M A T H E M A T I C A L   M E T H O D S   O F   S T A T I S T I C S

EXACT ADAPTIVE ESTIMATION OF THE SHAPE OF A PERIODIC FUNCTION WITH UNKNOWN PERIOD CORRUPTED BY WHITE NOISE

I. Castillo[1], C. Lévy-Leduc[2], and C. Matias[3]

[1]UMR CNRS 8628. Équipe Probabilités, Statistique et Modélisation
Bâtiment 425, Université de Paris-Sud, 91405 Orsay Cedex, France
E-mail: ismael.castillo@math.u-psud.fr

[2]UMR CNRS 5141. École Nationale Supérieure des Télécommunications
Département TSI, 37-39 Rue Dareau, 75014 Paris, France
E-mail: levyledu@enst.fr

[3]UMR CNRS 8071. Laboratoire Statistique et Génome
Tour Evry 2, 523 pl. des Terrasses de l'Agora, 91 000 Evry, France
E-mail: matias@genopole.cnrs.fr

We consider nonparametric estimation of the shape of a periodic function with unknown period θ observed in the presence of additive Gaussian white noise. In this semiparametric framework, estimators of the period with a parametric rate of convergence have been proposed in [11], [3], and [10]. The existence of such a preliminary estimator of θ enables us to introduce an estimation procedure of the shape of the periodic function, using Stein's blockwise method. This estimator is sharp minimax adaptive on a scale of a family of Sobolev classes. The results are illustrated by a simulation study and compared with blind methods which do not use the periodicity assumption on the signal.

Key words: periodic function, unknown period, Stein's blockwise method, adaptive estimation, semiparametric model, Sobolev class.

2000 Mathematics Subject Classification: 62G05, 62G20, 62G08, 62M09.

1. Introduction

Consider the observation of the process $\{X_t\}_{|t| \leq T/2}$ on the time interval $[-T/2; T/2]$ satisfying the diffusion equation

$$(1) \qquad dX_t = f(t/\theta)dt + dW_t,$$

where $\theta \in (0; +\infty)$ is the unknown period of the signal, the (rescaled) unknown function f is periodic with period 1 and $\{W_t\}$ is the standard Brownian motion. We assume that f belongs to $\mathbb{L}_2([0; 1])$, so that we can consider its Fourier coefficients

$$\forall k \in \mathbb{Z}, \quad c_k = c_k(f) = \int_0^1 f(x) e^{-2i\pi k x} \, dx$$

(in the following, the dependence of the Fourier coefficients on f is dropped).

Such a model arises in a wide variety of areas, e.g., in communication, radio location of objects, seismic signal processing, and computer assisted medical diagnosis. To our knowledge, previous works on this type of data only deal with the problem of estimation of the unknown period, in an entirely parametric or in a semiparametric framework.

The aim of this paper is to give an estimation procedure of the function f (the shape of the signal) in the presence of the nuisance parameter θ using a plug-in method of a preliminary estimator $\hat{\theta}_T$ of this unknown period θ.

Nonparametric estimation using the plug-in of a parametric component in a semiparametric framework is a complicated problem and, to our knowledge, there is no general theoretical solution regarding the convergence of such procedures (we refer to [23], Chapter 25, for a general presentation of estimation in semiparametric models). In fact, in some cases, convergence rates may be lowered by the plug-in operation. For instance, in a convolution setting, Butucea and Matias [2] studied the plug-in of an estimator of a scale parameter appearing in the additive noise into a kernel estimator of the deconvolution density. In this context, the unknown parameter acts as a real nuisance since the rates of convergence for the deconvolution density estimator are lowered as compared to the case of a known scale, those rates being nonetheless optimal in a minimax sense.

It may be also interesting to note that our problem is very similar to those arising in the framework of diffusion processes satisfying the stochastic differential equation $dX_t = g(X_t)dt + dW_t$. Indeed, estimation of the trend coefficient g in such models is an extensively studied subject. For instance, this problem has been solved for ergodic processes by Dalalyan and Kutoyants [5] relying on estimation of the invariant probability density of the process $\{X_t\}_t$ (see also [16]). More recently, Loukianov and Loukianova [18] have proposed a new approach in the whole recurrent case (ergodic plus null-recurrent) for Nadaraya–Watson type estimators, which is based on uniform deterministic equivalent for additive functionals of the process. Let us also mention that in the case of discrete-time observations of diffusion processes satisfying the differential equation $dX_t = g(X_t) \, dt + \sigma(X_t) \, dW_t$, another interesting problem is to estimate the diffusion coefficient σ (both in the parametric and non-parametric cases) in the presence of unknown trend coefficient g (see, for instance, [13]).

Let us come back to the observation of a periodic function in the presence of additive Gaussian white noise (1). In the classical case, where the period θ is known, the model simply reduces to a Gaussian sequence space model, using, for instance, a projection on the Fourier basis $\{t \to \exp(2i\pi k t/\theta)\}_{k \in \mathbb{Z}}$ of $\mathbb{L}_2([0, \theta])$. Indeed, let us introduce $[T/\theta]$, the integer part of T/θ. The observation of the process $\{X_t\}_{|t| \le T/2}$

induces the observation of its projection

$$(2) \quad x_k = \frac{1}{\theta[T/\theta]} \int_{-\theta/2[T/\theta]}^{\theta/2[T/\theta]} e^{-2i\pi kt/\theta} \, dX_t = c_k + \frac{1}{\sqrt{\theta[T/\theta]}} \, \xi_{k,[\,]}(\theta), \quad k \in \mathbb{Z},$$

where the random terms

$$\xi_{k,[\,]}(\theta) = (\theta[T/\theta])^{-1/2} \int_{-\theta/2[T/\theta]}^{\theta/2[T/\theta]} e^{-2i\pi kt/\theta} \, dW_t$$

are independent random variables. Moreover, $\xi_{k,[\,]}(\theta) = (1/\sqrt{2})(v + iw)$, where v and w are independent standard Gaussian random variables.

Nonparametric estimation in the presence of additive Gaussian white noise is a widely studied subject. Among many others, we shall refer to [20, 15, 9], as well as to [7, 8, 19, 4] for adaptive versions, and the references therein. We shall later discuss the approach of [20, 4], since our work relies on it. In the framework of nonparametric estimation, we also want to quote the very general results obtained by Golubev [12]. In this paper, he proposes a way to derive exact lower bounds for many nonparametric estimation problems based on a LAN (local asymptotic normality) property of the model at stake.

Since θ is unknown, model (2) is not observed. But the above projection can be replaced by a random projection onto the set of functions $\{t \to \exp(2i\pi kt/\hat{\theta}_T)\}_{k \in \mathbb{Z}}$, where $\hat{\theta}_T$ is any consistent estimator of θ based on the observation of $\{X_t\}_{|t| \leq T/2}$.

It is known since the seminal work of Golubev [11] that in this semiparametric framework there exist consistent estimators of θ. We will later discuss and use a modified version of Golubev's estimator introduced by Castillo [3]. We also mention that the same kind of estimator has been provided by Gassiat and Lévy-Leduc [10] in a discretized version of model (1).

Let us assume for a moment that a consistent estimator $\hat{\theta}_T$ of θ is chosen. We want to define the projection of the observation process onto the random set of functions $\{t \to \exp(2i\pi kt/\hat{\theta}_T)\}_{k \in \mathbb{Z}}$, namely

$$(3) \qquad \int_{-T/2}^{T/2} e^{-2i\pi kt/\hat{\theta}_T} \, dX_t.$$

Note that this quantity cannot be defined in Itô's integration framework, since the process $\{e^{-2i\pi kt/\hat{\theta}_T(\omega)}\}_t$ is generally not adapted with respect to $\{\mathcal{F}_t\}$, which is the σ-field generated by the random variables $\{W_s, -T/2 \leq s \leq t\}$. Indeed, the estimators $\hat{\theta}_T$ are built using the whole observation process. This difficulty is overcome using the fact that the process $\{e^{-2i\pi kt/\hat{\theta}_T(\omega)}\}_t$ has paths with finite total variation. Thus denoting

$$u_{\hat{\theta}_T}(t) = \exp(-2ik\pi t/\hat{\theta}_T),$$

we obtain that

$$\int_{-T/2}^{T/2} X_t \, u_{\hat{\theta}_T}(dt)$$

is well defined as a Stieltjes integral. This remark allows us to define (3) as

$$\int_{-T/2}^{T/2} e^{-2ik\pi t/\hat{\theta}_T}\, dX_t = \int_{-T/2}^{T/2} e^{-2ik\pi t/\hat{\theta}_T} f(t/\theta)\, dt + \int_{-T/2}^{T/2} u_{\hat{\theta}_T}(t)\, dW_t,$$

where

$$(4) \quad \int_{-T/2}^{T/2} u_{\hat{\theta}_T}(t)\, dW_t \triangleq u_{\hat{\theta}_T}(T/2)W_{T/2} - u_{\hat{\theta}_T}(-T/2)W_{-T/2} - \int_{-T/2}^{T/2} W_t\, u_{\hat{\theta}_T}(dt).$$

For further details, see subsection 4.3.

Finally, the observation of the process $\{X_t\}_t$ as in (1) induces an approximate Gaussian sequence space model with observations

$$(5) \qquad z_k = \frac{1}{T}\int_{-T/2}^{T/2} e^{-2i\pi kt/\hat{\theta}_T}\, dX_t = \gamma_k(\theta;\hat{\theta}_T) + \frac{1}{\sqrt{T}}\xi_k(\hat{\theta}_T), \quad k\in\mathbb{Z},$$

where

$$(6) \qquad\qquad \gamma_k(\theta;\hat{\theta}_T) = T^{-1}\int_{-T/2}^{T/2} f(t/\theta)e^{-2i\pi kt/\hat{\theta}_T}\, dt,$$

$$(7) \qquad\qquad \xi_k(\hat{\theta}_T) = T^{-1/2}\int_{-T/2}^{T/2} u_{\hat{\theta}_T}(t)\, dW_t.$$

Note that $\xi_k(\hat{\theta}_T)$ is not anymore a Gaussian random variable nor $\gamma_k(\theta;\hat{\theta}_T)$ is a deterministic term. We shall also consider another Gaussian sequence space model (still unobserved in our context)

$$(8) \qquad\qquad\qquad y_k = c_k + \frac{1}{\sqrt{T}}\xi_{k,[\ ]}(\theta),$$

which is similar to (2) except for the normalization factor. Our work is based on the fact that models (8), (5), and (2) are close enough (in a sense to be specified), so that the observed model (5) inherits classical results on the estimation of the signal from models (2) or (8).

Note that there is no restriction in using the Fourier basis. Actually, the use of any periodic orthonormal basis (for instance, periodic wavelets basis defined in [6], Chapter 9) would lead to the same kind of results as long as we can still define the projection model (5).

Estimation of the signal in model (1) relies on regularity assumptions on f. From now on, we assume that f belongs to a (periodic) Sobolev ball

$$W(\beta, L) = \Big\{ f\ 1\text{-periodic},\ f\in\mathbb{L}_2([0;1]);\ \sum_{k\in\mathbb{Z}} |2\pi k|^{2\beta}|c_k|^2 \le L \Big\}.$$

In the Gaussian sequence space model (8), classical linear estimates of the form $(\lambda_k y_k)_{k\in\mathbb{Z}}$ constructed with weight sequences $(\lambda_k)_{k\in\mathbb{Z}}$ in $\ell_2(\mathbb{Z})$, give consistent estimators of the Fourier coefficients $(c_k)_{k\in\mathbb{Z}}$ of f. Pinsker [20] introduces a particular

choice of the weights $(\lambda_k)_{k \in \mathbb{Z}}$ and establishes that the corresponding linear estimator is asymptotically minimax exact among all estimation procedures of the Fourier coefficients $(c_k)_{k \in \mathbb{Z}}$ of f in a Sobolev ball $W(\beta, L)$ (see also [1] or [22], Chapter 3, for a complete overview on this topic). Pinsker's weights depend on the smoothness parameter β and on the bound L of the Sobolev ball, and thus Pinsker's estimator fails to be adaptive.

A natural substitute for linear estimates is obtained when the weights $(\lambda_k)_{k \in \mathbb{Z}}$ are data-driven, leading to non-linear estimators. The optimal choice of such a sequence of weights is commonly considered in two different ways, using oracle inequalities or minimax adaptivity. On the one hand, oracle inequalities say that an estimator mimics the *oracle* within a fixed class of weights. The oracle consists in choosing weights giving the lowest risk but depending on the unknown parameters and thus not leading to an estimator. On the other hand, minimax adaptivity says that an estimator achieves the minimax rate of convergence on every set of a fixed class of sets. Cavalier and Tsybakov [4] construct a sequence of weights $(\lambda_k)_{k \in \mathbb{Z}}$ using Stein's blockwise method, such that the corresponding linear estimator satisfies both an exact oracle inequality within the class of monotone weights and sharp minimax adaptive property on the scale of Sobolev balls $\{W(\beta, L); \beta \geq 2, L > 0\}$.

