Average best m-term approximation

Jan Vybíral*

November 2, 2010

Abstract

We introduce the concept of average best m-term approximation widths with respect to a probability measure on the unit ball of ℓ_p^n. We estimate these quantities for the embedding $id : \ell_p^n \to \ell_q^n$ with $0 < p \le q \le \infty$ for the normalized cone and surface measure. Furthermore, we consider certain tensor product weights and show that a typical vector with respect to such a measure exhibits a strong compressible (i.e. nearly sparse) structure. This measure may be therefore used as a random model for sparse signals.

AMS subject classification (MSC 2010): Primary: 41A46, Secondary: 52A20, 60B11, 94A12.
Key words: nonlinear approximation, best m-term approximation, average widths, random sparse vectors, cone measure, surface measure.

1 Introduction

1.1 Best m-term approximation

Let $m \in \mathbb{N}_0$ and let Σ_m be the set of all sequences $x = \{x_j\}_{j=1}^\infty$ with

$$\|x\|_0 := \#\,\mathrm{supp}\, x = \#\{n \in \mathbb{N} : x_n \neq 0\} \le m.$$

Here stands $\#A$ for the number of elements of a set A. The elements of Σ_m are said to be *m-sparse*. Observe, that Σ_m is a non-linear subset of every $\ell_q := \{x = \{x_j\}_{j=1}^\infty : \|x\|_q < \infty\}$, where

$$\|x\|_q := \begin{cases} \left(\sum_{j=1}^\infty |x_j|^q\right)^{1/q}, & 0 < q < \infty,\\ \sup_{j\in\mathbb{N}} |x_j|, & q = \infty. \end{cases}$$

For every $x \in \ell_q$, we define its *best m-term approximation* error by

$$\sigma_m(x)_q := \inf_{y\in\Sigma_m} \|x - y\|_q.$$

*Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences, Altenbergerstrasse 69, A-4040 Linz, Austria, email: jan.vybiral@oeaw.ac.at, Tel: +43 732 2468 5250, Fax: +43 732 2468 5212.

Moreover for $0 < p \le q \le \infty$, we introduce the *best m-term approximation widths*

$$\sigma_m^{p,q} := \sup_{x:\|x\|_p \le 1} \sigma_m(x)_q.$$

The use of this concept goes back to Schmidt [31] and after the work of Oskolkov [27], it was widely used in the approximation theory, cf. [12, 14, 32]. In fact, it is the main prototype of nonlinear approximation [13]. It is well known, that

$$2^{-1/p}(m+1)^{1/q-1/p} \le \sigma_m^{p,q} \le (m+1)^{1/q-1/p}, \quad m = 0, 1, 2, \dots. \tag{1}$$

The proof of (1) is based on the simple fact, that (roughly speaking) the best m-term approximation error of $x \in \ell_p$ is realized by subtracting the m largest coefficients taken in absolute value. Hence,

$$\sigma_m(x)_q = \begin{cases} \left(\sum_{j=m+1}^{\infty} (x_j^*)^q\right)^{1/q}, & 0 < q < \infty, \\ x_{m+1}^* = \sup_{j \ge m+1} x_j^*, & q = \infty, \end{cases}$$

where $x^* = (x_1^*, x_2^*, \dots)$ denotes the so-called *non-increasing rearrangement* [4] of the vector $(|x_1|, |x_2|, |x_3|, \dots)$.

Let us recall the proof of (1) in the simplest case, namely $q = \infty$. The estimate from above then follows by

$$\sigma_m(x)_\infty = \sup_{j \ge m+1} x_j^* = x_{m+1}^* \le \left((m+1)^{-1} \sum_{j=1}^{m+1} (x_j^*)^p\right)^{1/p} \le (m+1)^{-1/p} \|x\|_p. \tag{2}$$

The lower estimate is supplied by taking

$$x = (m+1)^{-1/p} \sum_{j=1}^{m+1} e_j, \tag{3}$$

where $\{e_j\}_{j=1}^\infty$ are the canonical unit vectors.

For general q, the estimate from above in (1) may be obtained from (2) and Hölder's inequality

$$\|x\|_q \le \|x\|_p^\theta \cdot \|x\|_\infty^{1-\theta}, \quad \text{where} \quad \frac{1}{q} = \frac{\theta}{p}. \tag{4}$$

The estimate from below follows for all q's by simple modification of (3).
The discussion above exhibits two effects.

(i) Best m-term approximation works particularly well, when $1/p - 1/q$ is large, i.e. if $p < 1$ and $q = \infty$.

(ii) The elements used in the estimate from below (and hence the elements, where the best m-term approximation performs at worse) enjoy a very special structure.

Therefore, there is a reasonable hope, that the best m-term approximation could behave better, when considered in a certain average case. But first we point out two different interesting points of view on the subject.

1.2 Connection to compressed sensing

The interest in ℓ_p spaces (and especially in their finite-dimensional counterparts ℓ_p^n) with $0 < p < 1$ was recently stimulated by the impressive success of the novel and vastly growing area of *compressed sensing* as introduced by [6, 8, 9, 15]. Without going much into the details, we only note, that the techniques of compressed sensing allow to reconstruct a vector from an incomplete set of measurements utilizing the prior knowledge, that it is sparse, i.e. $\|x\|_0$ is small. Furthermore, this approach may be applied [11] also to vectors, which are *compressible*, i.e. $\|x\|_p$ is small for (preferably small) $0 < p < 1$. Indeed, (1) tells us, that such a vector x may be very well approximated by sparse vectors. We point to [7, 17, 18, 29] for the current state of the art of this field and for further references.

This leads in a very natural way to a question, which stands in the background of this paper, namely:

How does a typical vector of the ℓ_p^n unit ball look like?

or, posed in an exact way:

Let μ be a probability measure on the unit ball of ℓ_p^n. What is the mean value of $\sigma_m(x)_q$ with respect to this measure?

Of course, the choice of μ plays a crucial role. There are several standard probability measures, which are connected to the unit ball of ℓ_p^n in a natural way, namely (cf. Definitions 2 and 10)

(i) the normalized Lebesgue measure,

(ii) the $n-1$ dimensional Hausdorff measure restricted to the surface of the unit ball of ℓ_p^n and correspondingly normalized,

(iii) the so-called normalized cone measure.

Unfortunately, it turns out, that all these three measures are "bad" – a typical vector with respect to any of them does not involve much structure and corresponds rather to noise then signal (in the sense described in the next section). Therefore, we are looking for a new type of measures (cf. Definition 14), which would behave better from this point of view.

1.3 Random models of noise and signals

Random vectors play an important role in the area of signal processing. For example, if $n \in \mathbb{N}$ is a natural number, $\omega = (\omega_1, \dots, \omega_n)$ is a vector of independent Gaussian variables and $\varepsilon > 0$ is a real number, then $\varepsilon\omega$ is a classical model of noise, namely the *white noise*. This model is used in the theory but also in the real life applications of signal processing.

The random generation of a structured signal seems to be a more complicated task. Probably the most common random model to generate sparse vectors, cf. [5, 10, 20, 28], is the so-called *Bernoulli-Gaussian model*. Let again $n \in \mathbb{N}$ be a natural number and $\varepsilon > 0$ be a real number. Also $\omega = (\omega_1, \dots, \omega_n)$ stands for a vector of independent Gaussian variables. Furthermore, let $0 < p < 1$ be a real number and let $\varrho = (\varrho_1, \dots, \varrho_n)$ be a vector of independent Bernoulli variables defined as

$$\varrho_i = \begin{cases} 1, & \text{with probability } p, \\ 0, & \text{with probability } 1-p. \end{cases}$$

The random *Bernoulli-Gaussian vector* x is then defined through

$$x_i = \varepsilon \varrho_i \cdot \omega_i, \quad i = 1, \dots, n. \tag{5}$$

Obviously, the average number of non-zero components of x is $k := pn$. Unfortunately, if k is much smaller than n, then the concentration of the number of non-zero components of x around k is not very strong. This becomes better, if k gets larger. But in that case, the model (5) resembles more and more the model of white noise. In some sense, (5) represents rather a randomly filtered white noise then a structured signal. It is one of the main aims of this paper to find a new measure, such that a random vector with respect to this measure would show a nearly sparse structure without the need of random filtering.

1.4 Unit sphere

Let us describe the situation in the most prominent case, when $p = 2$, $m = 0$ and μ is the normalized surface measure on the unit sphere $\mathbb{S}^{n-1}$ of ℓ_2^n. Furthermore, we denote by γ_n the standard Gaussian measure on $\mathbb{R}^n$ with the density

$$\frac{1}{(2\pi)^{n/2}} e^{-\|x\|_2^2/2}, \quad x \in \mathbb{R}^n.$$

We use polar coordinates to calculate

$$\begin{aligned} \int_{\mathbb{R}^n} \max_{j=1,\dots,n} |x_j| \, d\gamma_n(x) &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \max_{j=1,\dots,n} |x_j| \cdot e^{-\|x\|_2^2/2} dx \\ &= \frac{\Omega_n}{(2\pi)^{n/2}} \int_0^\infty r^{n-1} \int_{\mathbb{S}^{n-1}} \max_{j=1,\dots,n} |r x_j| e^{-\|rx\|_2^2/2} d\mu(x) \, dr \\ &= \frac{\Omega_n}{(2\pi)^{n/2}} \int_0^\infty r^n e^{-r^2/2} dr \cdot \int_{\mathbb{S}^{n-1}} \max_{j=1,\dots,n} |x_j| d\mu(x) \qquad (6) \\ &= \frac{\Omega_n}{(2\pi)^{n/2}} \int_0^\infty r^n e^{-r^2/2} dr \cdot \int_{\mathbb{S}^{n-1}} \sigma_0(x)_\infty d\mu(x), \end{aligned}$$

where Ω_n denotes the area of $\mathbb{S}^{n-1}$. This formula connects the expected value of $\sigma_0(x)_\infty$ with the expected value of maximum of n independent Gaussian variables. Using that this quantity is known to be equivalent to $\sqrt{\log(n+1)}$, cf. [23, (3.14)],

$$\int_0^\infty r^n e^{-r^2/2} dr = 2^{(n-1)/2} \Gamma((n+1)/2) \quad \text{and} \quad \Omega_n = \frac{2\pi^{n/2}}{\Gamma(n/2)},$$

one obtains

$$\int_{\mathbb{S}^{n-1}} \sigma_0(x)_\infty d\mu(x) \approx \sqrt{\frac{\log(n+1)}{n}}, \quad n \in \mathbb{N}. \tag{7}$$

Several comments on (6) and (7) are necessary.

