Stochastic analysis on Gaussian space applied to drift estimation

Nicolas Privault*
Laboratoire de Mathématiques
Université de Poitiers
Téléport 2 - BP 30179
86962 Chasseneuil Cedex
France

Anthony Réveillac†
Laboratoire de Mathématiques
Université de La Rochelle
Avenue Michel Crépeau
17042 La Rochelle Cedex
France

august 29, 2007

Abstract

In this paper we consider the nonparametric functional estimation of the drift of Gaussian processes using Paley-Wiener and Karhunen-Loève expansions. We construct efficient estimators for the drift of such processes, and prove their minimaxity using Bayes estimators. We also construct superefficient estimators of Stein type for such drifts using the Malliavin integration by parts formula and stochastic analysis on Gaussian space, in which superharmonic functionals of the process paths play a particular role. Our results are illustrated by numerical simulations and extend the construction of James-Stein type estimators for Gaussian processes by Berger and Wolpert [2].

Key words: Nonparametric drift estimation, Stein estimation, Gaussian space, Malliavin calculus, harmonic analysis.
Mathematics Subject Classification: 62G05, 60H07, 31B05.

1 Introduction

The maximum likelihood estimator $\hat{\mu}$ of the mean $\mu \in \mathbb{R}^d$ of a Gaussian random vector X in $\mathbb{R}^d$ with covariance $\sigma^2 \mathrm{I}_{\mathbb{R}^d}$ under a probability $\mathbb{P}_\mu$ is well-known to be equal to X

*nicolas.privault@math.univ-poitiers.fr
†anthony.reveillac@univ-lr.fr

itself, and can be computed by maximizing the likelihood ratio

$$\frac{1}{(2\pi\sigma^2)^{d/2}} e^{-\frac{\|X-m\|_d^2}{2\sigma^2}}$$

with respect to m, where $\|\cdot\|_d$ denotes the Euclidean norm on $\mathbb{R}^d$. It is efficient in the sense that it attains the Cramer-Rao bound

$$\sigma^2 d = \mathbb{E}_\mu[\|X-\mu\|_d^2] = \inf_Z \mathbb{E}_\mu[\|Z-\mu\|_d^2], \qquad \mu \in \mathbb{R}^d,$$

over all unbiased estimators Z satisfying $\mathbb{E}_\mu[Z] = \mu$, for all $\mu \in \mathbb{R}^d$.

In [9], James and Stein have constructed superefficient estimators for the mean of $X \in \mathbb{R}^d$, of the form

$$\left(1 - \frac{d-2}{\|X\|_d^2}\right) X$$

whose risk is lower than the Cramer Rao bound $\sigma^2 d$ in dimension $d \geq 3$.

Drift estimation for Gaussian processes is of interest in several fields of application. For example in the decomposition

$$X_t = X_t^u + u_t, \qquad t \in [0,T],$$

the process $(X_t)_{t\in[0,T]}$ is interpreted as an observed output signal, the drift $(u_t)_{t\in[0,T]}$ is viewed as an input signal to be estimated and perturbed by a centered Gaussian noise $(X_t^u)_{t\in[0,T]}$, cf. e.g. [8], Ch. VII. Such results find applications in e.g. telecommunication (additive Gaussian channels) and finance (identification of market trends).

Berger and Wolpert [2], [17], have constructed estimators of James-Stein type for the drift of a Gaussian process $(X_t)_{t\in[0,T]}$ by applying the James-Stein procedure to the independent Gaussian random variables appearing in the Karhunen-Loève expansion of the process. In this context, $\hat{u} := (X_t)_{t\in\mathbb{R}_+}$ is seen as a minimax estimator of its own drift $(u_t)_{t\in\mathbb{R}_+}$.

Stein [15] has shown that the James-Stein estimators on $\mathbb{R}^d$ could be extended to a wider family of estimators, using integration by parts for Gaussian measures. Let us briefly recall Stein's argument, which relies on integration by parts with respect to the Gaussian density and on the properties of superharmonic functionals for the Laplacian on $\mathbb{R}^d$. Given an estimator of $\mu \in \mathbb{R}^d$ of the form $X + g(X)$, where $g : \mathbb{R}^d \to \mathbb{R}^d$ is sufficiently smooth, and applying the integration by parts formula

$$\mathbb{E}_\mu[(X_i - \mu_i) g_i(X)] = \sigma^2 \, \mathbb{E}_\mu[\partial_i g_i(X)], \tag{1.1}$$

$g = \sigma^2 \operatorname{grad} \log f = \sigma^2(\partial_1 \log f, \dots, \partial_d \log f)$, one obtains

$$\mathbb{E}_\mu[\|X + \sigma^2 \operatorname{grad} \log f(X) - \mu\|_d^2] = \sigma^2 d + 4\sigma^4 \sum_{i=1}^d \mathbb{E}_\mu \left[\frac{\partial_i^2 \sqrt{f}(X)}{\sqrt{f}(X)} \right],$$

i.e. $X + \sigma^2 \operatorname{grad} \log f(X)$ is a superefficient estimator if

$$\sum_{i=1}^d \partial_i^2 \sqrt{f}(x) < 0, \qquad dx - a.e.,$$

which is possible if $d \geq 3$. In this case, $X + \sigma^2 \operatorname{grad} \log f(X)$ improves in the mean square sense over the efficient estimator $\hat{u}$ which attains the Cramer-Rao bound $\sigma^2 d$ on unbiased estimators of μ.

In this paper we present an extension of Stein's argument to an infinite-dimensional setting using the Malliavin integration by parts formula, with application to the construction of Stein type estimators for the drift of a Gaussian process $(X_t)_{t\in[0,T]}$. Our approach applies to Gaussian processes such as Volterra processes and fractional Brownian motions. It also extends the results of Berger and Wolpert [2] in the same way that the construction of Stein [15] extends that of James and Stein [9], and this allows us to recover the estimators of James-Stein type introduced by Berger and Wolpert [2] as particular cases. Here we replace the Stein equation (1.1) with the integration by parts formula of the Malliavin calculus on Gaussian space. Our estimators are given by processes of the form

$$X_t + D_t \log F, \qquad t \in [0, T],$$

where F is a positive superharmonic random variable on Gaussian space and D_t is the Malliavin derivative indexed by $t \in [0, T]$. In contrast to the minimax estimator $\hat{u}$, such estimators are not only biased but also anticipating with respect to the Brownian filtration $(\mathcal{F}_t)_{t\in[0,T]}$. This however poses no problem when one has access to complete paths from time 0 to T.

For large values of σ it can be shown that the percentage gain of this estimator is at least equal to the universal constant

$$\frac{16}{\pi^4}\int_{\mathbb{R}^4} e^{-\frac{x^2+y^2+z^2+r^2}{2}}\frac{dxdydzdr}{x^2+9y^2+25z^2+49r^2} \tag{1.2}$$

which approximately represents 11.38%, see (7.4) below.

We proceed as follows. In Section 2 we use stochastic calculus in the independent increment case to derive a Cramer-Rao bound over all unbiased drift estimators. This bound is attained by the process $\hat{u} := (X_t)_{t\in[0,T]}$, which will be considered as an efficient drift estimator. In Section 3 we compute the Bayes estimators obtained under prior Gaussian distributions. We show that these Bayes estimators are admissible, and use them to prove that the drift estimator $\hat{u}$ is minimax. The tools and results presented in Sections 2 and 3 are not surprising, but we did not find any source covering them in the literature. In Section 4 we recall the elements of analysis and integration by parts on Gaussian space which will be needed in Section 5 to construct superefficient drift estimators for Gaussian processes using superharmonic random functionals on Gaussian space. The superefficiency of these estimators will show, as in the classical case, that the minimax estimator $\hat{u}$ is not admissible. In Section 6 we give examples of nonnegative superharmonic functionals using cylindrical functionals and potential theory on Gaussian space. Examples are considered in Section 7 in case u is deterministic. We show that the James-Stein estimators of Berger and Wolpert [2] can be recovered as particular cases in our approach, and we provide numerical simulations for the gain of such estimators. It turns out that in those examples, the gain obtained in comparison with the minimax estimator $(X_t)_{t\in[0,T]}$ is a function of σ^2/T, thus making σ and T play inverse roles, unlike in the usual setting of Brownian rescaling.

This paper is an extended version of [12] and provides proofs of the results presented in [13].

Notation

Let $T > 0$. Consider a real-valued centered Gaussian process $(X_t)_{t\in[0,T]}$ with covariance function

$$\gamma(s,t) = \mathbb{E}[X_s X_t], \qquad s,t \in [0,T],$$

on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathcal{F}$ is the σ-algebra generated by X. Recall that $(X_t)_{t\in[0,T]}$ can be represented in different ways as an isonormal Gaussian process on a real separable Hilbert space H, i.e. as an isometry $X : H \to L^2(\Omega, \mathcal{F}, P)$ such that $\{X(h) \ : \ h \in H\}$ is a family of centered Gaussian random variables satisfying

$$\mathbb{E}[X(h)X(g)] = \langle h, g\rangle_H, \qquad h, g \in H,$$

where $\langle \cdot, \cdot\rangle_H$ and $\|\cdot\|_H$ denote the scalar product and norm on H.

One can distinguish two main types of such isonormal representations of X_t, see e.g. [1] and [2] respectively for details.

(A) Paley-Wiener expansions. In this case, H is the completion of the linear space generated by the functions $\chi_t(s) = \min(s,t)$, $s,t \in [0,T]$, with respect to the norm

$$\langle \chi_t, \chi_s\rangle_H := \gamma(s,t), \qquad s,t \in [0,T],$$

and $X(\cdot)$ is constructed on H from $X(\chi_t) := X_t$, $t \in [0,T]$, i.e. we have

$$X(\chi_t) = \sum_{k=0}^{\infty} \langle \chi_t, h_k(t)\rangle_H X(h_k), \qquad t \in [0,T],$$

for any orthonormal basis $(h_k)_{k\in\mathbb{N}}$ of H. Assume in addition $\gamma(s,t)$ has the form

$$\gamma(s,t) = \int_0^{s\wedge t} K(t,r)K(s,r)dr, \qquad s,t \in [0,T],$$

where $K(\cdot,\cdot)$ is a deterministic kernel and

$$(Kh)(t) := \int_0^t K(t,s)\dot{h}(s)ds$$

is differentiable in $t \in [0, T]$, and let K^* denote the adjoint of K with respect to

$$\langle h, g \rangle := \langle \dot{h}, \dot{g} \rangle_{L^2([0,T],dt)}.$$

The scalar product in H then satisfies

$$\langle h, g \rangle_H = \langle K^* h, K^* g \rangle = \langle h, \Gamma g \rangle,$$

where $\Gamma = K K^*$, and we have the decomposition

$$X_t = \sum_{k=0}^{\infty} \langle \chi_t, h_k \rangle_H X(h_k) = \sum_{k=0}^{\infty} \langle 1_{[0,t]}, \dot{\Gamma} h_k \rangle_{L^2([0,T],dt)} X(h_k) = \sum_{k=0}^{\infty} \Gamma h_k(t) X(h_k),$$

$t \in [0, T]$. In this case we also have the representation

$$X_t = \int_0^t K(t, s) dW_s, \qquad t \in [0, T],$$

where $(W_s)_{s \in [0,T]}$ is a standard Brownian motion, cf. [1].

