

Small deviations for a family of smooth Gaussian processes

Frank Aurzada* Fuchang Gao[†] Thomas Kühn[‡] Wenbo V. Li[§]
Qi-Man Shao[¶]

August 4, 2011

Abstract

We study the small deviation probabilities of a family of very smooth self-similar Gaussian processes. The canonical process from the family has the same scaling property as standard Brownian motion and plays an important role in the study of zeros of random polynomials.

Our estimates are based on the entropy method, discovered in Kuelbs and Li (1992) and developed further in Li and Linde (1999), Gao (2004), and Aurzada et al. (2009). While there are several ways to obtain the result w.r.t. the L_2 norm, the main contribution of this paper concerns the result w.r.t. the supremum norm. In this connection, we develop a tool that allows to translate upper estimates for the entropy of an operator mapping into $L_2[0, 1]$ by those of the operator mapping into $C[0, 1]$, if the image of the operator is in fact a Hölder space.

The results are further applied to the entropy of function classes, generalizing results of Gao et al. (2010).

*Institut für Mathematik, Technische Universität Berlin, Germany, aurzada@math.tu-berlin.de, supported by the DFG Emmy Noether program.

[†]Department of Mathematics, University of Idaho, fuchang@uidaho.edu, supported in part by NSF grant DMS-0806211.

[‡]Mathematisches Institut, Universität Leipzig, Germany, kuehn@math.uni-leipzig.de, supported in part by Ministerio de Ciencia e Innovación (Spain), grant MTM2010-15814.

[§]Department of Mathematical Sciences, University of Delaware, wli@math.udel.edu, supported in part by NSF grant DMS-0805929, NSFC-6398100, CAS-2008DP173182.

[¶]Department of Mathematics, Hong Kong University of Science and Technology, Hong Kong, maqmshao@ust.hk, supported in part by Hong Kong RGC CERG 602608.

1 Introduction and main results

1.1 Introduction

The small deviation problem for a stochastic process $X = (X(t))_{t \geq 0}$ – also called small ball or small value problem – consists in determining the probability

$$-\log \mathbb{P}(\|X\| \leq \varepsilon), \quad \text{as } \varepsilon \rightarrow 0,$$

where $\|\cdot\|$ is for example the norm in some $L_p[0, 1]$ or in $C[0, 1]$. Small deviation probabilities play a fundamental role in many problems in probability and analysis, see the lecture notes [19] for details. This is why there has been a lot of interest in small deviation problems in recent years, cf. the survey [21] and the literature compilation [24]. There are many connections to other questions such as the law of the iterated logarithm of Chung type, strong limit laws in statistics, metric entropy properties of linear operators, quantization, and several other approximation quantities for stochastic processes. For Gaussian processes, a reasonable amount of theory has been developed up to date, see e.g. [21].

In this paper, we study a family of very smooth self-similar Gaussian processes and their respective small deviation probabilities. This family of processes does not seem to have drawn enough attention in the probability community. Yet, it is very important and appears naturally in physics [25].

Let us give a short motivation for studying this class of processes. Let $\phi(\mathbf{x}, t)$ be a scalar field in a d -dimensional space that evolves according to the deterministic diffusion equation

$$\frac{\partial}{\partial t} \phi(\mathbf{x}, t) = \nabla^2 \phi(\mathbf{x}, t)$$

with the initial condition $\phi(\mathbf{x}, 0) = \psi(\mathbf{x})$. The only randomness is the initial condition which could be a mean-zero Gaussian random field. For example, $\phi(\mathbf{x}, t)$ could represent the density fluctuation of a diffusing gas. In such a case, it is reasonable to assume that $\psi(\mathbf{x})$ is of zero mean and delta covariance $\mathbb{E} \psi(\mathbf{x}) \psi(\mathbf{y}) = \delta(\mathbf{x} - \mathbf{y})$. For a system of linear size L , the solution of the diffusion equation in the bulk of the system is

$$\phi(\mathbf{x}, t) = (4\pi t)^{-d/2} \int_{|\mathbf{y}| \leq L} \exp(-\|\mathbf{x} - \mathbf{y}\|_2^2 / 4t) \psi(\mathbf{y}) \, d\mathbf{y}.$$

Because of the linearity of the integral, $\phi(\mathbf{x}, t)$, $t \geq 0$, is a Gaussian process. It is customary to study the normalized process $X(t) = \phi(\mathbf{x}, t) / (\mathbb{E} \phi(\mathbf{x}, t)^2)^{1/2}$. When $L = \infty$, it is straightforward to calculate that the Gaussian process $X(t)$ has covariance structure

$$\mathbb{E} X(t) X(s) = \left(\frac{2ts}{(t+s)^2} \right)^{d/4}.$$

In modeling non-Fickian diffusion, it is common that the Green function

$$G(\mathbf{x}, t) = (4\pi t)^{-d/2} \exp(-\|\mathbf{x}\|_2^2 / 4t)$$

above needs to be replaced by a different Green function with a different scaling property, say $G_{p,q}(\mathbf{x}, t) = Ct^{-p}K(\mathbf{x}/t^q)$.

This naturally leads us to consider a centered Gaussian processes $X_{\alpha,\beta}(t)$, $t > 0$, defined by the covariance function

$$K(t, s) = \mathbb{E} X_{\alpha,\beta}(t)X_{\alpha,\beta}(s) = \frac{2^{2\beta+1}(ts)^\alpha}{(t+s)^{2\beta+1}}, \quad t, s > 0, \quad (1)$$

for $\alpha > 0$ and $\beta > -1/2$. Note that for $\alpha > \beta + 1/2$, we can define $X_{\alpha,\beta}(0) = 0$. It is also easy to check that $X_{\alpha,\beta}$ is an $(\alpha - \beta - 1/2)$ -self-similar process, i.e. $(X_{\alpha,\beta}(ct))$ has the same law as $(c^{\alpha-\beta-1/2}X_{\alpha,\beta}(t))$ for any $c > 0$. In particular, $X_{\alpha,\beta}$ has the same scaling property as Brownian motion for $\alpha - \beta = 1$. A useful stochastic integral representation is

$$X_{\alpha,\beta}(t) = \sqrt{\frac{2^{2\beta+1}}{\Gamma(2\beta+1)}} t^\alpha \int_0^\infty x^\beta e^{-xt} dB(x), \quad t > 0, \quad (2)$$

where B is a standard Brownian motion.