(i) There is a direct connection between the estimated value of a maximum of independent Gaussian variables and the estimated value of the largest coordinate of a random vector on $\mathbb{S}^{n-1}$. Due to the result of [30], this holds true also for other values of p, even for $p < 1$. We hope, that this comment justifies the extensive use of (modified) Gaussian variables throughout the paper.

(ii) Obviously, the average size of the coordinate of any $x \in \mathbb{S}^{n-1}$ is $1/\sqrt{n}$. The formula (7) says, that the average size of the largest coordinate is only slightly larger, namely in the logarithmic order. Intuitively, this means that many of the coordinates are typically of the same order, namely $1/\sqrt{n}$. This (and other similar aspects of the geometry of $\mathbb{S}^{n-1}$) is well known under the name *concentration of measure phenomena* [22].

(iii) The calculation (6) is based on the use of polar coordinates. For $p \neq 2$, the normalized cone measure is exactly that measure, for which a similar formula holds, cf. (12). The estimates for $n-1$ dimensional surface measure are later obtained using its density with respect to the cone measure, cf. Lemma 11.

(iv) As we want to keep the paper self-contained as much as possible and to make it readable also for readers without (almost) any stochastic background, we prefer to use simple and direct techniques. For example we use rather the simple estimates in Lemma 5, than any of their sophisticated improvements available in literature.

(v) The connection to random Gaussian variables explains, why a random point of $\mathbb{S}^{n-1}$ is sometimes referred to as *white (or Gaussian) noise*. It is usually not associated with any reasonable (i.e. structured) signal, rather it represents a good model for random noise.

Surprisingly enough, (7) has its direct counterpart for all $0 < p < \infty$ if μ is the normalized cone measure in ℓ_p^n, cf. Theorem 7. This means (as described above), that the coordinates of a "typical" element of the surface of the ℓ_p^n unit ball are well concentrated around the value $n^{-1/p}$. So, roughly speaking, it is only ℓ_p-normalized noise.

Therefore, we are looking for a new measure, which would "promote sparsity" in the sense, that the mean value of $\sigma_m(x)_q$ decays rapidly with m. Intuitively, a good candidate for this could be the normalized $n-1$ dimensional Hausdorff measure on the sphere of the unit ball of ℓ_p^n. It gives namely a bigger weight to those areas, where one (or more) of the components of x are approaching zero. This effect is mathematically described in Lemma 11. Unfortunately, it turns out (as shown in Theorem 13), that this does not effect essentially the results.

The search for a measure promoting sparsity is then finished in Definition 14 by introducing a new class of measures $\theta_{p,\beta}$. We show, that for an appropriate choice of β, namely $\beta = p/n - 1$, the estimated value of the m-th largest coefficient of elements of the ℓ_p^n-unit sphere decays exponentially with m. Furthermore, these results are in a certain way independent of n. This gives a hope, that one could apply this approach also to the infinite-dimensional spaces ℓ_p or, using a suitable discretization technique (like wavelet decomposition), also to some function spaces. This remains a subject of our further research.

The paper is structured as follows. The rest of Section 1 contains the definition of the concept of average best m-term width. Sections 2 and 3 provide estimates of this quantity with respect to the cone and surface measure, respectively. In Section 4, we study a new type of measures on the unit ball of ℓ_p^n. We show, that the typical element with respect to those measures behaves in a completely different way compared to the situations discussed before. Those results are illustrated by the numerical experiments described in Section 5.

1.5 Notation

We denote by $\mathbb{R}$ the set of real numbers, by $\mathbb{R}_+ := [0,\infty)$ the set of nonnegative real numbers and by $\mathbb{R}^n$ and $\mathbb{R}^n_+$ their n-fold tensor products. The components of $x \in \mathbb{R}^n$ are denoted by $x_1,\dots,x_n$. The symbol λ stands for the Lebesgue measure on $\mathbb{R}^n$ and $\mathcal{H}$ for the $n-1$ dimensional Hausdorff measure in $\mathbb{R}^n$. If $A \subset \mathbb{R}^n$ and $I \subset \mathbb{R}$ is an interval, we write $I \cdot A := \{tx : t \in I, x \in A\}$.

We shall use very often the *Gamma function*, defined by

$$\Gamma(s) := \int_0^\infty t^{s-1}e^{-t}dt, \quad s > 0. \tag{8}$$

In one case, we shall use also the *Beta function*

$$B(p,q) := \int_0^1 t^{p-1}(1-t)^{q-1}dt = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}, \quad p,q > 0 \tag{9}$$

and the *digamma function*

$$\Psi(s) := \frac{d}{ds}\log\Gamma(s) = \frac{\Gamma'(s)}{\Gamma(s)}, \quad s > 0.$$

We recommend [1, Chapter 6] as a standard reference for both basic and more advanced properties of these functions. We shall need the Stirling's approximation formula (which was implicitly used already in (7)) in its most simple form

$$\Gamma(x) = \sqrt{\frac{2\pi}{x}}\left(\frac{x}{e}\right)^x\left(1+\mathcal{O}\left(\frac{1}{x}\right)\right), \quad x > 0. \tag{10}$$

If $a = \{a_j\}_{j=1}^\infty$ and $b = \{b_j\}_{j=1}^\infty$ are real sequences, then $a_j \lesssim b_j$ means, that there is a constant $c > 0$, such that $a_j \le c\, b_j$ for all $j = 1,2,\dots$. Similar convention is used for $a_j \gtrsim b_j$ and $a_j \approx b_j$.

We may easily observe, that

$$\sigma_m((x_1,\dots,x_n))_q = \sigma_m((\varepsilon_1 x_1,\dots,\varepsilon_n x_n))_q = \sigma_m((|x_1|,\dots,|x_n|))_q$$

holds for every $x \in \mathbb{R}^n$ and $\varepsilon \in \{-1,+1\}^n$. Also all the measures, which we shall consider, are invariant under any of the mappings

$$(x_1,\dots,x_n) \to (\varepsilon_1 x_1,\dots,\varepsilon_n x_n), \quad \varepsilon \in \{-1,+1\}^n.$$

This explains, why we restrict our attention only to $\mathbb{R}^n_+$ in the following definition.

Definition 1. Let $0 < p \le q \le \infty$ and let $n \ge 2$ and $0 \le m \le n-1$ be natural numbers.

(i) We set

$$\Delta_p^n = \begin{cases} \{(t_1,\dots,t_n) \in \mathbb{R}_+^n : \sum_{j=1}^n t_j^p = 1\}, & p < \infty, \\ \{(t_1,\dots,t_n) \in \mathbb{R}_+^n : \max_{j=1,\dots,n} t_j = 1\}, & p = \infty. \end{cases}$$

(ii) Let μ be a Borel probability measure on Δ_p^n. Then

$$\sigma_m^{p,q}(\mu) = \int_{\Delta_p^n} \sigma_m(x)_q d\mu(x)$$

is called *average surface best m-term width of id* : $\ell_p^n \to \ell_q^n$ *with respect to* μ.

(iii) Let ν be a Borel probability measure on $[0,1] \cdot \Delta_p^n$. Then

$$\sigma_m^{p,q}(\nu) = \int_{[0,1]\cdot\Delta_p^n} \sigma_m(x)_q d\nu(x)$$

is called *average volume best m-term width of id* : $\ell_p^n \to \ell_q^n$ *with respect to* ν.

Remark 1. (i) Let us observe, that the estimates

$$\sigma_m^{p,q}(\mu) \le \sigma_m^{p,q} \quad \text{and} \quad \sigma_m^{p,q}(\nu) \le \sigma_m^{p,q}$$

follow trivially by Definition 1.

(ii) The mapping $x \to \sigma_m(x)_q$ is continuous and, therefore, measurable with respect to the Borel measure μ.

2 Normalized cone measure

In this section, we shall give estimates of average best m-term widths with respect to the so-called cone measure, which is well studied in the literature within the geometry of ℓ_p^n spaces, cf. [26, 2, 25, 3].

Definition 2. Let $0 < p \le \infty$ and $n \ge 2$. Then

$$\mu_p(\mathcal{A}) = \frac{\lambda([0,1]\cdot\mathcal{A})}{\lambda([0,1]\cdot\Delta_p^n)}, \quad \mathcal{A} \subset \Delta_p^n$$

is the normalized *cone measure* on Δ_p^n.

If ν_p denotes the p-normalized Lebesgue measure, i.e.

$$\nu_p(A) = \frac{\lambda(A)}{\lambda([0,1]\cdot\Delta_p^n)}, \quad A \subset \mathbb{R}_+^n,$$

then the connection between ν_p and μ_p is given by

$$\nu_p(A) = n\int_0^\infty r^{n-1}\mu_p\biggl(\frac{\{x\in A : \|x\|_p = r\}}{r}\biggr)dr. \tag{11}$$

The proof of (11) follows directly for sets of the type $[a,b]\cdot\mathcal{A}$ with $0<a<b<\infty$ and $\mathcal{A}\subset\Delta_p^n$ and is then finished by standard approximation arguments. The formula (11) may be generalized to the so-called *polar decomposition identity*, cf. [2],

$$\frac{\int_{\mathbb{R}_+^n} f(x)d\lambda(x)}{\lambda([0,1]\cdot\Delta_p^n)} = n\int_0^\infty r^{n-1}\int_{\Delta_p^n} f(rx)d\mu_p(x)dr, \tag{12}$$

which holds for every $f\in L_1(\mathbb{R}_+^n)$.

The formula (12) allows to transfer immediately the results for the average surface best m-term approximation with respect to μ_p to the average volume approximation with respect to ν_p.

