Electronic Journal of Statistics
Vol. 4 (2010) 781–821
ISSN: 1935-7524
DOI: 10.1214/10-EJS568

Nonparametric estimation of the volatility function in a high-frequency model corrupted by noise

Axel Munk[*,†] **and Johannes Schmidt-Hieber**[†]

Institut für Mathematische Stochastik
Goldschmidtstr. 7, 37077 Göttingen, Germany
e-mail: munk@math.uni-goettingen.de; schmidth@math.uni-goettingen.de

Abstract: We consider the models $Y_{i,n} = \int_0^{i/n} \sigma(s)dW_s + \tau(i/n)\epsilon_{i,n}$, and $\tilde{Y}_{i,n} = \sigma(i/n)W_{i/n} + \tau(i/n)\epsilon_{i,n}$, $i = 1, \dots, n$, where $(W_t)_{t\in[0,1]}$ denotes a standard Brownian motion and $\epsilon_{i,n}$ are centered i.i.d. random variables with $\mathrm{E}(\epsilon_{i,n}^2) = 1$ and finite fourth moment. Furthermore, σ and τ are unknown deterministic functions and $(W_t)_{t\in[0,1]}$ and $(\epsilon_{1,n}, \dots, \epsilon_{n,n})$ are assumed to be independent processes. Based on a spectral decomposition of the covariance structures we derive series estimators for σ^2 and τ^2 and investigate their rate of convergence of the MISE in dependence of their smoothness. To this end specific basis functions and their corresponding Sobolev ellipsoids are introduced and we show that our estimators are optimal in minimax sense. Our work is motivated by microstructure noise models. A major finding is that the microstructure noise $\epsilon_{i,n}$ introduces an additionally degree of ill-posedness of 1/2; irrespectively of the tail behavior of $\epsilon_{i,n}$. The performance of the estimates is illustrated by a small numerical study.

AMS 2000 subject classifications: Primary 62M09, 62M10; secondary 62G08, 62G20.
Keywords and phrases: Brownian motion, variance estimation, minimax rate, microstructure noise, Sobolev embedding.

Received April 2010.

1. Introduction

Consider the models

$$Y_{i,n} = \int_0^{i/n} \sigma(s)\, dW_s + \tau\left(\frac{i}{n}\right)\epsilon_{i,n} \quad i = 1, \dots, n, \tag{1.1}$$

and

$$\tilde{Y}_{i,n} = \sigma\left(\frac{i}{n}\right) W_{i/n} + \tau\left(\frac{i}{n}\right)\epsilon_{i,n} \quad i = 1, \dots, n \tag{1.2}$$

*Corresponding author.
†The research of Axel Munk and Johannes Schmidt-Hieber was supported by DFG Grant FOR 916 and GK 1023.

respectively, where $(W_t)_{t\in[0,1]}$ denotes a Brownian motion and $\epsilon_{i,n}$ is so called microstructure noise, i.e. we assume $\epsilon_{i,n}$ i.i.d., $\mathrm{E}(\epsilon_{i,n}^2) = 1$ and $\mathrm{E}(\epsilon_{i,n}^4) < \infty$. $(W_t)_{t\in[0,1]}$ and $(\epsilon_{1,n},\dots,\epsilon_{n,n})$ are assumed to be independent, and σ and τ are unknown, positive and deterministic functions.

Our models (1.1) and (1.2) are natural extensions of the situation when σ and τ are constant, which has been, in a slightly broader setting, previously considered by [9], [14], [15] and [27] among others. In the latter papers sharp minimax estimators were derived for σ^2 and τ^2. The minimax rate for σ^2 is $n^{-1/4}$ and for τ^2 it is $n^{-1/2}$, and the corresponding constants for quadratic loss (MSE) being $8\tau\sigma^3$ and $2\tau^4$, respectively. To estimate σ and τ, maximum likelihood is feasible (see [27]) and achieves these bounds. Other efficient estimators where given by [9], [14] or [15]. In our case, i.e. when σ and τ are functions these methods fail and techniques from nonparametric regression become necessary. We will postpone a more careful dicussion of models (1.1) and (1.2) to Section 2.

Both models incorporate, as usually in high-frequency financial models, an additional noise term, denoted as microstructure noise (cf. [2] and [19]) in order to model market frictions such as bid-ask spreads and rounding errors. In general, microstructure noise is often assumed as white noise process with bounded fourth moment. Therefore, we may interpret both models as obtaining data from transformed Brownian motions under additional measurement errors. Particularly, our assumptions cover the important case when $\epsilon_{i,n} \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$.

In this paper we try to understand how estimation of the functions σ^2 and τ^2 in (1.1) and (1.2) itself can be performed in an optimal way. To our knowledge, this issue has never been addressed before whereas the problem of estimating consistently the spot volatility, i.e. the time derivative of the integrated volatility has been discussed in other works, too (cf. [1, 18]). A remarkable work in this direction is [4] where a harmonic analysis technique is introduced in order to recover σ^2. A naive estimator of σ^2 would be the derivative of an estimator of $\int_0^s \sigma^2(x)dx$ with respect to s. However, (numerical) differentiation of $\int_0^s \sigma^2(x)dx$ with respect to s yields an additional degree of ill-posedness. Instead, we propose a regularized estimator for σ and τ that attains the minimax rate of convergence. Our estimator is a Fourier series estimator where we estimate the single cosine Fourier coefficients, $\int_0^1 \sigma^2(x)\cos(k\pi x)dx$, $k = 0,1,\dots$ by a particular spectral estimator which is specifically tailor suited to this problem. The difficulty to estimate σ^2 can be explained generically from the point of view of statistical inverse problem: Microstructure noise induces an additional degree of ill posedness -similar as in a deconvolution problem- which in our case leads to a reduction of the rate of convergence by a factor $1/2$. Surprisingly, and in contrast to deconvolution, this is only reflected in the behavior of the eigenvalues of the covariance operator of the process in (1.1) and (1.2) and not in the tail behavior of the Fourier transform of the error $\epsilon_{i,n}$.

We stress again that we are aware of the fact that our model assumes a deterministic function σ and τ, which only depends on time t and generalization to $\sigma(t, X_t)$ is not obvious and a challenge for further research. However, the purely deterministic case already helps us to reveal the daily pattern of the

volatility and finally we believe that our analysis is an important step into the understanding of these models from the view point of a statistical inverse problems.

Results: All results are obtained with respect to MISE-risk. Let α and β denote a certain smoothness of σ^2 and τ^2, respectively. Roughly speaking, these numbers correspond to the usual Sobolev indices, although in our situation, a particular choice of basis is required, leading us to the definition of Sobolev s-ellipsoids (see Definition 1). Then we show that τ^2 can be estimated at rate $n^{-\beta/(2\beta+1)}$ for $\beta > 1, \alpha > 1/2$ in model (1.1) and $\beta > 1, \alpha > 3/4$ in model (1.2). This corresponds to the classical minimax rates for the usual Sobolev ellipsoids without the Brownian motion term in (1.1) and (1.2). More interesting, we obtain for estimation of σ^2 the $n^{-\alpha/(4\alpha+2)}$ rate of convergence for $\alpha > 3/4, \beta > 5/4$ in model (1.1) and $\alpha > 3/2, \beta > 5/4$ in model (1.2). We will show that these rates are uniform for Sobolev s-ellipsoids. Lower bounds with respect to Hölder classes for estimation of σ^2 have been obtained in [20]. Here we will extend this result to Sobolev s-ellipsoids. It follows that the obtained rates are minimax, indeed.

To summarize, our major finding is that in contrast to ordinary deconvolution the difficulty of estimation σ^2 when corrupted by additional (microstructure) noise ϵ, is generically increased by a factor of $1/2$ within the s-ellipsoids. This is quite surprising because one might have expected that for instance Gaussian error leads to logarithmic convergence rates due to its exponential decay of the Fourier transform (see e.g. [5], [7], [8] and [12] for some results in this direction). We stress that for our method a minimal smoothness of σ in (1.1) of $\alpha > 1/2$ and in (1.2) of $\alpha > 3/2$ is required. Although convergence rates are half compared with usual nonparametric regression, it turns out that for large sample sizes we get reasonable estimates for smooth functions σ^2. Roughly speaking, the results imply that n data points for estimation of σ^2 can be compared to the situation, when we have $\sqrt{n}$ observation in usual nonparamteric regression.

The work is organized as follows. In Sections 2 and 3 we will discuss models (1.1) and (1.2) in more detail, introduce notation and define the required smoothness classes, Sobolev s-ellipsoids (details can be found in Appendix B). Section 4.1 and Section 4.2 are devoted to estimate σ^2 and τ^2, respectively, and to present the rates of convergence of the estimators (for a proof see Appendix A). Section 5 provides the minimax result. In Section 6 we briefly discuss some numerical results and illustrate the robustness of the estimator against non-normality and violations of the required smoothness assumptions for σ^2 and τ^2. Some further results and technicalities of Sections 4.1 and 4.2 are given in Appendices C and D.

2. Discussion of models (1.1) and (1.2)

In this subsection we briefly discuss the background from financial economics of model (1.1) and explore the differences between models (1.1) and (1.2). We may consider the processes $(\sigma(t)W_t)_{t\in[0,1]}$ and $\left(\int_0^t \sigma(s)dW_s\right)_{t\in[0,1]} \stackrel{\mathcal{D}}{=} (W(H(t)))_{t\in[0,1]}$, $H(t) := \int_0^t \sigma^2(s)\,ds$ as (inhomogeneously) scaled Brownian motions, where

scaling takes place in space and in time, respectively. Hence we will refer to $(\sigma(t) W_t)_{t\in[0,1]}$ and $\left(\int_0^t \sigma(s) dW_s\right)_{t\in[0,1]}$ in the future as space-transformed (sBM) and time-transformed (tBM) Brownian motion.

Model (1.1): In the financial econometrics literature variations of model (1.1) are often denoted as high-frequency models, since $(W_t)_{t\in[0,1]}$ is sampled on time points $t = i/n$ and nowadays there is a vast amount of literature on volatility estimation in high-frequency models with additional microstructure noise term (see [3], [16], [17], [30] and [31]). These kinds of models have attained a lot of attention recently, since the usual quadratic variation techniques for estimation of $\int_0^1 \sigma^2(x)dx$ lead to inconsistent estimators (cf. [30]).

We are aware of the fact, that in contrast to our model, volatility is modelled generally not only as time dependent but also depending on the process itself, i.e. $Y_{i,n} = X_{i/n} + \tau(i/n)\epsilon_{i,n}$, $i = 1,\dots,n$, $dX_t = \sigma(t, X_t) dW_t$. An overview over commonly used parametric forms of $\sigma(t, X_t)$ and a non-parametric treatment in the absence of microstructure noise, can be found in [13]. It is known that the same rates as for the case σ and τ constant hold true if we consider the model (1.1) and estimate the so called integrated volatility or realized volatility $\int_0^s \sigma^2(x)dx$ ($s \in [0,1]$) and $\int_0^s \tau^2(x)dx$ instead of σ^2 and τ^2, respectively (see [23] and [25] for a discussion on estimation of integrated volatility and related quantities). Recently, model (1.1) has been proven to be asymptotically equivalent to a Gaussian shift experiment (see [24]). σ^2 as a function of time corresponds in model (1.1) to the instantaneous volatility or spot volatility.

Model (1.2): Model (1.2) can be regarded as a nonparametric extension of the model with constant σ, τ as discussed for variogram estimation by [27]. In order to show how sBM generalizes Brownian motion, we give the following Lemma.

Lemma 1. *(i) Assume that σ, $0 < c \le \sigma$, is continuously differentiable. Then the corresponding sBM, $(\sigma(t) W_t)_{t\in[0,1]}$ is the unique solution of the SDE*

$$dX_t = X_t \, d(\log(\sigma(t))) + \sigma(t) dW_t, \quad X_0 = 0, \ 0 \le t \le T.$$

(ii) The variogram of sBM is given by

$$\begin{aligned}\gamma(s,t) &:= \mathrm{E}(X_t - X_s)^2 \\ &= \left(\sigma(t) t^{1/2} - \sigma(s) s^{1/2}\right)^2 + \sigma(t)\sigma(s)\left[|s-t| - \left(s^{1/2} - t^{1/2}\right)^2\right].\end{aligned}$$

Proof. (i) It is easy to check that sBM indeed is a solution. To establish uniqueness, we apply Theorem 9.1 in [26]. *(ii)* This follows by straightforward calculations. □

Comparison of the models: We remark that tBM can be related to sBM by partial integration $\int_0^t \sigma(s) dW_s = \sigma(t) W_t - \int_0^t \sigma'(s) W_s ds$. Thus, sBM can

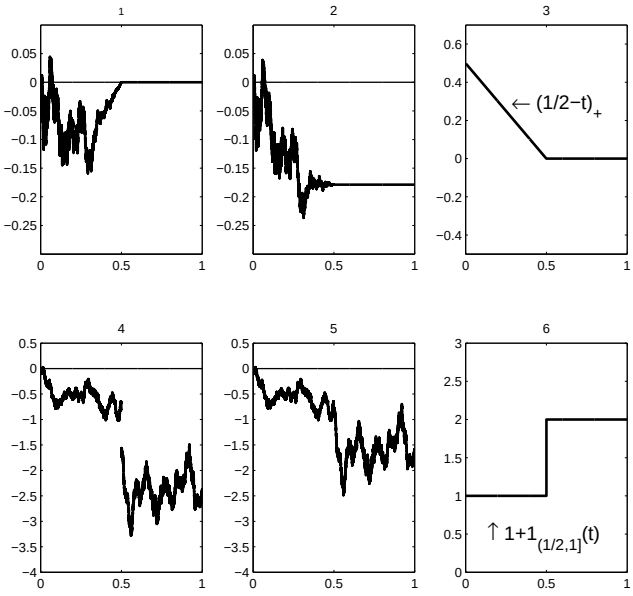

FIG 1. *Plots* 1 *and* 2 *display paths of sBM and tBM corresponding to* $\sigma(t) = (1/2 - t)_+$ *(Plot* 3*). Analogously, Plots* 4 *and* 5 *show paths of sBM and tBM with* $\sigma(t) = 1 + \mathbb{I}_{(1/2,1]}(t)$ *(Plot* 6*). For Plots* 1 *and* 2 *as well as Plots* 4 *and* 5 *we took the same realization* $(W_t)_{t\in[0,1]}$ *of the underlying Brownian motion. The first two plots show the different scaling behavior: sBM*$= 0$ *and tBM*$= \int_0^{1/2} \sigma(s)\, dW_s$ *for* $t > 1/2$*. On the other hand we see by Plots* 4 *and* 5 *that a jump induces a random shift, i.e. sBM=tBM for* $t \leq 1/2$ *and tBM*$+W_{1/2} =$*sBM for* $t > 1/2$.

be also interpreted as tBM plus a stochastic drift term. To see the differences between the processes, we compared in Figure 1 sBM and tBM in two typical situations: The case where $\sigma(t) = 0$ for $t > T$ and the case, where σ is non-continuous. If $\sigma(t) = 0$ for $t > T$, sBM tends to zero, whereas tBM tends to a constant, i.e. the random variable $\int_0^T \sigma(s)\, dW_s$. Furthermore, if σ is a jump function, sBM has a jump too, whereas tBM does not.

Unlike Model (1.1), which can be viewed as a price process, Model (1.2) has no direct application in financial mathematics. However, from the view point of nonparametric statistics it seems to be a natural extension of the situation when σ and τ are constant.

3. Introduction to Sobolev s-ellipsoids and technical preliminaries

In this section we shortly introduce the setup needed in order to define the estimators. First we define suitable smoothness classes, which are different, but related to well known Sobolev ellipsoids (see Definition B.1).

Definition 1. *For $\alpha > 0$, $C > 0$, we call the function space*

$$\Theta_s := \Theta_s(\alpha, C) := \left\{ f \in L^2[0,1] : \exists\, (\theta_n)_{n\in\mathbb{N}},\ s.\ t.\ f(x) = \theta_0 + 2\sum_{i=1}^{\infty} \theta_i \cos(i\pi x),\ \sum_{i=1}^{\infty} i^{2\alpha}\theta_i^2 \le C \right\}$$

a Sobolev s-ellipsoid. If there is a $C < \infty$ such that $f \in \Theta_s(\alpha, C)$, we say f has smoothness α. For $0 < l < u < \infty$, we further introduce the uniformly bounded Sobolev s-ellipsoid

$$\Theta_s^b(\alpha, C) := \Theta_s^b(\alpha, C, [l,u]) := \{ f \in \Theta_s(\alpha, C) : l \le f \le u \}.$$

Here the "s" refers to "*symmetry*" since the $L^2[0,1]$ basis

$$\{\psi_k, k = 0, \ldots\} := \left\{ 1, \sqrt{2}\cos(k\pi t), k = 1, \ldots \right\}, \tag{3.1}$$

can also be viewed as a basis of the symmetric $L^2[-1,1]$ functions

$$\left\{ f : f \in L^2[-1,1], f(x) = f(-x)\ \forall x \in [0,1] \right\}.$$

Usually, Sobolev ellipsoids are introduced with respect to the Fourier basis

$$\left\{ 1, \sqrt{2}\sin(2k\pi t), \sqrt{2}\cos(2k\pi t), k = 1, \ldots \right\}$$

on $L_2[0,1]$ (see Definition (B.1)). As will turn out later on, Sobolev s-ellipsoids appear naturally in our approach. If a function has a certain smoothness in one space, it might have a completely different smoothness with respect to the other basis. For instance the function $\cos((2l+1)\pi x)$, $l \in \mathbb{N}$ has smoothness α for all $\alpha < \infty$ with respect to basis (3.1), and as can be seen by direct calculations only smoothness $\alpha < 1/2$ for the Fourier basis. A more precise discussion can be found in Part B of the Appendix.

