A MEASURE ESTIMATE IN GEOMETRY OF NUMBERS AND IMPROVEMENTS TO DIRICHLET'S THEOREM

DMITRY KLEINBOCK, ANDREAS STRÖMBERGSSON, AND SHUCHENG YU

ABSTRACT. Let ψ be a continuous decreasing function defined on all large positive real numbers. We say that a real $m \times n$ matrix A is ψ-Dirichlet if for every sufficiently large real number t one can find $\boldsymbol{p} \in \mathbb{Z}^m$, $\boldsymbol{q} \in \mathbb{Z}^n \smallsetminus \{\mathbf{0}\}$ satisfying $\|A\boldsymbol{q} - \boldsymbol{p}\|^m < \psi(t)$ and $\|\boldsymbol{q}\|^n < t$. This property was introduced by Kleinbock and Wadleigh in 2018, generalizing the property of A being Dirichlet improvable which dates back to Davenport and Schmidt (1969). In the present paper, we give sufficient conditions on ψ to ensure that the set of ψ-Dirichlet matrices has zero or full Lebesgue measure. Our proof is dynamical and relies on the effective equidistribution and doubly mixing of certain expanding horospheres in the space of lattices. Another main ingredient of our proof is an asymptotic measure estimate for certain compact neighborhoods of the critical locus (with respect to the supremum norm) in the space of lattices. Our method also works for the analogous weighted problem where the relevant supremum norms are replaced by certain weighted quasi-norms.

CONTENTS

Date: August 13, 2021.

D.K. was supported by NSF grant DMS-1900560. A.S. and S.Y. were supported by the Knut and Alice Wallenberg Foundation.

1. Introduction

1.1. **Background.** Let m, n be two positive integers and let $M_{m,n}(\mathbb{R})$ be the space of m by n real matrices. The starting point of our work is the following higher dimensional generalization of the classical Dirichlet's Diophantine approximation theorem, see e.g. [2, §1.5].

Theorem 1.1. *For any $A \in M_{m,n}(\mathbb{R})$ and $t > 1$, there exists $(\boldsymbol{p}, \boldsymbol{q}) \in \mathbb{Z}^m \times (\mathbb{Z}^n \smallsetminus \{\mathbf{0}\})$ satisfying the following system of inequalities:*

$$\|A\boldsymbol{q} - \boldsymbol{p}\|^m \le \frac{1}{t} \quad \text{and} \quad \|\boldsymbol{q}\|^n < t. \tag{1.1}$$

Here $\|\cdot\|$ denotes the supremum norm on $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively.

A natural question to ask is whether one can improve (1.1) by replacing $1/t$ by a smaller function, that is, consider the following system of inequalities:

$$\|A\boldsymbol{q} - \boldsymbol{p}\|^m < \psi(t) \quad \text{and} \quad \|\boldsymbol{q}\|^n < t \tag{1.2}$$

where ψ is a positive, continuous, decreasing function which decays to zero at infinity. Historically there have been two directions to pursue in this regard: looking for solvability of (1.2) for an unbounded set of $t > 0$ vs. for all large enough t. The former is sometimes referred to as *asymptotic approximation*, and has culminated in definitive results such as the Khintchine-Groshev theorem. In this paper we are interested in the latter, less studied set-up of *uniform approximation*. Following the definition in Kleinbock and Wadleigh [25], we say that an m by n real matrix A is *ψ-Dirichlet* if the system of inequalities (1.2) has solutions in $(\boldsymbol{p}, \boldsymbol{q}) \in \mathbb{Z}^m \times (\mathbb{Z}^n \smallsetminus \{\mathbf{0}\})$ for all sufficiently large t. It is clear that $A \in M_{m,n}(\mathbb{R})$ is ψ-Dirichlet if and only if $A + A'$ is ψ-Dirichlet for any $A' \in M_{m,n}(\mathbb{Z})$. Thus with slight abuse of notation, we denote by $\mathbf{DI}_{m,n}(\psi) \subset M_{m,n}(\mathbb{R}/\mathbb{Z})$ the set of ψ-Dirichlet matrices.

Let $\psi_1(t) = 1/t$. The problem of improving Dirichlet's theorem was initiated by Davenport and Schmidt [10, 9] where they showed that the set

$$\mathbf{DI}_{m,n} := \bigcup_{0<c<1} \mathbf{DI}_{m,n}(c\psi_1) \tag{1.3}$$

of *Dirichlet improvable matrices* is of Lebesgue measure zero, while having full Hausdorff dimension mn. More recently, Kleinbock and Mirzadeh [22, Theorem 1.5] showed that for any fixed $0 < c < 1$, the Hausdorff dimension of $\mathbf{DI}_{m,n}(c\psi_1)$ is strictly smaller than mn. There have also been extensive studies on the Hausdorff dimensions of the (even smaller) set of the *singular matrices*,

$$\mathbf{Sing}_{m,n} := \bigcap_{0<c<1} \mathbf{DI}_{m,n}(c\psi_1).$$

After a series of breakthrough work, it is now known that the Hausdorff dimension of $\mathbf{Sing}_{m,n}$ is $mn - \frac{mn}{m+n}$ whenever $\max\{m, n\} > 1$; see [3, 4, 13, 7, 8].

On the other hand, for a general decreasing function ψ with $t \mapsto t\psi(t)$ increasing, Kleinbock and Wadleigh proved a zero-one law for the Lebesgue measure of $\mathbf{DI}_{1,1}(\psi)$ depending on the divergence or convergence of a certain series involving ψ [25, Theorem 1.8]. See also [12] for the relevant dimension theory of $\mathbf{DI}_{1,1}(\psi)$, [23] for a similar zero-one law with the supremum norm replaced by the Euclidean norm and [26, 15] for analogous results in the *inhomogeneous* setting.

The arguments in [25] rely on the theory of continued fractions and are not applicable for higher dimensions. Nevertheless, for general dimensions, building on ideas from [6, 19], a dynamical approach was proposed in [25, §4], reformulating the problem as a *shrinking target problem*, which asks whether a generic orbit in a dynamical system hits a given sequence of shrinking targets

infinitely often. To describe this dynamical interpretation, let us first fix some notation. Let $d = m + n$ and let $X_d := \mathrm{SL}_d(\mathbb{R})/\mathrm{SL}_d(\mathbb{Z})$ be the homogeneous space which parameterizes the space of unimodular lattices in $\mathbb{R}^d$ via $g\mathrm{SL}_d(\mathbb{Z}) \leftrightarrow g\mathbb{Z}^d$. We note that $\mathrm{SL}_d(\mathbb{R})$ acts on X_d naturally via the regular action: $g\Lambda = gh\mathbb{Z}^d$ for any $g \in \mathrm{SL}_d(\mathbb{R})$ and $\Lambda = h\mathbb{Z}^d \in X_d$. For any $s \in \mathbb{R}$, let a_s be the diagonal matrix

$$a_s := \begin{pmatrix} e^{s/m} I_m & 0 \\ 0 & e^{-s/n} I_n \end{pmatrix} \in \mathrm{SL}_d(\mathbb{R}).$$

Let $\Delta : X_d \to [0, \infty)$ be the function defined by

$$\Delta(\Lambda) := \sup_{\boldsymbol{v} \in \Lambda \smallsetminus \{\mathbf{0}\}} \log\Big(\frac{1}{\|\boldsymbol{v}\|}\Big). \tag{1.4}$$

Finally, let us denote

$$\mathcal{Y} := \left\{ \Lambda_A := \begin{pmatrix} I_m & A \\ 0 & I_n \end{pmatrix} \mathbb{Z}^d \in X_d \ : \ A \in M_{m,n}(\mathbb{R}) \right\}. \tag{1.5}$$

The submanifold $\mathcal{Y} \subset X_d$ can be naturally identified with the mn-dimensional torus $M_{m,n}(\mathbb{R}/\mathbb{Z})$ via $\Lambda_A \leftrightarrow A \in M_{m,n}(\mathbb{R}/\mathbb{Z})$. Throughout the paper, we denote by Leb the probability Lebesgue measure on $\mathcal{Y} \cong M_{m,n}(\mathbb{R}/\mathbb{Z})$; for simplicity of notation, for any function f on $\mathcal{Y}$ we will abbreviate the space average $\int_{\mathcal{Y}} f(\Lambda_A)\, \mathrm{d}\,\mathrm{Leb}(A)$ by either $\mathrm{Leb}(f)$ or $\int_{\mathcal{Y}} f(\Lambda_A)\ \mathrm{d}A$.

It was shown in [25, Proposition 4.5] that for any given ψ as above, there exists a unique continuous function $r = r_\psi : [s_0, \infty) \to (0, \infty)$ such that

$$A \in M_{m,n}(\mathbb{R}) \text{ is } \textit{not } \psi\text{-Dirichlet} \Leftrightarrow a_s\Lambda_A \in \Delta^{-1}[0, r(s)] \text{ for an unbounded set of } s > s_0. \tag{1.6}$$

This equivalence is usually called the *Dani Correspondence*. In view of this interpretation, our task is to find conditions which ensure that for almost every (or almost no) $A \in M_{m,n}(\mathbb{R})$, the orbit $\{a_s\Lambda_A\}_{s>s_0}$ hits the shrinking target $\Delta^{-1}[0, r(s)]$ for an unbounded set of s-values. We note that this was also the strategy used in [19] giving a dynamical proof of the classical Khintchine-Groshev Theorem, where the relevant shrinking targets are certain cusp neighborhoods in X_d. For our case, by Mahler's compactness criterion, the shrinking targets $\Delta^{-1}[0, r(s)]$ are compact neighborhoods of the critical locus $\Delta^{-1}\{0\}$, whose explicit description is given by Hajós's Theorem [11] (cf. Theorem 2.3 below). The fact that these shrinking targets are compact sets causes new difficulties compared to the situation in [19], see the discussion in Section 1.3.

1.2. **Main results.** In the present paper, based on the dynamical interpretation described above, we give sufficient conditions on ψ implying that $\mathbf{DI}_{m,n}(\psi)$ is of zero or full Lebesgue measure. In fact, with some modifications, our arguments also work for the analogous *weighted* problem where the supremum norms in (1.2) are replaced by certain *weighted quasi-norms*, as introduced in [16]. We thus prove our main result in that generality. We first introduce the relevant notation.

Let $\boldsymbol{\alpha} \in \mathbb{R}^m$ and $\boldsymbol{\beta} \in \mathbb{R}^n$ be two *weight vectors*, that is

$$\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_m) \in (\mathbb{R}_{>0})^m \quad \text{and} \quad \boldsymbol{\beta} = (\beta_1, \dots, \beta_n) \in (\mathbb{R}_{>0})^n$$

with $\sum_i \alpha_i = \sum_j \beta_j = 1$. We say that $A \in M_{m,n}(\mathbb{R})$ is $\psi_{\boldsymbol{\alpha},\boldsymbol{\beta}}$*-Dirichlet* if the system of inequalities

$$\|A\boldsymbol{q} - \boldsymbol{p}\|_{\boldsymbol{\alpha}} < \psi(t) \quad \text{and} \quad \|\boldsymbol{q}\|_{\boldsymbol{\beta}} < t \tag{1.7}$$

has solutions in $(\boldsymbol{p}, \boldsymbol{q}) \in \mathbb{Z}^m \times (\mathbb{Z}^n \smallsetminus \{\mathbf{0}\})$ for all sufficiently large t. Here

$$\|\boldsymbol{x}\|_{\boldsymbol{\alpha}} := \max\left\{ |x_i|^{1/\alpha_i} \ : \ 1 \le i \le m \right\} \quad \text{and} \quad \|\boldsymbol{y}\|_{\boldsymbol{\beta}} := \max\left\{ |y_j|^{1/\beta_j} \ : \ 1 \le j \le n \right\}$$

are the two quasi-norms associated with $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ respectively. Again it is easy to see that $A \in M_{m,n}(\mathbb{R})$ is $\psi_{\boldsymbol{\alpha},\boldsymbol{\beta}}$-Dirichlet if and only if $A + A'$ is $\psi_{\boldsymbol{\alpha},\boldsymbol{\beta}}$-Dirichlet for any $A' \in M_{m,n}(\mathbb{Z})$, and we denote by $\mathbf{DI}_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\psi) \subset M_{m,n}(\mathbb{R}/\mathbb{Z})$ the set of $\psi_{\boldsymbol{\alpha},\boldsymbol{\beta}}$-Dirichlet matrices. We note that when $\boldsymbol{\alpha} = (\frac{1}{m}, \dots, \frac{1}{m})$ and $\boldsymbol{\beta} = (\frac{1}{n}, \dots, \frac{1}{n})$, then $\mathbf{DI}_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\psi) = \mathbf{DI}_{m,n}(\psi)$.

We now state our main result which gives sufficient conditions on ψ determining when $\mathbf{DI}_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\psi)$ is of full or zero Lebesgue measure.

Theorem 1.2. *Fix $m, n \in \mathbb{N}$ and two weight vectors $\boldsymbol{\alpha} \in \mathbb{R}^m$ and $\boldsymbol{\beta} \in \mathbb{R}^n$. Let $d = m + n$, and let*

$$\varkappa_d = \frac{d^2 + d - 4}{2} \quad \text{and} \quad \lambda_d = \frac{d(d-1)}{2}.$$

Let $t_0 > 0$ and let $\psi : [t_0, \infty) \to (0, \infty)$ be a continuous, decreasing function such that

$$\text{the function } t \mapsto t\psi(t) \text{ is increasing} \tag{1.8}$$

and

$$\psi(t) < \psi_1(t) = 1/t \quad \text{for all } t \geq t_0. \tag{1.9}$$

Let $F_\psi(t) := 1 - t\psi(t)$. If the series

$$\sum_{k \geq t_0} k^{-1} F_\psi(k)^{\varkappa_d} \log^{\lambda_d}\left(\frac{1}{F_\psi(k)}\right) \tag{1.10}$$

converges, then $\mathbf{DI}_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\psi)$ is of full Lebesgue measure. Conversely, if the series (1.10) diverges, and

$$\liminf_{t_1 \to \infty} \frac{\sum_{t_0 \leq k \leq t_1} k^{-1} F_\psi(k)^{\varkappa_d} \log^{\lambda_d+1}\left(\frac{1}{F_\psi(k)}\right)}{\left(\sum_{t_0 \leq k \leq t_1} k^{-1} F_\psi(k)^{\varkappa_d} \log^{\lambda_d}\left(\frac{1}{F_\psi(k)}\right)\right)^2} = 0, \tag{1.11}$$

then $\mathbf{DI}_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\psi)$ is of zero Lebesgue measure.

Remark 1. When $m = n = 1$, Theorem 1.2 is not new; in fact [25, Theorem 1.8] is stronger in the sense that it gives a tight zero-one law without the extra assumption (1.11). We believe that an analogous tight zero-one law should also hold for general dimensions m, n, i.e. that Theorem 1.2 should hold with the assumption (1.11) removed. See Remark 8 for a discussion of why assumption (1.11) is needed in our proof.

Remark 2. The function $F_\psi(t) = 1 - t\psi(t)$ encodes ψ via $\psi(t) = \frac{1 - F_\psi(t)}{t}$. In view of the assumptions (1.8) and (1.9), F_ψ is a decreasing function and takes values in $(0, 1)$. In particular, the limit $\lim_{t\to\infty} F_\psi(t)$ exists and lies in $[0, 1)$.

If $\lim_{t\to\infty} F_\psi(t) > 0$, then the conclusion of Theorem 1.2 follows from the work of Kleinbock and Weiss [27]. Indeed, in this case the series (1.10) diverges, and for any fixed β we have $\sum_{t_0 \leq k \leq t_1} k^{-1} F_\psi(k)^{\varkappa_d} \log^{\beta}\left(\frac{1}{F_\psi(k)}\right) \asymp \log t_1$ as $t_1 \to \infty$, so that also the assumption (1.11) holds; moreover, $\lim_{t\to\infty} F_\psi(t) > 0$ implies that there exists some $c \in (0, 1)$ such that $\psi(t) < c/t$, implying that $\mathbf{DI}_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\psi) \subset \mathbf{DI}_{\boldsymbol{\alpha},\boldsymbol{\beta}}(c\psi_1)$; and by [27, Theorem 1.4], $\mathbf{DI}_{\boldsymbol{\alpha},\boldsymbol{\beta}}(c\psi_1)$ is a null set.

Remark 3. Let us give some explicit examples to illustrate our results. We note that each function ψ appearing below is strictly decreasing on $[t_0, \infty)$ for t_0 sufficiently large.

(1) Let $\psi(t) = \frac{1 - c(\log t)^{-\tau}}{t}$ ($\Leftrightarrow F_\psi(t) = c(\log t)^{-\tau}$) for some $c > 0$ and $\tau \geq 0$. In this case the series (1.10) diverges if and only if $\tau \leq \frac{1}{\varkappa_d}$. It is also easy to check that condition (1.11) is satisfied whenever $\tau \leq \frac{1}{\varkappa_d}$. Hence Theorem 1.2 implies that for such ψ, $\mathbf{DI}_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\psi)$ is of full measure if $\tau > \frac{1}{\varkappa_d}$, and of zero measure if $\tau \leq \frac{1}{\varkappa_d}$.

(2) Let $\psi(t) = \frac{1-c(\log t)^{-1/\varkappa_d}(\log\log t)^{-\tau}}{t}$ ($\Leftrightarrow F_\psi(t) = c(\log t)^{-1/\varkappa_d}(\log\log t)^{-\tau}$) for some $c > 0$ and $\tau \in \mathbb{R}$. In this case the series (1.10) diverges if and only if $\tau \le \frac{\lambda_d+1}{\varkappa_d}$, while the condition (1.11) is satisfied if and only if $\tau < \frac{\lambda_d}{\varkappa_d}$. Hence Theorem 1.2 implies that for such ψ, $\mathbf{DI}_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\psi)$ is of full measure if $\tau > \frac{\lambda_d+1}{\varkappa_d}$, and of zero measure if $\tau < \frac{\lambda_d}{\varkappa_d}$. However, for τ in the range $\frac{\lambda_d}{\varkappa_d} \le \tau \le \frac{\lambda_d+1}{\varkappa_d}$, Theorem 1.2 gives *no information* (although we believe that $\mathbf{DI}_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\psi)$ is of zero measure also for these τ; cf. Remark 1). We point out that in the special case $\tau = \frac{\lambda_d}{\varkappa_d}$, the quotient in (1.11) remains *bounded* as $t_1 \to \infty$; in this case our method of proof allows us to conclude that at least the set $\mathbf{DI}_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\psi)$ is not of full Lebesgue measure; see Remark 16.

Remark 4. Let us also point out that the assumption (1.9) is imposed only to avoid making the statement of Theorem 1.2 unnecessarily complicated (since otherwise $F_\psi(t)$ could be negative and then the series (1.10) is not well-defined). Indeed, if (1.9) fails but ψ satisfies the other assumptions in Theorem 1.2, then $\mathbf{DI}_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\psi)$ is certainly of full Lebesgue measure. This is true since in this case after possibly enlarging t_0, we have $\psi(t) \ge 1/t \ge \frac{1-(\log t)^{-\tau}}{t}$ for all $t \ge t_0$ and $\tau \ge 0$, and therefore $\mathbf{DI}_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\psi) \supset \mathbf{DI}_{\boldsymbol{\alpha},\boldsymbol{\beta}}\left(t \mapsto \frac{1-(\log t)^{-\tau}}{t}\right)$, where the last set is of full Lebesgue measure whenever $\tau > \frac{1}{\varkappa_d}$ by Remark 3(1).

One of the main ingredients in our proof of Theorem 1.2 is a measure estimate in geometry of numbers, which we believe is of independent interest. Let μ_d be the unique left $\mathrm{SL}_d(\mathbb{R})$-invariant probability measure on $X_d = \mathrm{SL}_d(\mathbb{R})/\mathrm{SL}_d(\mathbb{Z})$. We are interested in the sets $\Delta^{-1}[0,r]$ in X_d, as $r \to 0^+$. As we have discussed, these sets shrink toward the critical locus $\Delta^{-1}\{0\}$ as $r \to 0^+$, and by Hajós's Theorem [11] (cf. Theorem 2.3 below), the set $\Delta^{-1}\{0\}$ has a simple explicit description as a finite union of compact submanifolds of positive codimension $\frac{d^2+d}{2} - 1 = \varkappa_d + 1$ in X_d. In particular this implies that $\mu_d\big(\Delta^{-1}\{0\}\big) = 0$ and $\mu_d\left(\Delta^{-1}[0,r]\right) \to 0$ as $r \to 0^+$. The following theorem gives an asymptotic estimate on the exact rate of convergence in the limit just mentioned.

Theorem 1.3. *We have*

$$\mu_d\left(\Delta^{-1}[0,r]\right) \asymp_d r^{\frac{(d-1)(d+2)}{2}} \log^{\frac{d(d-1)}{2}}\left(\frac{1}{r}\right) = r^{\varkappa_d+1}\log^{\lambda_d}\left(\frac{1}{r}\right), \qquad \text{as } r \to 0^+. \tag{1.12}$$

Our proof of Theorem 1.3 proceeds by bounding the sets $\Delta^{-1}[0,r]$ from above and below by more explicit sets whose Haar measure we can estimate directly. In the proof of the upper bound we make crucial use of Hajós's Theorem. We remark that Hajós's proof (from 1941) of the theorem, which settled a conjecture of Minkowski from 1896, is surprisingly complicated, with the first step being a translation of the question into an algebraic statement about factorizations of finite abelian groups (see also [35] for a nice presentation). It seems difficult to extend this proof in any direct way from the case of $\Delta^{-1}\{0\}$ to deduce restrictions on the sets $\Delta^{-1}[0,r]$ which are sufficiently strong to imply the desired upper bound on $\mu_d\left(\Delta^{-1}[0,r]\right)$. Instead we apply Hajós's Theorem, in combination with a compactness argument, to obtain a convenient containment relation for $\Delta^{-1}[0,r]$ valid for *all* sufficiently small r (see Lemma 4.5 and Remark 11). This initial restriction serves as the starting point for our analysis where we use direct, geometric arguments to derive further, r-dependent restrictions on $\Delta^{-1}[0,r]$ for r small, strong enough to finally imply the desired upper bound on $\mu_d\left(\Delta^{-1}[0,r]\right)$.

Remark 5. In the case $d = 2$, the following explicit formula holds [36, p. 74]:

$$\mu_2\left(\Delta^{-1}[0,r]\right) = \begin{cases} \frac{12}{\pi^2}\left(e^{-2r} + 2r - 1 + \Psi(e^{-2r})\right) & \text{if } 0 \le r \le \frac{1}{2}\log 2, \\ 1 - \frac{12}{\pi^2}e^{-2r} & \text{if } r \ge \frac{1}{2}\log 2, \end{cases}$$

where the function $\Psi : (0,1] \to \mathbb{R}$ is defined by $\Psi(1) = 0$ and $\Psi'(x) = (x^{-1} - 1)\log(x^{-1} - 1)$. It follows that in this case we have an explicit asymptotic expansion sharpening (1.12):

$$\mu_2\left(\Delta^{-1}[0,r]\right) \sim \frac{24}{\pi^2} r^2 \log\Big(\frac{1}{r}\Big) + \frac{12(3-2\log 2)}{\pi^2} r^2 - \frac{16}{\pi^2} r^3 \log\Big(\frac{1}{r}\Big) + \cdots \qquad \text{as } r \to 0^+.$$

The explicit formula for $\mu_2\left(\Delta^{-1}[0,r]\right)$, stated in a different notation, was independently obtained in [28] using a different method.

Remark 6. Theorem 1.3 is also relevant for the study of the Hausdorff dimension of the set $\mathbf{DI}_{m,n}(c\psi_1)$. As we have mentioned, Kleinbock and Mirzadeh recently proved that the Hausdorff dimension of $\mathbf{DI}_{m,n}(c\psi_1)$ is less than mn for every $0 < c < 1$ [22, Theorem 1.5]. They derived this as an application of their main result, [22, Theorem 1.2], which gives an explicit upper bound on the Hausdorff dimension of a certain kind of dynamically defined subsets in the space X_d. It seems that by using Theorem 1.3 (cf. also Theorem 5.1 below), together with a further analysis of the quantities appearing in [22, Theorem 1.2], it should be possible to sharpen the conclusion of [22, Theorem 1.5] into a bound of the form

$$\dim_H\big(\mathbf{DI}_{m,n}(c\psi_1)\big) < mn - \delta(1-c)^{\varkappa_d} \log^{\lambda_d - 1}\big((1-c)^{-1}\big)$$

for all $c < 1$ sufficiently near 1, where $d = m + n$ and $\delta > 0$ is a constant which only depends on m, n.

1.3. **Discussion of the proof of Theorem 1.2.** We next give a more detailed outline of our proof of Theorem 1.2. For simplicity of presentation, we will only focus on the special case when $\boldsymbol{\alpha} = (\frac{1}{m}, \dots, \frac{1}{m})$ and $\boldsymbol{\beta} = (\frac{1}{n}, \dots, \frac{1}{n})$; we comment in Remark 9 below on the modifications needed to treat general weights.

