Universités de Paris 6 & Paris 7 - CNRS (UMR 7599)

PRÉPUBLICATIONS DU LABORATOIRE

DE PROBABILITÉS & MODÈLES ALÉATOIRES

4, place Jussieu - Case 188 - 75 252 Paris cedex 05

http://www.proba.jussieu.fr

Importance sampling for the ruin problem for general Gaussian processes

P. BALDI & B. PACCHIAROTTI

JANVIER 2004

Prépublication n^o 875

P. Baldi : Laboratoire de Probabilités et Modèles Aléatoires, CNRS-UMR 7599, Université Paris VI & Université Paris VII, 4 place Jussieu, Case 188, F-75252 Paris Cedex 05 & Dipartimento di Matematica, Università di Roma Tor Vergata, Via della Ricerca Scientifica 1. I-00133 Roma.

B. Pacchiarotti : Dipartimento di Matematica, Università di Roma Tor Vergata, Via della Ricerca Scientifica 1. I-00133 Roma.

Importance Sampling for the Ruin Problem for general gaussian Processes

Paolo Baldi * and Barbara Pacchiarotti *

Università di Roma Tor Vergata

Abstract

We study a family of importance sampling estimators for the problem of computing the probability of level crossing when the crossing level is large, or when the intensity of the noise is small. We give general results concerning centered gaussian processes with drift and develop a method which allows to compute explicitly the asymptotics of the second order moment, with a special mention for the fractional Brownian Motion case. The main tools are some refined versions of classical large deviations results and, for the fractional Brownian Motion, recent results on changes of probability by Norros et al.

AMS 1991 subject classifications. Primary 60F10; secondary 60J60, 60G15, 65C50.

Key words and phrases. Importance Sampling; Large Deviations; Ruin Probabilities; gaussian Processes; fractional Brownian Motion.

Aknowledgement The authors wish to thank Th.Jeulin for some illuminating remarks.

1 Introduction

A classical problem in risk theory is the investigation of the ruin probability

$$\mathrm{P}(\sup_{t>0}(X_t - \varphi_t) > B) \tag{1.1}$$

where $X = (X_t)_{t\geq 0}$ is a continuous centered process starting at 0; under suitable assumptions on the function φ (typically $\varphi_t \to +\infty$ as $t \to +\infty$) the probability above is very small. Thus its computation by simulation requires a large number of iterations in order to get a reasonable precision and becomes impossible in practice. A natural technique to tackle this question is the *importance sampling*, that is the simulation of the process under a new probability Q, for which the event considered is not rare any more, and to compensate by taking into account the density of P with respect to Q.

*Research performed within the project *Processi Stocastici e applicazioni a Filtraggio, Controllo, Simulazione e Finanza Matematica*, Ministero dell'Università e della Ricerca Scientifica 2001

In this paper we investigate the existence of good importance sampling distributions for this problem when the process $(X_t)_t$ is a fractional Brownian Motion, but some of our results hold for a general class of gaussian processes. We extend here the techniques developed in [1], where the case of the ordinary Brownian Motion was investigated. The fractional Brownian Motion case has been already investigated in Michna [8]; a comparison of the results is given in section §4.

The starting point is to reduce the problem to the determination of the good importance sampling distributions for the computation of the probability

$$p_\varepsilon \stackrel{def}{=} \mathrm{P}(\sup_{0\le t\le 1}(\varepsilon X_t - \varphi_t) > 1) \tag{1.2}$$

as $\varepsilon \to 0$. To this goal we consider, for the computation of the level crossing probability (1.2), the class of admissible importance distributions estimators of the form

$$1_{\{\tau_\varepsilon \le 1\}} Z_{\tau_\varepsilon}$$

where the process $(X_t)_t$ is simulated according to a distribution Q_ε, obtained by translating P with a path $\frac{1}{\varepsilon}\gamma$, γ belonging to the Reproducing Kernel Hilbert Space of the process $(X_t)_t$. Here τ_ε is the time at which the level 1 is attained (possibly $\tau_\varepsilon = +\infty$) and Z_{τ_ε} is the density of P with respect to Q_ε up to time τ_ε. For the ruin problem (1.1) the class of admissible importance sampling estimators is defined in a similar way, see §4.

The plan of the paper is as follows.

In §2 we investigate the asymptotics of the probability (1.2) as $\varepsilon \to 0$. These results are well known, but we develop a technique which is possibly new. This technique is also the main tool for the investigation of the importance sampling distributions for the estimation of p_ε, which is the object of §3. It should be pointed out that the results in §2 and §3 hold under general assumptions: $X = (X_t)_t$ needs to be a gaussian process, continuous, centered and starting at 0. We must also make an assumption on X (see Assumption 3.1), which is satisfied for many gaussian processes of interest (including the fractional Brownian Motion).

In §4 the results of §2 and §3 are used to study the importance sampling distributions for the ruin problem for the fractional Brownian Motion and for φ of the form $\varphi_t = ct^\alpha$, $c > 0$. We prove that an asymptotically efficient distribution *does not* exist (at least in the class of importance sampling distributions considered). This kind of phenomena (non existence of asymptotically efficient changes of probability, and in particular lack of efficiency by translating with the most likely path) have already been pointed out in the literature, see [6] and [13]. We are however able to compute the asymptotics of the variance of the estimator above (which, of course, depends on the choice of the path γ). Also it turns out that, al least if $H > \frac{1}{2}$, it is possible to find translations which are very near to asymptotic efficiency. Even if this means that the precision goes to zero exponentially fast as the level B goes to $+\infty$, it is natural to enquire whether these estimators are useful in practice and to determine the translation which produces the estimator having the smallest variance. These questions are

of a very different nature (the first one requires the performance of numerical experiments, the second one the resolution of a particular analytical problem) and will be addressed in future work.

The results described so far require to overcome a certain number of technical questions. These are discussed in §5, where also some well known facts about gaussian measures are recalled.

2 Estimates of level crossing

Let $X = (X_t)_{t\geq 0}$ be a *continuous centered* gaussian process starting at 0 and let $K(t,s) = \mathrm{Cov}(X_t, X_s)$ be its covariance function.

In this section we compute the limit $\lim_{\varepsilon\to 0} \varepsilon^2 \log p_\varepsilon$ for a fixed continuous path $\varphi \in \mathcal{C} = \mathcal{C}([0,1],\mathbb{R})$ (the space of continuous paths vanishing at 0, endowed with the uniform norm). We assume that the underlying probability space is $\Omega = \mathcal{C}$ and that $X = (X_t)_t$ is the canonical process, that is $X_t(\omega) = \omega_t$, $\omega \in \Omega$. We define, as usual, the filtration $(\mathcal{F}_t)_t$ by $\mathcal{F}_t = \sigma(X_u, u \leq t)$.

The computation above is actually quite simple as the process $t \to \varepsilon X_t - \varphi_t$ satisfies a Large Deviation Principle (see §5), so that, if we denote by I its rate function,

$$\lim_{\varepsilon\to 0} \varepsilon^2 \log p_\varepsilon = -\inf_{w\in\mathcal{A}} I(w), \tag{2.1}$$

where $\mathcal{A}$ denotes the set of paths $\{w \in \mathcal{C}; \sup_{0\leq t\leq 1} w_t - \varphi_t \geq 1\}$, I being the rate function of the process $(X_t)_{0\leq t\leq 1}$, that is $I(w) = \frac{1}{2}|w|^2_{\mathcal{H}}$ if w belongs to the Reproducing Kernel Hilbert Space (RKHS) $\mathcal{H}$ (see §5 again) and $I(w) = +\infty$ otherwise.

Proposition 2.1 *The infimum in (2.1) is equal to* $\inf_{0<t\leq 1} \frac{(1+\varphi_t)^2}{2K(t,t)}$, *i.e.,*

$$\lim_{\varepsilon\to 0} \varepsilon^2 \log \mathrm{P}\Big(\sup_{0\leq t\leq 1} (\varepsilon X_t - \varphi_t) \geq 1\Big) = -\inf_{0<t\leq 1} \frac{(1+\varphi_t)^2}{2K(t,t)} \stackrel{def}{=} -i_\varphi \tag{2.2}$$

Proof. Let $\mathcal{A}_t = \{w \in \mathcal{C}; w_t = 1 + \varphi_t\}$, so that $\mathcal{A} = \{w \in \mathcal{C}; \sup_{0\leq t\leq 1} w_t - \varphi_t \geq 1\} = \cup_{0\leq t\leq 1} \mathcal{A}_t$, and

$$\inf_{w\in\mathcal{A}} I(w) = \inf_{0\leq t\leq 1} \inf_{w\in\mathcal{A}_t} I(w).$$

Recall that the dual $\mathcal{C}'$ of $\mathcal{C}$ is given by the vector space of the measures with finite variation on $[0,1]$. As recalled in §5, a set of paths which is dense in $\mathcal{H}$ is the one which is formed by those paths which are barycenters of the r.v.'s belonging to $\mathcal{C}'$. That is

$$w_t = \mathrm{E}\Big(X_t \int_0^1 X_s \, d\alpha(s)\Big) = \int_0^1 K(t,s)\, d\alpha(s) \tag{2.3}$$

The square of the norm, $|w|^2_{\mathcal{H}}$, of w in $\mathcal{H}$ is given by the variance of the centered gaussian r.v. $\int_0^1 X_t \, d\alpha(t)$, that is

$$|w|^2_{\mathcal{H}} = \int_0^1 \int_0^1 K(u,v) d\alpha(u) d\alpha(v) \stackrel{def}{=} V(\alpha).$$

Recalling that $I(w) = \frac{1}{2}|w|^2_{\mathcal{H}}$,

$$\inf_{w\in\mathcal{A}_t} I(w) = \inf V(\alpha) = \inf \frac{1}{2}\int_0^1\int_0^1 K(u,v)d\alpha(u)d\alpha(v),$$

the infimum in the right hand side being taken among all $\alpha \in \mathcal{C}'$ such that

$$\int_0^1 K(t,s)\, d\alpha(s) = 1 + \varphi_t.$$

This is a constrained extremum problem and is easily handled by Lagrange multipliers: we need to find a $\alpha \in \mathcal{C}'$ such that

$$0 = V'(\alpha)\beta + \lambda\int_0^1 K(t,v)d\beta(v) = \int_0^1 K(v,u)d\alpha(v)d\beta(u) + \lambda\int_0^1 K(t,u)d\beta(u)$$

for some $\lambda \in \mathbb{R}$ and for every $\beta \in \mathcal{C}'$. This gives the equation

$$\int_0^1 K(u,v)d\alpha(v) + \lambda K(t,u) = 0, \qquad \text{for every } u, 0 \le u \le 1.$$

The solution is obviously

$$\alpha = \alpha_0 = -\lambda\delta_t$$

where λ is determined by the condition

$$w_t = -\lambda K(t,t) - \varphi_t = 1$$

that is

$$-\lambda = \frac{1+\varphi_t}{K(t,t)}, \qquad \alpha_0 = \frac{1+\varphi_t}{K(t,t)}\delta_t.$$

Therefore

$$\inf_{w\in A_t} I(w) = \frac{1}{2}\int_0^t\int_0^t K(u,v)d\alpha_0(u)d\alpha_0(v) = \frac{1}{2}\frac{(1+\varphi_t)^2}{K(t,t)}.$$

Remark that the minimizing path $w \in \mathcal{A}_t$ is

$$w_s = \frac{1+\varphi_t}{K(t,t)}K(t,s). \tag{2.4}$$

∎

Remark 2.2 Actually in the previous proof we assumed that in order to compute the infimum of I over $\mathcal{A}_t$ it is sufficient to consider only the paths w for which the representation (2.3) holds whereas these are only dense in the RKHS $\mathcal{H}$. Actually the proof of the previous Proposition 2.1 is completed as soon as we show that the paths of the form (2.3) and such that $w_t = 1 + \varphi_t$ are dense (in the topology of $\mathcal{H}$) in $\mathcal{H} \cap \{w_t = 1 + \varphi_t\}$. But this is immediate because the map $w \to w_t$ is continuous on $\mathcal{H}$: thus if $w \in \mathcal{H} \cap \{w_t = 1 + \varphi_t\}$, then there exists a sequence $(w_n)_n \subset \mathcal{H}$ of paths of the form (2.3) such that $w_n \to^{\mathcal{H}} w$. This implies that $w_n(t) \to w_t = 1 + \varphi_t$, so that if $\tilde{w}_n = w_n(1 + \varphi_t)/w_n(t)$, then still $w_n \to^{\mathcal{H}} w$ and $(\tilde{w}_n)_n \subset \mathcal{H} \cap \{w_t = 1 + \varphi_t\}$.

Remark 2.3 Remark that, since φ is supposed to be continuous, the infimum $\inf_{0<t\leq 1} \frac{(1+\varphi_t)^2}{2K(t,t)}$ is attained at some point t^*; therefore, whatever the path φ, the minimizing path w is of the form const $\cdot K(t^*, \cdot)$, for some $t^* \in]0,1]$.

Remark 2.4 (The fractional Brownian Motion) Assume that the process $(X_t)_t$ is a fractional Brownian Motion, that is

$$K(t,s) = \frac{1}{2}(t^{2H} + s^{2H} - |t-s|^{2H})$$

and thus $K(t,t) = t^{2H}$. Assume also that $\varphi_t = kt^\alpha$, $k > 0$. Then the infimum in (2.1) becomes

$$\inf_{0<t\leq 1} \frac{1}{2} \frac{(1+kt^\alpha)^2}{t^{2H}}.$$

The derivative is

$$\frac{1}{2t^{4H}} \left[2k\alpha t^{\alpha+2H-1}(1+kt^\alpha) - 2Ht^{2H-1}(1+kt^\alpha)^2\right]$$

which has the same sign as

$$k\alpha t^\alpha - H(1+kt^\alpha) = k(\alpha - H)t^\alpha - H.$$

If $\alpha \leq H$, this quantity is strictly negative for every $t > 0$. If $\alpha > H$, it is negative until

$$t^* = \Big(\frac{H}{k(\alpha - H)}\Big)^{1/\alpha} \tag{2.5}$$

and then positive. Thus, if the value in (2.5) is ≤ 1, then the infimum (minimum actually) is attained at t^* and is equal to

$$\frac{1}{2}\Big(\frac{\alpha}{\alpha - H}\Big)^2 \Big(\frac{k(\alpha - H)}{H}\Big)^{2H/\alpha}. \tag{2.6}$$

Let us define a^* so that $a^* K(t^*, t^*) = 1 + \varphi_{t^*}$, that is

$$a^* = \frac{1 + kt^{*\alpha}}{t^{*2H}} = \alpha\Big(\frac{k}{H}\Big)^{2H/\alpha} (\alpha - H)^{-1+\frac{2H}{\alpha}},$$

then the path at which the infimum is attained is

$$w_s = a^* K\big(t^*, s\big). \tag{2.7}$$

If the value t^* in (2.5) is larger than 1, then the infimum is attained at $t = 1$.