Proposition 3. *The identity*

$$\sigma_m^{p,q}(\nu_p) = \sigma_m^{p,q}(\mu_p)\cdot\frac{n}{n+1}$$

holds for all $0<p\le q\le\infty$*, all* $n\ge 2$ *and all* $0\le m\le n-1$*.*

Proof. We plug the function

$$f(x) = \sigma_m(x)_q\cdot\chi_{[0,1]\cdot\Delta_p^n}(x)$$

into (12) and obtain

$$\begin{aligned}\frac{\int_{[0,1]\cdot\Delta_p^n}\sigma_m(x)_q d\lambda(x)}{\lambda([0,1]\cdot\Delta_p^n)} &= \int_{[0,1]\cdot\Delta_p^n}\sigma_m(x)_q d\nu_p(x)\\ &= n\int_0^1 r^{n-1}\int_{\Delta_p^n}\sigma_m(rx)_q d\mu_p(x)dr = n\int_0^1 r^n dr\cdot\sigma_m^{p,q}(\mu_p),\end{aligned}$$

which gives the result. □

Remark 2. Proposition 3 shows, that the difference between approximation with respect to μ_p and ν_p is immaterial - i.e. of the order $1/n$. This justifies our interest in measures on Δ_p^n.

Let $p=2$ and let $\omega_1,\dots,\omega_n$ be independent normally distributed Gaussian random variables. Then

$$\varrho_p(\mathcal{A}) = \mu_p(\mathcal{A}) = \mathbb{P}\Biggl(\frac{(|\omega_1|,\dots,|\omega_n|)}{\bigl(\sum_{j=1}^n\omega_j^2\bigr)^{1/2}}\in\mathcal{A}\Biggr),\qquad \mathcal{A}\subset\Delta_p^n.$$

As noted in [30], this relation may be generalized to all values of p with $0 < p < \infty$. Let $\omega_1, \dots, \omega_n$ be independent random variables on $\mathbb{R}_+$ each with density

$$c_p e^{-t^p}, \quad t \geq 0$$

with respect to the Lebesgue measure, where $c_p = \frac{p}{\Gamma(1/p)} = \frac{1}{\Gamma(1/p+1)}$.

Then, cf. [30, Lemma 1],

$$\mu_p(\mathcal{A}) = \mathbb{P}\left(\frac{(\omega_1, \dots, \omega_n)}{\left(\sum_{j=1}^n \omega_j^p\right)^{1/p}} \in \mathcal{A} \right), \qquad \mathcal{A} \subset \Delta_p^n. \tag{13}$$

We shall fix $\omega_1, \dots, \omega_n$ to the end of this paper. Also the symbols $\mathbb{E}$ and $\mathbb{P}$ are always taken with respect to these variables.

2.1 The case $q = \infty$

In this section we deal with uniform approximation, i.e. with the case $q = \infty$.

Lemma 4. *Let $0 < p < \infty$ and let $n \geq 2$ and $1 \leq m \leq n$ be natural numbers. Then*

$$\int_{\Delta_p^n} x_m^* d\mu_p(x) = \frac{\Gamma(n/p)}{\Gamma(n/p + 1/p)} \cdot \mathbb{E}\, x_m^* \approx \frac{\mathbb{E}\, x_m^*}{n^{1/p}},$$

with constants of equivalence independent of m and n.

Proof. We put $f(x) = x_m^* e^{-x_1^p - \dots - x_n^p}$ and use the polar decomposition identity (12)

$$\begin{aligned} \frac{\int_{\mathbb{R}_+^n} x_m^* e^{-x_1^p - \dots - x_n^p} d\lambda(x)}{\lambda([0,1] \cdot \Delta_p^n)} &= n \int_0^\infty r^{n-1} \int_{\Delta_p^n} (r x_m^*) \cdot e^{-(rx_1)^p - \dots - (rx_n)^p} d\mu_p(x) dr \\ &= n \int_0^\infty r^{n-1} \cdot r e^{-r^p} dr \int_{\Delta_p^n} x_m^* d\mu_p(x) \end{aligned}$$

or, equivalently,

$$\int_{\Delta_p^n} x_m^* d\mu_p(x) = \frac{\int_{\mathbb{R}_+^n} x_m^* e^{-x_1^p - \dots - x_n^p} d\lambda(x)}{\lambda([0,1] \cdot \Delta_p^n) \cdot n \int_0^\infty r^n e^{-r^p} dr}. \tag{14}$$

The identity

$$\int_0^\infty r^n e^{-r^p} dr = \frac{\Gamma(n/p + 1/p)}{p},$$

follows by a simple substitution. Furthermore, we shall need the classical formula of Dirichlet for the volume of the unit ball $B_{\ell_p^n}$ of ℓ_p^n, cf. [16, p. 157],

$$\lambda([0,1] \cdot \Delta_p^n) = \frac{\lambda(B_{\ell_p^n})}{2^n} = \frac{\Gamma(1/p + 1)^n}{\Gamma(n/p + 1)}.$$

This allows us to reformulate (14) as

$$\int_{\Delta_p^n} x_m^* d\mu_p(x) = \frac{\Gamma(n/p+1)\,\mathbb{E}\,x_m^*}{c_p^n \cdot n/p \cdot \Gamma(n/p+1/p)\Gamma(1/p+1)^n} = \frac{\Gamma(n/p)\,\mathbb{E}\,x_m^*}{\Gamma(n/p+1/p)}.$$

Finally, we use Stirling's formula (10) to estimate

$$\frac{n^{1/p}\cdot\Gamma(n/p)}{\Gamma(n/p+1/p)} \approx \frac{n^{1/p}(n/p)^{n/p-1/2}}{(n/p+1/p)^{n/p+1/p-1/2}} \approx \left(\frac{n}{n+1}\right)^{n/p+1/p-1/2} \approx 1,$$

where the used constants of equivalence do not depend on n, but may depend on p. □

Lemma 5. *Let $\alpha \in \mathbb{R}$ and $\delta > 0$. Then*

$$\int_\delta^\infty u^\alpha e^{-u} du \le \delta^\alpha e^{-\delta} \cdot \begin{cases} 1, & \text{if} \quad \alpha \le 0, \\ \frac{1}{1-\alpha/\delta}, & \text{if} \quad \alpha > 0 \quad \text{and} \quad \frac{\alpha}{\delta} < 1, \\ \left(\frac{\alpha}{\delta}\right)^\alpha \cdot \frac{\alpha/\delta}{1-\delta/\alpha}, & \text{if} \quad \alpha > 0 \quad \text{and} \quad \frac{\alpha}{\delta} > 1. \end{cases}$$

Proof. If $\alpha \le 0$, we may estimate

$$\int_\delta^\infty u^\alpha e^{-u} du \le \delta^\alpha \int_\delta^\infty e^{-u} du = \delta^\alpha e^{-\delta}.$$

If $0 < \alpha \le 1$, we use partial integration and obtain

$$\int_\delta^\infty u^\alpha e^{-u} du = \delta^\alpha e^{-\delta} + \alpha \int_\delta^\infty u^{\alpha-1} e^{-u} du \le \delta^\alpha e^{-\delta}(1+\alpha\delta^{-1}).$$

This is smaller than

$$\delta^\alpha e^{-\delta}(1+\frac{\alpha}{\delta}+\frac{\alpha^2}{\delta^2}+\dots) = \delta^\alpha e^{-\delta}\cdot\frac{1}{1-\alpha/\delta}$$

if $\alpha/\delta < 1$ and smaller than

$$\delta^\alpha e^{-\delta}\frac{\alpha}{\delta}(1+\frac{\delta}{\alpha}+\frac{\delta^2}{\alpha^2}+\dots) = \delta^\alpha e^{-\delta}\frac{\alpha}{\delta}\cdot\frac{1}{1-\delta/\alpha}.$$

if $\alpha/\delta > 1$.

If $k-1 < \alpha \le k$ for some $k \in \mathbb{N}$, we iterate the partial integration and arrive at

$$\begin{aligned}\int_\delta^\infty u^\alpha e^{-u} du &\le \delta^\alpha e^{-\delta}(1+\alpha\delta^{-1}+\alpha(\alpha-1)\delta^{-2}+\dots+\alpha(\alpha-1)\dots(\alpha-k+1)\delta^{-k}) \\ &\le \delta^\alpha e^{-\delta}(1+\frac{\alpha}{\delta}+\frac{\alpha^2}{\delta^2}+\dots+\frac{\alpha^k}{\delta^k}) \\ &\le \delta^\alpha e^{-\delta}\begin{cases}\frac{1}{1-\alpha/\delta}, & \text{if} \quad \alpha/\delta < 1, \\ \left(\frac{\alpha}{\delta}\right)^{\alpha+1}\frac{1}{1-\delta/\alpha}, & \text{if} \quad \alpha/\delta > 1.\end{cases}\end{aligned}$$

□

Lemma 6. *Let $0 < p < \infty$ and let $1 \le m \le n$. Then*

$$\mathbb{E}\, x_m^* \lesssim \log^{1/p}\Big(\frac{en}{m}\Big).$$

Proof. We estimate

$$\begin{aligned}\mathbb{E}\, x_m^* &= \int_0^\infty \mathbb{P}(\omega_m^* > t)dt = \delta + \int_\delta^\infty \mathbb{P}(\omega_m^* > t)dt\\ &\le \delta + \binom{n}{m}\int_\delta^\infty \mathbb{P}(\omega_1 > t, \omega_2 > t, \dots, \omega_m > t)dt \qquad (15)\\ &= \delta + \binom{n}{m}\int_\delta^\infty \mathbb{P}(\omega_1 > t)^m dt.\end{aligned}$$

The parameter $\delta > \max(1, 3(1/p-1))^{1/p}$ is to be chosen later on. We substitute $v = u^p$ and obtain

$$\mathbb{P}(\omega_1 > t) = c_p \int_t^\infty e^{-u^p} du = \frac{c_p}{p}\int_{t^p}^\infty v^{1/p-1}e^{-v}dv.$$

Using the first two estimates of Lemma 5 (recall that $t^p \ge \delta^p > \max(1, 3(1/p-1))$), we arrive at

$$\mathbb{P}(\omega_1 > t) \le C_p t^{1-p} e^{-t^p},$$

where C_p depends only on p. We plug this estimate into (15) and obtain

$$\mathbb{E}\, x_m^* \le \delta + \binom{n}{m} C_p^m \int_\delta^\infty t^{m(1-p)} e^{-mt^p} dt. \qquad (16)$$

If $p \ge 1$, then

$$\int_\delta^\infty t^{m(1-p)}e^{-mt^p}dt \le \delta^{m(1-p)}\int_\delta^\infty e^{-mt^p}dt \le \delta^{m(1-p)}\int_{m\delta^p}^\infty e^{-u}u^{1/p-1}du \le e^{-m\delta^p}.$$

Altogether, we obtain

$$\mathbb{E}\, x_m^* \le \delta + \binom{n}{m} C_p^m e^{-m\delta^p}.$$

Using $\binom{n}{m} \le (\frac{en}{m})^m$ and choosing $\delta = c' \ln(\frac{en}{m})^{1/p}$ finishes the proof.

