# TIGHTNESS AND EXPONENTIAL TIGHTNESS OF GAUSSIAN PROBABILITIES[*]

## Paolo Baldi[**]

**Abstract.** We prove a simple criterion of exponential tightness for sequences of Gaussian r.v.'s with values in a separable Banach space from which we deduce a general result of Large Deviations which allows easily to obtain LD estimates in various situations.

## 1. Introduction

Let $(X_n)_n$ be a sequence of $E$-valued Gaussian r.v.'s, where $E$ is a separable Banach space. This article is concerned with the investigation of Large Deviation properties at speed $g$ of the sequence $(g_n^{-1/2} X_n)_n$, where $g : \mathbb{N} \to \mathbb{R}^+$ is a speed function *i.e.* such that $\lim_{n \to \infty} g_n = +\infty$.

This is a classical problem that is completely understood if the law of $X_n$ does not depend on $n$ (see [8], Sect. 5 or the exposition in [1], p. 39).

In [2] (see also [6], Sect. 4.5.3) it is proved that if, for every $\theta \in E'$,

$$\lim_{n \to \infty} \frac{1}{g_n} \log \mathrm{E}[\mathrm{e}^{g_n \langle \theta, Y_n \rangle}] := \Lambda(\theta) \tag{1.1}$$

*and* the sequence $(Y_n)_n$ is exponentially tight, then it enjoys some Large Deviations estimates at speed $g$ with a rate function $I$ which is the convex conjugate of $\Lambda$ (see the precise statement in Thm. 3.1).

It is therefore of interest to produce criteria ensuring the property of exponential tightness.

Recall that a sequence of r.v.'s $(X_n)_n$ with values in some metric space $E$ is said to be *tight* as $n \to +\infty$ if for every $\delta > 0$ there exists a compact set $K_\delta$ such that

$$\mathrm{P}(X_n \notin K_\delta) < \delta. \tag{1.2}$$

[**] Corresponding author: baldi@mat.uniroma2.it

A sequence of r.v.'s $(X_n)_n$ with values in some metric space is said to be *exponentially tight* as $n \to \infty$ at speed $g$ if for every $R > 0$ there exists a compact set $K_R$ such that

$$P(X_n \notin K_R) < e^{-g_n R}. \tag{1.3}$$

The main result of this paper (Thm. 2.4) is that if a sequence $(X_n)_n$ of Gaussian $E$-valued r.v.'s is tight, then, for every speed function $g$, the sequence $(g_n^{-1/2} X_n)_n$ is exponentially tight at speed $g$.

This result will allow us to prove the following result.

**Theorem 1.1.** *Let $(X_n)_n$ be a sequence of Gaussian r.v.'s with values in the separable Banach space $E$ and converging in law to a r.v. $X$. Then, for every speed function $g$, the sequence $(g_n^{-1/2} X_n)_n$ satisfies a Large Deviation Principle with speed $g$ and rate function*

$$I(x) = \frac{1}{2} \|x\|_{\mathcal{H}}^2 \tag{1.4}$$

*where $\| \ \|_{\mathcal{H}}$ denotes the Replicating Kernel Hilbert Space norm of the law of the r.v. $X - E[X]$, with the understanding $\|x\|_{\mathcal{H}} = +\infty$ if $x \notin \mathcal{H}$.*

In some sense Theorem 1.1 states that if $X_n \to X$ in law, then, in a Gaussian setting, the Large Deviations asymptotics of the sequences $(g_n^{-1/2} X_n)_n$ and $(g_n^{-1/2} X)_n$ are the same (*without any constraint* concerning the speed of convergence of $X_n \to X$).

Theorem 1.1 is well known if the law of $X_n$ does not depend on $n$ (see [8], Sect. 5 or the exposition in [1], p. 39). This, of course, suggests that Theorem 1.1 should be an immediate consequence of the above mentioned result for the sequence $(g_n^{-1/2} X)_n$, using the argument of exponential approximation, as explained in [6] Sect. 4.2.2. Actually it turns out that Theorem 1.1 can be proved using this argument, see Section 4, but this way of proof would also require the exponential tightness result of Theorem 2.4.

Recall that Fernique's theorem states that a Gaussian r.v. has bounded exponential moments. As a, possibly useful, by-product of our treatment, we obtain (Cor. 2.3) that a tight family of $E$-valued Gaussian r.v.'s has uniformly bounded exponential moments.

In Section 2 we prove our main result (Thm. 2.4) *i.e.* that if $(X_n)_n$ is tight then $(g_n^{-1/2} X_n)_n$ is exponentially tight at speed $g$.

In Sections 3 and 4 we give two different proofs of Theorem 1.1 whereas Section 5 is devoted to examples and counterexamples.

## 2. Exponential tightness of Gaussian families

Throughout this paper $E$ shall denote a separable Banach space, $E'$ its topological dual.

In order to investigate exponential tightness of Gaussian families we shall take advantage of the following result which is a particular case of Theorem 3.1 of De Acosta [4].

**Theorem 2.1.** ([4]) *Let $(\mu_\alpha)_{\alpha \in \mathscr{A}}$ be a family of probabilities on the separable Banach space $E$ such that*
  a) *$(\mu_\alpha)_{\alpha \in \mathscr{A}}$ is tight.*
  b) *There exists $t > 0$ such that*

$$\sup_{\alpha \in \mathscr{A}} \int_E e^{t|x|} \, d\mu_\alpha(x) < +\infty \ . \tag{2.1}$$

*Then there exists a convex compact well balanced set $K \subset E$ such that*

$$\sup_{\alpha \in \mathscr{A}} \int_E e^{q_K(x)} \, d\mu_\alpha(x) < +\infty$$

*where $q_K$ denotes the Minkowski functional of the set $K$.*

The definition of a Minkowski functional will be recalled shortly.

Remark that in Theorem 2.1 there is no assumption of Gaussianity. In order to take advantage of Theorem 2.1, we first prove that condition (2.1) is automatically satisfied if Gaussianity, in addition to the tightness condition of assumption (a) of the theorem, is enforced.

This fact will follow from the next result, whose proof follows the same line of reasoning of the classical Fernique's theorem (see [10, 11], p. 11).