The fact that the most likely path is as in (2.7) was already obtained in [10], but with very different approach.

3 Importance sampling

In this section we investigate the importance sampling distributions for the computation of the level crossing probability (1.2). We keep notations and settings of §2. Let $\gamma \in \mathcal{H}$ and Z the corresponding r.v. in the space $\mathcal{C}'_{\mathrm{P}}$ (see §5 for details). Thus, by Girsanov theorem (see again §5 for details), with respect to the probability $d\mathrm{Q}_\varepsilon = \mathrm{e}^{\frac{1}{\varepsilon}Z - \frac{1}{2\varepsilon^2}|\gamma|^2_{\mathcal{H}}} d\mathrm{P}$, the process εX has the same distribution as $t \to \varepsilon X_t + \gamma_t$ under P. Let $\varphi : [0,1] \to \mathbb{R}$ be a continuous function and define the stopping times

$$\tau(\omega) = \inf\{t; \omega_t - \varphi_t \geq 1\}, \qquad \tau_\varepsilon(\omega) = \tau(\varepsilon\omega)$$

with the usual understanding $\inf \emptyset = +\infty$. In the following we denote E and $\mathrm{E}^{\mathrm{Q}_\varepsilon}$ the expectations taken with respect to P and Q_ε respectively. Let $Z^\varepsilon_t = \mathrm{E}^{\mathrm{Q}_\varepsilon}(\mathrm{e}^{\frac{1}{\varepsilon}Z - \frac{1}{2\varepsilon^2}|\gamma|^2_{\mathcal{H}}} \,|\, \mathcal{F}_t)$, so that

$$Z^\varepsilon_t = \frac{d\mathrm{P}}{d\mathrm{Q}_\varepsilon}\Big|_{\mathcal{F}_t}.$$

The r.v.

$$1_{\{\tau_\varepsilon \leq 1\}} Z^\varepsilon_{\tau_\varepsilon} \tag{3.1}$$

is, under Q_ε, an unbiased estimator of $p_\varepsilon = \mathrm{P}(\sup_{t\leq 1}(\varepsilon X_t - \varphi_t) > 1)$. We investigate the existence of good importance sampling estimators in the class described by (3.1). We say, as usually, that the family of probabilities $(\mathrm{Q}_\varepsilon)_\varepsilon$ is *asymptotically efficient* if for the variances $\sigma^2_{\mathrm{Q}_\varepsilon}$ of these estimators it holds

$$\limsup_{\varepsilon\to 0} \varepsilon^2 \log \sigma^2_{\mathrm{Q}_\varepsilon} = -2i_\varphi \tag{3.2}$$

(i_φ was defined in (2.1)). By extension we shall also say that the translation γ is asymptotically efficient. The meaning of this definition is clear: the precision of the estimator can be defined as the ratio between the quantity p_ε to be estimated and the amplitude of the confidence interval; thus the precision is proportional to the ratio between p_ε and $\sigma_{\mathrm{Q}_\varepsilon}$. The condition of asymptotic efficiency thus ensures that the precision *does not go to* 0 *exponentially fast* as $\varepsilon \to 0$.

Since the mean of the estimator (3.1) is equal to p_ε, thanks to Proposition 2.1 in order to have asymptotic efficiency the second order moment $m_2(\varepsilon) = \mathrm{E}^{\mathrm{Q}_\varepsilon}(1_{\{\tau_\varepsilon\leq 1\}}(Z^\varepsilon_{\tau_\varepsilon})^2)$ must satisfy

$$\limsup_{\varepsilon\to 0} \varepsilon^2 \log m_2(\varepsilon) = -2i_\varphi \tag{3.3}$$

Remark that in (3.3) the left hand side is always equal to or greater than the right hand side, as the second order moment cannot be smaller than the square of the mean.

In order to get a reasonable representation of the martingale $(Z^\varepsilon_t)_t$, remark that, if

$$\zeta_t = \mathrm{E}(Z \,|\, \mathcal{F}_t)$$

then $(\zeta_t)_t$ is a gaussian P-martingale (the σ-field $\mathcal{F}_t$ is generated by $(X_s)_{s\le t}$ and $((X_s)_{s\le t}, Z)$ is a gaussian family). Hence it has independent increments and

$$\tilde{Z}_t^\varepsilon = \mathrm{e}^{\frac{1}{\varepsilon}\zeta_t - \frac{1}{2\varepsilon^2}v_t}, \quad v_t = \mathrm{E}(\zeta_t^2)$$

is also a P-martingale; as $\tilde{Z}_1^\varepsilon = \mathrm{e}^{\frac{1}{\varepsilon}Z - \frac{1}{2\varepsilon^2}|\gamma|^2_{\mathcal{H}}}$, then

$$\tilde{Z}_t^\varepsilon = \mathrm{E}(\mathrm{e}^{\frac{1}{\varepsilon}Z - \frac{1}{2\varepsilon^2}|\gamma|^2_{\mathcal{H}}} \,|\, \mathcal{F}_t).$$

Since

$$\tilde{Z}_t^\varepsilon = \frac{d\mathrm{Q}_\varepsilon}{d\mathrm{P}}\Big|_{\mathcal{F}_t},$$

it follows

$$(\tilde{Z}_t^\varepsilon)^{-1} = \frac{d\mathrm{P}}{d\mathrm{Q}_\varepsilon}\Big|_{\mathcal{F}_t} = Z_t^\varepsilon$$

and therefore

$$Z_t^\varepsilon = \mathrm{e}^{-\frac{1}{\varepsilon}\zeta_t + \frac{1}{2\varepsilon^2}v_t}.$$

The second order moment of the estimator (3.1) is therefore

$$m_2(\varepsilon) = \mathrm{E}^{\mathrm{Q}_\varepsilon}(1_{\{\tau_\varepsilon\le 1\}}(Z_{\tau_\varepsilon}^\varepsilon)^2) = \mathrm{E}^{\mathrm{P}}(1_{\{\tau_\varepsilon\le 1\}}Z_{\tau_\varepsilon}^\varepsilon) = \mathrm{E}^{\mathrm{P}}(1_{\{\tau_\varepsilon\le 1\}}\mathrm{e}^{-\frac{1}{\varepsilon}\zeta_{\tau_\varepsilon}+\frac{1}{2\varepsilon^2}v_{\tau_\varepsilon}}) =$$
$$= \int_{\{\tau_\varepsilon\le 1\}} \mathrm{e}^{-\frac{1}{\varepsilon}\zeta_{\tau_\varepsilon}(\omega)+\frac{1}{2\varepsilon^2}v_{\tau_\varepsilon(\omega)}}\, d\mathrm{P}(\omega) = \int 1_{\{\tau(\varepsilon\omega)\le 1\}}\mathrm{e}^{\frac{1}{\varepsilon^2}(-\zeta_\tau(\varepsilon\omega)+\frac{1}{2}v_{\tau(\varepsilon\omega)})}\, d\mathrm{P}(\omega).$$

(we used the fact that ζ_t is a linear function of ω, see §5).

Assumption 3.1 *We say that the probability* P *satisfies the property CGM (continuity of the gaussian martingales) if, for every r.v. $Z \in \mathcal{C}'_{\mathrm{P}}$, the gaussian martingale $(\zeta_t)_t$ admits a continuous version.*

The results of this section hold under Assumption 3.1. It is easy to check that this assumption is satisfied for many gaussian processes of interest and in particular for the Brownian Motion (of course) and the fractional Brownian Motion. It is possible to find examples of continuous gaussian processes that do not satisfy Assumption 3.1.

In the following we shall consider that these martingales are continuous.

In §5 we prove that, under Assumption 3.1 the functional

$$-\zeta_\tau + \tfrac{1}{2}v_\tau \tag{3.4}$$

satisfies the assumptions of Varadhan's Lemma (Theorem 5.3). This gives

$$\lim_{\varepsilon\to 0}\varepsilon^2\log m_2(\varepsilon) = \lim_{\varepsilon\to 0}\varepsilon^2\log \mathrm{E}^{\mathrm{P}}(1_{\{\tau_\varepsilon\le 1\}}\mathrm{e}^{-\frac{1}{\varepsilon}\zeta_{\tau_\varepsilon}+\frac{1}{2\varepsilon^2}v_{\tau_\varepsilon}}) =$$
$$= \sup_{\omega,\tau(\omega)\le 1}\{-\zeta_\tau(w)+\tfrac{1}{2}v_\tau - I(w)\} = \sup_{t\le 1}\sup_{w\in\mathcal{H};w_t-\varphi_t=1}\{-\zeta_t(w)+\tfrac{1}{2}v_t - \tfrac{1}{2}|w|^2_{\mathcal{H}}\} =$$
$$= -\inf_{t\le 1}\inf_{w\in\mathcal{H};w_t-\varphi_t=1}\{\zeta_t(w) - \tfrac{1}{2}v_t + \tfrac{1}{2}|w|^2_{\mathcal{H}}\}.$$

Proposition 3.2 *Let $t \in [0,1]$. It holds*

$$\inf_{w\in\mathcal{H};w_t-\varphi_t=1}\{\zeta_t(w) - \tfrac{1}{2}v_t + \tfrac{1}{2}|w|^2_{\mathcal{H}}\} = -v_t + \frac{1}{2K(t,t)}\,(1+\varphi_t+\gamma_t)^2.$$

Proof. $\zeta_t = \mathrm{E}(Z\,|\,\mathcal{F}_t)$ is a gaussian r.v. and the corresponding path in the RKHS is $\gamma^{(t)}_s = \mathrm{E}(X_s\zeta_t)$. If $w \in \mathcal{H}$, $\zeta_t(w) = \langle w, \gamma^{(t)}\rangle_{\mathcal{H}}$ (see §5, (5.6)). Thus we are looking for the minimum of the function $F(w) = \langle w, \gamma^{(t)}\rangle_{\mathcal{H}} + \frac{1}{2}|w|^2_{\mathcal{H}}$ on the hyperplane $w_t = 1 + \varphi_t$. This being, again, a constrained extremum, we use Lagrange multipliers. Since F is convex and tends to $+\infty$ as $|w|_{\mathcal{H}} \to +\infty$, the minimum exists and is certainly a critical point.

For $h \in \mathcal{H}$, $DF(w)h = \langle \gamma^{(t)}, h\rangle_{\mathcal{H}} + \langle w, h\rangle_{\mathcal{H}}$ and, if $G(w) = w_t$, $DG(w)h = h_t$. Thus the Lagrange multipliers relation for a critical point is

$$\langle \gamma^{(t)}, h\rangle_{\mathcal{H}} + \langle w, h\rangle_{\mathcal{H}} + \lambda h_t = 0.$$

The only vector $y \in \mathcal{H}$ such that $\langle y, h\rangle_{\mathcal{H}} = \lambda h_t$ for every $h \in \mathcal{H}$ is $y = K(t,\cdot)$. Thus the relation $\langle \gamma^{(t)} + w, h\rangle_{\mathcal{H}} = -\lambda h_t$ for every $h \in \mathcal{H}$ implies $\gamma^{(t)} + w = -\lambda K(t,\cdot)$, that is $w_s = -\gamma^{(t)}_s - \lambda K(t,s)$. The constraint $w_t = 1 + \varphi_t$ implies

$$-\lambda = \frac{1+\varphi_t+\gamma_t}{K(t,t)}$$

(we use the fact that $\gamma^{(t)}_t = \gamma_t$). Thus, the critical path $\tilde{w}$ is

$$\tilde{w}_s = -\gamma^{(t)}_s + \frac{1+\varphi_t+\gamma_t}{K(t,t)}\,K(t,s).$$

Replacing the value of $\tilde{w}$ and using the relationships

$$\langle \gamma^{(t)}, \gamma^{(t)}\rangle_{\mathcal{H}} = \mathrm{E}(\zeta_t^2) = v_t, \quad \langle \gamma^{(t)}, K(t,\cdot)\rangle_{\mathcal{H}} = \gamma_t, \quad \langle K(t,\cdot), K(t,\cdot)\rangle_{\mathcal{H}} = K(t,t)$$

one gets

$$\begin{aligned}
&\zeta_t(\tilde{w}) = \langle \tilde{w}, \gamma^{(t)}\rangle_{\mathcal{H}} = -v_t + \gamma_t\,\frac{1+\varphi_t+\gamma_t}{K(t,t)},\\
&|\tilde{w}|^2_{\mathcal{H}} = v_t + 2\lambda\gamma_t + \lambda^2 K(t,t) = v_t + 2\gamma_t\,\frac{1+\varphi_t+\gamma_t}{K(t,t)} + \frac{(1+\varphi_t+\gamma_t)^2}{K(t,t)}
\end{aligned}$$

and, finally,

$$\begin{aligned}
\inf_{w\in\mathcal{H};w_t-\varphi_t=1}\{\zeta_t(w) - \tfrac{1}{2}v_t + \tfrac{1}{2}|w|^2_{\mathcal{H}}\} &= \zeta_t(\tilde{w}) - \tfrac{1}{2}v_t + \tfrac{1}{2}|\tilde{w}|^2_{\mathcal{H}} =\\
&= -v_t + \frac{1}{2K(t,t)}\,(1+\varphi_t+\gamma_t)^2.
\end{aligned}$$

■

Let us define

$$H_\gamma(t) := -v_t + \frac{1}{2K(t,t)}\,(1+\varphi_t+\gamma_t)^2. \tag{3.5}$$

We shall call H_γ *the master function* associated to the translation γ. If we define

$$\rho(\gamma) = \inf_{0\le t\le 1} H_\gamma(t), \tag{3.6}$$

since, by Varadhan's Lemma,

$$\lim_{\varepsilon\to 0} \varepsilon \log m_2(\varepsilon) = -\inf_{t\le 1}\Big(-v_t + \frac{1}{2K(t,t)}\,(1+\varphi_t+\gamma_t)^2\Big) = -\rho(\gamma), \tag{3.7}$$

then γ is asymptotically efficient if and only if $\rho(\gamma) = 2i_\varphi$. The next statement gives information that is useful in order to determine whether $\gamma \in \mathcal{H}$ is asymptotically efficient and whether asymptotically efficient translations do exist.

