



**HAL**  
open science

## Representation of gaussian small ball probabilities in $l_2$

André Mas

► **To cite this version:**

| André Mas. Representation of gaussian small ball probabilities in  $l_2$ . 2008. hal-00349575v1

**HAL Id: hal-00349575**

**<https://hal.science/hal-00349575v1>**

Preprint submitted on 2 Jan 2009 (v1), last revised 1 Mar 2011 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Representation of gaussian small ball probabilities in $l_2$

Andre Mas\*

## Abstract

Let  $z = \sum_{i=1}^{+\infty} x_i^2/a_i^2$  where the  $x_i$ 's are i.d.d centered with unit variance gaussian random variables and  $(a_i)_{i \in \mathbb{N}}$  an increasing sequence such that  $\sum_{i=1}^{+\infty} a_i^{-2} < +\infty$ . We propose an exponential-integral representation theorem for the gaussian small ball probability  $\mathbb{P}(z < \varepsilon)$  when  $\varepsilon \downarrow 0$ . We start from a result by Meyer-Wolf, Zeitouni (1993) and Dembo, Meyer-Wolf, Zeitouni (1995) who computed this probability by means of series. We prove that  $\mathbb{P}(z < \varepsilon)$  belongs to a class of functions introduced by de Haan, well-known in extreme value theory, the class Gamma, for which an explicit exponential-integral representation is available. The converse implication holds under a mild additional assumption. Some applications are underlined in connection with statistical inference for random functions.

Keywords : de Haan's Gamma class, small ball problems, regular variations, gaussian random elements.

## 1 Introduction

### 1.1 About non-shifted and shifted small ball problems

Small ball problems could generally be stated the following way : consider a random variable  $X$  with values in a general normed space  $(E, \|\cdot\|)$  (which may not be finite-dimensional) and estimate  $\mathbb{P}(\|X\| < \varepsilon)$  for small values of  $\varepsilon$ . This issue may be viewed as a counterpart of the large deviations or concentration problems (where  $\mathbb{P}(\|X\| > M)$  is studied for large  $M$ ) and the term "small deviations" is sometimes encountered to name small ball problems. The core of the literature on small ball problems focuses on gaussian random variables. The nice monograph by Li and Shao (2001) is a complete state of the art, introducing the main concepts and providing numerous references. Much attention has been given to the Brownian motion (when  $(E, \|\cdot\|) = (C(0,1), |\cdot|_\infty)$ ) or its relatives (fractional Brownian motion, Brownian sheet, etc). Authors also investigated the case of stable random elements (see for instance Li, Linde (2004)). Another issue is related to the norm. Indeed in infinite dimensional spaces, norms or metrics are not equivalent and this may influence the local behaviour of  $\mathbb{P}(\|X\| < \varepsilon)$ .

A more general question could be the shifted small ball probability :

$$\mathbb{P}(\|X - x_0\| < \varepsilon) \tag{1}$$

---

\*I3M, Université Montpellier 2, CC051, Place Eugène Bataillon, 34095 Montpellier Cedex 5. Email : mas@math.univ-montp2.fr

for a fixed  $x_0$ . A stumbling stone arises from the shift  $x_0$ . It turns out that, in general, computations cannot be carried out for any  $x_0$ . Several works focus on expliciting the set of those  $x_0$  for which the shifted small ball probability may be computed from the non-shifted one (when  $x_0 = 0$ ). We refer to Borell (1976) or Li and Linde (1993) for instance. A classical example stems from the situation where  $\mathbb{P}_{X-x_0} \ll \mathbb{P}_X$  where  $\mathbb{P}_X$  denotes the probability distribution induced by the random element  $X$ . The classical Cameron-Martin's theorem for Brownian motion illustrates this case for instance. Then :

$$\mathbb{P}(\|X - x_0\| < \varepsilon) = \int_{B(0,\varepsilon)} \mathbb{P}_{X-x_0}(dx) = \int_{B(0,\varepsilon)} f_{x_0}(x) \mathbb{P}_X(dx)$$

where  $f_{x_0} = d\mathbb{P}_{X-x_0}/d\mathbb{P}_X$  and  $B(0,\varepsilon)$  stands for the ball centered at 0 with radius  $\varepsilon$ . When  $f_{x_0}$  is regular enough in a neighborhood of zero :

$$\mathbb{P}(\|X - x_0\| < \varepsilon) \sim f_{x_0}(0) \mathbb{P}(\|X\| < \varepsilon).$$

About this fact see Proposition 2.1 in de Acosta (1983). In general the sharpness of existing results may vary, depending on the triplet  $((E, \|\cdot\|), \mathbb{P}_X, x_0)$  under consideration. Here  $\mathbb{P}_X$  denotes the distribution of  $X$ . In fact there are only few spaces for which the local behaviour of  $\mathbb{P}(\|X - x_0\| < \varepsilon)$  is explicitly described. Quite often lower and upper bounds are computed so that :

$$\mathbb{P}(\|X - x_0\| < \varepsilon) \asymp \varphi_{x_0}(\varepsilon)$$

where  $\varphi_{x_0}$  is known and  $f \asymp g$  means here that the positive functions  $f$  and  $g$  satisfy :

$$\limsup_0 \frac{f}{g} < +\infty,$$

$$\liminf_0 \frac{f}{g} < +\infty.$$

Sometimes only one of these bounds is accessible or needed.

It is worth noting or recalling a few crucial features of small deviations techniques. The Laplace transform, as well as in large deviations problems, is a major tool when coupled with the saddlepoint method. Small deviations are intimately connected with the entropy of the unit ball of the reproducing kernel Hilbert space associated with  $X$ , with the  $l$ -approximation numbers of  $X$  (i.e. the rate of approximation of  $X$  by a finite dimensional random variable, see Li, Linde (1999) ) or to the degree of compactness of linear operators generating  $X$  (see Li, Linde (2004)). All these notions are clearly connected to the regularity of the process  $X$ , when  $X$  is a process.

Applications of small ball probabilities are numerous : they appear when studying rates of convergence in the Law of the Iterated Logarithm (see Talagrand (1992), Kuelbs, Li, Linde (1994)) or the rate of escape of the Brownian motion (see Erickson (1980)). They even surprisingly provide a sufficient condition for the CLT (see Ledoux, Talagrand (1991), Theorem 10.13 p.289). However small ball problems remained until nowadays a matter essentially reserved to probability theory. It turns out that this topic may be of major interest in a recent and soaring area of statistics : functional data modelling. The

situation may be summarized this way : since Lebesgue's density of an infinite-dimensional random  $X$  does not exist, all the inferential techniques based on the density cannot hold anymore. In this framework, the small ball probabilities appear as a natural counterpart and should be investigated with much care. We illustrate this fact by pointing out an elementary example in the next subsection below.

First let us precise the  $l_2$  framework. Consider  $X$  a random variable defined the following way :

$$X = \left( \frac{x_1}{a_1}, \frac{x_2}{a_2}, \dots \right) \quad (2)$$

where  $a_i$  is a real sequence such that  $\sum_{i=1}^{+\infty} 1/a_i^2 < +\infty$  and  $(x_i)_{1 \leq i \leq n}$  is a sequence of real independent and identically distributed random variables with  $\mathcal{N}(0,1)$  distribution. We assume throughout that the  $a_i$ 's are arranged in a non-decreasing order. From Kolmogorov's 0 - 1 law it is straightforward to see that  $X$  exists as a  $l_2$ -valued random element. The square norm of  $X$  is :

$$z = \sum_{i=1}^{+\infty} \frac{x_i^2}{a_i^2}. \quad (3)$$

The small ball problem consists here in estimating for different choices of the sequence  $(a_i)_{i \in \mathbb{N}}$  the probability  $\mathbb{P}(z < \varepsilon)$  when  $\varepsilon$  tends to zero. The latter probability is expected to depend on the  $a_i$ 's.