(B) Karhunen-Loève expansions. This framework is used in [2]. In this case, μ is a finite Borel measure on $[0, T]$ and H is defined from

$$\langle h, g \rangle_H = \langle h, \Gamma g \rangle,$$

where

$$\langle h, g \rangle := \langle h, g \rangle_{L^2([0,T],d\mu)},$$

and

$$(\Gamma g)(t) = \int_0^T g(s) \gamma(s, t) \mu(ds), \qquad t \in [0, T],$$

with

$$X(h) = \int_0^T X_s h(s) \mu(ds), \qquad h \in H.$$

Given $(h_k)_{k \in \mathbb{N}}$ an orthonormal basis of $L^2([0, T], d\mu)$, we have the expansion

$$X_t = \sum_{k=0}^{\infty} h_k(t) X(h_k), \qquad t \in [0, T].$$

In the sequel we will use mainly the framework (A) with $\langle h, g\rangle = \langle \dot{h}, \dot{g}\rangle_{L^2([0,T],dt)}$, which is better adapted to our approach, although some results valid in the general framework of Gaussian processes will be valid for (B) as well.

The Girsanov theorem for Gaussian processes, cf. e.g. [11], states that X^u defined as

$$X^u(g) := X(g) - \langle g, u\rangle = X(g) - \langle g, \Gamma^{-1}u\rangle_H, \qquad g \in H,$$

where $u \in H$ is deterministic, has same law as X under the probability $\mathbb{P}_u$ defined by

$$\frac{d\mathbb{P}_u}{d\mathbb{P}} = \exp\left(X(\Gamma^{-1}u) - \frac{1}{2}\|\Gamma^{-1}u\|_H^2 \right).$$

In other terms, in case (A) we have

$$\begin{aligned} X_t - u(t) &= X_t - \Gamma\Gamma^{-1}u(t) \\ &= \sum_{k=0}^{\infty} \Gamma h_k(t)(X(h_k) - \langle h_k, \Gamma^{-1}u\rangle_H) \\ &= \sum_{k=0}^{\infty} \Gamma h_k(t) X^u(h_k), \qquad t \in [0,T], \end{aligned}$$

where $(h_k)_{k\in\mathbb{N}}$ is orthonormal basis of H, and in case (B),

$$\begin{aligned} X_t - u(t) &= \sum_{k=0}^{\infty} h_k(t)(X(h_k) - \langle h_k, u\rangle_{L^2([0,T],d\mu)}) \\ &= \sum_{k=0}^{\infty} h_k(t)(X(h_k) - \langle h_k, \Gamma^{-1}u\rangle_H) \\ &= \sum_{k=0}^{\infty} h_k(t) X^u(h_k), \qquad t \in [0,T], \end{aligned}$$

where $(h_k)_{k\in\mathbb{N}}$ is orthonormal basis of $L^2([0,T], d\mu)$.

2 Efficient drift estimator

Here we work in the framework of (A), in the particular case where $(X_t)_{t\in[0,T]}$ has independent increments, i.e.

$$\gamma(s,t) = \int_0^{s\wedge t} \sigma_u^2 du,$$

where $\sigma \in L^2([0,T], dt)$ is an a.e. non-vanishing function,

$$(\dot{K}h)(t) = (\dot{K}^*h)(t) = \sigma_t \dot{h}(t), \qquad t \in [0,T],$$

with $K(t,r) = 1_{[0,t]}(r)\sigma_r$ and

$$\Gamma h(t) = \int_0^t \dot{h}_s \sigma_s^2 ds, \qquad t \in [0,T].$$

In other terms, $(X_t)_{t\in[0,T]}$ is a continuous Gaussian martingale with quadratic variation $\sigma_t^2 dt$, which can be represented as the time change

$$X_t = W_{\int_0^t \sigma_s^2 ds}, \qquad t \in [0,T],$$

of the standard Brownian motion $(W_t)_{t\in\mathbb{R}_+}$, or as the stochastic integral process $X_t = \int_0^t \sigma_s dW_s$, $t \in [0,T]$, and we have $X(h) = \int_0^T \dot{h}(s) dX_s$, $h \in H$, where

$$H = \left\{ v : [0,T] \to \mathbb{R} \ : \ v(t) = \int_0^t \dot{v}(s) ds, \ t \in [0,T], \ \dot{v} \in L^2([0,T], \sigma_t^2 dt) \right\}$$

is the Cameron-Martin space with inner product

$$\langle v_1, v_2 \rangle_H = \int_0^T \dot{v}_1(s) \dot{v}_2(s) \sigma_s^2 ds, \qquad v_1, v_2 \in H.$$

Let $(\mathcal{F}_t)_{t\in[0,T]}$ denote the filtration generated by $(X_t)_{t\in[0,T]}$, and for u an $\mathcal{F}_t$-adapted process, let $\mathbb{P}_u^\sigma$ denote the translation of the Wiener measure on Ω by u, i.e. $\mathbb{P}_u^\sigma$ is the measure on Ω under which

$$X_t^u := X_t - u_t, \qquad t \in [0,T],$$

is a continuous Gaussian martingale with quadratic variation

$$d\langle X^u, X^u \rangle_t = \sigma_t^2 dt.$$

Consider u an $\mathcal{F}_t$-adapted processes of the form

$$u_t = \int_0^t \dot{u}_s ds, \qquad t \in [0,T],$$

with

$$\mathbb{E}^\sigma \left[\int_0^T \frac{\dot{u}_s^2}{\sigma_s^2} ds \right] < \infty.$$

By the Girsanov theorem, $\mathbb{P}_u^\sigma$ is absolutely continuous with respect to $\mathbb{P}^\sigma$, with

$$d\mathbb{P}_u^\sigma = \Lambda(u) d\mathbb{P}^\sigma,$$

where

$$\Lambda(u) := \exp\left(\int_0^T \frac{\dot{u}_s}{\sigma_s^2} dX_s - \frac{1}{2}\int_0^T \frac{\dot{u}_s^2}{\sigma_s^2} ds\right)$$

denotes the Girsanov-Cameron-Martin density, the canonical process $(X_t)_{t\in[0,T]}$ becomes a continuous Gaussian semimartingale under $\mathbb{P}_u^\sigma$, with quadratic variation $\sigma_t^2 dt$ and drift $\dot{u}_t dt$. The expectation under $\mathbb{P}_u$ will be denoted by $\mathbb{E}_u$.

Definition 2.1. *A drift estimator ξ is called unbiased if*

$$\mathbb{E}_u[\xi_t] = \mathbb{E}_u[u_t], \quad t \in [0,T],$$

for all square-integrable $\mathcal{F}_t$-adapted process $(u_t)_{t\in[0,T]}$. It is called adapted if the process $(\xi_t)_{t\in[0,T]}$ is $\mathcal{F}_t$-adapted.

Here, the canonical process $(X_t)_{t\in[0,T]}$ will be considered as an unbiased estimator of own its drift $(u_t)_{t\in[0,T]}$ under $\mathbb{P}_u^\sigma$, with risk defined as

$$\mathbb{E}_u^\sigma\left[\|X-u\|^2_{L^2([0,T],d\mu)}\right] = \int_0^T \mathbb{E}_u^\sigma\left[|X_t^u|^2\right]\mu(dt) = \int_0^T\int_0^t \sigma_s^2 ds\mu(dt),$$

where μ is a finite Borel measure on $[0,T]$. Clearly this estimator is consistent as σ or T tend to 0: precisely, given N independent samples

$$(X_t^1)_{t\in[0,T]}, \ldots, (X_t^N)_{t\in[0,T]},$$

of $(X_t)_{t\in[0,T]}$, the process

$$\bar{X}_t := \frac{X_t^1 + \cdots + X_t^N}{N}, \qquad t \in [0,T], \tag{2.1}$$

is an unbiased estimator of $(u_t)_{t\in[0,T]}$ whose risk

$$\mathbb{E}_u^{\sigma/\sqrt{N}}\left[\|\bar{X}-u\|^2_{L^2([0,T],d\mu)}\right] = \frac{1}{N}\int_0^T\int_0^t \sigma_s^2 ds\mu(dt)$$

converges to zero as N goes to infinity.

The justification of the use of $\hat{u} = (X_t)_{t\in[0,T]}$ as an efficient estimator comes from the following proposition which allows us to compute a Cramer-Rao bound attained by $\hat{u}$. Here the parameter space is restricted to the space of adapted processes in $L^2(\Omega \times [0,T], \mathbb{P} \otimes \mu)$, which corresponds in a sense to a parametric estimation.

Proposition 2.2. *Cramer-Rao inequality. For any unbiased and adapted estimator ξ of u we have*

$$\mathbb{E}_u^\sigma \left[\int_0^T |\xi_t - u_t|^2 \mu(dt) \right] \geq \mathrm{R}(\sigma, \mu, \hat{u}), \tag{2.2}$$

where $u \in L^2(\Omega \times [0,T], \mathbb{P}_u^\sigma \otimes \mu)$ is adapted and the Cramer-Rao type bound

$$\mathrm{R}(\sigma, \mu, \hat{u}) := \int_0^T \int_0^t \sigma_s^2 ds \mu(dt)$$

is independent of u and attained by the efficient estimator $\hat{u} = X$.

Proof. Since ξ is unbiased, for all $\zeta \in H$ we have

$$\begin{aligned}
\mathbb{E}_{u+\varepsilon\zeta}^\sigma[\xi_t] &= \mathbb{E}_{u+\varepsilon\zeta}^\sigma[u_t + \varepsilon\zeta_t] \\
&= \mathbb{E}_{u+\varepsilon\zeta}^\sigma[u_t] + \varepsilon\, \mathbb{E}_{u+\varepsilon\zeta}^\sigma[\zeta_t] \\
&= \mathbb{E}_{u+\varepsilon\zeta}^\sigma[u_t] + \varepsilon\zeta_t, \quad t \in [0,T], \quad \varepsilon \in \mathbb{R},
\end{aligned}$$

hence

$$\begin{aligned}
\zeta_t &= \frac{d}{d\varepsilon}\, \mathbb{E}_{u+\varepsilon\zeta}^\sigma[\xi_t - u_t]_{|\varepsilon=0} \\
&= \frac{d}{d\varepsilon}\, \mathbb{E}^\sigma[(\xi_t - u_t)\Lambda(u+\varepsilon\zeta)]_{|\varepsilon=0} \\
&= \mathbb{E}^\sigma\left[(\xi_t - u_t) \frac{d}{d\varepsilon} \Lambda(u+\varepsilon\zeta)_{|\varepsilon=0} \right] \\
&= \mathbb{E}_u^\sigma\left[(\xi_t - u_t) \frac{d}{d\varepsilon} \log \Lambda(u+\varepsilon\zeta)_{|\varepsilon=0} \right] \\
&= \mathbb{E}_u^\sigma\left[(\xi_t - u_t) \left(\int_0^T \frac{\dot{\zeta}_s}{\sigma_s^2} dX_s - \int_0^T \frac{\dot{\zeta}_s \dot{u}_s}{\sigma_s^2} ds \right) \right] \\
&= \mathbb{E}_u^\sigma\left[(\xi_t - u_t) \int_0^T \frac{\dot{\zeta}_s}{\sigma_s^2} dX_s^u \right] \\
&= \mathbb{E}_u^\sigma\left[(\xi_t - u_t) \int_0^t \frac{\dot{\zeta}_s}{\sigma_s^2} dX_s^u \right],
\end{aligned}$$

where the exchange between expectation and derivative is justified by classical uniform integrability arguments. Thus, by the Cauchy-Schwarz inequality and the Itô isometry we have

$$\zeta_t^2 \leq \mathbb{E}_u^\sigma \left[\left(\int_0^t \frac{\dot{\zeta}_s}{\sigma_s^2} dX_s^u \right)^2 \right] \mathbb{E}_u^\sigma[|\xi_t - u_t|^2] = \int_0^t \frac{\dot{\zeta}_s^2}{\sigma_s^2} ds \, \mathbb{E}_u^\sigma[|\xi_t - u_t|^2], \qquad t \in [0, T].$$

It then suffices to take

$$\zeta_t = \int_0^t \sigma_s^2 ds, \qquad t \in [0, T],$$

to get

$$\mathrm{Var}_u^\sigma[\xi_t] = \mathbb{E}_u^\sigma[|\xi_t - u_t|^2] \geq \int_0^t \sigma_s^2 ds, \qquad t \in [0, T], \tag{2.3}$$

which leads to (2.2) after integration with respect to $\mu(dt)$. As noted above, $\hat{u} = (X_t)_{t\in[0,T]}$ is clearly unbiased under $\mathbb{P}_u^\sigma$ and it attains the lower bound $\mathrm{R}(\sigma, \mu, \hat{u})$. $\square$

Recall that the classical linear parametric estimation problem for the drift of a diffusion consists in estimating the coefficient θ appearing in

$$d\xi_t = \theta a_t(\xi_t)dt + dY_t, \qquad \xi_0 = 0,$$

with a maximum likelihood estimator $\hat{\theta}_T$ given by

$$\hat{\theta}_T = \frac{\int_0^T a_t(\xi_t) d\xi_t}{\int_0^T a_t^2(\xi_t) dt}, \tag{2.4}$$

cf. [10], [14] for Brownian motion and [16] for an extension to fractional Brownian motions.

