On the reduction of a random basis

Ali Akhavi, Jean-François Marckert, Alain Rouault

On the reduction of a random basis

Ali Akhavi * Jean-François Marckert † Alain Rouault ‡

April 14, 2006

Abstract

For $g < n$, let $b_1, \ldots, b_{n-g}$ be $n-g$ independent vectors in $\mathbb{R}^n$ with a common distribution invariant by rotation. Considering these vectors as a basis for the Euclidean lattice they generate, the aim of this paper is to provide asymptotic results when $n \to +\infty$ concerning the property that such a random basis is reduced in the sense of Lenstra, Lenstra & Lovász.
The proof passes by the study of the process $(r_{g+1}^{(n)}, r_{g+2}^{(n)}, \ldots, r_{n-1}^{(n)})$ where $r_j^{(n)}$ is the ratio of lengths of two consecutive vectors b_{n-j+1}^* and b_{n-j}^* built from $(b_1, \ldots, b_{n-g})$ by the Gram–Schmidt orthogonalization procedure, which we believe to be interesting in its own. We show that, as $n \to +\infty$, the process $(r_j^{(n)} - 1)_j$ tends in distribution in some sense to an explicit process $(\mathcal{R}_j - 1)_j$; some properties of this latter are provided.

1 Introduction.

We call *ambient space* the space $\mathbb{R}^n$ with its classical Euclidean structure. The Euclidean norm is denoted by $\|.\|$ and the scalar product by $\langle , \rangle$. Let $b_1^{(n)}, b_2^{(n)}, \ldots, b_p^{(n)}$ (for $p \leq n$) be a linearly independent system of p vectors of $\mathbb{R}^n$. The superscript $^{(n)}$ is used when needed to stress the dimension of the ambient space. The quantity

$$g = n - p,$$

is often used in this paper and referred to as the *codimension* of the independent system.

1.1 The Gram-Schmidt orthogonalization, the reduction level and the index of worst local reduction

To the independent system $b_1^{(n)}, b_2^{(n)}, \ldots, b_p^{(n)}$, the classical Gram-Schmidt orthogonalization procedure associates the orthogonal system $\widehat{b}_1^{(n)}, \cdots, \widehat{b}_p^{(n)}$ defined by the recursion

$$\widehat{b}_1^{(n)} = b_1^{(n)} \ , \ \widehat{b}_j^{(n)} = b_j^{(n)} - \sum_{i=1}^{j-1} \frac{\langle b_j^{(n)}, \widehat{b}_i^{(n)} \rangle}{\|\widehat{b}_i^{(n)}\|^2} \, \widehat{b}_i^{(n)} \qquad \text{for } j \geq 2. \tag{1.1}$$

*LIAFA, Université Denis Diderot- Case 7014, 2 place Jussieu, F-75251 Paris Cedex 05 [akhavi@liafa.jussieu.fr]
†LABRI, Université Bordeaux I, 351 cours de la Libération 33405-Talence cedex. [marckert@labri.fr]
‡LMV, UMR 8100, Université de Versailles-Saint-Quentin, 45 Avenue des Etats-Unis, 78035-Versailles. [rouault@math.uvsq.fr]

If $B = [b_1^{(n)}, \cdots, b_p^{(n)}]$ is the $n \times p$ matrix with column vectors $b_1^{(n)}, \cdots, b_p^{(n)}$ in the canonical basis, this orthogonalization corresponds to the QR decomposition $B = QR$ where

$$Q = \left[\widehat{b}_1^{(n)}, \cdots, \widehat{b}_p^{(n)}\right]$$

is an orthogonal $n \times p$ matrix and R is an upper triangular $p \times p$ matrix ($R_{k,j} = 0\ ,\ 1 \leq j < k \leq n$) and

$$R_{jj} = 1, \quad R_{k,j} = \frac{\langle \widehat{b}_k^{(n)}, b_j^{(n)} \rangle}{\|\widehat{b}_k^{(n)}\|^2}\ ,\ 1 \leq k < j \leq n\,. \tag{1.2}$$

Definition 1.1 *Let $b_1^{(n)}, b_2^{(n)}, \ldots, b_p^{(n)}$ be a linearly independent system of vectors of $\mathbb{R}^n$ whose codimension is $g = n - p$. Let $\widehat{b}_1^{(n)}, \cdots, \widehat{b}_p^{(n)}$ be the associated Gram-Schmidt orthogonalized system. We call* reduction level *of $b_1^{(n)}, b_2^{(n)}, \ldots, b_p^{(n)}$ the quantity*

$$\mathcal{M}_n^g := \min_{i \in \{1,\ldots,n-(g+1)\}} \frac{\|\widehat{b}_{i+1}^{(n)}\|^2}{\|\widehat{b}_i^{(n)}\|^2},$$

We call index of worst local reduction *of $b_1^{(n)}, b_2^{(n)}, \ldots, b_p^{(n)}$ the quantity*

$$\mathcal{I}_n^g := \min\left\{ i : \ \frac{\|\widehat{b}_{n-i}^{(n)}\|^2}{\|\widehat{b}_{n-i-1}^{(n)}\|^2} = \mathcal{M}_n^g \right\}.$$

The motivation of these definitions is explained in Section 1.3. When the vectors $b_1^{(n)}, b_2^{(n)}, \ldots, b_p^{(n)}$ are chosen at random, the reduction level and the index of worst local reduction are two random variables, well defined whenever $b_1^{(n)}, b_2^{(n)}, \ldots, b_p^{(n)}$ is a linearly independent system. This paper is essentially devoted to the study of these random variables. The next subsection details the distribution we consider for the vectors $b_1^{(n)}, b_2^{(n)}, \ldots, b_p^{(n)}$.

1.2 Models of random bases

In this paper we assume that the b_i's are picked up randomly in $\mathbb{R}^n$, independently, and with the same distribution ν_n. Moreover we require ν_n to be invariant by rotation and to satisfy $\nu_n(0) = 0$. It is then well known (see [13] Th. 1.5.6 p.38 and Letac [12]) that the radial part $\|b_i^{(n)}\|$ and the angular parts $\theta_i^{(n)} := b_i^{(n)}/\|b_i^{(n)}\|$ are independent, and that the angular parts are uniformly distributed on $\mathbb{S}^{n-1} := \{x \in \mathbb{R}^n : \|x\| = 1\}$. We call such a model a "simple spherical model". Since we are interested in the asymptotic behavior of a random basis in $\mathbb{R}^n$ when n goes to $+\infty$, a spherical model will be a sequence of distributions (ν_n), each ν_n being a simple spherical model in $\mathbb{R}^n$.

The uniform distribution $\mathbb{U}_n$ in the ball $\mathbb{B}^n := \{x \in \mathbb{R}^n : \|x\| \leq 1\}$ – called the "random ball model" – is a particular case of spherical model. Under $\mathbb{U}_n$, the distribution of the radial part is

$$\mathbb{U}_n(\{x : \|x\| \leq r\}) = \mathbb{U}_n(\|b_1^{(n)}\| \leq r) = r^n, \quad 0 \leq r \leq 1\,. \tag{1.3}$$

Under a spherical model, $b_1^{(n)}, b_2^{(n)}, \ldots, b_p^{(n)}$ (for $p \leq n$) are a.s. linearly independent. We call it a (p-dimensional) *random basis*.

Our main results hold under assumption (1.2). This is a technical condition on the distribution (ν_n) which allows to transfer results concerning the uniform distribution on $\mathbb{S}^{n-1}$ to more general spherical distributions.

Assumption 1.2 *There exists a deterministic sequence $(a_n)_n$ and constants $d_1, d_2, \alpha > 0$, $\rho_0 \in (0,1)$ such that, for every n and $\rho \in (0, \rho_0)$*

$$\nu_n \left(\left| \frac{\|b_1^{(n)}\|^2}{a_n} - 1 \right| \geq \rho \right) \leq d_1 e^{-n d_2 \rho^\alpha} . \tag{1.4}$$

This implies in particular that $\sup \left\{ \left| \frac{\|b_i^{(n)}\|^2}{a_n} - 1 \right|, i \in \{1, \ldots, n\} \right\} \xrightarrow[n]{proba} 0$.