If $\beta = 0$, it is easy to see, using integration by parts, that $X_{\alpha,0}$ has the same law as the process

$$\tilde{X}_{\alpha,0}(t) := \sqrt{2} t^{1+\alpha} \int_0^\infty e^{-xt} B(x) dx, \quad t > 0.$$

We also introduce, for $\alpha = 1$, $\beta = 0$, the canonical process

$$X(t) := X_{1,0}(t) = \sqrt{2} t \int_0^\infty e^{-xt} dB(x), \quad t \geq 0.$$

This process has the same scaling property as standard Brownian motion and plays an important role in the study of zeros of random polynomials, cf. [22, 23, 2], which gives another reason for studying this process. Yet a different motivation is the applicability of small deviation estimates of stationary processes in statistics, cf. e.g. [28]; we come back to this point in Section 1.3 below.

This paper is structured as follows. In Section 1.2, we present our main results concerning the small deviation rate w.r.t. $L_2[0, 1]$ and supremum norm. We compare the present findings to other known facts for smooth Gaussian processes in Section 1.3. The main tool for the proofs is the so-called entropy method, which is recalled in Section 1.4. We also present the functional analytic counterparts of our main theorems in that section.

There are different ways to obtain the small deviation rate for the L_2 -norm for our class of processes. Therefore, the main contribution of this paper is the transfer from L_2 to supremum norm estimates. One interesting tool in this connection is presented in Section 1.5, and we believe that it may be of independent interest.

The proofs of all statements are given in Section 2. Finally, our main theorems entail some new results for the entropy of function classes. This connection and the respective corollaries are presented in Section 3.

We remark that further properties of the class of processes $(X_{\alpha,\beta})$ and alternative proofs are contained in an earlier version of this paper available from <http://arxiv.org/abs/1009.5580>.

Let us fix some notation. We write $f \preceq g$ or $g \succeq f$ if $\limsup f/g < \infty$, and the asymptotic equivalence $f \asymp g$ means that we have both $f \preceq g$ and $g \preceq f$. Moreover, we write $f \lesssim g$ or $g \gtrsim f$, if $\limsup f/g \leq 1$. Finally, the strong equivalence $f \sim g$ means that $\lim f/g = 1$.

Acknowledgement: We are grateful to the American Institute of Mathematics (AIM) for supporting the workshop “Small ball inequalities in analysis, probability, and irregularities of distribution” (December 2008), where the work on this paper was started. We would like to thank two anonymous referees and an associate editor for their suggestions, that helped to improve the exposition of the paper.

1.2 Main results

We can clarify the small deviation order for the whole class of smooth processes introduced above for the L_2 -norm and the sup-norm.

For the L_2 -norm, there is the following result.

Theorem 1 *Let $\alpha > \beta > -1/2$. Let $X_{\alpha,\beta}$ be the process defined by (1). Then*

$$-\log \mathbb{P} \left(\int_0^1 |X_{\alpha,\beta}(t)|^2 dt \leq \varepsilon^2 \right) \sim \kappa_{\alpha,\beta} |\log \varepsilon|^3,$$

where the constant is given by

$$\kappa_{\alpha,\beta} := \frac{1}{3(\alpha - \beta)\pi^2}. \quad (3)$$

We remark that Theorem 1 follows from Proposition 4.3 in [13], which we got to know only after the submission of the present paper. We will outline their method of proof and the relation to our approach below.

For the sup-norm, we obtain the following theorem under optimal assumptions on the parameters α and β . However, the result is less precise with respect to the small deviation constant compared to the L_2 -norm result.

Theorem 2 *Let $\alpha > \beta + 1/2 > 0$. Let $X_{\alpha,\beta}$ be the process defined by (1) with $X_{\alpha,\beta}(0) = 0$. Then, with some constant $\tilde{\kappa}_{\alpha,\beta} > 0$, we have*

$$\tilde{\kappa}_{\alpha,\beta} |\log \varepsilon|^3 \lesssim -\log \mathbb{P} \left(\sup_{t \in [0,1]} |X_{\alpha,\beta}(t)| \leq \varepsilon \right) \lesssim \kappa_{\alpha-1/2,\beta} |\log \varepsilon|^3, \quad (4)$$

where $\kappa_{\alpha,\beta}$ is defined in (3) and $\tilde{\kappa}_{\alpha,\beta} > 0$. Further, $\tilde{\kappa}_{\alpha,\beta} \rightarrow \infty$ when $\alpha - \beta \rightarrow 1/2$.

The proofs of these two theorems use the connection between small deviations of Gaussian processes and the entropy numbers of a linear operator generating the process, cf. [16], [20], [3]. In fact, due to Corollaries 2.2 and 2.4 in [3], Theorem 1 and Theorem 2 are equivalent to Theorem 3 and Theorem 4, respectively, given in Section 1.4. Other interesting small deviation estimates for smooth Gaussian processes can be found in [3], [17], and [13], and we compare them to the present results in the next section.

1.3 Stationary version of our process

Let us shortly comment on a relation to [3] and [13], where, among other things, small deviation probabilities of stationary Gaussian processes are considered.

Our process $X_{\alpha,\beta}$ is $(\alpha - \beta - 1/2)$ -self-similar. We can thus consider the Lamperti transformed stationary Gaussian process: $Y_\beta(t) := e^{-(\alpha-\beta-1/2)t}X_{\alpha,\beta}(e^t)$, $t \in \mathbb{R}$, which has the correlation function

$$k(t) = \frac{2^{2\beta+1}}{(e^{t/2} + e^{-t/2})^{2\beta+1}} = \frac{1}{\cosh(t/2)^{2\beta+1}}, \quad t \geq 0.$$

One can find that for the spectral measure we have (for all $\beta > -1/2$)

$$k(t) =: \int_{-\infty}^{\infty} e^{itx} e^{-G(x)} dx, \quad \text{with} \quad G(x) \sim \pi x, \quad x \rightarrow \infty. \quad (5)$$

Therefore, by Proposition 4.2(3) in [13], for some numeric constant $c > 0$,

$$-\log \mathbb{P} \left(\int_0^1 |Y_\beta(t)|^2 dt \leq \varepsilon^2 \right) \sim c |\log \varepsilon|^2, \quad (6)$$

as for the process in [3] with $G(x) = |x|$. In comparison, note that our main theorems concern the infinite time horizon for the stationary process: Namely, for the L_2 norm, Theorem 1 reads in terms of Y_β :

$$-\log \mathbb{P} \left(\int_{-\infty}^0 |Y_\beta(t)|^2 e^{2(\alpha-\beta)t} dt \leq \varepsilon^2 \right) \sim \kappa_{\alpha,\beta} |\log \varepsilon|^3,$$

which makes sense since $\alpha - \beta > 0$.

