

# Large deviations for conditionally Gaussian processes: estimates of level crossing probability

Barbara Pacchiarotti<sup>a,\*</sup>, Alessandro Pigliacelli<sup>b,\*</sup>

<sup>a</sup>*Dept. of Mathematics, University of Rome “Tor Vergata”*

<sup>b</sup>*BIP-Business Integration Partners*

[pacchiar@mat.uniroma2.it](mailto:pacchiar@mat.uniroma2.it) (B. Pacchiarotti), [alex-matrix23@hotmail.it](mailto:alex-matrix23@hotmail.it) (A. Pigliacelli)

Received: 16 May 2018, Revised: 30 July 2018, Accepted: 1 October 2018,  
Published online: 12 October 2018

**Abstract** The problem of (pathwise) large deviations for conditionally continuous Gaussian processes is investigated. The theory of large deviations for Gaussian processes is extended to the wider class of random processes – the conditionally Gaussian processes. The estimates of level crossing probability for such processes are given as an application.

**Keywords** Conditionally Gaussian processes, large deviations, ruin problem

**2010 MSC** 60F10, 60G15, 60G07

## 1 Introduction

In this paper we study some large deviations principles for conditionally continuous Gaussian processes. Then we find estimates of level crossing probability for such processes. Large deviations theory is concerned with the study of probabilities of very “rare” events. There are events whose probability is very small, however these events are of great importance; they may represent an *atypical situation* (i.e. a deviation from the average behavior) that may cause disastrous consequences: an insurance company or a bank may bankrupt; a statistical estimator may give a wrong information; a physical or chemical system may show an atypical configuration. The aim of this paper is to extend the theory of large deviations for Gaussian processes to a wider class of

---

\*Corresponding authors.

random processes – the conditionally Gaussian processes. Such processes were introduced in applications in finance, optimization and control problems. See, for instance, [12, 16, 14] and [1]. More precisely, Doucet et al. in [12] considered modelling the behavior of latent variables in neural networks by Gaussian processes with random parameters; Lototsky in [16] studied stochastic parabolic equations with random coefficients; Gulisashvili in [14] studied large deviations principle for some particular stochastic volatility models where the log-price is, conditionally, a Gaussian process; in [1] probabilities of large extremes of conditionally Gaussian processes were considered, in particular sub-Gaussian processes i.e. Gaussian processes with a random variance. Let  $(Y, Z)$  be a random element on the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , where  $Z = (Z_t)_{t \in [0,1]}$  is a process taking values in  $\mathbb{R}$  and  $Y$  is an arbitrary random element (a process or a random variable). We say that  $Z$  is a *conditionally Gaussian process* if the conditional distribution of the process  $Z|Y$  is (almost surely) Gaussian. The theory of large deviations for Gaussian processes and for conditioned Gaussian processes is already well developed. See, for instance, Section 3.4 in [11] (and the references therein) for Gaussian processes, [7] and [13] for particular conditioned Gaussian processes. The extension of this theory is possible thanks to the results obtained by Chaganti in [8].

We consider a family of processes  $(Y^n, Z^n)_{n \in \mathbb{N}}$  on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ .  $(Y^n)_{n \in \mathbb{N}}$  is a family of processes taking values in a measurable space  $(E_1, \mathcal{E}_1)$  that satisfies a large deviation principle (LDP for short) and  $(Z^n)_{n \in \mathbb{N}}$  is a family of processes taking values in  $(E_2, \mathcal{E}_2)$  such that for every  $n \in \mathbb{N}$ ,  $Z^n|Y^n$  is a Gaussian process ( $\mathbb{P}$ -a.s.). We want to find a LDP for the family  $(Z^n)_{n \in \mathbb{N}}$ .

A possible application of LDPs is computing the estimates of level crossing probability (ruin problem). We will give the asymptotic behavior (in terms of large deviations) of the probability

$$p_n = \mathbb{P} \left( \sup_{0 \leq t \leq 1} (Z_t^n - \varphi(t)) > 1 \right),$$

where  $\varphi$  is a suitable function. We will consider the following families of conditionally Gaussian processes.

1) The class of Gaussian processes with random variance and random mean, i.e. the processes of the type  $(Z_t)_{t \in [0,1]} = (Y_1 X_t + Y_2)_{t \in [0,1]}$ , where  $X$  is a centered continuous Gaussian process with covariance function  $k$  and  $Y = (Y_1, Y_2)$  is a random element independent of  $X$ . Notice that  $Z|Y$  is Gaussian with

$$\text{Cov}(Z_t, Z_s|Y) = \mathbb{E}[(X_t Y_1)(X_s Y_1)|Y] = Y_1^2 k(t, s),$$

and

$$\mathbb{E}[Z_t|Y] = \mathbb{E}[Y_1 X_t + Y_2|Y] = Y_2.$$

2) The class of Ornstein–Uhlenbeck type processes with random diffusion coefficients. More precisely  $(Z_t)_{t \in [0,1]}$  is the solution of the following stochastic differential equation:

$$\begin{cases} dZ_t = (a_0 + a_1 Z_t) dt + Y dW_t, & 0 < t \leq 1, \\ Z_0 = x, \end{cases}$$

where  $x, a_0, a_1 \in \mathbb{R}$  and  $Y$  is a random element independent of the Brownian motion  $(W_t)_{t \in [0,1]}$ .

The paper is organized as follows. In Section 2 we recall some basic facts on large deviations theory for continuous Gaussian processes. In Section 3 we introduce the conditionally Gaussian processes and the Chaganty theory. In Section 4 and 5 we study the theoretical problem and we give the main results. Finally in Section 6 we investigate the ruin problem for such processes.

## 2 Large deviations for continuous Gaussian processes

We briefly recall some main facts on large deviations principles and reproducing kernel Hilbert spaces for Gaussian processes we are going to use. For a detailed development of this very wide theory we can refer, for example, to the following classical references: Chapitre II in Azencott [2], Section 3.4 in Deuschel and Strook [11], Chapter 4 (in particular Sections 4.1 and 4.5) in Dembo and Zeitouni [10], for large deviations principles; Chapter 4 (in particular Section 4.3) in [15], Chapter 2 (in particular Sections 2.2 and 2.3) in [5], for reproducing kernel Hilbert space. Without loss of generality, we can consider *centered* Gaussian processes.

### 2.1 Reproducing kernel Hilbert space

An important tool to handle continuous Gaussian processes is the associated reproducing kernel Hilbert space (RKHS).

Let  $U = (U_t)_{t \in [0,1]}$  be a continuous, centered, Gaussian process on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , with covariance function  $k$ . From now on, we will denote by  $\mathcal{C}([0, 1])$  the set of continuous functions on  $[0, 1]$  endowed with the topology induced by the sup-norm  $(\mathcal{C}([0, 1]), \|\cdot\|_\infty)$ . Moreover, we will denote by  $\mathcal{M}[0, 1]$  its dual, i.e. the set of signed Borel measures on  $[0, 1]$ . The action of  $\mathcal{M}[0, 1]$  on  $\mathcal{C}([0, 1])$  is given by

$$\langle \lambda, h \rangle = \int_0^1 h(t) d\lambda(t), \quad \lambda \in \mathcal{M}[0, 1], \quad h \in \mathcal{C}([0, 1]).$$

Consider the set

$$\mathcal{L} = \left\{ x \in \mathcal{C}([0, 1]) \mid x(t) = \int_0^1 k(t, s) d\lambda(s), \lambda \in \mathcal{M}[0, 1] \right\}.$$

The RKHS relative to the kernel  $k$  can be constructed as the completion of the set  $\mathcal{L}$  with respect to a suitable norm. Consider the set of (real) Gaussian random variables

$$\Gamma = \left\{ Y \mid Y = \langle \lambda, U \rangle = \int_0^1 U_t d\lambda(t), \lambda \in \mathcal{M}[0, 1] \right\} \subset L^2(\Omega, \mathcal{F}, \mathbb{P}).$$