**Theorem 2.2.** *Let $E$ be a separable Banach space and $\phi : E \to [0, +\infty]$ a measurable semi-norm, i.e. a measurable application such that*

$$\begin{aligned} \phi(\lambda x) &= |\lambda|\phi(x), \quad \lambda \in \mathbb{R} \\ \phi(x+y) &\leq \phi(x) + \phi(y). \end{aligned} \tag{2.2}$$

*Let $(\mu_\alpha)_{\alpha \in \mathscr{A}}$ be a family of centered Gaussian probabilities on $E$ such that $\mu_\alpha(\phi < +\infty) = 1$ for every $\alpha \in \mathscr{A}$. If there exists $s > 0$ such that $\mu_\alpha(\phi > s) \leq \beta < \frac{1}{2}$ for every $\alpha \in \mathscr{A}$, then there exists $a > 0$ such that*

$$\sup_{\alpha \in \mathscr{A}} \int_E e^{a\phi^2(x)} \mu_\alpha(\mathrm{d}x) < \infty .$$

*Proof.* From the relation

$$\mu_\alpha(\phi > s) \leq \beta < \frac{1}{2} , \tag{2.3}$$

following the lines of the proof of Fernique's theorem (see also [7], Sect. 1.3) if we define by recurrence the sequence $(t_n)_n$ by

$$t_0 = s \qquad t_n = \sqrt{2}\,t_{n-1} + s$$

*i.e.*

$$t_n = s\big(1 + \sqrt{2} + \cdots + (\sqrt{2})^n\big) = s\frac{(\sqrt{2})^{n+1} - 1}{\sqrt{2} - 1} \leq \underbrace{s\Big(\frac{\sqrt{2}}{\sqrt{2} - 1}\Big)}_{:= \sqrt{\zeta}} 2^{n/2}, \tag{2.4}$$

we obtain, for every $\alpha \in \mathscr{A}$,

$$\mu_\alpha(\phi > t_n) \leq \kappa^{2^n} \mu_\alpha(\phi < s) \tag{2.5}$$

where

$$\kappa = \frac{\beta}{1 - \beta} < 1$$

We can now split the integral

$$\int_E e^{a\phi^2(x)} \mu_\alpha(\mathrm{d}x) \leq \underbrace{\int_{\{\phi^2 \leq \zeta\}} e^{a\phi^2(x)} \mu_\alpha(\mathrm{d}x)}_{\leq e^{a\zeta}} + \int_{\{\phi^2 > \zeta\}} e^{a\phi^2(x)} \mu_\alpha(\mathrm{d}x).$$

But

$$\int_{\{\phi^2>\zeta\}} e^{a\phi^2(x)}\,\mu_\alpha(\mathrm{d}x) \le \sum_{n=0}^\infty \int_{\{\zeta\cdot 2^n<\phi^2\le\zeta\cdot 2^{n+1}\}} e^{a\phi^2(x)}\,\mu_\alpha(\mathrm{d}x)$$
$$\le \sum_{n=0}^\infty e^{a\zeta\cdot 2^{n+1}}\mu_\alpha(\phi^2>\zeta\cdot 2^n)\ .$$

Since $t_n^2 \le \zeta 2^n$, (2.5) gives

$$\int_E e^{a\phi^2(x)}\,\mu_\alpha(\mathrm{d}x) \le e^{a\zeta} + \sum_{n=0}^\infty e^{a\zeta\cdot 2^{n+1}}\mu_\alpha(\phi>t_n) \le e^{a\zeta} + \sum_{n=0}^\infty e^{a\zeta\cdot 2^{n+1}}\mu_\alpha(\phi<s)\cdot\kappa^{2^n}$$

which, for $a < \frac{\log\frac{1}{\kappa}}{2\zeta}$, gives a convergent series thus concluding the proof.  $\square$

**Corollary 2.3.** *Let $(\mu_\alpha)_{\alpha\in\mathscr{A}}$ be a family of centered Gaussian probabilities on the Banach space $E$, such that there exists $s > 0$ such that $\mu_\alpha(|x|>s)\le\beta<\frac{1}{2}$ for every $\alpha\in\mathscr{A}$. Then*

$$\sup_{\alpha\in\mathscr{A}}\int_E e^{t|x|}\,\mathrm{d}\mu_\alpha(x) < +\infty, \qquad \text{for every } t>0\ . \tag{2.6}$$

*In particular (2.6) holds if the family $(\mu_\alpha)_{\alpha\in\mathscr{A}}$ is tight.*

*Proof.* Let $a > 0$ be such that

$$\sup_{\alpha\in\mathscr{A}}\int_E e^{a|x|^2}\,\mathrm{d}\mu_\alpha(x) < +\infty$$

as guaranteed by Theorem 2.2, then

$$\int_E e^{t|x|}\,\mathrm{d}\mu_\alpha(x) = \int_{\{|x|\le t/a\}} e^{t|x|}\,\mathrm{d}\mu_\alpha(x) + \int_{\{|x|>t/a\}} e^{t|x|}\,\mathrm{d}\mu_\alpha(x) \le$$
$$\le e^{t^2/a} + \int_E e^{a|x|^2}\,\mathrm{d}\mu_\alpha(x) < +\infty\ .$$

Finally remark that the condition $\mu_\alpha(|x|>s)\le\beta<\frac{1}{2}$ for every $\alpha\in\mathscr{A}$ is automatically satisfied if $(\mu_\alpha)_{\alpha\in\mathscr{A}}$ is tight, as compact sets are bounded in $E$.  $\square$

Let $K\subset E$. Let us recall that its Minkowski functional is a map $q_K : E \to [0,+\infty]$ defined as

$$q_K(x) = \inf\{t>0; x\in tK\}\ .$$

It is well known that if $K$ is a convex set then $q_K$ is subadditive and that if $K$ is well balanced then $q_K$ is positively homogeneous. Well balanced means that if $x\in K$ then also $tx\in K$ for every $0\le t\le 1$.

The following is the main result of this section.

