SURE SHRINKAGE OF GAUSSIAN PATHS AND SIGNAL IDENTIFICATION [*]

Nicolas Privault[1] and Anthony Réveillac[2]

Abstract. Using integration by parts on Gaussian space we construct a Stein Unbiased Risk Estimator (SURE) for the drift of Gaussian processes, based on their local and occupation times. By almost-sure minimization of the SURE risk of shrinkage estimators we derive an estimation and denoising procedure for an input signal perturbed by a continuous-time Gaussian noise.

Mathematics Subject Classification. 93E10, 93E14, 60G35, 60H07.

1. Introduction

Let X be a Gaussian random vector on $\mathbb{R}^d$ with unknown mean m and known covariance matrix $\sigma^2 \mathbf{I}_d$ under a probability measure $\mathbb{P}_m$. It is well-known [19] that given $g : \mathbb{R}^d \to \mathbb{R}^d$ a sufficiently smooth function, the mean square risk $\|X + g(X) - m\|_{\mathbb{R}^d}^2$ of $X + g(X)$ to m can be estimated unbiasedly by

$$\mathrm{SURE} := \sigma^2 d + \sum_{i=1}^{d} g_i(X)^2 + 2 \sum_{i=1}^{d} \nabla_i g(X), \tag{1.1}$$

from the identity

$$\mathbb{E}_m \left[\|X + g(X) - m\|_{\mathbb{R}^d}^2 \right] = \sigma^2 d + \mathbb{E}_m \left[\sum_{i=1}^{d} g_i(X)^2 + 2 \sum_{i=1}^{d} \nabla_i g(X) \right] \tag{1.2}$$

which is obtained by Gaussian integration by parts under $\mathbb{P}_m$. The estimator (1.1), which is independent of m, is called the Stein Unbiased Risk Estimate (SURE).

When $(g^\lambda)_{\lambda \in \Lambda}$ is a family of functions it makes sense to almost surely minimize the *Stein Unbiased Risk Estimate* (1.1) of g^λ with respect to the parameter λ. This point of view has been developed by Donoho and

Johnstone [6] for the design of spatially adaptive estimators by shrinkage of wavelet coefficients of noisy data *via*

$$X + g^\lambda(X) = \lambda\eta(X/\lambda),$$

where $\eta(x)$ is a threshold function.

In this paper we construct a Stein type Unbiased Risk Estimator for the random adapted drift $(u_t)_{t\in\mathbb{R}_+}$ of a one dimensional Gaussian processes $(X_t)_{t\in[0,T]}$ *via* an extension of the identity (1.2) introduced in [14,15] on the Wiener space, *cf.* also [16] on the Poisson space. For example, given $\alpha(t)$ and $\lambda(t)$ two functions given in parametric form, we compute the SURE risk of the estimator

$$X_t + \xi_t^{\alpha,\lambda}(X_t) = \alpha(t) + \lambda(t)\eta\left(\frac{X_t - \alpha(t)}{\lambda(t)}\right), \qquad t \in [0,T],$$

where η is a threshold function as in (3.1) and (4.1) below. We apply this technique to de-noising and identification of the input signal in a Gaussian channel *via* the minimization of SURE $(X + \xi^{\alpha,\lambda}(X))$ in the parameter (α, λ). This yields in particular an estimator of the drift of X_t from the estimation of $\alpha(t)$, and an optimal noise removal threshold from the estimation of λ. This approach differs from classical signal detection techniques which usually rely on likelihood ratio tests, *cf. e.g.* [13], Chapter VI. It also requires an a priori hypothesis on the parametric form of $\alpha(t)$.

The plan of the paper is as follows. In Section 2 we recall our framework of functional estimation of drift trajectories and we derive Stein's unbiased risk estimate for the estimation of the drift of Gaussian processes. In Section 3 we discuss its application to soft thresholding for Gaussian processes using the local time and obtain an upper bound for the risk of such estimators, with some examples. We also show the existence of an optimal parameter and the smoothness of the risk function. In Section 4 we consider the case of hard thresholding and show that the SURE risk can be computed due to the smoothing effect of the integral over time. In Section 5 we consider several numerical examples where $\alpha(t)$ is given in parametric form. In Section 6 we recall some elements of stochastic analysis of Gaussian processes.

We close this section with some notation on the Gaussian model used in this paper. Given $T > 0$, let $X = (X_t)_{t\in[0,T]}$ be a centered real-valued Gaussian process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\Omega = \mathcal{C}([0,T])$ is a space of continuous functions on $\mathbb{R}_+$, and $(\mathcal{F}_t)_{t\in[0,T]}$ is the filtration generated by $(X_t)_{t\in[0,T]}$. In the sequel we observe the paths of $(X_t)_{t\in[0,T]}$ decomposed as

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

where $u = (u_t)_{t\in[0,T]}$ is a square integrable $\mathcal{F}$-adapted process and $(X_t^u)_{t\in[0,T]}$ is a centered Gaussian process under $\mathbb{P}_u$. The covariance function

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

of $(X_t^u)_{t\in[0,T]}$, where $\mathbb{E}_u$ denotes the expectation under $\mathbb{P}_u$, is assumed to satisfy the integrability condition

$$\int_0^T \int_0^T \left(1 - \frac{\gamma(s,t)}{\sqrt{\gamma(s,s)\gamma(t,t)}}\right)^{-1/2} \mathrm{d}s\mathrm{d}t < \infty, \tag{1.3}$$

with $\gamma(t,t) > 0$, $t \in [0,T]$, from Section 3 onwards.

Let Γ be the covariance operator defined as

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

on the Hilbert space H of functions on $[0,T]$ with the inner product

$$\langle h,g \rangle_H = \langle h, \Gamma g \rangle_{L^2([0,T],d\mu)}, \tag{1.4}$$

where μ is a finite measure on $[0,T]$. The process $(X_t^u)_{t\in[0,T]}$ can be used to construct an isometry $X^u : H \to L^2(\Omega, \mathcal{F}, P)$ as

$$X^u(h) = \int_0^T h(s)X_s^u\mu(\mathrm{d}s), \qquad h \in H.$$

Then $\{X^u(h) \ : \ h \in H\}$ is an isonormal Gaussian process on H, *i.e.* a family of centered Gaussian random variables satisfying

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

For any orthonormal basis $(h_k)_{k\in\mathbb{N}}$ of $L^2([0,T],\mathrm{d}\mu)$, we have the Karhunen-Loève expansion

$$X_t^u = \sum_{k=0}^\infty h_k(t)X^u(h_k), \qquad t \in [0,T], \tag{1.5}$$

where $(X^u(h_k))_{k\in\mathbb{N}}$ is a sequence of independent standard Gaussian random variables.

2. Stein's unbiased risk estimate

We start by recalling the setting of functional drift estimation to be used in this paper. Given a continuous-time observation of the process $(X_t)_{t\in[0,T]}$ we consider an estimator $\xi = (\xi_t)_{t\in[0,T]}$ of the unknown adapted drift process $u = (u_t)_{t\in[0,T]}$. Such an estimator $(\xi_t)_{t\in[0,T]}$ will be 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 processes $(u_t)_{t\in[0,T]}$. In the next definition the risk measure μ is the one used to determine the scalar product (1.4).