The inspection of the case  $E = l_2$  is motivated by the application to functional statistics mentioned in the paragraph above. Indeed random functions are often reconstructed by interpolation techniques, like splines or wavelets, in Hilbert spaces such as  $L^2([0, T])$  or the Sobolev space  $W^{m,2}([0, T])$ . Then the random element  $X$  is valued in a separable Hilbert space  $H$ ...and all these Hilbert spaces of functions are isometrically isomorphic to  $l_2$ . Any gaussian random function  $X$  will be represented by its coordinates in accordance with (2) and through the well-known Karhunen-Loève decomposition on the basis of the eigenelements of its covariance operator  $(a_i^{-1}, e_i)$  where the  $e_i$ 's are the eigenfunctions :

$$X = \sum_{i=1}^{+\infty} a_i^{-1} x_i e_i.$$

The notations of this article are the same as those of Meyer-Wolf, Zeitouni (1993) and Dembo, Meyer-Wolf, Zeitouni (1995). These authors completely solved the  $l_2$  problem. For the sake of completeness and since the main theorems of this work heavily rely on their results I recall them. In both above-mentioned articles they proved that :

$$\lim_{\varepsilon \rightarrow 0} e^{I(\theta)} \psi(\theta) \mathbb{P}(z < \varepsilon) = 1/\sqrt{2\pi} \quad (4)$$

where  $\theta_\varepsilon = \theta$  is defined by  $\mu(\theta_\varepsilon) = \varepsilon$  with

$$\mu(\theta) = \sum_{i=1}^{+\infty} \frac{1}{a_i^2 + 2\theta}, \quad (5)$$

$$\psi(\theta) = \sqrt{\sum_{i=1}^{+\infty} \frac{2\theta^2}{(a_i^2 + 2\theta)^2}},$$

$$I(\theta) = \frac{1}{2} \sum_{i=1}^{+\infty} \log\left(1 + \frac{2\theta}{a_i^2}\right) - \theta\mu(\theta). \quad (6)$$

Without further assumption on the  $a_i$ 's  $\mathbb{P}(z < \varepsilon)$  cannot be made more explicit.

The article is organized as follows. The next subsection develops some aspects of theoretical statistics which motivate this approach on small ball problems. Then a class of functions which appears in extremes value theory is introduced in the next section. This family of function turns out to be of major importance to assess our main results in section 3. The proofs are collected in the last part of the article.

## 1.2 Small ball problems and some issues in statistical inference on curves.

In statistics, local methods (like estimation of conditional expectation, regression functions at a fixed point, etc) make it necessary to compute

$$\mathbb{E}K(\|X - x_0\|/h) = \int K(\|u - x_0\|/h) d\mathbb{P}_X(u)$$

where  $x_0$  is a fixed point of the space,  $K$  is a kernel that is a measurable positive function with  $\int K = 1$  and  $h$  is a nonnegative number tending to 0 (the bandwidth).

In a multivariate setting, when  $X$  is a  $\mathbb{R}^d$  valued random variable, elementary computations lead to :

$$\mathbb{E}K\left(\frac{\|X - x_0\|}{h}\right) \sim C_d f_X(x_0) h^d = f_X(x_0) \cdot \text{Vol}[B(0, h)] \quad (7)$$

in many situations. The exponent  $d$  may vary, depending on the support of the distribution of  $X$ .

But Lebesgue's measure cannot be defined on a Banach space, for instance, and when  $X$  is a process, the density of  $X$  at  $x_0$  does not make sense anymore. A major issue is then to compute the preceding expectation without assuming that the density of  $X$  at  $x_0$  exists. Assuming that  $K$  has compact support (say  $[0, 1]$ ) with  $K(1) > 0$  and integrable derivative we see from :

$$K(s) = K(1) - \int_s^1 K'(s) ds$$

and by Fubini's Theorem that :

$$\mathbb{E}K\left(\frac{\|X - x_0\|}{h}\right) = K(1) \mathbb{P}(\|X - x_0\| < h) - \int_0^1 K'(s) \mathbb{P}(\|X - x_0\| < hs) ds.$$

Denoting  $\mathbb{P}(\|X - x_0\| < h) = F_{x_0}(h) = F(h)$  for the sake of simplicity we finally get :

$$\mathbb{E}K\left(\frac{\|X - x_0\|}{h}\right) = F(h)K(1) - \int_0^1 K'(s)F(hs)ds. \quad (8)$$

From both last displays it is plain that the evaluation of  $\mathbb{E}K\left(\frac{\|X - x_0\|}{h}\right)$  strictly depends on the small ball probability  $F(h)$ . When  $X$  is a random curve its behaviour at 0 is crucial and determines the rate of convergence to zero of the above expectation -what statisticians are truly interested in.

In fact more intricate problems appear in statistical inference such as computations of truncated moments like :

$$\mathbb{E}[\varphi_p(X - x_0)]K\left(\frac{\|X - x_0\|}{h}\right)$$

where  $p$  is a nonnegative integer and  $\varphi_p$  is a tensor of order  $p$  acting on the space supporting  $X$ . For example when  $p = 2$ , the local covariance operator may be defined by :

$$\mathbb{E}[(X - x_0) \otimes (X - x_0)]K\left(\frac{\|X - x_0\|}{h}\right). \quad (9)$$

Once again the local behaviour of the small ball probability is a major issue. We refer to Müller and Yan (2001) for some statistical results on local moments for finite-dimensional random variables and to Mas (2008) for some related results dealing with (9) and where random functions and small ball problems appear. Obviously the examples above are simplistic but shed some light on a new potential application of small deviations -out of the scope of pure probability theory- and which, to the author's knowledge, has not been pointed out until now.

The reader should notice that formula (4) does not seem to be suited to complete formula (9) or even (8) since the estimate of the small ball probability is purely implicit. On the other hand rewriting display (8) leads to :

$$\mathbb{E}K\left(\frac{\|X - x_0\|}{h}\right) = F(h)\left[K(1) - \int_0^1 K'(s)\frac{F(hs)}{F(h)}ds\right].$$

Assume that  $F$  is regularly varying at zero with index  $d$  (which is usually true when  $X$  is finite dimensional) then by definition  $F(h) = Ch^{dl}(h)$  where  $C$  is a constant,  $l$  is a slowly varying function at 0 and  $F(hs)/F(h) \rightarrow s^d$  when  $h \rightarrow 0$  which yields :

$$\mathbb{E}K\left(\frac{\|X - x_0\|}{h}\right) \sim Ch^{dl}(h)\left[K(1) - \int_0^1 s^d K'(s)ds\right].$$

Unfortunately when  $X$  lies in the space  $l_2$ , the most classical examples of  $F(h)$  are not regularly varying as will be seen below. But however we notice for further purpose that the theory of regular variation is of some help in the important special case mentioned just above.

## 2 The class $\Gamma_0$

This section is independent from the preceding.

The theory of extremes is another well-known topic connecting probability theory, mathematical statistics and real analysis through regular variation and Karamata's theory. The foundations of Extreme theory may be illustrated by the famous Fisher-Tippett theorem (see Fisher, Tippet (1928) and Gnedenko (1943)). This classical result assesses that whenever  $U_1, \dots, U_n$  is an i.i.d. sample of real random variables,  $M_n = \max\{U_1, \dots, U_n\}$  belongs to the domain of attraction of  $G$ , where  $G$  has same type as one of the three distributions Gumbel, Frechet and Weibull. The Gumbel law, also named double exponential distribution, with cumulative density function  $\Lambda(x) = \exp(-\exp(-x))$  defines the so-called "domain of attraction of the third type". Laurens de Haan (1971) characterized the (cumulative) distribution functions of  $U$  such that  $M_n$  belongs to the domain of attraction of  $\Lambda$ . We give this result below.

**Theorem (de Haan, 1971) :** *If  $F$  is the cumulative density function of a real random variable  $X$  which belongs to the domain of attraction of the third type (Gumbel) there exists a measurable function  $\rho : \mathbb{R} \rightarrow \mathbb{R}^+$ , called the auxiliary function of  $F$ , such that :*

$$\lim_{s \uparrow x_+} \frac{\overline{F}(s + x\rho(s))}{\overline{F}(s)} = \exp(-x)$$

where  $\overline{F}(s) = 1 - F(s)$ ,  $x_+ = \sup\{x : F(x) < 1\}$ .

This property was initially introduced by de Haan's as a "Form of Regular Variation" (see the title of his article). This class of distribution function is referred to as de Haan's Gamma class in the book by Bingham, Goldie and Teugels (1987) and within this article. In the latter book the definition is slightly different from the one given above. Gamma-variation is defined at infinity and for non-decreasing functions which comes down to taking  $x_+ = +\infty$  and taking  $\exp(x)$  instead of  $\exp(-x)$  in the display above. Surprisingly, in their book as well as in de Haan's article no examples of functions belonging to  $\Gamma$  is given. The cumulative density function of the gaussian distribution belongs to this class with  $x_+ = +\infty$  and  $\rho(s) = 1/s$ .

