Numer. Math. (2003) 94: 419–478
Digital Object Identifier (DOI) 10.1007/s00211-002-0385-1

Numerische Mathematik

On the Expected Condition Number of Linear Programming Problems

Felipe Cucker[1,⋆], **Mario Wschebor**[2]

[1] Department of Mathematics, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon, HONG KONG; e-mail: macucker@math.cityu.edu.hk
[2] Centro de Matemática. Facultad de Ciencias. Universidad de la República, Calle Iguá 4225. 11400 Montevideo, URUGUAY; e-mail: wscheb@fcien.edu.uy

Received June 12, 2001 / Revised version received October 29, 2001 /
Published online November 27, 2002 –

Summary. Let A be an $n \times m$ real matrix and consider the linear conic system

$$Ax \leq 0, x \neq 0.$$

In [Cheung and Cucker 2001] a condition number $\mathcal{C}(A)$ for this system is defined. In this paper we let the coefficients of A be independent identically distributed random variables with standard Gaussian distribution and we estimate the moments of the random variable $\ln \mathcal{C}(A)$. In particular, when n is sufficiently larger than m we obtain for its expected value $\mathbf{E}(\ln \mathcal{C}(A)) = \max\{\ln m, \ln \ln n\} + \mathcal{O}(1)$. Bounds for the expected value of the condition number introduced by Renegar [1994b, 1995a, 1995b] follow.

Mathematics Subject Classification (1991): 65F35, 65K05

1 Introduction

1.1 Condition of linear systems

The dawn of digital computers brought the possibility of mechanically solving a plethora of mathematical problems. It also rekindled the interest for round-off analysis. The issue here is that, within the standard floating point arithmetic, all computations are carried in a subset $\mathbb{F} \subset \mathbb{R}$ instead of on the whole set of real numbers $\mathbb{R}$. A characteristic property of floating point arithmetic is the existence of a number $0 < u < 1$, the *round-off unit*,

⋆ Partially supported by CERG grant City U 1085/02p.

and a function $r : \mathbb{R} \to \mathbb{F}$, *the rounding function*, such that, for all $x \in \mathbb{R}$, $|r(x) - x| \leq u|x|$. Arithmetic operations in $\mathbb{R}$ are then replaced by "rounded" versions in $\mathbb{F}$. The result of, for instance, multiplying $x, y \in \mathbb{F}$ is $r(xy) \in \mathbb{F}$.

During a computation these errors accumulate and the final result may be far away from what it should be. A prime concern when designing algorithms is thus to minimize the effects of this accumulation. Algorithms are consequently analyzed with this regard and compared between them in the same way they are compared regarding their running times. This practice, which today is done more or less systematically, was already present in Gauss' work.

> Since none of the numbers we take out from logarithmic or trigonometric tables admit of absolute precision, but are all to a certain extent approximate only, the results of all calculations performed by the aid of these numbers can only be approximately true. [. . .] It may happen, that in special cases the effect of the errors of the tables is so augmented that we may be obliged to reject a method, otherwise the best, and substitute another in its place.
>
> Carl Friedrich Gauss, *Theoria Motus* (cited in [Goldstine 1977] p. 258).

To study how errors accumulate during the execution of an algorithm it is convenient to first focus on a simplified situation namely, that in which errors occur only when reading the input. That is, an input $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$ is rounded to $r(a) = (r(a_1), \ldots, r(a_n)) \in \mathbb{F}^n$ and then the algorithm is executed with infinite precision over the input $r(a)$.

Let's see how this is done for the problem of linear equation solving. Let A be an invertible $n \times n$ real matrix and $b \in \mathbb{R}^n$. We are interested in solving the system

$$Ax = b$$

and want to study how the solution x is affected by perturbations in the input (A, b).

Early work by Turing [1948] and von Neumann and Goldstine [1947] identified that the key quantity was

$$\kappa(A) = \|A\| \|A^{-1}\|$$

where $\|A\|$ denotes the operator norm of A defined by

$$\|A\| = \max_{\|x\|=1} \|A(x)\|.$$

Here $\| \ \|$ denotes the Euclidean norm in $\mathbb{R}^n$ both as a domain and codomain of A. Turing called $\kappa(A)$ the *condition number* of A. A main result for $\kappa(A)$

states that, if $\kappa(A)\frac{\|\Delta A\|}{\|A\|} < 1$ then

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{\kappa(A)}{1-\kappa(A)\frac{\|\Delta A\|}{\|A\|}} \left(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta b\|}{\|b\|} \right).$$

Notice that the factor $\frac{\kappa(A)}{1-\kappa(A)\frac{\|\Delta A\|}{\|A\|}}$ tends to $\kappa(A)$ when $\|\Delta A\| \to 0$. In addition, $\kappa(A)$ is sharp in the sense that no smaller number will satisfy the inequality above for all A and b. Thus, $\kappa(A)$ measures how much the relative input error is amplified in the solution and $\log \kappa(A)$ measures the loss of precision. In Turing's words

> It is characteristic of ill-conditioned sets of equations that small percentage errors in the coefficients given may lead to large percentage errors in the solution.

When A is not invertible its condition number is not well defined. However, we can extend its definition by setting $\kappa(A) = \infty$ if A is singular. Matrices A with $\kappa(A)$ small are said to be *well-conditioned*, those with $\kappa(A)$ large are said to be *ill-conditioned*, and those with $\kappa(A) = \infty$ *ill-posed.*

Note that the set Σ of ill-posed matrices has Lebesgue measure zero in the space $\mathbb{R}^{n^2}$. The distance of a matrix A to this set is closely related to $\kappa(A)$.

Theorem 1 (Condition Number Theorem, [Eckart and Young 1936]) *For any $n \times n$ real matrix A one has*

$$\kappa(A) = \frac{\|A\|}{d_F(A, \Sigma)}$$

Here d_F means distance in $\mathbb{R}^{n^2}$ with respect ot the Frobenius norm $\|A\| = \sqrt{\sum a_{ij}^2}$.

The relationship between conditioning and distance to ill-posedness is a recurrent theme in numerical analysis (cf. [Demmel 1987]). It will play a central role in our understanding of the condition of a linear program.

Reasonably enough, $\kappa(A)$ will appear in more elaborate round-off analysis in which errors may occur in all the operations. As an example, we mention such an analysis for Cholesky's method. If A is symmetric and positive definite we may solve the linear system $Ax = b$ by using Cholesky's factorisation. If the computed solution is $(x + \Delta x)$ then one can prove that, for u sufficiently small,

$$\frac{\|\Delta x\|}{\|x\|} \leq 3n^3 u \kappa(A).$$

Remark 1 Note that a bound as the one above for the relative forward error of an input $a \in \mathbb{R}^n$ in the form of an expression in its size n, its condition number $\kappa(a)$, and the round-off unit u may not be computable for a particular input a since we may not know $\kappa(a)$. Yet, such bounds allow us to compare algorithms with respect to stability. The fastest the expression tends to zero with u, the more stable the algorithm is.

Numerical linear algebra is probably one of the best developped areas in numerical analysis. The interested reader can find more about it in, for instance, the introductory books [Demmel 1997; Trefethen and Bau III 1997] or in the more advanced [Higham 1996].

1.2 Condition of linear conic systems

To extend the ideas above to linear programming is not immediate. Let $A \in \mathbb{R}^{n\times m}$ be given and consider the two systems[1]

(1) $$Ax \le 0,\ \ x \ne 0$$

and

(2) $$A^{\mathrm{T}} y = 0,\ \ y \ge 0,\ \ y \ne 0.$$

It is well-known that one of these systems has a strict solution (one for which the satisfied inequality is strict in all coordinates) if and only if the other has no solutions at all. This is a generic property. Indeed, except for a set of Lebesgue measure zero, for any $A \in \mathbb{R}^{m\times n}$, one of the systems in the pair (1)–(2) has a strict solution, and this continues to hold even if the matrix A is slightly perturbed. A standard problem in linear programming is the following

> Given a $n \times m$ real matrix A, decide which of (1) or (2) is strictly feasible and return a strict solution for it.

Here the concept of "input error" still makes sense but that of "output error" becomes less clear. The answer of the "decision part" of the problem (finding which of the two systems is strictly feasible) will either be severely affected by input errors (the algorithm will return the wrong system) or will not be affected at all. No "small output error" is possible. Also, for the strictly feasible system there is not a unique strict solution and the algorithm may return any of them. So, it is less clear how to define a condition number.

[1] Usually, in the literature, one considers a $m \times n$ matrix A appearing in (2), the "primal system", and its transpose A^{T} appears in (1), the "dual system." We revert this notation here since in most of this paper we will deal with system (1) and we do not want to burden the notation with the transpose superscript.

A natural alternative to define condition numbers in the context above (and for other forms of conic systems as well) was proposed by Renegar [1994b, 1995a, 1995b]. The idea is to use the condition number theorem. More precisely, let $\mathcal{D}$ (resp. $\mathcal{P}$) denote the set of matrices A for which (1) (resp. (2)) is feasible and let Σ be the boundary of $\mathcal{D}$ and $\mathcal{P}$. Renegar defined

$$C_R(A) = \frac{\|A\|}{d(A, \Sigma)}$$

where both numerator and denominator are for the operator norm with respect to the Euclidean norm in both $\mathbb{R}^m$ and $\mathbb{R}^n$. It turns out that the condition number thus defined naturally appears in a variety of bounds related with the problem above —besides, of course, those in error analysis. For instance, it appears when studying the relative distance of solutions to the boundary of the solution set, the number of necessary iterations to solve the problem (for iterative algorithms such as interior-point or ellipsoid), or —as expected— in different error estimates in the round-off analysis of these algorithms. All these quantities may be bounded by expressions in which $C_R(A)$ appears as a parameter. References for the above are [Renegar 1995b; Freund and Vera 1999b; Freund and Vera 1999a; Vera 1998; Cucker and Peña 2001] (see also Section 4 for some of it).

If we are only interested in estimating complexity bounds for iterative algorithms with infinite precision, it is possible to consider some measures of condition which are always finite. This has been done for instance in [Ye 1994] where the condition measure $\sigma(A)$ for a matrix A with $n \geq m$ is defined. Then (cf. [Vavasis and Ye 1995]) this measure was used to show that a feasibility problem similar to the one above can be solved by an interior-point method with $\mathcal{O}(\sqrt{n}|\ln \sigma(A)| + \ln n)$ iterations.

Similarly, always under the assumption $n \geq m$, an algorithm was given in [Vavasis and Ye 1996] which solves optimization linear programs within $\mathcal{O}(n^{3.5}(\ln \overline{\chi}_A + \ln n))$ iterations each of them performing $\mathcal{O}(m^2 n)$ arithmetic operations. Here $\overline{\chi}_A$ is the condition measure introduced in [Stewart 1989] and [Todd 1990]. Both $|\ln \sigma(A)|$ and $\ln \overline{\chi}_A$ are finite for all matrices A.

Other recent measures of condition for linear programming, $\mathcal{C}(A)$ and $\mu(A)$, are defined in [Cheung and Cucker 2001] and [Freund and Epelman 2000]. The first one, a very close relative of $C_R(A)$, enjoys all the good properties of $C_R(A)$ and has some additional ones including some nice geometric interpretation. It is the central character in this paper. We give now its definition and delay exposing some of its main properties to Section 2 so that we can quickly state our main result. Let a_k denote the kth row of A, $k = 1, \ldots, n$, and $x \in S^{m-1}$, the unit sphere in $\mathbb{R}^m$. Define

$$f_k(x) = \frac{\langle a_k, x \rangle}{\|a_k\|}$$

(here, and in the rest of this paper, $\langle\ ,\ \rangle$ denotes the standard inner product in $\mathbb{R}^m$ and $\|\ \|$ its induced norm) and $D = \min_{x\in S^{m-1}} \max_{1\le k\le n} f_k(x)$. We define $\mathcal{C}(A)$ to be

$$\mathcal{C}(A) = \frac{1}{|D|}.$$

1.3 Averaging condition measures

Condition numbers have two probably disturbing features. One is that, in general, we don't have an a priori knowledge of the condition of an input. Unlike the size of a, which is straightforwardly obtained from a, the condition number of a seems to require a computation which is not easier than solving the problem for which a is an instance (see [Renegar 1994a] for a discussion on this). The other is that there are no bounds on their magnitude as a function of the input size. They may actually be infinite.

A reasonable way to cope with these features is to assume a probability measure on the space of inputs and to estimate the expected value (and if possible, other moments) of the condition number. The use of such idea in complexity analysis is pushed forward, for instance, in [Smale 1997]. Examples of this kind of result are the following. We say that a random matrix is *Gaussian* when its entries (real and imaginary parts of the entries for the complex case) are i.i.d. $N(0,1)$ random variables defined on a probability space $(\Omega, \mathcal{A}, P)$.

Theorem 2 ([Edelman 1988]) *Let A be a $n \times n$ Gaussian matrix. Then the expected value of* $\log(\kappa(A))$ *satisfies*

$$\mathbf{E}\,(\log(\kappa(A))) = \log\ n + c + o(1) \quad \textit{when} \ \ n \to \infty$$

where $c \approx 1.537$ for real matrices and $c \approx 0.982$ for complex matrices. □

Theorem 3 ([Todd, Tunçel, and Ye 2001]) *Let $n > m > 3$ and A be a Gaussian $n \times m$ matrix. Then*

$$\mathbf{E}(\ln \overline{\chi}_A),\ \mathbf{E}(\ln \sigma(A)) = \mathcal{O}(\min\{m \ln n, n\}).$$

□

Theorem 4 ([Cheung and Cucker 2002]) *Let A be a Gaussian $n \times m$ matrix. Then*

(i)

$$\mathbf{E}\,(\ln \mathcal{C}(A)) = \begin{cases} \mathcal{O}(\min\{m \ln n, n\}) & \textit{if } n > m \\ \mathcal{O}(\ln n) & \textit{otherwise.} \end{cases}$$

(ii)
$$\mathbf{E}(\ln C_R(A)) \leq \mathbf{E}(\ln \mathcal{C}(A)) + \frac{5\ln n}{2} + \frac{\ln m}{2} + 2\ln 2.$$
□

Remark 2 The restriction to the case $n > m$ in Theorem 3 (and in all the work related with $\overline{\chi}_A$ and $\sigma(A)$) is not severe. In the case $n \leq m$ system (1) is always feasible. The decision part of the problem at the beginning of the previous section is thus, in this case, empty of content. In addition in most of the occurrences in practice of that problem one actually has that n is some orders of magnitude larger than m. The case $n \gg m$ (i.e. n much larger than m) is actually the case of interest among researchers in linear programming.

1.4 Main result in this paper

The goal of this paper is to improve the bounds in Theorem 4 for the case $n \gg m$. We next state our main result. Let Φ denote the standard $N(0, 1)$ distribution,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$

and φ its density

$$\varphi(t) = \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}.$$

Let

$$D(n, m) = \frac{1}{6qm} \left(\frac{e}{m}\right)^{\frac{m}{2}} (\Phi(1))^{n-2}(1 + (m \wedge n)^{\frac{m}{2}}) \left(K_0 n^2 m^{3/2} q^2\right)^{(m \wedge n)}$$

where $m \wedge n$ denotes the minimum of m and n,

$$K_0 = \frac{e^3\sqrt{2}}{(e-1)\sqrt{\pi}\Phi(1)} \approx 11.086 \quad \text{and} \quad q = \max\left\{3.2, \sqrt{\frac{3\ln(2n)}{m}}\right\}.$$

Note that $D(n, m)$ becomes small when either n is sufficiently larger than m or m is sufficiently larger than n, and the larger one of this parameters becomes with respect to the other the smaller is the value of $D(n, m)$. For instance, if either $\frac{m(1+\ln n)}{n} \leq 1$ or $\frac{n(1+\ln m)}{m} \leq \frac{1}{14}$, we have $D(n, m) \leq 1$. It follows that if m or n (or both) tend to ∞ in such a way that either $\frac{m \ln n}{n}$ or $\frac{n \ln m}{m}$ tend to 0 then $D(n, m) \leq 1$ holds true for sufficiently large values of m or n. Our interpretation of $n \gg m$ will thus be $n > m$ and $D(n, m) \leq 1$. The relation $m \gg n$ is defined similarly.

Main Theorem *Let A be an $n \times m$ Gaussian matrix. Then*
(i)

$$\mathbf{E}(\ln \mathcal{C}(A)) \leq \begin{cases} \ln m + 2\ln q + 2D(n,m) & \textit{if } D(n,m) \leq 1 \\ \ln m + 2\ln q + 2\ln D(n,m) + 2 & \textit{if } D(n,m) > 1. \end{cases}$$

(ii) *For $\nu = 2, 3, \ldots$*

$$\mathbf{E}([\ln \mathcal{C}(A)]^\nu) \leq \begin{cases} 2\nu! \sum_{k=0}^{\nu} \frac{1}{k!} \left(\frac{1}{2}\ln m + \ln q\right)^k & \textit{if } D(n,m) \leq 1 \\ 2\nu! \sum_{k=0}^{\nu} \frac{1}{k!} \left(\ln D(n,m) + \frac{1}{2}\ln m + \ln q\right)^k & \textit{if } D(n,m) > 1. \end{cases}$$

Remark 3 Note that when $m \gg n$ then $q = 3.2$ and we get $\mathbf{E}(\ln \mathcal{C}(A)) = \ln m + \mathcal{O}(1)$. More interestingly, when $n \gg m$, we get $\mathbf{E}(\ln \mathcal{C}(A)) = \max\{\ln m, \ln\ln n\} + \mathcal{O}(1)$.

Corollary 1 *Let $m, n \geq 1$ be given. Then, for each $\lambda \in \mathbb{R}$, $0 < \lambda < 1$, and any $y > 0$*
(i) *If $D(n,m) \leq 1$,*

$$\mathbf{P}\left(\ln \mathcal{C}(A) > y\left(\frac{1}{2}\ln m + \ln q\right)\right) \leq \left(1 + 2e\frac{\lambda}{\lambda - 1}\right)e^{-\lambda y}.$$

(ii) *If $D(n,m) > 1$,*

$$\mathbf{P}\left(\ln \mathcal{C}(A) > y\left(\ln D(n,m) + \frac{1}{2}\ln m + \ln q\right)\right) \leq \left(1 + 2e\frac{\lambda}{\lambda - 1}\right)e^{-\lambda y}.$$

Proof. We only prove (i). Part (ii) is done similarly.

From the Main Theorem

$$\mathbf{E}\{(\ln \mathcal{C}(A))^\nu\} \leq 2e\nu!\alpha_0^\nu$$

with $\alpha_0 = \frac{1}{2}\ln m + \ln q$. It follows that, for $0 < \lambda < 1$,

$$\mathbf{E}\left\{e^{\lambda \frac{\ln \mathcal{C}(A)}{\alpha_0}}\right\} = 1 + \sum_{\nu=1}^{\infty} \frac{\lambda^\nu}{\nu!}\mathbf{E}\left\{\left(\frac{\ln \mathcal{C}(A)}{\alpha_0}\right)^\nu\right\} \leq 1 + 2e\sum_{\nu=1}^{\infty}\lambda^\nu = 1 + 2e\frac{\lambda}{\lambda - 1}$$

and using Markov's inequality, if $y > 0$,

$$\mathbf{P}(\ln \mathcal{C}(A) > y\alpha_0) = \mathbf{P}\left(e^{\lambda \frac{\ln \mathcal{C}(A)}{\alpha_0}} > e^{\lambda y}\right) \leq e^{-\lambda y}\left(1 + 2e\frac{\lambda}{\lambda - 1}\right). \quad \square$$

In the next section we review some of the basic properties of $\mathcal{C}(A)$. In Section 3 we prove, modulo a few technical results, the Main Theorem. The proofs of these technical results are given in Section 5. In doing so, we intended to isolate the central ideas behind the proof of the Main Theorem from the less conceptual, yet necessary, results allowing these central ideas to be applied. Section 4 is devoted to point out some consequences of the Main Theorem.

2 Main properties of $\mathcal{C}(A)$

Let A be a real $n \times m$ matrix. For $1 \le k \le n$ denote by a_k the kth row of A. For any vector $x \in \mathbb{R}^m$, $x \neq 0$, let $\theta_k(A, x) \in [0, \pi]$ be the angle between x and a_k, i.e. $\theta_k(A, x) = \arccos \frac{\langle x, a_k \rangle}{\|x\| \|a_k\|} = \arccos f_k\left(\frac{x}{\|x\|}\right)$ and

$$\theta(A, x) = \min_{k \le n} \theta_k(A, x).$$

Denote by $\overline{x}$ any vector satisfying

$$\theta(A) = \theta(A, \overline{x}) = \max_{x \in \mathbb{R}^m} \theta(A, x).$$

Then

$$\begin{aligned}\cos \theta(A) &= \cos\left(\max_{x \in \mathbb{R}^m} \min_{k \le n} \theta_k(A, x)\right)\\ &= \cos\left(\max_{\substack{x \in \mathbb{R}^m\\ x \neq 0}} \min_{k \le n} \arccos f_k\left(\frac{x}{\|x\|}\right)\right)\\ &= D\end{aligned}$$

and we conclude that

$$\mathcal{C}(A) = \frac{1}{|\cos(\theta(A))|}.$$

The number $\mathcal{C}(A)$ captures several features related with the feasibility of the system $Ax \le 0$. We next briefly state them. Proofs of these results can be found in [Cheung and Cucker 2001].

Let $\mathrm{Sol}(A) = \{x \mid Ax \le 0, x \neq 0\}$ and $\mathcal{D} = \{A \in \mathbb{R}^{n \times m} \mid \mathrm{Sol}(A) \neq \emptyset\}$.

Lemma 1 *Let $x \in \mathbb{R}^m$ and $\overline{x}$ as above. Then,*

(i) $\langle a_k, x \rangle \le 0 \iff \cos \theta_k(A, x) \le 0 \iff \theta_k(A, x) \ge \frac{\pi}{2}$,
(ii) $x \in \mathrm{Sol}(A) \iff \theta(A, x) \ge \frac{\pi}{2} \iff \cos \theta(A, x) \le 0$ *and*
(iii) $A \in \mathcal{D} \iff \overline{x} \in \mathrm{Sol}(A)$.

□

A version of the Condition Number Theorem holds for $\mathcal{C}(A)$. Let

$$\varrho(A) = \sup\left\{\Delta \;\middle|\; \max_{k \le n} \frac{\|a_k' - a_k\|}{\|a_k\|} < \Delta \Rightarrow (A \in \mathcal{D} \iff A' \in \mathcal{D})\right\}$$

where A' denotes the matrix with a_k' as its kth row.

Theorem 5 $\mathcal{C}(A) = \frac{1}{\varrho(A)}$. □

We already mentioned that $\mathcal{C}$ is a close relative to Renegar's condition number C_R. Actually, one can prove that, for every matrix A, $\mathcal{C}(A) \leq \sqrt{n}C_R(A)$. Moreoever, there is no converse of this in the sense that there is no function $f(m,n)$ such that $C_R(A) \leq f(m,n)\mathcal{C}(A)$ for all $n \times m$ matrices A. This shows that $C_R(A)$ can be arbitrarily larger than $\mathcal{C}(A)$. However, the following relation holds.