Here are three natural examples of model ν_n where such a sequence (a_n) exists:

- ν_n is the uniform distribution on $\mathbb{S}^{n-1}$. In this case $\|b_1^{(n)}\|^2 = 1$, and $a_n = 1$.
- $\nu_n = \mathbb{U}_n$. In this case, $a_n = 1$ and by (1.3),

$$\mathbb{U}_n(|\|b_1^{(n)}\|^2/a_n - 1| \geq \rho) = (1-\rho)^{n/2} \leq e^{-n\rho/2}.$$

- ν_n is the n-variate standard normal (the coordinates are i.i.d. $\mathcal{N}(0,1)$). Then $\|b_1^{(n)}\|^2/2$ is $\gamma_{n/2}$-distributed. For $a_n = n$,

$$\mathbb{P}(|\|b_1^{(n)}\|^2/n - 1| \geq \rho) \quad = \quad \mathbb{P}(\gamma(n/2) \geq (1+\rho)\frac{n}{2}) + \mathbb{P}(\gamma(n/2) \leq (1-\rho)\frac{n}{2}).$$

The Laplace transform $\mathbb{E}(e^{t\gamma(n/2)})$ of $\gamma(n/2)$ is $(1-t)^{-n/2}$, and its Cramèr transform is

$$H^{(n/2)}(x) = \sup_{\theta<1} \left\{ \theta x - \log \mathbb{E}(e^{\theta\gamma(n/2)}) \right\} = x - \frac{n}{2} + \frac{n}{2}\log(n/(2x)), \quad x \geq 0. \tag{1.5}$$

By Markov , $\mathbb{P}(\gamma(n/2) \geq (1+\rho)\frac{n}{2}) \leq e^{-H^{(n/2)}((1+\rho)n/2)} = e^{-\frac{n}{2}(\rho - \log(1+\rho))}$ and by an analogous calculus, $\mathbb{P}(\gamma(n/2) \leq (1-\rho)\frac{n}{2}) \leq e^{\frac{n}{2}(\rho + \log(1-\rho))}$. Hence assumption (1.2) holds in this case with $\alpha = 2$.

Notice that these three models are cited in the book of Knuth ([9, Section 3.4.1]).

The motivation to study the random variables $\mathcal{M}_n^g$ and $\mathcal{I}_n^g$ comes from the theory of "lattice basis reduction". The next section briefly describes this motivation and expresses our result in the vocabulary of this theory. The reader who is not interested by this theory may skip the next section.

1.3 LLL reduction of a random lattice

Let $b_1^{(n)}, b_2^{(n)}, \ldots, b_p^{(n)}$ (for $p \leq n$) be a linearly independent system of p vectors of $\mathbb{R}^n$. The set of all their integer linear combinations is an additive discrete subgroup of $\mathbb{R}^n$ called a lattice. The system $b_1^{(n)}, b_2^{(n)}, \ldots, b_p^{(n)}$ is then a *basis* of the lattice. The integer p is the *dimension* of the lattice or the dimension of the basis. The codimension of the lattice basis is the codimension $g = n - p$ of the linearly independent system $b_1^{(n)}, \ldots, b_p^{(n)}$. The basis is called *full* if $g = 0$.

The lattice basis reduction problem deals with finding a basis of a given lattice, whose vectors are "short" and "almost orthogonal". The problem is old and there are numerous notions of reduction.

For a general survey, see for example [8, 16, 7]. Solving even approximately the lattice basis reduction problem has numerous theoretical and practical applications in integer optimization [11], computational number theory [10] and cryptography [14].
In 1982, Lenstra, Lenstra and Lovász [10] introduced for the first time an efficient (polynomial with respect to the length of the input) approximation reduction algorithm. It depends on a real approximation parameter $s \in]0, \sqrt{3}/2[$ and is called LLL(s). The output basis of the LLL algorithm is called an LLL(s) reduced or s-reduced basis. In this paper we are concerned with the probability that a random basis under a spherical model is LLL(s) reduced, (i.e. is already an output basis of the LLL(s)-algorithm).
Roughly speaking the LLL reduction procedure is an approximation algorithm following a divide and conquer paradigm: Indeed for $i \in \{1 \dots p-1\}$, the following condition (1.6) ensures that some "local two dimensional basis" is s-reduced. This two dimensional basis is the projections of $b_i^{(n)}$ and $b_{i+1}^{(n)}$ into the orthogonal $H_i^{\perp}$ of the vector space H_i spanned by $b_1^{(n)}, b_2^{(n)}, \dots, b_{i-1}^{(n)}$. [10] showed that when all these two–dimensional bases are s-reduced then the whole basis has nice enough Euclidean properties. For instance, the length of the first vector of an LLL-reduced basis is not longer than $(1/s)^{(p-1)}$ times the length of a shortest vector in the lattice generated by $b_1^{(n)}, b_2^{(n)}, \dots, b_p^{(n)}$.
The next definition characterizes an LLL(s) reduced basis.

Definition 1.3 *Let $b_1^{(n)}, b_2^{(n)}, \dots, b_p^{(n)}$ (for $p \leq n$) be a linearly independent system of p vectors of $\mathbb{R}^n$. It is an LLL(s)-reduced basis of the lattice that it generates iff for all $1 \leq i \leq p-1$,*

$$\frac{\|\widehat{b}_{i+1}^{(n)}\|^2}{\|\widehat{b}_i^{(n)}\|^2} > s^2. \tag{1.6}$$

There are two minor differences between the definition of LLL reduction we consider here and the original definition introduced in [10].
Firstly in the original definition the basis has also to be *proper*, i.e. if $B = QR$ is the decomposition (1.2) associated with the Gram–Schmidt orthogonalization of the basis $b_1^{(n)}, b_2^{(n)}, \dots, b_p^{(n)}$, then

$$-1/2 \leq R_{k,j} < 1/2 \ , \ 1 \leq k < j \leq n \, . \tag{1.7}$$

But from any basis satisfying (1.6) one efficiently obtains a proper basis still satisfying (1.6) by a straightforward sequence of integer translations provided in appendix. Moreover considering the notion of *flag* [7] rather than basis for lattices, makes it possible to skip the notion of properness.
Secondly the approximation parameter of the original LLL in [10] is slightly different from the one we use here and the reduction we consider here is indeed Siegel reduction as called in [2, 1]. Our main Theorem 1.4 is still true with the original definition of a LLL reduced basis as detailed in appendix.

In this paper we study the asymptotics (with respect to the dimension n of the ambient space) of the random variables $\mathcal{M}_n^g$ and $\mathcal{I}_n^g$ under spherical models and for general codimensions of the random basis. The variable $\mathcal{M}_n^g$ is the supremum of the set of those s for which the basis is s^2-reduced. As mentioned earlier an LLL(s) reduced basis satisfies a set of local conditions. The second variable $\mathcal{I}_n^g$ is the place where the satisfied local condition is the weakest. This indicates where the limitation of the reduction comes from locally.

Theorem 1.4 *Let $b_1^{(n)}, b_2^{(n)}, \dots, b_{n-g}^{(n)}$ be a random basis with codimension g under a spherical model (ν_n) satisfying Assumption (1.2). Let $s \in (0,1)$ be a real parameter.*
(i) If $g = g(n)$ tends to infinity, then the probability that a random basis is s–reduced tends to 1.
(ii) If g is constant then the probability that a random basis is s–reduced converges to a constant in $(0,1)$ (depending on s and g).
(iii) If g is constant, the index of worst local reduction $\mathcal{I}_n^g$ converges in distribution.

Theorem 1.4 answers positively to a conjecture of Akhavi [2] (which says that for $c \in [0,1)$, $\mathcal{M}_n^{cn-1} \xrightarrow[n]{proba.} 1$). In his Lemma 3 p. 376, he proved that $\mathbb{P}(\mathcal{M}_n^{cn-1} \leq s) \to 0$, as soon as $s < \frac{1}{2}(1-c)^{\frac{1-c}{c}}(1+c)^{\frac{1}{c}}$, and that this convergence is exponentially fast. The proof of Theorem 1.4 relies on some properties of random basis under the spherical model which are of interest by their own; these results are overviewed in the next section.

Notice that in [6], Donaldson proved a phenomenon similar to the assertion (i) of Theorem 1.4. He considered a different random model: The basis $b_1^{(n)}, \cdots, b_{n-g}^{(n)}$ is picked up uniformly in the set $\{\|b_1^{(n)}\|^2 + \cdots + \|b_{n-g}^{(n)}\|^2 = 1\}$ (Euclidean sphere in $\mathbb{R}^{n\times(n-g)}$). He proved that as $n \to \infty$ with $n - g(n)$ a fixed constant , the basis is asymptotically reduced in the sense of Minkowski, i.e. each $b_i^{(n)}$ is a shortest vector among all vectors of the lattice that complete $b_1^{(n)}, \cdots, b_{i-1}^{(n)}$ to form a bigger subset of a lattice basis. So his result is about a stronger notion of reduction but he considered a much more restricted class of basis.

To finish this Section about lattice basis reduction, observe that our Theorem 1.4 about LLL reduction can be generalized to other reductions: In [15] Schnorr introduces a new type of reduction by segments. In this setting one fixes an integer k and partitions a basis whose vectors are in $\mathbb{R}^n$ and whose codimension is g into m segments of k consecutive basis vectors such that $n - g = km$. For a basis with codimension g, the reduction criterion is based on the quantity

$$M_{k,n}^g = \inf_{r:(k+1)r\leq n-g} \frac{\|\widehat{b}_{kr+1}^{(n)}\|^2 \cdots \|\widehat{b}_{(k+1)r}^{(n)}\|^2}{\|\widehat{b}_{k(r-1)+1}^{(n)}\|^2 \cdots \|\widehat{b}_{kr}^{(n)}\|^2} \tag{1.8}$$

Similarly to the assertions of Theorem 1.4, if $g = g(n)$ tends to infinity and the block size k is fixed, then for any $s \in [0,1]$ the probability that a random basis is s–reduced in the sense introduced by Schnorr tends to 1 with n. If g is constant then this probability tends to a constant in $[0,1]$ (depending on s, g and k). [1]

1.4 Random bases issued from spherical models

For any $j = 1, \dots, n$, let

$$Y_j^{(n)} := \|\widehat{b}_j^{(n)}\|^2 / \|b_j^{(n)}\|^2.$$

We denote by γ_a and $\beta_{a,b}$ respectively the gamma distribution with parameter a, and the beta distribution with parameter a and b. In the sequel $\gamma(a)$ and $\beta(a,b)$ stand for generic random

[1] Of course there is a choice of approximation parameters such that when a basis is LLL(s) reduced then for any fixed k, it is also s,k-reduced in the sense introduced by Schnorr. But our approach here shows the existence of limit probabilities (with n) for the reduceness of a random basis in the sense introduced by Schnorr.

variables with respective distribution γ_a and $\beta_{a,b}$. Some classical properties of these distributions are recalled in the appendix.