We start from the Dani Correspondence, (1.6), and discretize the shrinking target problem which appears there by introducing the following thickened targets:

$$B_k := \bigcup_{0 \le s < 1} a_{-s}\Delta^{-1}[0, r(k+s)], \qquad \text{for any integer } k > s_0.$$

It follows from this definition that for any $\Lambda \in X_d$, $a_k\Lambda \in B_k$ if and only if $a_s\Lambda \in \Delta^{-1}[0, r(s)]$ for some $k \le s < k+1$. In particular, by (1.6), $A \in M_{m,n}(\mathbb{R})$ is not ψ-Dirichlet if and only if $a_k\Lambda_A \in B_k$ for infinitely many integers k. For any $k > s_0$ let us define

$$E_k := \{\Lambda_A \in \mathcal{Y} \,:\, a_k\Lambda_A \in B_k\}.$$

Then, in view of the previous discussion and the identification $\mathcal{Y} \cong M_{m,n}(\mathbb{R}/\mathbb{Z})$, we have

$$\mathbf{DI}^c_{m,n}(\psi) = \limsup_{k\to\infty} E_k.$$

Hence, by the Borel-Cantelli lemma,

$$\begin{cases} \sum_k \mathrm{Leb}(E_k) < \infty & \Longrightarrow \mathrm{Leb}\left(\mathbf{DI}^c_{m,n}(\psi)\right) = 0; \\ \sum_k \mathrm{Leb}(E_k) = \infty \ \&\ \text{“quasi-independence conditions”} & \Longrightarrow \mathrm{Leb}\left(\mathbf{DI}^c_{m,n}(\psi)\right) = 1. \end{cases}$$

We thus need to understand when the sum $\sum_k \mathrm{Leb}(E_k)$ diverges or converges, respectively. It follows from our definitions that

$$\mathrm{Leb}(E_k) = \int_{\mathcal{Y}} \chi_{B_k}(a_k\Lambda_A)\, \mathrm{d}A.$$

Using Margulis's thickening argument [30], it is well known that the a_s-translates $a_s\mathcal{Y}$ equidistribute in X_d as $s \to \infty$, that is,

$$\int_{\mathcal{Y}} \chi_B(a_s \Lambda_A)\ \mathrm{d}A \to \mu_d(B), \qquad \text{as } s \to \infty \tag{1.13}$$

for any subset B in X_d with boundary of measure zero. However, since our shrinking target B_k varies in the parameter k, we need an *effective* version of (1.13). Such a result was first proved by Kleinbock and Margulis [18, Proposition 2.4.8]; we use the following explicit version (see Corollary 6.4 below): there exists $\delta > 0$ such that for any $f \in C_c^\infty(X_d)$ and any $s > 0$,

$$\int_{\mathcal{Y}} f(a_s \Lambda_A)\ \mathrm{d}A = \mu_d(f) + O\big(e^{-\delta s}\mathcal{N}(f)\big), \tag{1.14}$$

where the norm $\mathcal{N}(\cdot)$ is the maximum of a Lipschitz norm and a Sobolev L^2-norm (see Section 6.2). By approximating $\{\chi_{B_k}\}_{k>s_0}$ from above and below by smooth functions and applying (1.14) together with an explicit bound on the norm $\mathcal{N}(\cdot)$ (see Lemma 6.5), it follows that (see Lemma 7.1)

$$\sum_k \mathrm{Leb}(E_k) = \infty \qquad \Longleftrightarrow \qquad \sum_k \mu_d\left(B_k\right) = \infty.$$

Furthermore, it is not difficult to see from Theorem 1.3 that the series $\sum_k \mu_d\left(B_k\right)$ diverges if and only if the series in (1.10) diverges (see Theorem 5.1 and Lemma 6.1). This in particular settles the convergence case of Theorem 1.2.

For the divergence case, in addition to the assumption that the series in (1.10) diverges (which implies that $\sum_k \mathrm{Leb}(E_k) = \infty$), one also needs to establish a certain *quasi-independence* condition, see (7.8). Roughly speaking, we need to show that the quantities

$$\left|\mathrm{Leb}\left(E_i \cap E_j\right) - \mathrm{Leb}(E_i)\mathrm{Leb}(E_j)\right|, \qquad i \neq j > s_0$$

are small on average. Here note that

$$\mathrm{Leb}\left(E_i \cap E_j\right) = \int_{\mathcal{Y}} \chi_{B_i}(a_i \Lambda_A)\chi_{B_j}(a_j \Lambda_A)\ \mathrm{d}A.$$

We now apply the *effective doubly mixing* for the a_s-translates $\{a_s\mathcal{Y}\}_{s>0}$. This result is due to Kleinbock-Shi-Weiss [24, Theorem 1.2]; we use a more explicit version due to Björklund-Gorodnik [1, Corollary 2.4] which states that for any $f_1, f_2 \in C_c^\infty(X_d)$ and any $s_1, s_2 > 0$,

$$\int_{\mathcal{Y}} f_1(a_{s_1} \Lambda_A) f_2(a_{s_2} \Lambda_A)\ \mathrm{d}A = \mu_d(f_1)\mu_d(f_2) + O\left(e^{-\delta D(s_1,s_2)}\mathcal{N}(f_1)\mathcal{N}(f_2)\right), \tag{1.15}$$

where $D(s_1, s_2) := \min\{s_1, s_2, |s_1 - s_2|\}$. Combining this result with (1.14) we get

$$\left|\int_{\mathcal{Y}} f_1(a_{s_1} \Lambda_A) f_2(a_{s_2} \Lambda_A)\ \mathrm{d}A - \mathrm{Leb}(f_1)\mathrm{Leb}(f_2)\right| \ll e^{-\delta D(s_1,s_2)} \prod_{i=1}^{2} \max\left\{\mathcal{N}(f_i), \mu_d(f_i)\right\}.$$

Finally, by approximating $\{\chi_{B_k}\}_{k>s_0}$ from below by smooth functions, applying the above estimate (together with a trivial estimate when $D(s_1, s_2)$ is small, see (7.13)) and the bounds on the norm $\mathcal{N}(\cdot)$ (see Lemma 6.5), we show that the divergence of the series in (1.10) together with the additional technical assumption (1.11), implies that the required quasi-independence condition (7.8) is satisfied, thus concluding the proof of the divergence case of Theorem 1.2.

We end our discussion with a few remarks.

Remark 7. Our argument should be compared to that of Kleinbock and Margulis [19], where the shrinking targets are certain cusp neighborhoods: In [19] the relevant shrinking target problem is first solved for the case of a_s-orbits starting at μ_d-generic points in the ambient space X_d; for this

task it suffices to use, in place of (1.14) and (1.15) respectively, the invariance of the measure μ_d and the exponential mixing of the a_s-flow. Then by an approximation argument [19, §8.7], the shrinking target property for μ_d-generic points in X_d is shown to imply the same property for generic points in the submanifold $\mathcal{Y}$. A key observation in this approximation step is that *all* shrinking targets, by virtue of being cusp neighborhoods, remain essentially unaffected by perturbations from a *fixed* neighborhood of the identity in the neutral leaf of the a_s-flow, i.e. the centralizer of the a_s-flow in $\mathrm{SL}_d(\mathbb{R})$. This, however, is no longer the case in our setting, with the shrinking targets being compact sets. This is why we take the more direct approach using effective equidistribution and doubly mixing of the a_s-translates of $\mathcal{Y}$, that is, (1.14) and (1.15).

One potential advantage of this more direct approach is that if (1.14) could be refined by replacing the measure Leb by a natural measure on some *submanifold* of $\mathcal{Y}$, then by mimicking our analysis, one could establish the ψ-Dirichlet property for generic points in that submanifold, for any ψ such that (1.10) converges. See Remark 14 below for a discussion of the application along these lines of a recent effective equidistribution result obtained by Chow and Yang [5].

We note that the use of equidistribution of a_s-translates of $\mathcal{Y}$ in the study of the Dirichlet improvability problem is not new; it has been applied several times in the more well-studied setting of Dirichlet improvable vectors and matrices. For $\min\{m,n\}=1$ and $\mathcal{I}\subset\mathcal{Y}$ being an analytic curve in $\mathcal{Y}$ satisfying certain explicit conditions, Shah [32, Theorem 1.2] proved that the a_s-translates of $\mathcal{I}$ equidistribute in X_d with respect to μ_d as $s\to\infty$. Shah's proof relies on Ratner's classification of measures invariant under unipotent flows [31], and his equidistribution theorem is not effective; still it suffices for the deduction of the fact that generic points on the curve $\mathcal{I}$ are Dirichlet non-improvable, that is, lie outside of the set (1.3). (This is so since in this case, the relevant "shrinking" target is in fact a *fixed* set of positive measure.) Shah's results have been generalized and strengthened in various directions [33, 34, 37, 17]. In a recent breakthrough of Khalil and Luethi [14], the authors refined (1.14) (for the case when $n=1$) by replacing Leb with a certain fractal measure, from which they deduced a complete analogue of Khintchine's theorem with respect to this fractal measure.

Remark 8. Another difficulty, which also stems from the fact that our targets are shrinking compact sets, is the fact that the norm $\mathcal{N}(\cdot)$ unavoidably blows up (polynomially) for the smooth functions approximating the shrinking targets from above and below (see Lemma 6.5). While the impact of this blow-up of the norm can be eliminated in the convergence case due to the exponential decay in the parameter s (i.e. the factor $e^{-\delta s}$ in the error term in (1.14)), it causes serious problems in the divergence case, and this is exactly why we need to impose the extra assumption (1.11). Let us here also note that this assumption (1.11) can be rephrased in terms of the measure of the shrinking targets as follows:

$$\liminf_{s_1\to\infty}\frac{\sum_{s_0<k\le s_1}\mu_d(B_k)\log\left(\frac{1}{\mu_d(B_k)}\right)}{\left(\sum_{s_0<k\le s_1}\mu_d(B_k)\right)^2}=0.$$

Remark 9. In order to extend the argument outlined above to the case of general weight vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, we have to consider a more general one-parameter flow $\{g_s\}_{s>0}\subset\mathrm{SL}_d(\mathbb{R})$ associated to $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ (see (5.1)), and use a dynamical interpretation of $\psi_{\boldsymbol{\alpha},\boldsymbol{\beta}}$-Dirichlet matrices which involves this g_s-flow and generalizes (1.6); see Proposition 6.2 and Remark 13. We therefore need analogous effective equidistribution and doubly mixing results for the g_s-translates of $\mathcal{Y}$. Fortunately, such more general (and considerably more difficult) effective results are known to hold, thanks to the work of Kleinbock-Margulis [20, Theorem 1.3] and, again, Kleinbock-Shi-Weiss [24, Theorem 1.2] and Björklund-Gorodnik [1, Corollary 2.4] (see Theorem 6.3 below). In fact in [1] a uniform

treatment was given proving effective mixing of arbitrary order for the g_s-translates of $\mathcal{Y}$; however we will not make use of this.

Notation and conventions. Throughout the paper, the notation $\|\cdot\|$ denotes the supremum norm on various Euclidean spaces or matrix spaces (which can be viewed as Euclidean spaces on the matrix entries). Let $I \subset \mathbb{R}$ be an interval (not necessarily bounded). A function $f : I \to \mathbb{R}$ is called *increasing* (resp. *decreasing*) if $f(t_1) \le f(t_2)$ (resp. $f(t_1) \ge f(t_2)$) whenever $t_1 < t_2$. All the vectors in this paper are column vectors. For two positive quantities A and B, we will use the notation $A \ll B$ or $A = O(B)$ to mean that there is a constant $c > 0$ such that $A \le cB$, and we will use subscripts to indicate the dependence of the constant on parameters. We will write $A \asymp B$ for $A \ll B \ll A$.

2. Some preliminaries for Theorem 1.3

Fix an integer $d \ge 2$. In what follows we always denote $G = \mathrm{SL}_d(\mathbb{R})$, $\Gamma = \mathrm{SL}_d(\mathbb{Z})$ and $X_d = G/\Gamma$ the space of unimodular lattices in $\mathbb{R}^d$. Let μ_d be the unique G-invariant probability measure on X_d. Let $\Delta : X_d \to [0,\infty)$ be the function on X_d defined as in (1.4). In this section, we collect some preliminary results for our proof of Theorem 1.3. In fact, for simplicity of presentation we will prove an equivalent measure estimate result. For any $r \in [0,1)$ let $\mathcal{C}_r \subset \mathbb{R}^d$ be the open "$(1-r)$-cube", i.e.

$$\mathcal{C}_r := (r-1, 1-r)^d.$$

Let $K_r \subset X_d$ be the set of unimodular lattices having no nonzero points in $\mathcal{C}_r$, i.e.

$$K_r := \big\{\Lambda \in X_d \ : \ \Lambda \cap \mathcal{C}_r = \{\mathbf{0}\}\big\}.$$

We note that by definition of Δ, $K_r = \Delta^{-1}[0, -\log(1-r)]$, or equivalently, $\Delta^{-1}[0,r] = K_{1-e^{-r}}$. Since $1 - e^{-r} = r + O(r^2) \asymp r$ for all $r \in (0,1)$, Theorem 1.3 can be equivalently restated as follows.

Theorem 2.1. *Let $\varkappa_d = \frac{d^2+d-4}{2}$ and $\lambda_d = \frac{d(d-1)}{2}$ be as in Theorem 1.2. Then*

$$\mu_d(K_r) \asymp_d r^{\varkappa_d+1} \log^{\lambda_d}\Big(\frac{1}{r}\Big), \qquad \text{as } r \to 0^+.$$

We will prove Theorem 2.1 by proving a lower bound and an upper bound separately.

2.1. Haar measure and coordinates.

Let $P < G$ be the maximal parabolic subgroup fixing the line spanned by $\boldsymbol{e}_d \in \mathbb{R}^d$, and let $N < G$ be the transpose of the unipotent radical of P. Here and hereafter, $\{\boldsymbol{e}_i : 1 \le i \le d\}$ denotes the standard orthonormal basis of $\mathbb{R}^d$. Explicitly,

$$P = \{p \in G \ : \ p\boldsymbol{e}_d = t\boldsymbol{e}_d \text{ for some } t \ne 0\},$$

and

$$N = \left\{u_{\boldsymbol{x}} := \begin{pmatrix} I_{d-1} & \boldsymbol{x} \\ \mathbf{0}^t & 1 \end{pmatrix} \ : \ \boldsymbol{x} \in \mathbb{R}^{d-1}\right\}.$$

For any $p \in P$, let $\boldsymbol{b}_1 := p\boldsymbol{e}_1, \dots, \boldsymbol{b}_{d-1} := p\boldsymbol{e}_{d-1}$ be the first $d-1$ column vectors of p. We note that p is uniquely determined by $\boldsymbol{b}_1, \dots, \boldsymbol{b}_{d-1}$; we will sometimes denote $p \in P$ by $p_{\boldsymbol{b}_1,\dots,\boldsymbol{b}_{d-1}}$ to indicate this dependence. For any $g \in G$, let us denote by $\tilde{g} \in M_{d-1,d-1}(\mathbb{R})$ the top left $(d-1)\times(d-1)$ block of g. If $\det \tilde{g} \ne 0$, then g can be written uniquely as a product

$$g = p_{\boldsymbol{b}_1,\dots,\boldsymbol{b}_{d-1}} u_{\boldsymbol{x}} \quad \text{for some } p_{\boldsymbol{b}_1,\dots,\boldsymbol{b}_{d-1}} \in P \text{ and } u_{\boldsymbol{x}} \in N. \tag{2.1}$$

Let ν be the (left and right) Haar measure on G, normalized so that it agrees locally with μ_d. In terms of the coordinates in (2.1), ν is given by

$$\mathrm{d}\nu(g) = \frac{1}{\zeta(2)\cdots\zeta(d)}\,\mathrm{d}\boldsymbol{x}\prod_{1\le i\le d-1}\mathrm{d}\boldsymbol{b}_i, \tag{2.2}$$

where $\zeta(\cdot)$ is the Riemann zeta function, and where $\mathrm{d}\boldsymbol{x}$ and $\mathrm{d}\boldsymbol{b}_i$ denote Lebesgue measure on $\mathbb{R}^{d-1}$ and $\mathbb{R}^d$, respectively. For later purpose, we also note that the lattice Λ represented by $p_{\boldsymbol{b}_1,\dots,\boldsymbol{b}_{d-1}}u_{\boldsymbol{x}}$, i.e. $\Lambda = p_{\boldsymbol{b}_1,\dots,\boldsymbol{b}_{d-1}}u_{\boldsymbol{x}}\mathbb{Z}^d$, has a basis

$$\Lambda = \mathbb{Z}\boldsymbol{b}_1\oplus\cdots\oplus\mathbb{Z}\boldsymbol{b}_{d-1}\oplus\mathbb{Z}\boldsymbol{b}_d,$$

where $\boldsymbol{b}_d := \sum_{j=1}^{d-1}x_j\boldsymbol{b}_j + (\det\tilde{p})^{-1}\boldsymbol{e}_d$ is the d-th column vector of the matrix $p_{\boldsymbol{b}_1,\dots,\boldsymbol{b}_{d-1}}u_{\boldsymbol{x}}$. Here $\tilde{p}$ is the top left $(d-1)\times(d-1)$ block of $p_{\boldsymbol{b}_1,\dots,\boldsymbol{b}_{d-1}}$.

For our computation of the upper bounds, it will be more convenient to use another set of coordinates: For any $g = (g_{ij})_{1\le i,j\le d}\in G$ with $\det\tilde{g}\neq 0$, as mentioned above, we can write g uniquely as in (2.1). It is clear from this relation that g and $p_{\boldsymbol{b}_1,\dots,\boldsymbol{b}_{d-1}}$ share the same first $d-1$ column vectors, i.e. $g\boldsymbol{e}_j = \boldsymbol{b}_j$ for all $1\le j\le d-1$. Moreover, as noted above, for the d-th column vector we have

$$g\boldsymbol{e}_d = \sum_{j=1}^{d-1}x_j\boldsymbol{b}_j + (\det\tilde{p})^{-1}\boldsymbol{e}_d = \sum_{j=1}^{d-1}x_j\,(g\boldsymbol{e}_j) + (\det\tilde{g})^{-1}\boldsymbol{e}_d.$$

In particular, we have $(g_{1d},\dots,g_{d-1,d})^t = \tilde{g}\boldsymbol{x}$, which further implies

$$\mathrm{d}\boldsymbol{x} = (\det\tilde{g})^{-1}\prod_{1\le i\le d-1}\mathrm{d}g_{id}.$$

This relation, together with the relations $g\boldsymbol{e}_j = \boldsymbol{b}_j, 1\le j\le d-1$ and the Haar measure description (2.2), immediately implies the following:

Lemma 2.2. *For any (Borel) subset $\mathcal{K}$ of $\left\{g\in G\,:\,|\det\tilde{g}-1|<\frac{1}{2}\right\}$, we have*

$$\nu(\mathcal{K}) \asymp_d \int_{\mathcal{K}}\prod_{\substack{1\le i,j\le d\\(i,j)\neq(d,d)}}\mathrm{d}g_{ij}. \tag{2.3}$$

2.2. **A small parameter for the lower bound.** To prove the lower bound, we will construct a subset of K_r whose measure is of the same magnitude as K_r. For a lattice $\Lambda = g\mathbb{Z}^d\in X_d$, to show $\Lambda\in K_r$, by definition one needs to show $g\boldsymbol{m}\notin\mathcal{C}_r$ for all nonzero $\boldsymbol{m}\in\mathbb{Z}^d$. If $g\in G$ is sufficiently close to the identity element $I_d\in G$, so that Λ has a basis close to the standard basis $\{\boldsymbol{e}_i\,:\,1\le i\le d\}$, then one only needs to consider vectors $\boldsymbol{m}\in\mathbb{Z}^d$ with small supremum norms. For this reason, we will only focus on lattices that are close to $\mathbb{Z}^d$. Recall that the set K_r certainly does not get concentrated near the lattice $\mathbb{Z}^d$ as $r\to 0^+$; indeed, we have $\cap_{r>0}K_r = K_0 = \Delta^{-1}\{0\}$, which as we have mentioned is a finite union of compact submanifolds of positive codimension $\varkappa_d+1$ in X_d (see also Section 2.3). The fact that it still suffices to consider a small neighborhood of $\mathbb{Z}^d$ when proving the *lower* bound in Theorem 2.1 is related to the fact that the *mass* of K_r (with respect to μ_d) becomes concentrated near the lattice $\mathbb{Z}^d$ as $r\to 0^+$; see Remark 10.

Explicitly, we fix a small norm ball in X_d around $\mathbb{Z}^d$ as follows: For any $c>0$, let

$$G_c := \{g\in G\,:\,\|g-I_d\|<c\} \tag{2.4}$$

be the open ball in G of radius c, centered at I_d with respect to the supremum norm on $M_{d,d}(\mathbb{R})$. Let $\pi: G\to X_d$ be the natural projection from G to X_d. We fix a parameter $c_0\in(0,1)$ (which

only depends on d) so small that $\pi|_{G_{c_0}}$ is injective and, for any vectors $\boldsymbol{b}_1, \dots, \boldsymbol{b}_d \in \mathbb{R}^d$ satisfying $\|\boldsymbol{b}_i - \boldsymbol{e}_i\| < c_0$ for all $1 \le i \le d$, every hyperplane of the form

$$m\boldsymbol{b}_j + \sum_{\substack{1 \le \ell \le d \\ \ell \ne j}} \mathbb{R}\boldsymbol{b}_\ell$$

with $m \in \mathbb{R}$, $|m| \ge 2$ and $1 \le j \le d$, is disjoint from the cube $[-1,1]^d$. In particular, if $\Lambda = g\mathbb{Z}^d$ for some $g \in G_{c_0}$, then Λ has a basis $\{g\boldsymbol{e}_i : 1 \le i \le d\}$ satisfying $\|g\boldsymbol{e}_i - \boldsymbol{e}_i\| \le c_0$ for all $1 \le i \le d$. It follows that in order to prove that $\Lambda \in K_r$ for a given $r \in (0,1)$, it suffices to verify that $\sum_{1 \le i \le d} m_i (g\boldsymbol{e}_i) \notin \mathcal{C}_r$ for all $\boldsymbol{m} = (m_1, \dots, m_d) \in \{-1,0,1\}^d \smallsetminus \{\boldsymbol{0}\}$.

2.3. **Hajós's Theorem and its consequences.** Recall that

$$K_0 = \left\{ \Lambda \in X_d \,:\, \Lambda \cap (-1,1)^d = \{\boldsymbol{0}\} \right\}.$$

As we mentioned in the introduction, the explicit description of K_0 was conjectured (and proved in two and three dimensions) by Minkowski, and proved in full generality by Hajós in 1941 [11]:

Theorem 2.3 (Hajós)**.** *Let U be the subgroup of upper triangular unipotent matrices in G. Let W be the subgroup of permutation matrices in $GL_d(\mathbb{Z})$. Then*

$$K_0 = \bigcup_{w \in W} \left(wUw^{-1}\right) \mathbb{Z}^d.$$

If we set

$$U_0 = \left\{ (u_{ij}) \in U \,:\, -\tfrac{1}{2} < u_{ij} \le \tfrac{1}{2} \text{ for all } 1 \le i < j \le d \right\} \tag{2.5}$$

so that U_0 is a fundamental domain for $U/(\Gamma \cap U)$, then we have the following immediate corollary of Hajós's Theorem.

Corollary 2.4. *Given any $\Lambda \in K_0$, there exist $w \in W$ and $u \in U_0$ such that $\Lambda = wu\mathbb{Z}^d = wuw^{-1}\mathbb{Z}^d$.*

Proof. Since $\Lambda \in K_0$, by Theorem 2.3 we can find $u' \in U$ and $w \in W$ such that $\Lambda = wu'w^{-1}\mathbb{Z}^d$; but since $w^{-1}\mathbb{Z}^d = \mathbb{Z}^d$, we have $\Lambda = wu'\mathbb{Z}^d$. Now using the fact that U_0 is a fundamental domain for $U/(\Gamma \cap U)$, we can then find $u \in U_0$ such that $u\mathbb{Z}^d = u'\mathbb{Z}^d$. Thus $\Lambda = wu'\mathbb{Z}^d = wu\mathbb{Z}^d = wuw^{-1}\mathbb{Z}^d$, finishing the proof. □

There is a geometric interpretation of K_0 in terms of *lattice tilings by unit cubes* [35, Ch. 1.4]: Let us write $\frac{1}{2}\mathcal{C}_0 = (-\frac{1}{2}, \frac{1}{2})^d$ for the unit cube obtained by dilating $\mathcal{C}_0$ by a factor $\frac{1}{2}$. Then for any $\Lambda \in X_d$, the family of cubes $\boldsymbol{v} + \frac{1}{2}\mathcal{C}_0$, with $\boldsymbol{v}$ running through the lattice Λ, forms a tiling of $\mathbb{R}^d$ (modulo a null set) if and only if $\Lambda \in K_0$. More generally, for any $r \in [0,1)$ we write $\frac{1}{2}\mathcal{C}_r = \left(\frac{1}{2}(r-1), \frac{1}{2}(1-r)\right)^d$. Then for any $\Lambda \in X_d$, the condition $\Lambda \in K_r$, i.e. $\Lambda \cap \mathcal{C}_r = \{\boldsymbol{0}\}$, is equivalent to the condition that the cubes $\boldsymbol{v} + \frac{1}{2}\mathcal{C}_r$ ($\boldsymbol{v} \in \Lambda$) are *pairwise disjoint*. When this holds, we write

$$\mathfrak{C}_{\Lambda,r} := \Lambda + \tfrac{1}{2}\mathcal{C}_r$$

for the union of these disjoint cubes. This set is used in the statement of the following simple bound, which is of crucial importance in our proof of the upper bound in Theorem 2.1.

Lemma 2.5. *Let $\Lambda \in X_d$ and $r \in (0, \frac{1}{4})$ be such that $\Lambda \cap \mathcal{C}_r = \{\boldsymbol{0}\}$, and let $\mathcal{U}$ be a Borel subset of $\mathbb{R}^d$ which is disjoint from $\mathfrak{C}_{\Lambda,r}$ and which is contained in some translate of the cube $(0, \frac{3}{4})^d$. Then* $\mathrm{vol}(\mathcal{U}) < dr$.