Definition 2.1. The quadratic risk of an estimator $\xi := (\xi_t)_{t\in[0,T]}$ to $u = (u_t)_{t\in[0,T]}$ is defined as

$$R(\gamma,\mu,\xi) := \mathbb{E}_u\left[\int_0^T |\xi_t - u_t|^2\mu(\mathrm{d}t)\right]. \tag{2.1}$$

When decomposing $(\xi_t)_{t\in[0,T]}$ in the basis $(h_k)_{k\in\mathbb{N}}$ as

$$\xi_t = \sum_{k=0}^\infty h_k(t)\langle \xi, h_k \rangle_{L^2([0,T],\mu)}, \qquad t \in [0,T], \tag{2.2}$$

the risk (2.1) reads

$$R(\gamma,\mu,\xi) := \mathbb{E}_u\left[\sum_{k=0}^\infty |\langle \xi, h_k \rangle_{L^2([0,T],\mu)} - X^u(h_k)|^2\right].$$

The canonical process $(X_t)_{t\in[0,T]}$ can be viewed as an unbiased estimator $\hat{u} := (X_t)_{t\in[0,T]}$ of its random adapted drift $(u_t)_{t\in[0,T]}$ under $\mathbb{P}_u$, which is minimax in the sense that

$$\mathrm{R}(\gamma,\mu,\hat{u}) = \inf_{\xi} \sup_{v\in\Omega} \mathbb{E}_v\left[\int_0^T |\xi_t - v_t|^2 \mu(\mathrm{d}t)\right],$$

cf. Proposition 3.2 of [15], without restricting ξ as in (2.4). In this case we have

$$\begin{aligned}
\mathrm{R}(\gamma,\mu,\hat{u}) &= \mathbb{E}_u\left[\int_0^T |X_t - u_t|^2 \mu(\mathrm{d}t)\right] \\
&= \mathbb{E}_u\left[\sum_{k=0}^\infty |X^u(h_k)|^2\right] \\
&= \sum_{k=0}^\infty \langle h_k, \Gamma h_k\rangle^2_{L^2([0,T],\mathrm{d}\mu)} \\
&= \int_0^T \gamma(t,t)\mu(\mathrm{d}t),
\end{aligned}$$

and when in addition $(X_t)_{t\in[0,T]}$ has independent increments and $(u_t)_{t\in[0,T]}$ is $\mathbb{P}_u \otimes \mu$-square-integrable and adapted, $\hat{u} = (X_t)_{t\in[0,T]}$ becomes an efficient estimator of its own drift $u = (u_t)_{t\in[0,T]}$ as the Cramer-Rao bound

$$\mathbb{E}_u\left[\int_0^T |\xi_t - u_t|^2 \mu(\mathrm{d}t)\right] \geq \mathrm{R}(\gamma,\mu,\hat{u}) \tag{2.3}$$

holds for any unbiased and adapted estimator $\xi = (\xi_t)_{t\in[0,T]}$ of $u = (u_t)_{t\in[0,T]}$ and is attained by $\hat{u} = (X_t)_{t\in[0,T]}$, *cf.* Proposition 4.3 of [15].

Note that the standard framework of functional drift estimation, such as in Pinsker's theory, is usually concerned with the asymptotics of the minimax risk

$$\inf_{\xi} \sup_{v\in\Omega_r} \mathbb{E}_v\left[\int_0^T |\xi_t - v_t|^2 \mu(\mathrm{d}t)\right],$$

as $r \to \infty$, for a deterministic drift, where Ω_r is an ellipsoid constructed by constraining the coefficients of $(v_t)_{t\in[0,T]}$ to satisfy a condition of the form

$$\sum_{k=0}^\infty a_k |\langle v, h_k\rangle_{L^2([0,T],\mu)}|^2 \leq r. \tag{2.4}$$

This approach yields limiting results in small noise and large time or dimension, in relation to the smoothness of the drift function, *cf. e.g.* [2,8,10,12].

In this paper, instead of applying the procedures described above, we will estimate the drift of $(X_t)_{t\in[0,T]}$ by the almost sure minimization of a Stein Unbiased Risk Estimator for Gaussian processes, constructed in the next proposition by analogy with (1.1). For this we will use the gradient operator D_t whose definition and properties are recalled in the appendix, *cf.* Definition 6.2 and Lemma 6.3.

Proposition 2.2. *For any* $(\xi_t)_{t\in[0,T]} \in L^2(\Omega \times [0,T], \mathbb{P}_u \otimes \mu)$ *such that* $\xi_t \in \mathrm{Dom}\,(\nabla)$, $t \in [0,T]$, *and* $(D_t\xi_t)_{t\in[0,T]} \in L^1(\Omega \times [0,T], \mathbb{P}_u \otimes \mu)$, *the quantity*

$$\mathrm{SURE}\,_\mu(X + \xi) := \mathrm{R}(\gamma, \mu, \hat{u}) + \|\xi\|^2_{L^2([0,T],d\mu)} + 2\int_0^T D_t\xi_t\mu(\mathrm{d}t) \tag{2.5}$$

is an unbiased estimator of the mean square risk $\|X + \xi - u\|^2_{L^2([0,T],d\mu)}$.

Proof. From Lemma 6.3 we have

$$\mathbb{E}_u\left[\|X + \xi - u\|^2_{L^2([0,T],d\mu)}\right] = \mathbb{E}_u\left[\int_0^T \left|X_t^u + \xi_t\right|^2\mu(\mathrm{d}t)\right]$$

$$= \mathbb{E}_u\left[\int_0^T |X_t^u|^2\mu(\mathrm{d}t)\right] + \mathbb{E}_u\left[\|\xi\|^2_{L^2([0,T],d\mu)}\right] + 2\,\mathbb{E}_u\left[\int_0^T X_t^u\xi_t\mu(\mathrm{d}t)\right]$$

$$= \mathrm{R}(\gamma, \mu, \hat{u}) + \mathbb{E}_u\left[\|\xi\|^2_{L^2([0,T],d\mu)}\right] + 2\,\mathbb{E}_u\left[\int_0^T D_t\xi_t\mu(\mathrm{d}t)\right]$$

$$= \mathbb{E}_u\,[\mathrm{SURE}\,_\mu(X + \xi)]. \qquad \square$$

Unlike the pointwise mean square risk $\|X + \xi - u\|^2_{L^2([0,T],d\mu)}$, the SURE risk estimator does not depend on the estimated parameter $u = (u_t)_{t\in[0,T]}$.