Proposition 1

$$C_R(A) \leq \frac{\|A\|}{\min_k \| a_k \|}\mathcal{C}(A).$$

□

Therefore, restricted to the set of matrices A such that $\|a_k\| = 1$ for $k = 1, \ldots, n$, one has

$$\mathcal{C}(A) \leq \sqrt{n}C_R(A) \leq n\mathcal{C}(A).$$

We now note that if A is arbitrary and $\overline{A}$ is the matrix whose kth row is

$$\overline{a}_k = \frac{a_k}{\|a_k\|}$$

then $\mathcal{C}(\overline{A}) = \mathcal{C}(A)$ since $\mathcal{C}$ is invariant by row-scaling and $C_R(\overline{A}) \leq nC_R(A)$. Thus, $\mathcal{C}(A)$ is closely related to $C_R(\overline{A})$ for a normalization $\overline{A}$ of A which is easy to compute and does not increase too much $C_R(A)$.

The last feature of $\mathcal{C}(A)$ we mention in this section relates the probabilistic behaviour of $\mathcal{C}(A)$ (for random matrices A) with a classical problem in geometric probability. Before that, we do a remark on distributions for random matrices. Let $a_k \in \mathbb{R}^m$, $k = 1, \ldots, n$, denote the rows of a matrix A. The assumption that A is Gaussian is equivalent to say that the random variables $\left\{\frac{a_1}{\|a_1\|}, \ldots, \frac{a_n}{\|a_n\|}, \|a_1\|^2, \ldots, \|a_n\|^2\right\}$ are independent, the first n being uniformly distributed on the sphere S^{m-1} and the last n having the common distribution χ^2 with m degrees of freedom (i.e. $\|a_1\|^2$ can be written as $\|a_1\|^2 = \sum_{j=1}^m \zeta_j^2$ where $\zeta_1 \ldots \zeta_m$ are i.i.d. random variables in the real line with standard normal distribution). Thus, since $\mathcal{C}(A)$ is invariant under row scaling of A, its moments (or those of $\ln(\mathcal{C}(A))$) will be the same for both the Gaussian assumption on A or the assumption that $a_1, \ldots, a_n$ are independently drawn from S^{m-1} with the uniform distribution.

Assume that the rows of A have all norm 1. Then it is easy to prove that

$$\theta(A) = \inf\{\theta : \text{the union of the circular caps with centers } a_k \text{ and angular radius } \theta \text{ covers } S^{m-1}\}$$

Thus, if the rows a_k of A are randomly drawn from S^{m-1}, independently and with a uniform distribution, the random variable $\theta(A)$ (and a fortiori $\mathcal{C}(A)$) is related to the problem of covering the sphere with random circular caps. The latter is a classical problem in geometric probability (cf. [Hall 1988; Solomon 1978]). The aspect most studied of this problem is to estimate, for given θ and n, the probability that n circular caps of angular radius θ cover S^{m-1}. A full solution of this problem is, as today, unknown. Partial results and some asymptotics can be found in [Gilbert 1966; Miles 1969; Janson 1986]. In our case, for the estimation of the moments of the random variable $\ln \mathcal{C}(A)$ the main point is to understand the behaviour of the distribution of the random varible

$$\frac{1}{|\cos(\theta(A))|}$$

near the zero values of the denominator, that is, for the set of matrices A such that $\theta(A)$ is near $\frac{\pi}{2}$. The above mentioned results do not seem to be helpful for this purpose.

On the other hand, the covering problem can be explicitly solved for the special value $\theta = \frac{\pi}{2}$ in which case (cf. Theorem 1.5 in [Hall 1988]), the probability that n circular caps cover S^{m-1} is equal to

$$1 - \frac{1}{2^{n-1}} \sum_{k=0}^{m-1} \binom{n-1}{k}.$$

This has some immediate consequences in our context. In the following, by "$Ax \leq 0$ is feasible" we mean that $Ax \leq 0$ has non-zero solutions.

Proposition 2 *Let A be a random matrix whose n rows are randomly drawn in S^{m-1} independently and uniformly distributed. Then,*

$$\mathbf{P}(Ax \leq 0 \textit{ is feasible}) = \frac{1}{2^{n-1}} \sum_{k=0}^{m-1} \binom{n-1}{k}.$$

Consequently, $\mathbf{P}(Ax \leq 0$ is feasible$) = 1$ if $n \leq m$, $\mathbf{P}(Ax \leq 0$ is feasible$) \to 0$ if m is fixed and $n \to \infty$ and $\mathbf{P}(Ax \leq 0$ is feasible$) = \frac{1}{2}$ when $n = 2m$.

Proof. From (ii) and (iii) of Lemma 1 it follows that $A \in \mathcal{D}$ if and only if $\theta(A) \geq \frac{\pi}{2}$. And the latter is equivalent to say that the n circular caps with centers a_k and angular radius $\frac{\pi}{2}$ do not cover S^{m-1}. Thus, the first statement follows. The rest of the proposition is trivial. □

3 Proof of the Main Theorem

Let a_k, $k = 1, \dots, n$, be i.i.d. random variables with standard $(N(0, I_m))$ normal law. We want to give bounds (depending on n, m) for

$$\mathbf{E}\left\{\left(\ln \frac{1}{|D|}\right)^{\nu}\right\} \quad (\nu = 1, 2, \dots).$$

Note that $|D| \le 1$, so that $\ln \frac{1}{|D|} \ge 0$.

3.1 Replacing $\|a_k\|$ by $\sqrt{m}$ for $k = 1, \dots, n$

As we noted in Section 2 to compute $\mathbf{E}\left\{\left(\ln \frac{1}{|D|}\right)^{\nu}\right\}$ we may assume the a_k either uniformly distributed on S^{m-1} or standard normally distributed in $\mathbb{R}^m$. The first choice has the drawback of introducing dependencies among the coordinates of the a_k's. So, we will take the second one.

In addition, our computations will become less complicated if we replace $\|a_k\|$ by $\sqrt{m}$ for $k = 1, \dots, n$. To do so we now introduce a few objects which will be present during the whole development.

Let $q > 1$ and

$$E_q = \left\{\omega \in \Omega : \frac{\|a_k\|}{\sqrt{m}} \le q \text{ for } k = 1, \dots, n\right\}.$$

Also, let

$$\begin{aligned} \widetilde{f}(x) &= \max_{1\le k\le n} f_k(x), \\ Z(x) &= \max_{1\le k\le n} \langle x, a_k\rangle, \qquad \text{and} \\ \underline{Z} &= \min_{x\in S^{m-1}} Z(x). \end{aligned}$$

Lemma 2 *If $\omega \in E_q$ then*

$$\frac{1}{\sqrt{m}}|\underline{Z}| \le q|D|.$$

Proof. If $\omega \in E_q$ one has, for all $x \in S^{m-1}$ and all $k = 1, \dots, n$,

$$\langle x, a_k\rangle \ge 0 \Rightarrow \frac{\langle x, a_k\rangle}{\sqrt{m}} \le \frac{\langle x, a_k\rangle}{\|a_k\|} q$$

and

$$\langle x, a_k\rangle < 0 \Rightarrow q\frac{\langle x, a_k\rangle}{\|a_k\|} \le \frac{\langle x, a_k\rangle}{\sqrt{m}}.$$

Taking maxima over $1 \le k \le n$ it follows that

$$Z(x) \ge 0 \Rightarrow \frac{Z(x)}{\sqrt{m}} \le q\widetilde{f}(x)$$

and

$$Z(x) < 0 \Rightarrow q\widetilde{f}(x) \le \frac{Z(x)}{\sqrt{m}}.$$

Since $Z(x)$ and $\widetilde{f}(x)$ have the same sign for all $\omega \in E_q$, now taking minima over $x \in S^{m-1}$,

$$\underline{Z} \ge 0 \Rightarrow \frac{1}{\sqrt{m}}\underline{Z} \le qD$$

and

$$\underline{Z} < 0 \Rightarrow qD \le \frac{1}{\sqrt{m}}\underline{Z}$$

from which the conclusion follows. □

Lemma 3 *Let E_q^{c} denote the complement of the event E_q in the probability space. If*

$$q \ge \max\left\{3.2, \sqrt{\frac{3\ln(2n)}{m}}\right\}$$

then

$$\mathbf{P}(E_q^{\mathsf{c}}) \le \frac{1}{2}.$$

Proof. One has

$$\mathbf{P}(E_q^{\mathsf{c}}) \le n\mathbf{P}\left(\frac{\|a_1\|^2}{m} > q^2\right) = n\mathbf{P}\left(\frac{X_1 + \cdots + X_m}{m} > q^2 - 1\right)$$

where $X_1, \ldots, X_m$ are i.i.d. random variables with the distribution of $\xi^2 - 1$, ξ a normal standard random variable.

The logarithmic moment generating function of $\xi^2 - 1$ is

$$\Lambda(\lambda) = \ln \mathbf{E}\{e^{\lambda(\xi^2-1)}\} = \begin{cases} -\lambda - \frac{1}{2}\ln(1-2\lambda) & \text{if } \lambda < \frac{1}{2} \\ +\infty & \text{if } \lambda \ge \frac{1}{2} \end{cases}$$

and its Fenchel-Legendre transform

$$\Lambda^*(x) = \sup_{\lambda \in \mathbb{R}}(\lambda x - \Lambda(\lambda)) = \begin{cases} \frac{1}{2}(x - \ln(x+1)) & \text{if } x > -1 \\ +\infty & \text{if } x \le -1. \end{cases}$$

A basic result on large deviations (cf. Chapter 2 of [Dembo and Zeitouni 1998]) states that, for any $x > 0$,

$$\mathbf{P}\left(\frac{X_1 + \cdots + X_m}{m} \geq x\right) \leq e^{-m\Lambda^*(x)}.$$

Therefore, in our case,

$$\mathbf{P}(E_q^{\mathsf{c}}) \leq n\left(e^{-m\Lambda^*(q^2-1)}\right).$$

Using the form of Λ^* shown above, and the fact that if $q \geq 3.2$ then $q^2 - 1 - 2\ln q \geq \frac{2}{3}q^2$, an elementary computation shows that the right-hand side of this inequality is at most $1/2$ if q satisfies the hypothesis. □

Proposition 3 *Let* $\alpha \geq 0$ *and* $q \geq \max\left\{3.2, \sqrt{\frac{3\ln(2n)}{m}}\right\}$. *Then:*

$$\mathbf{E}\left\{\left(\ln\frac{1}{|D|}\right)^{\nu}\right\} \leq 2\left[\alpha + \int_{\alpha}^{+\infty} \mathbf{P}\left(\left\{|\underline{Z}| < \sqrt{m}qe^{-x^{1/\nu}}\right\} \cap E_q\right) dx\right].$$

Proof.

$$\text{(3)}\quad \mathbf{E}\left\{\left(\ln\frac{1}{|D|}\right)^{\nu}\right\} = \mathbf{E}\left\{\left(\ln\frac{1}{|D|}\right)^{\nu}\mathbb{I}_{E_q}\right\} + \mathbf{E}\left\{\left(\ln\frac{1}{|D|}\right)^{\nu}\mathbb{I}_{E_q^{\mathsf{c}}}\right\}$$

Here we denote by $\mathbb{I}_S$ the characteristic function of a set S.

Since D is a function of $\frac{a_k}{\|a_k\|}$, $k = 1, \ldots, n$, and $\mathbb{I}_{E_q}$ is a function of $\|a_k\|$, $k = 1, \ldots, n$, the random variables D and $\mathbb{I}_{E_q}$ are independent.

This implies that

$$\mathbf{E}\left\{\left(\ln\frac{1}{|D|}\right)^{\nu}\mathbb{I}_{E_q^{\mathsf{c}}}\right\} = \mathbf{E}\left\{\left(\ln\frac{1}{|D|}\right)^{\nu}\right\}\mathbf{P}(E_q^{\mathsf{c}}).$$

By Lemma 3, $\mathbf{P}(E_q^{\mathsf{c}}) \leq \frac{1}{2}$, and therefore it follows from (3)

$$\text{(4)}\quad \mathbf{E}\left\{\left(\ln\frac{1}{|D|}\right)^{\nu}\right\} \leq 2\mathbf{E}\left\{\left(\ln\frac{1}{|D|}\right)^{\nu}\mathbb{I}_{E_q}\right\}.$$

Consequently

$$\begin{aligned}
\text{(5)}\ \mathbf{E}\left\{\left(\ln\frac{1}{|D|}\right)^{\nu}\mathbb{I}_{E_q}\right\} &= \int_0^{+\infty} \mathbf{P}\left(\left(\ln\frac{1}{|D|}\right)^{\nu}\mathbb{I}_{E_q} > x\right) dx \\
&= \int_0^{+\infty} \mathbf{P}\left(\left\{|D| < e^{-x^{1/\nu}}\right\} \cap E_q\right) dx \\
&\leq \int_0^{+\infty} \mathbf{P}\left(\{|\underline{Z}| < \sqrt{m}qe^{-x^{1/\nu}}\} \cap E_q\right) dx \\
&\leq \alpha + \int_{\alpha}^{+\infty} \mathbf{P}\left(\left\{|\underline{Z}| < \sqrt{m}qe^{-x^{1/\nu}}\right\} \cap E_q\right) dx
\end{aligned}$$

the third line by Lemma 2. Replacing (5) in (4) we get

$$\mathbf{E}\left\{\left(\ln\frac{1}{|D|}\right)^{\nu}\right\} \le 2\left[\alpha + \int_{\alpha}^{+\infty} \mathbf{P}\left(\left\{|\underline{Z}| < \sqrt{m}qe^{-x^{1/\nu}}\right\} \cap E_q\right) dx\right].$$

□

In the rest of this paper we fix $q = \max\left\{3.2, \sqrt{\frac{3\ln(2n)}{m}}\right\}$. Thus

$$\ln q = \max\left\{\ln 3.2, \frac{1}{2}\left(\ln(\ln 2n) - \ln m + \ln 3\right)\right\}.$$

In the sequel we focus on bounding the probability

$$\mathbf{P}(\{|\underline{Z}| < a\} \cap E_q).$$

3.2 Bounding $\mathbf{P}(\{|\underline{Z}| < a\} \cap E_q)$*: preliminaries*

In what follows we face the difficulty that the random process $Z = \{Z(x) : x \in S^{m-1}\}$ does not have differentiable paths. We will deal with this by smoothing out Z through convolution with a deterministic kernel.

Let $\psi_\varepsilon : \mathbb{R}^m \to \mathbb{R}^+$ be a regular isotropic approximation of the unity. It is defined, for $\varepsilon > 0$, by

$$\psi_\varepsilon(y) = \frac{c_m}{\varepsilon^m}\psi\left(\frac{\|y\|}{\varepsilon}\right)$$

where $\psi : [0, +\infty) \to \mathbb{R}^+$ is of class C^∞, $\operatorname{supp}(\psi) \subseteq [0, 1]$,

$$c_m = \frac{1}{\sigma_{m-1}(S^{m-1})} = \frac{\Gamma(m/2)}{2\pi^{m/2}},$$

and $\int_{\mathbb{R}^m} \psi_\varepsilon(y)dy = 1$. Here σ_{m-1} denotes the standard surface measure in S^{m-1}.

Note that if $\int_{\mathbb{R}^m} \psi_\varepsilon(y)dy = 1$ holds then

$$1 = \int_{\mathbb{R}^m} \psi_\varepsilon(y)dy = c_m \int_{\mathbb{R}^m} \frac{1}{\varepsilon^m}\psi\left(\frac{\|y\|}{\varepsilon}\right) dy = \int_0^1 \rho^{m-1}\psi(\rho)\, d\rho.$$

Thus, to ensure that $\int_{\mathbb{R}^m} \psi_\varepsilon(y)dy = 1$ it suffices to choose ψ satisfying that $\int_0^1 \rho^{m-1}\psi(\rho)\, d\rho = 1$. This can be obtained with

$$\|\psi\|_\infty \le C_1 m \quad \text{and} \quad C_1 = \frac{e}{e-1}. \tag{6}$$

Indeed, notice that if we take $\psi(\rho) = bm\mathbb{I}_{(1-\frac{1}{m},1)}(\rho)$ we have:

$$\int_0^1 \rho^{m-1}\psi(\rho)\, d\rho = bm\left(\frac{\rho^m}{m}\right)\Bigg|_{\rho=1-\frac{1}{m}}^1$$

which is equal to 1 if we choose

$$b = \left[1 - \left(1 - \frac{1}{m}\right)^m\right]^{-1} < \frac{e}{e-1}.$$

This ψ verifies $\|\psi\|_\infty = bm$ but is not C^∞. At the cost of slightly increasing its norm we can change it into one such function that also verifies the remaining conditions.

Now define $Z_\varepsilon : \mathbb{R}^m \to \mathbb{R}$ by

$$Z_\varepsilon(x) = (\psi_\varepsilon * Z)(x) = \int_{\mathbb{R}^m} \psi_\varepsilon(x-y)Z(y)dy$$

and

$$\underline{Z}_\varepsilon = \min_{x \in S^{m-1}} Z_\varepsilon(x).$$

Since for every $\omega \in \Omega$

$$Z_\varepsilon(x) \underset{\varepsilon \to 0}{\longrightarrow} Z(x) \quad \text{uniformly on } S^{m-1}$$

we may deduce a bound for

$$\mathbf{P}(\{|\underline{Z}| < a\} \cap E_q)$$

by passing to the limit as $\varepsilon \to 0$ in a bound for $\mathbf{P}(\{|\underline{Z}_\varepsilon| < a\} \cap E_q)$. This will be our goal. To attain it we will use a simple geometric idea. For all $\varepsilon, a > 0$, the condition

$$\left|\min_{x \in S^{m-1}} Z_\varepsilon(x)\right| < a$$

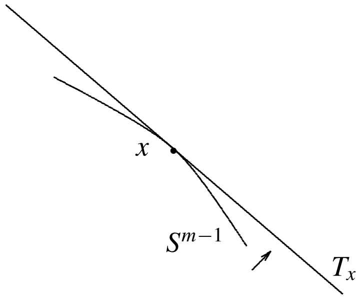
implies the existence of a local minimum ξ of Z_ε such that $|Z_\varepsilon(\xi)| < a$. The idea to bound $\mathbf{P}(\{|\underline{Z}_\varepsilon| < a\} \cap E_q)$ is to use a formula for counting the number of local minima ξ as above. We next give such formula.

Let $F : S^{m-1} \to \mathbb{R}$ and $a > 0$. Denote

$$\mathcal{M}^F(a) = \left\{x \in S^{m-1} \text{ s.t. } |F(x)| < a \text{ and } x \text{ is a local minimum of } F\right\}$$

and let $m^F(a) = \#\mathcal{M}^F(a)$, the cardinality of $\mathcal{M}^F(a)$ (if $a = +\infty$ we simply write $\mathcal{M}^F$ and m^F). We also use the following notations:

- $(\widetilde{D}F)$ and $(\widetilde{D}^2F)$ denote respectively the first and second derivatives of F. For their computation at a point $x \in S^{m-1}$ we will parametrize S^{m-1} locally by projection on the tangent hyperplane T_x.



- If $A : \mathbb{R}^{m-1} \to \mathbb{R}^{m-1}$ is a real symmetric linear transformation we write $A \succ 0$ to denote that A is positive definite.

Lemma 4 *Let F be smooth and assume that all the local minima of F are strict (in the sense that $\widetilde{D}^2F$ is non-singular at those points). Then,*

$$m^F = \lim_{\delta\to 0} \frac{1}{|B_{m-1}(0;\delta)|} \int_{S^{m-1}} \det(\widetilde{D}^2F)(x) \mathbb{I}_{\{\|(\widetilde{D}F)(x)\|<\delta\}} \mathbb{I}_{\{(\widetilde{D}^2F)(x)\succ 0\}} \sigma_{m-1}(dx).$$

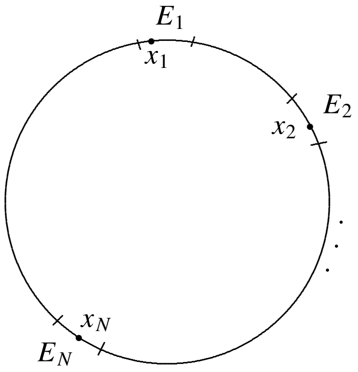
where $B_{m-1}(0;\delta)$ is the open ball of radius δ centered at the origin in $\mathbb{R}^{m-1}$ and $|B_{m-1}(0;\delta)|$ its Lebesgue measure in $\mathbb{R}^{m-1}$.

Proof. The hypothesis implies that the local minima are isolated points in S^{m-1}, hence m^F is finite. Let $m^F = N$ and $\mathcal{M}^F = \{x_1, \ldots, x_N\}$. Then, for $j = 1, \ldots, N$,

$$\left.\begin{array}{l}(\widetilde{D}F)(x_j) = 0 \\ (\widetilde{D}^2F)(x_j) \succ 0.\end{array}\right\}$$

For $\delta_0 > 0$, δ_0 small enough, there exist $E_1, \ldots, E_N$, pairwise disjoint neighbourhoods of $x_1, \ldots, x_N$ respectively, such that

$$\bigcup_{j=1}^{N} E_j = \left\{x : \|(\widetilde{D}F)(x)\| < \delta_0 \text{ and } (\tilde{D}^2F)(x) \succ 0\right\}$$



and, for each $j = 1, \ldots, N$, the map $x \mapsto (\widetilde{D}F)(x)$ is a diffeomorphism between E_j and $B_{m-1}(0; \delta_0)$. Thus, $\int_{E_j} \det(\widetilde{D}^2 F(x))\ \sigma_{m-1}(dx) = |B_{m-1}(0; \delta_0)|$. Consequently

$$m^F = \frac{1}{|B_{m-1}(0; \delta_0)|} \int_{S^{m-1}} \det(\widetilde{D}^2 F)(x) \mathbb{I}_{\{\|(\widetilde{D}F)(x)\| < \delta_0\}} \mathbb{I}_{\{(\widetilde{D}^2 F)(x) \succ 0\}} \sigma_{m-1}(dx).$$

Since this holds for every $\delta \leq \delta_0$ the conclusion follows. □

Remark 4 The same proof, *mutatis mutandis*, shows that

$$m^F(a) = \lim_{\delta \to 0} \frac{1}{|B_{m-1}(0;\delta)|} \int_{S^{m-1}} \det(\widetilde{D}^2 F)(x) \mathbb{I}_{\{\|(\widetilde{D}F)(x)\| < \delta\}} \mathbb{I}_{\{(\widetilde{D}^2 F)(x) \succ 0\}} \mathbb{I}_{\{|F(x)| < a\}} \sigma_{m-1}(dx).$$

Let

$$I^\bullet(\varepsilon, \delta, a) = \mathbf{E}\left\{\det(\widetilde{D}^2 Z_\varepsilon)(e_1) \mathbb{I}_{\{(\widetilde{D}^2 Z_\varepsilon)(e_1) \succ 0\} \cap \{\|(\widetilde{D} Z_\varepsilon)(e_1)\| < \delta\} \cap \{|Z_\varepsilon(e_1)| \leq a\} \cap E_q}\right\}$$

and

$$I(\varepsilon, \delta, a) = \frac{1}{|B_{m-1}(0; \delta)|} I^\bullet(\varepsilon, \delta, a).$$

Proposition 4 *For all $\varepsilon > 0$,*

$$\mathbf{P}(\{|\underline{Z}_\varepsilon| < a\} \cap E_q) \leq \sigma_{m-1}(S^{m-1}) \overline{\lim_{\delta \to 0}}\, I(\varepsilon, \delta, a).$$

To prove Proposition 4 we would like to apply Lemma 4 to the (random) function $F = Z_\varepsilon$. To be able to do this we need a result ensuring that, with probability one, the local minima of Z_ε are strict. The next lemma provides such a result.