In the same spirit, we can compare the results w.r.t. the sup-norm. Using (5) with Lemma 2.3 in [28] and (6) one obtains for the finite time horizon

$$-\log \mathbb{P} \left(\sup_{0 \leq t \leq 1} |Y_\beta(t)| \leq \varepsilon \right) \asymp |\log \varepsilon|^2,$$

while our main result for the sup-norm, Theorem 2, is again for the infinite time horizon and reads

$$-\log \mathbb{P} \left(\sup_{-\infty < t \leq 0} |Y_\beta(t) e^{(\alpha-\beta-1/2)t}| \leq \varepsilon \right) \asymp |\log \varepsilon|^3,$$

which makes sense since $\alpha - \beta - 1/2 > 0$.

1.4 The entropy method

In this section, we recall the entropy method for obtaining small deviation probabilities and formulate the respective counterparts for Theorems 1 and 2.

In order to state and use the connection to the entropy numbers, let us first define the entropy numbers. For a linear operator $u : E \rightarrow F$ between Banach spaces E and F and $n \in \mathbb{N}$, the entropy numbers are defined as follows:

$$e_n(u : E \rightarrow F) := \inf \left\{ \varepsilon > 0 : \exists f_1, \dots, f_{2^{n-1}} \in F \text{ s.t. } u(B_E) \subseteq \bigcup_{k=1}^{2^{n-1}} (f_k + B_F) \right\},$$

where B_E and B_F denote the closed unit balls in E and F , respectively. For elementary properties and further information see e.g. [7]. Since u is compact if and only if $\lim_{n \rightarrow \infty} e_n(u) = 0$, the decay rate of the entropy numbers is a measure for the “degree of compactness” of u .

It turns out that there is a close relation between the small deviation problem for a Gaussian process X attaining values in E and the entropy numbers of a compact operator $u : L_2(S) \rightarrow E$ related to X through its characteristic function:

$$\mathbb{E} e^{i\langle X, h \rangle} = \exp \left(-\frac{1}{2} \|u'(h)\|_{L_2(S)}^2 \right), \quad h \in E', \quad (7)$$

where $u' : E' \rightarrow L_2(S)$ is the dual operator and $(S, \mathcal{S}, \lambda)$ is some measure space. Here, $(E, \|\cdot\|)$ is some Banach space.

It can be checked easily that, up to an unimportant multiplicative constant, the process $X_{\alpha, \beta}$ defined in (1) is related – via (7) – to the operator

$$(uf)(t) = t^\alpha \int_0^\infty x^\beta e^{-xt} f(x) dx, \quad f \in L_2[0, \infty). \quad (8)$$

Note that the process defined in (1) can be considered with values in the Banach spaces $E = L_2[0, 1]$ for $\alpha > \beta > -1/2$. In this L_2 -setting, we obtain the precise behavior of the entropy numbers of the operator related to our process on the exponential scale.

Theorem 3 *Let $\alpha > \beta > -1/2$. Let $u : L_2[0, \infty) \rightarrow L_2[0, 1]$ be the operator given by (8). Then*

$$-\log e_n(u) \sim d_{\alpha, \beta} n^{1/3},$$

where

$$d_{\alpha, \beta} := (3(\alpha - \beta)\pi^2 \log 2)^{1/3}. \quad (9)$$

In the sup-norm case we get the following result under the optimal assumption on the parameters $\alpha - \beta > 1/2$, which is exactly when the process defined in (1) is almost surely in $E = C[0, 1]$.

Theorem 4 *Let $\alpha > \beta + 1/2 > 0$. Let $u : L_2[0, \infty) \rightarrow C[0, 1]$ be the operator given by (8). Then*

$$\tilde{d}_{\alpha, \beta} n^{1/3} \lesssim -\log e_n(u) \lesssim d_{\alpha-1/2, \beta} n^{1/3},$$

where $d_{\alpha, \beta}$ is defined in (9) and

$$\tilde{d}_{\alpha, \beta} := \frac{\min(\alpha - \beta - 1/2, 1)}{1/2 + \min(\alpha - \beta - 1/2, 1)} \cdot d_{\alpha, \beta}.$$

Note, in particular, that if $\alpha - \beta \rightarrow 1/2$, then

$$\tilde{d}_{\alpha, \beta} \rightarrow 0 \quad \text{and} \quad d_{\alpha-1/2, \beta} \rightarrow 0.$$

1.5 A lemma relating entropy in L_2 and L_∞ via Hölder continuity

In this section, we present an interpolation result which will be an essential tool in the proof of Theorem 4, and which might be of independent interest.

Roughly speaking, it provides a technique for deriving upper entropy estimates for an operator $u : E \rightarrow C[0, 1]$, where E is a Banach space, from entropy estimates of $u : E \rightarrow L_2[0, 1]$, i.e. the same operator, but considered as operator into a larger target space. The additional information we need for this argument is that u should map E even into a smaller space than $C[0, 1]$, namely into a Hölder space $C_\lambda[0, 1]$ for some $0 < \lambda \leq 1$. This space consists of all functions $f \in C[0, 1]$ such

$$\|f\|_{C_\lambda[0,1]} := \sup_{0 \leq s < t \leq 1} \frac{|f(t) - f(s)|}{|t - s|^\lambda} + \sup_{0 \leq t \leq 1} |f(t)| < \infty.$$

Moreover, $C_\lambda[0, 1]$ is a Banach space under the norm $\|\cdot\|_{C_\lambda[0,1]}$.