We have that, for  $Y_1, Y_2 \in \Gamma$ , say  $Y_i = \langle \lambda_i, U \rangle$ ,  $i = 1, 2$ ,

$$\begin{aligned} \langle Y_1, Y_2 \rangle_{L^2(\Omega, \mathcal{F}, \mathbb{P})} &= \text{Cov} \left( \int_0^1 U_t d\lambda_1(t), \int_0^1 U_t d\lambda_2(t) \right) \\ &= \int_0^1 \int_0^1 k(t, s) d\lambda_1(t) d\lambda_2(s). \end{aligned} \quad (1)$$

Define now

$$H = \overline{F}^{\|\cdot\|_{L^2(\Omega, \mathcal{F}, \mathbb{P})}}.$$

Since  $L^2$ -limits of Gaussian random variables are still Gaussian, we have that  $H$  is a closed subspace of  $L^2(\Omega, \mathcal{F}, \mathbb{P})$  consisting of real Gaussian random variables. Moreover, it becomes a Hilbert space when endowed with the inner product

$$\langle Y_1, Y_2 \rangle_H = \langle Y_1, Y_2 \rangle_{L^2(\Omega, \mathcal{F}, \mathbb{P})}, \quad Y_1, Y_2 \in H.$$

**Remark 1.** We remark that, since any signed Borel measure  $\lambda$  can be weakly approximated by a linear combination of Dirac deltas, the Hilbert space  $H$  above is nothing but the Hilbert space generated by the Gaussian process  $U$ , namely

$$\begin{aligned} H &= \overline{\text{sp}\{U_t, t \in [0, 1]\}}^{\|\cdot\|_{L^2(\Omega, \mathcal{F}, \mathbb{P})}} \\ &= \overline{\left\{ \sum_{j=1}^n a_j U_{t_j} \mid n \in \mathbb{N}, a_j \in \mathbb{R}, t_j \in [0, 1] \right\}}^{\|\cdot\|_{L^2(\Omega, \mathcal{F}, \mathbb{P})}}. \end{aligned}$$

Consider now the following mapping

$$\mathcal{S} : H \longrightarrow \mathcal{C}([0, 1]), \quad Y \mapsto (\mathcal{S}Y)_\cdot = \mathbb{E}(UY). \quad (2)$$

**Definition 1.** Let  $U = (U_t)_{t \in [0, 1]}$  be a continuous Gaussian process. We define the reproducing kernel Hilbert space relative to the Gaussian process  $U$  as

$$\mathcal{H} = \mathcal{S}(H) = \{h \in \mathcal{C}([0, 1]) \mid h(t) = (\mathcal{S}Y)_t, Y \in H\},$$

with an inner product defined as

$$\langle h_1, h_2 \rangle_{\mathcal{H}} = \langle \mathcal{S}^{-1}h_1, \mathcal{S}^{-1}h_2 \rangle_H = \langle \mathcal{S}^{-1}h_1, \mathcal{S}^{-1}h_2 \rangle_{L^2(\Omega, \mathcal{F}, \mathbb{P})}, \quad h_1, h_2 \in \mathcal{H}.$$

Then, we have

**Lemma 1.** (Theorem 35 in [5]). Let  $H$  be the Hilbert space of the continuous Gaussian process  $U$  defined above. Then  $H$  is isometrically isomorphic to the Reproducing Kernel Hilbert Space  $\mathcal{H}$  of  $U$ , and the corresponding isometry is given by (2).

The map  $\mathcal{S}$  defined in (2) is referred to as *Loève isometry*. Since the covariance function fully identifies, up to the mean, a Gaussian process, we can talk equivalently of RKHS associated with the process or with its covariance function.

## 2.2 Large deviations

**Definition 2.** (LDP) Let  $E$  be a topological space,  $\mathcal{B}(E)$  the Borel  $\sigma$ -algebra and  $(\mu_n)_{n \in \mathbb{N}}$  a family of probability measures on  $\mathcal{B}(E)$ ; let  $\gamma : \mathbb{N} \rightarrow \mathbb{R}^+$  be a function, such that  $\gamma(n) \rightarrow +\infty$  as  $n \rightarrow +\infty$ . We say that the family of probability measures  $(\mu_n)_{n \in \mathbb{N}}$  satisfies a large deviation principle (LDP) on  $E$  with the rate function  $I$  and the speed  $\gamma(n)$  if, for any open set  $\Theta$ ,

$$- \inf_{x \in \Theta} I(x) \leq \liminf_{n \rightarrow +\infty} \frac{1}{\gamma(n)} \log \mu_n(\Theta)$$

and for any closed set  $\Gamma$

$$\limsup_{n \rightarrow +\infty} \frac{1}{\gamma(n)} \log \mu_n(\Gamma) \leq - \inf_{x \in \Gamma} I(x). \quad (3)$$

A rate function is a lower semicontinuous mapping  $I : E \rightarrow [0, +\infty]$ . A rate function  $I$  is said *good* if the sets  $\{I \leq a\}$  are compact for every  $a \geq 0$ .

**Definition 3.** (WLDP) Let  $E$  be a topological space,  $\mathcal{B}(E)$  the Borel  $\sigma$ -algebra and  $(\mu_n)_{n \in \mathbb{N}}$  a family of probability measures on  $\mathcal{B}(E)$ ; let  $\gamma : \mathbb{N} \rightarrow \mathbb{R}^+$  be a function, such that  $\gamma(n) \rightarrow +\infty$  as  $n \rightarrow +\infty$ . We say that the family of probability measures  $(\mu_n)_{n \in \mathbb{N}}$  satisfies a weak large deviation principle (WLDP) on  $E$  with the rate function  $I$  and the speed  $\gamma(n)$  if the upper bound (3) holds for compact sets.

**Remark 2.** We say that a family of continuous processes  $((X_t^n)_{t \in [0,1]})_{n \in \mathbb{N}}$  satisfies a LDP if the associated family of laws satisfy a LDP on  $\mathcal{C}([0, 1])$ .

The following remarkable theorem (Proposition 1.5 in [2]) gives an explicit expression of the Cramér transform  $\Lambda^*$  of a continuous centered Gaussian process  $(U_t)_{t \in [0,1]}$  with covariance function  $k$ . Let us recall that

$$\Lambda(\lambda) = \log \mathbb{E}[\exp(\langle U, \lambda \rangle)] = \frac{1}{2} \int_0^1 \int_0^1 k(t, s) d\lambda(t) d\lambda(s),$$

for  $\lambda \in \mathcal{M}[0, 1]$ .

**Theorem 1.** Let  $(U_t)_{t \in [0,1]}$  be a continuous, centered Gaussian process with covariance function  $k$ . Let  $\Lambda^*$  denote the Cramér transform of  $\Lambda$ , that is,

$$\begin{aligned} \Lambda^*(x) &= \sup_{\lambda \in \mathcal{M}[0,1]} (\langle \lambda, x \rangle - \Lambda(\lambda)) \\ &= \sup_{\lambda \in \mathcal{M}[0,1]} \left( \langle \lambda, x \rangle - \frac{1}{2} \int_0^1 \int_0^1 k(t, s) d\lambda(t) d\lambda(s) \right). \end{aligned}$$

Then,

$$\Lambda^*(x) = \begin{cases} \frac{1}{2} \|x\|_{\mathcal{H}}^2, & x \in \mathcal{H}, \\ +\infty & \text{otherwise,} \end{cases} \quad (4)$$

where  $\mathcal{H}$  and  $\|\cdot\|_{\mathcal{H}}$  denote, respectively, the reproducing kernel Hilbert space and the related norm associated to the covariance function  $k$ .

In order to state a large deviation principle for a family of Gaussian processes, we need the following definition.

**Definition 4.** A family of continuous processes  $((X_t^n)_{t \in [0,1]})_{n \in \mathbb{N}}$  is exponentially tight at the speed  $\gamma(n)$  if, for every  $R > 0$  there exists a compact set  $K_R$  such that

$$\limsup_{n \rightarrow +\infty} \frac{1}{\gamma(n)} \log \mathbb{P}(X^n \notin K_R) \leq -R. \quad (5)$$

If the means and the covariance functions of an exponentially tight family of Gaussian processes have a good limit behavior, then the family satisfies a large deviation principle, as stated in the following theorem which is a consequence of the classic abstract Gärtner–Ellis Theorem (Baldi Theorem 4.5.20 and Corollary 4.6.14 in [10]) and Theorem 1.

