



Robust adaptive efficient estimation for semi-Markov nonparametric regression models

Vlad Stefan Barbu¹ · Slim Beltaief¹ · Sergey Pergamenshchikov^{1,2,3}

Received: 22 May 2016 / Accepted: 14 June 2018 / Published online: 22 June 2018
© Springer Nature B.V. 2018

Abstract

We consider the nonparametric robust estimation problem for regression models in continuous time with semi-Markov noises. An adaptive model selection procedure is proposed. Under general moment conditions on the noise distribution a sharp non-asymptotic oracle inequality for the robust risks is obtained and the robust efficiency is shown. It turns out that for semi-Markov models the robust minimax convergence rate may be faster or slower than the classical one.

Keywords Non-asymptotic estimation · Robust risk · Model selection · Sharp oracle inequality · Asymptotic efficiency

Mathematics Subject Classification Primary 62G08 · Secondary 62G05

1 Introduction

Let us consider a regression model in continuous time

$$d y_t = S(t) d t + d \xi_t, \quad 0 \leq t \leq n, \quad (1.1)$$

This work was done under financial support of the RSF Grant Number 14-49-00079 (National Research University “MPEI” 14 Krasnokazarmennaya, 111250 Moscow, Russia) and by the project XterM - Feder, University of Rouen, France.

✉ Sergey Pergamenshchikov
Serge.Pergamenchchikov@univ-rouen.fr

Vlad Stefan Barbu
barbu@univ-rouen.fr

Slim Beltaief
slim.beltaief1@univ-rouen.fr

¹ Laboratoire de Mathématiques Raphaël Salem, UMR 6085 CNRS-Université de Rouen, Saint-Étienne-du-Rouvray, France

² International Laboratory SSP & QF, National Research Tomsk State University, Tomsk, Russia

³ National Research University “MPEI”, Moscow, Russia

where $S(\cdot)$ is an unknown 1-periodic function from $\mathbf{L}_2[0, 1]$ defined on \mathbb{R} with values in \mathbb{R} , the noise process $(\xi_t)_{t \geq 0}$ is defined as

$$\xi_t = \varrho_1 w_t + \varrho_2 L_t + \varrho_3 z_t, \tag{1.2}$$

where ϱ_1, ϱ_2 and ϱ_3 are unknown coefficients, $(w_t)_{t \geq 0}$ is a standard Brownian motion, $(L_t)_{t \geq 0}$ is a jump Lévy process (with $\mathbf{E}L_t^2 = t$ cf. Eq. (2.3)) and the pure jump process $(z_t)_{t \geq 1}$, defined in (2.5), is assumed to be a semi-Markov process (see, for example, Barbu and Limnios 2008).

The problem is to estimate the unknown function S in the model (1.1) on the basis of observations $(y_t)_{0 \leq t \leq n}$. Firstly, this problem was considered in the framework of the “signal+white noise” models (see, for example, Ibragimov and Khasminskii 1981 or Pinsker 1981). Later, in order to study dependent observations in continuous time, were introduced “signal+color noise” regressions based on Ornstein-Uhlenbeck processes (cf. Höpfner and Kutoyants 2009, 2010; Konev and Pergamenshchikov 2003, 2010).

Moreover, to include jumps in such models, the papers Konev and Pergamenshchikov (2012) and Konev and Pergamenshchikov (2015) used non Gaussian Ornstein-Uhlenbeck processes introduced in Barndorff-Nielsen and Shephard (2001) for modeling of the risky assets in the stochastic volatility financial markets. Unfortunately, the dependence of the stable Ornstein-Uhlenbeck type decreases with a geometric rate. So, asymptotically when the duration of observations goes to infinity, we obtain very quickly the same “signal+white noise” model.

The main goal of this paper is to consider continuous time regression models with dependent observations for which the dependence does not disappear for a sufficient large duration of observations. To this end we define the noise in the model (1.1) through a semi-Markov process which keeps the dependence for any duration n . This type of models allows, for example, to estimate the signals observed under long impulse noise impact with a memory or in the presence of “against signals”.

In this paper we use the robust estimation approach introduced in Konev and Pergamenshchikov (2012) for such problems. To this end, we denote by Q the distribution of $(\xi_t)_{0 \leq t \leq n}$ in the Skorokhod space $\mathcal{D}[0, n]$. We assume that Q is unknown and belongs to some distribution family \mathcal{Q}_n specified in Sect. 4. In this paper we use the quadratic risk

$$\mathcal{R}_Q(\tilde{S}_n, S) = \mathbf{E}_{Q,S} \|\tilde{S}_n - S\|^2, \tag{1.3}$$

where $\|f\|^2 = \int_0^1 f^2(s)ds$ and $\mathbf{E}_{Q,S}$ is the expectation with respect to the distribution $\mathbf{P}_{Q,S}$ of the process (1.1) corresponding to the noise distribution Q . Since the noise distribution Q is unknown, it seems reasonable to introduce the robust risk of the form

$$\mathcal{R}_n^*(\tilde{S}_n, S) = \sup_{Q \in \mathcal{Q}_n} \mathcal{R}_Q(\tilde{S}_n, S), \tag{1.4}$$

which enables us to take into account the information that $Q \in \mathcal{Q}_n$ and ensures the quality of an estimate \tilde{S}_n for all distributions in the family \mathcal{Q}_n .

To summarize, the goal of this paper is to develop robust efficient model selection methods for the model (1.1) with the semi-Markov noise having unknown distribution, based on the approach proposed by Konev and Pergamenshchikov (2012, 2015) for continuous time regression models with semimartingale noises. Unfortunately, we cannot use directly this method for semi-Markov regression models, since their tool essentially uses the fact that the Ornstein-Uhlenbeck dependence decreases with geometrical rate and the “white noise” case is obtained sufficiently quickly.

Thus in the present paper we propose new analytical tools based on renewal methods to obtain the sharp non-asymptotic oracle inequalities. As a consequence, we obtain the robust efficiency for the proposed model selection procedures in the adaptive setting.

The rest of the paper is organized as follows. We start by introducing the main conditions in the next section. Then, in Sect. 3 we construct the model selection procedure on the basis of the weighted least squares estimates. The main results are stated in Sect. 4; here we also specify the set of admissible weight sequences in the model selection procedure. In Sect. 5 we derive some renewal results useful for obtaining other results of the paper. In Sect. 6 we develop stochastic calculus for semi Markov processes. In Sect. 7 we study some properties of the model (1.1). A numerical example is presented in Sect. 8. Most of the results of the paper are proved in Sect. 9. In “Appendix” some auxiliary propositions are given.

2 Main conditions

In the model (1.2) we assume that the jump Lévy process L_t is defined as

$$L_t = \int_0^t \int_{\mathbb{R}_*} x(\mu(ds, dx) - \tilde{\mu}(ds, dx)), \tag{2.1}$$

where $\mu(ds, dx)$ is the jump measure with the deterministic compensator $\tilde{\mu}(ds, dx) = ds\Pi(dx)$, where $\Pi(\cdot)$ is the Levy measure on $\mathbb{R}_* = \mathbb{R} \setminus \{0\}$ (see, for example Jacod and Shiryaev 2002; Cont and Tankov 2004 for details) for which we assume that

$$\Pi(x^2) = 1 \quad \text{and} \quad \Pi(x^8) < \infty, \tag{2.2}$$

where we use the usual notation $\Pi(|x|^m) = \int_{\mathbb{R}_*} |z|^m \Pi(dz)$ for any $m > 0$. Note that, using the Ito formula for the martingales (see, for example, Liptser and Shiryaev 1986, p.185) we can obtain directly that

$$\mathbf{E}_Q L_t^2 = \mathbf{E} \sum_{0 \leq s \leq t} (\Delta L_s)^2 = \mathbf{E}_Q \int_0^t \int_{\mathbb{R}_*} x^2 \mu(ds, dx) = \Pi(x^2)t = t, \tag{2.3}$$

where $\Delta L_s = L_s - L_{s-}$ and L_{s-} is the left limit to s in probability. Moreover, the last condition in (2.2) and the inequality (A.1) imply that for some positive constant C^* the expectation

$$\mathbf{E} L_t^8 \leq C^* (1 + \Pi(x^8)) t < \infty. \tag{2.4}$$

Note that $\Pi(\mathbb{R}_*)$ may be equal to $+\infty$. Moreover, we assume that the pure jump process $(z_t)_{t \geq 0}$ in (1.2) is a semi-Markov process with the following form

$$z_t = \sum_{i=1}^{N_t} Y_i, \tag{2.5}$$

where $(Y_i)_{i \geq 1}$ is an i.i.d. sequence of random variables with

$$\mathbf{E}_Q Y_i = 0, \quad \mathbf{E}_Q Y_i^2 = 1 \quad \text{and} \quad \mathbf{E}_Q Y_i^4 < \infty.$$

Here N_t is a general counting process (see, for example, Mikosch 2004) defined as

$$N_t = \sum_{k=1}^{\infty} \mathbb{1}_{\{T_k \leq t\}} \quad \text{and} \quad T_k = \sum_{l=1}^k \tau_l, \tag{2.6}$$

where $(\tau_l)_{l \geq 1}$ is an i.i.d. sequence of positive integrated random variables with distribution η and mean $\check{\tau} = \mathbf{E}_Q \tau_1 > 0$. We assume that the processes $(N_t)_{t \geq 0}$ and $(Y_i)_{i \geq 1}$ are independent between them and are also independent of $(L_t)_{t \geq 0}$.

Note that the process $(z_t)_{t \geq 0}$ is a special case of a semi-Markov process (see, e.g., Barbu and Limnios 2008; Limnios and Oprisan 2001).

Remark 2.1 It should be noted that if τ_j are exponential random variables, then $(N_t)_{t \geq 0}$ is a Poisson process and, in this case, $(\xi_t)_{t \geq 0}$ is a Lévy process for which this model has been studied in Konev and Pergamenschikov (2009a, b) and Konev and Pergamenschikov (2012). But, in the general case when the process (2.5) is not a Lévy process, this process has a memory and cannot be treated in the framework of semi-martingales with independent increments. In this case, we need to develop new tools based on renewal theory arguments, what we do in Sect. 5. This tools will be intensively used in the proofs of the main results of this paper.

Note that for any function f from $\mathbf{L}_2[0, n]$, $f : [0, n] \rightarrow \mathbb{R}$, for the noise process $(\xi_t)_{t \geq 0}$ defined in (1.2), with $(z_t)_{t \geq 0}$ given in (2.5), the integral

$$I_n(f) = \int_0^n f(s) d\xi_s \tag{2.7}$$

is well defined with $\mathbf{E}_Q I_n(f) = 0$. Moreover, as it is shown in Lemma 6.2,

$$\mathbf{E}_Q I_n^2(f) \leq \varkappa_Q \|f\|_n^2 \quad \text{and} \quad \varkappa_Q = \bar{q} + q_3^2 |\rho|_*, \tag{2.8}$$

where $\|f\|_n^2 = \int_0^n f^2(s) ds$, $\bar{q} = q_1^2 + q_2^2$ and $|\rho|_* = \sup_{t \geq 0} |\rho(t)| < \infty$. Here ρ is the density of the renewal measure $\check{\eta}$ defined as

$$\check{\eta} = \sum_{l=1}^{\infty} \eta^{(l)}, \tag{2.9}$$

where $\eta^{(l)}$ is the l th convolution power for η .

Remark 2.2 In Proposition 5.2 we will prove that, under Conditions (\mathbf{H}_1) – (\mathbf{H}_4) , the the renewal measure $\check{\eta}$ has a density ρ .

To study the series (2.9) we assume that the measure η has a density g which satisfies the following conditions.

(\mathbf{H}_1) Assume that, for any $x \in \mathbb{R}$, there exist the finite limits

$$g(x-) = \lim_{z \rightarrow x-} g(z) \quad \text{and} \quad g(x+) = \lim_{z \rightarrow x+} g(z)$$

and, for any $K > 0$, there exists $\delta = \delta(K) > 0$ for which

$$\sup_{|x| \leq K} \int_0^\delta \frac{|g(x+t) + g(x-t) - g(x+) - g(x-)|}{t} dt < \infty.$$

(\mathbf{H}_2) For any $\gamma > 0$,

$$\sup_{z \geq 0} z^\gamma |2g(z) - g(z-) - g(z+)| < \infty.$$

(\mathbf{H}_3) There exists $\beta > 0$ such that $\int_{\mathbb{R}} e^{\beta x} g(x) dx < \infty$.

Remark 2.3 It should be noted that Condition (\mathbf{H}_3) means that there exists an exponential moment for the random variable $(\tau_j)_{j \geq 1}$, i.e. these random variables are not too large. This is a natural constraint since these random variables define the intervals between jumps, i.e., the frequency of the jumps. So, to study the influence of the jumps in the model (1.1) one needs to consider the noise process (1.2) with “small” interval between jumps or large jump frequency.

For the next condition we need to introduce the Fourier transform of any function f from $\mathbf{L}_1(\mathbb{R})$, $f : \mathbb{R} \rightarrow \mathbb{R}$, defined as

$$\widehat{f}(\theta) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{i\theta x} f(x) dx. \tag{2.10}$$

(\mathbf{H}_4) There exists $t^* > 0$ such that the function $\widehat{g}(\theta - it)$ belongs to $\mathbf{L}_1(\mathbb{R})$ for any $0 \leq t \leq t^*$.

It is clear that Conditions (\mathbf{H}_1) – (\mathbf{H}_4) hold true for any continuously differentiable function g , for example for the exponential density.

Now we define the family of the noise distributions for the model (1.1) which is used in the robust risk (1.4). In our case the distribution family \mathcal{Q}_n consists in all distributions on the Skorokhod space $\mathcal{D}[0, n]$ of the process (1.2) with the parameters satisfying the conditions (2.11) and (2.12). Note that any distribution Q from \mathcal{Q}_n is defined by the unknown parameters in (1.2) and (2.1). We assume that

$$\varsigma_* \leq \sigma_Q \leq \varsigma^*, \tag{2.11}$$

where $\sigma_Q = \varrho_1^2 + \varrho_2^2 + \varrho_3^2/\check{\tau}$, the unknown bounds $0 < \varsigma_* \leq \varsigma^*$ are functions of n , i.e. $\varsigma_* = \varsigma_*(n)$ and $\varsigma^* = \varsigma^*(n)$, such that for any $\check{\epsilon} > 0$,

$$\lim_{n \rightarrow \infty} n^{\check{\epsilon}} \varsigma_*(n) = +\infty \quad \text{and} \quad \lim_{n \rightarrow \infty} \frac{\varsigma^*(n)}{n^{\check{\epsilon}}} = 0. \tag{2.12}$$

Remark 2.4 As we will see later, the parameter σ_Q is the limit for the Fourier transform of the noise process (1.2). Such limit is called variance proxy (see Konev and Pergamenschikov 2012).

Remark 2.5 Note that, generally (but it is not necessary) the parameters ϱ_1 , ϱ_2 and ϱ_3 can be dependent on n . Condition (2.12) means that we consider all possible cases, i.e. these parameters may go to the infinity or be constant or to zero as well. See, for example, the conditions (3.32) in Konev and Pergamenschikov (2015).

3 Model selection

Let $(\phi_j)_{j \geq 1}$ be an orthonormal uniformly bounded basis in $\mathbf{L}_2[0, 1]$, i.e., for some constant $\phi_* \geq 1$, which may be depend on n ,

$$\sup_{0 \leq j \leq n} \sup_{0 \leq t \leq 1} |\phi_j(t)| \leq \phi_* < \infty. \tag{3.1}$$

We extend the functions $\phi_j(t)$ by periodicity, i.e., we set $\phi_j(t) := \phi_j(\{t\})$, where $\{t\}$ is the fractional part of $t \geq 0$. For example, we can take the trigonometric basis defined as $\text{Tr}_1 \equiv 1$ and, for $j \geq 2$,

$$\text{Tr}_j(x) = \sqrt{2} \begin{cases} \cos(2\pi [j/2]x) & \text{for even } j; \\ \sin(2\pi [j/2]x) & \text{for odd } j, \end{cases} \tag{3.2}$$

where $[x]$ denotes the integer part of x .

To estimate the function S we use here the model selection procedure for continuous time regression models from Konev and Pergamenshchikov (2012) based on the Fourier expansion. We recall that for any function S from $L_2[0, 1]$ we can write

$$S(t) = \sum_{j=1}^{\infty} \theta_j \phi_j(t) \quad \text{and} \quad \theta_j = (S, \phi_j) = \int_0^1 S(t)\phi_j(t)dt. \tag{3.3}$$

So, to estimate the function S it suffices to estimate the coefficients θ_j and to replace them in this representation by their estimators. Using the fact that the function S and ϕ_j are 1-periodic we can write that

$$\theta_j = \frac{1}{n} \int_0^n \phi_j(t) S(t)dt.$$

If we replace here the differential $S(t)dt$ by the stochastic observed differential dy_t we obtain the natural estimate for θ_j on the time interval $[0, n]$

$$\widehat{\theta}_{j,n} = \frac{1}{n} \int_0^n \phi_j(t)dy_t, \tag{3.4}$$

which can be represented, in view of the model (1.1), as

$$\widehat{\theta}_{j,n} = \theta_j + \frac{1}{\sqrt{n}}\xi_{j,n}, \quad \xi_{j,n} = \frac{1}{\sqrt{n}}I_n(\phi_j). \tag{3.5}$$

Now (see, for example, Ibragimov and Khasminskii 1981) we can estimate the function S by the projection estimators, i.e.

$$\widehat{S}_m(t) = \sum_{j=1}^m \widehat{\theta}_{j,n} \phi_j(t), \quad 0 \leq t \leq 1, \tag{3.6}$$

for some number $m \rightarrow \infty$ as $n \rightarrow \infty$. It should be noted that Pinsker in Pinsker (1981) shows that the projection estimators of the form (3.6) are not efficient. For obtaining efficient estimation one needs to use weighted least square estimators defined as

$$\widehat{S}_\lambda(t) = \sum_{j=1}^n \lambda(j)\widehat{\theta}_{j,n}\phi_j(t), \tag{3.7}$$

where the coefficients $\lambda = (\lambda(j))_{1 \leq j \leq n}$ belong to some finite set Λ from $[0, 1]^n$. As it is shown in Pinsker (1981), in order to obtain efficient estimators, the coefficients $\lambda(j)$ in (3.7) need to be chosen depending on the regularity of the unknown function S . In this paper we consider the adaptive case, i.e. we assume that the regularity of the function S is unknown. In this case we chose the weight coefficients on the basis of the model selection procedure proposed in Konev and Pergamenshchikov (2012) for the general semi-martingale regression model in continuous time. These coefficients will be obtained later in (3.19). To the end, first we set

$$\check{n} = \#(\Lambda) \quad \text{and} \quad |\Lambda|_* = 1 + \max_{\lambda \in \Lambda} \check{L}(\lambda), \tag{3.8}$$

where $\#(\Lambda)$ is the cardinal number of Λ and $\check{L}(\lambda) = \sum_{j=1}^n \lambda(j)$. Now, to choose a weight sequence λ in the set Λ we use the empirical quadratic risk, defined as

$$\text{Err}_n(\lambda) = \|\widehat{S}_\lambda - S\|^2,$$

which in our case is equal to

$$\text{Err}_n(\lambda) = \sum_{j=1}^n \lambda^2(j) \widehat{\theta}_{j,n}^2 - 2 \sum_{j=1}^n \lambda(j) \widehat{\theta}_{j,n} \theta_j + \sum_{j=1}^{\infty} \theta_j^2. \tag{3.9}$$

Since the Fourier coefficients $(\theta_j)_{j \geq 1}$ are unknown, we replace the terms $\widehat{\theta}_{j,n} \theta_j$ by

$$\widetilde{\theta}_{j,n} = \widehat{\theta}_{j,n}^2 - \frac{\widehat{\sigma}_n}{n}, \tag{3.10}$$

where $\widehat{\sigma}_n$ is an estimate for the variance proxy σ_Q defined in (2.11). If it is known, we take $\widehat{\sigma}_n = \sigma_Q$; otherwise, we can choose it, for example, as in Konev and Pergamenshchikov (2012), i.e.

$$\widehat{\sigma}_n = \sum_{j=[\sqrt{n}]+1}^n \widehat{\tau}_{j,n}^2, \tag{3.11}$$

where $\widehat{\tau}_{j,n}$ are the estimators for the Fourier coefficients with respect to the trigonometric basis (3.2), i.e.

$$\widehat{\tau}_{j,n} = \frac{1}{n} \int_0^n Tr_j(t) dy_t. \tag{3.12}$$

Finally, in order to choose the weights, we will minimize the following cost function

$$J_n(\lambda) = \sum_{j=1}^n \lambda^2(j) \widehat{\theta}_{j,n}^2 - 2 \sum_{j=1}^n \lambda(j) \widetilde{\theta}_{j,n} + \delta P_n(\lambda), \tag{3.13}$$

where $\delta > 0$ is some threshold which will be specified later and the penalty term is

$$P_n(\lambda) = \frac{\widehat{\sigma}_n |\lambda|^2}{n}. \tag{3.14}$$

We define the model selection procedure as

$$\widehat{S}_* = \widehat{S}_{\widehat{\lambda}} \quad \text{and} \quad \widehat{\lambda} = \operatorname{argmin}_{\lambda \in \Lambda} J_n(\lambda). \tag{3.15}$$

We recall that the set Λ is finite so $\widehat{\lambda}$ exists. In the case when $\widehat{\lambda}$ is not unique, we take one of them.