Our approach relies on known results about nonparametric estimation of the function f when its period θ is known. Since one natural aim would be to get information about the whole signal $g = f(\cdot/\theta)$, we then use in practice (and for the second time, since $\hat{f}_T$ is already built on some estimator of θ) an estimator $\hat{\theta}_T$ of θ in order to finally obtain an estimator $\hat{g} = \hat{f}_T(\cdot/\hat{\theta}_T)$ of g. Note first that we do not theoretically assess the quality of the estimation of g itself. There is no natural way to define a risk for the estimation of g. Indeed, the function g does not naturally belong to some fixed normed vector space: g does not belong to $\mathbb{L}_2(\mathbb{R})$ but to the space $\mathbb{L}_2([0, \theta])$ where θ depends on g.

We also want to mention that another approach is to directly estimate the function g from the observation of $\{X_t = \int_0^t g(s)ds + W_t\}_{|t| \leq T/2}$ (i.e., without using its periodicity). This approach that we call the *blind method* is compared to ours in Section 3. The comparison of the two methods is based on an empirical criterion (see Section 3 for more details).

In this paper, we first prove that Pinsker's estimator combined with a consistent (and well-chosen, see assumptions in subsection 2.1) estimator $\hat{\theta}_T$ of the unknown period θ is asymptotically minimax exact in model (1) when the unknown function f belongs to the Sobolev ball $W(\beta, L)$ (Theorem 3). Then, applying Stein's blockwise method combined with the use of $\hat{\theta}_T$, we obtain an exact oracle inequality within the class of monotone sequences of weights (Proposition 2) and the sharp minimax adaptive property within the family of Sobolev balls $\{W(\beta, L); \beta \geq 2, L > 0\}$ (Theorem 4). Section 3 deals with practical implementation of our method. The proofs are postponed to Section 4.

2. Estimation Procedure

2.1. MAIN ASSUMPTIONS. We say that a function is a $o(1)$ (respectively $O(1)$) if it tends to zero (resp. is bounded) when T goes to infinity. The notation $\mathbb{E}_{\theta, f}$ denotes expectation with respect to the distribution of the process $\{X_t\}_{|t| \leq T/2}$ given by model (1) with unknown parameters θ and f.

Assumptions on f. For any real numbers $h > 0$ and $M > 0$, denote by $\mathcal{F}(h, M)$ the class of smooth functions whose Fourier coefficients (c_k) satisfy:

$$\forall r \geq 2, \quad \sum_{k \in \mathbb{Z}^*} |c_{rk}|^2 \leq (1-h) \sum_{k \in \mathbb{Z}^*} |c_k|^2 \quad \text{and} \quad \sum_{k \in \mathbb{Z}} |c_k| \leq M,$$

where r denotes an integer and $\mathbb{Z}^*$ denotes the set $\mathbb{Z} \backslash \{0\}$. In particular, this implies that f has period θ but not θ/r for some integer $r \geq 2$.

We assume that there exist constants $h > 0$, $M > 0$, $\beta \geq 2$, and $L > 0$ such that:

(**F**) $$f \in \mathcal{F}(h, M) \cap W(\beta, L).$$

Assumptions on the parameter θ. We assume that $\theta \in [\alpha_T, \beta_T]$ and that this set asymptotically covers $]0, +\infty[$ in the following way:

(**P1**) $$\lim_{T \to +\infty} \alpha_T = 0 \quad \text{and} \quad \frac{1}{\alpha_T} = O(T),$$

(**P2**) $$\lim_{T \to +\infty} \beta_T = +\infty \quad \text{and} \quad \beta_T = O(\log T).$$

We do not propose here any new method of estimation of the unknown period θ, but use one of the existing procedures, highlighting the main properties that are useful in our context. The first estimation procedure of θ in the semiparametric framework given by (1) is due to Golubev [11]. The idea is to use an approximate profile likelihood. The estimator $\hat{\theta}_T$ maximizes (in a proper way) with respect to τ the following criterion Λ, where $N(T)$ is a quantity to be well chosen and going to infinity as $T \to \infty$:

(9) $$\Lambda(\tau) = \sum_{k=1}^{N(T)} \frac{1}{T} \left| \int_{-T/2}^{T/2} e^{2i\pi kt/\tau} \, dX_t \right|^2.$$

Golubev defines an estimator $\hat{\theta}_T$, which is asymptotically efficient in the sense:

$$\lim_{T \to \infty} \mathbb{E}\left((\hat{\theta}_T - \theta)^2 I_T(\theta, f) \right) = 1,$$

where $I_T(\theta, f)$ is the Fisher Information in model (1) asymptotically given by

$$I_T(\theta, f) = (1 + o(1)) \frac{T^3}{12\theta^4} \sum_{k \in \mathbb{Z}} (2\pi k)^2 |c_k|^2,$$

see [14]. In [3], Castillo considers a weighted version of (9) and establishes second order properties of the quadratic risk for a large class of weighted estimators.

Here, we shall need to control the tail probabilities in the estimation of θ as follows.

Assumptions on the estimator $\hat{\theta}_T$.

(C)   There exist some $p > 23/5$ and some positive constant C such that:

$$\mathbb{P}\big(T^{3/2}\theta^{-2}|\hat{\theta}_T - \theta| \geq C\sqrt{\log T}\big) = o(T^{-p}).$$

Remark 1. There exist estimators satisfying condition **(C)**: for instance, the weighted estimators proposed by Castillo [3] satisfy Assumption **(C)** for any integer p (see Eq. (16) in [3]). However, to obtain this property, slightly stronger assumptions on the function f are needed (see [3] for more details). The constant $23/5$ comes from technical reasons and may not be optimal.

Let us mention that both the estimator $\hat{\theta}_T$ introduced in [3] and our estimation procedure of f do not depend on the parameters h and M introduced in Assumption **(F)** and thus, the method is adaptive with respect to those parameters.

2.2. LINEAR ESTIMATORS OF f. Recall that the main idea of the construction of our estimator of the function f is that the observed model (5) is not very far away from the non-observed models (2) or (8), where efficient estimators of f are known. Hence, we first follow the ideas of Pinsker [20] to construct a linear estimator $\hat{f}_T$ of f. Recall the definition of Pinsker's weights denoted by $(q_k)_{k\in\mathbb{Z}}$ in our setup. For any real number u, we denote by $(u)_+$ the quantity $\max(u,0)$, then Pinsker's weights are defined by:

$$(10) \qquad \forall\, k \in \mathbb{Z}, \quad q_k = \big(1 - w|2\pi k|^{2\beta}\big)_+,$$

where w is the solution of the equation

$$\frac{1}{wT} \sum_{k\in\mathbb{Z}} |2\pi k|^{2\beta}\big(1 - w|2\pi k|^{2\beta}\big)_+ = L.$$

Moreover, one can establish that, as T goes to infinity,

$$w = \left(\frac{\beta\pi^{2\beta}}{(2\beta+1)(\beta+1)L}\right)^{\frac{\beta}{2\beta+1}} T^{\frac{-\beta}{2\beta+1}}(1 + o(1)).$$

Remark 2. Note, in particular, that Pinsker's weights are equal to zero for $|k| > N_0$, where N_0 tends to infinity at the rate $T^{1/(2\beta+1)}$. Since $\beta \geq 2$, we have $N_0 = O(T^{1/5})$.

Let us introduce, for any sequence of weights $\lambda \in [0;1]^{\mathbb{Z}}$, he following functional:

$$(11) \qquad\qquad R(\lambda, f) \triangleq \sum_{k\in\mathbb{Z}}(1 - \lambda_k)^2|c_k|^2 + \frac{\lambda_k^2}{T}.$$

This quantity is nothing but the quadratic risk $\sum_{k\in\mathbb{Z}}\mathbb{E}\left(|\lambda_k y_k - c_k|^2\right)$ associated with the linear estimator $(\lambda_k y_k)_{k\in\mathbb{Z}}$ in model (8). Note that, as explained for instance in [22] (Section 3.5), there is no restriction in considering sequences of weights with

values in $[0;1]^{\mathbb{Z}}$, since projection of the weights onto $[0;1]^{\mathbb{Z}}$ obviously decreases the risk. Then, as T goes to infinity, Pinsker's Theorem (see [20]) gives:

$$\text{(12)} \qquad \lim_{T \to +\infty} \sup_{f \in W(\beta, L)} R(q, f) T^{2\beta/(2\beta+1)}$$

$$= \lim_{T \to +\infty} \inf_{\lambda \in [0;1]^{\mathbb{Z}}} \sup_{f \in W(\beta, L)} R(\lambda, f) T^{2\beta/(2\beta+1)}$$

$$= \lim_{T \to +\infty} \inf_{f_T} \sup_{f \in W(\beta, L)} \mathbb{E}\left(\|f - f_T\|^2\right) T^{2\beta/(2\beta+1)} = C^*,$$

where the second infimum is taken over all estimators f_T of f and C^* is Pinsker's constant:

$$\text{(13)} \qquad C^* = [L(2\beta+1)]^{\frac{1}{2\beta+1}} \left(\frac{\beta}{\pi(\beta+1)}\right)^{\frac{2\beta}{2\beta+1}}.$$

Unfortunately, model (8) is unobserved. However, we shall prove that we can achieve the same sharp rate using projection model (5).

Define the linear estimator $\hat{f}_T$ in the following way: if f belongs to $W(\beta, L)$ and $\hat{\theta}_T$ is a preliminary estimator of θ used to define the projection model (5), let

$$\text{(14)} \qquad \hat{f}_T(x) = \sum_{k \in \mathbb{Z}} q_k z_k e^{2i\pi k x},$$

where the q_k's are Pinsker's weights defined by (10) and associated with $W(\beta, L)$, and the z_k's are defined in (5). In the sequel, the Fourier coefficients of the function $\hat{f}_T$ are denoted by $\hat{c}_k$. Note that the estimator $\hat{f}_T$ in (14) depends on the plugged $\hat{\theta}_T$ only via z_k. From (14), we get: $\hat{c}_k = q_k z_k$, for any $k \in \mathbb{Z}$.

The linear estimator obtained with Pinsker's weights is asymptotically exact minimax in our setup, as ensured by the following theorem to be proved in Section 4.

Theorem 3. *Fix $\beta \geq 2$ and $L > 0$. Under the assumptions* **(F)**, **(P1)**, **(P2)**, **(C)**, *the estimator $\hat{f}_T$ defined by (14) satisfies*:

$$\lim_{T \to +\infty} \sup_{\theta \in [\alpha_T; \beta_T]} \sup_{f \in W(\beta, L)} T^{2\beta/(2\beta+1)} \mathbb{E}_{\theta, f} \|\hat{f}_T - f\|_2^2$$

$$= \lim_{T \to +\infty} \inf_{f_T} \sup_{\theta \in [\alpha_T; \beta_T]} \sup_{f \in W(\beta, L)} T^{2\beta/(2\beta+1)} \mathbb{E}_{\theta, f} \|f_T - f\|_2^2 = C^*,$$

where the infimum is taken over all estimators f_T based on the observation of $\{X_t\}_{|t| \leq T/2}$ described by model (1) and C^* is Pinsker's constant defined in (13).

Theorem 3 provides an efficient and minimax procedure of estimation of f in the class $W(\beta, L)$. However, in practice, the parameters β and L are unknown. Thus, we would like to construct an estimator of f which does not require the knowledge of the regularity parameters β and L. This issue is treated in the following subsection.

2.3. STEIN'S BLOCKWISE PROCEDURE. In order to construct an adaptive estimator of f, i.e., an estimator which does not require the knowledge of β and L,

we let the sequence of weights depend on the data $\{X_t\}_{|t|\leq T/2}$. Let us first recall the idea of Stein's blockwise procedure as explained in [4]. Define a partition of the interval $[-N_{\max}; N_{\max}]$, where $N_{\max} = T^{1/4}$, in subintervals B_j called blocks for $j = -J, \ldots, J$. We denote by T_j the cardinality of B_j. We shall make precise the construction of the B_j's in the sequel.

Given blocks B_j's with cardinalities T_j's, the positive James–Stein's weights are defined by:

$$(15) \qquad \psi_k(z) = \sum_{j=-J}^{J} \left(1 - \frac{T_j}{T\|z\|_j^2}\right)_+ \mathbf{1}_{k\in B_j}, \qquad \text{for} \quad k \in \mathbb{Z},$$

where $\|z\|_j$ denotes the $\mathbb{L}_2$-norm on the block B_j, i.e., $\|z\|_j^2 = \sum_{l\in B_j} |z_l|^2$ and $\mathbf{1}$ is the indicator function. The idea is to take into account only the observations such that the energy $\|z\|_j^2$ on the jth block is larger than the expected level of the noise T_j/T on the same block. Remark also that the weights $\psi_k(z)$ are constant over each block and that in fact $\psi_k(z) = 0$ for $|k| > T^{1/4}$.