Lemma 5 *Let $V : [0, 1]^d \to \mathbb{R}^d$, be a random vector field defined on the probability space $(\Omega, \mathcal{A}, P)$ and $y \in \mathbb{R}^d$ such that:*

(1) For almost every $\omega \in \Omega$ the function $\tau \mapsto V(\tau)$ is twice continuously differentiable.
(2) For each $\tau \in [0,1]^d$ the probability distribution in $\mathbb{R}^d$ of the random vector $V(\tau)$ has a density π_τ.
(3) There exist positive constants L and δ such that

$$\sup\left\{\pi_\tau(z) : \tau \in [0,1]^d, \|z-y\| \le \delta\right\} \le L.$$

Then the probability for y to be a critical value of V is equal to zero. That is,

$$\mathbf{P}\left(\left\{\tau : \tau \in [0,1]^d,\ V(\tau) = y,\ \det DV(\tau) = 0\right\} \neq \emptyset\right) = 0.$$

Proof. Suppose $y = 0$ (if this is not the case, replace V by $V - y$).

Denote by E_0 the set

$$E_0 = \left\{\tau : \tau \in [0,1]^d,\ V(\tau) = 0,\ \det DV(\tau) = 0\right\}$$

and for each positive integer N consider $[0,1]^d$ as the union of the cubes of sides equal to $\frac{1}{N}$ that are products of intervals of the form $[\frac{j_1}{N}, \frac{j_1+1}{N}] \times \cdots \times [\frac{j_d}{N}, \frac{j_d+1}{N}],\ 0 \le j_1, \ldots, j_d \le N-1$. Denote these cubes by $C_1, \ldots, C_{N^d}$.

In a similar way, consider each face of the boundary of each cube C_r as a union of $(d-1)$-dimensional cubes of sides equal to $\frac{1}{N^2}$. We denote these cubes by D_{rs}, $s = 1, \ldots, 2dN^{d-1}$.

In each set D_{rs} fix a point τ^*_{rs}. For instance, let τ^*_{rs} be the center of D_{rs}.

For given $\eta > 0$, using the hypothesis, we can find $B > 0$ (large enough) so that if we denote F_B the event

$$F_B = \left\{\left[\sup\left|\frac{\partial V_i}{\partial \tau_j}(\tau)\right|, \left|\frac{\partial^2 V_i}{\partial \tau_j \partial \tau_h}(\tau)\right| : i, j, h = 1, \ldots, d;\ \tau \in [0,1]^d\right] > B\right\}$$

where $V = (V_1, \ldots, V_d)$ and $\tau = (\tau_1, \ldots, \tau_d)$ one has

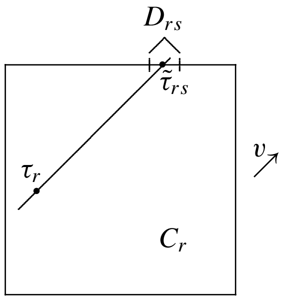
$$\mathbf{P}(F_B) < \eta.$$

Clearly,

$$\{E_0 \neq \emptyset\} = \bigcup_{r=1}^{N^d} \{E_0 \cap C_r \neq \emptyset\} \tag{7}$$

and

$$\begin{aligned}\{E_0 \cap C_r \neq \emptyset\} &\subset \{\exists \tau_r \in C_r \text{ s.t. } V(\tau_r) = 0, \det DV(\tau_r) = 0\}\\ &= \Big\{\exists \tau_r \in C_r, v \in \mathbb{R}^d \text{ s.t. } V(\tau_r) = 0, \|v\| = 1, \text{ and}\\ &\qquad DV(\tau_r)v = 0\Big\}.\end{aligned}$$



Let $\omega \in \{E_0 \cap C_r \neq \emptyset\} \cap F_B^{\mathsf{c}}$. Take an intersection point of the straight line through τ_r parallel to v with the boundary of C_r. This point belongs to some D_{rs} for some $s \in \{1, \dots, 2dN^{d-1}\}$. We denote it by $\tilde{\tau}_{rs}$. Let $i \le d$ and consider the Taylor expansion of V_i around the point τ_r, evaluated at $\tilde{\tau}_{rs}$. This is

$$V_i(\tilde{\tau}_{rs}) = V_i(\tau_r) + \sum_{j=1}^{d} \frac{\partial V_i}{\partial \tau_j}(\tau_r)(\tilde{\tau}_{rs,j} - \tau_{r,j})$$
$$+\frac{1}{2} \sum_{j,h=1}^{d} \frac{\partial^2 V_i}{\partial \tau_j \partial \tau_h}(\tau_r + \theta(\tilde{\tau}_{rs} - \tau_r))(\tilde{\tau}_{rs,j} - \tau_{r,j})(\tilde{\tau}_{rs,h} - \tau_{r,h})$$

with $0 < \theta < 1$. Since the first two terms of this sum are equal to zero, we deduce that, for $i = 1, \dots, d$,

$$|V_i(\widetilde{\tau}_{rs})| \le K_d B N^{-2}$$

where K_d is a constant depending only on the dimension d.

Since the diameter of each D_{rs} is bounded by $K_d N^{-2}$ the existence of $\widetilde{\tau}_{rs} \in D_{rs}$ s.t. $|V_i(\widetilde{\tau}_{rs})| \le K_d B N^{-2}$ implies that $\left\| V(\tau_{rs}^*) \right\| \le K N^{-2}$ for some constant K depending only on d and B. Here τ_{rs}^* is the point a priori chosen in D_{rs} (apply the mean value theorem on the line segment with extremities $\tau_{rs}^*, \widetilde{\tau}_{rs}$ and the fact that $\omega \in F_B^{\mathsf{c}}$).

Hence, if N is large enough so that

$$K \left(\frac{1}{N}\right)^2 < \delta$$

with δ given by hypothesis (3), one has

$$\mathbf{P}(E_0 \neq \emptyset) \le \mathbf{P}(F_B) + \mathbf{P}(\{E_0 \neq \emptyset\} \cap F_B^{\mathsf{c}})$$
$$< \eta + \mathbf{P}\left(\exists r \le N^d, s \le 2dN^{d-1} \text{ s.t. } \left\| V(\tau_{rs}^*) \right\| \le K N^{-2}\right)$$

$$
\begin{aligned}
&\le \eta + \sum_{r=1}^{N^d} \sum_{s=1}^{2dN^{d-1}} \mathbf{P}\left(\left\| V(\tau_{rs}^*)\right\| \le KN^{-2}\right)\\
&\le \eta + \sum_{r=1}^{N^d} \sum_{s=1}^{2dN^{d-1}} \int_{\|z\|\le KN^{-2}} \pi_{\tau_{rs}^*}(z)dz\\
&\le \eta + N^d N^{d-1} L K_1 N^{-2d}
\end{aligned}
$$

where K_1 is a new constant depending only on B and d. Letting $N \to +\infty$ and using that η is an arbitrary positive number, the result follows. □

Proof of proposition 4.

$$
\mathbf{P}\left(\{|\underline{Z}_\varepsilon| < a\} \cap E_q\right) \le \mathbf{P}\left(\left\{m^{Z_\varepsilon}(a) \ge 1\right\} \cap E_q\right) \le \mathbf{E}\left\{m^{Z_\varepsilon}(a)\mathbb{I}_{E_q}\right\}
$$

Note that the property in Lemma 5 is local and to apply it to the random field $V = \tilde{D}Z_\varepsilon$ and the value $y = 0$ one only needs to check conditions (2) and (3) in that lemma. The probability distribution of $\tilde{D}Z_\varepsilon(x)$, $x \in S^{m-1}$, is invariant under a linear linear isometry of $\mathbb{R}^m$ so that it suffices to check (2) and (3) for $x = e_1$. This is contained in Proposition 9 (take $g = 1$) which we state in §3.4 below and prove in Section 5.

So we may use Lemma 4 which, together with Fatou's Lemma, permit to bound the expectation above (for related formulae, see for example [Adler 1981; Brillinger 1972]). We get:

$$
\begin{aligned}
\mathbf{E}\left\{m^{Z_\varepsilon}(a)\mathbb{I}_{E_q}\right\} &= \mathbf{E}\left\{\lim_{\delta\to 0} \frac{1}{|B_{m-1}(0;\delta)|}\int_{S^{m-1}} \det(\widetilde{D}^2 Z_\varepsilon)(x)\mathbb{I}_{\{(\widetilde{D}^2 Z_\varepsilon)(x)\succ 0\}}\right.\\
&\qquad \left.\times \mathbb{I}_{\{\|(\widetilde{D}Z_\varepsilon)(x)\|<\delta\}\cap\{|Z_\varepsilon(x)|<a\}\cap E_q}\ \sigma_{m-1}(dx)\right\}\\
&\le \overline{\lim_{\delta\to 0}} \frac{1}{|B_{m-1}(0;\delta)|}\int_{S^{m-1}} \mathbf{E}\Big\{\det(\widetilde{D}^2 Z_\varepsilon)(x)\mathbb{I}_{\{(\widetilde{D}^2 Z_\varepsilon)(x)\succ 0\}}\\
&\qquad \times \mathbb{I}_{\{\|(\widetilde{D}Z_\varepsilon)(x)\|<\delta\}\cap\{|Z_\varepsilon(x)|<a\}\cap E_q}\Big\}\sigma_{m-1}(dx).
\end{aligned}
$$

Since the law of the random set of vectors $\{a_1, \dots, a_n\}$ and $\{Ta_1, \dots, Ta_n\}$ is the same for any isometry T of $\mathbb{R}^m$ it follows that the last integrand does not depend on x and the result follows. □

Using Proposition 4 we obtain

$$
\mathbf{P}(\{|\underline{Z}| < a\} \cap E_q) \le \sigma_{m-1}(S^{m-1})\overline{\lim_{\varepsilon\to 0}}\ \overline{\lim_{\delta\to 0}}\ I(\varepsilon, \delta, a).
$$

Consequently we will focus on estimating the limit in the right-hand side. The next section is devoted to the computation of the first and second derivatives of Z_ε and the following one to describe the joint distribution of Z_ε and its gradient, as well as its limiting behaviour as $\varepsilon \to 0$.

3.3 Computation of partial derivatives

We now compute the first and second derivatives of $Z_\varepsilon : \mathbb{R}^m \to \mathbb{R}$. Later we will be interested in computing the derivatives of the restriction $Z_\varepsilon : S^{m-1} \to \mathbb{R}$. To distinguish between both we will write a $\tilde{}$ over the latter (as we did above). Thus, for instance, DZ_ε denotes the first derivative of $Z_\varepsilon : \mathbb{R}^m \to \mathbb{R}$ and $\widetilde{D}Z_\varepsilon$ that of $Z_\varepsilon : S^{m-1} \to \mathbb{R}$.

Proposition 5 *For all* $x \in \mathbb{R}^m$ *and all* $j, l = 1, \ldots, m$,

(i) $\frac{\partial Z_\varepsilon}{\partial x_j}(x) = \int_{\mathbb{R}^m} \psi_\varepsilon(x-y)\left(\sum_{k=1}^n a_{kj}\mathbb{I}_{U_k}(y)\right) dy$

(ii) $\left|\frac{\partial^2 Z_\varepsilon}{\partial x_j \partial x_\ell}(x)\right| \le \frac{C_1\sqrt{m}}{\varepsilon}\sum_{k=1}^n |a_{kj}|$.

Proof. Recall that for $x \in \mathbb{R}^m$, we defined

$$Z_\varepsilon(x) = \int_{\mathbb{R}^m} \psi_\varepsilon(x-y)Z(y)dy.$$

Therefore, for $j = 1, \ldots, m$,

$$\begin{aligned}\frac{\partial Z_\varepsilon}{\partial x_j}(x) &= \int_{\mathbb{R}^m} \frac{\partial \psi_\varepsilon}{\partial x_j}(x-y)Z(y)dy \\ &= \int_{\mathbb{R}^{m-1}} \prod_{h\neq j} dy_h \int_{\mathbb{R}} \frac{\partial \psi_\varepsilon}{\partial x_j}(x-y)Z(y)dy_j \\ &= \int_{\mathbb{R}^{m-1}} \prod_{h\neq j} dy_h \int_{\mathbb{R}} \psi_\varepsilon(x-y)\frac{\partial Z}{\partial y_j}(y)dy_j \qquad (8)\end{aligned}$$

the last equality on integration by parts since $\frac{\partial \psi_\varepsilon}{\partial x_j}(x-y) = -\frac{\partial \psi_e}{\partial y_j}(x-y)$ and $\psi_\varepsilon(x-y)Z(y)\Big|_{y_j=-\infty}^{y_j=+\infty} = 0$ because ψ_ε has compact support (as a function of y_j).

Denote

$$U_k = \{y \in \mathbb{R}^m : \langle y, a_k\rangle > \langle y, a_l\rangle \text{ for all } l \neq k\}.$$

Note that almost surely, for almost all $y \in \mathbb{R}^m$ —in the sense of Lebesgue measure— there is exactly one k such that $y \in U_k$. In addition, if $y \in U_k$ one has

$$\frac{\partial Z}{\partial y_j}(y) = \frac{\partial}{\partial y_j}(\langle y, a_k\rangle) = a_{kj}.$$

Therefore

$$\frac{\partial Z_\varepsilon}{\partial x_j}(x) = \int_{\mathbb{R}^m} \psi_\varepsilon(x-y)\left(\sum_{k=1}^n a_{kj}\mathbb{I}_{U_k}(y)\right) dy.$$

We now proceed to part (ii). To do so, the following notation will be useful. For $x \in \mathbb{R}^m$ we denote by $\widetilde{x}^j$ the vector in $\mathbb{R}^{m-1}$ obtained from x by suppressing the jth coordinate.

For $j, \ell = 1, \ldots, m$,

$$\frac{\partial^2 Z_\varepsilon}{\partial x_j \partial x_\ell}(x) = \int_{\mathbb{R}^m} \frac{\partial \psi_\varepsilon}{\partial x_\ell}(x-y)\left[\sum_{k=1}^n a_{kj}\mathbb{I}_{U_k}(y)\right] dy$$
$$= \sum_{k=1}^n a_{kj} \int_{\mathbb{R}^{m-1}} \prod_{h\neq\ell} dy_h \int_{\mathbb{R}} \frac{\partial \psi_\varepsilon}{\partial x_\ell}(x-y)\mathbb{I}_{U_k}(y)\, dy_\ell.$$

Note that for each $k = 1, \ldots, n$, the set U_k is convex so that, fixing $\widetilde{y}_\ell \in \mathbb{R}^{m-1}$, the set

$$U_{k\ell} = \{y_\ell : (y_1, \ldots, y_\ell, \ldots, y_m) \in U_k\}$$

is also convex. Thus $U_{k\ell}$ is an interval (possibly infinite),

$$\bigcup_{k=1}^n \overline{U}_{k\ell} = \mathbb{R},$$

($\overline{U}_{k\ell}$ denotes the closure of $U_{k\ell}$) and, almost surely, the interiors of the $\overline{U}_{k\ell}$ are pairwise disjoint. Therefore, fixing $x \in \mathbb{R}^m$ and $\widetilde{y}_\ell \in \mathbb{R}^{m-1}$ we have

$$\int_{\mathbb{R}} \frac{\partial \psi_\varepsilon}{\partial x_\ell}(x-y)\mathbb{I}_{U_k}(y_\ell)\, dy_\ell = \int_{\alpha_{k\ell}}^{\beta_{k\ell}} \frac{\partial \psi_\varepsilon}{\partial x_\ell}(x-y)\, dy_\ell = G(\widetilde{y}_\ell; x)$$

where

$$G(\widetilde{y}_\ell; x) = \psi_\varepsilon(x_1 - y_1, \ldots, x_{\ell-1} - y_{\ell-1}, x_\ell - \alpha_{k\ell}, \ldots, x_m - y_m)$$
$$-\psi_\varepsilon(x_1 - y_1, \ldots, x_{\ell-1} - y_{\ell-1}, x_\ell - \beta_{k\ell}, \ldots, x_m - y_m)$$

and $\alpha_{k\ell} < \beta_{k\ell}$ are the extremities of $U_{k\ell}$. Hence

$$\frac{\partial^2 Z_\varepsilon}{\partial x_j \partial x_\ell}(x) = \sum_{k=1}^n a_{kj} \int_{B_{m-1}(\widetilde{x}^\ell;\varepsilon)} G(\widetilde{y}_\ell; x) \prod_{h\neq\ell} dy_h.$$

So,

$$\left|\frac{\partial^2 Z_\varepsilon}{\partial x_j \partial x_\ell}(x)\right| \leq \sum_{k=1}^n |a_{kj}||B_{m-1}(0;1)|\varepsilon^{m-1}\|\psi_\varepsilon\|_\infty$$
$$= \sum_{k=1}^n |a_{kj}|\frac{|B_{m-1}(0;1)|}{\varepsilon}\|\psi\|_\infty c_m \tag{9}$$

(this bound is asymmetric with respect to (j, ℓ) because of the way it has been obtained). Also:

$$|B_{m-1}(0;1)| = \sigma_{m-2}(S^{m-2}) \int_0^1 \rho^{m-2} d\rho = \frac{1}{(m-1)c_{m-1}}.$$

As a consequence, using (6) and Lemma 6 below, the bound in (9) implies that

$$\left| \frac{\partial^2 Z_\varepsilon}{\partial x_j \partial x_\ell}(x) \right| \le \frac{C_1\sqrt{m}}{\varepsilon} \sum_{k=1}^n |a_{kj}|. \tag{10}$$

□

Lemma 6 *For all* $m \ge 2$, $\frac{mc_m}{(m-1)c_{m-1}} \le \sqrt{m}$.

Proof. Use $c_m = \dfrac{\Gamma\left(\frac{m}{2}\right)}{2\pi^{\frac{m}{2}}}$ and Stirling's formula, $\Gamma(t) \ge \sqrt{2\pi} t^{t-\frac{1}{2}} e^{-t}$, (see [Ahlfors 1979], Chapter 5, Section 2.5, Exercise 2). □

We now compute the partial derivatives of the restriction of Z_ε to S^{m-1}. We denote by $\{e_1, \dots, e_m\}$ the canonical basis in $\mathbb{R}^m$, where $e_j = \left(\delta_{jh}\right)_{h=1,\dots,m}$, $j = 1, \dots, m$, and δ_{jh} is the Kronecker δ.

Proposition 6 *Choose* $\{e_2, \dots, e_m\}$ *as a basis for the tangent hyperplane to* S^{m-1} *at the point* e_1. *Then:*

(i) *For* $j = 2, \dots, m$

$$\frac{\widetilde{\partial} Z_\varepsilon}{\partial x_j}(e_1) = \frac{\partial Z_\varepsilon}{\partial x_j}(e_1) = \int_{\mathbb{R}^m} \psi_\varepsilon(e_1 - y) \left[\sum_{k=1}^n a_{kj} \mathbb{I}_{U_k}(y) \right] dy.$$

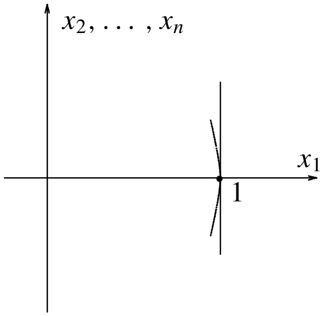
(ii) *For* $j, \ell = 2, \dots, m$

$$\frac{\widetilde{\partial}^2 Z_\varepsilon}{\partial x_j \partial x_\ell}(e_1) = -\frac{\partial Z_\varepsilon}{\partial x_1}(e_1)\delta_{j\ell} + \frac{\partial^2 Z_\varepsilon}{\partial x_j \partial x_\ell}(e_1).$$

Proof. Fix the parametrization of S^{m-1} defined in a small neighbourhood of e_1 by

$$\begin{aligned} x_1 &= \gamma(x_2, \dots, x_m) \\ &= (1 - (x_2^2 + \cdots + x_m^2))^{1/2}. \end{aligned}$$

Thus, $(x_2, \dots, x_m)$ varies in a small ball centered at $(0, \dots, 0) \in \mathbb{R}^{m-1}$.



To prove part (i), using Proposition 5 (i) we obtain, for $j = 2, \ldots, m$,

$$\begin{aligned}\frac{\widetilde{\partial} Z_\varepsilon}{\partial x_j}(x) &= \frac{\partial Z_\varepsilon}{\partial x_1}\frac{\partial \gamma}{\partial x_j}(x) + \frac{\partial Z_\varepsilon}{\partial x_j}(x)\\ &= \int_{\mathbb{R}^m} \psi_\varepsilon(x-y)\left[\sum_{k=1}^n a_{kj}\,\mathbb{I}_{U_k}(y)\right]dy\left(\frac{-x_j}{\gamma}\right)\\ &\quad + \int_{\mathbb{R}^m} \psi_\varepsilon(x-y)\left[\sum_{k=1}^n a_{kj}\mathbb{I}_{U_k}(y)\right]dy\end{aligned}$$

where γ denotes $\gamma(x_2, \ldots, x_m)$, and with $x = (\gamma(x_2, \ldots, x_m), x_2, \ldots, x_m)$. Evaluating at e_1, we get

$$\frac{\widetilde{\partial} Z_\varepsilon}{\partial x_j}(e_1) = \int_{\mathbb{R}^m} \psi_\varepsilon(e_1 - y)\left[\sum_{k=1}^n a_{kj}\mathbb{I}_{U_k}(y)\right]dy.$$

For part (ii), if $j = 2, \ldots, m$,

$$\begin{aligned}\left.\frac{\widetilde{\partial}^2 Z_\varepsilon}{\partial x_j^2}\right|_{x=e_1} &= \left[\left(-\frac{\partial^2 Z_\varepsilon}{\partial x_1^2}\cdot\frac{x_j}{\gamma} + \frac{\partial^2 Z_\varepsilon}{\partial x_1 \partial x_j}\right)(-\frac{x_j}{\gamma}) - \frac{\partial Z_\varepsilon}{\partial x_1}\left(\frac{\gamma - x_j(-\frac{x_j}{\gamma})}{\gamma^2}\right)\right.\\ &\quad \left.\left. + \frac{\partial^2 Z_\varepsilon}{\partial x_j \partial x_1}(-\frac{x_j}{\gamma}) + \frac{\partial^2 Z_\varepsilon}{\partial x_j^2}\right]\right|_{x=e_1} = -\frac{\partial Z_\varepsilon}{\partial x_1}(e_1) + \frac{\partial^2 Z_\varepsilon}{\partial x_j^2}(e_1).\end{aligned}$$

For $j, \ell = 2, \ldots, m$, $j \neq \ell$,

$$\begin{aligned}\left.\frac{\widetilde{\partial}^2 Z_\varepsilon}{\partial x_j \partial x_\ell}\right|_{x=e_1} &= \left[\left(-\frac{\partial^2 Z_\varepsilon}{\partial x_1^2}\frac{x_\ell}{\gamma} + \frac{\partial^2 Z_\varepsilon}{\partial x_1 \partial x_\ell}\right)(-\frac{x_j}{\gamma}) - \frac{\partial Z_\varepsilon}{\partial x_1}\left(x_j(-\frac{1}{2})\frac{1}{\gamma^3}(-2x_\ell)\right)\right.\\ &\quad \left.\left. + \frac{\partial^2 Z_\varepsilon}{\partial x_j \partial x_1}(-\frac{x_\ell}{\gamma}) + \frac{\partial^2 Z_\varepsilon}{\partial x_j \partial x_\ell}\right]\right|_{x=e_1} = \frac{\partial^2 Z_\varepsilon}{\partial x_j \partial x_\ell}(e_1).\end{aligned}$$

□

Remark 5 The computations in Propositions 5 and 6 remain valid if instead of the canonical basis $\{e_1, e_2, \ldots, e_m\}$ one uses a basis $\{e_1, v_2, \ldots, v_m\}$ with $\{v_2, \ldots, v_m\}$ any orthonormal basis of the orthogonal complement of e_1.