Theorem 5 *Let u be an operator from a Banach space E into a Hölder space $C_\lambda[0, 1]$ for some $0 < \lambda \leq 1$. Then we have for all $k \in \mathbb{N}$*

$$e_k(u : E \rightarrow C[0, 1]) \leq 2 \|u : E \rightarrow C_\lambda[0, 1]\|^{\frac{1}{2\lambda+1}} e_k(u : E \rightarrow L_2[0, 1])^{\frac{2\lambda}{2\lambda+1}}.$$

2 Proofs of the theorems

2.1 Proof of Theorems 1 and 3

There are many ways to obtain the small deviation estimates w.r.t. the L_2 norm. Even though Theorem 1 and Theorem 3 are equivalent due to Corollaries 2.2 and 2.4 in [3], we will give proofs for both of the theorems: On the one hand, the proof of Theorem 1 is simpler, on the other hand, for consistency with the $C[0, 1]$ case it seems reasonable to have a purely analytic proof of Theorem 3. Additionally, we will explicitly use Theorem 3 during the proof of Theorem 4, both for upper and lower bound.

We remark that the proof of Theorem 1 is as outlined in Proposition 4.3 in [13], which we got to know only after the submission of the present paper. However, the relation to the proof of our Theorem 3 is rather close so that it seems worthwhile to include it here. In particular, both proofs are based on Laptev's result [18].

Proof of Theorem 1. First note that the operator $uu^* : L_2[0, 1] \rightarrow L_2[0, 1]$, where u^* denotes the adjoint operator of u , is given by

$$(uu^*g)(t) = \Gamma(2\beta + 1) t^\alpha \int_0^1 \frac{x^\alpha}{(t+x)^{2\beta+1}} g(x) dx.$$

The exact asymptotic behavior of its singular numbers $s_n(uu^*)$ was found by Laptev [18] (based on Widom [29]). He showed that

$$-\log s_n(uu^*) \sim 2\pi(\alpha - \beta)^{1/2} n^{1/2}. \quad (10)$$

Similar operators were studied in [29], [14], and [4]. Since $s_n(uu^*) = s_n(u)^2$, this implies

$$-\log s_n(u) \sim \pi(\alpha - \beta)^{1/2} n^{1/2}. \quad (11)$$

It is well-known that, due to the Karhunen-Loève expansion, it suffices to study the small deviation behavior of the weighted sum

$$-\log \mathbb{P} \left(\sum_{n=1}^{\infty} s_n(u) \xi_n^2 \leq \varepsilon^2 \right),$$

which can be treated with Theorem 2 in [26] (cf. Proposition 4.1 in [13]) or Theorem 4.1 in [6]. \square

As explained above, we give an alternative proof in order to provide a fully analytic line of arguments for Theorem 3.

Proof of Theorem 3. Recall that we determined the behavior of the singular numbers $s_n(u)$ in (11). Since we are in the Hilbert space setting, the singular numbers and the entropy numbers of u coincide with those of the diagonal operator D_σ in ℓ_2 , $(x_n) \mapsto (\sigma_n x_n)$, with weight sequence $\sigma_n = s_n(u)$. A result of Gordon, König, and Schütt (Proposition 1.7 in [12]) says that

$$\sup_{n \geq 1} (2^{-k/n} (\sigma_1 \cdots \sigma_n)^{1/n}) \leq e_{k+1}(D_\sigma) \leq 6 \sup_{n \geq 1} (2^{-k/n} (\sigma_1 \cdots \sigma_n)^{1/n}).$$

This implies that

$$\log e_k(u) \sim \sup_{n \geq 1} \left(-\frac{k}{n} \log 2 + \frac{1}{n} \sum_{j=1}^n \log s_j(u) \right).$$

Using the asymptotics of $(s_n(u))$ from (11), we get

$$-\log e_k(u) \sim \inf_{n \geq 1} \left(\frac{k}{n} \log 2 + \frac{\pi(\alpha - \beta)^{1/2}}{n} \sum_{j=1}^n j^{1/2} \right).$$

Evaluating this expression we obtain the assertion. \square

2.2 Proofs of Theorems 4 and 5

Proof of Theorem 5. Clearly, since $C_\lambda[0, 1]$ is compactly embedded in $L_2[0, 1]$, the operator $u : E \rightarrow L_2[0, 1]$ is compact, whence its entropy numbers tend to zero.

For $0 < \delta \leq 1$ and $t \in [0, 1]$ we consider the interval $I_\delta(t) := [0, 1] \cap [t - \delta, t + \delta]$ and define for $g \in L_2[0, 1]$, the local averaging operator

$$(P_\delta g)(t) := \frac{1}{|I_\delta(t)|} \int_{I_\delta(t)} g(x) \, dx.$$

One can easily verify that the averaging operators P_δ map $L_2[0, 1]$ in $C[0, 1]$.

Step 1: We need the following simple norm estimates:

$$\|P_\delta : L_2[0, 1] \rightarrow C[0, 1]\| \leq \delta^{-1/2}; \quad (12)$$

$$\|\text{id} - P_\delta : C_\lambda[0, 1] \rightarrow C[0, 1]\| \leq \delta^\lambda. \quad (13)$$

For all $t \in [0, 1]$ and $g \in L_2[0, 1]$ we have, by the Cauchy-Schwarz inequality,

$$|(P_\delta g)(t)| \leq \frac{1}{|I_\delta(t)|} \int_{I_\delta(t)} |g(x)| \, dx \leq \frac{1}{|I_\delta(t)|} |I_\delta(t)|^{1/2} \|g\|_{L_2[0,1]} \leq \delta^{-1/2} \|g\|_{L_2[0,1]},$$

which implies (12).

Now let $t \in [0, 1]$ and $g \in C_\lambda[0, 1]$ with $\|g\|_{C_\lambda[0,1]} \leq 1$. Then

$$|g(t) - P_\delta g(t)| = \frac{1}{|I_\delta(t)|} \left| \int_{I_\delta(t)} (g(t) - g(x)) \, dx \right| \leq \frac{1}{|I_\delta(t)|} \int_{I_\delta(t)} |g(t) - g(x)| \, dx \leq \delta^\lambda,$$

since $|g(t) - g(x)| \leq |t - x|^\lambda \|g\|_{C_\lambda[0,1]} \leq \delta^\lambda$ for every $x \in I_\delta(t)$ and $\frac{1}{|I_\delta(t)|} \int_{I_\delta(t)} dx = 1$. To finish the proof of (13), we take the supremum over all $t \in [0, 1]$ and $g \in C_\lambda[0, 1]$ with $\|g\|_{C_\lambda[0,1]} \leq 1$.