Since we focus on the local behaviour at zero of the cumulative density function of a real valued random variable we have to modify again slightly the definitions above. We introduce the class  $\Gamma_0$  and feature some of its properties below. We share most of our notations with Bingham, Goldie and Teugels (1987) which differ from those of de Haan.

Let  $\mathcal{V}^+$  be a right-neighborhood of zero.

**Definition 1** *The class  $\Gamma_0$  consists of those functions  $F : \mathbb{R} \rightarrow \mathbb{R}^+$  null over  $(-\infty, 0]$ , non decreasing with  $F(0) = 0$  and right-continuous for which there exists a continuous non decreasing function  $\rho : \mathcal{V}^+ \rightarrow \mathbb{R}^+$ , the auxiliary function of  $F$ , such that  $\rho(0) = 0$  and for all  $x \in \mathbb{R}$ ,*

$$\lim_{s \downarrow 0^+} \frac{F(s + x\rho(s))}{F(s)} = \exp(x) \tag{10}$$

The properties of the auxiliary function are crucial.

**Proposition 1** From Definition 1 above we deduce that :  $\rho(s)/s \rightarrow 0$  as  $s \rightarrow 0$  and  $\rho$  is self-neglecting which means that :

$$\frac{\rho(s + x\rho(s))}{\rho(s)} \xrightarrow{s \rightarrow 0} 1$$

locally uniformly in  $x \in \mathbb{R}$ .

**Remark 1** When the property in the proposition above does not hold locally uniformly but only pointwise the function is called Beurling slowly varying. Assuming that  $\rho$  is continuous in Definition 1 yields local uniformity and enables to consider a self-neglecting  $\rho$ .

Let us also mention that Gaïffas (2005) proposed to model locally the density of sparse data by gamma-varying functions. This is another statistical application for  $\Gamma_0$ .

The class  $\Gamma_0$  is subject to an exponential-integral representation. In fact the following Theorem asserts that the local behaviour at 0 of any  $F$  in  $\Gamma_0$  depends only on the auxiliary mapping  $\rho$ .

**Theorem 1** Let  $F$  belong to  $\Gamma_0$  with auxiliary function  $\rho$  then when  $s \rightarrow 0$  :

$$F(s) = C(1 - \varepsilon(s)) \exp \left\{ - \int_s^1 \frac{1}{\rho(t)} dt \right\} / \rho^2(s) \quad (11)$$

with  $\varepsilon(s) \rightarrow 0$  and the auxiliary function  $\rho$  is unique up to asymptotic equivalence and may be taken as  $\int_0^s F(t) dt / F(s)$ . Besides

$$F(\lambda s) / F(s) \rightarrow \begin{cases} \infty & (\lambda > 1) \\ 1 & (\lambda = 1) \\ 0 & (\lambda < 1) \end{cases} \quad \text{as } s \rightarrow 0. \quad (12)$$

**Remark 2** The upper bound 1 in the integral in display (11) is unimportant and may be replaced by any positive number. Then the constant  $C$  will change as well.

The proof of Proposition 1 as well as Theorem 1 are inspired from the proofs of Lemma 3.10.1, Proposition 3.10.3 and Theorem 3.10.8 in Bingham et al (1987) and will be omitted. However the next Proposition seems to be new and specific from the class  $\Gamma_0$ .

**Proposition 2** Let  $F$  belong to  $\Gamma_0$ . Then for all integer  $p$   $F^{(p)}(0) = 0$  where  $F^{(p)}$  denotes the derivative of order  $p$  of  $F$ .

This means that functions from  $\Gamma_0$  are very flat in a neighborhood of zero. Proposition 2 may be intuitively explained. Let us start from the fact that the standard situation when  $X$  lies in  $\mathbb{R}^d$  is illustrated by the right hand side of (7). Then the  $d^{th}$  derivative of this power function is always non-null. Consequently in an infinite dimensional space we can expect that all the derivatives at 0 should be null. A more geometric way to understand this consists in considering the problem of the concentration of a probability measure. Let  $\mu$  be the measure associated with the random variable  $X$ . Once again starting from  $\mathbb{R}^d$  and letting  $d$  increase -even if this approach is not really fair- we see that  $\mu$  must allocate a constant mass of 1 to a space whose dimension increases. Then  $\mu$  gets more and more diffuse, allowing fewer mass to balls and visiting rarely fixed points such as  $x_0$  (and their neighborhoods), resulting in a very flat small ball probability.



## 3 Main results

### 3.1 The representation theorems

This article connects the two apparently distinct notions of probability of the two preceding sections : the class of gaussian small ball probabilities in  $l_2$  and de Haan's Gamma class of functions. Both families of functions are defined by their local behaviour around 0. The starting point is purely empirical and is explicated in the Proposition and subsequent comments below. This Proposition, which is nothing but an application of formula (4) in Dembo, Meyer-Wolf, Zeitouni (1995), aims at showing to the reader the process that lead to inferring the main results of this article.

**Lemma 1 (Dembo, Meyer-Wolf, Zeitouni (1995) and Mas (2008))**  
Take  $X$  as defined in (2) with  $a_i = i^{\beta/2}$  ( $\beta > 1$ ) then :

$$\mathbb{P} \left( \|X\|^2 < s \right) \underset{0}{\sim} C_1 s^{-1/(2\beta-2)} \exp \left( -C_2 s^{-1/(\beta-1)} \right). \quad (13)$$

where  $C_1$  and  $C_2$  are positive constants. By applying the formula of the previous authors, Mas (2008) proved that when  $a_i = \exp(\alpha \cdot i)$ ,  $\alpha > 0$

$$\mathbb{P} \left( \|X\|^2 < s \right) \underset{0}{\sim} \sqrt{\frac{\alpha}{-\pi \log(s)}} \exp \left( -\frac{1}{4\alpha} [\log(s)]^2 \right). \quad (14)$$

**Remark 3** This lemma has several by-products. Simple algebra proves that both functions on the right hand side of (13) and (14) have all their derivatives vanishing at 0. We notice that when  $\beta$  tends to 1 the r.h.s. of (13) is always flatter and that (13) is for all  $\beta$  flatter than (14) which in turn will always be flatter at 0 than any polynomial function (like  $C_d s^d$ ). However we notice the the degree of flatness is directly connected with the rate of increase of the  $a_i$ 's which quantifies, exactly like the  $l$ -numbers, the accuracy of a finite-dimensional approximation of  $X$ .

The next Proposition is crucial and is a first step in connecting both preceding sections.

**Proposition 3** Both functions on the right hand side of (13) and (14) belong to the class  $\Gamma_0$  with respective auxiliary functions :

$$\rho_1(s) = M_1 s^{1+1/(\beta-1)} \quad (15)$$

$$\rho_2(s) = -M_2 s / \log(s) \quad (16)$$

where  $M_1$  and  $M_2$  are positive constants.

**Remark 4** It is easily seen that  $\rho_1$  and  $\rho_2$  match Proposition 1 above. Besides letting  $\beta$  go to infinity we see that, in a way  $\rho_2$  may be viewed as a limit of  $\rho_1$ . In fact  $-1/\log(s)$  echoes the degeneracy of  $1/(\beta-1)$ . Remember that  $\rho(s)/s$  should tend to 0 with  $s$ .

Conversely turning to Theorem 1 and taking  $\rho(t) = t^m$  (with  $m > 1$ ) yields :

$$F(s) \underset{0}{\sim} C t^{-2m} \exp \left( -\frac{1}{t^{m-1}} \right) \quad (17)$$

and taking  $\rho(t) = -t/\log(t)$  yields :

$$F(s) \underset{0}{\sim} C \left( \frac{\log(t)}{t} \right)^2 \exp \left( - [\log(t)]^2 \right) \quad (18)$$

Clearly displays (13) and (17) on a one hand and (14) and (18) on the other hand are not equal but the dominating term (namely the exponential one) matches up to constants. The dissimilarity between the terms before the exponential will be developed and made clearer in the next subsection.

The Proposition 3 leads to our main concern : *is it possible to obtain a one to one representation, in a general framework, of the small ball probability of a gaussian random element in  $l_2$  -characterized by the sequence  $(a_i)_{i \in \mathbb{N}}$ - by a function in  $\Gamma_0$ , depending solely on its auxiliary function  $\rho$  ?* If so, what are the connections between  $\rho$  and the sequence  $(a_i)_{i \in \mathbb{N}}$  ? The next theorem answers positively to this question and assesses that any small ball probability belongs to  $\Gamma_0$ .