2.3.1. *Definition of B_j's and T_j's.* We choose weakly geometrically increasing blocks as introduced in [14] or in [22], Section 3.6 (here the construction slightly differs from the quoted references since we work with complex Fourier coefficients c_k's and not real ones). Let $\rho_T = \log^{-1}(T)$ and define the T_j's for $j \geq 0$ as follows:

$$T_1 = \lceil \rho_T^{-1}\rceil = \lceil \log(T)\rceil, \qquad T_2 = \lfloor T_1(1 + \rho_T)\rfloor, \qquad \ldots$$

$$(16)$$

$$T_{J-1} = \lfloor T_1(1 + \rho_T)^{J-2}\rfloor, \qquad T_J = N_{\max} - \sum_{j=1}^{J-1} T_j,$$

where

$$(17) \qquad J = \min\left\{m, \ T_1 + \sum_{j=2}^{m} \lfloor T_1(1 + \rho_T)^{j-1}\rfloor \geq N_{\max}\right\}.$$

Here, $\lfloor x\rfloor$ is the largest integer smaller than the real number x and $\lceil x\rceil$ is the smallest integer larger than x.

Now let us define blocks B_j^+'s as the partition of $\{1, 2, \ldots, N_{\max}\}$ such that $B_1^+ = \{1, \ldots, T_1\}$ and $\min\{k \in B_j^+\} > \max\{k \in B_{j-1}^+\}$ for any $2 \leq j \leq J$. Then for all $j \in \{1, \ldots, J\}$, define $B_j^- = \{-k, \ k \in B_j^+\}$. Finally, let

$$(18) \qquad B_1 = B_1^- \cup \{0\} \cup B_1^+ \quad \text{and} \quad B_j = B_j^- \cup B_j^+ \quad \text{for} \quad j \in \{2, \ldots, J\}.$$

2.3.2. *Properties of weakly geometrically increasing weights.* Let us recall some well-known properties of this system of blocks (see, for instance, [4] or [22], Section 3.6).

Lemma 1 (Lemma 3.11 in [22]). *Let $\{B_j\}$ be the system of blocks defined above. Then for T large enough, there exists a constant C such that the number J of blocks satisfies*:

$$(19) \qquad J \leq C\log^2(T).$$

Now recall the following inequality, which is satisfied in the non-observed model (8). Let f_T^* be the natural estimator if model (8) were observed:

$$(20) \qquad f_T^* = \sum_{k \in \mathbb{Z}} \psi_k(y) y_k e^{2ik\pi x}.$$

(Note that f_T^* is not an estimator in our setup). In the sequel, we shall say that a sequence of weights $(\lambda_j)_{j \in \mathbb{Z}} \in [0;1]^{\mathbb{Z}}$ is symmetric and decreasing when $\lambda_j = \lambda_{-j}$ for any $j \geq 1$ and the sequence $(\lambda_j)_{j \geq 1}$ is decreasing.

Lemma 2 (Theorem 2, Proposition 6 in [4], Theorem 3.6 of [22]) *Let Λ_{mon} be the class of symmetric and decreasing weights $(\lambda_j)_{j \in \mathbb{Z}} \in [0;1]^{\mathbb{Z}}$ such that $\lambda_j = 0$ for $|j| > N_{\max}$. Then for T large enough, if f_T^* is defined by (20), there exists a constant C such that*

$$(21) \qquad \mathbb{E}\|f_T^* - f\|^2 \leq (1 + 3\rho_T)\min_{\lambda \in \Lambda_{\mathrm{mon}}} R(\lambda, f) + \frac{C}{T}\log^2(T),$$

where $\rho_T = \log^{-1}(T)$.

The setup of this lemma is exactly the same as the one in [4] or [22], except for the fact that here, $N_{\max} = T^{1/4}$ (instead of $N_{\max} = T$). But one can easily see that this only changes the constant C in the above theorem by a multiplicative factor and hence the result still holds in our setup.

2.3.3. *Data-driven estimator of f.* If $\hat{\theta}_T$ is a preliminary estimator of θ used to define the projection model (5), let

$$(22) \qquad \tilde{f}_T(x) = \sum_{k \in \mathbb{Z}} \psi_k(z) z_k e^{2ik\pi x}.$$

In the sequel, the Fourier coefficients of the function $\tilde{f}_T$ are denoted by $\tilde{c}_k$. It follows from (22) that $\tilde{c}_k = \psi_k(z) z_k$ for any $k \in \mathbb{Z}$.

In the following theorem, we establish that $\tilde{f}_T$ is a sharp minimax adaptive estimator of f in model (1) among the family of Sobolev balls $\{W(\beta, L); \beta \geq 2, L > 0\}$.

Theorem 4 (sharp minimax adaptivity). *Under assumptions* **(F)**, **(P1)**, **(P2)**, **(C)**, *the estimator $\tilde{f}_T$ defined by (22) and with a system of blocks defined by (16)–(18), satisfies, for any $\beta \geq 2$ and $L > 0$:*

$$\lim_{T \to \infty} \sup_{\theta \in [\alpha_T; \beta_T]} \sup_{f \in W(\beta, L)} T^{2\beta/(2\beta+1)} \mathbb{E}_{\theta, f}\|\tilde{f}_T - f\|_2^2 = C^*,$$

where C^* is Pinsker's constant defined in (13).

Remark 5. This result means also that the data-driven method introduced in [4] for the Gaussian sequence model is stable with respect to the plug-in of a sufficiently well chosen estimator of the nuisance parameter θ, since we obtain the same sharp rate of convergence. It would be interesting to investigate if this is also

the case in a more general semiparametric setup, i.e., for a more general form (with respect to θ) of the model.

4. Practical Implementation of the Estimation Method

In this section, we deal with the following discrete-time model:

$$X_j = f\left(\frac{j}{n\theta}\right) + \varepsilon_j, \qquad j = 1, \dots, n,$$

where the ε_j's are i.i.d. Gaussian random variables with zero mean and unit variance and f is a periodic function with period 1. Indeed, when we are faced with an observed signal in a practical situation, the data at hand are sampled and the number of observations n is fixed. Then the estimator (22) is constructed using the discrete version of the $z_k's$:

$$z_k = \frac{1}{n} \sum_{j=1}^n e^{-2ik\pi j/(n\hat{\theta})} X_j, \qquad k = 1, 2, \dots$$

The estimation algorithm can be split into several steps, which show how to obtain an estimate of the shape function f by using the previous adaptive procedure.

• The first step consists in estimation of the period θ of the regression function. This can be done either by maximizing a penalized cumulative periodogram of the observations (this method is explained in [17]) or by maximizing a penalized weighted cumulative periodogram of the observations (this approach is developed in [3]).

• At the second step, we use the estimator of θ to compute the z_k's as explained above. Next we obtain Stein's blocks, which allow us to define $\tilde{f}_T$.

• An estimator of the periodic regression function $g = f(\cdot/\theta)$ is obtained by plugging in the estimator of θ obtained at the first step. This enables us to compare our procedure with the *blind method* (see subsection 3.2).

3.1. ILLUSTRATION ON SYNTHETIC DATA.

3.1.1. *Presentation of the synthetic data.* We present hereafter a periodic signal which we aim to estimate from noisy data. We shall estimate the following synthetic signal:

$$(23) \qquad s(t) = a \cos\left[c \cos\left(\frac{2\pi t}{\theta}\right)\right],$$

where $a = 0.1$, $\theta = 1/20$, $c = 15$ in the first case and $c = 150$ in the second one from the data:

$$(24) \qquad X_j = s\left(\frac{j}{n}\right) + \varepsilon_j, \qquad j = 1, \dots, n.$$

This specific form allows us to synthesize very easily signals having as many significant Fourier coefficients as we want.
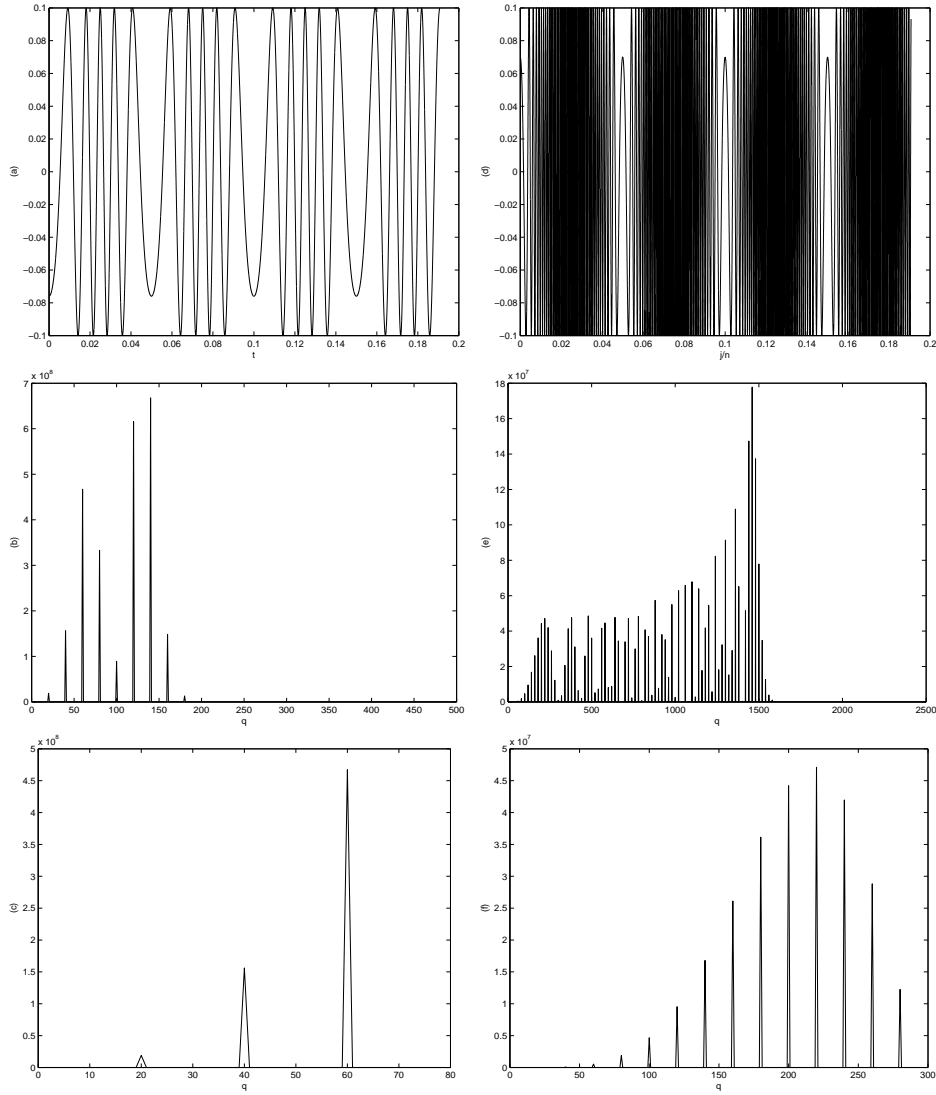
FIGURE 1. A synthetic signal s: (a),(d) $(s(j/n),\ 1 \leq j \leq 200000)$;
(b),(e) the DFT spectrum of s; (c),(f) zoom of (b) and (e)

Note that the parameter θ is assumed to be unknown in our practical estimation
procedure and that this type of signal for s is only used for simulating a periodic
signal but we do not use its parametric form in the algorithm.

With such a definition, s is a periodic function with frequency 20 Hz with
about 15 positive harmonics in the first case and 150 in the second one.

Figures 1 (a), (d) display $(s(j/n),\ j = 1, \dots, 200000)$ with $n = 2^{20}$ for $c = 15$
and $c = 150$ respectively. Figures 1 (b), (e) display, in these two cases, the squared
modulus of the DFT (Discrete Fourier Transform) of $(s(j/n),\ j = 1, \dots, n)$ defined

by

$$I_s(q) = \frac{1}{n}\left|\sum_{j=1}^{n} e^{-\frac{2i\pi qj}{n}} s\left(\frac{j}{n}\right)\right|^2, \qquad 0 \le q \le n-1.$$

We represent the sequence $I_s(q)$ only for $1 \le q \le 500$ and for $1 \le q \le 2500$ respectively in Figures 1 (b), (e).

Figures 1 (c), (f) are respective zooms of Figures 1 (b), (e).

The observed sequence is obtained by adding a Gaussian white noise to s:

$$X_j = s\left(\frac{j}{n}\right) + \varepsilon_j, \qquad j = 1, \dots, n,$$

where the ε_j's are independent Gaussian random variables with zero mean and unit variance.