3.4 Bounding $\overline{\lim}_{\varepsilon\to 0}\overline{\lim}_{\delta\to 0} I(\varepsilon,\delta,a)$

We now return to our purpose, stated at the end of §3.2, of bounding $\overline{\lim}_{\varepsilon\to 0}\overline{\lim}_{\delta\to 0} I(\varepsilon,\delta,a)$. The main statement in this section is the following.

Proposition 7 *For $a \geq 0$ the expression $\overline{\lim}_{\varepsilon\to 0}\overline{\lim}_{\delta\to 0} I(\varepsilon,\delta,a)$ is bounded by*

$$\sum_{t=1}^{m\wedge n}\Bigg[\binom{n}{t}\frac{(t-1)!}{\pi t}\left(\frac{e^2}{2}\right)^t \left(4C_1 n m^{3/2} q^2\right)^{t-1}\frac{1}{(2\pi)^{\frac{m+t-1}{2}}}$$
$$\int_{-a}^{a}|x|^{m-t}(\Phi(x))^{n-t-1}\left(\sqrt{\frac{2\pi}{t}}\Phi(\sqrt{t}x)(n-t)\varphi(x)+t^{\frac{m}{2}}e^{-\frac{t}{2}x^2}\Phi(x)\right)dx\Bigg].$$

To obtain bounds for the limiting behaviour of $I(\varepsilon,\delta,a)$ as $\delta\to 0$ and $\varepsilon\to 0$ (in that order) we intend to describe the joint distribution of the random vector

$$(Z_\varepsilon(e_1), (\tilde{D}Z_\varepsilon)(e_1))$$

in $\mathbb{R}^m$ as well as to bound the expression $\det(\tilde{D}^2 Z_\varepsilon)(e_1))\mathbb{I}_{E_q}$ appearing in $I^\bullet(\varepsilon,\delta,a)$ (or in $I(\varepsilon,\delta,a)$).

Note that almost surely one has

$$Z_\varepsilon(e_1)\underset{\varepsilon\to 0}{\longrightarrow} Z(e_1)=\max_{1\leq k\leq n} a_{k1}$$

and, for $j=2,\ldots,m$,

$$\frac{\tilde{\partial} Z_\varepsilon(e_1)}{\partial x_j}\underset{\varepsilon\to 0}{\longrightarrow}\sum_{\ell=1}^{n} a_{\ell j}\mathbb{I}_{\left\{a_{\ell 1}=\max_{1\leq k\leq n} a_{k1}\right\}}.$$

Denote, for $\ell=1,\ldots,n$,

$$\chi_\ell=\mathbb{I}_{\left\{a_{\ell 1}=\max_{1\leq k\leq n} a_{k1}\right\}}$$

and, for $j=2,\ldots,m$,

$$Y_j=\sum_{\ell=1}^{n} a_{\ell j}\chi_\ell.$$

Note that both the χ_ℓ and the Y_j are well-defined almost surely. With this notation, the joint distribution of the m-tuple

$$\left(Z_\varepsilon(e_1),\frac{\tilde{\partial} Z_\varepsilon(e_1)}{\partial x_2},\ldots,\frac{\tilde{\partial} Z_\varepsilon(e_1)}{\partial x_m}\right)$$

converges, as $\varepsilon \to 0$, to the distribution of

$$(Z(e_1), Y_2, \dots, Y_m)$$

which is easy to compute. Actually, conditionally on the n-tuple $(a_{11}, \dots, a_{n1})$ the random variables $Y_2, \dots, Y_m$ are independent (Y_j depends only on $a_{\ell j}$, $\ell = 1, \dots, n$), Gaussian with mean zero and variance

$$\mathbf{E}(Y_j^2) = \sum_{\ell=1}^{n} \chi_\ell^2 \mathbf{E}\{a_{\ell j}^2\} = 1.$$

Hence, the distribution of $(Z(e_1), Y_2, \dots, Y_m)$ is that of m independent random variables, $Z(e_1)$ being the maximum of n i.i.d. $N(0, 1)$ r.v.'s and $Y_2, \dots, Y_m$ being $N(0, 1)$. Thus, the joint density at the point $(z, y_2, \dots, y_m)$ is

$$n\varphi(z)(\Phi(z))^{n-1} \prod_{j=2}^{m} \varphi(y_j).$$

The actual picture is, however, more complicated since we have *first* to pass to the limit when $\delta \to 0$ and this implies that, as $\varepsilon \to 0$, we have to deal with convergence of densities, not only of distributions.

The other problem to overcome is that, due to the non-smoothness of Z, in some ω-sets $\det(\tilde{D}^z Z_\varepsilon)(e_1))$ becomes very large for small $\varepsilon > 0$. As a consequence, we will need to estimate the size of the probability of these ω-sets and check that it compensates the growth of $\det(\tilde{D}^z Z_\varepsilon(e_1))$.

These are the resons for the technical detour implied by Propositions 8, 9, and 10 which we state below in this section. We delay their proof to Section 5 to avoid breaking the main stream of our argument.

Towards the proof of Proposition 7 we introduce some additional notations. Let, for $1 \le i \ne j \le n$,

$$F_{ij} = \{y \in \mathbb{R}^m : \langle y, a_i \rangle = \langle y, a_j \rangle\}$$

and

$$\theta(\varepsilon) = \#\{k : B(e_1; \varepsilon) \cap U_k \ne \emptyset\}.$$

Clearly $\theta(\varepsilon)$ can take the values $1, 2, \dots, n$. Note that if $\theta(\varepsilon) = t$ then, out of a set of measure zero, one also has $\theta(\varepsilon') = t$ for some $\varepsilon' > \varepsilon$, ε' depending on $\omega \in \Omega$. This is because, for each pair k, l with $k \ne l$, the probability for the distance from e_1 to F_{kl} to be exactly equal to ε is equal to zero.

For $t = 1, 2, \dots, n$ let

$$\begin{aligned} &I_t^\bullet(\varepsilon, \delta, a) \\ &\quad = \mathbf{E}\Big\{ \det(\widetilde{D}^2 Z_\varepsilon)(e_1) \mathbb{I}_{\{(\widetilde{D}^2 Z_\varepsilon)(e_1) \succ 0\} \cap \{\|(\widetilde{D} Z_\varepsilon)(e_1)\| < \delta\} \cap \{|Z_\varepsilon(e_1)| \le a\} \cap E_q \cap \{\theta(\varepsilon) = t\}} \Big\} \end{aligned}$$

and

$$I_t(\varepsilon, \delta, a) = \frac{1}{|B_{m-1}(0; \delta)|} I_t^\bullet(\varepsilon, \delta, a).$$

Clearly

$$I^\bullet(\varepsilon, \delta, a) = \sum_{t=1}^{n} I_t^\bullet(\varepsilon, \delta, a)$$

and

$$I(\varepsilon, \delta, a) = \sum_{t=1}^{n} I_t(\varepsilon, \delta, a). \tag{11}$$

To prove Proposition 7 we study the limiting behaviour of each term of the sum in (11) as $\delta \to 0$ and $\varepsilon \to 0$ (in that order).

Before that we state as a lemma an observation that we will be using repeatedly.

Lemma 7 *Let $\omega \in E_q$. Then:*

$$\frac{\partial Z_\varepsilon}{\partial x_1}(e_1) = Z_\varepsilon(e_1) + R_1$$

where $|R_1| \le \varepsilon q \sqrt{m}$.

Proof. We have

$$\begin{aligned}\frac{\partial Z_\varepsilon}{\partial x_1}(e_1) &= \int_{\mathbb{R}^m} \psi_\varepsilon(e_1 - y) \left(\sum_{k=1}^{n} a_{k1} \mathbb{I}_{U_k}(y) \right) dy \\ &= Z_\varepsilon(e_1) + \int_{\mathbb{R}^m} \psi_\varepsilon(e_1 - y) \left(\sum_{k=1}^{n} \langle a_k, e_1 - y \rangle \, \mathbb{I}_{U_k}(y) \right) dy.\end{aligned}$$

Use now that, since $\omega \in E_q$, $\|a_k\| \le q\sqrt{m}$ and that $\sum_{k=1}^{n} \mathbb{I}_{U_k}(y) = 1$ almost everywhere, to bound the absolute value of the second term. □

Denote by $\mathcal{X}$ the random set of integers

$$\mathcal{X} = \{k : U_k \cap B(e_1; \varepsilon) \neq \emptyset\}$$

so that $\theta(\varepsilon) = \#\mathcal{X} = t$.

We will call a *chain on* $\mathcal{X}$ a set of $t - 1$ non-ordered pairwise different pairs $\{k_j, l_j\}$, $j = 1, \dots, t - 1$, such that $k_j \neq l_j$, and

$$\bigcup_{j=1}^{t-1} \{k_j, l_j\} = \mathcal{X}.$$

A chain is *proper* if there exists a permutation σ of the indices $\{1, \dots, t-1\}$ such that for all $j = 2, \dots, t-1$ exactly one of $k_{\sigma(j)}, l_{\sigma(j)}$ does not belong to $\{k_{\sigma(1)}, l_{\sigma(1)}, \dots, k_{\sigma(j-1)}, l_{\sigma(j-1)}\}$. Without loss of generality, when considering proper chains in the sequel, we will assume that σ is the identity and that the element which has not appeared previously is k_j.

We next give a coarse bound on the number of chains on $\mathcal{X}$.

Lemma 8 *There are at most* $\frac{(t-1)!}{\pi t}\left(\frac{e^2}{2}\right)^t$ *different chains on* $\mathcal{X}$.

Proof. Clearly, there are at most $t^{2(t-1)}$ sequences of length $2(t-1)$ with elements from $\mathcal{X}$.

Consider the subset of those sequences

$$(k_1, l_1, k_2, l_2, \dots, k_{t-1}, l_{t-1})$$

that verify $k_j \neq l_j$ for $j = 1, \dots, t-1$ and also that the 2−subsets of $\mathcal{X}$

$$\{k_1, l_1\}, \{k_2, l_2\}, \dots, \{k_{t-1}, l_{t-1}\}$$

are pairwise different.

Each such sequence induces a chain on $\mathcal{X}$ but the same chain is counted $(t-1)!2^{t-1}$ times since we can permute the $t-1$ pairs as well as the 2 elements of each of these pairs without altering the chain. Now use Stirling's formula to get the stated bound. □

Recall we defined, for $k, l = 1, \dots, n, k \neq l$,

$$F_{kl} = \{y : \langle y, a_k\rangle = \langle y, a_l\rangle\}.$$

Lemma 9 *Suppose* $\theta(\varepsilon) = t > 2$. *Then there exists a proper chain*

$$\{k_1, l_1\}, \{k_2, l_2\}, \dots, \{k_{t-1}, l_{t-1}\}$$

on $\mathcal{X}$ *such that*

$$d(e_1, F_{k_j l_j}) < \varepsilon \ \textit{for} \ \ j = 1, \dots, t-1.$$

Here d *denotes Euclidean distance in* $\mathbb{R}^m$.

Proof. The construction of such a chain is immediate. Start with $k_1, l_1 \in \mathcal{X}$, $k_1 \neq l_1$. Since $t > 2$, $(\overline{U}_{k_1} \cup \overline{U}_{l_1}) \cap B(e_1; \varepsilon)$ is not the whole $B(e_1; \varepsilon)$. Therefore, there exists $k \in \mathcal{X}$ such that $k \neq k_1, l_1$ and U_k has a common boundary either with U_{k_1} or with U_{l_1} with a non empty intersection with $B(e_1; \varepsilon)$. Say it is with U_{k_1}. In that case choose $k_2 = k_1, l_2 = k$. The procedure continues until one fills $B(e_1; \varepsilon)$. □

In what follows, let $\rho = q\sqrt{m}$ and $\overline{B}_{m-1}(0;\rho) = \{\zeta \in \mathbb{R}^{m-1} \mid \|\zeta\| \le \rho\}$. Also, let M_ρ be the set of $n \times (m-1)$ real matrices $M = (a_{kj})$, $k = 1, \dots, n$; $j = 2, \dots, m$ defined by

$$M_\rho = \left\{ M \mid \sum_{j=2}^{m} a_{kj}^2 \le \rho^2 \text{ for } k = 1, \dots, n \right\}.$$

Note that, for all $\omega \in \Omega$, $\omega \in E_q \Rightarrow \omega \in M_\rho$. That is, considered as events, $E_q \subset M_\rho$.

Proposition 8 **(i)** *For all $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$, $\varepsilon > 0$ and $g : \mathbb{R} \to \mathbb{R}$ continuous with compact support there exists a continuous function*

$$p^g_{\varepsilon,\alpha} : \overline{B}_{m-1}(0;\rho) \to \mathbb{R}^+$$

such that, for all continuous functions $G : \mathbb{R}^{m-1} \to \mathbb{R}$ with support contained in $\overline{B}_{m-1}(0;\rho)$,

$$\mathbf{E}_\alpha \left\{ g(Z_\varepsilon(e_1)) G\left(\frac{\partial Z_\varepsilon}{\partial x_2}(e_1), \dots, \frac{\partial Z_\varepsilon}{\partial x_m}(e_1) \right) \mathbb{I}_{M_\rho} \right\} = \int_{\overline{B}_{m-1}(0;\rho)} G(\zeta) p^g_{\varepsilon,\alpha}(\zeta) d\zeta.$$

Here $\mathbf{E}_\alpha\{\ \}$ denotes conditional expectation under $a_{11} = \alpha_1, \dots, a_{n1} = \alpha_n$. For fixed g, $p^g_{\varepsilon,\alpha}$ is uniformly bounded on $B_{m-1}(0;\rho)$, $0 < \varepsilon \le 1$, $\alpha \in \mathbb{R}^n$.

(ii) *Assume there exists $\bar{k} \le n$ such that $\alpha_{\bar{k}} > \alpha_l$ for $l = 1, \dots, n$, $l \ne \bar{k}$. Then $\lim_{\varepsilon \to 0} p^g_{\varepsilon,\alpha}(\zeta) = q^g_\alpha(\zeta)$ uniformly on $\overline{B}_{m-1}(0;\rho)$ where*

$$q^g_\alpha(\zeta) = g(\alpha_{\bar{k}}) \left(\mathbf{P}(\chi^2_{m-1} \le \rho^2) \right)^{n-1} \prod_{j=2}^{m} \varphi(\zeta_j).$$

Here $\zeta = (\zeta_2, \dots, \zeta_m)$ and χ^2_{m-1} is a random variable having χ^2-distribution with $m-1$ degrees of freedom.

Proposition 9 **(i)** *For fixed $\varepsilon > 0$ and $g : \mathbb{R} \to \mathbb{R}$ continuous with compact support, there exists a continuous function*

$$p^g_\varepsilon : \overline{B}_{m-1}(0;\rho) \to \mathbb{R}^+$$

such that for all continuous functions $G : \mathbb{R}^{m-1} \to \mathbb{R}$ with support contained in $\overline{B}_{m-1}(0;\rho)$ one has

$$\mathbf{E} \left\{ g(Z_\varepsilon(e_1)) G\left(\frac{\partial Z_\varepsilon}{\partial x_2}(e_1), \dots, \frac{\partial Z_\varepsilon}{\partial x_m}(e_1) \right) \mathbb{I}_{M_\rho} \right\} = \int_{\overline{B}_{m-1}(0;\rho)} G(\zeta)\, p^g_\varepsilon(\zeta)\, d\zeta.$$

That is, p_ε^g is the density of the Borel measure defined by

$$B \longmapsto \mathbf{E}\left\{ g(Z_\varepsilon(e_1))\mathbb{I}_{\left\{\left(\frac{\partial Z_\varepsilon}{\partial x_2}(e_1),\dots,\frac{\partial Z_\varepsilon}{\partial x_m}(e_1)\right)\in B\right\}}\mathbb{I}_{M_\rho}\right\}$$

where B is any Borel subset of $\overline{B}_{m-1}(0;\rho)$. For all g, p_ε^g is uniformly bounded on $\overline{B}_{m-1}(0;\rho)$.

(ii) $\lim_{\varepsilon\to 0} p_\varepsilon^g(\zeta) = p^g(\zeta)$ *for each $\zeta \in \overline{B}_{m-1}(0;\rho)$ where*

$$p^g(\zeta) = \frac{\left[\mathbf{P}(\chi_{m-1}^2 \le \rho^2)\right]^{n-1}}{(2\pi)^{\frac{m-1}{2}}} e^{-\frac{1}{2}(\zeta_2^2+\dots+\zeta_m^2)} \int_{-\infty}^{+\infty} g(x) n\varphi(x)(\Phi(x))^{n-1}\, dx.$$

Proposition 10 *Let $g : \mathbb{R} \to \mathbb{R}$, be continuous with compact support. Let $t \in \mathbb{N}$, $2 \le t \le n$ and let (k_j, l_j), $j = 1, \dots, t-1$, be a proper chain. Then, for all $\lambda > 0$,*

$$\begin{aligned}
&\overline{\lim_{\varepsilon\to 0}} \frac{1}{\varepsilon^{t-1}} \int_{\left\{\left|\alpha_{k_j}-\alpha_{l_j}\right|<\lambda\varepsilon \text{ for } j=1,\dots,t-1\right\}} p_{\varepsilon,\alpha}^g(0) \prod_{k=1}^n [\varphi(\alpha_k)d\alpha_k] \\
&\le \frac{(2\lambda)^{t-1}}{(2\pi)^{\frac{m+t-1}{2}}} \int_{-\infty}^{+\infty} g(x)(\Phi(x))^{n-t-1} \\
&\quad \times \left[\sqrt{\frac{2\pi}{t}}\Phi(x\sqrt{t})(n-t)\varphi(x) + t^{\frac{m}{2}} e^{-\frac{1}{2}tx^2}\Phi(x)\right] dx.
\end{aligned}$$

3.5 Proof of Proposition 7

Case $t = 1$ Let $k \le n$ be such that $e_1 \in U_k$. Then, for all $l \neq k$, $a_{k1} > a_{l1}$. In addition, because of the discussion above, almost surely, there exists $\varepsilon' > \varepsilon$ such that $B(e_1;\varepsilon') \subset U_k$.

So (use Proposition 5 (i)), if $\|x - e_1\| < \varepsilon' - \varepsilon$, for $j = 1, \dots, m$,

$$\frac{\partial Z_\varepsilon}{\partial x_j}(x) = a_{kj}$$

which implies that, for $j, \ell = 1, \dots, m$,

$$\frac{\partial^2 Z_\varepsilon}{\partial x_j \partial x_\ell}(e_1) = 0.$$

Therefore, by Proposition 6,

$$(\widetilde{D}^2 Z_\varepsilon)(e_1) = \left(-\frac{\partial Z_\varepsilon}{\partial x_1}(e_1)\right) I_{m-1}$$

where I_{m-1} denotes the identity map on $\mathbb{R}^{m-1}$ and, taking determinants,

$$\det(\widetilde{D}^2 Z_\varepsilon)(e_1) = \left(-\frac{\partial Z_\varepsilon}{\partial x_1}(e_1)\right)^{m-1}.$$

Observe also that Lemma 7 implies that if $Z_\varepsilon(e_1) > \varepsilon\rho$ then $(\widetilde{D}^2 Z_\varepsilon)(e_1)$ is not positive definite.

Let now $\Delta > 0$ be given and denote by $g_\Delta : \mathbb{R} \to \mathbb{R}$ a C^∞ function with support in the interval $[-a-\Delta, 2\Delta]$ and such that $0 \le g_\Delta(z) \le 1$ for all $z \in \mathbb{R}$ and $g_\Delta(z) = 1$ for $z \in [-a, \Delta]$.

Choose $\varepsilon > 0$ small enough so that $\varepsilon\rho < \Delta$.

Then:

$$\begin{aligned} I_1^\bullet(\varepsilon,\delta,a) &\le \mathbf{E}\left\{(|Z_\varepsilon(e_1)| + \varepsilon\rho)^{m-1}\,\mathbb{I}_{\{\|(\widetilde{D}Z_\varepsilon)(e_1)\|<\delta\}\cap\{-a\le Z_\varepsilon(e_1)\le\varepsilon\rho\}\cap E_q}\right\} \\ &\le \mathbf{E}\left\{g_{\Delta,1}(Z_\varepsilon(e_1))G_\delta((\widetilde{D}Z_\varepsilon)(e_1))\mathbb{I}_{M_\rho}\right\} \end{aligned}$$

with

$$g_{\Delta,1}(z) = (|z| + \Delta)^{m-1}\, g_\Delta(z)$$

and $G_\delta : \mathbb{R}^{m-1} \to \mathbb{R}$, C^∞ with support in the ball $\overline{B}_{m-1}(0;\delta+\delta^2)$ and satisfying $0 \le G_\delta(\zeta) \le 1$ for all $\zeta \in \mathbb{R}^{m-1}$ and $G_\delta(\zeta) = 1$ for $\zeta \in \overline{B}_{m-1}(0;\delta)$.