Step 2: Using elementary properties of entropy numbers (cf. [7]) and the above norm estimates (12) and (13), we obtain for all $k \in \mathbb{N}$ and $0 < \delta \leq 1$

$$\begin{aligned} e_k(u : E \rightarrow C[0, 1]) &\leq e_k(P_\delta u : E \rightarrow C[0, 1]) + \|u - P_\delta u : E \rightarrow C[0, 1]\| \\ &\leq e_k(u : E \rightarrow L_2[0, 1]) \cdot \|P_\delta : L_2[0, 1] \rightarrow C[0, 1]\| \\ &\quad + \|u : E \rightarrow C_\lambda[0, 1]\| \cdot \|\text{id} - P_\delta : C_\lambda[0, 1] \rightarrow C[0, 1]\| \\ &\leq e_k(u : E \rightarrow L_2[0, 1]) \cdot \delta^{-1/2} + \|u : E \rightarrow C_\lambda[0, 1]\| \cdot \delta^\lambda. \end{aligned}$$

Finally we choose δ such that $e_k(u : E \rightarrow L_2[0, 1]) \cdot \delta^{-1/2} = \|u : E \rightarrow C_\lambda[0, 1]\| \cdot \delta^\lambda$, i.e.

$$\delta = \left(\frac{e_k(u : E \rightarrow L_2[0, 1])}{\|u : E \rightarrow C_\lambda[0, 1]\|} \right)^{\frac{1}{\lambda+1/2}}.$$

Clearly $0 < \delta \leq 1$, and consequently this choice of δ gives the desired inequality

$$e_k(u : E \rightarrow C[0, 1]) \leq 2 \|u : E \rightarrow C_\lambda[0, 1]\|^{\frac{1}{2\lambda+1}} e_k(u : E \rightarrow L_2[0, 1])^{\frac{2\lambda}{2\lambda+1}}.$$

□

Proof of the upper bound for the entropy numbers in Theorem 4.

Step 1: First we show that u is a bounded operator from $L_2[0, \infty)$ into the Hölder space $C_\lambda[0, 1]$, where

$$\lambda := \min(\alpha - \beta - 1/2, 1).$$

Let $f \in L_2[0, \infty)$, let $0 \leq s < t \leq 1$, and set $h := t - s$. We consider

$$|(uf)(t) - (uf)(s)| \leq \int_0^\infty |t^\alpha x^\beta e^{-xt} - s^\alpha x^\beta e^{-xs}| \cdot |f(x)| \, dx.$$

Using the Cauchy-Schwartz inequality, this can be estimated by $A^{1/2} \|f\|_{L_2[0,\infty)}$, with

$$A := \int_0^\infty |t^\alpha x^\beta e^{-xt} - s^\alpha x^\beta e^{-xs}|^2 dx.$$

We have to show that $A \leq Ch^{2\lambda}$ for some constant C independent of t, s , since then one can take the supremum over all $0 \leq s < t \leq 1$ and finally over all f . Define

$$g(y) := \int_0^\infty x^{2\beta} e^{-2yx} dx = y^{-2\beta-1} 2^{-2\beta-1} \Gamma(2\beta+1)$$

and note that

$$\begin{aligned} A &= \int_0^\infty x^{2\beta} (t^{2\alpha} e^{-2xt} - 2(ts)^\alpha e^{-x(t+s)} + s^{2\alpha} e^{-2xs}) dx \\ &= t^{2\alpha} g(t) - 2(ts)^\alpha g((t+s)/2) + s^{2\alpha} g(s). \end{aligned}$$

Therefore, using the notation $\gamma := \alpha - \beta - 1/2$, we have

$$\begin{aligned} \frac{2^{2\beta+1} A}{\Gamma(2\beta+1)} &= t^{2\alpha-2\beta-1} - 2(ts)^\alpha \left(\frac{t+s}{2}\right)^{-2\beta-1} + s^{2\alpha-2\beta-1} \\ &= \left[t^{2\gamma} - 2 \left(\frac{t+s}{2}\right)^{2\gamma} + s^{2\gamma} \right] + 2 \left(\frac{t+s}{2}\right)^{-2\beta-1} \left[\left(\frac{t+s}{2}\right)^{2\alpha} - (ts)^\alpha \right]. \end{aligned} \quad (14)$$

Case 1, $h \geq s$: In this case we have $t = s + h \leq 2s$ and $s \leq h$, and therefore (14) implies

$$\frac{2^{2\beta+1} A}{\Gamma(2\beta+1)} \leq t^{2\gamma} + s^{2\gamma} \leq (2h)^{2\gamma} + h^{2\gamma} = Ch^{2\gamma}.$$

Case 2, $h \leq s$: We estimate the two terms in the second line of (14) separately. For the first term we use the Taylor expansion of the function $f(x) = x^{2\gamma}$ at $x_0 = (t+s)/2$ and obtain, with some $\xi \in (s, t)$,

$$\begin{aligned} t^{2\gamma} - 2 \left(\frac{t+s}{2}\right)^{2\gamma} + s^{2\gamma} &= f(x_0 + h/2) - 2f(x_0) + f(x_0 - h/2) \\ &= \frac{h^2}{4} f''(\xi) = h^2 \gamma(\gamma-1/2) \xi^{2\gamma-2} \leq C_1 h^2 s^{2\gamma-2}, \end{aligned} \quad (15)$$

where we used $s \leq \xi \leq t = s + h \leq 2s$ in the last inequality.

For the second term in (14) we apply the mean value theorem. Set $a := ts$ and $b := \left(\frac{t+s}{2}\right)^2$, and note that $a \leq b$ and $b - a = \frac{1}{4}(t^2 + 2ts + s^2 - 4ts) = (t-s)^2/4 = h^2/4$. We obtain, with some $\eta \in (a, b)$,

$$\left(\frac{t+s}{2}\right)^{2\alpha} - (ts)^\alpha = (b-a) \cdot \alpha \eta^{\alpha-1} \leq C_2 h^2 s^{2(\alpha-1)}. \quad (16)$$

Here we used that $s + t = 2s + h \leq 3s$, and therefore $s^2 \leq ts \leq \eta \leq (t+s)^2/4 \leq 9s^2/4$.