We first recall some facts concerning the spherical models, facts that are more or less part of the folklore, and which have been proved several times (e.g. [13], [2]).

Theorem 1.5 *For each n, under the simple spherical model, the variables $\|\widehat{b}_j^{(n)}\|^2$, $j = 1, \cdots, n$ are independent. For every $j = 2, \ldots, n$,*

$$Y_j^{(n)} \stackrel{(d)}{=} \beta\left(\frac{n-j+1}{2}, \frac{j-1}{2}\right), \tag{1.9}$$

and the random variables $Y_j^{(n)}, j \geq 1$, $\|b_j^{(n)}\|^2, j \geq 1$ are independent.

A probabilistic proof is given in Section 2.1 for the convenience of the reader.

Corollary 1.6 *Under the random ball model $\mathbb{U}_n$, the variables $\|\widehat{b}_j^{(n)}\|^2$, $j = 1, \cdots, n$ are independent and for $1 \leq j \leq n$*

$$\|\widehat{b}_j^{(n)}\|^2 \stackrel{(d)}{=} \beta\left(\frac{n-j+1}{2}, \frac{j+1}{2}\right). \tag{1.10}$$

As an easy consequence of the properties of the beta distribution, under $\mathbb{U}_n$,

$$\|\widehat{b}_{n-j}^{(n)}\|^2 \stackrel{(d)}{=} 1 - \|\widehat{b}_j^{(n)}\|^2. \tag{1.11}$$

The statement of Corollary 1.6 in this formulation is due to Daudé-Vallée ([5]). Actually, (1.10) is a consequence of Theorem 1.5 and identity (3.6), since (1.3) means that $\|b_i^{(n)}\|^2 \stackrel{(d)}{=} \beta(n/2, 1)$.

The random variable $\mathcal{M}_n^g$ has the representation :

$$\mathcal{M}_n^g = \min_{g+1 \leq j \leq n-1} r_j^{(n)} \ , \quad r_j^{(n)} := \frac{\|\widehat{b}_{n-j+1}^{(n)}\|^2}{\|\widehat{b}_{n-j}^{(n)}\|^2}. \tag{1.12}$$

As one can guess in view of Theorem 1.5, under ν_n, for each j, $r_j^{(n)}$ converges in distribution to $\gamma\left(\frac{j+1}{2}\right)/\gamma\left(\frac{j}{2}\right)$, where $\gamma\left(\frac{j+1}{2}\right)$ and $\gamma\left(\frac{j}{2}\right)$ are independent (see Proposition 2.1). By the strong law of large numbers, one sees that $\gamma\left(\frac{j+1}{2}\right)/\gamma\left(\frac{j}{2}\right) \xrightarrow[j]{a.s.} 1$; this allows to guess that the minimum $\mathcal{M}_n^g$ is reached by the firsts $r_j^{(n)}$; this motivates the time inversions done in (1.12).

The variable $\mathcal{M}_n^g$ is a function of the $(n-g)$-tuple $(r_{g+1}^{(n)}, \cdots r_{n-1}^{(n)})$, and then the convergence of each coordinate is not sufficient to yield that of $\mathcal{M}_n^g$. We have to take into account that the variables $(r_j^{(n)})_{j \leq n-1}$ are dependent, and that their number is growing. Since for the "last" indices ($n-i$ with i fixed), $r_{n-i}^{(n)} \xrightarrow[n]{(d)} 1$ (see (2.4)), it is convenient to embed the $(n-1)$-tuple $(r_1^{(n)}, \cdots r_{n-1}^{(n)})$ into $\mathbb{R}_+^{\mathbb{N}}$ (the set of infinite sequences of positive real numbers), setting

$$r_j^{(n)} := 1 \ , \ j \geq n. \tag{1.13}$$

Let $(\eta_i)_{i\geq 1}$ be a sequence of independent random variables such that $\eta_i \overset{(d)}{=} \gamma_{i/2}$ and set

$$\mathcal{R}_j = \eta_j/\eta_{j+1}\ ,\ j \geq 1\,. \tag{1.14}$$

We denote by $\|.\|_p$ the classical norm on the set ℓ_p of sequences of real numbers with finite pth moment: for any sequence of real numbers $\mathsf{x} = (\mathsf{x}_i)_{i\geq 1}$, $\|\mathsf{x}\|_p$ is $\left(\sum_{i\geq 1} |\mathsf{x}_i|^p\right)^{1/p}$ and ℓ_p is $\{\mathsf{x}, \|\mathsf{x}\|_p < +\infty\}$.

The following result states a limit behavior for the process $(r_j^{(n)})$ when n goes to$+\infty$.

Proposition 1.7 *For any $p > 2$, the following convergence in distribution holds in the metric space ℓ_p :*

$$(r_j^n - 1)_{j\geq 1} \xrightarrow[n]{(d)} (\mathcal{R}_j - 1)_{j\geq 1}.$$

The previous proposition is the key result here and it will entail all the convergence results given in the next theorem.
For $k \in \mathbb{N}$, set

$$\mathcal{M}^k := \inf\left\{\mathcal{R}_j, j \geq k+1\right\}.$$

The application $\mathsf{x} \mapsto 1 + \min_{i\geq k} \mathsf{x}_i$ is continuous from ℓ_p onto $\mathbb{R}$. It follows that $\mathcal{M}_n^g \wedge 1$ converges in distribution to $\mathcal{M}^g$. We will prove that

Theorem 1.8 *If ν_n is spherical and satisfies Assumption 1.2 then,*
(i) For each k, $\mathcal{M}_n^k \xrightarrow[n]{(d)} \mathcal{M}^k$.
(ii) Let $g : \mathbb{N} \to \mathbb{N}$ such that $g(n) \leq n$ and $g(n) \to \infty$. We have $\mathcal{M}_n^{g(n)} \xrightarrow[n]{proba.} 1$.
(iii) For any $k \geq 1$, $\mathcal{I}_n^k \xrightarrow[n]{(d)} \mathcal{I}^k$.

Notice that Proposition 1.7 and Theorem 1.8 have their analogous for the reduction introduced by Schnorr in [15]. By setting

$$\mathcal{M}_{k,n}^g = \min_{g+1\leq kr\leq n-1} r_{k,r}^{(n)}\ ,\quad r_{k,r}^{(n)} := \frac{\|\widehat{b}_{n-(r+1)k+1}^{(n)}\|^2 \cdots \|\widehat{b}_{n-rk}^{(n)}\|^2}{\|\widehat{b}_{n-(r+2)k+1}^{(n)}\|^2 \cdots \|\widehat{b}_{n-(r-1)k}^{(n)}\|^2} \text{ and } \quad r_{k,r}^{(n)} := 1 \text{ for } kr \geq n,$$

if we let $n \to \infty$, we have convergence of $(r_{k,r}^{(n)})_r$ to a process $(\mathcal{R}_{k,r})_r$ with

$$\mathcal{R}_{k,r} = \frac{\eta_{k,r}}{\eta_{k,r+1}}\ ,\ \eta_{k,r} \overset{(d)}{=} \gamma(r/2)\gamma((r+1)/2 \cdots \gamma((r+k-1)/2)\,,$$

where the $\eta_{k,r}, r \geq 1$ are independent, and the gamma variables too. Then by setting

$$\widetilde{\mathcal{M}}_k^g := \inf\left\{\mathcal{R}_{k,r}, kr \geq g+1\right\}.$$

one obtains also an analogous to Theorem 1.8.

A note on the proof of Proposition 1.7. The ambient spaces $\mathbb{R}^n, n \geq 1$ are not nested, and then we give up the geometrical consideration on $\mathbb{R}^n$ and focus on the representation of the processes r_j^n using the gamma distributions.