Proof. The set $\mathfrak{C}_{\Lambda,r}$ is invariant under translation by any vector in Λ, and if $\mathcal{F} \subset \mathbb{R}^d$ is any fundamental domain for $\mathbb{R}^d/\Lambda$, then

$$\mathrm{vol}(\mathcal{F} \cap \mathfrak{C}_{\Lambda,r}) = \sum_{\boldsymbol{v}\in\Lambda} \mathrm{vol}\left(\mathcal{F} \cap (\boldsymbol{v} + \tfrac{1}{2}\mathcal{C}_r)\right) = \sum_{\boldsymbol{v}\in\Lambda} \mathrm{vol}\left((\mathcal{F} - \boldsymbol{v}) + \tfrac{1}{2}\mathcal{C}_r\right) = \mathrm{vol}(\tfrac{1}{2}\mathcal{C}_r) = (1-r)^d,$$

where the first equality holds since the cubes $\boldsymbol{v} + \frac{1}{2}\mathcal{C}_r$ ($\boldsymbol{v} \in \Lambda$) are pairwise disjoint, and the third equality holds since the sets $\mathcal{F} - \boldsymbol{v}$ ($\boldsymbol{v} \in \Lambda$) form a partition of $\mathbb{R}^d$. Hence

$$\mathrm{vol}(\mathcal{F} \smallsetminus \mathfrak{C}_{\Lambda,r}) = 1 - (1-r)^d < dr.$$

Next, since $\mathcal{U}$ is contained in a translate of $(0, \frac{3}{4})^d$, the difference between any two vectors in $\mathcal{U}$ lies in $(-\frac{3}{4}, \frac{3}{4})^d = \mathcal{C}_{1/4} \subset \mathcal{C}_r$, and since $\Lambda \cap \mathcal{C}_r = \{\mathbf{0}\}$, this implies that the points in $\mathcal{U}$ are pairwise inequivalent modulo Λ. Hence the set $\big(\mathcal{F} \smallsetminus (\Lambda + \mathcal{U})\big) \cup \mathcal{U}$ is another fundamental domain for $\mathbb{R}^d/\Lambda$, and it contains $\mathcal{U}$. After replacing $\mathcal{F}$ by this set, we have $\mathcal{U} \subset \mathcal{F}$; thus $\mathcal{U} \subset \mathcal{F} \smallsetminus \mathfrak{C}_{\Lambda,r}$, and hence $\mathrm{vol}(\mathcal{U}) \le \mathrm{vol}(\mathcal{F} \smallsetminus \mathfrak{C}_{\Lambda,r}) < dr$. □

3. Proof of the lower bound

We keep the notation introduced in Section 2. In this section we prove the lower bound in Theorem 2.1. We will do this by constructing, for every sufficiently small r, an explicit subset $\underline{K}_r \subset X_d$ which we will show satisfies

$$\underline{K}_r \subset K_r \qquad \text{and} \qquad \mu_d(\underline{K}_r) \gg_d r^{\varkappa_d+1} \log^{\lambda_d}\Big(\frac{1}{r}\Big).$$

We start by giving a family of conditions which ensures that a given lattice is contained in K_r. Recall from Section 2.2 that $c_0 \in (0,1)$ is a small fixed parameter with the property that for any $g \in G_{c_0}$ and $0 < r < 1$, we have $g\mathbb{Z}^d \in K_r$ if and only if $\sum_{1\le j\le d} m_j\,(g\boldsymbol{e}_j) \notin \mathcal{C}_r$ for all $\boldsymbol{m} = (m_1,\dots,m_d)^t \in \{-1,0,1\}^d \smallsetminus \{\mathbf{0}\}$.

Lemma 3.1. *Let $c_0 \in (0,1)$ be as above, let $r \in (0, c_0/d)$ and let $\Lambda = p_{\boldsymbol{b}_1,\dots,\boldsymbol{b}_{d-1}} u_{\boldsymbol{x}} \mathbb{Z}^d \in X_d$ with $\boldsymbol{b}_j = (b_{1j},\dots,b_{dj})^t \in \mathbb{R}^d$ ($j = 1,\dots,d-1$) and $\boldsymbol{x} \in (0, c_0/d)^{d-1}$. Let $\tilde{p} = (b_{ij})_{1\le i,j\le d-1}$ be the top left $(d-1)\times(d-1)$ block of $p_{\boldsymbol{b}_1,\dots,\boldsymbol{b}_{d-1}}$ as before. Suppose $\boldsymbol{b}_1,\dots,\boldsymbol{b}_{d-1},\boldsymbol{x}$ further satisfies*

$$b_{ij}, -b_{ji} \in (-c_0, 0),\ \forall\ 1 \le j < i \le d-1;\ b_{d\ell} \in (-c_0, 0), b_{\ell\ell} \in (1-r, 1),\ \forall\ 1 \le \ell \le d-1, \tag{3.1}$$

$$b_{ij} < \sum_{k=1}^{j-1} b_{ik}, \qquad \forall\ 2 \le j < i \le d, \tag{3.2}$$

$$1 - r < \det \tilde{p} < (1-r)^{-1} \qquad \text{and} \qquad \sum_{j=1}^{d-1} |b_{dj}| x_j < (\det \tilde{p})^{-1} - (1-r), \tag{3.3}$$

and

$$b_{ii} + \sum_{j=1}^{d-1} b_{ij} x_j > 1 - r, \qquad \forall\ 1 \le i \le d-1. \tag{3.4}$$

Then $\Lambda \in K_r$.

Proof. Let us set $g := p_{\boldsymbol{b}_1,\dots,\boldsymbol{b}_{d-1}} u_{\boldsymbol{x}}$ and $\boldsymbol{b}_d = (b_{1d},\dots,b_{dd})^t := g\boldsymbol{e}_d$; then $\boldsymbol{b}_j = g\boldsymbol{e}_j$ for all $1 \le j \le d$, and $\boldsymbol{b}_d := \sum_{j=1}^{d-1} x_j \boldsymbol{b}_j + (\det \tilde{p})^{-1} \boldsymbol{e}_d$; in particular $b_{id} = \sum_{j=1}^{d-1} b_{ij} x_j$ for $1 \le i \le d-1$. We will start by proving that $g \in G_{c_0}$, i.e. $\|\boldsymbol{b}_j - \boldsymbol{e}_j\| < c_0$ for all $1 \le j \le d$. In fact, if $1 \le j \le d-1$ then $\|\boldsymbol{b}_j - \boldsymbol{e}_j\| < c_0$ is an immediate consequence of (3.1) and $0 < r < c_0/d < c_0$; thus it remains to show $\|\boldsymbol{b}_d - \boldsymbol{e}_d\| < c_0$. For each $1 \le i \le d-1$ we have

$$|b_{id}| = \left| \sum_{j=1}^{d-1} b_{ij} x_j \right| \le \sum_{j=1}^{d-1} |b_{ij}|\,|x_j| < \sum_{j=1}^{d-1} \frac{c_0}{d} < c_0,$$

where for the second inequality we used the assumption that $\boldsymbol{x} \in (0, c_0/d)^{d-1}$ and the fact that $|b_{ij}| < 1$ for all $1 \le i,j \le d-1$, which is immediate from (3.1). It remains to prove that $1 - c_0 < b_{dd} < 1 + c_0$. In fact, we have the following stronger bound:

$$1 - r < b_{dd} < 1 + c_0. \tag{3.5}$$

Indeed, using $b_{dj} < 0$ and $x_j > 0$ $(1 \le j \le d-1)$ in combination with (3.3), we get

$$b_{dd} = \sum_{j=1}^{d-1} b_{dj} x_j + (\det \tilde{p})^{-1} = (\det \tilde{p})^{-1} - \sum_{j=1}^{d-1} |b_{dj}|\, x_j > 1 - r$$

as well as

$$b_{dd} = \sum_{j=1}^{d-1} b_{dj} x_j + (\det \tilde{p})^{-1} < (\det \tilde{p})^{-1} < (1-r)^{-1} < 1 + 2r \le 1 + c_0.$$

(For the second to last inequality we used the fact that $0 < r < c_0/d < 1/2$.) This finishes the proof of (3.5), and hence $g \in G_{c_0}$.

Because of $g \in G_{c_0}$, in order to show $\Lambda \in K_r$, it suffices to prove that $g\boldsymbol{m} = \sum_{1 \le j \le d} m_j \boldsymbol{b}_j \notin \mathcal{C}_r$ for all $\boldsymbol{m} \in \{-1,0,1\}^d \smallsetminus \{\boldsymbol{0}\}$. Thus we now let the vector $\boldsymbol{m} = (m_1,\dots,m_d)^t \in \{-1,0,1\}^d \smallsetminus \{\boldsymbol{0}\}$ be given, and show that $g\boldsymbol{m} \notin \mathcal{C}_r$.

First assume that all the nonzero entries of $\boldsymbol{m}$ are of the same sign. After replacing $\boldsymbol{m}$ by $-\boldsymbol{m}$ if necessary, we may assume $m_j \ge 0$ for all $1 \le j \le d$. Let $k \in \{1,\dots,d\}$ be the smallest integer such that $m_k = 1$. If $k = d$ then $\boldsymbol{m} = \boldsymbol{e}_d$, and thus $g\boldsymbol{m} = \boldsymbol{b}_d \notin \mathcal{C}_r$ by (3.5). In the remaining case when $k < d$, the k-th coordinate of $g\boldsymbol{m}$ satisfies

$$m_k b_{kk} + \sum_{j=k+1}^{d} m_j b_{kj} \ge b_{kk} + m_d b_{kd} \ge \min\{b_{kk}, b_{kk} + b_{kd}\}, \tag{3.6}$$

where we used the fact that, by (3.1), $b_{kj} > 0$ for each $k < j \le d-1$. Furthermore,

$$\min\{b_{kk}, b_{kk} + b_{kd}\} > 1 - r, \tag{3.7}$$

since $b_{kk} > 1 - r$ by (3.1) and $b_{kk} + b_{kd} = b_{kk} + \sum_{j=1}^{d-1} b_{kj} x_j > 1 - r$ by (3.4). It follows from (3.6) and (3.7) that the k-th coordinate of $g\boldsymbol{m}$ is larger than $1 - r$, and so $g\boldsymbol{m} \notin \mathcal{C}_r$. This completes the proof in the case when all the nonzero entries of $\boldsymbol{m}$ are of the same sign.

It remains to treat the case when $\{-1,1\} \subset \{m_j : 1 \le j \le d\}$. Then let $1 \le i_1 < i_2 \le d$ be the indices which record the latest instance when the signs of the entries of $\boldsymbol{m}$ change, i.e. the unique indices such that $m_{i_1} m_{i_2} = -1$, $m_j = 0$ for $i_1 < j < i_2$ and $m_j \in \{m_{i_2}, 0\}$ for $i_2 < j \le d$ (the last two conditions are void if $i_1 + 1 = i_2$ or $i_2 = d$, respectively). Again after replacing $\boldsymbol{m}$ by $-\boldsymbol{m}$ if

necessary, we may assume $m_{i_1} = -1$, $m_{i_2} = 1$, and thus $m_j \geq 0$ for all $i_2 < j \leq d$. Now consider the i_2-th coordinate of $g\boldsymbol{m}$ which is

$$\sum_{1\leq j<i_1} m_j b_{i_2,j} - b_{i_2,i_1} + b_{i_2,i_2} + \sum_{i_2<j\leq d} m_j b_{i_2,j}.$$

Here we have

$$\sum_{1\leq j<i_1} m_j b_{i_2,j} - b_{i_2,i_1} \geq \sum_{1\leq j<i_1} b_{i_2,j} - b_{i_2,i_1} > 0, \tag{3.8}$$

where the first inequality holds since $|m_j| \leq 1$ and $b_{i_2,j} < 0$ for all $j < i_1$, and the second inequality holds by (3.2) (except if $i_1 = 1$; in that case (3.8) simply says that $b_{i_2,i_1} < 0$, which holds by (3.1)). Furthermore,

$$b_{i_2,i_2} + \sum_{i_2<j\leq d} m_j b_{i_2,j} \quad \begin{cases} = b_{dd} > 1 - r & \text{if } \ i_2 = d, \\ \geq b_{i_2,i_2} + m_d b_{i_2,d} > 1 - r & \text{if } \ i_2 < d, \end{cases} \tag{3.9}$$

where we used (3.5) in the case $i_2 = d$, and in the case $i_2 < d$ we used the fact that $m_j \geq 0$ and $b_{i_2,j} > 0$ for all $i_2 < j \leq d-1$ (if any), and then applied (3.7) with $k = i_2$. Adding the inequalities in (3.8) and (3.9), we conclude that the i_2-th coordinate of $g\boldsymbol{m}$ is larger than $1-r$; hence $g\boldsymbol{m} \notin \mathcal{C}_r$. This concludes the proof of the lemma. □

We next give another family of conditions, which implies the conditions in Lemma 3.1, and which is more suitable for the measure computations which we are going to carry out.

Proposition 3.2. *Let $c_0 \in (0,1)$ be as above, let $r \in (0, c_0/d)$ and let $\Lambda = p_{\boldsymbol{b}_1,\dots,\boldsymbol{b}_{d-1}} u_{\boldsymbol{x}} \mathbb{Z}^d \in X_d$ with $\boldsymbol{b}_j = (b_{1j}, \dots, b_{dj})^t \in \mathbb{R}^d$ $(j = 1, \dots, d-1)$ and $\boldsymbol{x} \in (0, c_0/d)^{d-1}$. Assume that*

$$b_{ij}, -b_{ji} \in (-c_0, 0), \ \forall\ 1 \leq j < i \leq d-1; \ b_{d\ell} \in (-c_0, 0), b_{\ell\ell} \in (1 - \tfrac{r}{2d}, 1), \ \forall\ 1 \leq \ell \leq d-1, \tag{3.10}$$

$$b_{ij} < d b_{i,j-1} \ (\Leftrightarrow |b_{ij}| > d|b_{i,j-1}|), \qquad \forall\ 2 \leq j < i \leq d, \tag{3.11}$$

$$|b_{ij} b_{ji}| < \frac{r}{d!}, \quad \forall 1 \leq j < i \leq d-1, \qquad \text{and} \qquad \sum_{j=1}^{d-1} |b_{dj}| x_j < \frac{r}{2}, \tag{3.12}$$

and

$$b_{kj} > b_{ij} \ (\Leftrightarrow |b_{kj}| < |b_{ij}|), \qquad \forall 1 \leq j < k < i \leq d. \tag{3.13}$$

Then $\Lambda \in K_r$.

Proof. In view of Lemma 3.1, it suffices to show that the conditions (3.1), (3.2), (3.3) and (3.4) are fulfilled. Among these, (3.1) is an immediate consequence of (3.10). Furthermore, (3.2) follows from (3.10) and (3.11) by the following computation, valid for any $2 \leq j < i \leq d$:

$$b_{ij} < d b_{i,j-1} < (j-1) b_{i,j-1} \leq \sum_{k=1}^{j-1} b_{ik} \quad (< 0),$$

where the last relation is an equality when $j = 2$, while for $j \geq 3$ it is a strict inequality which holds since (3.11) forces $b_{i,j-1} < b_{i,j-2} < \cdots < b_{i,1}$. Also (3.4) is easily proved: Let $1 \leq i \leq d-1$. Then for every $1 \leq j \leq d-1$ we have $x_j > 0$ and $b_{ij} > b_{dj}$ (the latter holds by (3.13) if $j < i$ and by (3.10) if $j \geq i$). Hence

$$b_{ii} + \sum_{j=1}^{d-1} b_{ij} x_j > b_{ii} + \sum_{j=1}^{d-1} b_{dj} x_j > 1 - \frac{r}{2d} - \frac{r}{2} > 1 - r,$$

where for the second last inequality we used (3.10) and the second part of (3.12).

It remains to prove (3.3). We first note that the ordering assumptions in (3.11) and (3.13) imply

$$|b_{i',j'}| \le |b_{ij}| \qquad \text{whenever } 1 \le j' < i' \le d,\ 1 \le j < i \le d \text{ and } i' \le i \text{ and } j' \le j. \tag{3.14}$$

Now let $\tilde{p} = (b_{ij})_{1 \le i,j \le d-1}$ be as in Lemma 3.1. Define

$$\varphi_{\tilde{p}} := \sum_{\substack{\sigma \in S_{d-1} \\ \sigma \neq \mathrm{id}}} \prod_{1 \le i \le d-1} \left| b_{\sigma(i)i} \right|,$$

where S_{d-1} is the symmetric group over the finite set $\{1, 2, \dots, d-1\}$ and $\mathrm{id} \in S_{d-1}$ is its identity element. Then

$$\prod_{i=1}^{d-1} b_{ii} - \varphi_{\tilde{p}} \le \det \tilde{p} \le \prod_{i=1}^{d-1} b_{ii} + \varphi_{\tilde{p}}. \tag{3.15}$$

If $d = 2$ then $S_{d-1} = \{\mathrm{id}\}$ and $\varphi_{\tilde{p}} = 0$. Now assume $d \ge 3$, and consider an arbitrary permutation $\sigma \in S_{d-1} \smallsetminus \{\mathrm{id}\}$. Let $(i_1 \dots i_\ell)$ be a cycle of σ of length $\ell \ge 2$, meaning that $\sigma(i_j) = i_{j+1}$ for all $1 \le j \le \ell - 1$ and $\sigma(i_\ell) = i_1$. Without loss of generality, we may assume $i_\ell = \max\{i_j : 1 \le j \le \ell\}$. Let $i_{j_0} := \min\{i_j : 1 \le j \le \ell\}$. Then (3.14) applies for the pairs $(i', j') = (\sigma(i_{j_0}), i_{j_0})$ and $(i, j) = (i_\ell, i_1)$, so that $\left|b_{\sigma(i_{j_0}),i_{j_0}}\right| \le \left|b_{i_\ell,i_1}\right|$. We also have $|b_{ij}| < 1$ for all $1 \le i, j \le d-1$, by (3.10). Hence

$$\prod_{1 \le i \le d-1} \left|b_{\sigma(i)i}\right| \le \left|b_{\sigma(i_{j_0}),i_{j_0}} b_{i_1,i_\ell}\right| \le |b_{i_\ell,i_1} b_{i_1,i_\ell}| < \frac{r}{d!},$$

where the last inequality holds by (3.12). The above holds for every $\sigma \in S_{d-1} \smallsetminus \{\mathrm{id}\}$; hence

$$\varphi_{\tilde{p}} < (d-1)! \frac{r}{d!} = \frac{r}{d}. \tag{3.16}$$

Note that (3.16) also holds when $d = 2$, trivially.

Using (3.15), (3.16), and the fact that $1 - \frac{r}{2d} < b_{ii} < 1$ for all $1 \le i \le d-1$ (cf. (3.10)), we get

$$\det \tilde{p} < \prod_{i=1}^{d-1} b_{ii} + \frac{r}{d} < 1 + \frac{r}{d} < (1-r)^{-1}, \tag{3.17}$$

and

$$\det \tilde{p} > \prod_{i=1}^{d-1} b_{ii} - \frac{r}{d} > \left(1 - \frac{r}{2d}\right)^{d-1} - \frac{r}{d} > 1 - \frac{r}{2} - \frac{r}{d} \ge 1 - r.$$

Hence we have proved the first condition in (3.3). For the second condition in (3.3), in view of the second condition in (3.12) it suffices to show $\frac{r}{2} < (\det \tilde{p})^{-1} - (1-r)$, or equivalently, that $\det \tilde{p}$ is smaller than $(1 - \frac{r}{2})^{-1}$. But this is true since by (3.17), $\det \tilde{p} < 1 + \frac{r}{d} < (1 - \frac{r}{2})^{-1}$. This finishes the proof. □

We can now give the

Proof of the lower bound in Theorem 2.1. Keep $0 < r < c_0/d$, and define

$$\underline{K}_r := \left\{ p_{\boldsymbol{b}_1,\dots,\boldsymbol{b}_{d-1}} u_{\boldsymbol{x}} \mathbb{Z}^d \in X_d \ : \ \begin{array}{c} (\boldsymbol{b}_1, \dots, \boldsymbol{b}_{d-1}, \boldsymbol{x}) \in (\mathbb{R}^d)^{d-1} \times (0, c_0/d)^{d-1} \\ \text{satisfies (3.10), (3.11), (3.12), (3.13)} \end{array} \right\}. \tag{3.18}$$

By Proposition 3.2 we have $\underline{K}_r \subset K_r$. It remains to bound $\mu_d(\underline{K}_r)$ from below. Set

$$\underline{\mathcal{K}}_r := \left\{ p_{\boldsymbol{b}_1,\dots,\boldsymbol{b}_{d-1}} u_{\boldsymbol{x}} \in G \ : \ \begin{array}{c} (\boldsymbol{b}_1,\dots,\boldsymbol{b}_{d-1},\boldsymbol{x}) \in (\mathbb{R}^d)^{d-1} \times (0,c_0/d)^{d-1} \\ \text{satisfies (3.10), (3.11), (3.12), (3.13)} \end{array} \right\},$$

so that $\pi(\underline{\mathcal{K}}_r) = \underline{K}_r$. Here $\pi : G \to X_d$ is the natural projection as before. By immediate inspection of the proofs of Proposition 3.2 and Lemma 3.1, we have $\underline{\mathcal{K}}_r \subset G_{c_0}$, and by our choice of c_0, $\pi|_{G_{c_0}}$ is injective (see Section 2.2). Hence $\mu_d(\underline{K}_r) = \nu(\underline{\mathcal{K}}_r)$, and by (2.2) we have

$$\nu(\underline{\mathcal{K}}_r) \asymp_d \prod_{1\le k\le d-1} \left(\int_{1-\frac{r}{2d}}^{1} \mathrm{d}b_{kk} \right) \int_{\mathcal{R}} \delta\big((b_{ij})_{1\le j<i\le d} \big) \prod_{1\le j<i\le d} \mathrm{d}b_{ij},$$

where

$$\mathcal{R} := \left\{ (b_{ij})_{1\le j<i\le d} \in (-c_0,0)^{d(d-1)/2} \ : \ (b_{ij}) \text{ satisfies (3.11) and (3.13)} \right\},$$

and

$$\begin{aligned} \delta\big((b_{ij})_{1\le j<i\le d} \big) :&= \prod_{1\le j<i\le d-1} \int_0^{\min\left\{c_0, \frac{r}{d!|b_{ij}|}\right\}} \mathrm{d}b_{ji} \times \int_{\left\{\boldsymbol{x}\in\left(0,\frac{c_0}{d}\right)^{d-1} : \sum_{j=1}^{d-1} |b_{dj}| x_j < \frac{r}{2}\right\}} \prod_{1\le j\le d-1} \mathrm{d}x_j \\ &\asymp_{d,c_0} \prod_{1\le j<i\le d-1} \min\left\{1, \frac{r}{|b_{ij}|}\right\} \times \prod_{1\le j\le d-1} \min\left\{1, \frac{r}{|b_{dj}|}\right\} \\ &= \prod_{1\le j<i\le d} \min\left\{1, \frac{r}{|b_{ij}|}\right\}. \end{aligned}$$

Hence

$$\mu_d(\underline{K}_r) = \nu(\underline{\mathcal{K}}_r) \asymp_{d,c_0} r^{d-1} \int_{\mathcal{R}} \left(\prod_{1\le j<i\le d} \min\left\{1, \frac{r}{|b_{ij}|}\right\} \mathrm{d}b_{ij} \right). \tag{3.19}$$

Now for each $1 \le j < i \le d$, we make a change of variable, $b_{ij} = -d^{j-1} z_{ij}$, to simplify the ordering condition (3.11). Then all the z_{ij}'s are positive, and the conditions (3.11) and (3.13) become

$$z_{i'j'} < z_{ij} \quad \text{whenever } 1 \le j' < i' \le d,\ 1 \le j < i \le d,\ i' \le i,\ j' \le j \text{ and } (i',j') \ne (i,j). \tag{3.20}$$

Moreover, for any $j < i$, the condition $b_{ij} \in (-c_0, 0)$ corresponds to $z_{ij} \in (0, c_0/d^{j-1})$, and we note that each of these intervals contains the fixed interval $(0, c_0/d^{d-1})$. In fact, let us restrict each z_{ij} to the even smaller interval $(r, c_0/d^{d-1})$, and assume $r < c_0/d^{d-1}$ so that this interval is non-empty; then $r/|b_{ij}| = rd^{1-j}/z_{ij} < d^{1-j} \le 1$, so that $\min\{1, r/|b_{ij}|\} = rd^{1-j}/z_{ij} \asymp_d r/z_{ij}$. Note also that $|\,\mathrm{d}b_{ij}| = d^{j-1}\,\mathrm{d}z_{ij} \asymp_d \mathrm{d}z_{ij}$; hence we get from (3.19):

$$\mu_d(\underline{K}_r) \gg_{d,c_0} r^{d-1+\frac{d(d-1)}{2}} \int_{\mathcal{R}'} \prod_{1\le j<i\le d} \frac{\mathrm{d}z_{ij}}{z_{ij}} = r^{\frac{(d-1)(d+2)}{2}} \int_{\mathcal{R}'} \prod_{1\le j<i\le d} \frac{\mathrm{d}z_{ij}}{z_{ij}},$$

where

$$\mathcal{R}' := \left\{ (z_{ij})_{1\le j<i\le d} \in (r, c_0/d^{d-1})^{d(d-1)/2} \ : \ (z_{ij}) \text{ satisfies (3.20)} \right\}.$$

The last integrand is invariant under every permutation of the variables $(z_{ij})_{1\le j<i\le d}$, and the integration regions $\sigma(\mathcal{R}')$ with σ running through all these permutations cover (modulo a null set)

the cube $(r, c_0/d^{d-1})^{d(d-1)/2}$; hence we obtain

$$\mu_d\left(\underline{K}_r\right) \gg_{d,c_0} r^{\frac{(d-1)(d+2)}{2}} \prod_{1\le j<i\le d}\left(\int_r^{c_0/d^{d-1}} \frac{\mathrm{d}z_{ij}}{z_{ij}}\right) \asymp_{d,c_0} r^{\frac{(d-1)(d+2)}{2}} \log^{\frac{d(d-1)}{2}}\Big(\frac{1}{r}\Big), \tag{3.21}$$

where the last estimate is valid e.g. for all $0 < r < c_0/(2d^{d-1})$. Recalling that c_0 only depends on d, we see that the lower bound in Theorem 2.1 is now proved. □

4. Proof of the upper bound

We keep the notation introduced in Section 2. In this section we prove the upper bound in Theorem 2.1. The main step in the proof is to show that for r small, K_r is contained in a certain set of more explicit nature; see Proposition 4.2 below. To prepare for the statement of this result, we start by introducing the following set, for any $r \in (0,1)$ and $C > 0$:

$$\overline{\mathcal{K}}_{r,C} := \left\{ g = (g_{ij}) \in G \ : \ \begin{array}{l} 1 - r \le g_{ii} \le 1 + Cr, \ \forall\ 1 \le i \le d \\ |g_{ij}| < 1, \ |g_{ij}g_{ji}| \le Cr, \ \forall\ 1 \le i \ne j \le d \\ |\det \tilde{g} - 1| < \frac{1}{2} \end{array} \right\}.$$

Here $\tilde{g}$ denotes the top left $(d-1)\times(d-1)$ block of g as before. In view of the Haar measure description (2.3), it is not difficult to compute the measure of $\overline{\mathcal{K}}_{r,C}$:

Lemma 4.1. *For any $r \in (0, \frac{1}{2})$ and $C > 0$ we have*

$$\nu\big(\overline{\mathcal{K}}_{r,C}\big) \ll_{d,C} r^{\varkappa_d+1} \log^{\lambda_d}\Big(\frac{1}{r}\Big).$$

Proof. Let us define

$$\overline{\mathcal{K}}'_{r,C} := \left\{ g = (g_{ij}) \in G \ : \ \begin{array}{l} 1 - r \le g_{ii} \le 1 + Cr, \ \forall\ 1 \le i \le d-1 \\ |g_{ij}| < 1, \ |g_{ij}g_{ji}| \le Cr, \ \forall\ 1 \le i \ne j \le d \\ |\det \tilde{g} - 1| < \frac{1}{2} \end{array} \right\}$$

by disregarding the restriction on g_{dd}. Then clearly $\overline{\mathcal{K}}_{r,C} \subset \overline{\mathcal{K}}'_{r,C}$. Moreover, in view of the Haar measure description (2.3) we have (noting also that $\varkappa_d + 1 = \lambda_d + d - 1$)

$$\begin{aligned} \nu\left(\overline{\mathcal{K}}'_{r,C}\right) &\ll_d \prod_{1\le i\le d-1} \int_{1-r}^{1+Cr} \mathrm{d}g_{ii} \prod_{1\le j<i\le d} \int_{-1}^{1}\int_{-1}^{1} \chi_{\{(g_{ij},g_{ji}) \,:\, |g_{ij}g_{ji}|<Cr\}}\, \mathrm{d}g_{ij}\, \mathrm{d}g_{ji} \\ &\asymp_C r^{d-1}\left(r\log\Big(\frac{1}{r}\Big)\right)^{\lambda_d} = r^{\varkappa_d+1}\log^{\lambda_d}\Big(\frac{1}{r}\Big). \end{aligned}$$

Thus

$$\nu\left(\overline{\mathcal{K}}_{r,C}\right) \le \nu\left(\overline{\mathcal{K}}'_{r,C}\right) \ll_{d,C} r^{\varkappa_d+1}\log^{\lambda_d}\Big(\frac{1}{r}\Big),$$

finishing the proof. □

Recall that $\pi : G \to X_d$ is the natural projection from G to X_d.