**Theorem 2.4.** *Let $(X_\alpha)_{\alpha\in\mathscr{A}}$ be a tight family of Gaussian $E$-valued r.v.'s. Then for every speed function $g$ the family $(g_n^{-1/2}X_\alpha)_n$ is uniformly exponentially tight, i.e., for every $R>0$ there exists a compact set $K_R\subset E$ such that*

$$\mathrm{P}(g_n^{-1/2}X_\alpha\notin K_R) \le e^{-g_n R}$$

*for every $\alpha \in \mathscr{A}$ and for every $n$. In particular, if a sequence $(X_n)_n$ of Gaussian $E$-valued r.v.'s is tight, then for every speed function $g$, $(g_n^{-1/2} X_n)_n$ is exponentially tight.*

*Proof.* Let us assume first that the r.v.'s $X_\alpha$ are centered. By Corollary 2.3 also condition (b) of Theorem 2.1 is satisfied. Thus, by Theorem 2.1 applied to the family $(\mu_\alpha)_{\alpha \in \mathscr{A}}$ of the laws of the $X_\alpha$'s, there exists a convex compact well balanced set $K$ such that

$$\sup_{\alpha \in \mathscr{A}} \mathrm{E}\big[\mathrm{e}^{q_K(X_\alpha)}\big] := C < +\infty .$$

For every $\alpha \in \mathscr{A}$ let $(Z_n)_n$ be a sequence of i.i.d. r.v.'s with the same law as $X_\alpha$. Let us assume for simplicity, at first, that $g_n$ is an integer number for every $n$. We have

$$g_n^{-1/2}(Z_1 + \cdots + Z_{g_n}) \sim X_\alpha \tag{2.7}$$

hence

$$\mathrm{E}\big[\mathrm{e}^{q_K(g_n^{1/2} X_\alpha)}\big] = \mathrm{E}\big[\mathrm{e}^{q_K(Z_1 + \cdots + Z_{g_n})}\big] \le \mathrm{E}\big[\mathrm{e}^{q_K(Z_1) + \cdots + q_K(Z_{g_n})}\big] = \big(\mathrm{E}\big[\mathrm{e}^{q_K(X_\alpha)}\big]\big)^{g_n} \le C^{g_n} .$$

Remark the relation

$$\{x \notin tK\} = \{x; q_K(x) > t\} . \tag{2.8}$$

Hence for every $t > 0$, by Markov inequality,

$$\mathrm{P}\big(g_n^{-1/2} X_\alpha \notin tK\big) = \mathrm{P}\big(g_n^{1/2} X_\alpha \notin tg_n K\big) = \mathrm{P}\big(q_K(g_n^{1/2} X_\alpha) \ge tg_n\big)$$
$$\le \mathrm{e}^{-tg_n} \mathrm{E}\big[\mathrm{e}^{q_K(g_n^{1/2} X_\alpha)}\big] \le \mathrm{e}^{(\log C - t)g_n}$$

and if we chose $t = R + \log C$, the compact $K_R = tK$ satisfies the requirement (1.3).

In general, if $g$ is not integer valued, (2.7) will become $g_n^{-1/2}(Z_1 + \cdots + Z_{\lfloor g_n \rfloor}) \sim \lfloor g_n \rfloor^{1/2} g_n^{-1/2} X_\alpha$, $\lfloor \ \rfloor$ denoting the integer part function and the remainder of the proof is to be modified accordingly.

If we do not assume the $X_\alpha$'s to be centered, let us denote by $x_\alpha = \mathrm{E}[X_\alpha]$ their means and let $\widetilde{X}_\alpha = X_\alpha - x_\alpha$. The assumption of tightness implies that the sequence $(x_\alpha)_\alpha$ is contained in some compact set $K$, that we can assume to be well balanced, and that also the sequence $(\widetilde{X}_\alpha)_\alpha$ is tight. Let $\widetilde{K}_R$ be a compact, convex and well balanced set such that $\mathrm{P}(g_n^{-1/2} \widetilde{X}_\alpha \notin \widetilde{K}_R) \le \mathrm{e}^{-g_n R}$. Then $K_R = \widetilde{K}_R + K$ is a compact set and we have

$$\mathrm{P}\big(g_n^{-1/2} X_\alpha \notin K_R\big) \le \mathrm{P}\big(g_n^{-1/2} \widetilde{X}_\alpha \notin \widetilde{K}_R\big) \le \mathrm{e}^{-g_n R} .$$

$\square$

It is natural, at this time, to inquire whether a converse of Theorem 2.4 holds. It is obvious that uniform exponential tightness, even with respect to a single speed function implies tightness.

Whether a Gaussian sequence $(X_n)_n$ that is exponentially tight with respect to every speed function is also tight is less obvious, as exponential tightness of the sequence is a weaker property than the uniform exponential tightness of the family $(X_n)_n$. A partial answer to this question is given in Example 5.4.

## 3. Large deviations

We give two proofs of Theorem 1.1. The second one, in the next section, is much shorter but maybe the first one gives more insight into the structure of Gaussian probabilities. Both proofs rely on Theorems 3.1 and 2.4.

In order to prove Theorem 1.1 we shall take advantage of the following result.

Let us recall that $x \in E$ is said to be an *exposed point* for the rate function $I$ if there exists $\eta \in E'$ such that, for $x \neq y$,

$$I(y) - I(x) > \langle \eta, y - x \rangle \tag{3.1}$$

which is a condition of strict convexity of the rate function $I$. $\eta$ is called an *exposing hyperplane*. The following theorem is an infinite dimensional extension of the Ellis-Gärtner theorem ([9, 12]).

**Theorem 3.1.** ([2], see also [6], Sect. 4.5.3) *Let $(\mu_n)_n$ be an exponentially tight sequence of probabilities at speed $g$ on the separable Banach space $E$, such that, for every $\theta \in E'$, the limit*

$$\lim_{n \to \infty} \frac{1}{g_n} \log \int_E e^{g_n \langle \theta, x \rangle} \, d\mu_n(x) := \Lambda(\theta) \tag{3.2}$$

*exists. Let us denote*

$$I(x) = \sup_{\theta \in E'} \left( \langle \theta, x \rangle - \Lambda(\theta) \right) \tag{3.3}$$

*the convex conjugate of $\Lambda$. Then for every closed set $F \subset E$*

$$\limsup_{n \to \infty} \frac{1}{g_n} \log \mu_n(F) \leq - \inf_{x \in F} I(x) \tag{3.4}$$

*and for every open set $G \subset X$,*

$$\liminf_{n \to \infty} \frac{1}{g_n} \log \mu_n(G) \geq - \inf_{x \in G \cap \mathscr{E}} I(x) \tag{3.5}$$

*where $\mathscr{E}$ denotes the set of all exposed points of $I$ such that for an exposing hyperplane $\eta$ it holds $\Lambda(s\eta) < +\infty$ for some $s > 1$.*

The application of Theorem 3.1, in order to obtain a Large Deviation Principle, is not always undemanding because it provides an incomplete lower bound. We shall see however that from (3.5) a true lower LDP bound can be easily obtained if the probabilities $\mu_n$ are Gaussian.

