\mathcal{N}_{16\eta}(\mathcal{C}(\mu,\sim))^{3/2}\mathcal{N}_{16\eta}(\mathcal{C}(\nu,\sim))^{3/2} \preccurlyeq E_{16\eta}(\mathcal{C}(\mu,\sim),\mathcal{C}(\nu,\sim));$$

and the claim follows. □

Proof of Theorem 2.12. Recall that μ and ν denote the distribution measures of random variables X and Y, respectively, and Z denotes a random variable independent of X and Y with uniform distribution over $1_{3\eta}$.

By Proposition 10.9, for $K\succcurlyeq 1$, we can apply Theorem 10.3 to the sets $A=\mathcal{C}(\mu;\sim)$ and $B=\mathcal{C}(\nu;\sim)$ to obtain $H\subseteq G$ and $x,y\in G$ such that

(1) (Approximate structure) H is an $K^{O(1)}$-approximate subgroup.
(2) (Controlling the metric entropy) $|h(H;16\eta)-\frac{h(\mathcal{C}(\mu;\sim);16\eta)+h(\mathcal{C}(\nu;\sim);16\eta)}{2}|\le\log K$.
(3) (Large intersection) $|h(\mathcal{C}(\mu;\sim)\cap xH;16\eta)-h(\mathcal{C}(\mu;\sim);16\eta)|\le\log K$ and

$$|h(\mathcal{C}(\nu;\sim)\cap Hy;16\eta)-h(\mathcal{C}(\nu;\sim);16\eta)|\le\log K.$$

We will show that Theorem 2.12 holds for these choices of $H \subseteq G$ and $x, y \in G$.

By Lemma 7.1 we have $|\log \mathcal{N}_{16\eta}(\mathcal{C}(\mu;\sim)) - \log(|\mathcal{C}(\mu;\sim)_\eta|/|1_\eta|)| \preccurlyeq 1$. Hence, Lemma 10.8 implies

$$|\log \mathcal{N}_{16\eta}(\mathcal{C}(\mu;\sim)) - (\log(1/|1_\eta|) - \log \|\mu_\eta\|_2^2)| \ll \log K$$

if $K \succcurlyeq 1$. Thus

$$|\log \mathcal{N}_{16\eta}(\mathcal{C}(\mu;\sim)) - H_2(\mu;\eta)| \ll \log K. \tag{10.25}$$

By (10.25), Lemma 7.1, and part (2) of Theorem 10.3 we have

$$\left|h(H;\eta) - \frac{H_2(\mu;\eta) + H_2(\nu;\eta)}{2}\right| \ll \log K$$

if $K \succcurlyeq 1$. We also notice that by (10.7) we have $|H_2(\mu;\eta) - H_2(\nu;\eta)| \ll \log K$. Combining these two fact we deduce that

$$|h(H;\eta) - H_2(\mu;\eta)| \ll \log K.$$

This proves the second property mentioned in Theorem 2.12 for the set H.

Finally, to prove the third property, note that

$$\mathcal{N}_\eta(\mathcal{C}(\mu;\sim) \cap xH) \succcurlyeq K^{-O(1)} \mathcal{N}_\eta(\mathcal{C}(\mu;\sim));$$

and so by (10.25) we get

$$\mathcal{N}_\eta(\mathcal{C}(\mu;\sim) \cap xH) \succcurlyeq K^{-O(1)} 2^{H_2(\mu;\eta)}. \tag{10.26}$$

On the other hand, by Lemma 7.1, Corollary 7.4, and the fact that $\mathcal{C}(\mu;\sim)$ is an η-separated set, we have

$$\mathcal{N}_\eta(\mathcal{C}(\mu;\sim) \cap xH) \approx \mathcal{N}_{\eta/2}(\mathcal{C}(\mu;\sim) \cap xH) = \#(\mathcal{C}(\mu;\sim) \cap xH).$$

Altogether we have

$$\#(\mathcal{C}(\mu;\sim) \cap xH) \succcurlyeq K^{-O(1)} 2^{H_2(\mu;\eta)}. \tag{10.27}$$

For every $z' \in \mathcal{C}(\mu;\sim)_\eta$ there exist $z \in \mathcal{C}(\mu;\sim)$ such that $z' \in z_\eta$. Since $\mu_{3\eta}(z') = \mu(z'_{3\eta})/|1_{3\eta}|$ and $\mu_{2\eta}(z) \geq K^{-10}\|\mu_\eta\|_2^2$, by part (1) of Lemma 10.4 we have

$$\mu_{3\eta}(z') \succcurlyeq \mu_{2\eta}(z) \geq K^{-10}\|\mu_\eta\|_2^2, \quad \text{and} \quad \mu(z'_{3\eta}) \geq \widehat{C} K^{-10} 2^{-H_2(\mu;\eta)}. \tag{10.28}$$

where $\widehat{C} = \Omega^{O(1)}$. Therefore

$$\begin{aligned}
\mathbb{P}(XZ \in (xH)_\eta) &\geq \int_{(\mathcal{C}(\mu;\sim)\cap xH)_\eta} \mu_{3\eta}(z')\, \mathrm{d}z' \\
&\succcurlyeq K^{-10}\|\mu_\eta\|_2^2 |(\mathcal{C}(\mu;\sim) \cap xH)_\eta| \\
&\approx K^{-10} 2^{-H_2(\mu;\eta)} \mathcal{N}_\eta(\mathcal{C}(\mu;\sim) \cap xH) \succcurlyeq K^{-O(1)}.
\end{aligned}$$