Here we consider the nonparametric functional estimation of the drift of a one-dimensional drifted Brownian motion $(X_t)_{t\in\mathbb{R}_+}$ with decomposition

$$dX_t = \dot{u}_t dt + dX_t^u, \tag{2.5}$$

where $(\dot{u}_t)_{t\in[0,T]} \in L^2(\Omega \times [0, T])$ is an adapted process and $(X_t^u)_{t\in\mathbb{R}_+}$ is a standard Brownian motion with quadratic variation σ_t^2 under a probability $\mathbb{P}_u^\sigma$.

In case u is constrained to have the form $u_t = \theta t$, $t \in [0,T]$, $\theta \in \mathbb{R}$, our efficient estimator $\hat{u}$ satisfies $\hat{u}_T = \hat{\theta}_T T$, where $\hat{\theta}_T$ is given by (2.4), $T > 0$, with the asymptotics $\hat{\theta}_T \to \theta$ in probability as T tends to infinity. The asymptotics is not in large time since T can be a fixed parameter, but the efficient estimator $\hat{u} = X$ converges to u as σ tends to 0, or equivalently as T tends to 0 by rescaling.

To close this section we note that, at least informally, $\hat{u} = (X_t)_{t\in[0,T]}$ can be viewed as a maximum likelihood estimator of its own adapted drift $(u_t)_{t\in[0,T]}$ under $\mathbb{P}_u^\sigma$. Indeed the functional differentiation of the Cameron-Martin density

$$\frac{d}{d\varepsilon}\Lambda(\hat{u} + \varepsilon\zeta)_{|\varepsilon=0} = 0, \qquad \zeta \in H,$$

implies

$$\int_0^T \frac{\dot{\zeta}_s}{\sigma_s^2} dX_s - \int_0^T \frac{\dot{\zeta}_s}{\sigma_s^2} d\hat{u}_s = 0, \qquad \zeta \in H,$$

which leads to $X = \hat{u}$.

3 Bayes estimators

In this section we consider Bayes estimators which will be useful in proving the minimaxity of the estimator $\hat{u} = X$ in the framework of (A) for Gaussian processes with non-necessarily independent increments. We will make use of the next lemma which is classical in the framework of Gaussian filtering and is proved in the Appendix.

Lemma 3.1. *Let Z be a Gaussian process with covariance operator Γ_τ and drift $v \in H$, and assume that X is a Gaussian process with drift Z and covariance operator Γ given Z. Then, conditionally to X, Z has drift*

$$f \mapsto \langle f, (\Gamma + \Gamma_\tau)^{-1}\Gamma v\rangle + X((\Gamma + \Gamma_\tau)^{-1}\Gamma_\tau f) \quad \textit{and covariance} \quad \Gamma_\tau(\Gamma + \Gamma_\tau)^{-1}\Gamma.$$

Note that unlike in Proposition 2.2, no adaptedness or unbiasedness restriction is made on ξ in the infimum taken in (3.2) below.

Proposition 3.2. *Bayes estimator. Let P_v^τ denote the Gaussian distribution on Ω with covariance operator Γ_τ and drift $v \in H$. The Bayes risk*

$$\int_\Omega \mathbb{E}_z \left[\int_0^T |\xi_t - z_t|^2 \mu(dt) \right] d\mathbb{P}_v^\tau(z) \tag{3.1}$$

of any estimator $(\xi_t)_{t\in[0,T]}$ on Ω under the prior distribution $\mathbb{P}_v^\tau$ is uniquely minimized by

$$\xi_t^{\tau,v} := \langle \chi_t, (\Gamma_\tau + \Gamma)^{-1}\Gamma v\rangle + X((\Gamma_\tau + \Gamma)^{-1}\Gamma_\tau\, \chi_t), \qquad t \in [0,T],$$

which has risk

$$\int_0^T \langle \chi_t, \Gamma(\Gamma_\tau + \Gamma)^{-1}\Gamma_\tau\, \chi_t\rangle \mu(dt) = \inf_\xi \int_\Omega \mathbb{E}_z \left[\int_0^T |\xi_t - z_t|^2 \mu(dt) \right] d\mathbb{P}_v^\tau(z). \tag{3.2}$$

Proof. Let Z denote a Gaussian process with drift $v \in H$ and covariance Γ_τ. Recall (cf. Lemma 3.1) that if X has drift Z and covariance Γ then, conditionally to X, $(Z_t)_{t\in[0,T]}$ has drift

$$t \mapsto \langle \chi_t, (\Gamma_\tau + \Gamma)^{-1}\Gamma v\rangle + X((\Gamma_\tau + \Gamma)^{-1}\Gamma_\tau\, \chi_t)$$

and covariance $\Gamma_\tau(\Gamma_\tau + \Gamma)^{-1}\Gamma$. Hence the Bayes risk of an estimator ξ under the prior distribution $\mathbb{P}_v^\tau$ is given by

$$\begin{aligned}
&\int_\Omega \mathbb{E}_z \left[\int_0^T |\xi_t - z_t|^2 \mu(dt) \right] d\mathbb{P}_v^\tau(z) = \mathbb{E}\left[\mathbb{E}\left[\int_0^T |\xi_t - Z_t|^2 \mu(dt) \Big| X \right]\right] \\
&= \mathbb{E}\left[\int_0^T |\xi_t - \mathbb{E}[Z_t \mid X]|^2 \mu(dt) \right] + \mathbb{E}\left[\int_0^T \mathrm{Var}(Z_t|X) \mu(dt) \right] \\
&= \mathbb{E}\left[\int_0^T \left|\xi_t - \langle \chi_t, (\Gamma_\tau + \Gamma)^{-1}\Gamma v\rangle - X((\Gamma_\tau + \Gamma)^{-1}\Gamma_\tau\, \chi_t)\right|^2 \mu(dt) \right] \\
&\quad + \int_0^T \langle \chi_t, \Gamma(\Gamma_\tau + \Gamma)^{-1}\Gamma_\tau\, \chi_t\rangle \mu(dt),
\end{aligned}$$

which is minimized by

$$\xi_t^{\tau,v} := \mathbb{E}[Z_t \mid X] = \langle \chi_t, (\Gamma_\tau + \Gamma)^{-1}\Gamma v\rangle - X((\Gamma_\tau + \Gamma)^{-1}\Gamma_\tau\, \chi_t), \qquad t \in [0,T].$$

□

Clearly $\xi^{\tau,v}$ is unique in the sense that it is the only estimator to minimize the Bayes risk (3.1). This shows in particular that every $\xi^{\tau,v}$ is admissible in the sense that if an estimator ξ satisfies

$$\mathbb{E}_z \left[\|\xi - z\|^2_{L^2([0,T],d\mu)} \right] \leq \mathbb{E}_z \left[\|\xi^{\tau,v} - z\|^2_{L^2([0,T],d\mu)} \right], \qquad z \in \Omega,$$

then

$$\begin{aligned} \int_\Omega \mathbb{E}_z \left[\|\xi - z\|^2_{L^2([0,T],d\mu)} \right] d\mathbb{P}^\tau_v &\leq \int_\Omega \mathbb{E}_z \left[\|\xi^{\tau,v} - z\|^2_{L^2([0,T],d\mu)} \right] d\mathbb{P}^\tau_v \\ &= \int_0^T \langle \chi_t, \Gamma(\Gamma_\tau + \Gamma)^{-1} \Gamma_\tau \, \chi_t \rangle \mu(dt), \end{aligned}$$

hence

$$\int_\Omega \mathbb{E}_z \left[\|\xi - z\|^2_{L^2([0,T],d\mu)} \right] d\mathbb{P}^\tau_v = \int_0^T \langle \chi_t, \Gamma(\Gamma_\tau + \Gamma)^{-1} \Gamma_\tau \, \chi_t \rangle \mu(dt), \tag{3.3}$$

and $\xi = \xi^{\tau,v}$ by Proposition 3.2.

The Bayes estimator $\xi^{\tau,v}$ is biased in general, and for deterministic $u \in H$ its mean square error under $\mathbb{P}_u$ is equal to

$$\begin{aligned} &\mathbb{E}_u \left[\int_0^T |\xi^{\tau,v}_t - u_t|^2 \mu(dt) \right] \qquad (3.4) \\ &= \mathbb{E}_u \left[\int_0^T \left| \xi^{\tau,v}_t - \mathbb{E}_u[\xi^{\tau,v}_t] \right|^2 \mu(dt) \right] + \mathbb{E}_u \left[\int_0^T | \, \mathbb{E}_u[\xi^{\tau,v}_t] - u_t|^2 \mu(dt) \right] \\ &= \mathbb{E}_u \left[\int_0^T \left| X^u((\Gamma_\tau + \Gamma)^{-1} \Gamma_\tau \, \chi_t) \right|^2 \mu(dt) \right] + \int_0^T \left| \langle \chi_t, (\Gamma_\tau + \Gamma)^{-1} \Gamma (v - u) \rangle \right|^2 \mu(dt) \\ &= \int_0^T \langle (\Gamma_\tau + \Gamma)^{-1} \Gamma_\tau \, \chi_t, \Gamma(\Gamma_\tau + \Gamma)^{-1} \Gamma_\tau \, \chi_t \rangle \mu(dt) \\ &\quad + \int_0^T \left| \langle \chi_t, (\Gamma_\tau + \Gamma)^{-1} \Gamma (v - u) \rangle \right|^2 \mu(dt), \end{aligned}$$

which shows that

$$\sup_{u \in H} \mathbb{E}_u \left[\int_0^T |\xi^{\tau,v}_t - u_t|^2 \mu(dt) \right] = +\infty,$$

hence $\xi^{\tau,v}$ is not minimax.

In the independent increment case of Section 2 we have, if $\Gamma_\tau f(s) = \tau_s^2 f(s)$, $s \in [0,T]$:

$$\xi^{\tau,v}_t := \int_0^t \frac{\sigma_s^2 \dot{v}_s}{\tau_s^2 + \sigma_s^2} ds + \int_0^t \frac{\tau_s^2}{\tau_s^2 + \sigma_s^2} dX_s, \qquad t \in [0,T],$$

with risk

$$\int_0^T \int_0^t \frac{\tau_s^2\sigma_s^2}{\tau_s^2+\sigma_s^2} ds\mu(dt) = \inf_\xi \int_\Omega \mathbb{E}_z \left[\int_0^T |\xi_t - z_t|^2 \mu(dt) \right] d\mathbb{P}_v^\tau(z). \tag{3.5}$$

Assuming now that $\Gamma_\tau f(t) = \tau^2 f(t)$, $t \in [0,T]$, the Bayes risk

$$\int_0^T \langle \chi_t, \Gamma_\tau(\Gamma_\tau + \Gamma)^{-1}\Gamma\, \chi_t\rangle \mu(dt) = \int_0^T \langle \chi_t, (I + \Gamma/\tau^2)^{-1}\Gamma\, \chi_t\rangle \mu(dt),$$

of $\xi^{\tau,v}$, $\tau \in \mathbb{R}$, converges as $\tau \to \infty$ to the bound

$$\mathrm{R}(\sigma, \mu, \hat{u}) = \int_0^T \langle \chi_t, \Gamma\, \chi_t\rangle \mu(dt),$$

hence it follows in the next proposition that, as in the finite dimensional Gaussian case, the estimator $\hat{u} = (X_t)_{t\in[0,T]}$ is minimax. Note again that unlike in Proposition 2.2, no adaptedness condition is imposed on ξ in the infima (3.2) and (3.6).