   Given a family $(\xi^\lambda)_{\lambda\in\Lambda}$ of estimators indexed by a parameter space Λ, we consider the estimator $X + \xi^{\lambda^*}$ that almost-surely minimizes the SURE risk, where λ^* is given by

$$\lambda^* = \mathrm{argmin}_{\lambda\in\Lambda}\mathrm{SURE}\,_\mu(X + \xi^\lambda).$$

For all values of $\lambda \in \Lambda$, the SURE risk of the estimator $X + \xi^{\lambda^*}$ improves on the mean square risk of $X + \xi^\lambda$, since for all $\nu \in \Lambda$ we have

$$
\begin{aligned}
\mathbb{E}_u[\mathrm{SURE}\,_\mu(X + \xi^{\lambda^*})] &\leq \mathbb{E}_u[\mathrm{SURE}\,_\mu(X + \xi^\nu)] \\
&= \mathbb{E}_u\left[\|\xi^\nu - u\|^2_{L^2([0,T],d\mu)}\right] \\
&\leq \inf_\lambda \mathbb{E}_u\left[\|\xi^\lambda - u\|^2_{L^2([0,T],d\mu)}\right].
\end{aligned}
$$

In the sequel we will apply the above to a process $(\xi_t)_{t\in[0,T]}$ given as a funtion

$$\xi_t = \xi_t(X_t) \quad \text{of} \quad X_t, \qquad t \in [0,T].$$

In particular we will discuss estimation and thresholding for estimators of the form

$$X_t + \xi_t^{\alpha,\lambda}(X_t) = \alpha(t) + \lambda(t)\eta\left(\frac{X_t - \alpha(t)}{\lambda(t)}\right), \tag{2.6}$$

where $\eta : \mathbb{R} \to \mathbb{R}$ is a threshold function with support in $(-\infty, -1]\cup[1,\infty)$ and $\lambda(t) \geq 0$ is a given level function.

   In particular we will apply our method to the joint estimation of parameters α, λ, successively in case $\alpha(t) = \alpha$, $\alpha(t) = \alpha t$, and $\lambda(t) = \lambda\sqrt{\gamma(t,t)}$.

3. Soft threshold

In this section we construct an example of SURE shrinkage by soft thresholding in the framework of Proposition 2.2, with application to identification and de-noising in the Gaussian signal $(X_t)_{t \in [0,T]}$. In case η is the soft threshold function

$$\eta_S(y) = \text{sign}(y)(|y| - 1)^+, \quad y \in \mathbb{R}, \tag{3.1}$$

the function $\xi_t^{\alpha,\lambda}$ in (2.6) becomes

$$\xi_t^{\alpha,\lambda}(x) = -\text{sign}(x - \alpha(t)) \min(\lambda(t), |x - \alpha(t)|), \qquad x \in \mathbb{R}.$$

In this setting the SURE risk can be computed as follows.

Proposition 3.1. *We have* $\mathbb{P}$*-a.s*

$$\text{SURE}_\mu(X + \xi^{\alpha,\lambda}(X)) = \text{R}(\gamma, \mu, \hat{u}) + \int_0^T |X_t - \alpha(t)|^2 \wedge \lambda^2(t)\mu(dt) - 2\int_0^T \mathbf{1}_{\{|X_t - \alpha(t)| \leq \lambda(t)\}}\gamma(t,t)\mu(dt). \tag{3.2}$$

Proof. Since $\frac{\partial}{\partial x}\xi_t^{\alpha,\lambda}(x) = -\mathbf{1}_{\{|x-\alpha(t)| \leq \lambda(t)\}}$, we have

$$
\begin{aligned}
\int_0^T D_t\xi_t^{\alpha,\lambda}(X_t)\mu(dt) &= -\int_0^T \mathbf{1}_{\{|X_t-\alpha(t)| \leq \lambda(t)\}}D_tX_t\mu(dt) \\
&= -\int_0^T \mathbf{1}_{\{|X_t-\alpha(t)| \leq \lambda(t)\}}\gamma(t,t)\mu(dt),
\end{aligned}
$$

hence the conclusion from Proposition 2.2. $\qquad\square$

In the simulations of Section 5 we will use the risk measure

$$\mu(dt) = \sum_{i=1}^n a_i\delta_{t_i}(dt), \qquad a_1, \ldots, a_n > 0, \tag{3.3}$$

associated to discrete values $(X_{t_1}, \ldots, X_{t_n})$ of the sample path observed at times $t_1, \ldots, t_n$, $n \geq 1$, in which case Relation (3.2) becomes

$$\text{SURE}(X + \xi^{\alpha,\lambda}(X)) = \text{R}(\gamma, \mu, \hat{u}) + \sum_{i=1}^n |X_{t_i} - \alpha(t_i)|^2 \wedge \lambda^2(t_i) - 2\sum_{i=1}^n \gamma(t_i, t_i)\mathbf{1}_{\{|X_{t_i}-\alpha(t_i)| \leq \lambda(t_i)\}},$$

which is analog to the finite dimensional SURE risk

$$\text{SURE}(X + g^\lambda(X)) = d + \sum_{i=1}^d (|x_i| \wedge \lambda)^2 - 2\#\{i; |x_i| \leq \lambda\} \tag{3.4}$$

of [5]. In addition, when $\mu(dt) = f(t)dt$ has a density $f(t)$ with respect to the Lebesgue measure and

$$\mu_n(dt) = \sum_{i=1}^{n-1} f(t_i)(t_{i+1} - t_i)\delta_{t_i}(dt),$$

Relation (3.2) shows that $\text{SURE}_{\mu_n}(X + \xi^{\alpha,\lambda}(X))$ becomes a consistent estimator of the risk $\text{SURE}_\mu(X + \xi^{\alpha,\lambda}(X))$ as n goes to infinity.

In the sequel we take

$$\mu(\mathrm{d}t) = \gamma^{-1}(t,t)\mathrm{d}t \quad \text{and} \quad \lambda(t) = \lambda\sqrt{\gamma(t,t)}, \qquad \lambda > 0, \quad t \in [0,T], \tag{3.5}$$

and letting

$$\bar{L}_T^\lambda := \int_0^T \mathbf{1}_{\{|X_t - \alpha(t)| \le \lambda\sqrt{\gamma(t,t)}\}}\,\mathrm{d}t \tag{3.6}$$

denote the occupation time of the process

$$Z_t^{\alpha,\gamma} := \frac{X_t - \alpha(t)}{\sqrt{\gamma(t,t)}}, \qquad t \in [0,T], \tag{3.7}$$

up to time T in the set $[-\lambda, \lambda]$, Proposition 3.1 yields the identity

$$\mathrm{SURE}_\mu(X + \xi^{\alpha,\lambda}(X)) = T + \int_0^T \left(|Z_t^{\alpha,\gamma}| \wedge \lambda\right)^2 \mathrm{d}t - 2\bar{L}_T^\lambda. \tag{3.8}$$

As a consequence we obtain the following bound for the risk of the thresholding estimator $X + \xi^{\alpha,\lambda}(X)$, which shows that $\mathrm{SURE}_\mu(X + \xi^{\alpha,\lambda}(X))$ is independent of large values $\|u - \alpha\|_{L^2([0,T])}$, and grows at most as $1 + \lambda^2$ in $\lambda \ge 0$.