Combining now (14) with (15) and (16) we get

$$\frac{2^{2\beta+1}A}{\Gamma(2\beta+1)} \leq C_1 h^2 s^{2\gamma-2} + 2C_2 s^{-2\beta-1} h^2 s^{2\alpha-2} = Ch^2 s^{2\gamma-2}.$$

If $\gamma \geq 1$, then $h^2 s^{2\gamma-2} \leq h^2$; and if $\gamma < 1$, we have $h^2 s^{2\gamma-2} = (h/s)^{2-2\gamma} h^{2\gamma} \leq h^{2\gamma}$, since we are in the case $h \leq s$. This proves, for all $0 < h \leq 1$, the desired estimate

$$\frac{2^{2\beta+1}A}{\Gamma(2\beta+1)} \leq Ch^{2\min(\gamma,1)} = Ch^{2\lambda}.$$

Step 2: Let $\varepsilon > 0$. Then, by Theorem 3, for large enough k ,

$$e_k(u : L_2[0, \infty) \rightarrow L_2[0, 1]) \leq e^{-(d_{\alpha,\beta}-\varepsilon)k^{1/3}}.$$

Now Theorem 5 implies that

$$e_k(u : L_2[0, \infty) \rightarrow C[0, 1]) \lesssim \exp\left(-\frac{2\lambda}{2\lambda+1}(d_{\alpha,\beta}-\varepsilon)k^{1/3}\right),$$

and consequently

$$-\log e_k(u : L_2[0, \infty) \rightarrow C[0, 1]) \gtrsim \frac{2\lambda}{2\lambda+1}(d_{\alpha,\beta}-\varepsilon)k^{1/3}.$$

In other words, we have

$$\liminf_{k \rightarrow \infty} \frac{-\log e_k(u : L_2[0, \infty) \rightarrow C[0, 1])}{k^{1/3}} \geq \frac{2\lambda}{2\lambda+1}(d_{\alpha,\beta}-\varepsilon).$$

Letting $\varepsilon \rightarrow 0$, we arrive at the desired upper bound for the entropy numbers. \square

Proof of the lower bound for the entropy numbers in Theorem 4. Obviously, $e_k(u : L_2[0, \infty) \rightarrow C[0, 1]) \geq e_k(u : L_2[0, \infty) \rightarrow L_2[0, 1])$. However, we can even gain a bit concerning the constant. For this purpose, let us stress the dependence on α and β in the definition (8) by denoting the operator $u_{\alpha,\beta}$. Further, for some fixed $\varepsilon > 0$, we let $v : C[0, 1] \rightarrow L_2[0, 1]$ denote the multiplication operator $(vf)(t) = t^{-1/2+\varepsilon}f(t)$. Note that $v : C[0, 1] \rightarrow L_2[0, 1]$ is bounded. Then one can observe that $vu_{\alpha,\beta} = u_{\alpha-1/2+\varepsilon}$. Therefore,

$$e_k(u_{\alpha-1/2+\varepsilon} : L_2[0, \infty) \rightarrow L_2[0, 1]) \leq e_k(u_{\alpha,\beta} : L_2[0, \infty) \rightarrow C[0, 1]) \cdot \|v : C[0, 1] \rightarrow L_2[0, 1]\|.$$

Using the L_2 estimate from Theorem 3 for the left-hand side, this shows

$$-\log e_k(u_{\alpha,\beta} : L_2[0, \infty) \rightarrow C[0, 1]) \lesssim d_{\alpha-1/2+\varepsilon,\beta}k^{1/3},$$

or in other words,

$$\limsup_{k \rightarrow \infty} \frac{-\log e_k(u_{\alpha,\beta} : L_2[0, \infty) \rightarrow C[0, 1])}{k^{1/3}} \leq d_{\alpha-1/2+\varepsilon,\beta}.$$

This holds for all $\varepsilon > 0$. Letting ε tend to zero yields the lower bound for the entropy numbers, since the constant $d_{\alpha,\beta}$ (defined in (9)) is continuous in the parameters. \square

3 Relation to the entropy of function classes

In this section, we relate the small deviation problem for $X_{\alpha,\beta}$ under the sup-norm to another small deviation problem, which in turn is related to a metric entropy problem of a certain function class. The function class related to the canonical case $\alpha = 1, \beta = 0$ is studied in Theorem 1.2 in [10]. The present proof not only generalizes the case $\alpha = 1, \beta = 0$ but also gives a simplified proof for that specific case.

Let us define the process

$$S(t) := t^{\alpha'} \int_0^1 x^\beta e^{-xt} dB(x), \quad t \geq 1,$$

where $\alpha' := 2\beta + 1 - \alpha$ and the natural restrictions are $\alpha > \beta + 1/2 > 0$. Note that exactly under these restrictions, S is bounded on $[1, \infty]$. Our main theorem concerning S is as follows.

Theorem 6 *Let $\alpha > \beta + 1/2 > 0$. Then*

$$-\log \mathbb{P} \left(\sup_{t \geq 1} |S(t)| \leq \varepsilon \right) \asymp |\log \varepsilon|^3. \quad (17)$$

Using the technique in [10], one finds that the associated convex hull for the process $S(t)$, $t \geq 1$, is the function class \mathcal{F} consisting of all the functions f on $[0, 1]$ corresponding to the kernel $K(t, x) = t^{\alpha'} x^\beta e^{-tx}$. More precisely, \mathcal{F} can be expressed as

$$\mathcal{F} := \left\{ f : [0, 1] \rightarrow \mathbb{R} \mid f(x) = x^\beta \int_1^\infty t^{\alpha'} e^{-tx} \mu(dt) : \|\mu\|_{TV} \leq 1 \right\}.$$

Under the $L_2[0, 1]$ norm $\|f\|_2 = (\int_0^1 f^2(x) dx)^{1/2}$, the class \mathcal{F} is compact and its metric entropy is denoted by $\log N(\varepsilon, \mathcal{F}, \|\cdot\|_2)$ where $N(\varepsilon, \mathcal{F}, \|\cdot\|_2)$ is the minimum number of ε -radii balls in the norm $\|\cdot\|_2$ to cover the class \mathcal{F} . Thus, as discussed in detail in [10], via the connection between the small deviation probability and the metric entropy, we obtain the following statement for the function class \mathcal{F} associated with S :

Corollary 1 *For the class \mathcal{F} defined above, and $\alpha' = 2\beta + 1 - \alpha$ with $\alpha > \beta + 1/2 > 0$,*

$$\log N(\varepsilon, \mathcal{F}, \|\cdot\|_2) \asymp |\log \varepsilon|^3. \quad (18)$$

The original proof (as conducted in [10] for the case $\alpha = 1, \beta = 0$) of the lower bounds for the estimates of the probability in (17) and (4) follows from covering estimates for the upper bound in (18) which is lengthy and unpleasant. The current approach for this part, which is turned around, is based on the simple and soft arguments summarized in Theorem 4.