Proposition 3.3. *The estimator $\hat{u} = X$ is minimax. For all $u \in \Omega$ we have*

$$\mathrm{R}(\gamma, \mu, \hat{u}) = \mathbb{E}_u \left[\int_0^T |X_t - u_t|^2 \mu(dt) \right] = \inf_\xi \sup_{v\in\Omega} \mathbb{E}_v \left[\int_0^T |\xi_t - v_t|^2 \mu(dt) \right]. \tag{3.6}$$

Proof. Clearly, taking $\xi = 0$ yields

$$\mathrm{R}(\gamma, \mu, \hat{u}) = \sup_{u\in\Omega} \mathbb{E}_u \left[\int_0^T |X_t - u_t|^2 \mu(dt) \right] \geq \inf_\xi \sup_{u\in\Omega} \mathbb{E}_u \left[\int_0^T |\xi_t - u_t|^2 \mu(dt) \right].$$

On the other hand, from Proposition 3.2, for all processes ξ we have

$$\begin{aligned}
\sup_{u\in\Omega} \mathbb{E}_u \left[\int_0^T |\xi_t - u_t|^2 \mu(dt) \right] &\geq \int_\Omega \mathbb{E}_z \left[\int_0^T |\xi_t - z_t|^2 \mu(dt) \right] d\mathbb{P}_0^\tau(z) \\
&\geq \int_0^T \langle \chi_t, (I + \Gamma/\tau^2)^{-1}\Gamma\, \chi_t\rangle \mu(dt),
\end{aligned}$$

for all $\tau > 0$, hence

$$\inf_\xi \sup_{u\in H} \mathbb{E}_u \left[\int_0^T |\xi_t - u_t|^2 \mu(dt) \right] \geq \int_0^T \langle \chi_t, \Gamma\, \chi_t\rangle \mu(dt) = \mathrm{R}(\gamma, \mu, \hat{u}).$$

□

4 Malliavin calculus on Gaussian space

Before proceeding to the construction of Stein type estimators, we need to introduce some elements of analysis on Gaussian space, see e.g. [11]. This construction is valid in both frameworks (A) and (B). Given $u \in H$, let

$$X^u = X - u.$$

We fix $(h_n)_{n\geq 1}$ a total subset of H and let $\mathcal{S}$ denote the space of cylindrical functionals of the form

$$F = f_n\left(X^u(h_1), \ldots, X^u(h_n)\right), \tag{4.1}$$

where f_n is in the space of infinitely differentiable rapidly decreasing functions on $\mathbb{R}^n$, $n \geq 1$.

Definition 4.1. *The H-valued Malliavin derivative is defined as*

$$\nabla_t F = \sum_{i=1}^{n} h_i(t)\partial_i f_n\left(X^u(h_1), \ldots, X^u(h_n)\right),$$

for $F \in \mathcal{S}$ of the form (4.1).

It is known that ∇ is closable, cf. Proposition 1.2.1 of [11], and its closed domain will be denoted by $\mathrm{Dom}\,(\nabla)$.

Definition 4.2. *Let D be defined on* $\mathrm{Dom}\,(\nabla)$ *as*

$$D_t F := (\Gamma\nabla F)(t), \quad t \in [0,T], \quad F \in \mathrm{Dom}\,(\nabla).$$

Let $\delta : L^2_u(\Omega; H) \to L^2(\Omega, \mathbb{P}_u)$ denote the closable adjoint of ∇, i.e. the divergence operator under $\mathbb{P}_u$, which satisfies the integration by parts formula

$$\mathbb{E}_u[F\delta(v)] = \mathbb{E}_u[\langle v, \nabla F\rangle_H], \qquad F \in \mathrm{Dom}\,(\nabla), \quad \mathrm{v} \in \mathrm{Dom}\,(\delta), \tag{4.2}$$

with the relation

$$\delta(hF) = FX(h) - \langle h, \nabla F\rangle_H,$$

cf. [11], for $F \in \mathrm{Dom}\,(\nabla)$ and $h \in H$ such that $hF \in \mathrm{Dom}\,(\delta)$. Note that (4.2) is an infinite-dimensional version of the integration by parts (1.1), which can be proved e.g. using the countable Gaussian random variables constructed from X.

Lemma 4.3. *We have*

$$\mathbb{E}_u[FX_t^u] = \mathbb{E}_u[D_t F], \qquad t \in [0,T], \quad F \in \mathrm{Dom}\,(\nabla).$$

Proof.

(A) In the case of Paley-Wiener expansions we have

$$\begin{aligned}
\mathbb{E}_u[FX_t^u] &= \mathbb{E}_u[FX^u(\chi_t)] \\
&= \mathbb{E}_u[F\delta(\chi_t)] \\
&= \mathbb{E}_u[\langle \chi_t, \nabla F\rangle_H] \\
&= \mathbb{E}_u[\langle 1_{[0,t]}, \dot{\Gamma}\nabla F\rangle_{L^2([0,T],dt)}] \\
&= \mathbb{E}_u[(\Gamma\nabla F)(t)], \qquad F \in \mathrm{Dom}\,(\nabla), \quad \mathrm{t} \in [0,\mathrm{T}].
\end{aligned}$$

(B) In the case of Karhunen-Loève expansions we have

$$\begin{aligned}
\mathbb{E}_u[FX_t^u] &= \sum_{k=0}^{\infty} h_k(t)\,\mathbb{E}_u[FX^u(h_k)] \\
&= \sum_{k=0}^{\infty} h_k(t)\,\mathbb{E}_u[F\delta(h_k)] \\
&= \sum_{k=0}^{\infty} h_k(t)\,\mathbb{E}_u[\langle h_k, \nabla F\rangle_H] \\
&= \sum_{k=0}^{\infty} h_k(t)\,\mathbb{E}_u[\langle h_k, \Gamma\nabla F\rangle_{L^2([0,T],\mu)}] \\
&= \mathbb{E}_u[(\Gamma\nabla F)(t)], \qquad F \in \mathrm{Dom}\,(\nabla), \quad \mathrm{t} \in [0,\mathrm{T}].
\end{aligned}$$

□

Definition 4.4. *We define the Laplacian* Δ *by*

$$\Delta F = \mathrm{trace}_{L^2([0,T],d\mu)^{\otimes 2}} DDF = \int_0^T D_t D_t F \mu(dt)$$

on the space $\mathrm{Dom}\,(\Delta)$ *made of all* $F \in \mathrm{Dom}\,(\nabla)$ *such that* $D_t F \in \mathrm{Dom}\,(\nabla)$, $t \in [0,T]$, *and* $(D_t D_t F)_{t\in[0,T]} \in L^2([0,T],\mu)$, $\mathbb{P}$-*a.s.*

If $F \in \mathcal{S}$ has the form (4.1) we have

$$\Delta F = \sum_{i,j=1}^{n} \langle \Gamma h_i, \Gamma h_j \rangle_{L^2([0,T],\mu)} \partial_i \partial_j f_n \left(X^u(h_1), \ldots, X^u(h_n) \right).$$

Unlike the Gross Laplacian Δ_G defined by

$$\Delta_G F = \mathrm{trace}_{H^{\otimes 2}} \nabla \nabla F,$$

the operator Δ is closable, as shown in the following proposition.

Proposition 4.5. *Closability of Δ. For any sequence $(F_n)_{n\in\mathbb{N}}$ of random variables converging to 0 in $L^2(\Omega, \mathbb{P}_u)$ and such that $(\Delta F_n)_{n\in\mathbb{N}}$ converges in $L^2(\Omega, \mathbb{P}_u)$, we have*

$$\lim_{n\to\infty} \Delta F_n = 0.$$

Proof. Let $(G_n)_{n\in\mathbb{N}}$ a sequence in $\mathcal{S}$ converging to 0 in $L^2(\Omega, \mathbb{P}_u)$, and such that $(\Delta G_n)_{n\in\mathbb{N}}$ converges to F in $L^2(\Omega, \mathbb{P}_u)$. For all $G \in \mathcal{S}$ we have, in the notation of (A):

$$\begin{aligned}
|\langle \Delta G_n, G \rangle_{L^2(\Omega,\mathbb{P}_u)}| &= \left| \mathbb{E}_u \left[G \int_0^T D_t D_t G_n \mu(dt) \right] \right| \\
&= \left| \int_0^T \mathbb{E}_u[\langle \nabla D_t G_n, \chi_t \, G \rangle_H] \, \mu(dt) \right| \\
&= \left| \int_0^T \mathbb{E}_u[D_t G_n \, \delta(\chi_t \, G)] \, \mu(dt) \right| \\
&= \left| \int_0^T \mathbb{E}_u[\langle \nabla G_n, \chi_t \, \delta(\chi_t \, G) \rangle_H] \, \mu(dt) \right| \\
&= \left| \int_0^T \mathbb{E}_u[G_n \delta(\chi_t \, \delta(\chi_t \, G))] \, \mu(dt) \right| \\
&\leq \|G_n\|_{L^2(\Omega,\mathbb{P}_u)} \int_0^T \|\delta(\chi_t \, \delta(\chi_t \, G))\|_{L^2(\Omega,\mathbb{P}_u)} \, \mu(dt),
\end{aligned}$$

hence $\langle F, G \rangle_{L^2(\Omega,\mathbb{P}_u)} = 0, \ G \in \mathcal{S}$, which implies $F = 0$. □

We will say that a random variable F in $\mathrm{Dom}\,(\Delta)$ is Δ-superharmonic on Ω if

$$\Delta F(\omega) \leq 0, \qquad \mathbb{P}(d\omega) - a.s. \tag{4.3}$$

Remark 4.6. *In the independent increment case where $\gamma(s,t)$ is given by*

$$\gamma(s,t) = \int_0^{s\wedge t} \sigma_u^2 dt, \qquad s,t \in [0,T],$$

we have

$$\delta(v) = \int_0^T \dot{v}_t dX_t^u, \tag{4.4}$$

for every $\mathcal{F}_t$-adapted process $v \in L^2(\Omega; H, \mathbb{P}_u)$.

5 Superefficient drift estimators

Our aim is to construct a superefficient estimator of u of the form $X+\xi$, whose mean square error is strictly smaller than the minimax risk $\mathrm{R}(\gamma,\mu,\hat{u})$ of Proposition 3.3 when $\xi \in L^2([0,T]\times\Omega, \mathbb{P}_u \otimes \mu)$ is a suitably chosen stochastic process. This estimator will be biased and anticipating with respect to the Brownian filtration. In the next lemma we follow Stein's argument which uses integration by parts but we replace (1.1) by the duality relation (4.2) between the gradient and divergence operators on Gaussian space. The results of this section are valid in both frameworks (A) and (B).