Instead of (3.1) it is convenient to introduce the functions $f_k : [0,1] \to \mathbb{R}$, $k \in \mathbb{N}$

$$f_k(x) := \psi_k\left(\frac{x}{2}\right).$$

Note that for $k \ge 1$, f_k^2 can be expanded in basis (3.1) by $f_k^2 = \psi_0 + 2^{-1/2}\psi_k$. For any function g we introduce the forward difference operator $\Delta_i g := g((i+1)/n) - g(i/n)$ and further the transformed variables $\Delta Y_{i,n}^{k,1} := (Y_{i+1,n} - Y_{i,n}) f_k(i/n)$ and $\Delta Y_{i,n}^{k,2} := \big(\tilde{Y}_{i+1,n} - \tilde{Y}_{i,n}\big) f_k(i/n)$, $i = 1, \ldots, n-1$ for models (1.1) and (1.2), respectively. In order to discuss the models simultaneously, we will write $\Delta Y_{i,n}^k = \Delta Y_{i,n}^{k,l}$, $l = 1, 2$. Throughout the paper we abbreviate first order differences of observations by

$$\Delta Y^k := \big(\Delta Y_{1,n}^k, \ldots, \Delta Y_{n-1,n}^k\big)^t.$$

We write $\mathbb{M}_{p,q}$, $\mathbb{M}_p$ and $\mathbb{D}_p$ for the space of $p \times q$ matrices, $p \times p$ matrices and $p \times p$ diagonal matrices over $\mathbb{R}$, respectively. Further let $D_{n-1} \in \mathbb{M}_{n-1}$ given by $(D_{n-1})_{i,j} = \sqrt{2/n} \sin(ij\pi/n)$ and define

$$\lambda_{i,n-1} := 4 \sin^2(i\pi/(2n)) \quad i = 1, \ldots, n-1 \,, \tag{3.2}$$

the eigenvalues of the covariance matrix $K_{n-1} \in \mathbb{M}_{n-1}$ of the MA(1) process $\Delta_i \epsilon_{i,n} := \epsilon_{i+1,n} - \epsilon_{i,n}$, $i = 1, \ldots, n-1$. More explicitly K_{n-1} is tridiagonal and

$$(K_{n-1})_{i,j} = \begin{cases} 2 & \text{for} \quad i = j \\ -1 & \text{for} \quad |i-j| = 1 \\ 0 & \text{else} \end{cases} \,. \tag{3.3}$$

Note that we can diagonalize K_{n-1} explicitly by $K_{n-1} = D_{n-1} \Lambda_{n-1} D_{n-1}$, where Λ_{n-1} is diagonal with diagonal entries given by (3.2).

We will suppress the index $n-1$ and write K, D, Λ, λ_i instead of K_{n-1}, D_{n-1}, Λ_{n-1}, and $\lambda_{i,n}$, respectively. We write $[x] := \max_{z \in \mathbb{Z}} \{z \leq x\}$, $x \in \mathbb{R}$, the integer part of x. log() is defined to be the binary logarithm and in order to define estimators properly, we assume throughout the paper additionally $n > 16$.

4. Estimators and rates of convergence

4.1. Estimation of τ^2

Before we will turn to the estimation of the volatility σ^2, we will first discuss estimation of the noise variance, i.e. τ^2. Let $J_n^\tau \in \mathbb{D}_{n-1}$ given by

$$(J_n^\tau)_{i,j} := \begin{cases} (n - n/\log n)^{-1} \lambda_i^{-1} \delta_{i,j}, & \text{for} \quad [n/\log n] \leq i,j \leq n-1 \\ 0 \quad \text{otherwise} \end{cases} \,,$$

where λ_i is defined by (3.2) and $\delta_{i,j}$ denotes the Kronecker delta. We consider models (1.1) and (1.2), simultaneously. Let

$$\hat{t}_{k,0} := \left(\Delta Y^k\right)^t D J_n^\tau D^t \left(\Delta Y^k\right). \tag{4.1}$$

In Lemma C.1 it will be shown that $\hat{t}_{k,0}$ is a $\sqrt{n}-$consistent estimator of

$$t_{k,0} := \int_0^1 \tau^2(x) f_k^2(x)\, dx.$$

Note that for $k \geq 1$ this means $t_{k,0} = \int_0^1 \tau^2(x)\psi_0(x)dx + 2^{-1/2} \int_0^1 \tau^2(x)\psi_k(x)dx$. Define $Z := D\left(\Delta Y^k\right)$ and denote by Z_i the i-th component of Z. Then

$$\hat{t}_{k,0} = (n - n/\log n)^{-1} \sum_{i=[n/\log n]}^{n-1} \lambda_i^{-1} Z_i^2. \tag{4.2}$$

Hence this also can be seen as a spectral filter in Fourier domain, where we cut off the first $n/\log n$ frequencies. Note that for $i \geq 1$, $2^{1/2}(t_{i,0} - t_{0,0}) = \int_0^1 \tau^2(x)\psi_i(x)dx$ is the i-th series coefficient with respect to basis (3.1). This observation suggests to construct the cosine series estimator

$$\hat{\tau}_N^2(t) := \hat{t}_{0,0} + 2\sum_{i=1}^{N}\left(\hat{t}_{i,0} - \hat{t}_{0,0}\right)\cos(i\pi t). \tag{4.3}$$

The next result provides the rate of convergence of $\hat{\tau}_N^2$ uniformly within Sobolev s-ellipsoids. To this end a version of the continuous Sobolev embedding theorem is required for non-integer indices α, β (see Lemma D.8). A proof of the following Theorem can be found in Appendix C.

Theorem 1 (MISE of $\hat{\tau}_N^2(t)$)**.** *Let $\hat{\tau}_N^2(t)$ as defined in (4.3). Assume $\beta > 1$, and $Q, \bar{Q} > 0$. Further suppose that $N = N_n = o\left(n^{1/2}/\log n\right)$. Assume either model (1.1) and $\alpha > 1/2$ or model (1.2) and $\alpha > 3/4$. Then it holds*

$$\sup_{\sigma^2 \in \Theta_s^b(\alpha,Q), \tau^2 \in \Theta_s^b(\beta,\bar{Q})} \mathrm{MISE}\left(\hat{\tau}_N^2\right) = O\left(N^{-2\beta} + Nn^{-1}\right).$$

Minimizing the r.h.s. yields $N^* = O\left(n^{1/(2\beta+1)}\right)$ and consequently

$$\sup_{\sigma^2 \in \Theta_s^b(\alpha,Q), \tau^2 \in \Theta_s^b(\beta,\bar{Q})} \mathrm{MISE}\left(\hat{\tau}_{N^*}^2\right) = O\left(n^{-2\beta/(2\beta+1)}\right).$$

Remark 1. *Note that for model (1.1) Theorem 1 holds, whenever $\alpha > 1/2$. Hence the Brownian motion part of the model can be viewed as a nuisance parameter, not affecting rates for estimation of τ^2. However, for model (1.2) $\alpha > 3/4$ is required here. This more restrictive assumption is essentially a consequence of the fact that the process $\sigma(i/n)W_{i/n}$ is in general no martingale.*

Remark 2. *The result from Theorem 1 can be extended to $1/2 < \beta \leq 1$ in model (1.1) and to $1/2 < \alpha \leq 3/4$, $1/2 < \beta \leq 1$ in model (1.2). Let $\tilde{t}_{k,0}$ be defined as $\hat{t}_{k,0}$ in (4.1) but J_n^τ is now replaced by $\tilde{J}_n^\tau \in \mathbb{D}_{n-1}$,*

$$\left(\tilde{J}_n^\tau\right)_{i,j} = \begin{cases} 2n^{-1}\lambda_i^{-1}\delta_{i,j}, & \text{for} \quad [n/2] \leq i,j \leq n-1 \\ 0 & \text{otherwise} \end{cases}.$$

Introduce further the estimator $\tilde{\tau}_N^2(t) = \tilde{t}_{0,0} + 2\sum_{i=1}^{N}\left(\tilde{t}_{i,0} - \tilde{t}_{0,0}\right)\cos(i\pi t)$. Further suppose that $N = O\left(n^{1/(2\beta+1)}\right)$. Then we obtain by slight modifications of the proof of Theorem 1 for $\beta > 1/2$, $\alpha > 1/2$ and $Q, \bar{Q} > 0$

(i) Assume model (1.1). Then it holds

$$\sup_{\sigma^2 \in \Theta_s^b(\alpha,Q), \tau^2 \in \Theta_s^b(\beta,\bar{Q})} \mathrm{MISE}\left(\tilde{\tau}_N^2\right) = O\left(N^{-2\beta} + Nn^{-1} + Nn^{1-2\beta}\right)$$

and $N^* = O\left(n^{(2\beta-1)/(2\beta+1)}\right)$ yields

$$\sup_{\sigma^2 \in \Theta_s^b(\alpha,Q), \tau^2 \in \Theta_s^b(\beta,\bar{Q})} \mathrm{MISE}\left(\tilde{\tau}_{N^*}^2\right) = O\left(n^{-\left(4\beta^2-2\beta\right)/(2\beta+1)}\right).$$

(ii) Assume model (1.2). Then we have the expansion

$$\sup_{\sigma^2\in\Theta_s^b(\alpha,Q),\tau^2\in\Theta_s^b(\beta,\bar{Q})} \mathrm{MISE}\left(\tilde{\tau}_N^2\right) = O\left(N^{-2\beta}+Nn^{-1}+Nn^{1-2\beta}+Nn^{2-4\alpha}\right),$$

and the choice

$$N^* = \begin{cases} O\left(n^{(2\beta-1)/(2\beta+1)}\right) & \text{for } \beta \le 1\wedge(2\alpha-1/2), \\ O\left(n^{(4\alpha-2)/(2\beta+1)}\right) & \text{for } \alpha \le 3/4\wedge(\beta/2+1/4) \end{cases}$$

yields

$$\sup_{\sigma^2\in\Theta_s^b(\alpha,Q),\tau^2\in\Theta_s^b(\beta,\bar{Q})} \mathrm{MISE}\left(\tilde{\tau}_{N^*}^2\right) = \begin{cases} O\left(n^{-\left(4\beta^2-2\beta\right)/(2\beta+1)}\right) & \text{for } \beta \le 1\wedge(2\alpha-1/2), \\ O\left(n^{-(2-2\alpha)/(2\beta+1)}\right) & \text{for } \alpha \le 3/4\wedge(\beta/2+1/4). \end{cases}$$

Remark 3. *It is also possible, although more technical, to compute the asymptotic constant of the estimator $\hat{\tau}_{N^*}^2$. Suppose that the microstructure noise is Gaussian and assume model (1.1) and $\beta > 1$ or (1.2) and $\beta > 1, \alpha > 3/4$, then we have more explicitly*

$$\mathrm{MISE}\left(\hat{\tau}_{N^*}^2\right) = \frac{2N^*}{n}\int_0^1 \tau^4(x)dx + \sum_{k=N^*+1}^{\infty}\left(\int_0^1 \tau^2(x)\psi_k(x)dx\right)^2 + o\left(N^*n^{-1}\right).$$

Remark 4. *There are of course simpler estimators for $t_{k,0}$. For instance if we replace J_n^τ in (4.1) by $(2n)^{-1}\mathbb{I}_{n-1}$, where $\mathbb{I}_{n-1}\in\mathbb{D}_{n-1}$ denotes the identity matrix, we obtain the quadratic variation estimator for $t_{k,0}$ (cf. [2]) and it is not difficult to show that this estimator attains the optimal rate of convergence. This approach could even be extended to a nonparametric estimator of the form (4.3). However, the single Fourier coefficients are not estimated efficiently, since in the case when the microstructure noise is Gaussian the asymptotic constant is $3n^{-1}\int\tau_k^4(x)dx$ (this is a straightforward extension of Theorem A.1 in [31]) whereas for our estimator we have $2n^{-1}\int\tau_k^4(x)dx$ (see Lemma C.1). If τ is constant it can be easily seen that estimators in (4.1) are efficient for $k=0$ whereas quadratic variation is not.*

Remark 5. *In practical application it would be more natural to use instead of $n/\log n$ in (4.2) other cut-off frequencies e.g. $n^\gamma/\log n$ or qn, where $1/2<\gamma\le 1$, $0<q<1$. Smaller γ decreases the variance while on the other hand increases the bias of the estimator.*

4.2. Estimation of σ^2

Define $J_n\in\mathbb{D}_{n-1}$ by

$$(J_n)_{i,j} = \begin{cases} \sqrt{n}\delta_{i,j}, & \text{for } \left[n^{1/2}\right]+1 \le i,j \le 2\left[n^{1/2}\right] \\ 0 & \text{otherwise} \end{cases}. \tag{4.4}$$

Similar, as for the estimation of τ^2 we first introduce an estimator of appropriate Fourier coefficients by

$$\hat{s}_{k,0} = \left(\Delta Y^k\right)^t D J_n D^t \left(\Delta Y^k\right) - 7\pi^2 \hat{t}_{k,0}/3. \tag{4.5}$$

The second part, i.e. $-7\pi^2\hat{t}_{k,0}/3$ is a bias correcting term, where the constant $7\pi^2/3$ is due to the choice of cut-off points $\lceil n^{1/2} \rceil + 1$ and $2\lceil n^{1/2} \rceil$ in (4.4). As we will see, the estimator of $\hat{t}_{k,0}$ has better convergence properties than the first term in $\hat{s}_{k,0}$, and hence does not affect the asymptotic variance. Similar to (4.3), we put

$$\hat{\sigma}^2_N(t) = \hat{s}_{0,0} + 2\sum_{i=1}^{N} \left(\hat{s}_{i,0} - \hat{s}_{0,0}\right) \cos\left(i\pi t\right). \tag{4.6}$$

Theorem 2 (MISE of $\hat{\sigma}^2_N$)**.** *Let $\hat{\sigma}^2_N$ as defined in (4.6). Suppose that $N = N_n = o\left(n^{1/4}\right)$, $\beta > 5/4$ and $Q, \bar{Q} > 0$. Assume model (1.1) and $\alpha > 3/4$ or model (1.2) and $\alpha > 3/2$. Then it holds*

$$\sup_{\sigma^2 \in \Theta^b_s(\alpha,Q), \tau^2 \in \Theta^b_s\left(\beta,\bar{Q}\right)} \mathrm{MISE}\left(\hat{\sigma}^2_N\right) = O\left(N^{-2\alpha} + Nn^{-1/2}\right)$$

and minimizing the r.h.s. yields

$$\sup_{\sigma^2 \in \Theta^b_s(\alpha,Q), \tau^2 \in \Theta^b_s\left(\beta,\bar{Q}\right)} \mathrm{MISE}\left(\hat{\sigma}^2_{N^*}\right) = O\left(n^{-\alpha/(2\alpha+1)}\right)$$

for $N^* = O\left(n^{1/(4\alpha+2)}\right)$.

The proof of Theorem 2 is given in Section A.2.

Remark 6. *It is also possible to extend this result for less smooth functions σ^2 and τ^2.*

(i) Assume model (1.1) and $\alpha > 1/2$, $\beta > 1$. Then it holds

$$\sup_{\sigma^2 \in \Theta^b_s(\alpha,Q), \tau^2 \in \Theta^b_s\left(\beta,\bar{Q}\right)} \mathrm{MISE}\left(\hat{\sigma}^2_N\right) = O\left(N^{-2\alpha} + Nn^{-1/2} + Nn^{2-2\beta} + Nn^{1-2\alpha}\right),$$

and

$$N^* = \begin{cases} O\left(n^{(2\alpha-1)/(2\alpha+1)}\right) & \text{for } \alpha \le 3/4 \wedge (\beta - 1/2), \\ O\left(n^{(2\beta-2)/(2\alpha+1)}\right) & \text{for } \beta \le 5/4 \wedge (\alpha + 1/2) \end{cases}$$

yields

$$\sup_{\sigma^2 \in \Theta^b_s(\alpha,Q), \tau^2 \in \Theta^b_s\left(\beta,\bar{Q}\right)} \mathrm{MISE}\left(\hat{\sigma}^2_{N^*}\right) = \begin{cases} O\left(n^{-2\alpha(2\alpha-1)/(2\alpha+1)}\right) & \text{for } \alpha \le 3/4 \wedge (\beta - 1/2), \\ O\left(n^{-2\alpha(2\beta-2)/(2\alpha+1)}\right) & \text{for } \beta \le 5/4 \wedge (\alpha + 1/2). \end{cases}$$

(ii) Assume model (1.2) and $\alpha > 3/2,\ \beta > 1$. *Then it holds*

$$\sup_{\sigma^2 \in \Theta_s^b(\alpha,Q), \tau^2 \in \Theta_s^b(\beta,\bar{Q})} \mathrm{MISE}\left(\hat{\sigma}_N^2\right) = O\left(N^{-2\alpha} + Nn^{-1/2} + Nn^{2-2\beta}\right),$$

and $N^* = O\left(n^{(2\beta-2)/(2\alpha+1)}\right)$ *yields*

$$\sup_{\sigma^2 \in \Theta_s^b(\alpha,Q), \tau^2 \in \Theta_s^b(\beta,\bar{Q})} \mathrm{MISE}\left(\hat{\sigma}_{N^*}^2\right) = O\left(n^{-2\alpha(2\beta-2)/(2\alpha+1)}\right).$$

Remark 7. *In analogy to (4.2), the estimator* $\hat{s}_{k,0}$ *can also be viewed as a spectral filter in Fourier domain, where essentially only the frequencies* $n^{1/2}, \ldots, 2n^{1/2}$ *play a role. For practical purposes one can generalize this to estimators where the frequencies* $k, \ldots, \lfloor cn^{1/2} \rfloor$, $c > 0$ *are used. If* σ *is assumed to be very smooth, one even may set* $k = 1$. *In this more general setting, the constant* $-7\pi^2/3$ *in the definition of the estimator has to be replaced by* $-n/\left(\lfloor cn^{1/2} \rfloor - k\right) \sum_{i=k}^{\lfloor cn^{1/2} \rfloor} \lambda_i$.