However, the argument used for this – that the upper bound of the metric entropy implies the lower bound of the small deviation probability, as discussed in Section 1.4 – is as in many other instances we know before: The key point is that it seems easier to find an upper bound of the metric entropy via analytic tools than a lower bound of the small deviation probability

via probabilistic tools, even though they are equivalent. It would be interesting to find a probabilistic proof for the probability lower bound in (17) or (4) for all parameters in the range $\alpha > \beta + 1/2 > 0$.

Proof of Theorem 6. The class of processes that we consider satisfies the following time inversion property: $(X_{\alpha,\beta}(1/\cdot))$ has the same law as $X_{\alpha',\beta}(\cdot)$, which can be easily checked from the covariances.

For simplicity, set $\rho := \sqrt{\Gamma(2\beta + 1)2^{-(2\beta+1)}}$. Theorem 2 yields that

$$-|\log \varepsilon/\rho|^3 \asymp \log \mathbb{P} \left(\sup_{t \leq 1} |\rho X_{\alpha,\beta}(t)| \leq \varepsilon \right) = \log \mathbb{P} \left(\sup_{t \geq 1} |\rho X_{\alpha',\beta}(t)| \leq \varepsilon \right).$$

Clearly, by Anderson's inequality and the integral representation (2), the last expression is smaller than

$$\log \mathbb{P} \left(\sup_{t \geq 1} |S(t)| \leq \varepsilon \right),$$

which already shows the lower bound of the small deviation probability of S in (17).

To see the opposite bound, note that

$$\rho X_{\alpha',\beta}(t) = t^{\alpha'} \int_0^{2/\varepsilon} x^\beta e^{-xt} dB(x) + t^{\alpha'} \int_{2/\varepsilon}^\infty x^\beta e^{-xt} dB(x) =: V(t) + U(t)$$

and thus

$$e^{-c|\log \varepsilon|^3} \geq \mathbb{P} \left(\sup_{t \geq 1} |\rho X_{\alpha',\beta}(t)| \leq \varepsilon \right) \geq \mathbb{P} \left(\sup_{t \geq 1} |V(t)| \leq \varepsilon/2 \right) \cdot \mathbb{P} \left(\sup_{t \geq 1} |U(t)| \leq \varepsilon/2 \right). \quad (19)$$

Since

$$\begin{aligned} \mathbb{P} \left(\sup_{t \geq 1} |V(t)| \leq \varepsilon/2 \right) &= \mathbb{P} \left(\sup_{t \geq 1} \left| \int_0^{2/\varepsilon} t^{\alpha'} x^\beta e^{-tx} dB(x) \right| \leq \varepsilon/2 \right) \\ &= \mathbb{P} \left(\sup_{t \geq 1} \left| (\varepsilon/2)^{\alpha'} \int_0^1 (2t/\varepsilon)^{\alpha'} (2x/\varepsilon)^\beta e^{-2xt/\varepsilon} (\varepsilon/2)^{-1/2} dB(x) \right| \leq \varepsilon/2 \right) \\ &= \mathbb{P} \left(\sup_{t \geq 2/\varepsilon} \left| \int_0^1 t^{\alpha'} x^\beta e^{-xt} dB(x) \right| \leq (\varepsilon/2)^{1+1/2+\beta-\alpha'} \right) \\ &\geq \mathbb{P}(\sup_{t \geq 1} |S(t)| \leq (\varepsilon/2)^{1+1/2+\beta-\alpha'}) \end{aligned}$$

and $1 + 1/2 + \beta - \alpha' > 0$, it is sufficient to show that the second term in (19) is bounded from below by a constant. This can be seen as follows: Note that the finite dimensional distributions of U are the same as of the following process

$$\begin{aligned} &t^{\alpha'} e^{-t/\varepsilon} \int_{1/\varepsilon}^\infty (x + 1/\varepsilon)^\beta e^{-tx} dB(x) \\ &= t^{\alpha'} \varepsilon^{-\beta} e^{-t/\varepsilon} \int_{1/\varepsilon}^\infty (\varepsilon x + 1)^\beta e^{-tx} dB(x) \\ &= (t/\varepsilon)^{\alpha'} \varepsilon^{\alpha'-\beta-1/2} e^{-t/\varepsilon} \int_1^\infty (u + 1)^\beta e^{-tu/\varepsilon} dB(u). \end{aligned}$$

Therefore, estimating $e^{t/\varepsilon} \geq e^{1/\varepsilon}$ in the first step, we obtain

$$\begin{aligned} \mathbb{P}(\sup_{t \geq 1} |U(t)| \leq \varepsilon/2) &\geq \mathbb{P}\left(\sup_{t \geq 1} \left| (t/\varepsilon)^{\alpha'} \int_1^\infty (u+1)^\beta e^{-tu/\varepsilon} dB(u) \right| \leq \varepsilon^{1+\beta+1/2-\alpha'} e^{1/\varepsilon}/2\right) \\ &= \mathbb{P}\left(\sup_{s \geq 1/\varepsilon} \left| s^{\alpha'} \int_1^\infty (u+1)^\beta e^{-su} dB(u) \right| \leq \varepsilon^{1+\beta+1/2-\alpha'} e^{1/\varepsilon}/2\right) \\ &\geq \mathbb{P}\left(\sup_{s \geq 1} \left| s^{\alpha'} \int_1^\infty (u+1)^\beta e^{-su} dB(u) \right| \leq 1\right), \end{aligned}$$

for small ε because $\varepsilon^{1+\beta+1/2-\alpha'} e^{1/\varepsilon} \rightarrow \infty$ as $\varepsilon \rightarrow 0^+$. Note that the Gaussian process

$$Z(s) = s^{\alpha'} \int_1^\infty (u+1)^\beta e^{-su} dB(u)$$

is sample bounded on $[1, \infty)$ under the assumption $\alpha - \beta > 1/2$. Indeed,

$$\begin{aligned} \mathbb{E} |Z(t) - Z(s)|^2 &= \int_1^\infty (u+1)^{2\beta} (t^{\alpha'} e^{-tu} - s^{\alpha'} e^{-su})^2 du \\ &\leq C \int_0^\infty u^{2\beta} (t^{\alpha'} e^{-tu} - s^{\alpha'} e^{-su})^2 du \\ &= C' \mathbb{E} |X_{\alpha', \beta}(t) - X_{\alpha', \beta}(s)|^2 \\ &= C' \mathbb{E} |X_{\alpha, \beta}(1/t) - X_{\alpha, \beta}(1/s)|^2. \end{aligned}$$

