ON THE MINIMUM OF A POSITIVE DEFINITE QUADRATIC FORM OVER NON–ZERO LATTICE POINTS. THEORY AND APPLICATIONS.

FAUSTIN ADICEAM AND EVGENIY ZORIN

Abstract. Let Σ_d^{++} be the set of positive definite matrices with determinant 1 in dimension $d \geqslant 2$. Identifying any two $SL_d(\mathbb{Z})$–congruent elements in Σ_d^{++} gives rise to the space of reduced quadratic forms of determinant one, which in turn can be identified with the locally symmetric space $X_d := SL_d(\mathbb{Z})\backslash SL_d(\mathbb{R})/SO_d(\mathbb{R})$. Equip the latter space with its natural probability measure coming from a Haar measure on $SL_d(\mathbb{R})$. In 1998, Kleinbock and Margulis [11] established sharp estimates for the probability that an element of X_d takes a value less than a given real number $\delta > 0$ over the non–zero lattice points $\mathbb{Z}^d\backslash\{\mathbf{0}\}$.

In this article, these estimates are extended to a large class of probability measures arising either from the spectral or the Cholesky decomposition of an element of Σ_d^{++}. The sharpness of the bounds thus obtained are also established (up to multiplicative constants) for a subclass of these measures.

Although of an independent interest, this theory is partly developed here with a view towards application to Information Theory. More precisely, after providing a concise introduction to this topic fitted to our needs, we lay the theoretical foundations of the study of some manifolds frequently appearing in the theory of Signal Processing. This is then applied to the recently introduced Integer–Forcing Receiver Architecture channel whose importance stems from its expected high performance. Here, we give sharp estimates for the probabilistic distribution of the so–called *Effective Signal–to–Noise Ratio*, which is an essential quantity in the evaluation of the performance of this model.

In honorem Henriettae Dickinsonis.

Contents

FA research is supported by EPSRC Programme Grant : EP/J018260/1 and EZ research is supported by EPSRC Grant : EP/M021858/1.

1. Introduction

Fix once and for all an integer $d \geqslant 2$. Let Q be a non–degenerate symmetric matrix in dimension d. Throughout, the matrix Q will be identified with the corresponding quadratic form $\boldsymbol{x} \in \mathbb{R}^d \mapsto {}^t\boldsymbol{x} \cdot Q \cdot \boldsymbol{x}$.

If Q is indefinite, the Oppenheim conjecture solved by Margulis states that the set of values taken by this quadratic form at non–zero integral points, viz.

$$\left\{ {}^t\boldsymbol{a} \cdot Q \cdot \boldsymbol{a} \ : \ \boldsymbol{a} \in \mathbb{Z}^d \backslash \{\mathbf{0}\} \right\},$$

is dense in the real line whenever $d \geqslant 3$. When $d = 2$ however (i.e. for indefinite binary quadratic forms), this set may exhibit very different structures : it may be dense or else closed and discrete, but it may also be not closed and/or not dense. For further details on the theory of values taken by an indefinite quadratic form, the reader is referred to [6, 7] and to the references therein.

In the case that Q is definite, say positive definite without loss of generality, it is easy to see that the quantity

$$M_d(Q) := \min_{\boldsymbol{a} \in \mathbb{Z}^d \backslash \{\mathbf{0}\}} {}^t\boldsymbol{a} \cdot Q \cdot \boldsymbol{a} \tag{1}$$

is well–defined. It is a result due to Hermite (see [2, p.43] for a proof) that one has always

$$M_d(Q) \leqslant \left(\frac{4}{3}\right)^{(d-1)/2} |Q|^{1/d}, \tag{2}$$

where $|Q|$ denotes the determinant of Q. It is known that the constant $(4/3)^{(d-1)/2}$ on the right–hand side of (2) is optimal only when $d = 2$. Denoting by $\mathcal{S}_d^{++}$ the set of positive definite matrices in dimension $d \geqslant 2$, this leads one to the definition of the *Hermite constant* γ_d :

$$\gamma_d := \frac{\sup_{Q \in \mathcal{S}_d^{++}} M_d(Q)}{|Q|^{1/d}}.$$

The supremum in this definition can actually be replaced with a maximum. Only the values of γ_d for $d = 2, 3, 4, 5, 6, 7, 8$ and $d = 24$ are exactly known. For other d's, several estimates have been established. See, e.g., [5] for proofs and further details on the Hermite constants. See also [4] for an algorithm to approximate $M_d(Q)$ for a *given* $Q \in \mathcal{S}_d^{++}$.

. It should be noted that the study of the quantity $M_d(Q)$ for a generic $Q \in \mathcal{S}_d^{++}$ underpins the more general problem of determining the minimum of such a quadratic form over non–zero elements of *any* full rank lattice Λ. Indeed, as such a lattice can be written in the form $\Lambda = L \cdot \mathbb{Z}^d$ for some $L \in GL_d(\mathbb{R})$, the minimum of Q over the elements of $\Lambda \backslash \{\mathbf{0}\}$ is given by $M_d({}^tLQL)$. Also, if $L' \in GL_d(\mathbb{R})$ is another matrix such that $\Lambda = L' \cdot \mathbb{Z}^d$, then there exists $Z \in SL_d(\mathbb{Z})$ such that $L' = LZ$. This implies in particular that $M_d(Q) = M_d({}^tZQZ)$ for any $Q \in \mathcal{S}_d^{++}$ and any $Z \in SL_d(\mathbb{Z})$, i.e. that the quantity $M_d(Q)$ is invariant under $SL_d(\mathbb{Z})$–congruent matrices.

. The problem of estimating $M_d(Q)$ is here considered from a probabilistic point of view. Given an estimate such as (2), even if it means renormalising in an obvious way the matrices under consideration, it is natural to focus on the case of positive definite matrices *with determinant one*. Let therefore

$$\Sigma_d^{++} := \left\{\Sigma \in \mathcal{S}_d^{++} \ : \ \det(\Sigma) = 1\right\}$$

denote such a set. In full generality, the main problem addressed in this work can loosely be summarised this way :

Problem 1 (Main Problem)**.** *For a* given *probability measure μ on the set Σ_d^{++}, estimate the probability $\mu\left(M_d(\Sigma) \leqslant \delta\right)$ as a function of $\delta > 0$.*

In order to take into account the $SL_d(\mathbb{Z})$–invariance of the problem, identify any two $SL_d(\mathbb{Z})$–congruent matrices in Σ_d^{++}. This defines the space of reduced quadratic forms with determinant one, which is henceforth denoted by $\Sigma_{d,red}^{++}$. It is easy to see that the map

$$\phi \ : \ \overline{g} \in X_d \mapsto g \cdot {}^tg \in \Sigma_{d,red}^{++} \tag{3}$$

is well–defined and bijective, where X_d denotes the locally symmetric space

$$X_d := SL_d(\mathbb{Z}) \backslash SL_d(\mathbb{R}) / SO_d(\mathbb{R})$$

and where $\overline{g} := SL_d(\mathbb{Z}) \cdot g \cdot SO_d(\mathbb{R})$ is the equivalence class in X_d of any $g \in SL_d(\mathbb{R})$ (the surjectivity of the map ϕ follows for instance from the Cholesky decomposition of an element of Σ_d^{++}). From now on, let

$$\Gamma := SL_d(\mathbb{Z}), \quad G := SL_d(\mathbb{R}) \quad \text{and} \quad H := SO_d(\mathbb{R})$$

(which are all unimodular groups) in such a way that $X_d := \Gamma\backslash G/H$.

The set X_d seen as a double coset space can be equipped with a natural G–invariant probability measure μ_{X_d} arising from the G–invariant probability measure $\mu_{\Gamma\backslash G}$ on the space of lattices $\Gamma\backslash G$. If one denotes by μ_H the Haar probability measure on H, the invariant measure μ_{X_d} is characterised by the fact that for any Borel measurable function $f \in \mathbb{L}^1(\mu_{\Gamma\backslash G})$, the following equation holds :

$$\int_{X_d} \left(\int_H f(gh) \cdot \mathrm{d}\mu_H(h) \right) \cdot \mathrm{d}\mu_{X_d}(gH) \ = \ \int_{\Gamma\backslash G} f(g) \cdot \mathrm{d}\mu_{\Gamma\backslash G}(g)$$

(see [13] for proofs and details). The probability measure $\mu_{\Gamma\backslash G}$ is itself obtained from any suitably normalised Haar measure μ_G on G. One can furthermore explicitly express the volume element $\mathrm{d}\mu_G(M)$ in terms of the Iwasawa decomposition of $M \in G$ — see [17, §2] for details.

With the help of the bijective map (3), the measure μ_{X_d} can be pushed forward to a probability measure $\phi_*\mu_{X_d}$ on the space $\Sigma_{d,red}^{++}$. In view of Problem 1, one is then concerned with the estimate of the probability

$$\begin{aligned} p_{X_d}(\delta) \ &= \ (\phi_*\mu_{X_d}) \left(\left\{ \overline{\Sigma} \in \Sigma_{d,red}^{++} \ : \ M_d(\overline{\Sigma}) \leqslant \delta \right\} \right) \\ &= \ \mu_{X_d} \left(\{ \overline{g} \in X_d \ : \ M_d(\phi(\overline{g})) \leqslant \delta \} \right) \end{aligned}$$

for any fixed $\delta > 0$ which may be assumed to be less than the Hermite constant γ_d for obvious reasons (note that the above equations are direct consequences of the change of variables formula for pushforward measures). This problem was emphatically solved by Kleinbock–Margulis who proved in [11, §7] the following result (see also [12, Theorem 1.3.5]). Before stating it, and in view of the statement of our own results, let from now

$$V_d \ = \ \frac{\pi^{d/2}}{\Gamma\left(\frac{d}{2}+1\right)} \qquad \text{and} \qquad A_d \ = \ \frac{2\pi^{d/2}}{\Gamma\left(\frac{d}{2}\right)} \tag{4}$$

denote respectively the volume and the area of the unit Euclidean ball in dimension $d \geqslant 2$ (here, $\Gamma(\,.\,)$ denotes the usual Euler Gamma function).

Theorem 1 (Kleinbock & Margulis, 1998)**.** *The following inequalities hold for any $\delta > 0$:*

$$\frac{V_d}{2\zeta(d)}\delta^{d/2} - c_d\frac{V_d^2}{4}\delta^d \ \leqslant \ p_{X_d}(\delta) \ \leqslant \ \frac{V_d}{2\zeta(d)}\delta^{d/2}. \tag{5}$$

Here, ζ denotes the Riemann zeta function and c_d a strictly positive constant which, when $d \geqslant 3$, can be taken to be

$$c_d \ = \ \frac{1}{\zeta(d) \cdot \zeta(d-1)}.$$

The implicit presence of the square root of δ on both sides of (5) is due to this easily verified equivalence valid for any $g \in G$:

$$(M_d(\phi(\overline{g})) \leqslant \delta) \quad \Longleftrightarrow \quad \left(g \cdot \mathbb{Z}^d \cap B_2(\mathbf{0}, \sqrt{\delta}) \neq \{\mathbf{0}\}\right),$$

where, given $\boldsymbol{x} \in \mathbb{R}^d$ and $r > 0$, $B_2(\boldsymbol{x}, r)$ is the closed Euclidean ball with radius r centered at $\boldsymbol{x}$.

Theorem 1 suggests that, as $\delta > 0$ tends to zero, one should expect the probability of the event $M_d(\Sigma) \leqslant \delta$ to grow like $\delta^{d/2}$ when the space Σ_d^{++} is equipped with a "typical" probability measure defined from the invariant measure μ_{X_d}. For the applications we have in mind however (see §4), the choice of any such measure is neither natural nor convenient. The primary theoretical goal of this work is thus to establish estimates in the likes of (5) for a larger class of probability measures on the space Σ_d^{++}. These probability measures will be defined from the spectral (§2) and then the Cholesky decomposition (§3) of an element of Σ_d^{++}.

Note that, although the problem of estimating the probability of the event $M_d(\Sigma) \leqslant \delta$ is well–defined in the space $\Sigma_{d,red}^{++}$ of reduced quadratic forms, there is no loss of information in working instead in the space Σ_d^{++}. Indeed, any probability measure on Σ_d^{++} defines a probability measure on $\Sigma_{d,red}^{++}$ after periodisation modulo $SL_d(\mathbb{Z})$–congruent matrices. Conversely, any probability measure on $\Sigma_{d,red}^{++}$ defines a probability measure on Σ_d^{++} supported on a fundamental domain of $\Sigma_{d,red}^{++}$ in Σ_d^{++}.

Before stating the main results, we mention that the latter may also be used to tackle the following less natural but nevertheless still relevant variant of the main problem stated above (namely, when the probability space is $\mathcal{S}_d^{++}$ instead of Σ_d^{++}) :

Problem 2 (Variant of the Main Problem)**.** *For a* given *probability measure μ' on the set $\mathcal{S}_d^{++}$, estimate the probability $\mu'(M_d(Q) \leqslant \delta)$ as a function of $\delta > 0$.*

The changes to make to the results dealing with Problem 1 in order to obtain their analogues for Problem 2 are straightforward when considering the approach via the spectral decomposition (§2). They will therefore not be explicitly stated. When considering the approach via the Cholesky decomposition however (§3), these changes will induce some technical difficulties and will therefore be explicitly stated.

Throughout, in order not interrupt the thread of the exposition, the lengthy proofs are postponed until the end of each section. They may be skipped at a first reading.

2. An Approach via the Spectral Decomposition.

Denote by $\mathcal{D}_d^{++}$ the set of diagonal matrices in dimension d with strictly positive entries. Let Δ_d^{++} be the subgroup of $\mathcal{D}_d^{++}$ consisting of all those matrices with determinant one :

$$\Delta_d^{++} := \mathcal{D}_d^{++} \cap SL_d(\mathbb{R}).$$

Throughout, $\mathcal{D}_d^{++}$ (resp. Δ_d^{++}) will be identified with $(\mathbb{R}_{>0})^d$ (resp. with $(\mathbb{R}_{>0})^{d-1}$ — in this case, one only considers the $d-1$ first diagonal entries of an element of Δ_d^{++} to define the identification). It will sometimes be more convenient to see an element of Δ_d^{++} as an element of $\mathcal{D}_d^{++}$, in which case it will also be represented as a d–tuple. This should not cause any confusion.

Let

$$\mathcal{O}_d := O_d(\mathbb{R})$$

denote the orthogonal group in dimension d. We first seek to equip the set Σ_d^{++} with a special class of probability measures defined from the spectral decomposition of an element therein. This class will play an important role in the forthcoming considerations : in short, Problem 1 will be addressed for probability measures lying in this class.

2.1. **Definition of a Suitable Class of Measures.** Let $\Sigma \in \Sigma_d^{++}$ be decomposed as $\Sigma = {}^tP\Delta P$ with $P \in \mathcal{O}_d$ and $\Delta \in \Delta_d^{++}$. Given $\boldsymbol{x} \in \mathbb{R}^d$, one has clearly ${}^t\boldsymbol{x} \cdot \Sigma \cdot \boldsymbol{x} = {}^t\boldsymbol{y} \cdot \boldsymbol{y}$ with $\boldsymbol{y} = \sqrt{\Delta} P \boldsymbol{x}$. This shows that the following equivalence holds for any $\delta > 0$:

$$(M_d(\Sigma) \leqslant \delta) \quad \Longleftrightarrow \quad \left(P \cdot \mathbb{Z}^d \cap \Delta^{-1/2} \cdot B_2(\mathbf{0}, \sqrt{\delta}) \neq \{\mathbf{0}\}\right). \tag{6}$$

This motivates the introduction of the surjective map

$$\Psi \ : \ (P, \Delta) \in \mathcal{O}_d \times \Delta_d^{++} \ \mapsto \ {}^tP\Delta^{-2}P \in \Sigma_d^{++} \tag{7}$$

which we now use to push forward to Σ_d^{++} a given measure defined on $\mathcal{O}_d \times \Delta_d^{++}$ (the exponent "-2" is just meant to simplify the formulae hereafter). It is important to keep in mind for what follows that the orthogonal matrix P appearing in the Spectral Decomposition of Σ as above is well–defined in the quotient $\mathcal{O}_d/\mathcal{I}_d$, where $\mathcal{I}_d$ is the subgroup of $\mathcal{O}_d$ consisting of all those diagonal matrices with entries ± 1. The equivalence (6) then still holds when P is seen as an element of $\mathcal{O}_d/\mathcal{I}_d$ in view of the fact that $P \cdot I \cdot \mathbb{Z}^d = P \cdot \mathbb{Z}^d$ for any $I \in \mathcal{I}_d$.

Let μ_d be the Haar probability measure on the compact group $\mathcal{O}_d$. Given $P \in \mathcal{O}_d$, the volume element $\mathrm{d}\mu_d(P)$ is explicitly described for instance in [20] in terms of $d(d-1)/2$ independent coordinates on $\mathcal{O}_d$. Let furthermore ν_d be a probability measure on Δ_d^{++}. Define then a measure on the product space $\mathcal{O}_d \times \Delta_d^{++}$ by setting

$$\tau_d \ := \ \mu_d \otimes \nu_d. \tag{8}$$

This can be pushed forward to a probability measure $\Psi_*\tau_d$ on Σ_d^{++}. Of course, the relevance of such a measure strongly relies on the properties of the map Ψ and of the measure τ_d. In this respect, the following lemma establishes a crucial property satisfied by Ψ :

Lemma 1. *Let $\Delta_{d,sub}^{++}$ be the subset of Δ_d^{++} consisting of all those elements in Δ_d^{++} whose entries are pairewise distinct :*

$$\Delta_{d,sub}^{++} \ := \ \left\{\Delta = (\alpha_1, \cdots, \alpha_d) \in \Delta_d^{++} \ : \ \forall i \neq j, \ \alpha_i \neq \alpha_j\right\}.$$

Then, the restriction of the map Ψ to the set $\mathcal{O}_d \times \Delta_{d,sub}^{++}$ is 2^d to 1.

More precisely, Ψ induces a bijection

$$\Psi' \ : \ (\mathcal{O}_d/\mathcal{I}_d) \times \Delta_{d,sub}^{++} \ \mapsto \ \Psi\left(\mathcal{O}_d \times \Delta_{d,sub}^{++}\right) \subset \Sigma_d^{++}. \tag{9}$$

Note that $\Psi\left(\mathcal{O}_d \times \Delta_{d,sub}^{++}\right)$ sits as a dense open set in Σ_d^{++}.

Proof. Let $Q \in \Sigma_d^{++}$ with spectral decomposition $Q = {}^tP\Delta^{-2}P$ for some $P \in \mathcal{O}_d$ and some $\Delta \in \Delta_{d,sub}^{++}$. The rows of the matrix P are then (normed) eigenvectors of Q. Since eigenvectors associated to distinct eigenvalues are orthogonal, these rows are determined up to their sign. The lemma follows. □

Let ρ_d be the Haar probability measure on $\mathcal{O}_d/\mathcal{I}_d$, which satisfies the property that for any function $f \in \mathbb{L}^1(\mu_d)$ defined over $\mathcal{O}_d$,

$$\int_{\mathcal{O}_d} f(P) \cdot \mathrm{d}\mu_d(P) \ = \ \frac{1}{2^d} \cdot \int_{\mathcal{O}_d/\mathcal{I}_d} \left(\sum_{I \in \mathcal{I}_d} f(PI) \right) \cdot \mathrm{d}\rho_d(P\mathcal{I}_d). \tag{10}$$

In view of Lemma 1, a dense open subset of Σ_d^{++} can be identified with the product space $(\mathcal{O}_d/\mathcal{I}_d) \times \Delta_{d,sub}^{++}$ via the map Ψ' defined in (9). We will be interested in probability measures supported on this dense open set. A natural class of such measures are obtained by taking the pushforward by Ψ' of a measure of the form $\rho_d \otimes \nu_d$ under the following assumption on ν_d which will be made throughout :

Assumption 1. *The complement of $\Delta_{d,sub}^{++}$ in Δ_d^{++} has zero ν_d–measure, i.e.*

$$\nu_d\left(\Delta_{d,sub}^{++}\right) = 1.$$

Thus, under this assumption, Ψ' establishes a bijection between a set of full $\rho_d \otimes \nu_d$– measure in $(\mathcal{O}_d/\mathcal{I}_d) \times \Delta_d^{++}$ and its image in Σ_d^{++}.

Note also that under Assumption 1, the two pushforward measures $\Psi'_*(\rho_d \otimes \nu_d)$ and $\Psi_*\tau_d$ (with τ_d defined in (8)) are exactly the same on Σ_d^{++}. Indeed, if $\Sigma \in \Sigma_d^{++}$ lies in the image of the restriction of the map Ψ to $\mathcal{O}_d \times \Delta_{d,sub}^{++}$, Lemma 1 implies that the preimage $\Psi^{-1}(\{\Sigma\})$ of Σ by Ψ is of the form $\Psi^{-1}(\{\Sigma\}) = \{(PI, \Delta) \ : \ I \in \mathcal{I}_d\}$ for some $P \in \mathcal{O}_d$ and $\Delta \in \Delta_d^{++}$. Since the orthogonal matrix P appearing in the the equivalence stated in (6) can be seen as an element of $\mathcal{O}_d/\mathcal{I}_d$, it follows from the definition of Ψ in (7) that either all or none of the 2^d elements (P, Δ) in this preimage satisfy/ies the relation

$$P \cdot \mathbb{Z}^d \ \cap \ \Delta \cdot B_2(\mathbf{0}, \sqrt{\delta}) \ \neq \ \{\mathbf{0}\}. \tag{11}$$

Together with (10), this establishes the claim.

Assumption 1 imposes a rather mild restriction on the measure ν_d, which is even allowed to be fractal. A natural class of measures satisfying this assumption is given by those probability measures which are absolutely continuous with respect to a Haar

measure ξ on Δ_d^{++}. Recall that, up to a multiplication constant, the volume element $\mathrm{d}\xi(\Delta)$ of any such invariant measure is given by

$$\mathrm{d}\xi(\Delta) \ = \ \prod_{i=1}^{d-1} \frac{\mathrm{d}\alpha_i'}{\alpha_i'}, \tag{12}$$

where $\Delta = (\alpha_1', \dots, \alpha_{d-1}') \in \Delta_d^{++}$.