Proposition 4.2. *There exist $r_0 > 0$ and $C > 0$ (depending only on d) such that for all $r \in (0, r_0)$*

$$K_r \subset \bigcup_{w\in W} \pi\big(w\,\overline{\mathcal{K}}_{r,C}\,w^{-1}\big). \tag{4.1}$$

Let us first give a quick

Proof of the upper bound in Theorem 2.1 assuming Proposition 4.2. The Haar measure ν on G is preserved by conjugation by any element $w \in W$ (even though w may be outside G). Hence it follows from (4.1) that, for any $r \in (0, r_0)$,

$$\mu_d(K_r) \le \sum_{w\in W} \mu_d\big(\pi(w\,\overline{\mathcal{K}}_{r,C}\,w^{-1})\big) \le \sum_{w\in W} \nu(w\,\overline{\mathcal{K}}_{r,C}\,w^{-1}) = d!\,\nu(\overline{\mathcal{K}}_{r,C}).$$

Using this inequality, the upper bound in Theorem 2.1 now follows from Lemma 4.1. □

Remark 10. Proposition 4.2, in combination with the lower bound in Theorem 2.1, also implies that as $r \to 0^+$, the mass of K_r with respect to μ_d becomes concentrated near the lattice $\mathbb{Z}^d$. In precise terms, if $\mathcal{O}$ is any fixed neighborhood of $\mathbb{Z}^d$ in X_d, then

$$\mu_d\left(K_r \smallsetminus \mathcal{O}\right) \ll_d r^{\varkappa_d+1} \log^{\lambda_d-1}\Big(\frac{1}{r}\Big), \quad \text{and thus} \quad \frac{\mu_d(K_r \cap \mathcal{O})}{\mu_d(K_r)} \to 1 \qquad \text{as } r \to 0^+. \tag{4.2}$$

Indeed, we can fix $\varepsilon > 0$ so that $\mathcal{O}$ contains the set $\pi(G_\varepsilon)$ with $G_\varepsilon \subset G$ the norm ball defined as in (2.4); then by arguing along the same lines as above, the first relation in (4.2) will follow from the following bound:

$$\nu\left(\overline{\mathcal{K}}_{r,C} \smallsetminus G_\varepsilon\right) \ll_{d,C} r^{\varkappa_d+1} \log^{\lambda_d-1}\Big(\frac{1}{r}\Big),$$

for $C > 0$ and r small. However, for $r < \varepsilon/C$, $g = (g_{ij}) \in \overline{\mathcal{K}}_{r,C}$ forces $|g_{ii} - 1| < \varepsilon$ for all i, and so the set $\overline{\mathcal{K}}_{r,C} \smallsetminus G_\varepsilon$ is contained in the union $\cup_{i'\neq j'}\big\{g = (g_{ij}) \in \overline{\mathcal{K}}_{r,C} \,:\, |g_{i'j'}| \ge \varepsilon\big\}$. Therefore, it suffices to prove that for any given $1 \le i' \ne j' \le d$ we have

$$\nu\big(\big\{g = (g_{ij}) \in \overline{\mathcal{K}}_{r,C} \,:\, |g_{i'j'}| \ge \varepsilon\big\}\big) \ll_{d,C} r^{\varkappa_d+1} \log^{\lambda_d-1}\Big(\frac{1}{r}\Big).$$

This is shown by following the proof of Lemma 4.1 and using

$$\int_{\varepsilon\le|x|\le 1} \int_{-1}^{1} \chi_{\{(x,y)\,:\,|xy|<Cr\}}\, \mathrm{d}y\, \mathrm{d}x \ll_{\varepsilon,C} r.$$

Finally, the second relation in (4.2) follows from the first relation combined with the lower bound in Theorem 2.1.

The remainder of this section is devoted to the proof of Proposition 4.2.

4.1. **Bounds for diagonal entries.** Recall that U_0 is the fixed fundamental domain for $U/(\Gamma \cap U)$ given in (2.5). The next lemma shows that if a lattice $\Lambda \in K_r$ has a representative sufficiently close to some element in U_0, then the diagonal entries of such a representative satisfy the desired bounds.

Lemma 4.3. *Let $g = (g_{ij}) \in G$ and $r \in (0, \frac{1}{8})$, and assume that $g\mathbb{Z}^d \cap \mathcal{C}_r = \{\mathbf{0}\}$ and $\|g - u\| < \frac{1}{8}$ for some $u \in U_0$. Then*

$$0 \le g_{ii} - (1 - r) \ll_d r, \qquad \forall\, 1 \le i \le d. \tag{4.3}$$

Proof. Note that it follows from $\|g - u\| < \frac{1}{8}$ and $u \in U_0$ that $|g_{ij}| < \frac{5}{8}$ for all $1 \le i < j \le d$, $|g_{ii} - 1| < \frac{1}{8}$ for all $1 \le i \le d$, and $|g_{ij}| < \frac{1}{8}$ for all $1 \le j < i \le d$. For each i we have $g\boldsymbol{e}_i \notin \mathcal{C}_r$, since $g\mathbb{Z}^d \cap \mathcal{C}_r = \{\mathbf{0}\}$. Combining this with the fact that $|g_{ji}| < \frac{5}{8} < 1 - r$ for all $j \ne i$, we conclude that $|g_{ii}| \ge 1 - r$. Since also $|g_{ii} - 1| < \frac{1}{8}$, we must in fact have $g_{ii} \ge 1 - r$, i.e. we have proved the left inequality in (4.3).

Next, let $i \in \{1, \dots, d\}$ be given, and let $\mathcal{U}$ be the open box

$$\mathcal{U} = \mathcal{I}_1 \times \cdots \times \mathcal{I}_d, \quad \text{with} \quad \begin{cases} \mathcal{I}_i = \big(\frac{1}{2}(1-r), g_{ii} - \frac{1}{2}(1-r)\big), \\ \mathcal{I}_j = \big(g_{ji}^+ - \frac{3}{8}, g_{ji}^- + \frac{3}{8}\big) \ (j \neq i), \end{cases}$$

where $g_{ji}^+ := \max\{g_{ji}, 0\}$ and $g_{ji}^- := \min\{g_{ji}, 0\}$. Note that each interval $\mathcal{I}_j$ $(j \neq i)$ has length $|\mathcal{I}_j| > \frac{1}{8}$, since $|g_{ji}| < \frac{5}{8}$; furthermore $|\mathcal{I}_i| = g_{ii} - (1-r) \geq 0$ (thus $\mathcal{I}_i$ and $\mathcal{U}$ are empty if $g_{ii} = 1-r$, but otherwise non-empty).

We claim that $\mathcal{U}$ is disjoint from $\mathfrak{C}_{g\mathbb{Z}^d,r} := g\mathbb{Z}^d + \frac{1}{2}\mathcal{C}_r$. Indeed, assume the opposite; then there is some $\boldsymbol{v} \in g\mathbb{Z}^d$ such that $\mathcal{U} \cap (\boldsymbol{v} + \frac{1}{2}\mathcal{C}_r) \neq \varnothing$. We must have $\boldsymbol{v} \neq \mathbf{0}$ and $\boldsymbol{v} \neq g\boldsymbol{e}_i$, since $\mathcal{U}$ is, by construction, disjoint both from $\frac{1}{2}\mathcal{C}_r$ and from $g\boldsymbol{e}_i + \frac{1}{2}\mathcal{C}_r$. Pick a point $\boldsymbol{x} \in \mathcal{U} \cap (\boldsymbol{v} + \frac{1}{2}\mathcal{C}_r)$. It follows from $\boldsymbol{x} \in \boldsymbol{v} + \frac{1}{2}\mathcal{C}_r$ and $\frac{1}{2}(1-r) > \frac{1}{4}$ that at least one of the points $\boldsymbol{x} - \frac{1}{4}\boldsymbol{e}_i$ or $\boldsymbol{x} + \frac{1}{4}\boldsymbol{e}_i$ *also* lies in $\boldsymbol{v} + \frac{1}{2}\mathcal{C}_r$. But we have

$$\boldsymbol{x} - \tfrac{1}{4}\boldsymbol{e}_i \in \tfrac{1}{2}\mathcal{C}_r,$$

since $x_j \in \mathcal{I}_j \subset \big(-\frac{1}{2}(1-r), \frac{1}{2}(1-r)\big)$ for all $j \neq i$ and $x_i - \frac{1}{4} < g_{ii} - \frac{1}{2}(1-r) - \frac{1}{4} < \frac{1}{2}(1-r)$. We also have

$$\boldsymbol{x} + \tfrac{1}{4}\boldsymbol{e}_i \in g\boldsymbol{e}_i + \tfrac{1}{2}\mathcal{C}_r,$$

since $x_j \in \mathcal{I}_j \subset g_{ji} + \big(-\frac{1}{2}(1-r), \frac{1}{2}(1-r)\big)$ for all $j \neq i$ and $x_i + \frac{1}{4} > \frac{1}{2}(1-r) + \frac{1}{4} > g_{ii} - \frac{1}{2}(1-r)$. Hence we have arrived at a contradiction against the fact that $\boldsymbol{v} + \frac{1}{2}\mathcal{C}_r$ is disjoint from both $\frac{1}{2}\mathcal{C}_r$ and $g\boldsymbol{e}_i + \frac{1}{2}\mathcal{C}_r$. This completes the proof of the fact that $\mathcal{U}$ is disjoint from $\mathfrak{C}_{g\mathbb{Z}^d,r}$.

Note also that $\mathcal{U}$ is contained in a translate of the cube $(0, \frac{3}{4})^d$, since each interval $\mathcal{I}_j$ has length at most $\frac{3}{4}$. Hence Lemma 2.5 applies, and yields that $\mathrm{vol}(\mathcal{U}) < dr$. But we have noted that $|\mathcal{I}_j| > \frac{1}{8}$ for each $j \neq i$, and $|\mathcal{I}_i| = g_{ii} - (1-r)$; hence $\mathrm{vol}(\mathcal{U}) \geq 8^{1-d}(g_{ii} - (1-r))$. Combining these facts, we obtain the right bound in (4.3) with the implied constant $8^{d-1}d$. □

4.2. **A technical choice of lattice representatives.** For any $\varepsilon > 0$ let us write

$$U_\varepsilon = \big\{(u_{ij}) \in U \ : \ -\tfrac{1}{2} + 4\varepsilon < u_{ij} \leq \tfrac{1}{2} + \varepsilon \text{ for all } 1 \leq i < j \leq d\big\}.$$

Lemma 4.4. *There exist constants $0 < a < 1$ and $A > 1$, which only depend on d, such that the following holds: given any $u \in U$ and $\varepsilon \in (0, a)$, there exist $\gamma \in \Gamma \cap U$ and $B \in [1, A]$ such that $u\gamma \in U_{B\varepsilon}$.*

Proof. We will show that the statement of this lemma holds with $A := 4^{2^d}$ and $a := 2^{-(3+2^{d+1})}$. Let us set

$$U[t] := \big\{(u_{ij}) \in U \ : \ -\tfrac{1}{2} + t < u_{ij} \leq \tfrac{1}{2} + t \text{ for all } 1 \leq i < j \leq d\big\} \qquad (t \in \mathbb{R}).$$

Note that $U_0 = U[0]$, and for each t, $U[t]$ is a fundamental domain for $U/(\Gamma \cap U)$.

For the given $u \in U$ and $\varepsilon \in (0, a)$, and for each $k \in \mathbb{Z}_{\geq 0}$, we let $u^{(k)}$ be the unique element in $u\Gamma \cap U[4^k\varepsilon]$, and set

$$P^{(k)} = \sum_{j=2}^{d} P_j^{(k)}, \quad \text{where} \quad P_j^{(k)} := \sum_{i=1}^{j-1} 2^{i-1} \cdot \chi_{(0,\infty)}(u_{ij}^{(k)}).$$

We now claim:

$$\forall\, k \in \mathbb{Z}_{\geq 0} : \quad k < \log_4(\tfrac{1}{8\varepsilon}) \text{ and } u^{(k)} \notin U_{4^k\varepsilon} \quad \Longrightarrow \quad P^{(k)} < P^{(k+1)}. \tag{4.4}$$

To prove (4.4), assume $0 \le k < \log_4(\frac{1}{8\varepsilon})$ and $u^{(k)} \notin U_{4^k\varepsilon}$. Then $u^{(k+1)} \neq u^{(k)}$, since otherwise $u^{(k)}$ would lie in the intersection $U[4^k\varepsilon] \cap U[4^{k+1}\varepsilon] = U_{4^k\varepsilon}$. Let us denote by $\boldsymbol{u}_j^{(k)}$ the jth column vector of $u^{(k)}$. It follows that $\boldsymbol{u}_j^{(k+1)} \neq \boldsymbol{u}_j^{(k)}$ for at least one $j \in \{2, \ldots, d\}$, and for each such index j we can argue as follows: Since $u^{(k)}, u^{(k+1)} \in u\Gamma \cap U = u(\Gamma \cap U)$, we have $\boldsymbol{u}_j^{(k+1)} = \boldsymbol{u}_j^{(k)} + \sum_{1 \le i < j} m_i \boldsymbol{u}_i^{(k)}$ for some $m_1, \ldots, m_{j-1} \in \mathbb{Z}$. Because of $\boldsymbol{u}_j^{(k+1)} \neq \boldsymbol{u}_j^{(k)}$, there exists $i \in \{1, \ldots, j\}$ such that $m_i \neq 0$; let us fix i to be the *largest* such index. Thus $u_{ij}^{(k+1)} = u_{ij}^{(k)} + m_i$; in particular $\left|u_{ij}^{(k+1)} - u_{ij}^{(k)}\right| \ge 1$. On the other hand, note that $u_{ij}^{(k)} \in \left(-\frac{1}{2} + 4^k\varepsilon, \frac{1}{2} + 4^k\varepsilon\right]$, $u_{ij}^{(k+1)} \in \left(-\frac{1}{2} + 4^{k+1}\varepsilon, \frac{1}{2} + 4^{k+1}\varepsilon\right]$ and $\left(\frac{1}{2} + 4^{k+1}\varepsilon\right) - \left(-\frac{1}{2} + 4^k\varepsilon\right) = 1 + 3 \cdot 4^k\varepsilon < 2$ (since $k < \log_4(\frac{1}{8\varepsilon})$). Hence we must have $m_i = 1$, implying that

$$\boldsymbol{u}_j^{(k+1)} \in \boldsymbol{u}_j^{(k)} + \boldsymbol{u}_i^{(k)} + \mathbb{Z}\boldsymbol{u}_{i-1}^{(k)} + \cdots + \mathbb{Z}\boldsymbol{u}_1^{(k)}.$$

Thus $u_{i'j}^{(k+1)} = u_{i'j}^{(k)}$ for all $i' > i$, and $u_{ij}^{(k+1)} = u_{ij}^{(k)} + 1 > 0$ while $u_{ij}^{(k)} \le -\frac{1}{2} + 4^{k+1}\varepsilon < 0$ (again since $k < \log_4(\frac{1}{8\varepsilon})$). It follows that $P_j^{(k+1)} - P_j^{(k)} \ge 2^{i-1} - \sum_{i'<i} 2^{i'-1} = 1$. On the other hand we clearly have $P_j^{(k+1)} = P_j^{(k)}$ for each $j \in \{2, \ldots, d\}$ such that $\boldsymbol{u}_j^{(k+1)} = \boldsymbol{u}_j^{(k)}$. Hence $\sum_{j=2}^d P_j^{(k+1)} > \sum_{j=2}^d P_j^{(k)}$, i.e. $P^{(k+1)} > P^{(k)}$. This finishes the proof of (4.4).

Next note that by definition, for each k, $P^{(k)}$ is a non-negative integer satisfying

$$P^{(k)} \le \sum_{j=2}^{d} \sum_{i=0}^{j-2} 2^i < 2^d.$$

This implies that we cannot have $P^{(k)} < P^{(k+1)}$ for all $k \in \{0, 1, \ldots, 2^d - 1\}$. But our assumption on ε implies that $\log_4(\frac{1}{8\varepsilon}) > 2^d$ (recall that $0 < \varepsilon < a = 2^{-(3+2^{d+1})}$); hence it now follows from (4.4) that $u^{(k)} \in U_{4^k\varepsilon}$ for at least one $k \in \{0, 1, \ldots, 2^d - 1\}$. Furthermore, for such k we have $4^k < 4^{2^d} = A$ and $u^{(k)} \in u(\Gamma \cap U)$ by construction. Hence the lemma is proved. □

Using Lemma 4.4, Corollary 2.4 and a compactness argument we have the following technical lemma, which gives us a good choice of lattice representatives for lattices in K_{r_0} for some small $r_0 > 0$.

Lemma 4.5. *There exist constants $0 < a < 1$ and $A > 1$, which only depend on d, such that the following holds: for any $\varepsilon_0 \in (0, a)$ there exists $r_0 > 0$ such that for every $\Lambda \in X_d$ satisfying $\Lambda \cap \mathcal{C}_{r_0} = \{\mathbf{0}\}$, there exist $g \in G$, $w \in W$, $B \in [1, A]$ and $u \in U_{B\varepsilon_0}$ such that $\Lambda = g\mathbb{Z}^d$ and $\|g - wuw^{-1}\| < \varepsilon_0$.*

Proof. Let a and A be as in Lemma 4.4; we will prove that the statement of the lemma holds with these a, A. The proof is by contradiction; thus we assume that the statement of the lemma is false, i.e. we assume that there exist some $\varepsilon_0 \in (0, a)$ and a sequence $r_1 > r_2 > \cdots$ in $(0, 1)$ with $r_j \to 0$, and a corresponding sequence $\Lambda_1, \Lambda_2, \ldots$ in X_d, such that $\Lambda_j \cap \mathcal{C}_{r_j} = \{\mathbf{0}\}$ for each j, and furthermore, for each j we have that there do *not* exist any $g \in G$, $w \in W$, $B \in [1, A]$ and $u \in U_{B\varepsilon_0}$ satisfying $\Lambda_j = g\mathbb{Z}^d$ and $\|g - wuw^{-1}\| < \varepsilon_0$. Now for every j we have $\mathcal{C}_{r_1} \subset \mathcal{C}_{r_j}$, and thus $\Lambda_j \cap \mathcal{C}_{r_1} = \{\mathbf{0}\}$. Hence by Mahler's Compactness Theorem, after passing to a subsequence we may assume that Λ_j tends to a limit point in X_d. Let us call this limit point Λ_0; thus $\Lambda_j \to \Lambda_0$ in X_d as $j \to \infty$. Let us also fix a representative $g_0 \in G$ such that $\Lambda_0 = g_0\mathbb{Z}^d$.

Recall that the standard topology on $X_d = G/\Gamma$ is given by the metric

$$\operatorname{dist}(g\Gamma, g'\Gamma) := \inf\{d(g\gamma, g') \,:\, \gamma \in \Gamma\} \qquad (g, g' \in G),$$

where $d(\cdot,\cdot)$ is any fixed right G-invariant Riemannian metric on G. Hence the fact that Λ_j converges to $\Lambda_0 = g_0\mathbb{Z}^d$ implies that there exist $g_1, g_2, \ldots \in G$ such that $\Lambda_j = g_j\mathbb{Z}^d$ for each j and $d(g_j, g_0) \to 0$ as $j \to \infty$.

Using $\Lambda_j \cap \mathcal{C}_{r_j} = \{\mathbf{0}\}$ for each j, and $r_j \to 0$, we claim that

$$\Lambda_0 \cap (-1,1)^d = \{\mathbf{0}\}. \tag{4.5}$$

Indeed, assume the opposite; this means that there exists some $\boldsymbol{m} \in \mathbb{Z}^d \smallsetminus \{\mathbf{0}\}$ such that $g_0\boldsymbol{m}$ belongs to $(-1,1)^d$, i.e. $\|g_0\boldsymbol{m}\| < 1$. Set $r := \frac{1}{2}(1 - \|g_0\boldsymbol{m}\|)$; then $g_0\boldsymbol{m} \in \mathcal{C}_r$. We have $g_j\boldsymbol{m} \to g_0\boldsymbol{m}$ as $j \to \infty$, since $d(g_j, g_0) \to 0$; hence for all sufficiently large j we have $g_j\boldsymbol{m} \in \mathcal{C}_r$ (since $\mathcal{C}_r$ is open). Also for all sufficiently large j we have $r_j < r$. Hence there exists some j for which $r_j < r$ and $g_j\boldsymbol{m} \in \mathcal{C}_r \subset \mathcal{C}_{r_j}$. This contradicts the fact that $\Lambda_j \cap \mathcal{C}_{r_j} = \{\mathbf{0}\}$ for all j. Hence (4.5) is proved.

It follows from (4.5) and Corollary 2.4 that $\Lambda_0 = wu'w^{-1}\mathbb{Z}^d$ for some $w \in W$ and $u' \in U_0$. Next, by Lemma 4.4 (and since $\varepsilon_0 < a$), there exist $\gamma \in \Gamma \cap U$ and $B \in [1, A]$ such that $u := u'\gamma \in U_{B\varepsilon_0}$. Using $w^{-1}\mathbb{Z}^d = \mathbb{Z}^d$ and $\gamma^{-1}\mathbb{Z}^d = \mathbb{Z}^d$, we then have $\Lambda_0 = wuw^{-1}\mathbb{Z}^d$. Hence $wuw^{-1} = g_0\gamma_0$ for some $\gamma_0 \in \Gamma$. Now $d(g_j\gamma_0, wuw^{-1}) = d(g_j\gamma_0, g_0\gamma_0) = d(g_j, g_0) \to 0$ as $j \to \infty$. This implies that every matrix entry of $g_j\gamma_0$ tends to the corresponding entry of wuw^{-1}, i.e. we have $\|g_j\gamma_0 - wuw^{-1}\| \to 0$ as $j \to \infty$. In particular there exists some j such that $\|g_j\gamma_0 - wuw^{-1}\| < \varepsilon_0$. Now we have a contradiction against our previous assumption; namely for our chosen j, if we set $g := g_j\gamma_0$ then $\Lambda_j = g\mathbb{Z}^d$ and $\|g - wuw^{-1}\| < \varepsilon_0$. This completes the proof of Lemma 4.5. □

Remark 11. By a similar compactness argument as in the proof of Lemma 4.5, one can also prove a more basic statement: for any $\varepsilon_0 > 0$ there exists $r_0 > 0$ such that every $\Lambda \in K_{r_0}$ has a representative $g \in G$ (i.e. $\Lambda = g\mathbb{Z}^d$) satisfying $\|g - wuw^{-1}\| < \varepsilon_0$ for some $w \in W$ and $u \in U_0$. The purpose of the choice of the more technical lattice representatives in Lemma 4.5 (with $U_{B\varepsilon_0}$ in place of U_0) is to ensure the following property, which is a key ingredient in the proof of the important Lemma 4.6 below: for any $g = (g_{ij}) \in G$ satisfying $\|g - u\| < \varepsilon_0$ for some $u \in U_{B\varepsilon_0}$, and any $1 \le i < j \le d$, we have (cf. (4.8))

$$|g_{ki} - g_{kj}| < 1 - B\varepsilon_0 \quad \text{for all } k \notin \{i, j\}.$$

A crucial consequence of this is that if $g\boldsymbol{e}_i - g\boldsymbol{e}_j \notin \mathcal{C}_r$ for some $r < B\varepsilon_0$, then either $|g_{ii} - g_{ij}| > 1 - r$ or $|g_{ji} - g_{jj}| > 1 - r$ must hold.