Lemma 5.1. *Unbiased risk estimate. For any $\xi \in L^2(\Omega\times[0,T], \mathbb{P}_u\otimes\mu)$ such that $\xi_t \in \mathrm{Dom}\,(\nabla)$, $t\in[0,T]$, and $(D_t\xi_t)_{t\in[0,T]} \in L^1(\Omega\times[0,T],\mathbb{P}_u\otimes\mu)$, we have*

$$\mathbb{E}_u\left[\|X+\xi-u\|^2_{L^2([0,T],\mu)}\right] = \mathrm{R}(\gamma,\mu,\hat{u}) + \|\xi\|^2_{L^2(\Omega\times[0,T],\mathbb{P}_u\otimes\mu)} + 2\,\mathbb{E}_u\left[\int_0^T D_t\xi_t\mu(dt)\right]. \tag{5.1}$$

Proof. We have

$$\begin{aligned}
\mathbb{E}_u\left[\|X+\xi-u\|^2_{L^2([0,T],d\mu)}\right] &= \mathbb{E}_u\left[\int_0^T \left|X_t^u+\xi_t\right|^2\mu(dt)\right] \\
&= \mathbb{E}_u\left[\int_0^T |X_t^u|^2\mu(dt)\right] + \|\xi\|^2_{L^2(\Omega\times[0,T],\mathbb{P}_u\otimes\mu)} + 2\,\mathbb{E}_u\left[\int_0^T X_t^u\xi_t\mu(dt)\right] \\
&= \mathrm{R}(\gamma,\mu,\hat{u}) + \|\xi\|^2_{L^2(\Omega\times[0,T],\mathbb{P}_u\otimes\mu)} + 2\,\mathbb{E}_u\left[\int_0^T X_t^u\xi_t\mu(dt)\right],
\end{aligned}$$

and apply Lemma 4.3 to obtain (5.1). □

The next proposition specializes the above lemma to processes ξ of the form

$$\xi_t = D_t \log F, \qquad t \in [0,T],$$

where F is an a.s. strictly positive and sufficiently smooth random variable.

Proposition 5.2. *Logarithmic gradient. Stein-type estimator. For any $\mathbb{P}$-a.s. positive random variable $F \in \mathrm{Dom}\,(\nabla)$ such that $D_t F \in \mathrm{Dom}\,(\nabla)$, $t \in [0,T]$, and $(D_t D_t F)_{t\in[0,T]} \in L^1(\Omega \times [0,T], \mathbb{P}_u \otimes \mu)$, we have*

$$\mathbb{E}_u\left[\|X + D\log F - u\|^2_{L^2([0,T],d\mu)}\right] = \mathrm{R}(\gamma,\mu,\hat{u}) - \mathbb{E}_u\left[\|D\log F\|^2_{L^2([0,T],\mu)}\right] + 2\,\mathbb{E}_u\left[\frac{\Delta F}{F}\right].$$

Proof. From (5.1) we have

$$\begin{aligned}
&\mathbb{E}_u\left[\|X + D\log F - u\|^2_{L^2([0,T],\mu)}\right] \\
&\quad = \mathrm{R}(\gamma,\mu,\hat{u}) + \|D\log F\|^2_{L^2(\Omega\times[0,T],\mathbb{P}_u\otimes\mu)} + 2\,\mathbb{E}_u\left[\int_0^T D_t D_t \log F \mu(dt)\right] \\
&\quad = \mathrm{R}(\gamma,\mu,\hat{u}) + \mathbb{E}_u\left[\int_0^T \left(\left|\frac{D_t F}{F}\right|^2 + 2 D_t D_t \log F\right)\mu(dt)\right],
\end{aligned}$$

and we use the relation

$$\left|\frac{D_t F}{F}\right|^2 + 2 D_t D_t \log F = 2\frac{D_t D_t F}{F} - \left|\frac{D_t F}{F}\right|^2, \qquad t \in [0,T].$$

□

From the above proposition it suffices that F be Δ-superharmonic for $X + D\log F$ to be superefficient. In this case we have

$$\mathbb{E}_u\left[\|X + D\log F - u\|^2_{L^2([0,T],d\mu)}\right] \leq \mathrm{R}(\gamma,\mu,\hat{u}) - \mathbb{E}_u\left[\|D\log F\|^2_{L^2([0,T],d\mu)}\right], \tag{5.2}$$

with equality in (5.2) when F is Δ-harmonic.

In the next proposition we show that the Δ-superharmonicity of F is not necessary for $X + D\log F$ to be superefficient, namely the Δ-superharmonicity of F can be replaced by the Δ-superharmonicity of $\sqrt{F}$, which is a weaker assumption, see [5] in the finite dimensional case. In particular, $X + D\log F$ is a superefficient estimator of u if $\Delta\sqrt{F} < 0$ on a set of strictly positive $\mathbb{P}$-measure.

Proposition 5.3. *Stein-type estimator. For any $\mathbb{P}$-a.s. positive random variable $F \in \mathrm{Dom}(\nabla)$ such that $D_t F \in \mathrm{Dom}(\nabla)$, $t \in [0,T]$, and $(D_t D_t F)_{t\in[0,T]} \in L^1(\Omega \times [0,T], \mathbb{P}_u \otimes \mu)$, we have*

$$\mathbb{E}_u\left[\|X + D\log F - u\|^2_{L^2([0,T],d\mu)}\right] = \mathrm{R}(\gamma,\mu,\hat{u}) + 4\,\mathbb{E}_u\left[\frac{\Delta\sqrt{F}}{\sqrt{F}}\right]. \tag{5.3}$$

Proof. For any $F \in \mathrm{Dom}(\nabla)$ such that $F > 0$, $\mathbb{P}$-a.s., and $\sqrt{F} \in \mathrm{Dom}(\Delta)$, we have

$$2\frac{D_t D_t F}{F} - \left|\frac{D_t F}{F}\right|^2 = \frac{2}{\sqrt{F}} D_t\left(\frac{D_t F}{\sqrt{F}}\right) = \frac{4}{\sqrt{F}} D_t D_t \sqrt{F}, \qquad t \in [0,T],$$

which implies

$$4\frac{\Delta\sqrt{F}}{\sqrt{F}} = 2\frac{\Delta F}{F} - \int_0^T |D_t \log F|^2 \mu(dt), \tag{5.4}$$

and allows us to conclude from Lemma 5.1. □

Relation (5.3) extends to any $F \in \mathrm{Dom}(\nabla)$ such that $\sqrt{F} \in \mathrm{Dom}(\Delta)$, and $F \geq 0$, $\Delta\sqrt{F} \leq 0$, $\mathbb{P}$-a.s.

In case $(X_t)_{t\in[0,T]}$ is a Brownian motion with constant variance $\sigma_t = \sigma$, $t \in [0,T]$, we have

$$\mathbb{E}_u\left[\|X + D\log F - u\|^2_{L^2([0,T])}\right] \leq \frac{\sigma^2 T^2}{2} + 4\,\mathbb{E}_u\left[\frac{\Delta\sqrt{F}}{\sqrt{F}}\right]. \tag{5.5}$$

Given $(X_t^{\sigma,1})_{t\in[0,T]}, \ldots, (X_t^{\sigma,N})_{t\in[0,T]}$ are N independent samples of $(X_t)_{t\in[0,T]}$, the process $\bar{X}$ defined in (2.1) satisfies

$$\mathbb{E}_u^{\sigma/N}\left[\|\bar{X} + D\log F - u\|^2_{L^2([0,T])}\right] = \frac{1}{N}\mathrm{R}(\sigma,\mu,\hat{u}) + \frac{4}{N^2}\,\mathbb{E}_u\left[\frac{\Delta\sqrt{F}}{\sqrt{F}}\right].$$

As in [15], the superefficient estimators constructed in this way are minimax in the sense that from Proposition 3.3 and Proposition 5.2, for all $u \in H$ we have

$$\mathbb{E}_u\left[\|X + D\log F - u\|^2_{L^2([0,T],\mu)}\right] < \mathrm{R}(\gamma,\mu,\hat{u}) = \inf_{\xi}\sup_{v\in\Omega}\mathbb{E}_v\left[\int_0^T |\xi_t - v_t|^2 dt\right],$$

provided $\Delta\sqrt{F} < 0$ on a set of strictly positive $\mathbb{P}$-measure, thus showing that the minimax estimator $\hat{u} = (X_t)_{t\in[0,T]}$ is inadmissible.

Both estimators $X_t + D_t \log F$ and $X_t + \mathbb{E}_u[D_t \log F \mid \mathcal{F}_t]$ have bias

$$b_t = \mathbb{E}_u[X_t + D_t \log F - u_t] = \mathbb{E}_u[D_t \log F], \qquad t \in [0,T],$$

which can be bounded as follows from (5.4):

$$\begin{aligned}
\|b\|^2_{L^2([0,T],\mu)} &= \int_0^T |\,\mathbb{E}_u[D_t \log F]|^2 dt \\
&\leq \mathbb{E}_u\left[\int_0^T |D_t \log F|^2 dt\right] \\
&= 2\,\mathbb{E}_u\left[\frac{\Delta F}{F}\right] - 4\,\mathbb{E}_u\left[\frac{\Delta\sqrt{F}}{\sqrt{F}}\right].
\end{aligned}$$

Remark 5.4. *In the independent increment case of Section 2, the formulas obtained in this section also hold for u an adapted process in $L^2(\Omega \times [0,T])$. However, in this case the computation of the gradient $D \log F$ requires in principle the knowledge of X^u, except when u is deterministic, in which case the knowledge of X is sufficient. Thus, assuming u to be deterministic will be necessary for the applications of Section 7.*

6 Superharmonic functionals

In this section we give examples of nonnegative superharmonic functionals with respect to the Laplacian Δ. We start by reviewing the construction of such functionals using potential theory on the Gaussian space $(\Omega, H, \mathbb{P})$, and next we turn to cylindrical functionals which will be used in the numerical applications of Section 7. We assume that $(\Gamma h_k)_{k\geq 1}$ is orthogonal in $L^2([0,T], d\mu)$, and we let

$$\lambda_k = \|\Gamma h_k\|_{L^2([0,T],\mu)}, \qquad k \geq 1.$$

The sequence $(h_k)_{k\geq 1}$ can be realized as the solution of the eigenvalue problem

$$\Gamma h_k = -\lambda_k^2 \ddot{h}_k, \qquad \dot{h}_k(T) = 0, \qquad k \geq 1, \tag{6.1}$$

in case (A), provided $\mu(dt) = dt$, and

$$\Gamma h_k = \lambda_k^2 h_k, \qquad k \geq 1,$$

in case (B) for general μ.

Potentials

We refer to [7] and [6] for the notion of harmonicity on the Wiener space with respect to the Gross Laplacian. From our orthonormality assumption on $(h_k)_{k\geq 1}$, the Laplacian Δ is written as

$$\Delta F = \sum_{i=1}^{n} \partial_i^2 f_n \left(\int_0^T \dot{h}_1(s) dX_s^u, \ldots, \int_0^T \dot{h}_n(s) dX_s^u \right)$$

on cylindrical functionals. Let $(W_t^\Omega)_{t\geq 0}$ denote the standard Ω-valued Wiener process with generator $\frac{1}{2}\Delta_G$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$, represented as

$$W_t^\Omega = \sum_{n=1}^{\infty} \int_0^{\cdot} \dot{h}_n(s) \sigma_s^2 ds \frac{\beta_n(t)}{\|h_n\|_H}, \qquad t \in \mathbb{R}_+, \tag{6.2}$$

where $(\beta_n(t))_{t\in\mathbb{R}_+}$, $n \geq 1$, are independent standard Brownian motions on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$, given as

$$\beta_n(t) = \int_0^T \frac{\dot{h}_n(r)}{\|h_n\|_H} dW_t^\Omega(r), \qquad t \in \mathbb{R}_+, \quad n \geq 1.$$

We have the covariance relation

$$\tilde{\mathbb{E}} \left[\int_0^T \dot{v}_1(r) dW_s^\Omega(r) \int_0^T \dot{v}_2(r) dW_t^\Omega(r) \right] = (s \wedge t) \langle v_1, v_2 \rangle_H, \qquad s, t \in \mathbb{R}_+, \quad v_1, v_2 \in H.$$

In other terms we have

$$\begin{aligned}
\tilde{\mathbb{E}}[W_t^\Omega(a) W_s^\Omega(b)] &= (s \wedge t) \sum_{n=1}^{\infty} \frac{\int_0^a \dot{h}_n(s) \sigma_s^2 ds \int_0^b \dot{h}_n(s) \sigma_s^2 ds}{\|h_n\|_H^2} \\
&= (s \wedge t) \left\langle \sum_{n=1}^{\infty} \frac{\dot{h}_n}{\|h_n\|_H^2} \int_0^a \dot{h}_n(s) \sigma_s^2 ds, \sum_{n=1}^{\infty} \frac{\dot{h}_n}{\|h_n\|_H^2} \int_0^b \dot{h}_n(s) \sigma_s^2 ds \right\rangle_{L^2([0,T], \sigma_t^2 dt)} \\
&= (s \wedge t) \langle \mathbf{1}_{[0,a]}, \mathbf{1}_{[0,b]} \rangle_{L^2([0,T], \sigma_t^2 dt)} \\
&= (s \wedge t) \int_0^{a\wedge b} \sigma_r^2 dr, \qquad 0 \leq a, b \leq T, \quad s, t \in \mathbb{R}_+,
\end{aligned}$$

which shows that $(W_t^\Omega(a))_{a\in[0,T]}$ is a continuous Gaussian martingale with quadratic variation $\sigma_a^2 da$ for fixed $t \in \mathbb{R}_+$.