**Theorem 2.** Let  $((X_t^n)_{t \in [0,1]})_{n \in \mathbb{N}}$  be an exponentially tight family of continuous Gaussian processes with respect to the speed function  $\gamma(n)$ . Suppose that, for any  $\lambda \in \mathcal{M}[0, 1]$ ,

$$\lim_{n \rightarrow +\infty} \mathbb{E}[\langle \lambda, X^n \rangle] = 0 \quad (6)$$

and the limit

$$\Lambda(\lambda) = \lim_{n \rightarrow +\infty} \gamma(n) \text{Var}(\langle \lambda, X^n \rangle) = \int_0^1 \int_0^1 \bar{k}(t, s) d\lambda(t) d\lambda(s) \quad (7)$$

exists for some continuous, symmetric, positive definite function  $\bar{k}$ , which is the covariance function of a continuous Gaussian process. Then  $((X_t^n)_{t \in [0,1]})_{n \in \mathbb{N}}$  satisfies a large deviation principle on  $\mathcal{C}([0, 1])$ , with the speed  $\gamma(n)$  and the good rate function

$$I(h) = \begin{cases} \frac{1}{2} \|h\|_{\bar{\mathcal{H}}}^2, & h \in \bar{\mathcal{H}}, \\ +\infty & \text{otherwise,} \end{cases} \quad (8)$$

where  $\bar{\mathcal{H}}$  and  $\|\cdot\|_{\bar{\mathcal{H}}}$  respectively denote the reproducing kernel Hilbert space and the related norm associated to the covariance function  $\bar{k}$ .

A useful result which can help in investigating the exponential tightness of a family of continuous centered Gaussian processes is the following proposition (Proposition 2.1 in [17]); the required property follows from Hölder continuity of the mean and the covariance function.

**Proposition 1.** Let  $((X_t^n)_{t \in [0,1]})_{n \in \mathbb{N}}$  be a family of continuous Gaussian processes with  $X_0^n = 0$  for all  $n \in \mathbb{N}$ . Denote  $m^n(t) = \mathbb{E}[X_t^n]$  and  $k^n(t, s) = \text{Cov}(X_t^n, X_s^n)$ . Suppose there exist constants  $\alpha, M_1, M_2 > 0$  such that for  $n \in \mathbb{N}$

$$\sup_{s, t \in [0,1], s \neq t} \frac{|m^n(t) - m^n(s)|}{|t - s|^\alpha} \leq M_1$$

and

$$\sup_{s, t \in [0,1], s \neq t} \gamma(n) \frac{|k^n(t, t) + k^n(s, s) - 2k^n(s, t)|}{|t - s|^{2\alpha}} \leq M_2. \quad (9)$$

Then the family  $((X_t^n)_{t \in [0,1]})_{n \in \mathbb{N}}$  is exponentially tight with respect to the speed function  $\gamma(n)$ .

### 3 Conditionally Gaussian processes

In this section we introduce conditionally Gaussian processes and the Chaganty theorem which allows us to find a LDP for families of such processes. We also recall, for sake of completeness, some results about conditional distributions in Polish spaces. We referred to Section 3.1 in [6] and Section 4.3 in [3].

Let  $Y$  and  $Z$  two random variables, defined on the same probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , with values in the measurable spaces  $(E_1, \mathcal{E}_1)$  and  $(E_2, \mathcal{E}_2)$  respectively, and let us denote by  $\mu_1, \mu_2$  the (marginal) laws of  $Y$  and  $Z$  respectively and by  $\mu$  the joint distribution of  $(Y, Z)$  on  $(E, \mathcal{E}) = (E_1 \times E_2, \mathcal{E}_1 \times \mathcal{E}_2)$ . A family of probabilities  $(\mu_2(dz|y))_{y \in E_1}$  on  $(E_2, \mathcal{E}_2)$  is a regular version of the conditional law of  $Z$  given  $Y$  if

1. For every  $B \in \mathcal{E}_2$ , the map  $y \mapsto \mu_2(B|y)$  is  $\mathcal{E}_1$ -measurable.
2. For every  $B \in \mathcal{E}_2$  and  $A \in \mathcal{E}_1$ ,  $\mathbb{P}(Y \in A, Z \in B) = \int_A \mu_2(B|y)\mu_1(dy)$ .

In this case we have

$$\mu(dy, dz) = \mu_2(dz|y)\mu_1(dy).$$

In this section we will use the notation  $(E, \mathcal{B})$  to indicate a Polish space (i.e. a complete separable metric space) with the Borel  $\sigma$ -field, and we say that a sequence  $(x_n)_{n \in \mathbb{N}} \subset \mathbb{E}$  converges to  $x \in E$ ,  $x_n \rightarrow x$ , if  $d_E(x_n, x) \rightarrow 0$ , as  $n \rightarrow \infty$ , where  $d_E$  denotes the metric on  $E$ . Regular conditional probabilities do not always exist, but they exist in many cases. The following result, that immediately follows from Corollary 3.2.1 in [6], shows that in Polish spaces the regular version of the conditional probability is well defined.

**Proposition 2.** *Let  $(E_1, \mathcal{B}_1)$  and  $(E_2, \mathcal{B}_2)$  be two Polish spaces endowed with their Borel  $\sigma$ -fields,  $\mu$  be a probability measure on  $(E, \mathcal{B}) = (E_1 \times E_2, \mathcal{B}_1 \times \mathcal{B}_2)$ . Let  $\mu_i$  be the marginal probability measure on  $(E_i, \mathcal{B}_i)$ ,  $i = 1, 2$ . Then there exists  $\mu_1$ -almost sure a unique regular version of the conditional law of  $\mu_2$  given  $\mu_1$ , i.e.*

$$\mu(dy, dz) = \mu_2(dz|y)\mu_1(dy).$$

In what follows we always suppose random variables taking values in a Polish space.

**Definition 5.** Let  $(Y, Z)$  be a random element on the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , where  $Z = (Z_t)_{t \in [0,1]}$  is a real process and  $Y$  is an arbitrary random element (a process or a random variable). We say that  $Z$  is a *conditionally Gaussian process* if the conditional distribution of the process  $Z|Y$  is (almost surely) Gaussian. We denote by  $(Z_t^y)_{t \in [0,1]}$  the Gaussian process  $Z|Y = y$ .

The main tool that we will use to study LDP for a family of conditionally Gaussian processes is provided by Chaganty Theorem (Theorem 2.3 in [8]). Let  $(E_1, \mathcal{B}_1)$  and  $(E_2, \mathcal{B}_2)$  be two Polish spaces. We denote by  $(\mu_n)_{n \in \mathbb{N}}$  a sequence of probability measures on  $(E, \mathcal{B}) = (E_1 \times E_2, \mathcal{B}_1 \times \mathcal{B}_2)$  (the sequence of *joint distributions*), by  $(\mu_{1n})_{n \in \mathbb{N}}$  the sequence of the *marginal distributions* on  $(E_1, \mathcal{B}_1)$  and by  $(\mu_{2n}(\cdot|x_1))_{n \in \mathbb{N}}$  the sequence of *conditional distributions* on  $(E_2, \mathcal{B}_2)$  ( $x_1 \in E_1$ ), given by Proposition 2, i.e.

$$\mu_n(B_1 \times B_2) = \int_{B_1} \mu_{2n}(B_2|x_1) \mu_{1n}(dx_1).$$

**Definition 6.** Let  $(E_1, \mathcal{B}_1)$ ,  $(E_2, \mathcal{B}_2)$  be two Polish spaces and  $x_1 \in E_1$ . We say that the sequence of conditional laws  $(\mu_{2n}(\cdot|x_1))_{n \in \mathbb{N}}$  on  $(E_2, \mathcal{B}_2)$  satisfies the LDP continuously in  $x_1$  with the rate function  $J(\cdot|x_1)$  and the speed  $\gamma(n)$ , or simply, the LDP continuity condition holds, if

- a) For each  $x_1 \in E_1$ ,  $J(\cdot|x_1)$  is a good rate function on  $E_2$ .
- b) For any sequence  $(x_{1n})_{n \in \mathbb{N}} \subset E_1$  such that  $x_{1n} \rightarrow x_1$ , the sequence of measures  $(\mu_{2n}(\cdot|x_{1n}))_{n \in \mathbb{N}}$  satisfies a LDP on  $E_2$  with the (same) rate function  $J(\cdot|x_1)$  and the speed  $\gamma(n)$ .
- c)  $J(\cdot)$  is lower semicontinuous as a function of  $(x_1, x_2) \in E_1 \times E_2$ .