4.3. **Bounds for off-diagonal symmetric pairs.** The next lemma shows that if a lattice Λ in K_r has a representative as in Lemma 4.5, then its entries satisfy the desired bounds for proving Proposition 4.2.

Lemma 4.6. *Let $A > 1$, $0 < \varepsilon_0 < (16A)^{-1}$ and $0 < r_0 < \frac{1}{2}\varepsilon_0$. Let $g = (g_{ij}) \in G$ and $r \in (0, r_0)$, and assume that $g\mathbb{Z}^d \cap \mathcal{C}_r = \{\mathbf{0}\}$ and that there exist $B \in [1, A]$ and $u \in U_{B\varepsilon_0}$ such that $\|g - u\| < \varepsilon_0$. Then*

$$0 \le g_{ii} - (1 - r) \ll_d r \quad \text{for all } 1 \le i \le d, \tag{4.6}$$

and

$$|g_{ij}g_{ji}| \ll_{d,\varepsilon_0} r \quad \text{for all } 1 \le i < j \le d. \tag{4.7}$$

Proof. Let us write $\varepsilon = B\varepsilon_0$. Note that $\|g - u\| < \varepsilon_0 \le \varepsilon$ and $u \in U_\varepsilon$ together imply that

$$-\tfrac{1}{2} + 3\varepsilon < g_{ij} < \tfrac{1}{2} + 2\varepsilon \quad \text{and} \quad |g_{ji}| < \varepsilon, \quad \forall 1 \le i < j \le d, \tag{4.8}$$

as well as $|g_{ii}-1|<\varepsilon$ for all i. The last inequality can be sharpened using Lemma 4.3. Indeed, we have $\varepsilon<\frac{1}{16}$ since $A\varepsilon_0<\frac{1}{16}$; hence the above inequalities imply that $\|g-u'\|<\frac{1}{8}$ for some $u'\in U_0$. Hence Lemma 4.3 applies, yielding that (4.6) holds.

Now let $1\le i<j\le d$ be given. We separate the proof of (4.7) into three cases. (Note that (4.7) holds trivially if $g_{ij}=0$ or $g_{ji}=0$; hence we may without loss of generality assume $g_{ij}g_{ji}\neq 0$.)

Case I: $g_{ij}>0$ and $g_{ji}>0$. In this case we will build the proof on the fact that $g\boldsymbol{e}_i-g\boldsymbol{e}_j\notin\mathcal{C}_r$, which holds since $g\mathbb{Z}^d\cap\mathcal{C}_r=\{\boldsymbol{0}\}$. Using (4.8) and $r<r_0<2r_0\le\varepsilon<\frac{1}{16}$, it follows that $|g_{ki}-g_{kj}|<1-\varepsilon<1-r$ for all $k\notin\{i,j\}$. Hence we must have either $|g_{ii}-g_{ij}|\ge 1-r$ or $|g_{ji}-g_{jj}|\ge 1-r$. If $|g_{ii}-g_{ij}|\ge 1-r$, then because of $g_{ii}\ge 1-r$ and $0<g_{ij}<\frac{1}{2}+2\varepsilon$ it follows that $g_{ii}-g_{ij}\ge 1-r$, and so by (4.6), $0<g_{ij}\ll r$. Similarly if $|g_{ji}-g_{jj}|\ge 1-r$ then $0<g_{ji}\ll r$. In both cases, it follows that (4.7) holds for our i,j.

Case II: $g_{ij}<0$ and $g_{ji}<0$. In this case we will prove the desired bound by proving the stronger assertion that either $|g_{ij}|\le Cr$ or $|g_{ji}|\le Cr$, with $C:=12d(\varepsilon_0/8)^{2-d}$. Assume the opposite, i.e. assume that

$$g_{ij}<-Cr\qquad\text{and}\qquad g_{ji}<-Cr\qquad(C:=12d(\varepsilon_0/8)^{2-d}).\tag{4.9}$$

We will prove that this leads to a contradiction.

Set $\mathcal{J}_r:=\big(-\frac{1}{2}(1-r),\frac{1}{2}(1-r)\big)$, so that $\frac{1}{2}\mathcal{C}_r=\mathcal{J}_r^d$. For each $k\notin\{i,j\}$ we introduce the following open interval:

$$\widetilde{\mathcal{I}}_k:=\mathcal{J}_r\cap(g_{ki}+\mathcal{J}_r)\cap(g_{kj}+\mathcal{J}_r).\tag{4.10}$$

Using $|g_{ki}|,|g_{kj}|,|g_{ki}-g_{kj}|<1-\varepsilon$ (see (4.8)) it follows that $\widetilde{\mathcal{I}}_k$ has length $|\widetilde{\mathcal{I}}_k|>(1-r)-(1-\varepsilon)>\frac{1}{2}\varepsilon_0$. Define $\mathcal{I}_k\subset\widetilde{\mathcal{I}}_k$ to be the open interval of length $\frac{1}{4}\varepsilon_0$ with the same center as $\widetilde{\mathcal{I}}_k$. Let us also set $\mathcal{I}_i=\big(\frac{1}{2}(1-r)+g_{ij},\frac{1}{2}(1-r)\big)$ and $\mathcal{I}_j=\big(g_{jj}-\frac{1}{2}(1-r),\frac{3}{4}\big)$. Then by construction,

$$\mathcal{I}_i\subset\mathcal{J}_r,\qquad\mathcal{I}_i\cap(g_{ii}+\mathcal{J}_r)=\varnothing,\qquad\mathcal{I}_i\cap(g_{ij}+\mathcal{J}_r)=\varnothing,\tag{4.11}$$

and

$$\mathcal{I}_j\subset g_{jj}+\mathcal{J}_r,\qquad\mathcal{I}_j\cap\mathcal{J}_r=\varnothing.\tag{4.12}$$

Furthermore, $|\mathcal{I}_i|=|g_{ij}|$ and $\frac{1}{8}<|\mathcal{I}_j|<\frac{1}{2}$. Now let $\mathcal{U}$ be the open box $\mathcal{U}=\mathcal{I}_1\times\cdots\times\mathcal{I}_d$. Then $\mathrm{vol}(\mathcal{U})>\frac{1}{8}|g_{ij}|\cdot(\frac{1}{4}\varepsilon_0)^{d-2}>dr$, where we used the first part of our assumption (4.9). Note also that $|\mathcal{I}_k|<\frac{3}{4}$ for all k. Hence by Lemma 2.5, $\mathcal{U}\cap\mathfrak{C}_{g\mathbb{Z}^d,r}\neq\varnothing$, i.e. there exists some $\boldsymbol{v}\in g\mathbb{Z}^d$ such that

$$\mathcal{U}\cap(\boldsymbol{v}+\tfrac{1}{2}\mathcal{C}_r)\neq\varnothing.$$

It follows from the disjointness relations in (4.11) and (4.12) that $\mathcal{U}$ is disjoint from the three cubes $g\boldsymbol{e}_i+\frac{1}{2}\mathcal{C}_r$, $g\boldsymbol{e}_j+\frac{1}{2}\mathcal{C}_r$ and $\frac{1}{2}\mathcal{C}_r$; hence $\boldsymbol{v}\notin\{\boldsymbol{0},g\boldsymbol{e}_j,g\boldsymbol{e}_i\}$. Let us fix a point $\boldsymbol{y}\in\mathcal{U}\cap(\boldsymbol{v}+\frac{1}{2}\mathcal{C}_r)$. Then the line $\boldsymbol{y}+\mathbb{R}\boldsymbol{e}_i$ goes through both the cubes $\boldsymbol{v}+\frac{1}{2}\mathcal{C}_r$ and $g\boldsymbol{e}_j+\frac{1}{2}\mathcal{C}_r$ (the latter holds since $\mathcal{I}_k\subset g_{kj}+\mathcal{J}_r$ for all $k\neq i$; see (4.10) and (4.12)); hence since these two cubes are disjoint, we must have $v_i\ge g_{ij}+1-r$. It also follows from $y_i\in v_i+\mathcal{J}_r$ and $y_i\in\mathcal{I}_i$ that $v_i<y_i+\frac{1}{2}(1-r)<1-r$. In summary:

$$g_{ij}+1-r\le v_i<1-r.\tag{4.13}$$

Similarly, using the fact that the line $\boldsymbol{y}+\mathbb{R}\boldsymbol{e}_j$ goes through the two disjoint cubes $\boldsymbol{v}+\frac{1}{2}\mathcal{C}_r$ and $\frac{1}{2}\mathcal{C}_r$, and also using $y_j\in v_j+\mathcal{J}_r$ and $y_j\in\mathcal{I}_j$, it follows that

$$1-r\le v_j<y_j+\tfrac{1}{2}(1-r)<\tfrac{5}{4}.\tag{4.14}$$

Next, for each $k \notin \{i,j\}$, since $y_k \in \mathcal{I}_k$ and $y_k \in v_k + \mathcal{J}_r$, we have that both of the intervals $(y_k - \frac{1}{8}\varepsilon_0, y_k]$ and $[y_k, y_k + \frac{1}{8}\varepsilon_0)$ are contained in $\widetilde{\mathcal{I}}_k$, and at least one of them is contained in $v_k + \mathcal{J}_r$; hence there exists an open subinterval $\mathcal{I}'_k$ of $\widetilde{\mathcal{I}}_k \cap (v_k + \mathcal{J}_r)$ of length $\frac{1}{8}\varepsilon_0$. Let us also set $\mathcal{I}'_i = (\frac{2}{3}, \frac{3}{4})$ and $\mathcal{I}'_j = \big(g_{ji} + \frac{1}{2}(1-r), v_j - \frac{1}{2}(1-r)\big)$, and then let $\mathcal{U}'$ be the open box $\mathcal{U}' = \mathcal{I}'_1 \times \cdots \times \mathcal{I}'_d$. Using (4.14) we have $|\mathcal{I}'_j| \geq -g_{ji}$; hence $\mathrm{vol}(\mathcal{U}') \geq (\frac{1}{8}\varepsilon_0)^{d-2} \cdot \frac{1}{12} \cdot |g_{ji}| > dr$, where we used the second part of our assumption (4.9). Note also that $|\mathcal{I}'_k| < \frac{3}{4}$ for all k; for $k = j$ this uses $v_j < \frac{5}{4}$ (see (4.14)) and $g_{ji} > -\varepsilon$ (see (4.8)). Hence, by Lemma 2.5, $\mathcal{U}' \cap \mathfrak{C}_{g\mathbb{Z}^d, r} \neq \varnothing$, i.e. there exists some $\boldsymbol{v}' \in g\mathbb{Z}^d$ such that

$$\mathcal{U}' \cap (\boldsymbol{v}' + \tfrac{1}{2}\mathcal{C}_r) \neq \varnothing.$$

Choose a point $\boldsymbol{z} \in \mathcal{U}' \cap (\boldsymbol{v}' + \frac{1}{2}\mathcal{C}_r)$. Note that for every $k \notin \{i,j\}$ we have $\mathcal{I}'_k \subset (g_{ki} + \mathcal{J}_r) \cap (v_k + \mathcal{J}_r)$ by construction. Furthermore, using $1 - r \leq g_{ii} < 1 + \varepsilon$ we have $\mathcal{I}'_i \subset g_{ii} + \mathcal{J}_r$, and using (4.13) we have $\mathcal{I}'_i \subset v_i + \mathcal{J}_r$. Hence the line $\boldsymbol{z} + \mathbb{R}\boldsymbol{e}_j$ goes through of the cubes $g\boldsymbol{e}_i + \frac{1}{2}\mathcal{C}_r$ and $\boldsymbol{v} + \frac{1}{2}\mathcal{C}_r$. Of course this line also goes through the cube $\boldsymbol{v}' + \frac{1}{2}\mathcal{C}_r$. Note also that $z_j \in \mathcal{I}'_j$, and by construction, $\mathcal{I}'_j$ is disjoint from and lies between the two intervals $g_{ji} + \mathcal{J}_r$ and $v_j + \mathcal{J}_r$. Hence we must have $\boldsymbol{v}' \notin \{g\boldsymbol{e}_i, \boldsymbol{v}\}$; thus the three cubes $g\boldsymbol{e}_i + \frac{1}{2}\mathcal{C}_r$, $\boldsymbol{v} + \frac{1}{2}\mathcal{C}_r$ and $\boldsymbol{v}' + \frac{1}{2}\mathcal{C}_r$ are pairwise disjoint, and the two intervals $g_{ji} + \mathcal{J}_r$ and $v_j + \mathcal{J}_r$ must lie at a distance $\geq 1 - r$ from each other. However, this is impossible, since $v_j - g_{ji} < \frac{5}{4} + \varepsilon < 2(1-r)$. This completes the proof in Case II.

Case III: $g_{ij}g_{ji} < 0$. If $g_{ji} < 0$ then let us swap the values of i and j; thus from now on we have $g_{ji} > 0$ and $g_{ij} < 0$, but either $i < j$ or $i > j$. If $g_{ji} \leq g_{jj} - (1-r)$, then $g_{ji} \ll r$ by (4.6), and so (4.7) holds for our i, j. Hence from now on we may assume $g_{ji} > g_{jj} - (1-r)$. Now set:

$$\begin{aligned}
\mathcal{I}_i &= \big(g_{ij} + \tfrac{1}{2}(1-r), g_{ii} - \tfrac{1}{2}(1-r)\big);\\
\mathcal{I}_j &= \big(g_{jj} - \tfrac{1}{2}(1-r), g_{ji} + \tfrac{1}{2}(1-r)\big);\\
\text{and } \widetilde{\mathcal{I}}_k &= (g_{ki} + \mathcal{J}_r) \cap (g_{kj} + \mathcal{J}_r) \quad \text{for } k \notin \{i,j\}.
\end{aligned}$$

These are non-empty intervals. Indeed, $\mathcal{I}_i$ is non-empty since $g_{ii} \geq 1 - r$ and $g_{ij} < 0$; $\mathcal{I}_j$ is non-empty because of our assumption $g_{ji} > g_{jj} - (1-r)$, and for each $k \notin \{i,j\}$, it follows from $|g_{ki} - g_{kj}| < 1 - \varepsilon$ (see (4.8)) that $\widetilde{\mathcal{I}}_k$ is non-empty with $|\widetilde{\mathcal{I}}_k| > (1-r) - (1-\varepsilon) \geq \frac{1}{2}\varepsilon_0$. Now for each $k \notin \{i,j\}$ we choose an open subinterval $\mathcal{I}_k$ of $\widetilde{\mathcal{I}}_k$ of length $\min\big\{\frac{3}{4}, |\widetilde{\mathcal{I}}_k|\big\}$, and then define $\mathcal{U}$ to be the open box $\mathcal{U} = \mathcal{I}_1 \times \cdots \times \mathcal{I}_d$. We claim that

$$\mathcal{U} \cap \mathfrak{C}_{g\mathbb{Z}^d, r} = \varnothing. \tag{4.15}$$

Indeed, assume the opposite; then there is some $\boldsymbol{v} \in g\mathbb{Z}^d$ with $\mathcal{U} \cap \big(\boldsymbol{v} + \frac{1}{2}\mathcal{C}_r\big) \neq \varnothing$. By construction, $\mathcal{I}_i$ is disjoint from both the intervals $g_{ij} + \mathcal{J}_r$ and $g_{ii} + \mathcal{J}_r$; hence $\mathcal{U}$ is disjoint from the two cubes $g\boldsymbol{e}_j + \frac{1}{2}\mathcal{C}_r$, $g\boldsymbol{e}_i + \frac{1}{2}\mathcal{C}_r$, and thus $\boldsymbol{v} \notin \{g\boldsymbol{e}_j, g\boldsymbol{e}_i\}$. Let $\boldsymbol{x}$ be a point in $\mathcal{U} \cap \big(\boldsymbol{v} + \frac{1}{2}\mathcal{C}_r\big)$. Then $x_i \in \mathcal{I}_i \cap (v_i + \mathcal{J}_r)$. Note also that the three intervals $g_{ij} + \mathcal{J}_r$, $\mathcal{I}_i$, $g_{ii} + \mathcal{J}_r$ are adjacent to each other in this order along the real line, with the length of $\mathcal{I}_i$ being

$$|\mathcal{I}_i| = g_{ii} - g_{ij} - (1-r) < 1 + \varepsilon + (\tfrac{1}{2} - 3\varepsilon) - (1-r) = \tfrac{1}{2} - 2\varepsilon + r < \tfrac{1}{2} < 1 - r.$$

But $v_i + \mathcal{J}_r$ has length $1 - r$; hence there exists a number $x'_i \in v_i + \mathcal{J}_r$ lying either in $g_{ii} + \mathcal{J}_r$ or $g_{ij} + \mathcal{J}_r$. Noticing also that $x_k \in \mathcal{I}_k \subset (g_{ki} + \mathcal{J}_r) \cap (g_{kj} + \mathcal{J}_r)$ for all $k \neq i$, it now follows that the point $\boldsymbol{x} + (x'_i - x_i)\boldsymbol{e}_i$ lies in the cube $\boldsymbol{v} + \frac{1}{2}\mathcal{C}_r$ and also in one of the two cubes $g\boldsymbol{e}_i + \frac{1}{2}\mathcal{C}_r$ or $g\boldsymbol{e}_j + \frac{1}{2}\mathcal{C}_r$. This is a contradiction against the fact that $\boldsymbol{v} + \frac{1}{2}\mathcal{C}_r$ is disjoint from both $g\boldsymbol{e}_i + \frac{1}{2}\mathcal{C}_r$ and $g\boldsymbol{e}_j + \frac{1}{2}\mathcal{C}_r$; hence we have completed the proof of (4.15).

We have $|\mathcal{I}_k| < \frac{3}{4}$ for all $1 \le k \le d$; hence Lemma 2.5 applies, giving $\mathrm{vol}(\mathcal{U}) < dr$. But $|\mathcal{I}_k| \ge \frac{1}{2}\varepsilon_0$ for all $k \notin \{i,j\}$, and $|\mathcal{I}_i| \ge |g_{ij}|$; thus

$$|g_{ij}|\big(g_{ji} - g_{jj} + 1 - r\big) = |g_{ij}||\mathcal{I}_j| \ll \mathrm{vol}(\mathcal{U}) \ll r.$$

Furthermore, $|g_{ij}|\big(g_{jj} - (1-r)\big) \ll r$ by (4.6). Adding the last two bounds, we conclude that (4.7) holds for our i, j. □

4.4. **Proof of Proposition 4.2.** Finally, we can give the

Proof of Proposition 4.2. Choose $0 < a < 1$ and $A > 1$ as in Lemma 4.5. Fix a number $0 < \beta < \frac{1}{16}$ so small that for every matrix $g \in G$ which has distance (w.r.t. $\|\cdot\|$) less than β to a matrix in U_0, we have $|\det \tilde{g} - 1| < \frac{1}{2}$, where $\tilde{g}$ is the top left $(d-1) \times (d-1)$ block of g. Fix a number $0 < \varepsilon_0 < \min\big\{a, \beta/(2A)\big\}$, and for this ε_0, take $r_0 > 0$ as in Lemma 4.5. Note that the defining property of r_0 trivially remains valid if we decrease r_0; hence we may assume that $0 < r_0 < \frac{1}{2}\varepsilon_0$. Let $C > 0$ be the maximum of the implied constants in the two "$\ll$" bounds in Lemma 4.6, for our fixed d and ε_0.

Now let $r \in (0, r_0)$ and $\Lambda \in K_r$ be given. This means that $\Lambda \cap \mathcal{C}_r = \{\mathbf{0}\}$, and a fortiori, $\Lambda \cap \mathcal{C}_{r_0} = \{\mathbf{0}\}$. Hence by our choice of A, a, ε_0, r_0 (see the statement of Lemma 4.5), there exist $g' \in G$, $w \in W$, $B \in [1, A]$ and $u \in U_{B\varepsilon_0}$ such that $\Lambda = g'\mathbb{Z}^d$ and $\|g' - wuw^{-1}\| < \varepsilon_0$. Note that the norm $\|\cdot\|$ is preserved by left and right multiplication by elements from W; hence letting $g = (g_{ij}) := w^{-1}g'w$ we have $\|g - u\| < \varepsilon_0$, and also $g\mathbb{Z}^d \cap \mathcal{C}_r = w^{-1}(g'\mathbb{Z}^d \cap \mathcal{C}_r) = \{\mathbf{0}\}$ (this is true since $w\mathbb{Z}^d = \mathbb{Z}^d$ and $\mathcal{C}_r = w^{-1}\mathcal{C}_r$). Hence by Lemma 4.6, and by our choice of C, we have $0 \le g_{ii} - (1-r) \le Cr$ for all i and $|g_{ij}g_{ji}| \le Cr$ for all $i \ne j$. Furthermore, it follows from $u \in U_{B\varepsilon_0}$ and $\|g - u\| < \varepsilon_0$ that $\|g - u'\| < (B+1)\varepsilon_0$ for some $u' \in U_0$; hence a fortiori $\|g - u'\| < 2A\varepsilon_0 < \beta$, which implies that $|\det \tilde{g} - 1| < \frac{1}{2}$ by our choice of β. It also follows that $|g_{ij}| < \frac{1}{2} + 2A\varepsilon_0 < 1$ for all $i \ne j$. Hence $g \in \overline{\mathcal{K}}_{r,C}$, and thus $g' = wgw^{-1} \in w\,\overline{\mathcal{K}}_{r,C}\,w^{-1}$ and $\Lambda = g'\mathbb{Z}^d \in \pi(w\,\overline{\mathcal{K}}_{r,C}\,w^{-1})$, finishing the proof. □

5. Measure estimates of the thickenings

Fix $m, n \in \mathbb{N}$ and let $d = m + n$. Let $\boldsymbol{\alpha} \in \mathbb{R}^m$ and $\boldsymbol{\beta} \in \mathbb{R}^n$ be two fixed weight vectors as in Theorem 1.2. As mentioned in Remark 9, in order to incorporate the case of general weights, we need to consider a more general one-parameter subgroup of G associated to $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. Explicitly, for any $s \in \mathbb{R}$ let us define

$$g_s = g_s^{\boldsymbol{\alpha},\boldsymbol{\beta}} := \mathrm{diag}(e^{\alpha_1 s}, \dots, e^{\alpha_m s}, e^{-\beta_1 s}, \dots, e^{-\beta_n s}) \in G. \tag{5.1}$$

Let $\Delta : X_d \to [0, \infty)$ be the function defined in (1.4). The main result of this section is an asymptotic estimate for the measure of the thickened set

$$\widetilde{\Delta}_r := \bigcup_{0 \le s < 1} g_{-s}\,\Delta^{-1}[0, r],$$

when $r > 0$ is small.

Theorem 5.1. *Let $\varkappa_d = \frac{d^2+d-4}{2}$ and $\lambda_d = \frac{d(d-1)}{2}$ be as in Theorem 1.2. Then*

$$\mu_d\big(\widetilde{\Delta}_r\big) \asymp_d r^{\varkappa_d} \log^{\lambda_d}\Big(\frac{1}{r}\Big), \qquad \text{as } r \to 0^+,$$

where the implicit constant is independent of r and the two weight vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.

Just as for Theorem 2.1, we prove Theorem 5.1 by proving the upper and lower bounds separately.

5.1. **Proof of the upper bound.** Note that for any $(\boldsymbol{x},\boldsymbol{y})\in\mathbb{R}^m\times\mathbb{R}^n$,

$$g_{-s}\left(\begin{smallmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{smallmatrix}\right)=(e^{-\alpha_1 s}x_1,\dots,e^{-\alpha_m s}x_m,e^{\beta_1 s}y_1,\dots,e^{\beta_n s}y_n)^t.$$

This implies that

$$|\Delta(g_{-s}\Lambda)-\Delta(\Lambda)|\le\max\{\alpha_1,\dots,\alpha_m,\beta_1,\dots,\beta_n\}|s|<|s|,\qquad\forall\ s\in\mathbb{R},\ \Lambda\in X_d.$$

Hence

$$g_{-s}\Delta^{-1}\big[0,r\big]\subset\Delta^{-1}\big[0,r+|s|\big],\qquad\forall\ s\in\mathbb{R}. \tag{5.2}$$

Given any $r\in(0,1)$, let $q=\lceil 1/r\rceil$. Using (5.2) and the fact that $1/q\le r$, we have

$$\widetilde{\Delta}_r=\bigcup_{0\le s<1}g_{-s}\Delta^{-1}[0,r]=\bigcup_{k=0}^{q-1}\bigcup_{0\le s<1/q}g_{-k/q}\,g_{-s}\,\Delta^{-1}[0,r]\subset\bigcup_{k=0}^{q-1}g_{-k/q}\,\Delta^{-1}[0,2r],$$

implying that (using the G-invariance of μ_d and $q\asymp r^{-1}$)

$$\mu_d\big(\widetilde{\Delta}_r\big)\le\sum_{k=0}^{q-1}\mu_d\big(g_{-k/q}\,\Delta^{-1}[0,2r]\big)\asymp r^{-1}\mu_d\big(\Delta^{-1}[0,2r]\big).$$

Finally, by Theorem 1.3 we get

$$\mu_d\big(\widetilde{\Delta}_r\big)\ll_d r^{-1}r^{\varkappa_d+1}\log^{\lambda_d}\Big(\frac{1}{r}\Big)=r^{\varkappa_d}\log^{\lambda_d}\Big(\frac{1}{r}\Big),\qquad\text{as } r\to0^+.$$

This finishes the proof of the upper bound in Theorem 5.1.

5.2. **Proof of the lower bound.** In this subsection we prove the lower bound in Theorem 5.1. By the discussion in the beginning of Section 2, we may replace the set $\Delta^{-1}[0,r]$ by K_r, that is, it suffices to prove the following lower bound

$$\mu_d\bigg(\bigcup_{0\le s<1}g_{-s}\,K_r\bigg)\gg_d r^{\varkappa_d}\log^{\lambda_d}\Big(\frac{1}{r}\Big),\qquad\text{as } r\to0^+. \tag{5.3}$$

The following lemma is the crucial ingredient in our proof of (5.3). Let c_0 be the small parameter which we fixed in Section 2.2; after possibly shrinking c_0, we may without loss of generality assume that $0<c_0<(3e)^{-1}$. For $r\in(0,c_0/d)$, let $\underline{K}_r\subset K_r$ be as in (3.18).