Let us now specify the weight coefficients $(\lambda(j))_{1 \leq j \leq n}$. Consider, for some fixed $0 < \varepsilon < 1$, a numerical grid of the form

$$\mathcal{A} = \{1, \dots, k^*\} \times \{\varepsilon, \dots, m\varepsilon\}, \tag{3.16}$$

where $m = [1/\varepsilon^2]$. We assume that both parameters $k^* \geq 1$ and ε are functions of n , i.e. $k^* = k^*(n)$ and $\varepsilon = \varepsilon(n)$, such that

$$\begin{cases} \lim_{n \rightarrow \infty} k^*(n) = +\infty, & \lim_{n \rightarrow \infty} \frac{k^*(n)}{\ln n} = 0, \\ \lim_{n \rightarrow \infty} \varepsilon(n) = 0 & \text{and} \quad \lim_{n \rightarrow \infty} n^{\delta} \varepsilon(n) = +\infty \end{cases} \tag{3.17}$$

for any $\delta > 0$. One can take, for example, for $n \geq 2$

$$\varepsilon(n) = \frac{1}{\ln n} \quad \text{and} \quad k^*(n) = k_0^* + \sqrt{\ln n}, \tag{3.18}$$

where $k_0^* \geq 0$ is some fixed constant. For each $\alpha = (\beta, \mathbf{I}) \in \mathcal{A}$, we introduce the weight sequence

$$\lambda_\alpha = (\lambda_\alpha(j))_{1 \leq j \leq n}$$

with the elements

$$\lambda_\alpha(j) = \mathbf{1}_{\{1 \leq j < j_*\}} + (1 - (j/\omega_\alpha)^\beta) \mathbf{1}_{\{j_* \leq j \leq \omega_\alpha\}}, \tag{3.19}$$

where $j_* = 1 + \lceil \ln v_n \rceil$, $\omega_\alpha = (d_\beta \mathbf{I} v_n)^{1/(2\beta+1)}$,

$$d_\beta = \frac{(\beta + 1)(2\beta + 1)}{\pi^{2\beta} \beta} \quad \text{and} \quad v_n = n/\zeta^*.$$

and the threshold $\zeta^*(n)$ is introduced in (2.11). Now we define the set Λ as

$$\Lambda = \{\lambda_\alpha, \alpha \in \mathcal{A}\}. \tag{3.20}$$

It will be noted that in this case the cardinal of the set Λ is

$$\check{l}_n = k^* m. \tag{3.21}$$

Moreover, taking into account that $d_\beta < 1$ for $\beta \geq 1$ we obtain for the set (3.20)

$$|\Lambda|_* \leq 1 + \sup_{\alpha \in \mathcal{A}} \omega_\alpha \leq 1 + (v_n/\varepsilon)^{1/3}. \tag{3.22}$$

Remark 3.1 Note that the form (3.19) for the weight coefficients in (3.7) was proposed by Pinsker in Pinsker (1981) for the efficient estimation in the nonadaptive case, i.e. when the regularity parameters of the function S are known. In the adaptive case these weight coefficients are used in Konev and Pergamenshchikov (2012, 2015) to show the asymptotic efficiency for model selection procedures.

4 Main results

In this section we obtain in Theorem 4.3 the non-asymptotic oracle inequality for the quadratic risk (1.3) for the model selection procedure (3.15) and in Theorem 4.4 the non-asymptotic oracle inequality for the robust risk (1.4) for the same model selection procedure (3.15), considered with the coefficients (3.19). We give the lower and upper bound for the robust risk in Theorems 4.5 and 4.7, and also the optimal convergence rate in Corollary 4.8.

Before stating the non-asymptotic oracle inequality, let us first introduce the following parameters which will be used for describing the rest term in the oracle inequalities. For the renewal density ρ defined in (2.9) we set

$$\Upsilon(x) = \rho(x) - \frac{1}{\check{\tau}} \quad \text{and} \quad |\Upsilon|_1 = \int_0^{+\infty} |\Upsilon(x)| dx, \tag{4.1}$$

where $\check{\tau} = \mathbf{E}_Q \tau_1$. In Proposition 5.2 we show that $|\rho|_* = \sup_{t \geq 0} |\rho(t)| < \infty$ and $|\Upsilon|_1 < \infty$. So, using this, we can introduce the following parameters

$$\Psi_Q = 4\kappa_Q \check{l}_n + 5\sigma_Q \check{\tau} \phi_{max}^2 |\Upsilon|_1 + \frac{4\check{l}}{\sigma_Q} \phi_{max}^4 (1 + \sigma_Q^2)^2 \check{\mathbf{I}} \tag{4.2}$$

and

$$\mathbf{c}_Q^* = \kappa_Q + \sigma_Q (1 + \check{\tau} \phi_{max}^2 |\Upsilon|_1) + \phi_{max}^2 (1 + \sigma_Q^2) \sqrt{\check{\mathbf{I}}}, \tag{4.3}$$

where $\check{\mathbf{I}} = 5(1 + \check{\tau})^2(1 + |\rho|_*^2) (2 + |\Upsilon|_1 + \mathbf{E}Y_1^4 + \Pi(x^4))$. We recall that \check{l}_n is cardinal of Λ , the noise variance σ_Q is defined in (2.11) and the parameter \varkappa_Q is given in (2.8). First, let us state the non-asymptotic oracle inequality for the quadratic risk for the model selection procedure (3.15) introduced in (1.3) by $\mathcal{R}_Q(\tilde{S}_n, S) = \mathbf{E}_{Q,S} \|\tilde{S}_n - S\|^2$.

Theorem 4.1 *Assume that Conditions (H₁)–(H₄) hold. Then, for any $n \geq 1$ and $0 < \delta < 1/6$, the estimator of S given in (3.15) satisfies the following oracle inequality*

$$\mathcal{R}_Q(\hat{S}_*, S) \leq \frac{1 + 3\delta}{1 - 3\delta} \min_{\lambda \in \Lambda} \mathcal{R}_Q(\hat{S}_\lambda, S) + \frac{\Psi_Q + 12|\Lambda|_* \mathbf{E}_S |\hat{\sigma}_n - \sigma_Q|}{n\delta}. \tag{4.4}$$

Now we study the estimate (3.11).

Proposition 4.2 *Assume that Conditions (H₁)–(H₄) hold and that the function $S(\cdot)$ is continuously differentiable. Then, for any $n \geq 2$,*

$$\mathbf{E}_{Q,S} |\hat{\sigma}_n - \sigma_Q| \leq \frac{5\|\dot{S}\|^2 + \mathbf{c}_Q^*}{\sqrt{n}}, \tag{4.5}$$

where \dot{S} is the differential of S .

Theorem 4.1 and Proposition 4.2 implies the following result.

Theorem 4.3 *Assume that Conditions (H₁)–(H₄) hold and that the function S is continuously differentiable. Then, for any $n \geq 1$ and $0 < \delta \leq 1/6$, the procedure (3.15), (3.11) satisfies the following oracle inequality*

$$\mathcal{R}_Q(\hat{S}_*, S) \leq \frac{1 + 3\delta}{1 - 3\delta} \min_{\lambda \in \Lambda} \mathcal{R}_Q(\hat{S}_\lambda, S) + \frac{60\tilde{\Lambda}_n \|\dot{S}\|^2 + \tilde{\Psi}_{Q,n}}{n\delta}, \tag{4.6}$$

where $\tilde{\Psi}_{Q,n} = 12\tilde{\Lambda}_n \mathbf{c}_Q^* + \Psi_Q$ and $\tilde{\Lambda}_n = |\Lambda|_*/\sqrt{n}$.

Remark 4.1 Note that the coefficient \varkappa_Q can be estimated as $\varkappa_Q \leq (1 + \check{\tau}|\rho|_*)\sigma_Q$. Therefore, taking into account that $\phi_{max}^4 \geq 1$, the remainder term in (4.6) can be estimated as

$$\tilde{\Psi}_{Q,n} \leq \mathbf{C}_* \left(1 + \sigma_Q^4 + \frac{1}{\sigma_Q} \right) (1 + \tilde{\Lambda}_n) \check{l}_n \phi_{max}^4, \tag{4.7}$$

where $\mathbf{C}_* > 0$ is some constant which is independent of the distribution Q .

Furthermore, let us study the robust risk (1.4) for the procedure (3.15). In this case, the distribution family \mathcal{Q}_n consists in all distributions on the Skorokhod space $\mathcal{D}[0, n]$ of the process (1.2) with the parameters satisfying the conditions (2.11) and (2.12).

Moreover, we assume also that the upper bound for the basis functions in (3.1) may depend on $n \geq 1$, i.e. $\phi_* = \phi_*(n)$, such that for any $\check{\epsilon} > 0$

$$\lim_{n \rightarrow \infty} \frac{\phi_*(n)}{n^{\check{\epsilon}}} = 0. \tag{4.8}$$

The next result presents the non-asymptotic oracle inequality for the robust risk (1.4) for the model selection procedure (3.15), considered with the coefficients (3.19).

Theorem 4.4 Assume that Conditions (\mathbf{H}_1) – (\mathbf{H}_4) hold and that the unknown function S is continuously differentiable. Then, for the robust risk defined by $\mathcal{R}_n^*(\widehat{S}_n, S) = \sup_{Q \in \mathcal{Q}_n} \mathcal{R}_Q(\widehat{S}_n, S)$, through the distribution family (2.11–2.12), the procedure (3.15) with the coefficients (3.19) for any $n \geq 1$ and $0 < \delta < 1/6$, satisfies the following oracle inequality

$$\mathcal{R}_n^*(\widehat{S}_*, S) \leq \frac{1 + 3\delta}{1 - 3\delta} \min_{\lambda \in \Lambda} \mathcal{R}_n^*(\widehat{S}_\lambda, S) + \frac{\mathbf{U}_n^*(S)}{n\delta}, \tag{4.9}$$

where the sequence $\mathbf{U}_n^*(S) > 0$ is such that, under the conditions (2.12), (3.17) and (4.8), for any $r > 0$ and $\check{\delta} > 0$,

$$\lim_{n \rightarrow \infty} \sup_{\|\check{S}\| \leq r} \frac{\mathbf{U}_n^*(S)}{n^{\check{\delta}}} = 0. \tag{4.10}$$

Now we study the asymptotic efficiency for the procedure (3.15) with the coefficients (3.19), with respect to the robust risk (1.4) defined by the distribution family (2.11–2.12). To this end, we assume that the unknown function S in the model (1.1) belongs to the Sobolev ball

$$W_r^k = \left\{ f \in \mathcal{C}_{per}^k[0, 1] : \sum_{j=0}^k \|f^{(j)}\|^2 \leq \mathbf{r} \right\}, \tag{4.11}$$

where $\mathbf{r} > 0$ and $k \geq 1$ are some unknown parameters, $\mathcal{C}_{per}^k[0, 1]$ is the set of k times continuously differentiable functions $f : [0, 1] \rightarrow \mathbb{R}$ such that $f^{(i)}(0) = f^{(i)}(1)$ for all $0 \leq i \leq k$. The function class W_r^k can be written as an ellipsoid in $L_2[0, 1]$, i.e.,

$$W_r^k = \left\{ f \in \mathcal{C}_{per}^k[0, 1] : \sum_{j=1}^{\infty} a_j \theta_j^2 \leq \mathbf{r} \right\}, \tag{4.12}$$

where $a_j = \sum_{i=0}^k (2\pi[j/2])^{2i}$ and $\theta_j = \int_0^1 f(v) \text{Tr}_j(v) dv$. We recall that the trigonometric basis $(\text{Tr}_j)_{j \geq 1}$ is defined in (3.2).

Similarly to Konev and Pergamenschikov (2012, 2015) we will show here that the asymptotic sharp lower bound for the robust risk (1.4) is given by

$$\mathbf{r}_k^* = ((2k + 1)\mathbf{r})^{1/(2k+1)} \left(\frac{k}{(k + 1)\pi} \right)^{2k/(2k+1)}. \tag{4.13}$$

Note that this is the well-known Pinsker constant obtained for the nonadaptive filtration problem in “signal + small white noise” model (see, for example, Pinsker 1981). Let Π_n be the set of all estimators \widehat{S}_n measurable with respect to the σ -field $\sigma\{y_t, 0 \leq t \leq n\}$ generated by the process (1.1).

The following two results give the lower and upper bound for the robust risk (1.4) defined for the distribution family (2.11–2.12).

Theorem 4.5 Under Conditions (2.11) and (2.12),

$$\liminf_{n \rightarrow \infty} v_n^{2k/(2k+1)} \inf_{\widehat{S}_n \in \Pi_n} \sup_{S \in W_r^k} \mathcal{R}_n^*(\widehat{S}_n, S) \geq \mathbf{r}_k^*, \tag{4.14}$$

where $v_n = n/\zeta^*$.

Note that if the parameters \mathbf{r} and k are known, i.e. for the non adaptive estimation case, then to obtain the efficient estimation for the “signal+white noise” model Pinsker (1981) proposed to use the estimate \widehat{S}_{λ_0} defined in (3.7) with the weights

$$\lambda_0 = \mathbf{1}_{\{1 \leq j < j_*\}} + (1 - (j/\omega_{\alpha_0})^\beta) \mathbf{1}_{\{j_* \leq j \leq \omega_{\alpha_0}\}}, \tag{4.15}$$

where $\alpha_0 = (k, \mathbf{l}_0)$ and $\mathbf{l}_0 = \lceil \mathbf{r}/\varepsilon \rceil$. For the model (1.1–1.2) we show the same result.

Proposition 4.6 *The estimator \widehat{S}_{λ_0} satisfies the following asymptotic upper bound*

$$\lim_{n \rightarrow \infty} \nu_n^{2k/(2k+1)} \sup_{S \in W_r^k} \mathcal{R}_n^*(\widehat{S}_{\lambda_0}, S) \leq \mathbf{r}_k^*. \tag{4.16}$$

Remark 4.2 Note that the inequalities (4.14) and (4.16) imply that the estimator \widehat{S}_{λ_0} is efficient. But we can’t use the weights (4.15) directly because the parameters k and \mathbf{r} are unknown. By this reason to obtain the efficient estimate in the adaptive setting we use the model selection procedure (3.15) over the estimate family (3.20) which includes the estimator (4.15). Then, using the oracle inequality (4.9) and the upper bound (4.16) we can obtain the efficient property for this model selection procedure.

For the adaptive estimation we use the model selection procedure (3.15) with the parameter δ defined as a function of n satisfying

$$\lim_n \delta_n = 0 \quad \text{and} \quad \lim_n n^{\delta} \delta_n = 0 \tag{4.17}$$

for any $\delta > 0$. For example, we can take $\delta_n = (6 + \ln n)^{-1}$.

Let \widehat{S}_* be the procedure (3.15) based on the trigonometric basis (3.2) with the coefficients (3.19) and the parameter $\delta = \delta_n$ satisfying (4.17).

Theorem 4.7 *Assume that Conditions (H₁)–(H₄) hold true. Then*

$$\limsup_{n \rightarrow \infty} \nu_n^{2k/(2k+1)} \sup_{S \in W_r^k} \mathcal{R}_n^*(\widehat{S}_*, S) \leq \mathbf{r}_k^*. \tag{4.18}$$

Theorems 4.5 and 4.7 allow us to compute the optimal convergence rate.

Corollary 4.8 *Under the assumptions of Theorem 4.7 the procedure \widehat{S}_* is efficient, i.e.*

$$\lim_{n \rightarrow \infty} \nu_n^{2k/(2k+1)} \inf_{\widehat{S}_n \in \Pi_n} \sup_{S \in W_r^k} \mathcal{R}_n^*(\widehat{S}_n, S) = \mathbf{r}_k^* \tag{4.19}$$

and

$$\lim_{n \rightarrow \infty} \frac{\inf_{\widehat{S}_n \in \Pi_n} \sup_{S \in W_r^k} \mathcal{R}_n^*(\widehat{S}_n, S)}{\sup_{S \in W_r^k} \mathcal{R}_n^*(\widehat{S}_*, S)} = 1.$$

Remark 4.3 It is well known that the optimal (minimax) risk convergence rate for the Sobolev ball W_r^k is $n^{2k/(2k+1)}$ (see, for example, Pinsker 1981; Nussbaum 1985). We see here that the efficient robust rate is $\nu_n^{2k/(2k+1)}$, i.e., if the distribution upper bound $\zeta^* \rightarrow 0$ as $n \rightarrow \infty$, we obtain a faster rate with respect to $n^{2k/(2k+1)}$, and, if $\zeta^* \rightarrow \infty$ as $n \rightarrow \infty$, we obtain a slower rate. In the case when ζ^* is constant, than the robust rate is the same as the classical non robust convergence rate. The same properties for the robust risks are obtained in Konev and Pergamenschikov (2010) and Konev and Pergamenschikov (2012) for the regression

model with Ornstein-Uhlenbeck noise process. So, this is typical situation, when we take the supremum over all noise distribution in (1.4). It is natural that we don't obtain the same convergence rate as for the usual risks (1.3) and the difference is given by the coefficient ζ^* which satisfies the “slowly changing” properties (2.12).

5 Renewal density

This section is concerned with results related to the renewal measure $\check{\eta} = \sum_{l=1}^\infty \eta^{(l)}$. We start with the following lemma.