For a signal of the form (23), we define the signal-to-noise ratio by

$$SNR = 10\log_{10}\left(\frac{a^2}{2}\right).$$

The quantity inside the parentheses is the ratio of the power of the signal to the variance of the noise. With the above values of the parameters, $SNR = -23\,\mathrm{dB}$ in these examples.
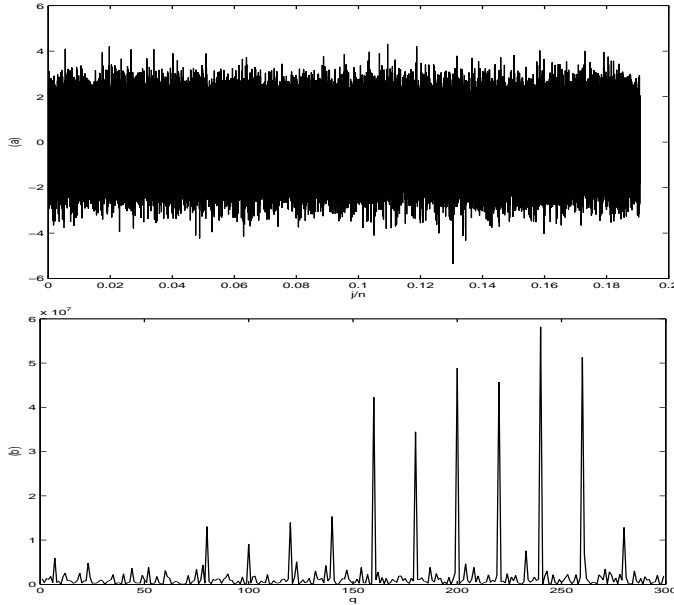

FIGURE 2. The observed signal X: (a) $(X_j,\ 1 \le j \le 200000)$; (b) the DFT spectrum of X

Figures 2 (a), (b) display the observations X_j in the case where $c = 150$ as well as its DFT spectrum. The signal-to-noise ratio is so low in this example that the original signal cannot be visually detected in Figure 2 (a) and the DFT spectrum of the observations is very different from the one obtained without noise displayed on Figure 1 (f).

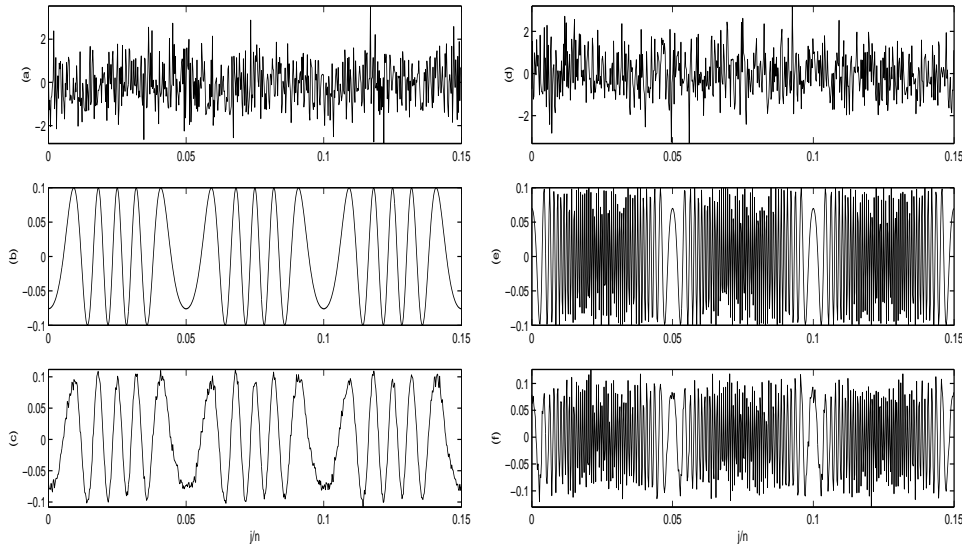3.1.2. *Results of our algorithm.* We give hereafter the denoised signals, that is estimators of the signal s obtained by using our algorithm.

FIGURE 3.   (a), (d) the original data X_j; (b), (e) the original signals s; (c), (f) the denoised signals

Figures 3 (a), (d) display the original data, Figures 3 (b), (e) display the original signals, while our estimators of s appear in Figures 3 (c), (f). In these examples, we can see that our algorithm provides denoised signals which appear to be visually close to the original ones despite the low signal-to-noise ratio. We shall make this 'closeness' more precise in the following section by computing the empirical risk of our estimator defined by (25) in various cases.

3.2. COMPARISON OF OUR METHOD WITH EXISTING ONES. In the following, we illustrate the performance of our algorithm by an example and compare it with that of a method using no periodicity assumption. We shall call it a *blind method*.

More precisely, we synthesize a signal of the same form as the one in (23) but with a parameter a chosen so that to have $SNR = -20$ dB. Then, we use the denoising method implemented in the *Wavemenu* of Matlab with a symlet 6 to the level 9 and a hard thresholding. This type of wavelets seem to be the most adapted to the signals at stake.

Figures 4 (a), (b) display respectively the original signal s with a plain line, the denoised signals obtained using the blind method with crosses and using our algorithm with points.

As can be seen in Figures 4 (a), (b), our estimator performs much better point-wise on the interval than the estimator constructed without using the periodicity assumption. We quantify this fact by performing some Monte-Carlo experiments.
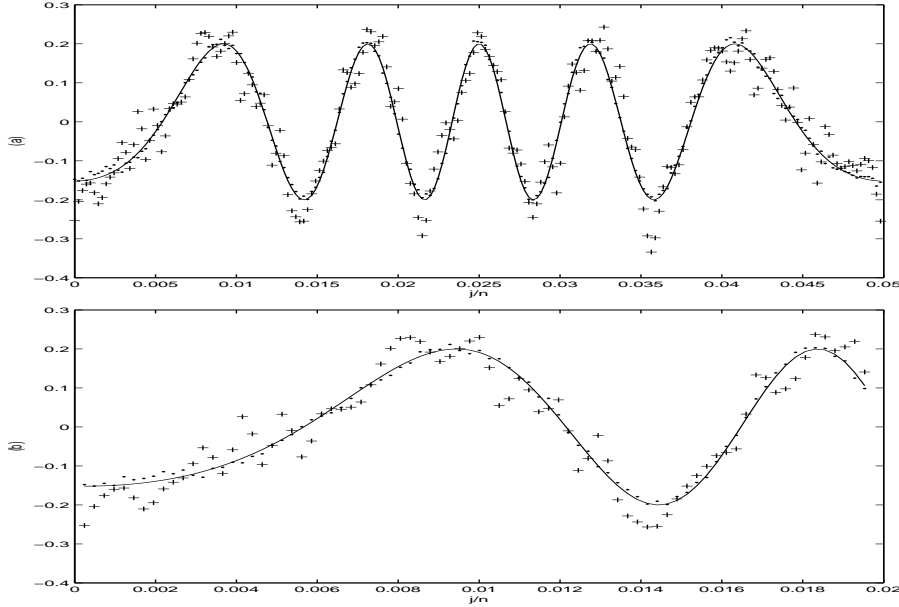

FIGURE 4. Comparison with a blind method; '−': the original signal s,
'+': denoising with a blind method, '·': denoising with our algorithm.
(b) is a zoom of (a)

We propose several numerical comparisons of the two approaches, obtained for a signal s defined by (23) for $c = 15$ and different values of a. For each value of a, we simulate $L = 10$ observed series (X_j) satisfying (24) and estimate s by the two methods proposed previously: our algorithm and the blind method.

The quality of an estimator is quantified by computing the root mean squared error (RMSE) on the whole observation interval defined by:

$$(25) \qquad \sqrt{\frac{1}{L} \sum_{l=1}^{L} \left[\frac{1}{n} \sum_{j=1}^{n} (\hat{s}_l(j/n) - s(j/n))^2 \right]},$$

where $(\hat{s}_l)$ are the L estimates obtained for s.

In Table 1, R_{per}, R_{blind} are the RMSE's for our algorithm and the blind method, respectively, and P_s is an approximation of the $\mathbb{L}_2$-norm of s (the square root of its power) computed by $\sqrt{n^{-1} \sum_{j=1}^{n} s^2(j/n)}$. These Monte-Carlo experiments show, as expected, that our algorithm provides better estimators than a method not using the periodicity of s.

TABLE 1. Comparison of the RMSE's given by a blind
method with our method, where a is defined in (23)

a	0.1	0.5	1	2	4
R_{per}	0.0083	0.0113	0.0162	0.0297	0.0576
R_{blind}	0.0733	0.0768	0.1008	0.0833	0.0795
P_s	0.0676	0.3379	0.6760	1.3519	2.7039
R_{per}/P_s	0.1226	0.0336	0.0241	0.0219	0.0212
R_{blind}/P_s	1.085	0.2272	0.1491	0.0616	0.0294

5. Proofs

In this section, C denotes an absolute nonnegative constant, which may change along the lines. Let us first mention that according to the assumption that f belongs to $W(\beta, L) \cap \mathcal{F}(h, M)$ with $\beta \geq 2$, we have the following uniform bounds:

$$\sum_{k \in \mathbb{Z}} |c_k| \leq M; \qquad \sum_{k \in \mathbb{Z}} |2\pi k|^4 |c_k|^2 \leq L.$$

4.1. PROOF OF THEOREM 3. We first prove here the upper bound part of Theorem 3, namely

$$\limsup_{T \to +\infty} \sup_{\theta \in [\alpha_T; \beta_T]} \sup_{f \in W(\beta, L)} T^{2\beta/(2\beta+1)} \mathbb{E}_{\theta, f} \|\hat{f}_T - f\|_2^2 \leq C^*.$$

Note first that $\mathbb{E}_{\theta, f} \|\hat{f}_T - f\|_2^2 = \mathbb{E} \sum_{k \in \mathbb{Z}} |\hat{c}_k - c_k|^2$, where we drop the dependence on θ and f in the expectation. Let us denote by $\mathrm{Re}(v)$ and $\bar{v}$ respectively the real part and the complex conjugate of v. Now,

$$\mathbb{E}\left[\sum_{k \in \mathbb{Z}} |\hat{c}_k - c_k|^2\right] = \mathbb{E}\left[\sum_{k \in \mathbb{Z}} |q_k z_k - c_k|^2\right] = \mathbb{E}\left[\sum_{k \in \mathbb{Z}} |q_k y_k - c_k|^2\right]$$
$$+ \mathbb{E}\left[\sum_{k \in \mathbb{Z}} |q_k(z_k - y_k)|^2\right] + 2\mathbb{E}\left[\sum_{k \in \mathbb{Z}} \mathrm{Re}(q_k y_k - c_k)\overline{(q_k(z_k - y_k))}\right].$$

We shall prove that the first term is the main one, where (as defined in the Introduction)

$$y_k = c_k + \frac{1}{\sqrt{T}} \xi_{k,[\ ]}(\theta) \quad \text{and} \quad \xi_{k,[\ ]}(\theta) = \frac{1}{\sqrt{\theta[T/\theta]}} \int_{-[T/\theta]\theta/2}^{[T/\theta]\theta/2} e^{-2ik\pi t/\theta} \, dW_t.$$

By the Cauchy–Schwarz inequality, the third term is negligible as soon as it is the case for the second one. Let us then focus on the second term. Introducing the random variable $\xi_k(\theta)$ defined in the same way as $\xi_k(\hat{\theta}_T)$ (see (7)) but using the deterministic period θ, we get

$$(26) \quad \mathbb{E}\left[\sum_{k \in \mathbb{Z}} |q_k(z_k - y_k)|^2\right] \leq 3 \Bigg\{ \mathbb{E}\left[\sum_{k \in \mathbb{Z}} q_k^2 |\gamma_k(\theta, \hat{\theta}_T) - c_k|^2\right]$$
$$+ T^{-1} \mathbb{E}\left[\sum_{k \in \mathbb{Z}} q_k^2 |\xi_k(\hat{\theta}_T) - \xi_k(\theta)|^2\right] + T^{-1} \mathbb{E}\left[\sum_{k \in \mathbb{Z}} q_k^2 |\xi_k(\theta) - \xi_{k,[\ ]}(\theta)|^2\right] \Bigg\}.$$

Consider the quantity $\xi_k(\tau)$, where τ is any deterministic point in $[\alpha_T, \beta_T]$. First, recall that by definition (4), we have

$$T^{1/2}\xi_k(\tau) = u_\tau(T/2)W_{T/2} - u_\tau(-T/2)W_{-T/2} - \int_{-T/2}^{T/2} W_t \, u_\tau(dt).$$

Here, the process $u_\tau(t) = \exp(-2ik\pi t/\tau)$ is deterministic, and we shall check that this quantity corresponds to the classical Itô's integral $\int_{-T/2}^{T/2} \exp(-2ik\pi t/\tau) \, dW_t$. Indeed, using the Integration by parts formula (see, for instance, [21], Proposition 3.1), we have for any continuous semimartingales Y and Z,

$$Y_t Z_t = Y_0 Z_0 + \int_0^t Y_s dZ_s + \int_0^t Z_s dY_s + \langle Y, Z \rangle_t,$$

where $\langle Y, Z \rangle_t$ is the bracket of Y and Z. Applying this formula to $Y_t = W_t$ and $Z_t = u_\tau(t)$, we get

(27) $\quad u_\tau(T/2)W_{T/2} - u_\tau(-T/2)W_{-T/2}$

$$= \int_{-T/2}^{T/2} W_s u_\tau \, (ds) + \int_{-T/2}^{T/2} u_\tau(s) \, dW_s + \langle u_\tau, W \rangle_{T/2} - \langle u_\tau, W \rangle_{-T/2}.$$

Since the process u_τ has finite variation, the bracket $\langle u_\tau, W \rangle$ vanishes, giving the expected result. Moreover, note that $\xi_k(\tau) = 2^{-1/2}(v + iw)$, where v and w are independent standard Gaussian random variables.