If $p < 1$, we use again the second estimate of Lemma 5

$$\begin{aligned}\int_\delta^\infty t^{m(1-p)}e^{-mt^p}dt &= \frac{1}{mp}\cdot m^{(1/p-1)(m+1)}\int_{m\delta^p}^\infty u^{(1/p-1)(m+1)}e^{-u}du\\ &\le \frac{1}{mp}\cdot \delta^{(1-p)(m+1)}e^{-m\delta^p}\cdot \frac{1}{1-\frac{2(1/p-1)}{\delta^p}} \le c_p'\delta^{(1-p)(m+1)}e^{-m\delta^p}.\end{aligned}$$

Using (16) and $\binom{n}{m} \le (\frac{en}{m})^m$ again, we get

$$\begin{aligned}\mathbb{E}\, x_1^* &\le \delta + \exp(-m\delta^p + m\ln(en/m) + (1-p)(m+1)\ln\delta + m\ln C_p + \ln c_p')\\ &\le \delta + \exp[-m(\delta^p + c\ln(en/m) + 2(1-p)\ln\delta)]\end{aligned}$$

The choice $\delta = c'\ln(\frac{en}{m})^{1/p}$ with c' large enough ensures, that

$$\frac{\delta^p}{2} \ge c\ln(en/m) \quad \text{and} \quad \frac{\delta^p}{2} \ge 2(1-p)\ln\delta$$

and finishes the proof. □

Lemma 4 and Lemma 6 imply immediately the following theorem if $p<\infty$. If $p=\infty$, the proof is trivial.

Theorem 7. *Let $0<p\le\infty$ and let $n\ge 2$ and $0\le m\le n-1$ be natural numbers. Then*

$$\sigma_m^{p,\infty}(\mu_p)\lesssim\left[\frac{\log\left(\frac{en}{m+1}\right)}{n}\right]^{1/p}.$$

Remark 3. (i) Theorem 7 may be interpreted in the sense of the discussion after (7). Namely, the average coordinate of $x\in\Delta_p^n$ is $n^{-1/p}$. Theorem 7 shows, that the average value of the largest coordinate is only slightly larger (namely $c[\ln(en)]^{1/p}$ times larger). In this sense, the average point of Δ_p^n is only slightly modified (and properly normalized) white noise.

(ii) Using the interpolation formula (4), one may immediately extend this result to all $0<p\le q<\infty$. But we shall see later on, that in the case $q<\infty$, one may prove slightly better estimates.

(iii) It is easy to see, that

$$\sigma_0^{p,\infty}(\mu_p)\ge\inf_{x\in\Delta_p^n}x_1^*=n^{-1/p}.$$

We shall show (at least in the case $m=0$), that the logarithmic factor is indispensable in Theorem 7. As far as we can tell, the exact behavior for $m>0$ is left open even for $p=2$.

Proposition 8. *Let $0<p\le\infty$ and $n\ge 2$. Then*

$$\sigma_0^{p,\infty}(\mu_p)\gtrsim\left[\frac{\log(n+1)}{n}\right]^{1/p},$$

where the constant may depend on p but not on n.

Proof. If $p=\infty$, then $x_1^*=1$ for all $x\in\Delta_p^n$ and the proof is trivial. Let us therefore assume, that $p<\infty$. According to Lemma 4, we have to estimate $\mathbb{E}\,x_1^*$ from below. This was done in [30, Lemma 2]. We include a slightly different proof for readers convenience. For every $t_0>0$, it holds

$$\mathbb{E}\,x_1^*\ge t_0\,\mathbb{P}(x_1^*>t_0)=t_0\,\mathbb{P}(\max_{1\le j\le n}x_j>t_0)\ge t_0[n\mathbb{P}(x_1>t_0)-\binom{n}{2}\mathbb{P}(x_1>t_0)^2].$$

We define t_0 by $\mathbb{P}(x_1>t_0)=\frac{1}{n}$ and obtain $\mathbb{E}\,x_1^*\ge t_0/2$.

From the simple estimate

$$\frac{c_p}{p}\int_{T^p}^{\infty}u^{1/p-1}e^{-u}du\ge C_pe^{-2T^p},\qquad T>1,$$

it follows, that there is an constant $\gamma_p>0$, such that

$$\mathbb{P}(x_1>\gamma_p(\log(n+1))^{1/p})\ge 1/n.$$

This gives $t_0\ge\gamma_p(\log(n+1))^{1/p}$ and $\mathbb{E}\,x_1^*\gtrsim(\log(n+1))^{1/p}$. □

2.2 The case $q < \infty$

We discuss briefly also the case when $q < \infty$. It turns out, that in this case the logarithmic term disappears.

Proposition 9. *Let $n \geq 2$ and $0 < p \leq q < \infty$. Then*

(i) $\mathbb{E}\, \|x\|_q \approx n^{1/q}$,

(ii)
$$\sigma_0^{p,q}(\mu_p) = \int_{\Delta_p^n} \|x\|_q d\mu_p(x) \approx \frac{\mathbb{E}\, \|x\|_q}{n^{1/p}},$$

(iii) $\sigma_0^{p,q}(\mu_p) \approx n^{1/q-1/p}$

and all three statements hold with constants of equivalence independent of n.

Proof. (i) The following two inequalities may be easily proved by Hölder's and Minkowski inequality.

$$\Big(\sum_{j=1}^n (\mathbb{E}x_j)^q\Big)^{1/q} \leq \mathbb{E}\big(\sum_{j=1}^n x_j^q\big)^{1/q} \leq \big(\sum_{j=1}^n \mathbb{E}x_j^q\big)^{1/q}, \qquad q \geq 1,$$
$$\big(\sum_{j=1}^n \mathbb{E}x_j^q\big)^{1/q} \leq \mathbb{E}\big(\sum_{j=1}^n x_j^q\big)^{1/q} \leq \Big(\sum_{j=1}^n (\mathbb{E}x_j)^q\Big)^{1/q}, \qquad q \leq 1.$$

This gives for $q \geq 1$

$$\mathbb{E}\|x\|_q \leq n^{1/q}(\mathbb{E}x_j^q)^{1/q} \quad \text{and} \quad \mathbb{E}\|x\|_q \geq n^{1/q}\mathbb{E}x_j$$

and for $q \leq 1$

$$\mathbb{E}\|x\|_q \leq n^{1/q}\mathbb{E}x_j \quad \text{and} \quad \mathbb{E}\|x\|_q \geq n^{1/q}(\mathbb{E}x_j^q)^{1/q}.$$

Let us note, that the value of $\mathbb{E}x_j$ and $(\mathbb{E}x_j^q)^{1/q}$ does not depend on n, only on p and q.

(ii) The proof of the second part resembles very much the proof of Lemma 4 and is left to the reader.

(iii) The last point follows immediately from (i) and (ii). □

Remark 4. A similar statement to Proposition 9 is included in [30, Lemma 2, point 4].

3 Normalized surface measure

Definition 10. Let $n \geq 2$ be a natural number. We denote by

$$\varrho_p(\mathcal{A}) = \frac{\mathcal{H}(\mathcal{A})}{\mathcal{H}(\Delta_p^n)}, \qquad \mathcal{A} \subset \Delta_p^n$$

the normalized $n-1$ dimensional Hausdorff measure on Δ_p^n.

Let us mention, that for $p \in \{1, 2, \infty\}$ the measure ϱ_p coincides with μ_p.

Lemma 11. *Let $0 < p < \infty$ and $n \geq 2$. Then ϱ_p is an absolutely continuous measure with respect to μ_p and for μ_p almost every $x \in \Delta_p^n$ it holds*

$$\frac{d\varrho_p}{d\mu_p}(x) = \frac{n\lambda([0,1]\cdot\Delta_p^n)}{\mathcal{H}(\Delta_p^n)}\Big\|\nabla(\|\cdot\|_p)(x)\Big\|_2 = C_{p,n}^{-1}\Big(\sum_{i=1}^n x_i^{2p-2}\Big)^{1/2},$$

where

$$C_{p,n} = \int_{\Delta_p^n}\Big(\sum_{i=1}^n x_i^{2p-2}\Big)^{1/2} d\mu_p(x)$$

is the normalizing constant.

Proof. The proof imitates the proof of [26, Lemma 1 and Lemma 2], where the statement was proven for $1 \leq p < \infty$. Hence, we may assume, that $0 < p < 1$. First, we introduce some notation.

We fix $x = (x_1, \ldots, x_n) \in \Delta_p^n$, such that

- the mapping $y \to \|y\|_p$ is differentiable at x,
- x is a density point of $\mathcal{H}$, i.e.

$$\lim_{\varepsilon\to 0+}\frac{\mathcal{H}(B(x,\varepsilon)\cap\Delta_p^n)}{\varepsilon^{n-1}V_{n-1}} = 1, \tag{17}$$

  where V_{n-1} denotes the Lebesgue volume of the $n-1$ dimensional Euclidean unit ball.

- $x_i > 0$ for all $i = 1, \ldots, n$.

Obviously, ϱ_p-almost every $x \in \Delta_p^n$ satisfies all the three properties (we refer for example to [24, Theorem 16.2] for the second one).

Furthermore, we put $z := \nabla(\|\cdot\|_p)(x)$. This means, that

$$\|x + y\|_p = 1 + \langle z, y\rangle + r(y), \tag{18}$$

where

$$\theta(\delta) := \sup\left\{\frac{|r(y)|}{\|y\|_2} : 0 < \|y\|_2 \leq \delta\right\}, \quad \delta > 0$$

tends to zero if δ tends to zero. Using (18) for $y = \delta x$, one observes, that $\langle z, x\rangle = 1$. We denote by $H = x + z^\perp$ the tangent hyperplane to Δ_p^n at x. Let us note, that for $0 < p < 1$ the set $\mathbb{R}_+^n \setminus [0,1)\cdot\Delta_p^n = [1,\infty)\cdot\Delta_p^n$ is convex. Next, we show, that $\langle z, y\rangle \geq 1$ for every $y \in [1,\infty)\cdot\Delta_p^n$. Indeed,

$$\begin{aligned}1 \leq \|x + \lambda(y-x)\|_p &= 1 + \langle z, \lambda(y-x)\rangle + r(\lambda(y-x))\\ &= 1 - \lambda + \lambda\langle z, y\rangle + r(\lambda(y-x))\end{aligned}$$

Dividing by $\lambda > 0$ and letting $\lambda \to 0$ gives the statement.