In all the sequel, the exponent  $-1$  is strictly reserved to denoting the generalized inverse of a function  $f$  denoted  $f^{-1}$ . Consequently in general  $f^{-1} \neq 1/f$ . This remark will be of much importance in the next section dedicated to the mathematical derivations.

**Theorem 2** *Let  $z$  be defined by (3) and set  $\mathbb{P}(z < s) = F(s)$  then :*

$$F \in \Gamma_0$$

*with auxiliary function :*

$$\rho = \frac{1}{\mu^{-1}} \quad (19)$$

*where  $\mu$  was defined at display (5).*

It turns out that, surprisingly, a partial converse holds. The only additional assumption relates to the local properties of  $\rho$  and regular variation is assumed in order to apply the direct part of Karamata's theorem's within the proof.

**Theorem 3** *Let  $F \in \Gamma_0$  with auxiliary function  $\rho$ . Assume that  $\rho$  is regularly varying at zero and define  $\mu = \left(\frac{1}{\rho}\right)^{-1}$  and  $\varphi(t) = t\mu(t)$  then  $F$  is the small ball probability of a  $l_2$ -valued gaussian random variable defined as above with, for some constant  $C$  :*

$$a_i^2 = C\varphi^{-1}(i)$$

**Remark 5** *Within Proposition 3 both auxiliary functions  $\rho_1$  and  $\rho_2$  are regularly varying and the assumption of Theorem 3 holds in this context. It is also possible to check - it takes some calculations however- that the relationship between  $\rho$  and the sequence  $(a_i)_{i \in \mathbb{N}}$  defined through (19) holds in the special cases examined in Lemma 1 at displays (13) and (14). An interesting open question would consist in finding examples of auxiliary functions which are not regularly varying with positive index, whenever it is possible.*

These theorems ensure that the exponential-integral representation (11) holds for any small ball probability in  $l_2$ . Comparing it with (4) we can expect some simplifications for computational purposes. This issue is developed in the Remark 6 below. But for the sake of completeness it is worth addressing a concern which may be misleading. Indeed we started from  $\mathbb{P}(z < \varepsilon)$  where  $z = \|X\|_{l_2}^2$  and the properties of this function may differ from those of what may be intended as the "true" small ball probability  $\mathbb{P}\left(\|X\|_{l_2}^2 < \varepsilon^2\right)$ . The next Proposition answers this question.

**Proposition 4** *Let  $F \in \Gamma_0$  with auxiliary function  $\rho_F$  and define  $G : s \rightarrow F(s^2)$ . Then  $G$  belongs to  $\Gamma_0$  as well with auxiliary function  $\rho_G$  defined by :*

$$\rho_G(s) = \frac{1}{2s} \rho_F(s^2).$$

**Remark 6** *Turning again to the statistical application mentioned at display (8) and applying Lebesgue's dominated convergence theorem, it is plain that :*

$$\mathbb{E}K\left(\frac{\|X - x_0\|}{h}\right) \sim K(1)F(h)$$

since  $F(hs)/F(h)$  tends to zero for fixed  $s$  and  $h$  going to 0 (see (12) within Theorem 1). Let us focus on truncated or smoothed moments mentioned in (9) above and consider the weak versions of the tensors mentioned there :

$$\mathbb{E}\langle X - x_0, u \rangle^p K\left(\frac{\|X - x_0\|}{h}\right) \text{ or } \mathbb{E}\langle X, u \rangle^p I_{\{\|X - x_0\| \leq h\}}.$$

where  $u \in l_2$ . It turns out that the rate of decrease to zero in the display above depends critically on the the auxiliary function  $\rho$  which appears in our representation theorem and was computed in the case  $p = 2$  by Mas (2008). This fact has serious consequences on the convergence of estimates in some statistical models (see Berline, Elamine, Mas (2008)).

### 3.2 Complementing the representation of the de Haan's Gamma class.

As mentioned earlier, just below display (18) we noted that the representations obtained from Dembo, Meyer-Wolf, Zeitouni (1995) and those from the class  $\Gamma_0$  do not match, due to a multiplicative term.

If we turn again to Definition 1 and pick a  $F$  in  $\Gamma_0$  it is plain to see that any function  $\phi^F$  where  $\phi(x + t\rho(x))/\phi(x) \rightarrow 1$  when  $x \rightarrow 0$  belongs to  $\Gamma_0$  with exactly the same auxiliary function  $\rho$ . This is a drawback of the representation proposed above. But the reader also noticed that the previous results assess that the function  $\rho$  is not unique and may be replaced by any  $\tilde{\rho}$  such that  $\tilde{\rho} \sim \rho$  at zero (see Theorem 1 just below display(11)). We prove now that these facts are intimately connected and that changing  $\rho$  to  $\tilde{\rho}$  finally comes down to changing the function  $\phi$  above.

We start with a definition which seems to be new.

**Definition 2** *Let  $\rho$  be a self-neglecting function. A function  $\phi$  is called  $\rho$ -self-neglecting if :*

$$\frac{\phi(x + t\rho(x))}{\phi(x)} \xrightarrow{x \rightarrow 0} 1.$$

We propose a representation theorem for  $\rho$ -self-neglecting functions. This theorem will be useful in the sequel.

**Theorem 4** *Let  $\rho$  be self-neglecting at 0 which does not vanish in a neighborhood of 0. A function  $\phi$  is  $\rho$ -self-neglecting iff :*

$$\phi(x) = c(x) \exp\left(\int_x^1 \frac{\varepsilon(u)}{\rho(u)} du\right)$$

where  $c(u) \rightarrow c \in ]0, +\infty)$  and  $\varepsilon(u) \rightarrow 0$  when  $u \rightarrow 0$  and  $\varepsilon$  has the same regularity as  $\rho$ .

This theorem generalizes the representation Theorem 2.11.3 for self-neglecting functions p.121 in Bingham et al. (1987) initially due to Bloom (1976). If one take  $\phi = \rho$  the representation above coincides with the one announced in this theorem.

The next result shows how an asymptotically invariant change of the auxiliary function  $\rho$  influences the overall expression of a function in  $\Gamma_0$  through the previous theorem.

**Theorem 5** *Let us take a function  $F$  in  $\Gamma_0$  with auxiliary function  $\rho$  with representation :*

$$F(x) = \exp\left\{\int_x^1 \frac{1}{\rho(t)} dt\right\}.$$

Then changing  $F$  into  $\phi F = \tilde{F}$  where  $\phi$  is a  $\rho$ -self neglecting function belongs again to  $\Gamma_0$ . The auxiliary function of  $\tilde{F}$  is  $\tilde{\rho} = \rho/(1 - \varepsilon)$  where  $\varepsilon$  is the function defined in the Theorem above, hence  $\tilde{\rho} \sim \rho$  at zero.

The proof of this last theorem is straightforward with Theorem 4 at hand and omitted.

**Remark 7** *The results of this section could be rephrased this way : the equivalent class of self-neglecting functions induced by the relationship  $\rho_1 \sim_0 \rho_2$  creates an equivalence class for functions in  $\Gamma_0$  through the relationship  $\Delta$  defined for all  $F$  and  $G$  in  $\Gamma_0$  by*

$$F \Delta G \Leftrightarrow \frac{F}{G} \text{ is } \rho - \text{self-neglecting}$$

And our results are rather a representation for functions in  $\Gamma_0 \setminus \Delta$ . Nevertheless from a more practical viewpoint these considerations vanish when considering the logarithm of the small ball probability in the following sense : if  $F$  (resp.  $G$ ) belongs to  $\Gamma_0$  with auxiliary function  $\rho_1$  (resp.  $\rho_2$ ) and  $\rho_1 \sim_0 \rho_2$  then  $\log F \sim_0 \log G$ .

### 3.3 Conclusion

The main results of this article are rather formal. They identify gaussian small ball probabilities in  $l_2$  with a class of rapidly varying functions involved in extreme value theory and whose derivatives at all orders vanish at zero. This

representation was obtained through previous works especially the seminal formula (4). But we hope and guess that this new formulation will be more convenient for modeling the small ball probabilities with some applied -especially statistical- purposes in mind. However many questions arise, most of them with rather theoretic aspects. It seems that the gaussian framework could be extended when the assumptions A1 – 4 p.279 of Dembo et al. (1995) hold. The generalization to random elements with values in  $l_p$  or in more general Banach spaces is an open and certainly intricate matter since the starting formulas (4) and followings seem to be intimately suited to the space  $l_2$ .