Lemma 5.2. *For any $r\in(0,c_0/d)$ and $s\in[r,1)$, the two sets $\underline{K}_r$ and $g_{-s}\,\underline{K}_r$ are disjoint.*

Proof. Assume the opposite; then there exist $r\in(0,c_0/d)$, $s\in[r,1)$ and $\Lambda\in\underline{K}_r$ such that $g_{-s}\Lambda\in\underline{K}_r$. By the definition of $\underline{K}_r$ in (3.18), we now have $\Lambda=p_{\boldsymbol{b}_1,\dots,\boldsymbol{b}_{d-1}}u_{\boldsymbol{x}}\mathbb{Z}^d$ for some vectors $\boldsymbol{b}_j=(b_{1j},\dots,b_{dj})^t\in\mathbb{R}^d$ $(j=1,\dots,d-1)$ and $\boldsymbol{x}\in(0,c_0/d)^{d-1}$ satisfying (3.10), (3.11), (3.12) and (3.13); moreover, we also have $g_{-s}\Lambda=p_{\boldsymbol{b}'_1,\dots,\boldsymbol{b}'_{d-1}}u_{\boldsymbol{x}'}\mathbb{Z}^d$ for some vectors $\boldsymbol{b}'_j=(b'_{1j},\dots,b'_{dj})^t\in\mathbb{R}^d$ $(j=1,\dots,d-1)$ and $\boldsymbol{x}'\in(0,c_0/d)^{d-1}$ which again satisfy (3.10), (3.11), (3.12) and (3.13).

Because of $\boldsymbol{\alpha}\in(\mathbb{R}_{>0})^m$ and $\sum_{i=1}^m\alpha_i=1$, there exists an index $1\le i\le m$ such that $\alpha_i\ge\frac{1}{m}$. Fixing such an i, we consider the vector

$$\boldsymbol{y}=(y_1,\dots,y_d)^t:=g_{-s}\boldsymbol{b}_i=\big(e^{-\alpha_1 s}b_{1i},\dots,e^{-\alpha_m s}b_{mi},e^{\beta_1 s}b_{m+1,i},\dots,e^{\beta_n s}b_{d,i}\big)^t.$$

By (3.10) we have $|b_{ji}|<c_0$ for all $j\ne i$, and $1-\frac{r}{2d}<b_{ii}<1$. Hence $|y_j|<ec_0$ for all $j\ne i$ (since $0<s<1$ and $\beta_\ell\le1$ for all $1\le\ell\le n$), and $0<y_i<e^{-\alpha_i s}<e^{-s/m}<1-\frac{s}{2m}<1-\frac{r}{2d}$. We have $\boldsymbol{b}_i\in\Lambda\smallsetminus\{\boldsymbol{0}\}$ and thus $\boldsymbol{y}=g_{-s}\boldsymbol{b}_i\in g_{-s}\Lambda\smallsetminus\{\boldsymbol{0}\}$; also $g_{-s}\Lambda\in\underline{K}_r\subset K_r$, and hence $\boldsymbol{y}\notin\mathcal{C}_r$. But for

all $j \neq i$ we have $|y_j| < ec_0 < \frac{1}{3} < 1 - r$ (indeed, recall that $0 < c_0 < (3e)^{-1}$); also $y_i > 0$; hence $\boldsymbol{y} \notin \mathcal{C}_r$ implies $y_i \geq 1 - r$. In summary:

$$|y_j| < ec_0 \quad (\forall\, j \neq i) \qquad \text{and} \quad 1 - r \leq y_i < 1 - \tfrac{r}{2d}.$$

Furthermore, $\boldsymbol{b}'_i \in g_{-s}\Lambda$, since $g_{-s}\Lambda = p_{\boldsymbol{b}'_1,\ldots,\boldsymbol{b}'_{d-1}} u_{\boldsymbol{x}'} \mathbb{Z}^d$; and by (3.10) we have $|b'_{ji}| < c_0$ for all $j \neq i$ and $1 - \frac{r}{2d} < b'_{ii} < 1$. It follows that $|b'_{ji} - y_j| < (e+1)c_0 < (e+1)(3e)^{-1} < \frac{1}{2} < 1 - r$ for all $j \neq i$, and $|b'_{ii} - y_i| < r < 1 - r$; hence $\boldsymbol{b}'_i - \boldsymbol{y} \in \mathcal{C}_r$. Note also that $\boldsymbol{b}'_i - \boldsymbol{y} \in g_{-s}\Lambda$, and $\boldsymbol{b}'_i - \boldsymbol{y} \neq \boldsymbol{0}$, since $y_i < 1 - \frac{r}{2d} < b'_{ii}$; hence we have obtained a contradiction against $g_{-s}\Lambda \in \underline{K}_r \subset K_r$. This completes the proof of the lemma. □

It follows from Lemma 5.2 that for any $r \in (0, c_0/d)$, the sets $g_{-kr}\,\underline{K}_r$, for k running through the integers in the interval $0 \leq k < 1/r$, are pairwise disjoint. (Indeed, if $g_{-kr}\,\underline{K}_r \cap g_{-k'r}\,\underline{K}_r \neq \varnothing$ for some $0 \leq k < k' < 1/r$ then $\underline{K}_r \cap g_{(k-k')r}\,\underline{K}_r \neq \varnothing$, contradicting Lemma 5.2.) Hence, using also $K_r \supset \underline{K}_r$, we have

$$\mu_d\Bigg(\bigcup_{0\leq s<1} g_{-s}\,K_r\Bigg) \geq \mu_d\Bigg(\bigcup_{0\leq k<1/r} g_{-kr}\,\underline{K}_r\Bigg) = \sum_{0\leq k<1/r} \mu_d\big(g_{-kr}\,\underline{K}_r\big) = \#\big(\mathbb{Z}\cap[0,1/r)\big)\cdot\mu_d\big(\underline{K}_r\big),$$

Here $\#\big(\mathbb{Z}\cap[0,1/r)\big) \gg r^{-1}$, and for r sufficiently small we have $\mu_d\big(\underline{K}_r\big) \gg r^{\varkappa_d+1}\log^{\lambda_d}\big(\frac{1}{r}\big)$ by (3.21). Hence we obtain the lower bound (5.3), and the proof of Theorem 5.1 is complete. □

6. Some preliminaries for Theorem 1.2

In this section we collect some preliminary results for our proof of Theorem 1.2.

6.1. **Dynamical interpretation of weighted ψ-Dirichlet matrices.** Let $m, n \in \mathbb{N}$ and let $\boldsymbol{\alpha} \in \mathbb{R}^m$ and $\boldsymbol{\beta} \in \mathbb{R}^n$ be two fixed weight vectors as before. Let $t_0 > 0$ and let $\psi : [t_0, \infty) \to (0, \infty)$ be a continuous decreasing function which tends to zero at infinity. In this subsection we give a dynamical interpretation of $\psi_{\boldsymbol{\alpha},\boldsymbol{\beta}}$-Dirichlet matrices which generalizes [25, Proposition 4.5]; see Proposition 6.2. Let us first introduce the following modified Dani Correspondence which is a special case of [19, Lemma 8.3].

Lemma 6.1. *Fix $m, n \in \mathbb{N}$ and let $d = m + n$. Let $t_0 > 0$, and let $\psi : [t_0, \infty) \to (0, \infty)$ be a continuous, decreasing function satisfying (1.8) and (1.9). Then there exists a unique continuous, decreasing function*

$$r = r_\psi : [s_0, \infty) \to (0, \infty), \quad \text{where } s_0 = \frac{m}{d}\log t_0 - \frac{n}{d}\log\psi(t_0),$$

such that

(6.1) *the function $s \mapsto s + mr(s)$ is increasing,*

and

(6.2) $$\psi\left(e^{s-nr(s)}\right) = e^{-s-mr(s)} \quad \text{for all } s \geq s_0.$$

Conversely, given $s_0 \in \mathbb{R}$ and a continuous, decreasing function $r : [s_0, \infty) \to (0, \infty)$ satisfying (6.1), there exists a unique continuous, decreasing function $\psi = \psi_r : [t_0, \infty) \to (0, \infty)$ with $t_0 = e^{s_0 - nr(s_0)}$ satisfying (1.8), (1.9) and (6.2). Furthermore, for any fixed $\alpha, \beta > 0$ the series

(6.3) $$\sum_{k\geq t_0} \frac{\big(1 - k\psi(k)\big)^\alpha \big(-\log\big(1 - k\psi(k)\big)\big)^\beta}{k}$$

diverges if and only if the series

$$\sum_{k\geq s_0} r(k)^\alpha \log^\beta\Big(1+\frac{1}{r(k)}\Big) \tag{6.4}$$

diverges.

Proof. The ψ and r-functions determine each other uniquely via the relation

$$\psi(t)^{1/m}e^{s/m} = t^{1/n}e^{-s/n} = e^{-r(s)}, \tag{6.5}$$

which captures the moment when the a_s-flow transforms the long and thin 'rectangle'

$$\big\{(\boldsymbol{x},\boldsymbol{y})\in\mathbb{R}^m\times\mathbb{R}^n \,:\, \|\boldsymbol{x}\|^m<\psi(t),\ \|\boldsymbol{y}\|^n<t\big\}$$

determined by (1.2) into a cube (with side length $2e^{-r(s)}$). Here $a_s=\operatorname{diag}(e^{s/m}I_m, e^{-s/n}I_n)$, as defined in the introduction. This correspondence between $\psi(\cdot)$ and $r(\cdot)$ is a special case of [19, Lemma 8.3], as here we assume that $\psi(\cdot)$ additionally satisfies (1.8) and (1.9), which on the r-function side corresponds respectively to the assumptions that $r(\cdot)$ is decreasing and $r(s)>0$ for all $s\geq s_0$. The equivalence of these additional assumptions is easily checked using the following three relations, which follow from (6.5):

$$e^{-dr(s)}=t\psi(t), \qquad s=\frac{m}{d}\log t-\frac{n}{d}\log\psi(t), \qquad \text{and} \qquad t=e^{s-nr(s)}. \tag{6.6}$$

Finally we prove the equivalence of the divergence of the two series. If $\lim_{t\to\infty}t\psi(t)<1$ then both the functions $1-t\psi(t)$ and $r(s)$ are bounded away from zero (and positive), which implies that the two series in (6.3) and (6.4) are divergent. Hence from now on we may assume that $\lim_{t\to\infty}t\psi(t)=1$. Then $\lim_{s\to\infty}r(s)=0$ (by (6.6)), and $F_\psi(t):=1-t\psi(t)$ is a decreasing function taking values in the interval $(0,1)$ and satisfying $\lim_{t\to\infty}F_\psi(t)=0$. After enlarging s_0 (thus also enlarging t_0) we may assume that $0<r(s)<1/d$ for all $s\geq s_0$. Then by (6.6) we have, with $t=t(s)=e^{s-nr(s)}$:

$$\frac{d}{2}r(s)<F_\psi(t)<d\,r(s) \quad\text{and}\quad e^{s-1}<t<e^s, \qquad \forall\ s\geq s_0. \tag{6.7}$$

It follows that $r(s)^\alpha\log^\beta\big(1+\frac{1}{r(s)}\big)\asymp_{d,\alpha,\beta} F_\psi(t)^\alpha\log^\beta\big(\frac{1}{F_\psi(t)}\big)$ for all $s\geq s_0$. Hence, using also $e^{k-1}<t(k)<e^k$ (see (6.7)), the fact that $F_\psi(t)$ is decreasing, and $\sum_{e^k\leq j<e^{k+1}}\frac{1}{j}\asymp 1$ $(\forall k\geq 1)$, we have for all sufficiently large integers k:

$$r(k)^\alpha\log^\beta\Big(1+\frac{1}{r(k)}\Big)\ll_{d,\alpha,\beta} F_\psi(e^{k-1})^\alpha\log^\beta\Big(\frac{1}{F_\psi(e^{k-1})}\Big)\ll_{\alpha,\beta}\sum_{e^{k-2}\leq j<e^{k-1}}\frac{1}{j}F_\psi(j)^\alpha\log^\beta\Big(\frac{1}{F_\psi(j)}\Big),$$

and similarly

$$r(k)^\alpha\log^\beta\Big(1+\frac{1}{r(k)}\Big)\gg_{d,\alpha,\beta} F_\psi(e^{k})^\alpha\log^\beta\Big(\frac{1}{F_\psi(e^{k})}\Big)\gg_{\alpha,\beta}\sum_{e^{k}\leq j<e^{k+1}}\frac{1}{j}F_\psi(j)^\alpha\log^\beta\Big(\frac{1}{F_\psi(j)}\Big).$$

It follows that the series in (6.4) diverges if and only if $\sum_j j^{-1}F_\psi(j)^\alpha\log^\beta\big(\frac{1}{F_\psi(j)}\big)$ diverges, that is, if and only if the series in (6.3) diverges. □

Remark 12. Let ψ and r be as in Lemma 6.1 with $\lim_{t\to\infty}t\psi(t)=1$. Let $F_\psi(t)=1-t\psi(t)$ be as above. Assume that the series (6.3) (and thus also the series (6.4)) diverges for some $\alpha,\beta>0$. It is then not difficult to see from the proof of Lemma 6.1 that for any $\beta'\geq\beta$ and for all large $s_1>s_0$,

$$\sum_{s_0\leq k\leq s_1} r(k)^\alpha\log^{\beta'}\Big(1+\frac{1}{r(k)}\Big)\asymp_{d,\alpha,\beta'}\sum_{t_0\leq k\leq e^{s_1}} k^{-1}F_\psi(k)^\alpha\log^{\beta'}\Big(\frac{1}{F_\psi(k)}\Big).$$

In particular it follows that for any $\beta' > \beta$ we have the following equivalence:

$$\liminf_{s_1\to\infty} \frac{\sum_{s_0<k\le s_1} r(k)^\alpha \log^{\beta'}\big(1+\frac{1}{r(k)}\big)}{\Big(\sum_{s_0<k\le s_1} r(k)^\alpha \log^{\beta}\big(1+\frac{1}{r(k)}\big)\Big)^2} = 0$$

$$\iff \quad \liminf_{t_1\to\infty} \frac{\sum_{t_0\le k\le t_1} k^{-1} F_\psi(k)^\alpha \log^{\beta'}\left(\frac{1}{F_\psi(k)}\right)}{\Big(\sum_{t_0\le k\le t_1} k^{-1} F_\psi(k)^\alpha \log^{\beta}\left(\frac{1}{F_\psi(k)}\right)\Big)^2} = 0.$$

Similarly, the above two limits inferior remain bounded simultaneously.

We now state the dynamical interpretation of $\psi_{\boldsymbol{\alpha},\boldsymbol{\beta}}$-Dirichlet matrices.

Proposition 6.2. *Let ψ be as in Theorem 1.2, and let $r = r_\psi$ be as in Lemma 6.1. Let $\{g_s\}_{s\in\mathbb{R}}$ be the one-parameter subgroup associated to the two fixed weight vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ as in (5.1). Set*

$$\omega_1 := \max\{m\alpha_i, n\beta_j \,:\, 1\le i\le m, 1\le j\le n\} \quad \text{and} \quad \omega_2 := \min\{m\alpha_i, n\beta_j \,:\, 1\le i\le m, 1\le j\le n\}.$$

Then for any $A \in M_{m,n}(\mathbb{R})$ we have, with Λ_A as in (1.5):

1. *if $\Delta(g_s\Lambda_A) > \omega_1 r(s)$ for all sufficiently large s, then A is $\psi_{\boldsymbol{\alpha},\boldsymbol{\beta}}$-Dirichlet;*
2. *if $\Delta(g_s\Lambda_A) \le \omega_2 r(s)$ for an unbounded set of s, then A is not $\psi_{\boldsymbol{\alpha},\boldsymbol{\beta}}$-Dirichlet.*

Remark 13. When $\boldsymbol{\alpha} = (\frac{1}{m},\dots,\frac{1}{m}) \in \mathbb{R}^m$ and $\boldsymbol{\beta} = (\frac{1}{n},\dots,\frac{1}{n}) \in \mathbb{R}^n$, then $\omega_1 = \omega_2 = 1$ and Proposition 6.2 recovers [25, Proposition 4.5].

Proof of Proposition 6.2. For any $t > \max\{t_0, 1\}$, define

$$\mathcal{R}_t = \mathcal{R}_t^{\boldsymbol{\alpha},\boldsymbol{\beta}} := \big\{(\boldsymbol{x},\boldsymbol{y}) \in \mathbb{R}^m\times\mathbb{R}^n \,:\, \|\boldsymbol{x}\|_{\boldsymbol{\alpha}} < \psi(t), \|\boldsymbol{y}\|_{\boldsymbol{\beta}} < t\big\},$$

so that $(\boldsymbol{p},\boldsymbol{q}) \in \mathbb{Z}^m\times(\mathbb{Z}^n\smallsetminus\{\boldsymbol{0}\})$ is a solution to (1.7) if and only if $(A\boldsymbol{q}-\boldsymbol{p},\boldsymbol{q}) \in \mathcal{R}_t$. On the other hand, the lattice Λ_A consists exactly of the points

$$\begin{pmatrix} I_m & A \\ 0 & I_n \end{pmatrix}\begin{pmatrix} -\boldsymbol{p} \\ \boldsymbol{q}\end{pmatrix} = \begin{pmatrix} A\boldsymbol{q}-\boldsymbol{p} \\ \boldsymbol{q}\end{pmatrix} \qquad \text{for } (\boldsymbol{p},\boldsymbol{q}) \in \mathbb{Z}^m\times\mathbb{Z}^n.$$

Moreover, if $(A\boldsymbol{q}-\boldsymbol{p},\boldsymbol{q}) \in \Lambda_A\cap\mathcal{R}_t$ is nonzero for some $(\boldsymbol{p},\boldsymbol{q}) \in \mathbb{Z}^m\times\mathbb{Z}^n$, then we must have $\boldsymbol{q}\neq\boldsymbol{0}$. Indeed, otherwise we would have $\|A\boldsymbol{q}-\boldsymbol{p}\|_{\boldsymbol{\alpha}} = \|\boldsymbol{p}\|_{\boldsymbol{\alpha}} \ge 1$, but $(A\boldsymbol{q}-\boldsymbol{p},\boldsymbol{q}) \in \mathcal{R}_t$ implies that $\|A\boldsymbol{q}-\boldsymbol{p}\|_{\boldsymbol{\alpha}} < \psi(t) < 1/t < 1$ (since $t > \max\{t_0,1\}$), contradicting $\|A\boldsymbol{q}-\boldsymbol{p}\|_{\boldsymbol{\alpha}} \ge 1$. Thus there exists a solution $(\boldsymbol{p},\boldsymbol{q}) \in \mathbb{Z}^m\times(\mathbb{Z}^n\smallsetminus\{\boldsymbol{0}\})$ to (1.7) if and only if $\Lambda_A\cap\mathcal{R}_t \neq \{\boldsymbol{0}\}$, implying that $A \in M_{m,n}(\mathbb{R})$ is $\psi_{\boldsymbol{\alpha},\boldsymbol{\beta}}$-Dirichlet if and only if $\Lambda_A\cap\mathcal{R}_t\neq\{\boldsymbol{0}\}$ for all sufficiently large t. Now let $s = s(t) = \frac{m}{d}\log t - \frac{n}{d}\log\psi(t)$; then $s\to\infty$ if and only if $t\to\infty$, and by (6.5) we have

$$g_s\mathcal{R}_t = \Big\{(\boldsymbol{x}',\boldsymbol{y}') \in \mathbb{R}^m\times\mathbb{R}^n \,:\, \|\boldsymbol{x}'\|_{\boldsymbol{\alpha}} < e^{-mr(s)},\ \|\boldsymbol{y}'\|_{\boldsymbol{\beta}} < e^{-nr(s)}\Big\} =: \mathcal{E}_s.$$

It follows that A is $\psi_{\boldsymbol{\alpha},\boldsymbol{\beta}}$-Dirichlet if and only if $g_s\Lambda_A\cap\mathcal{E}_s\neq\{\boldsymbol{0}\}$ for all sufficiently large s. Next, note that we have the following simple relation:

$$\big(-e^{-\omega_1 r(s)}, e^{-\omega_1 r(s)}\big)^d \subset \mathcal{E}_s \subset \big(-e^{-\omega_2 r(s)}, e^{-\omega_2 r(s)}\big)^d, \tag{6.8}$$

with ω_1,ω_2 defined as in the statement of the proposition. Note also that $\Delta(g_s\Lambda_A) > \omega_1 r(s)$ is equivalent with $g_s\Lambda_A\cap\big(-e^{-\omega_1 r(s)}, e^{-\omega_1 r(s)}\big)^d \neq \{\boldsymbol{0}\}$. Hence, using the first inclusion relation in (6.8) we have

$$\Delta(g_s\Lambda_A) > \omega_1 r(s) \text{ for all sufficiently large } s \ \Rightarrow g_s\Lambda_A\cap\mathcal{E}_s\neq\{\boldsymbol{0}\} \text{ for all sufficiently large } s,$$

and the latter condition implies that A is $\psi_{\boldsymbol{\alpha},\boldsymbol{\beta}}$-Dirichlet. We have thus proved part (1) of the proposition. Similarly, part (2) follows using the second inclusion relation in (6.8). □

Let ψ and $r = r_\psi$ be as above. For any integer $k > s_0$, let us define

$$\underline{B}_k := \bigcup_{0\le s<1} g_{-s}\Delta^{-1}[0,\omega_2 r(k+s)] \quad \text{and} \quad \overline{B}_k := \bigcup_{0\le s<1} g_{-s}\Delta^{-1}[0,\omega_1 r(k+s)].$$

It follows that for any $\Lambda \in X_d$, we have $g_k\Lambda \in \underline{B}_k$ (respectively $g_k\Lambda \in \overline{B}_k$) if and only if there is some $k \le s < k+1$ such that $g_s\Lambda \in \Delta^{-1}[0,\omega_2 r(s)]$ (respectively $g_s\Lambda \in \Delta^{-1}[0,\omega_1 r(s)]$). In particular, in view of Proposition 6.2, a given matrix $A \in M_{m,n}(\mathbb{R})$ is $\psi_{\boldsymbol{\alpha},\boldsymbol{\beta}}$-Dirichlet if $g_k\Lambda_A \notin \overline{B}_k$ for all sufficiently large k, or equivalently, if $g_k\Lambda_A \in \overline{B}_k$ holds only finitely often. Similarly, A is not $\psi_{\boldsymbol{\alpha},\boldsymbol{\beta}}$-Dirichlet if $g_k\Lambda_A \in \underline{B}_k$ holds infinitely often.

6.2. **Effective equidistribution and doubly mixing for certain g_s-translates.** Let $m, n \in \mathbb{N}$ and $d = m+n$ be as before. Let

$$\mathcal{Y} = \{\Lambda_A : A \in M_{m,n}(\mathbb{R})\} \cong M_{m,n}(\mathbb{R}/\mathbb{Z})$$

be defined as in (1.5), and recall that $\mathcal{Y}$ is equipped with the probability Lebesgue measure, Leb.

As mentioned in Remark 9, we will need an effective equidistribution and doubly mixing result for the g_s-translates $\{g_s\mathcal{Y}\}_{s>0}$ which is analogous to (1.14) and (1.15) respectively. In fact, we will state a corollary of a more general effective mixing result of arbitrary order proved by Björklund and Gorodnik [1, Theorem 2.2]. To state their result, let us first fix some notation.

Let $\mathfrak{g} = \mathfrak{sl}_d(\mathbb{R})$ be the Lie algebra of G. For each $Y \in \mathfrak{g}$, let us denote by $\mathcal{D}_Y$ the corresponding Lie derivative (a first order differential operator) on $C^\infty(G)$ defined by

$$\mathcal{D}_Y(f)(g) := \frac{\mathrm{d}}{\mathrm{d}t} f(\exp(tY)g)\big|_{t=0}, \qquad f \in C^\infty(G).$$

Here $\exp : \mathfrak{g} \to G$ denotes the usual exponential map from $\mathfrak{g}$ to G. Note that this definition naturally extends to the function space $C_c^\infty(X_d)$ since we can view elements in $C_c^\infty(X_d)$ as right Γ-invariant smooth functions on G. Fix an ordered basis $\{Y_1,\dots,Y_a\}$ of $\mathfrak{g}$. Then every monomial $Z = Y_1^{\ell_1}\cdots Y_a^{\ell_a}$ defines a differential operator of degree $\deg(Z) := \ell_1+\cdots+\ell_a$, via

$$\mathcal{D}_Z := \mathcal{D}_{Y_1}^{\ell_1} \circ \cdots \circ \mathcal{D}_{Y_a}^{\ell_a}.$$

Now for each $\ell \in \mathbb{N}$ we define the "L^2, degree ℓ" Sobolev norm on $C_c^\infty(X_d)$ by

$$\|f\|_{L^2_\ell} := \left(\sum_{\deg(Z)\le\ell} \int_{X_d} |\mathcal{D}_Z(f)|^2 \,\mathrm{d}\mu_d\right)^{1/2},$$

where the summation is over all the monomials Z in $\{Y_1,\dots,Y_a\}$ with degree no greater than ℓ. Fix a metric $\mathrm{dist}(\cdot,\cdot)$ on $X_d = G/\Gamma$ which is induced from a right G-invariant Riemannian metric on G. We also define the following Lipschitz (semi-)norm on $C_c^\infty(X_d)$ with respect to this metric:

$$\|f\|_{\mathrm{Lip}} := \sup\left\{\frac{|f(x_1)-f(x_2)|}{\mathrm{dist}(x_1,x_2)} \;:\; x_1,x_2 \in X_d,\ x_1 \ne x_2\right\}, \qquad f \in C_c^\infty(X_d).$$

Let us also write $\|\cdot\|_{C^0}$ for the uniform norm on $C_c(X_d)$. Finally, for any $f \in C_c^\infty(X_d)$ we define

$$\mathcal{N}_\ell(f) := \max\left\{\|f\|_{C^0}, \|f\|_{\mathrm{Lip}}, \|f\|_{L^2_\ell}\right\}.$$

We can now state the result which we need from [1].