The proof of the lemma is based on the following two inclusions, namely

$$[0,1]\cdot\Big(B(x,\varepsilon(1-\theta(\varepsilon)))\cap H\Big)\subset[0,1]\cdot\Big(B(x,\varepsilon)\cap\Delta_p^n\Big) \tag{19}$$

and

$$[0,1]\cdot\Big(B(x,\varepsilon)\cap\Delta_p^n\Big)\subset[0,1+\varepsilon\theta(\varepsilon)]\cdot\Big(B(x,\varepsilon(1+\theta(\varepsilon)\|x\|_2))\cap H\Big), \tag{20}$$

which hold for all $\varepsilon>0$ small enough.

First, we prove (19). To given $0\le s\le 1$ and $v\in B(x,\varepsilon(1-\theta(\varepsilon))\cap H$ we need to find $0\le t\le 1$ and $w\in B(x,\varepsilon)\cap\Delta_p^n$, such that $sv=tw$. To do this, we set

$$w:=\frac{v}{\|v\|_p}\in\Delta_p^n\quad\text{and}\quad t:=s\|v\|_p.$$

We need to show, that $t\le 1$ and $\|x-w\|_2\le\varepsilon$.

We choose $0<\varepsilon\le\min_i x_i$. Then

$$x_i\le|x_i-v_i|+v_i\le\|x-v\|_2+v_i\le\varepsilon+v_i$$

for every $i=1,\dots,n$, which implies, that $v_i\ge 0$ and $v\in\mathbb{R}_+^n$. From $v\in H$ and $v\in\mathbb{R}_+^n$ we deduce, that $\|v\|_p\le 1$. Hence $t=s\|v\|_p\le\|v\|_p\le 1$.

Next, we write

$$\begin{aligned}\|x-w\|_2&=\Big\|x-\frac{v}{\|v\|_p}\Big\|_2\le\|x-v\|_2+\Big\|v-\frac{v}{\|v\|_p}\Big\|_2\\&\le\varepsilon(1-\theta(\varepsilon))+\|v\|_2\cdot\frac{1-\|v\|_p}{\|v\|_p}\le\varepsilon(1-\theta(\varepsilon))+1-\|v\|_p\\&=\varepsilon(1-\theta(\varepsilon))+1-\{1+\langle v-x,z\rangle+r(v-x)\}\\&=\varepsilon(1-\theta(\varepsilon))+r(v-x)\le\varepsilon.\end{aligned}$$

Next, we prove (20). We need to find to given $0\le t\le 1$ and $w\in B(x,\varepsilon)\cap\Delta_p^n$ some $0\le s\le 1+\varepsilon\theta(\varepsilon)$ and $v\in B(x,\varepsilon(1+\theta(\varepsilon)\|x\|_2))\cap H$, such that $tw=sv$. We put

$$s:=t\langle w,z\rangle\quad\text{and}\quad v:=\frac{w}{\langle w,z\rangle}.$$

Let us recall, that we have shown above, that $w\in\Delta_p^n$ implies that $\langle w,z\rangle\ge 1$.

Of course, $tw=sv$ and $v\in H$ (as $\langle v,z\rangle=1$). Hence, it remains to show, that $s\le 1+\varepsilon\theta(\varepsilon)$ and $\|v-x\|_2\le\varepsilon(1+\theta(\varepsilon)\|x\|_2)$.

The application of (18) gives

$$1=\|w\|_p=\|x+(w-x)\|_p=1+\langle w-x,z\rangle+r(w-x),$$

which again forces $\langle w,z\rangle\le 1+\varepsilon\theta(\varepsilon)$. Then $s=t\langle w,z\rangle\le\langle w,z\rangle\le 1+\varepsilon\theta(\varepsilon)$.

Finally, we write

$$\begin{aligned}\|v-x\|_2&=\Big\|\frac{w}{\langle w,z\rangle}-x\Big\|_2\le\Big\|\frac{w}{\langle w,z\rangle}-\frac{x}{\langle w,z\rangle}\Big\|_2+\Big\|\frac{x}{\langle w,z\rangle}-x\Big\|_2\\&\le\frac{\|w-x\|_2}{\langle w,z\rangle}+\|x\|_2\frac{\langle w,z\rangle-1}{\langle w,z\rangle}\le\varepsilon+\varepsilon\theta(\varepsilon)\|x\|_2.\end{aligned}$$

Equipped with (19) and (20), we may finish the proof of the lemma. We write

$$\begin{aligned}\lim_{\varepsilon\to 0}\frac{\varrho_p(B(x,\varepsilon)\cap\Delta_p^n)}{\mu_p(B(x,\varepsilon)\cap\Delta_p^n)} &= \lim_{\varepsilon\to 0}\frac{\mathcal{H}(B(x,\varepsilon)\cap\Delta_p^n)}{\mathcal{H}(\Delta_p^n)}\cdot\frac{\varepsilon^{n-1}V_{n-1}}{\varepsilon^{n-1}V_{n-1}}\cdot\frac{\lambda([0,1]\cdot\Delta_p^n)}{\lambda([0,1]\cdot[B(x,\varepsilon)\cap\Delta_p^n])}\\ &= \frac{\lambda([0,1]\cdot\Delta_p^n)}{\mathcal{H}(\Delta_p^n)}\cdot\lim_{\varepsilon\to 0}\frac{\varepsilon^{n-1}V_{n-1}}{\lambda([0,1]\cdot[B(x,\varepsilon)\cap\Delta_p^n])},\end{aligned}\tag{21}$$

where we have used (17). As the perpendicular distance between zero and H is equal to $1/\|z\|_2$, we observe, that

$$\mathrm{vol}(B(x,a)\cap H)=\frac{a^{n-1}V_{n-1}}{n\|z\|_2}$$

holds for every $a>0$. Using this, we get from (19) and (20)

$$\begin{aligned}\lambda\Big([0,1]\cdot\Big(B(x,\varepsilon(1-\theta(\varepsilon)))\cap H\Big)\Big) &= \frac{[\varepsilon(1-\theta(\varepsilon))]^{n-1}V_{n-1}}{n\|z\|_2}\\ &\le \lambda\Big([0,1]\cdot\Big(B(x,\varepsilon)\cap\Delta_p^n\Big)\Big)\\ &\le \lambda\Big([0,1+\varepsilon\theta(\varepsilon)]\cdot\Big(B(x,\varepsilon(1+\theta(\varepsilon)\|x\|_2))\cap H\Big)\Big)\\ &= [1+\varepsilon\theta(\varepsilon)]^n\cdot\frac{[\varepsilon(1+\theta(\varepsilon)\|x\|_2)]^{n-1}V_{n-1}}{n\|z\|_2}.\end{aligned}$$

Combining these estimates with (21) gives the result. □

Lemma 12. *Let $0<p<\infty$ and $n\ge 2$. Then*

$$\sigma_0^{p,\infty}(\varrho_p)=\int_{\Delta_p^n}x_1^*d\varrho_p=\frac{\displaystyle\int_{\Delta_p^n}x_1^*\Big(\sum_{i=1}^n x_i^{2p-2}\Big)^{1/2}d\mu_p(x)}{\displaystyle\int_{\Delta_p^n}\Big(\sum_{i=1}^n x_i^{2p-2}\Big)^{1/2}d\mu_p(x)}\approx\frac{\mathbb{E}\,x_1^*\Big(\sum_{i=1}^n x_i^{2p-2}\Big)^{1/2}}{\mathbb{E}\Big(\sum_{i=1}^n x_i^{2p-2}\Big)^{1/2}}\cdot n^{-1/p}.\tag{22}$$

Proof. Only the last equivalence needs a proof. It resembles the proof of Lemma 4 and is again based on the polar decomposition formula (12).

We plug the functions

$$f_1(x)=x_1^*\Big(\sum_{i=1}^n x_i^{2p-2}\Big)^{1/2}e^{-x_1^p-\dots-x_n^p}\quad\text{and}\quad f_2(x)=\Big(\sum_{i=1}^n x_i^{2p-2}\Big)^{1/2}e^{-x_1^p-\dots-x_n^p}$$

into (12) and obtain

$$\sigma_0^{p,\infty}(\varrho_p) = \frac{\displaystyle\int_{\mathbb{R}^n_+} f_1(x)dx \cdot \int_0^\infty r^{n+p-2}e^{-r^p}dr}{\displaystyle\int_{\mathbb{R}^n_+} f_2(x)dx \cdot \int_0^\infty r^{n+p-1}e^{-r^p}dr}$$

$$= \frac{\mathbb{E}\, x_1^* \Big(\sum_{i=1}^n x_i^{2p-2}\Big)^{1/2}}{\mathbb{E}\Big(\sum_{i=1}^n x_i^{2p-2}\Big)^{1/2}} \cdot \frac{\Gamma(n/p+1-1/p)}{\Gamma(n/p+1)}.$$

By Stirling's formula, the last expression is equivalent to $n^{-1/p}$. □

Theorem 13. *Let $0 < p < \infty$ and $n \geq 2$. Then*

$$\sigma_0^{p,\infty}(\varrho_p) \lesssim \left[\frac{\log(n+1)}{n}\right]^{1/p}. \tag{23}$$

Proof. We define a probability measure $\alpha_{p,n}$ on $\mathbb{R}_n^+$ by the density

$$\tilde{C}_{p,n}^{-1} \cdot \left(\sum_{i=1}^n x_i^{2p-2}\right)^{1/2} e^{-x_1^p-\cdots-x_n^p}, \quad \tilde{C}_{p,n} := \int_{\mathbb{R}^n_+} \left(\sum_{i=1}^n x_i^{2p-2}\right)^{1/2} e^{-x_1^p-\cdots-x_n^p}dx$$

with respect to the Lebesgue measure. Let us note, that due to the inequality

$$\left(\sum_{i=1}^n x_i^{2p-2}\right)^{1/2} \leq \sum_{i=1}^n x_i^{p-1}$$

the integral in the definition of $\tilde{C}_{p,n}$ really converges and $\alpha_{p,n}$ is well defined.