Using Proposition 9 (i) with $g = g_{\Delta,1}$ and $G = G_\delta$ we get

$$\overline{\lim_{\delta\to 0}}\, I_1(\varepsilon,\delta,a) \le \overline{\lim_{\delta\to 0}} \frac{1}{\left|\overline{B}_{m-1}(0;\delta)\right|} \int_{\overline{B}_{m-1}(0;\rho)} G_\delta(\zeta)\, p_\varepsilon^{g_{\Delta,1}}(\zeta)\, d\zeta = p_\varepsilon^{g_{\Delta,1}}(0)$$

the last by the continuity of $p_\varepsilon^{g_{\Delta,1}}$ proved in that proposition. Therefore, using Proposition 9 (ii),

$$\begin{aligned} \overline{\lim_{\varepsilon\to 0}}\,\overline{\lim_{\delta\to 0}}\, I_1(\varepsilon,\delta,a) &\le \overline{\lim_{\varepsilon\to 0}}\, p_\varepsilon^{g_{\Delta,1}}(0) \\ &= \frac{\left[\mathbf{P}(\chi^2_{m-1} \le \rho^2)\right]^{n-1}}{(2\pi)^{\frac{m-1}{2}}} \int_{-\infty}^{+\infty} g_{\Delta,1}(x) n\varphi(x)(\Phi(x))^{n-1} dx. \end{aligned}$$

Since Δ is an arbitrary positive number, using the rough bound $\mathbf{P}(\chi^2_{m-1} \le \rho^2) \le 1$, it follows that

$$\overline{\lim_{\varepsilon\to 0}}\,\overline{\lim_{\delta\to 0}} I_1(\varepsilon,\delta,a) \le \frac{1}{(2\pi)^{\frac{m-1}{2}}} \int_{-a}^{0} |x|^{m-1}\, n\varphi(x)(\Phi(x))^{n-1}\, dx.$$

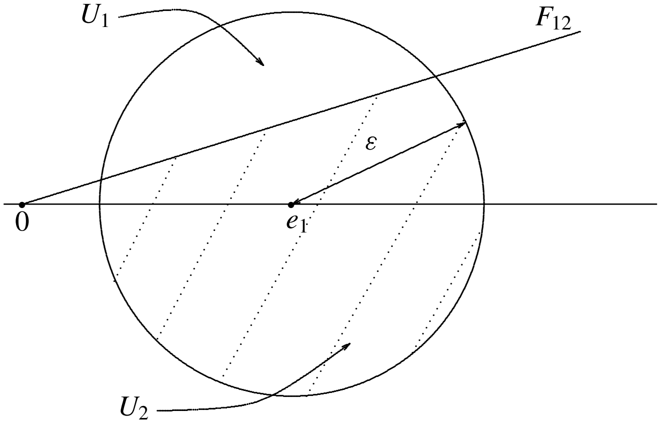
This inequality yields the term corresponding to $t = 1$ in Proposition 7.

Case $t = 2$ The extension of the above result to the case $\theta(\varepsilon) = 2$ requires some additional computations that will also suggest the general procedure for $\theta(\varepsilon) > 2$.

Without loss of generality assume that the U_k's having non-empty intersection with the ball of radius ε centered at the point e_1 are exactly U_1 and U_2, that is:

$$\begin{cases} U_k \cap B(e_1; \varepsilon) \neq \emptyset \text{ for } k = 1, 2 \\ U_k \cap B(e_1; \varepsilon) = \emptyset \text{ for } k = 3, \dots, n. \end{cases} \tag{12}$$

Then $B(e_1; \varepsilon)$ is partitioned into the three sets: $B(e_1; \varepsilon) \cap U_1$, $B(e_1; \varepsilon) \cap U_2$, and $B(e_1; \varepsilon) \cap F_{12}$.



Let $E_{12} = F_{12} \cap \{x_1 = 0\}$. Then, almost surely, E_{12} has dimension $m-2$. Choose an orthonormal basis $\{v_2, \cdots, v_m\}$ of the subspace $\{x_1 = 0\}$ so that $\{v_3, \dots, v_m\} \subseteq E_{12}$ (note that $v_2, \dots, v_m$ depend on ω).

As in the previous case, almost surely, there exists $\varepsilon' > \varepsilon$ such that $B(e_1; \varepsilon') \subset U_1 \cup U_2 \cup F_{12}$. For $y \in B(e_1; \varepsilon')$ write the orthogonal decomposition

$$y = y^* + y^{**} \qquad \text{with} \qquad y^* \in E_{12} \quad \text{and} \quad y^{**} \in E_{12}^{\perp}.$$

Then

$$\langle y, a_1 \rangle = \langle y^*, a_1 \rangle + \langle y^{**}, a_1 \rangle$$

and

$$\langle y, a_2 \rangle = \langle y^*, a_2 \rangle + \langle y^{**}, a_2 \rangle$$

where $\langle y^*, a_1 \rangle = \langle y^*, a_2 \rangle$ since $y^* \in E_{12}$. It follows that

$$Z(y) = \max_{1 \le k \le n} \langle y, a_k \rangle = \max_{1 \le k \le 2} \langle y, a_k \rangle = \langle y^*, a_1 \rangle + \max\{\langle y^{**}, a_1 \rangle, \langle y^{**}, a_2 \rangle\}.$$

For $v \in E_{12}$ and $\delta \in \mathbb{R}$, $|\delta|$ small enough so that $y + \delta v \in B(e_1; \varepsilon')$, one has

$$Z(y + \delta v) - Z(y) = \langle y^* + \delta v, a_1 \rangle - \langle y^*, a_1 \rangle = \delta \langle v, a_1 \rangle$$

which implies

$$\frac{\partial Z}{\partial v_j}(y) = \langle v_j, a_1 \rangle \quad \text{for} \quad j = 3, \dots, m \quad \text{and all} \quad y \in B(e_1; \varepsilon').$$

Using Proposition 5 (i) and the fact that $\varepsilon' > \varepsilon$, it follows that, for $j = 3, \dots, m$ and $x \in B(e_1; \varepsilon' - \varepsilon)$,

$$\frac{\partial Z_\varepsilon}{\partial v_j}(x) = \langle v_j, a_1 \rangle$$

which implies, for $j = 3, \dots, m$, and $\ell = 2, \dots, m$,

$$\frac{\partial^2 Z_\varepsilon}{\partial v_j \partial v_\ell}(e_1) = 0. \tag{13}$$

In the basis $\{e_1, v_2, \dots, v_m\}$ the matrix of $(\widetilde{D^2} Z_\varepsilon)(e_1)$ is diagonal (use Proposition 6, Remark 5, and (13)) and

$$\det(\widetilde{D^2} Z_\varepsilon)(e_1) = \left(-\frac{\partial Z_\varepsilon}{\partial x_1}(e_1) + \frac{\partial^2 Z_\varepsilon}{\partial v_2^2}(e_1) \right) \left(-\frac{\partial Z_\varepsilon}{\partial x_1}(e_1) \right)^{m-2}.$$

In addition, by Proposition 5 (ii),

$$\left| \frac{\partial^2 Z_\varepsilon}{\partial v_j \partial v_\ell} \right| \le \frac{C_1 \sqrt{m}}{\varepsilon} \sum_{k=1}^n \left| \langle a_k, v_j \rangle \right|$$

for $j, \ell = 2, \dots, m$. From here it follows that, for $\omega \in E_q$,

$$\left| \frac{\partial^2 Z_\varepsilon}{\partial v_j \partial v_\ell} \right| \le \frac{C_1 \sqrt{m}}{\varepsilon} \sum_{k=1}^n \|a_k\| \le \frac{C_1 \sqrt{m}}{\varepsilon} n\rho \qquad \text{for } j, \ell = 2, \dots, m. \tag{14}$$

Denote with V_{12} the event (the set of ω's) such that (12) holds. If $\omega \in V_{12}$, the Euclidean distance from e_1 to F_{12} is smaller than ε, that is

$$\frac{|a_{11} - a_{21}|}{\|a_1 - a_2\|} < \varepsilon$$

which implies, if in addition $\omega \in E_q$, that

$$|a_{11} - a_{21}| < 2\varepsilon\rho.$$

Hence, taking into account the $\binom{n}{2}$ forms in which $\theta(\varepsilon) = 2$ may occur, corresponding to all pairs (k, l) in the place of $(1, 2)$, we obtain, with the same notations as for $t = 1$:

$$I_2^{\bullet}(\varepsilon, \delta, a) \leq \binom{n}{2} \mathbf{E} \left\{ \begin{array}{l} (|Z_\varepsilon(e_1)| + \varepsilon\rho)^{m-2} \left(|Z_\varepsilon(e_1)| + \varepsilon\rho + \frac{C_1\sqrt{m}}{\varepsilon} n\rho \right) \\ \mathbb{I}_{\{\|(\widetilde{D}Z_\varepsilon)(e_1)\| < \delta\} \cap \{-a \leq Z_\varepsilon(e_1) \leq \varepsilon\rho\} \cap E_q \cap \{|a_{11} - a_{21}| < 2\varepsilon\rho\}} \end{array} \right\}$$
$$\leq \binom{n}{2} \mathbf{E} \left\{ G_\delta((\widetilde{D}Z_\varepsilon)(e_1)) g_{\Delta,2}(Z_\varepsilon(e_1)) \mathbb{I}_{\{|a_{11} - a_{21}| < 2\varepsilon\rho\}} \mathbb{I}_{M_\rho} \right\}$$

where

$$g_{\Delta,2}(z) = g_{\Delta,2}^{(1)}(z) + \frac{C_1\sqrt{m}}{\varepsilon} n\rho \; g_{\Delta,2}^{(2)}(z)$$

with

$$g_{\Delta,2}^{(1)}(z) = g_\Delta(z) \, (|z| + \Delta)^{m-1}, \quad g_{\Delta,2}^{(2)}(z) = g_\Delta(z) \, (|z| + \Delta)^{m-2}.$$

We apply Proposition 8 (i) to express the right-hand member of the last inequality:

$$I_2^{\bullet}(\varepsilon, \delta, a) \leq \binom{n}{2} \int_{\overline{B}_{m-1}(0;\rho)} G_\delta(\zeta) d\zeta \int_{|\alpha_1 - \alpha_2| < 2\varepsilon\rho} p_{\varepsilon,\alpha}^{g_{\Delta,2}}(\zeta) \prod_{k=1}^{n} [\varphi(\alpha_k) d\alpha_k]$$
$$= \binom{n}{2} \int_{\overline{B}_{m-1}(0;\rho)} G_\delta(\zeta) d\zeta$$
$$\int_{|\alpha_1 - \alpha_2| < 2\varepsilon\rho} \left[p_{\varepsilon,\alpha}^{g_{\Delta,2}^{(1)}}(\zeta) + \frac{C_1\sqrt{m}}{\varepsilon} n\rho \; p_{\varepsilon,\alpha}^{g_{\Delta,2}^{(2)}}(\zeta) \right] \prod_{k=1}^{n} [\varphi(\alpha_k) d\alpha_k].$$

Letting $\delta \to 0$ and applying dominated convergence:

$$\overline{\lim_{\delta \to 0}} I_2(\varepsilon, \delta, a) \leq \binom{n}{2} \int_{|\alpha_1 - \alpha_2| < 2\varepsilon\rho} \left[p_{\varepsilon,\alpha}^{g_{\Delta,2}^{(1)}}(0) + \frac{C_1\sqrt{m}}{\varepsilon} n\rho \; p_{\varepsilon,\alpha}^{g_{\Delta,2}^{(2)}}(0) \right]$$
$$\prod_{k=1}^{n} [\varphi(\alpha_k) d\alpha_k].$$

We want to obtain a bound for the upper limit of the right-hand side above as $\varepsilon \to 0$. Note that this right-hand side splits as the sum of two integrals (corresponding to the two terms $p_{\varepsilon,\alpha}^{g_{\Delta,2}^{(1)}}(0)$ and $\frac{C_1\sqrt{m}}{\varepsilon} n\rho \; p_{\varepsilon,\alpha}^{g_{\Delta,2}^{(2)}}(0)$ in it). To bound the upper limit of each of these integrals we use, for each of them, Proposition 10 with $t = 2$, $k_1 = 1$, $l_1 = 2$, and $\lambda = 2\rho = 2q\sqrt{m}$.

It is easy to see that the upper limit for the first of these integrals, the one corresponding to $p_{\varepsilon,\alpha}^{g_{\Delta,2}^{(1)}}(0)$, is zero. This is due to the term $1/\varepsilon$ in the left-hand side of the inequality in Proposition 10.

Thus, using Proposition 10 on the second integral, we obtain:

$$\overline{\lim_{\varepsilon\to 0}}\,\overline{\lim_{\delta\to 0}} I_2(\varepsilon,\delta,a) \le \binom{n}{2} 4C_1 n m^{3/2} q^2 \frac{1}{(2\pi)^{\frac{m+1}{2}}} \int_{-\infty}^{+\infty} g_{\Delta,2}^{(2)}(x)\,[\Phi(x)]^{n-3}$$
$$\cdot\left[\sqrt{\pi}(n-2)\Phi(x\sqrt{2})\varphi(x) + 2^{\frac{m}{2}} e^{-x^2}\Phi(x)\right] dx$$

and therefore, since Δ may be any positive number,

$$\overline{\lim_{\varepsilon\to 0}}\,\overline{\lim_{\delta\to 0}} I_2(\varepsilon,\delta,a) \le \binom{n}{2} 4C_1 n m^{3/2} q^2 \frac{1}{(2\pi)^{\frac{m+1}{2}}} \int_{-a}^{0} |x|^{m-2}\,[\Phi(x)]^{n-3}$$
$$\cdot\left[\sqrt{\pi}(n-2)\Phi(x\sqrt{2})\varphi(x) + 2^{\frac{m-1}{2}} e^{-x^2}\Phi(x)\right] dx.$$

This inequality yields the term corresponding to $t = 2$ in Proposition 7.

Case $t > 2$ The method follows closely the previous calculations.

Lemma 10 *Assume* $\omega \in \{\theta(\varepsilon) = t\} \cap \left\{(\widetilde{D}^2 Z_\varepsilon)(e_1) \succ 0\right\} \cap E_q$, $t > 2$. *Then there exists an orthonormal basis* $\{v_2, \dots, v_n\}$ *of* $\{x_1 = 0\}$ *such that*

$$\det(\widetilde{D}^2 Z_\varepsilon)(e_1) \le \begin{cases} \left|-\frac{\partial Z_\varepsilon}{\partial x_1}(e_1)\right|^{m-t} \left|-\frac{\partial Z_\varepsilon}{\partial x_1}(e_1) + \frac{C_1\sqrt{m}}{\varepsilon} n\rho\right|^{t-1} & \text{if } t < m \\ \left|-\frac{\partial Z_\varepsilon}{\partial x_1}(e_1) + \frac{C_1\sqrt{m}}{\varepsilon} n\rho\right|^{m-1} & \text{if } t \ge m. \end{cases}$$

Proof. We only prove the case $t < m$. The other case is proved in the same way. Applying Lemma 9, there exists a proper chain $\mathcal{C} = \left\{\{k_1, l_1\}, \{k_2, l_2\}, \dots, \{k_{t-1}, l_{t-1}\}\right\}$ on $\mathcal{X}$ such that $d(e_1, F_{k_j,l_j}) < \varepsilon$ for $j = 1, \dots, t-1$. Let $E_{\mathcal{C}} \subset \mathbb{R}^m$ be the subspace defined by

$$E_{\mathcal{C}} = \left(\bigcap_{j=1}^{t-1} F_{k_j l_j}\right) \cap \{x_1 = 0\}$$

(which plays the role E_{12} had in the case $t = 2$). Almost surely, $E_{\mathcal{C}}$ has dimension $m - t$, and the argument in the case $t = 2$ on the differentiation in directions parallel to $E_{\mathcal{C}}$ can be extended to this case without significant changes. That is, we can construct an orthonormal basis $\{v_2, \dots, v_m\}$ of $\{x_1 = 0\}$, so that $v_{t+1}, \dots, v_m \in E_{\mathcal{C}}$ and, for $j = t+1, \dots, m$, and a certain $\varepsilon' > \varepsilon$,

$$\frac{\partial Z}{\partial v_j} = \text{const} \quad \text{on} \quad B(e_1; \varepsilon' - \varepsilon).$$

Therefore, for $j = t+1, \dots, m$ and $\ell = 2, \dots, m$,

$$\frac{\partial^2 Z_\varepsilon}{\partial v_j \partial v_\ell}(e_1) = 0. \tag{15}$$

Write $(\widetilde{D}^2 Z_\varepsilon)(e_1)$ in the basis $\{v_2, \ldots, v_m\}$ of $\{x_1 = 0\}$. Then, it follows from Proposition 6, Remark 5, and (15) that its matrix has the form

$$\begin{pmatrix} \left[\frac{-\partial Z_\varepsilon}{\partial x_1}(e_1)\right] I_{t-1} + M & 0 \\ 0 & \left[\frac{-\partial Z_\varepsilon}{\partial x_1}(e_1)\right] I_{m-t} \end{pmatrix}$$

where I_k denotes the $k \times k$ identity matrix and M is the $(t-1) \times (t-1)$ matrix given by

$$M = \begin{pmatrix} \frac{\partial^2 Z_e}{\partial v_2^2}(e_1) & \ldots & \ldots & \frac{\partial^2 Z_e}{\partial v_2 \partial v_t}(e_1) \\ \frac{\partial^2 Z_e}{\partial v_2 \partial v_3}(e_1) & \frac{\partial^2 Z_e}{\partial v_3^2}(e_1) & \ldots & \ldots \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 Z_e}{\partial v_2 \partial v_t}(e_1) & \ldots & \ldots & \frac{\partial^2 Z_e}{\partial v_t^2}(e_1) \end{pmatrix}.$$

Then,

$$\det(\widetilde{D}^2 Z_\varepsilon)(e_1) = \left(-\frac{\partial Z_e}{\partial x_1}(e_1)\right)^{m-t} \det\left(\left[\frac{-\partial Z_\varepsilon}{\partial x_1}(e_1)\right] I_{t-1} + M\right).$$

Since $(\widetilde{D}^2 Z_\varepsilon)(e_1)$ is positive definite,

$$\det\left(\left[\frac{-\partial Z_\varepsilon}{\partial x_1}(e_1)\right] I_{t-1} + M\right) \le \left(\frac{\operatorname{tr}\left(\left[\frac{-\partial Z_\varepsilon}{\partial x_1}(e_1)\right] I_{t-1} + M\right)}{t-1}\right)^{t-1}$$

where tr( ) denotes the trace. Therefore

$$\det(\widetilde{D}^2 Z_\varepsilon)(e_1) \le \left|-\frac{\partial Z_\varepsilon}{\partial x_1}(e_1)\right|^{m-t} \left|-\frac{\partial Z_\varepsilon}{\partial x_1}(e_1) + \frac{1}{t-1}\sum_{j=2}^{t} \frac{\partial^2 Z_\varepsilon}{\partial v_j^2}(e_1)\right|^{t-1}.$$

Use now inequality (14) to finish the proof. □

Reasoning as in the case $t = 2$ we obtain for $\overline{\lim_{\delta \to 0}} I_t(\varepsilon, \delta, a)$ an upper bound of the form

$$\sum_{\{\overline{\mathcal{X}}:\#\overline{\mathcal{X}}=t\}} \sum\nolimits^* \int_{\{|\alpha_{k_j}-\alpha_{\ell_j}|<2\varepsilon\rho \text{ for } j=1,\ldots,t-1\}} \left[\sum_{h=1}^{t} \frac{A_h}{\varepsilon^{h-1}} p_{\varepsilon,\alpha}^{g_{\Delta,t}^{(h)}}\right] \prod_{k=1}^{n} [\varphi(\alpha_k) d\alpha_k]$$

where the outer sum is over all subsets $\overline{\mathcal{X}}$ of $\{1, \ldots, n\}$ having t elements, $\sum^*$ extends over all proper chains on $\overline{\mathcal{X}}$, and in each term the notation for the chain is that of Proposition 10. Also, the functions $g_{\Delta,t}^{(h)}$, $h = 1, \ldots, t$, are defined as in the case $t = 2$ and the A_h are real numbers depending on n and m but independent of ε. Again, when taking the limit for $\varepsilon \to 0$, only the term corresponding to $h = t$ in the last summation has a nonzero limit and its corresponding coefficient, it can be shown with the same arguments used in the case $t = 2$, is $A_t = (C_1 n\sqrt{m}\rho)^{t-1}$.

Since the value of the integral above does not depend on $\overline{\mathcal{X}}$ or on the considered chain, we may use Lemma 8 to obtain for $\overline{\lim}_{\delta\to 0} I_t(\varepsilon, \delta, a)$ the bound

$$\binom{n}{t}\frac{(t-1)!}{\pi t}\left(\frac{e^2}{2}\right)^t \int_{\{|\alpha_{k_j}-\alpha_{\ell_j}|<2\varepsilon\rho \text{ for } j=1,\dots,t-1\}} \left[\sum_{h=1}^{t} \frac{A_h}{\varepsilon^{h-1}} p_{\varepsilon,\alpha}^{g_{\Delta,t}^{(h)}}\right] \prod_{k=1}^{n}[\varphi(\alpha_k)d\alpha_k]$$

where now $\mathcal{C} = \left\{k_j, \ell_j\right\}_{j=1,\dots,t-1}$ is an arbitrary —but fixed— proper chain and

$$\overline{\mathcal{X}} = \bigcup_{j=1}^{t-1}\{k_j, \ell_j\}$$

satisfies that $\#\overline{\mathcal{X}} = t$.

Using Proposition 10 we obtain, for $1 \le t \le m$,

$$\overline{\lim_{\varepsilon\to 0}}\ \overline{\lim_{\delta\to 0}} I_t(\varepsilon, \delta, a) \le \binom{n}{t}\frac{(t-1)!}{\pi t}\left(\frac{e^2}{2}\right)^t \left(4C_1 n m^{3/2} q^2\right)^{t-1} \frac{1}{(2\pi)^{\frac{m+t-1}{2}}} \times \int_{-a}^{a} |x|^{m-t}\,(\Phi(x))^{n-t-1} \left[\sqrt{\frac{2\pi}{t}}\Phi(\sqrt{t}x)(n-t)\varphi(x) + t^{\frac{m}{2}} e^{-\frac{t}{2}x^2}\Phi(x)\right] dx.$$

Actually for $t < m$ one can replace $\int_{-a}^{a}$ by $\int_{-a}^{0}$ but this is irrelevant in what follows.

Also note that, if $t > m$ we obtain

$$\overline{\lim_{\varepsilon\to 0}}\ \overline{\lim_{\delta\to 0}}\ I_t(\varepsilon, \delta, a) = 0$$

by Proposition 10 and using the bound

$$\det(\widetilde{D^2} Z_\varepsilon)(e_1) \le \frac{K}{\varepsilon^{m-1}}$$

which is valid for $\omega \in \left\{(\widetilde{D^2} Z_\varepsilon)(e_1) \succ 0\right\} \cap \{|Z_\varepsilon(e_1)| \le 1\} \cap E_q$ and some constant K.