**Theorem 3** (Theorem 2.3 in [8]). *Let  $(E_1, \mathcal{B}_1)$ ,  $(E_2, \mathcal{B}_2)$  be two Polish spaces. For  $i = 1, 2$  let  $(\mu_{in})_{n \in \mathbb{N}}$  be a sequence of measures on  $(E_i, \mathcal{B}_i)$ . For  $x_1 \in E_1$ , let  $(\mu_{2n}(\cdot|x_1))_{n \in \mathbb{N}}$  be the sequence of the conditional laws (of  $\mu_{2n}$  given  $\mu_{1n}$ ) on  $(E_2, \mathcal{B}_2)$ . Suppose that the following two conditions are satisfied:*

- i)  $(\mu_{1n})_{n \in \mathbb{N}}$  satisfies a LDP on  $E_1$  with the good rate function  $I_1(\cdot)$  and the speed  $\gamma(n)$ .
- ii) For every  $x_1 \in E_1$ , the sequence  $(\mu_{2n}(\cdot|x_1))_{n \in \mathbb{N}}$  satisfies the LDP continuity condition on  $E_2$  with the rate function  $J(\cdot|x_1)$  and the speed  $\gamma(n)$ .

*Then the sequence of joint distributions  $(\mu_n)_{n \in \mathbb{N}}$  satisfies a WLDP on  $E = E_1 \times E_2$  with the speed  $\gamma(n)$  and the rate function*

$$I(x_1, x_2) = I_1(x_1) + J(x_2|x_1), \quad x_1 \in E_1, x_2 \in E_2.$$

*The sequence of marginal distributions  $(\mu_{2n})_{n \in \mathbb{N}}$  defined on  $(E_2, \mathcal{B}_2)$ , satisfies a LDP with the speed  $\gamma(n)$  and the rate function*

$$I_2(x_2) = \inf_{x_1 \in E_1} I(x_1, x_2).$$

*Moreover, if  $I(\cdot, \cdot)$  is a good rate function then  $(\mu_n)_{n \in \mathbb{N}}$  satisfies a LDP and  $I_2(\cdot)$  is a good rate function.*

#### 4 Gaussian process with random mean and random variance

Let  $\alpha > 0$  and define  $\mathcal{C}_\alpha([0, 1]) = \{y \in \mathcal{C}([0, 1]) : y(t) \geq \alpha, t \in [0, 1]\}$  (with the uniform norm on compact sets).  $\mathcal{C}_\alpha([0, 1])$  is a Polish space. Consider the family of processes  $(Y^n, Z^n)_{n \in \mathbb{N}}$ , where  $(Y^n)_{n \in \mathbb{N}} = (Y_1^n, Y_2^n)_{n \in \mathbb{N}}$  is a family of processes with paths in  $\mathcal{C}_\alpha([0, 1]) \times \mathcal{C}([0, 1])$  and for  $n \in \mathbb{N}$ ,  $Z^n = X^n Y_1^n + Y_2^n$  with  $(Y^n)_{n \in \mathbb{N}}$  independent of  $(X^n)_{n \in \mathbb{N}}$ . Suppose  $((X_t^n)_{t \in [0, 1]})_{n \in \mathbb{N}}$  is family of continuous centered Gaussian processes which satisfy the hypotheses of Theorem 2 and suppose that  $(Y^n)_{n \in \mathbb{N}}$  satisfies a LDP with the good rate function  $I_Y$  and the speed  $\gamma(n)$ . We want to prove a LDP principle for  $(Z^n)_{n \in \mathbb{N}}$ .

**Proposition 3.** Let  $((X_t^n)_{t \in [0, 1]})_{n \in \mathbb{N}}$  be a family of continuous Gaussian processes which satisfies the hypotheses of Theorem 2 and let  $y = (y_1, y_2) \in \mathcal{C}_\alpha([0, 1]) \times \mathcal{C}([0, 1])$ . Then the family  $((X_t^n y_1(t) + y_2(t))_{t \in [0, 1]})_{n \in \mathbb{N}}$  is still a family of continuous Gaussian processes which satisfies the hypotheses of Theorem 2 with the same speed function and limit covariance function (depending only on  $y_1$ )  $k^{y_1}$  given by

$$k^{y_1}(s, t) = y_1(s)y_1(t)\bar{k}(s, t). \quad (10)$$

Therefore, also  $((X_t^n y_1(t) + y_2(t))_{t \in [0, 1]})_{n \in \mathbb{N}}$  satisfies a LDP with the good rate function

$$\Lambda_y^*(z) = \begin{cases} \frac{1}{2} \|z\|_{\mathcal{H}_{y_1}}^2, & z \in \bar{\mathcal{H}}_{y_1}, \\ +\infty & \text{otherwise,} \end{cases} = \begin{cases} \frac{1}{2} \left\| \frac{z-y_2}{y_1} \right\|_{\mathcal{H}}^2, & \frac{z-y_2}{y_1} \in \bar{\mathcal{H}}, \\ +\infty & \text{otherwise,} \end{cases} \quad (11)$$

where  $\bar{\mathcal{H}}_{y_1}$  is the RKHS associated to the covariance function defined in (10).

**Proof.** This is a simple application of the contraction principle.  $\square$

**Remark 3.** If  $y_1(t) = y_1 > 0$  for all  $t \in [0, 1]$ , then we have

$$\Lambda_y^*(z) = \begin{cases} \frac{1}{2y_1^2} \|z - y_2\|_{\mathcal{H}}^2, & z - y_2 \in \tilde{\mathcal{H}}, \\ +\infty & \text{otherwise.} \end{cases}$$

**Definition 7.** Let  $(E, d_E)$  be a metric space, and let  $(\mu_n)_{n \in \mathbb{N}}, (\tilde{\mu}_n)_{n \in \mathbb{N}}$  be two families of probability measures on  $E$ . Then  $(\mu_n)_{n \in \mathbb{N}}$  and  $(\tilde{\mu}_n)_{n \in \mathbb{N}}$  are exponentially equivalent (at the speed  $\gamma(n)$ ) if there exist a family of probability spaces  $((\Omega, \mathcal{F}^n, \mathbb{P}^n))_{n \in \mathbb{N}}$  and two families of  $E$ -valued random variables  $(Z^n)_{n \in \mathbb{N}}$  and  $(\tilde{Z}^n)_{n \in \mathbb{N}}$  such that, for any  $\delta > 0$ , the set  $\{\omega : d_E(\tilde{Z}^n(\omega), Z^n(\omega)) > \delta\}$  is  $\mathcal{F}^n$ -measurable and

$$\limsup_{n \rightarrow +\infty} \gamma(n) \log \mathbb{P}^n(d_E(\tilde{Z}^n(\omega), Z^n(\omega)) > \delta) = -\infty.$$

As far as the LDP is concerned exponentially equivalent measures are indistinguishable. See Theorem 4.2.13 in [10].