According to Lemma 12, we need to estimate

$$\int_{\mathbb{R}^n_+} x_1^* d\alpha_{p,n}(x).$$

We calculate for $\delta > 1$, which is to be chosen later on,

$$\int_{\mathbb{R}^n_+} x_1^* d\alpha_{p,n}(x) = \int_0^\infty \alpha_{p,n}(x_1^* > t)dt \leq \delta + \int_\delta^\infty \alpha_{p,n}(x_1^* > t)dt$$

$$\leq \delta + n\int_\delta^\infty \alpha_{p,n}(x_1 > t)dt.$$

We write $x' = (x_2, \dots, x_n) \in \mathbb{R}_+^{n-1}$. Then

$$\begin{aligned}
\alpha_{p,n}(x_1 > t) &= \tilde{C}_{p,n}^{-1} \int_t^\infty e^{-x_1^p} \int_{\mathbb{R}_+^{n-1}} \left(\sum_{i=1}^n x_i^{2p-2} \right)^{1/2} e^{-x_2^p - \dots - x_n^p} dx' dx_1 \\
&\le \tilde{C}_{p,n}^{-1} \int_t^\infty e^{-x_1^p} \int_{\mathbb{R}_+^{n-1}} \left[x_1^{p-1} + \left(\sum_{i=2}^n x_i^{2p-2} \right)^{1/2} \right] e^{-x_2^p - \dots - x_n^p} dx' dx_1 \\
&= \tilde{C}_{p,n}^{-1} \int_t^\infty e^{-x_1^p} x_1^{p-1} dx_1 \cdot \int_{\mathbb{R}_+^{n-1}} e^{-x_2^p - \dots - x_n^p} dx' \\
&\quad + \tilde{C}_{p,n}^{-1} \int_t^\infty e^{-x_1^p} dx_1 \cdot \int_{\mathbb{R}_+^{n-1}} \left(\sum_{i=2}^n x_i^{2p-2} \right)^{1/2} e^{-x_2^p - \dots - x_n^p} dx' \\
&:= I_1 + I_2.
\end{aligned}$$

The inequality

$$\begin{aligned}
c_p^n \tilde{C}_{p,n} &= c_p^n \int_{\mathbb{R}_+^n} \left(\sum_{i=1}^n x_i^{2p-2} \right)^{1/2} e^{-x_1^p - \dots - x_n^p} dx \\
&\ge c_p^n \int_{\mathbb{R}_+^n} \left(\sum_{i=2}^n x_i^{2p-2} \right)^{1/2} e^{-x_1^p - \dots - x_n^p} dx \qquad (24) \\
&= c_p^n \int_0^\infty e^{-x_1^p} dx_1 \int_{\mathbb{R}_+^{n-1}} \left(\sum_{i=2}^n x_i^{2p-2} \right)^{1/2} e^{-x_2^p - \dots - x_n^p} dx' = c_p^{n-1} \tilde{C}_{p,n-1}
\end{aligned}$$

shows, that

$$I_1 = \frac{c_p \int_t^\infty x_1^{p-1} e^{-x_1^p} dx_1}{c_p^n \tilde{C}_{p,n}} \le \frac{c_p \int_t^\infty x_1^{p-1} e^{-x_1^p} dx_1}{c_p \tilde{C}_{p,1}} = \tilde{C}_{p,1}^{-1} \cdot \frac{e^{-t^p}}{p}.$$

Using (24) again, we get also

$$I_2 = \tilde{C}_{p,n}^{-1} \cdot \tilde{C}_{p,n-1} \int_t^\infty e^{-x_1^p} dx_1 \le c_p \int_t^\infty e^{-x_1^p} dx_1 = \frac{c_p}{p} \cdot \int_{t^p}^\infty s^{1/p-1} e^{-s} ds.$$

If $p \ge 1$, we get

$$I_1 + I_2 \le c e^{-t^p}, \quad t > 1 \qquad (25)$$

and

$$\int_{\mathbb{R}_+^n} x_1^* d\alpha_{p,n}(x) \le \delta + cn \int_\delta^\infty e^{-t^p} dt \le \delta + c' n e^{-\delta^p}.$$

By choosing $\delta \approx \log(n+1)^{1/p}$, we get the result.

If $p < 1$, we use the second estimate of Lemma 5 and replace (25) with

$$I_1 + I_2 \le c t^{1-p} e^{-t^p}, \quad t > t_0$$

for $t_0 > 1$ large enough and the result again follows by the choice of δ.

□

Remark 5. (i) Theorem 13 shows, that the average size of the largest coordinate of $x \in \Delta_p^n$ taken with respect to the normalized Hausdorff measure is again only slightly larger than $n^{-1/p}$. Hence, also in this case, the typical element of Δ_p^n seems to be far from being sparse and resembles rather properly normalized white noise in the sense described in Introduction.

(ii) Using interpolation inequality (4), one may again obtain a similar estimate also for $0 < p \le q < \infty$, namely

$$\sigma_0^{p,q}(\varrho_p) \lesssim \left[\frac{\log(n+1)}{n}\right]^{1/p-1/q}.$$

It would be probably possible to avoid the logarithmic terms and provide improved estimates also for $m > 0$, but we shall not go into this direction. Our main aim of this section was to show, that normalized Hausdorff measure does not prefer sparse (or nearly sparse) vectors, and this was clearly demonstrated by Theorem 13.

4 Tensor product measures

As discussed already in the Introduction and proved in Theorem 7 and Theorem 13, the average vector of Δ_p^n with respect to the cone measure μ_p and with respect to surface measure ϱ_p behaves "badly" meaning that (roughly speaking) many of its coordinates are approximately of the same size. As promised before, we shall now introduced a new class of measures, for which the random vector behaves in a completely different way. This measures are defined through their density with respect to the cone measure μ_p. This density has a strong singularity near the points with vanishing coordinates.

Definition 14. Let $0 < p < \infty$, $\beta > -1$ and $n \ge 2$. Then we define the probability measure $\theta_{p,\beta}$ on Δ_p^n by

$$\frac{d\theta_{p,\beta}}{d\mu_p}(x) = C_{p,\beta}^{-1} \cdot \prod_{i=1}^n x_i^\beta, \quad x \in \Delta_p^n, \tag{26}$$

where

$$C_{p,\beta} = \int_{\Delta_p^n} \prod_{i=1}^n x_i^\beta d\mu_p(x). \tag{27}$$

Remark 6. (i) If $0 > \beta > -1$, then (26) defines the density of $\theta_{p,\beta}$ with respect to μ_p only for points, where $x_i \neq 0$ for all $i = 1, \dots, n$. That means, that this density is defined μ_p-almost everywhere. The definition is then complemented by the statement, that $\theta_{p,\beta}$ is absolutely continuous with respect to μ_p.

(ii) We shall see later on, that the condition $\beta > -1$ ensures, that (27) is finite.

Lemma 15. *Let $0 < p < \infty$, $\beta > -1$ and $n \ge 2$.*

(i) Let $1 \le m \le n$. *Then*

$$\sigma^{p,\infty}_{m-1}(\theta_{p,\beta}) = \int_{\Delta^n_p} x^*_m d\theta_{p,\beta} = \frac{\mathbb{E}\, x^*_m \prod_{i=1}^n x_i^\beta}{\mathbb{E} \prod_{i=1}^n x_i^\beta} \cdot \frac{\Gamma(n(\beta+1)/p)}{\Gamma(n(\beta+1)/p+1/p)}.$$

(ii)

$$\mathbb{E} \prod_{i=1}^n x_i^\beta = \left[\frac{c_p}{p} \cdot \Gamma((\beta+1)/p) \right]^n.$$

Proof. The proof of the first part follows again by (12), this time used for the functions

$$f_1(x) = x^*_m \Big(\prod_{i=1}^n x_i^\beta \Big) e^{-x_1^p - \dots - x_n^p} \quad \text{and} \quad f_2(x) = \Big(\prod_{i=1}^n x_i^\beta \Big) e^{-x_1^p - \dots - x_n^p}.$$

The proof of the second part is straightforward. □

It follows directly from (8), that $\Gamma(s)$ tends to infinity, when s tends to zero. The following lemma quantifies this phenomenon. Although the statement seems to be well known, we were not able to find a reference and we therefore provide at least a sketch of the proof.

Lemma 16. *Let* $C \simeq 0.577\dots$ *denote the Euler constant. Then*

$$\lim_{n\to\infty} \left(\frac{\Gamma(1/n)}{n} \right)^n = e^{-C}.$$

Proof. It is enough to show, that

$$\lim_{n\to\infty} n \cdot \log(\Gamma(1+1/n)) = -C,$$

which (by using the l'Hospital rule) follows from

$$\lim_{n\to\infty} \frac{\int_0^\infty s^{1/n} e^{-s} \log s \, ds}{\int_0^\infty s^{1/n} e^{-s} ds} = -C.$$

But the numerator of this fraction is equal to $\Gamma'(1+1/n)$ and its denominator to $\Gamma(1+1/n)$. The whole fraction is therefore equal to $\Psi(1+1/n)$ and $\Psi(1+1/n) \to \Psi(1) = -C$ as n tends to infinity, cf. [1, Section 6.3.2, p. 258]. □

Next theorem shows, that if $\beta = p/n - 1$, then the measure $\theta_{p,\beta}$ promotes sparsity and one may even consider limiting behavior of n growing to infinity.

Theorem 17. *Let* $0 < p < \infty$ *and let* $n \ge 2$ *and* $1 \le m \le n$ *be integers. Then*

$$\sigma^{p,\infty}_{m-1}(\theta_{p,p/n-1}) \gtrsim \frac{\Gamma(n+1)}{\Gamma(n-m+1)} \cdot \frac{\Gamma(n/p+n-m+1)}{\Gamma(n/p+n+1)}, \tag{28}$$

and

$$\sigma_{m-1}^{p,\infty}(\theta_{p,p/n-1}) \lesssim \frac{\Gamma(n+1)}{\Gamma(n-m+1)} \left\{ \frac{\Gamma(n/p+n-m+1)}{\Gamma(n/p+n+1)} + \frac{1}{m!} \cdot \left(\frac{e^{-1}}{\Gamma(1/n)} \right)^m \right\} \quad (29)$$

where the constants do not depend on n and m, but may depend on p.