Remark 8. *Since the matrix* D *in the definition of* $\hat{s}_{k,0}$ *is a discrete sine transform (for a definition see [6]) the estimator* $\hat{\sigma}_N^2$ *can be calculated explicitly taking* $O(Nn \log n)$ *steps.*

5. Minimax

In this section we will discuss the optimality of the proposed estimators. To this end we establish lower bounds with respect to Sobolev s-ellipsoids and Gaussian microstructure noise.

Theorem 3. *Assume model (1.1) or model (1.2),* $\alpha \in \mathbb{N} \setminus \{0\}$. *Further assume* τ *constant. Then there exists a* $C > 0$ *(depending only on* α, Q, l, u*), such that*

$$\underline{\lim}_{n\to\infty} \inf_{\hat{\sigma}_n^2} \sup_{\sigma^2 \in \Theta_s^b(\alpha,Q)} \mathrm{E}\left(n^{\frac{\alpha}{2\alpha+1}} \left\|\hat{\sigma}_n^2 - \sigma^2\right\|_2^2\right) \geq C.$$

Proof. The proof relies on a multiple hypothesis testing argument and is almost the same as the proof given in [20], Theorem 2.1. However, the lower bounds there are established with respect to the space of Hölder continuous functions of index α on the interval $[0,1]$, i.e. for $l < u$

$$\mathcal{C}^b(\alpha, L) := \mathcal{C}^b(\alpha, L, [l,u]) := \Big\{ f : f^{(p)} \text{ exists for } p = [\alpha], \\ \left| f^{(p)}(x) - f^{(p)}(y) \right| \leq L |x-y|^{\alpha-p},\ \forall x, y \in I,\ 0 < l \leq f \leq u < \infty \Big\}.$$

Due to boundary effects $\mathcal{C}^b(\alpha, L) \subsetneq \Theta_s^b(\alpha, Q)$ and therefore the statement above does not follow immediately from [20], Theorem 2.1. In fact, the only difference to the proof of [20] is to show that the constructed functions $\sigma_{i,n}^2$ are also elements of $\Theta_s^b(\alpha, Q)$. To be more precise, write $\sigma_{\min}$, $\sigma_{\max}$ for the lower and upper bound of σ^2, respectively, i.e. $\sigma^2 \in \Theta_s^b(\alpha, Q, [\sigma_{\min}, \sigma_{\max}])$. Without loss

of generality, we may assume that $\sigma_{\min} = 1$. For the multiple hypothesis testing argument (cf. [29]) a specific choice of functions $\sigma^2_{i,n}$ is required. For a construction see [20], proof of Theorem 2.1 where $L := (2\pi^{2\alpha}Q)^{1/2}/\|K^{(\alpha)}\|_\infty$. As mentioned above, it remains to show

$$\sigma^2_{i,n} \in \Theta^b_s(\alpha, Q), \quad i = 0, 1, \ldots, M.$$

Due to the construction of $\sigma^2_{i,n}$, we have $\sigma^2_{i,n}(t) = 1$ for $t \in [0, 1/4] \cup [3/4, 1]$ and $\sigma^{2\ (l)}_{i,n}(0) = \sigma^{2\ (l)}_{i,n}(1) = 0$ for $l \in \{0, 1, \ldots, \alpha\}$. Thus, $\sigma^2_{i,n} \in \mathcal{W}(\alpha, L^2\|K^{(\alpha)}\|^2_\infty)$ (for a definition see Equation (B.1)), $\alpha \in \{1, 2, 3, \ldots\}$, $j = 0, \ldots, M$. Hence by Theorem B.1 it follows $\sigma^2_{i,n} \in \Theta_s(\alpha, Q)$ for $i = 0, \ldots, M$. □

6. Simulations

In this section we briefly illustrate the performance of our estimators. Our aim is not to give a comprehensive simulation study, rather we would like to illustrate the behaviour of the estimator when assumptions of Theorems 1 and 2 are violated. In the following we plotted our estimator to simulated data, where we always set $n = 25.000$. From the point of view of financial statistics this is approximately the sample size obtained over a trading day (6.5 hours) if log-returns are sampled at every second. For simplicity, we will choose N in (4.3) and (4.6) as the minimizer of $\|\hat\tau^2 - \tau^2\|^2_n$ and $\|\hat\sigma^2 - \sigma^2\|^2_n$, respectively, which is in practice unknown. Of course, proper selection of the threshold N^* is of major importance for the performance of the estimator. To this end various methods are available, among others, cross validation techniques, balancing principles, and variants thereof could be employed (see e.g. [10], [11], [21] and [22]). A thorough investigation is postponed to a separate paper. Throughout our simulations we assumed $\tau = 0.01$ and concentrated mainly on estimation of σ^2, as it is the more challenging task.

In Figure 2 we have displayed the estimator for $\sigma(t) = (2 + \cos(2\pi t))^{1/2}$. Note that by Definition 1, σ^2 has "infinite" smoothness, i.e. for any $\alpha > 0$, we can find a $Q < \infty$, such that $\sigma^2 \in \Theta_s(\alpha, Q)$. The reconstruction shows that estimation of τ^2 can be done much easier than estimation of σ^2 although it is of smaller magnitude. In Figure 3, we are interested in the behavior of the estimators if heavy-tailed microstructure noise is present. This was simulated by generating $\epsilon_{i,n} \sim 3^{-1/2}t(3)$, $i = 1, \ldots, n$, i.i.d., where $t(3)$ denotes a t-distribution with 3 degrees of freedom. We can see from Plot 1 in Figure 3 that the resulting microstructure noise has some severe outliers according to the tail x^{-4} of the density of $t(3)$. Nevertheless, estimation of τ^2 and σ^2 is not visibly affected by the distribution of the noise.

In the subsequent figures we illustrate the behaviour of the estimator when the required smoothness assumptions on σ^2 and τ^2 are violated. To this end, we investigate in Figure 4 the situation when σ is random itself, i.e. a realization from a Brownian motion, $\sigma(t) = 3|\tilde W_t|$. The Brownian motion $(\tilde W_t)_{t\in[0,1]}$ was modelled as independent from the Brownian motion in (1.1) and the microstructure noise process. It is of course not possible to reconstruct the complete path

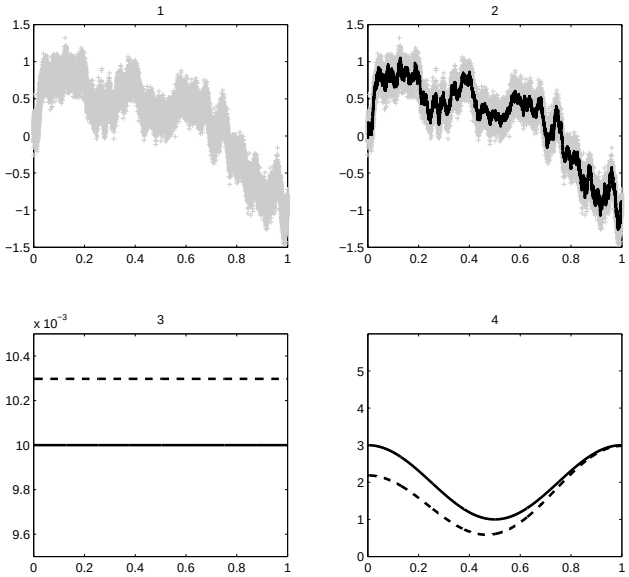

FIG 2. $n = 25000$ *data points from model (1.1),* $\epsilon_{i,n} \sim \mathcal{N}(0,1)$*, i.i.d.,* $\tau = 0.1$*,* $\sigma(t) = (2 + \cos(2\pi t))^{1/2}$*. Plot* 1 *shows the data. Additionally to the data, we plotted the path of the tBM in Plot* 2*. The reconstruction of* τ^2 *and* σ^2 *(dashed lines) as well as the true function (solid lines) are given in Plot* 3 *and* 4*, respectively. The threshold parameters were selected as* $N^* = 1$ *for estimation of* τ^2 *and* $N^* = 3$ *for estimation of* σ^2*.*

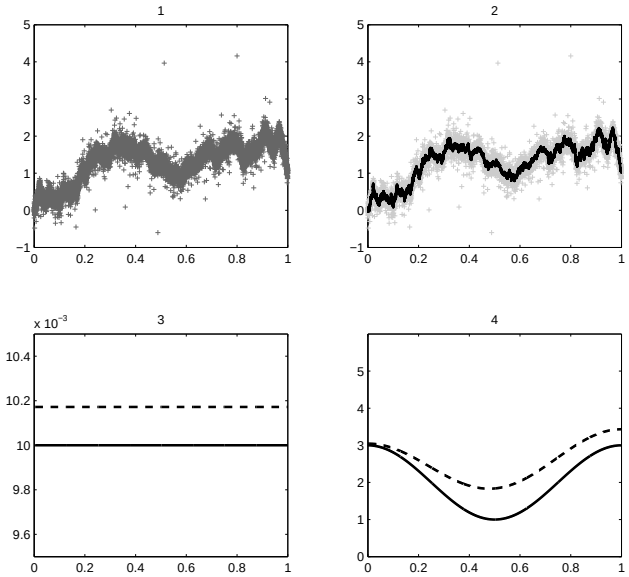

FIG 3. *(Heavy-tailed microstructure noise) As Figure 2 but instead of Gaussian errors we assumed that the noise follows a normalized Student's t-distribution with* 3 *degrees of freedom. We observe that performance of* $\hat{\tau}^2$ *and* $\hat{\sigma}^2$ *is quite robust to heavy-tailed noise. The threshold parameters* N^* *were selected as* 1 *and* 3 *for estimation of* τ^2 *and* σ^2*, respectively.*

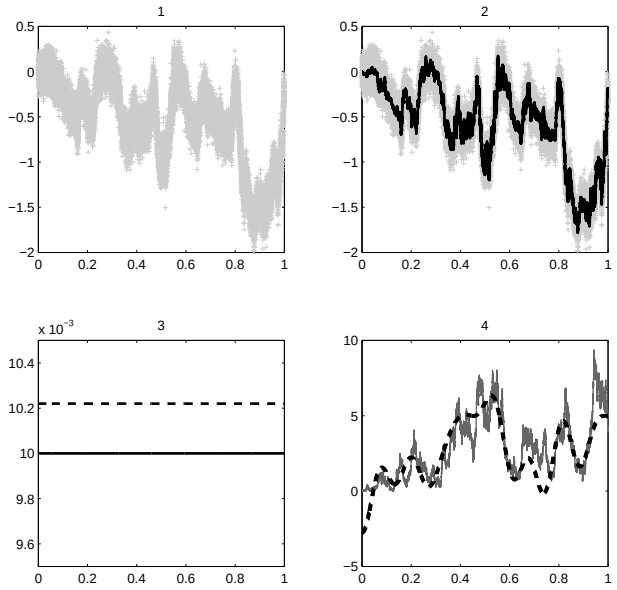

FIG 4. *(Low-smoothness) As Figure 2 but we chose* $\sigma(t) = 3|\tilde{W}_t|$, *where* $(\tilde{W}_t)_{t\in[0,1]}$ *denotes a Brownian motion independent of the noise and the Brownian motion in (1.1). The estimator returns a smoothed version of the path. The threshold parameters* N^* *were selected as* 1 *and* 17 *for estimation of* τ^2 *and* σ^2, *respectively.*

of σ^2, but as Figure 4 indicates, the estimators at least detects the smoothed shape of the path and so our estimator might already reveal some parts of the pattern of volatility also in case σ is non-deterministic, which is certainly more realistic in most applications.

Finally, in Figure 5 we investigated the case of σ being a jump-function. We put $\sigma(t) = 1+\mathbb{I}_{(1/2,1]}(t)$, a function with jump at $t = 1/2$. Fourier series usually show a Gibbs phenomenon, i.e. an oscillating behavior at discontinuities. This behavior is also clearly visible in the graph of $\hat{\sigma}^2$. In order to reconstruct jumps in volatility other methods certainly will be more suitable and are postponed to a separate paper.

Computational tasks: We implemented the estimators in Matlab using the routine fft() for the discrete sine transform (see Remark 8). Calculation of the estimators for a sample size of $n = 25.000$ took around 2-3 seconds on a Intel Celeron 1.7 GHz processor. As mentioned in Remark 8, the estimator can be calculated in $O(Nn\log n)$ steps. If we choose N with the optimal scale, i.e. $N \sim n^{1/(4\alpha+2)}$, we have for the complexity $O(Nn\log n) = o(n^{5/4}\log n)$, whenever $\alpha > 1/2$.

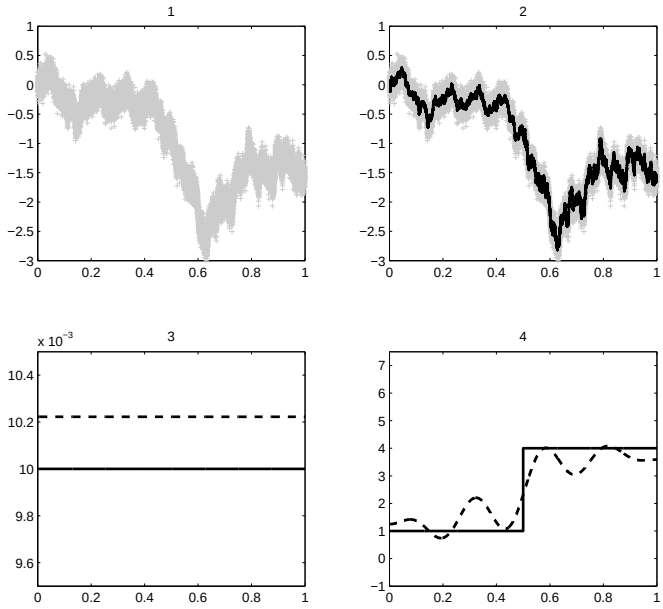

FIG 5. *(Jump function) As Figure 2 but we chose* $\sigma(t) = 1 + \mathbb{I}_{(1/2,1]}(t)$. *The Gibbs phenomenon is clearly visible. The threshold parameters* N^* *were selected as* 1 *and* 10 *for estimation of* τ^2 *and* σ^2, *respectively.*

Appendix A Convergence rate of $\hat{\sigma}^2$

In this section we will give a proof of Theorem 2. To this end we first introduce some notation and then prove a Lemma in order to get uniform estimates of bias and variance of the single estimators $\hat{s}_{k,0}$.

A.1 Preliminary results and notation

Proofs of the upper bounds are based on a decomposition of ΔY^k. In this subsection we present some further notation. Let $\sigma_k(t) := \sigma(t) f_k(t)$ and $\tau_k(t) := \tau(t) f_k(t)$, $t \in [0,1]$. Let throughout the following for the Sobolev s-ellipsoids in Definition 1 for σ^2 the constants being $l = \sigma_{\min}$ and $u = \sigma_{\max}$ and for τ^2, $l = \tau_{\min}$, $u = \tau_{\max}$. We define

$$\begin{aligned}\phi_n &:= \sup_{\sigma^2 \in \Theta_s^b(\alpha,Q)} \max_{i=1,\dots,n-1} \sup_{\xi \in [i/n,(i+1)/n]} \left| \sigma(\xi) - \sigma\left(\frac{i}{n}\right) \right|, \\ \bar{\phi}_n &:= \sup_{\tau^2 \in \Theta_s^b(\beta,\bar{Q})} \max_{k \le n^{1/4}} \max_{i=1,\dots,n} |\Delta_i \tau_k|. \end{aligned} \tag{A.1}$$

In order to do the proofs for model (1.1) and model (1.2) simultaneously, we first define the more general process

$$V_{k,l} := X_{1,k} + X_{2,k} + Z_{1,k,l} + Z_{2,k}, \quad l = 1, 2, \tag{A.2}$$

where $X_{1,k}, X_{2,k}, Z_{1,k,l}$ and $Z_{2,k}$ are $n-1$ dimensional random vectors with components

$$\begin{aligned}
(X_{1,k})_i &:= \sigma_k(i/n)\,\Delta_i W_{i/n}, \\
(X_{2,k})_i &:= \tau_k(i/n)\,\Delta_i \epsilon_{i,n}, \\
(Z_{1,k,1})_i &:= f_k(i/n) \int_{i/n}^{(i+1)/n} \left(\sigma(s) - \sigma\left(\frac{i}{n}\right)\right) dW_s, \\
(Z_{1,k,2})_i &:= f_k(i/n)\,(\Delta_i \sigma)\, W_{(i+1)/n}, \\
(Z_{2,k})_i &:= f_k(i/n)\,(\Delta_i \tau)\,\epsilon_{i+1,n}, \qquad i = 1, \ldots, n-1.
\end{aligned}$$

Obviously, $\Delta Y^k = V_{k,1}$ and $\Delta Y^k = V_{k,2}$ if model (1.1) and (1.2) holds, respectively. Define the generalized estimators $\hat{t}_{k,0,l} := V_{k,l}^t D J_n^\tau D^t V_{k,l}$ and $\hat{s}_{k,0,l} := V_{k,l}^t D J_n D^t V_{k,l} - 7\pi^2 \hat{t}_{k,0,l}/3$. Further there exists a decomposition with $C_{1,k,l}, C_{2,k} \in \mathbb{M}_{n-1,n}$ such that

$$V_{k,l} = C_{1,k,l}\xi + C_{2,k}\epsilon, \tag{A.3}$$

where $\epsilon = (\epsilon_{1,n}, \ldots, \epsilon_{n,n})^t$ and $\xi = \xi_n$ is standard n-variate normal, ϵ, ξ independent and $C_{1,k,l}\xi = X_{1,k} + Z_{1,k,l}$, $C_{2,k}\epsilon = X_{2,k} + Z_{2,k}$. Now, let

$$s_{k,p} := \int_0^1 \sigma_k^2(x)\cos(p\pi x)dx, \quad t_{k,p} := \int_0^1 \tau_k^2(x)\cos(p\pi x)dx \tag{A.4}$$

be the scaled p-th Fourier coefficients of the cosine series of σ_k^2 and τ_k^2, respectively. Define the sums $A(\sigma_k^2, r)$ by

$$A\left(\sigma_k^2, r\right) = \begin{cases} s_{k,0} + 2\sum_{m=1}^\infty s_{k,2nm} & \text{for } r \equiv 0 \bmod 2n, \\ \sum_{m=0}^\infty s_{k,2nm+n} & \text{for } r \equiv n \bmod 2n, \\ \sum_{q \equiv \pm r \bmod 2n,\ q \geq 0} s_{k,q} & \text{for } r \not\equiv 0 \bmod n, \end{cases}$$

and analogously $A(\tau_k^2, r)$ with $s_{k,p}$ replaced by $t_{k,p}$. Some properties of these variables are given in Lemma D.1 and Lemma D.2.