Now, let us return to the control of (26). The last term of the upper bound in (26) can be rewritten as follows:

$$T^{-1}\mathbb{E}\left[\sum_{k \in \mathbb{Z}} q_k^2 |\xi_k(\theta) - \xi_{k,[\,]}(\theta)|^2 \right]$$

$$= \sum_{k \in \mathbb{Z}} \mathbb{E}\left[\frac{q_k^2}{T} \left| \frac{1}{\sqrt{T}} \int_{-T/2}^{T/2} e^{\frac{-2ik\pi t}{\theta}} \, dW_t - \frac{1}{\sqrt{\theta[T/\theta]}} \int_{-[T/\theta]\theta/2}^{[T/\theta]\theta/2} e^{\frac{-2ik\pi t}{\theta}} \, dW_t \right|^2 \right]$$

$$\leq 2 \sum_{k \in \mathbb{Z}} \mathbb{E}\left[\frac{q_k^2}{T^2} \left| \int_{-T/2}^{T/2} e^{\frac{-2ik\pi t}{\theta}} \, dW_t - \int_{-[T/\theta]\theta/2}^{[T/\theta]\theta/2} e^{\frac{-2ik\pi t}{\theta}} \, dW_t \right|^2 \right]$$

$$+ 2 \sum_{k \in \mathbb{Z}} \mathbb{E}\left[\frac{q_k^2}{T} \left(\frac{1}{\sqrt{T}} - \frac{1}{\sqrt{\theta[T/\theta]}} \right)^2 \left| \int_{-[T/\theta]\theta/2}^{[T/\theta]\theta/2} e^{\frac{-2ik\pi t}{\theta}} \, dW_t \right|^2 \right].$$

Since

$$\int_{[T/\theta]\theta/2}^{T/2} e^{-2ik\pi t/\theta} \, dW_t$$

is a centered random variable with variance $T/2 - [T/\theta]\theta/2$, which is less than $\theta/2$, the first term in the right-hand side is bounded by $4\theta(\sum_k q_k^2)/T^2$. In the same way, using the bound

$$\left| \frac{1}{\sqrt{T}} - \frac{1}{\sqrt{\theta[T/\theta]}} \right|^2 \leq \frac{|\sqrt{\theta[T/\theta]} - \sqrt{T}|^2}{T\theta[T/\theta]} \leq \frac{|\theta[T/\theta] - T|}{T\theta[T/\theta]} \leq \frac{1}{T[T/\theta]},$$

we get that the second term is less than $2\theta(\sum_k q_k^2)/T^2$. Finally,

$$(28) \qquad T^{-1}\mathbb{E}\Big[\sum_{k\in\mathbb{Z}} q_k^2|\xi_k(\theta) - \xi_{k,[\,]}(\theta)|^2\Big] \le \frac{6\beta_T}{T^2}\Big(\sum_{k\in\mathbb{Z}} q_k^2\Big).$$

Let us now control the first and the second terms of the right-hand side of inequality (26). For this, we shall use the following two lemmas.

Lemma 3. *For $p = 1$ or 2 and for $(\lambda_k)_{k\in\mathbb{Z}} \in [0;1]^{\mathbb{Z}}$ a sequence of weights such that $\lambda_k = 0$ for $|k| > T^{1/4}$, there exists an absolute nonnegative constant C such that*

$$\mathbb{E}\left[\Big(\sum_{k\in\mathbb{Z}} \lambda_k^2|\gamma_k(\theta,\hat\theta_T) - c_k|^2\Big)^p\right]$$
$$\le C\left[\Big(\sum_{k\in\mathbb{Z}} |k|^4\lambda_k^2|c_k|^2\Big)^p\Big(\frac{\log^2 T}{T^2}\Big)^p + \Big(\sum_{k\in\mathbb{Z}} \lambda_k^2\Big)^p\Big(\frac{\beta_T}{T}\Big)^{2p}\right].$$

Lemma 4. *Under the assumptions of Lemma 3, there exists an absolute non-negative constant C such that*

$$\mathbb{E}\left[\Big(\sum_{k\in\mathbb{Z}} \lambda_k^2|\xi_k(\hat\theta_T) - \xi_k(\theta)|^2\Big)^p\right] \le C\Big(\sum_{k\in\mathbb{Z}} \lambda_k^2\Big)^{p-1}\Big(\sum_{k\in\mathbb{Z}} |k|^{2p}\lambda_k^{2p}\Big)\Big(\frac{\log T}{T}\Big)^p.$$

The proofs of these lemmas are postponed to the end of this subsection.

From (26), (28) and Lemmas 3 and 4, which can be applied according to Remark 2, we deduce that

$$\mathbb{E}\left[\sum_{k\in\mathbb{Z}} |q_k(z_k - y_k)|^2\right] \le C\left[\Big(\sum_{k\in\mathbb{Z}} |k|^4 q_k^2|c_k|^2\Big)\Big(\frac{\log^2 T}{T^2}\Big) + \Big(\sum_{k\in\mathbb{Z}} q_k^2\Big)\Big(\frac{\beta_T}{T}\Big)^2\right.$$
$$\left. + \frac{1}{T}\Big(\sum_{k\in\mathbb{Z}} |k|^2 q_k^2\Big)\Big(\frac{\log T}{T}\Big) + \frac{\beta_T}{T^2}\Big(\sum_{k\in\mathbb{Z}} q_k^2\Big)\right].$$

Now we use assumption **(P2)** and Remark 2 to obtain an explicit upper bound for the preceding quantity:

$$\mathbb{E}\left[\sum_{k\in\mathbb{Z}} |q_k(z_k - y_k)|^2\right] \le C\left[\frac{\log^2 T}{T^2} + \frac{\log^2 T}{T^{7/4}} + \frac{\log T}{T^{5/4}}\right] \le C\frac{\log T}{T^{5/4}}.$$

Let us now turn to the first term in the expansion of $\mathbb{E}\left[\sum_{k\in\mathbb{Z}} |\hat c_k - c_k|^2\right]$:

$$\mathbb{E}\left[\sum_{k\in\mathbb{Z}} |q_k y_k - c_k|^2\right] = \sum_{k\in\mathbb{Z}}\left\{(1 - q_k)^2|c_k|^2 + \frac{q_k^2}{T}\right\} \ge T^{-1}\sum_{k\in\mathbb{Z}} q_k^2.$$

But since $q_0 = 1$, we have $\sum_{k \in \mathbb{Z}} q_k^2 \geq 1$. From this and the bounds for the second and third terms in the right-hand side of inequality (26), we deduce that:

$$\mathbb{E}\left[\sum_{k \in \mathbb{Z}} |\hat{c}_k - c_k|^2\right] = \left\{\sum_{k \in \mathbb{Z}} (1 - q_k)^2 |c_k|^2 + \frac{q_k^2}{T}\right\}(1 + o(1)).$$

According to Pinsker's Theorem [20], the last quantity

$$R(q, f) = \sum_{k \in \mathbb{Z}} (1 - q_k)^2 |c_k|^2 + \frac{q_k^2}{T}$$

satisfies

$$\sup_{f \in W(\beta, L)} R(q, f) \leq C^* T^{-2\beta/(2\beta+1)} (1 + o(1)).$$

Since in our setup we have

$$\sup_{\theta \in [\alpha_T; \beta_T]} \sup_{f \in W(\beta, L)} \mathbb{E}_{\theta, f} \|\hat{f}_T - f\|_2^2 = \left\{\sup_{f \in W(\beta, L)} R(q, f)\right\}(1 + o(1)),$$

we finally obtain the upper bound in Theorem 3.

Let us now focus on the lower bound part of this Theorem, namely

$$\liminf_{T \to +\infty} \inf_{f_T} \sup_{\theta \in [\alpha_T; \beta_T]} \sup_{f \in W(\beta, L)} T^{2\beta/(2\beta+1)} \mathbb{E}_{\theta, f} \|f_T - f\|_2^2 \geq C^*.$$

Obviously, the left-hand side is larger than the same term taken at a fixed θ, say $\theta = 1$. So, it is enough to establish

$$\liminf_{T \to +\infty} \inf_{f_T} \sup_{f \in W(\beta, L)} T^{2\beta/(2\beta+1)} \mathbb{E}_f \|f_T - f\|_2^2 \geq C^*,$$

in the model given by $dX_t = f(t)\, dt + dW_t$, $t \in [-T/2; T/2]$. But according to Girsanov's Theorem (see Appendix II in [14]),

$$\inf_{f_T} \sup_f \mathbb{E}_f \|f_T - f\|_2 = \inf_{\bar{c}_T} \sup_c \mathbb{E}_c \|\bar{c}_T - c\|_2,$$

where the first infimum concerns any estimator f_T based on the observation of $\{X_t\}_{|t| \leq T/2}$ satisfying $dX_t = f(t)\, dt + dW_t$, while the function f ranges over the Sobolev ball $W(\beta, L)$, whereas the second one concerns any estimator $\bar{c}_T$ based on the observation of the corresponding projection $y_k = c_k + T^{-1/2}\xi_{k,[\ ]}(1)$, for any $k \in \mathbb{Z}$, and the sequence of Fourier coefficients c ranges over the set

$$\left\{c = (c_k)_{k \in \mathbb{Z}}; \sum_{k \in \mathbb{Z}} |2\pi k|^{2\beta} |c_k|^2 \leq L\right\}.$$

Now, Pinsker's lower bound [20] establishes

$$\inf_{\bar{c}_T} \sup_c \mathbb{E}_c \|\bar{c}_T - c\|_2^2 \geq C^* T^{-2\beta/(2\beta+1)} (1 + o(1)),$$

which quite achieves the proof. The last thing to see is that Assumption **(F)** (which is necessary in our setup since the period is unknown) does not affect Pinsker's proof. But we may assume without restriction that the finite set of sequences c used in Pinsker's proof to obtain the lower bound satisfy $\sum_k |c_k| \leq M$ and for any integer $r \geq 2$ we choose c such that $\sum_{k \in \mathbb{Z}^*} |c_{rk}|^2 \leq (1-h) \sum_{k \in \mathbb{Z}^*} |c_k|^2$. $\square$

4.1.1. *Proof of Lemma* 3. This proof and the next one rely on the convergence property of the estimator $\hat{\theta}_T$ given by Assumption **(C)**. Let us denote by φ_T the rate

$$\varphi_T = \frac{\theta^2 \sqrt{\log T}}{T^{3/2}}.$$

Now, we write

$$(30) \quad \mathbb{E}\left[\left(\sum_{k \in \mathbb{Z}} \lambda_k^2 |\gamma_k(\theta, \hat{\theta}_T) - c_k|^2\right)^p\right]$$

$$= \mathbb{E}\left[\left(\sum_{k \in \mathbb{Z}} \lambda_k^2 \left|\sum_{p \in \mathbb{Z}} c_p \Phi\left[\frac{T}{\theta}\left(p - \frac{k\theta}{\hat{\theta}_T}\right)\right] - c_k\right|^2\right)^p \mathbf{1}_{\{|\hat{\theta}_T - \theta| \leq \varphi_T\}}\right]$$

$$+ \mathbb{E}\left[\left(\sum_{k \in \mathbb{Z}} \lambda_k^2 \left|\sum_{p \in \mathbb{Z}} c_p \Phi\left[\frac{T}{\theta}\left(p - \frac{k\theta}{\hat{\theta}_T}\right)\right] - c_k\right|^2\right)^p \mathbf{1}_{\{|\hat{\theta}_T - \theta| > \varphi_T\}}\right],$$

where

$$\Phi(t) = \int_{-1/2}^{1/2} e^{2i\pi tu}\, du = \frac{\sin(\pi t)}{\pi t}.$$

Note that we shall abundantly use the following properties of Φ.