**Proposition 4.** Let  $((X_t^n)_{t \in [0,1]})_{n \in \mathbb{N}}$  be an exponential tight (at the speed  $\gamma(n)$ ) family of continuous Gaussian processes. Let  $(y^n)_{n \in \mathbb{N}} \subset \mathcal{C}_\alpha([0, 1]) \times \mathcal{C}([0, 1])$  such that  $y^n \rightarrow y$  in  $\mathcal{C}_\alpha([0, 1]) \times \mathcal{C}([0, 1])$ . Then, the family of processes  $((y_2^n(t) + y_1^n(t)X^n(t)))_{t \in [0,1]})_{n \in \mathbb{N}}$  is exponentially equivalent to  $((y_2(t) + y_1(t)X^n(t)))_{n \in \mathbb{N}}$ .

**Proof.** Let  $Z^n(t) = y_2(t) + y_1(t)X^n(t)$  and  $\tilde{Z}^n(t) = y_2^n(t) + y_1^n(t)X^n(t)$  for  $t \in [0, 1]$ ,  $n \in \mathbb{N}$ . Then, for any  $\delta > 0$ ,

$$\mathbb{P}(\|Z^n - \tilde{Z}^n\|_\infty > \delta) \leq \mathbb{P}\left(\|X^n\|_\infty \|y_1^n - y_1\|_\infty > \frac{\delta}{2}\right) + \mathbb{P}\left(\|y_2^n - y_2\|_\infty > \frac{\delta}{2}\right).$$

For  $n$  large enough  $\|y_2^n - y_2\|_\infty \leq \frac{\delta}{2}$  and thanks to (5)

$$\limsup_{n \rightarrow +\infty} \frac{1}{\gamma(n)} \log \mathbb{P}\left(\|X^n\|_\infty > \frac{\delta}{2\|y_1^n - y_1\|_\infty}\right) = -\infty,$$

therefore

$$\limsup_{n \rightarrow +\infty} \frac{1}{\gamma(n)} \log \mathbb{P}(\|Z^n - \tilde{Z}^n\|_\infty > \delta) = -\infty. \quad \square$$

Let us denote  $J(z|y) = \Lambda_y^*(z)$ , for  $z \in \mathcal{C}([0, 1])$  and  $y \in \mathcal{C}_\alpha([0, 1]) \times \mathcal{C}([0, 1])$ . We want to prove the lower semicontinuity of  $J(\cdot|\cdot)$ .

**Proposition 5.** If  $(z^n, y^n) \rightarrow (z, y)$  in  $\mathcal{C}([0, 1]) \times \mathcal{C}_\alpha([0, 1]) \times \mathcal{C}([0, 1])$ , then

$$\liminf_{n \rightarrow +\infty} J(z^n|y^n) \geq J(z|y).$$

**Proof.** Thanks to the lower semicontinuity of  $\|\cdot\|_{\tilde{\mathcal{H}}}^2$

$$\begin{aligned} \liminf_{(y^n, z^n) \rightarrow (y, z)} J(z^n|y^n) &= \liminf_{(y^n, z^n) \rightarrow (y, z)} \frac{1}{2} \left\| \frac{z^n - y_2^n}{y_1^n} \right\|_{\tilde{\mathcal{H}}}^2 \\ &= \liminf_{h^n \rightarrow h} \frac{1}{2} \|h^n\|_{\tilde{\mathcal{H}}}^2 \geq \frac{1}{2} \|h\|_{\tilde{\mathcal{H}}}^2 = \frac{1}{2} \left\| \frac{z - y_2}{y_1} \right\|_{\tilde{\mathcal{H}}}^2 = J(z|y) \end{aligned}$$

where,  $h^n = \frac{z^n - y_2^n}{y_1^n} \xrightarrow{\mathcal{C}([0,1])} h = \frac{z - y_2}{y_1}$ . □

**Theorem 4.** Consider the family of processes  $(Y^n, Z^n)_{n \in \mathbb{N}}$ , where  $(Y^n)_{n \in \mathbb{N}} = (Y_1^n, Y_2^n)_{n \in \mathbb{N}}$  is a family of processes with paths in  $\mathcal{C}_\alpha([0, 1]) \times \mathcal{C}([0, 1])$  and for  $n \in \mathbb{N}$ ,  $Z^n = X^n Y_1^n + Y_2^n$  with  $(Y^n)_{n \in \mathbb{N}}$  independent of  $(X^n)_{n \in \mathbb{N}}$ . Suppose  $((X_t^n)_{t \in [0, 1]})_{n \in \mathbb{N}}$  is family of continuous centered Gaussian processes which satisfy the hypotheses of Theorem 2 and suppose that  $(Y^n)_{n \in \mathbb{N}}$  satisfies a LDP with the good rate function  $I_Y$  and the speed  $\gamma(n)$ . Then  $(Y^n, Z^n)_{n \in \mathbb{N}}$  satisfies the WLDP with the speed  $\gamma(n)$  and the rate function

$$I(y, z) = I_Y(y) + J(z|y),$$

and  $(Z^n)_{n \in \mathbb{N}}$  satisfies the LDP with the speed  $\gamma(n)$  and the rate function

$$I_Z(z) = \inf_{y \in \mathcal{C}_\alpha([0, 1])} \{I_Y(y) + J(z|y)\}.$$

**Proof.** Thanks to Propositions 3, 4 and 5 the family of processes  $(Y^n, Z^n)_{n \in \mathbb{N}}$  satisfies the hypotheses of Theorem 3, therefore the theorem holds.  $\square$

## 5 Ornstein–Uhlenbeck processes with random diffusion coefficient

Let  $\alpha > 0$  and let again  $\mathcal{C}_\alpha([0, 1]) = \{y \in \mathcal{C}([0, 1]) : y(t) \geq \alpha, t \in [0, 1]\}$ . Consider the family of processes  $(Y^n, Z^n)_{n \in \mathbb{N}}$ , where  $(Y^n)_{n \in \mathbb{N}}$  is a family of processes with paths in  $\mathcal{C}_\alpha([0, 1])$  and for  $n \in \mathbb{N}$ ,  $Z^n$  is the solution of the following stochastic differential equation.

$$\begin{cases} dZ_t^n = (a_0 + a_1 Z_t^n) dt + \frac{1}{\sqrt{n}} Y_t^n dW_t, & 0 < t \leq 1, \\ Z_0^n = x, \end{cases} \quad (12)$$

where,  $x, a_0, a_1 \in \mathbb{R}$  and  $(Y^n)_{n \in \mathbb{N}}$  is a family of random processes independent of the Brownian motion  $(W_t)_{t \in [0, 1]}$ .

Suppose that  $(Y^n)_{n \in \mathbb{N}}$  satisfies a LDP with the good rate function  $I_Y$ , and the speed  $\gamma(n) = n$ . We want to prove a LDP principle for  $(Z^n)_{n \in \mathbb{N}}$ .

Let  $Z^{n,y}, y \in \mathcal{C}_\alpha([0, 1])$ , be the solution of the following stochastic differential equation,

$$\begin{cases} dZ_t^{n,y} = (a_0 + a_1 Z_t^{n,y}) dt + \frac{1}{\sqrt{n}} y(t) dW_t, & 0 < t \leq 1, \\ Z_0^{n,y} = x, \end{cases} \quad (13)$$

that is

$$\begin{aligned} Z_t^{n,y} &= e^{a_1 t} \left( x + \frac{a_0}{a_1} [1 - e^{-a_1 t}] + \frac{1}{\sqrt{n}} \int_0^t e^{-a_1 s} y(s) dW_s \right) \\ &= m(t) + e^{a_1 t} \frac{1}{\sqrt{n}} \int_0^t e^{-a_1 s} y(s) dW_s, \end{aligned} \quad (14)$$