2.2. **Estimation of the Probability that a Non–Zero Integer Vector should lie in a Random Ellipsoid Centered at the Origin.** We adopt here a geometric approach in order to address Problem 1 within the framework developed thus far. Part of the ideas behind this approach have been applied in [17] to problems in mathematical physics. However, unlike here, the focus in the latter work was rather on the probability that a *large* convex set should contain a non–zero lattice point. Furthermore, the multiplicative constants appearing in the formulae proved in [17] are not explicit while it will be one of our objectives to obtain fully explicit estimates.

From the change of variables formula for pushforward measures and in view of (6), (7) and (11), the objective boils down to estimating, for a given $\delta > 0$, the quantity

$$(\Psi_* \tau_d) \left(\left\{ \Sigma \in \Sigma_d^{++} \ : \ M_d(\Sigma) \ \leqslant \ \delta \right\} \right) \ = \ \tau_d \left(\mathfrak{F}_d(\delta) \right),$$

where

$$\mathfrak{F}_d(\delta) \ := \ \left\{ (P, \Delta) \in \mathcal{O}_d \times \Delta_d^{++} \ : \ P \cdot \mathbb{Z}^d \ \cap \ \Delta \cdot B_2(\mathbf{0}, \sqrt{\delta}) \ \neq \ \{\mathbf{0}\} \right\}.$$

To avoid cumbersome notation, the set $\mathfrak{F}_d(\delta)$ will from now on be denoted by $\mathfrak{F}(\delta)$ whenever there is no risk of confusion.

In order to state the results regarding the estimate of the probability $\tau_d(\mathfrak{F}(\delta))$, a good deal of notation is first introduced.

Throughout, a vector in $\mathbb{R}^d$ will be seen as the datum of a d–tuple represented in *column* (that is, we consider the right action of d–dimensional matrices on $\mathbb{R}^d$). Whenever this does not induce any ambiguity, such a vector shall indifferently be written in row for convenience. Given a vector $\boldsymbol{\alpha} := (\alpha_1, \dots, \alpha_d) \in (\mathbb{R}_{>0})^d$, $\mathcal{E}_d(\boldsymbol{\alpha})$ will denote the *full* ellipsoid

$$\mathcal{E}_d(\boldsymbol{\alpha}) \ := \ \left\{ \boldsymbol{x} \in \mathbb{R}^d \ : \ \sum_{i=1}^{d} \left(\frac{x_i}{\alpha_i} \right)^2 \ \leqslant \ 1 \right\} \tag{13}$$

($\alpha_1, \dots, \alpha_d$ are thus the lengths of the semi–principal axes of this ellipsoid). If there is no risk of confusion, one shall also write more simply $\mathcal{E}(\boldsymbol{\alpha})$ for $\mathcal{E}_d(\boldsymbol{\alpha})$.

Let $\mathbb{S}^{d-1}$ denote the unit sphere in dimension d. Let also σ_{d-1} be the spherical probability measure on $\mathbb{S}^{d-1}$. This measure is given by a volume element denoted by $\mathrm{d}\boldsymbol{v}$ which is such that for any σ_{d-1}–measurable surface $\mathcal{A} \subset \mathbb{S}^{d-1}$,

$$\sigma_{d-1}(\mathcal{A}) \ := \ \frac{1}{A_d} \int_{\mathcal{A}} \mathrm{d}\boldsymbol{v}$$

(we have chosen not to include the factor A_d in the volume element as otherwise any use of our results will unavoidably involve the computation of constants involving this factor). If $\mathcal{A}$ is any subset of $\mathbb{R}^d$ such that its intersection $\mathcal{A} \cap \mathbb{S}^{d-1}$ with the unit sphere is σ_{d-1}–measurable, set

$$\widetilde{\sigma}_{d-1}(\mathcal{A}) \ := \ \sigma_{d-1}\left(\mathcal{A} \cap \mathbb{S}^{d-1}\right).$$

Given a vector $\boldsymbol{v} \in \mathbb{S}^{d-1}$, $\boldsymbol{v}^{\perp}$ shall denote the hyperplane in $\mathbb{R}^d$ passing through the origin with unit normal vector $\boldsymbol{v}$. Also, the notation $\|\,.\,\|_2$ and $\|\,.\,\|_\infty$ shall refer to the usual Euclidean and sup norms in $\mathbb{R}^d$. The set of points in $\mathbb{Z}^d$ visible from the origin shall be denoted by $\mathcal{P}(\mathbb{Z}^d)$:

$$\mathcal{P}(\mathbb{Z}^d) \ := \ \left\{\boldsymbol{a} \in \mathbb{Z}^d \ : \ \gcd(\boldsymbol{a}) = 1\right\}.$$

Finally, given a closed convex set $\mathcal{C} \subset \mathbb{R}^d$ centered at the origin, define

$$p_d(\mathcal{C}) \ := \ \mu_d\left(\left\{P \in \mathcal{O}_d \ : \ P \cdot \mathbb{Z}^d \cap \mathcal{C} \neq \{\mathbf{0}\}\right\}\right).$$

Note that in the case $d = 1$, $\mathcal{O}_1 = \{\pm 1\}$, the convex body $\mathcal{C}$ is an interval $\mathcal{J}$ and

$$p_1(\mathcal{J}) \ = \ \begin{cases} 1 & \text{if } \lambda(\mathcal{J}) \geqslant 2 \\ 0 & \text{if } \lambda(\mathcal{J}) < 2, \end{cases} \tag{14}$$

where $\lambda(\mathcal{J})$ denotes the length of $\mathcal{J}$.

The main result in this section can now be stated as follows.

Theorem 2. *Let $\delta > 0$. Then,*

$$\tau_d(\mathfrak{F}(\delta)) \ = \ \int_{\Delta_d^{++}} p_d\left(\mathcal{E}(\sqrt{\delta}\Delta)\right) \cdot \mathrm{d}\nu_d(\Delta). \tag{15}$$

Furthermore, the quantity $p_d\left(\mathcal{E}(\sqrt{\delta}\Delta)\right)$ satisfies the estimates

$$g_d(\Delta, \delta) \ \leqslant \ p_d\left(\mathcal{E}(\sqrt{\delta}\Delta)\right) \ \leqslant \ f_d(\Delta, \delta), \tag{16}$$

where

$$g_d(\Delta, \delta) \ := \ \max\left\{\widetilde{\sigma}_{d-1}\left(\mathcal{E}_d(\sqrt{\delta}\Delta)\right), \ \int_{\mathbb{S}^{d-1}} p_{d-1}\left(\mathcal{E}_d(\sqrt{\delta}\Delta) \cap \boldsymbol{v}^{\perp}\right) \cdot \frac{\mathrm{d}\boldsymbol{v}}{A_d}\right\}$$

and

$$f_d(\Delta, \delta) \ := \ \min\left\{1, \ \sum_{\substack{\boldsymbol{n} \in \mathcal{P}(\mathbb{Z}^d) \\ \|\boldsymbol{n}\|_2 \leqslant \sqrt{\delta}\|\Delta\|_\infty}} \widetilde{\sigma}_{d-1}\left(\mathcal{E}_d\left(\frac{\sqrt{\delta}}{\|\boldsymbol{n}\|_2}\Delta\right)\right)\right\}.$$

Here, the base case for the recursive formula induced by the integral in $g_d(\Delta, \delta)$ is given by (14) *and the sum in $f_d(\Delta, \delta)$ is to be seen as equal to zero when $\sqrt{\delta}\,\|\Delta\|_\infty < 1$.*

In view of such a statement, we now seek to determine, one the one hand the intersection of an ellipsoid with a hyperplane and on the other the spherical measure of the intersection of a (full) ellipsoid with the unit sphere. The former question is addressed in this proposition :

Proposition 1. *Let $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_d) \in (\mathbb{R}_{>0})^d$ and $\boldsymbol{v} = (v_1, \dots, v_d) \in \mathbb{S}^{d-1}$. Assume that $v_d \neq 0$.*

Then, the intersection $\mathcal{E}_d(\boldsymbol{\alpha}) \cap \boldsymbol{v}^\perp$ of the d–dimensional ellipsoid $\mathcal{E}_d(\boldsymbol{\alpha})$ with the hyperplane $\boldsymbol{v}^\perp$ is a $(d-1)$–dimensional ellipsoid $\mathcal{E}_{d-1}(\boldsymbol{\alpha}, \boldsymbol{v})$. Furthermore, one has

$$\mathcal{E}_{d-1}(\boldsymbol{\alpha}, \boldsymbol{v}) \ = \ \left\{ \boldsymbol{y} \in \mathbb{R}^{d-1} \ : \ {}^t\boldsymbol{y} \cdot Q \cdot \boldsymbol{y} \ \leqslant \ 1 \right\}, \tag{17}$$

where

$$Q \ := \ D\left(I_{d-1} + \boldsymbol{u} \cdot {}^t\boldsymbol{u}\right) D \ \in \ \mathcal{S}_d^{++} \tag{18}$$

with I_{d-1} the identity matrix in dimension $d-1$,

$$D \ := \ \left(\alpha_1^{-1}, \cdots, \alpha_d^{-1}\right) \in \mathcal{D}_d^{++} \quad \text{and} \quad {}^t\boldsymbol{u} \ := \ \left(\frac{\alpha_i v_i}{\alpha_d v_d}\right)_{1 \leqslant i \leqslant d-1} \in \mathbb{R}^{d-1}.$$

Also, if the lengths of the semi–principal axes of $\mathcal{E}_d(\boldsymbol{\alpha})$ are ordered increasingly in the sense that $\alpha_1 \leqslant \cdots \leqslant \alpha_d$, then the lengths $\beta_1, \dots, \beta_{d-1}$ of the semi–principal axes of $\mathcal{E}_{d-1}(\boldsymbol{\alpha}, \boldsymbol{v})$ ordered increasingly satisfy the inequalities

$$\alpha_1 \ \leqslant \ \beta_1 \ \leqslant \ \alpha_2 \ \leqslant \ \ldots \ \leqslant \ \alpha_{d-1} \ \leqslant \ \beta_{d-1} \ \leqslant \ \alpha_d.$$

Note that, even if it means relabelling the axes, there is no loss of generality in assuming that the lengths of the semi–principal axes of $\mathcal{E}_d(\boldsymbol{\alpha})$ are ordered increasingly. Also, the condition $v_d \neq 0$ is not restrictive at all as formula (17) holds *mutatis mutandis* with any other non–zero coordinate v_j in place of v_d — see the proof in §2.4 for details.

We now turn to the estimate of the spherical measure of the intersection of the ellipsoid $\mathcal{E}_d(\boldsymbol{\alpha})$ with the unit sphere (where $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_d) \in (\mathbb{R}_{>0})^d$). To this end, it may be assumed, without loss of generality in view of Assumption 1, that

$$0 \ < \ \alpha_1 \ < \ \alpha_2 \ < \ \ldots \ < \ \alpha_{d-1} \ < \ \alpha_d. \tag{19}$$

Whenever $\alpha_d > 1$, define then

$$\underset{\sim}{\boldsymbol{\alpha}} \ := \ \left(\underset{\sim}{\alpha}_1, \ \ldots \ , \underset{\sim}{\alpha}_{d-1}\right) \ \in \ \mathcal{D}_{d-1}^{++}, \tag{20}$$

where for $i = 1, \dots, d-1$,

$$\underset{\sim}{\alpha}_i \ := \ \sqrt{\alpha_i^2 \cdot \frac{\alpha_d^2 - 1}{\alpha_d^2 - \alpha_i^2}}.$$

The following statement provides an inductive formula for $\widetilde{\sigma}_{d-1}\left(\mathcal{E}_d(\boldsymbol{\alpha})\right)$. The quantity

$$W_k \ = \ \int_0^{\pi/2} \sin^k \theta \cdot \mathrm{d}\theta \ = \ \frac{\sqrt{\pi}}{2} \cdot \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k+2}{2}\right)} \tag{21}$$

appearing therein denotes the Wallis integral of order $k \geqslant 0$.

Proposition 2. *Assuming* (19)*, one has*

$$\widetilde{\sigma}_{d-1}\left(\mathcal{E}_d(\boldsymbol{\alpha})\right) \ = \ \begin{cases} 1 & \text{if } \alpha_1 \geqslant 1 \\ 0 & \text{if } \alpha_d \leqslant 1. \end{cases} \tag{22}$$

Moreover, if $\alpha_1 < 1 < \alpha_d$, *then*

$$\widetilde{\sigma}_{d-1}\left(\mathcal{E}_d\left(\boldsymbol{\alpha}\right)\right) \;=\; \frac{1}{2W_{d-2}} \cdot \int_0^{\pi} \widetilde{\sigma}_{d-2}\left(\mathcal{E}_{d-1}\left(\frac{\underset{\sim}{\boldsymbol{\alpha}}}{\sin\theta}\right)\right) \cdot (\sin\theta)^{d-2} \cdot \mathrm{d}\theta \tag{23}$$

with base case

$$\widetilde{\sigma}_0\left(\mathcal{E}_1\left(\alpha\right)\right) \;=\; \begin{cases} 1 & \text{if } \alpha \geqslant 1 \\ 0 & \text{if } \alpha < 1 \end{cases}$$

for any $\alpha > 0$.

Although providing an exact theoretical formula, equation (23) may lead to lengthy calculations for a given ellipsoid. In order to overcome this difficulty, the next proposition provides rather accurate estimates for the quantity $\widetilde{\sigma}_{d-1}\left(\mathcal{E}_d\left(\boldsymbol{\alpha}\right)\right)$ when $\alpha_1 < 1 < \alpha_d$. Before stating it, we introduce some additional notation : given $x \geqslant 0$, let

$$b(x) \;:=\; \arccos\left(\min\left\{1, x\right\}\right) \;=\; \begin{cases} \arccos(x) \in [0,\, \pi/2] & \text{if } x \in [0,\, 1], \\ 0 & \text{if } x \geqslant 1. \end{cases}$$

Under (19), define

$$\mathfrak{I}_d(\boldsymbol{\alpha}) \;:=\; \frac{2^d}{A_d} \cdot \prod_{i=2}^{d} \int_{b(\alpha_{d-i+1})}^{\pi/2} \sin^{i-2}\theta \cdot \mathrm{d}\theta \quad \text{whenever} \quad \alpha_d \geqslant 1. \tag{24}$$

We leave this quantity undefined when $\alpha_d < 1$. For $i = 1, \ldots, d-1$, assuming $\alpha_d \geqslant 1$, set furthermore

$$\underset{\sim}{\alpha}_i^* \;:=\; \min\left\{1, \underset{\sim}{\alpha}_i\right\} \;=\; \begin{cases} \underset{\sim}{\alpha}_i & \text{if } \alpha_i \leqslant 1, \\ 1 & \text{if } \alpha_i \geqslant 1 \end{cases}$$

and let $\underset{\sim}{\boldsymbol{\alpha}}^* = \left(\underset{\sim}{\alpha}_1^*, \ldots, \underset{\sim}{\alpha}_{d-1}^*\right)$.

Proposition 3. *Assume that* (19) *holds and that* $\alpha_1 < 1 < \alpha_d$. *Then, with the notation above, one has*

$$\mathfrak{I}_d\left(\frac{\underset{\sim}{\boldsymbol{\alpha}}^*}{\sqrt{d-1}}, 1\right) \;\leqslant\; \widetilde{\sigma}_{d-1}\left(\mathcal{E}_d\left(\boldsymbol{\alpha}\right)\right) \;\leqslant\; \mathfrak{I}_d\left(\underset{\sim}{\boldsymbol{\alpha}}^*, 1\right).$$

The following cruder but easier–to–estimate inequalities also hold :

$$\mathfrak{I}_d\left(\frac{\boldsymbol{\alpha}}{\sqrt{d}}\right) \;\leqslant\; \widetilde{\sigma}_{d-1}\left(\mathcal{E}_d\left(\boldsymbol{\alpha}\right)\right) \;\leqslant\; \mathfrak{I}_d\left(\boldsymbol{\alpha}\right),$$

where the lower bound is defined whenever $\alpha_d \geqslant \sqrt{d}$.

Here, given a generic vector $\boldsymbol{\alpha} \in \left(\mathbb{R}_{>0}\right)^d$ *satisfying* (19) *and* $\alpha_d \geqslant 1$, *the quantity* $\mathfrak{I}_d(\boldsymbol{\alpha})$ *can be estimated as follows :*

$$a(d) \cdot \prod_{j=1}^{d-1} \min\left\{\alpha_j,\, 1\right\} \;\leqslant\; \mathfrak{I}_d\left(\boldsymbol{\alpha}\right) \;\leqslant\; a'(d) \cdot \prod_{j=1}^{d-1} \min\left\{\alpha_j,\, 1\right\}$$

with

$$a(d) \;=\; \frac{2^d}{(d-1)! \cdot A_d} \cdot \left(\frac{\pi}{2}\right)^{(d-2)(d-3)/2} \qquad \textit{and} \qquad a'(d) = \frac{2^d}{A_d} \cdot \left(\frac{\pi}{2}\right)^{d(d-1)/2}.$$

With the help of Propositions 1, 2 and 3, one may now answer the question as to whether Theorem 2 leads to sharp estimates for the probability $\tau_d\left(\mathfrak{F}(\delta)\right)$ as expressed in (15). To this end, one must focus on a relevant subclass of probability measures ν_d. A natural choice is to restrict the attention to compactly supported measures. Indeed, such measures can approximate a large class of measures and appear naturally in practical problems (see §4). Assume therefore without loss of generality that ν_d seen as a measure on $(\mathbb{R}_{>0})^{d-1}$ is absolutely continuous with respect to the Haar measure (12) with density supported on the hypercube $\left[\epsilon,\, \epsilon^{-1}\right]^{d-1}$. Denote by $\chi_\epsilon^{(d)} : \mathbb{R}^{d-1} \to \mathbb{R}$ the characteristic function of the latter set.

To simplify the calculations, we will further require that the density of ν_d with respect to the Haar measure ξ is uniform, i.e. that ξ–almost everywhere, the density $\mathrm{d}\nu_d/\mathrm{d}\xi$ is proportional to $\chi_\epsilon^{(d)}$. In view of (12), given $\boldsymbol{\alpha}' = (\alpha'_1, \ldots, \alpha'_{d-1}) \in \Delta_d^{++}$, one has explicitly

$$\mathrm{d}\nu_d^{(\epsilon)}(\boldsymbol{\alpha}') \;=\; \frac{1}{\left|2\log\epsilon\right|^{d-1}} \cdot \chi_\epsilon^{(d)}(\boldsymbol{\alpha}') \cdot \prod_{i=1}^{d-1} \frac{\mathrm{d}\alpha'_i}{\alpha'_i}, \tag{25}$$

where $\nu_d^{(\epsilon)} = \nu_d$. Inasmuch as one is working up to multiplicative constants, one can reduce to this case any measure whose density with respect to $\nu_d^{(\epsilon)}$ is almost everywhere bounded above on the hypercube $K_\varepsilon(d) = \left[\epsilon,\, \epsilon^{-1}\right]^{d-1}$ and almost everywhere bounded below by a strictly positive constant on a sub–hypercube of $K_\varepsilon(d)$.

The next proposition shows that, for any given $\epsilon > 0$, the estimates of the probability $\tau_d^{(\epsilon)}\left(\mathfrak{F}(\delta)\right) := \tau_d\left(\mathfrak{F}(\delta)\right)$ obtained from Theorem 2 are essentially sharp in δ.

Theorem 3. *Fix $\epsilon > 0$ and assume that $\delta \in (0,\, 1)$. Let $\tau_d^{(\epsilon)}$ be the probability measure defined as in (8) from the measure $\nu_d^{(\epsilon)}$ given by (25).*

Then,

$$\tau_d^{(\epsilon)}\left(\mathfrak{F}(\delta)\right) \;=\; 0 \quad \text{ if } \quad \delta \leqslant \epsilon^{2(d-1)}. \tag{26}$$

Moreover, if $\delta > \epsilon^{2(d-1)}$, then

$$c_d(\epsilon) \cdot s_d\left(\epsilon, \delta\right) \;\leqslant\; \tau_d^{(\epsilon)}\left(\mathfrak{F}(\delta)\right) \;\leqslant\; C_d(\epsilon) \cdot S_d\left(\epsilon, \delta\right) \tag{27}$$

for some constants $c_d(\epsilon), C_d(\epsilon) > 0$. Here,

$$s_d\left(\epsilon, \delta\right) \;:=\; \int_{J_d(\epsilon,\delta)} \prod_{i=1}^{d-1} \min\left\{\sqrt{\delta},\, \frac{1}{\alpha_i}\right\} \cdot \mathrm{d}\alpha_i$$

and

$$S_d\left(\epsilon, \delta\right) \;:=\; \delta^{d/2} \cdot \int_{J_d(\epsilon,\delta)} \prod_{i=1}^{d-1} \frac{\mathrm{d}\alpha_i}{\alpha_i},$$

where the domain of integration $J_d(\epsilon, \delta)$ is defined by the set of inequalities

$$\epsilon \leqslant \alpha_1 < \cdots < \alpha_{d-1} \leqslant \epsilon^{-1} \quad \text{ and } \quad \max\left\{\delta^{-1/2},\, \alpha_{d-1}\right\} < \left(\alpha_1 \ldots \alpha_{d-1}\right)^{-1}.$$

These quantities $s_d(\epsilon,\delta)$ *and* $S_d(\epsilon,\delta)$ *satisfy the estimates*

$$s_d(\epsilon,\delta) \geqslant \min\left\{\sqrt{\delta},\epsilon\right\}^{d-1} \cdot \frac{|2\log\epsilon|^{d-2}}{(d-2)!} \cdot \left(\min\left\{\sqrt{\delta},\epsilon\right\} - \epsilon^{d-1}\right). \tag{28}$$

and

$$S_d(\epsilon,\delta) \leqslant \delta^{d/2} \cdot \log\left(\frac{\sqrt{\delta}}{\epsilon^{d-1}}\right) \cdot \frac{|2\log\epsilon|^{d-2}}{(d-2)!}. \tag{29}$$

One can furthermore choose

$$c_d(\epsilon) = \frac{a(d)\cdot(d-1)!}{(d\cdot|2\log\epsilon|)^{d-1}}$$

and

$$C_d(\epsilon) = \frac{3^{d-1}\cdot a'(d)\cdot d!\cdot d}{|2\cdot\log\epsilon|^{d-1}},$$

where $a(d)$ *and* $a'(d)$ *are defined in Proposition 3.*

Theorem 3 implies for instance the existence of two positive constants $\kappa(d)$ and $K(d)$ depending *only* on the dimension d such that for any δ lying in the interval $\left[\epsilon^{2(d-1)},\,\epsilon^2\right]$,

$$\kappa(d)\cdot\frac{\delta^{d/2}}{|\log\epsilon|}\cdot\left(1-\frac{\epsilon^{d-1}}{\sqrt{\delta}}\right) \leqslant \tau_d^{(\epsilon)}\left(\mathfrak{F}(\delta)\right) \leqslant K(d)\cdot\frac{\delta^{d/2}}{|\log\epsilon|}\cdot\left(\frac{\sqrt{\delta}}{\epsilon^{d-1}}-1\right)$$

(the upper bound is a direct consequence of the convexity inequality $\log(1+x)\leqslant x$ valid for all $x\geqslant 0$). We thus recover in this case also the growth in $\delta^{d/2}$ appearing in Theorem 1.