(B1)   There exist absolute nonnegative constants M_1 and M_2 such that

$$|\Phi(u)| \leq \frac{M_1}{|u|} \quad \text{for} \quad |u| > 1/4, \qquad \|\Phi''\|_\infty \leq M_2.$$

By using assumption **(C)** on the estimator $\hat{\theta}_T$, we can deduce that the last term in the right-hand side of (30) satisfies for any integer q and large enough T:

$$(32) \quad \mathbb{E}\left[\left(\sum_{k \in \mathbb{Z}} \lambda_k^2 \left|\sum_{p \in \mathbb{Z}} c_p \Phi\left[\frac{T}{\theta}\left(p - \frac{k\theta}{\hat{\theta}_T}\right)\right] - c_k\right|^2\right)^p \mathbf{1}_{\{|\hat{\theta}_T - \theta| > \varphi_T\}}\right]$$

$$\leq \left(2 \sum_{k \in \mathbb{Z}} |c_k|\right)^{2p} \left(\sum_{k \in \mathbb{Z}} \lambda_k^2\right)^p P(|\hat{\theta}_T - \theta| > \varphi_T) \leq \frac{C}{T^q} \left(\sum_{k \in \mathbb{Z}} \lambda_k^2\right)^p.$$

Let us now turn to the first term:

$$(33) \quad \mathbb{E}\left[\left(\sum_{k \in \mathbb{Z}} \lambda_k^2 \left|\sum_{p \in \mathbb{Z}} c_p \Phi\left[\frac{T}{\theta}\left(p - \frac{k\theta}{\hat{\theta}_T}\right)\right] - c_k\right|^2\right)^p \mathbf{1}_{\{|\hat{\theta}_T - \theta| \leq \varphi_T\}}\right]$$

$$= \mathbb{E}\left[\left(\sum_{k \in \mathbb{Z}} \lambda_k^2 \left|c_k \left\{\Phi\left[\frac{kT}{\theta}\left(1 - \frac{\theta}{\hat{\theta}_T}\right)\right] - 1\right\}\right.\right.\right.$$

$$\left.\left.\left. + \sum_{p \neq k} c_p \Phi\left[\frac{T}{\theta}\left(p - \frac{k\theta}{\hat{\theta}_T}\right)\right]\right|^2\right)^p \mathbf{1}_{\{|\hat{\theta}_T - \theta| \leq \varphi_T\}}\right].$$

First, note that

$$\left|\frac{T}{\theta}\left(p - \frac{k\theta}{\hat{\theta}_T}\right)\right| = \left|\frac{T}{\theta}(p - k) + \frac{Tk}{\hat{\theta}_T\theta}(\hat{\theta}_T - \theta)\right|,$$

where

$$\left|\frac{Tk}{\hat{\theta}_T\theta}(\hat{\theta}_T - \theta)\right|\mathbf{1}_{\{|\hat{\theta}_T-\theta|\leq\varphi_T\}} \leq \frac{T|k|}{\theta^2}\frac{\theta^2\sqrt{\log T}}{T^{3/2}}(1 + o(1)) \leq \sqrt{\frac{\log T}{T}}|k|(1 + o(1)).$$

Since $\lambda_k = 0$ for $|k| > T^{1/4}$, by using that $\left|\frac{T}{\theta}(p - k)\right| \geq \frac{T}{\theta}$ when $p \neq k$, we get

$$\left|\frac{T}{\theta}\left(p - \frac{k\theta}{\hat{\theta}_T}\right)\right| \geq \frac{T}{\theta}(1 + o(1)) \geq \frac{T}{\beta_T}(1 + o(1)),$$

ensuring that for large enough T, we have (using property **(B1)**),

$$\left|\sum_{p\neq k}c_p\Phi\left[\frac{T}{\theta}\left(p - \frac{k\theta}{\hat{\theta}_T}\right)\right]\right|\mathbf{1}_{\{|\hat{\theta}_T-\theta|\leq\varphi_T\}} \leq C\frac{\theta}{T}.$$

Secondly, a Taylor expansion of Φ gives, for some ζ (depending on T and θ) in a neighborhood of zero:

$$\left|\Phi\left[\frac{kT}{\theta}\left(1 - \frac{\theta}{\hat{\theta}_T}\right)\right] - 1\right|\mathbf{1}_{\{|\hat{\theta}_T-\theta|\leq\varphi_T\}} = \frac{|k|^2T^2}{2\theta^2}\left|1 - \frac{\theta}{\hat{\theta}_T}\right|^2|\Phi''(\zeta)|\mathbf{1}_{\{|\hat{\theta}_T-\theta|\leq\varphi_T\}}$$

$$\leq C\frac{|k|^2\log T}{T}.$$

We can deduce from the previous inequalities an upper bound for (33):

$$\mathbb{E}\left[\left(\sum_{k\in\mathbb{Z}}\lambda_k^2\left|\sum_{p\in\mathbb{Z}}c_p\Phi\left[\frac{T}{\theta}\left(p - \frac{k\theta}{\hat{\theta}_T}\right)\right] - c_k\right|^2\right)^p\mathbf{1}_{\{|\hat{\theta}_T-\theta|\leq\varphi_T\}}\right]$$

$$\leq C\left(\sum_{k\in\mathbb{Z}}|k|^4\lambda_k^2|c_k|^2\right)^p\left(\frac{\log^2 T}{T^2}\right)^p + C\left(\sum_{k\in\mathbb{Z}}\lambda_k^2\right)^p\left(\frac{\beta_T}{T}\right)^{2p}.$$

This result together with (32) leads to the expected upper bound for the quantity at stake. $\square$

4.1.2. *Proof of Lemma* 4. We have

$$(34) \qquad \mathbb{E}\left[\left(\sum_{k\in\mathbb{Z}}\lambda_k^2|\xi_k(\hat{\theta}_T) - \xi_k(\theta)|^2\right)^p\right]$$

$$= \mathbb{E}\left[\left(\sum_{k\in\mathbb{Z}}\lambda_k^2|\xi_k(\hat{\theta}_T) - \xi_k(\theta)|^2\right)^p\mathbf{1}_{\{|\hat{\theta}_T-\theta|\leq\varphi_T\}}\right]$$

$$+ \mathbb{E}\left[\left(\sum_{k\in\mathbb{Z}}\lambda_k^2|\xi_k(\hat{\theta}_T) - \xi_k(\theta)|^2\right)^p\mathbf{1}_{\{|\hat{\theta}_T-\theta|>\varphi_T\}}\right].$$

Let us focus on the first term. Remember that $p = 1$ or 2, so that

$$\mathbb{E}\left[\left(\sum_{k\in\mathbb{Z}}\lambda_k^2|\xi_k(\hat{\theta}_T) - \xi_k(\theta)|^2\right)^p \mathbf{1}_{\{|\hat{\theta}_T - \theta|\leq\varphi_T\}}\right]$$

$$\leq \left(\sum_{k\in\mathbb{Z}}\lambda_k^2\right)^{p-1}\sum_{k\in\mathbb{Z}}\lambda_k^{2p}\mathbb{E}\left[|\xi_k(\hat{\theta}_T) - \xi_k(\theta)|^{2p}\,\mathbf{1}_{\{|\hat{\theta}_T - \theta|\leq\varphi_T\}}\right]$$

$$\leq \left(\sum_{k\in\mathbb{Z}}\lambda_k^2\right)^{p-1}\sum_{k\in\mathbb{Z}}\lambda_k^{2p}\mathbb{E}\left[\sup_{\tau\in\mathcal{V}_{\varphi_T}(\theta)}|\xi_k(\tau) - \xi_k(\theta)|^{2p}\right]$$

$$\leq \left(\sum_{k\in\mathbb{Z}}\lambda_k^2\right)^{p-1}\sum_{k\in\mathbb{Z}}\lambda_k^{2p}\mathbb{E}\left[\left(\sup_{\tau\in\mathcal{V}_{\varphi_T}(\theta)}|\xi_k(\tau) - \xi_k(\theta)|\right)^{2p}\right],$$

where

(35) $$\mathcal{V}_{\varphi_T}(\theta) = \left\{\tau \in [\alpha_T, \beta_T],\ |\tau - \theta| \leq \varphi_T\right\}.$$

To control this last expected value, we use the following equality, which holds for a positive random variable X and an integer r:

$$\mathbb{E}\left(X^{2r}\right) = 2r\int_0^{+\infty} P(X \geq y)\, y^{2r-1} dy.$$

This leads to

$$\mathbb{E}\left[\left(\sup_{\tau\in\mathcal{V}_{\varphi_T}(\theta)}|\xi_k(\tau) - \xi_k(\theta)|\right)^{2p}\right] = 2p\int_0^{+\infty} P\left(\sup_{\tau\in\mathcal{V}_{\varphi_T}(\theta)}|\xi_k(\tau) - \xi_k(\theta)| \geq y\right)y^{2p-1}dy.$$

We now use, without proving it, the following classical result (see, e.g., [11]).

Lemma 5 (Generalized Markov inequality). *Let $\mathcal{L}$ be a stochastic process and $A = [\alpha; \beta]$, then for all $\mu > 0$ and for all $R > 0$,*

$$\mathbb{P}\left(\sup_{\tau\in A}\mathcal{L}(\tau) > R\right) \leq \exp(-\mu R)\sup_{\tau\in A}\left(\sqrt{\mathbb{E}\left[e^{2\mu\mathcal{L}(\tau)}\right]}\right)\left[1 + \mu\int_{\tau\in A}\sqrt{\mathbb{E}\left[|\mathcal{L}'(\tau)|^2\right]}\,d\tau\right],$$

as soon as the quantities at stake are well defined.

We deduce that for $\mu, y > 0$

$$\mathbb{P}\left(\sup_{\tau\in\mathcal{V}_{\varphi_T}(\theta)}|\xi_k(\tau) - \xi_k(\theta)| \geq y\right)$$

$$\leq e^{-\mu y}\sup_{\tau\in V_{\varphi_T}(\theta)}\left\{\left[\mathbb{E}\left(e^{2\mu(\xi_k(\tau)-\xi_k(\theta))}\right)\right]^{1/2}\right\}\left(1 + \mu\int_{\tau\in\mathcal{V}_{\varphi_T}(\theta)}\left[\mathbb{E}\left(|\xi_k'(\tau)|^2\right)\right]^{1/2}d\tau\right)$$

$$\leq e^{-\mu y}\,\exp\left(Ck^2\mu^2\frac{\log T}{T}\right)\left(1 + \mu\int_{\tau\in\mathcal{V}_{\varphi_T}(\theta)}\frac{|k|\pi T}{\sqrt{3}\tau^2}\,d\tau\right)$$

$$\leq e^{-\mu y}\exp\left(C\mu^2 k^2\frac{\log T}{T}\right)\left(1 + C\mu|k|\sqrt{\frac{\log T}{T}}\right)$$

$$\leq \exp\left(-\frac{y^2 T}{4Ck^2\log T}\right)\left(1 + \frac{y\sqrt{T}}{2|k|\sqrt{\log T}}\right),\quad \text{by taking } \mu = \frac{yT}{2Ck^2\log T}.$$

This leads to the bound

$$(36) \qquad \mathbb{E}\left[\left(\sup_{\tau \in \mathcal{V}_{\varphi_T}(\theta)} |\xi_k(\tau) - \xi_k(\theta)|\right)^{2p}\right] \leq C\left(\frac{|k|\sqrt{\log T}}{\sqrt{T}}\right)^{2p}.$$

Thus,

$$\mathbb{E}\left[\left(\sum_{k\in\mathbb{Z}} \lambda_k^2 |\xi_k(\hat{\theta}_T) - \xi_k(\theta)|^2\right)^p \mathbf{1}_{\{|\hat{\theta}_T - \theta| \leq \varphi_T\}}\right]$$
$$\leq C\left(\sum_{k\in\mathbb{Z}} \lambda_k^2\right)^{p-1}\left(\sum_{k\in\mathbb{Z}} k^{2p}\lambda_k^{2p}\right)\left(\frac{\log T}{T}\right)^p.$$

Let us now turn to the second term of (34). By using the Cauchy–Schwarz inequality, we have

$$\mathbb{E}\left[\left(\sum_{k\in\mathbb{Z}} \lambda_k^2 |\xi_k(\hat{\theta}_T) - \xi_k(\theta)|^2\right)^p \mathbf{1}_{\{|\hat{\theta}_T - \theta| > \varphi_T\}}\right]$$
$$\leq \left(\mathbb{E}\left[\left(\sum_{k\in\mathbb{Z}} \lambda_k^2 |\xi_k(\hat{\theta}_T) - \xi_k(\theta)|^2\right)^{2p}\right]\right)^{1/2} \mathbb{P}(|\hat{\theta}_T - \theta| > \varphi_T)^{1/2}$$
$$\leq \left(\sum_{k\in\mathbb{Z}} \lambda_k^2\right)^{p-1/2}\left\{\sum_{k\in\mathbb{Z}} \lambda_k^{4p} \mathbb{E}\left[\left(\sup_{\tau \in [\alpha_T, \beta_T]} |\xi_k(\tau) - \xi_k(\theta)|\right)^{4p}\right]\right\}^{1/2}$$
$$\times \mathbb{P}(|\hat{\theta}_T - \theta| > \varphi_T)^{1/2}$$
$$\leq C(N_0)^{p-1/2}\left(\sum_{k\in\mathbb{Z}} |k|\lambda_k^{4p}\right)^{1/2}\left(\frac{T}{\alpha_T}\right)^{1/2} \mathbb{P}(|\hat{\theta}_T - \theta| > \varphi_T)^{1/2},$$

where the last inequality is obtained by the same method as the one applied in (36), except that $\mathcal{V}_{\varphi_T}(\theta)$ is replaced by $[\alpha_T, \beta_T]$ and N_0 is defined in Remark 2.