3.6 Putting the pieces together

Recall

$$D(n,m) = \frac{1}{6qm}\left(\frac{e}{m}\right)^{\frac{m}{2}} (\Phi(1))^{n-2}(1+(m\wedge n)^{\frac{m}{2}}) \left(K_0 n^2 m q^2\right)^{(m\wedge n)}.$$

Proposition 11 *For* $0 \le a \le 1$, *we have:*

$$\mathbf{P}(\{|\underline{Z}| < a\} \cap E_q) \le a D(n,m).$$

Proof. First note that if $a \le 1$, the bound for $\overline{\lim}_{\varepsilon\to 0}\ \overline{\lim}_{\delta\to 0}\ I(\varepsilon,\delta,a)$ in Proposition 7 is at most

$$\sum_{t=1}^{m\wedge n}\binom{n}{t}\frac{(t-1)!}{\pi t}\left(\frac{e^2}{2}\right)^t \left(4C_1 n m^{3/2} q^2\right)^{t-1}$$
$$\frac{1}{(2\pi)^{\frac{m+t-1}{2}}} 2a\,(\Phi(1))^{n-t-1}\left(\frac{n-t}{\sqrt{t}} + t^{\frac{m}{2}}\right)$$
$$= ae^2(\Phi(1))^{n-2}(2\pi)^{-\frac{m}{2}} \sum_{t=1}^{m\wedge n}\binom{n}{t}\frac{(t-1)!}{\pi t}$$
$$\times\left(\frac{C_1\sqrt{2}e^2 n m^{3/2} q^2}{\Phi(1)\sqrt{\pi}}\right)^{t-1}\left(\frac{n-t}{\sqrt{t}} + t^{\frac{m}{2}}\right)$$
$$\le ae^2(\Phi(1))^{n-2}(2\pi)^{-\frac{m}{2}} \sum_{t=1}^{m\wedge n}\frac{n^t}{\pi t^2}\left(\frac{C_1\sqrt{2}e^2 n m^{3/2} q^2}{\Phi(1)\sqrt{\pi}}\right)^{t-1}\left(\frac{n-t}{\sqrt{t}} + t^{\frac{m}{2}}\right)$$
$$\le ae^2(\Phi(1))^{n-2}(2\pi)^{-\frac{m}{2}}\frac{n}{\pi}\left(n + (m\wedge n)^{\frac{m}{2}} - 1\right)\sum_{t=1}^{m\wedge n}\left(K_0 n^2 m^{3/2} q^2\right)^{t-1}$$
$$= ae^2(\Phi(1))^{n-2}(2\pi)^{-\frac{m}{2}}\frac{n}{\pi}\left(n + (m\wedge n)^{\frac{m}{2}} - 1\right)\frac{\left(K_0 n^2 m^{3/2} q^2\right)^{(m\wedge n)} - 1}{K_0 n^2 m^{3/2} q^2 - 1}$$

where we have denoted $K_0 = \dfrac{C_1\sqrt{2}e^2}{\sqrt{\pi}\Phi(1)}$. Therefore,

$$\mathbf{P}(\{|\underline{Z}| < a\} \cap E_q) \le \overline{\lim_{\varepsilon\to 0}}\ \mathbf{P}(\{|\underline{Z}_\varepsilon| < a\} \cap E_q)$$
$$\le \sigma_{m-1}(S^{m-1})\overline{\lim_{\varepsilon\to 0}}\ \overline{\lim_{\delta\to 0}} I(\varepsilon,\delta,a)$$
$$\le \frac{2\pi^{\frac{m}{2}}}{\Gamma(\frac{m}{2})} ae^2(\Phi(1))^{n-2}(2\pi)^{-\frac{m}{2}}\frac{n}{\pi}\left(n + (m\wedge n)^{\frac{m}{2}} - 1\right)\frac{\left(K_0 n^2 m^{3/2} q^2\right)^{(m\wedge n)}}{K_0 n^2 m^{3/2} q^2 - 1}.$$

By Stirling's inequality, the last expression is bounded by

$$\frac{2\pi^{\frac{m}{2}}e^{\frac{m}{2}}}{\sqrt{2\pi}\left(\frac{m}{2}\right)^{\frac{m-1}{2}}}ae^2(\Phi(1))^{n-2}(2\pi)^{-\frac{m}{2}}$$
$$\frac{n}{\pi}\left(n+(m\wedge n)^{\frac{m}{2}}-1\right)\frac{\left(K_0n^2m^{3/2}q^2\right)^{(m\wedge n)}}{K_0n^2m^{3/2}q^2-1}$$
$$=a\left(\frac{e}{m}\right)^{\frac{m}{2}}(\Phi(1))^{n-2}e^2n\sqrt{\frac{m}{\pi^3}}\left(n+(m\wedge n)^{\frac{m}{2}}-1\right)\frac{\left(K_0n^2m^{3/2}q^2\right)^{(m\wedge n)}}{K_0n^2m^{3/2}q^2-1}$$
$$\leq\frac{a}{6qm}\left(\frac{e}{m}\right)^{\frac{m}{2}}(\Phi(1))^{n-2}(1+(m\wedge n)^{\frac{m}{2}})\left(K_0n^2m^{3/2}q^2\right)^{(m\wedge n)}$$
$$=aD(n,m).$$

Note in the last inequality we used that $\frac{\sqrt{\frac{1}{\pi^3}}e^2n^2m^{3/2}q}{K_0n^2m^{3/2}q^2-1}\leq\frac{1}{6}$ for all $m,n\geq 1$. □

Proof of the main theorem. Consider first the first moment of $\ln\mathcal{C}(A)$, i.e. $\nu=1$. If $\sqrt{m}qe^{-\alpha}\leq 1$, i.e. $\alpha\geq\alpha_0=\frac{1}{2}\ln m+\ln q$, we may apply Propositions 3 and 11 to obtain

$$\begin{aligned}\mathbf{E}(\ln\mathcal{C}(A))&=\mathbf{E}\left(\ln\frac{1}{|D|}\right)\\&\leq 2\left(\alpha+\int_\alpha^{+\infty}\mathbf{P}\left(\left\{|\underline{Z}|<\sqrt{m}qe^{-x}\right\}\cap E_q\right)dx\right)\\&\leq 2\left(\alpha+D(n,m)\sqrt{m}q\int_\alpha^{+\infty}e^{-x}dx\right)\\&=2\left(\alpha+D(n,m)\sqrt{m}qe^{-\alpha}\right).\end{aligned}$$

This bound is minimized if one chooses

$$\alpha=\begin{cases}\alpha_0 & \text{if } D(n,m)\leq 1\\ \alpha_0+\ln D(n,m) & \text{if } D(n,m)>1.\end{cases}$$

Hence,

$$\mathbf{E}(\ln\mathcal{C}(A))\leq\begin{cases}\ln m+2\ln q+2D(n,m) & \text{if } D(n,m)\leq 1\\ \ln m+2\ln q+2\ln D(n,m)+2 & \text{if } D(n,m)>1.\end{cases}$$

For higher moments we proceed in a similar way. Consider $\nu\in\mathbb{N}$, $\nu\geq 2$. Reasoning as above and using the (easy to check) formula

$$\int_\alpha^{+\infty}e^{-x^{\frac{1}{\nu}}}dx=e^{-\alpha^{\frac{1}{\nu}}}\sum_{k=0}^{\nu-1}\frac{\nu!}{k!}\alpha^{k/\nu}$$

we deduce

$$\mathbf{E}\left((\ln \mathcal{C}(A))^{\nu}\right) \leq 2\left(\alpha + D(n,m)\sqrt{m}qe^{-\alpha^{1/\nu}}\sum_{k=0}^{\nu-1}\frac{\nu!}{k!}\alpha^{k/\nu}\right).$$

If $D(n,m) \leq 1$ taking $\alpha = \alpha_0^{\nu}$ we obtain

$$\begin{aligned}\mathbf{E}\left((\ln \mathcal{C}(A))^{\nu}\right) &\leq 2\left(\alpha_0^{\nu} + D(n,m)\sqrt{m}qe^{-\alpha_0}\sum_{k=0}^{\nu-1}\frac{\nu!}{k!}\alpha_0^{k}\right)\\ &\leq 2\nu!\sum_{k=0}^{\nu}\frac{1}{k!}\left(\frac{1}{2}\ln m + \ln q\right)^{k}.\end{aligned}$$

Else, if $D(n,m) > 1$, taking $\alpha = (\alpha_0 + \ln D(n,m))^{\nu}$ so that $D(n,m)$ $\sqrt{m}qe^{-\alpha^{1/\nu}} = 1$ we get

$$\begin{aligned}\mathbf{E}\left((\ln \mathcal{C}(A))^{\nu}\right) &\leq 2\left(\alpha + \sum_{k=0}^{\nu-1}\frac{\nu!}{k!}\alpha^{k/\nu}\right) = 2\nu!\sum_{k=0}^{\nu}\frac{\alpha^{\frac{k}{\nu}}}{k!}\\ &\leq 2\nu!\sum_{k=0}^{\nu}\frac{1}{k!}\left(\ln D(n,m) + \frac{1}{2}\ln m + \ln q\right)^{k}.\end{aligned}$$

□

4 Some consequences

4.1 On the expected value of $\ln C_R(A)$

A first consequence of the Main Theorem is that it also yields bounds for $C_R(A)$. From Theorem 4 (ii) we have

$$\mathbf{E}(\ln C_R(A)) \leq \mathbf{E}(\ln \mathcal{C}(A)) + \frac{5\ln n}{2} + \frac{\ln m}{2} + 2\ln 2.$$

Therefore, the following result follows.

Theorem 6 *Let A be an* $n \times m$ *Gaussian matrix. Then*

$$\mathbf{E}(\ln C_R(A)) \leq \begin{cases} 2\ln q + 2D(n,m) + \frac{5\ln n}{2} + \frac{3\ln m}{2} + 2\ln 2 & \textit{if } D(n,m) \leq 1\\ 2\ln q + 2\ln D(n,m) + \frac{5\ln n}{2} + \frac{3\ln m}{2} + 2\ln 2 + 2 & \textit{if } D(n,m) > 1.\end{cases}$$

□

Note that in the most interesting case, when $n \gg m$, the bound for $\mathbf{E}(\ln C_R(A))$ is significantly worse than that for $\mathbf{E}(\ln \mathcal{C}(A))$. We don't know whether this is inherent to $C_R(A)$ or a better bound for $\mathbf{E}(\ln C_R(A))$ can be obtained (for instance, improving the first inequality in this section).

4.2 Running time and round-off analysis of an interior-point algorithm

Recall, in §1.2 we considered the problem

> Given a $n \times m$ real matrix A, decide which of (1) or (2) is strictly feasible and return a strict solution for it

where (1) and (2) are the systems

$$Ax \leq 0, \quad x \neq 0$$

and

$$A^{\mathrm{T}} y = 0, \quad y \geq 0, \ y \neq 0$$

respectively. We remarked there that one of these systems is strictly feasible if and only if the other is not feasible.

In [Cucker and Peña 2001] a finite (but variable) precision algorithm for solving the problem above is described. Due to the finite precision assumption, if the system having a strict solution is (2), there is no hope to exactly compute one such solution y since the set of solutions is thin in $\mathbb{R}^n$ (i.e., has empty interior). One can however (and the algorithm in [Cucker and Peña 2001] does) compute good approximations.

Definition 1 *Let* $\gamma \in (0, 1)$. *A point* $y \in \mathbb{R}^n$ *is a* γ-forward solution *of the system* $Ay = 0$, $y \geq 0$, $y \neq 0$, *if there exists* $\bar{y} \in \mathbb{R}^n$, $\bar{y} \neq 0$, *such that*

$$A\bar{y} = 0, \quad \bar{y} \geq 0$$

and, for $i = 1, \ldots, n$,

$$|y_i - \bar{y}_i| \leq \gamma y_i.$$

The point $\bar{y}$ *is said to be an* associated solution *for* y.

The *algebraic complexity* of an algorithm is the number of arithmetic operations performed by the algorithm. If the precision is fixed, this is a good measure of the amount of work realized by the algorithm. If the precision is variable the cost of each operation needs to be considered as well. In §1.1 we defined the *round-off unit* of an algorithm to be a number $u \in \mathbb{R}$, $0 < u < 1$, such that during the execution of the algorithm real numbers x are systematically replaced by approximations $r(x)$ satisfying $|r(x) - x| \leq u|x|$. Roughly, $|\log u|$ corresponds with the number of bits (digits if the log is in base 10) of the mantissa in the floating-point representation of $r(x)$. The cost of an arithmetic operation with machine precision u is quadratic in $|\log u|$. The *total cost* of an algorithm with variable precision is the addition of the costs of all the operations performed by the algorithm.

The main result of [Cucker and Peña 2001] can be stated as follows.

Theorem 7 *There exists a round-off algorithm which, with input a matrix $A \in \mathbb{R}^{m\times n}$ and a number $\gamma \in (0,1)$, finds either a strict γ-forward solution $y \in \mathbb{R}^n$ of $A^{\mathrm{T}}y = 0$, $y > 0$, or a solution $x \in \mathbb{R}^m$ of the system $Ax \le 0$, $x \ne 0$. The machine precision varies during the execution of the algorithm. The finest required precision is*

$$u = \frac{1}{\mathbf{c}(m+n)^{12} C_R(A)^2},$$

where $\mathbf{c}$ *is a universal constant. The algebraic complexity of the algorithm is bounded by $\mathcal{O}\left((m+n)^{3.5}(\ln(m+n) + \ln C_R(A) + |\ln\gamma|)\right)$. The total cost of the algorithm is bounded by*

$$\mathcal{O}\left((m+n)^{3.5}(\ln(m+n) + \ln C_R(A) + |\ln\gamma|)^3\right).$$

The bounds above are in case (2) is strictly feasible. If (1) is, then similar bounds hold with the $|\ln\gamma|$ terms removed. □

The complexity bounds in Theorem 7 cannot be written as a function of m and n solely due to the unboundedness of $C_R(A)$. One can, however, eliminate the ocurrences of $\ln C_R(A)$ in these bounds at the cost of trading worst-case by average-case complexity. This is done using Theorem 6.

5 Proof of Propositions 8, 9 and 10

5.1 Proof of Proposition 8

We introduce the following notations:

$$H^{(1)}_{\varepsilon,\alpha}(M) = Z_\varepsilon(e_1),$$
$$H^{(h)}_{\varepsilon,\alpha}(M) = \frac{\partial Z_\varepsilon}{\partial x_h}(e_1) \quad \text{for } h = 2, \ldots, m.$$

We are writing the matrix $A = ((a_{kj}))_{\substack{k=1,\ldots,n\\ j=1,\ldots,m}}$ as

$$A = \begin{pmatrix} \alpha_1 & a_{12} & \ldots & a_{1m} \\ \alpha_2 & a_{22} & \ldots & a_{2m} \\ & & \ddots & \\ \alpha_n & a_{n2} & \ldots & a_{nm} \end{pmatrix} = (\alpha, M)$$

that is, we fix $a_{11}, \ldots, a_{n1}$ equal to $\alpha_1, \ldots, \alpha_n$ respectively and M is the $n \times (m-1)$ matrix $M = ((a_{kj}))_{\substack{k=1,\ldots,n\\ j=2,\ldots,m}}$, and we consider the functions

$Z_\varepsilon(e_1)$ and $\frac{\partial Z_\varepsilon}{\partial x_h}(e_1)$ as functions of M for each fixed $\alpha \in \mathbb{R}^n$. We also identify the set of these matrices M with $\mathbb{R}^{n(m-1)}$ and denote

$$\begin{aligned} L_{\varepsilon,\alpha} : \mathbb{R}^{n(m-1)} &\to \mathbb{R}^{m-1} \\ M &\mapsto \left(H^{(2)}_{\varepsilon,\alpha}(M), \dots , H^{(m)}_{\varepsilon,\alpha}(M)\right). \end{aligned}$$

Also, $\bar{\varphi}(M)$ denotes the density

$$\bar{\varphi}(M) = \prod_{k=1}^{n} \prod_{j=2}^{m} \varphi(a_{kj})$$

in $\mathbb{R}^{n(m-1)}$.

The proof of (i) is divided in two parts. Firstly we exhibit the function $p^g_{\varepsilon,\alpha}$ that verifies, for any continuous $G : \mathbb{R}^{m-1} \to \mathbb{R}$ with support contained in $\overline{B}_{m-1}(0; \rho)$

$$\begin{aligned} &\mathbf{E}_\alpha \left\{ g(Z_\varepsilon(e_1)) G\left(\frac{\partial Z_\varepsilon}{\partial x_2}(e_1), \dots , \frac{\partial Z_\varepsilon}{\partial x_m}(e_1)\right) \mathbb{I}_{M_\rho} \right\} \\ &\quad = \int_{\overline{B}_{m-1}(0;\rho)} G(\zeta)\, p^g_{\varepsilon,\alpha}(\zeta)\, d\zeta. \end{aligned} \tag{16}$$

Secondly, we show the continuity of $p^g_{\varepsilon,\alpha}$.

For the first part we consider the Borel measure $\mu_{\varepsilon,\alpha}$ in $M_\rho \subset \mathbb{R}^{n(m-1)}$ given by

$$\mu_{\varepsilon,\alpha}(dM) = g\left(H^{(1)}_{\varepsilon,\alpha}(M)\right) \bar{\varphi}(M) dM.$$

One can express the left-hand side of (16) as

$$\begin{aligned} &\mathbf{E}_\alpha \left\{ g(Z_\varepsilon(e_1)) G\left(\frac{\partial Z_\varepsilon}{\partial x_2}(e_1), \dots , \frac{\partial Z_\varepsilon}{\partial x_m}(e_1)\right) \mathbb{I}_{M_\rho} \right\} \\ &= \int_{\mathbb{R}^{n(m-1)}} g\left(H^{(1)}_{\varepsilon,\alpha}(M)\right) G\left(L_{\varepsilon,\alpha}(M)\right) \mathbb{I}_{M_\rho}(M) \bar{\varphi}(M) dM \\ &= \int_{M_\rho} G\left(L_{\varepsilon,\alpha}(M)\right) \mu_{\varepsilon,\alpha}(dM). \end{aligned}$$

Then, according to the coarea formula, if $L_{\varepsilon,\alpha} : \mathbb{R}^{n(m-1)} \to \mathbb{R}^{m-1}$ is C^1 and $DL_{\varepsilon,\alpha}(M)$ has full rank at every point $M \in M_\rho$, the image measure of $\mu_{\varepsilon,\alpha}$ by $L_{\varepsilon,\alpha}$ has the density $p^g_{\varepsilon,\alpha}$ given by

$$\begin{aligned} p^g_{\varepsilon,\alpha}(\zeta) = &\int_{L^{-1}_{\varepsilon,\alpha}(\{\zeta\}) \cap M_\rho} g\left(H^{(1)}_{\varepsilon,\alpha}(M)\right) \bar{\varphi}(M) \\ &\cdot \left[\det(DL_{\varepsilon,\alpha}(M))((DL_{\varepsilon,\alpha}(M))^{\mathrm{T}})\right]^{-1/2} \sigma_{\zeta,\varepsilon,\alpha}(dM). \end{aligned} \tag{17}$$

For each $\zeta \in \mathbb{R}^{m-1}$, $L_{\varepsilon,\alpha}^{-1}(\{\zeta\}) \cap M_\rho$ is a compact C^1 manifold of dimension $(n-1)\times(m-1)$ embedded in $\mathbb{R}^{n(m-1)}$ and $\sigma_{\zeta,\varepsilon,\alpha}(dM)$ denotes the standard geometric measure on it.

Thus, to prove the first part of (i) it is enough to show that $L_{\varepsilon,\alpha}$ is C^1 and that $DL_{\varepsilon,\alpha}(M)$ has full rank at every point $M \in M_\rho$. With this aim let us compute the partial derivatives of the coordinates of $L_{\varepsilon,\alpha}$ with respect to the coordinates of M in $\mathbb{R}^{n(m-1)}$.

Lemma 11 *For $k = 1, \dots, n$ and $j = 1, \dots, m$,*

$$\frac{\partial Z_\varepsilon}{\partial a_{kj}}(x) = \int_{\mathbb{R}^m} \psi_\varepsilon(x-y) y_j \mathbb{I}_{U_k}(y)dy.$$

For $h, j = 2, \dots, m$ and $k = 1, \dots, n$,

$$\frac{\partial H_{\varepsilon,\alpha}^{(h)}}{\partial a_{kj}} = c_m \int_{\mathbb{R}^m} \psi'(\|w\|)\frac{w_h w_j}{\|w\|}\mathbb{I}_{U_k}(e_1 - \varepsilon w)dw.$$

Proof. It is useful to start with

$$\frac{\partial Z_\varepsilon}{\partial a_{kj}}(x) = \int_{\mathbb{R}^m} \psi_\varepsilon(x-y)\frac{\partial Z}{\partial a_{kj}}(y)dy = \int_{\mathbb{R}^m} \psi_\varepsilon(x-y) y_j \mathbb{I}_{U_k}(y)dy \tag{18}$$

which is valid for $k = 1, \dots, n$ and $j = 1, \dots, m$. The first equality follows from the fact that $a_{kj} \mapsto Z(y)$ is absolutely continuous, toghether with Fubini's Theorem. The second follows from the simple computation

$$\frac{\partial Z}{\partial a_{kj}}(y) = y_j \mathbb{I}_{U_k}(y)$$

which holds for almost all $y \in \mathbb{R}^m$.

For the computation of

$$\frac{\partial H_{\varepsilon,\alpha}^{(h)}}{\partial a_{kj}} = \frac{\partial}{\partial a_{kj}} \frac{\partial Z_\varepsilon}{\partial x_h}\Bigg|_{x=e_1} \qquad h, j = 2, \dots, m \text{ and } k = 1, \dots, n$$

it is convenient to reverse the order of differentiation and to compute

$$\frac{\partial}{\partial x_h}\frac{\partial Z_\varepsilon}{\partial a_{kj}}\Bigg|_{x=e_1}.$$

One must check that

$$\frac{\partial}{\partial x_h}\frac{\partial Z_\varepsilon}{\partial a_{kj}}, \ \frac{\partial Z_\varepsilon}{\partial x_h}, \ \text{and } \frac{\partial Z_\varepsilon}{\partial a_{kj}}$$

are continuous functions of the pair (x_h, a_{kj}) which follows from the foregoing formulae ([Courant 1988] p.56).