Proposition 3.3 *It holds, for every $\gamma \in \mathcal{H}$ and $0 < t \le 1$,*

$$H_\gamma(t) \le \frac{(1+\varphi_t)^2}{K(t,t)}. \tag{3.8}$$

Moreover, if γ is asymptotically efficient, the infimum in (3.6) is attained at every $t^* \in \operatorname{argmin}_{0\le s\le 1} \frac{(1+\varphi_t)^2}{2K(t,t)}$ and γ must be of the form $\gamma_s = a^* K(t^*, s)$ for $s \le t^*$, where $a^* = (1+\varphi_{t^*})/K(t^*,t^*)$.

Proof. One can write, for every t,

$$\begin{aligned} H_\gamma(t) &= -v_t + \frac{1}{2K(t,t)}\,(1+\varphi_t+\gamma_t)^2 = \\ &= -v_t + \frac{1}{2K(t,t)}\,\gamma_t^2 + \frac{(1+\varphi_t)^2}{2K(t,t)} + \frac{1}{K(t,t)}\,\gamma_t(1+\varphi_t) = \\ &-\Big(v_t - \frac{1}{K(t,t)}\,\gamma_t^2\Big) - \frac{1}{2K(t,t)}\,(1+\varphi_t-\gamma_t)^2 + \frac{(1+\varphi_t)^2}{K(t,t)}. \end{aligned} \tag{3.9}$$

If Z is the r.v. associated to γ in $\mathcal{C}'_{\mathrm{P}}$, then $\gamma_t = \mathrm{E}(X_t Z) = \mathrm{E}(X_t \mathrm{E}(Z \,|\, \mathcal{F}_t))$, so that

$$\gamma_t^2 \le \mathrm{E}(\mathrm{E}(Z \,|\, \mathcal{F}_t)^2)\mathrm{E}(X_t^2) = v_t K(t,t) \tag{3.10}$$

which implies

$$v_t - \frac{1}{K(t,t)}\,\gamma_t^2 \ge 0.$$

Therefore the first two terms in the left hand side of (3.9) are both negative, which proves (3.8). Rewriting (3.9) for $t = t^*$, one gets

$$H_\gamma(t^*) = -\Big(v_{t^*} - \frac{1}{K(t^*,t^*)}\,\gamma_{t^*}^2\Big) - \frac{1}{2K(t^*,t^*)}\,(1+\varphi_{t^*}-\gamma_{t^*})^2 + \underbrace{\frac{(1+\varphi_{t^*})^2}{K(t^*,t^*)}}_{=2i_\varphi}.$$

Therefore a necessary condition for γ to be asymptotically efficient is that both the relations

$$v_{t^*} - \frac{1}{K(t^*,t^*)}\,\gamma_{t^*}^2 = 0, \qquad 1 + \varphi_{t^*} - \gamma_{t^*} = 0$$

hold. The second inequality implies $\gamma_{t^*} = 1 + \varphi_{t^*}$. As for the first one, a closer look to (3.10) shows that the inequality is strict unless the r.v.'s $\mathrm{E}(Z \,|\, \mathcal{F}_{t^*})$ and X_{t^*} are co-linear, that is unless $\gamma_{t^*} = \lambda \mathrm{E}(X_{t^*}^2) = \lambda K(t^*,t^*)$. The condition $\gamma_{t^*} = 1 + \varphi_{t^*}$ implies $\lambda = \frac{1+\varphi_{t^*}}{K(t^*,t^*)} = a^*$. Thus, if $s \le t^*$,

$$\gamma_s = \mathrm{E}(X_s Z) = \mathrm{E}[X_s \mathrm{E}(Z \mid \mathcal{F}_{t^*})] = a^* \mathrm{E}(X_s X_{t^*}) = a^* K(t^*, s)$$

∎

Proposition 3.3 gives only a necessary condition for asymptotic efficiency and an asymptotically efficient translation may not exist, as already pointed out in [1], §3. In [1] one can also find a discussion about existence of asymptotically efficient importance sampling distributions, when the process $(\mathcal{C}, (X_t)_t, \mathrm{P})$ is a Brownian Motion. The next example takes care of a much more complicated situation.

Example 3.4 (Fractional Brownian Motion) Assume now that the process $(\mathcal{C}, (X_t)_{t\le 1}, \mathrm{P})$ is a fractional Brownian Motion. That is a continuous centered gaussian process with covariance function $K(t,s) = \frac{1}{2}(t^{2H} + s^{2H} - |t-s|^{2H})$. The RKHS of this process is well described using the stochastic integral, whose main features we recall here. For more details see Norros et al [9] e.g.

Assume first $H > \frac{1}{2}$ and consider the vector space, Λ_H^2 of the functions $f : [0,1] \to \mathbb{R}$ such that

$$H(2H-1)\int_0^1 \int_0^1 f(u) f(v) |u-v|^{2H-2}\, du dv < +\infty$$

endowed with the obvious scalar product. For every $f \in \Lambda_H^2$ one can define the stochastic integral

$$\int_0^1 f(u)\, dX_u \tag{3.11}$$

first imposing that

$$\int_0^1 f(u)\, dX_u = X_b - X_a$$

if $f = 1_{[a,b]}$, $0 \le a \le b \le 1$, and then extending it to an isometry between Λ_H^2 and $L^2(\mathrm{P})$. In particular it holds

$$\mathrm{E}\Big(\int_0^1 f(u)\, dX_u \int_0^1 g(v)\, dX_v\Big) = H(2H-1)\int_0^1 \int_0^1 f(u) g(v) |u-v|^{2H-2}\, du dv.$$

If $H < \frac{1}{2}$, one can also define the stochastic integral with respect to the fractional Brownian Motion, but in a different way, owing to lack of integrability properties of the function $(u,v) \to |u-v|^{2H-2}$. Again see [9] for details. It is

clear that the r.v.'s of the form (3.11) are gaussian and belong to the gaussian space $\mathcal{C}'_P$.

Norros et al [9] proved that, if

$$w(t,s) = \begin{cases} c_1 s^{\frac{1}{2}-H}(t-s)^{\frac{1}{2}-H} & \text{if } 0 \le s < t, \\ 0 & \text{if } s \ge t. \end{cases} \tag{3.12}$$

where

$$c_1 = \frac{1}{2H\,\Gamma(H+\frac{1}{2})\Gamma(\frac{3}{2}-H)},$$

then

$$M_t = \int_0^t w(t,s)\,dX_s$$

is a gaussian martingale (the *fundamental martingale*) such that $\mathrm{E}(M_t^2) = c_2^2 t^{2-2H}$ where

$$c_2^2 = \frac{\Gamma(\frac{3}{2}-H)}{2H(2-2H)\,\Gamma(H+\frac{1}{2})\Gamma(2-2H)}. \tag{3.13}$$

Figure 1 gives the graph of c_2^2 as a function of H, its behaviour being relevant in the sequel. Norros & al. prove two important properties of the fundamental

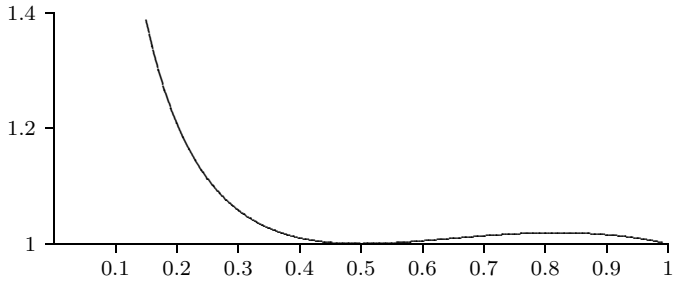

Figure 1 Behaviour of c_2^2 as H varies. For $H > \frac{1}{2}$, the values remain very close to 1.

martingale, namely, first, the representation formula

$$X_t = \int_0^t \tilde{z}(t,s)\,dM_s \tag{3.14}$$

where the integral is an ordinary stochastic integral with respect to the continuous square integrable martingale $(M_t)_t$ and

$$\tilde{z}(t,s) = 2H\Big(t^{H-\frac{1}{2}}(t-s)^{H-\frac{1}{2}} - (H-\tfrac{1}{2})\int_s^t u^{H-\frac{3}{2}}(u-s)^{H-\frac{1}{2}}\,du\Big). \tag{3.15}$$

If $H > \frac{1}{2}$ this expression can be slightly simplified by integrating by parts, which gives

$$\tilde{z}(t,s) = 2H(H-\tfrac{1}{2})\int_s^t u^{H-\frac{1}{2}}(u-s)^{H-\frac{3}{2}}\,du. \tag{3.16}$$

This is Theorem 5.2 and formula (5.1) of [9]. As a consequence, the fundamental martingale generates the filtration $(\mathcal{F}_t)_t$ and every r.v. in the gaussian space $\mathcal{C}'_{\mathrm{P}}$ can be written in the form

$$\int_0^1 g(s)\, dM_s \tag{3.17}$$

as g varies in the Hilbert space L^2_H of the functions such that $\int_0^1 g(s)^2\, d\langle M\rangle_s = c_2^2(2H-2)\int_0^1 g(s)^2 s^{1-2H}\, ds < +\infty$ (whereas, if $H > \frac{1}{2}$, not every r.v. of $\mathcal{C}'_{\mathrm{P}}$ admits the representation (3.11), see [11]).

The representation formula (3.17) implies that property CGM of Assumption 3.1 is satisfied for the fractional Brownian Motion. Actually, if $Z \in \mathcal{C}'_{\mathrm{P}}$, then it is of the form (3.17), so that

$$\zeta_t = \mathrm{E}(Z \,|\, \mathcal{F}_t) = \int_0^t g(s)\, dM_s.$$

The second important feature of the fundamental martingale is that

$$\mathrm{E}(X_t M_1) = t. \tag{3.18}$$

This implies that the path $\gamma_t = t$ belongs to the RKHS and that $|\gamma|^2_{\mathcal{H}} = c_2^2$.

Let us now show that an asymptotically efficient importance sampling distribution cannot exist. We consider first the case $H > \frac{1}{2}$. Proposition 3.3 states that the corresponding translation γ should be of the form $\gamma_t = a^* K(t^*, t) = a^* \mathrm{E}(X_t X_{t^*})$, where t^* is a point at which the infimum in (2.2) is attained. We prove now that the associated master function H_γ has a minimum that is strictly smaller than $2i_\varphi$. This will follow from the fact that $H'_\gamma(t^*-) = 0$, $H''_\gamma(t^*-) = -\infty$. This implies that H_γ is increasing in a left neighbourhood of t^*, and therefore has a minimum that is strictly smaller than $H_\gamma(t^*) = 2i_\phi$.

The term v_t appearing in the master function is, for $t \le t^*$,

$$v_t = a^{*2}\mathrm{E}(\mathrm{E}(Z_{t^*} \,|\, \mathcal{F}_t)^2) = a^{*2}(2-2H)c_2^2 \int_0^t \tilde{z}(t^*, s)^2 s^{1-2H}\, ds.$$

Therefore

$$\begin{aligned} v'_t &= a^{*2}(2-2H)c_2^2 \tilde{z}(t^*, t)^2 t^{1-2H} \\ v''_t &= a^{*2}(2-2H)c_2^2 \Big(2t^{1-2H} \frac{\partial \tilde{z}(t^*, t)}{\partial t} \tilde{z}(t^*, t) + (1-2H)\tilde{z}(t^*, t)^2 t^{-2H} \Big) \end{aligned}$$

The derivative of $\tilde{z}(t^*, t)$ is

$$\frac{\partial \tilde{z}(t^*, t)}{\partial t} = 2H(H-\tfrac{1}{2})\Big(-t^{*H-\frac{1}{2}}(t^*-t)^{H-\frac{3}{2}} + (H-\tfrac{1}{2}) \int_t^{t^*} u^{H-\frac{3}{2}}(u-t)^{H-\frac{3}{2}}\, du \Big).$$

The integral inside the parenthesis is of order $(t^*-t)^{H-\frac{1}{2}}$, thus

$$\frac{\partial \tilde{z}(t^*, t)}{\partial t} \simeq -2H(H-\tfrac{1}{2})\, t^{*H-\frac{1}{2}}(t^*-t)^{H-\frac{3}{2}}$$

as $t \uparrow t^*$. Since $\tilde{z}(t^*,t) \simeq 2Ht^{*H-\frac{1}{2}}(t^*-t)^{H-\frac{1}{2}}$ and

$$\tilde{z}(t^*,t)^2 \simeq 4H^2 t^{*2H-1}(t^*-t)^{2H-1},$$

replacing one finds $v'_{t^*-} = 0$ (recall that $H > \frac{1}{2}$). Also

$$\frac{\partial \tilde{z}(t^*,t)}{\partial t}\tilde{z}(t^*,t) \simeq -4H^2(H-\tfrac{1}{2})t^{*2H-1}(t^*-t)^{2H-2}$$

and, finally,

$$v''_t \simeq -a^{*2}(2-2H)8H^2(H-\tfrac{1}{2})c_2^2(t^*-t)^{2H-2}.$$

As for the other terms appearing in the master function, since $\gamma_t = \frac{1}{2}a^*(t^{2H} + t^{*2H} - (t^*-t)^{2H})$,

$$\begin{aligned}\gamma'_t &= a^*H(t^{2H-1} + (t^*-t)^{2H-1}),\\ \gamma''_t &= a^*H(2H-1)(t^{2H-2} - (t^*-t)^{2H-2}).\end{aligned}$$