Thanks to the convergence property **(C)** of $\hat{\theta}_T$, the expectation

$$\mathbb{E}\left[\left(\sum_{k\in\mathbb{Z}} \lambda_k^2 |\xi_k(\hat{\theta}_T) - \xi_k(\theta)|^2\right)^p \mathbf{1}_{\{|\hat{\theta}_T - \theta| \leq \varphi_T\}}\right]$$

is then the main term of (34).

The expected result thus follows. $\square$

    4.2. PROOF OF THEOREM 4. We want to bound from above the quantity $\mathbb{E}\left[\sum_{|k|\leq N_{\max}} |\tilde{c}_k - c_k|^2\right]$, $\tilde{c}_k$ being defined by

$$\tilde{c}_k = \psi_k(z)z_k, \qquad \text{where} \quad \psi_k(z) = \left(1 - \frac{T_j}{T\|z\|_j^2}\right)_+ \quad \text{for} \quad k \in B_j,$$

with $\|z\|_j^2 = \sum_{l\in B_j} |z_l|^2$ and $x_+ = \max(0, x)$ for all $x \in \mathbb{R}$.

Indeed, the term $\mathbb{E}\sum_{|k|>N_{\max}}|c_k|^2 = O(T^{-1})$ is negligible compared to this main term

$$\mathbb{E}\Big[\sum_{|k|\leq N_{\max}}|\tilde{c}_k - c_k|^2\Big] = \sum_{j=1}^{J}\mathbb{E}\Big[\sum_{k\in B_j}|\psi_k(z)z_k - c_k|^2\Big].$$

From now on, we shall focus on

$$(37)\qquad \mathbb{E}\Big[\sum_{k\in B_j}|\psi_k(z)z_k - c_k|^2\Big]$$
$$= \mathbb{E}\Big[\sum_{k\in B_j}|\psi_k(y)y_k - c_k|^2\Big] + \mathbb{E}\Big[\sum_{k\in B_j}|\psi_k(z)z_k - \psi_k(y)y_k|^2\Big]$$
$$+ 2\operatorname{Re}\mathbb{E}\Big[\sum_{k\in B_j}\big(\psi_k(z)z_k - \psi_k(y)y_k\big)\overline{\big(\psi_k(y)y_k - c_k\big)}\Big],$$

where y_k was introduced in (8). We shall prove that the first term in the right-hand side is the main one. To do this, we just have to prove that the second term is negligible thanks to the Cauchy–Schwarz inequality.

Let us then focus on the second term,

$$\mathbb{E}\Big[\sum_{k\in B_j}|\psi_k(z)z_k - \psi_k(y)y_k|^2\Big]$$
$$\leq 2\mathbb{E}\Big[\sum_{k\in B_j}\big|\big(\psi_k(z)-\psi_k(y)\big)z_k\big|^2\Big] + 2\mathbb{E}\Big[\sum_{k\in B_j}|\psi_k(y)r_k|^2\Big]$$
$$\leq 2\mathbb{E}\Big[\sum_{k\in B_j}\big|\big(\psi_k(z)-\psi_k(y)\big)z_k\big|^2\Big] + 2\mathbb{E}\Big[\sum_{k\in B_j}|r_k|^2\Big],$$

since $\forall k,\ \psi_k(y)\in[0,1]$ and $r_k = z_k - y_k$.

Let us remark that for $k\in B_j$,

$$|\psi_k(z)-\psi_k(y)|^2 \leq |\psi_k(z)-\psi_k(y)| = \Big|\Big(1 - \frac{T_j}{T\|z\|_j^2}\Big)_+ - \Big(1 - \frac{T_j}{T\|y\|_j^2}\Big)_+\Big|$$
$$\leq \Big|\Big(1 - \frac{T_j}{T\|z\|_j^2}\Big) - \Big(1 - \frac{T_j}{T\|y\|_j^2}\Big)\Big| = \frac{T_j}{T}\Big|\frac{\|z\|_j^2 - \|y\|_j^2}{\|y\|_j^2\|z\|_j^2}\Big|$$
$$\leq \frac{T_j}{T}\Big|\frac{\|r\|_j^2 + 2\sum_{k\in B_j}\operatorname{Re}(r_k\overline{y_k})}{\|y\|_j^2\|z\|_j^2}\Big|.$$

Finally, the second term in the right-hand side of (37) is bounded from above as follows:

$$\mathbb{E}\Big[\sum_{k\in B_j}|\psi_k(z)z_k - \psi_k(y)y_k|^2\Big] \leq 2\mathbb{E}\Big[\frac{T_j}{T}\Big|\frac{\|r\|_j^2 + 2\sum_{k\in B_j}\operatorname{Re}(r_k\overline{y_k})}{\|y\|_j^2}\Big|\Big] + 2\mathbb{E}\big(\|r\|_j^2\big)$$
$$\leq 2\frac{T_j}{T}\mathbb{E}\Big(\frac{\|r\|_j^2}{\|y\|_j^2} + 2\frac{\|r\|_j}{\|y\|_j}\Big) + 2\mathbb{E}\big(\|r\|_j^2\big).$$

Now, using the Cauchy–Schwarz inequality and a sequence $\{\delta_T\}$ of positive numbers going to zero as $T \to \infty$, we have

$$(38) \qquad \mathbb{E}\Big[\sum_{k \in B_j} |\psi_k(z)z_k - \psi_k(y)y_k|^2 \Big]$$

$$\leq 2\frac{T_j}{T}\Big[2\mathbb{E}\left(\|r\|_j^2\right)^{1/2}\Big\{ \mathbb{E}\left(\frac{1}{\|y\|_j^2}\mathbf{1}_{\|y\|_j \leq \delta_T} \right)^{1/2} + O(\delta_T^{-1}) \Big\}$$

$$+ \mathbb{E}\left(\|r\|_j^4\right)^{1/2}\mathbb{E}\left(\frac{1}{\|y\|_j^4} \right)^{1/2}\Big] + 2\mathbb{E}\left(\|r\|_j^2\right).$$

To control the upper bound of (38), we shall use the following lemmas.

Lemma 6. *There exists a positive constant C such that*

$$\mathbb{E}\left(\frac{1}{\|y\|_j^2}\mathbf{1}_{\|y\|_j \leq \delta_T} \right) \leq C\frac{T^{T_j/2}\delta_T^{T_j-2}}{[T_j/2]!}.$$

Lemma 7. *There exists a positive constant C such that*

$$\mathbb{E}\left(\frac{1}{\|y\|_j^4} \right) \leq CT^2.$$

Lemma 8. *For $p = 1$ or 2, there exists a positive constant C such that*

$$\mathbb{E}\left(\|r\|_j^{2p}\right) \leq C\Big[T_j^{p-1}\Big(\sum_{k \in B_j} k^{2p} \Big)\frac{\log^p T}{T^{2p}} + \left(\frac{\log^2 T}{T^2} \right)^p + T_j^p\left(\frac{\beta_T}{T} \right)^{2p} \Big].$$

The proofs of these lemmas are postponed to the end of the proof of Theorem 4. Thus,

$$\mathbb{E}\Big[\sum_{k \in B_j} |\psi_k(z)z_k - \psi_k(y)y_k|^2 \Big]$$

$$\leq C\frac{T_j}{T}\left(\frac{T^{T_j/4}\delta_T^{T_j/2-1}}{\sqrt{[T_j/2]!}} + \delta_T^{-1} \right)\Big[\Big(\sum_{k \in B_j} k^2 \Big)^{1/2}\frac{\sqrt{\log T}}{T} + \frac{\log T}{T} + T_j^{1/2}\left(\frac{\beta_T}{T} \right) \Big]$$

$$+ CT_j\Big[T_j^{1/2}\Big(\sum_{k \in B_j} k^4 \Big)^{1/2}\frac{\log T}{T^2} + \frac{\log^2 T}{T^2} + T_j\left(\frac{\beta_T}{T} \right)^2 \Big]$$

$$+ C\Big[\Big(\sum_{k \in B_j} k^2 \Big)\frac{\log T}{T^2} + \left(\frac{\log^2 T}{T^2} \right) + T_j\left(\frac{\beta_T}{T} \right)^2 \Big].$$

We choose

$$\delta_T = \frac{1}{T^{1/2}}\{[T_j/2]!\}^{1/T_j}$$

in order to obtain

$$\frac{T^{T_j/4}\delta_T^{T_j/2-1}}{\sqrt{[T_j/2]!}} + \delta_T^{-1} = 2T^{1/2}\{[T_j/2]!\}^{-1/T_j}.$$

Using the bounds

$$\left(\sum_{k\in B_j} k^\alpha\right) = O(T_j^{\alpha+1}) \text{ for any integer } \alpha \geq 1,$$

$$\sum_j T_j^\alpha \leq N_{\max}^\alpha = O(T^{\alpha/4}) \text{ for any } \alpha \geq 1, \text{ and } J \leq C\log^2(T), \text{ by Lemma 1,}$$

we finally obtain:

$$\sum_j \mathbb{E}\left[\sum_{k\in B_j} |\psi_k(z)z_k - \psi_k(y)y_k|^2\right]$$
$$\leq C\frac{\log T}{T^{3/2}} \sum_j \exp\left\{-\frac{1}{T_j}\log\left(\frac{T_j}{2}\right)! + \frac{5}{2}\log T_j\right\} + C\frac{\log^2 T}{T}.$$

Now, Stirling's Inequality gives the bound $x! > (x/e)^x\sqrt{2\pi x}$ for large enough x, leading to

$$\sum_j \mathbb{E}\left[\sum_{k\in B_j} |\psi_k(z)z_k - \psi_k(y)y_k|^2\right]$$
$$\leq C\frac{\log T}{T^{3/2}} \sum_j \exp\left\{-\frac{1}{T_j}\log\left[\left(\frac{T_j}{2e}\right)^{T_j/2}(\pi T_j)^{1/2}\right] + \frac{5}{2}\log T_j\right\} + C\frac{\log^2 T}{T}$$
$$\leq C\frac{\log T}{T^{3/2}} \sum_j \exp\left\{-\frac{1}{2}\log T_j - \frac{1}{2T_j}\log T_j + \frac{5}{2}\log T_j\right\} + C\frac{\log^2 T}{T}$$
$$\leq C\frac{\log T}{T^{3/2}} \sum_j T_j^2 + C\frac{\log^2 T}{T} \leq C\frac{\log^2 T}{T}.$$

Hence

$$\sum_{j=1}^J \mathbb{E}\left[\sum_{k\in B_j} |\psi_k(z)z_k - c_k|^2\right] \leq \sum_{j=1}^J \mathbb{E}\left[\sum_{k\in B_j} |\psi_k(y)y_k - c_k|^2\right] + C\frac{\log^2(T)}{T}.$$

Now remark that if f_T^* is defined by (20), we have:

$$\mathbb{E}\|f_T^* - f\|^2 = \sum_{k\in\mathbb{Z}} |\psi_k(y)y_k - c_k|^2 = \sum_{j=1}^J \mathbb{E}\left[\sum_{k\in B_j} |\psi_k(y)y_k - c_k|^2\right] + O\left(\frac{1}{T}\right).$$

This together with Lemma 2 gives:

$$\mathbb{E}\|\tilde{f}_T - f\|^2 \leq (1 + 3\rho_T)\min_{\lambda\in\Lambda_{\mathrm{mon}}} R(\lambda, f) + \frac{C}{T}\log^2(T) + O\left(\frac{1}{T}\right).$$

Now using the fact that Pinsker's weights fulfill (12) and that they belong to Λ_{mon}, we deduce:

$$\lim_{T\to+\infty} \sup_{\theta\in[\alpha_T;\beta_T]} \sup_{f\in W(\beta,L)} T^{2\beta/(2\beta+1)} \mathbb{E}_{\theta,f} \|\tilde{f}_T - f\|_2^2 \le C^*,$$

which concludes the proof of the theorem using the minimax lower bound. $\square$

4.2.1. *Proof of Lemma* 6. To prove this lemma, we shall compute more generally $\mathbb{E}\left(\|y\|_{\mathbb{R}^d}^{-2} \mathbf{1}_{\|y\|_{\mathbb{R}^d} \le \delta_T}\right)$, where $\|\cdot\|_{\mathbb{R}^d}$ is the classical Euclidean norm in $\mathbb{R}^d$ and $d \ge 3$. For notational simplicity, $\|\cdot\|_{\mathbb{R}^d}$ will be replaced in the sequel by $\|\cdot\|$. Let us recall that $y = c + \xi/\sqrt{T}$, where ξ is a standard Gaussian random variable in $\mathbb{R}^d$. Denoting $\mathcal{B}(a;\rho)$ the $\mathbb{R}^d$-ball centered at a with radius ρ, we have

$$\mathbb{E}\left(\|y\|^{-2} \mathbf{1}_{\|y\|\le\delta_T}\right) = \frac{1}{(2\pi)^{d/2}} \int_{\mathcal{B}(-c;\delta_T\sqrt{T})} \frac{e^{-\|x\|^2/2}}{\|c+x/\sqrt{T}\|^2}\, dx \le \frac{T^{d/2}}{(2\pi)^{d/2}} \int_{\mathcal{B}(0;\delta_T)} \frac{dy}{\|y\|^2}$$
$$= \frac{T^{d/2}\delta_T^{d-2}}{(2\pi)^{d/2}} \operatorname{Vol}(\mathcal{B}_d(0;1)) \frac{d}{d-2} \le C\frac{T^{d/2}\delta_T^{d-2}}{[d/2]!},$$

where the last inequality follows from the classical expression for $\operatorname{Vol}(\mathcal{B}_d(0;1))$. This quantity is equal to $(2\pi^{d/2})/(d\Gamma(d/2))$, where Γ is the Gamma function. $\square$