$m(t) = e^{a_1 t} (x + \frac{a_0}{a_1} [1 - e^{-a_1 t}])$ .  $((Z_t^{n,y})_{t \in [0, 1]})_{n \in \mathbb{N}}$  is a family of Gaussian processes and a family of diffusions. It is well known from the Wentzell–Friedlin theory that  $((Z_t^{n,y})_{t \in [0, 1]})_{n \in \mathbb{N}}$  satisfies the LDP in  $\mathcal{C}([0, 1])$  with the speed  $\gamma(n) = n$  and the good rate function

$$J(f|y) = \begin{cases} \frac{1}{2} \int_0^1 \left( \frac{\dot{f}(t) - (a_0 + a_1 f(t))}{y(t)} \right)^2 dt, & f \in H_1^x, \\ +\infty, & f \notin H_1^x, \end{cases} \quad (15)$$

where

$$H_1^x := \left\{ f : f(t) = x + \int_0^t \phi(s) ds, \phi \in L^2([0, 1]) \right\}.$$

And it is well known from the theory of Gaussian processes that the family  $((Z_t^{n,y})_{t \in [0,1]})_{n \in \mathbb{N}}$  satisfies the LDP in  $\mathcal{C}([0, 1])$  with the speed  $\gamma(n) = n$  and the good rate function

$$J(f|y) = \begin{cases} \frac{1}{2} \|f - m\|_{\mathcal{H}_y}^2, & f - m \in \mathcal{H}_y, \\ +\infty, & f - m \notin \mathcal{H}_y, \end{cases} \quad (16)$$

where  $\mathcal{H}_y$  is the reproducing kernel Hilbert space associate to the covariance function

$$k^y(s, t) = e^{a_1(s+t)} \int_0^{s \wedge t} e^{-2a_1 u} y^2(u) du.$$

The two rate functions (for the unicity of the rate function) are the same rate function. So we can deduce a LDP for the family  $(Z^n)_{n \in \mathbb{N}}$  in two different ways. First let  $(Z_t^{n,y})_{t \in [0,1]}$  be a family of diffusions.

**Remark 4.** For  $y^n \in \mathcal{C}_\alpha([0, 1])$  let  $Z^{n,y^n}$  denote the solution of equation (13) with  $y$  replaced by  $y^n$ . Then if  $y^n \rightarrow y$  in  $\mathcal{C}_\alpha([0, 1])$ , from the generalized Wentzell–Friedlin theory (Theorem 1 in [9]), we have that  $((Z_t^{n,y^n})_{t \in [0,1]})_{n \in \mathbb{N}}$  satisfies the same LDP in  $\mathcal{C}([0, 1])$  as  $((Z_t^{n,y})_{t \in [0,1]})_{n \in \mathbb{N}}$ .

We now want to prove the lower semicontinuity of  $J(\cdot|y)$  on  $\mathcal{C}([0, 1]) \times \mathcal{C}_\alpha([0, 1])$ .

**Proposition 6.** *If  $(f^n, y^n) \rightarrow (f, y)$  in  $\mathcal{C}([0, 1]) \times \mathcal{C}_\alpha([0, 1])$ , then*

$$\liminf_{n \rightarrow +\infty} J(f^n|y^n) \geq J(f|y).$$

**Proof.** If  $y^n \xrightarrow{\mathcal{C}_\alpha([0,1])} y$ , then for any  $\varepsilon > 0$ , eventually  $\inf_{t \in [0,1]} \left| \frac{y(t)}{y^n(t)} \right|^2 \geq (1 - \varepsilon)$ , and by the lower semicontinuity of  $J(\cdot|y)$ ,

$$\begin{aligned} & \liminf_{(y^n, f^n) \rightarrow (y, f)} J(f^n|y^n) \\ &= \liminf_{(y^n, f^n) \rightarrow (y, f)} \frac{1}{2} \int_0^1 \left| \frac{\dot{f}^n(t) - (a_0 + a_1 f^n(t))}{y^n(t)} \right|^2 dt \\ &= \liminf_{(y^n, f^n) \rightarrow (y, f)} \frac{1}{2} \int_0^1 \left| \frac{\dot{f}^n(t) - (a_0 + a_1 f^n(t))}{y(t)} \right|^2 \cdot \left| \frac{y(t)}{y^n(t)} \right|^2 dt \\ &\geq \liminf_{(y^n, f^n) \rightarrow (y, f)} \frac{1}{2} \int_0^1 \left| \frac{\dot{f}^n(t) - (a_0 + a_1 f^n(t))}{y(t)} \right|^2 dt \cdot \inf_{t \in [0,1]} \left| \frac{y(t)}{y^n(t)} \right|^2 \\ &= (1 - \varepsilon) \liminf_{f^n \rightarrow f} J(f^n|y), \end{aligned}$$

and the proposition holds.  $\square$

**Theorem 5.** Consider the family of processes  $(Y^n, Z^n)_{n \in \mathbb{N}}$ , where  $(Y^n)_{n \in \mathbb{N}}$  is a family of processes with paths in  $\mathcal{C}_\alpha([0, 1])$  and for  $n \in \mathbb{N}$ ,  $Z^n$  is the solution of (12). Suppose that  $(Y^n)_{n \in \mathbb{N}}$  is independent from the Brownian motion and satisfies a LDP with the good rate function  $I_Y$  and the speed  $\gamma(n) = n$ . Then  $(Y^n, Z^n)_{n \in \mathbb{N}}$  satisfies the WLDP with the speed  $\gamma(n) = n$  and rate function

$$I(y, z) = I_Y(y) + J(z|y),$$

and  $(Z^n)_{n \in \mathbb{N}}$  satisfies the LDP with the speed  $\gamma(n)$  and the rate function

$$I_Z(z) = \inf_{y \in \mathcal{C}_\alpha([0, 1])} \{I_Y(y) + J(z|y)\}.$$

**Proof.** The family of processes  $(Y^n, Z^n)_{n \in \mathbb{N}}$ , thanks to Remark 4 and Proposition 6, satisfies the hypotheses of Theorem 3, therefore the theorem holds.  $\square$

Now let  $(Z_t^{n,y})_{t \in [0, 1]}$  be a family of continuous Gaussian processes. We have to prove that  $(Z_t^{n,y^n})_{t \in [0, 1]}$  satisfies the same LDP as  $(Z_t^{n,y})_{t \in [0, 1]}$  when  $y^n \xrightarrow{\mathcal{C}_\alpha([0, 1])} y$ . Let  $\tilde{Z}_t^{n,y^n} = Z_t^{n,y^n} - m(t)$  for every  $n \in \mathbb{N}$  and  $t \in [0, 1]$ .

Straightforward calculations show that there exists  $L > 0$ , such that

$$\begin{aligned} & \sup_{s, t \in [0, 1], s \neq t} n \cdot \frac{|k^{y^n}(t, t) + k^{y^n}(s, s) - 2k^{y^n}(t, s)|}{|t - s|^{2\alpha}} \\ & \leq L \sup_{s, t \in [0, 1], s \neq t} \frac{(e^{a_1(t-s)} - 1)}{|t - s|^{2\alpha}} < +\infty, \quad \text{for } 2\alpha = 1. \end{aligned}$$

Therefore the family  $(\tilde{Z}^{n,y^n})_{n \in \mathbb{N}}$  is exponentially tight at the speed  $n$ . Furthermore, conditions (6) and (7) of Theorem 2 are fulfilled, in fact

$$\lim_{n \rightarrow +\infty} \mathbb{E}[\langle \lambda, \tilde{Z}^{n,y^n} \rangle] = 0$$

and

$$\lim_{n \rightarrow +\infty} \text{Var}(\langle \lambda, \tilde{Z}^{n,y^n} \rangle) \cdot n = \int_0^1 \int_0^1 k^y(s, t) d\lambda(t) d\lambda(s),$$

where  $k^y(s, t) = e^{a_1(s+t)} \int_0^{s \wedge t} e^{-2a_1 u} y^2(u) du$ . Therefore  $(\tilde{Z}^{n,y^n})_{n \in \mathbb{N}}$  satisfies a LDP on  $\mathcal{C}([0, 1])$ . Finally, thanks to the contraction principle, the family  $(Z^{n,y^n})_{n \in \mathbb{N}}$  satisfies a LDP on  $\mathcal{C}([0, 1])$  with the rate function  $J(\cdot|y)$  defined in (16).