The remainder of this section is devoted to the proofs of the various results stated above.

2.3. **Proof of Theorem 2.** Note that equation (15) follows immediately from Fubini's Theorem applied to the probability measure τ_d. The upper and lower bounds in (16) will now be established separately. To this end, we first make the following crucial remark : if $\mathcal{A}\subset\mathbb{S}^{d-1}$ is a σ_{d-1}–measurable set and $\boldsymbol{x_0}\in\mathbb{S}^{d-1}$, then

$$\sigma_{d-1}(\mathcal{A}) = \mu_d\left(\{G\in\mathcal{O}_d \,:\, G\boldsymbol{x_0}\in\mathcal{A}\}\right). \tag{30}$$

Indeed, each of the measures involved in this equation is clearly Borelian and uniformly distributed on the unit sphere (in the sense that the measure of a ball on the sphere depends only on the radius of the ball but not on the position of its centre). Now, a result of Christensen [3] states that two Borelian measures uniformly distributed in a separable metric space must be proportional. As the measures under consideration have been normalised to become probability measures, they must be equal — see [14, Chap. 3] for details.

Proof of the upper bound in (16). Let $\delta > 0$ and $\Delta \in \Delta_d^{++}$. The symmetry with respect of the origin and the convexity of the ellipsoid $\mathcal{E}_d\left(\sqrt{\delta}\Delta\right)$ imply that

$$\left\{P \in \mathcal{O}_d \ : \ P \cdot \mathbb{Z}^d \ \cap \ \mathcal{E}_d\left(\sqrt{\delta}\Delta\right) \ \neq \ \{\mathbf{0}\}\right\} \\ = \left\{P \in \mathcal{O}_d \ : \ P \cdot \mathcal{P}\left(\mathbb{Z}^d\right) \ \cap \ \mathcal{E}_d\left(\sqrt{\delta}\Delta\right) \ \neq \ \varnothing\right\}.$$

Given an event $\mathfrak{E}$, let $\chi_{\mathfrak{E}}$ denote the Boolean function

$$\chi_{[\mathfrak{E}]} \ = \ \begin{cases} 1 & \text{if } \mathfrak{E} \text{ holds} \\ 0 & \text{if } \mathfrak{E} \text{ does not holds.} \end{cases}$$

Then, denoting by $\#S$ the cardinality of a finite set S, one has

$$\begin{aligned} p_d\left(\mathcal{E}(\sqrt{\delta}\Delta)\right) \ &= \ \int_{\mathcal{O}_d} \mathrm{d}\mu_d(P) \cdot \chi_{\left[P \cdot \mathcal{P}(\mathbb{Z}^d) \, \cap \, \mathcal{E}_d\left(\sqrt{\delta}\Delta\right) \neq \varnothing\right]} \\ &\leqslant \ \int_{\mathcal{O}_d} \mathrm{d}\mu_d(P) \cdot \#\left(P \cdot \mathcal{P}\left(\mathbb{Z}^d\right) \ \cap \ \mathcal{E}_d\left(\sqrt{\delta}\Delta\right)\right) \\ &= \ \int_{\mathcal{O}_d} \mathrm{d}\mu_d(P) \cdot \left(\sum_{\boldsymbol{n} \in \mathcal{P}(\mathbb{Z}^d)} \chi_{\left[P\boldsymbol{n} \, \in \, \mathcal{E}_d\left(\sqrt{\delta}\Delta\right)\right]}\right). \end{aligned} \tag{31}$$

Now, given $P \in \mathcal{O}_d$ and $\boldsymbol{n} \in \mathcal{P}(\mathbb{Z}^d)$, it should be clear that

$$P\boldsymbol{n} \ \in \ \mathcal{E}_d\left(\sqrt{\delta}\Delta\right) \qquad \Longleftrightarrow \qquad P\frac{\boldsymbol{n}}{\|\boldsymbol{n}\|_2} \ \in \ \mathcal{E}_d\left(\frac{\sqrt{\delta}}{\|\boldsymbol{n}\|_2} \cdot \Delta\right) \cap \mathbb{S}^{d-1}.$$

For either of these statements to be true, it is furthermore necessary that

$$\|\boldsymbol{n}\|_2 \ \leqslant \ \sqrt{\delta} \cdot \|\Delta\|_\infty .$$

Therefore,

$$\begin{aligned} p_d\left(\mathcal{E}(\sqrt{\delta}\Delta)\right) \ &\leqslant \ \sum_{\substack{\boldsymbol{n} \in \mathcal{P}(\mathbb{Z}^d) \\ \|\boldsymbol{n}\|_2 \leqslant \sqrt{\delta}\|\Delta\|_\infty}} \mu_d\left(\left\{P \in \mathcal{O}_d \ : \ P\frac{\boldsymbol{n}}{\|\boldsymbol{n}\|_2} \ \in \ \mathcal{E}_d\left(\frac{\sqrt{\delta}}{\|\boldsymbol{n}\|_2} \cdot \Delta\right) \cap \mathbb{S}^{d-1}\right\}\right) \\ &\underset{(30)}{=} \ \sum_{\substack{\boldsymbol{n} \in \mathcal{P}(\mathbb{Z}^d) \\ \|\boldsymbol{n}\|_2 \leqslant \sqrt{\delta}\|\Delta\|_\infty}} \widetilde{\sigma}_{d-1}\left(\mathcal{E}_d\left(\frac{\sqrt{\delta}}{\|\boldsymbol{n}\|_2} \cdot \Delta\right)\right), \end{aligned}$$

hence the claim. □

Proof of the lower bound in (16). Let $\boldsymbol{e_1} = {}^t(1, 0, \ldots, 0) \in \mathbb{R}^d$ be the first element of the standard vector basis in $\mathbb{R}^d$. It then follows from (31) that

$$p_d\left(\mathcal{E}(\sqrt{\delta}\Delta)\right) \ \geqslant \ \mu_d\left(\left\{P \in \mathcal{O}_d \ : \ P\boldsymbol{e_1} \in \mathcal{E}_d\left(\sqrt{\delta}\Delta\right)\right\}\right) \\ \underset{(30)}{=} \ \widetilde{\sigma}_{d-1}\left(\mathcal{E}_d\left(\sqrt{\delta}\Delta\right)\right),$$

which establishes the first of the two inequalities to be proved.

The proof of the second one is more involved. Let $\boldsymbol{e_d} = {}^t(0,\dots,0,1) \in \mathbb{R}^d$ denote the last element of the standard vector basis in $\mathbb{R}^d$. Letting the group $\mathcal{O}_d$ act on the sphere $\mathbb{S}^{d-1}$, the stabiliser of $\boldsymbol{e_d}$ is isomorphic to $\mathcal{O}_{d-1}$ identified with the subgroup

$$\begin{pmatrix} \mathcal{O}_{d-1} & \mathbf{0} \\ {}^t\mathbf{0} & 1 \end{pmatrix} \subset \mathcal{O}_d.$$

With this identification, given $R, S \in \mathcal{O}_d$, the product $S^{-1}R$ lies in $\mathcal{O}_{d-1}$ if, and only if the last columns of R and S are the same, i.e.

$$S^{-1}R \in \mathcal{O}_{d-1} \quad \Longleftrightarrow \quad R\boldsymbol{e_d} = S\boldsymbol{e_d} \in \mathbb{S}^{d-1}.$$

This implies the well–known fact that the quotient $\mathcal{O}_d/\mathcal{O}_{d-1}$ is isomorphic to the sphere $\mathbb{S}^{d-1}$. Fix now a measurable function $f : \mathbb{S}^{d-1} \to \mathcal{O}_d$ such that

$$\forall \boldsymbol{v} \in \mathbb{S}^{d-1}, \quad f(\boldsymbol{v}) \cdot \boldsymbol{e_d} = \boldsymbol{v}. \tag{32}$$

Any $S \in \mathcal{O}_d$ can then be written uniquely in the form

$$S = f(\boldsymbol{v}) \cdot \begin{pmatrix} S' & \mathbf{0} \\ {}^t\mathbf{0} & 1 \end{pmatrix}, \tag{33}$$

where $S' \in \mathcal{O}_{d-1}$ and $\boldsymbol{v} \in \mathbb{S}^{d-1}$ (in particular, the last column of S is then $\boldsymbol{v}$).

Furthermore, if $R, S \in \mathcal{O}_d$ are respectively represented by $(R', \boldsymbol{u})$ and $(S', \boldsymbol{v})$ in these coordinates (where $R', S' \in \mathcal{O}_{d-1}$ and $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{S}^{d-1}$), then RS is represented by $(T'S', R\boldsymbol{v})$ for some $T' \in \mathcal{O}_{d-1}$ depending only on R and $\boldsymbol{v}$. Indeed, this follows from the uniqueness of the representation (33) together with (32) which implies that the last column of $R \cdot f(\boldsymbol{v})$ is $R\boldsymbol{v}$. Thus, identifying $\mathcal{O}_d$ with $\mathcal{O}_{d-1} \times \mathbb{S}^{d-1}$, left multiplication on $\mathcal{O}_d$ by some $R \in \mathcal{O}_d$ induces a left multiplication on $\mathcal{O}_{d-1}$ by some $T' \in \mathcal{O}_{d-1}$ (depending only on R and $\boldsymbol{v}$) and the orthogonal transformation on $\mathbb{S}^{d-1}$ induced by the action of R. This implies (see, e.g., [20] for details) that for any $S \in \mathcal{O}_d$, the volume element $\mathrm{d}\mu_d(S)$ is given in the coordinates $(S', \boldsymbol{v})$ by

$$\mathrm{d}\mu_d(S) = \frac{\mathrm{d}\boldsymbol{v}}{A_d} \cdot \mathrm{d}\mu_{d-1}(S') \tag{34}$$

(recall that $\mathrm{d}\boldsymbol{v}/A_d$ is the volume element of the uniform probability measure on the unit sphere).

Consider now the immersion

$$\iota : \boldsymbol{x} \in \mathbb{R}^{d-1} \mapsto {}^t({}^t\boldsymbol{x}, 0) \in \mathbb{R}^d.$$

Let $P = (P', \boldsymbol{w}) \in \mathcal{O}_d$ (with $P' \in \mathcal{O}_{d-1}$ and $\boldsymbol{w} \in \mathbb{S}^{d-1}$). It is then easily seen that

$$P \cdot \mathbb{Z}^d = \mathbb{Z}\boldsymbol{w} + f(\boldsymbol{w}) \cdot \iota\left(P' \cdot \mathbb{Z}^{d-1}\right) \supset f(\boldsymbol{w}) \cdot \iota\left(P' \cdot \mathbb{Z}^{d-1}\right).$$

This implies that

$$
\begin{aligned}
&p_d\left(\mathcal{E}(\sqrt{\delta}\Delta)\right) \underset{(34)}{=} \\
&\frac{1}{A_d}\cdot\int_{\mathbb{S}^{d-1}}\mathrm{d}\boldsymbol{w}\cdot\mu_{d-1}\left(\left\{P'\in\mathcal{O}_{d-1}\ :\ \left(\mathbb{Z}\boldsymbol{w}+f(\boldsymbol{w})\cdot\iota\left(P'\cdot\mathbb{Z}^{d-1}\right)\right)\cap\mathcal{E}_d\left(\sqrt{\delta}\Delta\right)\ \neq\ \{\mathbf{0}\}\right\}\right)\\
&\quad\geqslant\ \frac{1}{A_d}\cdot\int_{\mathbb{S}^{d-1}}\mathrm{d}\boldsymbol{w}\cdot\mu_{d-1}\left(\left\{P'\in\mathcal{O}_{d-1}\ :\ \left(f(\boldsymbol{w})\cdot\iota\left(P'\cdot\mathbb{Z}^{d-1}\right)\right)\cap\mathcal{E}_d\left(\sqrt{\delta}\Delta\right)\ \neq\ \{\mathbf{0}\}\right\}\right)\\
&\quad=\ \frac{1}{A_d}\cdot\int_{\mathbb{S}^{d-1}}\mathrm{d}\boldsymbol{w}\cdot\mu_{d-1}\left(\left\{P'\in\mathcal{O}_{d-1}\ :\ P'\cdot\mathbb{Z}^{d-1}\cap\mathcal{E}_d^{(\boldsymbol{w})}\left(\sqrt{\delta}\Delta\right)\ \neq\ \{\mathbf{0}\}\right\}\right),
\end{aligned}
$$

where

$$
\mathcal{E}_d^{(\boldsymbol{w})}\left(\sqrt{\delta}\Delta\right)\ :=\ \iota^{-1}\left(f(\boldsymbol{w})^{-1}\cdot\mathcal{E}_d\left(\sqrt{\delta}\Delta\right)\right).
$$

Since the set $\mathcal{E}_d\left(\sqrt{\delta}\Delta\right)\cap\boldsymbol{w}^{\perp}$ is sent to $\mathcal{E}_d^{(\boldsymbol{w})}\left(\sqrt{\delta}\Delta\right)$ by the linear isomorphism $\boldsymbol{x}\in\boldsymbol{w}^{\perp}\mapsto\iota^{-1}\left(f(\boldsymbol{w})^{-1}\cdot\boldsymbol{x}\right)$ which preserves μ_{d-1}–volumes, one obtains that

$$
p_d\left(\mathcal{E}(\sqrt{\delta}\Delta)\right)\ \geqslant\ \int_{\mathbb{S}^{d-1}}\frac{\mathrm{d}\boldsymbol{v}}{A_d}\cdot p_{d-1}\left(\mathcal{E}_d(\sqrt{\delta}\Delta)\cap\boldsymbol{v}^{\perp}\right).
$$

This concludes the proof of Theorem 2. □

2.4. **Proof of Proposition 1.** The proof of Proposition 1 is rather elementary and will be done in two steps.

We first seek to prove (17). To this end, it will be convenient to use the Kronecker symbol δ_{ij} which is equal to 1 if the integers i and j are equal and zero otherwise. Then, with the notation of Proposition 1, given $\boldsymbol{x}=(x_1,\ldots,x_d)\in\mathbb{R}^d$,

$$
\begin{aligned}
\boldsymbol{x}\in\mathcal{E}_{d-1}(\boldsymbol{\alpha},\boldsymbol{v})\ &\Longleftrightarrow\ \left(\sum_{i=1}^{d}\left(\frac{x_1}{\alpha_i}\right)^2\leqslant 1\right)\wedge\left(x_d=\frac{-1}{v_d}\cdot\sum_{i=1}^{d-1}x_iv_i\right)\\
&\Longleftrightarrow\ \frac{1}{(v_d\cdot\alpha_d)^2}\cdot\left(\sum_{i=1}^{d-1}x_iv_i\right)^2+\sum_{i=1}^{d-1}\left(\frac{x_i}{\alpha_i}\right)^2\ \leqslant\ 1\\
&\Longleftrightarrow\ \sum_{1\leqslant i,j\leqslant d-1}\left(\frac{\delta_{ij}}{\alpha_i^2}+\frac{v_iv_j}{(v_d\cdot\alpha_d)^2}\right)x_ix_j\ \leqslant\ 1\\
&\Longleftrightarrow\ {}^t\boldsymbol{y}\cdot Q\cdot\boldsymbol{y}\ \leqslant\ 1,
\end{aligned}
$$

where $\boldsymbol{y}={}^t(x_1,\ldots,x_{d-1})\in\mathbb{R}^{d-1}$ and where the matrix Q is defined in (18). Since Q is clearly definite positive, this establishes the first claim in Proposition 1.

To prove the second claim, denote by $R_{\boldsymbol{v}}\in SO_d(\mathbb{R})$ a rotation in $\mathbb{R}^d$ which maps the first vector $\boldsymbol{e_1}$ in the standard basis of $\mathbb{R}^d$ to $\boldsymbol{v}$. Let furthermore $Q_{\boldsymbol{\alpha}}:=(\alpha_1^{-2},\ldots,\alpha_d^{-2})\in\mathcal{D}_d^{++}$. Then, the d–dimensional ellipsoid $\mathcal{E}_d(\boldsymbol{\alpha})$ is congruent to the ellipsoid

$$
\widetilde{\mathcal{E}}_d^{(\boldsymbol{v})}(\boldsymbol{\alpha})\ :=\ \left\{\boldsymbol{x}\in\mathbb{R}^d\ :\ {}^t\boldsymbol{x}\cdot({}^tR_{\boldsymbol{v}}Q_{\boldsymbol{\alpha}}R_{\boldsymbol{v}})\cdot\boldsymbol{x}\leqslant 1\right\}
$$

and the $(d-1)$–dimensional ellipsoid $\mathcal{E}_{d-1}(\boldsymbol{\alpha}, \boldsymbol{v})$ becomes congruent to the ellipsoid $\widetilde{\mathcal{E}}_d^{(\boldsymbol{v})}(\boldsymbol{\alpha}) \cap \{x_1 = 0\}$ given by a positive definite matrix $Q_{\boldsymbol{\alpha}}^{(\boldsymbol{v})} \in \mathcal{S}_{d-1}^{++}$. This matrix $Q_{\boldsymbol{\alpha}}^{(\boldsymbol{v})}$ is obtained by stripping off the matrix ${}^tR_{\boldsymbol{v}} Q_{\boldsymbol{\alpha}} R_{\boldsymbol{v}}$ from its first row and first column. Let $\beta_{d-1}^{-2} \leqslant \cdots \leqslant \beta_1^{-2}$ denote the eigenvalues of $Q_{\boldsymbol{\alpha}}^{(\boldsymbol{v})}$ (in other words, $\beta_1, \ldots, \beta_{d-1}$ are the lengths of the semi–principal axes of the ellipsoid $\widetilde{\mathcal{E}}_d^{(\boldsymbol{v})}(\boldsymbol{\alpha}) \cap \{x_1 = 0\}$). It then follows from a direct application of the Cauchy Interlacing Inequalities that

$$\frac{1}{\alpha_d^2} \leqslant \frac{1}{\beta_{d-1}^2} \leqslant \cdots \leqslant \frac{1}{\beta_1^2} \leqslant \frac{1}{\alpha_1^2},$$

which completes the proof of Proposition 1.