Further define

$$\Sigma_k := \begin{pmatrix} \sigma_k(1/n) & & \\ & \ddots & \\ & & \sigma_k(1-1/n) \end{pmatrix}. \tag{A.5}$$

We put $\mathrm{Cum}_4(\epsilon) := \mathrm{Cum}_4(\epsilon_{1,n})$ for the fourth cumulant of $\epsilon_{1,n}$. If X, Y are independent random vectors, we write $X \perp Y$.

A.2 Proofs for estimation of σ^2

Lemma A.1. *Let $\hat{s}_{k,0}$ be defined as in (4.5). Further assume $\beta > 1$, $Q, \bar{Q} > 0$, $0 < \sigma_{\min} \leq \sigma_{\max} < \infty$, $0 < \tau_{\min} \leq \tau_{\max} < \infty$ and $k = k_n \in \mathbb{N}$.*

(i) Assume model (1.1), $\alpha > 1/2$. Then it holds

$$\sup_{\sigma^2 \in \Theta_s^b(\alpha,Q),\, \tau^2 \in \Theta_s^b(\beta,\bar{Q})} \max_{k \leq n^{1/4}} \left|\mathrm{E}\left(\hat{s}_{k,0}\right) - s_{k,0}\right| = O\left(n^{-1/4} + n^{1-\beta} + n^{1/2-\alpha}\right), \tag{A.6}$$

$$\sup_{\sigma^2 \in \Theta_s^b(\alpha,Q),\, \tau^2 \in \Theta_s^b(\beta,\bar{Q})} \max_{k \leq n^{1/4}} \mathrm{Var}\left(\hat{s}_{k,0}\right) = O\left(n^{-1/2} + n^{4-4\beta}\right). \tag{A.7}$$

(ii) Assume model (1.2), $\alpha > 5/4$. Then it holds

$$\sup_{\sigma^2 \in \Theta_s^b(\alpha,Q),\, \tau^2 \in \Theta_s^b(\beta,\bar{Q})} \max_{k \leq n^{1/4}} \left|\mathrm{E}\left(\hat{s}_{k,0}\right) - s_{k,0}\right| = O\left(n^{1-\beta} + n^{5/2-2\alpha} + n^{-1/4}\right), \tag{A.8}$$

$$\sup_{\sigma^2 \in \Theta_s^b(\alpha,Q),\, \tau^2 \in \Theta_s^b(\beta,\bar{Q})} \max_{k \leq n^{1/4}} \mathrm{Var}\left(\hat{s}_{k,0}\right) = O\left(n^{-1/2} + n^{4-4\beta}\right). \tag{A.9}$$

Proof. The proof mainly uses the generalized estimators as introduced in Section A.1. It is clear that for two centered random vectors P and Q

$$\langle P, Q \rangle_\sigma := \mathrm{E}\left(P^t D J_n D Q\right)$$

defines a semi-inner product and by Lemma D.5, $P \perp Q \Rightarrow \langle P, Q \rangle_\sigma = 0$. Hence

$$\begin{aligned} \mathrm{E}\,\hat{s}_{k,0,l} &= \langle X_{1,k}, X_{1,k} \rangle_\sigma + \langle X_{2,k}, X_{2,k} \rangle_\sigma + \langle Z_{1,k,l}, Z_{1,k,l} \rangle_\sigma + \langle Z_{2,k}, Z_{2,k} \rangle_\sigma \\ &\quad + 2\langle X_{1,k}, Z_{1,k,l} \rangle_\sigma + 2\langle X_{2,k}, Z_{2,k} \rangle_\sigma - \frac{7\pi^2}{3} \mathrm{E}\left(\hat{t}_{k,0,l}\right). \end{aligned} \tag{A.10}$$

Clearly with (iii) in Lemma D.1 and $r_n := n^{-1/2}\left[n^{1/2}\right]$,

$$\begin{aligned} \langle X_{1,k}, X_{1,k} \rangle_\sigma &= \frac{1}{n} \mathrm{tr}\left(\Sigma_k D J_n D \Sigma_k\right) = \frac{1}{n} \mathrm{tr}\left(J_n D \Sigma_k^2 D\right) \\ &= n^{-1/2} \sum_{i=\left[n^{1/2}\right]+1}^{2\left[n^{1/2}\right]} \left(A\left(\sigma_k^2, 0\right) - A\left(\sigma_k^2, 2i\right)\right) \\ &= r_n A\left(\sigma_k^2, 0\right) - n^{-1/2} \sum_{i=\left[n^{1/2}\right]+1}^{2\left[n^{1/2}\right]} A\left(\sigma_k^2, 2i\right). \end{aligned}$$

Hence due to $r_n \leq 1$ and $|r_n - 1| \leq n^{-1/2}$

$$\left|\langle X_{1,k}, X_{1,k} \rangle_\sigma - s_{k,0}\right| \leq 2 \sum_{m=n}^{\infty} |s_{k,m}| + \frac{2}{\sqrt{n}} \sum_{i=0}^{\infty} |s_{k,i}|,$$

and with Lemma D.2

$$\sup_{\sigma^2\in\Theta_s^b(\alpha,Q)} \max_{k\leq n^{1/4}} \left|\langle X_{1,k}, X_{1,k}\rangle_\sigma - s_{k,0}\right| = O\left(n^{1/2-\alpha}\right).$$

Next we will bound $\langle X_{2,k}, X_{2,k}\rangle_\sigma$. In order to do this let $T_k \in \mathbb{D}_{n-1}$ with entries $(T_k)_{i,j} = \tau_k(i/n)\,\delta_{i,j}$. Further we define $\tilde{T}_k \in \mathbb{M}_{n-1}$

$$\left(\tilde{T}_k\right)_{i,j} = \begin{cases} (\Delta_i\tau_k)^2 & \text{for} \quad i=j-1, \\ (\Delta_j\tau_k)^2 & \text{for} \quad i=j+1, \\ 0 & \text{otherwise.} \end{cases} \tag{A.11}$$

Note the relation

$$\mathrm{Cov}(X_{2,k}) = T_kKT_k = 1/2T_k^2K + 1/2KT_k^2 + 1/2\tilde{T}_k. \tag{A.12}$$

Using Lemma D.5 yields

$$\begin{aligned}\langle X_{2,k}, X_{2,k}\rangle_\sigma &= \mathrm{E}\left(X_{2,k}^tDJ_nDX_{2,k}\right) = \mathrm{tr}\left(DJ_nDT_kKT_k\right) \\ &= \frac{1}{2}\mathrm{tr}\left(J_nDT_k^2KD\right) + \frac{1}{2}\mathrm{tr}\left(J_nDKT_k^2D\right) + \frac{1}{2}\mathrm{tr}\left(J_nD\tilde{T}_kD\right) \\ &= \mathrm{tr}\left(\Lambda J_nDT_k^2D\right) + \frac{1}{2}\mathrm{tr}\left(J_nD\tilde{T}_kD\right), \end{aligned} \tag{A.13}$$

and further

$$\begin{aligned}\mathrm{tr}\left(\Lambda J_nDT_k^2D\right) &= n^{1/2}\sum_{i=\lceil n^{1/2}\rceil+1}^{2\lceil n^{1/2}\rceil} \lambda_i\left(A\left(\tau_k^2,0\right) - A\left(\tau_k^2,2i\right)\right) \\ &= A\left(\tau_k^2,0\right)n^{1/2}\sum_{i=\lceil n^{1/2}\rceil+1}^{2\lceil n^{1/2}\rceil}\lambda_i - n^{1/2}\sum_{i=\lceil n^{1/2}\rceil+1}^{2\lceil n^{1/2}\rceil}\lambda_iA\left(\tau_k^2,2i\right). \end{aligned} \tag{A.14}$$

Because $\max_{i=\lceil n^{1/2}\rceil+1,\ldots,2\lceil n^{1/2}\rceil}\lambda_i = \lambda_{2\lceil n^{1/2}\rceil} \leq 4\pi^2n^{-1}$, it holds

$$\begin{aligned}\left|\sqrt{n}\sum_{i=\lceil n^{1/2}\rceil+1}^{2\lceil n^{1/2}\rceil}\lambda_iA\left(\tau_k^2,2i\right)\right| &\leq n^{1/2}\sum_{i=\lceil n^{1/2}\rceil+1}^{2\lceil n^{1/2}\rceil}\lambda_i\sum_{q\equiv\pm 2i \bmod 2n,\ q\geq 0}|t_{k,q}| \\ &\leq 4\pi^2n^{-1/2}\sum_{i=0}^{\infty}|t_{k,i}| \leq 8\pi^2n^{-1/2}\sum_{i=0}^{\infty}|t_{0,i}|. \end{aligned}$$

Therefore, (A.14) can be written as

$$\left|\mathrm{tr}\left(\Lambda J_nDT_k^2D\right) - t_{k,0}n^{1/2}\sum_{i=\lceil n^{1/2}\rceil+1}^{2\lceil n^{1/2}\rceil}\lambda_i\right| \leq 8\pi^2\sum_{m=n}^{\infty}|t_{k,m}| + 8\pi^2n^{-1/2}\sum_{i=0}^{\infty}|t_{0,i}|.$$

This gives by Lemma D.7 and Lemma D.2

$$\sup_{\tau^2\in\Theta_s^b\left(\beta,\bar{Q}\right)}\max_{k\leq n^{1/4}}\left|\operatorname{tr}\left(\Lambda J_n D T_k^2 D\right)-\frac{7\pi^2}{3}t_{k,0}\right|=O\left(n^{-1/2}\right).$$

Recall that $\operatorname{tr}(J_n)=O(n)$. It follows

$$\left|\operatorname{tr}\left(J_n D\tilde{T}_k D\right)\right|\leq \operatorname{tr}(J_n)\max_{i,j}\left|\left(D\tilde{T}_k D\right)_{i,j}\right|\leq 4\operatorname{tr}(J_n)\max_i\left(\Delta_i\tau_k\right)^2.$$

So,

$$\sup_{\tau^2\in\Theta_s^b\left(\beta,\bar{Q}\right)}\max_{k\leq n^{1/4}}\operatorname{tr}\left(J_n D\tilde{T}_k D\right)=O\left(n\bar{\phi}_n^2\right) \tag{A.15}$$

and therefore

$$\sup_{\tau^2\in\Theta_s^b\left(\beta,\bar{Q}\right)}\max_{k\leq n^{1/4}}\left|\langle X_{2,k},X_{2,k}\rangle_\sigma-\frac{7\pi^2}{3}t_{k,0}\right|=O\left(n^{-1/2}+n\bar{\phi}_n^2\right).$$

We bound the remaining terms of (A.10). Note

$$\langle Z_{1,k,1},Z_{1,k,1}\rangle_\sigma=\operatorname{tr}\left(DJ_nD\operatorname{Cov}\left(Z_{1,k,1}\right)\right)\leq\lambda_1\left(\operatorname{Cov}\left(Z_{1,k,1}\right)\right)\operatorname{tr}(J_n)\leq 2\phi_n^2$$

implying

$$\sup_{\sigma^2\in\Theta_s^b(\alpha,Q)}\max_{k\leq n^{1/4}}\langle Z_{1,k,1},Z_{1,k,1}\rangle_\sigma=O\left(\phi_n^2\right).$$

In order to bound $\langle Z_{1,k,2},Z_{1,k,2}\rangle_\sigma$ define

$$L:=\left(\frac{(i\wedge j)+1}{n}\right)_{i,j=1,\ldots,n-1} \tag{A.16}$$

and $\Delta\Sigma_k\in\mathbb{D}_{n-1}$ by

$$\left(\Delta\Sigma_k\right)_{i,j}:=f_k\left(\frac{i}{n}\right)\left(\Delta_i\sigma\right)\delta_{i,j}.$$

We obtain

$$\langle Z_{1,k,2},Z_{1,k,2}\rangle_\sigma=\operatorname{tr}\left(DJ_nD\left(\Delta\Sigma_k\right)L\left(\Delta\Sigma_k\right)\right)\leq 2n^{3/2}\phi_n^2 \tag{A.17}$$

and hence

$$\sup_{\sigma^2\in\Theta_s^b(\alpha,Q)}\max_{k\leq n^{1/4}}\langle Z_{1,k,2},Z_{1,k,2}\rangle_\sigma=O\left(n^{3/2}\phi_n^2\right). \tag{A.18}$$

Similarly to (A.16), $\langle Z_{2,k},Z_{2,k}\rangle_\sigma\leq\bar{\phi}_n^2 n$ and thus

$$\sup_{\tau^2\in\Theta_s^b\left(\beta,\bar{Q}\right)}\max_{k\leq n^{1/4}}\langle Z_{2,k},Z_{2,k}\rangle_\sigma=O\left(\bar{\phi}_n^2 n\right). \tag{A.19}$$

Note that

$$\mathrm{Cov}\left(X_{1,k}, Z_{1,k,2}\right)_{i,j} = \begin{cases} 0 & \text{for} \quad j < i \\ n^{-1} f_k\left(i/n\right) f_k\left(j/n\right) \sigma\left(i/n\right) \left(\Delta_j \sigma\right) & \text{for} \quad j \geq i \end{cases}.$$

Hence by Proposition D.1, we obtain

$$\sup_{\sigma^2 \in \Theta_s^b(\alpha,Q)} \max_{k \leq n^{1/4}} \left|\langle X_{1,k}, Z_{1,k,2}\rangle_\sigma\right| = O\left(n^{1/2} \log n \ \phi_n\right)$$

Applying the CS-inequality gives

$$\sup_{\sigma^2 \in \Theta_s^b(\alpha,Q)} \max_{k \leq n^{1/4}} \left|\langle X_{1,k}, Z_{1,k,1}\rangle_\sigma\right| = O\left(\phi_n\right),$$
$$\left|\langle X_{2,k}, Z_{2,k}\rangle_\sigma\right| \leq \langle X_{2,k}, X_{2,k}\rangle_\sigma^{1/2} \langle Z_{2,k}, Z_{2,k}\rangle_\sigma^{1/2}.$$

Using Proposition C.1 this yields (A.6) and (A.8). In order to give an upper bound for the variance of $\hat{s}_{k,0,l}$ note

$$\mathrm{Var}\left(\hat{s}_{k,0,l}\right) \leq 2\,\mathrm{Var}\left(V_{k,l}^t D J_n D V_{k,l}\right) + 2\frac{7^2\pi^4}{9}\,\mathrm{Var}\left(\hat{t}_{k,0,l}\right).$$

Furthermore we have using (A.3) and Lemma D.3 (vi)

$$\begin{aligned} V_{k,l}^t D J_n D V_{k,l} &= \xi^t C_{1,k,l}^t D J_n D C_{1,k,l}^t \xi + 2\xi^t C_{1,k,l}^t D J_n D C_{2,k} \epsilon + \epsilon^t C_{2,k}^t D J_n D C_{2,k} \epsilon \\ &\leq 2\xi^t C_{1,k,l}^t D J_n D C_{1,k,l}^t \xi + 2\epsilon^t C_{2,k}^t D J_n D C_{2,k} \epsilon. \end{aligned}$$

Hence

$$\mathrm{Var}\left(V_{k,l}^t D J_n D V_{k,l}\right) \leq 8\,\mathrm{Var}\left(\xi^t C_{1,k,l}^t D J_n D C_{1,k,l} \xi\right) + 8\,\mathrm{Var}\left(\epsilon^t C_{2,k}^t D J_n D C_{2,k} \epsilon\right).$$

Finally, we bound $\mathrm{Var}(\xi^t C_{1,k,l}^t D J_n D C_{1,k,l} \xi)$ and $\mathrm{Var}(\epsilon^t C_{2,k}^t D J_n D C_{2,k} \epsilon)$ in two steps, which will be denoted by (a) and (b).