Proposition 3.2. *Assume that $u \in L^2([0,T], \mathrm{d}\mu)$ is a deterministic function and that (3.5) holds. Then for all fixed $\lambda \ge 0$ we have*

$$\mathbb{E}_u[\|X + \xi^{\alpha,\lambda}(X) - u\|_{L^2([0,T],\mathrm{d}\mu)}^2] \le (1 + \lambda^2)\left(T \wedge \int_0^T |u(t) - \alpha(t)|^2\mu(\mathrm{d}t)\right) + T(1 + \lambda)\mathrm{e}^{-\frac{\lambda^2}{2}}.$$

Proof. From Proposition 3.1 we have

$$\mathrm{SURE}_\mu(X + \xi^{\alpha,\lambda}(X)) = T + \int_0^T \left(|Z_t^{\alpha,\gamma}| \wedge \lambda\right)^2 \mathrm{d}t - 2\int_0^T \mathbf{1}_{\{|X_t - \alpha(t)| \le \lambda\sqrt{\gamma(t,t)}\}}\,\mathrm{d}t, \tag{3.9}$$

hence

$$\mathbb{E}_u[\mathrm{SURE}_\mu(X + \xi^{\alpha,\lambda}(X))] \le T(1 + \lambda^2),$$

and

$$
\begin{aligned}
\mathbb{E}_u[\mathrm{SURE}_\mu(X + \xi^{\alpha,\lambda}(X))] &\le \int_0^T (1 + \mathbb{E}_u[|Z_t^{\alpha,\gamma}|^2] \wedge \lambda^2 - 2\mathbb{P}_u(|Z_t^{\alpha,\gamma}| \le \lambda))\mathrm{d}t \\
&\le \int_0^T (1 + \lambda^2)\left(\mathrm{e}^{-\frac{\lambda^2}{2}} + \frac{|u(t) - \alpha(t)|^2}{\gamma(t,t)}\right)\mathrm{d}t \\
&\le (1 + \lambda^2)\int_0^T |u(t) - \alpha(t)|^2\mu(\mathrm{d}t) + T(1 + \lambda^2)\mathrm{e}^{-\frac{\lambda^2}{2}},
\end{aligned}
$$

where we recall that from [5], Appendix 1, we have

$$1 + \mathbb{E}_u[|Z_t^{\alpha,\gamma}|^2] \wedge \lambda^2 - 2\mathbb{P}_u(|Z_t^{\alpha,\gamma}| \le \lambda) \le (1 + \lambda^2)\left(\mathrm{e}^{-\frac{\lambda^2}{2}} + \frac{|u(t) - \alpha(t)|^2}{\gamma(t,t)}\right)$$

for every t in $[0,T]$, and we conclude from Proposition 2.2. $\square$

Since $\lambda \mapsto \mathrm{SURE}_\mu(X + \xi^{\alpha,\lambda}(X))$ in (3.8) is lower bounded by $-T$ and equal to 0 when $\lambda = 0$, the optimal threshold

$$\lambda^* := \mathrm{argmin}_\lambda \mathrm{SURE}_\mu(X + \xi^{\alpha,\lambda}(X)) \tag{3.10}$$

exists almost surely in $[0, \infty)$.

In addition we have the following proposition which is important for the numerical search of the optimal parameter value λ^*.

Proposition 3.3. *Under* (3.5) *and Condition* (1.3)*, the function* $\lambda \mapsto \mathrm{SURE}_\mu(X + \xi^{\alpha,\lambda}(X))$ *is continuously differentiable.*

Proof. Letting

$$
\begin{aligned}
\Delta(s,t) &= \mathrm{Var}_u[Z_t^{\alpha,\gamma} - Z_s^{\alpha,\gamma}] \\
&= \mathrm{Var}_u\left[\frac{X_t^u}{\sqrt{\gamma(t,t)}} - \frac{X_s^u}{\sqrt{\gamma(s,s)}} \right] \\
&= 2 - 2\frac{\gamma(s,t)}{\sqrt{\gamma(s,s)\gamma(t,t)}}, \qquad 0 \le s, t \le T,
\end{aligned}
$$

by (3.7), the local time

$$\bar{\ell}_T^\lambda := \frac{d}{d\lambda}\bar{L}_T^\lambda$$

of $(|Z_t^{\alpha,\gamma}|)_{t\in[0,T]}$ exists almost surely under Condition (1.3), *cf.* [3,7] and Section 6.2, and by (3.9) we have

$$
\begin{aligned}
\frac{\partial}{\partial\lambda}\mathrm{SURE}_\mu(X + \xi^{\alpha,\lambda}(X)) &= \frac{\partial}{\partial\lambda}\int_0^T (|Z_t^{\alpha,\gamma}| \wedge \lambda)^2\, \mathrm{d}t - 2\bar{\ell}_T^\lambda \\
&= 2\lambda\int_0^T \mathbf{1}_{\{|X_t - \alpha(t)| \ge \lambda\sqrt{\gamma(t,t)}\}}\mathrm{d}t - 2\bar{\ell}_T^\lambda \\
&= 2\lambda(T - \bar{L}_T^\lambda) - 2\bar{\ell}_T^\lambda,
\end{aligned}
$$

which is a continuous function of λ since the covariance $\gamma(s,t)$ does not vanish, *cf. e.g.* Theorem 26.1 of [7]. $\quad\square$

Consequently we have

$$\frac{\partial}{\partial\lambda}\mathrm{SURE}_\mu(X + \xi^{\alpha,\lambda}(X))_{|\lambda=0} = -2\bar{\ell}_T^0,$$

hence $\lambda^* > 0$ a.s. when ℓ_T^0 is a.s. positive, which is the case for example when $(X_t)_{t\in[0,T]}$ is a Brownian motion, see Corollary 2.2, page 240 of [18], Chapter VI.

In practice we will compute λ^* numerically by minimization of $\lambda \mapsto \mathrm{SURE}_\mu(X + \xi^{\alpha,\lambda}(X))$ over λ in a range $\Lambda = [0, C(T)]$ where $C(T)$ is such that

$$\lim_{T\to\infty}\mathbb{P}_u\left(\sup_{t\in[0,T]} |Z_t^{\alpha,\gamma}| \le C(T) \right) = 1, \tag{3.11}$$

where $(Z_t^{\alpha,\gamma})_{t\in[0,T]}$ is defined in (3.7). This condition is analog to Condition (31) in [5] and allows us to restrict the range of λ when searching for an optimal threshold using Proposition 6.4 in the appendix.

The function $\alpha(t)$ can be given in parametric form, in which case the parameters will be used to minimize $(\alpha, \lambda) \mapsto \mathrm{SURE}_\mu(X + \xi^{\alpha,\lambda}(X))$, *cf.* Section 5.

Examples of Gaussian processes X^u satisfying Condition (1.3) include Ornstein Uhlenbeck processes with covariance function $\gamma(s,t) = \dfrac{\sigma^2}{2a}\mathrm{e}^{-a|t-s|}$, $s, t \in [0, T]$, used in the simulations of Section 5, *cf.* Proposition 6.4

below, and fractional Brownian processes with covariance function

$$\gamma(s,t) = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H}), \qquad s, t > 0.$$

Note that not all noise processes satisfy the above condition, for example the sine-cosine process

$$X_t^u := \xi_1 \cos(\omega t) + \xi_2 \sin(\omega t), \quad t \in [0, T],$$

cf. [1], where $\omega \neq 0$ is fixed and ξ_1, ξ_2 are two independent standard Gaussian random variables, does not have a local time at all levels, and its covariance function

$$\gamma(s,t) = \cos(\omega(t-s))), \qquad s, t > 0,$$

does not satisfy (1.3).