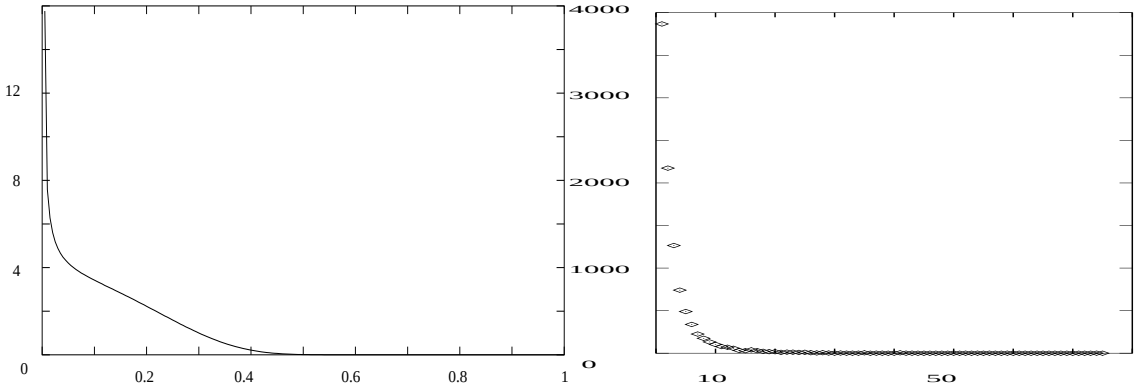

Figure 1: On the first picture, simulation of the density of $\mathcal{M}_\infty^0$ with 10^8 data. On the second, the histogram provided by 10000 simulations of $\mathcal{I}_\infty$. The sequence $k \mapsto P(\mathcal{I}^g = k)$ seems to be decreasing.

We end this section by stating some properties of the limiting process $(\mathcal{R}_k)_{k\geq 1}$. First of all, in statistics the distribution of $\frac{j+1}{j}\mathcal{R}_j$ is known as the Fisher $F_{j,j+1}$-distribution (its distribution is recall in (3.5)); the mean of $\mathcal{R}_j$ is $j/(j-1)$ and, as said above, $\mathcal{R}_k \xrightarrow[k]{a.s.} 1$. Here are some sharper results (see also simulations on Figure 1).

Proposition 1.9 *(i) For each k, the distribution of $\mathcal{M}^k$ has a density, which is positive on $(0,1)$ and zero outside.*

(ii) For each k,

$$\lim_{x\downarrow 0} x^{-(k+1)/2}\mathbb{P}(\mathcal{M}^k \leq x) = 1/\Gamma\left(\frac{k+2}{2}\right).$$

(iii) There exists $\tau > 0$ such that for each $k \geq 0$,

$$\limsup_{y\uparrow 1} e^{\frac{\tau}{(1-y)^2}}\,\mathbb{P}(\mathcal{M}^k \geq y) < \infty\,.$$

(iv) For each k, there is a.s. a unique random index $\mathcal{I}^k$ such that $\mathcal{R}_{\mathcal{I}^k} = \mathcal{M}^k$.

2 Proofs

2.1 Additional information on random basis

We first give a proof of Theorem 1.5 for convenience.

Proof of Theorem 1.5: Let us skip the superscript (n) in this proof. We have $b_i = \theta_i\|b_i\|$ and from (1.1), we see that $\widehat{b}_i = \|b_i\|\widehat{\theta}_i$, where the $\widehat{\theta}_i$'s are obtained by the Gram-Schmidt algorithm applied to the θ_i's. The independence of $(\widehat{\theta}_1, \cdots, \widehat{\theta}_n)$ and $(\|b_1\|^2, \cdots, \|b_n\|^2)$ is then a direct consequence of the radial-angular independence. Notice that $\widehat{\theta}_1 = 1$.

Now, fix $k \geq 2$. Conditionally upon $\widehat{\theta}_1, \cdots, \widehat{\theta}_{k-1}$, the variable $\|\widehat{\theta}_k\|$ is distributed as the norm of the projection of a random vector uniformly distributed on $\mathbb{S}^{n-1}$ on $span\{\widehat{\theta}_1, \cdots, \widehat{\theta}_{k-1}\}$. Since the problem is invariant by rotation, this distribution is independent of $(\widehat{\theta}_1, \cdots, \widehat{\theta}_{k-1})$ which proves (recursively) that the $\|\widehat{\theta}_i\|$'s are independent. Moreover, $\|\widehat{\theta}_k\|$ is distributed as the norm of the projection of θ_k (or θ_1) on the subspace generated by the $n-k+1$ last vectors of the canonical basis. From Muirhead [13, Theorem 1.5.7, p. 38-39] the distribution of $\|\widehat{\theta}_k\|^2$ is $\beta_{\frac{n-k+1}{2}, \frac{k-1}{2}}$. $\square$

Here are some information on the asymptotic behavior of the random variables $Y_j^{(n)}$:

Proposition 2.1 *Under a spherical model, for each $j \geq 1$,*

$$\frac{n}{2}\, Y_{n-j}^{(n)} \quad \xrightarrow[n]{(d)} \quad \gamma_{\frac{j+1}{2}}\,, \tag{2.1}$$

$$Y_j^{(n)} \quad \xrightarrow[n]{(d)} \quad 1\,. \tag{2.2}$$

In view of Theorem 1.5 this yields:

Proposition 2.2 *Under a spherical model, if $\|b_1^{(n)}\|^2/a_n \xrightarrow[n]{(d)} 1$ for some deterministic sequence a_n, then for each $j \geq 1$,*

$$\frac{n}{2a_n}\, \|\widehat{b}_{n-j}^{(n)}\|^2 \quad \xrightarrow[n]{(d)} \quad \gamma_{\frac{j+1}{2}}\,, \tag{2.3}$$

$$\frac{1}{a_n}\, \|\widehat{b}_j^{(n)}\|^2 \quad \xrightarrow[n]{(d)} \quad 1\,. \tag{2.4}$$

Remark 2.3 *Under the same assumptions, we have also:*

If $h(n) \to \infty$ and $h(n)/n \to 0$, then

$$\frac{n}{h(n)a_n}\, \|\widehat{b}_{n-h(n)}^{(n)}\|^2 \xrightarrow[n]{proba.} 1\,. \tag{2.5}$$

If $0 < \alpha < 1$ et $k(n)/n \to 0$, then

$$\frac{1}{a_n}\, \|\widehat{b}_{\alpha n+k(n)}^{(n)}\|^2 \xrightarrow[n]{proba.} 1-\alpha\,. \tag{2.6}$$

This result stated under $\mathbb{U}_n$ can be found in [2, Theorem 8]. Let us give a new proof which prefigures the main arguments used to prove the convergences in Section 2.3.

Proof of Propositions 2.1 and 2.2 From Theorem 1.5 we have the decomposition,

$$\|\widehat{b}_{n-j}^{(n)}\|^2 \overset{(d)}{=} Y_{n-j}^{(n)}\, \|b_1^{(n)}\|^2\,, \tag{2.7}$$

with $Y_{n-j}^{(n)} \overset{(d)}{=} \beta\left(\frac{j+1}{2}, \frac{n-j-1}{2}\right)$. Let $(\xi_j)_{j\geq 1}$ be a sequence of i.i.d. $\gamma_{1/2}$-distributed random variables. From (3.3) and (3.2), we can write

$$Y_{n-j}^{(n)} \overset{(d)}{=} \frac{\sum_{m=1}^{j+1} \xi_m}{\sum_{m=1}^{n} \xi_m}. \tag{2.8}$$

By the strong law of large numbers,

$$\frac{\sum_{m=1}^{n}\xi_m}{n} \xrightarrow[n]{a.s.} \frac{1}{2},$$

and for each j, $\sum_{m=1}^{j+1}\xi_m \overset{(d)}{=} \gamma((j+1)/2)$, which yields (2.1). From this and the additional assumption $\|b_1^{(n)}\|^2/a_n \xrightarrow[n]{(d)} 1$, we see that (2.3) holds true. For (2.2), notice that $(1-Y_j^{(n)}) \overset{(d)}{=} Y_{n-j+2}^{(n)}$, and that $Y_{n-j+2}^{(n)} \xrightarrow[n]{proba.} 0$ by (2.1). To end, (2.4) is a consequence of (2.2) and $\|b_1^{(n)}\|^2/a_n \overset{(d)}{\to} 1$. □

The following lemma will be used to transfer results from the uniform distribution on $\mathbb{S}^{n-1}$ to more general spherical distributions.