*First proof of Theorem 1.1.* By Theorem 2.4 the sequence $(g_n^{-1/2} X_n)_n$ is exponentially tight at speed $g$. Let us denote $x_n = \mathrm{E}[X_n]$, $\widetilde{X}_n = X_n - x_n$, $x = \mathrm{E}[X]$, $\widetilde{X} = X - x$. We have

$$\log \mathrm{E}[e^{g_n \langle \theta, g_n^{-1/2} X_n \rangle}] = g_n^{1/2} \langle \theta, x_n \rangle + \log \mathrm{E}[e^{g_n \langle \theta, g_n^{-1/2} \widetilde{X}_n \rangle}]$$

Hence condition (3.2) is satisfied as the r.v.'s $\langle \theta, X_n \rangle$, $\theta \in E'$, are Gaussian so that

$$\begin{aligned}
&\lim_{n \to \infty} \frac{1}{g_n} \log \mathrm{E}[e^{g_n \langle \theta, g_n^{-1/2} X_n \rangle}] = \lim_{n \to \infty} \frac{1}{g_n} \left( g_n^{1/2} \langle \theta, x_n \rangle + \log \mathrm{E}[e^{g_n \langle \theta, g_n^{-1/2} \widetilde{X}_n \rangle}] \right) \\
&= \lim_{n \to \infty} \frac{1}{2g_n} \mathrm{Var}(g_n^{1/2} \langle \theta, \widetilde{X}_n \rangle) = \frac{1}{2} \mathrm{Var}(\langle \theta, \widetilde{X} \rangle) = \log \mathrm{E}[e^{\langle \theta, \widetilde{X} \rangle}] := \Lambda(\theta) \ .
\end{aligned} \tag{3.6}$$

In order to conclude the proof of Theorem 1.1 we must compute the convex conjugate $I$ of $\Lambda$. This is well known (see [1], Chap II *e.g.*), however, in order to investigate the exposed points of $I$ we recall some facts about this computation and about Gaussian probabilities $\mu$ on a separable Banach space $E$.

For every $\theta \in E'$, the function $E \ni z \mapsto \langle \theta, z \rangle$ defines a r.v. on the probability space $(E, \mu)$. Hence $E'$ can be considered as a Gaussian space of r.v.'s on $(E, \mu)$. Let us denote by $\overline{E}'_\mu$ the closure of $E'$ in the Hilbert space $L^2(\mu)$. $\overline{E}'_\mu$ is a space of Gaussian r.v.'s and for every $\phi \in \overline{E}'_\mu$ we can define its barycenter

$$h = \int_E z \phi(z) \, \mathrm{d}\mu(z) . \tag{3.7}$$

The convergence of the integral in (3.7) is a consequence of well known facts about the existence of moments of Gaussian probabilities. The elements $h \in E$ of this form constitute a vector space $\mathscr{H}$ that becomes an Hilbert space when endowed with the norm

$$\|h\|_{\mathscr{H}} = \|\phi\|_{L^2(\mu)} = \int_E \phi(z)^2 \, \mathrm{d}\mu(z) \tag{3.8}$$

which makes $\mathscr{H}$ isometric with $\overline{E}'_\mu$. Actually it is easy to see that (3.8) is a good definition: if $h$ is the barycenter of $\phi$ and of $\phi'$, $\phi, \phi' \in \overline{E}'_\mu$, then $\|\phi\|_{L^2(\mu)} = \|\phi'\|_{L^2(\mu)}$ necessarily.

It is important to remark that the Hilbert space $\mathscr{H}$ is dense in the closed subspace $\widetilde{E} = \mathrm{supp}\,\mu$. Actually if $\theta \in E'$ is such that $\langle \theta, h \rangle = 0$ for every $h \in \mathscr{H}$, then, if $h$ is the barycenter of $\phi \in \overline{E}'_\mu$, we would have

$$0 = \langle \theta, h \rangle = \int_E \langle \theta, x \rangle \phi(x) \, \mathrm{d}\mu(x) .$$

The function $x \mapsto \langle \theta, x \rangle$ being orthogonal to every $\phi \in \overline{E}'_\mu$, necessarily $\langle \theta, x \rangle = 0$ $\mu$-a.s.

The functional $\Lambda : E' \to \mathbb{R}$ defined in (3.6) is convex and, as we already remarked, the computation of its convex conjugate $I$, *i.e.*

$$I(x) = \sup_{\theta \in E'} \left( \langle \theta, x \rangle - \Lambda(\theta) \right) \tag{3.9}$$

is a classical fact (see [1], Prop. 1.5, p. 53 and the literature therein). Actually, if $x \in \mathscr{H}$ and $x = \int z \phi(z) \, \mathrm{d}\mu(z)$ for some $\phi \in \overline{E}'_\mu$, we have

$$
\begin{aligned}
I(x) &= \sup_{\theta \in E'} \left( \int_E \langle \theta, z \rangle \phi(z) \, \mathrm{d}\mu(z) - \frac{1}{2} \int_E \langle \theta, z \rangle^2 \, \mathrm{d}\mu(z) \right) \\
&= \sup_{\theta \in E'} \left( -\frac{1}{2} \int (\langle \theta, z \rangle - \phi(z))^2 \, \mathrm{d}\mu(z) \right) + \frac{1}{2} \int_E \phi(z)^2 \, \mathrm{d}\mu(z)
\end{aligned}
\tag{3.10}
$$

from which, by considering a sequence $(\theta_n)_n \subset E'$ converging to $\phi$ in $\overline{E'_\mu}$, we have

$$\sup_{\theta \in E'} \left( -\frac{1}{2} \int (\langle \theta, z \rangle - \phi(z))^2 \, \mathrm{d}\mu(z) \right) = 0$$

and

$$I(x) = \frac{1}{2} \int_E \phi(z)^2 \, \mathrm{d}\mu(z) = \frac{1}{2} \|x\|_{\mathscr{H}}^2 .$$

Moreover if $\phi \in E'$ (which is a stronger assumption than $\phi \in \overline{E}'_\mu$) the supremum in (3.10) is attained at $\theta = \phi$ so that in this case we have also $I(x) = \langle \phi, x \rangle - \Lambda(\phi) = \Lambda(\phi)$. See [1] Proposition 1.5 p. 53 for a proof that