Lemma 5.1 *Let τ be a positive random variable with a density g , such that $\mathbf{E}e^{\beta\tau} < \infty$ for some $\beta > 0$. Then there exists a constant $\beta_1, 0 < \beta_1 < \beta$ for which,*

$$\mathbf{E}e^{(\beta_1+i\omega)\tau} \neq 1 \quad \forall \omega \in \mathbb{R}.$$

Proof We will show this lemma by contradiction, i.e. assume there exists some sequence of positive numbers going to zero $(\gamma_k)_{k \geq 1}$ and a sequence $(w_k)_{k \geq 1}$ such that

$$\mathbf{E}e^{(\gamma_k+i\omega_k)\tau} = 1 \tag{5.1}$$

for any $k \geq 1$. Firstly assume that $\limsup_{k \rightarrow \infty} w_k = +\infty$, i.e. there exists $(l_k)_{k \geq 1}$ for which $\lim_{k \rightarrow \infty} w_{l_k} = +\infty$. Note that in this case, for any $N \geq 1$,

$$\begin{aligned} \left| \int_0^N e^{\gamma_k t} \cos(w_{l_k} t) g(t) dt \right| &\leq \left| \int_0^N \cos(w_{l_k} t) g(t) dt \right| \\ &\quad + \left| \int_0^N (e^{\gamma_k t} - 1) \cos(w_{l_k} t) g(t) dt \right|, \end{aligned}$$

i.e., in view of Lemma A.5, for any fixed $N \geq 1$

$$\lim_{k \rightarrow \infty} \int_0^N e^{\gamma_k t} \cos(w_{l_k} t) g(t) dt = 0.$$

Since for some $\beta > 0$ the integral $\int_0^{+\infty} e^{\beta t} g(t) dt < \infty$, we get

$$\lim_{k \rightarrow \infty} \int_0^{+\infty} e^{\gamma_k t} \cos(w_{l_k} t) g(t) dt = 0.$$

Let now $\limsup_{k \rightarrow \infty} w_k = \omega_\infty$ and $|\omega_\infty| < \infty$. In this case there exists a sequence $(l_k)_{k \geq 1}$ such that $\lim_{k \rightarrow \infty} w_{l_k} = \omega_\infty$, i.e.

$$1 = \lim_{k \rightarrow \infty} \mathbf{E}e^{\gamma_k \tau} \cos(\tau w_{l_k}) = \mathbf{E}_Q \cos(\tau \omega_\infty).$$

It is clear that, for random variables having density, the last equality is possible if and only if $\omega_\infty = 0$. In this case, i.e. when $\lim_{k \rightarrow \infty} w_{l_k} = 0$, the Eq. (5.1) implies

$$\lim_{k \rightarrow \infty} \mathbf{E}_Q e^{\gamma_k \tau} \frac{\sin(\tau w_{l_k})}{w_{l_k}} = \mathbf{E} \tau = 0.$$

But, $\mathbf{E} \tau > 0$, under our conditions. These contradictions imply the desired result. □

Proposition 5.2 *Let τ be a positive random variable with the distribution η having a density g which satisfies Conditions (\mathbf{H}_1) – (\mathbf{H}_4) . Then the renewal measure (2.9) is absolutely continuous with density ρ , for which*

$$\rho(x) = \frac{1}{\check{\tau}} + \Upsilon(x), \tag{5.2}$$

where $\check{\tau} = \mathbf{E}\tau_1$ and $\Upsilon(\cdot)$ is some function defined on \mathbb{R}_+ with values in \mathbb{R} such that

$$\sup_{x \geq 0} x^\gamma |\Upsilon(x)| < \infty \text{ for all } \gamma > 0.$$

Proof First note, that we can represent the renewal measure $\check{\eta}$ as $\check{\eta} = \eta * \eta_0$ and $\eta_0 = \sum_{j=0}^\infty \eta^{(j)}$. It is clear that in this case the density ρ of $\check{\eta}$ can be written as

$$\rho(x) = \int_0^x g(x-y) \sum_{n \geq 0} g^{(n)}(y) dy. \tag{5.3}$$

Now we use the arguments proposed in the proof of Lemma 9.5 from Goldie (1991). For any $0 < \epsilon < 1$ we set

$$\rho_\epsilon(x) = \int_0^x g(x-y) \left(\sum_{n \geq 0} (1-\epsilon)^n g^{(n)}(y) - \frac{(1-\epsilon)}{\check{\tau}} g_0(y) \right) dy - g(x), \tag{5.4}$$

where $g_0(y) = e^{-\epsilon y/\check{\tau}} 1_{\{y>0\}}$. It is easy to deduce that for any $x \in \mathbb{R}$

$$\lim_{\epsilon \rightarrow 0} \rho_\epsilon(x) = \rho(x) - \frac{1}{\check{\tau}} \int_0^x g(z) dz - g(x). \tag{5.5}$$

Moreover, in view of Condition (\mathbf{H}_1) we obtain that the function $\rho_\epsilon(x)$ satisfies Condition \mathbf{D}) from Section A.3. So, through Proposition A.6 we get

$$\rho_\epsilon(x+) + \rho_\epsilon(x-) = \frac{1}{\pi} \int_{\mathbb{R}} e^{-ix\theta} \widehat{\rho}_\epsilon(\theta) d\theta,$$

where $\widehat{\rho}_\epsilon(\theta) = \int_{\mathbb{R}} e^{i\theta x} \rho_\epsilon(x) dx$. Note that by the Bunyakovskii–Cauchy–Schwarz inequality

$$|\widehat{g}(\theta)| = \left| \int_{\mathbb{R}} e^{i\theta x} g(x) dx \right| \leq \int_{\mathbb{R}} g(x) dx = 1. \tag{5.6}$$

It should be noted that this inequality becomes an equality if and only if $\theta = 0$. Therefore, for any $0 < \epsilon < 1$ the module $|(1-\epsilon)\widehat{g}(\theta)| < 1$ and

$$\sum_{n=0}^\infty (1-\epsilon)^n (\widehat{g}(\theta))^n = \frac{1}{1 - (1-\epsilon)\widehat{g}(\theta)}.$$

So, taking into account that

$$\widehat{g}_0(\theta) = \int_{\mathbb{R}} e^{i\theta x} g_0(x) dx = \frac{\check{\tau}}{\epsilon - i\check{\tau}\theta},$$

we obtain

$$\widehat{\rho}_\epsilon(\theta) = \widehat{g}(\theta) \sum_{n=0}^\infty (1-\epsilon)^n (\widehat{g}(\theta))^n - \left(\frac{1-\epsilon}{\check{\tau}} \right) \widehat{g}(\theta) \widehat{g}_0(\theta) - \widehat{g}(\theta) = \widehat{g}(\theta) G_\epsilon(\theta)$$

where

$$G_\epsilon(\theta) = \frac{1}{1 - (1 - \epsilon)\widehat{g}(\theta)} - \frac{1 - i\check{\tau}\theta}{\epsilon - i\check{\tau}\theta},$$

i.e.

$$\rho_\epsilon(x-) + \rho_\epsilon(x+) = \frac{1}{\pi} \int_{\mathbb{R}} e^{-ix\theta} \widehat{g}(\theta) G_\epsilon(\theta) d\theta. \tag{5.7}$$

In section A.5 we show that

$$\sup_{0 < \epsilon < 1, \theta \in \mathbb{R}} |G_\epsilon(\theta)| < \infty. \tag{5.8}$$

Therefore, using Condition (\mathbf{H}_4) and Lebesgue’s dominated convergence theorem, we can pass to limit as $\epsilon \rightarrow 0$ in (5.7), i.e., we obtain that

$$\rho(x+) + \rho(x-) - \frac{2}{\check{\tau}} \int_0^x g(z) dz - g(x+) - g(x-) = \frac{1}{\pi} \int_{\mathbb{R}} e^{-ix\theta} \widehat{g}(\theta) G_0(\theta) d\theta,$$

where

$$G_0(\theta) = \frac{1}{1 - \widehat{g}(\theta)} + \frac{1 - i\check{\tau}\theta}{i\check{\tau}\theta}.$$

Using here again Proposition A.6 we deduce that

$$\rho(x+) + \rho(x-) = \frac{2}{\check{\tau}} \int_0^x g(z) dz + \frac{1}{\pi} \int_{\mathbb{R}} e^{-ix\theta} \widehat{g}(\theta) \check{G}(\theta) d\theta \tag{5.9}$$

and

$$\check{G}(\theta) = \frac{1}{1 - \widehat{g}(\theta)} + \frac{1}{i\check{\tau}\theta}.$$

Note now that we can represent the density (5.3) as

$$\rho(x) = g * \sum_{n \geq 0} g^{(n)} = \sum_{n \geq 1} g^{(n)}(x) = g(x) + \sum_{n \geq 2} g^{(n)}(x) =: g(x) + \rho_c(x)$$

and the function $\rho_c(x)$ is continuous for all $x \in \mathbb{R}$. This means that

$$\widetilde{\rho}(x) = \frac{\rho(x+) + \rho(x-)}{2} - \rho(x) = \frac{g(x+) + g(x-)}{2} - g(x)$$

and, therefore, Condition (\mathbf{H}_2) implies that, for any $\gamma > 0$,

$$\sup_{x \geq 0} x^\gamma |\widetilde{\rho}(x)| < \infty.$$

Now we can rewrite (5.9) as

$$\rho(x) - \frac{1}{\check{\tau}} = \frac{1}{\check{\tau}} \int_x^{+\infty} g(z) dz + \frac{1}{2\pi} \int_{\mathbb{R}} e^{-ix\theta} \widehat{g}(\theta) \check{G}(\theta) d\theta - \widetilde{\rho}(x). \tag{5.10}$$

Taking into account that $\mathbf{E}_Q e^{\beta\tau} < \infty$ for some $\beta > 0$ we can obtain that

$$\sup_{x \geq 0} x^\gamma \int_x^{+\infty} g(z) dz < \infty.$$

To study the second term in (5.10) we will apply Proposition A.4 to the function $\widehat{g}(\theta)\check{G}(\theta)$. First, note that the function $\widehat{g}(\theta) = \mathbf{E}_Q e^{i\tau\theta}$ is holomorphic for any $\theta \in \mathbb{C}$ with $\text{Im}(\theta) > -\beta$, due to Condition (\mathbf{H}_3) ; in view of Lemma A.8, there exists $0 < \beta_* < \beta$ for which the function $\check{G}(\theta)$ is holomorphic. Second, Condition (\mathbf{H}_4) applied to the function $\widehat{g}(\theta)\check{G}(\theta)$ implies the first condition in Eq. (A.2). The second condition of (A.2) follows directly from Lemma A.5.

Therefore, the conditions of Proposition A.4 hold with $\beta_2 = +\infty$. Thus Proposition A.4 implies that for some $0 < \beta_0 < \beta_*$

$$\int_{\mathbb{R}} e^{-ix\theta} \widehat{g}(\theta) \check{G}(\theta) d\theta = e^{-\beta_0 x} \int_{\mathbb{R}} e^{-ix\theta} \widehat{g}(\theta - i\beta_0) \check{G}(\theta - i\beta_0) d\theta .$$

Taking into account here Condition (\mathbf{H}_4) and the bound for \check{G} given in (A.8), we obtain

$$\sup_{x \geq 0} e^{\beta_0 x} \left| \int_{\mathbb{R}} e^{-ix\theta} \widehat{g}(\theta) \check{G}(\theta) d\theta \right| < \infty .$$

Hence Proposition 5.2. □

Using this proposition we can study the renewal process $(N_t)_{t \geq 0}$ introduced in (2.6).

Corollary 5.3 *Assume that Conditions (\mathbf{H}_1) – (\mathbf{H}_4) hold true. Then, for any $t > 0$,*

$$\mathbf{E} N_t \leq |\rho|_* t, \quad \mathbf{E} N_t^2 \leq |\rho|_* t + |\rho|_*^2 t^2 \tag{5.11}$$

and, moreover, $\mathbf{E} N_t^m < \infty$ for any $m \geq 3$.

Proof First, by means of Proposition 5.2, note that we get

$$\mathbf{E} N_t = \mathbf{E} \sum_{k \geq 1} \mathbf{1}_{\{T_k \leq t\}} = \int_0^t \rho(v) dv \leq |\rho|_* t .$$

To estimate the second moment of N_t note that,

$$\begin{aligned} \mathbf{E} N_t^2 &= \mathbf{E} \sum_{k \geq 1} \mathbf{1}_{\{T_k \leq t\}} + 2\mathbf{E} \sum_{k \geq 1} \mathbf{1}_{\{T_k \leq t\}} \sum_{j \geq k+1} \mathbf{1}_{\{T_j \leq t\}} \\ &= \mathbf{E} N_t + 2\mathbf{E} \sum_{k \geq 1} \mathbf{1}_{\{T_k \leq t\}} \mathbf{E} \left(\sum_{j \geq k+1} \mathbf{1}_{\{T_j \leq t\}} / T_k \right) \\ &= \mathbf{E} N_t + 2\mathbf{E} \sum_{k \geq 1} \mathbf{1}_{\{T_k \leq t\}} \Theta(t - T_k) = \mathbf{E} N_t + 2 \int_0^t \Theta(t - v) \rho(v) dv , \end{aligned}$$

i.e. we obtain that

$$\mathbf{E} N_t^2 \leq |\rho|_* t + 2|\rho|_* \int_0^t \Theta(t - v) dv ,$$

where $\Theta(v) = \mathbf{E} \sum_{j \geq k+1} \mathbf{1}_{\{\sum_{i=k+1}^j \tau_i \leq v\}}$. Taking into account that $(\tau_k)_{k \geq 1}$ is i.i.d. sequence, this term can be represented as

$$\Theta(v) = \mathbf{E} \sum_{m \geq 1} \mathbf{1}_{\{T_m \leq v\}} = \mathbf{E} N_v \leq |\rho|_* v . \tag{5.12}$$

This implies the second inequality in (5.11). Similarly, for $m \geq 3$ we obtain that for some constant $C_m > 0$

$$\begin{aligned} \mathbf{E}N_t^m &\leq C_m \mathbf{E} \sum_{k_1 < \dots < k_m} \prod_{j=1}^m \mathbf{1}_{\{T_{k_j} \leq t\}} \\ &= C_m \mathbf{E} \sum_{k_1 < \dots < k_{m-1}} \prod_{j=1}^{m-1} \mathbf{1}_{\{T_{k_j} \leq t\}} \mathbf{E} \left(\sum_{l=k_{m-1}+1}^m \mathbf{1}_{\{T_l \leq t\}} \mid T_{k_{m-1}} \right) \\ &= C_m \mathbf{E} \sum_{k_1 < \dots < k_{m-1}} \prod_{j=1}^{m-1} \mathbf{1}_{\{T_{k_j} \leq t\}} \Theta(t - T_{k_{m-1}}) \leq C_m |\rho|_* t \mathbf{E}N_t^{m-1}. \end{aligned}$$

Therefore, by induction we obtain that $\mathbf{E}N_t^m < \infty$ for any $m \geq 3$. Hence Corollary 5.3. \square

6 Stochastic calculus for semi-Markov processes

In this section we give some results of stochastic calculus for the process $(\xi_t)_{t \geq 0}$ given in (1.2), needed all along this paper. As this process is the combination of a Lévy process and a semi-Markov process, these results are not standard and need to be provided.

Lemma 6.1 *Let f and g be any non-random functions from $\mathbf{L}_2[0, n]$ and $(I_t(f))_{t \geq 0}$ be the process defined in (2.7). Then, for any $0 \leq t \leq n$,*

$$\mathbf{E}_Q I_t(f) I_t(g) = \bar{\varrho}(f, g)_t + \varrho_3^2(f, g\rho)_t, \tag{6.1}$$

where $(f, g)_t = \int_0^t f(s)g(s)ds$ and ρ is the density of the renewal measure $\check{\eta} = \sum_{l=1}^\infty \eta^{(l)}$.

Proof First, note that the noise process (1.2) is square integrated martingale which can be represented as

$$\xi_t = \xi_t^c + \xi_t^d, \tag{6.2}$$

where $\xi_t^c = \varrho_1 w_t$ and $\xi_t^d = \varrho_2 L_t + \varrho_3 z_t$. Note that the process $(\xi_t^c)_{t \geq 0}$ is the square integrated continuous martingale with the quadratic characteristic $\langle \xi^c \rangle_t = \varrho_1^2 t$. Therefore, the quadratic variance $[\xi]_t$ is the following:

$$[\xi]_t = \langle \xi^c \rangle_t + \sum_{0 \leq s \leq t} (\Delta \xi_s^d)^2 = \varrho_1^2 t + \sum_{0 \leq s \leq t} (\Delta \xi_s^d)^2,$$

where $\Delta \xi_s = \xi_s - \xi_{s-}$ (see, for example, Liptser and Shiryaev 1986). Recalling that the processes $(L_t)_{t \geq 0}$ and $(z_t)_{t \geq 0}$ are independent, we obtain that $\Delta L_t \Delta z_t = 0$ for any $t > 0$, i.e.

$$[\xi]_t = \varrho_1^2 t + \varrho_2^2 \sum_{0 \leq s \leq t} (\Delta L_s)^2 + \varrho_3^2 \sum_{0 \leq s \leq t} (\Delta z_s)^2. \tag{6.3}$$

Moreover, note that we can represent the stochastic integral $I_t(f)$ as

$$I_t(f) = \varrho_1 I_t^w(f) + \varrho_2 I_t^L(f) + \varrho_3 I_t^z(f), \tag{6.4}$$

where the stochastic integrals $I_t^w(f) = \int_0^t f(s)dw_s$, $I_t^L(f) = \int_0^t f(s)dL_s$ and $I_t^z(f) = \int_0^t f(s)dz_s$ are independent square integrated martingales. Therefore,

$$\mathbf{E}_Q I_t(f)I_t(g) = \varrho_1^2 \mathbf{E} I_t^w(f)I_t^w(g) + \varrho_2^2 \mathbf{E} I_t^L(f)I_t^L(g) + \varrho_3^2 \mathbf{E} I_t^z(f)I_t^z(g).$$

Taking into account that $\mathbf{E} I_t^w(f)I_t^w(g) = (f, g)_t$ and that the expectation of the product of square integrated martingales equals to the expectation of their mutual covariance, i.e. $\mathbf{E} I_t^L(f)I_t^L(g) = \mathbf{E} [I^L(f), I^L(g)]_t$ and $\mathbf{E} I_t^z(f)I_t^z(g) = \mathbf{E} [I^z(f), I^z(g)]_t$, we obtain that

$$\mathbf{E}_Q I_t(f)I_t(g) = \varrho_1^2 (f, g)_t + \varrho_2^2 \mathbf{E} [I^L(f), I^L(g)]_t + \varrho_3^2 \mathbf{E} [I^z(f), I^z(g)]_t.$$

In view of (6.3) the mutual covariances may be calculated as

$$\left[I^L(f), I^L(g) \right]_t = \sum_{0 \leq s \leq t} f(s)g(s) (\Delta L_s)^2$$

and

$$[I^z(f), I^z(g)]_t = \sum_{0 \leq s \leq t} f(s)g(s) (\Delta z_s)^2 = \sum_{l=1}^{\infty} f(T_l)g(T_l)Y_l^2 1_{\{T_l \leq t\}}.$$

Taking into account that $\Pi(x^2) = 1$ and that the sequences $(Y_k)_{k \geq 1}$ and $(T_k)_{k \geq 1}$ are independent, we find

$$\mathbf{E} \left[I^L(f) I^L(g) \right]_t = \Pi(x^2) (f, g)_t = (f, g)_t \tag{6.5}$$

and

$$\mathbf{E} [I^z(f) I^z(g)]_t = \mathbf{E} \sum_{l=1}^{\infty} f(T_l)g(T_l)1_{\{T_l \leq t\}} = \int_0^t f(s)g(s)\rho(s)ds = (f, g\rho)_t.$$

Hence the conclusion follows. □

Corollary 6.2 *Assume that Conditions (H₁)–(H₄) hold true. Then, for any $n \geq 1$ and for any non random function f from $\mathbf{L}_2[0, n]$, the stochastic integral (2.7) exists and satisfies the inequality (2.8).*

Proof This lemma follows directly from Lemma 6.1 with $f = g$ and Proposition 5.2 which ensures that $\sup_{t \geq 0} \rho(t) < \infty$. □

Lemma 6.3 *Let f and g be bounded functions defined on $[0, \infty) \times \mathbb{R}$. Then, for any $k \geq 1$,*

$$\mathbf{E}_Q \left(I_{T_k^-}(f) I_{T_k^-}(g) \mid \mathcal{G} \right) = \bar{\varrho}(f, g)_{T_k} + \varrho_3^2 \sum_{l=1}^{k-1} f(T_l) g(T_l),$$

where \mathcal{G} is the σ -field generated by the sequence $(T_l)_{l \geq 1}$, i.e., $\mathcal{G} = \sigma\{T_l, l \geq 1\}$.