The lower bound for $\mathbb{P}(ZY \in (Hy)_\eta)$ can be proved by a similar argument. Finally, to prove the last claim, we have

$$\begin{aligned}
|\{h \in H_\eta|\, \mathbb{P}(X \in (xh)_{3\eta}) \geq \widehat{C} K^{-10} 2^{-H_2(X;\eta)}\}| &= |\{z' \in (xH)_\eta | \mu(z'_{3\eta}) \geq \widehat{C} K^{-10} 2^{-H_2(X;\eta)}\}| \\
&\geq |\{z' \in (\mathcal{C}(\mu;\sim) \cap xH)_\eta | \mu(z'_{3\eta}) \geq \widehat{C} K^{-10} 2^{-H_2(X;\eta)}\}| \\
&= |(\mathcal{C}(\mu;\sim) \cap xH)_\eta| \\
&\succcurlyeq K^{-O(1)} 2^{H_2(\mu;\eta)} \cdot |1_\eta| = K^{-O(1)} |H_\eta|.
\end{aligned}$$

This proves the claim. □

References

[1] László Babai, Nikolay Nikolov, and László Pyber. Product growth and mixing in finite groups. In *Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 248–257. ACM, New York, 2008.

[2] Antal Balog and Endre Szemerédi. A statistical theorem of set addition. *Combinatorica*, 14(3):263–268, 1994.

[3] Hyman Bass, Alexander Lubotzky, Andy R. Magid, and Shahar Mozes. The proalgebraic completion of rigid groups. In *Proceedings of the Conference on Geometric and Combinatorial Group Theory, Part II (Haifa, 2000)*, volume 95, pages 19–58, 2002.

[4] Yves Benoist and Nicolas de Saxcé. Convolution in perfect Lie groups. *Math. Proc. Cambridge Philos. Soc.*, 161(1):31–45, 2016.

[5] Yves Benoist and Nicolas de Saxcé. A spectral gap theorem in simple Lie groups. *Invent. Math.*, 205(2):337–361, 2016.

[6] J. Bourgain and A. Gamburd. A spectral gap theorem in $\mathrm{SU}(d)$. *J. Eur. Math. Soc. (JEMS)*, 14(5):1455–1511, 2012.

[7] Jean Bourgain and Alex Gamburd. On the spectral gap for finitely-generated subgroups of $\mathrm{SU}(2)$. *Invent. Math.*, 171(1):83–121, 2008.

[8] 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.

[9] Jean Bourgain and Alex Gamburd. Expansion and random walks in $\mathrm{SL}_d(\mathbb{Z}/p^n\mathbb{Z})$. II. *J. Eur. Math. Soc. (JEMS)*, 11(5):1057–1103, 2009. With an appendix by Bourgain.

[10] Rémi Boutonnet, Adrian Ioana, and Alireza Salehi Golsefidy. Local spectral gap in simple Lie groups and applications. *Invent. Math.*, 208(3):715–802, 2017.

[11] F. R. K. Chung, R. L. Graham, and R. M. Wilson. Quasi-random graphs. *Combinatorica*, 9(4):345–362, 1989.

[12] Nicolas de Saxcé. Trou dimensionnel dans les groupes de Lie compacts semisimples via les séries de Fourier. *J. Anal. Math.*, 120:311–331, 2013.

[13] Alex Gamburd. On the spectral gap for infinite index "congruence" subgroups of $\mathrm{SL}_2(\mathbf{Z})$. *Israel J. Math.*, 127:157–200, 2002.

[14] A. Salehi Golsefidy and Péter P. Varjú. Expansion in perfect groups. *Geom. Funct. Anal.*, 22(6):1832–1891, 2012.

[15] W. T. Gowers. A new proof of Szemerédi's theorem. *Geom. Funct. Anal.*, 11(3):465–588, 2001.

[16] W. T. Gowers. Quasirandom groups. *Combin. Probab. Comput.*, 17(3):363–387, 2008.

[17] Elon Lindenstrauss and Nicolas de Saxcé. Hausdorff dimension and subgroups of $SU(2)$. *Israel J. Math.*, 209(1):335–354, 2015.

[18] V. P. Platonov. The theory of algebraic linear groups and periodic groups. *Izv. Akad. Nauk SSSR Ser. Mat.*, 30:573–620, 1966.

[19] 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.

[20] Peter Sarnak and Xiao Xi Xue. Bounds for multiplicities of automorphic representations. *Duke Math. J.*, 64(1):207–227, 1991.

[21] Elias M. Stein. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, volume 43 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.

[22] Terence Tao. Product set estimates for non-commutative groups. *Combinatorica*, 28(5):547–594, 2008.

[23] 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.

[24] Péter Pál Varjú. Random walks in compact groups. *Doc. Math.*, 18:1137–1175, 2013.

Jacobs University, Campus Ring I, 28759, Bremen, Germany
Email address: k.mallahikarai@jacobs-university.de

Mathematics Department, University of California, San Diego, CA 92093-0112, USA
Email address: ammohammadi@ucsd.edu

Mathematics Department, University of California, San Diego, CA 92093-0112, USA
Email address: golsefidy@ucsd.edu