Theorem 6.3 ([1, Corollary 2.4]). *There exist $\ell \in \mathbb{N}$ and $\delta > 0$ such that for every $b \in \mathbb{N}$ and any $f_0 \in C^\infty(\mathcal{Y})$, $f_1, \dots, f_b \in C_c^\infty(X_d)$ and $s_1, \dots, s_b > 0$, we have*

$$\int_{\mathcal{Y}} f_0(\Lambda_A) \left(\prod_{i=1}^{b} f_i(g_{s_i}\Lambda_A) \right) \mathrm{d}A = Leb(f_0) \prod_{i=1}^{b} \mu_d(f_i) + O_{b,f_0}\left(e^{-\delta D(s_1,\dots,s_b)} \prod_{i=1}^{b} \mathcal{N}_\ell(f_i) \right),$$

where $D(s_1, \dots, s_b) := \min\{s_i, |s_i - s_j| \ : \ 1 \le i \ne j \le b\}$.

(In fact, in [1, Corollary 2.4], the error term is also explicit in terms of f_0.)

Taking $f_0 \equiv 1$ on $\mathcal{Y}$ and $b = 1, 2$, we get the following effective equidistribution and doubly mixing of the family of g_s-translates $\{g_s\mathcal{Y}\}_{s>0}$ in X_d.

Corollary 6.4. *Let $\ell \in \mathbb{N}$ and $\delta > 0$ be as in Theorem 6.3. Then for any $f, f_1, f_2 \in C_c^\infty(X_d)$ and $s, s_1, s_2 > 0$,*

$$\int_{\mathcal{Y}} f(g_s\Lambda_A) \ \mathrm{d}A = \mu_d(f) + O\left(e^{-\delta s} \mathcal{N}_\ell(f) \right), \tag{6.9}$$

and

$$\int_{\mathcal{Y}} f_1(g_{s_1}\Lambda_A) f_2(g_{s_2}\Lambda_A) \ \mathrm{d}A = \mu_d(f_1)\mu_d(f_2) + O\left(e^{-\delta \min\{s_1, s_2, |s_1 - s_2|\}} \mathcal{N}_\ell(f_1)\mathcal{N}_\ell(f_2) \right). \tag{6.10}$$

6.3. **Smooth approximations and estimates on norms.** In this subsection we prove the existence of smooth functions $\phi \in C_c^\infty(X_d)$ bounding our shrinking targets from above and below in an appropriate sense, with control on the norm $\mathcal{N}_\ell(\phi)$. We follow the strategy of [21, Theorem 1.1] while allowing the small identity neighborhoods of G (against which we convolve) to shrink.

Recall that

$$\widetilde{\Delta}_r := \bigcup_{0 \le s < 1} g_{-s}\Delta^{-1}[0, r] \qquad (0 < r < 1).$$

Lemma 6.5. *Let $\varepsilon > 0$. For any $0 < r < 1$, there exists $\phi_r \in C_c^\infty(X_d)$ satisfying $\chi_{\widetilde{\Delta}_r} \le \phi_r \le \chi_{\widetilde{\Delta}_{2r}}$ and, for any $\ell \in \mathbb{N}$:*

$$\mathcal{N}_\ell(\phi_r) \ll_{\ell,d,\varepsilon} r^{-L}, \qquad \text{with } \ L := 1 + \max\left\{0, \ell + \varepsilon - \tfrac{d^2}{4} - \tfrac{d}{4}\right\}. \tag{6.11}$$

(Note that the implied constant in the bound in (6.11) is independent of r.)

To prepare for the proof, let us define, for any $r > 0$,

$$\mathcal{O}_r := \left\{ g \in G \ : \ \max\{\|g - I_d\|, \|g^{-1} - I_d\|\} < \frac{r}{10d} \right\}.$$

Here the norm is the supremum norm on the matrix space $M_{d,d}(\mathbb{R})$. Clearly, $\mathcal{O}_r$ is an open neighborhood of the identity element in G and it is invariant under inversion. Let ν be the normalized Haar measure of G as in Section 2.1; recall that ν locally agrees with μ_d.

We will need the following auxiliary lemma.

Lemma 6.6. *For every $0 < r < 1$, there exists a function $\theta_r \in C_c^\infty(G)$ satisfying $\theta_r \ge 0$, $\mathrm{supp}(\theta_r) \subset \mathcal{O}_{r/2}$, $\int_G \theta_r(g)\, d\nu(g) = 1$ and $\|\mathcal{D}_Z(\theta_r)\|_{L^\infty(G)} \ll_{\ell,d} r^{1-d^2-\ell}$ for every monomial $Z = Y_1^{\ell_1} \cdots Y_a^{\ell_a}$, where $\ell = \deg(Z) = \ell_1 + \cdots + \ell_a$.*

(The implied constant in the bound on $\|\mathcal{D}_Z(\theta_r)\|_{L^\infty(G)}$ depends only on ℓ and d, and not on r.)

Proof (sketch). Let $\varphi : \Omega \to \mathbb{R}^a$ ($a = d^2 - 1$) be an arbitrary C^∞ coordinate chart of an open neighbourhood Ω of I_d in G, with $\varphi(I_d) = \mathbf{0}$. Let $\eta \in C_c^\infty(\mathbb{R}^a)$ be a fixed bump function in $\mathbb{R}^a$, i.e. a function satisfying $\eta \geq 0$ and $\int_{\mathbb{R}^a} \eta \, \mathrm{d}\boldsymbol{x} = 1$. We may assume that the support of η is contained in the unit ball centered at the origin, $\mathcal{B}_1^a$. For each $t > 0$ define $\eta_t \in C_c^\infty(\mathbb{R}^a)$ through

$$\eta_t(\boldsymbol{x}) = t^{-a}\eta(t^{-1}\boldsymbol{x}) \qquad (\boldsymbol{x} \in \mathbb{R}^a),$$

and note that $\mathrm{supp}(\eta_t) \subset \mathcal{B}_t^a$ and $\int_{\mathbb{R}^a} \eta_t \, \mathrm{d}\boldsymbol{x} = 1$. Let us choose the constant $c > 0$ so small that $\mathcal{B}_c^a \subset \varphi(\Omega)$ and $\varphi^{-1}(\mathcal{B}_{cr}^a) \subset \mathcal{O}_{r/2}$ for all $0 < r < 1$. This is possible since the matrix entries of g and g^{-1} are C^∞ functions of $g \in G$. Now we may simply set, for each $0 < r < 1$,

$$\theta_r := v_r \cdot (\eta_{cr} \circ \varphi),$$

where $v_r > 0$ is chosen so as to make $\int_G \theta_r \, \mathrm{d}\nu = 1$. One verifies that the limit $\lim_{r \to 0^+} v_r$ exists and is a positive real number. Using this fact, and recalling $a = d^2 - 1$, all the properties stated in the lemma are straightforward to verify. □

Proof of Lemma 6.5. We claim that for any $r_1, r_2 > 0$,

$$\mathcal{O}_{r_1}\widetilde{\Delta}_{r_2} \subset \widetilde{\Delta}_{r_1+r_2}. \tag{6.12}$$

First we note that for any $h \in M_{d,d}(\mathbb{R})$ and $\boldsymbol{v} \in \mathbb{R}^d$, $\|h\boldsymbol{v}\| \leq d\|h\|\|\boldsymbol{v}\|$. This implies that for any $r > 0$ and any $g \in \mathcal{O}_{10r}$ and $\boldsymbol{v} \in \mathbb{R}^d$,

$$\|g\boldsymbol{v}\| \leq \|\boldsymbol{v}\| + \|(g - I_d)\boldsymbol{v}\| \leq (1 + r)\|\boldsymbol{v}\|.$$

Hence for all $r > 0$,

$$\Delta(\Lambda) - \Delta(g\Lambda) \leq \log(1 + r) < r, \qquad \forall\, g \in \mathcal{O}_{10r},\ \Lambda \in X_d.$$

Similarly, since $\mathcal{O}_{10r}$ is invariant under inversion, we also have

$$\Delta(g\Lambda) - \Delta(\Lambda) = \Delta(g\Lambda) - \Delta(g^{-1}g\Lambda) < r, \qquad \forall\, g \in \mathcal{O}_{10r},\ \Lambda \in X_d.$$

Thus

$$\mathcal{O}_{10r_1}\Delta^{-1}[0, r_2] \subset \Delta^{-1}[0, r_1 + r_2], \qquad \forall\, r_1, r_2 > 0.$$

Now to prove the relation (6.12), in view of the definition of $\widetilde{\Delta}_r$, it suffices to show that for any $g \in \mathcal{O}_{r_1}$, $0 \leq s < 1$ and $\Lambda \in \Delta^{-1}[0, r_2]$ there exists some $0 \leq s' < 1$ such that

$$gg_{-s}\Lambda \in g_{-s'}\Delta^{-1}[0, r_1 + r_2],$$

or equivalently, $g_{s'}gg_{-s}\Lambda \in \Delta^{-1}[0, r_1 + r_2]$. We take $s' = s$. By direct computation and using $\alpha_i, \beta_j \in (0, 1)$ for all $1 \leq i \leq m, 1 \leq j \leq n$, we have

$$\begin{aligned}\max\left\{\|g_s g g_{-s} - I_d\|, \|(g_s g g_{-s})^{-1} - I_d\|\right\} &= \max\left\{\|g_s(g - I_d)g_{-s}\|, \|g_s(g^{-1} - I_d)g_{-s}\|\right\} \\ &< e^2 \max\left\{\|g - I_d\|, \|g^{-1} - I_d\|\right\} < e^2\frac{r_1}{10d} < \frac{r_1}{d}.\end{aligned}$$

Thus $g_s g g_{-s} \in \mathcal{O}_{10r_1}$, implying that

$$g_s g g_{-s}\Lambda \in \mathcal{O}_{10r_1}\Delta^{-1}[0, r_2] \subset \Delta^{-1}[0, r_1 + r_2].$$

This finishes the proof of (6.12).

Given any $0 < r < 1$, we now choose θ_r as in Lemma 6.6, and then define our approximating function $\phi_r \in C_c^\infty(X_d)$ as the convolution

$$\phi_r(x) := \theta_r * \chi_{\widetilde{\Delta}_{3r/2}}(x) = \int_G \theta_r(g)\chi_{\widetilde{\Delta}_{3r/2}}(g^{-1}x) \, \mathrm{d}\nu(g).$$

It follows from $\theta_r \geq 0$ and $\int_G \theta_r \, d\nu = 1$ that ϕ_r takes values in $[0,1]$. Moreover, for any $g \in \mathrm{supp}(\theta_r) \subset \mathcal{O}_{r/2}$ (so that g^{-1} is also contained in $\mathcal{O}_{r/2}$ since $\mathcal{O}_{r/2}$ is invariant under inversion) and for any $x \in \widetilde{\Delta}_r$, we have by (6.12)

$$g^{-1}x \in \mathcal{O}_{r/2}\widetilde{\Delta}_r \subset \widetilde{\Delta}_{3r/2},$$

implying that for any $x \in \widetilde{\Delta}_r$,

$$\phi_r(x) = \int_{\mathrm{supp}(\theta_r)} \theta_r(g)\chi_{\widetilde{\Delta}_{3r/2}}(g^{-1}x)\, \mathrm{d}\nu(g) = \int_{\mathrm{supp}(\theta_r)} \theta_r(g)\, \mathrm{d}\nu(g) = 1.$$

Thus $\chi_{\widetilde{\Delta}_r} \leq \phi_r$. Next, we claim that $\mathrm{supp}(\phi_r) \subset \widetilde{\Delta}_{2r}$. To prove this, note that since $\mathrm{supp}(\theta_r)$ is compact and contained in $\mathcal{O}_{r/2}$, there exists some $\epsilon \in (0, r/2)$ such that $\mathrm{supp}(\theta_r) \subset \mathcal{O}_{r/2-\epsilon}$. Now if $x \in \mathrm{supp}(\phi_r)$ then there exists some $x' \in \mathcal{O}_\epsilon\, x$ with $\phi_r(x') > 0$; and by the definition of ϕ_r there then exists some $g \in \mathrm{supp}(\theta_r) \subset \mathcal{O}_{r/2-\epsilon}$ such that $g^{-1}x' \in \widetilde{\Delta}_{3r/2}$. Hence $x' \in \mathcal{O}_{r/2-\epsilon}\widetilde{\Delta}_{3r/2} \subset \widetilde{\Delta}_{2r-\epsilon}$, and (since $\mathcal{O}_\epsilon$ is invariant under inversion) $x \in \mathcal{O}_\epsilon\, x' \subset \mathcal{O}_\epsilon\, \widetilde{\Delta}_{2r-\epsilon} \subset \widetilde{\Delta}_{2r}$. We have thus proved that $\mathrm{supp}(\phi_r) \subset \widetilde{\Delta}_{2r}$. Using this inclusion together with the fact that ϕ_r takes values in $[0,1]$, we conclude that $\phi_r \leq \chi_{\widetilde{\Delta}_{2r}}$. (Note that $\phi_r \leq \chi_{\widetilde{\Delta}_{2r}}$ follows already from the easier fact that for any $x' \in X_d$, $\phi_r(x') > 0$ implies $x' \in \widetilde{\Delta}_{2r}$. However we need some control on $\mathrm{supp}(\phi_r)$ below when we discuss derivatives of ϕ_r.)

For the norm bounds, we first note that using the invariance of the Haar measure, for any $Y \in \mathfrak{g}$ we have $\mathcal{D}_Y(\phi_r) = \big(\mathcal{D}_Y(\theta_r)\big) * \chi_{\widetilde{\Delta}_{3r/2}}$. More generally, for any monomial Z in $\{Y_1, \ldots, Y_a\}$,

$$\mathcal{D}_Z(\phi_r) = \big(\mathcal{D}_Z(\theta_r)\big) * \chi_{\widetilde{\Delta}_{3r/2}}. \tag{6.13}$$

Recall from Lemma 6.6 that $\mathrm{supp}(\theta_r) \subset \mathcal{O}_{r/2}$ and $\|\mathcal{D}_Z(\theta_r)\|_{L^\infty(G)} \ll_{\ell_Z,d} r^{1-d^2-\ell_Z}$, where ℓ_Z is the degree of Z. Furthermore, it is easily verified that $\nu(\mathcal{O}_{r/2}) \asymp_d r^{d^2-1}$. Using these facts, we have, for every $x \in X_d$ and every monomial Z of degree $\ell_Z \leq \ell$,

$$|\mathcal{D}_Z(\phi_r)(x)| = \left| \int_{\mathrm{supp}(\theta_r)} \mathcal{D}_Z(\theta_r)(g)\chi_{\widetilde{\Delta}_{3r/2}}(g^{-1}x)\, \mathrm{d}\nu(g) \right| \ll_{d,\ell} r^{1-d^2-\ell_Z}\nu(\mathcal{O}_{r/2}) \ll_d r^{-\ell_Z}. \tag{6.14}$$

Hence, using also $\mathrm{supp}(\phi_r) \subset \widetilde{\Delta}_{2r}$ and Theorem 5.1, we get

$$\|\phi_r\|_{L^2_\ell} \ll_{d,\ell} \mu_d(\widetilde{\Delta}_{2r})^{1/2} r^{-\ell} \ll_{d,\varepsilon} r^{\frac{d^2}{4}+\frac{d}{4}-1-\ell-\varepsilon}.$$

Finally, using the fact that for any $0 < r < 1$, the support of ϕ_r is contained in the fixed precompact set $\widetilde{\Delta}_2$, we have $\|\phi_r\|_{\mathrm{Lip}} \ll_d \sup_{x\in X_d} \sup_{j\in\{1,\ldots,a\}} \big|\mathcal{D}_{Y_j}(\phi_r)(x)\big|$, and hence by (6.14),

$$\|\phi_r\|_{\mathrm{Lip}} \ll_d r^{-1}.$$

Now the bound in (6.11) follows, via the definition of the norm $\mathcal{N}_\ell$. □

7. Proof of Theorem 1.2

In this section, building on the analysis developed in the previous section, we give the proof of Theorem 1.2. We keep the notation as in the previous section. In particular, throughout this section, we fix constants $\delta > 0$ and $\ell \in \mathbb{N}$ as in Corollary 6.4, and for each $0 < r < 1$ we fix a function $\phi_r \in C_c^\infty(X_d)$ as in Lemma 6.5. Taking $\varepsilon = 1$ in Lemma 6.5 we have that the norm bound in (6.11) holds for $L := 1 + \max\left\{0, \ell + 1 - \frac{d^2}{4} - \frac{d}{4}\right\}$.

7.1. **Application of effective equidistribution.** For any $0 < r < 1$, taking $f = \phi_r$ in the effective equidistribution result (6.9) and applying the norm estimate (6.11), we get for any $s > 0$,

$$\int_{\mathcal{Y}} \phi_r(g_s \Lambda_A)\, \mathrm{d}A = \mu_d(\phi_r) + O_d\left(e^{-\delta s} r^{-L}\right). \tag{7.1}$$

When r is small, the above integral should be expected to be small as well; however, the error term in (7.1) blows up as $r \to 0^+$. To remedy this issue, for r very small we will instead prove an upper bound on the integral, obtained by applying (7.1) for a suitable enlargened r-value. The result is as follows:

Lemma 7.1. *Let $\eta := \frac{\delta}{\varkappa_d + L}$. There exists $r_0 \in (0, \frac{1}{2})$ such that for any $0 < r < r_0$ and $s > 0$,*

$$\int_{\mathcal{Y}} \phi_r(g_s \Lambda_A)\, \mathrm{d}A \begin{cases} \asymp_d r^{\varkappa_d} \log^{\lambda_d}\left(\frac{1}{r}\right) & \text{if } r > e^{-\eta s}, \\ \ll_d e^{-\varkappa_d \eta s/2} & \text{if } r \le e^{-\eta s}. \end{cases} \tag{7.2}$$

In particular, for any sequence $\{\rho_k\}_{k \in \mathbb{N}} \subset (0, \frac{1}{2})$ with $\lim_{k\to\infty} \rho_k = 0$, we have

$$\sum_k \int_{\mathcal{Y}} \phi_{\rho_k}(g_k \Lambda_A)\, \mathrm{d}A = \infty \iff \sum_k \rho_k^{\varkappa_d} \log^{\lambda_d}\left(\frac{1}{\rho_k}\right) = \infty. \tag{7.3}$$

Proof. First we note that by Theorem 5.1 and the relation $\chi_{\widetilde{\Delta}_r} \le \phi_r \le \chi_{\widetilde{\Delta}_{2r}}$ (see Lemma 6.5), we have $\mu_d(\phi_r) \asymp_d r^{\varkappa_d} \log^{\lambda_d}\left(\frac{1}{r}\right)$. Furthermore, if $r > e^{-\eta s}$, then the ratio of the main term and the error term in (7.1) satisfies:

$$\frac{\mu_d(\phi_r)}{e^{-\delta s} r^{-L}} \asymp_d \frac{r^{\varkappa_d} \log^{\lambda_d}\left(\frac{1}{r}\right)}{e^{-\delta s} r^{-L}} > \log^{\lambda_d}\left(\tfrac{1}{r}\right),$$

which we can force to be as large as we like by taking the constant r_0 sufficiently small (in a way which only depends on d). Hence it follows from (7.1) that (7.2) holds in the case $r > e^{-\eta s}$.

Next assume $r \le e^{-\eta s}$. Set $\rho := 2e^{-\eta s}$. If $\rho < r_0$, then by what we proved in the previous paragraph,

$$\int_{\mathcal{Y}} \phi_\rho(g_s \Lambda_A)\, \mathrm{d}A \asymp_d \rho^{\varkappa_d} \log^{\lambda_d}\left(\tfrac{1}{\rho}\right) \ll_d e^{-\varkappa_d \eta s/2},$$

and hence the bound in (7.2) follows, since $\phi_r \le \chi_{\widetilde{\Delta}_{2r}} \le \chi_{\widetilde{\Delta}_\rho} \le \phi_\rho$ (again see Lemma 6.5). In the remaining case when $\rho \ge r_0$, we have $s \ll_d 1$ and $e^{-\varkappa_d \eta s/2} \gg_d 1$, and hence the bound in (7.2) holds simply because of $\phi_r \le 1$. This completes the proof of (7.2).

For the last part of the lemma, since $\lim_{k\to\infty} \rho_k = 0$, after possibly deleting finitely many terms from the two sums in (7.3), we may assume $\rho_k < r_0$ for all appearing terms. Next, using the second bound in (7.2) and the fact that both of the series $\sum_k e^{-\varkappa_d \eta k/2}$ and $\sum_{\rho_k \le e^{-\eta k}} \rho_k^{\varkappa_d} \log^{\lambda_d}\left(\frac{1}{\rho_k}\right)$ converge, it follows that the two divergence statements in (7.3) remain unaffected if all the terms for which $\rho_k \le e^{-\eta k}$ are removed from the respective series. After this operation, the equivalence in (7.3) is an immediate consequence of the first relation in (7.2). □

7.2. **The convergence case.** This case is now easily handled using Lemma 7.1.

Proof of the convergence case of Theorem 1.2. Let $r = r_\psi : [s_0, \infty) \to (0, \infty)$ be the continuous, decreasing function corresponding to ψ as in Lemma 6.1. First note that since the series (1.10) converges, we have $\lim_{t\to\infty} t\psi(t) = 1$ (or equivalently, $\lim_{s\to\infty} r(s) = 0$ as seen from the proof of

Lemma 6.1). Moreover, by the last part of Lemma 6.1, the fact that the series (1.10) converges implies that the series $\sum_k r(k)^{\varkappa_d} \log^{\lambda_d}\big(1+\frac{1}{r(k)}\big)$ also converges.

Now for each $k > s_0$ let us define

$$\overline{B}_k := \bigcup_{0\le s<1} g_{-s}\Delta^{-1}[0,\omega_1 r(k+s)] \quad \text{and} \quad \overline{E}_k := \big\{\Lambda_A \in \mathcal{Y} \,:\, g_k\Lambda_A \in \overline{B}_k\big\},$$

where $\omega_1 := \max\{m\alpha_i, n\beta_j \,:\, 1\le i\le m, 1\le j\le n\}$ is as in Proposition 6.2. In view of Proposition 6.2 (and the paragraph after it), it suffices to show that for Leb-a.e. $\Lambda_A \in \mathcal{Y}$, $g_k\Lambda_A \in \overline{B}_k$ for only finitely many $k > s_0$, or equivalently, that the limsup set $\limsup_{k\to\infty} \overline{E}_k$ is of zero measure. Thus in view of the Borel-Cantelli lemma, it suffices to show that $\sum_k \mathrm{Leb}(\overline{E}_k) < \infty$.

To prove this, we will approximate the shrinking targets $\{\overline{B}_k\}_{k>s_0}$ from above. Since $r(s)\to 0$ as $s\to\infty$, by enlarging s_0 if necessary (equivalently, enlarging t_0 as in Lemma 6.1), we may assume $\omega_1 r(s) \in (0,\frac{1}{2})$ for all $s > s_0$. Moreover, by Lemma 6.1, $r(\cdot)$ is decreasing; thus with $\rho_k := \omega_1 r(k)$, we have

$$\overline{B}_k \subset \widetilde{\Delta}_{\rho_k}, \qquad \forall\, k > s_0. \tag{7.4}$$

Recall that for each $0 < r < 1$ we have fixed a function $\phi_r \in C_c^\infty(X_d)$ as in Lemma 6.5. Now for each $k > s_0$, we have $\chi_{\overline{B}_k} \le \chi_{\widetilde{\Delta}_{\rho_k}} \le \phi_{\rho_k}$ (by (7.4) and Lemma 6.5), implying that

$$\mathrm{Leb}(\overline{E}_k) = \int_{\mathcal{Y}} \chi_{\overline{B}_k}(g_k\Lambda_A)\,\mathrm{d}A \le \int_{\mathcal{Y}} \phi_{\rho_k}(g_k\Lambda_A)\,\mathrm{d}A. \tag{7.5}$$

Next, it follows from $\rho_k = \omega_1 r(k)$ and the convergence of $\sum_k r(k)^{\varkappa_d}\log^{\lambda_d}\big(1+\frac{1}{r(k)}\big)$ that the series $\sum_k \rho_k^{\varkappa_d}\log^{\lambda_d}\big(\frac{1}{\rho_k}\big)$ also converges; in addition, $\lim_{k\to\infty}\rho_k = 0$ since $\lim_{k\to\infty} r(k) = 0$. Hence by the last part of Lemma 7.1 combined with (7.5), we have $\sum_k \mathrm{Leb}\left(\overline{E}_k\right) < \infty$, finishing the proof. □

Remark 14. Let $(m,n) = (2,1)$; thus $d = 3$. In [5, Theorem 1.1], Chow and Yang proved an effective equidistribution result for certain Diophantine lines in $\mathcal{Y}$ translated under the full (two dimensional) diagonal subgroup of G along certain restricted directions. In particular, their result implies the following: Let $(a,b) \in \mathbb{R}^2$ be a *Diophantine* vector (see [5, p. 2] for the definition), and let $J \subset \mathbb{R}$ be a compact subinterval. Then these exist constants $\ell' \in \mathbb{N}$, $c \in (0,1)$ and $\delta' > 0$ such that for any pair of weights $\boldsymbol{\alpha} = (\frac{c'}{1+c'}, \frac{1}{1+c'})$ with $0 < c' \le c$, for any $f \in C_c^\infty(X_3)$ and for any $s > 0$

$$\frac{1}{|J|}\int_J f\left(g_s\Lambda_{\boldsymbol{v}(x)}\right)\mathrm{d}x = \mu_3(f) + O\left(e^{-\delta' s}\|f\|_{L^\infty_{\ell'}}\right), \tag{7.6}$$

where $g_s = g_s^{\boldsymbol{\alpha},1} = \mathrm{diag}(e^{\alpha_1 s}, e^{\alpha_2 s}, e^{-s})$ with $\boldsymbol{\alpha}$ as above, $\boldsymbol{v}(x) := (ax+b, x)^t \in \mathbb{R}^2$ and $\|\cdot\|_{L^\infty_{\ell'}}$ the "L^∞, degree ℓ'" Sobolev norm defined by

$$\|f\|_{L^\infty_{\ell'}} := \sum_{\deg(Z)\le\ell'} \|\mathcal{D}_Z(f)\|_{C^0}.$$

Here $\|\cdot\|_{C^0}$ is the uniform norm on $C_c^\infty(X_3)$ as before. On the other hand, it is easy to see from Lemma 6.6 and the relation (6.13) that there exists $L' > 1$ such that

$$\|\phi_r\|_{L^\infty_{\ell'}} \ll_{\ell'} r^{-L'}, \qquad \forall\ 0 < r < 1. \tag{7.7}$$

Using a similar analysis with (7.6) and (7.7) in place of (6.9) and (6.11) respectively, we can conclude that if the series (1.10) (with $d = 3$) converges, then for Leb-a.e. $x \in J$ the column vector $\boldsymbol{v}(x) = (ax+b, x)^t$ is $\psi_{\boldsymbol{\alpha},1}$-Dirichlet.