**Remark 5.** The lower semicontinuity of  $J(\cdot|\cdot)$  on  $\mathcal{C}([0, 1]) \times \mathcal{C}_\alpha([0, 1])$  follows from Proposition 6.

We have proved that the hypotheses of Theorem 3 are verified, so the LDP for  $(Z^n)_{n \in \mathbb{N}}$  follows.

## 6 Estimates of level crossing probability

In this section we will study the probability of level crossing for a family of conditionally Gaussian processes. In particular, we will study the probability

$$p_n = \mathbb{P}\left(\sup_{0 \leq t \leq 1} (Z_t^n - \varphi(t)) > 1\right), \quad (17)$$

as  $n \rightarrow \infty$ , where  $(Z^n)_{n \in \mathbb{N}}$  is a family of conditionally Gaussian process. In this situation the probability  $p_n$  has a large deviation limit

$$\lim_{n \rightarrow \infty} \frac{1}{\gamma(n)} \log(p_n) = -I_\varphi.$$

The main reference in this section is [4]. We now compute  $\lim_{n \rightarrow \infty} \frac{1}{\gamma(n)} \log(p_n)$ , for a fixed continuous path  $\varphi \in \mathcal{C}([0, 1])$ . The computation is simple, in fact, since  $(Z^n)_{n \in \mathbb{N}}$  satisfies a LDP with the rate function

$$I_Z(z) = \inf_{y \in C} \{I_Y(y) + J(z|y)\}, \quad (18)$$

where  $I_Y(\cdot)$  is the rate function associated to the family of conditioning processes  $(Y^n)_{n \in \mathbb{N}}$ ,  $C$  is the Polish set where  $(Y^n)_{n \in \mathbb{N}}$  takes values, and  $J(\cdot|y)$  is the good rate function of the family of Gaussian processes  $(Z^{n,y})_{n \in \mathbb{N}}$ . If we denote

$$\mathcal{A} = \left\{w \in \mathcal{C}([0, 1]) : \sup_{0 \leq t \leq 1} (w(t) - \varphi(t)) > 1\right\},$$

we have that

$$-\inf_{w \in \mathcal{A}} I_Z(w) \leq \liminf_{n \rightarrow +\infty} \frac{1}{\gamma(n)} \log(p_n) \leq \limsup_{n \rightarrow +\infty} \frac{1}{\gamma(n)} \log(p_n) \leq -\inf_{w \in \mathcal{A}} I_Z(w)$$

where

$$\mathcal{A} = \left\{w \in \mathcal{C}([0, 1]) : \sup_{0 \leq t \leq 1} (w(t) - \varphi(t)) \geq 1\right\}$$

and

$$\mathcal{A} = \mathcal{A} = \left\{w \in \mathcal{C}([0, 1]) : \sup_{0 \leq t \leq 1} (w(t) - \varphi(t)) > 1\right\}.$$

It is a simple calculation to show that  $\inf_{w \in \mathcal{A}} I_Z(w) = \inf_{w \in \mathcal{A}} I_Z(w)$ . Therefore,

$$\lim_{n \rightarrow \infty} \frac{1}{\gamma(n)} \log(p_n) = -\inf_{w \in \mathcal{A}} I_Z(w). \quad (19)$$

For every  $t \in [0, 1]$  let  $\mathcal{A}_t = \{w \in \mathcal{C}([0, 1]) : w(t) = 1 + \varphi(t)\}$ , then  $\mathcal{A} = \bigcup_{0 \leq t \leq 1} \mathcal{A}_t$  and so

$$\inf_{w \in \mathcal{A}} I_Z(w) = \inf_{y \in C} \inf_{0 \leq t \leq 1} \inf_{w \in \mathcal{A}_t} \{I_Y(y) + J(w|y)\}.$$

### 6.1 Gaussian process with random mean and variance

For every  $n \in \mathbb{N}$ , let  $Z^n = X^n Y_1^n + Y_2^n$  as in Section 4. In this case we know that

$$J(z|y) = \begin{cases} \frac{1}{2} \left\| \frac{z-y_2}{y_1} \right\|_{\mathcal{H}}^2, & \frac{z-y_2}{y_1} \in \mathcal{H}, \\ +\infty & \text{otherwise.} \end{cases}$$

Therefore, we have

$$\begin{aligned} \inf_{w \in \mathcal{A}} I_Z(w) &= \inf_{y \in \mathcal{C}_\alpha([0,1]) \times \mathcal{C}([0,1])} \inf_{0 \leq t \leq 1} \inf_{w \in \mathcal{A}_t} \{I_Y(y) + J(w|y)\} \\ &= \inf_{y \in \mathcal{C}_\alpha([0,1]) \times \mathcal{C}([0,1])} \inf_{0 \leq t \leq 1} \inf_{w \in \mathcal{A}_t} \left\{ I_Y(y) + \frac{1}{2} \left\| \frac{w - y_2}{y_1} \right\|_{\mathcal{H}}^2 \right\}. \end{aligned}$$

The set of paths of the form

$$h(u) = \int_0^1 \bar{k}(u, v) d\lambda(v), \quad u \in [0, 1], \quad \lambda \in \mathcal{M}[0, 1],$$

is dense in  $\mathcal{H}$  and, therefore, the infimum  $\inf_{w \in \mathcal{A}_t} \{I_Y(y) + \frac{1}{2} \left\| \frac{w}{y} \right\|_{\mathcal{H}}^2\}$  is the same as that over the functions  $w$  such that

$$w(u) - y_2(u) = y_1(u) \cdot \int_0^1 \bar{k}(u, v) d\lambda(v), \quad u \in [0, 1],$$

for some  $\lambda \in \mathcal{M}[0, 1]$ . For such kind of paths, recalling the expression of their norms in the RKHS, the functional we aim to minimize is given by

$$I_Y(y) + \frac{1}{2} \left\| \frac{w - y_2}{y_1} \right\|_{\mathcal{H}}^2 = I_Y(y) + \frac{1}{2} \int_0^1 \int_0^1 \bar{k}(u, v) d\lambda(u) d\lambda(v)$$

therefore, it is enough to minimize the right-hand side of the above equation with respect to the measure  $\lambda$ , with the additional constraint that

$$w(t) = 1 + \varphi(t),$$

which we can write in the equivalent form

$$\int_0^1 \bar{k}(t, v) d\lambda(v) - \frac{1 + \varphi(t) - y_2(t)}{y_1(t)} = 0.$$

This is a constrained extremum problem, and thus we are led to use the method of Lagrange multipliers. The measure  $\lambda$  must be such that

$$\int_0^1 \int_0^1 \bar{k}(u, v) d\lambda(u) d\mu(v) = \beta \int_0^1 \bar{k}(t, v) d\mu(v), \quad \mu \in \mathcal{M}[0, 1],$$

for some  $\beta \in \mathbb{R}$ . We find

$$\beta = \frac{\int_0^1 \bar{k}(t, u) d\lambda(u)}{\bar{k}(t, t)} = \frac{1 + \varphi(t) - y_2(t)}{y_1(t) \bar{k}(t, t)}$$

and

$$\bar{\lambda} = \frac{1 + \varphi(t) - y_2(t)}{y_1(t)\bar{k}(t, t)}\delta_{\{t\}}.$$

Such measure satisfies the Lagrange multipliers problem, and it is therefore a critical point for the functional we want to minimize. Since this is a strictly convex functional restricted on a linear subspace of  $\mathcal{M}[0, 1]$ , it is still strictly convex, and thus the critical point  $\bar{\lambda}$  is actually its unique point of minimum. Hence, we have

$$\inf_{w \in \mathcal{A}} I_Z(w) = \inf_{y \in \mathcal{C}_\alpha([0, 1]) \times \mathcal{C}([0, 1])} \inf_{0 \leq t \leq 1} \left\{ I_Y(y) + \frac{(1 + \varphi(t) - y_2(t))^2}{2y_1^2(t)\bar{k}(t, t)} \right\}.$$

## 6.2 Ornstein–Uhlenbeck processes with random diffusion coefficient

In this case

$$\begin{aligned} J(f|y) &= \begin{cases} \frac{1}{2} \|f - m\|_{\mathcal{H}_y}^2, & f - m \in \mathcal{H}_y, \\ +\infty, & f - m \notin \mathcal{H}_y, \end{cases} \\ &= \begin{cases} \frac{1}{2} \int_0^1 \left| \frac{\dot{f}(t) - (a_0 + a_1 f(t))}{y(t)} \right|^2 dt, & f \in H_1^x, \\ +\infty, & f \notin H_1^x, \end{cases} \end{aligned}$$

where  $m(t) = e^{a_1 t} (x + \frac{a_0}{a_1} [1 - e^{-a_1 t}])$ ,  $t \in [0, 1]$ ,  $\mathcal{H}_y$  is the RKHS associated to

$$k^y(s, t) = e^{a_1(s+t)} \int_0^{s \wedge t} e^{-2a_1 u} y^2(u) du,$$

and  $H_1^x = m + \mathcal{H}_y$ .