(a) By Lemma D.4 (iii), we have

$$\begin{aligned} \mathrm{Var}\left(\xi^t C_{1,k,l}^t D J_n D C_{1,k,l} \xi\right) &= 2\left\|J_n^{1/2} D\,\mathrm{Cov}\left(X_{1,k} + Z_{1,k,l}\right) D J_n^{1/2}\right\|_F^2 \\ &\leq 8\left\|J_n^{1/2} D\left(\mathrm{Cov}\left(X_{1,k}\right) + \mathrm{Cov}\left(Z_{1,k,l}\right)\right) D J_n^{1/2}\right\|_F^2 \\ &\leq 16\left\|J_n^{1/2} D\,\mathrm{Cov}\left(X_{1,k}\right) D J_n^{1/2}\right\|_F^2 + 16\left\|J_n^{1/2} D\,\mathrm{Cov}\left(Z_{1,k,l}\right) D J_n^{1/2}\right\|_F^2. \end{aligned} \tag{A.20}$$

Firstly,

$$\begin{aligned} \left\|J_n^{1/2} D\,\mathrm{Cov}\left(Z_{1,k,1}\right) D J_n^{1/2}\right\|_F^2 &\leq \lambda_1^2\left(\mathrm{Cov}\left(Z_{1,k,1}\right)\right) \mathrm{tr}\left(J_n^2\right) \leq 4n^{-1/2}\phi_n^4, \\ \left\|J_n^{1/2} D\,\mathrm{Cov}\left(Z_{1,k,2}\right) D J_n^{1/2}\right\|_F^2 &\leq \lambda_1^2\left(D J_n D\right) \mathrm{tr}\left(\mathrm{Cov}\left(Z_{1,k,2}\right)^2\right) \\ &\leq 4n\phi_n^4 \left\|L\right\|_F^2 \leq 4n^3\phi_n^4, \end{aligned}$$

and also

$$\left\| J_n^{1/2} D \operatorname{Cov}(X_{1,k}) D J_n^{1/2} \right\|_F^2 \leq \lambda_1^2 (\operatorname{Cov}(X_{1,k})) \operatorname{tr}\left(J_n^2\right) \leq \sigma_{\max} n^{-1/2}.$$

Therefore,

$$\sup_{\sigma^2 \in \Theta_s^b(\alpha,Q)} \max_{k \leq n^{1/4}} \operatorname{Var}\left(\xi^t C_{1,k,l}^t D J_n D C_{1,k,l} \xi\right) = O\left(n^{-1/2}\right).$$

(b) Next, we see with the same arguments as in (A.20)

$$\begin{aligned}
\operatorname{Var}\left(\epsilon^t C_{2,k}^t D J_n D C_{2,k} \epsilon\right) &\leq (2 + \operatorname{Cum}_4(\epsilon)) \left\| J_n^{1/2} D \operatorname{Cov}(X_{2,k} + Z_{2,k}) D J_n^{1/2} \right\|_F^2 \\
&\leq 8 (2 + \operatorname{Cum}_4(\epsilon)) \left\| J_n^{1/2} D \operatorname{Cov}(X_{2,k}) D J_n^{1/2} \right\|_F^2 \\
&\quad + 8 (2 + \operatorname{Cum}_4(\epsilon)) \left\| J_n^{1/2} D \operatorname{Cov}(Z_{2,k}) D J_n^{1/2} \right\|_F^2.
\end{aligned}$$

We obtain

$$\left\| J_n^{1/2} D \operatorname{Cov}(Z_{2,k}) D J_n^{1/2} \right\|_F^2 \leq 4 \bar{\phi}_n^4 \operatorname{tr}\left(J_n^2\right) = 4 \bar{\phi}_n^4 n^{3/2}.$$

From (A.11) follows now

$$\left\| J_n^{1/2} D \operatorname{Cov}(X_{2,k}) D J_n^{1/2} \right\|_F^2 \leq \frac{3}{2} \left\| J_n^{1/2} D T_k^2 K D J_n^{1/2} \right\|_F^2 + \frac{3}{4} \left\| J_n^{1/2} D \tilde{T}_k D J_n^{1/2} \right\|_F^2.$$

Let I_{n-1} be the $n-1 \times n-1$ identity matrix. Due to $\Lambda J_n \Lambda \leq I_n \lambda_{2[n^{1/2}]}^2 n^{1/2}$ we have

$$\begin{aligned}
\left\| J_n^{1/2} D T_k^2 K D J_n^{1/2} \right\|_F^2 &= \operatorname{tr}\left(J_n^{1/2} D T_k^2 \Lambda J_n \Lambda T_k^2 D J_n^{1/2} \right) \\
&\leq \lambda_{2[n^{1/2}]}^2 n^{1/2} \operatorname{tr}\left(J_n^{1/2} D T_k^4 D J_n^{1/2} \right) \\
&\leq \lambda_{2[n^{1/2}]}^2 n^{1/2} \tau_{\max}^2 \operatorname{tr}(J_n).
\end{aligned}$$

Also

$$\left\| J_n^{1/2} D \tilde{T}_k D J_n^{1/2} \right\|_F^2 \leq \lambda_1^2(J_n) \left\| \tilde{T}_k \right\|_F^2 \leq 2 n^2 \bar{\phi}_n^4$$

and therefore

$$\sup_{\tau^2 \in \Theta_s^b(\beta,\bar{Q})} \max_{k \leq n^{1/4}} \operatorname{Var}\left(\epsilon^t C_{2,k}^t D J_n D C_{2,k} \epsilon\right) = O\left(n^{-1/2} + n^2 \bar{\phi}_n^4\right).$$

Combining (a) and (b) gives

$$\sup_{\sigma^2 \in \Theta_s^b(\alpha,Q),\ \tau^2 \in \Theta_s^b(\beta,\bar{Q})} \max_{k \leq n^{1/4}} \operatorname{Var}\left(V_{k,l}^t D J_n D V_{k,l}\right) = O\left(n^{-1/2} + n^2 \bar{\phi}_n^4\right),$$

so (A.7) and (A.9) follow with Lemma C.1 and Propositon C.1. □

Proof of Theorem 2. We decompose

$$\mathrm{MISE}\left(\hat{\sigma}_N^2\right) = \int_0^1 \mathrm{Bias}^2\left(\hat{\sigma}_N^2(t)\right) dt + \int_0^1 \mathrm{Var}\left(\hat{\sigma}_N^2(t)\right) dt.$$

We have that $\sigma^2(t) = \sum_{i=0}^{\infty} \left\langle \psi_k, \sigma^2 \right\rangle \psi_k(t)$, where $\langle \, . \, , \, . \, \rangle$ denotes the standard scalar product on $L^2[0,1]$. Let $\eta_{k,n,l} := \mathrm{E}\left(\hat{s}_{k,0,l}\right) - s_{k,0}$. Then for $i \geq 1$, $\mathrm{E}\left(2\hat{s}_{i,0,l} - 2\hat{s}_{0,0,l}\right) = 2^{1/2}\left\langle \psi_i, \sigma^2 \right\rangle + 2\eta_{i,n,l} - 2\eta_{0,n,l}$. Hence

$$\int_0^1 \mathrm{Bias}^2\left(\hat{\sigma}^2(t)\right) dt \quad = \quad \eta_{0,n,l}^2 + 2\sum_{i=1}^{N}\left(\eta_{i,n,l} - \eta_{0,n,l}\right)^2 + \sum_{i=N+1}^{\infty}\left\langle \psi_i, \sigma^2 \right\rangle^2$$

and we obtain

$$\eta_{0,n,l}^2 + 2\sum_{i=1}^{N}\left(\eta_{i,n,l} - \eta_{0,n,l}\right)^2 \quad \leq \quad (8N+1)\max_{i=0,\ldots,N}\eta_{i,n,l}^2.$$

Because $\sigma^2 \in \Theta_s(\alpha, Q)$ it holds

$$\sum_{i=N+1}^{\infty}\left\langle \psi_i, \sigma^2 \right\rangle^2 \leq \frac{1}{(N+1)^{2\alpha}}\sum_{i=1}^{\infty} i^{2\alpha}\left\langle \psi_i, \sigma^2 \right\rangle^2 \leq 2Q(N+1)^{-2\alpha}. \qquad \text{(A.21)}$$

Therefore,

$$\sup_{\sigma^2 \in \Theta_s^b(\alpha,Q),\ \tau^2 \in \Theta_s^b(\beta,\bar{Q})} \int_0^1 \mathrm{Bias}^2\left(\hat{\sigma}^2(t)\right) dt$$
$$= O\left(N \sup_{\sigma^2 \in \Theta_s^b(\alpha,Q),\ \tau^2 \in \Theta_s^b(\beta,\bar{Q})} \max_{i=0,\ldots,N} \eta_{i,n,l}^2 + N^{-2\alpha}\right).$$

Assume model (1.1), then by Lemma A.1

$$\sup_{\sigma^2 \in \Theta_s^b(\alpha,Q), \tau^2 \in \Theta_s^b(\beta,\bar{Q})} \int_0^1 \mathrm{Bias}^2\left(\hat{\sigma}^2(t)\right) dt$$
$$= O\left(Nn^{-1/2} + Nn^{2-2\beta} + Nn^{1-2\alpha} + N^{-2\alpha}\right).$$

and for model (1.2),

$$\sup_{\sigma^2 \in \Theta_s^b(\alpha,Q),\ \tau^2 \in \Theta_s^b(\beta,\bar{Q})} \int_0^1 \mathrm{Bias}^2\left(\hat{\sigma}^2(t)\right) dt$$
$$= O\left(Nn^{2-2\beta} + Nn^{5-4\alpha} + Nn^{-1/2} + N^{-2\alpha}\right).$$

For the variance note

$$\int_0^1 \operatorname{Var}\left(\hat{\sigma}^2(t)\right) dt = \operatorname{Var}\left(\hat{s}_{0,0,l}\right) + \frac{1}{2}\sum_{i=1}^{N} \operatorname{Var}\left(2\hat{s}_{i,0,l} - 2\hat{s}_{0,0,l}\right)$$
$$\leq (4N+1)\operatorname{Var}\left(\hat{s}_{0,0,l}\right) + 4\sum_{i=1}^{N} \operatorname{Var}\left(\hat{s}_{i,0,l}\right).$$

Hence

$$\sup_{\sigma^2 \in \Theta_s^b(\alpha,Q),\ \tau^2 \in \Theta_s^b\left(\beta,\bar{Q}\right)} \int_0^1 \operatorname{Var}\left(\hat{\sigma}^2(t)\right) dt$$
$$= O\left(N \sup_{\sigma^2 \in \Theta_s^b(\alpha,Q), \tau^2 \in \Theta_s^b\left(\beta,\bar{Q}\right)} \max_{0 \leq k \leq N} \operatorname{Var}\left(\hat{s}_{k,0,l}\right)\right).$$

Using Lemma A.1 yields the result. □

Appendix B Sobolev s-ellipsoids

In this chapter we will shortly discuss the function space introduced in Section 3 and provide a theorem needed for the lower bound. First recall the classical definition of Sobolev ellipsoids (cf. Proposition 1.14 in [29]).

Definition B.1. *Define*

$$a_{j,\alpha} = \begin{cases} j^{\alpha}, & \text{for even } j, \\ (j-1)^{\alpha}, & \text{for odd } j \end{cases}.$$

Let $\{\varphi_j\}_{j=1}^{\infty}$, $\varphi_1(x) := 1$, $\varphi_{2j}(x) := \sqrt{2}\cos(2\pi j x)$, $\varphi_{2j+1}(x) := \sqrt{2}\sin(2\pi j x)$ *denote the trigonometric basis on* $[0,1]$. *Then we call the function space*

$$\Theta(\alpha, C) := \left\{ f \in L^2[0,1] : \exists\, (\theta_n)_{n=1}^{\infty},\ s.\ t.\ f(x) = \sum_{i=1}^{\infty} \theta_i \varphi_i(x), \sum_{i=1}^{\infty} a_{i,\alpha}^2 \theta_i^2 \leq C \right\}$$

a Sobolev ellipsoid.

Interesting characterizations arise if we put Sobolev s-ellipsoids into relation with Sobolev ellipsoids:

Remark B.1. *Let* $\mathcal{S}$ *be the class of all functions* $f \in L^2[0,1]$ *such that* $f(x) = f(1-x)$, $\forall x \in [0,1]$. *Further let* $\Theta(\alpha, C)$ *be a Sobolev ellipsoid as in Definition B.1. Then careful calculations show that a function belongs to* $\Theta_s(\alpha, C) \cap \mathcal{S}$ *if and only if it belongs to* $\Theta(\alpha, C2^{1-2\alpha}) \cap \mathcal{S}$.

Let

$$\mathcal{W}(\alpha,\bar{C}) :=\mathcal{W}(\alpha,\bar{C},[0,1]) := \Big\{ f \in L^2[0,1] : f^{(l)}(0) = f^{(l)}(1) = 0 \\ \text{for } l \text{ odd, } l < \alpha, \int_0^1 \left(f^{(\alpha)}(x)\right)^2 dx \leq \bar{C} \Big\}. \tag{B.1}$$

For positive integer values of α, we have the following equivalence.

Theorem B.1. *Assume $\alpha \in \{1,2,\ldots\}$, $C>0$. Let $\bar{C} = 2\pi^{2\alpha}C$. Then a function is in $\mathcal{W}(\alpha,\bar{C})$ if and only if it is in $\Theta_s(\alpha,C)$.*

Proof. First we show that if a function $f \in \mathcal{W}(\alpha,\bar{C})$ then also $f \in \Theta_s(\alpha,C)$. Let $\tilde{f}$ be defined on $[-1,1]$ by

$$\tilde{f}(x) := \begin{cases} f(x) & \text{for} \quad x \in [0,1] \\ f(-x) & \text{for} \quad x \in [-1,0] \end{cases}.$$

Note that $\tilde{f}$ is an α-times differentiable function, $\tilde{f}^{(l)}$ is even if l is even and $\tilde{f}^{(l)}$ is odd if l is odd. Let

$$s_k(j) = \begin{cases} \int_{-1}^1 \tilde{f}^{(j)}(x)dx & \text{for} \quad k=0 \\ \int_{-1}^1 \tilde{f}^{(j)}(x)\cos(k\pi x)dx & \text{for} \quad k \geq 1, j \text{ even} \\ \int_{-1}^1 \tilde{f}^{(j)}(x)\sin(k\pi x)dx & \text{for} \quad k \geq 1, j \text{ odd} \end{cases}.$$

It holds for $j \geq 1$

$$s_0(j) = \int_{-1}^1 \tilde{f}^{(j)}(x)dx = \tilde{f}^{(j-1)}(1) - \tilde{f}^{(j-1)}(-1) = 0.$$

Hence we have the Parseval type equality

$$\int_0^1 \left(f^{(\alpha)}(x)\right)^2 dx = \frac{1}{2}\sum_{k=1}^{\infty} s_k^2(\alpha). \tag{B.2}$$

Further for $k \geq 1$, j even, it follows by partial integration

$$\begin{aligned} s_k(j) &= \int_{-1}^1 \tilde{f}^{(j)}(x)\cos(k\pi x)dx \\ &= \tilde{f}^{(j-1)}(x)\cos(k\pi x)\Big|_{-1}^1 + k\pi \int_{-1}^1 \tilde{f}^{(j-1)}(x)\sin(k\pi x)dx = k\pi s_k(j-1) \end{aligned}$$

and for $k \geq 1$ and j odd

$$\begin{aligned} s_k(j) &= \int_{-1}^1 \tilde{f}^{(j)}(x)\sin(k\pi x)dx \\ &= \tilde{f}^{(j-1)}(x)\sin(k\pi x)\Big|_{-1}^1 - k\pi \int_{-1}^1 \tilde{f}^{(j-1)}(x)\cos(k\pi x)dx = -k\pi s_k(j-1). \end{aligned}$$

With $\theta_k = \int_0^1 f(x)\cos(k\pi x)dx$ it follows for $k \geq 1$ $s_k^2(\alpha) = k^{2\alpha}\pi^{2\alpha}s_k^2 = 4k^{2\alpha}\pi^{2\alpha}\theta_k^2$. Combining this result with (B.2) yields

$$\int_0^1 \left(f^{(\alpha)}(x)\right)^2 dx = 2\pi^{2\alpha}\sum_{k=1}^{\infty} k^{2\alpha}\theta_k^2$$

and hence proves the first part of the theorem. The other direction follows in a straightforward way by differentiation and is thus omitted. □

Appendix C Convergence rate of $\hat{\tau}^2$

C.1 Preliminary results and notation

First we recall some notation. Let $\sigma_k(i/n) := \sigma(i/n)f_k(i/n)$ and $\tau_k(i/n) := \tau(i/n)f_k(i/n)$. Let throughout the following for the Sobolev s-ellipsoids in Definition 1 for σ^2 the constants being $l = \sigma_{\min}$ and $u = \sigma_{\max}$ and for τ^2, $l = \tau_{\min}$, $u = \tau_{\max}$. We define $K_n := n^{1/2}/\log n$ and

$$\phi_n := \sup_{\sigma^2\in\Theta_s^b(\alpha,Q)} \max_{i=1,\ldots,n-1} \sup_{\xi\in[i/n,i+1/n]} \left|\sigma(\xi) - \sigma\left(\frac{i}{n}\right)\right|,$$

$$\bar{\phi}_{n,1/2} := \sup_{\tau^2\in\Theta_s^b(\beta,\bar{Q})} \max_{k\leq K_n} \max_{i=1,\ldots,n} \sup_{\xi_i\in[(i-1)/n,i/n]} \left|\tau_k(\xi_i) - \tau_k\left(\frac{i}{n}\right)\right|.$$

Proposition C.1. *Assume $\alpha, \beta > 1/2$. It holds for any $\delta > 0$,*

$$\phi_n = O\left(n^{1/2-\alpha} + n^{\delta-1}\right)$$
$$\bar{\phi}_n = O\left(n^{1/2-\beta} + n^{-3/4}\right)$$
$$\bar{\phi}_{n,1/2} = O\left(n^{1/2-\beta} + n^{-1/2}\log^{-1} n\right).$$