2.5. **Proof of Proposition 2.** Before proving Proposition 2, we make a crucial remark which will be used several times hereafter. Fix $\boldsymbol{\alpha} \in \mathbb{R}^d$ satisfying (19). Let

$$\mathcal{A}_d(\boldsymbol{\alpha}) \,:=\, \mathcal{E}_d(\boldsymbol{\alpha}) \cap \mathbb{S}^{d-1} \tag{35}$$

and $\boldsymbol{x} := (x_1, \ldots, x_d) \in \mathbb{R}^d$. Then,

$$\begin{aligned} \boldsymbol{x} \in \mathcal{A}_d(\boldsymbol{\alpha}) \iff & \left(\sum_{i=1}^d \left(\frac{x_1}{\alpha_i} \right)^2 \leqslant 1 \right) \wedge \left(\sum_{i=1}^d x_i^2 = 1 \right) \\ \iff & \left(\sum_{i=1}^{d-1} x_i^2 \cdot \left(\frac{1}{\alpha_i^2} - \frac{1}{\alpha_d^2} \right) \leqslant 1 - \frac{1}{\alpha_d^2} \right) \wedge \left(\sum_{i=1}^d x_i^2 = 1 \right). \end{aligned}$$

Given $\boldsymbol{\mu} \in (\mathbb{R}_{>0})^{d-1}$, let $\mathcal{C}_d(\boldsymbol{\mu})$ denote the *full* cylinder with axis spanned by $\boldsymbol{e_d}$ whose section with the hyperplane $\{x_d = 0\}$ is the $(d-1)$–dimensional ellipsoid $\mathcal{E}_{d-1}(\boldsymbol{\mu})$. With the notation of Proposition 2, the above chain of equivalences thus amounts to claiming that

$$\mathcal{A}_d(\boldsymbol{\alpha}) \,=\, \mathcal{C}_d(\underset{\sim}{\boldsymbol{\alpha}}) \cap \mathbb{S}^{d-1}. \tag{36}$$

Proof of Proposition 2. Note first that the relations (22) are trivial. Indeed, under (19), $\mathcal{A}_d(\boldsymbol{\alpha}) = \mathbb{S}^{d-1}$ if $\alpha_1 \geqslant 1$ and $\#\mathcal{A}_d(\boldsymbol{\alpha}) \leqslant 2$ if $\alpha_d \leqslant 1$. Assume therefore that $\alpha_1 < 1 < \alpha_d$. Parameter a dense open set in $\mathbb{S}^{d-1}$ as follows :

$$\boldsymbol{v} \,=\, (\boldsymbol{u} \cdot \sin\theta,\, \cos\theta)\,,$$

where $\boldsymbol{u} \in \mathbb{S}^{d-2}$ and $\theta \in (0, \pi)$ (θ is thus the angle between $\boldsymbol{u}$ and $\boldsymbol{e_d}$). A standard calculation shows that, in these coordinates, the volume element $\mathrm{d}\boldsymbol{v}$ reads $\mathrm{d}\boldsymbol{v} = (\sin\theta)^{d-2} \cdot \mathrm{d}\theta \cdot \mathrm{d}\boldsymbol{u}$ (if $d = 2$, $\mathrm{d}\boldsymbol{u}$ is the counting probability measure on $\mathbb{S}^0 = \{\pm 1\}$). Therefore,

$$\widetilde{\sigma}_{d-1}\left(\mathcal{E}_d(\boldsymbol{\alpha})\right) \,=\, \frac{1}{A_d} \int_0^\pi \mathrm{d}\theta \cdot (\sin\theta)^{d-2} \int_{\mathbb{S}^{d-2}} \chi_{[(\boldsymbol{u}\cdot\sin\theta,\, \cos\theta) \,\in\, \mathcal{A}_d(\boldsymbol{\alpha})]} \cdot \mathrm{d}\boldsymbol{u}.$$

In view of (35) and (36), the intersection of $\mathcal{A}_d(\boldsymbol{\alpha})$ with the hyperplane $\{x_d = \cos\theta\}$ is obtained as the intersection of the $(d-1)$–dimensional ellipsoid $\mathcal{E}_{d-1}(\underset{\sim}{\boldsymbol{\alpha}})$ with the

$(d-1)$–dimensional unit sphere centred at the origin with radius $\sin\theta$:

$$\boldsymbol{x} \in \mathcal{A}_d(\boldsymbol{\alpha}) \cap \{x_d = \cos\theta\} \iff \left(\sum_{i=1}^{d-1}\left(\frac{x_1}{\underset{\sim}{\alpha}_i}\right)^2 \leqslant 1\right) \wedge \left(\sum_{i=1}^{d-1} x_i^2 = \sin^2\theta\right) \wedge (x_d = \cos\theta)\,.$$

This implies that, given $\boldsymbol{u} \in \mathbb{S}^{d-2}$ and $\theta \in (0,\pi)$,

$$(\boldsymbol{u}\cdot\sin\theta,\,\cos\theta) \,\in\, \mathcal{A}_d(\boldsymbol{\alpha}) \iff \boldsymbol{u} \,\in\, \mathcal{E}_{d-1}\left(\frac{\underset{\sim}{\boldsymbol{\alpha}}}{\sin\theta}\right).$$

Thus :

$$\begin{aligned}\widetilde{\sigma}_{d-1}\left(\mathcal{E}_d(\boldsymbol{\alpha})\right) &= \frac{1}{A_d}\int_0^\pi \mathrm{d}\theta\cdot(\sin\theta)^{d-2}\int_{\mathbb{S}^{d-2}} \chi_{\left[\boldsymbol{u}\,\in\,\mathcal{E}_{d-1}\left(\frac{\underset{\sim}{\alpha}}{\sin\theta}\right)\right]}\cdot\mathrm{d}\boldsymbol{u}\\ &= \frac{A_{d-1}}{A_d}\cdot\int_0^\pi \mathrm{d}\theta\cdot(\sin\theta)^{d-2}\cdot\widetilde{\sigma}_{d-2}\left(\mathcal{E}_{d-1}\left(\frac{\boldsymbol{\alpha}}{\underset{\sim}{\sin\theta}}\right)\right).\end{aligned}$$

The result then follows from (4) and (21). □

2.6. **Proof of Proposition 3.** The proof of Proposition 3 rests on the following lemma. Throughout, we adopt the notation introduced before the statement of Proposition 3 and fix $\boldsymbol{\alpha} \in \mathbb{R}^d$ satisfying (19) and the inequalities $\alpha_1 < 1 < \alpha_d$. Let furthermore

$$K_d(\boldsymbol{\alpha}) := \prod_{i=1}^{d}[-\alpha_i,\,\alpha_i]\,.$$

Lemma 2. *The following equation holds :*

$$\widetilde{\sigma}_{d-1}\left(K_d(\boldsymbol{\alpha})\right) = \mathfrak{I}_d\left(\boldsymbol{\alpha}\right).$$

Furthermore, one has also the estimates

$$\mathfrak{L}_d\left(\boldsymbol{\alpha}\right)\cdot\left(\frac{2}{\pi}\right)^{d-2} \leqslant \mathfrak{I}_d\left(\boldsymbol{\alpha}\right) \leqslant \mathfrak{L}_d\left(\boldsymbol{\alpha}\right)$$

with

$$\mathfrak{L}_d\left(\boldsymbol{\alpha}\right) := \frac{2^d}{(d-1)!\cdot A_d}\cdot\prod_{j=1}^{d-1}\left(\left(\frac{\pi}{2}\right)^j - b\left(\alpha_{d-j}\right)^j\right).$$

Proof. Parametrise the unit sphere in spherical coordinates by defining the coordinates of $\boldsymbol{v} := \boldsymbol{v_d} \in \mathbb{S}^{d-1}$ by induction in the following way :

$$\boldsymbol{v_d} = (\cos\theta_1,\,\boldsymbol{v_{d-1}}\cdot\sin\theta_1)\,,$$

where $\boldsymbol{v_k} \in \mathbb{S}^{k-1}$ for $k = 2,\ldots,d-1$. Here, the base case is $\boldsymbol{v_2} = (\cos\theta_{d-1},\,\sin\theta_{d-1}) \in \mathbb{S}^1$. Thus, given $i = 1,\ldots,d-1$, the real number θ_i is the angle between $\boldsymbol{v}$ and the i^{th} standard vector basis $\boldsymbol{e_i}$ of $\mathbb{R}^d$. These angles θ_i are unique upon requiring that $\theta_i \in [0,\pi]$ for $i = 1,\ldots,d-2$ and $\theta_{d-1} \in [0,2\pi)$. Upon taking into account the notation convention

adopted here to label the angles, the volume element $\mathrm{d}\boldsymbol{v}$ is then given by the usual formula

$$\mathrm{d}\boldsymbol{v} \;=\; \frac{1}{A_d}\cdot\prod_{i=2}^{d}\sin^{i-2}\theta_{d-i+1}\cdot\mathrm{d}\theta_{d-i+1}.$$

Thus, given $\boldsymbol{v}\in\mathbb{R}^d$ with (cartesian) coordinates $(x_1,\ldots,x_d)$,

$$\begin{aligned}\boldsymbol{v}\in K_d(\boldsymbol{\alpha})\cap\mathbb{S}^{d-1} &\iff \forall i\in[\![1,d]\!],\ |x_i|=|\cos\theta_i|\leqslant\alpha_i\\ &\underset{(\alpha_d>1)}{\iff} \forall i\in[\![1,d-1]\!],\ |\cos\theta_i|\leqslant\alpha_i\\ &\iff \begin{cases}\forall i\in[\![1,d-2]\!],\ \theta_i\in\big[b(\alpha_i),\,\pi-b(\alpha_i)\big],\\ \theta_{d-1}\in\big[b(\alpha_{d-1}),\,\pi-b(\alpha_{d-1})\big]\cup\big[\pi+b(\alpha_{d-1}),\,2\pi-b(\alpha_{d-1})\big]\end{cases}\end{aligned}$$

(with obvious changes for the bounds of the latter intervals when $b(\alpha_{d-1})=0$). Therefore,

$$\begin{aligned}\widetilde{\sigma}_{d-1}\left(K_d\left(\boldsymbol{\alpha}\right)\right) &= \frac{1}{A_d}\cdot 2(\pi-2b(\alpha_{d-1}))\cdot\prod_{i=3}^{d}\int_{b(\alpha_{d-i+1})}^{\pi-b(\alpha_{d-i+1})}\sin^{i-2}\theta\cdot\mathrm{d}\theta\\ &= \frac{2^d}{A_d}\cdot\Big(\frac{\pi}{2}-b(\alpha_{d-1})\Big)\cdot\prod_{i=3}^{d}\int_{b(\alpha_{d-i+1})}^{\pi/2}\sin^{i-2}\theta\cdot\mathrm{d}\theta\\ &\underset{(24)}{=}\ \mathfrak{I}_d(\boldsymbol{\alpha}).\end{aligned}$$

The estimates involving $\mathfrak{L}_d(\boldsymbol{\alpha})$ follow now straightforwardly from the definition of $\mathfrak{I}_d(\boldsymbol{\alpha})$ and from the convexity inequalities $(2/\pi)\cdot t\leqslant\sin t\leqslant t$ valid for any $t\in[0,\pi/2]$. □

Proof of Proposition 3. It plainly follows from the definition of the ellipsoid $\mathcal{E}_d\left(\boldsymbol{\alpha}\right)$ in (13) that

$$\prod_{i=1}^{d}\left[-\frac{\alpha_i}{\sqrt{d}},\,\frac{\alpha_i}{\sqrt{d}}\right]\;\subset\;\mathcal{E}_d\left(\boldsymbol{\alpha}\right)\;\subset\;\prod_{i=1}^{d}\left[-\alpha_i,\,\alpha_i\right]. \tag{37}$$

Also, relations (35) and (36) imply that

$$\left(\prod_{i=1}^{d-1}\left[-\frac{\underset{\sim}{\alpha}_i^*}{\sqrt{d-1}},\,\frac{\underset{\sim}{\alpha}_i^*}{\sqrt{d-1}}\right]\right)\times[-1,\,1]\;\subset\;\mathcal{A}_d\left(\boldsymbol{\alpha}\right)\;\subset\;\left(\prod_{i=1}^{d-1}\left[-\underset{\sim}{\alpha}_i^*,\,\underset{\sim}{\alpha}_i^*\right]\right)\times[-1,\,1] \tag{38}$$

(this is because the basis of the cylinder $\mathcal{C}_d\left(\underset{\sim}{\boldsymbol{\alpha}}\right)$ is the ellipsoid $\mathcal{E}_{d-1}\left(\underset{\sim}{\boldsymbol{\alpha}}\right)$).

Thus, the estimates for $\widetilde{\sigma}_{d-1}\left(\mathcal{E}_d\left(\boldsymbol{\alpha}\right)\right)$ in Proposition 3 become straightforward consequences of relations (37) and (38) and of Lemma 2. As for the bounds for $\mathfrak{I}_d\left(\boldsymbol{\alpha}\right)$ therein, they also follow from Lemma 2 and from the inequalities

$$\left(\frac{\pi}{2}\right)^{j-1}\cdot\min\left\{1,\alpha_{d-j}\right\}\;\leqslant\;\left(\frac{\pi}{2}\right)^{j}-b(\alpha_{d-j})^j\;\leqslant\;j\cdot\left(\frac{\pi}{2}\right)^{j}\cdot\min\left\{1,\alpha_{d-j}\right\}.$$

The latter is a direct consequence of the convexity inequalities

$$x\leqslant\frac{\pi}{2}-\arccos x\leqslant\frac{\pi}{2}x$$

valid for all $x \in [0,1]$ and of the factorisation identity

$$\left(\frac{\pi}{2}\right)^j - b(\alpha_{d-j})^j \;=\; \left(\frac{\pi}{2} - b(\alpha_{d-j})\right) \cdot \sum_{k=0}^{j-1} \left(\frac{\pi}{2}\right)^{j-1-k} b(\alpha_{d-j})^k.$$

□

2.7. **Proof of Theorem 3.** Let $\epsilon > 0$ and let $\Delta := (\alpha_1, \dots, \alpha_d) \in \Delta_d^{++}$ be such that the vector $\boldsymbol{\alpha}' := (\alpha_1, \dots, \alpha_{d-1})$ lies in the support of the measure $\nu_d^{(\epsilon)}$ as defined in (25) (i.e. $\epsilon \leqslant \alpha_i \leqslant \epsilon^{-1}$ for all $i = 1, \dots, d-1$). This clearly implies that $\|\Delta\|_\infty \leqslant \epsilon^{-d+1}$. In particular, in view of the upper bound in (16), the probability $\tau_d^{(\epsilon)}\left(\mathfrak{F}(\delta)\right)$ vanishes whenever $\sqrt{\delta} \cdot \epsilon^{-d+1} < 1$, i.e. whenever $\delta < \epsilon^{2\cdot(d-1)}$. Since $\nu_d^{(\epsilon)}\left(\Delta_d^{++} \backslash \Delta_{d,sub}^{++}\right) = 0$, the same conclusion holds if $\delta = \epsilon^{2\cdot(d-1)}$. This establishes (26).

Assume from now on that $\delta > \epsilon^{2(d-1)}$. The goal is to bound from below and above the probability

$$\tau_d^{(\epsilon)}\left(\mathfrak{F}_d(\delta)\right) \;=\; \frac{1}{\left|2\log\epsilon\right|^{d-1}} \cdot \int_{[\epsilon,\,\epsilon^{-1}]^{d-1}} \prod_{i=1}^{d-1} \frac{\mathrm{d}\alpha_i}{\alpha_i} \cdot p_d\left(\mathcal{E}\left(\sqrt{\delta}\Delta\right)\right).$$

Upon reordering the coordinates of the vector Δ as defined above, it follows from the invariance of the quantity $p_d\left(\mathcal{E}\left(\sqrt{\delta}\Delta\right)\right)$ under such permutation that

$$\begin{aligned}\frac{(d-1)!}{\left|2\log\epsilon\right|^{d-1}} \cdot \int_{\substack{\epsilon \leqslant \alpha_1 < \dots < \alpha_{d-1} \leqslant \epsilon^{-1} \\ \alpha_{d-1} < \alpha_d := (\alpha_1 \dots \alpha_{d-1})^{-1}}} \prod_{i=1}^{d-1} \frac{\mathrm{d}\alpha_i}{\alpha_i} \cdot p_d\left(\mathcal{E}\left(\sqrt{\delta}\Delta\right)\right) \;&\leqslant\; \tau_d^{(\epsilon)}\left(\mathfrak{F}_d(\delta)\right) \\ &\leqslant\; \frac{d!}{\left|2\log\epsilon\right|^{d-1}} \cdot \int_{\substack{\epsilon \leqslant \alpha_1 < \dots < \alpha_{d-1} \leqslant \epsilon^{-1} \\ \alpha_{d-1} < \alpha_d := (\alpha_1 \dots \alpha_{d-1})^{-1}}} \prod_{i=1}^{d-1} \frac{\mathrm{d}\alpha_i}{\alpha_i} \cdot p_d\left(\mathcal{E}\left(\sqrt{\delta}\Delta\right)\right).\end{aligned}$$

Here, we are using two facts to obtain the upper bound : on the one hand, if σ is a permutation of $[\![1,\, d]\!]$ such that, given a d–tuple $(\alpha_1, \dots, \alpha_d)$, $\alpha_{\sigma(1)} \leqslant \dots \leqslant \alpha_{\sigma(d)}$, then $\prod_{i=1}^{d-1} \alpha_i^{-1} \leqslant \prod_{i=i}^{d-1} \alpha_{\sigma(i)}^{-1}$; on the other, given a d-tuple $(\beta_1, \dots, \beta_d)$ such that $\beta_1 < \dots < \beta_d$, there are $d!$ d–tuples $(\alpha_1, \dots, \alpha_d)$ for which there exists a permutation σ such that $\alpha_{\sigma(1)} = \beta_1, \dots, \alpha_{\sigma(d)} = \beta_d$. The lower bound follows from a similar argument : given a d-tuple $(\beta_1, \dots, \beta_d)$ such that $\beta_1 < \dots < \beta_d$, there are $(d-1)!$ d–tuples $(\alpha_1, \dots, \alpha_d)$ for which there exists a permutation σ of $[\![1, d-1]\!]$ such that $\alpha_{\sigma(1)} = \beta_1, \dots, \alpha_{\sigma(d-1)} = \beta_{d-1}$ and $\alpha_d = \beta_d = \max_{1 \leqslant i \leqslant d} \beta_i$.

Note that in the domain of integration,

$$\|\Delta\|_\infty = \alpha_d = (\alpha_1 \dots \alpha_{d-1})^{-1}. \tag{39}$$

Since from Proposition 2, $p_d\left(\mathcal{E}\left(\sqrt{\delta}\Delta\right)\right) = 0$ whenever $\sqrt{\delta}\cdot\alpha_d \leqslant 1$, one has also

$$\frac{(d-1)!}{\left|2\log\epsilon\right|^{d-1}}\cdot\int\limits_{\substack{\epsilon\leqslant\alpha_1<\cdots<\alpha_{d-1}\leqslant\epsilon^{-1}\\ \max\left\{\delta^{-1/2},\,\alpha_{d-1}\right\}<(\alpha_1\cdots\alpha_{d-1})^{-1}}}\prod_{i=1}^{d-1}\frac{\mathrm{d}\alpha_i}{\alpha_i}\cdot p_d\left(\mathcal{E}\left(\sqrt{\delta}\Delta\right)\right)\leqslant\tau_d^{(\epsilon)}\left(\mathfrak{F}_d(\delta)\right)\tag{40}$$

$$\leqslant\frac{d!}{\left|2\log\epsilon\right|^{d-1}}\cdot\int\limits_{\substack{\epsilon\leqslant\alpha_1<\cdots<\alpha_{d-1}\leqslant\epsilon^{-1}\\ \max\left\{\delta^{-1/2},\,\alpha_{d-1}\right\}<(\alpha_1\cdots\alpha_{d-1})^{-1}}}\prod_{i=1}^{d-1}\frac{\mathrm{d}\alpha_i}{\alpha_i}\cdot p_d\left(\mathcal{E}\left(\sqrt{\delta}\Delta\right)\right).\tag{41}$$

We now call on Theorem 2 to bound the probability $p_d\left(\mathcal{E}\left(\sqrt{\delta}\Delta\right)\right)$ as follows :

$$\widetilde{\sigma}_{d-1}\left(\mathcal{E}_d(\sqrt{\delta}\Delta)\right)\leqslant p_d\left(\mathcal{E}\left(\sqrt{\delta}\Delta\right)\right)\leqslant\sum_{\substack{\boldsymbol{n}\in\mathbb{Z}^d\setminus\{\mathbf{0}\}\\ \|\boldsymbol{n}\|_\infty\leqslant\sqrt{\delta}\cdot\|\Delta\|_\infty}}\widetilde{\sigma}_{d-1}\left(\mathcal{E}\left(\frac{\sqrt{\delta}}{\|\boldsymbol{n}\|_2}\cdot\Delta\right)\right).\tag{42}$$

Furthermore, from Proposition 3,

$$\widetilde{\sigma}_{d-1}\left(\mathcal{E}_d(\sqrt{\delta}\Delta)\right)\geqslant a(d)\cdot\prod_{i=1}^{d-1}\min\left\{\frac{\sqrt{\delta}\cdot\alpha_i}{d},\,1\right\}\geqslant\frac{a(d)}{d^{d-1}}\cdot\prod_{i=1}^{d-1}\min\left\{\sqrt{\delta}\cdot\alpha_i,\,1\right\}.\tag{43}$$

Given the domain of integration of the integrals above, one has also

$$\begin{aligned}\sum_{\substack{\boldsymbol{n}\in\mathbb{Z}^d\setminus\{\mathbf{0}\}\\ \|\boldsymbol{n}\|_\infty\leqslant\sqrt{\delta}\cdot\|\Delta\|_\infty}}\widetilde{\sigma}_{d-1}\left(\mathcal{E}\left(\frac{\sqrt{\delta}}{\|\boldsymbol{n}\|_2}\cdot\Delta\right)\right)&\leqslant\sum_{\substack{\boldsymbol{n}\in\mathbb{Z}^d\setminus\{\mathbf{0}\}\\ \|\boldsymbol{n}\|_\infty\leqslant\sqrt{\delta}\cdot\|\Delta\|_\infty}}a'(d)\cdot\prod_{i=1}^{d-1}\min\left\{\frac{\sqrt{\delta}\cdot\alpha_i}{\|\boldsymbol{n}\|_2},\,1\right\}\\ &\leqslant a'(d)\cdot\delta^{(d-1)/2}\cdot\left(\prod_{i=1}^{d-1}\alpha_i\right)\cdot\left(\sum_{\substack{\boldsymbol{n}\in\mathbb{Z}^d\setminus\{\mathbf{0}\}\\ \|\boldsymbol{n}\|_\infty\leqslant\sqrt{\delta}\cdot\|\Delta\|_\infty}}\frac{1}{\|\boldsymbol{n}\|_\infty^{d-1}}\right)\\ &\leqslant a'(d)\cdot\delta^{(d-1)/2}\cdot\left(\prod_{i=1}^{d-1}\alpha_i\right)\cdot\left(\sum_{k=1}^{\sqrt{\delta}\cdot\|\Delta\|_\infty}d\cdot\frac{(2k+1)^{d-1}}{k^{d-1}}\right)\\ &\leqslant a'(d)\cdot\delta^{(d-1)/2}\cdot\left(\prod_{i=1}^{d-1}\alpha_i\right)\cdot\left(3^{d-1}\cdot d\cdot\sqrt{\delta}\cdot\|\Delta\|_\infty\right)\\ &\underset{(39)}{\leqslant}3^{d-1}\cdot a'(d)\cdot d\cdot\delta^{d/2}.\end{aligned}\tag{44}$$

Inequalities (27) thus turn out to be a rephrasing of the relations (40)—(44) with the constants $c_d(\epsilon)$ and $C_d(\epsilon)$ stated in the theorem.