Thus $\gamma'_{t^*-} = a^*Ht^{*2H-1}$, whereas $\gamma''_t \simeq -a^*H(2H-1)(t^*-t)^{2H-2}$ as $t \uparrow t^*$. Since

$$H'_\gamma(t) = -v'_t + \frac{(1+\varphi_t+\gamma_t)[(\varphi'_t+\gamma'_t)t - H(1+\varphi_t+\gamma_t)]}{t^{2H+1}}$$

one gets $H'_\gamma(t^*-) = 0$, as $\gamma'_{t^*} = \varphi'_{t^*} = \frac{H}{t^*}(1+\varphi_{t^*})$. As for the second order derivative, denoting by $\dots$ the terms that have a finite limit in t^*,

$$H''_\gamma(t) = -v''_t + \frac{(1+\varphi_t+\gamma_t)}{t^{2H}}\,\gamma''_t + \dots \simeq$$
$$\simeq (t^*-t)^{2H-2}\Big(a^{*2}(2-2H)8H^2(H-\tfrac{1}{2})c_2^2 - a^*H(2H-1)\,\frac{(1+\varphi_{t^*}+\gamma_{t^*})}{t^{*2H}}\Big).$$

Recalling that $(1+\varphi_{t^*}+\gamma_{t^*})t^{*-2H} = 2a^*$ and replacing c_2^2 with its expression (see (3.13)), we find

$$\begin{aligned}H''_\gamma(t) &\simeq Ha^{*2}(t^*-t)^{2H-2}\Big((2-2H)8H(H-\tfrac{1}{2})c_2^2 + 2(2H-1)\Big) =\\ &= 2(2H-1)Ha^{*2}(t^*-t)^{2H-2}\underbrace{\Big(\frac{\Gamma(\frac{3}{2}-H)}{\Gamma(H+\frac{1}{2})\Gamma(2-2H)} - 1\Big)}_{:=K_H}.\end{aligned}$$

The quantity K_H is strictly negative for every $H \neq \frac{1}{2}$ (see Figure 2). Therefore $H''_\gamma(t^*-) = -\infty$ and the master function is increasing in a left neighbourhood of t^* (see also Figure 6).

Similar computations show that, if $H < \frac{1}{2}$, the derivative of the master function, H_γ, of $\gamma_t = a^*K(t^*,t)$, converges to $+\infty$ as $t \uparrow t^*$. Thus, also in this case, an asymptotically efficient translation cannot exist.

Remark 3.5 The computations of the this section raise two questions.
• Concerning the fractional Brownian Motion, there is no importance sampling distribution which is asymptotically efficient *in the class considered*. Is there

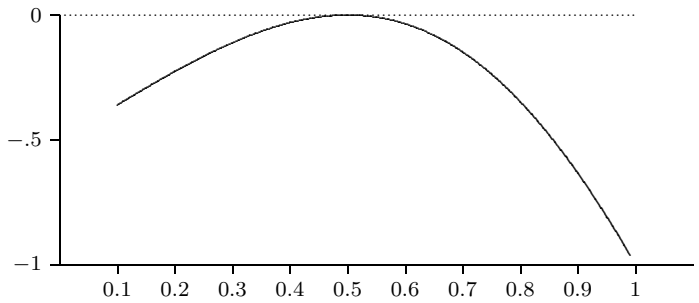

Figure 2 Behaviour of the discrepancy function K_H as H varies. K_H is strictly negative unless $H = \frac{1}{2}$.

a reasonable, larger, class of importance sampling distributions in which an asymptotically efficient one can be found?

• As for now, if an asymptotically efficient importance sampling distribution does not exist, is it possible to find a i.s. distribution that is optimal in the class considered (even if it fails to be asymptotically efficient)? In other words, if we denote $J(\gamma) = \inf_{0<t\leq 1} H_\gamma(t)$, does the functional J attain a maximum as γ varies in $\mathcal{H}$? If this is the case, which are these maxima?

Remark 3.6 The use of Varadhan's Lemma gives an intuition of the reason why the translation using the most likely path given by the large deviations estimates of §2 is not asymptotically efficient: the main contribution to the asymptotics of the second order moment is given by paths which give a very small contribution to the probability $\mathrm{P}(\sup_{0\leq t\leq 1}(\varepsilon X_t - \varphi_t) > 1)$, as $\varepsilon \to 0$, but for which the weight $\mathrm{e}^{-\frac{1}{\varepsilon}\zeta_{\tau_\varepsilon}+\frac{1}{\varepsilon^2}v_{\tau_\varepsilon}}$ is very large.

4 The ruin probability

In this section we apply the results of §2 and §3 to give an estimate of the behaviour of the ruin probability (1.1) as $B \to +\infty$ and to the determination of good (or not bad, at least...) importance sampling distributions. Whereas the results of §2 and §3 hold true for general gaussian processes (albeit centered and continuous and satisfying Assumption 3.1), in this section we deal with the fractional Brownian Motion. Let $\tilde{X} = (\tilde{X}_s)_{s\geq 0}$ be a fractional Brownian Motion and $c > 0$, $\alpha > 0$. We want to investigate the behaviour, as $B \to +\infty$, of

$$\mathrm{P}\Big(\sup_{s\geq 0}(\tilde{X}_s - cs^\alpha) > B\Big) \tag{4.1}$$

where $\tau_B = \inf\{s > 0; \tilde{X}_s - cs^\alpha > B\}$, with the understanding $\tau_B = +\infty$ if the above set of times is empty. Let d be a positive number and let us first study the behaviour as $B \to +\infty$ of

$$\mathrm{P}\Big(\sup_{0\leq s\leq (Bd)^{1/\alpha}}(\tilde{X}_s - cs^\alpha) > B\Big).$$

As we shall see, this term gives the main contribution to the ruin probability (4.1), provided d is large enough. Remark that,

$$\mathrm{P}\Big(\sup_{0\le s\le (Bd)^{1/\alpha}}(\tilde X_s - cs^\alpha) > B\Big) = \mathrm{P}\Big(\sup_{0\le t\le 1}(\tilde X_{t(Bd)^{1/\alpha}} - c((Bd)^{1/\alpha}t)^\alpha) > B\Big) =$$
$$\mathrm{P}\Big(\sup_{0\le t\le 1}((Bd)^{\frac{H}{\alpha}} X_t - cBd\,t^\alpha) > B\Big) = \mathrm{P}\Big(\sup_{0\le t\le 1}(B^{-1+\frac{H}{\alpha}} d^{\frac{H}{\alpha}} X_t - cdt^\alpha) > 1\Big).$$

Here we set $X_t = (Bd)^{-\frac{H}{\alpha}}\tilde X_{t(Bd)^{1/\alpha}}$. By scaling, $X = (X_t)_t$ is again a fractional Brownian Motion with Hurst exponent equal to H. If $\alpha > H$ and we set $\varepsilon = B^{-1+\frac{H}{\alpha}} d^{\frac{H}{\alpha}}$ and $k = cd$, the problem is reduced to the estimation of the probability

$$p_\varepsilon = \mathrm{P}\Big(\sup_{0\le t\le 1}(\varepsilon X_t - kt^\alpha) > 1\Big),$$

which was already considered in the previous sections. Let

$$s^* = \Big(\frac{H}{c(\alpha - H)}\Big)^{1/\alpha}$$

and assume that d is such that $d > s^{*\alpha}$. Then

$$\begin{aligned}
&\lim_{B\to+\infty} B^{-2+\frac{2H}{\alpha}} \log \mathrm{P}\Big(\sup_{0\le s\le (Bd)^{1/\alpha}}(\tilde X_s - cs^\alpha) > B\Big) = \\
&d^{-\frac{2H}{\alpha}} \lim_{B\to+\infty} B^{-2+\frac{2H}{\alpha}} d^{\frac{2H}{\alpha}} \log \mathrm{P}\Big(\sup_{0\le s\le (Bd)^{1/\alpha}}(\tilde X_s - cs^\alpha) > B\Big) = \\
&= d^{-\frac{2H}{\alpha}} \lim_{\varepsilon\to 0} \varepsilon^2 \log \mathrm{P}\Big(\sup_{0\le t\le 1}(\varepsilon X_t - kt^\alpha) > 1\Big) = \\
&= d^{-\frac{2H}{\alpha}} \inf_{0\le t\le 1} \frac{(1+kt^\alpha)^2}{2t^{2H}} \overset{def}{=} -\tilde i.
\end{aligned} \tag{4.2}$$

With the assumptions made on d, it holds

$$t^* = \Big(\frac{H}{k(\alpha - H)}\Big)^{1/\alpha} = \Big(\frac{H}{cd(\alpha - H)}\Big)^{1/\alpha} < 1.$$

Thus the critical point t^* belongs to $]0,1[$ and, thanks to (2.6),

$$\tilde i = d^{-\frac{2H}{\alpha}} \frac{1}{2}\Big(\frac{\alpha}{\alpha - H}\Big)^2 \Big(\frac{k(\alpha - H)}{H}\Big)^{\frac{2H}{\alpha}} = \frac{1}{2}\Big(\frac{\alpha}{\alpha - H}\Big)^2 \Big(\frac{c(\alpha - H)}{H}\Big)^{\frac{2H}{\alpha}} \tag{4.3}$$

which does not depend on d (provided that $d > s^{*\alpha}$). This term gives the main contribution to the ruin probability $\mathrm{P}(\sup_{s\ge 0}(\tilde X_s - cs^\alpha) > B)$. Indeed it is sufficient to show that

$$\limsup_{B\to+\infty} B^{-2+\frac{2H}{\alpha}} \log \mathrm{P}\Big(\sup_{s>(Bd)^{1/\alpha}}(\tilde X_s - cs^\alpha) > B\Big) < -\tilde i$$

if $d > s^{*\alpha}$. This is a rather straightforward application of Fernique's theorem, as outlined in [5] (see also [1] for more details).

We now describe how the results of the previous section can be used in order to derive information concerning the second order moment of importance

sampling estimators. If $\tilde\gamma \in \mathcal{H}$, then $\tilde\gamma_t = \mathrm{E}(\tilde Z \tilde X_t)$, where $\tilde Z$ belongs to the gaussian space generated by the r.v.'s $(\tilde X_t)_t$. As pointed out in Example 3.4, a nice way of describing this gaussian space is provided by the fundamental martingale of Norros & al. [9]. Let

$$\tilde M_t = \int_0^t w(t,s)\, d\tilde X_s$$

where w is as in (3.12); then every r.v. belonging to the gaussian space generated by $(\tilde X_t)_{t\le (Bd)^{1/\alpha}}$ is of the form

$$\tilde Z = \int_0^{(Bd)^{1/\alpha}} \tilde g(s)\, d\tilde M_s$$

for some function $\tilde g$ such that $\int_0^{(Bd)^{1/\alpha}} \tilde g(s)^2 d\langle \tilde M\rangle_s < +\infty$. We denote $\tilde L^2_H$ the Hilbert space of these functions, endowed with the scalar product

$$\langle \tilde g, \tilde h\rangle_{\tilde{\mathcal{H}}} = \int_0^{(Bd)^{1/\alpha}} \tilde g(s)\tilde h(s)\, d\langle \tilde M\rangle_s = (2H-2)c_2^2 \int_0^{(Bd)^{1/\alpha}} \tilde g(s)\tilde h(s) s^{1-2H}\, ds.$$

To every function $\tilde g \in \tilde L^2_H$ one can therefore associate a path $\tilde\gamma \in \mathcal{H}$ by setting

$$\begin{aligned}\tilde\gamma_s = \mathrm{E}\Big(\tilde X_s \int_0^{(Bd)^{1/\alpha}} \tilde g(u)\, d\tilde M_u\Big) = \mathrm{E}\Big(\tilde X_s \int_0^{s} \tilde g(u)\, d\tilde M_u\Big) = \\ = \int_0^s \tilde z(s,u)\tilde g(u)\, d\langle \tilde M\rangle_u.\end{aligned} \tag{4.4}$$

Moreover the relation (4.4) between $\tilde L^2_H$ and $\mathcal{H}$ is one to one.

Lemma 4.1 *Let X be a fractional Brownian Motion and $\tilde X_t = \lambda^{-H} X_{\lambda t}$, so that $\tilde X$ is still a fractional Brownian Motion. If $g \in L^2_H([0,T])$, then it holds*

$$\int_0^T g(s)\, dX_s = \lambda^H \int_0^{T/\lambda} g(\lambda s)\, d\tilde X_s \qquad \text{a.s.}$$

Proof. All the r.v.'s considered belong to the gaussian space generated by $(X_t)_t$. Thus it is sufficient to prove that the two r.v.'s above have the same covariance with respect to each of the r.v.'s $X_t, t>0$. Indeed, with the change of variable $u=\lambda r$, $v=\lambda s$,

$$\begin{aligned}\mathrm{E}\Big(X_t \int_0^T g(s)\, dX_s\Big) &= H(2H-1)\int_0^T\int_0^T 1_{[0,t]}(u) g(v) |u-v|^{2H-2}\, du dv = \\ &= \lambda^2 H(2H-1)\int_0^{T/\lambda}\int_0^{T/\lambda} 1_{[0,t]}(\lambda r) g(\lambda s) |\lambda r - \lambda s|^{2H-2}\, dr ds = \\ &= \lambda^{2H} H(2H-1)\int_0^{T/\lambda}\int_0^{T/\lambda} 1_{[0,t]}(\lambda r) g(\lambda s) |r-s|^{2H-2}\, dr ds\end{aligned}$$

whereas

$$\lambda^H \mathrm{E}\Big(X_t \int_0^{T/\lambda} g(\lambda s)\, d\tilde{X}_s\Big) = \lambda^{2H} \mathrm{E}\Big(\tilde{X}_{t/\lambda} \int_0^{T/\lambda} g(\lambda s)\, d\tilde{X}_s\Big) =$$
$$\lambda^{2H} H(2H-1) \int_0^{T/\lambda} \int^{T/\lambda} 1_{[0,t/\lambda]}(u) g(\lambda v) |u-v|^{2H-2} =$$
$$= \lambda^{2H} H(2H-1) \int_0^{T/\lambda} \int_0^{T/\lambda} 1_{[0,t]}(\lambda u) g(\lambda v) |u-v|^{2H-2}\, du dv.$$

■

As a consequence of the previous lemma, let us look at the transformation leading the fundamental martingale of X to the one of $\tilde{X}$. These are defined by

$$M_t = \int_0^t w(t,s)\, dX_s, \qquad \tilde{M}_t = \int_0^t w(t,s)\, d\tilde{X}_s.$$

Thus

$$M_t = \int_0^t w(t,s)\, dX_s = \lambda^H \int_0^{t/\lambda} w(t,\lambda s)\, d\tilde{X}_s =$$
$$= c_1 \lambda^H \int_0^{t/\lambda} (\lambda s)^{\frac{1}{2}-H} (t-\lambda s)^{\frac{1}{2}-H}\, d\tilde{X}_s = c_1 \lambda^{1-H} \int_0^{t/\lambda} s^{\frac{1}{2}-H} (\tfrac{t}{\lambda} - s)^{\frac{1}{2}-H}\, d\tilde{X}_s =$$
$$= \lambda^{1-H} \tilde{M}_{t/\lambda}.$$

This relation is easily extended to

$$\int_0^t g(s)\, dM_s = \lambda^{1-H} \int_0^{t/\lambda} g(\lambda s)\, d\tilde{M}_s \tag{4.5}$$

for every function $g \in L^2_H$.