Lemma 2.4 *Assume that Assumption 1.2 holds. If U_1 and U_2 be independent and $U_1 \overset{(d)}{=} U_2 \overset{(d)}{=} \|b_1^{(n)}\|^2$, then there exist $d_1', d'2, \alpha > 0$ and $\rho_0 \in$ such that for any $k \geq 1, n \geq 1$ and $\rho \in (0, \rho_0)$*

$$\mathbb{P}\left(\left|\frac{U_1}{U_2} - 1\right| \geq \rho\right) \leq d_1' \exp(-n d_2' \rho^{\alpha}). \tag{2.9}$$

Proof: We have

$$\begin{aligned}\mathbb{P}\left(\frac{U_1}{U_2} \geq 1+\rho\right) &\leq \mathbb{P}(U_2 \leq (1-\rho/2)) + \mathbb{P}(U_1 \geq (1+\rho)(1-\rho/2)) \\ &\leq \mathbb{P}(U_2 \leq (1-\rho/2)) + \mathbb{P}(U_1 \geq (1+\rho/4)))\end{aligned}$$

as soon as $\rho \leq 1/2$. Similarly

$$\begin{aligned}\mathbb{P}\left(\frac{U_1}{U_2} \leq 1-\rho\right) &\leq \mathbb{P}(U_2 \geq (1+\rho/2)) + \mathbb{P}(U_1 \leq (1-\rho)(1+\rho/2)) \\ &\leq \mathbb{P}(U_2 \geq (1+\rho/2)) + \mathbb{P}(U_1 \leq (1-\rho/2))\end{aligned}$$

With the help of assumption (1.4), this yields

$$\mathbb{P}\left(\left|\frac{U_1}{U_2} - 1\right| \geq \rho\right) \leq d_1' \exp(-n d_2' \rho^{\alpha}).$$

□

2.2 The process $(\mathcal{R}_k)$: estimates and proof of Proposition 1.9

Lemma 2.5 and Proposition 2.6 first state some properties concerning the fluctuations and large deviations of the distribution $\mathcal{R}_k$

Lemma 2.5 *The following convergence in distribution holds*

$$\sqrt{k}\,(\mathcal{R}_k - 1) \xrightarrow[k]{(d)} \mathcal{N}(0,4)\,.$$

Proof : Setting

$$\xi_k = \frac{\eta_k - k/2}{\sqrt{k}} \quad \text{and} \quad \xi'_k = \frac{\eta_{k+1} - (k+1)/2}{\sqrt{k}}$$

the CLT gives $(\xi_k, \xi'_k) \xrightarrow[k]{(d)} \mathcal{N}(0,1/2) \otimes \mathcal{N}(0,1/2)$ hence $\xi_k - \xi'_k \xrightarrow[k]{(d)} \mathcal{N}(0,1)$. Since

$$\sqrt{k}\,(\mathcal{R}_k - 1) = \frac{k}{\eta_{k+1}} \left(\xi_k - \xi'_k - \frac{1}{2\sqrt{k}} \right) ,$$

and $\eta_{k+1}/k \to 1/2$ a.s., we get the result. □

Proposition 2.6 *Let $f_{\mathcal{R}_k}$ be the density of $\mathcal{R}_k$ and*

$$\Phi_k(x) = (4x)^{\frac{k}{2}-1}(1+x)^{-k-\frac{1}{2}} .$$

1. *For $A < 2\pi^{-1/2} < B$ we can find an integer K such that*

$$A\sqrt{k}\ \Phi_k(x) \le f_{\mathcal{R}_k}(x) \le B\sqrt{k}\ \Phi_k(x) \tag{2.10}$$

   for every $x \in (0,\infty)$ and every $k \ge K$.

2. *There exists a constant C such that for every $k \ge 1$ and $\rho \in [0,1]$*

$$\mathbb{P}(\mathcal{R}_k < 1-\rho) \le C \left(1 - \frac{\rho^2}{(2-\rho)^2}\right)^{k/2} \tag{2.11}$$

$$\mathbb{P}(\mathcal{R}_k > 1+\rho) \le C \left(1 - \frac{\rho^2}{(2+\rho)^2}\right)^{k/2} . \tag{2.12}$$

3. *Assertion 2 holds true when $\mathcal{R}_k$ is replaced by $\mathcal{R}'_k := \dfrac{S_k^{(2)}}{S_k^{(1)}}$.*

Notice that the distribution of $\mathcal{R}'_k$ is known in statistics as the Fisher $F_{k,k}$.

Proof: 1) We have $f_{\mathcal{R}_k}(x) = C_k \Phi_k(x)$ where

$$C_k = 4^{1-\frac{k}{2}} \frac{\Gamma\left(k+\frac{1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{k+1}{2}\right)} = \frac{2}{\sqrt{\pi}} \frac{\Gamma\left(k+\frac{1}{2}\right)}{\Gamma(k)} \sim \sqrt{k}\frac{2}{\sqrt{\pi}} .$$

2) The bounds may be obtained by integration, but also by writing the beta variables as ratios of gamma variables and using Chernov's bounds. Noticing that $\mathcal{R}_k$ and $\mathcal{R}'_k$ are Fisher-distributed, the above results are related to section 4 of [3]. Since we need bounds holding for ρ depending on k, we use the classical Chernov's method :

$$\begin{aligned}
\mathbb{P}(\mathcal{R}_k > 1+\rho) &= \mathbb{P}\left(\eta_k - (1+\rho)\eta_{k+1} > 0\right) \\
&\le E\exp\left(\theta\eta_k - \theta(1+\rho)\eta_{k+1}\right) = \left(Ee^{\theta\eta_1}\right)^k \left(Ee^{-\theta(1+\rho)\eta_1}\right)^{k+1} \\
&= (1-\theta)^{-k/2}\left(1+\theta(1+\rho)\right)^{-(k+1)/2} \\
&= \left(1+\theta(1+\rho)\right)^{-1/2}\left((1-\theta)(1+\theta(1+\rho))\right)^{-k/2} .
\end{aligned}$$

The function $\theta \mapsto (1-\theta)(1+\theta(1+\rho))$ reaches its maximum for $\theta = \frac{\rho}{2(1+\rho)} \in (0,1)$, so that :

$$\mathbb{P}(\mathcal{R}_k > 1+\rho) \leq \left(1 - \frac{\rho^2}{(2+\rho)^2}\right)^{k/2}. \tag{2.13}$$

Similarly

$$\begin{aligned}
\mathbb{P}(\mathcal{R}_k < 1-\rho) &\leq E\exp\left(\theta(1-\rho)\eta_{k+1} - \theta\eta_k\right) \\
&= ((1+\theta))(1-\theta(1-\rho))^{-k/2}\left(1-\theta(1-\rho)\right)^{-1/2} \\
&\leq \sqrt{2}\left(1 - \frac{\rho^2}{(2+\rho)^2}\right)^{k/2}.
\end{aligned}$$

3) For $\mathcal{R}'_k$ the proof needs similar evaluations and is left to the reader. □

Thanks to these bounds on the deviation of the process $(\mathcal{R}_k)$ around the value 1, one may establish the following corollary.

Corollary 2.7 *For any $p > 2$, the process $(\mathcal{R}_k - 1)$ is a.s. in ℓ_p, i.e. $\sum_k |\mathcal{R}_k - 1|^p < \infty$ a.s..*

Proof Thanks to the Borel-Cantelli lemma, it is enough to find $v = (v_k)_{k\geq 1} \in \ell_p$, such that

$$\sum_k \mathbb{P}(|\mathcal{R}_k - 1| \geq v_k) < \infty. \tag{2.14}$$

Taking $\rho = k^{-\mu}$ in the bounds (2.12) and (2.11), we have $\sum_k \mathbb{P}\left(|\mathcal{R}_k - 1| > k^{-\mu}\right) < \infty$ if $1-2\mu > 0$. For $p > 2$, one may choose $\mu \in]1/p, 1/2[$ and $v_k = k^{-\mu}$. Then $(v_k)_{k\geq 1} \in \ell_p$ and satisfies (2.14). □

Proof of Proposition 1.9

Proof of (i)**.** We give a proof in the case $k = 0$, but the argument is the same for any $k > 0$.

First, since for any j, $\mathcal{R}_j > 0$ a.s. and since a.s., $\lim_j \mathcal{R}_j = 1$, the support of $\mathcal{M}^0$ is included in $[0,1]$. For the same reason, the sequence $(\mathcal{R}_k)$ does not accumulate at 0, which yields that the distribution of $\mathcal{M}^0$ has no atom at 0.

Using Lemma 2.5 write

$$\mathbb{P}(\mathcal{R}_j < 1) = \mathbb{P}\left(\sqrt{2j}\,(\mathcal{R}_j - 1) < 0\right) \xrightarrow[j\to\infty]{} 1/2.$$

Hence by the reverse Borel Cantelli lemma, a.s. there exists an infinite sequence of j such that $\mathcal{R}_{2j} < 1$, which yields that $\mathcal{M}^0$ has no atom at 1.

It remains to check that the support of $\mathcal{M}^0$ is exactly $[0,1]$ (see (1) below) and that $\mathcal{M}^0$ has a density (see (2) below).

(1) Let us prove that $\mathbb{P}(\inf_j \mathcal{R}_j \in [a,b]) > 0$, for every $[a,b] \subset [0,1]$. It is enough to find a sequence of (independent) events $B_j := \{\eta_j \in (\alpha_j, \beta_j)\}, j \geq 0$ such that

$$\bigcap_{j=1}^{\infty} B_j \subset \left\{\inf_j \mathcal{R}_j \in [a,b]\right\} \quad \text{and} \quad \prod_{j=1}^{\infty} \mathbb{P}(B_j) > 0. \tag{2.15}$$

Let $j_a = \inf\{j : j > 2(1+a)/(1-a)\}$, $A := j_a(1+a)/4$ and $c_1 < c_2$ in (a,b). Choose

$$\alpha_1 = Ac_1,\ \alpha_2 = A \quad , \quad \alpha_j = A \text{ for } 3 \le j \le j_a \ , \ \alpha_j = \frac{j(1+a)}{4} \qquad \text{for } j \ge j_a + 1\,,$$
$$\beta_1 = Ac_2,\ \beta_2 = \frac{Ac_1}{a} \quad , \quad \beta_j = \frac{A}{a} \text{ for } 3 \le j \le j_a \ , \ \beta_j = \frac{(j-1)(1+a)}{4a} \text{ for } j \ge j_a + 1\,.$$

We check easily that $B_1 \cap B_2 \subset \{\mathcal{R}_1 \in (a,c_2)\}$, and $B_j \cap B_{j+1} \subset \{\mathcal{R}_j \in (a,\infty)\}$ for $j \ge 2$. This proves the first claim of (2.15).