Now, Theorem 2 implies that when $\alpha - \beta > 1/2$, $Z(t)$ is sample bounded on $[1, \infty)$. Therefore,

$$\mathbb{P}\left(\sup_{s \geq 1} \left| s^{\alpha'} \int_1^\infty (u+1)^\beta e^{-su} dB(u) \right| \leq 1\right)$$

is a positive constant, as required. □

References

- [1] H. Alzer and C. Berg. Some classes of completely monotonic functions. *Ann. Acad. Sci. Fenn. Math.* 27 (2002), 445–460.
- [2] F. Aurzada and S. Dereich. Universality of the asymptotics of the one-sided exit problem for integrated processes. To appear in: *Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques*, preprint available from: <http://arxiv.org/abs/1008.0485>
- [3] F. Aurzada, I.A. Ibragimov, M.A. Lifshits and J.H. van Zanten. Small deviations of smooth stationary Gaussian processes. *Teor. Veroyatn. Primen.* 53 (2008), 788–798 (Russian); translation in *Theory Probab. Appl.* 53 (2009), 697–707 (English).
- [4] E. Belinsky and W. Linde. Compactness properties of certain integral operators related to fractional integration. *Math. Z.* 252 (2006), 669–686.

- [5] R. Blei, F. Gao and W.V. Li. Metric entropy of high dimensional distributions. *Proc. Amer. Math. Soc.* 135 (2007), 4009–4018.
- [6] A.A. Borovkov and R.S. Ruzankin. On small deviations of series of weighted random variables. *J. Theoret. Probab.* 21 (2008), 628–649.
- [7] B. Carl and I. Stephani. *Entropy, compactness and the approximation of operators*, volume 98 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1990.
- [8] F. Gao. Entropy of absolute convex hulls in Hilbert spaces. *Bull. London Math. Soc.* 36 (2004), 460–468.
- [9] F. Gao. Entropy estimate for k -monotone functions via small ball probability of integrated Brownian motion. *Electron. Commun. Probab.* 13 (2008), 121–130.
- [10] F. Gao, W.V. Li and J. Wellner. How many Laplace transforms of probability measures are there? *Proc. Amer. Math. Soc.* 138 (2010), 4331–4344.
- [11] F. Gao and J. Wellner. On the rate of convergence of the maximum likelihood estimator of a k -monotone density. *Science in China, Series A: Mathematics* 52 (2009), 1525–1538.
- [12] Y. Gordon, H. König and C. Schütt. Geometric and probabilistic estimates for entropy and approximation numbers of operators. *J. Approx. Theory* 49 (1987), 219–239.
- [13] A.I. Karol’ and A.I. Nazarov. Small Ball Probabilities for Smooth Gaussian fields and Tensor Products of Compact Operators. Preprint available from: <http://arxiv.org/abs/1009.4412>
- [14] H. König and S. Richter. Eigenvalues of integral operators defined by analytic kernels. *Math. Nachr.* 119 (1984), 141–155.
- [15] C. Krattenthaler. Advanced determinant calculus. The Andrews Festschrift (Maratea, 1998). *Sém. Lothar. Combin.* 42 (1999), Art. B42q, 67 pp. (electronic).
- [16] J. Kuelbs and W.V. Li. Metric entropy and the small ball problem for Gaussian measures. *J. Funct. Anal.* 116 (1993), 133–157.
- [17] T. Kühn. Covering numbers of Gaussian reproducing kernel Hilbert spaces. *J. Complexity* 27 (2011) 489–499.
- [18] A.A. Laptev. Spectral asymptotic behavior of a class of integral operators. *Mat. Zametki* 16 (1974), 741–750 (Russian); translation in: *Math. Notes* 16 (1974), 1038–1043 (English).
- [19] W.V. Li. Small Value Probabilities: Techniques and Applications. Lecture notes. <http://www.math.udel.edu/~wli/svp.html>

- [20] W.V. Li and W. Linde. Approximation, metric entropy and small ball estimates for Gaussian measures. *Ann. Probab.* 27 (1999), 1556–1578.
- [21] W.V. Li and Q.-M. Shao. *Gaussian processes: inequalities, small ball probabilities and applications*, Stochastic processes: theory and methods, Handbook of Statist., vol. 19, pp. 533–597, North-Holland, Amsterdam, 2001.
- [22] W.V. Li and Q.-M. Shao. Lower tail probabilities for Gaussian processes. *Ann. Probab.* 32 (2004), 216–242.
- [23] W.V. Li and Q.-M. Shao. Recent developments on lower tail probabilities for Gaussian processes. *COSMOS, the Journal of the Singapore National Academy of Science* 1 (2005), 95–106.
- [24] M.A. Lifshits. *Bibliography compilation on small deviation probabilities*, available from: <http://www.proba.jussieu.fr/pageperso/smalldev/biblio.html>, 2010.
- [25] S.N. Majumdar. Persistence in nonequilibrium systems. *Curr. Sci.*, 77 (1999), 370–375. <http://arxiv.org/abs/cond-mat/9907407>.
- [26] A. Nazarov. Log-level comparison principle for small ball probabilities. *Stat. Probab. Letters* 79 (2009), 481–486.
- [27] A. Pietsch. *Eigenvalues and s-numbers*, volume 13 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1987.
- [28] A.W. van der Vaart and J.H. van Zanten. Rates of contraction of posterior distributions based on Gaussian process priors. *Ann. Statist.* 36 (2008), 1435–1463.
- [29] H. Widom. Asymptotic behavior of the eigenvalues of certain integral equations. II. *Arch. Rat. Mech. Anal.* 17 (1964), 215–229.