Denote by $(B_t)_{t\in\mathbb{R}_+}$ the H-valued Wiener process represented as

$$B_t = \sum_{n=0}^{\infty} \frac{h_n}{\|h_n\|_H^2} \beta_n(t),$$

with $\beta_n(t) = \langle B_t, h_n\rangle_H$, $n \geq 1$, and covariance

$$\tilde{\mathbb{E}}[\langle B_s, h_n\rangle_H \langle B_t, h_m\rangle_H] = \mathbf{1}_{\{n=m\}}(s \wedge t), \qquad s, t \in \mathbb{R}_+,$$

i.e.

$$\tilde{\mathbb{E}}[\langle B_t, v_1\rangle_H \langle B_s, v_2\rangle_H] = (s \wedge t)\langle Q v_1, v_2\rangle_H, \qquad s, t \in \mathbb{R}_+, \quad v_1, v_2 \in H,$$

where $Q : H \to H$ is the operator with eigenvalues $\{\|h_n\|_H^{-2} \ : \ n \geq 1\}$ in the Hilbert basis $(h_n)_{n\geq 1}$. Itô's formula for Hilbert-valued Wiener processes, cf. Theorem 4.17 of [3], shows that

$$F(B_t) = F(B_0) + \int_0^t \langle DF(B_s), dB_s\rangle_H + \frac{1}{2}\int_0^t \Delta F(B_s) ds, \qquad F \in \mathcal{S},$$

hence $(B_t)_{t\in\mathbb{R}_+}$ has generator $\frac{1}{2}\Delta$.

Dynkin's formula, cf. [4], Theorem 5.1, shows that for all stopping time τ such that $\tilde{\mathbb{E}}[\tau \mid B_0 = \omega] < \infty$ we have, $\mathbb{P}^\sigma(d\omega)$-a.s.:

$$\tilde{\mathbb{E}}[F(B_\tau) \mid B_0 = \omega] - F(\omega) = \frac{1}{2}\tilde{\mathbb{E}}\left[\int_0^\tau \Delta F(B_s) ds \mid B_0 = \omega\right],$$

hence $\Delta F \leq 0$ implies

$$F(\omega) \geq \tilde{\mathbb{E}}[F(B_\tau) \mid B_0 = \omega].$$

For $r > 0$, let

$$\tau_r = \inf\{t \in \mathbb{R}_+ \ : \ B_t \notin \mathfrak{B}_r(B_0)\}$$

denotes the first exit time of $(B_t)_{t\in[0,T]}$ from the open ball $\mathfrak{B}_r(\omega)$ of radius $r > 0$, centered at $B_0 = \omega \in \Omega$. We have the following converse.

Proposition 6.1. *Let $F \in \mathrm{Dom}\,(\Delta)$ be such that ΔF is continuous on Ω, and assume that there exists $r_0 > 0$ such that*

$$F(\omega) \geq \tilde{\mathbb{E}}[F(B_{\tau_r}) \mid B_0 = \omega], \qquad \mathbb{P}_u^\sigma(d\omega) - a.s., \quad 0 < r < r_0. \tag{6.3}$$

Then F is Δ-superharmonic on Ω in the sense of Relation (4.3).

Proof. From Remark 3, page 134 of [4], we have

$$\frac{1}{2}\Delta F(\omega) = \lim_{n\to\infty} \frac{\tilde{\mathbb{E}}[F(B_{\tau_{1/n}}) \mid B_0 = \omega] - F(\omega)}{\tilde{\mathbb{E}}[\tau_{1/n} \mid B_0 = \omega]}, \tag{6.4}$$

which shows that $\Delta F \leq 0$ when (6.3) is satisfied. □

This yields in particular the following class of Δ-superharmonic functionals.

Proposition 6.2. *Let the potential of $F \geq 0$ be defined by*

$$G(\omega) = \int_0^{+\infty} \tilde{\mathbb{E}}[F(B_t) \mid B_0 = \omega]dt, \qquad \mathbb{P}_u^\sigma(d\omega) - a.s., \tag{6.5}$$

assume that $G \in \mathrm{Dom}\,(\Delta)$ and that ΔG is continuous on Ω. Then G is a Δ-superharmonic on Ω.

Proof. For all $r > 0$ we have

$$\begin{aligned} G(\omega) &= \tilde{\mathbb{E}}\left[\int_0^{\tau_r} F(B_t)dt \middle| B_0 = \omega\right] + \tilde{\mathbb{E}}[G(B_{\tau_r}) \mid B_0 = \omega] \\ &\geq \tilde{\mathbb{E}}[G(B_{\tau_r}) \mid B_0 = \omega], \end{aligned}$$

which shows that G is Δ-superharmonic. □

Note that if F is bounded with bounded support in Ω then G is bounded on Ω, see e.g. Remark 3.5 of [7].

Convolution

Positive superharmonic functionals can also be obtained by convolution, i.e. if F is Δ-superharmonic and G is positive and sufficiently integrable, then

$$\omega \mapsto \int_\Omega G(\tilde{\omega})F(\omega - \tilde{\omega})\mathbb{P}^\sigma(d\tilde{\omega})$$

is positive and Δ-superharmonic.

Cylindrical functionals

Superharmonic functionals on Gaussian space can also be constructed as cylindrical functionals, by composition with finite-dimensional functions. Here we use the expansions of case (A). From the expression of Δ on cylindrical functionals

$$\Delta F = \sum_{i=1}^{n} \partial_i^2 f_n \left(\lambda_1^{-1} X^u(h_1), \ldots, \lambda_n^{-1} X^u(h_n) \right),$$

we check that

$$F = f_n \left(\lambda_1^{-1} X^u(h_1), \ldots, \lambda_n^{-1} X^u(h_n) \right)$$

is superharmonic on Ω if and only if f_n is superharmonic on $\mathbb{R}^n$. Given $a \in \mathbb{R}$ and $b \in \mathbb{R}^n$, let $f_{n,a,b} : \mathbb{R}^n \to \mathbb{R}$ be defined as

$$f_{n,a,b}(x_1, \ldots, x_n) = \|x + b\|^a = ((x_1 + b_1)^2 + \cdots + (x_n + b_n)^2)^{a/2},$$

then $\sqrt{f_{n,a,b}}$ is superharmonic on $\mathbb{R}^n$, $n \geq 3$, if and only if $a \in [4 - 2n, 0]$. Let

$$F_{n,a,b} = f_{n,a,b} \left(\lambda_1^{-1} X^u(h_1), \ldots, \lambda_n^{-1} X^u(h_n) \right).$$

We have

$$D_t \log F_{n,a,b} = a \sum_{i=1}^{n} \frac{\lambda_i^{-1} \Gamma h_i(t) \left(b_i + \lambda_i^{-1} X^u(h_i) \right)}{\left| b_1 + \lambda_1^{-1} X^u(h_1) \right|^2 + \cdots + \left| b_n + \lambda_n^{-1} X^u(h_n) \right|^2},$$

and

$$\Delta \sqrt{F_{n,a,b}} \;=\; \sum_{i=1}^{n} \partial_i^2 \sqrt{f_{n,a,b}} \left(\lambda_1^{-1} X^u(h_1), \ldots, \lambda_n^{-1} X^u(h_n) \right),$$

since $(\Gamma h_k)_{k \geq 1}$ is orthogonal in $L^2([0,T], dt)$, hence

$$\frac{\Delta \sqrt{F_{n,a,b}}}{\sqrt{F_{n,a,b}}} \;=\; \frac{a(n - 2 + a/2)/2}{\left| b_1 + \lambda_1^{-1} X^u(h_1) \right|^2 + \cdots + \left| b_n + \lambda_n^{-1} X^u(h_n) \right|^2},$$

is negative if $4 - 2n \leq a \leq 0$, which is minimal for $a = 2 - n$. We also have

$$\frac{\Delta F_{n,a,b}}{F_{n,a,b}} \;=\; \frac{a(n + a - 2)}{\left| b_1 + \lambda_1^{-1} X^u(h_1) \right|^2 + \cdots + \left| b_n + \lambda_n^{-1} X^u(h_n) \right|^2},$$

which is negative for $a \in [2-n, 0]$ and vanishes for $a = 2-n$. In this case the estimator is given by

$$D_t \log F_{n,2-n,b} = -(n-2) \sum_{i=1}^{n} \frac{\lambda_i^{-1} \left(b_i + \lambda_i^{-1} X^u(h_i)\right) \Gamma h_i(t)}{\left|b_1 + \lambda_1^{-1} X^u(h_1)\right|^2 + \cdots + \left|b_n + \lambda_n^{-1} X^u(h_n)\right|^2},$$

and from Proposition 5.2, inequality (5.2) actually also holds as an equality:

$$\mathbb{E}_u \left[\|X + D \log F_{n,2-n,b} - u\|^2_{L^2([0,T],dt)} \right] = \mathrm{R}(\sigma, \mu, \hat{u}) - \mathbb{E}_u \left[\int_0^T |D_t \log F_{n,2-n,b}|^2 dt \right], \tag{6.6}$$

with

$$\|D \log F_{n,2-n,b}\|^2_{L^2([0,T],dt)} = \frac{(n-2)^2}{\left|b_1 + \lambda_1^{-1} X^u(h_1)\right|^2 + \cdots + \left|b_n + \lambda_n^{-1} X^u(h_n)\right|^2}. \tag{6.7}$$

Note that when u is deterministic, any superharmonic functional of the form

$$f_n \left(\lambda_1^{-1} X^u(h_1), \ldots, \lambda_n^{-1} X^u(h_n) \right),$$

can be replaced with

$$f_n \left(\lambda_1^{-1} X(h_1), \ldots, \lambda_n^{-1} X(h_n) \right),$$

which retains the same harmonicity property, and can be directly computed from an observation of X.