If $X_t = (Bd)^{-\frac{H}{\alpha}} \tilde{X}_{t(Bd)^{1/\alpha}}$ and $k = cd$, let us define as before

$$\tau_\varepsilon = \inf\{t; \varepsilon X_t - k t^\alpha > 1\}$$

with $\varepsilon = B^{-1+\frac{H}{\alpha}} d^{\frac{H}{\alpha}}$, $k = cd$. Then

$$\tau_\varepsilon = \inf\{t; B^{-1+\frac{H}{\alpha}} d^{\frac{H}{\alpha}} X_t - k t^\alpha > 1\} = \inf\{t; (Bd)^{\frac{H}{\alpha}} X_t - kB t^\alpha > B\} =$$
$$= \inf\{s; \tilde{X}_{s(Bd)^{1/\alpha}} - kB s^\alpha > B\} = \inf\{s; \tilde{X}_{s(Bd)^{1/\alpha}} - \tfrac{k}{d}((Bd)^{1/\alpha} s)^\alpha > B\} =$$
$$= (Bd)^{-1/\alpha} \inf\{s; \tilde{X}_s - c\, s^\alpha > B\} = (Bd)^{-1/\alpha} \tilde{\tau}_B.$$

Moreover, thanks to formula (4.5) with $\lambda = (Bd)^{-1/\alpha}$,

$$\int_0^{\tau_\varepsilon} g(t)\, dM_t = \int_0^{(Bd)^{-1/\alpha}\tilde{\tau}_B} g(t)\, dM_t = (Bd)^{-\frac{1}{\alpha}(1-H)} \int_0^{\tilde{\tau}_B} g(t(Bd)^{-1/\alpha})\, d\tilde{M}_t$$

whereas

$$\int_0^{\tau_\varepsilon} g(t)^2\, d\langle M\rangle_t = (2-2H) c_2^2 \int_0^{\tau_\varepsilon} g(t)^2 t^{1-2H}\, dt =$$
$$= (2-2H)(Bd)^{-\frac{1}{\alpha}(1-2H+1)} c_2^2 \int_0^{\tilde{\tau}_B} g(s(Bd)^{-1/\alpha})^2 s^{1-2H}\, ds =$$
$$= (Bd)^{-\frac{1}{\alpha}(2-2H)} \int_0^{\tilde{\tau}_B} g(s(Bd)^{-1/\alpha})^2\, d\langle \tilde{M}\rangle_s.$$

Therefore, recalling that $\varepsilon = B^{-1+\frac{H}{\alpha}} d^{\frac{H}{\alpha}}$,

$$-\frac{1}{\varepsilon}\int_0^{\tau_\varepsilon} g(t)\,dM_t + \frac{1}{2\varepsilon^2}\int_0^{\tau_\varepsilon} g(t)^2\,d\langle M\rangle_t =$$

$$= -B^{1-\frac{H}{\alpha}} d^{-\frac{H}{\alpha}} (Bd)^{-\frac{1}{\alpha}(1-H)} \int_0^{\tilde\tau_B} g(s(Bd)^{-1/\alpha})\,d\tilde M_s +$$

$$+\frac{1}{2} B^{2-\frac{2H}{\alpha}} d^{-\frac{2H}{\alpha}} (Bd)^{-\frac{1}{\alpha}(2-2H)} \int_0^{\tilde\tau_B} g(s(Bd)^{-1/\alpha})^2\,d\langle\tilde M\rangle_s =$$

$$= -B^{1-\frac{1}{\alpha}} d^{-\frac{1}{\alpha}} \int_0^{\tilde\tau_B} g(s(Bd)^{-1/\alpha})\,d\tilde M_s + \frac{1}{2} B^{2-\frac{2}{\alpha}} d^{-\frac{2}{\alpha}} \int_0^{\tilde\tau_B} g(s(Bd)^{-1/\alpha})^2\,d\langle\tilde M\rangle_s.$$

Thus, if we define $\tilde g_B(s) = B^{1-\frac{1}{\alpha}} d^{-\frac{1}{\alpha}} g(s(Bd)^{-1/\alpha})$, the two r.v.'s

$$\begin{aligned} &1_{\{\tau_\varepsilon\le 1\}} \exp\Big(-\frac{1}{\varepsilon}\int_0^{\tau_\varepsilon} g(t)\,dM_t + \frac{1}{2\varepsilon^2}\int_0^{\tau_\varepsilon} g(t)^2\,d\langle M\rangle_t\Big) \\ &1_{\{\tilde\tau_B\le (Bd)^{1/\alpha}\}} \exp\Big(-\int_0^{\tilde\tau_B} \tilde g_B(s)\,d\tilde M_s + \frac{1}{2}\int_0^{\tilde\tau_B} g_B(s)^2\,d\langle\tilde M\rangle_s\Big) \end{aligned} \tag{4.6}$$

are P-a.s. equal.

The relationship between the paths γ and $\tilde\gamma_B$, associated to the integrands g and $\tilde g_B$ respectively can be made explicit. Actually

$$\tilde\gamma_B(s) = \mathrm{E}\Big[\tilde X_s \int_0^{(Bd)^{1/\alpha}} \tilde g_B(s)\,d\tilde M_s\Big] = B^{1-\frac{1}{\alpha}} d^{-\frac{1}{\alpha}} \mathrm{E}\Big[\tilde X_s \int_0^{(Bd)^{1/\alpha}} \tilde g(s(Bd)^{-1/\alpha})\,d\tilde M_s\Big]$$

But

$$\tilde X_s = (Bd)^{\frac{H}{\alpha}} X_{s(Bd)^{-1/\alpha}}$$

$$\int_0^{(Bd)^{1/\alpha}} \tilde g_B(s)\,d\tilde M_s = \int_0^{(Bd)^{1/\alpha}} \tilde g(s(Bd)^{-1/\alpha})\,d\tilde M_s = (Bd)^{\frac{1}{\alpha}(1-H)} \int_0^1 g(s)\,dM_s$$

so that

$$\tilde\gamma_B(s) = B\mathrm{E}\Big[X_{s(Bd)^{-1/\alpha}} \int_0^1 g(s)\,dM_s\Big] = B\gamma(s(Bd)^{-1/\alpha}) \tag{4.7}$$

We consider the family of importance sampling estimators for the ruin probability of the form

$$1_{\{\tilde\tau_B\le (Bd)^{1/\alpha}\}} \tilde Z_{\tilde\tau_B} \tag{4.8}$$

where $\tilde\tau_B = \inf\{s; \tilde X_s - c\,s^\alpha \ge B\}$, the r.v.'s $\tilde X_s$ are simulated under the distribution, $\tilde{\mathrm{Q}}_B$, of a fractional Brownian Motion translated with a path $\tilde\gamma_B$ of the form (4.7) for some $\gamma \in \mathcal{H}$ and $(Z_t)_t$ is the martingale (under $\tilde{\mathrm{Q}}_B$) that gives the density of the change of probability, that is

$$\tilde Z_t = \exp\Big(-\int_0^t \tilde g_B(s)\,dM_s + \frac{1}{2}\int_0^t \tilde g_B(s)^2\,d\langle M\rangle_s\Big).$$

Here the function $\tilde g_B \in \tilde L^2_H$ represents the path $\tilde\gamma_B$ as in (4.4). The second order moment of this estimator is

$$\tilde m_2(B) = \mathrm{E}\Big[1_{\{\tilde\tau_B\le (Bd)^{1/\alpha}\}} \exp\Big(-\int_0^t \tilde g_B(s)\,dM_s + \frac{1}{2}\int_0^t \tilde g_B(s)^2\,d\langle M\rangle_s\Big)\Big].$$

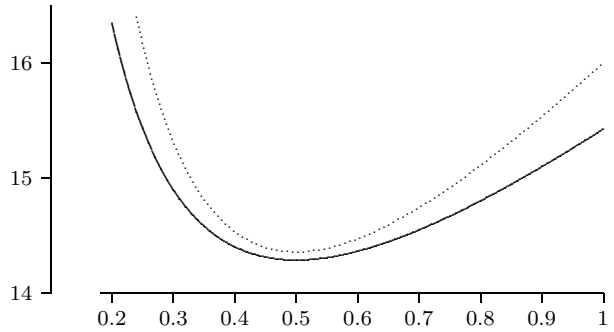

Figure 3 The dotted line is the graph of $t \to (1+kt)^2 t^{-2H}$ whereas the solid one is the graph of H_γ, with $\gamma_t = at$. It is apparent that the infimum of the latter is strictly smaller. Thus the translation is not asymptotically efficient. Here $H = 0.6$, $k = 3$, $a = \frac{1+kt^*}{(2c_2^2-1)t^*} = 4.95$; the infimum of H_γ is equal to 14.2854 instead of 14.3588 for asymptotic efficiency.

Similarly to the definition given in §3, we say that the estimators defined through a family of paths $(\tilde{\gamma}_B)_B$ is *asymptotically efficient* if

$$\limsup_{B\to+\infty} B^{-2+\frac{2H}{\alpha}} \log \tilde{m}_2(B) = -2\tilde{i}.$$

Thanks to (4.6), it holds

$$\begin{aligned} &\tilde{m}_2(B) = \\ &= \mathrm{E}\Big[1_{\{\tilde{\tau}_B \le (Bd)^{1/\alpha}\}} \exp\Big(-\int_0^{\tilde{\tau}_B} \tilde{g}_B(s)\, d\tilde{M}_s + \frac{1}{2}\int_0^{\tilde{\tau}_B} g_B(s)^2\, d\langle \tilde{M}\rangle_s\Big)\Big] = \\ &= \mathrm{E}\Big[1_{\{\tau_\varepsilon \le 1\}} \exp\Big(-\frac{1}{\varepsilon}\int_0^{\tau_\varepsilon} g(t)\, dM_t + \frac{1}{2\varepsilon^2}\int_0^{\tau_\varepsilon} g(t)^2\, d\langle M\rangle_t\Big)\Big] = \\ &= m_2(\varepsilon) \end{aligned} \tag{4.9}$$

and

$$\lim_{B\to\infty} B^{2-\frac{2H}{\alpha}} \log \tilde{m}_2(B) = d^{\frac{2H}{\alpha}} \lim_{\varepsilon\to 0} \varepsilon^2 \log m_2(\varepsilon).$$

Comparing with (4.2) and (4.3), it is clear that the family of translations $(\tilde{\gamma}_B)_B$ is asymptotically efficient for the simulation of the probability (1.1) if and only if the translation γ is asymptotically efficient for the simulation of the probability (1.2).

Example 4.2 In [8] Z.Michna investigates the efficiency of the importance sampling estimator for the ruin probability

$$\mathrm{P}\Big(\sup_{0\le t\le 1}(X_t - ct) > B\Big)$$

obtained by a translation with a path of the form $\gamma_t = at$. He studied, using extensive simulations, the efficiency of this kind of i.s. distributions. This choice of a translation corresponds, with the notations introduced above, to $\tilde{g}_B \equiv 1$ and also $g \equiv 1$. This importance sampling estimator is asymptotically efficient if and only if

$$\inf_{0<t\le 1} H_\gamma(t) = \inf_{0<t\le 1} \frac{(1+kt)^2}{t^{2H}}.$$

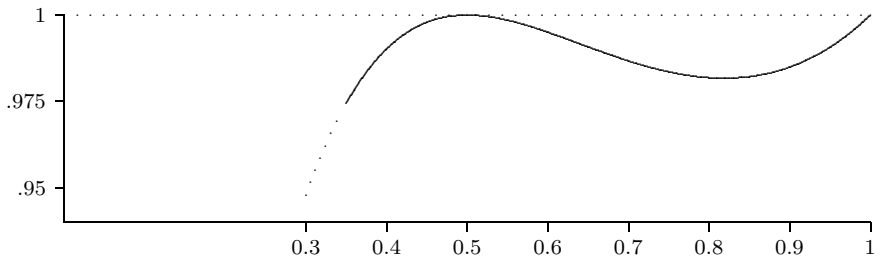

Figure 4 The graph of $\frac{c_2^2}{2c_2^2-1}$, as H varies. Between $\frac{1}{2}$ and 1 the minimum is attained at $\simeq 0.82$ and takes a value $\simeq 0.9817$.

Thanks to Proposition 3.3 we know that this cannot be, as in order to produce an asymptotically efficient translation the path γ must be of the form $\gamma_t = a^*K(t^*,t)$, for $t \le t^*$. One can however investigate for the best value of a and whether this estimator is far from asymptotical efficiency.