$I(x) = +\infty$ if $x \notin \mathcal{H}$. Then, thanks to Theorem 3.1, the Large Deviation upper bound (3.4) holds with respect to the rate function (1.4). In order to complete the proof of Theorem 1.1 we must show that the exposed points, $\mathcal{E}$, of $I$ satisfying the condition of Theorem 3.1 are such that, for every open set $A \subset E$,

$$\inf_{x \in A \cap \mathcal{E}} I(x) = \inf_{x \in A} I(x) .$$

This is the consequence of the following lemma.                                                    $\square$

**Lemma 3.2.** *Let us denote by $\widetilde{\mathcal{H}}$ the elements of $\mathcal{H}$ which are of the form (3.7) with $\phi(z) = \langle \lambda, z \rangle$, $\lambda \in E'$. Then the elements of $\widetilde{\mathcal{H}}$ are exposed points with exposing hyperplane $\lambda$ and such that $\Lambda(s\lambda) < +\infty$ for some $s > 1$. Moreover for every open set $A \subset E$ we have*

$$\inf_{A \cap \mathcal{E}} I = \inf_A I \tag{3.11}$$

*where $\mathcal{E}$ is as in the statement of Theorem 3.1.*

*Proof.* Remember, from (3.10), that if $x \in \widetilde{\mathcal{H}}$ and $x = \int_E \langle \lambda, z \rangle z \, d\mu(z)$ then the supremum in (3.9) is attained at $\theta = \lambda$ and that $I(x) = \Lambda(\lambda) = \langle \lambda, x \rangle - \Lambda(\lambda)$. Let us prove that such an $x$ is an exposed point with exposing hyperplane $\lambda$.

Let $y \in \mathcal{H}$, $y \neq x$, be such that $y = \int_E z\phi(z) \, d\mu(z)$ for some $\phi \in \overline{E}'_\mu$. Then

$$I(y) = \frac{1}{2} \int_E \phi(z)^2 \, d\mu(z) > \frac{1}{2} \int_E \phi(z)^2 \, d\mu(z) - \frac{1}{2} \underbrace{\int_E (\langle \lambda, z \rangle - \phi(z))^2 \, d\mu(z)}_{>0}$$

$$= \int_E \langle \lambda, z \rangle \phi(z) \, d\mu(z) - \frac{1}{2} \int_E \langle \lambda, z \rangle^2 \, d\mu(z) = \langle \lambda, y \rangle - \Lambda(\lambda).$$

Hence, as $I(x) = \langle \lambda, x \rangle - \Lambda(\lambda)$,

$$I(y) - I(x) > \langle \lambda, y \rangle - \Lambda(\lambda) - (\langle \lambda, x \rangle - \Lambda(\lambda)) = \langle \lambda, y - x \rangle$$

so that every $x \in \widetilde{\mathcal{H}}$ is an exposed point. The condition $\Lambda(s\lambda) < +\infty$ for some $s > 1$ is obviously satisfied as $\Lambda(s\lambda) = s^2 \Lambda(\lambda)$.

In order to prove (3.11) we need only to show that $\inf_{A \cap \mathcal{E}} I \leq \inf_A I$. As $I \equiv +\infty$ outside $\mathcal{H}$,

$$\inf_A I = \inf_{A \cap \mathcal{H}} I .$$

Let us assume first that $\operatorname{supp}\mu = E$, so that $\mathcal{H}$ is dense in $E$. As $\widetilde{\mathcal{H}}$ is dense in $\mathcal{H}$, for every $x \in A \cap \mathcal{H}$ there exists a sequence $(x_n)_n \subset \widetilde{\mathcal{H}}$ converging to $x$ in the topology of $\mathcal{H}$, hence such that $I(x_n) \to I(x)$. As the topology of $\mathcal{H}$ is stronger than the topology of $E$, $(x_n)_n$ also converges to $x$ in $E$, whence $x_n \in A$ for $n$ large (recall that $A$ is an open set). Therefore, for every $x \in \mathcal{H} \cap A$ and $n > 0$ there exists $y \in \widetilde{\mathcal{H}} \cap A$ such that $I(y) \leq I(x) + \frac{1}{n}$. As $\widetilde{\mathcal{H}} \subset \mathcal{E}$, we can conclude as

$$\inf_A I = \inf_{A \cap \mathcal{H}} I = \inf_{A \cap \widetilde{\mathcal{H}}} I \geq \inf_{A \cap \mathcal{E}} I.$$

If $\operatorname{supp}\mu \neq E$ then $\widetilde{E} := \operatorname{supp}\mu$ is a proper closed subspace of $E$ and it is immediate that $I \equiv +\infty$ on $E \setminus \widetilde{E}$. Actually if $x \in E \setminus \widetilde{E}$, by the Hahn-Banach theorem there exists $\lambda \in E'$ such that $\langle \lambda, z \rangle = 0$ for every $z \in \widetilde{E}$

and $\langle \lambda, x \rangle \neq 0$. Hence

$$I(x) = \sup_{\theta \in E'} \big( \langle \theta, x \rangle - \Lambda(\theta) \big) \geq \sup_{t \in \mathbb{R}} \big( t \langle \lambda, x \rangle - \Lambda(t\lambda) \big) = \sup_{t \in \mathbb{R}} t \langle \lambda, x \rangle = +\infty$$

Moreover, considering $\mu$ as a Gaussian probability on $\widetilde{E}$, $\mathscr{H}$ is dense in $\widetilde{E}$ (in the topology of $\widetilde{E}$) and by a repetition of the previous argument we have

$$\inf_A I = \inf_{A \cap \widetilde{E}} I = \inf_{A \cap \widetilde{E} \cap \mathscr{H}} I = \inf_{A \cap \widetilde{E} \cap \widetilde{\mathscr{H}}} I \geq \inf_{A \cap \widetilde{E} \cap \mathscr{E}} I = \inf_{A \cap \mathscr{E}} I.$$

$\square$

Remark that the argument above provides a proof of the LDP of Theorem 1.1 which is new also in the classical case, *i.e.* when $X_n \equiv X$. At least to the author's knowledge.