The Stein type estimator of u is given by

$$X_t + D_t \log F_{n,2-n,b}, \qquad t \in [0,T],$$

with

$$b_i = \lambda_i^{-1} \langle u, h_i \rangle, \qquad i = 1, \ldots, n,$$

i.e.

$$D_t \log F_{n,2-n,b} = -(n-2) \frac{[\Pi_n X]_t}{\|\Pi_n X\|^2_{L^2([0,T],dt)}},$$

where Π_n denotes the orthogonal projection

$$\Pi_n X(t) := \sum_{k=1}^{n} \lambda_k^{-1} X(h_k) \Gamma h_k(t) = \sum_{k=1}^{n} \lambda_k^{-1} \left(b_k + \lambda_k^{-1} X^u(h_k) \right) \Gamma h_k(t).$$

We have

$$
\begin{aligned}
\|D \log F_{n,2-n,b}\|^2_{L^2([0,T]\times\Omega,\mathbb{P}_u\otimes dt)} &= -4\,\mathbb{E}_u\left[\frac{\Delta\sqrt{F_{n,2-n,b}}}{\sqrt{F_{n,2-n,b}}}\right] \\
&= (n-2)^2\,\mathbb{E}_u\left[\frac{1}{\left|\lambda_1^{-1}X(h_1)\right|^2+\cdots+|\lambda_n^{-1}X(h_n)|^2}\right] \\
&= (n-2)^2\,\mathbb{E}_u\left[\|\Pi_n X\|^{-2}_{L^2([0,T],dt)}\right],
\end{aligned}
$$

and

$$
\mathbb{E}_u\left[\|X + D\log F_{n,2-n,b} - u\|^2_{L^2([0,T],dt)}\right] = \mathrm{R}(\sigma,\mu,\hat{u}) - (n-2)^2\,\mathbb{E}_u\left[\|\Pi_n X\|^{-2}_{L^2([0,T],dt)}\right].
$$

Note that the estimator

$$
X_t - (n-2)\frac{[\Pi_n X]_t}{\|\Pi_n X\|^2_{L^2([0,T],dt)}}, \qquad t \in [0,T],
$$

is of James-Stein type, but it is not a shrinkage operator. Another difference with James-Stein estimators is that here the denominator consists in a sum of squared Gaussians with different variances.

Given $(X_t^1)_{t\in[0,T]},\ldots,(X_t^N)_{t\in[0,T]}$, N independent samples of $(X_t)_{t\in[0,T]}$, the process

$$
\bar{X}_t = \frac{1}{N}\left(X_t^1+\cdots+X_t^N\right)
$$

is a Brownian motion with drift u and quadratic variation $\sigma_t^2 dt/N$ under $\mathbb{P}_u$, and can be used for both efficient and Stein type estimation.

7 Numerical application

In this section we present numerical simulations which allow us to measure the efficiency of our estimators. We use the framework of case (A) and the superharmonic functionals constructed as cylindrical functionals in the previous section, and we assume that $u \in H$ is deterministic.

We work in the independent increment framework of Section 2 and we additionally assume that $\sigma_t = \sigma$ is constant, $t \in [0,T]$, i.e. $(X_t)_{t\in[0,T]}$ is a Brownian motion with variance σ^2, $\Gamma h(t) = \sigma^2 h(t)$, $t \in [0,T]$, and

$$\mathrm{R}(\sigma,\mu,\hat{u}) = \frac{\sigma^2}{2}T^2.$$

Letting

$$h_n(t) = \frac{\sqrt{2T}}{\sigma\pi\,(n-1/2)}\sin\left(\left(n-\frac{1}{2}\right)\frac{\pi t}{T}\right), \qquad t\in[0,T], \quad n\geq 1,$$

i.e.

$$\dot{h}_n(t) = \frac{1}{\sigma}\sqrt{\frac{2}{T}}\cos\left(\left(n-\frac{1}{2}\right)\frac{\pi t}{T}\right), \qquad t\in[0,T], \quad n\geq 1,$$

provides an orthonormal basis $(h_n)_{n\geq 1}$ of H such that $(\Gamma h_k)_{k\geq 1}$ is orthogonal in $L^2([0,T],dt)$, with

$$\lambda_n = \frac{\sigma T}{\pi(n-1/2)}, \qquad n\geq 1,$$

solution of (6.1). The estimator of u will be given by

$$D_t \log F_{n,2-n,b} = -(n-2)\sqrt{\frac{2}{T}}\sum_{k=1}^{n}\frac{X(h_k)}{\left|\lambda_1^{-1}X(h_1)\right|^2+\cdots+|\lambda_n^{-1}X(h_n)|^2}\sin\left(\left(k-\frac{1}{2}\right)\frac{\pi t}{T}\right),$$

For simulation purposes we will use $X + D\log F$, and construct the (nondrifted) Brownian motion $(X_t^u)_{t\in[0,T]}$ via the Paley-Wiener expansion

$$X_t^u = \sigma^2\sum_{n=1}^{\infty}\eta_n h_n(t) = \sigma\frac{\sqrt{2T}}{\pi}\sum_{n=1}^{\infty}\eta_n\frac{\sin\left(\left(n-\frac{1}{2}\right)\frac{\pi t}{T}\right)}{\left(n-\frac{1}{2}\right)}, \tag{7.1}$$

where $(\eta_n)_{n\geq 1}$ are independent standard Gaussian random variables with unit variance under $\mathbb{P}_u$ and

$$\eta_n = \int_0^T \dot{h}_n(s)dX_s^u, \qquad n\geq 1.$$

In this case we have

$$\begin{aligned} & D_t \log F_{n,2-n,b} \\ = \; & -(n-2)\sqrt{\frac{2}{T}}\sum_{k=1}^{n}\frac{\eta_k+\langle u,h_k\rangle}{\sum_{l=1}^{n}\lambda_l^{-2}\left(\eta_l+\langle u,h_l\rangle\right)^2}\sin\left(\left(k-\frac{1}{2}\right)\frac{\pi t}{T}\right). \end{aligned} \tag{7.2}$$

Recall that the improvement obtained in comparison with the efficient estimator $\hat{u}$ is not obtained pathwise, but in expectation. The gain of the superefficient estimator $X + D \log F_{n,2-n,b}$ compared to the efficient estimator $\hat{u}$ is given by

$$G(u, \sigma, T, n) := -\frac{4}{\mathrm{R}(\sigma, \mu, \hat{u})} \mathbb{E}_u \left[\frac{\Delta \sqrt{F_{n,2-n,b}}}{\sqrt{F_{n,2-n,b}}} \right]$$

as a function of $n \geq 3$. From (6.6) and (7.2) we have

$$G(u, \sigma, T, n) = 2(n-2)^2 \, \mathbb{E} \left[\left(\sum_{l=1}^{n} \left(\pi \left(l - \frac{1}{2} \right) (\eta_l + \langle u, h_l \rangle) \right)^2 \right)^{-1} \right], \qquad (7.3)$$

hence $G(u, \sigma, T, n)$ converges to

$$(n-2)^2 \frac{8}{\pi^2} \mathbb{E} \left[\left(\sum_{l=1}^{n} (2l-1)^2 \, \eta_l^2 \right)^{-1} \right], \qquad (7.4)$$

as σ tends to infinity. The quantity (7.4) can be evaluated as a Gaussian integral to yield (1.2). Unlike in the classical Stein method, we stress that here n becomes a free parameter and there is some interest in determining the values of n which yield the best performance.

Proposition 7.1. *For all $\sigma, T > 0$, and $u \in H$ we have*

$$G(u, \sigma, T, n) \simeq \frac{6}{n\pi^2}$$

as n goes to infinity.

Proof. Let

$$S_n = \sum_{l=1}^{n} \left(\pi \left(n - l + \frac{1}{2} \right) (\eta_l + \langle u, h_l \rangle) \right)^2, \qquad n \geq 1.$$

We have

$$G(\alpha, \sigma, T, n) = 2(n-2)^2 \, \mathbb{E} \left[\frac{1}{S_n} \right],$$

and by the strong law of large numbers, $n(n-2)^2 {S_n}^{-1}$ converges to $3/\pi^2$ as n goes to infinity, since

$$\lim_{n \to \infty} \frac{\mathbb{E}[S_n]}{n^3} = \frac{\pi^2}{4} \lim_{n \to \infty} \frac{1}{n^3} \sum_{i=1}^{n} (2i-1)^2 = \frac{\pi^2}{3}.$$

Now for all $n > 10$ we have

$$\begin{aligned}
\mathbb{E}_u\left[\left(\frac{(n-2)^3}{S_n}\right)^2\right] &= \mathbb{E}\left[\Lambda(u)\left(\frac{(n-2)^3}{S_n}\right)^2\right] \\
&\leq n^2\pi^2\,\mathbb{E}\left[\Lambda(u)^2\right]^{1/2}\mathbb{E}\left[\left(\sum_{l=1}^{[n/2]}\left(1-\frac{l}{n}+\frac{1}{2n}\right)^2\eta_l^2\right)^{-4}\right]^{1/2} \\
&\leq n^2\frac{4}{\pi^4}\,\mathbb{E}\left[\Lambda(u)^2\right]^{1/2}\mathbb{E}\left[\left(\sum_{l=1}^{[n/2]}\eta_l^2\right)^{-4}\right]^{1/2} \\
&\leq \frac{4n^2}{\pi^4}\,\mathbb{E}\left[\Lambda(u)^2\right]^{1/2}\left(\prod_{k=1}^{4}([n/2]-2k)\right)^{-1/4},
\end{aligned}$$

hence n^3/S_n is uniformly integrable in $n > 16$, where $[n/2]$ denotes the integer part of $n/2$. This concludes the proof. $\square$

In the sequel we choose $u_t = \alpha t$, $t \in [0,T]$, $\alpha \in \mathbb{R}$. Figure 7.1 gives a sample path representation of the process $X + D\log F$.

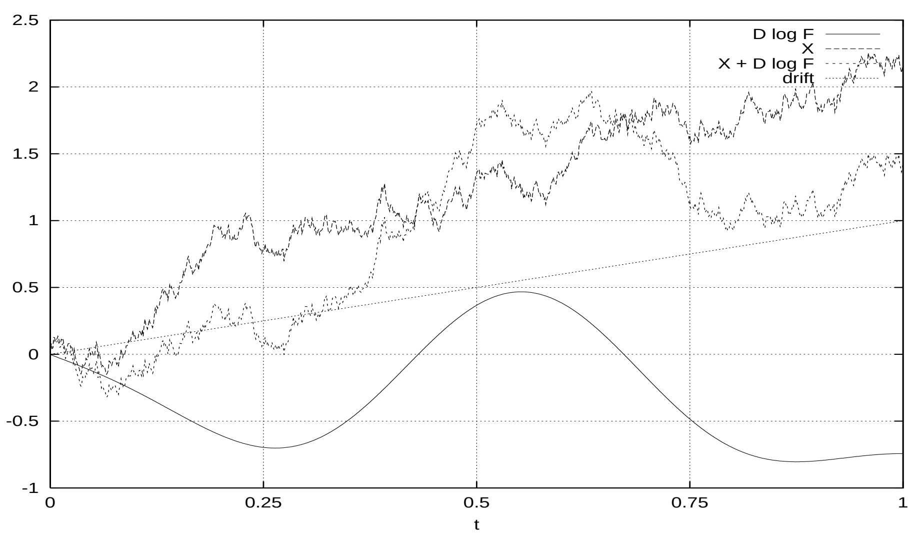

Figure 7.1: $u(t) = t,\ t \in [0,T];\ n = 5.$

In this case, from (7.3) we have

$$G(\alpha,\sigma,T,n) = 2(n-2)^2\,\mathbb{E}\left[\left(\sum_{l=1}^{n}\left(\pi\left(l-\frac{1}{2}\right)\eta_l - \alpha\frac{\sqrt{2T}}{\sigma}(-1)^l\right)^2\right)^{-1}\right],$$

from which it follows that $G(\alpha, \sigma, T, n)$ converges to

$$(n-2)^2 \frac{8}{\pi^2} \mathbb{E}\left[\left(\sum_{l=1}^{n} (2l-1)^2 \eta_l^2\right)^{-1}\right],$$

when $\alpha^{-2}\sigma^2/T$ tends to infinity, and is equivalent to

$$\left(1 - \frac{2}{n}\right)^2 \frac{\sigma^2}{\alpha^2 T}$$

as $\alpha^{-2}\sigma^2/T$ tends to 0. Figure 7.2 represents the gain in percentage of the super-efficient estimator $X + \sigma^2 D \log F_{n,2-n,b}$ compared to the efficient estimator $\hat{u}$ using Monte-Carlo simulations, i.e. we represent $100 \times G(\alpha, \sigma, T, n)$ as a function of $n \geq 3$.