As for inequalities (28) and (29), note first that, on the one hand,

$$s_d(\epsilon,\delta) \;\geqslant\; \min\left\{\sqrt{\delta},\epsilon\right\}^{d-1} \cdot \int_{\substack{\epsilon\leqslant\alpha_1<\cdots<\alpha_{d-1}\leqslant\epsilon^{-1}\\ \max\left\{\epsilon^{-1},\delta^{-1/2}\right\}<(\alpha_1\ldots\alpha_{d-1})^{-1}}} \mathrm{d}\alpha_1\ldots\mathrm{d}\alpha_{d-1}$$

and that, on the other,

$$S_d(\epsilon,\delta) \;\leqslant\; \delta^{d/2} \cdot \int_{\substack{\epsilon\leqslant\alpha_1<\cdots<\alpha_{d-1}\leqslant\epsilon^{-1}\\ \delta^{-1/2}<(\alpha_1\ldots\alpha_{d-1})^{-1}}} \prod_{i=1}^{d-1}\frac{\mathrm{d}\alpha_i}{\alpha_i}.$$

Now, given any $c>0$, the change of variables $y_i=\alpha_i$ for $1\leqslant i\leqslant d-2$ and $y_{d-1}=\prod_{i=1}^{d-1}\alpha_i$ shows that

$$\begin{aligned}\int_{\substack{\epsilon\leqslant\alpha_1<\cdots<\alpha_{d-1}\leqslant\epsilon^{-1}\\ c<(\alpha_1\ldots\alpha_{d-1})^{-1}}} \mathrm{d}\alpha_1\ldots\mathrm{d}\alpha_{d-1} \;&=\; \int_{\substack{\epsilon\leqslant y_1<\cdots<y_{d-2}\leqslant\epsilon^{-1}\\ \epsilon^{d-1}<y_{d-1}<c^{-1}}} \mathrm{d}y_{d-1}\cdot\prod_{i=1}^{d-2}\frac{\mathrm{d}y_i}{y_i}\\ &=\; \frac{\left|2\log\epsilon\right|^{d-2}}{(d-2)!}\cdot\left(c^{-1}-\epsilon^{d-1}\right)\end{aligned}$$

and that

$$\begin{aligned}\int_{\substack{\epsilon\leqslant\alpha_1<\cdots<\alpha_{d-1}\leqslant\epsilon^{-1}\\ c<(\alpha_1\ldots\alpha_{d-1})^{-1}}} \prod_{i=1}^{d-1}\frac{\mathrm{d}\alpha_i}{\alpha_i} \;&=\; \int_{\substack{\epsilon\leqslant y_1<\cdots<y_{d-2}\leqslant\epsilon^{-1}\\ \epsilon^{d-1}<y_{d-1}<c^{-1}}} \prod_{i=1}^{d-1}\frac{\mathrm{d}y_i}{y_i}\\ &=\; \frac{\left|2\log\epsilon\right|^{d-2}}{(d-2)!}\cdot\log\left(\frac{c^{-1}}{\epsilon^{d-1}}\right).\end{aligned}$$

This completes the proof of Theorem 3.

3. An Approach via the Cholesky Decomposition.

The probabilistic approach via the spectral decomposition exposed in §2 requires that the probability measures under consideration be essentially defined from the set of eigenvalues of a given element in Σ_d^{++}. While this should not be seen as a big restriction in view of the spectral decomposition and of the fact that the orthogonal group is compact, the determination of the eigenvalues of a matrix is known to be a hard task. We therefore adopt here an alternative approach based on the Cholesky decomposition of a quadratic form in Σ_d^{++} or, in view of Problem 2, on the Cholesky decomposition of a quadratic form in $\mathcal{S}_d^{++}$.

Let $\mathcal{T}_d^{++}$ be the group of upper triangular matrices with strictly positive diagonal entries. Let Θ_d^{++} be the subgroup of $\mathcal{T}_d^{++}$ consisting of all those matrices with determinant one :

$$\Theta_d^{++} := \mathcal{T}_d^{++}\cap SL_d(\mathbb{R}). \tag{45}$$

Let

$$p := \frac{d(d-1)}{2}. \tag{46}$$

The set $\mathcal{T}_d^{++}$ shall be identified with $(\mathbb{R}_{>0})^d \times \mathbb{R}^p$ by splitting a matrix therein between its d diagonal terms and the remaining p off–diagonal upper coefficients. A generic element in $\mathcal{T}_d^{++}$ shall thus be represented as $(\boldsymbol{\beta}, \boldsymbol{u})$ with $\boldsymbol{\beta} \in (\mathbb{R}_{>0})^d$ and $\boldsymbol{u} \in \mathbb{R}^p$, in which case it will be convenient to adopt the notation

$$\boldsymbol{\beta} := (\beta_1, \underset{\sim}{\boldsymbol{\beta}})$$

with $\beta_1 \in \mathbb{R}$ and $\underset{\sim}{\boldsymbol{\beta}} \in \mathbb{R}^{d-1}$ (this notation is independent from (20)). In the same way, the set Θ_d^{++} shall be identified with $(\mathbb{R}_{>0})^{d-1} \times \mathbb{R}^p$. A generic element of Θ_d^{++} shall thus be represented as $(\boldsymbol{\beta}', \boldsymbol{u})$ with $\boldsymbol{\beta}' \in (\mathbb{R}_{>0})^{d-1}$ and $\boldsymbol{u} \in \mathbb{R}^p$, in which case it will be convenient to adopt the notation

$$\boldsymbol{\beta}' := (\beta_1', \underset{\sim}{\boldsymbol{\beta}}')$$

with $\beta_1' \in \mathbb{R}$ and $\underset{\sim}{\boldsymbol{\beta}}' \in \mathbb{R}^{d-2}$. When a matrix in Θ_d^{++} is seen as an element of $\mathcal{T}_d^{++}$, it shall also be given as a vector from $(\mathbb{R}_{>0})^d \times \mathbb{R}^p$. This should not cause any confusion.

The Cholesky decomposition of a positive definite matrix amounts to claiming that the map

$$\varphi_{chol} \ : \ L \in \mathcal{T}_d^{++} \ \mapsto \ {}^tLL \in \mathcal{S}_d^{++} \tag{47}$$

is bijective. This implies in particular that the map

$$\widetilde{\varphi}_{chol} \ : \ L \in \Theta_d^{++} \ \mapsto \ {}^tLL \in \Sigma_d^{++} \tag{48}$$

is also bijective. Determining the Cholesky decomposition of a given positive definite matrix is a problem which has been extensively studied from an algorithmic point of view and which can be implemented in a very efficient way — see, e.g., [19] for details.

3.1. **Definition of a Suitable Class of Measures.** Note that $\mathcal{S}_d^{++}$ sits as an open cone in the space of symmetric matrices in dimension d. It is a $(p+d)$–dimensional manifold (with p as defined in (46)) and any matrix therein can be identified with a vector in $\mathbb{R}^{p+d}$ by considering its upper triangular part. Similarly, Σ_d^{++} sits as a $(p+d-1)$–dimensional manifold in $\mathcal{S}_d^{++}$ which can be identified with a subset of $\mathbb{R}^{p+d-1}$ by considering the upper triangular part of a matrix therein minus the bottom right coefficient. For a rigorous justification of the fact that this indeed gives a system of independent coordinates, see (the proof of) Lemma 3 in §3.4 below.

With the help of these identifications, we will be concerned with measures supported on $\mathcal{S}_d^{++}$ (resp. on Σ_d^{++}) absolutely continuous with respect to the $(p+d)$–dimensional Lebesgue measure λ_{p+d} (resp. with respect to the $(p+d-1)$–dimensional Lebesgue measure λ_{p+d-1}).

Let then $f : \mathcal{S}_d^{++} \to \mathbb{R}_+$ (resp. $\widetilde{f} : \Sigma_d^{++} \to \mathbb{R}_+$) be a density function supported on $\mathcal{S}_d^{++}$ (resp. on Σ_d^{++}). The corresponding measure is denoted by ν_f (resp. by $\widetilde{\nu}_{\widetilde{f}}$).

3.2. **The Main Estimates.** Given $\delta > 0$, the quantities of interest are

$$m_f(\delta) := \nu_f\left(\left\{Q \in \mathcal{S}_d^{++} \ : \ M_d(Q) \leqslant \delta\right\}\right) \tag{49}$$

and

$$\widetilde{m}_{\widetilde{f}}(\delta) := \widetilde{\nu}_{\widetilde{f}}\left(\left\{\Sigma \in \Sigma_d^{++} \ : \ M_d(\Sigma) \leqslant \delta\right\}\right).$$

Given any $\boldsymbol{\beta} \in (\mathbb{R}_{>0})^d$, define

$$G_f(\boldsymbol{\beta}) := 2^d \cdot \prod_{i=1}^{d} \beta_i^{d-i+1} \cdot \int_{\mathbb{R}^p} (f \circ \varphi_{chol})(\boldsymbol{\beta}, \boldsymbol{u}) \cdot \mathrm{d}\lambda_p(\boldsymbol{u})$$

and, given any $\beta_1 > 0$, let

$$g_f(\beta_1) := \int_{(\mathbb{R}_{>0})^{d-1}} G_f(\beta_1, \underset{\sim}{\boldsymbol{\beta}}) \cdot \mathrm{d}\lambda_{d-1}(\underset{\sim}{\boldsymbol{\beta}}). \tag{50}$$

Similarly, given any $\boldsymbol{\beta}' \in (\mathbb{R}_{>0})^{d-1}$, define

$$\widetilde{G}_{\widetilde{f}}(\boldsymbol{\beta}') := 2^{d-1} \cdot \prod_{i=1}^{d-1} \beta_i^{d-i+1} \cdot \int_{\mathbb{R}^p} \left(\widetilde{f} \circ \widetilde{\varphi}_{chol}\right)(\boldsymbol{\beta}', \boldsymbol{u}) \cdot \mathrm{d}\lambda_p(\boldsymbol{u})$$

and, given any $\beta_1' > 0$, let

$$\widetilde{g}_{\widetilde{f}}(\beta_1') := \int_{(\mathbb{R}_{>0})^{d-2}} \widetilde{G}_{\widetilde{f}}(\beta_1', \underset{\sim}{\boldsymbol{\beta}'}) \cdot \mathrm{d}\lambda_{d-2}(\underset{\sim}{\boldsymbol{\beta}'}).$$

With these definitions, the main theorem in this section reads as follows :

Theorem 4. *Let $\delta \in (0,\, 1)$. Then,*

$$0 \leqslant 1 - \int_{\sqrt{\delta}}^{\infty} g_f \leqslant m_f(\delta) \leqslant 1 - \int_{I_d(\delta)} G_f \leqslant 1, \tag{51}$$

where

$$I_d(\delta) := \left(\sqrt{\delta},\, +\infty\right)^d.$$

Furthermore, one has also the estimates

$$0 \leqslant 1 - \int_{\sqrt{\delta}}^{\infty} \widetilde{g}_{\widetilde{f}} \leqslant \widetilde{m}_{\widetilde{f}}(\delta) \leqslant 1 - \int_{\Delta_{d-1}(\delta)} \widetilde{G}_{\widetilde{f}} \leqslant 1, \tag{52}$$

where

$$\Delta_{d-1}(\delta) := \left\{\boldsymbol{\beta}' \in (\mathbb{R}_{>0})^{d-1} \ : \ \left(\forall i \in [\![1,\, d-1]\!],\, \beta_i > \sqrt{\delta}\right) \wedge \left(\prod_{i=1}^{d-1} \beta_i < \frac{1}{\sqrt{\delta}}\right)\right\}.$$

Both sets of inequalities (51) and (52) provide non–trivial lower and upper bounds for the probabilities $m_f(\delta)$ and $\widetilde{m}_{\widetilde{f}}(\delta)$, although the former bounds are doomed to be cruder than the latter (see the proof in §3.4 for details). In fact, we will mostly be interested in obtaining accurate upper bounds. In this respect, it is worth pointing out that those obtained above amount to finding short lattice vectors in a ball with respect

to the sup–norm in $\mathbb{R}^d$ centered at the origin rather than in the largest Euclidean ball contained in it (see the proof of Lemma 4 below for details). For "not too wild" density functions, the loss of accuracy in doing so should be seen as involving a multiplicative constant depending only on the dimension d.

3.3. **A Numerical Example.** A most standard distribution supported on the set of positive definite matrices is the so–called *Wishart distribution*. It is used in various fields such as the spectral theory of random matrices, multidimensional bayesian analysis and more generally in statistics, where its importance stems from the fact that it is a multidimensional generalisation of the chi–squared distribution which appears naturally in the likelihood–test ratio. The Wishart distribution is also commonly used to analyse the problem of wave fading in wireless communication, which is of particular interest to us in view of the results presented in §4 below. For further details on this probability distribution, see, e.g., [8]. We only mention here the few definitions and properties needed for our purpose.

Let X be a random $n \times d$ matrix. Assume that the rows $\boldsymbol{x_i}$ $(1 \leqslant i \leqslant n)$ of X are independent random vectors distributed according to a d–variate normal distribution $\mathcal{N}_d(\mathbf{0}, V)$ with zero mean and covariance matrix $V \in \mathcal{S}_d^{++}$. The Wishart distribution in dimension $d \geqslant 1$ with n *degrees of freedom* with respect to the *scale matrix* V is then the probability distribution of the matrix ${}^tXX$. It is usually denoted by $\mathcal{W}_d(V, n)$. Whenever $n \geqslant d$, the matrix ${}^tXX$ is invertible with probability one and the Wishart distribution admits a density function given by

$$f_{\mathcal{W}_d(V,n)}(Q) \;=\; \frac{1}{2^{nd/2} \cdot |V|^{n/2} \cdot \Gamma_d\left(\frac{n}{2}\right)} \cdot |Q|^{(n-d-1)/2} \cdot \exp\left(-\frac{1}{2} \cdot \mathrm{Tr}\left(V^{-1}Q\right)\right).$$

Here, $Q \in \mathcal{S}_d^{++}$, $|V|$ and $|Q|$ are shorthand notation for the determinant of V and Q respectively, $\mathrm{Tr}(\,.\,)$ is the usual trace operator over the space of matrices and

$$\Gamma_d\left(\frac{n}{2}\right) \;:=\; \pi^{d(d-1)/4} \prod_{j=1}^{d} \Gamma\left(\frac{n}{2} + \frac{1-j}{2}\right)$$

is the multivariate Gamma function.

Let $\delta > 0$. Denote by $m_{\mathcal{W}_d(V,n)}(\delta)$ the probability corresponding to the Wishart distribution defined as in (49). With the notation of Theorem 4, one has then the estimates

$$1 - \int_{\sqrt{\delta}}^{\infty} g_{\mathcal{W}_d(V,n)} \;\leqslant\; m_{\mathcal{W}_d(V,n)}(\delta) \;\leqslant\; 1 - \int_{I_d(\delta)} G_{\mathcal{W}_d(V,n)}, \tag{53}$$

where the function $G_{\mathcal{W}_d(V,n)}$ is explicitly given for any $\boldsymbol{\beta} \in (\mathbb{R}_{>0})^d$ by

$$G_{\mathcal{W}_d(V,n)}(\boldsymbol{\beta}) \;=\; \frac{\prod_{i=1}^{d} \beta_i^{n-i}}{2^{d(n/2-1)} \cdot |V|^{n/2} \cdot \Gamma_d\left(\frac{n}{2}\right)} \cdot \int_{\mathbb{R}^p} \exp\left(-\frac{1}{2} \cdot \mathrm{Tr}\left(V^{-1} \cdot \varphi_{chol}\left(\boldsymbol{\beta}, \boldsymbol{u}\right)\right)\right) \mathrm{d}\lambda_p(\boldsymbol{u})$$

and where the function $g_{\mathcal{W}_d(V,n)}$ is defined as in (50).

For the sake of concreteness, assume from now on that

$$n = d = 2 \quad \text{and} \quad V = I_2.$$

Then,

$$\varphi_{chol} \ : \ \begin{pmatrix} \beta_1 & u \\ 0 & \beta_2 \end{pmatrix} \mapsto \begin{pmatrix} \beta_1^2 & u\beta_1 \\ u\beta_1 & u^2+\beta_2^2 \end{pmatrix}$$

and, after calculations,

$$G_{\mathcal{W}_2(I_2,2)}(\beta_1,\beta_2) \ = \ \sqrt{\frac{2}{\pi}} \cdot \beta_1 \cdot \exp\left(-\frac{1}{2}\left(\beta_1^2+\beta_2^2\right)\right)$$

and

$$g_{\mathcal{W}_2(I_2,2)}\left(\beta_1\right) \ = \ \beta_1 \cdot \exp\left(-\frac{1}{2}\beta_1^2\right).$$

Inequalities (53) now read :

$$\begin{aligned} J_1(\delta) \ &:= \ 1 - \int_{\sqrt{\delta}}^{+\infty} \beta_1 \cdot \exp\left(-\frac{1}{2}\beta_1^2\right) \cdot \mathrm{d}\beta_1 \ \leqslant \ m_{\mathcal{W}_2(I_2,2)}(\delta) \\ &\leqslant \ 1 - \sqrt{\frac{2}{\pi}} \cdot \left(\int_{\sqrt{\delta}}^{+\infty} \beta_1 \cdot \exp\left(-\frac{1}{2}\beta_1^2\right) \cdot \mathrm{d}\beta_1\right) \cdot \left(\int_{\sqrt{\delta}}^{+\infty} \exp\left(-\frac{1}{2}\beta_2^2\right) \cdot \mathrm{d}\beta_2\right) \ := \ J_2(\delta). \end{aligned}$$

Some values taken by the functions J_1 and J_2 are represented in the following table :

δ	0.2	0.1	0.01	0.001
$J_1(\delta)$	0.095	0.049	$4.99 \cdot 10^{-3}$	$5.0 \cdot 10^{-4}$
$J_2(\delta)$	0.41	0.28	$8.42 \cdot 10^{-2}$	$2.6 \cdot 10^{-2}$

If the space of two dimensional positive definite matrices is equipped with the probability distribution $\mathcal{W}_2(I_2,2)$, the numerical values above imply for instance that at most 8.42% of these matrices admit a minimum over $\mathbb{Z}^2\backslash\{\mathbf{0}\}$ less than 0.01. Conversely, such a minimum is bigger than 0.2 for at least 9.5% of these matrices.

The remainder of this section is devoted to the proof of Theorem 4.

3.4. **Proof of Theorem 4.** We first prove two preliminary lemmata. The first one is presented in a context slightly more general than the one imposed by Theorem 4 : this more general statement will be needed in §4 below. It involves the set

$$\mathcal{M}_d^*\left(\gamma,c\right) \ := \ \left\{H \in \mathcal{T}_d^{++} \ : \ \det\left(\gamma I_d + {}^tH \cdot H\right) = c\right\}. \tag{54}$$

Here, I_d is the identity matrix in dimension d and γ and c are non–negative real numbers. It is easily seen (with the help of the spectral decomposition for instance) that the set $\mathcal{M}_d^*\left(\gamma,c\right)$ is non–empty if, and only if, $c > \gamma^d$.

Lemma 3. *The map φ_{chol} as defined in* (47) *is a $\mathcal{C}^1$–diffeomorphism with Jacobian determinant*

$$\mathrm{Jac}_L\left(\varphi_{chol}\right) \ = \ 2^d \cdot \prod_{i=1}^{d} l_{ii}^{d-i+1} \tag{55}$$

for any $L \in \mathcal{T}_d^{++}$ *with diagonal entries* $(l_{11}, \ldots, l_{dd})$.

Also, assuming $c > \gamma^d$, *the map*

$$\Psi_{(\gamma,c)}^{(d)} \ : H \in \mathcal{M}_d^*(\gamma, c) \mapsto c^{-1/d} \cdot (\gamma I_d + {}^tH \cdot H) \in \Sigma_d^{++}$$

is a $\mathcal{C}^1$*–diffeomorphism between* $\mathcal{M}_d^*(\gamma, c)$ *and its image with Jacobian determinant*

$$\mathrm{Jac}_H\left(\Psi_{(\gamma,c)}^{(d)}\right) = 2^{d-1} \cdot c^{-(d-1)(d+2)/(2d)} \cdot \prod_{i=1}^{d-1} h_{ii}^{d-i+1} \tag{56}$$

for any $H \in \mathcal{M}_d^*(\gamma, c)$ *with diagonal entries* $(h_{11}, \ldots, h_{dd})$.

Proof. Only equation (56) will be established hereafter as equation (55) can be deduced (in an easier way) from the argument presented below.

We first seek to determine a system of independent coordinates in $\mathcal{M}_d^*(\gamma, c)$ and in its image $\Psi_{(\gamma,c)}^{(d)}(\mathcal{M}_d^*(\gamma, c))$. To this end, given $c > \gamma^d$, define the auxiliary polynomial map

$$\widetilde{\Psi}_{\gamma}^{(d)} \ : H \in \mathcal{T}_d^{++} \mapsto \det(\gamma I_d + {}^tH \cdot H)$$

is such a way that $\mathcal{M}_d^*(\gamma, c) = \left(\widetilde{\Psi}_{\gamma}^{(d)}\right)^{-1}(\{c\})$. Since the differential of the determinant map at a square matrix A is the map $X \mapsto \mathrm{Tr}({}^t\mathrm{com}(A) \cdot X)$ (where $\mathrm{com}(A)$ is the comatrix of A), an elementary calculation shows that, at any $H \in \mathcal{M}_d^*(\gamma, c)$, the differential $\mathrm{d}_H\widetilde{\Psi}_{\gamma}^{(d)}$ of $\widetilde{\Psi}_{\gamma}^{(d)}$ is the linear map

$$\mathrm{d}_H\widetilde{\Psi}_{\gamma}^{(d)} \ : X \in \mathcal{T}_d^{++} \mapsto 2 \cdot \mathrm{Tr}\left[c \cdot (\gamma I_d + {}^tH \cdot H)^{-1} \cdot {}^tHX\right].$$

This map has clearly rank one. From the Regular Value Theorem (see [15, Lemma 1 p.11]), the fibre $\mathcal{M}_d^*(\gamma, c)$ is therefore a manifold of dimension $\dim \mathcal{T}_d^{++} - 1 = (d-1)(d+2)/2$.