Now, using (18),

$$\frac{\partial}{\partial x_h}\frac{\partial Z_\varepsilon}{\partial a_{kj}}\bigg|_{x=e_1} = \frac{\partial}{\partial x_h}\int_{\mathbb{R}^m} \psi_\varepsilon(x-y)y_j\mathbb{I}_{U_k}(y)dy\,\bigg|_{x=e_1}$$
$$= \int_{\mathbb{R}^m} \frac{\partial}{\partial x_h}\psi_\varepsilon(x-y)y_j\mathbb{I}_{U_k}(y)dy\,\bigg|_{x=e_1}.$$

So, for $h, j = 2, \dots, m$ and $k = 1, \dots, n$,

$$\frac{\partial H^{(h)}_{\varepsilon,\alpha}}{\partial a_{kj}} = \frac{\partial}{\partial x_h}\frac{\partial Z_\varepsilon}{\partial a_{kj}}\bigg|_{x=e_1}$$
$$= \frac{c_m}{\varepsilon^{m+1}}\int_{\mathbb{R}^m} \psi'\left(\frac{\|e_1-y\|}{\varepsilon}\right)\left(\frac{-y_h}{\|e_1-y\|}\right) y_j\mathbb{I}_{U_k}(y)\,dy$$
$$= c_m\int_{\mathbb{R}^m} \psi'(\|w\|)\frac{w_h w_j}{\|w\|}\mathbb{I}_{U_k}(e_1-\varepsilon w)dw$$

where the last equality comes from the change of variables

$$\frac{1-y_1}{\varepsilon} = w_1,\ \frac{-y_2}{\varepsilon} = w_2,\ \dots,\ \frac{-y_m}{\varepsilon} = w_m. \qquad \square$$

The continuity of $\frac{\partial H^{(h)}_{\varepsilon,\alpha}}{\partial a_{kj}}$ as a function of M is now immediate.

Lemma 12 *For all values of $\varepsilon > 0$, α and M one has*

$$\det\left[(DL_{\varepsilon,\alpha})(M)\left((DL_{\varepsilon,\alpha})(M)\right)^{\mathrm{T}}\right] \geq \left(\frac{1}{n}\right)^{m-1}.$$

In particular, $(DL_{\varepsilon,\alpha})(M)$ has full rank.

Proof. Write $\theta = (\theta_2, \dots, \theta_m) \in \mathbb{R}^{m-1}$ and consider the quadratic form

$$(19)\quad \theta(DL_{\varepsilon,\alpha})(M)\left[(DL_{\varepsilon,\alpha})(M)\right]^{\mathrm{T}}\theta^{\mathrm{T}} = \sum_{h,h'=2}^{m}\theta_h\theta_{h'}\sum_{k=1}^{n}\sum_{j=2}^{m}\frac{\partial H^{(h)}_{\varepsilon,\alpha}}{\partial a_{kj}}\frac{\partial H^{(h')}_{\varepsilon,\alpha}}{\partial a_{kj}}$$
$$= \sum_{k=1}^{n}\sum_{j=2}^{m}\left(\sum_{h=2}^{m}\theta_h\frac{\partial H^{(h)}_{\varepsilon,\alpha}}{\partial a_{kj}}\right)^2$$
$$\geq \sum_{j=2}^{m}\frac{1}{n}\left(\sum_{k=1}^{n}\sum_{h=2}^{m}\theta_h\frac{\partial H^{(h)}_{\varepsilon,\alpha}}{\partial a_{kj}}\right)^2$$

the last by the Cauchy-Schwartz inequality. Using Lemma 11

$$\sum_{k=1}^{n}\sum_{h=2}^{m}\theta_h\frac{\partial H^{(h)}_{\varepsilon,\alpha}}{\partial a_{kj}} = c_m\sum_{h=2}^{m}\theta_h\int_{\mathbb{R}^m}\psi'(\|w\|)\frac{w_h w_j}{\|w\|}dw$$

because $\sum_{k=1}^{n}\mathbb{I}_{U_k}(y) = 1$ for almost all $y \in \mathbb{R}^m$. The last integral is easily evaluated:

- If $h \neq j$ it is zero by a symmetry argument.
- If $h = j$ integration by parts and the compactness of the support of ψ gives

$$\begin{aligned}\int_{\mathbb{R}^m} \psi'(\|w\|) \frac{w_j^2}{\|w\|} dw &= \int_{\mathbb{R}^{m-1}} \prod_{\substack{\ell=1\\ \ell\neq j}}^m dw_\ell \int_{\mathbb{R}} \psi'(\|w\|) \frac{w_j^2}{\|w\|} dw_j \\ &= \int_{\mathbb{R}^{m-1}} \prod_{\substack{\ell=1\\ \ell\neq j}}^m dw_\ell \left(- \int_{\mathbb{R}} \psi(\|w\|) dw_j \right) \\ &= - \int_{\mathbb{R}^m} \psi(\|w\|) dw \\ &= -\frac{1}{c_m}.\end{aligned}$$

Hence, (19) implies that, for all $\theta \in \mathbb{R}^{m-1}$,

$$\theta (DL_{\varepsilon,\alpha})(M) \left[(DL_{\varepsilon,\alpha})(M) \right]^{\mathrm{T}} \theta^{\mathrm{T}} \geq \frac{1}{n} \sum_{j=2}^m \theta_j^2$$

which implies that the determinant

$$\det \left[(DL_{\varepsilon,\alpha})(M) \left[(DL_{\varepsilon,\alpha})(M) \right]^{\mathrm{T}} \right] \geq \left(\frac{1}{n} \right)^{m-1}$$

for any values of ε, α and M. □

It now remains to prove the second part of (i) namely, the continuity of $p^g_{\varepsilon,\alpha}$.

Let us denote by $K_{\varepsilon,\alpha}$ the (compact) image of M_ρ by the function $L_{\varepsilon,\alpha}$, i.e. $K_{\varepsilon,\alpha} = L_{\varepsilon,\alpha}(M_\rho)$.

Lemma 13 *For all* $\alpha \in \mathbb{R}^n$,

$$\overline{B}_{m-1}(0; \rho) \subseteq \bigcap_{\varepsilon > 0} K_{\varepsilon,\alpha}.$$

Proof. Let $\zeta = (\zeta_2, \ldots, \zeta_m)$ be such that $\|\zeta\| \leq \rho$. Choose, $M = (a_{kj})$, $k = 1, \ldots, n$, $j = 2, \ldots, m$ such that $a_{kj} = \zeta_j$ for $k = 1, \ldots, n$, $j = 2, \ldots, m$. Then, for all $\varepsilon > 0$ and $h = 2, \ldots, m$,

$$H^{(h)}_{\varepsilon,\alpha}(M) = \zeta_h \int_{\mathbb{R}^m} \psi_\varepsilon(e_1 - y) \left(\sum_{k=1}^n \mathbb{I}_{U_k}(y) \right) dy = \zeta_h$$

i.e. $L_{\varepsilon,\alpha}(M) = \zeta$. □

Now, a close look at the implicit function theorem shows that for fixed g, ε and α the function

$$\zeta \mapsto p^g_{\varepsilon,\alpha}(\zeta)$$

is continuous on the set $K_{\varepsilon,\alpha}$. This is due to the C^1 character of $L_{\varepsilon,\alpha}$, the lower bound in Lemma 12, the continuity of the function $M \mapsto \bar{\varphi}(M)$, and the compactness of $L^{-1}_{\varepsilon,\alpha}(\{\zeta\}) \cap M_\rho$ for $\zeta \in K_{\varepsilon,\alpha}$. In addition, for fixed g, the value of $p^g_{\varepsilon,\alpha}(\zeta)$ is bounded by a constant which does not depend on $\varepsilon, \alpha, \zeta$ when $0 < \varepsilon \leq 1$, $\alpha \in \mathbb{R}^n$ and $\zeta \in \overline{B}_{m-1}(0;\rho)$. This follows from ((17)) and the uniform boundedness of $\frac{\partial H^{(h)}_{\varepsilon,\alpha}}{\partial a_{kj}}$ which is implied by the formulae in Lemma 11.

This finishes the proof of Proposition 8 (i).

We now focus on part (ii).

Lemma 14 *Let $\alpha \in \mathbb{R}^n$ satisfy*

$$\alpha_{\overline{k}} > \alpha_\ell \quad \textit{for } \ell = 1, \dots, n, \ \ell \neq \bar{k}$$

for some $\overline{k}$. Then,

(i)

$$H^{(1)}_{\varepsilon,\alpha}(M) \underset{\varepsilon\to 0}{\to} \alpha_{\overline{k}}$$

and, for $h, j = 2, \dots, m$ and $k = 1, \dots, n$,

(ii)

$$H^{(h)}_{\varepsilon,\alpha}(M) \underset{\varepsilon\to 0}{\to} a_{\overline{k}h}$$

and

(iii)

$$\frac{\partial H^{(h)}_{\varepsilon,\alpha}}{\partial a_{kj}}(M) \underset{\varepsilon\to 0}{\to} -\delta_{jh}\delta_{k\overline{k}}.$$

The convergence in (i), (ii) and (iii) is uniform on $M \in M_\rho$.

Proof. We only prove (i). For (ii) and (iii) the same arguments, *mutatis mutandis*, can be used.

To prove (i) note that

$$\begin{aligned} H^{(1)}_{\varepsilon,\alpha}(M) &= \int_{\mathbb{R}^m} \psi_\varepsilon(e_1 - y) Z(y) dy \\ &= \sum_{k=1}^{n} \int_{\mathbb{R}^m} \psi_\varepsilon(e_1 - y) \left(\alpha_k y_1 + \sum_{j=2}^{m} a_{kj} y_j \right) \\ &\quad \mathbb{I}_{\left\{\alpha_k y_1 + \sum_{j=2}^m a_{kj} y_j > \alpha_\ell y_1 + \sum_{j=2}^m a_{\ell j} y_j \text{ for all } \ell \neq k\right\}}. \end{aligned} \tag{20}$$

Denote $\eta = \alpha_{\overline{k}} - \max\{\alpha_\ell : 1 \le \ell \le n, \ell \ne \overline{k}\}$. If $M \in M_\rho$ and one chooses

$$\varepsilon < \frac{\eta}{\eta + 2\rho}$$

then, for all $y \in \mathbb{R}^m$ such that $\|e_1 - y\| \le \varepsilon$ one has that, for $\ell \ne \overline{k}$,

$$(\alpha_{\overline{k}} - \alpha_\ell)y_1 + \sum_{j=2}^{m}(a_{\overline{k}j} - a_{\ell j})y_j \ge \eta(1-\varepsilon) - 2\rho\varepsilon > 0$$

which means that in the integrand in the right-hand side of (20) the indicator function is equal to 1 if $k = \overline{k}$ and vanishes for $k \ne \overline{k}$.

So, if $\varepsilon < \frac{\eta}{\eta+2\rho}$,

$$H^{(1)}_{\varepsilon,\alpha}(M) = \int_{\mathbb{R}^m} \psi_\varepsilon(e_1 - y)\left(\alpha_{\overline{k}} y_1 + \sum_{j=2}^{m} a_{\overline{k}j} y_j\right) dy$$

for all $M \in M_\rho$ and

$$\begin{aligned}\left|H^{(1)}_{\varepsilon,\alpha}(M) - \alpha_{\overline{k}}\right| &= \left|\int_{\mathbb{R}^m} \psi_\varepsilon(e_1 - y)\left(\alpha_{\overline{k}}(y_1 - 1) + \sum_{j=2}^{m} a_{\overline{k}j} y_j\right) dy\right| \\ &\le \left|\int_{B(e_1;\varepsilon)} \psi_\varepsilon(e_1 - y) dy\right| \sup_{y \in B(e_1;\varepsilon)} \left|\alpha_{\overline{k}}(y_1 - 1) + \sum_{j=2}^{m} a_{\overline{k}j} y_j\right| \\ &\le \varepsilon(|\alpha_{\overline{k}}| + \rho).\end{aligned}$$

The remainder is plain. □

Remark 6 We have already used the uniform boundedness of the functions $\frac{\partial H^{(h)}_{\varepsilon,\alpha}}{\partial a_{kj}}$ to prove the uniform boundedness of $p^g_{\varepsilon,\alpha}$. We note now that it follows from Lemma 14 that the functions $\frac{\partial H^{(h)}_{\varepsilon,\alpha}}{\partial a_{kj}}$ are not uniformly continuous in α. This produces some technical problems we will have to deal with in our proof.

Lemma 15 *Let $\alpha \in \mathbb{R}^n$ satisfy the hypotesis of Lemma 14. Then*

$$\bigcap_{\varepsilon>0} K_{\varepsilon,\alpha} = \overline{B}_{m-1}(0;\rho).$$

Proof. We have already shown one of the inclusions in Lemma 13. To show the other, note that Lemma 14 (ii) implies that

$$\sup_{M \in M_\rho} \sum_{j=2}^{m} \left(H^{(j)}_{\varepsilon,\alpha}(M) - a_{\overline{k}j}\right)^2 \underset{\varepsilon \to 0}{\longrightarrow} 0$$

so that

$$\overline{\lim_{\varepsilon \to 0}} \sup_{M \in M_\rho} \|L_{\varepsilon,\alpha}(M)\| \leq \left(\sum_{j=2}^{m} a_{\overline{k}j}^2 \right)^{1/2} \leq \rho .$$

From this inequality it follows that, for every $\eta > 0$, if $\varepsilon > 0$ is small enough, $K_{\varepsilon,\alpha} \subseteq \overline{B}_{m-1}(0; \rho + \eta)$. This finishes the proof. □

End of the Proof of Proposition 8 (ii). Parts (i), (ii) and (iii) in Lemma 14 and the argument used to show that $p^g_{\varepsilon,\alpha}$ is continuous show that, actually, for fixed g and α the set of functions

$$\{p^g_{\varepsilon,\alpha}\}_{0<\varepsilon\leq 1} \tag{21}$$

defined on $\overline{B}_{m-1}(0; \rho)$ is equibounded and equicontinuous. Using this, we next show that for fixed g and α, the functions $p^g_{\varepsilon,\alpha}$ converge uniformly on $\overline{B}_{m-1}(0; \rho)$, as $\varepsilon \to 0$, to a continuous function which we will denote by q^g_α.

To do so, by the Arzela-Ascoli theorem applied to the family of functions (21) on $\overline{B}_{m-1}(0; \rho)$, it suffices to prove that, for all sequence $\{\varepsilon_\nu\}$, $0 < \varepsilon_\nu < 1$, if $\varepsilon_\upsilon \to 0$ and

$$\{p^g_{\varepsilon_\upsilon,\alpha}\}_{\upsilon=1,2,\ldots}$$

is a uniformly convergent sequence of functions, then the limit function does not depend on the sequence $\{\varepsilon_\upsilon\}$. Let $\{\varepsilon_\upsilon\}$ be such a sequence and put

$$q(\zeta) = \lim_{\upsilon\to\infty} \{p^g_{\varepsilon_\upsilon,\alpha}\}(\zeta), \tag{22}$$

take any function $G : \mathbb{R}^{m-1} \to \mathbb{R}$ continuous with support in $\overline{B}_{m-1}(0; \rho)$, replace ε by ε_υ in (16), and let $\upsilon \to \infty$. The limit on the right-hand side of (16) is equal to

$$\int_{\mathbb{R}^{m-1}} G(\zeta) q(\zeta) d\zeta .$$

For the left-hand side, use Lemma 14 (i) and (ii) plus the fact that g and G are continuous and bounded to obtain the limit

$$\begin{aligned} &\mathbf{E}_\alpha \left\{ g(\alpha_{\overline{k}}) G(a_{\overline{k}2}, \ldots, a_{\overline{k}m}) \mathbb{I}_{M_\rho} \right\} \\ &= g(\alpha_{\overline{k}}) \left[\mathbf{P}(\chi^2_{m-1} \leq \rho^2) \right]^{n-1} \mathbf{E} \left\{ G(a_{\overline{k}2}, \ldots, a_{\overline{k}m})\, \mathbb{I}_{\left\{\sum_{j=2}^m a_{\overline{k}j}^2 \leq \rho^2\right\}} \right\} \\ &= g(\alpha_{\overline{k}}) \left[\mathbf{P}(\chi^2_{m-1} \leq \rho^2) \right]^{n-1} \mathbf{E} \left\{ G(a_{\overline{k}2}, \ldots, a_{\overline{k}m}) \right\} \end{aligned} \tag{23}$$

since the random variables a_k, $k = 1, \ldots, n$, are independent and the support of G is contained in $\overline{B}_{m-1}(0; \rho)$.

We conclude that

$$g(\alpha_{\overline{k}})\left[\mathbf{P}(\chi^2_{m-1} \leq \rho^2)\right]^{n-1} \int_{\mathbb{R}^{m-1}} G(\zeta) \prod_{j=2}^{m} \varphi(\zeta_j)\, d\zeta = \int_{\mathbb{R}^{m-1}} G(\zeta) q(\zeta) d\zeta.$$

So,

$$q(\zeta) = g(\alpha_{\overline{k}})\left[\mathbf{P}(\chi^2_{m-1} \leq \rho^2)\right]^{n-1} \prod_{j=2}^{m} \varphi(\zeta_j) \tag{24}$$

almost everywhere in $\overline{B}_{m-1}(0;\rho)$ since both sides are continuous on this ball. This proves that the limit q does not depend on the sequence $\{\varepsilon_\upsilon\}$ and at the same time identifies this limit, that we have denoted

$$q^g_\alpha(\zeta)$$

as the right-hand member of (24). □

5.2 Proof of Proposition 9

Integrate both sides of the equality in Proposition 8 (i) with respect to the distribution of $(a_{11}, \ldots, a_{n1})$ a random vector which is independent of the event M_ρ. We obtain

$$\mathbf{E}\left\{g(Z_\varepsilon(e_1))G\left(\frac{\partial Z_\varepsilon}{\partial x_2}(e_1), \ldots, \frac{\partial Z_\varepsilon}{\partial x_m}(e_1)\right)\, \mathbb{I}_{M_\rho}\right\}$$
$$= \int_{\overline{B}_{m-1}(0;\rho)} G(\zeta) p^g_\varepsilon(\zeta) d\zeta$$

with

$$p^g_\varepsilon(\zeta) = \int_{\mathbb{R}^n} p^g_{\varepsilon,\alpha}(\zeta)\left(\prod_{k=1}^{n} \varphi(\alpha_k)\right) d\alpha. \tag{25}$$

The uniform boundedness of p^g_ε on $\overline{B}_{m-1}(0;\rho)$ for $\varepsilon > 0$ follows from the uniform boundedness of $p^g_{\varepsilon,\alpha}$. To prove part (ii) apply dominated convergence as $\varepsilon \to 0$ in (25). We have, using Proposition 8 (ii),

$$
\begin{aligned}
p_\varepsilon^g(\zeta) \longrightarrow & \int_{\mathbb{R}^n} q_\alpha^g(\zeta)\left(\prod_{k=1}^n \varphi(\alpha_k)\right) d\alpha \\
= & \int_{\mathbb{R}^n}\left(\prod_{k=1}^n \varphi(\alpha_k)\right) g\left(\max_{1\le k\le n} \alpha_k\right)\left[\mathbf{P}(\chi_{m-1}^2 \le \rho^2)\right]^{n-1} \prod_{j=2}^m \varphi(\zeta_j) d\alpha \\
= & \ \mathbf{E}\left\{g\left(\max_{1\le k\le n} a_{k1}\right)\right\}\left[\mathbf{P}(\chi_{m-1}^2 \le \rho^2)\right]^{n-1} \prod_{j=2}^m \varphi(\zeta_j) \\
= & \left(\int_{\mathbb{R}} g(r) n\varphi(r)(\Phi(r))^{n-1} dr\right)\left[\mathbf{P}(\chi_{m-1}^2 \le \rho^2)\right]^{n-1} \prod_{j=2}^m \varphi(\zeta_j)
\end{aligned}
$$

since the random variable $\max_{1\le k\le n} a_{k1}$ has density $n\varphi(r)(\Phi(r))^{n-1}$ by a standard calculation. □

5.3 Proof of Proposition 10

As a first step towards the proof of Proposition 10 we first show the following result.

Proposition 12 *In the hypothesis of Proposition 10,*

$$
\begin{aligned}
&\overline{\lim_{\varepsilon\to 0}} \frac{1}{\varepsilon^{t-1}} \int_{\left\{\left|\alpha_{k_j}-\alpha_{l_j}\right|<\lambda\varepsilon \text{ for } j=1,\dots,t-1\right\}} p_{\varepsilon,\alpha}^g(0) \prod_{k=1}^n [\varphi(\alpha_k) d\alpha_k] \\
&\le \frac{(2\lambda)^{t-1}}{(2\pi)^{\frac{m-1}{2}}}\left[\begin{array}{c} \int_{-\infty}^{+\infty} (\varphi(x))^t dx \int_x^{+\infty} g(y)(n-t)\varphi(y)\,(\Phi(y))^{n-t-1}\, dy \\ +t^{\frac{m}{2}} \int_{-\infty}^{+\infty} g(x)(\varphi(x))^t\,(\Phi(x))^{n-t}\, dx \end{array}\right]
\end{aligned}
$$

We prove the statement first in the case $t=2, k_1=1, l_1=2$. This is done in Lemma 16 below. The general case will follow using similar arguments.

Lemma 16 *Let $g:\mathbb{R}\to\mathbb{R}$ be continuous with compact support and λ a positive constant. Then*

$$
\begin{aligned}
&\overline{\lim_{\varepsilon\to 0}} \frac{1}{\varepsilon} \int_{|\alpha_1-\alpha_2|<\lambda.\varepsilon} p_{\varepsilon,\alpha}^g(0) \prod_{k=1}^n [\varphi(\alpha_k) d\alpha_k] \\
&\le \frac{\lambda}{\pi} \frac{1}{(2\pi)^{\frac{m-1}{2}}} \int_{-\infty}^{+\infty} g(x)\,[\Phi(x)]^{n-3} \\
&\quad \times\left[\sqrt{\pi}(n-2)\varphi(x)\Phi(x\sqrt{2}) + 2^{\frac{m-1}{2}} e^{-x^2}\Phi(x)\right] dx
\end{aligned}
$$

Proof. Perform in the left-hand side of the inequality in the statement the change of variables $\alpha_1 = \alpha_2 + \varepsilon\tau$ so that it becomes

$$\int_{\mathbb{R}^{n-1}} \prod_{k=2}^{n} [\varphi(\alpha_k)d\alpha_k] \int_{-\lambda}^{\lambda} p^{g}_{\varepsilon,\alpha(\varepsilon)}(0)\varphi(\alpha_2 + \varepsilon\tau)d\tau \tag{26}$$

where

$$\alpha(\varepsilon) = (\alpha_2 + \varepsilon\tau, \alpha_2, \dots, \alpha_n)$$

To pass to the limit in (26) as $\varepsilon \to 0$ we want to proceed again as in the proof of Proposition 8 (ii). For that purpose we have a difficulty, namely that in the proof of this proposition we have assumed that α is fixed and $\alpha_{\overline{k}} = \max_{1\le k\le n} \alpha_k > \alpha_\ell$ for $\ell \neq \overline{k}$. It is not hard to see that the proof of Proposition 8 (ii) still works when α is not fixed but depends on ε and converges to a limit with the additional condition that the difference between $\alpha_{\overline{k}}$ and the remaining α_k's is bounded below by a positive number. This will be the case when we fix $\alpha_2, \dots, \alpha_n, \tau$ in the integral (26) so that $\alpha_2 < \max\{\alpha_3, \dots, \alpha_n\}$ but it is not the case if $\alpha_2 \geq \max\{\alpha_3, \dots, \alpha_n\}$. So, we divide the integral into two parts.