4. HARD THRESHOLD

Here we use the threshold function

$$\eta_H(y) = y\mathbf{1}_{\{|y|>1\}}, \quad y \in \mathbb{R}, \tag{4.1}$$

hence $\xi_t^{\alpha,\lambda}$ in (2.6) becomes

$$\xi_t^{\alpha,\lambda}(x) = -(x - \alpha(t))\mathbf{1}_{\{|x-\alpha(t)|<\lambda\sqrt{\gamma(t,t)}\}}, \qquad x \in \mathbb{R},$$

where $\lambda \geq 0$ is a level parameter.

In finite dimensions [5] the SURE estimator (1.1) cannot be computed due to the non-differentiability of η_H, however a deterministic optimal threshold equal to $\sqrt{2\log d}$ can be obtained by other methods, *cf.* Theorem 4 of [5].

In continuous time the situation is different due to the smoothing effect of the integral over time, and the SURE risk can be computed as in the next proposition, using the local time of Gaussian processes when $\mu(\mathrm{d}t) = \gamma^{-1}(t,t)\mathrm{d}t$.

Proposition 4.1. *Under Condition* (1.3) *we have* $\mathbb{P}$-*a.s*

$$\mathrm{SURE}_\mu(X + \xi^{\alpha,\lambda}(X)) = T + \int_0^T \frac{(X_t - \alpha(t))^2}{\gamma(t,t)}\mathbf{1}_{\{|X_t-\alpha(t)|\leq\lambda\sqrt{\gamma(t,t)}\}}\mathrm{d}t + 2\lambda\bar{\ell}_T^\lambda - 2\bar{L}_T^\lambda. \tag{4.2}$$

Proof. Let $\phi \in \mathcal{C}_c^\infty([-1,1])$, $\phi \geq 0$ be symmetric around the origin, such that $\int_{-1}^1 \phi(x)\mathrm{d}x = 1$, and let

$$\phi_\varepsilon(x) = \varepsilon^{-1}\phi(\varepsilon^{-1}x), \qquad x \in \mathbb{R}, \quad \varepsilon > 0.$$

Let

$$\xi_t^{\alpha,\lambda,\varepsilon}(x) = \phi_{\varepsilon\sqrt{\gamma(t,t)}} * \xi_t^{\alpha,\lambda}(x) = \int_{-\infty}^\infty \phi_{\varepsilon\sqrt{\gamma(t,t)}}(y)\xi_t^{\alpha,\lambda}(x-y)\mathrm{d}y,$$

denote the convolution of $\phi_{\varepsilon\sqrt{\gamma(t,t)}}$ with $\xi_t^{\alpha,\lambda}$, with

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\phi_{\varepsilon\sqrt{\gamma(t,t)}} * \xi_t^{\alpha,\lambda}(x) &= \phi_{\varepsilon\sqrt{\gamma(t,t)}} * \frac{\mathrm{d}}{\mathrm{d}x}\xi_t^{\alpha,\lambda}(x) \\
&= \lambda\sqrt{\gamma(t,t)}\phi_{\varepsilon\sqrt{\gamma(t,t)}}(-\lambda\sqrt{\gamma(t,t)} + x - \alpha(t)) \\
&\quad + \lambda\sqrt{\gamma(t,t)}\phi_{\varepsilon\sqrt{\gamma(t,t)}}(\lambda\sqrt{\gamma(t,t)}) + x - \alpha(t) \\
&\quad - \int_{-\infty}^{\infty}\phi_{\varepsilon\sqrt{\gamma(t,t)}}(y)\mathbf{1}_{\{|x-y-\alpha(t)|<\lambda\sqrt{\gamma(t,t)}\}}\,\mathrm{d}y.
\end{aligned}
$$

From the occupation time density formula (6.3) below we have

$$
\begin{aligned}
\int_0^T D_t\xi_t^{\alpha,\lambda,\varepsilon}(X_t)\mu(\mathrm{d}t) &= \lambda\int_0^T\sqrt{\gamma(t,t)}\phi_{\varepsilon\sqrt{\gamma(t,t)}}(-\lambda\sqrt{\gamma(t,t)} + X_t - \alpha(t))\mathrm{d}t \\
&\quad + \lambda\int_0^T\sqrt{\gamma(t,t)}\phi_{\varepsilon\sqrt{\gamma(t,t)}}(\lambda\sqrt{\gamma(t,t)} + X_t - \alpha(t))\mathrm{d}t \\
&\quad - \int_0^T\int_{-\infty}^{\infty}\phi_{\varepsilon\sqrt{\gamma(t,t)}}(y)\mathbf{1}_{\{|x-y-\alpha(t)|<\lambda\sqrt{\gamma(t,t)}\}}\,\mathrm{d}y\mathrm{d}t \\
&= \lambda\int_{-\infty}^{\infty}(\phi_\varepsilon(-\lambda + Z_t^{\alpha,\gamma}) + \phi_\varepsilon(-\lambda - Z_t^{\alpha,\gamma}))\mathrm{d}t \\
&\quad - \int_0^T\int_{-\infty}^{\infty}\phi_{\varepsilon\sqrt{\gamma(t,t)}}(y)\mathbf{1}_{\{|x-y-\alpha(t)|<\lambda\sqrt{\gamma(t,t)}\}}\,\mathrm{d}y\mathrm{d}t \\
&= \lambda\int_{-\infty}^{\infty}\phi_\varepsilon(a - \lambda)\bar{\ell}_T^a\,\mathrm{d}a \\
&\quad - \int_0^T\int_{-\infty}^{\infty}\phi_{\varepsilon\sqrt{\gamma(t,t)}}(y)\mathbf{1}_{\{|x-y-\alpha(t)|<\lambda\sqrt{\gamma(t,t)}\}}\,\mathrm{d}y\mathrm{d}t,
\end{aligned}
$$

which converges in $L^2(\Omega, \mathbb{P}_u)$ to

$$
\lambda\bar{\ell}_T^\lambda - \int_0^T\mathbf{1}_{\{|X_t-\alpha(t)|<\lambda\sqrt{\gamma(t,t)}\}}\,\mathrm{d}t
$$

as ε tends to zero. $\square$

5. Numerical examples

In this section present some numerical estimations of the parameters α, λ in (2.6), successively in case $\alpha(t) = \alpha$, $\alpha(t) = \alpha t$, and $\lambda(t) = \lambda\sqrt{\gamma(t,t)}$. We take X^u to be a centered stationary Ornstein-Uhlenbeck process solution of

$$
\mathrm{d}X_t^u = -aX_t^u\mathrm{d}t + \sigma\mathrm{d}B_t, \qquad t \in [0,T], \tag{5.1}
$$

with $X_0^u \sim \mathcal{N}\left(0, \frac{\sigma^2}{2a}\right)$ and covariance function

$$
\gamma(s,t) = \frac{\sigma^2}{2a}\mathrm{e}^{-a|t-s|}, \qquad s,t \in [0,T],
$$