It remains to prove that the infinite product is convergent, i.e. that

$$\sum_{k > j_a} \mathbb{P}(B_k^c) < \infty\,. \tag{2.16}$$

For $j > j_a$, the interval (α_j, β_j) straddles the mean $j/2$ of η_j :

$$\alpha_j = \frac{j(1+a)}{4} < \frac{j}{2} \quad , \quad \beta_j \ge \frac{j(1+3a)}{8a} > \frac{j}{2}\,,$$

so that the large deviations inequalities hold:

$$\log \mathbb{P}(\eta_j < \alpha_j) \le -jH^{(1/2)}\left(\frac{1+a}{4}\right) \ , \ \log \mathbb{P}(\eta_j > \beta_j) \le -jH^{(1/2)}\left(\frac{1+3a}{8a}\right)$$

where $H^{(1/2)}$, the Cramér transform of $\gamma_{1/2}$ is given in (1.5). This yields a positive constant M such that for $j > j_a$

$$\mathbb{P}(B_j^c) = \mathbb{P}(\eta_j < \alpha_j) + \mathbb{P}(\eta_j > \beta_j) \le 2e^{-jM}$$

and the series is convergent, which proves (2.16) and $\mathbb{P}(\inf_j \mathcal{R}_j \in [a,b]) > 0$.

(2) According to Radon-Nikodym's theorem, it suffices to find a positive integrable function f on $(0,1)$, such that for any $[a,b] \subset (0,1)$,

$$\mathbb{P}(\mathcal{M}^0 \in [a,b]) \le \int_{[a,b]} f(x)dx$$

By the union bound, we have for every $b' \in (b,1)$:

$$\mathbb{P}(\mathcal{M}^0 \in [a,b]) = \mathbb{P}(\inf_{k\ge 1} \mathcal{R}_k \in [a,b]) \le \mathbb{P}\left(\cup_k \{\mathcal{R}_k \in [a,b']\}\right) \le \sum_{k\ge 1} \mathbb{P}\big(\mathcal{R}_k \in [a,b']\big)\,.$$

For $B > 2/\sqrt{\pi}$, thanks to formula (2.10), there exists $K \ge 1$ such that

$$\begin{aligned}
\sum_{k\ge K} \mathbb{P}\big(\mathcal{R}_k \in [a,b]\big) &\le B\int_a^{b'} \left(\sum_{k\ge K} \sqrt{k}\ \Phi_k(x)\right) dx \\
&\le \frac{B}{2}\int_a^{b'} \left[\sum_{k\ge 1} k\left(\frac{2\sqrt{x}}{1+x}\right)^{k-1}\right] \frac{dx}{\sqrt{x}(1+x)^{3/2}} \\
&= \frac{B}{2}\int_a^{b'} \frac{\sqrt{1+x}}{\sqrt{x}(1-\sqrt{x})^4}\ dx\,.
\end{aligned}$$

Since every $\mathcal{R}_k$ has a density, one may bound the $K-1$ first terms of the sum by $\int_a^{b'} f_1(x)dx$ for some integrable f_1. Then, since the bound holds true for any $b' > b$, we can let $b' \downarrow b$ and we get the result.

Proof of (ii) We have $\mathbb{P}(\mathcal{R}_{k+1} \leq x) \leq \mathbb{P}(\mathcal{M}^k \leq x) \leq \sum_{j \geq k+1} \mathbb{P}(\mathcal{R}_j \leq x)$. Using (3.5), one obtains, for $x \to 0$,

$$\mathbb{P}(\mathcal{R}_{k+1} \leq x) = \frac{\Gamma\left(\frac{k+3}{2}\right)\ x^{(k+1)/2}}{\frac{(k+1)}{2}\Gamma\left(\frac{k+1}{2}\right)\Gamma\left(\frac{k+2}{2}\right)}\ (1+o(1)) = \frac{x^{(k+1)/2}}{\Gamma\left(\frac{k+2}{2}\right)}\ (1+o(1))\,.$$

On the other hand, a simple computation shows that, when $x \to 0$,

$$\sum_{j \geq k+2} \mathbb{P}(\mathcal{R}_j \leq x) = O(x^{(k+3)/2}).$$

Proof of (iii) We have, for $j \geq k$

$$\mathbb{P}\left(\mathcal{M}^k > 1 - j^{-1/2}\right) \leq \prod_{i=j}^{2j} \mathbb{P}\left(\mathcal{R}_{2i} > 1 - j^{-1/2}\right) \leq \prod_{i=j}^{2j} \mathbb{P}\left(\mathcal{R}_{2i} > 1 - i^{-1/2}\right)\,.$$

From Lemma 2.5, we know that $\lim_k \mathbb{P}\left(\mathcal{R}_{2k} > 1 - k^{-1/2}\right) = \mathbb{P}(N > -\sqrt{2})$ where N is $\mathcal{N}(0,4)$. Taking $\tau > 0$ with $e^{-\tau} > \mathbb{P}(N > -\sqrt{2})$ we see that for j large enough

$$\mathbb{P}\left(\mathcal{M}^k > 1 - j^{-1/2}\right) \leq e^{-\tau j}$$

which ends the proof of (iii).

Proof of (iv) The support of $\mathcal{M}^k$ is $[0,1]$ and $\lim \mathcal{R}_j = 1$ a.s. so that the set $\{j \geq k+1, \mathcal{R}_j = \mathcal{M}^k\}$ is not empty. Moreover since there are no ties ($\mathbb{P}(\mathcal{R}_i = \mathcal{R}_j) = 0$ a.s. for $i \neq j$) this set is a.s. a singleton. □

2.3 The proofs of convergence (Theorem 1.8 and Proposition 1.7)

In order to prove Proposition 1.7 and Theorem 1.8, we build a probability space on which are defined some copies of the variables $\|b_i^{(n)}\|$, $i \geq 0$, $n \geq 0$ (and then also $r_j^{(n)}$) and the process $(\mathcal{R}_k)$. This space is not related with some embedding of $\mathbb{R}^n$ in some larger space: the proof is not geometrical. Thanks to that procedure, we will be able to use the strong law of large numbers obtaining in such a way strong versions of the convergences in distribution stated in Proposition 1.7 and Theorem 1.8.

From Theorem 1.5 and the representation (3.3) we see that

$$\begin{aligned} \|\widehat{b}_{n-k+1}^{(n)}\|^2 &= Y_{n-k+1}^{(n)} \|b_{n-k+1}^{(n)}\|^2 \\ Y_{n-k+1}^{(n)} \stackrel{(d)}{=} \frac{\sum_{m=1}^k \xi_m}{\sum_{m=1}^n \xi_m} \quad &, \quad \|b_{n-k+1}^{(n)}\|^2 \stackrel{(d)}{=} \|b_1^{(n)}\|^2 \end{aligned} \tag{2.17}$$

where the ξ_m's are $\gamma_{1/2}$ distributed, and $\|b^{(n)}_{n-k+1}\|^2$ is independent of the ξ_m's. Since the $\|\widehat{b}^{(n)}_{n-k+1}\|^2$ for $1 \leq k \leq n-1$ are independent, we may consider two double arrays $(\xi^k_i, i \geq 1, k \geq 1)$, $(\zeta^k_j, j \geq 1, k \geq 1)$ of independent random variables (and independent together), such that

a) for every $j \geq 1$ and $k \geq 1$, $\xi^k_j \overset{(d)}{=} \gamma(1/2)$,

b) for every $j \geq 1$ and $k \geq 1$, $\zeta^k_j \overset{(d)}{=} \|b^{(j)}_1\|^2$.

The common probability space on which are defined all the variables ξ^k_j and ζ^k_j is denoted by Ω. From now on we work exclusively on Ω.