For

$$\int_{\alpha_2<\max\{\alpha_3,\dots,\alpha_n\}} \prod_{k=2}^{n} [\varphi(\alpha_k)d\alpha_k] \int_{-\lambda}^{\lambda} p^{g}_{\varepsilon,\alpha(\varepsilon)}(0)\varphi(\alpha_2 + \varepsilon\tau)d\tau$$

we pass to the limit as $\varepsilon \to 0$ and use Proposition 8 (ii) to obtain

$$\begin{aligned}
&\int_{\alpha_2<\max\{\alpha_3,\dots,\alpha_n\}} \prod_{k=2}^{n} [\varphi(\alpha_k)d\alpha_k] \int_{-\lambda}^{\lambda} q^{g}_{\alpha(0)}(0)\varphi(\alpha_2)d\tau \\
&\quad = \frac{\left[\mathbf{P}(\chi^2_{m-1} \le \rho^2)\right]^{n-1}}{(2\pi)^{\frac{m-1}{2}}} 2\lambda \int_{\alpha_2<\max\{\alpha_3,\dots,\alpha_n\}} g(\max\{\alpha_3, \dots, \alpha_n\})\varphi(\alpha_2) \\
&\qquad \times \prod_{k=2}^{n} [\varphi(\alpha_k)d\alpha_k] \\
&\quad = \frac{\left[\mathbf{P}(\chi^2_{m-1} \le \rho^2)\right]^{n-1}}{(2\pi)^{\frac{m-1}{2}}} 2\lambda \mathbf{E}\left\{g(\max\{\xi_3, \dots, \xi_n\})\varphi(\xi_2)\mathbb{I}_{\{\xi_2<\max\{\xi_3,\dots,\xi_n\}\}}\right\}
\end{aligned}$$

where $\xi_2, \dots, \xi_n$ are i.i.d. standard normal. The expectation in the right-hand side can be written as

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-y^2} dy \int_{y}^{+\infty} g(x)(n-2)\varphi(x)(\Phi(x))^{n-3}dx.$$

For the integral over the set $\Lambda = \{\alpha_2 > \max\{\alpha_3, \ldots, \alpha_n\}\}$ we need to go back to the proof of Proposition 9 (ii) to understand the behaviour of $p^g_{\varepsilon,\alpha(\varepsilon)}(\zeta)$, $\zeta \in \overline{B}_{m-1}(0;\rho)$ as $\varepsilon \to 0$, when $(\alpha_2, \alpha_3, \ldots, \alpha_n) \in \Lambda$.

As $\varepsilon \to 0$ the statement that $p^g_{\varepsilon,\alpha(\varepsilon)}$ converges uniformly on $\overline{B}_{m-1}(0;\rho)$ remains valid. To see this, applying the same method as in Proposition 9 (ii), it suffices to show that the existence of the limits (i), (ii), (iii) in Lemma 14 hold true, only that these limits are not the same. This is next Lemma.

Lemma 17 *Let* $\alpha(\varepsilon) = (\alpha_2 + \varepsilon\tau, \alpha_2, \ldots, \alpha_n)$ *and* τ, $(\alpha_2, \ldots, \alpha_n)$ *be fixed,* $\alpha_2 > \max\{\alpha_3, \ldots, \alpha_n\}$. *Then,*

(i)

$$H^{(1)}_{\varepsilon,\alpha(\varepsilon)}(M) \underset{\varepsilon\to 0}{\longrightarrow} \alpha_2 \tag{27}$$

(ii)

$$L_{\varepsilon,\alpha(\varepsilon)}(M) \underset{\varepsilon\to 0}{\longrightarrow} \gamma_1\widetilde{a}_1 + \gamma_2\widetilde{a}_2, \tag{28}$$

with

$$\gamma_1 = c_m \int_{\mathbb{R}^m} \psi(\|w\|)\mathbb{I}_{E(\tau)}dw, \qquad \gamma_2 = c_m \int_{\mathbb{R}^m} \psi(\|w\|)\mathbb{I}_{[E(\tau)]^c}dw$$

where we denote $\widetilde{w} = (w_2, \ldots, w_m)$,

$$E(\tau) = \left\{w \in \mathbb{R}^m \mid \langle \widetilde{a}_1 - \widetilde{a}_2, \widetilde{w}\rangle < \tau\right\}$$

and $\langle\ ,\ \rangle$ *denotes here scalar product in* $\mathbb{R}^{m-1}$.

(iii) *For* $j, h = 2, \ldots, m$

$$\frac{\partial H^{(h)}_{\varepsilon,\alpha(\varepsilon)}}{\partial a_{1j}}(M) \underset{\varepsilon\to 0}{\longrightarrow} c_m \int_{\mathbb{R}^m} \psi'(\|w\|)\frac{w_h w_j}{\|w\|}\mathbb{I}_{E(\tau)}dw = \ell_{1,j,h}$$

$$\frac{\partial H^{(h)}_{\varepsilon,\alpha(\varepsilon)}}{\partial a_{2j}}(M) \underset{\varepsilon\to 0}{\longrightarrow} c_m \int_{\mathbb{R}^m} \psi'(\|w\|)\frac{w_h w_j}{\|w\|}\mathbb{I}_{[E(\tau)]^c}dw = \ell_{2,j,h}$$

$$\frac{\partial H^{(h)}_{\varepsilon,\alpha(\varepsilon)}}{\partial a_{kj}}(M) \underset{\varepsilon\to 0}{\longrightarrow} 0 \quad \text{for } k = 3, \ldots, n.$$

Note that $\ell_{1,j,h} + \ell_{2,j,h} = \delta_{jh}$.

The convergence in (i), (ii) and (iii) is uniform on $M \in M_\rho$.

Proof. Note first that, as in the proof of Lemma 14, there exists $\varepsilon_0 > 0$ such that if $0 < \varepsilon < \varepsilon_0$, simultaneously for all $M \in M_\rho$ one has $U_k \cap B(e_1; \varepsilon) = \emptyset$ for $k = 3, \dots, n$. Hence:

$$H^{(1)}_{\varepsilon,\alpha}(M) = c_m \sum_{k=1}^{2} \int_{\mathbb{R}^m} \psi(\|w\|) \left(\alpha_k(1-\varepsilon w_1) - \varepsilon \sum_{j=2}^{m} a_{kj} w_j \right) \mathbb{I}_{\left\{\alpha_k(1-\varepsilon w_1) - \varepsilon \sum_{j=2}^{m} a_{kj} w_j > \alpha_\ell(1-\varepsilon w_1) - \varepsilon \sum_{j=2}^{m} a_{\ell j} w_j \text{ for all } \ell \neq k\right\}} dw$$

where $\alpha_1 = \alpha_2 + \varepsilon\tau$.

Almost everywhere in Ω,

$$\mathbb{I}_{\left\{\alpha_1(1-\varepsilon w_1) - \varepsilon \sum_{j=2}^{m} a_{1j} w_j > \alpha_\ell(1-\varepsilon w_1) - \varepsilon \sum_{j=2}^{m} a_{\ell j} w_j \text{ for all } \ell \neq 1\right\}} \to \mathbb{I}_{E(\tau)}$$

$$\mathbb{I}_{\left\{\alpha_2(1-\varepsilon w_1) - \varepsilon \sum_{j=2}^{m} a_{2j} w_j > \alpha_\ell(1-\varepsilon w_1) - \varepsilon \sum_{j=2}^{m} a_{\ell j} w_j \text{ for all } \ell \neq 2\right\}} \to \mathbb{I}_{[E(\tau)]^{\mathrm{c}}}.$$

This shows that $H^{(1)}_{\varepsilon,\alpha(\varepsilon)}(M) \underset{\varepsilon \to 0}{\longrightarrow} \alpha_2$. It is easy to see that the convergence is uniform on $A \in M_\rho$.

Similar computations show (ii) and (iii). □

We go back to the proof of Lemma 16. Replace α by $\alpha(\varepsilon)$ in (16) to obtain

$$\mathbf{E}_{\alpha(\varepsilon)} \left\{ g(Z_\varepsilon(e_1)) G\left(\frac{\partial Z_\varepsilon}{\partial x_2}(e_1), \dots, \frac{\partial Z_\varepsilon}{\partial x_m}(e_1) \right) \mathbb{I}_{M_\rho} \right\} = \int_{\overline{B}_{m-1}(0;\rho)} G(\zeta) p^g_{\varepsilon,\alpha(\varepsilon)}(\zeta) d\zeta.$$

Let $\varepsilon \to 0$ in both sides of this equality.

We use the same type of reasoning as in the proof of Proposition 9 (ii). Suppose that $p^g_{\varepsilon,\alpha(\varepsilon)}(\zeta)$ converges uniformly on some sequence $\varepsilon_\upsilon \to 0$ on $\zeta \in \overline{B}_{m-1}(0; \rho)$ to some function $\widetilde{q}_{\alpha_2,..,\alpha_n}(\zeta)$. The right-hand side in the equality above tends to

$$\int_{\overline{B}_{m-1}(0;\rho)} G(\zeta) \widetilde{q}_{\alpha_2,\dots,\alpha_n}(\zeta) d\zeta.$$

On the left-hand side, according to Lemma 17 (i), $Z_\varepsilon(e_1) \to \alpha_2$ a.s. and the random vector $\left(\frac{\partial Z_\varepsilon}{\partial x_2}(e_1), \dots, \frac{\partial Z_\varepsilon}{\partial x_m}(e_1) \right) \to \widetilde{\zeta}$, the limit in (ii) of the same lemma. A standard Gaussian regression shows that (use $\gamma_1 + \gamma_2 = 1$) the conditional distribution of the random vector $\widetilde{\zeta}$ given that $\widetilde{a}_1 - \widetilde{a}_2 = v \in \mathbb{R}^{m-1}$ is Gaussian with mean $\frac{1}{2} v(\gamma_1 - \gamma_2)$ and covariance matrix $\frac{1}{2} I_{m-1}$. Note that γ_1, γ_2 are functions of v.

Hence, the density of the random vector $\widetilde{\zeta}$ has the form:

$$p_{\widetilde{\zeta}}(\zeta) = \frac{1}{(\pi)^{\frac{m-1}{2}}} \int_{\mathbb{R}^{m-1}} e^{-\frac{1}{2}.2^{m-1}\left\|\zeta - \frac{1}{2}v(\gamma_1-\gamma_2)\right\|^2} p_{\widetilde{a}_1-\widetilde{a}_2}(v)dv$$

where $p_{\widetilde{a}_1-\widetilde{a}_2}$ denotes the density of the random vector $\widetilde{a}_1 - \widetilde{a}_2$. It follows that

$$p_{\widetilde{\zeta}}(\zeta) \leq \frac{1}{(\pi)^{\frac{m-1}{2}}} \quad \text{for all } \zeta.$$

So,

$$\mathbf{E}_{(\alpha_2,\dots,\alpha_n)}\left\{g(\alpha_2)G\left(\widetilde{\zeta}\right)\mathbb{I}_{M_\rho}\right\} = \int_{\overline{B}_{m-1}(0;\rho)} G(\zeta)\widetilde{q}_{\alpha_2,\dots,\alpha_n}(\zeta)d\zeta$$

which obviously implies

$$\begin{aligned} g(\alpha_2)\mathbf{E}_{(\alpha_2,\dots,\alpha_n)}\{G(\widetilde{\zeta})\} &= g(\alpha_2)\int_{\overline{B}_{m-1}(0;\rho)} G(\zeta)\ p_{\widetilde{\zeta}}(\zeta)d\zeta \\ &= \int_{\overline{B}_{m-1}(0;\rho)} G(\zeta)\widetilde{q}_{\alpha_2,\dots,\alpha_n}(\zeta)d\zeta \end{aligned}$$

because $\widetilde{\zeta}$ is a convex combination of $\widetilde{a}_1$ and $\widetilde{a}_2$.

The same type of arguments as in Proposition 9 (ii) imply that

$$\widetilde{q}_{\alpha_2,\dots,\alpha_n}(\zeta) = g(\alpha_2)\ p_{\widetilde{\zeta}}(\zeta), \qquad \zeta \in \overline{B}_{m-1}(0;\rho)$$

is the uniform limit of $p^g_{\varepsilon,\alpha(\varepsilon)}(\zeta)$ as $\varepsilon \to 0$.

It follows that

$$\begin{aligned} &\lim_{\varepsilon\to 0}\int_{\alpha_2\geq\max\{\alpha_3,\dots,\alpha_n\}} \prod_{k=2}^{n}[\varphi(\alpha_k)d\alpha_k]\int_{-\lambda}^{\lambda} p^g_{\varepsilon,\alpha(\varepsilon)}(0)\varphi(\alpha_2+\varepsilon\tau)d\tau \\ &= \int_{\alpha_2\geq\max\{\alpha_3,\dots,\alpha_n\}} g(\alpha_2)p_{\widetilde{\zeta}}(0)\varphi(\alpha_2).2\lambda \prod_{k=2}^{n}[\varphi(\alpha_k)d\alpha_k] \\ &\leq 2\lambda\frac{1}{(\pi)^{\frac{m-1}{2}}}\int_{\alpha_2\geq\max\{\alpha_3,\dots,\alpha_n\}} g(\alpha_2)\ \varphi(\alpha_2)\prod_{k=2}^{n}[\varphi(\alpha_k)d\alpha_k]. \end{aligned}$$

The proof of Lemma 16 follows now by grouping the two terms in the integral. □

Proof of Proposition 12 (continued). We can now proceed with the case $t > 2$. As we have already mentioned, the proof goes through the same lines as for $t = 2$ and we only give a brief sketch of it. Denote

$$\overline{\mathcal{X}} = \bigcup_{j=1}^{t-1}\{k_j, l_j\}.$$

Note that $\overline{\mathcal{X}}$ is a deterministic set of integers that one should consider as a realization of the random set $\mathcal{X}$. Clearly, $\overline{\mathcal{X}}$ has t different elements.

In the left-hand side of Proposition 12 make the change of variables

$$\alpha_{k_j} - \alpha_{l_j} = \varepsilon\tau_j, \quad j = 1, \dots, t-1 \tag{29}$$

We obtain:

$$\int_{[-\lambda,\lambda]^{t-1}} d\tau_1 \dots d\tau_{t-1} \int_{\mathbb{R}^{n-t+1}} p^g_{\varepsilon,\alpha(\varepsilon)}(0) \prod_{j=1}^{t-1} \varphi(\alpha_{k_j(\varepsilon)}) \prod_{\substack{k=1\\ k\neq k_j(j=1,\dots,t-1)}}^{n} [\varphi(\alpha_k)d\alpha_k]$$

where $\alpha(\varepsilon) = (\alpha_1(\varepsilon), \dots, \alpha_n(\varepsilon))$ and the $\alpha_k(\varepsilon)$ are computed from (29).

Clearly

$$\alpha_k(\varepsilon) \to \alpha_{l_1} \quad \text{for} \quad k \in \overline{\mathcal{X}}.$$

There is again the easy part that corresponds to the integral over the set of α's such that $\alpha_{l_1} < \max_{k\notin\overline{\mathcal{X}}} \alpha_k$ which converges, as $\varepsilon \to 0$, to the first term in the right-hand side of Proposition 12.

As for the integral over $\alpha_{l_1} > \max_{k\notin\overline{\mathcal{X}}} \alpha_k$ one proves that on $\alpha = \alpha(\varepsilon),\ \varepsilon \to 0$, almost surely:

$$\left(\frac{\partial Z_\varepsilon}{\partial x_2}(e_1), \dots, \frac{\partial Z_\varepsilon}{\partial x_m}(e_1)\right) \to \widetilde{\zeta}$$

where the conditional distribution of the random vector $\widetilde{\zeta}$ given that $\widetilde{a}_k - \widetilde{a}_l = v_{kl}$ for $k, l \in \overline{\mathcal{X}},\ k \neq l$ is normal with mean

$$\frac{1}{t} \sum_{\substack{k,l\in\overline{\mathcal{X}},\\ k\neq l}} \gamma_k v_{kl}$$

and variance matrix

$$\frac{1}{t} I_{m-1}.$$

Here the γ_k's are non-negative functions of the v_{kl}'s such that $\sum_{k\in\overline{\mathcal{X}}} \gamma_k = 1$. This allows one to pass to the limit in the second term, as we did for $t = 2$.

□

Proof of Proposition 10. The expression between the square brackets in the statement of Proposition 12 can be written as $Q_1 + Q_2$ where

$$Q_1 = \int_{-\infty}^{+\infty} (\varphi(x))^t dx \int_x^{+\infty} g(y)(n-t)\varphi(y)\,(\Phi(y))^{n-t-1}\, dy$$

and

$$Q_2 = t^{\frac{m}{2}} \int_{-\infty}^{+\infty} g(x)(\varphi(x))^t \, (\Phi(x))^{n-t} \, dx.$$

Write

$$H(x) = \sqrt{\frac{2\pi}{t}} \Phi(x\sqrt{t}) \text{ and } G(x) = \int_x^{+\infty} g(y)(n-t)\varphi(y) \, (\Phi(y))^{n-t-1} \, dy.$$

Then,

$$H'(x) = \sqrt{\frac{2\pi}{t}} \varphi(x\sqrt{t})\sqrt{t} = e^{-\frac{1}{2}tx^2}$$

and we have $H(+\infty) = \sqrt{\frac{2\pi}{t}}$, $H(-\infty) = 0$, $G(+\infty) = 0$, and $G(-\infty) \in \mathbb{R}$ due to the convergence of the integral. Now note

$$\begin{aligned} Q_1 &= \frac{1}{(2\pi)^{t/2}} \int_{-\infty}^{+\infty} H'(x)G(x) \, dx \\ &= \frac{1}{(2\pi)^{t/2}} \left[H(+\infty)G(+\infty) - H(-\infty)G(+\infty) - \int_{-\infty}^{+\infty} H(x)G'(x) \, dx \right] \\ &= \frac{1}{(2\pi)^{t/2}} \int_{-\infty}^{+\infty} \sqrt{\frac{2\pi}{t}} \Phi(x\sqrt{t}) g(x)(n-t)\varphi(x)(\Phi(x))^{n-t-1} \, dx \end{aligned}$$

the second equality by integration by parts. We conclude

$$\begin{aligned} Q_1 + Q_2 &= \frac{1}{(2\pi)^{t/2}} \int_{-\infty}^{+\infty} g(x)(\Phi(x))^{n-t-1} \\ &\quad \times \left[\sqrt{\frac{2\pi}{t}} \Phi(x\sqrt{t})(n-t)\varphi(x) + t^{\frac{m}{2}} e^{-\frac{1}{2}tx^2} \Phi(x) \right] dx \end{aligned}$$

from which the proposition follows. □

References

R. Adler: The Geometry of Random Fields. John Wiley & Sons 1981

L. Ahlfors: Complex Analysis (3 ed.). McGraw-Hill 1979

D. Brillinger: On the number of solutions of systems of random equations. Ann. of Math. Statistics **43**, 534–540 (1972)

D. Cheung, F. Cucker: A new condition number for linear programming. Math. Program. **91**, 163–174 (2001)

D. Cheung, F. Cucker: Probabilistic analysis of condition numbers for linear programming. Journal of Optimization Theory and Applications **114**, 55–67 (2002)

R. Courant: Differential and Integral Calculus, Volume II. John Wiley & Sons, 1988

F. Cucker, J. Peña: A primal-dual algorithm for solving polyhedral conic systems with a finite-precision machine. SIAM Journal on Optimization **12**, 522–554 (2001)

A. Dembo, O. Zeitouni: Large Deviations Techniques and Applications (2 ed.). Springer-Verlag. 1998

J. Demmel: On condition numbers and the distance to the nearest ill-posed problem. Numer. Math. **51**, 251–289 (1987)

J. Demmel: Applied Numerical Linear Algebra. SIAM, 1997

C. Eckart, G. Young: The approximation of one matrix by another of lower rank. Psychometrika **1**, 211–218 (1936)

A. Edelman: Eigenvalues and condition numbers of random matrices. SIAM J. of Matrix Anal. and Applic. **9**, 543–556 (1988)

R. Freund, M. Epelman: Pre-conditioners and relations between different measures of conditioning for conic linear systems. MIT Operations Research Center Working Paper OR-344-00. (2000)

R. Freund, J. Vera: Condition-based complexity of convex optimization in conic linear form via the ellipsoid algorithm. SIAM Journal on Optimization **10**, 155–176 (1999a)

R. Freund, J. Vera: Some characterizations and properties of the "distance to ill-posedness" and the condition measure of a conic linear system. Math. Program. **86**, 225–260 (1999b)

E. Gilbert: The probability of covering a sphere with *N* circular caps. Biometrika **52**, 323–330 (1966)

H. Goldstine: A History of Numerical Analysis from the 16th through the 19th Century. Springer-Verlag. 1977

P. Hall: Introduction to the Theory of Coverage Processes. John Wiley & Sons. 1988

N. Higham: Accuracy and Stability of Numerical Algorithms. SIAM. 1996

S. Janson: Random coverings in several dimensions. Acta Math. **156**, 83–118 (1986)

R. Miles: The asymptotic values of certain coverage probabilities. Biometrika **56**, 661–680 (1969)

J. Renegar: Is it possible to know a problem instance is ill-posed? Journal of Complexity **10**, 1–56 (1994a)

J. Renegar: Some perturbation theory for linear programming. Math. Program. **65**, 73–91 (1994b)

J. Renegar: Incorporating condition measures into the complexity theory of linear programming. SIAM J. on Optimization **5**, 506–524 (1995a)

J. Renegar: Linear programming, complexity theory and elementary functional analysis. Math. Program. **70**, 279–351 (1995b)

S. Smale: Complexity theory and numerical analysis. In A. Iserles (Ed.), Acta Numerica, pp. 523–551. Cambridge University Press. 1997

H. Solomon: Geometric Probability, Volume 28 of Regional Conference Series in Applied Mathematics. SIAM. 1978

G. Stewart: On scaled projections and pseudoinverses. Linear Algebra and its Applications **112**, 189–193 (1989)

M. Todd: A Dantzig-Wolfe-like variant of Karmarkar's interior-point linear programming algorithm. Operations Research **38**, 1006–1018 (1990)

M. Todd, L. Tunçel, Y. Ye: Characterizations, bounds and probabilistic analysis of two complexity measures for linear programming problems. Math. Program. **90**, 59–69 (2001)

L. Trefethen, D. Bau III: Numerical Linear Algebra. SIAM. 1997

A. Turing: Rounding-off errors in matrix processes. Quart. J. Mech. Appl. Math. **1**, 287–308 (1948)

S. Vavasis, Y. Ye: Condition numbers for polyhedra with real number data. Oper. Res. Lett. **17**, 209–214 (1995)

S. Vavasis, Y. Ye: A primal-dual interior point method whose running time depends only on the constraint matrix. Math. Program. **74**, 79–120 (1996)

J. Vera: On the complexity of linear programming under finite precision arithmetic. Math. Program. **80**, 91–123 (1998)

J. von Neumann, H. Goldstine: Numerical inverting matrices of high order. Bulletin of the Amer. Math. Soc. **53**, 1021–1099 (1947)

Y. Ye: Toward probabilistic analysis of interior-point algorithms for linear programming. Math. of Oper. Res. **19**, 38–52 (1994)