## 4. Second proof: exponential approximation

The second proof of Theorem 1.1 is based on the notion of exponential equivalence (see [6], Sect. 4.2.2). Theorem 4.2.13 there states that if, for every $\delta > 0$,

$$\limsup_{n \to 0} \frac{1}{g_n} \log P(|g_n^{-1/2} Y_n - g_n^{-1/2} X_n| \geq \delta) = -\infty \tag{4.1}$$

then if $(g_n^{-1/2} Y_n)_n$ satisfies a Large Deviation Principle with speed $g$ and rate function $I$, the same is true for $(g_n^{-1/2} X_n)_n$. We shall apply this criterion with $Y_n \equiv X$.

By Skorokhod representation theorem, we can assume that the r.v.'s in the statement of Theorem 1.1 are defined on the same probability space and that $X_n \to X$ a.s. Hence

$$\lim_{n \to \infty} \frac{1}{g_n} \log E[e^{g_n^{1/2} \langle \theta, X_n - X \rangle}] = \lim_{n \to \infty} \frac{1}{2} \mathrm{Var}(\langle \theta, X_n - X \rangle) = 0. \tag{4.2}$$

By Theorem 2.4 the sequence $(g_n^{-1/2}(X_n - X))_n$ is exponentially tight at speed $g$ and by (4.2) it satisfies the relation (3.2) with $\Lambda \equiv 0$. We can therefore apply the upper bound (3.4) of Theorem 3.1 with respect to the convex conjugate $I$ of $\Lambda \equiv 0$, *i.e.* $I(0) = 0$, $I(x) = +\infty$ for every $x \neq 0$ and to the closed set $F = \{|x| \geq \delta\}$ in order to obtain

$$\limsup_{n \to \infty} \frac{1}{g_n} \log P(|g_n^{-1/2} X_n - g_n^{-1/2} X| \geq \delta) = -\inf_F I = -\infty \ . \tag{4.3}$$

Hence $(g_n^{-1/2} X_n)_n$ satisfies a Large Deviation principle with the same rate function as the sequence $(g_n^{-1/2} X)_n$ and the result follows from the classical Large Deviation results of [8]. $\square$

It does not seem immediate to verify the exponential equivalence relation (4.1) under the hypotheses of Theorem 1.1 without taking advantage of Theorems 2.4 and 3.1.

## 5. Examples and concluding remarks

**Example 5.1.** Let $B$ be a real Brownian motion and, for every integer $n$ and $t \leq T$,

$$X_t^{(n)} = \sqrt{2} \int_0^t \sin(ns) \, \mathrm{d}B_s \ .$$

We have $X_t^{(n)} = W_{A_n(t)}^{(n)}$ where $A_n(t) = 2 \int_0^t \sin^2(ns)\,ds$ and $W^{(n)}$ is another Brownian motion. As $A_n(t) \to t$ uniformly as $n \to \infty$, the sequence of processes $(X^{(n)})_n$ converges in law to the Wiener measure. Hence $(X^{(n)})_n$ is a Gaussian tight sequence taking values in the Banach space $E = \mathscr{C}([0,T], \mathbb{R})$ of continuous paths endowed with the topology of uniform convergence and, by Theorem 1.1, $(g_n^{-1/2} X^{(n)})_n$ satisfies a Large Deviation Principle at speed $g$ with the same rate function as in Schilder's theorem ([13]) for the Brownian motion.

**Example 5.2.** Let again $E = \mathscr{C}([0,T], \mathbb{R})$ and let $X_n$ be a fractional Brownian motion with parameter $H_n$ with $\lim_{n \to \infty} H_n = H$, $0 < H < 1$. It is well known that $X_n \to_{n \to \infty} X$, where $X$ is a fractional Brownian motion with Hurst index $H$, the convergence being in law. This follows easily: the convergence of the covariance functions implies the convergence of the finite dimensional distributions and the equality

$$\mathrm{E}[(X_n(t) - X_n(s))^{2p}] = \frac{(2p)!}{2^p p!}\,|t - s|^{2pH_n}$$

implies the tightness of the sequence thanks to Billingley's criterion ([3], Thm. 12.3). Hence Theorem 1.1 implies immediately that, for any speed function $g$, the sequence $(g_n^{-1/2} X_n)_n$ enjoys a Large Deviation estimate with respect to the rate function given by the RKHS of the fractional Brownian motion with Hurst parameter equal to $H$, as can be found in [5] *e.g.*

It is well known that, if $E$ is finite dimensional, (3.2) with the assumption that the limit $\Lambda$ is finite in a neighborhood of the origin implies exponential tightness, even without Gaussianity. The author has been wondering whether (3.2) might imply exponential tightness for Gaussian sequences in infinite dimensions. Here we produce a counterexample.

**Example 5.3.** Let us denote, as usual, by $c_0$ the space of real sequences $x = (x_n)_n$ such that $\lim_{n \to \infty} x_n = 0$ which is a separable Banach space with respect to the sup norm.

Let $(\xi_n)_n$ be a sequence of i.i.d. $N(0,1)$-distributed r.v.'s and

$$X_n = (\xi_1, \xi_2, \ldots, \xi_n, 0, 0, \ldots)$$

be a corresponding $c_0$-valued sequence of r.v.'s. We know that $(X_n)_n$ does not converge, hence it is not tight in $c_0$, because this would require $\lim_{n \to \infty} \xi_n = 0$ a.s.

We prove first that $(\frac{1}{\sqrt{n}} X_n)_n$ is not exponentially tight at speed $g_n = n$, *i.e.* that, for $R > 0$ fixed, there exist no compact set $K = K_R \subset c_0$ such that, for $n \geq n_0$,

$$\mathrm{P}\Big(\frac{1}{\sqrt{n}} X_n \notin K\Big) \leq \mathrm{e}^{-nR}\ .$$

Actually we prove that, for every compact set $K \subset c_0$,

$$\liminf_{n \to \infty} \frac{1}{n} \log \mathrm{P}\Big(\frac{1}{\sqrt{n}} X_n \notin K\Big) = 0\ . \tag{5.1}$$

It is easy to see that every such compact set $K$ is contained in a set of the form $\{x, |x_k| \leq a_k$ for every $k = 1, 2, \ldots\}$, where $(a_k)_k$ is a positive sequence converging to 0 as $k \to +\infty$. We have, for every $k_0 > 0$ and $n \geq k_0$,

$$\mathrm{P}\Big(\frac{1}{\sqrt{n}} X_n \notin K\Big) \geq \mathrm{P}\big(|\xi_k| > a_k \sqrt{n} \text{ for at least a } k = 1, \ldots, n\big)$$

$$\geq \mathrm{P}(|\xi_{k_0}| > a_{k_0} \sqrt{n}) = 2\big(1 - \Phi(a_{k_0} \sqrt{n}\,)\big)$$

where by $\Phi$ we denote the $N(0,1)$ partition function. Now, thanks to classical estimates on the behavior of $\Phi$ at infinity,

$$\liminf_{n\to\infty} \frac{1}{n} \log\Big(2\big(1 - \Phi(a_{k_0}\sqrt{n}\,)\big)\Big) \geq -\frac{a_{k_0}^2}{2}$$

hence, $k_0$ being arbitrary and $(a_k)_k$ infinitesimal, we have (5.1).