If $\gamma_t = at$, $v_t = a^2\mathrm{E}(\mathrm{E}(M_1 \,|\, \mathcal{F}_t)^2) = a^2\mathrm{E}(M_t^2) = a^2c_2^2t^{2-2H}$. Therefore the master function is easily computed:

$$H_\gamma(t) = -a^2c_2^2t^{2-2H} + \frac{1}{2t^{2H}}(1+(k+a)t)^2$$

and one must determine the value of a such that $\inf_{0<t\le 1} H_\gamma(t)$ is largest. It can be shown that in this case sup and inf can be interchanged and (for a fixed t this is a second order polynomial in a)

$$\max_{a>0}\Big\{-a^2c_2^2t^{2-2H} + \frac{1}{2t^{2H}}(1+(k+a)t)^2\Big\} = \frac{c_2^2}{2c_2^2-1}\frac{(1+kt)^2}{t^{2H}} \tag{4.10}$$

the maximum being attained at $a = a(t) = \frac{1+kt}{(2c_2^2-1)t}$. The infimum of the right hand side of (4.10) is attained at $t^* = \frac{H}{c(\alpha-H)}$ and we obtain, for the value of the infimum,

$$\frac{c_2^2}{2c_2^2-1}\frac{(1+kt^*)^2}{t^{*2H}} = \frac{c_2^2}{2c_2^2-1}\,2i_\varphi.$$

Thus the estimator would be asymptotically efficient, but by the presence of the factor

$$\frac{c_2^2}{2c_2^2-1}. \tag{4.11}$$

The graph of the function c_2^2 given in Figure 1 tells us that, for values of H between $\frac{1}{2}$ and 1, c_2^2 is very much close to 1, as well as $c_2^2(2c_2^2-1)^{-1}$ (regardless the value of k), and the graph of $H \to c_2^2(2c_2^2-1)^{-1}$ in Figure 4 confirms this fact. Therefore, for $\frac{1}{2} < H \le 1$, this importance sampling estimator, whereas it is not asymptotically efficient, is not far from being so (see also Figure 3). If $H = 0.6$, $c_2^2(2c_2^2-1)^{-1} = 0.995$. Unfortunately this means that the precision of the estimator tends to 0 as $\varepsilon \to 0$ exponentially fast. It might be interesting to check, numerically, whether this estimator is useful in practice for the usual values of the level of B. Its performance is moreover much poorer for values of H smaller than $\frac{1}{2}$.

The method developed here allows also to determine the best value of a. Actually here

$$a^* = \frac{1 + kt^*}{(2c_2^2 - 1)t^*} = \frac{k}{(2c_2^2 - 1)H}.$$

The path $\gamma_t = a^* t$ gives rise to the lowest second moment order among the linear paths for the simulation of the probability (1.2). The corresponding path for the simulation of the probability (1.1) is therefore, using (4.7),

$$\tilde{\gamma}_B(s) = B\,\gamma(s(Bd)^{-1}) = \frac{c}{(2c_2^2 - 1)H}\, s.$$

In [8], numerical evidence indicates that, for $H = 0.6$ and $c = 1$, the best path of this form is $\gamma_t = at$ with a between 1.5 and 1.9 (beware, we call a here what is written $1+a$ in [8]). The computation above gives $a = ((2c_2^2 - 1)H)^{-1} = 1.65$.

Remark 4.3 If the translation $\gamma_t = at$ is not asymptotically efficient, one can however investigate whether it is optimal, in the sense of the second point of Remark 3.5. Also the answer to this question is negative, as we are able to produce a translation γ such that $\inf_{0<t\le 1} H_\gamma(t)$ is larger. Actually, let us choose a path γ associated to a r.v. of the form $Z = a^* X_{t^*} + Y$, where Y is independent of $\mathcal{F}_{t^*}$. Thus Y is of the form

$$Y = \int_{t^*}^1 g(s)\, dM_s$$

with $g \in L^2_H$. Z being of the form above, this ensures that

$$\gamma_t = \mathrm{E}(X_t Z) = a^* \mathrm{E}(X_t X_{t^*}) = a^* K(t^*, t), \qquad t \le t^*.$$

whereas, if $t^* \le t \le 1$,

$$\gamma_t = \mathrm{E}(X_t Z) = \mathrm{E}[X_t(a^* X_{t^*} + Y)] = a^* K(t^*, t) + \int_{t^*}^t \tilde{z}(t, s) g(s)\, d\langle M\rangle_s$$

Assuming for simplicity that $H > \frac{1}{2}$, so that we are enabled to use the simpler form (3.16), for $t \ge t^*$,

$$\gamma_t = a^* K(t^*, t) + 2H(2-2H)(H-\tfrac{1}{2})c_2^2 \int_{t^*}^t g(s) s^{1-2H}\, ds \int_s^t u^{H-\frac{1}{2}} (u-s)^{H-\frac{3}{2}}\, du =$$

$$= a^* K(t^*, t) + 2H(2-2H)(H-\tfrac{1}{2})c_2^2 \int_{t^*}^t u^{H-\frac{1}{2}} \int_{t^*}^u g(s) s^{1-2H} (u-s)^{H-\frac{3}{2}}\, ds.$$

The corresponding conditional variance $v_t = \mathrm{E}(\mathrm{E}(Z \mid \mathcal{F}_t)^2)$ takes the form

$$v_t = \begin{cases} a^*(2-2H)c_2^2 \displaystyle\int_0^t \tilde{z}(t, s)^2 s^{1-2H}\, ds & \text{if } t \le t^*, \\ v_t = a^* K(t^*, t^*) + (2-2H)c_2^2 \displaystyle\int_{t^*}^t g_s^2 s^{1-2H}\, ds & \text{if } t \ge t^*. \end{cases}$$

If one chooses g of the form $g(s) = c_0(s-t^*)^\beta s^{2H-1}$ for some values of c_0, β to be chosen later, one finds, making simplifications with the value of c_2^2,

$$\gamma_t = a^*K(t^*,t) + \frac{c_0\Gamma(\beta+1)\Gamma(\frac{3}{2}-H)}{\Gamma(H+\beta+\frac{1}{2})\Gamma(2-2H)}\times \\ \times \int_{t^*}^t u^{H-\frac{1}{2}}(u-t^*)^{H+\beta-\frac{1}{2}}\,du. \tag{4.12}$$

for $t \geq t^*$, whereas, for the same values of t,

$$v_t = a^*K(t^*,t^*) + (2-2H)c_0^2c_2^2\int_{t^*}^t (s-t^*)^{2\beta}s^{2H-1}\,ds$$

The master function H_γ associated to this path can be easily computed numerically. Figures 5 and 6 show that, for a good choice of c_0 and β, the translation

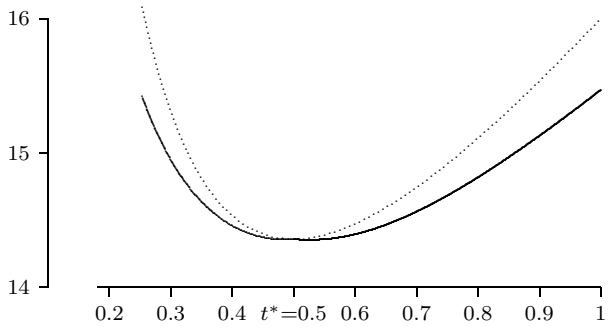

Figure 5 Here again $H = 0.6$, $k = 3$ and the dotted line is the graph of $t \to (1+kt)^2t^{-2H}$. The solid one is the graph of H_γ, with $\gamma = a^*K(t^*,t)$ for $t \leq t^*$ and defined in (4.12) for $t > t^*$. Here $c = 4.7$ and $\beta = 0.01$. The infimum of H_γ is now equal to 14.3527 which is very close to asymptotic efficiency (the required value is still 14.3588). The minimum is attained at $t = .526$. See the next figure for more information of the behaviour of H_γ near t^*.

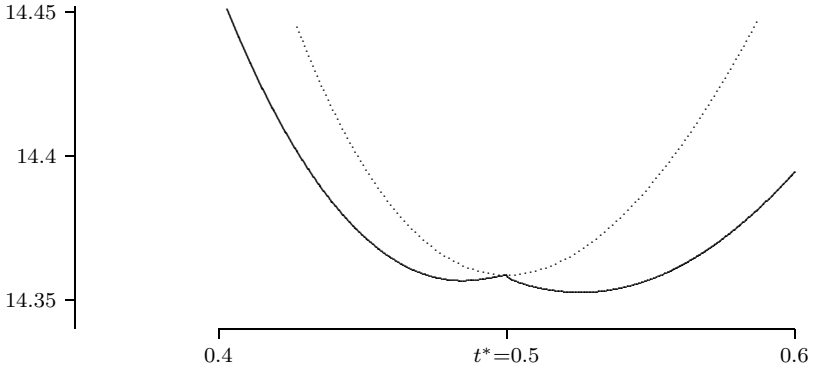

Figure 6 This is an enlargement of the previous graph, in order to enhance the behaviour near t^*. Recall that, at the left of t^*, the path is $\gamma_t = a^*K(t^*,t)$. It holds $H'_\gamma(t^*-) = 0$, but the graph shows clearly, as proved in Example 3.4, that the associated master function attains a local minimum at a point $t' < t^*$.

by such a path γ is closer to asymptotic efficiency than Michna's estimator.

5 Appendix: continuous Gaussian Processes

In this section we review some more or less well known properties of the structure of continuous gaussian processes and prove some specific facts that have been used in the previous sections. Many references are available in the literature. We refer mainly to Lifshits's book [7] and to §2 of de Acosta [3]. Let $\mathcal{C}_t = \mathcal{C}([0,t], \mathbb{R}^m)$ be the Banach space of real continuous paths on $[0,t]$, endowed with the uniform norm $|w|_{t,\infty} = \sup_{0\le s\le t} |w_s|$. Its dual $\mathcal{C}'_t$ is formed by all signed measures on $[0,t]$ with finite variation, through the duality

$$\langle \alpha, w\rangle = \int_0^t w_s \, d\alpha_s.$$

A probability P on $\mathcal{C}_t$ is said to be gaussian if the canonical random variables $X_s : \mathcal{C}_t \to \mathbb{R}$, $s \le t$, defined by $X_s(w) = w_s$ form a gaussian family. In the following we shall assume that P is *full*, that is that it gives strictly positive probability to all open sets of $\mathcal{C}$. The family $(X_s)_s$ then is said to be a (continuous) gaussian process and the function $K(u,v) = \mathrm{Cov}(X_u, X_v)$ is the *covariance function* of the process. If $\alpha \in \mathcal{C}'_t$, it is easy to check that the r.v.'s $X_\alpha(w) = \langle \alpha, w\rangle$ are also gaussian. Let $\mathcal{C}'_{t,\mathrm{P}}$ be the completion in $L^2(\mathcal{C}_t, \mathrm{P})$ of the r.v.'s of $\mathcal{C}'_t$. It is a closed vector space of $L^2(\mathcal{C}, \mathrm{P})$, whose elements form a gaussian family (the *gaussian space* of $(X_t)_t$).

For every r.v. $Z \in \mathcal{C}'_{t,\mathrm{P}}$ let us define, for $s \le t$,

$$w_s = \mathrm{E}(X_s Z). \tag{5.1}$$

With this definition, w is a continuous path, that is $w \in \mathcal{C}_t$. It is easy to prove that the application $Z \to w$ is one to one. We denote by $\mathcal{H}_t$ the set of the paths w of this form. Endowed with the norm

$$|w|_{\mathcal{H}_t} \stackrel{def}{=} \|Z\|_{L^2},$$

$\mathcal{H}_t$ is a Hilbert space, the *Reproducing Kernel Hilbert Space* (RKHS) of P. It is useful to remark that the r.v.'s $X_r, r \le t$, certainly belong to $\mathcal{C}'_{t,\mathrm{P}}$ and that the corresponding path is

$$w_s = \mathrm{E}(X_s X_r) = K(r,s).$$

Thus $K(r,\cdot)$ always belongs to $\mathcal{H}_t$ and $|K(r,\cdot)|_{\mathcal{H}_t} = K(r,r)^{1/2}$. More generally, if Z is of the form

$$Z = \int_0^t X_s \, d\alpha_s,$$

then the corresponding path in the RKHS is

$$w_s = \mathrm{E}(X_s Z) = \int_0^t K(s,u) \, d\alpha_u. \tag{5.2}$$

By construction, these paths are dense in $\mathcal{H}_t$. If $w \in \mathcal{H}_t$ and $Z \in \mathcal{C}'_{t,\mathrm{P}}$ is the corresponding r.v., then the r.v.

$$\mathrm{e}^{Z - \frac{1}{2}|w|^2_{\mathcal{H}_t}} = \mathrm{e}^{Z - \frac{1}{2}\|Z\|^2_{L^2}}$$

has mean 1 with respect to P, so that $d\mathrm{Q} = \mathrm{e}^{Z-\frac{1}{2}|w|^2_{\mathcal{H}_t}} d\mathrm{P}$ is a probability measure on $\mathcal{C}_t$. Let us denote P_w the probability obtained by a translation by w: $\mathrm{P}_w(A) = \mathrm{P}(A - w)$. It is easy to check that, for every $\alpha \in \mathcal{C}_t$,

$$\int \mathrm{e}^{\langle \alpha, x\rangle}\, d\mathrm{Q}(x) = \int \mathrm{e}^{\langle \alpha, x\rangle}\, d\mathrm{P}_w(x).$$

Thus, since the Laplace transforms coincide, $\mathrm{Q} = \mathrm{P}_w$.

If $T > 0$ and P is a gaussian probability on $\mathcal{C}_T$, let $\gamma \in \mathcal{H}_T$ and let Z be the r.v. in the gaussian space $\mathcal{C}'_{T,P}$ corresponding to γ through (5.1). Then the r.v. $\tilde{Z} = \mathrm{E}(Z \,|\, \mathcal{F}_t)$ is still gaussian, as $\{Z, X_u, u \le t\}$ is a gaussian family. The corresponding path $\tilde{\gamma}_s = \mathrm{E}(X_s \tilde{Z})$ enjoys the property that, for $s \le t$,

$$\tilde{\gamma}_s = \mathrm{E}(X_s \tilde{Z}) = \mathrm{E}(X_s Z) = \gamma_s.$$

Thus it coincides with γ up to time t.