Proof Using (6.5) and, taking into account that the process $(L_t)_{t \geq 0}$ is independent of \mathcal{G} , we obtain

$$\mathbf{E}_Q \left(I_{T_k^-}(f) I_{T_k^-}(g) \mid \mathcal{G} \right) = \bar{\varrho}(f, g)_{T_k} + \varrho_3^2 \mathbf{E} \left(I_{T_k^-}^z(f) I_{T_k^-}^z(g) \mid \mathcal{G} \right).$$

Moreover,

$$\begin{aligned} \mathbf{E} \left(I_{T_k^-}^z(f) I_{T_k^-}^z(g) \mid \mathcal{G} \right) &= \mathbf{E} \left(\left(\sum_{l=1}^{k-1} f(T_l) Y_l \right) \left(\sum_{l=1}^{k-1} g(T_l) Y_l \right) \mid \mathcal{G} \right) \\ &= \sum_{l=1}^{k-1} f(T_l) g(T_l). \end{aligned}$$

Then we obtain the desired result. □

Lemma 6.4 *Assume that Conditions (\mathbf{H}_1) – (\mathbf{H}_4) hold true. Then, for any measurable bounded non-random functions f and g , we have*

$$\left| \mathbf{E}_Q \int_0^n I_{t^-}^2(f) g(t) dm_t \right| \leq 2\varrho_3^2 |g|_* |f|_*^2 |\Upsilon|_1 n,$$

where $m_t = \sum_{0 \leq s \leq t} (\Delta z_s)^2 - \int_0^t \rho(s) ds$ and the norm $|\Upsilon|_1$ is given in (4.1).

Proof Using the definition of the process $(m_t)_{t \geq 0}$ we can represent this integral as

$$\begin{aligned} \int_0^n I_{t^-}^2(f) g(t) dm_t &= \sum_{k \geq 1} I_{T_k^-}^2(f) g(T_k) Y_k^2 \mathbf{1}_{\{T_k \leq n\}} \\ &\quad - \int_0^n I_t^2(f) g(t) \rho(t) dt =: V_n - U_n. \end{aligned} \tag{6.6}$$

Note now that

$$\mathbf{E}_Q V_n = \mathbf{E}_Q \sum_{k \geq 1} g(T_k) \mathbf{E} \left(I_{T_k^-}^2(f) \mid \mathcal{G} \right) \mathbf{1}_{\{T_k \leq n\}}.$$

Now, using Lemma 6.3 we can represent the last expectation as

$$\mathbf{E}_Q V_n = \bar{\varrho} \mathbf{E} V'_n + \varrho_3^2 \mathbf{E} V''_n, \tag{6.7}$$

where

$$V'_n = \sum_{k \geq 1} g(T_k) \|f\|_{T_k}^2 \mathbf{1}_{\{T_k \leq n\}} \quad \text{and} \quad V''_n = \sum_{k \geq 2} g(T_k) \mathbf{1}_{\{T_k \leq n\}} \sum_{l=1}^{k-1} f^2(T_l).$$

The term $\mathbf{E} V'_n$ can be represented as

$$\mathbf{E} V'_n = \int_0^n g(t) \|f\|_t^2 \rho(t) dt.$$

We recall that the norm $\|\cdot\|_t$ is defined in (2.8). To estimate $\mathbf{E} V_n''$, note that in view of Fubini’s theorem

$$\begin{aligned} \mathbf{E} V_n'' &= \mathbf{E} \sum_{l \geq 1} f^2(T_l) \sum_{k \geq l+1} g(T_k) \mathbf{1}_{\{T_k \leq n\}} \mathbf{1}_{\{T_l \leq n\}} \\ &= \mathbf{E} \sum_{l \geq 1} f^2(T_l) \mathbf{E}_Q \left(\sum_{k \geq l+1} g(T_k) \mathbf{1}_{\{T_k \leq n\}} | T_l \right) \mathbf{1}_{\{T_l \leq n\}} \\ &= \sum_{l \geq 1} \mathbf{E} f^2(T_l) \bar{g}(T_l) \mathbf{1}_{\{T_l \leq n\}} = \int_0^n f^2(v) \bar{g}(v) \rho(v) dv, \end{aligned}$$

where, similarly to (5.12), the function $\bar{g}(\cdot)$ can be represented as

$$\begin{aligned} \bar{g}(v) &= \mathbf{E} \sum_{k \geq l+1} g \left(v + \sum_{j=l+1}^k \tau_j \right) \mathbf{1}_{\{\sum_{j=l+1}^k \tau_j \leq n-v\}} \\ &= \mathbf{E} \sum_{m \geq 1} g(v + T_m) \mathbf{1}_{\{T_m \leq n-v\}} = \int_0^{n-v} g(v+s) \rho(s) ds = \int_v^n g(t) \rho(t-v) dt. \end{aligned}$$

Moreover, using Lemma 6.1, we calculate the expectation of the last term in (6.6), i.e.

$$\mathbf{E}_Q U_n = \bar{\varrho} \int_0^n \|f\|_t^2 g(t) \rho(t) dt + \varrho_3^2 \int_0^n \|f\sqrt{\rho}\|_t^2 g(t) \rho(t) dt.$$

This implies that

$$|\mathbf{E}_Q \int_0^n I_{t-}^2(f) g(t) dm_t| = \varrho_3^2 \left| \int_0^n g(t) \delta(t) dt \right| \leq \varrho_3^2 |g|_* \int_0^n |\delta(t)| dt,$$

where $\delta(t) = \int_0^t f^2(v) (\rho(t-v) - \rho(t)) \rho(v) dv$. Note here that, in view of Proposition 5.2 the function $\delta(t)$ can be estimated for any $0 \leq t \leq n$ as

$$|\delta(t)| \leq |f|_*^2 |\rho|_* \int_0^t |\Upsilon(t-v) - \Upsilon(t)| dv \leq |f|_*^2 |\rho|_* (|\Upsilon|_1 + n|\Upsilon(t)|),$$

with $\Upsilon(x) = \rho(x) - 1/\check{\tau}$. So,

$$\int_0^n |\delta(t)| dt \leq n |f|_*^2 |\rho|_* \left(|\Upsilon|_1 + \int_0^n |\Upsilon(t)| dt \right) \leq 2n |f|_*^2 |\rho|_* |\Upsilon|_1,$$

and, therefore,

$$\left| \mathbf{E}_Q \int_0^n I_{t-}^2(f) g(t) dm_t \right| \leq 2\varrho_3^2 |g|_* |f|_*^2 |\Upsilon|_1 n$$

and this finishes the proof. □

Lemma 6.5 *Assume that Conditions (\mathbf{H}_1) – (\mathbf{H}_4) hold true. Then, for any measurable bounded non-random functions f and g , one has*

$$\mathbf{E}_Q \int_0^n I_{t-}^2(f) I_{t-}(g) g(t) d\xi_t = 0. \tag{6.8}$$

Proof Putting $\check{h}_t = I_{t-}^2(f)I_{t-}(g)g(t)$, we can represent the integral in (6.8) as

$$I_n(\check{h}) = \sqrt{\check{\varrho}} I_n^{\check{L}}(\check{h}) + \varrho_3 I_n^z(\check{h}). \tag{6.9}$$

where $\check{L} = \check{\varrho}_1 w_t + \check{\varrho}_2 L_t$, $\check{\varrho}_1 = \varrho_1/\sqrt{\check{\varrho}}$ and $\check{\varrho}_2 = \varrho_2/\sqrt{\check{\varrho}}$. First, we will show that

$$\mathbf{E}_Q I_n^{\check{L}}(\check{h}) = 0. \tag{6.10}$$

Using the notations (6.4), we set

$$J_1 = \int_0^n I_{t-}^2(f)I_{t-}^{\check{L}}(g)g(t)d\check{L}_t \quad \text{and} \quad J_2 = \int_0^n I_{t-}^2(f)I_{t-}^z(g)g(t)d\check{L}_t,$$

we obtain that

$$\int_0^n I_{t-}^2(f)I_{t-}(g)g(t)d\check{L}_t = \sqrt{\check{\varrho}} J_1 + \varrho_3 J_2. \tag{6.11}$$

Taking into account that, for any non random square integrated function g the integral $(\int_0^t g(s)dw_s)$ is Gaussian with the parameters $(0, \int_0^t g^2(s)ds)$, we obtain

$$\sup_{0 \leq t \leq n} \mathbf{E} (I_t^w(g))^8 < \infty. \tag{6.12}$$

By applying inequality (A.1) for the non-random function $h(s, x) = g(s)x$, and, recalling that $\Pi(x^8) < \infty$, we obtain,

$$\sup_{0 \leq t \leq n} \mathbf{E} (I_t^L(g))^8 < \infty.$$

Therefore, we obtain that

$$\sup_{0 \leq t \leq n} \mathbf{E}_Q (I_t^{\check{L}}(g))^8 < \infty.$$

Finally, by using the Cauchy inequality, we can estimate for any $0 < t \leq n$ the following expectation as

$$\mathbf{E}_Q (I_t^{\check{L}}(f))^4 (I_t^{\check{L}}(g))^2 < \sqrt{\mathbf{E}_Q (I_t^{\check{L}}(f))^8} \sqrt{\mathbf{E}_Q (I_t^{\check{L}}(g))^4}$$

i.e.,

$$\sup_{0 \leq t \leq n} \mathbf{E}_Q (I_t^{\check{L}}(f))^4 (I_t^{\check{L}}(g))^2 < \infty.$$

Moreover, taking into account that the processes $(\check{L}_t)_{t \geq 0}$ and $(z_t)_{t \geq 0}$ are independent, we obtain that

$$\mathbf{E}_Q (I_t^z(f))^4 (I_t^{\check{L}}(g))^2 = \mathbf{E}_Q (I_t^z(f))^4 \mathbf{E}_Q (I_t^{\check{L}}(g))^2 = \int_0^t g^2(s)ds \mathbf{E} (I_t^z(f))^4.$$

One can check directly here that, for $t > 0$,

$$\mathbf{E} |I_t^z(f)|^4 \leq |f|_*^4 \mathbf{E} Y_1^4 \mathbf{E} N_t^2.$$

where $|f|_* = \sup_{0 \leq t \leq n} |f(t)|$. Note that Corollary 5.3 yields $\mathbf{E} N_t^2 < \infty$, therefore $\sup_{0 \leq t \leq n} \mathbf{E}_Q (I_t^z(f))^4 < \infty$ and we obtain,

$$\sup_{0 \leq t \leq n} \mathbf{E}_Q (I_t(f))^4 (I_t^{\check{L}}(g))^2 < \infty.$$

Taking into account that the process $(\check{L}_t)_{t \geq 0}$ is square integrated martingale with the quadratic characteristic $\langle \check{L} \rangle_t = t$, we obtain that $\mathbf{E} J_1 = 0$. As to the last term in (6.11) note that similarly to the previous reasoning we obtain that

$$\mathbf{E}_Q \int_0^n (I_{t-}^{\check{L}}(f))^2 I_{t-}^z(g) g(t) d\check{L}_t = 0 \quad \text{and} \quad \mathbf{E}_Q \int_0^n I_{t-}^{\check{L}}(f) I_{t-}^z(f) I_{t-}^z(g) g(t) d\check{L}_t = 0.$$

Therefore, to show (6.10) one needs to check that

$$\mathbf{E}_Q I_n^{\check{L}}(\check{h}^z) = 0, \tag{6.13}$$

where $\check{h}_t^z = (I_{t-}^z(f))^2 I_{t-}^z(g) g(t)$. To this end, note that, for any $0 < t \leq n$

$$I_t^z(f) = \sum_{k=1}^{\infty} f(T_k) Y_k \mathbf{1}_{\{T_k \leq t\}} = \sum_{k=1}^{N_n} f(T_k) Y_k \mathbf{1}_{\{T_k \leq t\}}, \tag{6.14}$$

i.e.,

$$I_n^L(\check{h}^z) = \sum_{k=1}^{N_n} \sum_{l=1}^{N_n} \sum_{j=1}^{N_n} f(T_k) f(T_l) g(T_j) Y_j Y_l Y_k I_{klj},$$

where $I_{klj} = \int_0^n \mathbf{1}_{\{T_k \leq t\}} \mathbf{1}_{\{T_l \leq t\}} \mathbf{1}_{\{T_j \leq t\}} g(t) d\check{L}_t$. Taking into account that the process $(\check{L}_t)_{t \geq 0}$ is independent of the field $\mathcal{G}_z = \sigma\{z_t, t \geq 0\}$, we obtain that $\mathbf{E}_Q (I_{klj} | \mathcal{G}_z) = 0$ and

$$\mathbf{E}_Q (I_{klj}^2 | \mathcal{G}_z) = \int_0^n \mathbf{1}_{\{T_k \leq t\}} \mathbf{1}_{\{T_l \leq t\}} \mathbf{1}_{\{T_j \leq t\}} g^2(t) dt \leq \|g\|_n^2 < \infty.$$

Moreover,

$$\begin{aligned} \mathbf{E}_Q |I_n^{\check{L}}(\check{h}^z)| &\leq |g|_* |f|_*^2 \|g\|_n \mathbf{E} \sum_{k=1}^{N_n} \sum_{l=1}^{N_n} \sum_{j=1}^{N_n} |Y_j| |Y_l| |Y_k| \\ &\leq \mathbf{E} |Y_1|^3 |g|_* |f|_*^2 \|g\|_n \mathbf{E} N_n^3. \end{aligned}$$

Corollary 5.3 implies that $\mathbf{E} N_n^3 < \infty$, i.e. $\mathbf{E}_Q |I_n^{\check{L}}(\check{h}^z)| < \infty$, and, therefore,

$$\mathbf{E}_Q I_n^{\check{L}}(\check{h}^z) = \mathbf{E}_Q \sum_{k=1}^{N_n} \sum_{l=1}^{N_n} \sum_{j=1}^{N_n} f(T_k) f(T_l) g(T_j) Y_j Y_l Y_k \mathbf{E}_Q (I_{klj} | \mathcal{G}_z) = 0.$$

So, we obtain (6.13). Furthermore, to study the last term in (6.9) note that the $(z_t)_{t \geq 0}$ is the martingale with the bounded variation. Moreover, in view of the definition of \check{h} we get

$$\begin{aligned} \mathbf{E}_Q \int_0^n |\check{h}_t| d[z]_t &= \mathbf{E}_Q \sum_{k=1}^{N_n} |\check{h}_{T_k}| Y_k^2 = \mathbf{E}_Q \sum_{k=1}^{N_n} |\check{h}_{T_k}| \\ &\leq \frac{|g|_*}{2} \left(\sum_{k \geq 1} \mathbf{E}_Q I_{T_k^-}^4(f) \mathbf{1}_{\{T_k \leq n\}} + \sum_{k \geq 1} \mathbf{E}_Q I_{T_k^-}^2(g) \mathbf{1}_{\{T_k \leq n\}} \right). \end{aligned}$$

Reminding here, that the processes $(\check{L}_t)_{t \geq 0}$ and $(z_t)_{t \geq 0}$ are independent and using (6.12), we obtain that for $1 \leq m \leq 2$ and any bounded function f

$$\begin{aligned} \sum_{k \geq 1} \mathbf{E}_Q \left(I_{T_k^-}^{\check{L}}(f) \right)^{2m} \mathbf{1}_{\{T_k \leq n\}} &\leq \sup_{0 \leq t \leq n} \mathbf{E}_Q \left(I_t^{\check{L}}(f) \right)^{2m} \mathbf{E} \sum_{k \geq 1} \mathbf{1}_{\{T_k \leq n\}} \\ &= \sup_{0 \leq t \leq n} \mathbf{E}_Q \left(I_t^{\check{L}}(f) \right)^{2m} \mathbf{E} N_n < \infty. \end{aligned}$$

Moreover, from (6.14) through the Hölder inequality we obtain that

$$\left(I_{T_k^-}^z(f) \right)^{2m} = \left(\sum_{j=1}^{k-1} f(T_j) Y_j \right)^{2m} \leq |f|_*^{2m} (k-1)^{2m-1} \sum_{j=1}^{k-1} Y_j^{2m}.$$

Taking into account that the sequences $(Y_j)_{j \geq 1}$ and $(T_k)_{k \geq 1}$ are independent we obtain through Corrolary 5.3 that

$$\begin{aligned} \sum_{k \geq 1} \mathbf{E} \left(I_{T_k^-}^z(f) \right)^{2m} \mathbf{1}_{\{T_k \leq n\}} &\leq |f|_*^{2m} \mathbf{E} Y_1^{2m} \mathbf{E} \sum_{k \geq 1} k^{2m} \mathbf{1}_{\{T_k \leq n\}} \\ &= |f|_*^{2m} \mathbf{E} Y_1^{2m} \mathbf{E} \sum_{k \geq 1}^{N_n} k^{2m} \leq |f|_*^{2m} \mathbf{E} Y_1^{2m} \mathbf{E} N_n^{2m+1} < \infty. \end{aligned}$$

Thus $\mathbf{E}_Q \int_0^n |\check{h}_t| d[z]_t < \infty$, therefore, $\mathbf{E}_Q I_n^z(\check{h}) = 0$ and we get the equality (6.8). □

7 Properties of the regression model (1.1)

In order to prove the non-asymptotic sharp oracle inequalities we use the method proposed in Konev and Pergamenschikov (2009a) and Konev and Pergamenschikov (2012) for the general semi-martingale model (1.1). To this end we need to study the following functions of $x \in \mathbb{R}^n$

$$B_{1,Q,n}(x) = \sum_{j=1}^n x_j \left(\mathbf{E}_Q \xi_{j,n}^2 - \sigma_Q \right) \quad \text{and} \quad B_{2,Q,n}(x) = \sum_{j=1}^n x_j \tilde{\xi}_{j,n}, \tag{7.1}$$

where σ_Q is defined in (2.11), $\tilde{\xi}_{j,n} = \xi_{j,n}^2 - \mathbf{E}_Q \xi_{j,n}^2$ and $\xi_{j,n}$ is given in (3.5). These functions describe the behavior for the total noise intensity and variance respectively for the chosen Fourier coefficients in the estimators (3.7).

Remark 7.1 The propositions 7.1 – 7.2 given below are used to obtain the oracle inequalities in Sect. 4 (see, for example, Konev and Pergamenschikov 2012).

Proposition 7.1 Assume that Conditions (H₁)–(H₄) hold. Then

$$\sup_{x \in [-1, 1]^n} |B_{1,Q,n}(x)| \leq C_{1,Q,n}, \tag{7.2}$$

where $C_{1,Q,n} = \sigma_Q \check{\tau} \phi_{max}^2 |\Upsilon|_1$, $\sigma_Q = \bar{\varrho} + \varrho_3^2/\check{\tau}$ and $\Upsilon(x) = \rho(x) - 1/\check{\tau}$.