Proof. We only prove the third equality the other two can be deduced similarly. Note that for $\tau^2 \in \Theta^b(\beta, Q)$,

$$\begin{aligned}\left|\tau_k(\xi_i) - \tau_k\left(\frac{i}{n}\right)\right| &\leq \sqrt{2}\left|\tau(\xi_i) - \tau\left(\frac{i}{n}\right)\right| + \tau_{\max}^{1/2}k\pi/(2n)\\ &\leq \frac{1}{\sqrt{2}\tau_{\min}}\left|\tau^2(\xi_i) - \tau^2\left(\frac{i}{n}\right)\right| + \tau_{\max}^{1/2}k\pi/(2n).\end{aligned}$$

Taking supremum and applying Lemma D.8 gives the result. □

C.2 Proofs for estimation of τ^2

Lemma C.1. *Let $\hat{t}_{k,0}$ be defined as in (4.1). Further assume $\alpha, \beta > 1/2$, $Q, \bar{Q} > 0$, $0 < \sigma_{\min} \leq \sigma_{\max} < \infty$, $0 < \tau_{\min} \leq \tau_{\max} < \infty$ and $k = k_n \in \mathbb{N}$. Assume model*

(1.1). Then

$$\sup_{\sigma^2\in\Theta_s^b(\alpha,Q),\tau^2\in\Theta_s^b(\beta,\bar{Q})}\max_{k\le K_n}\left|\mathrm{E}\,\hat{t}_{k,0}-t_{k,0}\right| = O\left(n^{1/2-\beta}\log^{1/2}n\right)+o\left(n^{-1/2}\right), \tag{C.1}$$

$$\sup_{\sigma^2\in\Theta_s^b(\alpha,Q),\ \tau^2\in\Theta_s^b(\beta,\bar{Q})}\max_{k\le K_n}\mathrm{Var}\left(\hat{t}_{k,0}\right) = O\left(n^{-1}\right). \tag{C.2}$$

If further ϵ is n-variate standard normal, then

$$\sup_{\sigma^2\in\Theta_s^b(\alpha,Q),\ \tau^2\in\Theta_s^b(\beta,\bar{Q})}\max_{k\le K_n}\left|\mathrm{Var}\left(\hat{t}_{k,0}\right)-\frac{2}{n}\int_0^1\tau_k^4(x)dx\right| = O\left(n^{-1}\log^{-1}n\right), \tag{C.3}$$

$$n^{1/2}\left(\hat{t}_{k,0}-t_{k,0}\right)\xrightarrow{L}\mathcal{N}\left(0,2\int_0^1\tau_k^4(x)dx\right)\quad \textit{for } \beta>1,\ k\le K_n. \tag{C.4}$$

Assume model (1.2). Then it holds

$$\sup_{\sigma^2\in\Theta_s^b(\alpha,Q),\ \tau^2\in\Theta_s^b(\beta,\bar{Q})}\max_{k\le K_n}\left|\mathrm{E}\,\hat{t}_{k,0}-t_{k,0}\right| = O\left(n^{1/2-\beta}\log^{1/2}n+n^{1-2\alpha}\log^2 n\right)+o\left(n^{-1/2}\right), \tag{C.5}$$

$$\sup_{\sigma^2\in\Theta_s^b(\alpha,Q),\ \tau^2\in\Theta_s^b(\beta,\bar{Q})}\max_{k\le K_n}\mathrm{Var}\left(\hat{t}_{k,0}\right) = O\left(n^{2-4\alpha}\log^4 n+n^{-1}\right). \tag{C.6}$$

If further ϵ is n-variate standard normal, then

$$\sup_{\sigma^2\in\Theta_s^b(\alpha,Q),\tau^2\in\Theta_s^b(\beta,\bar{Q})}\max_{k\le K_n}\left|\mathrm{Var}\left(\hat{t}_{k,0}\right)-\frac{2}{n}\int_0^1\tau_k^4(x)dx\right| = O\left(n^{2-4\alpha}\log^4 n+n^{-1}\log^{-1}n\right), \tag{C.7}$$

$$n^{1/2}\left(\hat{t}_{k,0}-t_{k,0}\right)\xrightarrow{L}\mathcal{N}\left(0,2\int_0^1\tau_k^4(x)dx\right)\quad \textit{for } \alpha>3/4,\ \beta>1,\ k\le K_n. \tag{C.8}$$

Proof. Again we work with the generalized estimators as introduced in Section A.1. As in the proof of Lemma A.1 we introduce for two centered random vectors P and Q a semi-inner product defined by $\langle P,Q\rangle_\tau := \mathrm{E}\left(P^tDJ_n^\tau D^tQ\right)$ and obtain

$$\begin{aligned}\mathrm{E}\,\hat{t}_{k,0,l} &= \langle X_{1,k},X_{1,k}\rangle_\tau+\langle X_{2,k},X_{2,k}\rangle_\tau+\langle Z_{1,k,l},Z_{1,k,l}\rangle_\tau+\langle Z_{2,k},Z_{2,k}\rangle_\tau\\ &\quad+2\langle X_{1,k},Z_{1,k,l}\rangle_\tau+2\langle X_{2,k},Z_{2,k}\rangle_\tau.\end{aligned} \tag{C.9}$$

First we bound $\langle X_{2,k},X_{2,k}\rangle_\tau$, which will turn out to be the leading term. Similar to (A.13) we have

$$\langle X_{2,k},X_{2,k}\rangle_\tau = \mathrm{tr}\left(\Lambda J_n^\tau D^tT_k^2D\right)+\frac{1}{2}\mathrm{tr}\left(J_n^\tau D^t\tilde{T}_kD\right),$$

and due to

$$\operatorname{tr}\left(J_n^\tau\right) = O\left(\log n\right) \tag{C.10}$$

the same argument as for (A.15) gives

$$\sup_{\tau^2\in\Theta_s^b\left(\beta,\bar{Q}\right)} \max_{k\le K_n} \operatorname{tr}\left(J_n^\tau D\tilde{T}_k D\right) = O\left(\bar{\phi}_{n,1/2}^2 \log n\right).$$

Hence this is a negligible term. Using Lemma D.1 (iii)

$$\begin{aligned}\operatorname{tr}\left(\Lambda J_n^\tau D^t T_k^2 D\right) &= \left(n - n/\log n\right)^{-1} \sum_{i=[n/\log n]}^{n-1} \left(A\left(\tau_k^2, 0\right) - A\left(\tau_k^2, 2i\right)\right) \\ &= \bar{r}_n A\left(\tau_k^2, 0\right) - \left(n - n/\log n\right)^{-1} \sum_{i=[n/\log n]}^{n-1} A\left(\tau_k^2, 2i\right),\end{aligned}$$

where $\bar{r}_n = \left(n - [n/\log n]\right)/\left(n - n/\log n\right)$. Note $1 - \bar{r}_n \le 1/\left(n - n/\log n\right)$. By Lemma D.2

$$\begin{aligned}&\sup_{\tau^2\in\Theta_s^b\left(\beta,\bar{Q}\right)} \max_{k\le K_n} \left|\operatorname{tr}\left(\Lambda J_n^\tau D^t T_k^2 D\right) - t_{k,0}\right| \\ &\le \sup_{\tau^2\in\Theta_s^b\left(\beta,\bar{Q}\right)} \max_{k\le K_n} \left(\left(1-\bar{r}_n\right)\left|t_{k,0}\right| + \sum_{m=n}^{\infty}\left|t_{k,m}\right| + 2\left(n - n/\log n\right)^{-1}\sum_{i=0}^{\infty}\left|t_{k,i}\right|\right) \\ &\le 2C_{\beta,Q}n^{1/2-\beta} + 6\left(n - n/\log n\right)^{-1} \sup_{\tau^2\in\Theta_s^b\left(\beta,\bar{Q}\right)} \sum_{i=0}^{\infty}\left|t_{0,i}\right| = O\left(n^{-1} + n^{1/2-\beta}\right).\end{aligned}$$

This shows that

$$\sup_{\tau^2\in\Theta_s^b\left(\beta,\bar{Q}\right)} \max_{k\le K_n} \left|\left\langle X_{2,k}, X_{2,k}\right\rangle_\tau - t_{0,k}\right| = O\left(n^{-1} + \bar{\phi}_{n,1/2}^2 \log n + n^{1/2-\beta}\right).$$

The remaining part of the proofs of (C.5) and (C.1) is concerned with bounding the other expressions in (C.9). Applying Lemma D.3, we obtain

$$\left\langle X_{1,k}, X_{1,k}\right\rangle_\tau = \frac{1}{n}\operatorname{tr}\left(D J_n^\tau D^t \Sigma_k^2\right) \le 2\sigma_{\max}\frac{1}{n}\operatorname{tr}\left(J_n^\tau\right)$$

implying

$$\sup_{\sigma^2\in\Theta_s^b(\alpha,Q)} \max_{k\le K_n} \left\langle X_{1,k}, X_{1,k}\right\rangle_\tau = O\left(n^{-1}\log n\right).$$

We obtain with Lemma D.6 in the same way as in (A.16), (A.18) and (A.19)

$$\begin{aligned}\sup_{\sigma^2\in\Theta_s^b(\alpha,Q)} \max_{k\le K_n} \left\langle Z_{1,k,1}, Z_{1,k,1}\right\rangle_\tau &= O\left(n^{-1}\log n\ \phi_n^2\right), \\ \sup_{\sigma^2\in\Theta_s^b(\alpha,Q)} \max_{k\le K_n} \left\langle Z_{1,k,2}, Z_{1,k,2}\right\rangle_\tau &= O\left(\log^2 n\ \phi_n^2\right), \\ \sup_{\tau^2\in\Theta_s^b\left(\beta,\bar{Q}\right)} \max_{k\le K_n} \left\langle Z_{2,k}, Z_{2,k}\right\rangle_\tau &= O\left(\bar{\phi}_{n,1/2}^2 \log n\right).\end{aligned}$$

From the Cauchy-Schwarz inequality follows that

$$\begin{aligned}\left|\langle X_{1,k}, Z_{1,k,l}\rangle_\tau\right| &\leq \langle X_{1,k}, X_{1,k}\rangle_\tau^{1/2} \langle Z_{1,k,l}, Z_{1,k,l}\rangle_\tau^{1/2} \\ &\leq \langle X_{1,k}, X_{1,k}\rangle_\tau + \langle Z_{1,k,l}, Z_{1,k,l}\rangle_\tau ,\end{aligned}$$

$$\left|\langle X_{2,k}, Z_{2,k}\rangle_\tau\right| \leq \langle X_{2,k}, X_{2,k}\rangle_\tau^{1/2} \langle Z_{2,k}, Z_{2,k}\rangle_\tau^{1/2} .$$

This yields

$$\begin{aligned}&\sup_{\sigma^2\in\Theta_s^b(\alpha,Q),\ \tau^2\in\Theta_s^b(\beta,\bar{Q})} \max_{k\leq K_n} \left|E\hat{t}_{k,0} - t_{k,0}\right| \\ &\quad= \begin{cases} O\left(n^{-1}\log n + n^{1/2-\beta} + \bar{\phi}_{n,1/2}\log^{1/2} n\right) & \text{for } l=1, \\ O\left(n^{-1}\log n + n^{1/2-\beta} + \bar{\phi}_{n,1/2}\log^{1/2} n + \phi_n^2\log^2 n\right) & \text{for } l=2, \end{cases}\end{aligned}$$

and therefore (C.5) and (C.1) holds by Proposition C.1. In order to calculate the covariance we use the decomposition (A.3). We have

$$\hat{t}_{k,0,l} = \xi^t C_{1,k,l}^t D J_n^\tau D C_{1,k,l}\xi + 2\xi^t C_{1,k,l}^t D J_n^\tau D C_{2,k}\epsilon + \epsilon^t C_{2,k}^t D J_n^\tau D C_{2,k}\epsilon.$$

Using the CS-inequality repeatedly, we can write

$$\begin{aligned}&\left|\mathrm{Var}\left(\hat{t}_{k,0,l}\right) - \mathrm{Var}\left(\epsilon^t C_{2,k}^t D J_n^\tau D C_{2,k}\epsilon\right)\right| \qquad \text{(C.11)} \\ &\quad\leq \left(\mathrm{Var}^{1/2}\left(\xi^t C_{1,k,l}^t D J_n^\tau D C_{1,k,l}\xi\right) + 2\,\mathrm{Var}^{1/2}\left(\xi^t C_{1,k,l}^t D J_n^\tau D C_{2,k}\epsilon\right)\right)^2 \\ &\qquad + \left(\mathrm{Var}^{1/2}\left(\xi^t C_{1,k,l}^t D J_n^\tau D C_{1,k,l}\xi\right) + 2\,\mathrm{Var}^{1/2}\left(\xi^t C_{1,k,l}^t D J_n^\tau D C_{2,k}\epsilon\right)\right) \\ &\qquad \cdot 2\,\mathrm{Var}^{1/2}\left(\epsilon^t C_{2,k}^t D J_n^\tau D C_{2,k}\epsilon\right)\end{aligned}$$

We subdivide the remaining part of the proofs of (C.6) and (C.2) into three steps (a), (b) and (c), where we calculate $\mathrm{Var}(\epsilon^t C_{2,k}^t D J_n^\tau D C_{2,k}\epsilon)$, $\mathrm{Var}(\xi^t C_{1,k,l}^t D J_n^\tau D C_{1,k,l}\xi)$ and $\mathrm{Var}(\xi^t C_{1,k,l}^t D J_n^\tau D C_{2,k}\epsilon)$, respectively.

(a) Let $\mathrm{TrSq}(A) := \sum_{i=1}^n A_{i,i}^2$ for $A\in\mathbb{M}_n$. Then by Lemma D.5 it follows

$$\begin{aligned}&\mathrm{Var}\left(\epsilon^t C_{2,k}^t D J_n^\tau D C_{2,k}\epsilon\right) \\ &\quad= 2\left\|C_{2,k}^t D J_n^\tau D C_{2,k}\right\|_F^2 + \mathrm{Cum}_4(\epsilon)\,\mathrm{TrSq}\left(C_{2,k}^t D J_n^\tau D C_{2,k}\right) \\ &\quad\leq (2+\mathrm{Cum}_4(\epsilon))\left\|(J_n^\tau)^{1/2} D\,\mathrm{Cov}\left(X_{2,k}+Z_{2,k}\right) D\,(J_n^\tau)^{1/2}\right\|_F^2 ,\end{aligned}$$

where equality holds if $\mathrm{Cum}_4(\epsilon)=0$. By Proposition D.2 we see that

$$\begin{aligned}&\sup_{\tau^2\in\Theta_s^b(\beta,\bar{Q})} \max_{k\leq K_n} \left|\mathrm{Var}\left(\epsilon^t C_{2,k}^t D J_n^\tau D C_{2,k}\epsilon\right) - \frac{2}{n}\int_0^1 \tau_k^4(x)dx\right| \\ &\qquad= O\left(\mathrm{Cum}_4(\epsilon)\,n^{-1} + n^{-1}\log^{-1} n\right).\end{aligned}$$

(b) In this part of the proof we will bound $\mathrm{Var}\left(\xi^t C^t_{1,k,l} D J^\tau_n D C_{1,k,l}\xi\right)$. Similar to part (a) in the proof of Lemma A.1 it holds

$$\begin{aligned}\mathrm{Var}\left(\xi^t C^t_{1,k,l} D J^\tau_n D C_{1,k,l}\xi\right) &\le 16\lambda_1^2\left(J^\tau_n\right)\left(\|\mathrm{Cov}\left(X_{1,k}\right)\|_F^2 + \|\mathrm{Cov}\left(Z_{1,k,l}\right)\|_F^2\right)\\ &\le \begin{cases} 2^8 n^{-2}\log^4 n\left(n^{-1}\sigma^2_{\max} + 4n^{-1}\phi_n^4\right), & l=1,\\ 2^8 n^{-2}\log^4 n\left(n^{-1}\sigma^2_{\max} + 4n^{2}\phi_n^4\right), & l=2,\end{cases}\end{aligned}$$

where we used Lemma D.6 in the second inequality. Hence we get

$$\sup_{\sigma^2\in\Theta^b_s(\alpha,Q)}\max_{k\le K_n}\mathrm{Var}\left(\xi^t C^t_{1,k,l} D J^\tau_n D C_{1,k,l}\xi\right) = \begin{cases} O\left(n^{-3}\log^4 n\right), & l=1,\\ O\left(\log^4 n\left(\phi_n^4 + n^{-3}\right)\right), & l=2.\end{cases} \tag{C.12}$$

(c) Using Lemma D.5 (ii)

$$\mathrm{Var}\left(\xi^t C^t_{1,k,l} D J^\tau_n D C_{2,k}\epsilon\right) \le \frac{1}{\sqrt{2}}\mathrm{Var}^{1/2}\left(\xi^t C^t_{1,k,l} D J^\tau_n D C_{1,k,l}\xi\right)\left\|C^t_{2,k} D J^\tau_n D C_{2,k}\right\|_F$$

and hence

$$\begin{aligned}&\sup_{\sigma^2\in\Theta^b_s(\alpha,Q),\ \tau^2\in\Theta^b_s\left(\beta,\bar{Q}\right)}\max_{k\le K_n}\mathrm{Var}\left(\xi^t C^t_{1,k,l} D J^\tau_n D C_{2,k}\epsilon\right)\\ &\qquad = \begin{cases} O\left(n^{-2}\log^2 n\right), & l=1,\\ O\left(\log^2 n\left(\phi_n^2 + n^{-3/2}\right)\right) O\left(n^{-1/2}\right), & l=2.\end{cases}\end{aligned}$$

Combining (a), (b) and (c) in (C.11) yields

$$\begin{aligned}&\sup_{\sigma^2\in\Theta^b_s(\alpha,Q),\ \tau^2\in\Theta^b_s\left(\beta,\bar{Q}\right)}\max_{k\le K_n}\left|\mathrm{Var}\left(\hat{t}_{k,0}\right) - \frac{2}{n}\int_0^1 \tau_k^4(x)\right|\\ &= O\left(\frac{\mathrm{Cum}_4\left(\epsilon\right)}{n} + \frac{1}{n\log n}\right) + \begin{cases} O\left(\phi_n^4\log^4 n + \phi_n n^{-3/4}\log n\right), & l=1,\\ 0, & l=2,\end{cases}\end{aligned} \tag{C.13}$$

and hence (C.2), (C.3), (C.6) and (C.7) follow using Proposition C.1.