We have

$$\begin{aligned} \inf_{w \in \mathcal{A}} I_Z(w) &= \inf_{y \in \mathcal{C}_\alpha[0, 1]} \inf_{0 \leq t \leq 1} \inf_{w \in \mathcal{A}_t} \{ I_Y(y) + J(w|y) \} \\ &= \inf_{y \in \mathcal{C}_\alpha[0, 1]} \inf_{0 \leq t \leq 1} \inf_{w \in \mathcal{A}_t} \left\{ I_Y(y) + \frac{1}{2} \|w - m\|_{\mathcal{H}_y}^2 \right\}. \end{aligned}$$

The set of paths of the form

$$h(u) = \int_0^1 k^y(u, v) d\lambda(v), \quad u \in [0, 1], \quad \lambda \in \mathcal{M}[0, 1],$$

is dense in  $\mathcal{H}_y$ , therefore, the infimum

$$\inf_{w \in \mathcal{A}_t} \left\{ I_Y(y) + \frac{1}{2} \|w - m\|_{\mathcal{H}_y}^2 \right\},$$

is the same as that over the functions of the form

$$w(u) = m(u) + \int_0^1 k^y(u, v) d\lambda(v), \quad u \in [0, 1],$$

for some  $\lambda \in \mathcal{M}[0, 1]$ . For paths of such kind, recalling the expression of their norms in the RKHS, the functional we aim to minimize is given by

$$I_Y(y) + \frac{1}{2} \|w - m\|_{\mathcal{H}_y}^2 = I_Y(y) + \frac{1}{2} \int_0^1 \int_0^1 k^y(u, v) d\lambda(u) d\lambda(v)$$

therefore, it is enough to minimize the right-hand side of the above equation with respect to the measure  $\lambda$ , with the additional constraint

$$w(t) = 1 + \varphi(t),$$

which can be written in the equivalent form

$$\int_0^1 k^y(t, v) d\lambda(v) + m(t) - (1 + \varphi(t)) = 0.$$

This is a constrained extremum problem, and thus we are led to use the method of Lagrange multipliers. We find

$$\beta = \frac{\int_0^1 k^y(t, u) d\lambda(u)}{k^y(t, t)} = \frac{1 + \varphi(t) - m(t)}{k^y(t, t)}$$

and

$$\bar{\lambda} = \frac{1 + \varphi(t) - m(t)}{k^y(t, t)} \delta_{\{t\}},$$

$\delta_{\{t\}}$  standing for the Dirac mass in  $t$ . Such measure satisfies the Lagrange multipliers problem, and it is therefore a critical point for the functional we want to minimize. Since this functional is a strictly convex restricted on a linear subspace of  $\mathcal{M}[0, 1]$ , it is still strictly convex, and thus the critical point  $\bar{\lambda}$  is actually its unique point of minimum. Hence, we have

$$\inf_{w \in \mathcal{A}_t} I_Y(y) + \frac{1}{2} \|w - m\|_{\mathcal{H}_y}^2 = I_Y(y) + \frac{(1 + \varphi(t) - m(t))^2}{2k^y(t, t)},$$

and therefore

$$\inf_{w \in \mathcal{A}} I_Z(w) = \inf_{y \in \mathcal{C}_\alpha} \inf_{0 \leq t \leq 1} \left\{ I_Y(y) + \frac{(1 + \varphi(t) - m(t))^2}{2k^y(t, t)} \right\}.$$

## References

- [1] Adler, R.J., Samorodnitsky, G., Gadrich, T.: The expected number of level crossings for stationary, harmonizable, symmetric, stable processes. *Ann. Appl. Probab.* **3**, 553–575 (1993) [MR1221165](#)
- [2] Azencott, R.: Grande déviations et applications. In: *École D'été de Probabilités de St. Flour VIII*, L.N.M. Volume 774, (1980) [MR0590626](#)
- [3] Baldi, P.: *Stochastic Calculus*. Springer (2017) [MR3726894](#). <https://doi.org/10.1007/978-3-319-62226-2>

- [4] Baldi, P., Pacchiarotti, B.: Explicit computation of second order moments of importance sampling estimators for fractional Brownian motion. *Bernoulli* **12**(4), 663–688 (2006) [MR2248232](#). <https://doi.org/10.3150/bj/1155735931>
- [5] Berline, A., Thomas-Agnan, C.: *Reproducing Kernel Hilbert Spaces in Probability and Statistics*. Kluwer Academic Publishers (2004) [MR2239907](#). <https://doi.org/10.1007/978-1-4419-9096-9>
- [6] Borkar, S.V.: *Probability Theory*. Springer (1995) [MR1367959](#). <https://doi.org/10.1007/978-1-4612-0791-7>
- [7] Caramellino, L., Pacchiarotti, B., Salvadei, S.: Large deviation approaches for the numerical computation of the hitting probability for Gaussian processes. *Methodol. Comput. Appl. Probab.* **17**(2), 383–401 (2015) [MR3343412](#). <https://doi.org/10.1007/s11009-013-9364-5>
- [8] Chaganty, N.R.: Large deviations for joint distributions and statistical applications. *Sankhya* **59**(2), 147–166 (1997) [MR1665683](#)
- [9] Chiarini, A., Fischer, M.: On large deviations for small noise Itô processes. *Adv. Appl. Probab.* **46**(4), 1126–1147 (2014) [MR3290432](#). <https://doi.org/10.1239/aap/1418396246>
- [10] Dembo, A., Zeitouni, O.: *Large Deviations Techniques and Applications*. Jones and Bartlett, Boston, Ma (1998) [MR1202429](#)
- [11] Deuschel, J.D., Stroock, D.W.: *Large Deviations*. Academic Press, Boston, MA (1989) [MR0997938](#)
- [12] Doucet, A., de Freitas, N., Murphy, K., Russel, S.: Rao-Blackwellized particle filtering for dynamic Bayesian networks. In: *16th Conference on Uncertainty in AI*, pp. 176–183 (2000)
- [13] Giorgi, F., Pacchiarotti, B.: Large deviations for conditional Volterra processes. *Stoch. Anal. Appl.* **35**(2), 191–210 (2017) [MR3597612](#). <https://doi.org/10.1080/07362994.2016.1237291>
- [14] Gulisashvili, A.: Large deviation principle for Volterra type fractional stochastic volatility models. <https://arxiv.org/abs/1710.10711> [MR3858803](#). <https://doi.org/10.1137/17M116344X>
- [15] Hida, T., Hitsuda, M.: *Gaussian Processes*. AMS Translations (1993) [MR1216518](#)
- [16] Lototsky, S.V.: Optimal filtering of stochastic parabolic equations. In: *Recent Development in Stochastic Analysis and Related Topics, Proceedings of the First Sino-German Conference on Stochastic Analysis (A Satellite Conference of ICM 2002)*, pp. 330–353 (2002)
- [17] Macci, C., Pacchiarotti, B.: Exponential tightness for Gaussian processes with applications to some sequences of weighted means. *Stochastic* **89**(2), 469–484 (2017) [MR3590430](#). <https://doi.org/10.1080/17442508.2016.1248968>