Proof First, taking into account that $\xi_{j,n} = n^{-1/2} I_n(\phi_j)$ and $\|\phi_j\|_n^2 = n$, we obtain through Lemma 6.1 that

$$\mathbf{E}_Q \xi_{j,n}^2 = \bar{\varrho} + \frac{\varrho_3^2}{n} \int_0^n \phi_j^2(x) \rho(x) dx = \sigma_Q + \frac{\varrho_3^2}{n} \int_0^n \phi_j^2(x) \Upsilon(x) dx .$$

So, in view of Condition (3.1) and the Eq. (4.1), we obtain

$$\left| \mathbf{E}_Q \xi_{j,n}^2 - \sigma_Q \right| = \frac{\varrho_3^2}{n} \left| \int_0^n \phi_j^2(x) \Upsilon(x) dx \right| \leq \frac{\varrho_3^2}{n} \phi_{max}^2 |\Upsilon|_1 .$$

We bound here ϱ_3^2 by $\sigma_Q \check{\tau}$ we obtain the inequality (7.2) and hence the conclusion follows. □

Proposition 7.2 Assume that Conditions (H₁)–(H₄) hold. Then

$$\sup_{|x| \leq 1} \mathbf{E}_Q B_{2,Q,n}^2(x) \leq C_{2,Q,n}, \tag{7.3}$$

where $|x|^2 = \sum_{j=1}^n x_j^2$, $C_{2,Q,n} = \phi_{max}^4 (1 + \sigma_Q^2)^2 \check{\mathbf{I}}$ and $\check{\mathbf{I}}$ is given in (4.3).

Proof By Ito’s formula one gets

$$dI_t^2(f) = 2I_{t-}(f)dI_t(f) + \varrho_1^2 f^2(t)dt + d \sum_{0 \leq s \leq t} f^2(s)(\Delta \xi_s^d)^2. \tag{7.4}$$

Using the representations (6.2) and (6.3), we can rewrite this differential as

$$\begin{aligned} dI_t^2(f) &= 2I_{t-}(f)dI_t(f) + \varrho_1^2 f^2(t)dt \\ &+ \varrho_2^2 d \sum_{0 \leq s \leq t} f^2(s)(\Delta L_s)^2 + \varrho_3^2 d \sum_{0 \leq s \leq t} f^2(s)(\Delta z_s)^2. \end{aligned} \tag{7.5}$$

From Lemma 6.1 it follows that $\mathbf{E}_Q I_t^2(f) = \bar{\varrho} \|f\|_t^2 + \varrho_3^2 \|f \sqrt{\bar{\rho}}\|_t^2$. Therefore, putting

$$\tilde{I}_t(f) = I_t^2(f) - \mathbf{E}_Q I_t^2(f), \tag{7.6}$$

we obtain

$$d\tilde{I}_t(f) = 2I_{t-}(f)f(t)d\xi_t + f^2(t)d\tilde{m}_t, \quad \tilde{m}_t = \varrho_2^2 \check{m}_t + \varrho_3^2 m_t, \tag{7.7}$$

where $\check{m}_t = \sum_{0 \leq s \leq t} (\Delta L_s)^2 - t$ and $m_t = \sum_{0 \leq s \leq t} (\Delta z_s)^2 - \int_0^t \rho(s) ds$. For any non-random vector $x = (x_j)_{1 \leq j \leq n}$ with $\sum_{j=1}^n x_j^2 \leq 1$, we set

$$\bar{I}_t = \bar{I}_t(x) = \sum_{j=1}^n x_j \tilde{I}_t(\phi_j). \tag{7.8}$$

Denoting

$$A_t = \sum_{j=1}^n x_j I_t(\phi_j) \phi_j(t) \quad \text{and} \quad B_t = \sum_{j=1}^n x_j \phi_j^2(t), \tag{7.9}$$

we get the following stochastic differential equation for (7.8)

$$d\bar{I}_t = 2A_{t-} d\xi_t + B_t d\tilde{m}_t, \quad \bar{I}_0(x) = 0.$$

Now, similarly to (7.4) the Ito formula and representation (6.2) yield

$$\begin{aligned} d\bar{I}_t^2 &= 2\bar{I}_{t-} d\bar{I}_t + d \langle \bar{I}^c \rangle_t + d \sum_{0 \leq s \leq t} (\Delta \bar{I}_s)^2 \\ &= 2\bar{I}_{t-} d\bar{I}_t + 4\varrho_1^2 A_t^2 dt + d \sum_{0 \leq s \leq t} \left(2A_{s-} \Delta \xi_s^d + B_s \Delta \tilde{m}_s \right)^2. \end{aligned}$$

Taking into account that the processes $(L_t)_{t \geq 0}$ and $(z_t)_{t \geq 0}$ are independent, we obtain that $\Delta L_t \Delta z_t = 0$, therefore, for any $t \geq 0$

$$\begin{aligned} \left(2A_{t-} \Delta \xi_t^d + B_t \Delta \tilde{m}_t \right)^2 &= (2\varrho_2 A_{t-} \Delta L_t + \varrho_2^2 B_t (\Delta L_t)^2)^2 \\ &\quad + (2\varrho_3 A_{t-} \Delta z_t + \varrho_3^2 B_t (\Delta z_t)^2)^2. \end{aligned}$$

This implies

$$\begin{aligned} \mathbf{E}_Q \bar{I}_n^2 &= 2\mathbf{E}_Q \int_0^n \bar{I}_{t-} d\bar{I}_t + 4\varrho_1^2 \mathbf{E}_Q \int_0^n A_t^2 dt \\ &\quad + \varrho_2^2 \mathbf{E}_Q \check{D}_n + \varrho_3^2 \mathbf{E}_Q D_n, \end{aligned} \tag{7.10}$$

where $\check{D}_n = \sum_{0 \leq t \leq n} (2A_{t-} \Delta L_t + \varrho_2 B_t (\Delta L_t)^2)^2$ and

$$D_n = \sum_{0 \leq t \leq n} (2A_{t-} \Delta z_t + \varrho_3 B_t (\Delta z_t)^2)^2.$$

It should be noted here that

$$\mathbf{E}_Q B_{2,Q,n}^2(x) = \frac{1}{n^2} \mathbf{E}_Q \bar{I}_n^2(x). \tag{7.11}$$

Let us now show that

$$\left| \mathbf{E}_Q \int_0^n \bar{I}_{t-} d\bar{I}_t \right| \leq 2\varrho_3^4 \phi_{max}^3 |\Upsilon|_1 n^2. \tag{7.12}$$

To this end, note that by (7.7)

$$\begin{aligned} \int_0^n \bar{I}_{t-} d\bar{I}_t &= 2 \sum_{1 \leq j, l \leq n} x_j x_l \int_0^n \tilde{I}_{t-}(\phi_j) I_{t-}(\phi_l) \phi_l(t) d\xi_t \\ &\quad + \sum_{j=1}^n x_j \int_0^n \tilde{I}_{t-}(\phi_j) B_t d\tilde{m}_t. \end{aligned}$$

Using here Lemma 6.5, we get $\mathbf{E}_Q \int_0^n \tilde{I}_{t-}(\phi_j) I_{t-}(\phi_i) \phi_i(t) d\xi_t = 0$. Moreover, the process $(\check{m}_t)_{t \geq 0}$ is a martingale, i.e. $\mathbf{E}_Q \int_0^n \tilde{I}_{t-}(\phi_j) B_t d\check{m}_t = 0$. Therefore,

$$\mathbf{E}_Q \int_0^n \bar{I}_{t-} d\bar{I}_t = \varrho_3^2 \sum_{j=1}^n x_j \mathbf{E}_Q \int_0^n \tilde{I}_{t-}(\phi_j) B_t dm_t.$$

Taking into account here that for any non-random bounded function f

$$\mathbf{E}_Q \int_0^n f(t) dm_t = 0,$$

we obtain $\mathbf{E}_Q \int_0^n \tilde{I}_{t-}(\phi_j) B_t dm_t = \mathbf{E}_Q \int_0^n I_{t-}^2(\phi_j) B_t dm_t$. So, Lemma 6.4 yields

$$\begin{aligned} \left| \mathbf{E}_Q \int_0^n \tilde{I}_{t-}(\phi_j) B_t dm_t \right| &= \left| \sum_{l=1}^n x_l \mathbf{E}_Q \int_0^n I_{t-}^2(\phi_j) \phi_l^2(t) dm_t \right| \\ &\leq 2 \varrho_3^2 \phi_{max}^3 |\Upsilon|_1 \sum_{l=1}^n |x_l| n. \end{aligned}$$

Therefore,

$$\begin{aligned} \left| \mathbf{E}_Q \int_0^n \bar{I}_{t-} d\bar{I}_t \right| &\leq 2 \varrho_3^4 \phi_{max}^3 |\Upsilon|_1 n \sum_{1 \leq l, j \leq n} |x_l| |x_j| \\ &= 2 \varrho_3^4 \phi_{max}^3 |\Upsilon|_1 n \left(\sum_{l=1}^n |x_l| \right)^2. \end{aligned}$$

Taking into account here that $(\sum_{l=1}^n |x_l|)^2 \leq n \sum_{l \geq 1} x_l^2 \leq n$, we obtain (7.12). Recall that $\Pi(x^2) = 1$. Using the definition (2.1) and the properties of the jump measures (see, for example, Liptser and Shiryaev 1986, chapter 3) we obtain that

$$\mathbf{E}_Q \sum_{0 \leq t \leq n} (\Delta L_t)^4 = \mathbf{E}_Q \int_0^n \int_{\mathbb{R}_*} x^4 \mu(dt, dx) = \mathbf{E}_Q \int_0^n \int_{\mathbb{R}_*} x^4 \tilde{\mu}(dt, dx) = n \Pi(x^4)$$

and

$$\begin{aligned} \mathbf{E}_Q \sum_{0 \leq t \leq n} B_t A_{t-} (\Delta L_t)^3 &= \mathbf{E}_Q \int_0^n \int_{\mathbb{R}_*} B_t A_{t-} x^3 \mu(dt, dx) \\ &= \mathbf{E}_Q \int_0^n \int_{\mathbb{R}_*} B_t A_t x^3 \tilde{\mu}(dt, dx) = \Pi(x^3) \int_0^n B_t \mathbf{E}_Q A_t dt = 0. \end{aligned}$$

From this it follows directly that

$$\mathbf{E}_Q \check{D}_n = 4 \mathbf{E}_Q \int_0^n A_t^2 dt + \varrho_2^4 \Pi(x^4) \int_0^n B_t^2 dt. \tag{7.13}$$

Note that, thanks to Lemma 6.1, we obtain

$$\begin{aligned} \mathbf{E}_Q \int_0^n A_t^2 dt &= \sum_{i,j} x_i x_j \int_0^n \phi_i(t) \phi_j(t) \mathbf{E}_Q I_t \phi_i(t) I_t \phi_j(t) dt \\ &= \sum_{i,j} x_i x_j \int_0^n \phi_i(t) \phi_j(t) \int_0^t \phi_i(v) \phi_j(v) (\bar{\varrho} + \varrho_3^2 \rho(v)) dv \\ &= \frac{1}{2} \bar{\varrho} \sum_{i,j} x_i x_j \left(\int_0^n \phi_i(t) \phi_j(t) dt \right)^2 + \varrho_3^2 A_{1,n} \leq \frac{n^2}{2} \bar{\varrho} + \varrho_3^2 A_{1,n}, \end{aligned}$$

where $A_{1,n} = \sum_{i,j} x_i x_j \int_0^n \phi_i(t) \phi_j(t) \left(\int_0^t \phi_i(v) \phi_j(v) \rho(v) dv \right) dt$. This term can be estimated through Proposition 5.2 as

$$\begin{aligned} |A_{1,n}| &= \left| \frac{n^2}{2\check{\tau}} + \sum_{i,j} x_i x_j \int_0^n \phi_i(t) \phi_j(t) \left(\int_0^t \phi_i(v) \phi_j(v) \Upsilon(v) dv \right) dt \right| \\ &\leq \frac{n^2}{2\check{\tau}} + n \phi_{max}^4 |\Upsilon|_1 \sum_{i,j} |x_i| |x_j| \leq \left(\frac{1}{2\check{\tau}} + \phi_{max}^4 |\Upsilon|_1 \right) n^2. \end{aligned}$$

So, reminding that $\sigma_Q = \bar{\varrho} + \varrho_3^2/\check{\tau}$ and that $\phi_{max} \geq 1$, we obtain that

$$\begin{aligned} \mathbf{E}_Q \int_0^n A_t^2 dt &\leq \left(\frac{\sigma_Q}{2} + \varrho_3^2 \phi_{max}^4 |\Upsilon|_1 \right) n^2 \\ &\leq \left(\frac{1}{2} + \check{\tau} |\Upsilon|_1 \right) \phi_{max}^4 \sigma_Q n^2. \end{aligned} \tag{7.14}$$

Taking into account that

$$\sup_{t \geq 0} B_t^2 \leq \phi_{max}^4 \left(\sum_{j=1}^n |x_j| \right)^2 \leq \phi_{max}^4 n, \tag{7.15}$$

that $\phi_{max} \geq 1$, and that $\bar{\varrho}^2 \leq \sigma_Q^2$ we estimate the expectation in (7.13) as

$$\mathbf{E}_Q \check{D}_n \leq \phi_{max}^4 (1 + \sigma_Q^2) (1 + 2\check{\tau} |\Upsilon|_1 + \Pi(x^4)) n^2. \tag{7.16}$$

Moreover, taking into account that the random variable Y_k is independent of $A_{T_{k-}}$ and of the field $\mathcal{G} = \sigma\{T_j, j \geq 1\}$ and that $\mathbf{E}_Q (A_{T_{k-}} | \mathcal{G}) = 0$, we get

$$\begin{aligned} \mathbf{E}_Q \sum_{k=1}^{+\infty} B_{T_k} A_{T_{k-}} Y_k^3 1_{\{T_k \leq n\}} &= \sum_{k=1}^{+\infty} \mathbf{E}_Q \left(\mathbf{E}_Q (B_{T_k} A_{T_{k-}} Y_k^3 1_{\{T_k \leq n\}} | \mathcal{G}) \right) \\ &= \mathbf{E}_Q Y_1^3 \mathbf{E}_Q \sum_{k=1}^{+\infty} B_{T_k} 1_{\{T_k \leq n\}} \mathbf{E}_Q (A_{T_{k-}} | \mathcal{G}) = 0. \end{aligned}$$

Therefore,

$$\mathbf{E}_Q D_n = \varrho_3^2 D_{1,n} \mathbf{E} Y_1^4 + 4D_{2,n}, \tag{7.17}$$

where

$$D_{1,n} = \sum_{k=1}^{+\infty} \mathbf{E} B_{T_k}^2 \mathbf{1}_{\{T_k \leq n\}} \quad \text{and} \quad D_{2,n} = \sum_{k=1}^{+\infty} \mathbf{E}_Q A_{T_{k-}}^2 \mathbf{1}_{\{T_k \leq n\}}.$$

Using the bound (7.15) we can estimate the term $D_{1,n}$ as $D_{1,n} \leq \phi_{max}^4 n \mathbf{E} N_n$. Using here Corollary 5.3, we obtain

$$D_{1,n} \leq |\rho|_* \phi_{max}^4 n^2. \tag{7.18}$$

Now, to estimate the last term in (7.17), note that the process A_t can be rewritten as

$$A_t = \int_0^t \mathbf{K}(t, s) d\xi_s, \quad \mathbf{K}(t, s) = \sum_{j=1}^n x_j \phi_j(s) \phi_j(t). \tag{7.19}$$

Applying Lemma 6.3 again, we obtain for any $k \geq 1$

$$\mathbf{E}_Q \left(A_{T_{k-}}^2 | \mathcal{G} \right) = \bar{\varrho} \int_0^{T_k} \mathbf{K}^2(T_k, s) ds + \varrho_3^2 \sum_{j=1}^{k-1} \mathbf{K}^2(T_k, T_j).$$

So, we can represent the last term in (7.17) as

$$D_{2,n} = \bar{\varrho} D_{2,n}^{(1)} + \varrho_3^2 D_{2,n}^{(2)}, \tag{7.20}$$

where

$$D_{2,n}^{(1)} = \sum_{k=1}^{+\infty} \mathbf{E} \mathbf{1}_{\{T_k \leq n\}} \int_0^{T_k} \mathbf{K}^2(T_k, s) ds \quad \text{and} \quad D_{2,n}^{(2)} = \sum_{k=1}^{+\infty} \mathbf{E} \mathbf{1}_{\{T_k \leq n\}} \sum_{j=1}^{k-1} \mathbf{K}^2(T_k, T_j).$$

Thanks to Proposition 5.2 we obtain

$$D_{2,n}^{(1)} = \int_0^n \int_0^t \mathbf{K}^2(t, s) ds \rho(t) dt \leq |\rho|_* \int_0^n \int_0^n \mathbf{K}^2(t, s) ds dt.$$

In view of the definition of $\mathbf{K}(\cdot, \cdot)$ in (7.19), we can rewrite the last integral as

$$\begin{aligned} \int_0^n \mathbf{K}^2(t, s) ds &= \sum_{1 \leq i, j \leq n} x_i x_j \phi_i(t) \phi_j(t) \int_0^n \phi_i(s) \phi_j(s) ds \\ &= \sum_{i=1}^n x_i^2 \phi_i^2(t) \int_0^n \phi_i^2(s) ds = n \sum_{i=1}^n x_i^2 \phi_i^2(t). \end{aligned}$$

Since $\sum_{j=1}^n x_j^2 \leq 1$, we obtain that,

$$\int_0^n \mathbf{K}^2(t, s) ds \leq \phi_{max}^2 n \quad \text{and} \quad D_{2,n}^{(1)} \leq \phi_{max}^2 |\rho|_* n^2. \tag{7.21}$$

Let us estimate now the last term in (7.20). This term can be represented as

$$\begin{aligned}
 D_{2,n}^{(2)} &= \sum_{j=1}^{\infty} \mathbf{E} \mathbf{1}_{\{T_j \leq n\}} \mathbf{E} \left(\sum_{k=j+1}^{\infty} \mathbf{K}^2(T_k, T_j) \mathbf{1}_{\{T_k \leq n\}} | T_j \right) \\
 &= \sum_{j=1}^{\infty} \mathbf{E} \mathbf{1}_{\{T_j \leq n\}} \bar{\mathbf{K}}(T_j) = \int_0^n \bar{\mathbf{K}}(t) \rho(t) dt,
 \end{aligned}$$

where

$$\begin{aligned}
 \bar{\mathbf{K}}(t) &= \mathbf{E} \left(\sum_{k=j+1}^{\infty} \mathbf{K}^2(T_k, T_j) \mathbf{1}_{\{T_k \leq n\}} | T_j = t \right) = \sum_{l=1}^{+\infty} \mathbf{E} \mathbf{1}_{\{T_l+t \leq n\}} \mathbf{K}^2(t + T_l, t) \\
 &= \int_0^{n-t} \mathbf{K}^2(t + v, t) \rho(v) dv = \int_t^n \mathbf{K}^2(u, t) \rho(u - t) du \leq |\rho|_* \int_0^n \mathbf{K}^2(v, t) dv.
 \end{aligned}$$

In view of the inequality (7.21) we obtain $\bar{\mathbf{K}}(t) \leq |\rho|_* \phi_{max}^2 n$ and, therefore, $D_{2,n}^{(2)} \leq |\rho|_*^2 \phi_{max}^2 n^2$. So, bounding in (7.20) q_3^2 by $\check{\tau} \sigma_Q$ we obtain that

$$D_{2,n} \leq n^2 \sigma_Q (1 + \check{\tau}) |\rho|_*^2 \phi_{max}^2.$$

Therefore, taking into account in (7.17) that $\mathbf{E} Y_1^4 \geq 1$,

$$\mathbf{E}_Q D_n \leq 5 (1 + \check{\tau}) \phi_{max}^4 \mathbf{E} Y_1^4 (1 + |\rho|_*^2) n^2 \sigma_Q. \tag{7.22}$$

Using all these bound in (7.10) and taking into account that

$$\mathbf{E}_Q B_{2,Q,n}^2(x) = \frac{1}{n^2} \mathbf{E}_Q \bar{I}_n^2(x),$$

we obtain (7.3) and thus the conclusion follows. □

8 Simulation

In this section we report the results of a Monte Carlo experiment in order to assess the performance of the proposed model selection procedure (3.15). In (1.1) we chose a 1-periodic function which is defined, for $0 \leq t \leq 1$, as

$$S(t) = t \sin(2\pi t) + t^2(1 - t) \cos(4\pi t). \tag{8.1}$$

We simulate the model

$$dy_t = S(t)dt + d\xi_t,$$

where $\xi_t = 0.5w_t + 0.5z_t$.