A more promising track could be to explore the links between the auxiliary function  $\rho$ , which inherits all the seminal information on the regularity of  $X$ , with the metric entropy of the unit ball of the reproducing kernel Hilbert space of  $X$  as explored in Li, Linde (1999) or with the degree of compactness of the operator  $v$  in Li, Linde (2004) for instance, the latter operator  $v$  being obviously close to the covariance operator of  $X$  hence in connection with the  $a_i$ 's of this article.

The connections between regular variations and small ball probabilities have been known since de Bruin in 1959, and his theorem on Laplace transforms (see Theorem 4.12.9 in Bingham et al. (1987)). This work confirms that both Tauberian and extreme value theory may provide tools complementing large deviations techniques to derive new results in this area. This deserves certainly more attention.

## 4 Proofs

We start with the proof of Proposition 1.

### Proof of Proposition 1 :

Suppose that  $\rho(s)/s$  does not tend to zero when  $s$  does. Then we may pick an  $\varepsilon > 0$  such that for infinitely many  $s_k \downarrow 0$  when  $k \uparrow +\infty$ ,  $\rho(s_k)/s_k > \varepsilon$ . Now fix  $x < -\varepsilon^{-1}$  then  $s_k + x\rho(s_k) < 0$  and  $F(s_k + x\rho(s_k)) = 0$  for all  $k$  and  $F(s_k + x\rho(s_k))/F(s_k)$  cannot converge to  $\exp(x)$ . The second part of the proof, namely ensuring the  $\rho$  is self-neglecting, follows the lines of the proof of Proposition 3.10.6 in Bingham et al. (1987). ■

**Proof of Proposition 2 :** Suppose that for some  $p$   $F^{(p)}(0) \neq 0$  and take  $p^* = \inf \{p \in \mathbb{N} : F^{(p)}(0) \neq 0\}$ . It is plain that  $F^{(p^*)}(0) > 0$  since  $F$  is positive. Then we should consider two cases. First if  $F^{(p^*)}(0) = c < +\infty$  then  $F(s) \sim cs^{p^*}$ . Taking :

$$\frac{F(s + \rho(s))}{F(s)} = \frac{F(s + \rho(s))}{(s + \rho(s))^{p^*}} \frac{s^{p^*}}{F(s)} \frac{(s + \rho(s))^{p^*}}{s^{p^*}}$$

we see that the left hand side of the display above tends to  $\exp(1)$  whereas the right hand side tends to 1.

Second if  $F^{(p^*)}(0) = +\infty$  we clearly have  $F(s)/s^{p^*} \rightarrow +\infty$  when  $s \rightarrow 0$ . Take  $\varepsilon$  such that  $1/\varepsilon > p^* + 2$ . Since  $\rho'(0) = 0$  and  $\rho$  is positive we may pick

an  $s_0$  such that  $\sup_{0 \leq u \leq s_0} \rho'(u) \leq \varepsilon$ . From (11) we get

$$\begin{aligned} \frac{F(s)}{s^p} &\leq \frac{C}{s^p \rho^2(s)} \exp \left\{ - \int_s^1 \frac{1}{\rho(t)} dt \right\} \leq \frac{C}{\rho^{2+p}(s)} \exp \left\{ - \int_s^1 \frac{1}{\rho(t)} dt \right\} \\ &\leq \frac{C'}{\rho^{2+p}(s)} \exp \left\{ - \int_s^{s_0} \frac{1}{\rho(t)} dt \right\} \end{aligned}$$

where we assume that  $s \leq s_0$ . Then we have

$$\begin{aligned} \exp \left\{ - \int_s^{s_0} \frac{1}{\rho(t)} dt \right\} &= \exp \left\{ - \int_s^{s_0} \frac{\rho'(t)}{\rho(t)} \frac{1}{\rho'(t)} dt \right\} \leq \exp \left\{ - \frac{1}{\varepsilon} \int_s^{s_0} \frac{\rho'(t)}{\rho(t)} dt \right\} \\ &= \exp \left\{ \frac{1}{\varepsilon} \ln \rho(s) - \frac{1}{\varepsilon} \ln \rho(s_0) \right\}. \end{aligned}$$

At last

$$\frac{F(s)}{s^p} \leq C'' [\rho(s)]^{\frac{1}{\varepsilon} - p^* - 2}$$

which contradicts the fact that  $F(s)/s^{p^*} \rightarrow +\infty$ . ■

Before turning to the proof of Theorem 2 we go on with a lemma. Remind that the function  $\psi$  was defined just below display (5).

**Lemma 2** *Function  $\psi$  (resp.  $\psi \circ \mu^{-1}$ ) is increasing (resp. decreasing), and*

$$\frac{\psi'(u)}{\psi(u)} = \frac{1 \sum_{i=1}^{+\infty} \frac{a_i^2}{(a_i^2 + 2u)^3}}{u \sum_{i=1}^{+\infty} \frac{1}{(a_i^2 + 2u)^2}}. \quad (20)$$

Besides,

$$\theta \mu(\theta) \rightarrow +\infty \quad (21)$$

when  $\theta$  goes to  $+\infty$ .

**Proof :**

Display (20) is obtained by straightforward calculations, we leave it to the reader. It implies that  $\psi$  is increasing and that  $\psi \circ \mu^{-1}$  is decreasing since  $\mu$  is. We prove (21). We have :

$$\theta \mu(\theta) = \sum_{i=1}^{+\infty} \frac{\theta}{a_i^2 + 2\theta} \geq \sum_{i=1}^{N(\theta)} \frac{\theta}{a_i^2 + 2\theta}$$

where  $N(\theta) = \sup \{i : a_i^2 \leq \theta\}$ . Then

$$\theta \mu(\theta) \geq \frac{1}{3} N(\theta).$$

Obviously  $N(\theta)$  goes to infinity when  $\theta$  does because the  $a_i$ 's do and are arranged in a non-decreasing order. ■

**Proof of Theorem 2 :**

From Definition 1 and (4) we see that Theorem 2 holds whenever for all  $x \in \mathbb{R}$  :

$$\lim_{s \rightarrow 0} \frac{\psi(\mu^{-1}(s)) \exp(I(\mu^{-1}(s)))}{\psi(\mu^{-1}(s + x\rho(s))) \exp(I(\mu^{-1}(s + x\rho(s))))} = \exp x.$$

We focus on proving that :

$$\lim_{s \rightarrow 0} \frac{\psi(\mu^{-1}(s))}{\psi(\mu^{-1}(s + x\rho(s)))} \exp(I(\mu^{-1}(s)) - I(\mu^{-1}(s + x\rho(s))) - x) = 1.$$

The two next lemmas are dedicated to showing that, in the above display the fraction as well as the exponential both tend to 1 when  $s$  goes to zero and  $\rho$  is chosen as in the Theorem.

**Lemma 3** Take  $\rho(s) = 1/\mu^{-1}(s)$ , then :

$$\lim_{s \rightarrow 0} \exp(I(\mu^{-1}(s)) - I(\mu^{-1}(s + x\rho(s))) - x) = 1.$$

**Remark 8** Obviously  $\mu$  has at least two (we do not need more) continuous derivatives on a neighborhood of infinity (here  $]1, +\infty[$  for instance). It is also straightforward to see that  $\mu$ , which is strictly decreasing on  $]1, +\infty[$ , is also a  $C^1$  diffeomorphism on this set. Clearly  $\lim_{s \rightarrow 0} \rho(s) = 0$  but from (21) it is plain that  $\rho(s)/s$  also tends to zero when  $s$  does which implies that  $\rho'(0) = 0$ . Indeed proving that  $\rho(s)/s$  tends to zero comes down to proving that  $\theta\mu(\theta) \rightarrow +\infty$ .