Finally we will show the asymptotic normality (C.8) and (C.4). Because of the decomposition (A.3), we have

$$\hat{t}_{k,0,l} = \xi^t C^t_{1,k,l} D J^\tau_n D C_{1,k,l}\xi + 2\xi^t C^t_{1,k,l} D J^\tau_n D C_{2,k}\epsilon + \epsilon^t C^t_{2,k} D J^\tau_n C_{2,k}\epsilon.$$

As proved above $n^{1/2}(\xi^t C^t_{1,k,l} D J^\tau_n D C_{1,k,l}\xi + 2\xi^t C^t_{1,k,l} D J^\tau_n D C_{2,k}\epsilon) \xrightarrow{P} 0$ for $\beta > 1, \alpha > 3/4$ if $l=1$ and $\beta > 1$ if $l=2$. Hence by Slutzky's Lemma it suffices to show that

$$n^{1/2}\left(\epsilon^t C^t_{2,k} D J^\tau_n D C_{2,k}\epsilon - \mathrm{E}\left(\epsilon^t C^t_{2,k} D J^\tau_n D C_{2,k}\epsilon\right)\right) \xrightarrow{L} \mathcal{N}\left(0, 2\int_0^1 \tau_k^4(x)dx\right).$$

In order to apply Theorem D.1, it remains to show

$$n^{1/2}\lambda_1\left(C_{2,k}^t D J_n^\tau D C_{2,k}\right) \to 0.$$

Using Corollary D.1, we see that

$$\begin{aligned} n^{1/2}\lambda_1\left(C_{2,k}^t D J_n^\tau D C_{2,k}\right) &\le 4n^{-1/2}\log^2 n\lambda_1\left(\mathrm{Cov}\left(X_{2,k}+Z_{2,k}\right)\right) \\ &\le 8n^{-1/2}\log^2 n\ \lambda_1\left(\mathrm{Cov}\left(X_{2,k}\right)\right) + 8n^{-1/2}\log^2 n\lambda_1\left(\mathrm{Cov}\left(Z_{2,k}\right)\right) \\ &\le 8n^{-1/2}\log^2 n\sup_{t\in[0,1]}\tau_k^2(t)\max_i\lambda_i + 8n^{-1/2}\log^2 n\phi_{n,1/2}^2 = o\left(1\right), \end{aligned}$$

which yields the last statement of the lemma. □

Proof of Theorem 1. The proof is close to the one of Theorem 2. We obtain

$$\begin{aligned} &\sup_{\sigma^2\in\Theta_s^b(\alpha,Q),\ \tau^2\in\Theta_s^b\left(\beta,\bar{Q}\right)}\int_0^1 \mathrm{Bias}\left(\hat{\tau}_N^2(t)\right)dt \\ &= O\left(N\sup_{\sigma^2\in\Theta_s^b(\alpha,Q),\ \tau^2\in\Theta_s^b\left(\beta,\bar{Q}\right)}\max_{0\le k\le N}\left|\mathrm{E}\left(\hat{t}_{k,0,l}\right)-t_{k,0}\right|^2 + N^{-2\beta}\right), \\ &\sup_{\sigma^2\in\Theta_s^b(\alpha,Q),\ \tau^2\in\Theta_s^b\left(\beta,\bar{Q}\right)}\int_0^1 \mathrm{Var}\left(\hat{\tau}_N^2(t)\right)dt \\ &= O\left(N\sup_{\sigma^2\in\Theta_s^b(\alpha,Q),\ \tau^2\in\Theta_s^b\left(\beta,\bar{Q}\right)}\max_{0\le k\le N}\mathrm{Var}\left(\hat{t}_{k,0,l}\right)\right). \end{aligned}$$

□

Appendix D Technical results

Proposition D.1. *Let $A\in\mathbb{M}_{n-1}$. Then*

$$\mathrm{tr}\left(J_n D A D\right) \le \left(n+5n^{3/2}+8n^{3/2}\left(1+\log n\right)\right)\max_{i,j}\left|\left(A\right)_{i,j}\right|.$$

Proof. Write $A=\left(a_{i,j}\right)_{i,j=1,\ldots,n-1}$. Note that

$$\left(DAD\right)_{i,j} = \frac{2}{n}\sum_{p,q=1}^{n-1}\sin\left(\frac{ip\pi}{n}\right)\sin\left(\frac{qj\pi}{n}\right)a_{p,q}.$$

For $i=j$ we have further

$$\left(DAD\right)_{i,i} = \frac{1}{n}\sum_{p,q=1}^{n-1}a_{p,q}\cos\left(\left(p-q\right)\frac{i\pi}{n}\right) - \frac{1}{n}\sum_{p,q=1}^{n-1}a_{p,q}\cos\left(\left(p+q\right)\frac{i\pi}{n}\right). \tag{D.1}$$

In order to bound the r.h.s. we need bounds for

$$\begin{aligned}\left|\sum_{i=\lfloor n^{1/2}\rfloor+1}^{2\lfloor n^{1/2}\rfloor}\cos\left(r\frac{i\pi}{n}\right)\right| &\le \left|\mathrm{Dir}_{2\lfloor n^{1/2}\rfloor}(r\pi/n)-\mathrm{Dir}_{\lfloor n^{1/2}\rfloor}(r\pi/n)\right| \\ &\le \left|\frac{\sin\left(\left(2\lfloor n^{1/2}\rfloor+1/2\right)r\pi/n\right)}{2\sin\left(r\pi/(2n)\right)}\right|+\left|\frac{\sin\left(\left(\lfloor n^{1/2}\rfloor+1/2\right)r\pi/n\right)}{2\sin\left(r\pi/(2n)\right)}\right| \\ &\le \left|\frac{1}{\sin\left(r\pi/(2n)\right)}\right|.\end{aligned}$$

Let $B_1 := \{1,\dots,n\}$ and $B_2 := \{n+1,\dots,2n-2\}$. Then

$$\left|\sum_{i=\lfloor n^{1/2}\rfloor+1}^{2\lfloor n^{1/2}\rfloor}\cos\left(r\frac{i\pi}{n}\right)\right| \le \begin{cases} n^{1/2} & \text{for } r=0,\\ 4n/r & \text{for } r\in B_1,\\ n/(2n-r) & \text{for } r\in B_2.\end{cases}$$

Therefore, we can bound the first term of the r.h.s. of (D.1) by

$$\begin{aligned}&\left|\sum_{i=\lfloor n^{1/2}\rfloor+1}^{2\lfloor n^{1/2}\rfloor}\frac{1}{n}\sum_{p,q=1}^{n-1}a_{p,q}\cos\left((p-q)\frac{i\pi}{n}\right)\right| \\ &\quad\le \frac{1}{n}\sum_{p,q=1}^{n-1}|a_{p,q}|\left|\sum_{i=\lfloor n^{1/2}\rfloor+1}^{2\lfloor n^{1/2}\rfloor}\cos\left((p-q)\frac{i\pi}{n}\right)\right| \\ &\quad\le n^{-1}\max_{p,q=1,\dots,n-1}|a_{p,q}|\left(n^{3/2}+2\sum_{\substack{p,q=1\\ q-p\in B_1}}^{n-1}\frac{4n}{q-p}\right)\end{aligned}$$

and the second term by

$$\begin{aligned}&\left|\sum_{i=\lfloor n^{1/2}\rfloor+1}^{2\lfloor n^{1/2}\rfloor}\frac{1}{n}\sum_{p,q=1}^{n-1}a_{p,q}\cos\left((p+q)\frac{i\pi}{n}\right)\right| \\ &\le n^{-1}\max_{p,q=1,\dots,n-1}|a_{p,q}|\left(\sum_{\substack{p,q=1\\ p+q\in B_1}}^{n-1}\frac{4n}{p+q}+\sum_{\substack{p,q=1\\ p+q\in B_2}}^{n-1}\frac{n}{2n-(p+q)}\right) \\ &\le 5n\max_{p,q=1,\dots,n-1}|a_{p,q}|.\end{aligned}$$

Due to

$$\sum_{\substack{p,q=1\\ q-p\in B_1}}^{n-1}\frac{1}{q-p}\le n\sum_{r=1}^{n}\frac{1}{r}\le n(1+\log n)$$

and

$$\begin{aligned}\operatorname{tr}(J_n DAD) &= \sqrt{n} \sum_{i=\lceil n^{1/2}\rceil+1}^{2\lceil n^{1/2}\rceil} (DAD)_{i,i} \\ &\leq \left(n + 5n^{3/2} + 8n^{3/2}(1+\log n)\right) \max_{p,q=1,\ldots,n-1} |a_{p,q}|\end{aligned}$$

we obtain the result. □

Proposition D.2. *It holds*

$$\begin{aligned}\sup_{\tau^2\in\Theta_s(\beta,Q)} \max_{k\leq n^{1/2}} &\left| \left\| (J_n^\tau)^{1/2} D \operatorname{Cov}(X_{2,k}+Z_{2,k}) D (J_n^\tau)^{1/2} \right\|_F^2 - \frac{2}{n}\int_0^1 \tau_k^4(x)dx \right| \\ &= O\left(n^{-1}\log^{-1} n\right).\end{aligned}$$

Proof. We obtain with (A.12) $\operatorname{Cov}(X_{2,k}+Z_{2,k}) = 1/2T_k^2K + 1/2KT_k^2 + S_k$, where $S_k := 1/2\tilde{T}_k + \operatorname{Cov}(X_{2,k},Z_{2,k}) + \operatorname{Cov}(Z_{2,k},X_{2,k}) + \operatorname{Cov}(Z_{2,k})$. Application of the triangle inequality gives

$$\begin{aligned}&\frac{1}{2}\left\|(J_n^\tau)^{1/2} D\left(T_k^2K + KT_k^2\right) D (J_n^\tau)^{1/2}\right\|_F - \left\|(J_n^\tau)^{1/2} DS_k D (J_n^\tau)^{1/2}\right\|_F \\ &\quad\leq \left\|(J_n^\tau)^{1/2} D \operatorname{Cov}(X_{2,k}+Z_{2,k}) D (J_n^\tau)^{1/2}\right\|_F \\ &\quad\leq \frac{1}{2}\left\|(J_n^\tau)^{1/2} D\left(T_k^2K + KT_k^2\right) D (J_n^\tau)^{1/2}\right\|_F + \left\|(J_n^\tau)^{1/2} DS_k D (J_n^\tau)^{1/2}\right\|_F.\end{aligned} \tag{D.2}$$

Note that because of Lemma D.4 (iii) it holds

$$\begin{aligned}\operatorname{tr}\left(\left((J_n^\tau)^{1/2} DT_k^2 KD (J_n^\tau)^{1/2}\right)^2\right) &\leq \frac{1}{4}\left\|(J_n^\tau)^{1/2} D\left(T_k^2K + KT_k^2\right) D (J_n^\tau)^{1/2}\right\|_F^2 \\ &\leq \left\|(J_n^\tau)^{1/2} DT_k^2 KD (J_n^\tau)^{1/2}\right\|_F^2.\end{aligned} \tag{D.3}$$

Now we will bound

$$\begin{aligned}\operatorname{tr}\left(\left((J_n^\tau)^{1/2} DT_k^2 KD (J_n^\tau)^{1/2}\right)^2\right) &= \operatorname{tr}\left(\left[(J_n^\tau\Lambda)^{1/2} DT_k^2 D (\Lambda J_n^\tau)^{1/2}\right]^2\right) \\ &= \sum_{i=1}^{n-1} \lambda_i^2\left(DT_k^2 D\Lambda J_n^\tau\right)\end{aligned}$$

from below. We obtain with Lemma D.3

$$\lambda_i\left(DT_k^2 D\Lambda J_n^\tau\right) \geq \begin{cases}\lambda_{n-[n/\log n]}(\Lambda J_n^\tau)\,\lambda_{[n/\log n]-1+i}\left(DT_k^2 D\right), & i \leq n-[n/\log n], \\ 0, & i > n-[n/\log n].\end{cases}$$

Denote by $\tau_{k,(i)}$ the i-th largest component of the vector

$$(\tau_k(1/n), \ldots, \tau_k(1-1/n)).$$

Then

$$\begin{aligned}\operatorname{tr}\left(\left((J_n^\tau)^{1/2} DT_k^2 KD (J_n^\tau)^{1/2}\right)^2\right) &= \sum_{i=1}^{n-1} \lambda_i^2 \left(DT_k^2 D\Lambda J_n^\tau\right) \\ &\geq \sum_{i=1}^{n-[n/\log n]} (n - n/\log n)^{-2} \tau_{k,([n/\log n]-1+i)}^4 \\ &\geq (n - n/\log n)^{-2} \sum_{i=1}^{n-1} \tau_k^4\left(\frac{i}{n}\right) - \tau_{\max}^2 \frac{n}{\log n} (n - n/\log n)^{-2}. \end{aligned} \tag{D.4}$$

Next we will derive an upper bound for the r.h.s. of (D.3). Let analogously to the Definition (A.11) $\bar{T}_k$ be a tridiagonal matrix with entries

$$\left(\bar{T}_k\right)_{i,j} := \begin{cases} \left(\Delta_i \tau_k^2\right)^2 & \text{for} \quad i = j-1, \\ \left(\Delta_j \tau_k^2\right)^2 & \text{for} \quad i = j+1, \\ 0 & \text{otherwise.} \end{cases}$$

Note that $\max_i \left|\Delta_i \tau_k^2\right| \leq 2\tau_{\max}^{1/2} \bar{\phi}_{n,1/2}$. It is easy to check that $T_k^2 K T_k^2 = 1/2 T_k^4 K + 1/2 K T_k^4 + 1/2 \bar{T}_k$ holds. Clearly, $J_n^\tau \leq (n - n/\log n)^{-1} \Lambda^{-1}$, and therefore we have for the upper bound in (D.3)

$$\begin{aligned}\left\|(J_n^\tau)^{1/2} DT_k^2 KD (J_n^\tau)^{1/2}\right\|_F^2 &\leq (n - n/\log n)^{-1} \left\|(J_n^\tau)^{1/2} DT_k^2 KD\Lambda^{-1/2}\right\|_F^2 \\ &\leq (n - n/\log n)^{-1} \operatorname{tr}\left((J_n^\tau)^{1/2} DT_k^2 KT_k^2 D (J_n^\tau)^{1/2}\right) \\ &\leq (n - n/\log n)^{-1} \operatorname{tr}\left(T_k^4 KDJ_n^\tau D\right) + \frac{1}{2}(n - n/\log n)^{-1} \operatorname{tr}\left(D\bar{T}_k DJ_n^\tau\right) \\ &\leq (n - n/\log n)^{-2} \operatorname{tr}\left(T_k^4\right) + 2(n - n/\log n)^{-1} \max_{i,j=1,\ldots,n-1} \left|\bar{T}_k\right|_{i,j} \operatorname{tr}(J_n^\tau), \end{aligned}$$

where we used in the last inequality an argument as for (A.15). Combining this with (D.4) and Proposition C.1 yields

$$\begin{aligned}\sup_{\tau^2 \in \Theta_s^b(\beta,\bar{Q})} \max_{k \leq n^{1/2}} &\left| \left\|(J_n^\tau)^{1/2} D\left(T_k^2 K + KT_k^2\right) D (J_n^\tau)^{1/2}\right\|_F^2 - \frac{2}{n}\int_0^1 \tau_k^4(x)dx \right| \\ &= O\left(n^{-1}\bar{\phi}_{n,1/2} + n^{-1}\log n \bar{\phi}_{n,1/2}^2 + n^{-1}\log^{-1} n\right). \end{aligned} \tag{D.5}$$

Now we will bound the remainder term in (D.2). Using Lemma D.6 gives

$$\begin{aligned}\left\|(J_n^\tau)^{1/2} DS_k D (J_n^\tau)^{1/2}\right\|_F^2 &\leq \lambda_1^2 (J_n^\tau) \|S_k\|_F^2 \\ &\leq 16n^{-2}\log^4 n \left(\left\|\tilde{T}_k\right\|_F^2 + 4\|\mathrm{Cov}(Z_{2,k})\|_F^2 \right. \\ &\qquad \left. +8\|\mathrm{Cov}(X_{2,k}, Z_{2,k})\|_F^2\right).\end{aligned}$$

Because $\mathrm{Cov}(X_{2,k}, Z_{2,k})$ is tridiagonal it holds with Lemma D.4 (i)

$$\|\mathrm{Cov}(X_{2,k}, Z_{2,k})\|_F^2 = \sum_{i,j=1}^{n-1} \left(\mathrm{Cov}(X_{2,k}, Z_{2,k})_{i,j}\right)^2 \leq 8n\tau_{\max}\phi_{n,1/2}^2$$

and therefore

$$\begin{aligned}&\left\|(J_n^\tau)^{1/2} DS_k D (J_n^\tau)^{1/2}\right\|_F^2 \\ &\quad \leq 16n^{-2}\log^4 n \left(2n\bar{\phi}_{n,1/2}^4 + 16n\bar{\phi}_{n,1/2}^4 + 64n\bar{\phi}_{n,1/2}^2\tau_{\max}\right)\end{aligned}$$