Here z_t is the semi-Markov process defined in (2.5) with a Gaussian $\mathcal{N}(0, 1)$ sequence $(Y_j)_{j \geq 1}$ and $(\tau_k)_{k \geq 1}$ used in (2.6) taken as $\tau_k \sim \chi_3^2$.

We use the model selection procedure (3.15) with the weights (3.19) in which $k^* = 100 + \sqrt{\ln(n)}$, $t_i = i / \ln(n)$, $m = \lfloor \ln^2(n) \rfloor$ and $\delta = (3 + \ln(n))^{-2}$. We define the empirical

risk as

$$\bar{\mathbf{R}} = \frac{1}{p} \sum_{j=1}^p \hat{\mathbf{E}} (\hat{S}_n(t_j) - S(t_j))^2, \tag{8.2}$$

where the observation frequency $p = 100001$ and the expectation was taken as an average over $N = 10000$ replications, i.e.,

$$\hat{\mathbf{E}} (\hat{S}_n(\cdot) - S(\cdot))^2 = \frac{1}{N} \sum_{l=1}^N (\hat{S}_n^l(\cdot) - S(\cdot))^2.$$

We set the relative quadratic risk as

$$\bar{\mathbf{R}}_* = \bar{\mathbf{R}} / \|S\|_p^2, \quad \text{with} \quad \|S\|_p^2 = \frac{1}{p} \sum_{j=0}^p S^2(t_j). \tag{8.3}$$

In our case $\|S\|_p^2 = 0.1883601$.

Table 1 gives the values for the sample risks (8.2) and (8.3) for different numbers of observations n .

Figures 1, 2, 3 and 4 show the behaviour of the regression function and its estimates by the model selection procedure (3.15) depending on the values of observation periods n . The black full line is the regression function (8.1) and the red dotted line is the associated estimator.

Table 1 Empirical risks

n	$\bar{\mathbf{R}}$	$\bar{\mathbf{R}}_*$
20	0.04430	0.235
100	0.01290	0.068
200	0.00812	0.043
1000	0.00196	0.010

Fig. 1 Estimator of S for $n = 20$

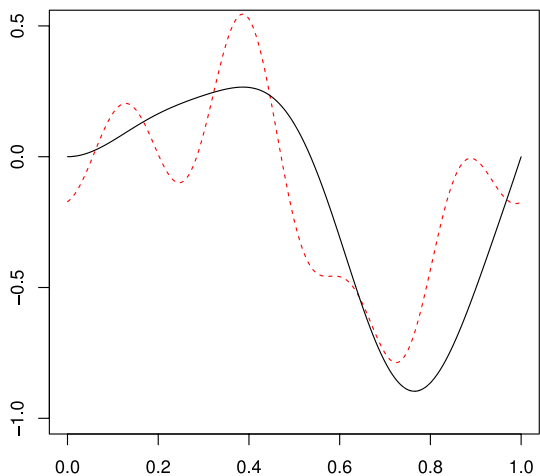


Fig. 2 Estimator of S for $n = 100$

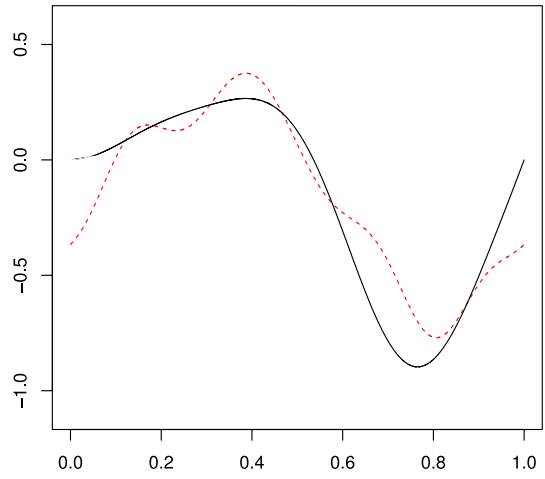


Fig. 3 Estimator of S for $n = 200$

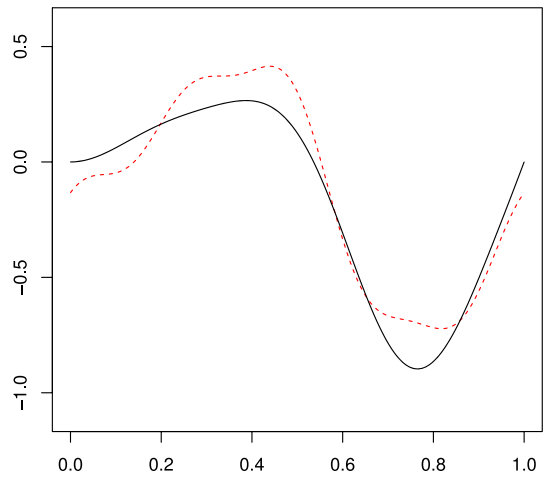
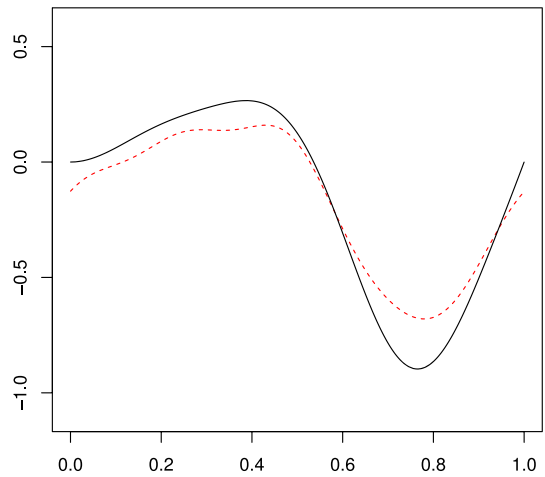


Fig. 4 Estimator of S for $n = 1000$



Remark 8.1 From numerical simulations of the procedure (3.15) with various observation numbers n we may conclude that the quality of the proposed procedure: (i) is good for practical needs, i.e. for reasonable (non large) number of observations; (ii) improves as the number of observations increases.

9 Proofs

We will prove here most of the results of this paper.

9.1 Proof of Theorem 4.1

First, note that we can rewrite the empirical squared error in (3.9) as follows

$$\text{Err}_n(\lambda) = J_n(\lambda) + 2 \sum_{j=1}^{\infty} \lambda(j) \check{\theta}_{j,n} + \|S\|^2 - \delta P_n(\lambda), \tag{9.1}$$

where the cost function $J_n(\lambda)$ and the penalty terms are defined in (3.13) and (3.14) respectively, $\check{\theta}_{j,n} = \tilde{\theta}_{j,n} - \theta_j \hat{\theta}_{j,n}$,

$$\hat{\theta}_{j,n} = \theta_j + \frac{1}{\sqrt{n}} \xi_{j,n} \quad \text{and} \quad \tilde{\theta}_{j,n} = \hat{\theta}_{j,n}^2 - \frac{\hat{\sigma}_n}{n}.$$

Using the definition of $\tilde{\theta}_{j,n}$ in (3.10) we obtain that

$$\begin{aligned} \check{\theta}_{j,n} &= \hat{\theta}_{j,n}^2 - \frac{\hat{\sigma}_n}{n} - \theta_j \left(\theta_j + \frac{1}{\sqrt{n}} \xi_{j,n} \right) = \frac{1}{\sqrt{n}} \theta_j \xi_{j,n} + \frac{1}{n} \tilde{\xi}_{j,n}^2 - \frac{\hat{\sigma}_n}{n} \\ &= \frac{1}{\sqrt{n}} \theta_j \xi_{j,n} + \frac{1}{n} \tilde{\xi}_{j,n} + \frac{1}{n} \varsigma_{j,n} + \frac{\sigma_Q - \hat{\sigma}_n}{n}, \end{aligned}$$

where $\varsigma_{j,n} = \mathbf{E}_Q \xi_{j,n}^2 - \sigma_Q$ and $\tilde{\xi}_{j,n} = \xi_{j,n}^2 - \mathbf{E}_Q \xi_{j,n}^2$. Putting

$$M(\lambda) = \frac{1}{\sqrt{n}} \sum_{j=1}^n \lambda(j) \theta_j \xi_{j,n} \quad \text{and} \quad P_n^0 = \frac{\sigma_Q |\lambda|^2}{n}, \tag{9.2}$$

we can rewrite (9.1) as

$$\begin{aligned} \text{Err}_n(\lambda) &= J_n(\lambda) + 2 \frac{\sigma_Q - \hat{\sigma}_n}{n} \check{L}(\lambda) + 2M(\lambda) + \frac{2}{n} B_{1,Q,n}(\lambda) \\ &\quad + 2\sqrt{P_n^0(\lambda)} \frac{B_{2,Q,n}(e(\lambda))}{\sqrt{\sigma_Q n}} + \|S\|^2 - \delta P_n(\lambda), \end{aligned} \tag{9.3}$$

where $e(\lambda) = \lambda/|\lambda|$, $\check{L}(\lambda) = \sum_{j=1}^n \lambda(j)$ and the functions $B_{1,Q,n}(\cdot)$ and $B_{2,Q,n}(\cdot)$ are given in (7.1).

Let $\lambda_0 = (\lambda_0(j))_{1 \leq j \leq n}$ be a fixed sequence in Λ and $\widehat{\lambda} = \operatorname{argmin}_{\lambda \in \Lambda} J_n(\lambda)$. Substituting λ_0 and $\widehat{\lambda}$ in Eq. (9.3), we obtain

$$\begin{aligned} \operatorname{Err}_n(\widehat{\lambda}) - \operatorname{Err}_n(\lambda_0) &= J(\widehat{\lambda}) - J(\lambda_0) + 2 \frac{\sigma_Q - \widehat{\sigma}_n}{n} \check{L}(\varpi) + \frac{2}{n} B_{1,Q,n}(\varpi) + 2M(\varpi) \\ &\quad + 2\sqrt{P_n^0(\widehat{\lambda})} \frac{B_{2,Q,n}(\widehat{e})}{\sqrt{\sigma_Q n}} - 2\sqrt{P_n^0(\lambda_0)} \frac{B_{2,Q,n}(e_0)}{\sqrt{\sigma_Q n}} \\ &\quad - \delta P_n(\widehat{\lambda}) + \delta P_n(\lambda_0), \end{aligned} \tag{9.4}$$

where $\varpi = \widehat{\lambda} - \lambda_0$, $\widehat{e} = e(\widehat{\lambda})$ and $e_0 = e(\lambda_0)$. Note that, by (3.8),

$$|\check{L}(\varpi)| \leq \check{L}(\widehat{\lambda}) + \check{L}(\lambda_0) \leq 2|\Lambda|_*.$$

Using the inequality

$$2|ab| \leq \delta a^2 + \delta^{-1} b^2 \tag{9.5}$$

and taking into account that $P_n^0(\lambda) \geq 0$ we obtain that for any $\lambda \in \Lambda$

$$2\sqrt{P_n^0(\lambda)} \frac{|B_{2,Q,n}(e(\lambda))|}{\sqrt{\sigma_Q n}} \leq \delta P_n^0(\lambda) + \frac{B_{2,Q,n}^2(e(\lambda))}{\delta \sigma_Q n}.$$

Taking into account the bound (7.2) and that $J(\widehat{\lambda}) \leq J(\lambda_0)$, we get

$$\begin{aligned} \operatorname{Err}_n(\widehat{\lambda}) &\leq \operatorname{Err}_n(\lambda_0) + 4 \frac{|\widehat{\sigma}_n - \sigma_Q|}{n} |\Lambda|_* + 2M(\varpi) + \frac{2C_{1,Q,n}}{n} + \frac{2B_{2,Q,n}^*}{\delta \sigma_Q n} \\ &\quad + \frac{1}{n} |\widehat{\sigma}_n - \sigma_Q| (|\widehat{\lambda}|^2 + |\lambda_0|^2) + 2\delta P_n^0(\lambda_0), \end{aligned}$$

where $B_{2,Q,n}^* = \sup_{\lambda \in \Lambda} B_{2,Q,n}^2(e(\lambda))$. Moreover, noting that in view of (3.8) $\sup_{\lambda \in \Lambda} |\lambda|^2 \leq |\Lambda|_*$, we can rewrite the previous bound as

$$\begin{aligned} \operatorname{Err}_n(\widehat{\lambda}) &\leq \operatorname{Err}_n(\lambda_0) + 2M(\varpi) + \frac{2C_{1,Q,n}}{n} + \frac{2B_{2,Q,n}^*}{\delta \sigma_Q n} \\ &\quad + \frac{6|\Lambda|_*}{n} |\widehat{\sigma}_n - \sigma_Q| + 2\delta P_n^0(\lambda_0). \end{aligned} \tag{9.6}$$

To estimate the second term in the right side of this inequality we set

$$S_x = \sum_{j=1}^n x(j) \theta_j \phi_j, \quad x = (x(j))_{1 \leq j \leq n} \in \mathbb{R}^n.$$

Taking into account that $M(x) = n^{-1} I_n(S_x)$, we can estimate this term through the inequality (2.8) shown in Corollary 6.2 for any $x \in \mathbb{R}^n$ as

$$\mathbf{E}_Q M^2(x) \leq \varkappa_Q \frac{\|S_x\|^2}{n}, \tag{9.7}$$

where, taking into account that the functions $(\phi_j)_{j \geq 1}$ are orthogonal, the norm $\|S_x\|^2 = \int_0^1 S_x^2(t) dt = \sum_{j=1}^n x^2(j)\theta_j^2$. To estimate this function for a random vector $x \in \mathbb{R}^n$ we set

$$Z^* = \sup_{x \in \Lambda_1} \frac{nM^2(x)}{\|S_x\|^2}, \quad \Lambda_1 = \Lambda - \{\lambda_0\} = \{\lambda - \lambda_0, \lambda \in \Lambda\}.$$

So, as we did for proving (9.6) and (9.7), through Inequality (9.5), we get

$$2|M(x)| \leq \delta \|S_x\|^2 + \frac{Z^*}{n\delta}. \tag{9.8}$$

It is clear that the last term here can be estimated as

$$\mathbf{E}_Q Z^* \leq \sum_{x \in \Lambda_1} \frac{n\mathbf{E}_Q M^2(x)}{\|S_x\|^2} \leq \sum_{x \in \Lambda_1} \varkappa_Q = \varkappa_Q \check{l}, \tag{9.9}$$

where $\check{l} = \text{card}(\Lambda)$. Using the equality (3.5), we obtain that for any $x \in \Lambda_1$,

$$\begin{aligned} \|S_x\|^2 - \|\widehat{S}_x\|^2 &= \sum_{j=1}^n x^2(j)(\theta_j^2 - \widehat{\theta}_j^2) \\ &= -\sum_{j=1}^n x^2(j) \left(2\frac{1}{\sqrt{n}}\theta_j \xi_{j,n} + \frac{\xi_{j,n}^2}{n} \right) \leq -2M_1(x), \end{aligned} \tag{9.10}$$

where $M_1(x) = n^{-1/2} \sum_{j=1}^n x^2(j)\theta_j \xi_{j,n}$. Taking into account that, for any $x \in \Lambda_1$ the components $|x(j)| \leq 1$, we can estimate this term as in (9.7), i.e.,

$$\mathbf{E}_Q M_1^2(x) \leq \varkappa_Q \frac{\|S_x\|^2}{n}.$$

Similarly to the previous reasoning we set

$$Z_1^* = \sup_{x \in \Lambda_1} \frac{nM_1^2(x)}{\|S_x\|^2}$$

and we get

$$\mathbf{E}_Q Z_1^* \leq \varkappa_Q \check{l}. \tag{9.11}$$

Using the same type of arguments as in (9.8), we can derive

$$2|M_1(x)| \leq \delta \|S_x\|^2 + \frac{Z_1^*}{n\delta}. \tag{9.12}$$

From here and (9.10), we get

$$\|S_x\|^2 \leq \frac{\|\widehat{S}_x\|^2}{1-\delta} + \frac{Z_1^*}{n\delta(1-\delta)} \tag{9.13}$$

for any $0 < \delta < 1$. Using this bound in (9.8) yields

$$2M(x) \leq \frac{\delta \|\widehat{S}_x\|^2}{1-\delta} + \frac{Z^* + Z_1^*}{n\delta(1-\delta)}.$$

Taking into account that

$$\|\widehat{S}_w\|^2 = \|\widehat{S}_\lambda - \widehat{S}_{\lambda_0}\|^2 = \|(\widehat{S}_\lambda - S) - (\widehat{S}_{\lambda_0} - S)\|^2 \leq 2(\text{Err}_n(\widehat{\lambda}) + \text{Err}_n(\lambda_0)),$$

we obtain

$$2M(\varpi) \leq \frac{2\delta(\text{Err}_n(\widehat{\lambda}) + \text{Err}_n(\lambda_0))}{1 - \delta} + \frac{Z^* + Z_1^*}{n\delta(1 - \delta)}.$$

Using this bound in (9.6) we obtain

$$\begin{aligned} \text{Err}_n(\widehat{\lambda}) &\leq \frac{1 + \delta}{1 - 3\delta} \text{Err}_n(\lambda_0) + \frac{Z^* + Z_1^*}{n\delta(1 - 3\delta)} + \frac{2\mathbf{C}_{1,Q,n}}{n(1 - 3\delta)} + \frac{2B_{2,Q,n}^*}{\delta(1 - 3\delta)\sigma_Q n} \\ &\quad + \frac{6|\Lambda|_*}{n(1 - 3\delta)} |\widehat{\sigma} - \sigma_Q| + \frac{2\delta}{(1 - 3\delta)} P_n^0(\lambda_0). \end{aligned}$$

Moreover, for $0 < \delta < 1/6$, we can rewrite this inequality as

$$\begin{aligned} \text{Err}_n(\widehat{\lambda}) &\leq \frac{1 + \delta}{1 - 3\delta} \text{Err}_n(\lambda_0) + \frac{2(Z^* + Z_1^*)}{n\delta} + \frac{4\mathbf{C}_{1,Q,n}}{n} + \frac{4B_{2,Q,n}^*}{\delta\sigma_Q n} \\ &\quad + \frac{12|\Lambda|_*}{n} |\widehat{\sigma}_n - \sigma_Q| + \frac{2\delta}{(1 - 3\delta)} P_n^0(\lambda_0). \end{aligned}$$

In view of Proposition 7.2 we bound the expectation of the term $B_{2,Q,n}^*$ in (9.6) as

$$\mathbf{E}_Q B_{2,Q,n}^* \leq \sum_{\lambda \in \Lambda} \mathbf{E}_Q B_{2,Q,n}^2(e(\lambda)) \leq i\mathbf{C}_{2,Q,n}.$$