Furthermore, for every fixed $m \in \mathbb{N}$,

$$\frac{1}{\left(\frac{1}{p}+1\right)^m} \lesssim \liminf_{n\to\infty} \sigma_{m-1}^{p,\infty}(\theta_{p,p/n-1}) \quad (30)$$

$$\leq \limsup_{n\to\infty} \sigma_{m-1}^{p,\infty}(\theta_{p,p/n-1}) \lesssim \frac{1}{\left(\frac{1}{p}+1\right)^m} + \frac{e^{-m}}{m!},$$

where the constants do not depend on m, but may depend on p.

Proof. First observe, that $n(\beta+1)/p = 1$ for $\beta = p/n - 1$ and therefore

$$\frac{\Gamma(n(\beta+1)/p)}{\Gamma(n(\beta+1)/p+1/p)} = \frac{1}{\Gamma(1+1/p)}$$

depends only on p. Due to Lemma 15, we have to estimate

$$\mathbb{E}\, x_m^* \left(\prod_{i=1}^n x_i^{p/n-1} \right) = c_p^n \int_{\mathbb{R}_+^d} x_m^* \prod_{i=1}^n x_i^{p/n-1} e^{-x_1^p - \cdots - x_n^p} dx. \quad (31)$$

Let $t = x_m^*$ and let us assume, that there is only one coordinate $j = 1, \ldots, n$, such that $x_j = t$. Obviously, this assumption holds almost everywhere. Of course, we have n possibilities for j. Furthermore, $m-1$ from the remaining $n-1$ components of x are bigger than t and the remaining $n-m$ components are smaller. This allows to rewrite (31) as

$$c_p^n\, n \binom{n-1}{m-1} \int_0^\infty t^{p/n} e^{-t^p} \left(\int_0^t u^{p/n-1} e^{-u^p} du \right)^{n-m} \times$$

$$\times \left(\int_t^\infty u^{p/n-1} e^{-u^p} du \right)^{m-1} dt$$

$$= \frac{c_p^n n}{p^n} \binom{n-1}{m-1} \int_0^\infty \omega^{1/p+1/n-1} e^{-\omega} \left(\int_0^\omega s^{1/n-1} e^{-s} ds \right)^{n-m} \times$$

$$\times \left(\int_\omega^\infty s^{1/n-1} e^{-s} ds \right)^{m-1} d\omega.$$

Let us denote

$$\gamma = \Gamma(1/n) = \int_0^\infty s^{1/n-1} e^{-s} ds \quad \text{and} \quad y(\omega) = \gamma^{-1} \cdot \int_0^\omega s^{1/n-1} e^{-s} ds. \quad (32)$$

Then $y(\omega)$ is a monotone function of y, $y(0) = 0$ and $\lim_{\omega\to\infty} y(\omega) = 1$. We denote by $\omega(y)$ its inverse function, i.e.

$$y = \gamma^{-1} \cdot \int_0^{\omega(y)} s^{1/n-1} e^{-s} ds, \quad 0 \leq y \leq 1. \quad (33)$$

Using this notation, we obtain

$$\mathbb{E}\, x_m^* \left(\prod_{i=1}^n x_i^{p/n-1} \right) = \frac{c_p^n\, \gamma^n}{p^n}\, n \binom{n-1}{m-1} \int_0^1 \omega(y)^{1/p} y^{n-m} (1-y)^{m-1} dy$$

and

$$\sigma_{m-1}^{p,\infty}(\theta_{p,p/n-1}) = \frac{\Gamma(n+1)}{\Gamma(m)\Gamma(n-m+1)} \int_0^1 \omega(y)^{1/p} y^{n-m} (1-y)^{m-1} dy, \tag{34}$$

where $\omega(y)$ is given by (33).

Step 1. Estimate from below

The estimate

$$\gamma y = \int_0^{\omega(y)} s^{1/n-1} e^{-s} ds \le \int_0^{\omega(y)} s^{1/n-1} ds = n\omega(y)^{1/n}$$

implies together with Lemma 16

$$\omega(y) \ge \left(\frac{\gamma y}{n} \right)^n \ge c y^n$$

with c independent of n. This gives finally

$$\begin{aligned} \sigma_{m-1}^{p,\infty}(\theta_{p,p/n-1}) &\ge c^{1/p} \cdot \frac{\Gamma(n+1)}{\Gamma(m)\Gamma(n-m+1)} \cdot \int_0^1 y^{n/p+n-m} (1-y)^{m-1} dy \\ &= c^{1/p} \cdot \frac{\Gamma(n+1)}{\Gamma(m)\Gamma(n-m+1)} \cdot B(n/p+n-m+1, m) \\ &= c^{1/p} \cdot \frac{\Gamma(n+1)}{\Gamma(n-m+1)} \cdot \frac{\Gamma(n/p+n-m+1)}{\Gamma(n/p+n+1)}, \end{aligned}$$

where we used the Beta function (9) and the proof of (28) is complete.

Step 2. Estimate from above

Let us first take y, such that $1 - e^{-1}/\gamma \le y \le 1$. Then $-\ln(\gamma(1-y)) \ge 1$ and

$$\int_{-\ln(\gamma(1-y))}^{\infty} s^{1/n-1} e^{-s} ds \le \int_{-\ln(\gamma(1-y))}^{\infty} e^{-s} ds = \gamma(1-y).$$

Hence,

$$\omega(y) \le -\ln(\gamma(1-y)), \quad 1 - e^{-1}/\gamma \le y \le 1. \tag{35}$$

Finally, we observe, that

$$f : y \to \int_{Cy^n}^{\infty} s^{1/n-1} e^{-s} ds$$

is a convex function on $\mathbb{R}_+$, $f(0) = \gamma$ and

$$\begin{aligned} f(1 - e^{-1}/\gamma) &= \int_{C(1-e^{-1}/\gamma)^n}^{\infty} s^{1/n-1} e^{-s} ds \\ &\le \int_1^{\infty} s^{1/n-1} e^{-s} ds \le e^{-1}, \end{aligned}$$

if we choose C so large, that $C(1-e^{-1}/\gamma)^n \geq 1$ for all $n \in \mathbb{N}$. This is indeed possible, while a byproduct of Lemma 16 is also a relation $\lim_{n\to\infty} \gamma/n = 1$. Using the convexity of f, we obtain

$$f(y) \leq \gamma(1-y), \quad 0 \leq y \leq 1-e^{-1}/\gamma,$$

which further leads to

$$\omega(y) \leq Cy^n, \quad 0 \leq y \leq 1-e^{-1}/\gamma. \tag{36}$$

We insert (35) and (36) into (34) and obtain

$$\sigma^{p,\infty}_{m-1}(\theta_{p,p/n-1}) \leq \frac{\Gamma(n+1)}{\Gamma(m)\Gamma(n-m+1)} \left\{ C^{1/p} I_1 + I_2 \right\}, \tag{37}$$

where

$$I_1 := \int_0^{1-e^{-1}/\gamma} y^{n/p+n-m}(1-y)^{m-1}dy$$

and

$$I_2 := \int_{1-e^{-1}/\gamma}^1 |\ln(\gamma(1-y))|^{1/p} y^{n-m}(1-y)^{m-1}dy.$$

The first integral may be estimated again using the Beta function, which gives

$$I_1 \leq B(n/p+n-m+1, m). \tag{38}$$

We denote by k the uniquely defined integer, such that $1/p \leq k < 1/p+1$ holds, and estimate

$$I_2 \leq \int_{1-e^{-1}/\gamma}^1 |\ln(\gamma(1-y))|^{1/p}(1-y)^{m-1}dy \leq I_{k,m} := \int_0^{e^{-1}/\gamma} |\ln(\gamma y)|^k y^{m-1}dy.$$

Next, we use partial integration to estimate $I_{k,m}$. We obtain

$$I_{k,m} = \frac{1}{m}\left(\frac{e^{-1}}{\gamma}\right)^m + \frac{k}{m} \cdot I_{k-1,m}.$$

Together with $I_{0,m} = 1/m \cdot (e^{-1}/\gamma)^m$, this leads finally to

$$I_{k,m} \leq \frac{(k+1)!}{m}\left(\frac{e^{-1}}{\gamma}\right)^m.$$

This, together with (37) and (38) finishes the proof of (29).

The proof of (30) then follows directly by Stirling's formula (10). □

Remark 7. (i) Let us take $m = 0$. Then the formula (30) describes an essentially different behavior compared to the normalized cone and surface measure. Namely, the expected value of the largest coordinate of $x \in \Delta^n_p$ with respect to $\theta_{p,p/n-1}$ does not decay to zero with n growing to infinity. We shall demonstrate this effect also numerically in next section.

(ii) If $m > 0$, then (30) shows, that $\sigma_m^{p,\infty}(\theta_{p,p/n-1})$ decays exponentially fast with m, as soon as n is large enough. That means, that for n large enough, the average vector of Δ_p^n exhibits a strong sparsity-like structure. Namely, its m-th largest component decays exponentially with m.

(iii) We have chosen in (26) a different β for each n, namely $\beta_n = p/n - 1 > -1$. This was of course a crucial ingredient in the proof of Theorem 17. It is not difficult to modify the analysis of the proof of Theorem 17 to the situation, when $\beta > -1$ is fixed for all $n \in \mathbb{N}$. In this case we obtain again, that (up to logarithmic factors) $\sigma_0^{p,\infty}(\theta_{p,\beta})$ is equivalent to $n^{-1/p}$ with constants of equivalence depending on $p > 0$ and $\beta > -1$.

(iv) Last, but not least, we observe, that one may choose $p = 1$ or even $p = 2$ in Theorem 17 and still obtains the exponential decay of coordinates as described by (30). It seems, that there is no significant connection between sparsity of an average vector of $x \in \Delta_p^n$ and the size of $p > 0$.

5 Numerical experiments

5.1 Cone measure

We would like to demonstrate the most significant effects of the theory also by numerical experiments. We start with the case of the cone measure. The key role is played by (13). It may be interpreted in the following way. To generate a random point on Δ_p^n with respect to the normalized cone measure, it is enough to generate $\omega_1, \dots, \omega_n$ with respect to the density $c_p e^{-t^p}, t > 0$ and then calculate

$$\frac{(\omega_1, \dots, \omega_n)}{\left(\sum_{j=1}^n \omega_j^p\right)^{1/p}} \in \Delta_p^n.$$

This method is very practical, as the running time of this algorithm depends only linearly on n.