However the limit (3.2) exists in this case. Recall that here $E' = \ell_1$, the space of the sequences $\theta = (\theta_n)_n$ such that $\sum_{n=1}^{\infty} |\theta_n| < +\infty$. As the r.v.

$$\frac{1}{\sqrt{n}} \langle \theta, X_n \rangle = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \theta_k \xi_k$$

is Gaussian, we have

$$\log \mathrm{E}[\mathrm{e}^{n^{1/2}\langle\theta,X_n\rangle}] = \frac{n}{2} \mathrm{Var}(\langle\theta,X_n\rangle) = \frac{n}{2} \sum_{k=1}^{n} \theta_k^2$$

hence the limit in (3.2) exists and takes the value $\frac{1}{2} \sum_{k=1}^{\infty} |\theta_k|^2 < +\infty$ for every $\theta \in E'$.

**Example 5.4.** Let $(\xi_n)_n$ be a sequence of i.i.d. $N(0,1)$-distributed r.v.'s and let

$$X_n = (a_1\xi_1, \ldots, a_n\xi_n, 0, 0, \ldots)$$

be a $c_0$-valued r.v. where $(a_n)_n$ is a positive sequence converging to 0. In this example we show that for some choice of the sequence $(a_n)_n$ the sequence $(X_n)_n$ is not tight but is exponentially tight, at least for some choices of the speed function.

Let us investigate the tightness of $(X_n)_n$. Let $K = \{x \in c_0, |x_k| \leq b_k\}$, where $(b_k)_k$ is a positive sequence converging to 0, be a compact set of $c_0$. We have

$$\mathrm{P}(X_n \in K^c) = 1 - \mathrm{P}(X_n \in K) = 1 - \prod_{k=1}^{n} \mathrm{P}\Big(|\xi_k| \leq \frac{b_k}{a_k}\Big).$$

The quantity $\mathrm{P}(|\xi_k| \leq \frac{b_k}{a_k})$ appearing in the infinite product is always $\leq 1$, hence it is well known that the infinite product converges to a number that is $> 0$ if and only if the series

$$\sum_{k=1}^{\infty} \Big(\mathrm{P}(|\xi_k| \leq \frac{b_k}{a_k}) - 1\Big)$$

is convergent. But $\mathrm{P}(|\xi_k| \leq \frac{b_k}{a_k}) = 2F(\frac{b_k}{a_k}) - 1$, with $F$ denoting the partition function of the standard Gaussian distribution. But recalling that, for large $x$,

$$1 - F(x) \geq \frac{1}{2\pi} \Big(x + \frac{1}{x}\Big)^{-1} \mathrm{e}^{-x^2/2} \geq \sqrt{\frac{2}{\pi}} \frac{1}{x} \mathrm{e}^{-x^2/2}$$

we have

$$1 - \mathrm{P}\Big(|\xi_k| \leq \frac{b_k}{a_k}\Big) \geq const \cdot \frac{1}{b_k \sqrt{\log(k+1)}} \frac{1}{(k+1)^{b_k^2/2}}$$

and this is the general term of a divergent series for every sequence $(b_k)_k$ such that $b_k \to_{k\to 0} 0$. Hence we have $\sup_{n\geq 1} P(X_n \in K^c) = 1$ and $(X_n)_n$ is not tight.

Let us turn now to the exponential tightness of this sequence. Let $K = \{x \in c_0, |x_k| \leq b_k\}$ be a compact set of $c_0$ and $R > 0$ as above. We must find a sequence $(b_k)_k$ such that, for every $n$,

$$P(g_n^{-1/2} X_n \in K^c) \leq e^{-Rg_n} \ .$$

Now

$$P(g_n^{-1/2} X_n \in K^c) = 1 - P(g_n^{-1/2} X_n \in K) = 1 - \prod_{k=1}^{n} P\Big(|\xi_k| \leq g_n^{1/2} \frac{b_k}{a_k}\Big)$$

$$= 1 - \exp\Big\{\sum_{k=1}^{n} \log P\Big(|\xi_k| \leq g_n^{1/2} \frac{b_k}{a_k}\Big)\Big\} \ .$$

As $1 - e^x \leq -x$,

$$\cdots \leq -\sum_{k=1}^{n} \log P\Big(|\xi_k| \leq g_n^{1/2} \frac{b_k}{a_k}\Big) = -\sum_{k=1}^{n} \log\Big(2F\Big(g_n^{1/2} \frac{b_k}{a_k}\Big) - 1\Big)$$

$$= -\sum_{k=1}^{n} \log\Big(1 - 2\Big(1 - F\Big(g_n^{1/2} \frac{b_k}{a_k}\Big)\Big)\Big).$$

As $\log(1-x) \geq -x - \frac{1}{2} x^2$,

$$\cdots \leq 2\sum_{k=1}^{n} \Big(1 - F\Big(g_n^{1/2} \frac{b_k}{a_k}\Big)\Big) + \frac{1}{2}\Big(2\sum_{k=1}^{n}\Big(1 - F\Big(g_n^{1/2} \frac{b_k}{a_k}\Big)\Big)\Big)^2 \ . \tag{5.2}$$

With the well known inequality $1 - F(x) \leq \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, $x \geq 0$,

$$\sum_{k=1}^{n} \Big(1 - F\Big(g_n^{1/2} \frac{b_k}{a_k}\Big)\Big) \leq \frac{1}{\sqrt{2\pi}} \sum_{k=1}^{n} e^{-g_n \frac{b_k^2}{2a_k^2}} \ .$$

If $g_n \geq M \log n$, $n \geq 2$, for some $M > 0$, let us choose $b_k^2 = 2(R + \frac{1}{M})a_k^2$. We have

$$\sum_{k=1}^{n} \Big(1 - F\Big(g_n^{1/2} \frac{b_k}{a_k}\Big)\Big) \leq \frac{1}{\sqrt{2\pi}} \frac{1}{n} \sum_{k=1}^{n} e^{-Rg_n} = \frac{1}{\sqrt{2\pi}} e^{-Rg_n}.$$

Going back to (5.2), we find finally

$$P(g_n^{-1/2} X_n \in K^c) \leq const \cdot e^{-Rg_n} \ .$$

so that $(X_n)_n$ is exponentially tight, for every speed function increasing to $+\infty$ faster than $n \mapsto M \log n$.