which satisfies (1.3) for $\sigma, a > 0$. From (3.11) above and Proposition 6.4 in the appendix, we take $\Lambda = [0, \sqrt{2\log T}]$ as parameter range for λ when T is large.
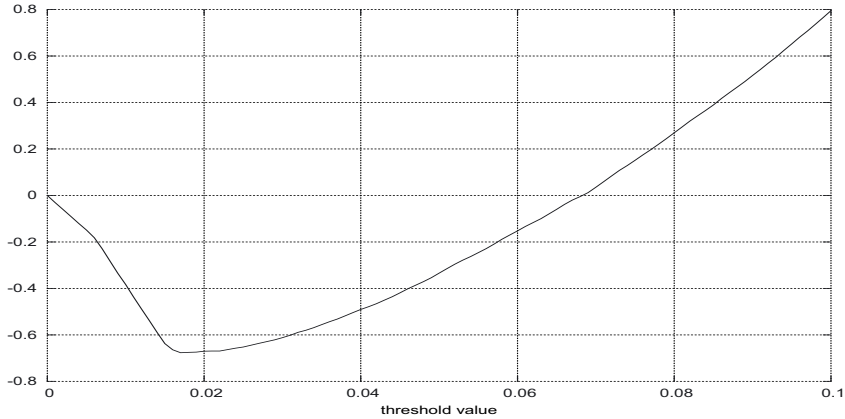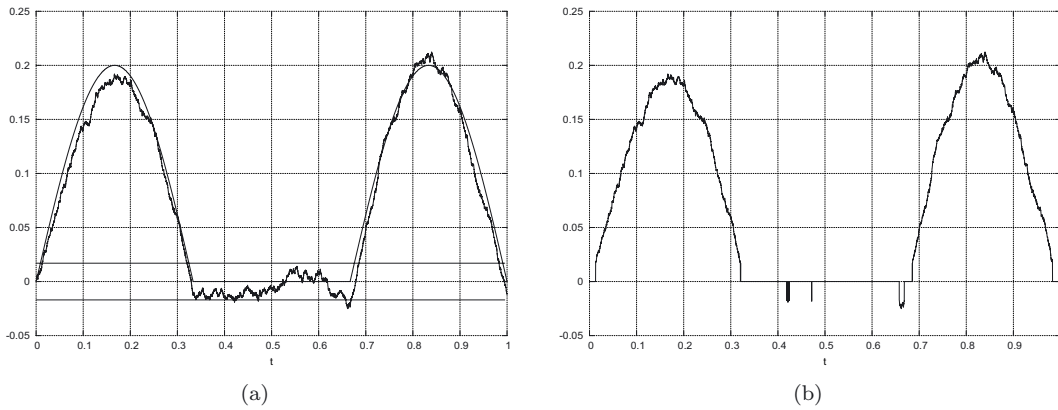
FIGURE 1.   Risk function $\lambda \mapsto \mathrm{SURE}_\mu(X + \xi^{0,\lambda}(X))$.

FIGURE 2.   (a) Process trajectory. (b) Estimated trajectory.

In the next figures we present some numerical simulations when the signal $(X_t)_{t \in [0,T]}$ is a deterministic function $(u(t))_{t \in [0,T]}$ perturbed by $(X_t^u)_{t \in [0,T]}$ a centered Ornstein-Uhlenbeck process, with parameters $a = 0.5$, $\sigma = 0.05$, $T = 1$.

In each of the following 3 examples we represent

– the risk function $(\alpha, \lambda) \mapsto \mathrm{SURE}_\mu(X + \xi^{\alpha,\lambda}(X))$ whose minimum gives the optimal parameter value, *cf.* Figures 1, 3 and 5.

– a simulated sample path with the optimal thresholds obtained by soft thresholding, and the de-noised signal after hard thresholding, *cf.* Figures 2, 4 and 6.

The hard threshold function has not been used for estimation due to increased numerical instabilities linked to the simulation of the local time in (4.2).

5.1. **Simple thresholding**

Here we take

$$u_t = 0.2 \times \max(0, \sin(3\pi t)),$$

with $\lambda(t) = \lambda\sqrt{\gamma}$, and we aim at de-noising the signal around the level $\alpha(t) = 0$, $t \in [0,T]$.
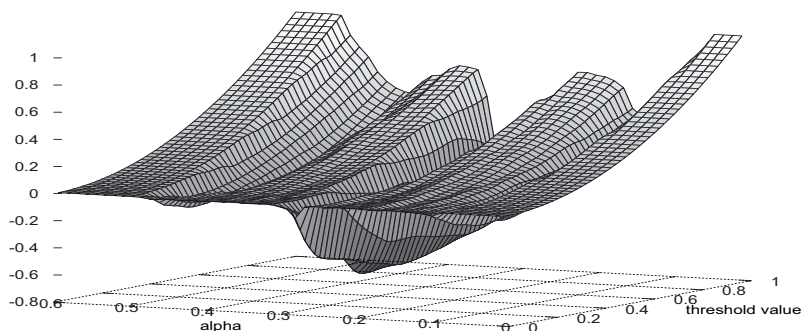
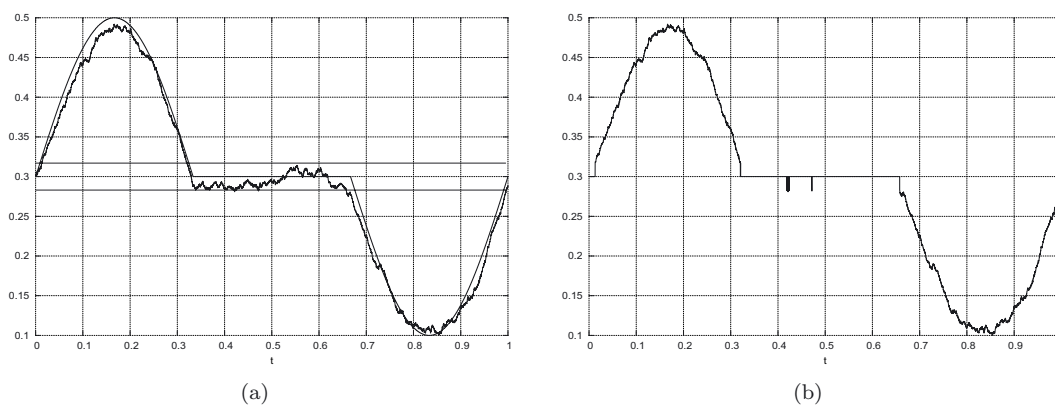FIGURE 3.   Risk function $(\alpha, \lambda) \mapsto \mathrm{SURE}_\mu(X + \xi^{\alpha,\lambda}(X))$.

FIGURE 4.   (a) Process trajectory. (b) Estimated trajectory.

From Figure 1 we estimate the optimal threshold to $\lambda^*\sqrt{\gamma} = 0.018$, after numerical minimization on a grid, which leads to the thresholding described in Figure 2.

5.2. Level detection and thresholding

We apply our method to the joint estimation of parameters α, λ, in case

$$u_t = 0.3 + 0.2 \times \mathrm{sign}(\sin(2\pi t)) \times \max(0, \sin(3\pi t)),$$

with $\alpha(t) = \alpha$ and $\lambda(t) = \lambda\sqrt{\gamma}$, *i.e.* we aim at detecting simultaneously the level $\alpha = 0.3$ and the threshold $\lambda\sqrt{\gamma}$ at which the noise can be removed. For this we have the following proposition that completes Proposition 3.3 for the search of an optimal parameter value.