If $H = (h_{ij})_{1 \leqslant i \leqslant j \leqslant d} \in \mathcal{M}_d^*(\gamma, c)$, choose for a system of coordinates in $\mathcal{M}_d^*(\gamma, c)$ the $(d-1)(d+2)/2$ variables $\widetilde{h} := (h_{ij})_{1 \leqslant i \leqslant j \leqslant d-1}$ (i.e. excluding h_{dd}). Let $\Sigma := (\sigma_{ij})_{1 \leqslant i,j \leqslant d}$ lie in the image of $\mathcal{M}_d^*(\gamma, c)$ by $\Psi_{(\gamma,c)}^{(d)}$. Let $\widetilde{\sigma} := (\sigma_{ij})_{1 \leqslant i \leqslant j \leqslant d-1}$ (this is the upper triangular part of Σ excluding the term σ_{dd}). In order to show that $\widetilde{\sigma}$ is a system of $(d-1)(d+2)/2$ independent coordinates parametrised by $\widetilde{h}$, express Σ as $\Sigma = c^{-1/d} \cdot (\gamma I_d + {}^tH \cdot H)$ for some $H \in \mathcal{M}_d^*(\gamma, c)$. Note then that when the elements of $\widetilde{\sigma}$ are listed row by row, each new entry

$$\sigma_{ij} = c^{-1/d} \cdot \left(\gamma \delta_{ij} + \sum_{k=1}^{i} h_{ki} h_{kj}\right) \tag{57}$$

$(1 \leqslant i \leqslant j \leqslant d-1)$ depends on an entry of H which has not appeared previously. However, $\sigma_{dd} = c^{-1/d} \cdot \left(\gamma + h_{dd}^2 + \sum_{k=1}^{d} h_{kd}^2\right)$ can be expressed as a function of $\widetilde{h}$ and

h_{dd}. For example, when $d = 3$,

$$\Sigma \;=\; c^{-1/d}\gamma I_d + c^{-1/d}\cdot \begin{pmatrix} h_{11}^2 & h_{11}h_{12} & h_{11}h_{13} \\ & h_{12}^2+h_{22}^2 & h_{12}h_{13}+h_{22}h_{23} \\ * & & h_{33}^2+h_{13}^2+h_{23}^2 \end{pmatrix}.$$

This legitimates $\widetilde{h}$ and $\widetilde{\sigma}$ as systems of coordinates respectively for $\mathcal{M}_d^*(\gamma, c)$ and for its image by $\Psi^{(d)}_{(\gamma,c)}$.

In order to compute the Jacobian determinant in (56), we now adapt the argument developed in [1, Chap. 7] to our purpose. Fix $H = (h_{ij})_{1\leqslant i\leqslant j\leqslant d} \in \mathcal{M}_d^*(\gamma, c)$ and denote by $\left(\mathrm{d}_{\Psi(H)}\,\sigma_{ij}\right)_{i,j}$ (resp. by $(\mathrm{d}_H\, h_{ij})_{i,j}$) the canonical basis of the tangent space to $\Psi^{(d)}_{(\gamma,c)}\left(\mathcal{M}_d^*(\gamma, c)\right)$ at $\Psi^{(d)}_{(\gamma,c)}(H)$ with respect to the system of coordinates $\widetilde{\sigma}$ (resp. of the tangent space to $\mathcal{M}_d^*(\gamma, c)$ at H with respect to the system of coordinates $\widetilde{h}$). For the sake of simplicity of notation, set further $\mathrm{d}\sigma_{ij} := \mathrm{d}_{\Psi(H)}\,\sigma_{ij}$ and $\mathrm{d}h_{ij} := \mathrm{d}_H\, h_{ij}$. The latter tangent vectors then satisfy the property that for any i, j,

$$\mathrm{d}h_{ij} \wedge \mathrm{d}h_{ij} = 0. \tag{58}$$

Moreover, the change of coordinates induced by $\Psi^{(d)}_{(\gamma,c)}$ implies that

$$\bigwedge_{1\leqslant i,j\leqslant d-1} \mathrm{d}\sigma_{ij} \;=\; \mathrm{Jac}_H\left(\Psi^{(d)}_{(\gamma,c)}\right)\cdot \bigwedge_{1\leqslant i,j\leqslant d-1} \mathrm{d}h_{ij}$$

(see [1, Chap. 7] for details). In view of (57), one has

$$\mathrm{d}\sigma_{ij} = c^{-1/d}\sum_{k=1}^{i}\left(h_{kj}\cdot \mathrm{d}h_{ki} + h_{ki}\cdot \mathrm{d}h_{kj}\right),$$

i.e.

$$\begin{aligned}
c^{1/d}\mathrm{d}\sigma_{11} &= 2h_{11}\cdot \mathrm{d}h_{11}, \\
c^{1/d}\mathrm{d}\sigma_{12} &= h_{11}\cdot \mathrm{d}h_{12} + \ldots, \qquad \ldots \qquad , \\
c^{1/d}\mathrm{d}\sigma_{1d} &= h_{11}\cdot \mathrm{d}h_{1d} + \ldots, \\
c^{1/d}\mathrm{d}\sigma_{22} &= 2h_{22}\cdot \mathrm{d}h_{22} + \ldots, \qquad \ldots \qquad , \\
c^{1/d}\mathrm{d}\sigma_{2d} &= h_{22}\cdot \mathrm{d}h_{2d} + \ldots, \qquad \ldots \qquad , \\
&\vdots \\
c^{1/d}\mathrm{d}\sigma_{d-1,d-1} &= 2h_{d-1,d-1}\cdot \mathrm{d}h_{d-1,d-1} \;+\; \ldots
\end{aligned}$$

The point to write these expressions this way is that, in view of (58), as soon as $\mathrm{d}h_{ij}$ appears in one of the terms in $\mathrm{d}\sigma_{ij}$, it may be ignored in all the others. All in all, this leads to

$$c^{(d-1)(d+2)/(2d)} \bigwedge_{1\leqslant i,j\leqslant d-1} \mathrm{d}\sigma_{ij} \;=\; \left(2^{d-1}\cdot \prod_{i=1}^{d-1} h_{ii}^{d-i+1}\right)\cdot \bigwedge_{1\leqslant i,j\leqslant d-1} \mathrm{d}h_{ij},$$

which completes the proof of the lemma. □

The second lemma needed to prove Theorem 4 is more elementary.

Lemma 4. *Let $L = (\boldsymbol{\beta}, \boldsymbol{u}) \in \mathcal{T}_d^{++}$ and $\eta > 0$. Write $\boldsymbol{\beta} = (\beta_1, \dots, \beta_d) \in (\mathbb{R}_{>0})^d$. The following holds :*

- *if $\beta_i > \eta$ for all $i = 1, \dots, d$, then*

$$L \cdot \mathbb{Z}^d \cap B_2(\mathbf{0}, \eta) = \{\mathbf{0}\}; \tag{59}$$

- *conversely, if $L \cdot \mathbb{Z}^d \cap B_2(\mathbf{0}, \eta) = \{\mathbf{0}\}$, then $\beta_1 > \eta$.*

Proof. The second claim is immediate upon noticing that $\beta_1 = \|L\boldsymbol{e_1}\|_2$. Assume therefore that $\beta_i > \eta$ for all $i = 1, \dots, d$ and note that conclusion (59) is trivial when $d = 1$. Let $d \geqslant 2$. Decompose the matrix $L := L_d$ in the following way :

$$L_d = \begin{pmatrix} L_{d-1} & \boldsymbol{u_{d-1}} \\ \mathbf{0} & \beta_d \end{pmatrix}.$$

Here, $L_{d-1} \in \mathcal{T}_{d-1}^{++}$ and $\boldsymbol{u_{d-1}} \in \mathbb{R}^{d-1}$. It is then readily seen that

$$L \cdot \mathbb{Z}^d = \bigcup_{n \in \mathbb{Z}} A_{L_d}(n), \qquad \text{where} \qquad A_{L_d}(n) = \begin{pmatrix} L_{d-1} \cdot \mathbb{Z}^{d-1} + n\boldsymbol{u_{d-1}} \\ n\beta_d \end{pmatrix}.$$

Proceeding by induction on $d \geqslant 2$, given $\boldsymbol{x} \in A_{L_d}(n)$, the inequality $\|\boldsymbol{x}\|_\infty > \eta$ follows by the induction hypothesis if $n = 0$ and is otherwise a direct consequence of the fact that $\|\boldsymbol{x}\|_\infty \geqslant \beta_d > \eta$. This completes the proof of the lemma. □

Proof of Theorem 4. Only the estimates (52) will be established hereafter as inequalities (51) follow from the argument presented below in a similar way.

Let $\Sigma \in \Sigma_d^{++}$ decomposed in its Cholesky form as $\Sigma = {}^tLL$, where $L = (\boldsymbol{\beta}', \boldsymbol{u}) \in \Theta_d^{++}$ with $\boldsymbol{\beta}' = (\beta_1', \dots, \beta_{d-1}') \in (\mathbb{R}_{>0})^{d-1}$ and $\boldsymbol{u} \in \mathbb{R}^p$. Set furthermore

$$\beta_d' = \left(\prod_{k=1}^{d-1} \beta_k' \right)^{-1}.$$

It should be clear that, given $\delta > 0$,

$$(M_d(\Sigma) > \delta) \quad \Longleftrightarrow \quad \left(L \cdot \mathbb{Z}^d \cap B_2\left(\mathbf{0}, \sqrt{\delta}\right) = \{\mathbf{0}\} \right).$$

From Lemma 4, if either statement in this equivalence holds, then $\beta_1' > \delta$. Conversely, it also follows from Lemma 4 that if $\min_{1 \leqslant i \leqslant d} \beta_i' > \sqrt{\delta}$, that is, if $\boldsymbol{\beta}' \in \Delta_{d-1}(\delta)$, then any of the statements in this equivalence holds.

Since

$$\begin{aligned} 1-\widetilde{m}_{\widetilde{f}}(\delta) &= \int_{\Sigma_d^{++}} \widetilde{f}(\Sigma)\cdot\chi_{[M_d(\Sigma)>\delta]}\cdot \mathrm{d}\Sigma \\ &= \int_{\Theta_d^{++}} \left(\widetilde{f}\circ\widetilde{\varphi}_{chol}\right)(L)\cdot\left|\mathrm{Jac}_L\left(\widetilde{\varphi}_{chol}\right)\right|\cdot\chi_{\left[L\cdot\mathbb{Z}^d\cap B_2\left(\mathbf{0},\sqrt{\delta}\right)=\{\mathbf{0}\}\right]}\cdot \mathrm{d}L, \end{aligned}$$

one thus obtains the estimates

$$\begin{aligned} &\int_{\Delta_{d-1}(\delta)} \mathrm{d}\lambda_{d-1}(\boldsymbol{\beta}')\int_{\mathbb{R}^p}\left(\widetilde{f}\circ\widetilde{\varphi}_{chol}\right)(\boldsymbol{\beta}',\boldsymbol{u})\cdot\left|\mathrm{Jac}_{(\boldsymbol{\beta}',\boldsymbol{u})}\left(\widetilde{\varphi}_{chol}\right)\right|\cdot\mathrm{d}\lambda_p(\boldsymbol{u}) \\ &\qquad\qquad \leqslant 1-\widetilde{m}_{\widetilde{f}}(\delta)\leqslant \\ &\int_{\sqrt{\delta}}^{+\infty}\mathrm{d}\lambda(\beta_1')\int_{(\mathbb{R}_{>0})^{d-2}}\mathrm{d}\lambda_{d-2}(\underset{\sim}{\boldsymbol{\beta}'})\int_{\mathbb{R}^p}\left(\widetilde{f}\circ\widetilde{\varphi}_{chol}\right)(\beta_1',\underset{\sim}{\boldsymbol{\beta}'},\boldsymbol{u})\cdot\left|\mathrm{Jac}_{(\beta_1',\underset{\sim}{\boldsymbol{\beta}'},\boldsymbol{u})}\left(\widetilde{\varphi}_{chol}\right)\right|\cdot\mathrm{d}\lambda_p(\boldsymbol{u}) \end{aligned}$$

(recall that $\boldsymbol{\beta}'=\left(\beta_1',\underset{\sim}{\boldsymbol{\beta}'}\right)$). The upper and lower bounds for $\widetilde{m}_{\widetilde{f}}(\delta)$ in (52) now follow directly from Lemma 3 (with $\gamma=0$ and $c=1$). Furthermore, to prove that these bounds always lie in the interval $[0,\,1]$, it is enough to notice that, from the definitions of the functions $\widetilde{G}_{\widetilde{f}}$ and $\widetilde{g}_{\widetilde{f}}$,

$$\int_{(\mathbb{R}_{>0})^{d-1}}\widetilde{G}_{\widetilde{f}}=\int_0^{+\infty}\widetilde{g}_{\widetilde{f}}=\int_{\Sigma_d^{++}}\widetilde{f}(\Sigma)\cdot\mathrm{d}\Sigma=1.$$

□

4. Application to Signal Processing

The initial motivation of this work was to address a fundamental problem that emerged very recently in Information Theory. The latter is related to a new model of communication channel (the so called *Integer–Forcing Architecture*) which has been receiving considerable attention in the literature due to its expected high performance. The precise estimation of this performance involves the probability that a quadratic form admits a minimum over non–zero lattice points less than a given constant.

In what follows, we first present the very basic tools from Information Theory that will enable one to understand the importance and the position of the problem under consideration — for a deeper introduction to the topic, see [18], especially Chapter 5. The theory developed in the previous sections will then allow one to bound accurately the probability to estimate.

4.1. **Position of the Problem.** Assume that two *users* (or *transmitters*) S_1 and S_2 want to *transmit* messages (or *signals*) x_1 (for S_1) and x_2 (for S_2) along a communication channel (e.g., a cable or a radio channel) simultaneously to two *receivers* R_1 and R_2 ([1]). Independently of the familiar concept of noise, the signal is distorted during

[1]This configuration, widely studied in Information Theory, is known as an "X–Channel" with a reference to the shape of Figure 4.1 below.

transmission up to a certain degree of *fading*. This may be due for instance to the distance between the users and the receivers or else to reflections on obstacles such as buildings in the path of the signals. This phenomenon is modelled by the so–called *channel coefficients*. For the message sent by S_i to R_j $(i, j \in \{1, 2\})$ the corresponding channel coefficient is denoted by h_{ij}. Thus, in the simplest case of an additive channel, the message y_i received by R_i $(i \in \{1, 2\})$ is represented by the system of equations

$$\begin{cases} y_1 = h_{11}x_1 + h_{12}x_2 + z_1 \\ y_2 = h_{21}x_1 + h_{22}x_2 + z_2, \end{cases} \tag{60}$$

where z_1 and z_2 are the noise — see also the figure below.

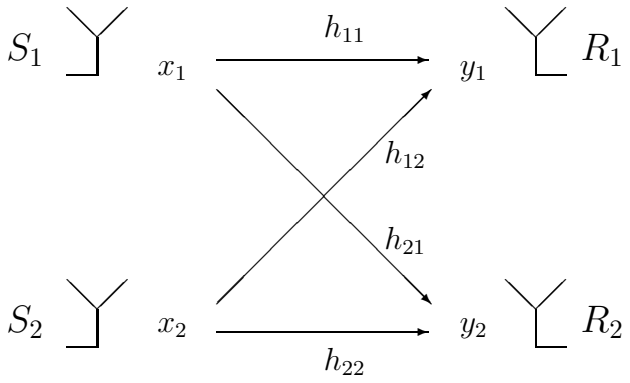

FIGURE 1. Channel of communication corresponding to the configuration in (60)

. Matricially, the system of equations (60) reads

$$\boldsymbol{y} \,=\, H\boldsymbol{x} + \boldsymbol{z} \tag{61}$$

with

$$\boldsymbol{y} \,=\, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \quad H \,=\, \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix}, \quad \boldsymbol{x} \,=\, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \text{and} \quad \boldsymbol{z} \,=\, \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}.$$

Of course, it is obvious to generalise this model to the case when there are $m \geqslant 1$ users and $n \geqslant 1$ receivers. Then, the matrix H in (61) is rectangular with dimensions $n \times m$, the vectors $\boldsymbol{y}$ and $\boldsymbol{z}$ are n–dimensional and the vector $\boldsymbol{x}$ is m–dimensional. From the receiver's point of view, it is natural to consider $\boldsymbol{x}$ and $\boldsymbol{z}$ as random vectors, in which case the entries of the noise vector $\boldsymbol{z}$ are often taken as independent with Gaussian distribution with zero mean and unit variance. As for the input $\boldsymbol{x}$, it satisfies a power constraint of the form

$$\mathbb{E}\,({}^t\boldsymbol{x} \cdot \boldsymbol{x}) \leqslant m \cdot \mathrm{SNR}, \tag{62}$$

where $\mathbb{E}(\,.\,)$ denotes the expectation and where SNR stands for the *Signal–to–Noise Ratio*, a fundamental strictly positive quantity which will be discussed later. In the

standard case when each entry of $\boldsymbol{x}$ is a sum of binary digits (*bits*), the power constraint (62) reflects the fact that the number of bits that can be sent through the channel is limited by some of its physical properties.

It is important to point out here that the seemingly simple model with two users and two receivers exposed above underpins some of the most fundamental features of the more general model with m users and n receivers. Thus, some channel architectures with $m = 2$ users and $n = 2$ receivers have been at the heart of deep theoretical problems in Information Theory — see, e.g., [18, §5.4.3].

The most basic problem when considering a channel of communication is to determine whether the received information is reliable; that is, to what extent the noise affects the quality of the signal. In order to make the probability error small, an obvious guess is that one has to reduce the rate of new data sent by the users (for instance, by repeating each string of message several times). In 1948, Shannon proved that this intuition is surprisingly incorrect : it is actually possible to exchange information at a *strictly positive* data rate keeping at the same time the error probability as small as desired. There is nevertheless a maximal rate, the *capacity* of the channel, above which this cannot be done any more. The latter quantity is usually expressed in bits.

As the proof of the result established by Shannon is non–effective (i.e. it does not provide a way to code the information in order to approach the capacity), from an engineering standpoint, the problem to determine the capacity of a channel and then to provide a way to get as close as possible to this capacity remains open.

There is no single expression for the capacity of a channel; rather, it depends on its intrinsic architecture. It nevertheless always involves the Signal–to–Noise Ratio (SNR). This quantity, often expressed in decibels, compares the level of a desired signal to the level of background noise : the bigger this ratio, the better the quality of the signal. For the model represented by the equations in (61) and (62) (with any $m, n \geqslant 1$), it is shown in [16] that the capacity C can be expressed as

$$C \ = \ \log\det\left(I_m + \mathrm{SNR}\cdot {}^t\!H\cdot H\right). \tag{63}$$

Note also the following important point : the performances of a channel depend heavily on whether or not the transmitter knows the channel coefficients matrix H. Indeed, if such information is available, they can for instance allocate more power to the stronger antennas to minimise the effect of fading. In most cases however (for instance in wireless systems), this information is not known to the transmitter, in which case a reasonable strategy is to allocate equal power to each of the antennas. In the latter configuration, the capacity of the channel is rather referred to as the *mutual information*.

4.2. **Channels with Integer–Forcing Receiver Architecture.** Recently, an important breakthrough has been achieved in Information Theory. Indeed, Zhan & *alii* introduced in [21] a new architecture of channel, the so–called *Integer–Forcing Receiver Architecture*, which has been receiving considerable attention in the literature

(see [16] and the references therein for further details). It is not our goal to describe the channel precisely : if interested, the reader is referred to [21]. Here is however the main ingredient from which follow all the properties of this new model : in a standard communication channel, the receivers obtain the message $\boldsymbol{x}$ sent to them by first eliminating interferences from the vector $\boldsymbol{y}$ (especially the noise $\boldsymbol{z}$) and then by decoding each individual data stream (i.e. each component of the vector $\boldsymbol{y}$). The idea introduced by Zhan & *alii* is, first to decode integer linear combinations of data stream and, then, to eliminate the interference.

The near optimality of this strategy has been verified by extensive *ad hoc* calculations (see [16, §I.A.] for details). As for a theoretical proof of this fact, this task has been started in [16] in the following general set–up which, as explained in the paper, appears in several important communication scenarios.

Assume that each transmitter wishes to send the same message to all the receivers (this is for instance the case for TV broadcast). They all are aware of the characteristics of the channel, namely its SNR coefficient and also the mutual information C_0. However, they ignore the actual channel matrix H modelling the transmission as in (61). Without any more information and in view of (63), this matrix H is considered as being randomly and "uniformly" chosen[2] from the set

$$\mathcal{H}_{m,n}\left(C_0, \mathrm{SNR}\right) \,:=\, \left\{H \in \mathbb{R}^{n\times m} \;:\; \log\det\left(I_m + \mathrm{SNR}\cdot {}^tH\cdot H\right) \,=\, C_0\right\}. \tag{64}$$

It is proved in [16] that the performance of the channel under consideration after applying the integer–forcing technique is actually determined by the so–called *Effective Signal–to–Noise Ratio* $\mathrm{SNR}_{\mathrm{eff}}$. We shall not be concerned with the actual definition of this quantity, which is rather technical — for details, see [16, §II.B.]. The crucial point formulated with our notation is the following estimates satisfied by the $\mathrm{SNR}_{\mathrm{eff}}$ coefficient (see [16, Theorem 2] for a proof) :

$$\frac{1}{4m^2}\cdot M_m\left(I_m + \mathrm{SNR}\cdot {}^tH\cdot H\right) \,<\, \mathrm{SNR}_{\mathrm{eff}} \,\leqslant\, M_m\left(I_m + \mathrm{SNR}\cdot {}^tH\cdot H\right). \tag{65}$$

For the quality of communication to be best possible, one wishes to obtain a $\mathrm{SNR}_{\mathrm{eff}}$ coefficient as large as possible. Inequalities (65) show that the order of magnitude of this coefficient is dictated by the minimum of the positive definite quadratic form $I_m + \mathrm{SNR}\cdot {}^tH\cdot H$ over non–zero elements of $\mathbb{Z}^m$. In view of the probabilistic model developed so far, the main problem which emerges from this theory can be formulated as follows :

Problem 3 (Main Problem of Application)**.** *Assume that the channel matrix H is chosen randomly and "uniformly" from the set* (64)*. Let $\kappa \in (0,1)$.*

Find the best possible value of $s \geqslant 0$ such that the event $SNR_{\mathit{eff}} \geqslant s$ is realised with probability greater than κ; equivalently, determine the cumulative distribution function of the quantity SNR_{eff} seen as a random variable.

[2]As will be shown later, this concept of uniformity, understood here intuitively, needs to be clarified.