**Proof of the Lemma :**

Taylor's formula gives :

$$I(\mu^{-1}(s)) - I(\mu^{-1}(s + x\rho(s))) = -x\rho(s) \frac{I'(\mu^{-1}(s))}{\mu'(\mu^{-1}(s))} + \frac{x^2}{2} \rho^2(s) [I(\mu^{-1})]''(c_{s,x}) \quad (22)$$

where  $c_{s,x} = c$  lies somewhere in  $[s, s + x\rho(s)]$  if  $x \geq 0$  and in  $[s + x\rho(s), s]$  if  $x < 0$ . From (6) we see that :

$$\begin{aligned} I'(\theta) &= \frac{1}{2} \sum_{i=1}^{+\infty} \log \left( 1 + \frac{2\theta}{a_i^2} \right) - \theta\mu(\theta) \\ &= \mu(\theta) - \mu(\theta) - \theta\mu'(\theta) \\ &= -\theta\mu'(\theta) \end{aligned}$$

which immediately yields  $\frac{I'(\mu^{-1}(s))}{\mu'(\mu^{-1}(s))} = \mu^{-1}(s)$ . Hence (22) may be rewritten :

$$I(\mu^{-1}(s)) - I(\mu^{-1}(s + x\rho(s))) = x\rho(s) \mu^{-1}(s) + \frac{x^2}{2} \rho^2(s) [I(\mu^{-1})]''(c_s)$$

Choosing  $\rho(\cdot) = 1/\mu^{-1}(\cdot)$  we have :

$$I(\mu^{-1}(s)) - I(\mu^{-1}(s + x\rho(s))) - x = \frac{x^2}{2} \rho^2(s) [I(\mu^{-1})]''(c)$$

and the Lemma will be proved if we get

$$\lim_{s \rightarrow 0} \rho^2(s) [I(\mu^{-1})]''(c) = 0. \quad (23)$$

But for all  $t$  :

$$\begin{aligned} [I(\mu^{-1})]''(t) &= \frac{d \left\{ [I(\mu^{-1})]'(t) \right\}}{dt} = \frac{d\mu^{-1}}{dt}(t) \\ &= \left( \frac{1}{\rho} \right)'(t) = -\frac{\rho'(t)}{\rho^2(t)}. \end{aligned}$$

Hence

$$\left| \rho(s) [I(\mu^{-1})]''(c) \right| = \rho^2(s) \frac{\rho'(c)}{\rho^2(c)}. \quad (24)$$

We first show that  $\rho^2(s)/\rho^2(c)$  is bounded above. We may always write  $c = s + t_x(s)\rho(s)$  where  $-x \leq t_x(s) \leq x$  for all  $s$ . Taylor's formula yields

$$\rho(s + t_x(s)\rho(s)) = \rho(s) + t_x(s)\rho(s)\rho'(d) = \rho(s)(1 + t_x(s)\rho'(d))$$

where  $d$  lies between  $s$  and  $s + t_x(s)\rho(s)$ . Hence

$$\frac{\rho(s)}{\rho(c)} = \frac{1}{1 + t_x(s)\rho'(d)} \leq \frac{1}{1 - |x|\rho'(d)}$$

The continuity of  $\rho'$  at 0 and its nullity at 0 (see Remark 8) implies on a one hand that the display above is bounded above for fixed  $x$  and  $s$  (hence  $d$ ) going to zero and also that (24) holds as well as (23). This finishes the proof of the Lemma. ■

**Lemma 4** When  $\rho(s) = 1/\mu^{-1}(s)$ ,

$$\lim_{s \rightarrow 0} \frac{\psi(\mu^{-1}(s))}{\psi(\mu^{-1}(s + x\rho(s)))} = 1$$

**Proof :**

Assume that  $x > 0$ . We will prove that :

$$\lim_{s \rightarrow 0} \frac{\psi(\mu^{-1}(s + x\rho(s)))}{\psi(\mu^{-1}(s))} = 1$$

Once again Taylor's formula leads to :

$$\frac{\psi(\mu^{-1}(s + x\rho(s)))}{\psi(\mu^{-1}(s))} = 1 + x\rho(s) \frac{(\psi \circ \mu^{-1})'(c^*)}{\psi(\mu^{-1}(s))}$$

where  $c^* \in [s, s + x\rho(s)]$  since  $x \geq 0$ . So we focus on proving that :

$$\lim_{s \rightarrow 0} \rho(s) \left| \frac{(\psi \circ \mu^{-1})'(c^*)}{\psi(\mu^{-1}(s))} \right| = 0$$



where :

$$(\psi \circ \mu^{-1})'(c^*) = \frac{(\psi' \circ \mu^{-1})(c^*)}{\mu' \circ \mu^{-1}(c^*)}$$

and

$$\begin{aligned} \rho(s) \frac{(\psi \circ \mu^{-1})'(c^*)}{\psi(\mu^{-1}(s))} &= \rho(s) \frac{(\psi' \circ \mu^{-1})(c^*)}{\mu' \circ \mu^{-1}(c^*) \psi(\mu^{-1}(s))} \\ &= \frac{\rho(s) \rho(c^*)}{\mu' \circ \mu^{-1}(c^*)} \cdot \frac{\psi(\mu^{-1}(c^*))}{\psi(\mu^{-1}(s))} \cdot \frac{(\psi' \circ \mu^{-1})(c^*)}{\rho(c^*) \psi(\mu^{-1}(c^*))}. \end{aligned}$$

We set

$$\begin{aligned} M_1^x(s) &= \left| \frac{\rho(s) \rho(c^*)}{\mu' \circ \mu^{-1}(c^*)} \right|, \\ M_2^x(s) &= \frac{\psi(\mu^{-1}(c^*))}{\psi(\mu^{-1}(s))}, \\ M_3^x(s) &= \frac{(\psi' \circ \mu^{-1})(c^*)}{\rho(c^*) \psi(\mu^{-1}(c^*))}. \end{aligned}$$

We keep in mind that  $x > 0$  is fixed and  $s$  goes to zero, and prove that :

$$\begin{aligned} M_1^x(s) &\rightarrow 0 \\ \sup_s \{M_2^x(s), M_3^x(s)\} &\leq 1. \end{aligned}$$

We start with :

$$\begin{aligned} M_3^x(s) &= \frac{(\psi' \circ \mu^{-1})(c^*)}{\rho(c^*) \psi(\mu^{-1}(c^*))} \\ &= \frac{\mu^{-1}(c^*) (\psi' \circ \mu^{-1})(c^*)}{\psi(\mu^{-1}(c^*))} \\ &= \frac{\theta^* \psi'(\theta^*)}{\psi(\theta^*)} \end{aligned}$$

where  $\theta^* = \theta^*(s) = \mu^{-1}(c^*)$  and clearly  $\theta^* \rightarrow +\infty$  when  $s$  goes to zero. We prove that for large  $\theta^*$ ,  $\theta^* \psi'(\theta^*) / \psi(\theta^*)$  is bounded. From (20) it is plain that :

$$\frac{\theta^* \psi'(\theta^*)}{\psi(\theta^*)} = \frac{\sum_{i=1}^{+\infty} \frac{a_i^2}{(a_i^2 + 2\theta^*)^3}}{\sum_{i=1}^{+\infty} \frac{1}{(a_i^2 + 2\theta^*)^2}} \leq \sup_i \frac{a_i^2}{a_i^2 + 2\theta^*} \leq 1 \quad (25)$$

for all  $i$  and  $\theta^*$ .

We turn to  $M_2^x(s)$ . Since  $\psi \circ \mu^{-1}$  is decreasing and  $c^* > s$ ,  $M_2^x(s) \leq 1$ .

At last we deal with  $M_1^x(s)$  :

$$M_1^x(s) = \left| \frac{\rho(s) \rho(c^*)}{\mu' \circ \mu^{-1}(c^*)} \right| = \left| -\frac{\rho'(c^*) \rho(s) \rho(c^*)}{\rho^2(c^*)} \right| = |\rho'(c^*)| \frac{\rho(s)}{\rho(c^*)}$$

since  $\rho(\cdot) = 1/\mu^{-1}(\cdot)$ . Now as  $\rho$  is increasing

$$0 \leq \frac{\rho(s)}{\rho(c^*)} \leq 1$$

and since  $\rho'$  is continuous at 0 with  $\rho'(0) = 0$  we have

$$\lim_{s \rightarrow 0} M_1^x(s) = 0$$

hence

$$\lim_{s \rightarrow 0} \rho(s) \frac{(\psi \circ \mu^{-1})'(c^*)}{\psi(\mu^{-1}(s))} = 0$$

and Lemma 4 is proved when  $x > 0$ .

Now when  $x < 0$ , all the steps carried out above lead to the same result when starting from :

$$\lim_{s \rightarrow 0} \frac{\psi(\mu^{-1}(s))}{\psi(\mu^{-1}(s + x\rho(s)))} = 1$$

instead of dealing with  $\frac{\psi(\mu^{-1}(s+x\rho(s)))}{\psi(\mu^{-1}(s))}$ . In fact everything holds just replacing  $s + x\rho(s)$  (with  $x > 0$ ) by  $s$  and  $s$  by  $s + x\rho(s)$  (with  $x < 0$ ). ■

Now we turn to the proof of the converse part, Theorem 3. It takes two steps.