Let us set

$$S^k_p = \sum_{m=1}^{p} \xi^k_m, \quad k \geq 1, \quad p \geq 1\,.$$

Now, the processes $(S^k_j)_{j\geq 1}$ for $k = 1, \cdots$ are independent copies of $(S^1_j)_{j\geq 1}$, and for each $n \geq 1$, we have the following distributional representation :

$$\{\|\widehat{b}^{(n)}_{n-k+1}\|^2\,, \quad 1 \leq k \leq n-1\} \overset{(d)}{=} \left\{ \frac{S^k_k}{S^k_n}\, \zeta^k_n, \quad 1 \leq k \leq n-1 \right\}. \tag{2.18}$$

For $n \geq 2$, set

$$R^{(n)}_k \quad = \quad \begin{cases} \dfrac{S^k_k S^{k+1}_n}{S^{k+1}_{k+1} S^k_n}\, \dfrac{\zeta^k_n}{\zeta^{k+1}_n} & \text{if } 1 \leq k \leq n-1 \\ 1 & \text{if } k \geq n \end{cases} \tag{2.19}$$

we have now, (see (1.12) and (1.13))

$$r^{(n)} \overset{(d)}{=} R^{(n)}\,. \tag{2.20}$$

The processes $r^{(n)}, n \geq 2$ are not defined on a unique probability space, since the ambient spaces are not nested. On the contrary, the sequence $R^{(n)}, n \geq 2$ is defined on the unique probability space Ω. For each $k \geq 1$, the strong law of large numbers yields

$$\frac{S^{k+1}_n}{n} \xrightarrow[n]{a.s.} \frac{1}{2}\,, \quad \frac{S^k_n}{n} \xrightarrow[n]{a.s.} \frac{1}{2}\,,$$

Besides, Lemma 2.4, with the help of Borel-Cantelli's lemma yields

$$\frac{\zeta^k_n}{\zeta^{k+1}_n} \xrightarrow[n]{a.s.} 1\,,$$

so that if we set

$$\mathcal{R}_k := \frac{S^k_k}{S^{k+1}_{k+1}}\,, \tag{2.21}$$

we get for any $k \geq 1$

$$R^n_k \xrightarrow[n]{a.s.} \mathcal{R}_k.$$

Notice that $\mathcal{R}$ is defined in (1.14). Hence, letting $\mathcal{R}_k$ be $\frac{S_k^k}{S_{k+1}^{k+1}}$ here is a slight abuse of notation but this is consistent in terms of distribution and allows to avoid a new symbols. From now on $(\mathcal{R}_k)$ is then a random variable on Ω. Setting, for any $g \geq 0$,

$$M_n^g = \min_{g+1 \leq k \leq n-1} R_k^{(n)} \quad \text{and} \quad \mathcal{M}^g = \min_{k \geq g+1} \mathcal{R}_k \,, \tag{2.22}$$

we get

$$\mathcal{M}_n^g \quad \stackrel{(d)}{=} \quad M_n^g \,, \tag{2.23}$$

and want to prove a convergence (in probability) of M_n^g to $\mathcal{M}^g$. Since the convergence of the coordinates of $R^{(n)}$ to those of $(\mathcal{R}_k)$ is not sufficient to this aim, we need a uniform control.

Set

$$\widetilde{M}_n^g := \inf_{k \geq g+1} R_k^{(n)},$$

so that $\widetilde{M}_n^g = M_n^g \wedge 1$. This yields $0 \leq M_n^g - \widetilde{M}_n^g = (M_n^g - 1)^+ \leq (R_{n-1}^n - 1)^+$. Since $R_{n-1}^{(n)} \xrightarrow[n]{proba.} 1$ (by Theorem 2.2), we get

$$M_n^g - \widetilde{M}_n^g \xrightarrow[n]{proba.} 0, \tag{2.24}$$

and so, M_n^g and $\widetilde{M}_n^g$ have the same limit behavior.

To prove Theorem 1.8, we first assume that the following lemma which is a strong form of Proposition 1.7 holds true

Lemma 2.8 *For any $p > 2$, $(R_k^{(n)} - \mathcal{R}_k)$ converge a.s. (in Ω) to 0 in ℓ_p, i.e.*

$$\sum_{k=1}^{\infty} |R_k^{(n)} - \mathcal{R}_k|^p \xrightarrow[n]{a.s.} 0. \tag{2.25}$$

Proof of Theorem 1.8

(i) From (2.25) and Lemma 2.7, the sequence $\left(R_k^{(n)} - 1\right)_{k \geq 1}$ converges a.s. in ℓ_p to $(\mathcal{R}_k - 1)_{k \geq 1}$. Let K be a fixed integer. Since the mapping $(c_k)_{k \geq 1} \in \ell_p \longmapsto \inf_{k \geq K} c_k$ is continuous, one has

$$\widetilde{M}_n^K \quad \xrightarrow[n]{a.s.} \quad \mathcal{M}^K. \tag{2.26}$$

Thanks to (2.24), we obtain $M_n^K \xrightarrow[n]{proba.} \mathcal{M}^K$ and then, by (2.23) $\mathcal{M}_n^K \xrightarrow[n]{(d)} \mathcal{M}^K$.

(ii) Let $\epsilon > 0$ and $\epsilon' > 0$ be fixed. Since $(\mathcal{R}_k - 1)_{k \geq 1} \in \ell_p$, there exists K such that

$$\mathbb{P}(\mathcal{M}^K \leq 1 - \epsilon/2) \leq \epsilon'.$$

For n large enough, one then has, by (2.26) and (2.24),

$$\mathbb{P}(M_n^K \leq 1 - \epsilon) \leq 2\epsilon'.$$

Since the function $k \mapsto M_n^k$ is non-decreasing, one has, for n large enough such that $g(n) \geq K$,

$$\mathbb{P}(M_n^{g(n)} \leq 1 - \epsilon) \leq 2\epsilon'.$$

(iii) Take $k = 0$ for the sake of simplicity. For $a \in \mathbb{R}^{\mathbb{N}}$, let $\operatorname{argmin} a = \{i : \inf_{j \geq 1} a_j = a_i\}$ and as usual set $\min \emptyset = \infty$. Denote by $I_n^0 = \min \operatorname{argmin}\{R_j^n, j \geq 1\}$, the natural version of $\mathcal{I}_n^0$ on Ω :

$$I_n^0 \overset{(d)}{=} \mathcal{I}_n^0 \tag{2.27}$$

We know that a.s. $\mathcal{M}^0 < 1$ so that for n large enough, we have $M_n^0 < 1$, hence $\operatorname{argmin} \widetilde{R}^n = \operatorname{argmin} R^n$. Now, from Proposition 2.8(ii), a.s $\lim \widetilde{R}^n = \mathcal{R}$ in ℓ_p. Now, the convergence of y_n to y in ℓ_p implies the convergence of $\min \operatorname{argmin}(y_n)$ to $\operatorname{argmin}(y)$ if $\# \operatorname{argmin}(y) = 1$. Hence, a.s. $\lim I_n^0 = \mathcal{I}^0$. Thanks to (2.27), we deduce $\mathcal{I}_n^0 \xrightarrow[n]{(d)} \mathcal{I}^0$. □

Proof of Lemma 2.8.

Set $V_n := \sum_k |R_k^{(n)} - \mathcal{R}_k|^p = V_n' + V_n''$ where

$$V_n' := \sum_{1 \leq k \leq n-1} |R_k^{(n)} - \mathcal{R}_k|^p \text{ and } V_n'' := \sum_{k \geq n} |1 - \mathcal{R}_k|^p.$$

According to Lemma 2.7, $V_n'' \xrightarrow[n]{a.s.} 0$. Then, it is enough to prove that $V_n' \xrightarrow[n]{a.s.} 0$. Since

$$R_k^{(n)} = \mathcal{R}_k \frac{S_n^{k+1}}{S_n^k} \frac{\zeta_n^k}{\zeta_n^{k+1}}$$

and since $\sup_{k \geq 1} \mathcal{R}_k$ is a.s. finite, it is enough to prove that

$$\sum_{k=1}^{n-1} \left| \frac{S_n^{k+1}}{S_n^k} \frac{\zeta_n^k}{\zeta_n^{k+1}} - 1 \right|^p \xrightarrow[n]{a.s.} 0\,. \tag{2.28}$$

Let $\delta > 0$. By the union bound and the identity of distributions, we have

$$\begin{aligned} \mathbb{P}\left(\sum_{k=1}^{n-1} \left| \frac{S_n^{k+1}}{S_n^k} \frac{\zeta_n^k}{\zeta_n^{k+1}} - 1 \right|^p > \delta \right) &\leq \sum_{k=1}^{n-1} \mathbb{P}\left(\left| \frac{S_n^{k+1}}{S_n^k} \frac{\zeta_n^k}{\zeta_n^{k+1}} - 1 \right|^p > \frac{\delta}{n} \right) \\ &= (n-1)\mathbb{P}\left(\left| \frac{S_n^{(2)}}{S_n^{(1)}} \frac{\zeta_n^1}{\zeta_n^2} - 1 \right| > \frac{\delta^{1/p}}{n^{1/p}} \right). \end{aligned}$$

Splitting this event, we get easily for $\varepsilon = \frac{\delta^{1/p}}{n^{1/p}}$

$$\mathbb{P}\left(\left| \frac{S_n^{(2)}}{S_n^{(1)}} \frac{\zeta_n^k}{\zeta_n^{k+1}} - 1 \right| > \varepsilon \right) \leq \mathbb{P}\left(\left| \frac{S_n^{(2)}}{S_n^{(1)}} - 1 \right| > \varepsilon/3 \right) + \mathbb{P}\left(\left| \frac{\zeta_n^1}{\zeta_n^2} - 1 \right| > \varepsilon/2 \right)$$

With the notation of the preliminaries, the first probability is $\mathbb{P}(|\mathcal{R}_n' - 1| > \varepsilon/3)$. By a simple calculation using (2.12),(2.11) and lemma 2.4, we can find c_1 and $c_2 > 0$ such that for every n

$$\mathbb{P}\left(\sum_{k=1}^{n} \left| \frac{S_n^{k+1}}{S_n^k} \frac{\zeta_n^k}{\zeta_n^{k+1}} - 1 \right|^p > \delta \right) \leq c_1 n \exp\left(-c_2 n^{1-\frac{2}{p}} \right).$$

For $p > 2$, we get a convergent series, so (2.28) holds true, which ends the proof of $ii)$. □

3 Appendix

3.1 LLL(δ)-reduced basis versus Siegel(s)-reduced basis

As mentioned in Section 1.3, the definition 1.3 is slightly different from the original definition of an LLL reduced basis as defined in [10]. Here we make precise this difference and show that our main result (Theorem 1.4) is still true with the original definition.