Proposition 5.1. *Under Condition* (1.3)*, the function* $(\alpha, \lambda) \mapsto \mathrm{SURE}_\mu(X + \xi^{\alpha,\lambda}(X))$ *is continuously differentiable.*
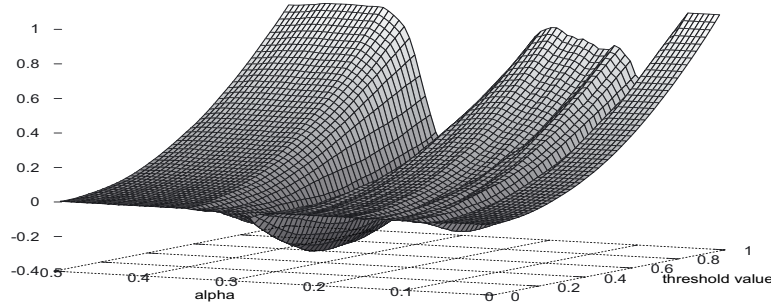
FIGURE 5.   Risk function $(\alpha, \lambda) \mapsto \mathrm{SURE}_\mu(X + \xi^{\alpha,\lambda}(X))$.
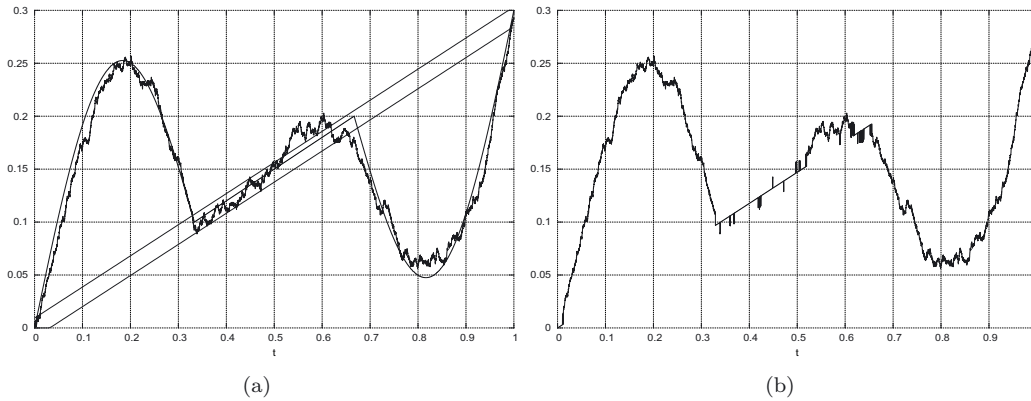

FIGURE 6.   (a) Process trajectory. (b) Estimated trajectory.

Proof. We have

$$\frac{\partial}{\partial \alpha} \mathrm{SURE}_\mu(X + \xi^{\alpha,\lambda}(X)) = -2 \int_0^T \frac{X_t - \alpha}{\gamma(t,t)} \mathbf{1}_{\{|X_t - \alpha| \leq \lambda \sqrt{\gamma(t,t)}\}} \mathrm{d}t + 2\ell_T^{\alpha,\lambda} - 2\ell_T^{\alpha,-\lambda},$$

where $\ell_T^{\alpha,\lambda}$ denotes the local time at level α of the process $(X_t + \lambda \sqrt{\gamma(t,t)})_{t \in [0,T]}$.   □

From Figure 3 we estimate the optimal threshold and shift parameters at $\lambda^* \sqrt{\gamma} = 0.017$ and $\alpha^* = 0.30$, which leads to the thresholding described in Figure 4 below.

Figure 3 also shows that the values 0.5 and 0.1 are other candidates to an estimation of α. These values correspond to the extrema in the sample trajectory of Figure 4.

5.3. **Drift detection and thresholding**

We apply our method to the joint estimation of parameters α, λ, in case

$$u_t = 0.3t + 0.2 \times \mathrm{sign}(\sin(2\pi t)) \times \max(0, \sin(3\pi t)),$$

with $\alpha(t) = \alpha t$, and $\lambda(t) = \lambda \sqrt{\gamma}$, *i.e.* we aim at locating noise with threshold $\lambda \sqrt{\gamma}$ around a line of slope $\alpha = 0.3$. Analogously to Propositions 3.3 and 5.1 we have the following result for the search of an optimal parameter value.

Proposition 5.2. *Under Condition* (1.3), *the function* $(\alpha, \lambda) \mapsto \mathrm{SURE}_\mu(X + \xi^{\alpha,\lambda}(X))$ *is continuously differentiable.*

Proof. We have

$$\frac{\partial}{\partial \alpha}\text{SURE}_\mu(X + \xi^{\alpha,\lambda}(X)) = -2\int_0^T \frac{X_t - \alpha t}{\gamma(t,t)}\mathbf{1}_{\{|X_t - \alpha t| \leq \lambda\sqrt{\gamma(t,t)}\}}t\mathrm{d}t + 2\ell_T^{\alpha,\lambda} - 2\ell_T^{\alpha,-\lambda},$$

where $\ell_T^{\alpha,\lambda}$ denotes the local time at level α of the process $((X_t + \lambda\sqrt{\gamma(t,t)})/t)_{t\in[0,T]}$. □

The optimal threshold and slope parameters are numerically estimated at $\lambda^*\sqrt{\gamma} = 0.0093$ and $\alpha^* = 0.294$.

In Figure 6 the threshold and slope appear to have been slightly underestimated, as the larger noise at the right end of the slope line has been interpreted as being part of the signal.

6. APPENDIX

In this section we review three aspects of stochastic analysis for Gaussian processes, including local time and the Malliavin calculus calculus.

6.1. **Malliavin calculus on Gaussian space**

Here we recall some elements of the Malliavin calculus on Gaussian space for the centered Gaussian process $(X_t^u)_{t\in[0,T]}$ under $\mathbb{P}_u$, see *e.g.* [9].

Let $\mathcal{S}$ denote the space of cylindrical functionals of the form

$$F = f_n(X^u(h_1), \dots, X^u(h_n)), \tag{6.1}$$

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

Definition 6.1. The H-valued Malliavin derivative is defined as

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

for $F \in \mathcal{S}$ of the form (6.1).

It is known that ∇ is closable, *cf.* Proposition 1.2.1 of [9], and its closed domain will be denoted by $\text{Dom}(\nabla)$.

Definition 6.2. Let D_t be defined on $F \in \text{Dom}(\nabla)$ as

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

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

$$\mathbb{E}_u[F\delta(v)] = \mathbb{E}_u[\langle v, \nabla F\rangle_H], \qquad F \in \text{Dom}(\nabla), \quad v \in \text{Dom}(\delta), \tag{6.2}$$

where $\mathbb{E}_u$ denotes the expectation under $\mathbb{P}_u$, with the relation

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

cf. [9], for $F \in \text{Dom}(\nabla)$ and $h \in H$ such that $hF \in \text{Dom}(\delta)$. The next lemma will be needed in Proposition 2.2 below to establish Stein's Unbiased Risk Estimate for Gaussian processes.