First we should make sure that when  $a_i^2 = \varphi^{-1}(i)$ ,

$$\sum \frac{1}{a_i^2} < +\infty$$

which will ensure that the random element defined by  $X = \left(\frac{x_1}{a_1}, \frac{x_2}{a_2}, \dots\right)$  is well-defined. Then we will prove that when  $X$  is defined this way its small ball probability is  $F$  with auxiliary function  $\rho$ .

**Lemma 5** When  $a_i^2 = \varphi^{-1}(i)$ ,  $\sum \frac{1}{a_i^2} < +\infty$ .

**Proof :** It is easily seen that  $\varphi^{-1}$  is non decreasing in a neighborhood of  $+\infty$ . Indeed it suffices to prove that  $\varphi$  is, which may be deduced from its definition by studying its derivative. By the way one may also see that  $\varphi$  is concave. Now since  $\varphi^{-1}$  is non decreasing it is enough to prove that :

$$\int^{+\infty} \frac{dx}{\varphi^{-1}(x)} < +\infty$$

where the notation above means "the improper integral converges at infinity". Set  $u = \varphi^{-1}(x)$  above then we should examine :

$$\int^{+\infty} \frac{\varphi'(u)}{u} du.$$

Integrating by part this comes down to ensuring first that  $\frac{\varphi(A)}{A}$  tends to a finite limit as  $A$  tends to infinity which is plain with  $\frac{\varphi(A)}{A} = \mu(A) \rightarrow 0$  and second investigating

$$\int^{+\infty} \frac{\varphi(u)}{u^2} du = \int^{+\infty} \frac{\mu(u)}{u} du = \int^{+\infty} \frac{1}{u} \left(\frac{1}{\rho}\right)^{-1}(u) du$$

Once again let us change the variable and set  $t = \left(\frac{1}{\rho}\right)^{-1}(u)$ . The integral above becomes :

$$\int_0 \frac{t\rho'(t)}{\rho(t)} dt.$$

Now we are in a position to apply Karamata's theorem to  $\rho'$  : since  $\rho$  is regularly varying at 0 with index  $d \geq 1$  (since  $\rho'(0) = 0$ ), and monotone in a right neighborhood of zero,  $\rho'$  is also regularly varying with index  $\geq 0$  (see Theorem 1.732.b p.39 in Bingham et al.(1987)). Then we can apply the direct part of Karamata's Theorem to  $\rho'$  (see ibid. Theorem 1.5.11 (i) p.28 where the limit should be taken here at zero) and

$$\lim_{t \rightarrow 0} \frac{t\rho'(t)}{\rho(t)} < +\infty$$

which ensures that the integral above converges and finally that  $\sum \frac{1}{a_i^2} < +\infty$ . this completes the proof of Lemma 5. ■

**Proof of Theorem 3 :**

Take  $X = \left(\frac{x_1}{a_1}, \frac{x_2}{a_2}, \dots\right)$  where the  $a_i$ 's were defined at Lemma (5) and where the  $x_i$ 's are i.i.d.  $\mathcal{N}(0, 1)$ . By Theorem 2 we know that the small ball probability of  $X$ , denoted  $\tilde{F}$  is Gamma-varying at 0 and by (4) that :

$$\lim_{\varepsilon \rightarrow 0} e^{I(\theta)} \psi(\theta) \tilde{F}(\varepsilon) = 1/\sqrt{2\pi}.$$

For the sake of notational simplicity we identify  $\tilde{F}$  with the function  $\varepsilon \rightarrow \psi \circ \theta(\varepsilon) \times e^{I \circ \theta(\varepsilon)}$ . In order to complete the proof of the Theorem we just have to check that the auxiliary function of  $\tilde{F}$  is, up to asymptotic equivalence,  $\rho$ . We follow Corollary 3.10.5 (b) p. 177 in Bingham et al. (1987) and claim that it is enough to prove that :

$$\lim_{\varepsilon \rightarrow 0} \frac{\tilde{F}(\varepsilon)}{\left(\tilde{F}\right)'(\varepsilon) \rho(\varepsilon)} \rightarrow 1.$$

It takes basic algebra to get :

$$\begin{aligned} \frac{\tilde{F}(\varepsilon)}{\left(\tilde{F}\right)'(\varepsilon)} &= \frac{-(\mu' \circ \theta)(\varepsilon)}{(I' \circ \theta)(\varepsilon) + \frac{\psi' \circ \theta(\varepsilon)}{\psi \circ \theta(\varepsilon)}} \\ &= \frac{1}{\theta(\varepsilon)} \frac{1}{1 - \frac{\psi' \circ \theta(\varepsilon)}{\psi \circ \theta(\varepsilon)} \frac{1}{\theta(\varepsilon)(\mu' \circ \theta)(\varepsilon)}}. \end{aligned}$$

We are going to show that :

$$\frac{\psi' \circ \theta(\varepsilon)}{\psi \circ \theta(\varepsilon)} \frac{1}{\theta(\varepsilon) (\mu' \circ \theta)(\varepsilon)} \rightarrow 0$$

when  $\varepsilon \rightarrow 0$  which comes down to proving :

$$\frac{\psi'(\theta)}{\psi(\theta)} \frac{1}{\theta \mu'(\theta)} = \frac{\theta \psi'(\theta)}{\psi(\theta)} \frac{1}{\theta^2 \mu'(\theta)} \rightarrow 0$$

when  $\theta \rightarrow +\infty$ . From display (25) within Lemma 4 we know that :

$$\sup_{\theta} \left| \frac{\theta \psi'(\theta)}{\psi(\theta)} \right| \leq 1$$

and finally we focus on

$$|\theta^2 \mu'(\theta)| = 2 \sum_{i=1}^{+\infty} \left( \frac{\theta}{a_i^2 + 2\theta} \right)^2 \geq 2 \sum_{i=1}^{N_\theta} \left( \frac{\theta}{a_i^2 + 2\theta} \right)^2$$

where  $N_\theta = \max \{i : a_i^2 \leq \theta\}$  and  $|\theta^2 \mu'(\theta)| \geq \frac{2}{9} N_\theta$ . At last since  $N_\theta \rightarrow +\infty$  when  $\theta$  does we get

$$\lim_{\varepsilon \rightarrow 0} \frac{\tilde{F}(\varepsilon) \theta(\varepsilon)}{\left(\tilde{F}\right)'(\varepsilon)} \rightarrow 1$$

but  $\mu \circ \theta(\varepsilon) = \varepsilon$  hence  $\theta = \mu^{-1} = 1/\rho$  which finishes the proof of Theorem 3. ■

**Proof of Proposition 4 :**

First we prove that  $G$  belongs to  $\Gamma_0$ . It suffices to show that for a well-chosen  $\rho_G$ ,

$$\frac{G(s + x\rho_G(s))}{G(s)} \rightarrow \exp(x)$$

when  $s$  tends to 0. But

$$\frac{G(s + x\rho_G(s))}{G(s)} = \frac{F(s^2 + x^2\rho_G^2(s) + 2xs\rho_G(s))}{F(s^2)} \sim \frac{F(s^2 + 2xs\rho_G(s))}{F(s^2)}$$

since  $\rho_G^2(s) = o(s\rho_G(s))$  and by applying the counterpart of Proposition 3.10.2 in Bingham et al. (1987) in the case of  $\Gamma_0$ . By the way this property of uniform convergence in (10) clearly holds for  $\Gamma_0$  as well and its proof is omitted (Dini's theorem may be applied since  $F$  as monotone). Now setting  $2s\rho_G(s) = \rho_F(s^2)$  yields

$$\frac{G(s + x\rho_G(s))}{G(s)} \sim \frac{F(s^2 + x\rho_F(s^2))}{F(s^2)} \rightarrow \exp(x).$$

It is simple to check that  $\rho_G(s)/s \rightarrow 0$  and that  $\rho_G(s)$  is self-neglecting. ■

Finally we turn to the proof of Theorem 4 and start with a Lemma. This Lemma, its proof and the subsequent proof of the theorem adapt the derivation of Lemma 2.11.2 and Theorem 2.11.3 of Bingham et al. (1987).