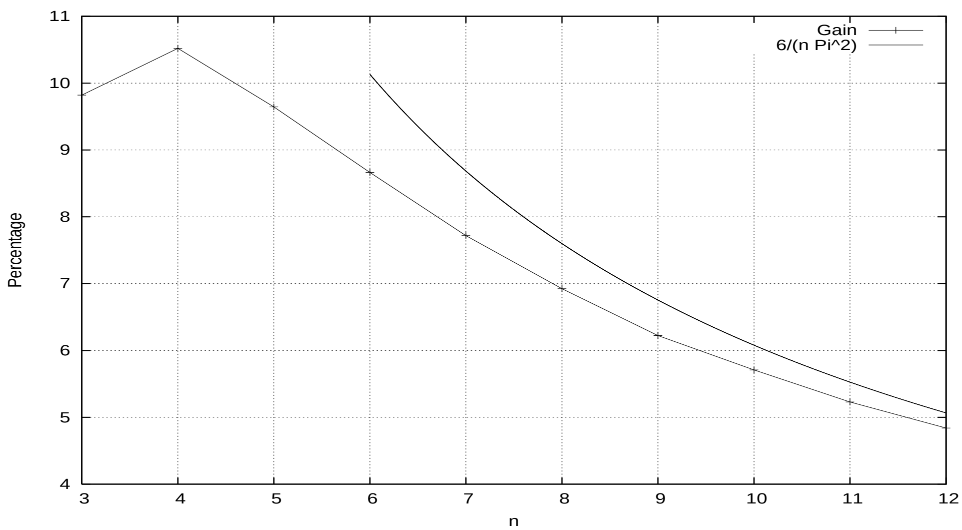

Figure 7.2: Percentage gain as a function of n for 10000 samples and $\alpha = \sigma = T = 1$.

An optimal value

$$n_{\text{opt}} = \operatorname{argmax} \{G(\alpha, \sigma, T, n) \ : \ n \geq 3\}$$

of n exists in general and is equal to 4 when $\alpha = \sigma = T = 1$. Figure 7.3 shows the variation of the gain as a function of n and T for $\alpha = \sigma = 1$.

Figure 7.4 represents the variation of the gain as a function of n and σ.

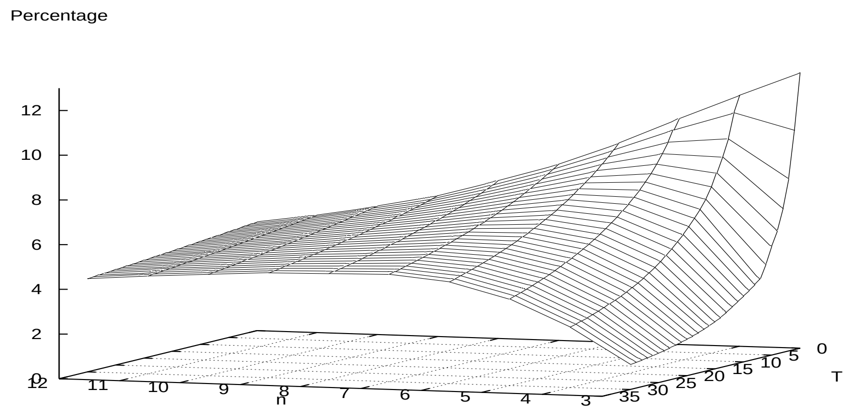

Figure 7.3: Gain as a function of n and T.

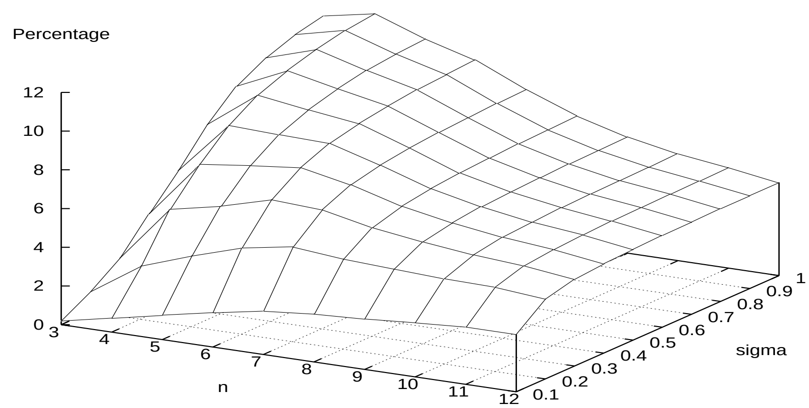

Figure 7.4: Gain as a function of n and σ.

8 Appendix

The next Proposition is classical in the framework of Gaussian filtering and is needed in Section 2 for Bayes estimation. Its proof is stated for completeness since we did not find it in the literature.

Proposition 8.1. *Let Z be a Gaussian process with covariance operator Γ_τ and drift*

$v \in H$, and assume that X is a Gaussian process with drift Z and quadratic covariance operator Γ given Z. Then, conditionally to X, Z has drift

$$f \mapsto \langle \chi_t, (\Gamma + \Gamma_\tau)^{-1}\Gamma v\rangle + X((\Gamma + \Gamma_\tau)^{-1}\Gamma_\tau f\ \chi_t) \quad \textit{and covariance} \quad \Gamma_\tau(\Gamma + \Gamma_\tau)^{-1}\Gamma.$$

Proof. For convenience of notation, let

$$V(f) = \langle f, (\Gamma + \Gamma_\tau)^{-1}\Gamma v\rangle, \qquad f \in H.$$

For all $f, g \in H$ we have:

$$\begin{aligned}
\mathbb{E}\left[\exp(iX(f))\right] &= \mathbb{E}\left[\mathbb{E}\left[\exp(iX(f)) \Big| Z\right]\right] \\
&= \mathbb{E}\left[\exp\left(iZ(f) - \frac{1}{2}\langle f, \Gamma f\rangle\right)\right] \\
&= \exp\left(iV(f) - \frac{1}{2}\langle f, (\Gamma_\tau + \Gamma) f\rangle\right),
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}\left[\exp(iX(g))\exp(iZ(f))\right] \\
&= \mathbb{E}\left[\exp(iZ(f))\,\mathbb{E}\left[\exp(iX(g)) \Big| Z\right]\right] \\
&= \mathbb{E}\left[\exp\left(iZ(f+g) - \frac{1}{2}\langle g, \Gamma g\rangle\right)\right] \\
&= \exp\left(-\frac{1}{2}\langle f+g, \Gamma_\tau(f+g)\rangle - \frac{1}{2}\langle g, \Gamma g\rangle + iV(f+g)\right) \\
&= \exp\left(iV((\Gamma + \Gamma_\tau)^{-1}\Gamma_\tau f) + iV(g + (\Gamma + \Gamma_\tau)^{-1}\Gamma f) - \frac{1}{2}\langle \Gamma_\tau f, (\Gamma + \Gamma_\tau)^{-1}\Gamma f\rangle \right.\\
&\qquad \left. -\frac{1}{2}\langle g + (\Gamma + \Gamma_\tau)^{-1}\Gamma_\tau f, (\Gamma + \Gamma_\tau)(g + (\Gamma + \Gamma_\tau)^{-1}\Gamma_\tau f)\rangle\right) \\
&= \mathbb{E}\left[\exp(iX(g)) \right.\\
&\qquad \left.\exp\left(iX((\Gamma + \Gamma_\tau)^{-1}\Gamma_\tau f) + iV((\Gamma + \Gamma_\tau)^{-1}\Gamma f) - \frac{1}{2}\langle \Gamma_\tau f, (\Gamma_\tau + \Gamma)^{-1}\Gamma f\rangle\right)\right],
\end{aligned}$$

which shows that

$$\begin{aligned}
&\mathbb{E}\left[\exp(iZ(f)) \Big| X\right] \\
&\quad = \exp\left(iV((\Gamma + \Gamma_\tau)^{-1}\Gamma f) + iX((\Gamma + \Gamma_\tau)^{-1}\Gamma_\tau f) - \frac{1}{2}\langle \Gamma_\tau f, (\Gamma + \Gamma_\tau)^{-1}\Gamma f\rangle\right).
\end{aligned}$$

□

In particular we get the following corollary which is classical in the framework of Gaussian filtering.

Proposition 8.2. *Let* $(Z_t)_{t\in[0,T]}$ *be a Brownian motion with quadratic variation* $\tau_t^2 dt$, $\tau \in L^2([0,T],dt)$, *and drift* $(v_t)_{t\in[0,T]}$, $v \in H$, *and let* $(X_t)_{t\in[0,T]}$ *have drift* $(Z_t)_{t\in[0,T]}$ *and quadratic variation* $(\sigma_t^2)_{t\in[0,T]}$, *given* Z. *Then, conditionally to* X, *the process* $(Z_t)_{t\in[0,T]}$ *has drift*

$$\int_0^t \frac{\sigma_s^2}{\tau_s^2+\sigma_s^2} dv_s + \int_0^t \frac{\tau_s^2}{\tau_s^2+\sigma_s^2} dX_s \quad \text{and variance} \quad \int_0^t \frac{\tau_s^2\sigma_s^2}{\tau_s^2+\sigma_s^2} ds, \qquad t\in[0,T].$$

References

[1] E. Alòs, O. Mazet, and D. Nualart. Stochastic calculus with respect to Gaussian processes. *Ann. Probab.*, 29(2):766–801, 2001.

[2] J. Berger and R. Wolpert. Estimating the mean function of a Gaussian process and the Stein effect. *J. Multivariate Anal.*, 13(3):401–424, 1983.

[3] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, 1992.

[4] E.B. Dynkin. *Markov processes. Vols. I, II*. Die Grundlehren der Mathematischen Wissenschaften. Academic Press Inc., New York, 1965.

[5] D. Fourdrinier, W.E. Strawderman, and M.T. Wells. On the construction of Bayes minimax estimators. *Ann. Stat.*, 26(2):660–671, 1998.

[6] V. Goodman. Harmonic functions on Hilbert space. *J. Funct. Anal.*, 10:451–470, 1972.

[7] L. Gross. Potential theory on Hilbert space. *J. Funct. Anal.*, 1:123–181, 1967.

[8] I.A. Ibragimov and Y.A. Rozanov. *Gaussian random processes*, volume 9 of *Applications of Mathematics*. Springer-Verlag, New York, 1978.

[9] W. James and C. Stein. Estimation with quadratic loss. In *Proc. 4th Berkeley Sympos. Math. Statist. and Prob., Vol. I*, pages 361–379. Univ. California Press, Berkeley, Calif., 1961.

[10] R.S. Liptser and A.N. Shiryaev. *Statistics of random processes. II*, volume 6 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 2001.

[11] D. Nualart. *The Malliavin calculus and related topics*. Probability and its Applications (New York). Springer-Verlag, Berlin, second edition, 2006.

[12] N. Privault and A. Réveillac. Stein estimation for the drift of Gaussian processes using the Malliavin calculus. Preprint, 2006, to appear in the Annals of Statistics.

[13] N. Privault and A. Réveillac. Superefficient drift estimation on the Wiener space. *C. R. Acad. Sci. Paris Sér. I Math.*, 343:607–612, 2006.

[14] B.L.S. Prakasa Rao. *Statistical inference for diffusion type processes*, volume 8 of *Kendall's Library of Statistics*. Edward Arnold, London, 1999.

[15] C. Stein. Estimation of the mean of a multivariate normal distribution. *Ann. Stat.*, 9(6):1135–1151, 1981.

[16] C.A. Tudor and F.G. Viens. Statistical aspects of the fractional stochastic calculus. *Ann. Stat.*, 35(3):1183–1212, 2007.

[17] R. Wolpert and J. Berger. Incorporating prior information in minimax estimation of the mean of a Gaussian process. In *Statistical decision theory and related topics, III, Vol. 2 (West Lafayette, Ind., 1981)*, pages 451–464. Academic Press, New York, 1982.