7.3. **The divergence case.** In this subsection we prove the divergence case of Theorem 1.2. We first record from [29] the following divergence Borel-Cantelli lemma which we will use.

Lemma 7.2. *Let $(\mathcal{X}, \mu)$ be a probability space. Let $\{h_k\}_{k\in\mathbb{N}}$ be a sequence of measurable functions on $\mathcal{X}$ taking values in $[0,1]$. Let $b_k := \mu(h_k)$. Suppose $\sum_k b_k = \infty$ and*

$$\liminf_{k_2\to\infty} \frac{\int_{\mathcal{X}} \left(\sum_{i=k_1}^{k_2} h_i(x) - \sum_{i=k_1}^{k_2} b_i\right)^2 \mathrm{d}\mu(x)}{\left(\sum_{i=k_1}^{k_2} b_i\right)^2} = 0 \quad \text{for some } k_1 \in \mathbb{N}. \tag{7.8}$$

Then for μ-a.e. $x \in \mathcal{X}$, $h_k(x) > 0$ infinitely often.

Proof. Let $Y_1, Y_2, \dots$ be the sequence of random variables defined by $Y_k(x) = \sum_{i=k_1}^{k+k_1-1} h_i(x)$ ($k \in \mathbb{N}$). Note that

$$\int_{\mathcal{X}} \left(\sum_{i=k_1}^{k_2} h_i(x) - \sum_{i=k_1}^{k_2} b_i\right)^2 \mathrm{d}\mu(x) = \mu\big(Y_{k_2-k_1+1}^2\big) - \mu\big(Y_{k_2-k_1+1}\big)^2;$$

hence (7.8) implies that $\limsup_{k\to\infty} \mu(Y_k)^2/\mu(Y_k^2) = 1$. Therefore by part (iii) of the main theorem in [29], for μ-a.e. x we have $\limsup_{k\to\infty} Y_k(x)/\mu(Y_k) > 0$. Also $\mu(Y_k) = \sum_{i=k_1}^{k+k_1-1} b_i \to \infty$ as $k \to \infty$. Hence it follows that for μ-a.e. x we have $\lim_{k\to\infty} Y_k(x) = +\infty$, and in particular $h_i(x) > 0$ for infinitely many i. □

Remark 15. If one replaces the assumption (7.8) by the weaker assumption that

$$\liminf_{k_2\to\infty} \frac{\int_{\mathcal{X}} \left(\sum_{i=k_1}^{k_2} h_i(x) - \sum_{i=k_1}^{k_2} b_i\right)^2 \mathrm{d}\mu(x)}{\left(\sum_{i=k_1}^{k_2} b_i\right)^2} =: C < \infty \quad \text{for some } k_1 \in \mathbb{N},$$

then by the application of part (iii) of the main theorem in [29] we get instead

$$\mu\left(\left\{x \in X \,:\, \limsup_{k\to\infty} \frac{Y_k(x)}{\mu(Y_k)} > 0\right\}\right) \geq \frac{1}{1+C}.$$

In particular, there is a positive measure set of $x \in X$ such that $h_i(x) > 0$ infinitely often.

Proof of the divergence case of Theorem 1.2. First we note that in view of Remark 2 we may assume $\lim_{t\to\infty} t\psi(t) = 1$. Let $r = r_\psi$ be the continuous, decreasing function corresponding to ψ as in Lemma 6.1; then from the proof of that lemma we have $\lim_{s\to\infty} r(s) = 0$. Also by Lemma 6.1, since the series (1.10) diverges, the series $\sum_k r(k)^{\varkappa_d} \log^{\lambda_d}\big(1 + \frac{1}{r(k)}\big)$ also diverges. Moreover, by Remark 12, condition (1.11) is equivalent to

$$\liminf_{s_1\to\infty} \frac{\sum_{s_0<k\leq s_1} r(k)^{\varkappa_d} \log^{\lambda_d+1}\big(1 + \frac{1}{r(k)}\big)}{\left(\sum_{s_0<k\leq s_1} r(k)^{\varkappa_d} \log^{\lambda_d}\big(1 + \frac{1}{r(k)}\big)\right)^2} = 0. \tag{7.9}$$

Now for any $k > s_0$ let

$$\underline{B}_k = \bigcup_{0\leq s<1} g_{-s}\Delta^{-1}[0, \omega_2 r(k+s)]$$

be as before with $\omega_2 := \min\{m\alpha_i, n\beta_j \,:\, 1 \leq i \leq m, 1 \leq j \leq n\}$ as in Proposition 6.2. In view of Proposition 6.2 (and the paragraph after it) it suffices to show that for Leb-a.e. $\Lambda_A \in \mathcal{Y}$, $g_k\Lambda_A \in \underline{B}_k$ infinitely often.

In this case we will approximate the shrinking targets $\{\underline{B}_k\}_{k>s_0}$ from below. Recall that for each $0<r<1$ we have fixed a function $\phi_r \in C_c^\infty(X_d)$ as in Lemma 6.5. Again we may assume $\omega_2 r(s) \in (0, \frac{1}{2})$ for all $s > s_0$, and since $r(\cdot)$ is decreasing, we have $\widetilde{\Delta}_{\omega_2 r(k+1)} \subset \underline{B}_k$, implying that, with $\rho_k := \omega_2 r(k+1)/2$:

$$f_k := \phi_{\rho_k} \le \chi_{\widetilde{\Delta}_{\omega_2 r(k+1)}} \le \chi_{\underline{B}_k}, \qquad \forall\, k > s_0. \tag{7.10}$$

Let us set, for each $k > s_0$,

$$b_k := \int_{\mathcal{Y}} f_k(g_k \Lambda_A)\, \mathrm{d}A = \int_{\mathcal{Y}} \phi_{\rho_k}(g_k \Lambda_A)\, \mathrm{d}A.$$

Then, similarly to the proof of the convergence case, by applying the last part of Lemma 7.1 and using the relation $\rho_k = \omega_2 r(k+1)/2$ and the facts that the series $\sum_k r(k)^{\varkappa_d} \log^{\lambda_d}\big(1+\frac{1}{r(k)}\big)$ diverges and $\lim_{s\to\infty} r(s) = 0$, it follows that the series $\sum_k b_k$ also diverges.

Now for each $k > s_0$, let h_k be the function on $\mathcal{Y}$ defined by $h_k(\Lambda_A) := f_k(g_k \Lambda_A)$. Then in view of the definition of $f_k := \phi_{\rho_k}$ and the relation $\chi_{\widetilde{\Delta}_{\rho_k}} \le \phi_{\rho_k} \le \chi_{\widetilde{\Delta}_{2\rho_k}}$, the function h_k takes values in $[0,1]$, and

$$\int_{\mathcal{Y}} h_k(\Lambda_A)\, \mathrm{d}A = \int_{\mathcal{Y}} f_k(g_k \Lambda_A)\, \mathrm{d}A = b_k.$$

We will apply Lemma 7.2 to the probability space $(\mathcal{Y}, \mathrm{Leb})$ and the sequence $\{h_k\}_{k>s_0}$. We have already seen that $\sum_k b_k = \infty$; thus in view of Lemma 7.2 it suffices to show that $\{h_k\}_{k>s_0}$ satisfies condition (7.8).

Let us take $C > 0$ sufficiently large so that for all $k > C$, $\rho_k \in (0, r_0)$, where r_0 is the constant as in Lemma 7.1. For any $k_2 > k_1 > C$, let us denote

$$Q_{k_1,k_2} := \int_{\mathcal{Y}} \left(\sum_{i=k_1}^{k_2} h_i(\Lambda_A) - \sum_{i=k_1}^{k_2} b_i \right)^2 \mathrm{d}A = \sum_{k_1 \le i,j \le k_2} \int_{\mathcal{Y}} \big(h_i(\Lambda_A) h_j(\Lambda_A) - b_i b_j\big)\, \mathrm{d}A.$$

Using the fact that for each $k_1 \le i \le k_2$,

$$\int_{\mathcal{Y}} \big(h_i^2(\Lambda_A) - b_i^2\big)\, \mathrm{d}A \le \int_{\mathcal{Y}} h_i(\Lambda_A)\, \mathrm{d}A = b_i,$$

we have

$$Q_{k_1,k_2} \le \sum_{i=k_1}^{k_2} b_i + 2 \sum_{k_1 \le i < j \le k_2} \int_{\mathcal{Y}} \big(h_i(\Lambda_A) h_j(\Lambda_A) - b_i b_j\big)\, \mathrm{d}A.$$

Fix $k_1 \le i < j \le k_2$; we will use two different estimates for the term $\int_{\mathcal{Y}} \big(h_i(\Lambda_A) h_j(\Lambda_A) - b_i b_j\big)\, \mathrm{d}A$ depending on whether $\min\{i, j-i\}$ is large or small. First, applying the effective doubly mixing (6.10) to the pair (f_i, f_j), we get

$$\int_{\mathcal{Y}} f_i(g_i \Lambda_A) f_j(g_j \Lambda_A)\, \mathrm{d}A = \mu_d(f_i)\mu_d(f_j) + O\left(e^{-\delta \min\{i,j-i\}} \mathcal{N}_\ell(f_i) \mathcal{N}_\ell(f_j)\right). \tag{7.11}$$

On the other hand, by (6.9) we have

$$b_k = \mu_d(f_k) + O\left(e^{-\delta k} \mathcal{N}_\ell(f_k)\right), \quad \forall\, k > C. \tag{7.12}$$

Combining (7.12), (7.11), the norm estimate $\mathcal{N}_\ell(f_k) = \mathcal{N}_\ell(\phi_{\rho_k}) \ll_d \rho_k^{-L}$ (by (6.11)) and noting that $\int_{\mathcal{Y}} h_i(\Lambda_A)h_j(\Lambda_A)\,\mathrm{d}A = \int_{\mathcal{Y}} f_i(g_i\Lambda_A)f_j(g_j\Lambda_A)\,\mathrm{d}A$, we have

$$\left|\int_{\mathcal{Y}} (h_i(\Lambda_A)h_j(\Lambda_A) - b_ib_j)\,\mathrm{d}A\right| \ll_d e^{-\delta\min\{i,j-i\}}\rho_i^{-L}\rho_j^{-L} < e^{-\delta\min\{i,j-i\}}\rho_j^{-2L}.$$

On the other hand, using the trivial estimate $|h_ih_j - b_ib_j| \le h_ih_j + b_ib_j \le h_j + b_j$ we have

$$\left|\int_{\mathcal{Y}} (h_i(\Lambda_A)h_j(\Lambda_A) - b_ib_j)\,\mathrm{d}A\right| \le \int_{\mathcal{Y}} (h_j(\Lambda_A) + b_j)\,\mathrm{d}A = 2b_j. \tag{7.13}$$

Combining these two bounds, we conclude:

$$Q_{k_1,k_2} \ll_d \sum_{i=k_1}^{k_2} b_i + \sum_{j=k_1+1}^{k_2}\sum_{i=k_1}^{j-1} \min\left\{e^{-\delta\min\{i,j-i\}}\rho_j^{-2L}, b_j\right\}. \tag{7.14}$$

In order to bound the above inner sum we replace k_1 by 1 and use the symmetry $i \mapsto j-i$ to get

$$\sum_{i=k_1}^{j-1} \min\left\{e^{-\delta\min\{i,j-i\}}\rho_j^{-2L}, b_j\right\} \le 2\sum_{i=1}^{j-1} \min\left\{e^{-\delta\, i}\rho_j^{-2L}, b_j\right\}. \tag{7.15}$$

In the last sum, all terms are $\le b_j$, and there are at most $O_d\big(\log\big(2 + b_j^{-1}\rho_j^{-2L}\big)\big)$ terms which are equal to b_j (indeed, remember that δ depends only on d). Furthermore, if there are any terms which are less than b_j, then these are bounded above by $b_j, b_je^{-\delta}, b_je^{-2\delta}, \ldots$, and so their sum is $O_d(b_j)$. It follows that the last sum in (7.15) is $O_d\left(b_j \min\left\{j, \log\big(2 + b_j^{-1}\rho_j^{-2L}\big)\right\}\right)$, and hence from (7.14) we get

$$Q_{k_1,k_2} \ll_d \sum_{i=k_1}^{k_2} b_i + \sum_{j=k_1+1}^{k_2} b_j \min\left\{j, \log\big(2 + b_j^{-1}\rho_j^{-2L}\big)\right\} \ll \sum_{j=k_1}^{k_2} b_j \min\left\{j, \log\big(2 + b_j^{-1}\rho_j^{-2L}\big)\right\}.$$

Let $\eta = \frac{\delta}{\varkappa_d + L}$ be as in Lemma 7.1 and set $\kappa := \frac{\varkappa_d\eta}{2}$. Then by (7.2) we have for each $k_1 \le j \le k_2$:

$$b_j = \int_{\mathcal{Y}} \phi_{\rho_j}(g_j\Lambda_A)\,\mathrm{d}A \begin{cases} \asymp_d \rho_j^{\varkappa_d}\log^{\lambda_d}\big(\frac{1}{\rho_j}\big) & \text{if } \rho_j > e^{-\eta j}, \\ \ll_d e^{-\kappa j} & \text{if } \rho_j \le e^{-\eta j}. \end{cases} \tag{7.16}$$

Thus for any $k_2 > k_1 > C$ we have (recalling that $\rho_j = \omega_2 r(j+1)/2$)

$$\begin{aligned} Q_{k_1,k_2} &\ll_d \sum_{\substack{k_1\le j\le k_2\\(\rho_j \le e^{-\eta j})}} e^{-\kappa j}j + \sum_{\substack{k_1\le j\le k_2\\(\rho_j > e^{-\eta j})}} \rho_j^{\varkappa_d}\log^{\lambda_d}\Big(\frac{1}{\rho_j}\Big)\log\left(2 + \rho_j^{-\varkappa_d-2L}\log^{-\lambda_d}\Big(\frac{1}{\rho_j}\Big)\right) \\ &\ll_d 1 + \sum_{j=k_1}^{k_2} \rho_j^{\varkappa_d}\log^{\lambda_d+1}\Big(\frac{1}{\rho_j}\Big) \ll 1 + \sum_{j=k_1+1}^{k_2+1} r(j)^{\varkappa_d}\log^{\lambda_d+1}\Big(1 + \frac{1}{r(j)}\Big). \end{aligned}$$

Similarly, by (7.16), for any fixed $k_1 > C$ we have as $k_2 \to \infty$:

$$\sum_{j=k_1}^{k_2} b_j \gg_d \sum_{\substack{k_1\le j\le k_2\\(\rho_j > e^{-\eta j})}} \rho_j^{\varkappa_d}\log^{\lambda_d}\Big(\frac{1}{\rho_j}\Big) \gg \sum_{j=k_1}^{k_2} \rho_j^{\varkappa_d}\log^{\lambda_d}\Big(\frac{1}{\rho_j}\Big) \gg_{\omega_2} \sum_{j=k_1+1}^{k_2+1} r(j)^{\varkappa_d}\log^{\lambda_d}\Big(1 + \frac{1}{r(j)}\Big),$$

where the second relation holds since the series $\sum_{j=k_1}^{\infty} \rho_j^{\varkappa_d}\log^{\lambda_d}\big(\frac{1}{\rho_j}\big)$ diverges (this follows from the relation $\rho_j = \omega_2 r(j+1)/2$ and the fact that the series $\sum_j r(j)^{\varkappa_d}\log^{\lambda_d}\big(1 + \frac{1}{r(j)}\big)$ diverges), while the same sum restricted to those j for which $\rho_j \le e^{-\eta j}$ is convergent.

Combining the last two bounds, we conclude that for any $k_1 > C$, and for k_2 sufficiently large,

$$\frac{Q_{k_1,k_2}}{\left(\sum_{j=k_1}^{k_2} b_j\right)^2} \ll_{d,\omega_2} \frac{1+\sum_{j=k_1+1}^{k_2+1} r(j)^{\varkappa_d} \log^{\lambda_d+1}\left(1+\frac{1}{r(j)}\right)}{\left(\sum_{j=k_1+1}^{k_2+1} r(j)^{\varkappa_d} \log^{\lambda_d}\left(1+\frac{1}{r(j)}\right)\right)^2}. \tag{7.17}$$

Since the series $\sum_j r(j)^{\varkappa_d} \log^{\lambda_d}\left(1+\frac{1}{r(j)}\right)$ diverges, condition (7.9) implies that the limit inferior of the expression in the right hand side of (7.17) tends to zero as $k_2 \to \infty$. Hence (7.8) holds. We have also noted that $\sum_k b_k = \infty$. Hence by Lemma 7.2, for Leb-a.e. $A \in M_{m,n}(\mathbb{R}/\mathbb{Z})$ we have $h_k(\Lambda_A) = f_k(g_k\Lambda_A) > 0$ infinitely often. Together with (7.10), this implies that for Leb-a.e. $A \in M_{m,n}(\mathbb{R}/\mathbb{Z})$, the lattice $g_k\Lambda_A$ belongs to $\mathrm{supp}(f_k) \subset \underline{B}_k$ for infinitely many $k \in \mathbb{N}$. This finishes the proof. □

Remark 16. For the divergence case in Theorem 1.2, we note that if one replaces the assumption (1.11) by the weaker assumption that

$$\liminf_{t_1\to\infty} \frac{\sum_{t_0\le k\le t_1} k^{-1} F_\psi(k)^{\varkappa_d} \log^{\lambda_d+1}\left(\frac{1}{F_\psi(k)}\right)}{\left(\sum_{t_0\le k\le t_1} k^{-1} F_\psi(k)^{\varkappa_d} \log^{\lambda_d}\left(\frac{1}{F_\psi(k)}\right)\right)^2} < \infty,$$

then, in view of Remark 12, Remark 15 and the estimate (7.17), we can conclude that, under this weaker assumption, $\mathbf{DI}_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\psi)^c$ is of positive Lebesgue measure.

References

[1] M. Björklund and A. Gorodnik. Central limit theorems for Diophantine approximants. *Math. Ann.*, 374(3-4):1371–1437, 2019.

[2] J. W. S. Cassels. *An introduction to Diophantine approximation*. Cambridge Tracts in Mathematics and Mathematical Physics, No. 45. Cambridge University Press, New York, 1957.

[3] Y. Cheung. Hausdorff dimension of the set of singular pairs. *Ann. of Math. (2)*, 173(1):127–167, 2011.

[4] Y. Cheung and N. Chevallier. Hausdorff dimension of singular vectors. *Duke Math. J.*, 165(12):2273–2329, 2016.

[5] S. Chow and L. Yang. An effective ratner equidistribution theorem for multiplicative diophantine approximation on planar lines. *arXiv preprint arXiv:1902.06081*, 2019.

[6] S. G. Dani. Divergent trajectories of flows on homogeneous spaces and Diophantine approximation. *J. Reine Angew. Math.*, 359:55–89, 1985.

[7] T. Das, L. Fishman, D. Simmons, and M. Urbański. A variational principle in the parametric geometry of numbers, with applications to metric Diophantine approximation. *C. R. Math. Acad. Sci. Paris*, 355(8):835–846, 2017.

[8] T. Das, L. Fishman, D. Simmons, and M. Urbański. A variational principle in the parametric geometry of numbers. arXiv:1901.06602, 2019.

[9] H. Davenport and W. M. Schmidt. Dirichlet's theorem on diophantine approximation. II. *Acta Arith.*, 16:413–424, 1969/70.

[10] H. Davenport and W. M. Schmidt. Dirichlet's theorem on diophantine approximation. In *Symposia Mathematica, Vol. IV (INDAM, Rome, 1968/69)*, pages 113–132. Academic Press, London, 1970.

[11] G. Hajós. Über einfache und mehrfache Bedeckung des n-dimensionalen Raumes mit einem Würfelgitter. *Math. Z.*, 47:427–467, 1941.

[12] M. Hussain, D. Y. Kleinbock, N. Wadleigh, and B. Wang. Hausdorff measure of sets of Dirichlet non-improvable numbers. *Mathematika*, 64(2):502–518, 2018.

[13] S. Kadyrov, D. Y. Kleinbock, E. Lindenstrauss, and G. A. Margulis. Singular systems of linear forms and non-escape of mass in the space of lattices. *J. Anal. Math.*, 133:253–277, 2017.

[14] O. Khalil and M. Luethi. Random walks, spectral gaps, and khintchine's theorem on fractals. *arXiv preprint arXiv:2101.05797*, 2021.

[15] T. Kim and W. Kim. Hausdorff measure of sets of dirichlet non-improvable affine forms. *arXiv preprint arXiv:2006.05727*, 2020.

[16] D. Y. Kleinbock. Flows on homogeneous spaces and Diophantine properties of matrices. *Duke Math. J.*, 95(1):107–124, 1998.

[17] D. Y. Kleinbock, N. de Saxcé, N. A. Shah, and P. Yang. Equidistribution in the space of 3-lattices and dirichlet-improvable vectors on planar lines. *arXiv preprint arXiv:2106.08860*, 2021.
[18] D. Y. Kleinbock and G. A. Margulis. Bounded orbits of nonquasiunipotent flows on homogeneous spaces. In *Sinaĭ's Moscow Seminar on Dynamical Systems*, volume 171 of *Amer. Math. Soc. Transl. Ser. 2*, pages 141–172. Amer. Math. Soc., Providence, RI, 1996.
[19] D. Y. Kleinbock and G. A. Margulis. Logarithm laws for flows on homogeneous spaces. *Invent. Math.*, 138(3):451–494, 1999.
[20] D. Y. Kleinbock and G. A. Margulis. On effective equidistribution of expanding translates of certain orbits in the space of lattices. In *Number theory, analysis and geometry*, pages 385–396. Springer, New York, 2012.
[21] D. Y. Kleinbock and G. A. Margulis. Erratum to: Logarithm laws for flows on homogeneous spaces. *Invent. Math.*, 211(2):855–862, 2018.
[22] D. Y. Kleinbock and S. Mirzadeh. On the dimension drop conjecture for diagonal flows on the space of lattices. arXiv:2010.14065, 2020.
[23] D. Y. Kleinbock and A. Rao. A zero-one law for uniform diophantine approximation in euclidean norm. *arXiv preprint arXiv:1910.00126*, 2019. to appear in Int. Math. Res. Not. IMRN.
[24] D. Y. Kleinbock, R. Shi, and B. Weiss. Pointwise equidistribution with an error rate and with respect to unbounded functions. *Math. Ann.*, 367(1-2):857–879, 2017.
[25] D. Y. Kleinbock and N. Wadleigh. A zero-one law for improvements to Dirichlet's Theorem. *Proc. Amer. Math. Soc.*, 146(5):1833–1844, 2018.
[26] D. Y. Kleinbock and N. Wadleigh. An inhomogeneous Dirichlet theorem via shrinking targets. *Compos. Math.*, 155(7):1402–1423, 2019.
[27] D. Y. Kleinbock and B. Weiss. Dirichlet's theorem on Diophantine approximation and homogeneous flows. *J. Mod. Dyn.*, 2(1):43–62, 2008.
[28] D. Y. Kleinbock and S. Yu. A dynamical Borel–Cantelli lemma via improvements to Dirichlet's theorem. *Mosc. J. Comb. Number Theory*, 9(2):101–122, 2020.
[29] S. Kochen and C. Stone. A note on the Borel-Cantelli lemma. *Illinois J. Math.*, 8:248–251, 1964.
[30] G. A. Margulis. *On some aspects of the theory of Anosov systems*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2004. With a survey by Richard Sharp: Periodic orbits of hyperbolic flows, Translated from the Russian by Valentina Vladimirovna Szulikowska.
[31] M. Ratner. On Raghunathan's measure conjecture. *Ann. of Math. (2)*, 134(3):545–607, 1991.
[32] N. A. Shah. Equidistribution of expanding translates of curves and Dirichlet's theorem on Diophantine approximation. *Invent. Math.*, 177(3):509–532, 2009.
[33] N. A. Shah. Expanding translates of curves and Dirichlet-Minkowski theorem on linear forms. *J. Amer. Math. Soc.*, 23(2):563–589, 2010.
[34] N. A. Shah and L. Yang. Equidistribution of curves in homogeneous spaces and Dirichlet's approximation theorem for matrices. *Discrete Contin. Dyn. Syst.*, 40(9):5247–5287, 2020.
[35] S. K. Stein and S. Szabó. *Algebra and tiling*, volume 25 of *Carus Mathematical Monographs*. Mathematical Association of America, Washington, DC, 1994. Homomorphisms in the service of geometry.
[36] A. Strömbergsson and A. Venkatesh. Small solutions to linear congruences and Hecke equidistribution. *Acta Arith.*, 118(1):41–78, 2005.
[37] P. Yang. Equidistribution of expanding translates of curves and Diophantine approximation on matrices. *Invent. Math.*, 220(3):909–948, 2020.

Brandeis University, Waltham MA, USA, 02454-9110
E-mail address: kleinboc@brandeis.edu

Department of Mathematics, Uppsala University, Box 480, SE-75106, Uppsala, SWEDEN
E-mail address: astrombe@math.uu.se

Department of Mathematics, Uppsala University, Box 480, SE-75106, Uppsala, SWEDEN
E-mail address: shucheng.yu@math.uu.se