It is worth noting that the techniques developed here in order to tackle this problem can also be used to solve other questions appearing in the literature dealing with the Integer–Forcing Architecture. An example of such questions is the estimate of the probability that the so–called effective noise variance as defined in [21, §IV.E.] should be less than a given constant. Another more general example is the estimate of the so–called probability of outage of some channels — see [18, 21]. In all cases, the main ingredient is Theorem 4 (more precisely, the upper bounds appearing therein). Also, it must be pointed out that the manifold (64) is ubiquitous in the literature related to Signal Processing. Some of its topological properties playing a crucial role in the study of the performance of various channels are established in §4.3 below.

4.3. **Formalisation of the Concept of a "Uniformly" Distributed Measure on the Set $\mathcal{H}_{m,n}(C_0, \mathbf{SNR})$.** For convenience, set from now on

$$\gamma := (\mathrm{SNR})^{-1} \quad \text{and} \quad c_0 := \gamma^m e^{C_0} \tag{66}$$

in such a way that

$$\mathcal{H}_{m,n}(C_0, \mathrm{SNR}) \;=\; \left\{H \in \mathbb{R}^{n\times m} \;:\; \det\left(\gamma I_m + {}^tH \cdot H\right) \;=\; c_0\right\}.$$

For the sake of simplicity of notation, the dependency of the various quantities on γ and c_0 will not be marked hereafter. The reader should however keep in mind that almost all the constants, sets and functions introduced hereafter depend on these two parameters.

A crucial remark is that Sylvester's determinant identity immediately implies that

$$\det\left(\gamma I_m + {}^tH \cdot H\right) \;=\; \det\left(\gamma I_n + H \cdot {}^tH\right).$$

Therefore, even if it means working throughout with ${}^tH$ instead of H to obtain the analogues in the case $n \geqslant m$ of the results stated below, it may be assumed *without loss of generality* that

$$d := \min\{m, n\} = m. \tag{67}$$

In order to address Problem 3 as stated above, one needs first to formalise the idea of a "uniform" measure on the set $\mathcal{H}_{m,n}(C_0, \mathrm{SNR})$. If one understands this concept in the usual mathematical meaning of a Borelian measure in a complete metric space such that the measure of a ball depends only on its radius but not on the position of its center, this is problematic. Indeed, as shown in Lemma 5 below, the set $\mathcal{H}_{m,n}(C_0, \mathrm{SNR})$ is compact. Now, it is proved in [10, Proposition 1.7] that a bounded subset of an Euclidean space carries a uniform measure only if it is contained in a sphere. It is not hard to see that this never happens for the set $\mathcal{H}_{m,n}(C_0, \mathrm{SNR})$ as soon as $d \geqslant 2$. In view of this and in order to render this idea of uniform distribution in a different way, we first establish some properties of the set $\mathcal{H}_{m,n}(C_0, \mathrm{SNR})$.

Given an integer $k \in [\![0, d]\!]$, let $\mathcal{R}_{m,n}^{(k)}$ be the subset of $\mathcal{H}_{m,n}(C_0, \mathrm{SNR})$ consisting of all those matrices with rank k :

$$\mathcal{R}_{m,n}^{(k)} \;:=\; \{H \in \mathcal{H}_{m,n}(C_0, \mathrm{SNR}) \;:\; \mathrm{rank}(H) = k\}.$$

Note that any of the sets $\mathcal{R}_{m,n}^{(k)}$ is invariant under a map of the form $H \mapsto U \cdot H$, where $U \in \mathcal{O}_n$ is an orthogonal transformation. This legitimate the focus on a fundamental domain for the left action of $\mathcal{O}_n$ on $\mathcal{R}_{m,n}^{(k)}$. As shown in Lemma 5 below, such a fundamental domain is naturally be related to the set

$$\mathcal{M}_d^{(k)} := \left\{ T \in \mathcal{T}_d^+ \ : \ \mathrm{rank}(T) = k \quad \text{and} \quad \det\left(\gamma I_d + {}^t T \cdot T\right) = c_0 \right\},$$

where $\mathcal{T}_d^+$ is the set of all those upper triangular d–dimensional square matrices with non–negative diagonal entries. Note that when $k = d$, the set $\mathcal{M}_d^{(k)}$ coincides with the set $\mathcal{M}_d^*(\gamma, c_0)$ defined in (54). In what follows, we will adopt the simpler notation

$$\mathcal{M}_d^* := \mathcal{M}_d^*(\gamma, c_0).$$

It is not hard to see that a necessary and sufficient condition for the subset $\mathcal{M}_d^*$ to be non–empty is that

$$c_0 > \gamma^d. \tag{68}$$

In this case, the zero matrix cannot belong to the set

$$\widetilde{\mathcal{M}}_d := \bigcup_{k=0}^{d} \mathcal{M}_d^{(k)} = \left\{ T \in \mathcal{T}_d^+ \ : \ \det\left(\gamma I_d + {}^t T \cdot T\right) = c_0 \right\} \tag{69}$$

(if $c_0 = \gamma^d$, the latter set only contains the zero matrix and if $c_0 < \gamma^d$, it is empty — see §3.4 or the proof of Lemma 5 for details). The relation (68) will be assumed to hold throughout.

Lemma 5. *The following two points hold :*

- *The set $\mathcal{H}_{m,n}(C_0, SNR)$ is compact.*
- *Given an integer $k \in [\![0, d]\!]$, a fundamental domain for the left action of the orthogonal group $\mathcal{O}_n$ on $\mathcal{R}_{m,n}^{(k)}$ can naturally be identified with a subset of $\mathcal{M}_d^{(k)}$. Furthermore, when $k = d$, a fundamental domain for the left action of the orthogonal group $\mathcal{O}_n$ on $\mathcal{R}_{m,n}^{(d)}$ can naturally be identified with the set $\mathcal{M}_d^*$ itself.*

Proof. The second point is a direct consequence of the QR decomposition : any matrix $H \in \mathcal{R}_{m,n}^{(k)}$ can be decomposed as $H = QR$, where $Q \in \mathcal{O}_n$ and where the matrix R has rank k and is of the form

$$R = \begin{pmatrix} T \\ \mathbf{0} \end{pmatrix}$$

with $T \in \mathcal{T}_d^+$. Furthermore, this decomposition is unique when R has full rank.

As for the first point, note that the set $\mathcal{H}_{m,n}(C_0, \mathrm{SNR})$ is clearly closed. To show that it is also bounded, we will adopt the following notation : given a $n \times m$ rectangular matrix M, $\|M\|_\infty$ will denote the sup–norm of the vector in $\mathbb{R}^{nm}$ determined by its entries. Also, $|||M|||_2$ (resp. $|||M|||_\infty$) will stand for the operator norm of M induced by the Euclidean norms (resp. the sup–norms). Given two positive real numbers a and b, the Vinogradov symbol $a \ll b$ will as usual indicate the existence of a positive constant $c > 0$ such that $a \leqslant cb$.

Let then $H \in \mathcal{H}_{m,n}(C_0, \mathrm{SNR})$. By looking at the diagonal elements in ${}^tH \cdot H$, it is plain that

$$\|H\|_\infty \leqslant \sqrt{\|{}^tH \cdot H\|_\infty}.$$

Let ${}^tH \cdot H = {}^tP \cdot D \cdot P$ be the spectral decomposition of the positive matrix ${}^tH \cdot H$, where $P \in \mathcal{O}_m$ and where D is a diagonal matrix with entries $\lambda_1, \ldots, \lambda_m \geqslant 0$. From the equivalence of norms in finite dimension and from the fact that $|||P|||_2 = 1$, one thus obtains :

$$\begin{aligned}
\|{}^tH \cdot H\|_\infty &\ll |||\,{}^tH \cdot H|||_2 = |||\,{}^tPDP|||_2 \\
&\leqslant |||\,{}^tP|||_2 |||D|||_2 |||P|||_2 \\
&= |||D|||_2 \\
&\ll \|D\|_\infty := \max \mathrm{Spect}\,({}^tH \cdot H),
\end{aligned}$$

where $\mathrm{Spect}\,({}^tH \cdot H)$ denotes the spectrum of the matrix ${}^tH \cdot H$. From the definition of the set $\mathcal{H}_{m,n}(C_0, \mathrm{SNR})$, one has furthermore that

$$c_0 = \det(\gamma I_m + {}^tH \cdot H) = \det(\gamma I_m + D) = \prod_{i=1}^{m} (\gamma + \lambda_i).$$

Since $\lambda_i \geqslant 0$ for all $i = 1, \ldots, m$, this implies that $\mathrm{Spect}\,({}^tH \cdot H) \subset [0,\, \gamma\,(c_0\gamma^{-m} - 1)]$ (which set is empty if $c_0 < \gamma^m$). This completes the proof. □

Remark 1. *We would like to point out here that the first point in Lemma 5 rules out an assumption often made in the literature related to Information Theory (see, among many other examples,* [22, Problem 13.12]*); namely, the coefficients of a matrix H lying in $\mathcal{H}_{m,n}(C_0, SNR)$ cannot have a Gaussian distribution.*

Remark 2. *A much more involved argument presented in the proof of Lemma 6 below implies that the Euclidean norm of a matrix lying in the set $\widetilde{\mathcal{M}}_d$ and viewed as a vector in $\mathbb{R}^{d(d+1)/2}$ is at most $\sqrt{(c_0 - \gamma^d)/\gamma^{d-1}}$ and at least $\sqrt{c_0^{1/d} - \gamma}$ — see the end of §4.5 for details. From the QR decomposition, this also holds for a matrix lying in $\mathcal{H}_{m,n}(C_0, SNR)$.*

If one understands the concept of a "uniform" measure as a measure "evenly" distributed (in some intuitive sense), in view of the invariance of the set $\mathcal{H}_{m,n}(C_0, \mathrm{SNR})$ under the left action of the orthogonal group, it is natural to define such a measure from a fundamental domain of $\mathcal{H}_{m,n}(C_0, \mathrm{SNR})$ for this action. Thus, if one is able to equip the set $\widetilde{\mathcal{M}}_d$ as defined in (69) with a "uniform" probability measure $\widetilde{\nu}_d$ which satisfies furthermore the property that

$$\widetilde{\nu}_d(\mathcal{M}_d^*) = 1 \tag{70}$$

(that is, the measure $\widetilde{\nu}_d$ is only supported on those matrices of full rank), then, in view of Lemma 5, $\widetilde{\nu}_d$ would be a relevant candidate for our purpose[3].

A natural choice for $\widetilde{\nu}_d$ is a measure which takes into account the geometry of the manifold $\widetilde{\mathcal{M}}_d$. Setting

$$p' = \frac{d(d+1)}{2} - 1,$$

this leads one to define $\widetilde{\nu}_d$ from the infinitesimal volume element $\mathrm{d}\,\mathrm{vol}_{p'}(T)$ on the hypersurface $\widetilde{\mathcal{M}}_d \subset \mathbb{R}^{p'+1}$. More precisely, for any measurable subset $\mathcal{B} \subset \widetilde{\mathcal{M}}_d$,

$$\widetilde{\nu}_d(\mathcal{B}) := \frac{\int_{\mathcal{B}} \mathrm{d}\,\mathrm{vol}_{p'}(T)}{\int_{\widetilde{\mathcal{M}}_d} \mathrm{d}\,\mathrm{vol}_{p'}(T)}. \tag{71}$$

Note that this is a well–defined probability measure as $\widetilde{\mathcal{M}}_d$ is compact.

Let

$$f : T \in \mathcal{T}_d^+ \mapsto c_0^{-1/d} \cdot (\gamma I_d + {}^t T \cdot T)$$

and

$$g := \det \circ f \tag{72}$$

in such a way that

$$\widetilde{\mathcal{M}}_d = g^{-1}(\{1\}).$$

Given $T = (t_{ij})_{1 \leqslant i,j, \leqslant d} \in \mathcal{T}_d^+$ and given indices i and j such that $1 \leqslant i \leqslant j \leqslant d$, set

$$\partial_{ij} := \frac{\partial}{\partial t_{ij}}$$

and define furthermore the charts

$$\mathcal{B}_{ij} := \left\{ T \in \mathcal{T}_d^+ \ : \ (\partial_{ij} g)(T) \neq 0 \right\}. \tag{73}$$

The relevance of this definition follows from this lemma :

Lemma 6. *Assume* (68). *Then :*

- *The gradient ∇g of g never vanishes on $\widetilde{\mathcal{M}}_d$. In other words,*

$$\widetilde{\mathcal{M}}_d = \bigcup_{1 \leqslant i \leqslant j \leqslant d} \left(\mathcal{B}_{ij} \cap \widetilde{\mathcal{M}}_d \right).$$

- *On each of the charts $\mathcal{B}_{ij}$, the volume element $\mathrm{d}\,vol_{p'}(T)$ can be expressed as follows :*

$$\mathrm{d}\,vol_{p'}(T) = \left(\frac{\|\nabla g\|_2}{|\partial_{ij} g|} \right)(T) \cdot \mathrm{d}t_{11} \ldots \widehat{\mathrm{d}t_{ij}} \ldots \mathrm{d}t_{dd} \tag{74}$$

(as usual, the hat means that the corresponding index is removed from the list).

[3]It must be pointed out here that, from an engineering standpoint, it is often assumed that the channel matrix has full rank not to have to deal with redundant information. Lemma 7 below shows that we will not have to make such an assumption here.

- *The subset of matrices of full rank in $\widetilde{\mathcal{M}}_d$ is contained in $\mathcal{B}_{dd}$:*
$$\mathcal{M}_d^* \subset \mathcal{B}_{dd}.$$

With the help of this lemma, one can now prove that the measure $\widetilde{\nu}_d$ defined in (71) satisfies (70) :

Lemma 7. *Let $k \in [\![0, d-1]\!]$. Then, under* (68)*,*
$$\widetilde{\nu}_d\left(\mathcal{M}_d^{(k)}\right) = 0$$

Proof. It follows from Lemma 6 that $\widetilde{\mathcal{M}}_d$ can be covered by a finite number of subsets $\left(\mathcal{B}'_{ij}\right)_{1\leqslant i\leqslant j\leqslant d}$ such that, within each $\mathcal{B}'_{ij}$, the function $\partial_{ij}g$ never vanishes. Also, within each $\mathcal{B}'_{ij}$, the measure determined by the volume element $\mathrm{d}\,\mathrm{vol}_{p'}(T)$ is absolutely continuous with respect to the p'–dimensional Lebesgue measure $\lambda_{p'}$. In order to prove the lemma, it is therefore enough to establish that for all $0 \leqslant k \leqslant d-1$ and all $1 \leqslant i \leqslant j \leqslant d$,
$$\lambda_{p'}\left(\mathcal{M}_d^{(k)} \cap \mathcal{B}'_{ij}\right) = 0. \tag{75}$$
To this end, note that $\bigcup_{k=0}^{d-1}\mathcal{M}_d^{(k)}$ sits as an algebraic subvariety in $\widetilde{\mathcal{M}}_d \subset \mathcal{T}_d^+$; it is defined as the intersection of $\widetilde{\mathcal{M}}_d$ with the hypersurface
$$\mathcal{L} := \left\{T \in \mathcal{T}_d^+ \ : \ \det(T) = 0\right\}.$$
Since the hypersurface $\mathcal{L}$ defines an irreducible variety, any variety intersects it properly (with the possibility of an empty intersection) or is contained in it. It is easily seen (with the help of the spectral decomposition for instance) that the set $\mathcal{M}_d^*$ is non–empty under (68); in other words, that there are points in $\widetilde{\mathcal{M}}_d$ not contained in $\mathcal{L}$. Thus, the intersection $\widetilde{\mathcal{M}}_d \cap \mathcal{L}$ has codimension at least one in $\widetilde{\mathcal{M}}_d$, which readily implies (75) and completes the proof. □

4.4. **Estimation of the Cumulative Distribution Function of the Effective Signal–to–Noise Ratio.** In view of (65), Problem 3 boils down to finding, for a given $s \geqslant 0$, a lower bound for the event $M_d\left(I_d + \mathrm{SNR}\cdot {}^tH\cdot H\right) \geqslant 4sd^2$ when H is chosen randomly from the set $\mathcal{H}_{m,n}\left(C_0, \mathrm{SNR}\right)$ according to the distribution of the probability measure $\widetilde{\nu}_d$. From the change of variables operated in (66) and from Lemma 7, this amounts to bounding from below the quantity
$$\mathfrak{m}_d(\delta) := \widetilde{\nu}_d\left(\left\{H \in \mathcal{M}_d^* \ : \ M_d\left(c_0^{-1/d}\left(\gamma I_d + {}^tH\cdot H\right)\right) > \delta\right\}\right),$$
where we have set
$$\delta := 4d^2 s\gamma^d c_0^{-1/d} \tag{76}$$
(note that the definitions of $\mathfrak{m}_d(\delta)$ above and of $m_f(\delta)$ in (49) differ inasmuch as the inequalities defining each of these quantities are reversed. The definition of $\mathfrak{m}_d(\delta)$ is here motivated by the statement of Problem 3). Note that when $H \in \mathcal{M}_d^*$,
$$c_0^{-1/d}\left(\gamma I_d + {}^tH\cdot H\right) \in \Sigma_d^{++}.$$

It follows immediately from the definition of the the function M_d in (1) that $M_d\left(c_0^{-1/d}\left(\gamma I_d + {}^tH\cdot H\right)\right) \geqslant \gamma c_0^{-1/d}$ in such a way that

$$\mathfrak{m}_d(\delta) \;=\; 1 \qquad \text{whenever} \qquad \delta \leqslant \frac{\gamma}{c_0^{1/d}}.$$

In what follows, it will therefore be assumed without loss of generality that

$$\delta \;>\; \frac{\gamma}{c_0^{1/d}} \;:=\; \delta_d^*. \tag{77}$$

In order to call on Theorem 4 under this assumption, one needs to push forward the measure $\widetilde{\nu}_d$ from $\mathcal{M}_d^*$ to the space Θ_d^{++} as defined in (45) via the maps

$$\mathcal{M}_d^* \xrightarrow{f} \Sigma_d^{++} \xrightarrow{\widetilde{\varphi}_{chol}^{-1}} \Theta_d^{++} \tag{78}$$

(cf. (48) for the definition of $\widetilde{\varphi}_{chol}$). The main apparent difficulty in doing so is that the Cholesky decomposition of the matrix $\gamma I_d + {}^tH\cdot H$ cannot be straightforwardly deduced from to the Cholesky form ${}^tH\cdot H$ when $H\in\mathcal{M}_d^*$. However, explicit expressions can be given from the general Cholesky algorithm which, as mentioned in §3, can be implemented in an very efficient way. Thus, given $H = (h_{ij})_{1\leqslant i\leqslant j\leqslant d} \in \mathcal{M}_d^*$, if $L = (l_{ij})_{1\leqslant i\leqslant j\leqslant d}\in\Theta_d^{++}$ is the Cholesky form of the matrix $c_0^{-1/d}\left(\gamma I_d + {}^tH\cdot H\right)\in\Sigma_d^{++}$ (that is, if ${}^tL\cdot L = c_0^{-1/d}\left(\gamma I_d + {}^tH\cdot H\right)$), one can express recursively the coefficients h_{ij} as functions of l_{ij} (which is what is needed to apply Theorem 4) as follows : for all $1\leqslant i\leqslant d$,

$$h_{ii} \;=\; \sqrt{\sum_{k=1}^{i} c_0^{1/d} l_{ki}^2 - \gamma - \sum_{k=1}^{i-1} h_{ki}^2} \tag{79}$$

and, for all $1\leqslant i<j\leqslant d$,

$$h_{ij} \;=\; \frac{1}{h_{ii}}\left(\sum_{k=1}^{i} c_0^{1/d} l_{ki}l_{kj} - \sum_{k=1}^{i-1} h_{ki}h_{kj}\right) \tag{80}$$

(this is just the classical algorithm giving the Cholesky decomposition applied to the positive definite matrix $c_0^{1/d}\cdot{}^tL\cdot L - \gamma I_d$ — see [19] for details).

In order to transport the measure $\widetilde{\nu}_d$ to the space Θ_d^{++}, one will also need to compute the Jacobian J_d of the map $f^{-1}\circ\widetilde{\varphi}_{chol}\;:\mathcal{N}_d^*\to\mathcal{M}_d^*$ obtained from (78), where

$$\mathcal{N}_d^* \;:=\; \left(\widetilde{\varphi}_{chol}^{-1}\circ f\right)\left(\mathcal{M}_d^*\right). \tag{81}$$

To this end, note that, with the notation of Lemma 3, one has $\widetilde{\varphi}_{chol} = \Psi_{(0,1)}^{(d)}$ and $f = \Psi_{(\gamma,c_0)}^{(d)}$ in such a way that (56) implies that

$$J_d \;=\; c_0^{(d-1)(d+2)/(2d)}\prod_{i=1}^{d-1}\left(\frac{l_{ii}}{h_{ii}}\right)^{d-i+1}.$$

Also, it follows from Lemmata 6 and 7 that it is enough to consider the restriction of the measure $\widetilde{\nu}_d$ to the chart $\mathcal{B}_{dd}$ defined from (73). It is given therein by the volume element (74) with $i = j = d$.