Lemma 6.3. *For any $F \in \text{Dom}(\nabla)$ and $u \in H$ we have*

$$\mathbb{E}_u[FX_t^u] = \mathbb{E}_u[D_t F], \qquad t \in [0,T].$$

Proof. We have

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

Note that since $u \in H$ we have $\nabla_s X_t(h) = \nabla_s X_t^u(h) = h(s)$ and

$$
\begin{aligned}
D_t X_t &= (\Gamma\nabla X_t)(t) \\
&= \int_0^T \gamma(s,t)\nabla_s X_t \mu(\mathrm{d}s) \\
&= \sum_{k=0}^{\infty} h_k(t)\int_0^T \gamma(s,t)\nabla_s X(h_k)\mu(\mathrm{d}s) \\
&= \sum_{k=0}^{\infty} h_k(t)\langle \gamma(\cdot,t), h_k\rangle_{L^2([0,T],\mathrm{d}\mu)} \\
&= \gamma(t,t), \quad t \in [0,T].
\end{aligned}
$$

6.2. **Local time and supremum of Gaussian processes**

Given $(Z_t)_{t\in[0,T]}$ a Gaussian process let

$$
\Delta(s,t) = \mathrm{Var}\,(Z_t - Z_s), \qquad 0 \le s,t \le T,
$$

and denote by

$$
L_T^\lambda := \int_0^T \mathbf{1}_{\{Z_t \le \lambda\}}\mathrm{d}t
$$

the occupation time of $(Z_t)_{t\in[0,T]}$ up to T in the set $(-\infty, \lambda]$.

Recall that a classical result of Berman [3], see Theorems 21.9 and 22.1 of [7], shows that under the condition

$$
\int_0^T \int_0^T (\Delta(s,t))^{-1/2}\mathrm{d}s\mathrm{d}t < \infty,
$$

the local time

$$
\ell_T^\lambda := \frac{\partial}{\partial\lambda}L_T^\lambda
$$

of $(Z_t)_{t\in[0,T]}$ exists at all levels $\lambda \in \mathbb{R}$ and the occupation time density formula

$$
\int_0^T f(Z_t)\mathrm{d}t = \int_{\mathbb{R}} f(\lambda)\ell_T^\lambda\mathrm{d}\lambda \tag{6.3}
$$

holds for every positive measurable function f on $\mathbb{R}$. The local time $\bar{\ell}_T^\lambda$ of $|Z_t|$ is given by $\bar{\ell}_T^\lambda = \ell_T^{-a} + \ell_T^a$ and the related occupation time formula can be obtained under the same condition from the relation

$$\int_0^T f(|Z_t|)\mathrm{d}t = \int_{-\infty}^\infty f(|a|)\ell_T^a \mathrm{d}a = \int_0^\infty f(a)\bar{\ell}_T^a \mathrm{d}a.$$

In case X^u is a centered stationary Ornstein-Uhlenbeck process solution of (5.1) the supremum of $(X_t^u)_{t\in[0,T]}$ over $[0,T]$ admits the following asymptotic estimate.

Proposition 6.4. *Assume that $\|\alpha\|_{L^\infty([0,\infty))} < \infty$ and $\|u\|_{L^\infty([0,\infty))} < \infty$. Then for any $r > 1$ we have*

$$\lim_{T\to\infty} \mathbb{P}_u\left(\sup_{t\in[0,T]} |Z_t| \le \sqrt{2r\log T}\right) = 1.$$

Proof. From Theorem 1.1 of [20] (see also [11], Thm. 2.1 of [17], and [4], p. 488) there exists a universal constants $c_1, c_2 > 0$ such that for all $\lambda, T > 0$,

$$\mathbb{P}_u\left(\sup_{t\in[0,T]} |Z_t| > \lambda\right) \le c_1 M(2aT, c_2/\lambda)\Psi(\lambda),$$

where $\Psi(x) = \int_x^\infty \mathrm{e}^{-y^2/2}\mathrm{d}y/\sqrt{2\pi}$ and $M(2aT, c_2/\lambda)$ is the maximal cardinal of all sequences $\mathcal{S}$ in $[0, 2aT]$ such that

$$\|Z_t - Z_s\|_{L^2(\Omega)} = \sigma\sqrt{\frac{1 - \mathrm{e}^{-a|t-s|}}{a}} > \frac{c_2}{\lambda}, \qquad s, t \in \mathcal{S}.$$

Setting $\lambda = \sqrt{2r\log T}$, $r > 0$, $T > 1$, and using the bound $\Psi(\lambda) \le \mathrm{e}^{-\lambda^2/2}/(\lambda\sqrt{2\pi})$ this yields, for all T large enough:

$$\mathbb{P}\left(\sup_{t\in[0,T]} |Z_t| \le \sqrt{2r\log T}\right) \ge 1 - c\frac{r}{\sqrt{a}}T^{1-r},$$

which tends to 1 as $T \to \infty$ provided $r > 1$. $\qquad\square$

References

[1] J.M. Azaïs and M. Wschebor, *Level Sets and Extrema of Random Processes and Fields*. Wiley, Hoboken (2009).

[2] E.N. Belitser and B.Y. Levit, On minimax filtering over ellipsoids. *Math. Methods Statist.* **4** (1995) 259–273.

[3] S.M. Berman, Local times and sample function properties of stationary Gaussian processes. *Trans. Amer. Math. Soc.* **137** (1969) 277–299.

[4] J. Cuzick, Boundary crossing probabilities for stationary Gaussian processes and Brownian motion. *Trans. Amer. Math. Soc.* **263** (1981) 469–492.

[5] D.L. Donoho and I.M. Johnstone, Ideal spatial adaptation by wavelet shrinkage. *Biometrika* **81** (1994) 425–455.

[6] D.L. Donoho and I.M. Johnstone, Adapting to unknown smoothness *via* wavelet shrinkage. *J. Amer. Statist. Assoc.* **90** (1995) 1200–1224.

[7] D. Geman and J. Horowitz, Occupation densities. *Ann. Probab.* **8** (1980) 1–67.

[8] G.K. Golubev, Minimax filtration of functions in L_2. *Probl. Inf. Transm.* **18** (1982) 272–278.

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

[10] M. Nussbaum, Minimax risk, Pinsker bound, in *Encyclopedia of Statistical Sciences*, S. Kotz Ed. Wiley, New York (1999).

[11] J. Pickands, Upcrossing probabilities for stationary Gaussian processes. *Trans. Amer. Math. Soc.* **145** (1969) 51–73.

[12] M.S. Pinsker, Optimal filtration of square-integrable signals in Gaussian noise. *Probl. Inf. Transm.* **16** (1980) 52–68.

[13] H.V. Poor, *An introduction to signal detection and estimation*. Springer Texts in Electrical Engineering. Springer-Verlag, New York, second edition (1994).

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

[15] N. Privault and A. Réveillac, Stein estimation for the drift of Gaussian processes using the Malliavin calculus. *Ann. Stat.* **35** (2008) 2531–2550.

[16] N. Privault and A. Réveillac, Stein estimation of Poisson process intensities. *Stat. Inference Stoch. Process.* **12** (2009) 37–53.

[17] C. Qualls and H. Watanabe, Asymptotic properties of Gaussian processes. *Ann. Math. Statist.* **43** (1972) 580–596.

[18] D. Revuz and M. Yor, *Continuous martingales and Brownian motion*, Vol. 293 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, third edition (1999).

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

[20] M. Weber, The supremum of Gaussian processes with a constant variance. *Prob. Th. Rel. Fields* **81** (1989) 585–591.