**Lemma 6** *Let  $\rho$  be self-neglecting at 0. For  $x_0 > 0$  sufficiently small the sequence  $x_n = x_{n-1} - \rho(x_{n-1})$  tends to 0.*

**Proof :** First note that the sequence  $x_n$  is decreasing since  $\rho \geq 0$  and notice from the properties of self-neglecting functions (namely  $\rho(s)/s \rightarrow 0$  when  $s \rightarrow 0$ ) that for a sufficiently small  $x_0 > 0$ ,  $x_n \geq 0$  for all  $n$ . The limit of  $x_n$  exists, is denoted  $l$ . Suppose that  $l > 0$ . Then  $\rho(l) > 0$  and since  $\rho$  is a non decreasing function  $\rho(x_k) \geq \rho(l)$  for all  $k$ . At last

$$\begin{aligned} x_n &= x_{n-1} - \rho(x_{n-1}) = x_0 - \sum_{k=0}^{n-1} \rho(x_k) \\ &\leq x_0 - n\rho(l). \end{aligned}$$

Letting  $n$  go to infinity  $x_n$  goes to  $-\infty$  which contradicts  $x_n \geq 0$  hence the Lemma. ■

**Proof of Theorem 4:**

Let  $x_n$  be as in the preceding Lemma. Let  $p$  be a  $C^\infty$  probability density on  $[0, 1]$  and set for  $x_{n+1} \leq u \leq x_n$

$$\varepsilon(u) = \frac{\ln \phi(x_{n+1}) - \ln \phi(x_n)}{x_n - x_{n+1}} p\left(\frac{x_n - u}{x_n - x_{n+1}}\right) \rho(u).$$

The proof takes three steps.

We prove first that for all  $x_n$ ,  $\phi(x_n) = \exp\left(\int_{x_n}^1 \frac{\varepsilon(u)}{\rho(u)} du\right)$ . In fact we may always define  $\varepsilon(u)$ ,  $x_0 \leq u \leq 1$  such that  $\phi(x_0) = \exp\left(\int_{x_0}^1 \frac{\varepsilon(u)}{\rho(u)} du\right)$ . Then assume that  $\phi(x_k) = \exp\left(\int_{x_k}^1 \frac{\varepsilon(u)}{\rho(u)} du\right)$  for  $k = 0, 1, \dots, n$ . We have :

$$\begin{aligned} \int_{x_{n+1}}^1 \frac{\varepsilon(u)}{\rho(u)} du &= \int_{x_{n+1}}^{x_n} \frac{\varepsilon(u)}{\rho(u)} du + \int_{x_n}^1 \frac{\varepsilon(u)}{\rho(u)} du \\ &= \ln \phi(x_n) + \frac{\ln \phi(x_{n+1}) - \ln \phi(x_n)}{x_n - x_{n+1}} \int_{x_{n+1}}^{x_n} p\left(\frac{x_n - u}{x_n - x_{n+1}}\right) du \\ &= \ln \phi(x_n) - (\ln \phi(x_{n+1}) - \ln \phi(x_n)) \int_1^0 p(t) dt \\ &= \ln \phi(x_{n+1}) \end{aligned}$$

Second we prove that for  $x_{n+1} \leq x \leq x_n$   $\lim_{x \rightarrow 0} \phi(x) / \phi(x_n) = 1$ . We note that  $x = x_n - \lambda_x \rho(x_n)$  where  $\lambda_x \in [0, 1]$  hence

$$\lim_{x \rightarrow 0} \frac{\phi(x_n - \lambda_x \rho(x_n))}{\phi(x_n)} = 1$$

uniformly with respect to  $\lambda_x \in [0, 1]$ .

The third and last step is devoted to proving that  $|\varepsilon(u)| \rightarrow 0$  when  $u \rightarrow 0$ . Indeed for all  $x_{n+1} \leq u \leq x_n$ ,

$$|\varepsilon(u)| \leq |p|_\infty \left| \frac{\ln \phi(x_{n+1}) - \ln \phi(x_n)}{x_n - x_{n+1}} \right| \rho(u)$$

We focus on

$$\begin{aligned} \left| \frac{\ln \phi(x_{n+1}) - \ln \phi(x_n)}{x_n - x_{n+1}} \right| \rho(u) &= \frac{\rho(u)}{\rho(x_n)} \ln \frac{\phi(x_n)}{\phi(x_{n+1})} \\ &= \frac{\rho(x_n - \lambda_u \rho(x_n))}{\rho(x_n)} \ln \frac{\phi(x_n)}{\phi(x_n - \rho(x_n))} \end{aligned}$$

Just like above  $\rho(x_n - \lambda_u \rho(x_n)) / \rho(x_n) \rightarrow 1$  since  $\rho$  is self-neglecting. Finally by the definition of  $\phi$  we get

$$\ln \frac{\phi(x_n)}{\phi(x_n - \rho(x_n))} \rightarrow 0$$

which finishes the proof of the Theorem. ■

## References

- [1] DE ACOSTA A. (1983) : Small deviations in the functional central limit theorem with applications to functional laws of the iterated logarithm, *Ann. Probab.*, **11**, 78-101.
- [2] BERLINET A., ELAMINE A., MAS A. (2008) : Local linear regression for functional data, submitted, <http://arxiv.org/abs/0710.5218v1>.
- [3] BINGHAM N.H., GOLDIE C.M. and TEUGELS J.L. (1987) : *Regular Variations*. Encyclopedia of Mathematics and Its Applications, Cambridge University Press.
- [4] BLOOM S. (1976) : A characterization of B-slowly varying functions, *Proc. Amer. Math. Soc.*, **54**, 243-250.
- [5] BORELL C. (1976) : Gaussian Radon measures on locally convex spaces, *Math. Scand.*, **38**, 265-285.
- [6] DEMBO A., MEYER-WOLF E. and ZEITOUNI O. (1995) : Exact behavior of gaussian semi-norms, *Statist. Probab. Lett.*, **23**, 275-280.
- [7] ERICKSON K. (1980) : Rates of escape of infinite-dimensional Brownian motion, *Ann. Probab.*, **8**, 325-338.
- [8] FISHER R.A. and TIPPETT L.H.C. (1928) : Limiting forms of the frequency distribution of the largest or smaller member of a sample, *Proceedings of the Cambridge Philosophical Society*, **24**, 180-190.
- [9] GAIFFAS S. (2005) : Convergence rates for pointwise curve estimation with a degenerate design, *Math. Methods Statist.*, **14**, 1-27.
- [10] GNEDENKO B.V.(1943) : Sur la distribution limite du terme maximum d'une série aléatoire, *Ann. of Math.*, **44**, 423-453.
- [11] DE HAAN L. (1971) : A form of regular variation and its application to the domain of attraction of the double exponential distribution, *Z. Wahrscheinlichkeitstheorie. verw. Geb.*, **17**, 241-258.
- [12] DE HAAN L. (1974) : Equivalence classes of regularly varying functions, *Stochastic Process. Appl.*, **2**, 243-259.
- [13] KUELBS J., LI W.V. and TALAGRAND M. (1994) : Lim inf results for Gaussian samples and Chung's functional LIL, *Ann. Probab.*, **22**, 1879-1903.
- [14] LEDOUX M. and TALAGRAND M. (1991) : *Probability in Banach Spaces. Isoperimetry and Processes*, Ergebnisse der Mathematik und ihrer Grenzgebiete, **23**, Springer-Verlag.
- [15] LI W.V. and LINDE W. (1993) : Small ball problems for non centered gaussian measures, *Probab. Math. Statist.*, **14**, 231-251.
- [16] LI W.V. and LINDE W. (1999) : Approximation, metric entropy and small ball estimates for Gaussian measures. *Ann. Probab.*, **27**, 1556-1578.

- [17] LI W.V. and LINDE W. (2004) : Small deviations of stable processes via metric entropy, *J. Theoret. Probab.*, **17**, 261-284.
- [18] LI, W.V. and SHAO Q.-M. (2001) : Gaussian processes : Inequalities, small ball probabilities and applications, In Rao and Shanbhag (eds.), *Handbook of Statistics*, **19**, Elsevier, New-York, 533-597.
- [19] MAS A. (2008) : Local functional principal component analysis, *Complex Analysis and Operator Theory*, **2**, 135-167.
- [20] MEYER-WOLF E. and ZEITOUNI O (1993) : The probability of small gaussian ellipsoids, *Ann. Probab.*, **21**, 14-24.
- [21] MULLER H.G. and YAN X. (2001) : On local moments, *J. Multivariate Anal.*, **76**, 90-109.
- [22] TALAGRAND M. (1992) : On the rate of convergence in Strassen's LIL, *Progr. Probab.*, Birkhauser, Boston, 339-351.