In view of formulae (79) and (80), any expression involving the coefficients h_{ij} of a matrix $H \in \mathcal{M}_d^*$ can be viewed as a function of the coefficients l_{ij} of the matrix L as defined above. With this in mind, define two auxiliary functions $\widetilde{J}_d$ and $\widetilde{\Gamma}_d$ over the space $\mathcal{N}_d^*$ by setting

$$\widetilde{J}_d(L) \,:=\, J_d \qquad \text{and} \qquad \widetilde{\Gamma}_d(L) \,:=\, \left(\frac{\|\nabla g\|_2}{|\partial_{dd} g|}\right)(H). \tag{82}$$

Furthermore, if $L \in \Theta_d^{++}$ is decomposed as $L = (\boldsymbol{\beta}', \boldsymbol{u})$ with $\boldsymbol{\beta}' \in (\mathbb{R}_{>0})^{d-1}$ and $\boldsymbol{u} \in \mathbb{R}^p$ as in §3 (see Equation (45) sqq. for the notation), it will be convenient to set

$$\widetilde{J}_d(\boldsymbol{\beta}', \boldsymbol{u}) \,:=\, \widetilde{J}_d(L) \qquad \text{and} \qquad \widetilde{\Gamma}_d(\boldsymbol{\beta}', \boldsymbol{u}) \,:=\, \widetilde{\Gamma}_d(L).$$

The main result of this section, which is a direct consequence of the upper bound in (52), can now be stated as follows :

Theorem 5. *Assume* (68), (77) *and also that* $\delta < 1$. *Then,*

$$\mathfrak{m}_d(\delta) \,\geqslant\, \kappa_d^{-1} \cdot \int_{\mathcal{N}_d^*[\delta]} \widetilde{J}_d(\boldsymbol{\beta}', \boldsymbol{u}) \cdot \widetilde{\Gamma}_d(\boldsymbol{\beta}', \boldsymbol{u}) \cdot \mathrm{d}\lambda_{p+d-1}\,(\boldsymbol{\beta}', \boldsymbol{u})\,. \tag{83}$$

Here,

$$\mathcal{N}_d^*\,[\delta] \,:=\, \{(\boldsymbol{\beta}', \boldsymbol{u}) \in \mathcal{N}_d^* \;:\; \boldsymbol{\beta}' \in \Delta_{d-1}(\delta)\}$$

is a subset of $\mathcal{N}_d^*$, $\Delta_{d-1}(\delta)$ *is defined as in* (52) *and*

$$\kappa_d \,:=\, \int_{\widetilde{\mathcal{M}}_d} \mathrm{d}\,\mathrm{vol}_{p'}(H)$$

is the area of the hypersurface $\widetilde{\mathcal{M}}_d$.

In view of Lemmata 6 and 7, the constant κ_d can also be computed with the help of any of the following formulae :

$$\kappa_d \;=\; \int_{\mathcal{M}_d^*} \left(\frac{\|\nabla g\|_2}{|\partial_{dd} g|}\right)(H) \cdot \mathrm{d}h_{11} \ldots \mathrm{d}h_{d,d-1} \tag{84}$$

$$\;=\; \int_{\mathcal{N}_d^*} \widetilde{J}_d(\boldsymbol{\beta}', \boldsymbol{u}) \cdot \widetilde{\Gamma}_d(\boldsymbol{\beta}', \boldsymbol{u}) \cdot d\lambda_{p+d-1}\,(\boldsymbol{\beta}', \boldsymbol{u})\,. \tag{85}$$

A direct use of (84) requires that the coefficient h_{dd} be expressed as a function of the other entries of the matrix H. To this end, it should be mentioned that, as established in the course of the proof of Lemma 6 below, the coefficient h_{dd} appears only once (in the form h_{dd}^2) in the determinant defining the set $\widetilde{\mathcal{M}}_d$ in (69) — see §4.5 for details.

If one wants cruder but simpler–to–obtain estimates for the right–hand side of (83), it should first be noted that the density function $\widetilde{\Gamma}_d$ defined in (82) and appearing in (83)

and (85) as a function of L and in (84) as a function of H is clearly bounded below by 1. In order to bound it from above, one can bound the gradient therein from above with the help of Remark 2. Also, the explicit formula given in Equation (93) below for the partial derivative $(\partial_{dd} g)(H)$ can easily be used to bound the latter quantity from below as a function of h_{dd}, γ and c_0.

The lower bound appearing in Theorem 5 involves the computation of the integral of an algebraic function (more precisely : the square root of some rational function) over an algebraic domain (which can be explicitly defined with the help of inequalities involving polynomials). This can certainly be done numerically in such a way that Theorem 5 can be seen as a way to obtain numerical values for the quantity $\mathfrak{m}_d(\delta)$. A more theoretical approach would necessarily require involved calculations which can nevertheless be carried out for a fixed value of d.

As mentioned in §4.1, the case of $d = m = 2$ users and $n = 2$ receivers is already of interest in the theory of Signal Processing. We explicitly work out the estimates that can be obtained from Theorem 5 in this case. In order to put the emphasis on the behaviour of the probability $\mathfrak{m}_2(\delta)$ as a function of δ and in order not to introduce unnecessary cumbersome definitions, we present the result in the following way, where an explicit expression for the function χ follows immediately from the proof presented in §4.6 (see Equation (94) below) :

Corollary 1. *Assume that $c_0 > \gamma^2$ and that $\delta_2^* := \gamma / c_0^{1/2} < \delta < 1$. Then, there exists a function χ such that*

$$\mathfrak{m}_2(\delta) \;\geqslant\; \gamma^{-1} c_0^{-1/2} \cdot \int_{\sqrt{\delta}}^{1/\sqrt{\delta}} \frac{\mathrm{d}a}{\sqrt{c_0^{1/2} a^2 - \gamma}} \int_{-\theta(a)}^{\theta(a)} \mathrm{d}b \cdot \frac{\chi(a,b)}{\sqrt{\theta(a)^2 - b^2}} \;:=\; \mathfrak{n}_2(\delta), \tag{86}$$

where

$$\theta(a) \;:=\; \sqrt{\frac{1}{\gamma c_0^{1/2}} \cdot \left(\frac{c_0^{1/2}}{a^2} - \gamma \right) \cdot \left(c_0^{1/2} a^2 - \gamma \right)} \tag{87}$$

and where the right–hand side is equal to 1 when $\delta = \delta_2^*$.

Furthermore, the function χ takes its values in a interval of the form $[\omega_1, \omega_2]$, where the constants ω_1 and ω_2 are such that $0 < \omega_1 < \omega_2 < +\infty$ and depend only on γ and c_0.

The corollary implies that the probability $\mathfrak{m}_2(\delta)$ tends to 1 as δ tends to the critical value δ_2^* with an error term governed by the size of the difference $\mathfrak{n}_2(\delta_2^*) - \mathfrak{n}_2(\delta)$. Note that upon bounding the function χ from above by the constant ω_2, the inner integral in (86) becomes independent of the variable a. This shows that the error term in the difference $1 - \mathfrak{m}_2(\delta)$ is, up to a multiplicative constant, bounded above by

$$\left(\int_{\sqrt{\delta_2^*}}^{1/\sqrt{\delta_2^*}} - \int_{\sqrt{\delta}}^{1/\sqrt{\delta}} \right) \frac{\mathrm{d}a}{\sqrt{c_0^{1/2} a^2 - \gamma}} \;=\; O\left(\delta - \delta_2^*\right)$$

(this relation follows from a direct evaluation of the integral in the left–hand side. Details of the calculations are left as an exercise for the interested reader). We thus recover when $d = 2$ the growth in $\delta^{d/2}$ as in Theorem 1.

Typical values for the capacity C_0 of a channel and for the Signal–to–Noise Ratio SNR can be taken as $C_0 = 30$ bits and $SNR = 5$ dB. From the expression for the function χ deduced from the proof of Corollary 1, one can find an explicit lower bound for the probability that the Effective Signal–to–Noise Ratio SNR_{eff} should be bigger than a given value $s \geqslant 0$. From the discussion held at the beginning of §4.4, this amounts to bounding from below the quantity $\mathfrak{m}_2(\delta)$ when δ (hereafter denoted by δ_s) is viewed as a function of s according to (76). Note that with such choices, $\gamma = 1/5$ and $c_0 = e^{30}/25$. Furthermore, $\delta_2^* = e^{-15} \approx 3.06 \cdot 10^{-7}$ arises from the limit value $s_2^* = 5/16 = 0.3125$. Some numerical values are recorded in the following table.

s	$s_2^* = 0.3125$	1	1.5	2
$\delta_s \approx$	$3.06 \cdot 10^{-7}$	$9.79 \cdot 10^{-7}$	$1.47 \cdot 10^{-6} \cdot 10^{-7}$	$1.96 \cdot 10^{-6} \cdot 10^{-7}$
$\mathfrak{m}_2(\delta_s) \geqslant$	1	0.672723	0.560289	0.489859

s	5	10	30
δ_s	$4.90 \cdot 10^{-6} \cdot 10^{-7}$	$9.79 \cdot 10^{-6} \cdot 10^{-7}$	$2.94 \cdot 10^{-5}$
$\mathfrak{m}_2(\delta_s) \geqslant$	0.314961	0.223899	0.12972

Thus, for instance, to ensure that the event $SNR_{eff} \geqslant s$ occurs with probability at least 45%, it is enough to choose $s = 2$. Also, the initial value of $SNR = 5$ is recovered with probability at least 31%.

As a concluding remark, we would like to mention here that, from a numerical point of view, the computation of the Cholesky transforms required to estimate the integrals in Theorem 5 can be implemented in a much more efficient and stable way than using Equations (79) and (80). For further details, the interested reader is referred to [19] and to the references therein.

The rest of this section is devoted to the proofs of Lemma 6 and Corollary 1.

4.5. **Proof of Lemma 6.** The second point is proved in [9, Chap. 11, §C].

As for the first point, given $T := (t_{ij})_{1 \leqslant i \leqslant j \leqslant d} \in \widetilde{\mathcal{M}}_d$ and $\beta > 0$, consider the homogeneous polynomial F of degree $2d$ defined as

$$F(T, \beta) \ := \ \det\left(\beta^2 I_d + {}^t T \cdot T\right).$$

Note that

$$F(T, \gamma^{1/2}) \underset{(72)}{=} c_0 \cdot g(T) \tag{88}$$

and assume for a contradiction that

$$\partial_{ij} F(T, \gamma^{1/2}) = 0 \tag{89}$$

for all $1 \leqslant i \leqslant j \leqslant d$.

It follows from Euler's formula for the derivative of a homogeneous function that

$$2d \cdot F(T,\beta) \;=\; \sum_{1 \leqslant i \leqslant j \leqslant d} t_{ij} \cdot \partial_{ij} F(T,\beta) + \beta \cdot \partial_\beta F(T,\beta)$$

(here, ∂_β obviously denotes the partial derivative with respect to the last variable β). Under (89), this implies that

$$2d \cdot F\left(T, \gamma^{1/2}\right) \;=\; \gamma^{1/2} \cdot \partial_\beta F(T, \gamma^{1/2}). \tag{90}$$

Let $[\![1,d]\!]$ denote the interval of positive integers less than d. Given $K \subset [\![1,d]\!]$, denote furthermore by $|K|$ the cardinality of K and by m_K the $|K| \times |K|$ matrix obtained by considering the rows and columns indexed by K in the matrix ${}^tT \cdot T$. Set conventionally

$$\det m_\varnothing \;:=\; 1.$$

As m_K is the Gramian matrix of the columns of T indexed by K, $\det m_K$ is non–negative. Furthermore, the definition of the determinant readily implies that

$$F(T,\beta) \;=\; \sum_{K \subset [\![1,d]\!]} \beta^{2d-2|K|} \det m_K. \tag{91}$$

Differentiating with respect to β and multiplying throughout by β then yields

$$\beta \cdot \partial_\beta F(T,\beta) \;=\; \sum_{K \subset [\![1,d]\!]} (2d - 2\,|K|) \beta^{2d-2|K|} \det m_K. \tag{92}$$

On combining (90), (91) and (92), one thus obtains the relation

$$2d \sum_{K \subset [\![1,d]\!]} \gamma^{d-|K|} \det m_K \;=\; \sum_{K \subset [\![1,d]\!]} (2d - 2\,|K|) \gamma^{d-|K|} \det m_K,$$

i.e.

$$\sum_{K \subset [\![1,d]\!]} 2\,|K|\, \gamma^{d-|K|} \det m_K \;=\; 0.$$

Since each term on the left–hand side of this equation is positive, this implies that $\det m_K = 0$ for all non–empty $K \subset [\![1,d]\!]$, i.e. $T = \mathbf{0}$. Under assumption (68), this contradicts the fact that $T \in \widetilde{\mathcal{M}}_d$ and thus concludes the proof of the first point.

The third point is elementary : given $T \in \mathcal{M}_d^*$, the coefficient t_{dd} appears only in the bottom right corner in the matrix $\gamma I_d + {}^tT \cdot T$, where it is present as t_{dd}^2. Thus, after expanding the determinant $g(T)$ following the last column, one obtains that

$$\left(\partial_{dd} g\right)(T) \;=\; c_0^{-1} \cdot 2t_{dd} \cdot \det\left(\gamma I_{d-1} + {}^tT' \cdot T'\right), \tag{93}$$

where the matrix T' is obtained by stripping off the matrix T from its last column and row. Clearly, the latter quantity does not vanish under the assumption that T has full rank. This concludes the proof of the lemma.

. The claims made in Remark 2 can now be justified as follows : given $T \in \widetilde{\mathcal{M}}_d$ denote by $\boldsymbol{t_i}$ $(1 \leqslant i \leqslant d)$ the i^{th} column of the matrix T and by $\boldsymbol{t}$ this matrix viewed as a vector in $\mathbb{R}^{d(d+1)/2}$. Upon isolating the terms corresponding to $K = \varnothing$ and $K = \{i\}$ $(1 \leqslant i \leqslant d)$ from the others in (91), this equation together with (88) readily implies that $\|\boldsymbol{t}\|_2^2 \leqslant (c_0 - \gamma^d)/\gamma^{d-1}$. Conversely, it follows from Hadamard's inequality that the determinant of the positive definite matrix $\gamma I_d + {}^tT \cdot T$ is less than or equal to the product of its diagonal entries. Thus,

$$c_0 \;=\; \det\left(\gamma I_d + {}^tT \cdot T\right) \;\leqslant\; \prod_{i=1}^{d} \left(\gamma + \|\boldsymbol{t_i}\|_2^2\right) \;\leqslant\; \left(\gamma + \|\boldsymbol{t}\|_2^2\right)^d,$$

hence the fact that $\|\boldsymbol{t}\|_2^2 \geqslant c_0^{1/d} - \gamma$.

4.6. **Proof of Corollary 1.** Let

$$H \;:=\; \begin{pmatrix} u & v \\ 0 & w \end{pmatrix} \;\in\; \mathcal{M}_2^*$$

and

$$L \;:=\; \begin{pmatrix} a & b \\ 0 & 1/a \end{pmatrix} \;\in\; \Theta_2^{++}$$

be such that

$${}^tL \cdot L \;=\; c_0^{-1/2}\left(\gamma I_2 + {}^tH \cdot H\right).$$

Formulae (79) and (80) then read

$$u \;=\; \sqrt{c_0^{1/2}a^2 - \gamma}, \qquad v \;=\; \frac{c_0^{1/2}ab}{\sqrt{c_0^{1/2}a^2 - \gamma}}$$

and

$$w \;=\; \sqrt{c_0^{1/2}b^2 + \frac{c_0^{1/2}}{a^2} - \frac{c_0a^2b^2}{c_0^{1/2}a^2 - \gamma} - \gamma} \;=\; \sqrt{\frac{\left(\frac{c_0^{1/2}}{a^2} - \gamma\right)\cdot\left(c_0^{1/2}a^2 - \gamma\right) - \gamma c_0^{1/2}b^2}{c_0^{1/2}a^2 - \gamma}}.$$

This is easily seen to imply that the set $\mathcal{N}_2^*$ defined in (81) can be explicitly expressed as follows :

$$\mathcal{N}_2^* \;=\; \left\{(a,b) \in \mathbb{R}_{>0} \times \mathbb{R} \;:\; \sqrt{\delta_2^*} < a < (\sqrt{\delta_2^*})^{-1} \quad \text{and} \quad |b| < \theta(a)\right\},$$

where the quantity $\theta(a)$ has been defined in (87).

Furthermore, the function g defined in (72) reads in this case

$$g(u,v,w) \;=\; c_0^{-1} \cdot \left(\left(u^2 + \gamma\right)\cdot\left(w^2 + \gamma\right) + \gamma v^2\right)$$

and, with the notation of Theorem 5,

$$\widetilde{J}_2(a,b) \cdot \widetilde{\Gamma}_2(a,b) \;=\; \left(c_0 \cdot \frac{a^2}{u^2(a,b)}\right) \cdot \left(\frac{\widetilde{g}(a,b)}{2c_0^{-1} \cdot w(a,b) \cdot (u^2(a,b) + \gamma)}\right).$$

In this equation, the variables u and w are seen as functions of a and b and $\widetilde{g}$ is the norm of the gradient of g (with respect to u, v and w) also expressed as a function of the parameters a and b; that is, with obvious notation,

$$\widetilde{g}(a,b) \ := \ \left(\left\|\nabla_{(u,v,w)}\, g\right\|_2\right)(a,b).$$

Set

$$\chi(a,b) \ := \ \frac{c_0^2}{2\kappa_2} \cdot \frac{a^2 \cdot \widetilde{g}(a,b)}{u^2(a,b)+\gamma}, \tag{94}$$

where κ_2 is the constant defined for instance in (85).

The existence of the constants ω_1 and ω_2 is then guaranteed by the fact the parameter a stays bounded away from zero (see the expression of u above) and the fact that the gradient of g is continuous and never vanishes on the compact set $\widetilde{\mathcal{M}}_d$ (see Lemma 5 and Remark 2).

Note also that

$$u^2(a,b)\cdot w(a,b) \ = \ \gamma c_0^{1/2} \cdot \sqrt{c_0^{1/2}a^2-\gamma} \cdot \sqrt{\theta^2(a)-b^2}.$$

In order to conclude the proof, one needs to show that the right–hand side of (86) is equal to 1 when $\delta = \delta_2^*$; that is, that $\mathfrak{n}_2(\delta_2^*) = 1$. With the notation of Theorem 5, this readily follows from the fact that

$$\mathcal{N}_2^*\left[\delta_2^*\right] \ = \ \mathcal{N}_2^*$$

(such a relation does not hold any more in dimension $d \geqslant 3$).

Acknowledgement. The main catalyst for this work was the International Workshop on Interactions between Number Theory and Wireless Communication held at the University of York between 9–23 May 2014. The authors would like to thank the engineers, especially Uri Erez, Bobak Nazer and Or Ordentlich, for providing them with such an interesting topic of research which has turned out to be related to deep theoretical questions. The authors hope that this work will contribute to foster further collaboration between Number Theorists and Engineers.

References

[1] O.E. Barndorff–Nielsen, P. Blaesild and P. Svante Eriksen. *Decomposition and Invariance of Measures, and Statistical Transformation Models.* New–York : Springer, 1989.

[2] J.W.S. Cassels. *An introduction to the geometry of numbers.* Die Grundlehren der mathematischen Wissenschaften. Bd. 99. Berlin-Göttingen-Heidelberg: Springer–Verlag, 1959.

[3] J.P.R Christensen. On some measures analogous to Haar measure. *Math. Scand.*, 26 : 103–106, 1970.

[4] H. Cohen. *A course in computational algebraic number theory.* Berlin: Springer-Verlag, 1993.

[5] J.H. Conway and N.J.A. Sloane. *Sphere packings, lattices, and groups* (2nd Ed.). Die Grundlehren der mathematischen Wissenschaften. New–York : Springer–Verlag, 1993.

[6] G. Courtois. Sur les valeurs aux entiers des formes quadratiques reélles. In *Sur la dynamique des groupes de matrices et applications arithmétiques*, pp. 111–140. Palaiseau: Les Éditions de l'École Polytechnique, 2007.

[7] F. Dal'Bo. Points de vue sur les valeurs aux entiers des formes quadratiques binaires. In *Sur la dynamique des groupes de matrices et applications arithmétiques*, pp. 7–45. Palaiseau: Les Éditions de l'École Polytechnique, 2007.
[8] M.L. Eaton. The Wishart Distribution. In *Multivariate Statistics. A Vector Space Approach*, chap. 8 Institute of Mathematical Statistics, Beachwood, Ohio, USA, 2007.
[9] F. Jones. Lectures notes in Calculus. Available at : `http://www.owlnet.rice.edu/ fjones/`.
[10] B. Kirchheim and D. Preiss. Uniformly Distributed Measures in Euclidean Spaces. *Math. Scand.*, 90(1) : 152–160, 2002.
[11] D.Y. Kleinbock and G.A. Margulis. Logarithm laws for flows on homogeneous spaces. *Invent. Math.*, 138(3) : 451–494, 1998.
[12] D. Kleinbock, N. Shah and A. Starkov. Dynamics of subgroup actions on homogeneous spaces of Lie groups and applications to number theory. In *Handbook of dynamical systems. Volume 1A,* pp.813–930. Amsterdam: North-Holland, 2002.
[13] T.-S. Liu. Invariant measures on double coset spaces. *J. Aust. Math. Soc.*, 5 : 495–505, 1965.
[14] P. Mattila. *Geometry of sets and measures in Euclidean spaces. Fractals and rectifiability.* Cambridge : Univ. Press, 1995.
[15] J.M. Milnor. *Topology from the differentiable viewpoint. Based on notes by David W. Weaver. Revised 2nd ed.* Princeton, NJ: Princeton University Press, 1997.
[16] O. Ordentlich and U. Erez. Precoded Integer–Forcing Universally Achieves the MIMO Capacity to Within a Constant Gap. *IEEE Transactions on Information Theory*, 61(3) : 323–340, 2015.
[17] A. Strömbergsson. On the probability of a random lattice avoiding a large convex set. *Proc. Lond. Math. Soc. (3)*, 103(6) : 950–1006, 2011.
[18] D. Tse and P. Viswanath. Fundamentals of Wireless Communication. Cambridge University Press, 103(6) : 950–1006, 2011. 2005.
[19] D. Watkins. *Fundamentals of Matrix Computations.* New York : Wiley, 1991.
[20] Y. Yamasaki. Projective limit of Haar measures on O(n). *Publ. Res. Inst. Math. Sci.*, 2 : 141–149, 1972.
[21] J. Zhan, B. Nazer, U. Erez and M. Gastpar. Integer–Forcing Linear Receivers. *IEEE Transactions on Information Theory*, 60(12) : 7661–7685, 2014.
[22] K.Q.T. Zhang. *Wireless Communications: Principles, Theory and Methodology.* New York : Wiley, 2015.

FA, EZ: Department of Mathematics, University of York, York, YO10 5DD, UK

E-mail address: `faustin.adiceam@york.ac.uk, evgeniy.zorin@york.ac.uk,`