Let us note, that the values of ω_i may be generated very easily. For example the package *GNU Scientific Library* [19] implements a random number generator with respect to the gamma distribution using the method described in the classical work of Knuth [21]. Using this package, we generated 10^8 random points $x \in \Delta_p^n$ for $n = 100$ and $p \in \{1/2, 1, 2\}$ to approximate numerically the value of $n^{1/p} \cdot \int_{\Delta_p^n} x_m^* d\mu_p(x)$. The result may be found in the Figure 1.

5.2 Tensor measures

It was observed already in [2], that the measures $\theta_{p,\beta}$ allow a formula similar to (13). We plug the function $f(x) = \chi_{[0,\infty)\cdot A} \prod_{i=1}^n x_i^\beta e^{-\|x\|_p^p}$ into (12), where A is any μ_p-measurable subset of Δ_p^n, and obtain

$$\int_{[0,\infty)\cdot A} \prod_{i=1}^n x_i^\beta e^{-\|x\|_p^p} d\lambda(x) = \lambda([0,1]\cdot\Delta_p^n)\cdot n \cdot \int_0^\infty r^{n-1+n\beta} e^{-r^p} dr \cdot \int_A \prod_{i=1}^n x_i^\beta d\mu_p(x).$$

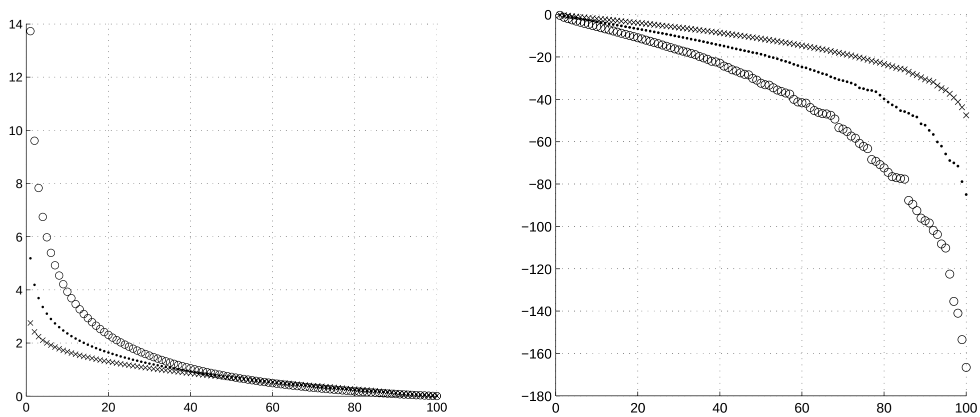
We use a similar formula also for $A=\Delta_p^n$, which leads to

$$\int_A 1 d\,\theta_{p,\beta} = \frac{\displaystyle\int_A \prod_{i=1}^n x_i^\beta d\mu_p(x)}{\displaystyle\int_{\Delta_p^n} \prod_{i=1}^n x_i^\beta d\mu_p(x)} = \frac{\displaystyle\int_{[0,\infty)\cdot A} \prod_{i=1}^n x_i^\beta e^{-\|x\|_p^p} dx}{\displaystyle\int_{\mathbb{R}_+^n} \prod_{i=1}^n x_i^\beta e^{-\|x\|_p^p} dx}.$$

A random point on Δ_p^n with respect to $\theta_{p,\beta}$ may therefore be generated in the following way. We generate $\omega_1',\dots,\omega_n'$ with respect to the density $c_{p,\beta}t^\beta e^{-t^p}, t>0$, where $c_{p,\beta}^{-1}=\int_0^\infty t^\beta e^{-t^p}dt$ is a normalizing constant and we consider the vector

$$\frac{(\omega_1',\dots,\omega_n')}{\left(\sum_{j=1}^n (\omega_j')^p\right)^{1/p}} \in \Delta_p^n.$$

Also this may be easily done with the help of [19]. We generated again 10^8 random points $x\in\Delta_p^n$ with respect to $\theta_{p,p/n-1}$ for $n=100$ and $p\in\{1/2,1,2\}$. Then we used those points to numerically approximate the expression $\log_{10}(\int_{\Delta_p^n} x_m^* d\theta_{p,p/n-1})$.



(a) $n^{1/p}\cdot\int_{\Delta_p^n} x_m^* d\mu_p(x)$ (b) $\log_{10}(\int_{\Delta_p^n} x_m^* d\theta_{p,p/n-1})$

Figure 1: Approximations of $n^{1/p}\cdot\int_{\Delta_p^n} x_m^* d\mu_p(x)$ (left) and $\log_{10}(\int_{\Delta_p^n} x_m^* d\theta_{p,p/n-1})$ (right) for $n=100$, $p=1/2(\circ)$, $p=1(\bullet)$ and $p=2(\times)$ based on sampling of 10^8 random points.

Acknowledgments

I would like to thank to Stephan Dahlke, Massimo Fornasier, Aicke Hinrichs, Erich Novak and Henryk Woźniakowski for their interest in this topic. I acknowledge the financial support provided by the START-award "Sparse Approximation and Optimization in High Dimensions" of the Fonds zur Förderung der wissenschaftlichen Forschung (FWF, Austrian Science Foundation).

References

[1] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, U.S. Government Printing Office, Washington, D.C. 1964.

[2] F. Barthe, M. Csörnyei and A. Naor, *A note on simultaneous polar and Cartesian decomposition*, in: Geometric Aspects of Functional Analysis, Lecture Notes in Mathematics, Springer, Berlin, 2003.

[3] F. Barthe, O. Guédon, S. Mendelson and A. Naor, *A probabilistic approach to the geometry of the l_p^n-ball*, Ann. Probab. 33 (2005), no. 2, 480–513.

[4] C. Bennett and R. Sharpley, Interpolation of operators, Pure and Applied Mathematics, 129, Academic Press, Boston, 1988.

[5] J. Bobin, J.-L. Starck, J. M. Fadili, Y. Moudden and D. L. Donoho, *Morphological Component Analysis: An Adaptive Thresholding Strategy*, IEEE Trans. Image Process. 16 (2007), no. 11, 2675 – 2681.

[6] E. J. Candés, J. K. Romberg and T. Tao, *Stable signal recovery from incomplete and inaccurate measurements*, Comm. Pure Appl. Math. 59 (2006), no. 8, 1207–1223.

[7] E. J. Candés, *Compressive sampling*, In Proceedings of the International Congress of Mathematicians, Madrid, Spain, 2006.

[8] E. J. Candés, J. K. Romberg and T. Tao, *Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information*, IEEE Trans. Inform. Theory 52 (2006), no. 2, 489–509.

[9] E. J. Candés and T. Tao, *Decoding by linear programming*, IEEE Trans. Inform. Theory 51 (2005), no. 12, 4203–4215.

[10] F. Champagnat, Y. Goussard and J. Idier, *Unsupervised deconvolution of sparse spike trains using stochastic approximation*, IEEE Trans. Signal Process. 44 (1996), no. 12, 2988 – 2998.

[11] A. Cohen, W. Dahmen, and R. DeVore, *Compressed sensing and best k-term approximation*, J. Amer. Math. Soc. 22 (2009), no. 1, 211–231.

[12] S. Dahlke, E. Novak and W. Sickel, *Optimal approximation of elliptic problems by linear and nonlinear mappings I*, J. Complexity 22 (2006), no. 1, 29-49.

[13] R. A. DeVore, *Nonlinear approximation*, Acta Num. 51–150, (1998).

[14] R. A. DeVore, B. Jawerth and V. Popov, *Compression of wavelet decompositions*, Amer. J. Math. 114 (1992), no. 4, 737-785.

[15] D. L. Donoho, *Compressed sensing*, IEEE Trans. Inform. Theory 52 (2006), no. 4, 1289–1306.

[16] J. Edwards, *A treatise on the integral calculus*, Vol. II, Chelsea Publishing Company, New York, 1922.

[17] M. Fornasier, *Numerical methods for sparse recovery*, Theoretical Foundations and Numerical Methods for Sparse Recovery, (Massimo Fornasier Ed.) Radon Series on Computational and Applied Mathematics 9, 2010.

[18] S. Foucart and H. Rauhut, *A mathematical introduction to compressive sensing*, Appl. Numer. Harmon. Anal., Birkhuser, Boston, in preparation.

[19] GNU Scientific Library, `http://www.gnu.org/software/gsl/`

[20] R. Gribonval and K. Schnass, *Dictionary identification - sparse matrix factorisation via ℓ_1 minimisation*, IEEE Trans. Infor. Theory, to appear.

[21] D. E. Knuth, The Art of Computer Programming, Vol. 2: *Seminumerical Algorithms*, 3rd ed., Addison-Wesley 1998.

[22] M. Ledoux, The concentration of measure phenomenon, AMS, 2001.

[23] M. Ledoux and M. Talagrand, Probability in Banach spaces. Springer-Verlag, Berlin, 1991.

[24] P. Mattila, *Geometry of sets and measures in Euclidean Spaces*, Cambridge University Press, 1995.

[25] A. Naor, *The surface measure and cone measure on the sphere of l_p^n.* Trans. Amer. Math. Soc. 359 (2007), no. 3, 1045–1079.

[26] A. Naor and D. Romik, *Projecting the surface measure of the sphere of l_p^n*, Ann. Inst. H. Poincaré Probab. Statist. 39 (2003), no. 2, 241–261.

[27] K. Oskolkov, *Polygonal approximation of functions of two variables*, Math. USSR Sbornik 35, 851–861, (1979).

[28] J. C. Pesquet, H. Krim, D. Leporini and E. Hamman, *Bayesian approach to best basis selection*, In Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Proc., pages 2634 – 2637, 1996.

[29] H. Rauhut, *Compressive sensing and structured random matrices*, Theoretical Foundations and Numerical Methods for Sparse Recovery, (Massimo Fornasier Ed.) Radon Series on Computational and Applied Mathematics 9, 2010.

[30] G. Schechtman and J. Zinn, *On the volume of the intersection of two L_p^n balls*, Proc. AMS 110 (1), 217–224, (1990).

[31] E. Schmidt, *Zur Theorie der linearen und nichtlinearen Integralgleichungen I*, Math. Anal. 63, 433–476, (1907).

[32] V. N. Temlyakov, *Nonlinear methods of approximation*, Found. Comput. Math. 3 (2003), no. 1, 33-107.