Taking into account that $|\Lambda|_* \geq 1$, we get

$$\begin{aligned} \mathcal{R}(\widehat{S}_*, S) &\leq \frac{1 + \delta}{1 - 3\delta} \mathcal{R}(\widehat{S}_{\lambda_0}, S) + \frac{4\kappa_Q i}{n\delta} + \frac{4\mathbf{C}_{1,Q,n}}{n} + \frac{4i\mathbf{C}_{2,Q,n}}{\delta\sigma_Q n} \\ &\quad + \frac{12|\Lambda|_*}{n} \mathbf{E}_Q |\widehat{\sigma}_n - \sigma_Q| + \frac{2\delta}{(1 - 3\delta)} P_n^0(\lambda_0). \end{aligned}$$

Using the upper bound for $P_n^0(\lambda_0)$ in Lemma A.2, one obtains (4.1), that finishes the proof. □

9.2 Proof of Proposition 4.2

We use here the same method as in Konev and Pergamenschikov (2009a). First of all note that the definition (3.12) implies that

$$\widehat{t}_{j,n} = t_j + \frac{1}{\sqrt{n}} \eta_{j,n}, \tag{9.14}$$

where

$$t_j = \int_0^1 S(t) \text{Tr}_j(t) dt \quad \text{and} \quad \eta_{j,n} = \frac{1}{\sqrt{n}} \int_0^n \text{Tr}_j(t) d\xi_t.$$

So, we have

$$\widehat{\sigma}_n = \sum_{j=[\sqrt{n}]+1}^n t_j^2 + 2M_n + \frac{1}{n} \sum_{j=[\sqrt{n}]+1}^n \eta_{j,n}^2, \tag{9.15}$$

where

$$M_n = \frac{1}{\sqrt{n}} \sum_{j=[\sqrt{n}]+1}^n t_j \eta_{j,n}.$$

Note that, for continuously differentiable functions (see, for example, Lemma A.6 in Konev and Pergamenschikov 2009a), the Fourier coefficients (t_j) satisfy the following inequality, for any $n \geq 1$,

$$\sum_{j=[\sqrt{n}]+1}^{\infty} t_j^2 \leq \frac{4 \left(\int_0^1 |\dot{S}(t)| dt \right)^2}{\sqrt{n}} \leq \frac{4 \|\dot{S}\|^2}{\sqrt{n}}. \tag{9.16}$$

In the same way as in (9.7) we estimate the term M_n , i.e.,

$$\mathbf{E}_Q M_n^2 \leq \frac{\varkappa_Q}{n} \sum_{j=[\sqrt{n}]+1}^n t_j^2 \leq \frac{4\varkappa_Q \|\dot{S}\|^2}{n\sqrt{n}},$$

while the absolute value of this term for $n \geq 1$ can be estimated as

$$\mathbf{E}_Q |M_n| \leq \frac{\varkappa_Q + \|\dot{S}\|^2}{\sqrt{n}}.$$

Moreover, using the functions (7.1) for the trigonometric basis (3.2), i.e. with $\xi_{j,n} = \eta_{j,n}$ and $\tilde{\xi}_{j,n} = \eta_{j,n}^2 - \mathbf{E}\eta_{j,n}^2$ we can represent the last term in (9.15) as

$$\begin{aligned} \frac{1}{n} \sum_{j=[\sqrt{n}]+1}^n \eta_{j,n}^2 &= \frac{\sigma_Q(n - \sqrt{n})}{n} + \frac{\sum_{j=[\sqrt{n}]+1}^n (\mathbf{E}\eta_{j,n}^2 - \sigma_Q + \tilde{\xi}_{j,n})}{n} \\ &= \frac{\sigma_Q(n - \sqrt{n})}{n} + \frac{B_{1,Q,n}(x')}{n} + \frac{B_{2,Q,n}(x'')}{\sqrt{n}}, \end{aligned}$$

with $x'_j = \mathbf{1}_{\{\sqrt{n} < j \leq n\}}$ and $x''_j = \mathbf{1}_{\{\sqrt{n} < j \leq n\}}/\sqrt{n}$. Therefore, using Propositions 7.1 and 7.2, we obtain

$$\mathbf{E}_Q \left| \frac{1}{n} \sum_{j=[\sqrt{n}]+1}^n \eta_{j,n}^2 - \sigma_Q \right| \leq \frac{\sigma_Q}{\sqrt{n}} + \frac{\mathbf{C}_{1,Q,n}}{n} + \frac{\sqrt{\mathbf{C}_{2,Q,n}}}{\sqrt{n}}.$$

From here we obtain the bound (4.5) and hence the desired result. □

9.3 Proof of Theorem 4.4

Note, that Theorem 4.1 implies directly the oracle inequality (4.9) with

$$\mathbf{U}_n^*(S) = 60 \tilde{\Lambda}_n \|\dot{S}\|^2 + \tilde{\Psi}_n^* \quad \text{and} \quad \tilde{\Psi}_n^* = \sup_{Q \in \mathcal{Q}_n} \tilde{\Psi}_{Q,n}.$$

Using the bound (4.7) and the conditions (2.11) we obtain that for some positive constant \mathbf{C}_*

$$\tilde{\Psi}_n^* \leq \mathbf{C}_* \left(1 + (\varsigma^*)^4 + \frac{1}{\varsigma^*} \right) (1 + \tilde{\Lambda}_n) \check{\iota}_n \phi_{max}^4. \tag{9.17}$$

Moreover, note, that in view of (3.21) and (3.17)

$$\lim_{n \rightarrow \infty} \frac{\check{\gamma}}{n^{\check{\epsilon}}} = \lim_{n \rightarrow \infty} \frac{k^*m}{n^{\check{\epsilon}}} = 0 \quad \text{for any } \check{\epsilon} > 0,$$

where $m = [1/\varepsilon^2]$. Furthermore, the bound (3.22) and Conditions (2.12) and (3.17) yield

$$\lim_{n \rightarrow \infty} \frac{|\Lambda|_*}{n^{1/3+\check{\epsilon}}} = 0 \quad \text{for any } \check{\epsilon} > 0,$$

where $|\Lambda|_* = 1 + \max_{\lambda \in \Lambda} \check{L}(\lambda)$, i.e.

$$\tilde{\Lambda}_n = \frac{|\Lambda|_*}{n^{1/2}} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

So, taking into account in (9.17) the condition (4.8) we obtain the convergence (4.10). \square

9.4 Proof of Theorem 4.5

First, we denote by Q_0 the distribution of the noise (1.2) and (2.1) with the parameter $q_1 = \varsigma^*$ and $q_1 = q_2 = 0$, i.e. the distribution for the “signal + white noise” model, i.e. for any $n \geq 1$ this distribution belongs to the family \mathcal{Q}_n defined in (2.11–2.12) and, therefore, for any $n \geq 1$

$$\mathcal{R}_n^*(\tilde{S}_n, S) = \sup_{Q \in \mathcal{Q}_n} \mathbf{E}_{S,Q} \|\tilde{S}_n - S\|^2 \geq \mathbf{E}_{S,Q_0} \|\tilde{S}_n - S\|^2.$$

Now, taking into account the conditions (2.12) Theorem A.9 yields the lower bound (4.14). Hence this finishes the proof. \square

9.5 Proof of Proposition 4.6

Putting $\lambda_0(j) = 0$ for $j \geq n$ we can represent the quadratic risk for the estimator (3.7) as

$$\|\widehat{S}_{\lambda_0} - S\|^2 = \sum_{j=1}^{\infty} (1 - \lambda_0(j))^2 \theta_j^2 - 2H_n + \frac{1}{n} \sum_{j=1}^n \lambda_0^2(j) \xi_{j,n}^2,$$

where $H_n = n^{-1/2} \sum_{j=1}^n (1 - \lambda_0(j)) \lambda_0(j) \theta_j \xi_{j,n}$. Note that $\mathbf{E}_Q H_n = 0$ for any $Q \in \mathcal{Q}_n$, therefore,

$$\mathbf{E}_Q \|\widehat{S}_{\lambda_0} - S\|^2 = \sum_{j=1}^{\infty} (1 - \lambda_0(j))^2 \theta_j^2 + \frac{1}{n} \mathbf{E}_Q \sum_{j=1}^n \lambda_0^2(j) \xi_{j,n}^2.$$

Proposition 7.1 and the last inequality in (2.11) imply that for any $Q \in \mathcal{Q}_n$

$$\mathbf{E}_Q \sum_{j=1}^n \lambda_0^2(j) \xi_{j,n}^2 \leq \varsigma^* \sum_{j=1}^n \lambda_0^2(j) + \mathbf{C}_{1,n}^*,$$

where $\mathbf{C}_{1,n}^* = \phi_{max}^2 \varsigma^* \check{\tau} |\Upsilon|_1$. Therefore,

$$\mathcal{R}_n^*(\widehat{S}_{\lambda_0}, S) \leq \sum_{j=j_*}^{\infty} (1 - \lambda_0(j))^2 \theta_j^2 + \frac{1}{v_n} \sum_{j=1}^n \lambda_0^2(j) + \frac{\mathbf{C}_{1,n}^*}{n},$$

where j_* and v_n are defined in (3.19). Setting

$$\Upsilon_{1,n}(S) = v_n^{2k/(2k+1)} \sum_{j=j_*}^{\infty} (1 - \lambda_0(j))^2 \theta_j^2 \quad \text{and} \quad \Upsilon_{2,n} = \frac{1}{v_n^{1/(2k+1)}} \sum_{j=1}^n \lambda_0^2(j),$$

we rewrite the last inequality as

$$v_n^{2k/(2k+1)} R_n^*(\widehat{S}_{\lambda_0}, S) \leq \Upsilon_{1,n}(S) + \Upsilon_{2,n} + \check{C}_n, \tag{9.18}$$

where $\check{C}_n = v_n^{2k/(2k+1)} \mathbf{C}_{1,n}^*/n$. Note, that the conditions (2.12) and (4.8) imply that $\mathbf{C}_{1,n}^* = o(n^{\check{\delta}})$ as $n \rightarrow \infty$ for any $\check{\delta} > 0$; therefore, $\check{C}_n \rightarrow 0$ as $n \rightarrow \infty$. Putting

$$u_n = v_n^{2k/(2k+1)} \sup_{j \geq j_*} (1 - \lambda_0(j))^2 / a_j,$$

with $a_j = \sum_{i=0}^k (2\pi [j/2])^{2i}$, we estimate the first term in (9.18) as

$$\sup_{S \in W_r^k} \Upsilon_{1,n}(S) \leq \sup_{S \in W_r^k} u_n \sum_{j \geq 1} a_j \theta_j^2 \leq u_n \mathbf{r}.$$

We remind that ω_{α_0} is defined in (3.19) with $\alpha_0 = (k, \mathbf{l}_0)$ and $\mathbf{l}_0 = [\mathbf{r}/\varepsilon]\varepsilon$. So, taking into account that $a_j/(\pi^{2k} j^{2k}) \rightarrow 1$ as $j \rightarrow \infty$ and $\mathbf{l}_0 \rightarrow \mathbf{r}$ as $\varepsilon \rightarrow 0$ we obtain that

$$\begin{aligned} \limsup_{n \rightarrow \infty} u_n &\leq \lim_{n \rightarrow \infty} v_n^{2k/(2k+1)} \sup_{j \geq j_*} \frac{(1 - \lambda_0(j))^2}{(\pi j)^{2k}} \\ &= \lim_{n \rightarrow \infty} \frac{v_n^{2k/(2k+1)}}{\pi^{2k} \omega_{\alpha_0}^{2k}} = \frac{1}{\pi^{2k} (\mathbf{d}_k \mathbf{r})^{2k/(2k+1)}}, \end{aligned}$$

where

$$\mathbf{d}_k = \frac{(k + 1)(2k + 1)}{\pi^{2k} k}.$$

Therefore,

$$\limsup_{n \rightarrow \infty} \sup_{S \in W_r^k} \Upsilon_{1,n}(S) \leq \frac{r^{1/(2k+1)}}{\pi^{2k} (\mathbf{d}_k)^{2k/(2k+1)}} =: \Upsilon_1^*. \tag{9.19}$$

As to the second term in (9.18), note that in vue of the definition (3.19) and taking into account that $j_* = o(\omega_{\alpha_0})$ as $n \rightarrow \infty$ we deduce that

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{\omega_{\alpha_0}} \sum_{j=1}^n \lambda_0^2(j) &= \lim_{\omega_{\alpha_0} \rightarrow \infty} \frac{1}{\omega_{\alpha_0}} \sum_{1 \leq j \leq \omega_{\alpha_0}} \left(1 - \left(\frac{j}{\omega_{\alpha_0}} \right)^k \right)^2 \\ &= \int_0^1 (1 - t^k)^2 dt = \frac{2k^2}{(k + 1)(2k + 1)}. \end{aligned}$$

So, taking into account that $\omega_{\alpha_0}/v_n^{1/(2k+1)} \rightarrow (\mathbf{d}_k \mathbf{r})^{1/(2k+1)}$ as $n \rightarrow \infty$, the limit of $\Upsilon_{2,n}$ can be calculated as

$$\lim_{n \rightarrow \infty} \Upsilon_{2,n} = \frac{2(\mathbf{d}_k \mathbf{r})^{1/(2k+1)} k^2}{(k + 1)(2k + 1)} =: \Upsilon_2^*.$$

Moreover, since $\Upsilon_1^* + \Upsilon_2^* =: \mathbf{r}_k^*$, we obtain

$$\lim_{n \rightarrow \infty} v_n^{2k/(2k+1)} \sup_{S \in W_r^k} \mathcal{R}_n^*(\widehat{S}_{\lambda_0}, S) \leq \mathbf{r}_k^*$$

and get the desired result. □

9.6 Proof of Theorem 4.7

Combining Proposition 4.6 and Theorem 4.4 yields Theorem 4.7. □

Acknowledgements The last author is partially supported by RFBR Grant 16-01-00121, by the Ministry of Education and Science of the Russian Federation in the framework of the research Project No 2.3208.2017/4.6 and by the Russian Federal Professor program (Project No 1.472.2016/1.4, Ministry of Education and Science of the Russian Federation).

10 Appendix

A.1 Inequalities for the purely discontinuous martingales

Now let us recall the Novikov inequalities, Novikov (1975), also referred to as the Bichteler–Jacod inequalities (see Bichteler and Jacod 1983; Marinelli and Röckner 2014)

Lemma A.1 *Let μ be a jump measure and $\tilde{\mu}$ its compensator on the $\mathbb{R}_* = \mathbb{R} \setminus \{0\}$. Then for any $T > 0$, any predictable function h and any $p \geq 2$*

$$\mathbf{E} \sup_{0 \leq t \leq T} \left| \int_{[0, T] \times \mathbb{R}_*} h \, d(\mu - \tilde{\mu}) \right|^p \leq C_p^* \mathbf{E}_Q \check{J}_{p, T}(h), \tag{A.1}$$

where C_p^* is some positive constant and

$$\check{J}_{p, T}(h) = \left(\int_{[0, T] \times \mathbb{R}_*} h^2 \, d\tilde{\mu} \right)^{p/2} + \int_{[0, T] \times \mathbb{R}_*} h^p \, d\tilde{\mu}.$$

A.2 Property of the penalty term

Lemma A.2 *For any $n \geq 1$ and $\lambda \in \Lambda$,*

$$P_n^0(\lambda) \leq \mathbf{E}_Q \text{Err}_n(\lambda) + \frac{\mathbf{C}_{1, Q, n}}{n},$$

where the coefficient $P_n^0(\lambda)$ was defined in (9.2).

Proof By the definition of $\text{Err}_n(\lambda)$ one has

$$\begin{aligned} \mathbf{E}_Q \text{Err}_n(\lambda) &= \mathbf{E}_Q \sum_{j=1}^n \left((\lambda(j) - 1)\theta_j + \frac{\lambda(j)}{\sqrt{n}} \xi_{j, n} \right)^2 \\ &= \sum_{j=1}^n (\lambda(j) - 1)^2 \theta_j^2 + \frac{1}{n} \sum_{j=1}^n \lambda^2(j) \mathbf{E}_Q \xi_{j, n}^2. \end{aligned}$$

So, denoting by $\lambda^2 = (\lambda^2(j))_{1 \leq j \leq n}$ we obtain that

$$\mathbf{E}_Q \text{Err}_n(\lambda) \geq \frac{1}{n} \sum_{j=1}^n \lambda^2(j) \mathbf{E}_Q \xi_{j,n}^2 = P_n^0(\lambda) + \mathbf{B}_{1,Q,n}(\lambda^2).$$

where $\lambda^2 = (\lambda^2(j))_{1 \leq j \leq n}$. Now Proposition 7.1 implies the desired result, i.e.

$$\mathbf{E}_Q \text{Err}_n(\lambda) \geq P_n^0(\lambda) - \frac{\mathbf{C}_{1,Q,n}}{n}.$$

Hence lemma A.2. □

A.3 Properties of the Fourier transform

Theorem A.3 *Cauchy (1825)* Let U be a simply connected open subset of \mathbb{C} , let $g : U \rightarrow \mathbb{C}$ be a holomorphic function, and let γ be a rectifiable path in U whose start point is equal to its end point. Then

$$\oint_{\gamma} g(z) dz = 0.$$

Proposition A.4 Let $g : \mathbb{C} \rightarrow \mathbb{C}$ be a holomorphic function in $U = \{z \in \mathbb{C} : -\beta_1 < \text{Im} z < \beta_2\}$ for some $0 < \beta_1 \leq \infty$ and $0 < \beta_2 \leq \infty$. Assume that, for any $-\beta_1 < t \leq 0$,

$$\int_{\mathbb{R}} |g(\theta + it)| d\theta < \infty \quad \text{and} \quad \lim_{|\theta| \rightarrow \infty} g(\theta + it) = 0. \tag{A.2}$$

Then, for any $x \in \mathbb{R}$ and for any $0 < \beta < \beta_1$,

$$\int_{\mathbb{R}} e^{i\theta x} g(\theta) d\theta = e^{-\beta x} \int_{\mathbb{R}} e^{i\theta x} g(\theta - i\beta) d\theta. \tag{A.3}$$

Proof First note that the conditions of this theorem imply that

$$\int_{\mathbb{R}} e^{i\theta x} g(\theta) d\theta = \lim_{N \rightarrow \infty} \int_{-N}^N e^{i\theta x} g(\theta) d\theta.$$

We fix now $0 < \beta < \beta_1$ and we set for any $N \geq 1$

$$\begin{aligned} \gamma &= \{z \in \mathbb{C} : -N \leq \text{Re}z \leq N, \text{Im}z = 0\} \cup \{z \in \mathbb{C} : -N \leq \text{Im}z \leq N, \text{Re}z = N\} \\ &\cup \{z \in \mathbb{C} : -N \leq \text{Re}z \leq N, \text{Im}z = -\beta\} \cup \{z \in \mathbb{C} : -\beta \leq \text{Im}z \leq 0, \text{Re}z = -N\}. \end{aligned}$$

Now, in view of the Cauchy theorem, we obtain that for any $N \geq 1$

$$\begin{aligned} \oint_{\gamma} e^{izx} g(z) dz &= \int_{-N}^N e^{i\theta x} g(\theta) d\theta + \int_0^{-\beta} e^{i(N+it)x} g(N + it) dt \\ &\quad + \int_N^{-N} e^{i(-i\beta+\theta)x} g(-i\beta + \theta) d\theta \\ &\quad + \int_{-\beta}^0 e^{i(-N+it)x} g(-N + it) dt = 0. \end{aligned} \tag{A.4}$$

The conditions (A.2) provide that

$$\lim_{N \rightarrow \infty} \int_0^{-\beta} e^{i(N+it)x} g(N+it) dt = \lim_{N \rightarrow \infty} \int_{-\beta}^0 e^{i(-N+it)x} g(-N+it) dt = 0.$$

Therefore, letting $N \rightarrow \infty$ in (A.4) we obtain (A.3). Hence we get the desired result. \square

The following technical lemma is also needed in this paper.