Moreover, if $\bar{\gamma} \in \mathcal{H}_T$ is a path coinciding with γ up to time t, by repeating the previous argument, its corresponding gaussian r.v., $\bar{Z}$, must satisfy $\tilde{Z} = \mathrm{E}(Z \,|\, \mathcal{F}_t) = \mathrm{E}(\bar{Z} \,|\, \mathcal{F}_t)$; this means that $\bar{Z} = \tilde{Z} + W$, for some r.v. $W \in \mathcal{C}'_{T,P}$ and independent of $\mathcal{F}_t$. Thus

$$|\tilde{\gamma}|^2_{\mathcal{H}_T} = \mathrm{E}(\tilde{Z}^2) \le \mathrm{E}(\tilde{Z}^2) + \mathrm{E}(W^2) = \mathrm{E}(\bar{Z}^2) = |\bar{\gamma}|^2_{\mathcal{H}_T}. \tag{5.3}$$

Moreover in the previous relation the equal sign holds if and only if $W = 0$ and if and only if $\bar{\gamma} = \tilde{\gamma}$. We have proved the following

Proposition 5.1 *If $\gamma \in \mathcal{H}_T$ and $t \le T$, then there exists a unique path $\gamma^{(t)} \in \mathcal{H}_T$ coinciding with γ up to time t and such that*

$$|\gamma^{(t)}|^2_{\mathcal{H}_T} = \inf |w|^2_{\mathcal{H}_T} \tag{5.4}$$

the infimum being taken over all paths $w \in \mathcal{H}_T$ whose restriction to $[0,t]$ coincides with $\tilde{\gamma}$.

The paths $\gamma^{(t)}$ will play an important role in the following. As for now, remark that, if we denote $\tilde{\mathrm{P}}$ the restriction of P to $\mathcal{F}_t$, $\tilde{\mathrm{P}}$ is a gaussian probability on $\mathcal{C}_t$ and another way of looking at Proposition 5.1 is to say that, if $\gamma \in \mathcal{H}_T$, then its restriction to $[0,t]$, $\tilde{\gamma}$, belongs to $\mathcal{H}_t$ and $|\tilde{\gamma}|_{\mathcal{H}_t} = \inf |w|^2_{\mathcal{H}_T}$, the infimum being taken among the paths $w \in \mathcal{H}_T$ which coincide with γ up to time t.

Remark also that, if $s \le t$, $\langle \gamma^{(t)}, \gamma^{(s)} \rangle_{\mathcal{H}_T} = \langle \gamma^{(s)}, \gamma^{(s)} \rangle_{\mathcal{H}_T}$.

The following statement gives an important, albeit elementary, consequence of the property CGM (see Assumption 3.1).

Lemma 5.2 *If the property CGM is satisfied, then the function $t \to \gamma^{(t)}$ is continuous from $[0,T]$ to $\mathcal{H}_T$.*

Proof.

Remark first that, since for gaussian r.v.'s the a.s. convergence implies the convergence in L^2, the mapping $t \to \zeta_t$ is continuous in L^2.

Thus it is sufficient to remark that, for $s \le t$, $|\gamma^{(t)} - \gamma^{(s)}|^2_{\mathcal{H}_T} = \mathrm{E}[(\zeta_t - \zeta_s)^2]$.

∎

In the following when speaking of the martingale $(\zeta_t)_t$, we will still understand the continuous version whose existence is ensured by Assumption 3.1.

Let P^ε be the probability obtained by scaling with parameter ε: $\mathrm{P}^\varepsilon(A) = \mathrm{P}(\frac{1}{\varepsilon}A)$. Then the following large deviations result is well known: for every closed set $F \subset \mathcal{C}_t$ and every open set $G \subset \mathcal{C}_t$,

$$\begin{aligned} \limsup_{\varepsilon\to 0} \varepsilon^2 \log \mathrm{P}^\varepsilon(F) &\le -\inf_{w\in F} I(w) \\ \liminf_{\varepsilon\to 0} \varepsilon^2 \log \mathrm{P}^\varepsilon(G) &\ge -\inf_{w\in G} I(w) \end{aligned} \tag{5.5}$$

where $I(w) = \frac{1}{2}|w|^2_{\mathcal{H}_t}$ if $w \in \mathcal{H}_t$ and $I(w) = +\infty$ otherwise. See, e.g., [7] §12 or [4] §3.4.

Let $\alpha \in \mathcal{C}'_t$, then the r.v. $Z(x) = \int_0^t x_s \, d\alpha_s$ is a continuous function of the path x. If $w \in \mathcal{H}_t$ is the path corresponding to Z through (5.1), then it is easy to check that, if $\gamma \in \mathcal{H}_t$, $Z(\gamma) = \langle w, \gamma\rangle_{\mathcal{H}_t}$. Conversely, if $Z \in \mathcal{C}'_{t,\mathrm{P}}$ but $Z \notin \mathcal{C}'_t$, then Z is not a continuous function on $\mathcal{C}_t$ (it is a.s. linear and does not belong to $\mathcal{C}'_t$). However Z is still *quasi continuous* (see Definition 4.2 in [1]) with respect to the Large Deviations Principle (5.5), provided we define, for $\gamma \in \mathcal{H}_t$,

$$Z(\gamma) = \langle w, \gamma\rangle_{\mathcal{H}_t} \tag{5.6}$$

where w is the path in the RKHS corresponding to Z. This means that, for every $\rho > 0$, $R > 0, \eta > 0$ there exist $\delta > 0$, $\varepsilon_0 > 0$ such that for every $\varepsilon \le \varepsilon_0$

$$\mathrm{P}^\varepsilon\big(x; |Z(x) - Z(\gamma)| > \eta, d(x,\gamma) < \delta\big) \le \mathrm{e}^{-R/\varepsilon^2} \tag{5.7}$$

uniformly for $\gamma \in \{I \le \rho\}$, where we denote d the distance in the uniform norm: $d(x,y) = ||x - y||_{t,\infty}$.

Quasi-continuity is important in order to apply Varadhan's Lemma, in the following form

Theorem 5.3 *Assume that the family of probability measures $(\mu_\varepsilon)_\varepsilon$ on the complete metric space (E,d) satisfies a Large Deviation Principle with rate function $I$and speed ε^2. Let $\Phi : E \to \mathbb{R}$ a function which is quasi continuous with respect to this Large Deviation Principle, that is satisfying (5.7). Assume moreover that for every $R > 0$ there exists $M > 0$ such that*

$$\limsup_{\varepsilon\to 0} \varepsilon^2 \log \int_E \mathrm{e}^{\Phi(x)/\varepsilon^2} 1_{\{|\Phi(x)|>M\}} \, d\mu_\varepsilon(x) \le -R. \tag{5.8}$$

Then

i) If $F \subset E$ is a closed set, then it holds

$$\limsup_{\varepsilon\to 0} \varepsilon^2 \log \int_F \mathrm{e}^{\Phi(x)/\varepsilon^2} \, d\mu_\varepsilon \le \sup_{x\in F}[\Phi(x) - I(x)].$$

ii) If $G \subset E$ is open, then it holds

$$\liminf_{\varepsilon\to 0} \varepsilon^2 \log \int_G \mathrm{e}^{\Phi(x)/\varepsilon^2} \, d\mu_\varepsilon \ge \sup_{x\in G}[\Phi(x) - I(x)].$$

Theorem 5.3 is proved, e.g., in [1] Theorems 4.3 and 4.4. Let us prove (5.7). Actually, as $\mathcal{C}'_t$ is dense in $\mathcal{C}'_{t,\mathrm{P}}$, there exists $\tilde{Z} \in \mathcal{C}'_t$ such that $\|Z - \tilde{Z}\|_{L^2} < \eta \min((3\sqrt{2\rho})^{-1}, R^{-1/2})$. If $\tilde{w}, w \in \mathcal{H}_t$ are the paths in the RKHS corresponding to $\tilde{Z}, Z$ respectively, then, whenever $I(\gamma) = \frac{1}{2}|\gamma|^2_{\mathcal{H}_t} \le \rho$,

$$|\tilde{Z}(\gamma) - Z(\gamma)| = |\langle \tilde{w}, \gamma\rangle - \langle w, \gamma\rangle| \le |\tilde{w} - w|_{\mathcal{H}_t}|\gamma|_{\mathcal{H}_t} \le \tfrac{\eta}{3}.$$

Thus

$$\mathrm{P}^\varepsilon\big(x; |Z(x) - Z(\gamma)| > \eta, d(x,\gamma) < \delta\big) \le$$
$$\le \mathrm{P}^\varepsilon\big(x; |Z(x) - \tilde{Z}(x)| > \tfrac{\eta}{3}, d(x,\gamma) < \delta\big) + \mathrm{P}^\varepsilon\big(x; |\tilde{Z}(x) - \tilde{Z}(\gamma)| > \tfrac{\eta}{3}, d(x,\gamma) < \delta\big).$$

By the continuity of $\tilde{Z}$, the last term in the right-hand side is equal to 0 if δ is small enough. The remaining term is easily majorized, keeping in mind that, under P^ε, the r.v. $Z - \tilde{Z}$ is gaussian with a variance that is smaller than $\frac{\eta^2}{R}\varepsilon^2$.

We now prove that the functional $-\zeta_\tau + \frac{1}{2}v_\tau$, appearing in (3.4) satisfies the assumptions of Theorem 5.3. First we take care of (5.8). The term v_τ is bounded. If we set $\zeta^* = \sup_{0\le t\le 1}|\zeta_t|$, then obviously $\zeta^* \ge -\zeta_\tau$, and it is sufficient to prove (5.8) for $\Phi = \zeta^*$. This is simple since $(\zeta_t)_t$ is a continuous gaussian process, so that the r.v. ζ^* is finite and Borell's inequality (see [12], e.g.) gives

$$\mathrm{P}(\zeta^* \ge r) \le 2\mathrm{e}^{-c^* r^2}$$

for some $c^* > 0$ (actually one can choose $c^* = (8\mathrm{E}[(\zeta^*)^2])^{-1}$).

Lemma 5.4 *For every $R > 0$ there exists $M > 0$ such that*

$$\limsup_{\varepsilon\to 0} \varepsilon^2 \log \int_{\mathcal{C}} \mathrm{e}^{\frac{1}{\varepsilon^2}\zeta^*(x)} 1_{\{x;\zeta^*(x)>M\}}\, d\mathrm{P}^\varepsilon(x) \le -R.$$

Proof. First remark that the distribution of ζ^* under P_ε is the same as the distribution of $\zeta^*(\varepsilon\,\cdot)$ under P, and, ζ_t being a a.s. linear functional of the path x, $\zeta^*(\varepsilon x) = \varepsilon\zeta^*(x)$ P-a.s. Therefore

$$\begin{aligned}\int_{\mathcal{C}} \mathrm{e}^{\frac{1}{\varepsilon^2}\zeta^*(x)} 1_{\{x;\zeta^*(x)>M\}}\, d\mathrm{P}^\varepsilon(x) &= \int_{\mathcal{C}} \mathrm{e}^{\frac{1}{\varepsilon}\zeta^*(x)} 1_{\{x;\varepsilon\zeta^*(x)>M\}}\, d\mathrm{P}(x) \le \\ &\le \mathrm{P}(\zeta^* \ge \tfrac{M}{\varepsilon})^{1/2}\mathrm{E}(\mathrm{e}^{\frac{2}{\varepsilon}\zeta^*})^{1/2}.\end{aligned} \tag{5.9}$$

By Borell's inequality,

$$\mathrm{E}(\mathrm{e}^{\frac{2}{\varepsilon}\zeta^*}) = \int_0^{+\infty} \mathrm{P}(\mathrm{e}^{\frac{2}{\varepsilon}\zeta^*} \ge r)\, dr = 1 + \int_1^{+\infty} \mathrm{P}(\mathrm{e}^{\frac{2}{\varepsilon}\zeta^*} \ge r)\, dr =$$
$$= 1 + \int_0^{+\infty} \mathrm{P}(\mathrm{e}^{\frac{2}{\varepsilon}\zeta^*} \ge \mathrm{e}^t)\mathrm{e}^t\, dt = 1 + \int_0^{+\infty} \mathrm{P}(\zeta^* \ge \tfrac{\varepsilon}{2}\, t)\mathrm{e}^t\, dt \le$$
$$\le 1 + \int_0^{+\infty} \mathrm{e}^{-\frac{c^*\varepsilon^2}{4}t^2 + t}\, dt \le 1 + \mathrm{e}^{\frac{1}{c^*\varepsilon^2}} \int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{c^*\varepsilon^2}{4}(t - \frac{2}{c^*\varepsilon^2})^2}\, dt = 1 + \tfrac{1}{\varepsilon}\,\mathrm{e}^{\frac{1}{c^*\varepsilon^2}}\sqrt{\tfrac{\pi}{c^*}}.$$

Going back to (5.9),

$$\int_{\mathcal{C}} \mathrm{e}^{\frac{1}{\varepsilon^2}\zeta^*(x)} 1_{\{x;\zeta^*(x)>M\}}\, d\mathrm{P}^\varepsilon(x) \le \Big(1 + \tfrac{1}{\varepsilon}\,\mathrm{e}^{\frac{1}{c^*\varepsilon^2}}\sqrt{\tfrac{\pi}{c^*}}\Big)^{1/2} \mathrm{e}^{-\frac{c^*M^2}{2\varepsilon^2}}$$

which gives

$$\limsup_{\varepsilon\to 0} \varepsilon^2 \log \int_{\mathcal{C}} \mathrm{e}^{\frac{1}{\varepsilon^2}\zeta^*(x)} 1_{\{x;\zeta^*(x)>M\}}\, d\mathrm{P}^\varepsilon(x) \le \frac{1}{2c^*} - \frac{c^*M^2}{2}$$

which allows to conclude.