Let $(b) := b_1^{(n)}, b_2^{(n)}, \ldots, b_p^{(n)}$ (for $p \leq n$) be a linearly independent system of p vectors of $\mathbb{R}^n$ and recall the definition of the matrix R given in Section 1.1.

Definition 3.1 *Let $0 < \delta < 1$ be a real parameter. The basis (b) is called truly–LLL(δ) reduced if it is proper (1.7) and if*

$$\forall i \in \{1 \ldots, n-1\}, \quad \|\widehat{b}_{i+1}\|^2 + R_{i,i+1}^2 \|\widehat{b}_i\|^2 > \delta^2 \|\widehat{b}_i\|^2. \tag{3.29}$$

From the above definition and the definition of a LLL(s)-reduced basis (1.3), and since $R_{i+1,i}^2 \leq 1/4$ (thanks to the properness) one deduces immediately:

Fact 3.2 *(i) If a basis is LLL(s) reduced and proper then it is truly–LLL(s) reduced.*

(ii) If a basis is truly–LLL(δ) reduced then it is LLL($\sqrt{\delta^2 - 1/4}$) reduced.

3.2 How to make a basis proper while preserving its LLL reduceness

Here is a simple enunciation of the LLL(δ) algorithm:

The Make–proper algorithm:

Input: A basis $b = (\mathbf{b}_1, \ldots, \mathbf{b}_p)$ of a lattice L.
Output: A proper basis b of the lattice L.
Initialization: Compute the orthogonalized system $\widehat{b}$ and the matrix R.
For i from 2 to n do
  For j from (i-1) downto 1 do
    $\mathbf{b}_i := \mathbf{b}_i - \lfloor R_{j,i} \rceil \mathbf{b}_j$ ($\lfloor x \rceil$ is the integer nearest to x).

Clearly the Gram-Schmidt basis associated with the input basis is preserved under the integer translations of the above algorithm. So the Gram Schmidt orthogonalized basis associated with the output basis is the same as the one associated with the input basis and the Make–proper algorithm preserves LLL(s)-reduceness and truly–LLL(s)-reduceness.

3.3 A brief description of the LLL algorithm

In this subsection, we provide a simple enunciation of the LLL(δ) algorithm. Clearly if the input basis is LLL(s)-reduced and proper then it is also truly LLL(s)-reduced. So in this case the following algorithm will stop after one iteration of the **while loop** (which makes the basis proper).

The LLL(δ)-reduction algorithm:

Input: A basis $b = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ of a lattice L.
Output: A $LLL(\delta)$-reduced basis b (or a truly LLL(s)-reduced basis) of the lattice L.
Initialization: Compute the orthogonalized system $\widehat{b}$ and the matrix R.
i := **1**;
While i < n do
  $\mathbf{b}_{i+1} := \mathbf{b}_{i+1} - \lfloor R_{i,i+1} \rceil \mathbf{b}_i$ ($\lfloor x \rceil$ is the integer nearest to x).
  Test: $\|\widehat{b}_{i+1}\| > s\|\widehat{b}_i\|$? (or $\|\widehat{b}_{i+1}\|^2 + R_{i,i+1}^2\|\widehat{b}_i\|^2 > \delta\|\widehat{b}_i\|^2$?)
    If true, make $(\mathbf{b}_1, \ldots, \mathbf{b}_{i+1})$ proper by **Make-proper**; **set i** := **i** + **1**;
    If false, swap $\mathbf{b}_i$ and $\mathbf{b}_{i+1}$; update $\widehat{b}$ and R; if $i \neq 1$ then **set i** := **i** − **1**;

3.4 The Beta–Gamma algebra

We recall some properties of the Gamma and Beta distribution, used all along the lines of the paper. They can be found in [4] pp. 93-94. For $a > 0$, the gamma distribution of parameter a is

$$\gamma_a(dx) = \frac{e^{-x}x^{a-1}}{\Gamma(a)}\ \mathbb{1}_{[0,\infty)}(x)\ dx\,,$$

and its mean is a.

For $(a, b) \in \mathbb{R}^{+\star}$, the beta distribution of parameters (a, b) denoted by $\beta_{a,b}$ is

$$\beta_{a,b}(dx) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1}\ \mathbb{1}_{(0,1)}(x)\ dx.$$

In the following, $\gamma(a)$ denotes a variable with distribution γ_a, and $\beta(a, b)$ denotes a variable with distribution $\beta_{a,b}$. The first relation is

$$\big(\gamma(a), \gamma(b)\big) \stackrel{(d)}{=} \big(\beta(a,b)\gamma(a+b), (1-\beta(a,b))\gamma(a+b)\big)\,, \tag{3.1}$$

where, on the left hand side the random variables $\gamma(a)$ and $\gamma(b)$ are independent and on the right hand side the random variables $\beta(a, b)$ and $\gamma(a + b)$ are independent. It entails

$$\gamma(a) + \gamma(b) \stackrel{(d)}{=} \gamma(a+b)\,, \tag{3.2}$$

$$\frac{\gamma(a)}{\gamma(a)+\gamma(b)} \stackrel{(d)}{=} \beta(a,b)\,, \tag{3.3}$$

and

$$\frac{\gamma(a)}{\gamma(b)} \stackrel{(d)}{=} \frac{\beta(a,b)}{1-\beta(a,b)}\,, \tag{3.4}$$

which gives

$$\mathbb{P}\big(\gamma(a)/\gamma(b) \in dx\big) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\frac{x^{a-1}}{(1+x)^{a+b}}\ \mathbb{1}_{[0,\infty[}(x)\ dx\,. \tag{3.5}$$

The second relation is

$$\beta(a,b)\beta(c,a-c) \stackrel{(d)}{=} \beta(c,a+b-c)\,, \tag{3.6}$$

where on the left hand side the random variables are independent.

References

[1] A. Akhavi. *Analyse comparative d'algorithmes de réduction sur les réseaux aléatoires*. PhD thesis, Université de Caen, 1999.

[2] A. Akhavi. Random lattices, threshold phenomena and efficient reduction algorithms. *Theoretical Computer Science*, 287:359–385, 2002.

[3] N.R. Chaganthy. Large deviations for joint distributions and statistical applications. *Sankhya*, 59:147–166, 1997.

[4] L. Chaumont and M. Yor. *Exercises in probability*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2003. A guided tour from measure theory to random processes, via conditioning.

[5] H. Daudé and B. Vallée. An upper bound on the average number of iterations of the LLL algorithm. *Theor. Comput. Sci.*, 123(1):95–115, 1994.

[6] J.L. Donaldson. Minkowski reduction of integral matrices. *Mathematics of Computation*, 33(145):201–216, 1979.

[7] Jr. H.W. Lenstra. Flags and lattice basis reduction. In *European Congress of Mathematics, Vol. I (Barcelona, 2000)*, volume 201 of *Progr. Math.*, pages 37–51. Birkhäuser, Basel, 2001.

[8] R. Kannan. Algorithmic geometry of numbers. In *Annual review of computer science, Vol. 2*, pages 231–267. Annual Reviews, Palo Alto, CA, 1987.

[9] D. E. Knuth. *The art of computer programming. Vol. 2*. Addison-Wesley Publishing Co., Reading, Mass., second edition, 1981. Seminumerical algorithms, Addison-Wesley Series in Computer Science and Information Processing.

[10] A. K. Lenstra, H. W. Lenstra, Jr., and L. Lovász. Factoring polynomials with rational coefficients. *Math. Ann.*, 261(4):515–534, 1982.

[11] H. W. Lenstra, Jr. Integer programming and cryptography. *Math. Intelligencer*, 6(3):14–19, 1984.

[12] G. Letac. Isotropy and sphericity: some characterisations of the normal distribution. *Ann. Statist.*, 9(2):408–417, 1981.

[13] R. J. Muirhead. *Aspects of multivariate statistical theory*. John Wiley, 1982.

[14] P. Q. Nguyen and J. Stern. The two faces of lattices in cryptology. In *Cryptography and lattices (Providence, RI, 2001)*, volume 2146 of *Lecture Notes in Comput. Sci.*, pages 146–180. Springer, 2001.

[15] C.P. Schnorr. Fast LLL-Type Lattice Reduction. *Information and Computation*, 204:1–25, 2006.

[16] B. Vallée. Un problème central en géometrie algorithmique des nombres: la réduction des réseaux. Autour de l'algorithme de Lenstra Lenstra Lovasz. In *Informatique Théorique et Applications*, volume 3, pages 345–376. 1989. English translation by E. Kranakis CWI-Quaterly - 1990 - 3.