For this class of speed functions we can deduce that the sequence $(g_n^{-1/2} X_n)_n$ enjoys a Large Deviations Principle at speed $g$. Actually the dual of $c_0$ is the space of summable sequences $\ell_1$ and, for every $\theta \in \ell_1$ we

have,

$$\lim_{n\to\infty} \frac{1}{g_n} \log \mathrm{E}(\mathrm{e}^{g_n \langle \theta, g_n^{-1/2} X_n \rangle}) = \lim_{n\to\infty} \frac{1}{2} \mathrm{Var}(\langle \theta, X_n \rangle) = \frac{1}{2} \sum_{k=1}^{\infty} \theta_k^2 a_k^2 := \Lambda(\theta)$$

and we find easily for the convex conjugate, $I$, of $\Lambda$, *i.e.*

$$I(x) = \sup_{\theta \in \ell_1} \big( \langle \theta, x \rangle - \Lambda(\theta) \big) \ ,$$

that

$$I(x) = \frac{1}{2} \sum_{k=1}^{\infty} \frac{x_k^2}{a_k^2}$$

with the understanding $I(x) = +\infty$ if the sum is not convergent.

It is now easy to complete this example by showing that actually the sequence $(g_n^{-1/2} X_n)_n$ enjoys a Large Deviations Principle at speed $g$ and with respect to the rate function $I$. As the sequence $(a_k)_k$ tends to 0 at infinity, it is immediate that the level sets of $I$ are compact in $c_0$. Moreover Theorem 3.1 guarantees that the upper bound holds with respect to $I$ (this is Eq. (3.4)) and also the partial lower bound (3.5). In order to conclude we must only prove that the exposed points for $I$ are dense in $c_0$. Let us consider the set $\widetilde{\mathscr{H}} \subset c_0$ of the sequences $x = (x_k)_k$ that are $= 0$ for $k \geq N$ for some $N > 0$. Of course $\widetilde{\mathscr{H}}$ is dense in $c_0$. The elements of $\widetilde{\mathscr{H}}$ are exposed points. Actually, let $\eta \in \ell_1$ be given by

$$\eta_k = \frac{x_k}{a_k^2} \ , \qquad k \leq N$$

and $\eta_k = 0$ for $k > N$. Then if $x \in \widetilde{\mathscr{H}}$ we have for every $y \in c_0$ such that $I(y) < +\infty$, $y \neq x$,

$$\begin{aligned}
I(y) - I(x) - \langle \eta, y - x \rangle &= \frac{1}{2} \sum_{k=1}^{\infty} \frac{y_k^2}{a_k^2} - \frac{1}{2} \sum_{k=1}^{\infty} \frac{x_k^2}{a_k^2} - \sum_{k=1}^{\infty} \frac{x_k}{a_k^2} (y_k - x_k) \\
&= \frac{1}{2} \sum_{k=1}^{\infty} \frac{y_k^2}{a_k^2} + \frac{1}{2} \sum_{k=1}^{\infty} \frac{x_k^2}{a_k^2} - \sum_{k=1}^{\infty} \frac{y_k x_k}{a_k^2} = \frac{1}{2} \sum_{k=1}^{\infty} \frac{1}{a_k^2} (y_k - x_k)^2 > 0
\end{aligned}$$

hence $\eta$ is an exposing hyperplane and we are allowed to conclude.

## References

[1] R. Azencott, Grandes déviations et applications, Eighth Saint Flour Probability Summer School – 1978, Saint Flour, 1978. Vol. 774 of *Lecture Notes in Math.* Springer, Berlin (1980), 1–176.

[2] P. Baldi, Large deviations and stochastic homogenization. *Ann. Mat. Pura Appl.* **151** (1988) 161–177.

[3] P. Billingsley, Convergence of probability measures. John Wiley & Sons, Inc., New York-London-Sydney (1968).

[4] A. de Acosta, Upper bounds for large deviations of dependent random vectors. *Z. Wahrsch. Verw. Gebiete* **69** (1985) 551–565.

[5] L. Decreusefond and A. S. Üstünel, Stochastic analysis of the fractional Brownian motion. *Potential Anal.* **10** (1999) 177–214.

[6] A. Dembo and O. Zeitouni, Large deviations techniques and applications. Vol. 38 of *Stochastic Modelling and Applied Probability.* Springer-Verlag, Berlin (2010).

[7] J.-D. Deuschel and D.W. Stroock, Large deviations. Vol. 137 of *Pure and Applied Mathematics.* Academic Press, Inc., Boston, MA (1989).

[8] M.D. Donsker and S.R.S. Varadhan, Asymptotic evaluation of certain Markov process expectations for large time. III. *Commun. Pure Appl. Math.* **29** (1976) 389–461.

[9] R.S. Ellis, Large deviations for a general class of random vectors. *Ann. Probab.* **12** (1984) 1–12.

[10] X Fernique, Intégrabilité des vecteurs gaussiens. *C. R. Acad. Sci. Paris Sér. A-B* **270** (1970) A1698–A1699.

[11] X. Fernique, Regularité des trajectoires des fonctions aléatoires gaussiennes. École d'Été de Probabilités de Saint-Flour, IV-1974. Vol. 480 of *Lecture Notes in Math.* Springer, Berlin (1975) 1–96.

[12] J. Gärtner, On large deviations from an invariant measure. *Teor. Verojatnost. i Primenen.* **22** (1977) 27–42.

[13] M. Schilder, Some asymptotic formulas for Wiener integrals. *Trans. Am. Math. Soc.* **125** (1966) 63–85.