∎

Let us now check that the functional $-\zeta_\tau + \frac{1}{2}v_\tau$ is quasi-continuous, so that Varadhan's Lemma can be applied. The functional v_τ is even continuous a.s., since τ is continuous P^ε-a.s. for every $\varepsilon > 0$, as proved in the next Proposition 5.8. The quasi-continuity of ζ_τ follows from the following statement. From now on we assume $t = 1$ and drop the subscript t, in order to simplify the notations. Thus $\mathcal{C}, \mathcal{C}', \mathcal{H}, \ldots$ are as $\mathcal{C}_1, \mathcal{C}'_1, \mathcal{H}_1, \ldots$ until now. Recall that $\zeta_t = \mathrm{E}(Z \mid \mathcal{F}_t)$, where Z is the gaussian r.v. associated to the path $\gamma \in \mathcal{H}$ and that we assume that $(\zeta_t)_t$ is continuous.

Proposition 5.5 *For every $\rho > 0$, $R > 0$, $\eta > 0$ there exist $\varepsilon_0 > 0$, $\delta > 0$ such that, for every $w \in \mathcal{H}$,*

$$\mathrm{P}\Big(x;\ \sup_{0\le t\le 1} |\varepsilon\zeta_t(x) - \langle w, \gamma^{(t)}\rangle_{\mathcal{H}}| \ge \eta, \|\varepsilon x - w\|_\infty \le \delta\Big) \le \mathrm{e}^{-R/\varepsilon^2}, \tag{5.10}$$

for every $\varepsilon \le \varepsilon_0$, uniformly for $|w|_{\mathcal{H}} \le \rho$.

Proof. The two-dimensional process $(X_t, \zeta_t)_t$ takes its values in $\mathcal{C}([0,1], \mathbb{R}^2)$ and is gaussian. It is an easy matter to check that its RKHS $\tilde{\mathcal{H}}$ is formed by the paths of the form (w, g), where $w \in \mathcal{H}$ and $g_t = \langle w, \gamma^{(t)}\rangle_{\mathcal{H}}$, endowed with the norm

$$|(w,g)|_{\tilde{\mathcal{H}}} = |w|_{\mathcal{H}}.$$

The rest of the proof consists in majorizing the quantity in (5.10) using large deviations. Actually (5.10) is more or less equivalent to the previous large deviations statement, as developed in Baldi-Sanz [2], whose arguments we reproduce here. The large deviations result for gaussian processes, states that,

$$\begin{aligned}\limsup_{\varepsilon\to 0} \mathrm{P}\Big(x;\ \sup_{0\le t\le 1} |\varepsilon\zeta_t(x) - \langle w, \gamma^{(t)}\rangle_{\mathcal{H}}| \ge \eta, \|\varepsilon x - w\|_\infty \le \delta\Big) \le \\ \le -\inf \tfrac{1}{2}|(y,g)|^2_{\tilde{\mathcal{H}}}\end{aligned} \tag{5.11}$$

the infimum being taken among those paths (y, g) such that $\|y - w\|_\infty \le \delta$ and $\|g - \langle w, \gamma^{(\cdot)}\rangle_{\mathcal{H}}\|_\infty \ge \eta$. Let us fix $R > 0$. Thanks to the following Lemma 5.6, for every $\eta > 0$ there exists $\delta > 0$ such, if $|\tilde{w}|^2_{\mathcal{H}} < 2R$ and $\|\tilde{w} - w\|_\infty \le \delta$, then $\|\langle \tilde{w}, \gamma^{(\cdot)}\rangle_{\mathcal{H}} - \langle w, \gamma^{(\cdot)}\rangle_{\mathcal{H}}\|_\infty \le \eta$. Thus with this choice of δ, if (y, g) is such that $\|y - w\|_\infty \le \delta$ and $\|g - \langle w, \gamma^{(\cdot)}\rangle_{\mathcal{H}}\|_\infty \ge \eta$, then $|(y,g)|^2_{\tilde{\mathcal{H}}} \ge 2R$ and the right hand side in (5.11) is smaller than $-R$, which completes the proof.

∎

Lemma 5.6 *The application $w \to (t \to \langle w, \gamma^{(t)}\rangle_{\mathcal{H}})$ is continuous from $K_R = \{w; I(w) \le R\}$ to $\mathcal{C}$, both endowed with the $\| \ \|_\infty$ topology.*

Proof. Let $(w_n)_n \subset K_R$ such that $\|w_n - w\|_\infty \to 0$. Then, by Proposition 5.1,

$$\langle w_n, \gamma^{(t)}\rangle_{\mathcal{H}} \le |w_n|_{\mathcal{H}}\, |\gamma^{(t)}|_{\mathcal{H}} \le |w_n|_{\mathcal{H}}\, |\gamma|_{\mathcal{H}} \le \sqrt{2R}\, |\gamma|_{\mathcal{H}}$$

so that the functions $\langle w_n, \gamma^{(\cdot)}\rangle_{\mathcal{H}}$, $n = 1, \dots,$ are equi-bounded. Let us prove that they are also equi-continuous:

$$|\langle w_n, \gamma^{(t)}\rangle_{\mathcal{H}} - \langle w_n, \gamma^{(s)}\rangle_{\mathcal{H}}| = |\langle w_n, \gamma^{(t)} - \gamma^{(s)}\rangle_{\mathcal{H}}\,| \le |w_n|_{\mathcal{H}} |\gamma^{(t)} - \gamma^{(s)}|_{\mathcal{H}} \le \\ \le \sqrt{2R}\, |\gamma^{(t)} - \gamma^{(s)}|_{\mathcal{H}}.$$

But, if $s \le t$, it holds $\langle \gamma^{(s)}, \gamma^{(t)}\rangle_{\mathcal{H}} = \langle \gamma^{(s)}, \gamma^{(s)}\rangle_{\mathcal{H}}$; therefore $|\gamma^{(t)} - \gamma^{(s)}|^2_{\mathcal{H}} = |\gamma^{(t)}|^2_{\mathcal{H}} - |\gamma^{(s)}|^2_{\mathcal{H}}$ and

$$|\langle w_n, \gamma^{(t)}\rangle_{\mathcal{H}} - \langle w_n, \gamma^{(s)}\rangle_{\mathcal{H}}| \le \sqrt{R}\, \sqrt{|\gamma^{(t)}|^2_{\mathcal{H}} - |\gamma^{(s)}|^2_{\mathcal{H}}}.$$

Since the function

$$t \to |\gamma^{(t)}|^2_{\mathcal{H}} = \mathrm{E}(\zeta_t^2)$$

is continuous and therefore uniformly continuous on $[0,1]$ (Proposition 5.2), the set $(\langle w_n, \gamma^{(\cdot)}\rangle_{\mathcal{H}})_n$ is equi-bounded and equi-continuous and by the Ascoli-Arzelà theorem, it is relatively compact. Since, by the following Lemma 5.7 $\langle w_n, \gamma^{(t)}\rangle_{\mathcal{H}} \to \langle w, \gamma^{(t)}\rangle_{\mathcal{H}}$ for every t, all the limit points in $\mathcal{C}$ of the sequence $(\langle w_n, \gamma^{(\cdot)}\rangle_{\mathcal{H}})_n$ must coincide with $\langle w, \gamma^{(\cdot)}\rangle_{\mathcal{H}}$ and the proof is complete.
∎

Lemma 5.7 *With the notations of Proposition 5.6, if $w, w_n \in K_R$, $n = 1, \dots$ and $\|w_n - w\|_\infty \to 0$, then $(w_n)_n$ converges to w in the weak topology of $\mathcal{H}$.*

Proof. Let $\gamma \in \mathcal{H}$. Assume first that $w_n, w, \gamma \in \mathcal{H}$ are of the form

$$w_n(t) = \int K(t,s)\, d\alpha_n(s), \quad w(t) = \int K(t,s)\, d\alpha(s), \\ \gamma(t) = \int K(t,s)\, d\beta(s), \tag{5.12}$$

where $\alpha, \alpha_n, \beta \in \mathcal{C}'$.

Then, by the assumptions,

$$\int K(\cdot, s)\, d\alpha_n(s) \underset{n\to\infty}{\to} \int K(\cdot, s)\, d\alpha(s)$$

uniformly. It follows, by the Lebesgue theorem, that

$$\langle w_n, \gamma\rangle_{\mathcal{H}} = \int K(t,s)\, d\alpha_n(s) d\beta(t) \underset{n\to\infty}{\to} \int K(t,s)\, d\alpha(s) d\beta(t) = \langle w, \gamma\rangle_{\mathcal{H}}.$$

As already remarked, the paths of the form (5.12) are dense in $\mathcal{H}$. Let $\tilde{w}_n, \tilde{w}, \tilde{\gamma}$ be paths of this form such that $|\tilde{w} - w|_{\mathcal{H}} < \varepsilon$, $|\tilde{w}_n - w_n|_{\mathcal{H}} < \varepsilon$, $|\tilde{\gamma} - \gamma|_{\mathcal{H}} < \varepsilon$. Then, by the usual trick of adding and subtracting and the Schwartz inequality,

$$\begin{aligned} &|\langle w_n, \gamma\rangle_{\mathcal{H}} - \langle w, \gamma\rangle_{\mathcal{H}}| \leq \\ \leq\ &|\langle w_n, \gamma\rangle_{\mathcal{H}} - \langle \tilde{w}_n, \gamma\rangle_{\mathcal{H}}| + |\langle \tilde{w}_n, \gamma\rangle_{\mathcal{H}} - \langle \tilde{w}_n, \tilde{\gamma}\rangle_{\mathcal{H}}| + \\ &+|\langle \tilde{w}_n, \tilde{\gamma}\rangle_{\mathcal{H}} - \langle \tilde{w}, \tilde{\gamma}\rangle_{\mathcal{H}}| + |\langle \tilde{w}, \tilde{\gamma}\rangle_{\mathcal{H}} - \langle \tilde{w}, \gamma\rangle_{\mathcal{H}}| + |\langle \tilde{w}, \gamma\rangle_{\mathcal{H}} - \langle w, \gamma\rangle_{\mathcal{H}}| \leq \\ &\leq 2\varepsilon(|\gamma|_{\mathcal{H}} + R) |\langle \tilde{w}_n, \tilde{\gamma}\rangle_{\mathcal{H}} - \langle \tilde{w}, \tilde{\gamma}\rangle_{\mathcal{H}}| \end{aligned}$$

which allows to conclude.

∎

Finally we need to prove that the crossing time is a functional of the path which is a.s. continuous. Let, for $a > 0$,

$$\tau_a(w) = \inf\{t \geq 0; X_t(w) \geq a\}.$$

Proposition 5.8 *τ_a is a functional of the path w which is P^ε-a.s. continuous for every $\varepsilon > 0$.*

Proof. τ_a is not a continuous functional of the path. However it is easy to check that it is lower semi-continuous. Similarly the stopping time

$$\tau'_a(w) = \inf\{t \geq 0; X_t(w) > a\}$$

is upper semi-continuous. The statement then follows if we prove that $\tau_a = \tau'_a$ P-a.s. Let $q > 0$, then

$$\mathrm{P}^\varepsilon(\tau_a \leq q < \tau'_a) = \mathrm{P}^\varepsilon\Big(\sup_{s \leq q} X_s = a\Big) = 0$$

as, by Tsyrelson's theorem (see Lifshits's book [7], p.136) the r.v. $\sup_{s \leq q} X_s$ has a density. Thus the event $\{\tau_a < \tau'_a\}$ is negligible, since

$$\{\tau_a < \tau'_a\} = \bigcup_{q \in \mathbb{Q} \cap [0,1]} \{\tau_a \leq q < \tau'_a\}.$$

∎

References

[1] Baldi P., Pacchiarotti B., Importance sampling for the level crossing probability of a Brownian Motion with a negative drift. Preprint 2003, available at http://mat.uniroma2.it/~processi/ISBROW.pdf.

[2] Baldi P., Sanz M., Une remarque sur la théorie des grandes déviations. Séminaire de Probabilités XXV, 345–348, Lecture Notes in Math. 1485, Springer, Berlin, 1991.

[3] de Acosta A., Small deviations in the functional central limit theorem with applications to functional laws of the iterated logarithm. Ann. Probab. **11**, 78–101 (1983).

[4] Deuschel J.-D., Strook D.W., Large Deviations. Academic Press, San Diego 1989.

[5] Duffield N.G., O'Connell N., Large deviations and overflow probabilities for the general single-server queue, with applications. *Math. Proc. Camb. Phil. Soc.*, **118**, 363–374 (1995).

[6] Glasserman P., Yashan W., Counterexamples in importance sampling for large deviations probabilities. *Ann. Appl. Probab*, **7**, 731–746 (1997).

[7] Lifshits M.A., Gaussian Random Functions. Kluwer Academic Publications, Dordrecht–Boston–London 1995.

[8] Michna Z., On tail probability and first passage times for fractional Brownian Motion. Math. Methods Oper. Res. **49**, 335–354 (1999).

[9] Norros I., Valkeila E., Virtamo J., An elementary approach to a Girsanov formula and other analytical results on fractional Brownian motions. *Bernoulli*, **5**, 571–587 (1999).

[10] O'Connell N., Procissi G., On the Build-Up of Large Queues in a Queuing Model with Fractional Brownian Motion Input. Technical Report HPL-BRIMS-98-18, BRIMS, HP Labs. Bristol (U.K.), August 1998.

[11] Pipiras V., Taqqu M.S., Are classes of deterministic integrands for fractional Brownian motion on an interval complete? *Bernoulli*, **7**, 873–897 (2001).

[12] van der Vaart A.W., Wellner J.A. Weak Convergence and Empirical Processes. Springer-Verlag, New York-Berlin-Heidelberg 1996.

[13] Sadowsky J.S., On Monte Carlo estimation of large deviations probabilities. *Ann. Appl. Probab.* **6** 399–422 (1996).

Paolo Baldi
Dipartimento di Matematica
Università di Roma Tor Vergata
Via della Ricerca Scientifica 1
I-00133 Rome, Italy
fax: + 39 0672594699;
e-mail: `baldi@mat.uniroma2.it`
http://160.80.10.11/∼processi/baldi.htm.

Barbara Pacchiarotti
Dipartimento di Matematica
Università di Roma Tor Vergata
Via della Ricerca Scientifica 1
I-00133 Rome, Italy
fax: + 39 0672594699;
e-mail: `pacchiar@mat.uniroma2.it`
http://160.80.10.11/∼pacchiar.