4.2.2. *Proof of Lemma* 7. In the same way as in the preceding proof,

$$\mathbb{E}\left(\|y\|^{-4}\right) = \frac{T^2}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \frac{1}{\|\sqrt{T}c+x\|^4} \exp\left(-\frac{\|x\|^2}{2}\right) dx$$
$$= \frac{T^2}{(2\pi)^{d/2}} \left[\int_{x\in\mathcal{B}_d(-\sqrt{T}c,1)} \frac{1}{\|\sqrt{T}c+x\|^4} \exp\left(-\frac{\|x\|^2}{2}\right) dx \right.$$
$$\left. + \int_{x\notin\mathcal{B}_d(-\sqrt{T}c,1)} \frac{1}{\|\sqrt{T}c+x\|^4} \exp\left(-\frac{\|x\|^2}{2}\right) dx\right]$$
$$\le \frac{T^2}{(2\pi)^{d/2}} \left[\int_{x\in\mathcal{B}_d(0,1)} \frac{1}{\|x\|^4} dx + (2\pi)^{d/2}\right]$$
$$\le \frac{T^2}{(2\pi)^{d/2}} \left[\operatorname{Vol}(\mathcal{B}_d(0,1))\, d \int_0^1 \frac{1}{r^4} r^{d-1} dr + (2\pi)^{d/2}\right]$$
$$= T^2 + \frac{dT^2}{(2\pi)^{d/2}(d-4)} \operatorname{Vol}(\mathcal{B}_d(0,1)) \le CT^2,$$

since the volume $\operatorname{Vol}(\mathcal{B}_d(0,1))$ of the unit ball in $\mathbb{R}^d$ is uniformly bounded with respect to d. $\square$

4.2.3. *Proof of Lemma* 8. Recall that for $k \in B_j$, r_k can be rewritten as follows:

$$r_k = \left(\gamma_k(\theta,\hat{\theta}_T) - c_k\right) + \frac{1}{\sqrt{T}}\left(\xi_k(\hat{\theta}_T) - \xi_k(\theta)\right) + \frac{1}{\sqrt{T}}\left(\xi_k(\theta) - \xi_{k,[\,]}(\theta)\right).$$

Thus, for $p = 1$ or 2,

$$(39) \quad \mathbb{E}\left(\|r\|_j^{2p} \right) = \mathbb{E}\left[\left(\sum_{k \in B_j} |r_k|^2 \right)^p \right] \leq C \mathbb{E}\left[\left(\sum_{k \in B_j} |\gamma_k(\theta, \hat{\theta}_T) - c_k|^2 \right)^p \right]$$

$$+ \frac{C}{T^p} \mathbb{E}\left[\left(\sum_{k \in B_j} |\xi_k(\hat{\theta}_T) - \xi_k(\theta)|^2 \right)^p \right] + \frac{C}{T^p} \mathbb{E}\left[\left(\sum_{k \in B_j} |\xi_k(\theta) - \xi_{k,[\,]}(\theta)|^2 \right)^p \right].$$

Let us first control the last term of the upper bound (39) using that there are T_j elements in B_j:

$$\frac{1}{T^p} \mathbb{E}\left[\left(\sum_{k \in B_j} |\xi_k(\theta) - \xi_{k,[\,]}(\theta)|^2 \right)^p \right] \leq \frac{1}{T^p} \left(\sum_{k \in B_j} 1 \right)^{p-1} \mathbb{E}\left[\sum_{k \in B_j} |\xi_k(\theta) - \xi_{k,[\,]}(\theta)|^{2p} \right]$$

$$\leq \frac{T_j^{p-1}}{T^p} \mathbb{E}\left[\sum_{k \in B_j} |\xi_k(\theta) - \xi_{k,[\,]}(\theta)|^{2p} \right].$$

But,

$$\mathbb{E}\left[|\xi_k(\theta) - \xi_{k,[\,]}(\theta)|^{2p} \right]$$

$$= \mathbb{E}\left[\left| \frac{1}{\sqrt{T}} \int_{-T/2}^{T/2} e^{-2ik\pi t/\theta} \, dW_t - \frac{1}{\sqrt{\theta[T/\theta]}} \int_{-[T/\theta]\theta/2}^{[T/\theta]\theta/2} e^{-2ik\pi t/\theta} \, dW_t \right|^{2p} \right]$$

$$\leq \frac{C}{T^p} \mathbb{E}\left[\left| \int_{-T/2}^{T/2} e^{-2ik\pi t/\theta} \, dW_t - \int_{-[T/\theta]\theta/2}^{[T/\theta]\theta/2} e^{-2ik\pi t/\theta} \, dW_t \right|^{2p} \right]$$

$$+ C \left(\frac{1}{\sqrt{\theta[T/\theta]}} - \frac{1}{\sqrt{T}} \right)^{2p} \mathbb{E}\left[\left| \int_{-[T/\theta]\theta/2}^{[T/\theta]\theta/2} e^{-2ik\pi t/\theta} \, dW_t \right|^{2p} \right] \leq C \frac{\theta^p}{T^p} \leq C \frac{\beta_T^p}{T^p}.$$

Finally, the last term in (39) can be bounded as follows:

$$(40) \qquad \frac{1}{T^p} \mathbb{E}\left[\left(\sum_{k \in B_j} |\xi_k(\theta) - \xi_{k,[\,]}(\theta)|^2 \right)^p \right] \leq C \frac{T_j^p \beta_T^p}{T^{2p}}.$$

Let us now control the first term of the upper bound (39). For this, we shall use Lemma 3, which gives:

$$(41) \qquad \mathbb{E}\left[\left(\sum_{k \in B_j} |\gamma_k(\theta, \hat{\theta}_T) - c_k|^2 \right)^p \right]$$

$$\leq C \left[\left(\sum_{k \in B_j} k^4 |c_k|^2 \right)^p \left(\frac{\log^2 T}{T^2} \right)^p + \left(\sum_{k \in B_j} 1 \right)^p \left(\frac{\beta_T}{T} \right)^{2p} \right]$$

$$\leq C \left[\left(\frac{\log^2 T}{T^2} \right)^p + T_j^p \left(\frac{\beta_T}{T} \right)^{2p} \right].$$

Now we address the second term of (39) by using Lemma 4, which provides the following upper bound:

$$(42) \quad \frac{1}{T^p} \mathbb{E}\left[\left(\sum_{k \in B_j} |\xi_k(\hat{\theta}_T) - \xi_k(\theta)|^2\right)^p\right] \leq \frac{C}{T^p} \left(\sum_{k \in B_j} 1\right)^{p-1} \left(\sum_{k \in B_j} k^{2p}\right) \left(\frac{\log T}{T}\right)^p$$

$$\leq C T_j^{p-1} \left(\sum_{k \in B_j} k^{2p}\right) \frac{\log^p T}{T^{2p}}.$$

The inequalities (40), (41), and (42) lead to the expected result. $\square$

4.3. ADDITIONAL COMMENTS ON EQUATION (3). Recall that for any real-valued function G with finite total variation, the Stieltjes integral with respect to G is defined in the following way. There exists a decomposition $G = G_1 - G_2$, where each G_i is an increasing function. For any continuous function ψ and any subdivision $\pi = \{-T/2 = t_0 < t_1 < \ldots < t_m = T/2\}$ of the time interval $[-T/2; T/2]$, with $|\pi| = \max(t_{i+1} - t_i)$, we define

$$\int_{-T/2}^{T/2} \psi(t) G_i(dt) = \lim_{|\pi| \to 0} \sum_{j=0}^{m-1} \psi(t_j)\big(G_i(t_{j+1}) - G_i(t_j)\big), \quad i = 1, 2,$$

and

$$\int_{-T/2}^{T/2} \psi(t) G(dt) = \int_{-T/2}^{T/2} \psi(t) \, G_1(dt) - \int_{-T/2}^{T/2} \psi(t) \, G_2(dt)$$

(the first limit is independent of the choice of the subdivision π as is the resulting Stieltjes integral from the choice of the G_i's). We then denote

$$\int_{-T/2}^{T/2} e^{-2ik\pi t/\hat{\theta}_T} \, dX_t \triangleq u_{\hat{\theta}_T}(T/2) X_{T/2} - u_{\hat{\theta}_T}(-T/2) X_{-T/2} - \int_{-T/2}^{T/2} X_t \, u_{\hat{\theta}_T}(dt).$$

Note that

$$X_t = \int_0^t f(s/\theta) \, ds + W_t = F(t) + W_t,$$

and a simple integration by parts formula gives that

$$\int_{-T/2}^{T/2} e^{-2ik\pi t/\hat{\theta}_T} f(t/\theta) \, dt$$

$$= u_{\hat{\theta}_T}(T/2) F(T/2) - u_{\hat{\theta}_T}(-T/2) F(-T/2) - \int_{-T/2}^{T/2} F(t) \, u_{\hat{\theta}_T}(dt).$$

References

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

[2] C. Butucea and C. Matias, *Minimax estimation of the noise level and of the deconvolution density in a semiparametric convolution model*, Bernoulli, 11 (2005), 309–340.

[3] I. Castillo, *Semiparametric Second-Order Efficient Estimation of the Period of a Signal and Applications*, Preprint, 2005,
http://www.math.u-psud.fr/~castillo/prepub/castillo.pdf

[4] L. Cavalier and A. B. Tsybakov, *Penalized blockwise Stein's method, monotone oracles and sharp adaptive estimation*, Math. Methods Statist., 10 (2001), 247–282.

[5] A. S. Dalalyan and Yu. A. Kutoyants, *Asymptotically efficient trend coefficient estimation for ergodic diffusion*, Math. Methods Statist., 11 (2002), 402–427.

[6] I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Appl. Math., Vol. 61, SIAM, Philadelphia, PA, 1992.

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

[8] D. L. Donoho and I. M. Johnstone, *Neo-classical minimax problems, thresholding and adaptive function estimation*, Bernoulli, 2 (1996), 39–62.

[9] D. L. Donoho and M. G. Low, *Renormalization exponents and optimal pointwise rates of convergence*, Ann. Statist., 20 (1992), 944–970.

[10] E. Gassiat and C. Lévy-Leduc, *Efficient semi-parametric estimation of the periods in a superposition of periodic functions with unknown shape.* To appear in J. Time Series Anal., 2006.

[11] G. K. Golubev, *Estimating the period of a signal of unknown shape corrupted by white noise*, Problems Inform. Transmission, 24 (1988), 288–299.

[12] G. K. Golubev, *Local asymptotic normality in problems of nonparametric estimation of functions, and lower bounds for quadratic risks*, Theory Probab. Appl., 36 (1991), 152–157.

[13] M. Hoffmann, L_p *estimation of the diffusion coefficient*, Bernoulli, 5 (1999), 447–481.

[14] I. A. Ibragimov and R. Z. Has'minskii, *Statistical Estimation*: *Asymptotic Theory*, Springer-Verlag, 1981.

[15] I. A. Ibragimov and R. Z. Has'minskii, *Nonparametric estimation of the value of a linear functional in Gaussian white noise*, Theory Probab. Appl., 29 (1984), 18–32.

[16] Yu. A. Kutoyants, *Statistical Inference for Ergodic Diffusion Processes*, Springer Series in Statist., 2004.

[17] M. Lavielle and C. Lévy-Leduc, *Semiparametric estimation of the frequency of unknown periodic functions: Application to laser vibrometry signals*, IEEE Trans. Signal Processing, 53 (2005), 2306–2314.

[18] O. Loukianov and D. Loukianova, *Uniform deterministic equivalent for AF of recurrent diffusion and non-parametric drift estimation*, Preprint, 2006.

[19] A. Nemirovski, *Lectures on Probability Theory and Statistics*, Lecture Notes in Mathematics, Vol. 1738, Springer-Verlag, Berlin, 2000.

[20] M. S. Pinsker, *Optimal filtration of square-integrable signals in Gaussian noise*, Probl. Inform. Transmission, 16 (1980), 120-133.

[21] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, 3rd. ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Vol. 293, Springer-Verlag, Berlin, 1999.

[22] A. B. Tsybakov, *Introduction à l'estimation non-paramétrique* (*Introduction to Nonparametric Estimation*), Mathématiques & Applications (Paris), Vol. 41, Springer, Paris, 2004.

[23] A. W. van der Vaart, *Asymptotic Statistics*, Cambridge University Press, Cambridge, 1998.