Lemma A.5 *Let $g : [a, b] \rightarrow \mathbb{R}$ be a function from $\mathbf{L}_1[a, b]$. Then, for any fixed $-\infty \leq a < b \leq +\infty$,*

$$\lim_{N \rightarrow \infty} \int_a^b g(x) \sin(Nx) dx = 0 \quad \text{and} \quad \lim_{N \rightarrow \infty} \int_a^b g(x) \cos(Nx) dx = 0. \quad (\text{A.5})$$

Proof Let first $-\infty < a < b < +\infty$. Assume that g is continuously differentiable, i.e. $g \in \mathbf{C}^1[a, b]$. Then integrating by parts gives us

$$\int_a^b g(x) \sin(Nx) dx = \frac{1}{N} \left(g(b) \sin(Nb) - g(a) \sin(Na) - \int_a^b g'(x) \cos(Nx) dx \right).$$

So, from this we obtain that

$$\left| \int_a^b g(x) \sin(Nx) dx \right| \leq \frac{|g(b)| + |g(a)| + (b-a) \max_{a \leq x \leq b} |g'(x)|}{N}.$$

This implies the first limit in (A.5) for this case. The second one is obtained similarly. Let now g be any absolutely integrated function on $[a, b]$, i.e. $g \in \mathbf{L}_1[a, b]$. In this case there exists a sequence $g_n \in \mathbf{C}^1[a, b]$ such that

$$\lim_{n \rightarrow \infty} \int_a^b |g(x) - g_n(x)| dx = 0.$$

Therefore, taking into account that for any $n \geq 1$

$$\lim_{N \rightarrow \infty} \int_a^b g_n(x) \sin(Nx) dx = 0,$$

we obtain that

$$\limsup_{n \rightarrow \infty} \left| \int_a^b g(x) \sin(Nx) dx \right| \leq \int_a^b |g(x) - g_n(x)| dx.$$

So, letting in this inequality $n \rightarrow \infty$ we obtain the first limit in (A.5) and, similarly, we obtain the second one. Let now $b = +\infty$ and $a = -\infty$. In this case we obtain that for any $-\infty < a < b < +\infty$

$$\begin{aligned} \left| \int_{-\infty}^{+\infty} g(x) \sin(Nx) dx \right| &\leq \left| \int_{-\infty}^{+\infty} g(x) \sin(Nx) dx \right| + \int_b^{+\infty} |g(x)| dx \\ &\quad + \int_{-\infty}^a |g(x)| dx. \end{aligned}$$

Using here the previous results we obtain that for any $-\infty < a < b < +\infty$

$$\limsup_{N \rightarrow \infty} \left| \int_{-\infty}^{+\infty} g(x) \sin(Nx) dx \right| \leq \int_b^{+\infty} |g(x)| dx + \int_{-\infty}^a |g(x)| dx.$$

Passing here to limit as $b \rightarrow +\infty$ and $a \rightarrow -\infty$ we obtain the first limit in (A.5). Similarly, we can obtain the second one. \square

Let us now study the inverse Fourier transform. To this end, we need the following local Dini condition.

(D) Assume that, for some fixed $x \in \mathbb{R}$, there exist the finite limits

$$g(x-) = \lim_{z \rightarrow x-} g(z) \quad \text{and} \quad g(x+) = \lim_{z \rightarrow x+} g(z)$$

and there exists $\delta = \delta(x) > 0$ for which

$$\int_0^\delta \frac{|\tilde{g}(x, t) - \tilde{g}_0(x)|}{t} dt < \infty,$$

where $\tilde{g}(x, t) = g((x + t) + g(x - t))/2$ and $\tilde{g}_0(x) = (g(x+) + g(x-))/2$.

Remark 10.1 Note that the condition (H₁) is the “uniform” version of the condition D).

Proposition A.6 Let $g : \mathbb{R} \rightarrow \mathbb{R}$ be a function from $\mathbf{L}_1(\mathbb{R})$. If, for some $x \in \mathbb{R}$, this function satisfies Condition D, then

$$\tilde{g}_0(x) = \frac{g(x+) + g(x-)}{2} = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\theta x} \widehat{g}(\theta) d\theta, \tag{A.6}$$

where $\widehat{g}(\theta) = \int_{\mathbb{R}} e^{i\theta t} g(t) dt$.

Proof First, for any fixed $N > 0$ we set

$$J_N(x) = \frac{1}{2\pi} \int_{-N}^N e^{-i\theta x} \widehat{g}(\theta) d\theta = \frac{1}{\pi} \int_{\mathbb{R}} g(z) \int_0^N \cos(\theta(z - x)) d\theta dz.$$

Here, the inner integral may be represented as

$$J_N(x) = \frac{1}{\pi} \int_{\mathbb{R}} g(z) \frac{\sin(N(z - x))}{z - x} dz = \frac{2}{\pi} \int_0^\infty \tilde{g}(x, t) \frac{\sin(Nt)}{t} dt.$$

Taking into account that for any $N > 0$ the integral

$$\frac{2}{\pi} \int_0^\infty \frac{\sin(Nt)}{t} dt = 1 \tag{A.7}$$

we obtain that

$$J_N(x) - \tilde{g}_0(x) = \frac{2}{\pi} \int_0^\infty \frac{(\tilde{g}(x, t) - \tilde{g}_0(x)) \sin(Nt)}{t} dt.$$

Now we represent the last integral as

$$\int_0^\infty \frac{(\tilde{g}(x, t) - \tilde{g}_0(x)) \sin(Nt)}{t} dt = I_{1,N} + I_{2,N} - \tilde{g}_0(x) I_{3,N},$$

where $I_{1,N} = \int_0^\delta (\tilde{g}(x, t) - \tilde{g}_0(x)) \sin(Nt) t^{-1} dt$,

$$I_{2,N} = \int_\delta^\infty t^{-1} \tilde{g}(x, t) \sin(Nt) dt \quad \text{and} \quad I_{3,N} = \int_\delta^\infty t^{-1} \sin(Nt) dt.$$

Condition D) and Lemma A.5 imply that $I_{1,N} \rightarrow 0$ as $N \rightarrow \infty$. Now note that, since $g \in \mathbf{L}_1(\mathbb{R})$, then the function $t^{-1} \tilde{g}(x, t)$ is absolutely integrated. Therefore, in view of

Lemma A.5, $I_{2,N} \rightarrow 0$ as $N \rightarrow \infty$. As to the last integral we use the property (A.7), i.e., the changing of the variables gives

$$I_{3,N} = \int_{\delta N}^{\infty} \frac{\sin t}{t} dt \rightarrow 0 \text{ as } N \rightarrow \infty.$$

Hence we have the desired result. □

A.4 Properties of the analytic functions

We need the following identity theorem (see, for example, Sveshnikov and Tikhonov 1978, Corollary 1, p. 78)

Theorem A.7 *Let function $f(z) \neq 0$ and be analytic in the domain D . In any closed bounded subdomain $D' \subset D$ it has only a finite numbers of zeros.*

Lemma A.8 *Assume that the condition (\mathbf{H}_3) – (\mathbf{H}_4) hold. Then there exists some $0 < \beta_* < \beta$ such that the function \check{G} defined in (5.9) is holomorphic in*

$$U = \{\theta \in \mathbb{C} : \text{Im}(\theta) \geq -\beta_*\}$$

and for any $x \geq -\beta_*$

$$\sup_{\theta \in \mathbb{R}} |\check{G}(\theta + ix)| < \infty. \tag{A.8}$$

Proof Firstly, we remind that thanks to the condition (\mathbf{H}_3) the function $\widehat{g}(\theta)$ is holomorphic in $D = \{\theta \in \mathbb{C} : \text{Im}(\theta) > -\beta\}$. Therefore, the function $\check{G}(\theta)$ is holomorphic for any θ from $D \cap \{\theta \in \mathbb{C} : \widehat{g}(\theta) \neq 1\}$. In this case its derivative can be calculated as

$$\check{G}'(\theta) = \frac{\widehat{g}'(\theta)}{(1 - \widehat{g}(\theta))^2} - \frac{1}{i\check{\tau}\theta^2},$$

where $\widehat{g}'(\theta)$ is the derivative of the function $\widehat{g}(\theta)$. Moreover, using the Taylor expansion for e^z , the function $\widehat{g}(\theta)$ and its derivative $\widehat{g}'(\theta)$ can be represent as

$$\widehat{g}(\theta) = \mathbf{E}_Q e^{i\tau\theta} = 1 + i\check{\tau}\theta - \frac{\check{\tau}_1\theta^2}{2} + \theta^3 A_1(\theta),$$

$$\widehat{g}'(\theta) = i\mathbf{E}_Q \tau e^{i\tau\theta} = i\check{\tau} - \check{\tau}_1\theta + \theta^2 A_2(\theta), \tag{A.9}$$

where $\check{\tau}_1 = \mathbf{E}_Q \tau^2$ and the terms $A_1(\theta)$ and $A_2(\theta)$ are such that for any $L > 0$

$$\sup_{|\theta| \leq L} |A_1(\theta)| < \infty \text{ and } \sup_{|\theta| \leq L} |A_2(\theta)| < \infty.$$

Therefore, there exists $0 < \delta < 1$ for which

$$\inf_{|\theta| \leq \delta} \frac{|1 - \widehat{g}(\theta)|}{|\theta|} > 0. \tag{A.10}$$

So, the expansion (A.9) implies, that the function $\check{G}(\theta)$ is holomorphic at the point $\theta = 0$ and for such $\delta > 0$

$$\sup_{|\theta| \leq \delta} |\check{G}(\theta)| < \infty \text{ and } \sup_{|\theta| \leq \delta} |\check{G}'(\theta)| < \infty.$$

Moreover, note that in view of lemma A.5 for any $0 < r < \beta$

$$\lim_{|\operatorname{Re}(\theta)| \rightarrow \infty, |\operatorname{Im}(\theta)| \leq r} \widehat{g}(\theta) = 0. \tag{A.11}$$

Taking into account here that, thanks to (5.6), the module $|\widehat{g}(\theta)| < 1$ for any $\theta \in \mathbb{R} \setminus \{0\}$, we obtain

$$\mathbf{g}_\delta = \sup_{|\theta| > \delta} |\widehat{g}(\theta)| < 1 \text{ for any } \delta > 0. \tag{A.12}$$

Therefore, the function $\check{G}(\theta)$ is holomorphic at the line $\{z \in \mathbb{C} : \operatorname{Im}(z) = 0\}$ and

$$\sup_{\theta \in \mathbb{R}} |\check{G}(\theta)| < \infty. \tag{A.13}$$

In the general case note that, due to Lemma 5.1, the function $1 - \widehat{g}(\theta)$ has no zeros on the line $\{z \in \mathbb{C} : \operatorname{Im}(z) = -\beta_1\}$ for some $0 < \beta_1 < \beta$. Moreover, in view of Theorem A.7 for any $N > 1$ there can be only finitely many zeros of the function $1 - \widehat{g}(\theta)$ in

$$\{z \in \mathbb{C} : -\beta_1 \leq \operatorname{Im}(z) \leq 0, |\operatorname{Re}(z)| \leq N\} \subset D.$$

The limiting equality (A.11) implies that there exists $N > 0$ such that the function $1 - \widehat{g}(\theta) \neq 0$ on the set

$$\{z \in \mathbb{C} : -\beta_1 \leq \operatorname{Im}(z) \leq 0, |\operatorname{Re}(z)| > N\}.$$

So, there can be only finitely many zeros of the function $1 - \widehat{g}(\theta)$ in

$$\{z \in \mathbb{C} : -\beta_1 \leq \operatorname{Im}(z) \leq 0\}.$$

Moreover, taking into account the Taylor expansion (A.9) one can check that $\theta = 0$ is an isolated zero of the $1 - \widehat{g}(\theta)$. Therefore, there exists some $\beta_* > 0$ for which the function $1 - \widehat{g}(\theta)$ has no zeros in $\{z \in \mathbb{C} : -\beta_* \leq \operatorname{Im}(z) < 0\}$. Note also that the function $1 - \widehat{g}(\theta)$ has no zeros in $\{z \in \mathbb{C} : \operatorname{Im}(z) > 0\}$. Since we already shown that the function $\check{G}(\theta)$ is holomorphic at the line $\{z \in \mathbb{C} : \operatorname{Im}(z) = 0\}$ we obtain that it is holomorphic for any $\theta \in \mathbb{C}$ with $\operatorname{Im}(\theta) \geq -\beta_*$ and in view of (A.13) we obtain the upper bound (A.8). Hence Lemma A.8. □

A.5 Proof of (5.8)

Fist note, that for any $|\theta| > \delta$ the inequality (A.12) implies

$$|G_\epsilon(\theta)| \leq \frac{1}{|1 - (1 - \epsilon)\widehat{g}(\theta)|} + \frac{|1 - i\check{\tau}\theta|}{|\epsilon - i\check{\tau}\theta|} \leq \frac{1}{1 - \mathbf{g}_\delta} + \frac{1}{\check{\tau}\delta},$$

i.e.

$$\sup_{0 < \epsilon < 1} \sup_{|\theta| > \delta} |G_\epsilon(\theta)| < \infty.$$

Moreover, from (A.9) we can obtain that

$$G_\epsilon(\theta) = \frac{-\epsilon_1 + \epsilon_1(1 - i\check{\tau}\theta)\widehat{g}(\theta)}{(1 - \epsilon_1\widehat{g}(\theta))(\epsilon - i\check{\tau}\theta)} = \frac{\epsilon_1\check{\tau}^2\theta^2 + \epsilon_1\theta^2(1 - i\check{\tau}\theta)A_0(\theta)}{(1 - \epsilon_1\widehat{g}(\theta))(\epsilon - i\check{\tau}\theta)},$$

where $\epsilon_1 = 1 - \epsilon$ and $A_0(\theta) = -\check{\tau}_1/2 + \theta A_1(\theta)$. Note that $G_\epsilon(0) = 0$ for any $\epsilon > 0$. Let now $0 < |\theta| \leq \delta$ for some δ for which the inequality (A.10) holds. In this case for some $0 < C_* < \infty$

$$|G_\epsilon(\theta)| = \frac{|\epsilon_1 \check{\tau}^2 \theta^2 + \epsilon_1 \theta^2 (1 - i \check{\tau} \theta) A_0(\theta)|}{|1 - \epsilon_1 \widehat{g}(\theta)| \sqrt{\epsilon^2 + \check{\tau}^2 \theta^2}} \leq \frac{C_* |\theta|}{|1 - \epsilon_1 \widehat{g}(\theta)|}.$$

Moreover, taking into account here the lower bound (A.10) and that

$$|1 - \epsilon_1 \widehat{g}(\theta)| = \sqrt{(1 - \epsilon_1 \operatorname{Re}(\widehat{g}(\theta)))^2 + \epsilon_1^2 (\operatorname{Im}(\widehat{g}(\theta)))^2} \geq \epsilon_1 |\operatorname{Im}(\widehat{g}(\theta))|,$$

we obtain that

$$\sup_{0 < \epsilon < 1} \sup_{|\theta| \leq \delta} |G_\epsilon(\theta)| < \infty.$$

Hence the inequality (5.8). □

A.6 Lower bound for the robust risks

In this section we give the lower bound obtained in Konev and Pergamenschikov (2009b) (Theorem 6.1) for the robust risks (1.3) for the “signal+white noise” model, i.e. the model (1.1) in which the noise process $(\xi_t)_{0 \leq t \leq n}$ is defined as $\xi_t = \varsigma^* w_t$ for $0 \leq t \leq n$. To study the efficiency properties we need to obtain the lower bound for the mean square accuracy in the set of all possible estimators \widehat{S}_n denoted by Π_n .

Theorem A.9 *Under Conditions (2.12)*

$$\liminf_{n \rightarrow \infty} \left(\frac{n}{\varsigma^*} \right)^{2k/(2k+1)} \inf_{\widehat{S}_n \in \Pi_n} \sup_{S \in W_r^k} \mathbf{E}_{Q_0, S} \|\widehat{S}_n - S\|^2 \geq \mathbf{r}_k^*, \tag{A.14}$$

where Q_0 is the distribution of the noise process $(\varsigma^* w_t)_{0 \leq t \leq n}$ and \mathbf{r}_k^* is given in (4.13).

References

Barbu VS, Limnios N (2008) Semi-Markov Chains and Hidden Semi-Markov Models toward Applications - Their use in Reliability and DNA Analysis. Lecture Notes in Statistics, 191, Springer, New York

Barndorff-Nielsen OE, Shephard N (2001) Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial mathematics. J R Stat Soc B 63:167–241

Bichteler K, Jacod J (1983) Calcul de Malliavin pour les diffusions avec sauts: existence d’une densité dans le cas unidimensionnel. Séminaire de probabilité, vol XVII, lecture notes in Math., 986, Springer, Berlin, pp 132–157

Cont R, Tankov P (2004) Financial modelling with jump processes. Chapman & Hall, London

Goldie CM (1991) Implicit renewal theory and tails of solutions of random equations. Ann Appl Probab 1(1):126–166

Höpfner R, Kutoyants YuA (2009) On LAN for parametrized continuous periodic signals in a time inhomogeneous diffusion. Stat Decis 27(4):309–326

Höpfner R, Kutoyants YuA (2010) Estimating discontinuous periodic signals in a time inhomogeneous diffusion. Stat Infer Stoch Process 13(3):193–230

Ibragimov IA, Khasminskii RZ (1981) Statistical estimation: asymptotic theory. Springer, Berlin

Jacod J, Shiryaev AN (2002) Limit theorems for stochastic processes, 2nd edn. Springer, Berlin

Konev VV, Pergamenschikov SM (2003) Sequential estimation of the parameters in a trigonometric regression model with the gaussian coloured noise. Stat Infer Stoch Process 6:215–235

- Konev VV, Pergamenshchikov SM (2009) Nonparametric estimation in a semimartingale regression model. Part 1. Oracle inequalities. *J Math Mech Tomsk State Univ* 3:23–41
- Konev VV, Pergamenshchikov SM (2009) Nonparametric estimation in a semimartingale regression model. Part 2. Robust asymptotic efficiency. *J Math Mech Tomsk State Univ* 4:31–45
- Konev VV, Pergamenshchikov SM (2010) General model selection estimation of a periodic regression with a Gaussian noise. *Ann Inst Stat Math* 62:1083–1111
- Konev VV, Pergamenshchikov SM (2012) Efficient robust nonparametric estimation in a semimartingale regression model. *Ann Inst Henri Poincaré Probab Stat* 48(4):1217–1244
- Konev VV, Pergamenshchikov SM (2015) Robust model selection for a semimartingale continuous time regression from discrete data. *Stoch Processes Their Appl* 125:294–326
- Limnios N, Oprisan G (2001) Semi-Markov processes and reliability. Birkhäuser, Boston
- Liptser R Sh, Shiryaev AN (1986) Theory of martingales. Springer, Berlin
- Marinelli C, Röckner M (2014) On maximal inequalities for purely discontinuous martingales in infinite dimensions. *Séminaire de Probabilités, Lect. Notes Math.*, vol XLVI, pp 293–315
- Mikosch T (2004) Non-life insurance mathematics. An introduction with stochastic processes. Springer, Berlin
- Novikov AA (1975) On discontinuous martingales. *Theory Probab Appl* 20(1):11–26
- Nussbaum M (1985) Spline smoothing in regression models and asymptotic efficiency in L_2 . *Ann Statist* 13:984–997
- Pinsker MS (1981) Optimal filtration of square integrable signals in gaussian white noise. *Probl Transm Inf* 17:120–133
- Sveshnikov AG, Tikhonov AN (1978) The theory of functions of a complex variable, Translated from the Russian, English translation, Mir Publishers