This leads to

$$\sup_{\tau^2\in\Theta_s^b(\beta,\bar{Q})} \max_{k\leq n^{1/2}} \left\|(J_n^\tau)^{1/2} DS_k D (J_n^\tau)^{1/2}\right\|_F^2 = O\left(n^{-1}\log^4 n\bar{\phi}_{n,1/2}^2\right)$$

and with (D.2) and (D.5) completes the proof. □

Lemma D.1. *Let $s_{k,p}$ and $t_{k,p}$ as defined in (A.4). Then it holds*

(i)

$$s_{k,p} = \frac{1}{2}s_{0,p} + \frac{1}{4}s_{0,p-k} + \frac{1}{4}s_{0,p+k}, \quad t_{k,p} = \frac{1}{2}t_{0,p} + \frac{1}{4}t_{0,p-k} + \frac{1}{4}t_{0,p+k}.$$

(ii)

$$\frac{1}{n}\sum_{r=1}^{n-1} \sigma_k^2\left(\frac{r}{n}\right)\cos\left(\frac{pr\pi}{n}\right) = A(\sigma_k^2, p) - \frac{1}{2n}\left((-1)^p\sigma_k^2(1) + \sigma_k^2(0)\right).$$

(iii) Let Σ_k as defined in (A.5). Then

$$\left(D\Sigma_k^2 D\right)_{i,j} = A(\sigma_k^2, i-j) - A(\sigma_k^2, i+j).$$

Remark D.1. *In (iii), for $|i-j| \ll i+j$, the r.h.s. behaves like $s_{k,i-j}$. In the same way we obtain the equivalent result if we replace σ^2 by τ^2.*

Proof. (ii) Note that we can write

$$\sigma_k^2\left(\frac{r}{n}\right) = s_{k,0} + 2\sum_{q=1}^{\infty} s_{k,q}\cos(q\pi r/n)$$

and hence it holds

$$\frac{1}{n}\sum_{r=1}^{n-1}\sigma_k^2\left(\frac{r}{n}\right)\cos\left(\frac{pr\pi}{n}\right)$$
$$=\frac{1}{n}s_{k,0}\sum_{r=1}^{n-1}\cos\left(\frac{pr\pi}{n}\right)+\frac{2}{n}\sum_{q=1}^{\infty}s_{k,q}\sum_{r=1}^{n-1}\cos\left(\frac{q\pi r}{n}\right)\cos\left(\frac{pr\pi}{n}\right).$$

Let $\mathbb{I}_{\{A\}}$ denote the indicator function on the set A. We have the identities

$$\sum_{r=1}^{n-1}\cos\left(\frac{pr\pi}{n}\right)=n\mathbb{I}_{\{p\equiv 0 \bmod 2n\}}-\frac{1}{2}\left(1+(-1)^p\right)$$

and

$$2\sum_{r=1}^{n-1}\cos\left(\frac{q\pi r}{n}\right)\cos\left(\frac{pr\pi}{n}\right)=\sum_{r=1}^{n-1}\cos\left(\frac{(q-p)\,\pi r}{n}\right)+\sum_{r=1}^{n-1}\cos\left(\frac{(q+p)\,\pi r}{n}\right).$$

From this it follows

$$\frac{1}{n}\sum_{r=1}^{n-1}\sigma_k^2\left(\frac{r}{n}\right)\cos\left(\frac{pr\pi}{n}\right)$$
$$=\frac{1}{n}\left[-\frac{1}{2}\left(1+(-1)^p\right)s_{k,0}-\sum_{q=1}^{\infty}s_{k,q}\left(1+(-1)^{q-p}\right)\right]+A\left(\sigma_k^2,p\right),$$

which yields the result.

(iii) This follows by applying (ii) to

$$\left(D\Sigma_k^2D\right)_{i,j}=\frac{2}{n}\sum_{r=1}^{n-1}\sigma_k^2\left(\frac{r}{n}\right)\sin\left(\frac{ir\pi}{n}\right)\sin\left(\frac{rj\pi}{n}\right)$$
$$=\frac{1}{n}\sum_{r=1}^{n-1}\sigma_k^2\left(\frac{r}{n}\right)\cos\left(\frac{(i-j)\,r\pi}{n}\right)-\frac{1}{n}\sum_{r=1}^{n-1}\sigma_k^2\left(\frac{r}{n}\right)\cos\left(\frac{(i+j)\,r\pi}{n}\right).$$

□

The next Lemma gives a bound of the absolute values of Fourier coefficients of σ_k^2 in Sobolev s-ellipsoids. In particular the result shows that the Fourier series is absolute summable.

Lemma D.2. *Let $s_{k,p}$ be as defined in (A.4). Assume $k\le cn^\gamma$, where $0<c<1$ is a constant and either $\gamma>0,\alpha>1/2$ or $k=0,\gamma=0$ and $\alpha>1/2$. Then it holds for n large enough*

$$\sup_{\sigma^2\in\Theta_s(\alpha,Q)}\sum_{m=[n^\gamma]}^{\infty}|s_{k,m}|\le C_{\gamma,\alpha,Q,c}n^{\gamma(1/2-\alpha)},$$

where $C_{\gamma,\alpha,Q,c}$ is independent of n.

Proof. Consider the case $\gamma > 0, \alpha > 1/2$. Using Lemma D.1 (i), we see that for n large enough

$$\begin{aligned}\sum_{m=[n^\gamma]}^{\infty} |s_{k,m}| &\leq \sum_{m=[(1-c)n^\gamma]}^{\infty} |s_{0,m}| = \sum_{m=1}^{\infty} |s_{0,m}| \, I_{\{m \geq [(1-c)n^\gamma]\}} \\ &\leq 2\left(\sum_{i=1}^{\infty} i^{2\alpha} \frac{s_{0,i}^2}{4}\right)^{1/2} \left(\sum_{i=[(1-c)n^\gamma]}^{\infty} i^{-2\alpha}\right)^{1/2} \leq C_{\gamma,\alpha,Q,c} n^{\gamma(1/2-\alpha)},\end{aligned}$$

where we used the definition of a Sobolev s-ellipsoid in the last step. If $k = 0$, $\gamma = 0$ and $\alpha > 1/2$ we can argue similarly. □

In the next lemma we collect some important facts about positive semidefinite matrices and trace calculation.

Lemma D.3. *(i) Let $A \in \mathbb{M}_n$ be symmetric. A is positive semidefinite iff $A = B^t B$ for some $B \in \mathbb{M}_n$.*

(ii) If A, B are positive semidefinite matrices. Denote by $\lambda_1(A)$ the largest eigenvalue of A. Then $\operatorname{tr}(AB) \leq \lambda_1(A) \operatorname{tr}(B)$.

(iii) Let $A, B \in \mathbb{M}_{n-1}$ be positive semidefinite. Then

$$\begin{aligned}\lambda_{r+s+1}(AB) &\leq \lambda_{r+1}(A)\lambda_{s+1}(B) \quad 0 \leq r+s \leq n-2 \\ \lambda_{n-r-s+1}(AB) &\geq \lambda_{n-r}(A)\lambda_{n-s}(B) \quad 2 \leq r+s \leq n.\end{aligned}$$

(iv) Let A and B symmetric matrices. Then

$$\lambda_{r+s+1}(A+B) \leq \lambda_{r+1}(A) + \lambda_{s+1}(B) \quad 0 \leq r+s \leq n-2.$$

(v) (CS inequality for trace operator) Let A and B matrices of the same size. Then

$$\left|\operatorname{tr}\left(AB^t\right)\right| \leq \operatorname{tr}^{1/2}\left(AA^t\right) \operatorname{tr}^{1/2}\left(BB^t\right).$$

(vi) Let A, B matrices of the same size. Then

$$A^t B + B^t A \leq A^t A + B^t B.$$

Corollary D.1. *Let A and B matrices of the same size. Then*

$$\lambda_1\left(AB^t + BA^t\right) \leq \lambda_1\left(AA^t\right) + \lambda_1\left(BB^t\right).$$

Proof. By Lemma D.3 *(vi)* $A^t B + AB^t \leq A^t A + B^t B$. Applying Lemma D.3 *(iv)* for $r = s = 0$ yields the result. □

In the following Lemma, we summarize some facts on Frobenius norms.

Lemma D.4. *Let $A \in \mathbb{M}_{n-1}$. Then*

(i)

$$\|A\|_F^2 := \operatorname{tr}\left(AA^t\right) = \sum_{i=1}^{n-1} \lambda_i\left(AA^t\right) = \sum_{i,j=1}^{n-1} a_{i,j}^2$$

and whenever $A = A^t$ *also* $\|A\|_F^2 = \sum_{i=1}^{n-1} \lambda_i^2(A)$.

(ii) It holds

$$4\operatorname{tr}\left(A^2\right) \le \left\|A + A^t\right\|_F^2 \le 4\|A\|_F^2 .$$

(iii) Let A, B *be positive semidefinite matrices of the same size and* $0 \le A \le B$. *Further let* X *be another matrix of the same size. Then*

$$\left\|X^t A X\right\|_F \le \left\|X^t B X\right\|_F .$$

Proof. (i) and (ii) are well known and omitted. (iii) By assumptions it holds $0 \le X^t A X \le X^t B X$. Hence $\lambda_i^2(X^t A X) \le \lambda_i^2(X^t B X)$ and the result follows. □

Lemma D.5. *Let* $V = (V_1, \ldots, V_n), W = (W_1, \ldots, W_m)$ *be two independent, centered random vectors. Let* $A = (a_{i,j})_{i,j=1,\ldots,n} \in \mathbb{M}_n$ *and* $B \in \mathbb{M}_{n,m}$.

(i) Then $\mathrm{E}(V^t A V) = \operatorname{tr}(A \operatorname{Cov}(V))$ *and* $\mathrm{E}(V^t B W) = 0$.

(ii) Assume further that $V_i \perp V_j$ *for all* $i, j = 1, \ldots, n$, $i \ne j$ *and* $W_k \perp W_l$ *for all* $k, l = 1, \ldots, m$, $k \ne l$ *and* $\operatorname{Var}(V_i) = \operatorname{Var}(W_k) = 1$ *for* $i = 1, \ldots, n$ *and* $j = 1, \ldots, m$. *We set* $\operatorname{TrSq}(A) := \sum_{i=1}^n a_{i,i}^2$. *Then*

$$\begin{aligned}
\operatorname{Var}\left(V^t A V\right) &= \operatorname{Cum}_4(V) \sum_{i=1}^n a_{ii}^2 + \operatorname{tr}\left(A^2 + AA^t\right) \\
&\le \operatorname{Cum}_4(V) \sum_{i=1}^n a_{ii}^2 + 2\|A\|_F^2 \\
&\le \left(2 + \operatorname{Cum}_4(V)\right)\|A\|_F^2 , \qquad \text{(D.6)} \\
\operatorname{Var}\left(V^t B W\right) &= \|B\|_F^2 , \\
\operatorname{Var}\left(V^t A B W\right) &\le \left\|AA^t\right\|_F \left\|BB^t\right\|_F . \qquad \text{(D.7)}
\end{aligned}$$

Proof. We only proof the first and the last statement in *(ii)*. Note that

$$\operatorname{Var}\left(V^t A V\right) = \sum_{i,j,k,l=1}^n a_{ij} a_{kl} \operatorname{Cov}\left(V_i V_j, V_k V_l\right) .$$

If $i = j = k = l$ then $\operatorname{Cov}(V_i V_j, V_k V_l) = 2 + \operatorname{Cum}_4(V)$; if $i = k$, $j = l$, $i \ne j$ or $i = l$, $j = k$, $i \ne j$ then $\operatorname{Cov}(V_i V_j, V_k V_l) = 1$. Otherwise $\operatorname{Cov}(V_i V_j, V_k V_l) = 0$ and this gives (D.6).

In order to see (D.7) note that by Lemma D.3 (v)

$$\begin{aligned}
\operatorname{Var}\left(V^t A B W\right) &= \left\|B^t A\right\|_F^2 = \operatorname{tr}\left(\left(BB^t\right)\left(AA^t\right)\right) \\
&\le \operatorname{tr}^{1/2}\left(\left(BB^t\right)^2\right) \operatorname{tr}^{1/2}\left(\left(AA^t\right)^2\right) = \left\|BB^t\right\|_F \left\|AA^t\right\|_F .
\end{aligned}$$

□

Theorem D.1. *Let $\xi \sim \mathcal{N}(0, I_n)$ and A be a positive semidefinite matrix. Then*

$$\mathrm{Var}^{-1/2}\left(\xi^t A\xi\right)\left(\xi^t A\xi - \mathrm{E}\,\xi^t A\xi\right) \to \mathcal{N}(0,1)$$

if and only if $\mathrm{Var}^{-1/2}\left(\xi^t A\xi\right)\lambda_1\left(A\right) \to 0$.

Lemma D.6. *Let $n \geq 4$. Then*

$$\lambda_1\left(J_n^\tau\right) \leq 4n^{-1}\log^2 n.$$

Proof. Let $r = [n/\log n]$ and note that $\sin(x)^{-1} \leq 2/x$ for $x \in (0, \pi/2]$. Then

$$\lambda_r^{-1} \leq \left(\frac{2n}{r\pi}\right)^2 \leq \frac{4}{\pi^2}n^2\left(\frac{1}{2}\frac{n}{\log n}\right)^{-2} \leq 2\log^2 n$$

and

$$\lambda_1\left(J_n^\tau\right) = \left(n - n/\log n\right)^{-1}\lambda_r^{-1} \leq \frac{4}{n}\log^2 n$$

□

Lemma D.7. *Let λ_i be as defined in (3.2). Then it holds*

$$\sqrt{n}\sum_{i=\lceil n^{1/2}\rceil+1}^{2\lceil n^{1/2}\rceil}\lambda_i = \frac{7\pi^2}{3} + O\left(n^{-1/2}\right).$$

Proof. Let $x_i = i\pi/\left(2n\right)$. Note that $\sin^2\left(x_i\right) = x_i^2 - \xi_i^4/3$, where $\xi_i \in (0, x_i)$. Further $\max_{i=\lceil n^{1/2}\rceil+1,\ldots,2\lceil n^{1/2}\rceil} x_i \leq n^{-1/2}\pi$. Hence

$$n^{1/2}\sum_{i=\lceil n^{1/2}\rceil+1}^{2\lceil n^{1/2}\rceil}\xi_i^4 \leq n\max_{i=\lceil n^{1/2}\rceil+1,\ldots,2\lceil n^{1/2}\rceil} x_i^4 = O\left(n^{-1}\right)$$

and thus

$$\begin{aligned} n^{1/2}\sum_{i=\lceil n^{1/2}\rceil+1}^{2\lceil n^{1/2}\rceil}\lambda_i &= 4n^{1/2}\sum_{i=\lceil n^{1/2}\rceil+1}^{2\lceil n^{1/2}\rceil}\frac{i^2\pi^2}{4n^2} + \frac{1}{3}\xi_i^4 \\ &= \pi^2 n^{-3/2}\sum_{i=\lceil n^{1/2}\rceil+1}^{2\lceil n^{1/2}\rceil} i^2 + O\left(n^{-1}\right) = \frac{7\pi^2}{3} + O\left(n^{-1/2}\right). \end{aligned}$$

□

Lemma D.8 (Continuous Sobolev Embedding)**.** *Let $\mathcal{C}(q)$, $q > 0$ denote the space of Hölder continuous functions on $[0,1]$ equipped with the canonical norm $\|.\|_{\mathcal{C}(q)}$ and define*

$$\eta : (1/2,\infty) \times [0,\infty) \to \mathbb{R}, \quad \eta(\alpha,\delta) := \begin{cases} \alpha - 1/2 & \alpha \in (1/2, 3/2), \\ 1-\delta & \alpha = 3/2, \\ 1 & \alpha > 3/2. \end{cases}$$

Suppose $\alpha > 1/2$. Then for any $\delta > 0$ the embedding

$$\iota : \ \Theta_s^b(\alpha, Q) \hookrightarrow \mathcal{C}(\eta(\alpha,\delta))$$

is continuous and in particular

$$\sup_{f \in \Theta_s^b(\alpha,Q)} \|f\|_{\mathcal{C}(\eta(\alpha,\delta))} < \infty.$$

Proof. For a given function $f : [0,1] \to \mathbb{R}$ define $\tilde{f} : [-1,1] \to \mathbb{R}$,

$$\tilde{f}(x) := \begin{cases} f(x) & x \in [0,1], \\ f(-x) & x \in [-1,0]. \end{cases}$$

Let for $s > 0$, $W^{s,2}[-1,1]\big|_{[0,1]}$ denote the (fractional) Sobolev space on $[-1,1]$, where the domain of functions is restricted to $[0,1]$ equipped with the norm

$$\|f\|_{W^{s,2}[-1,1]|_{[0,1]}} := \left\|\tilde{f}\right\|_{W^{s,2}[-1,1]}.$$

Note this is a function space on $[0,1]$ and $W^{s,2}[-1,1]\big|_{[0,1]} \neq W^{s,2}[0,1]$. By the Sobolev embedding theorem (see [28], Proposition 8.5) we have for $\alpha > 1/2$ that

$$\iota : \ \Theta_s^b(\alpha, Q) \subseteq W^{\alpha,2}[-1,1]\big|_{[0,1]} \hookrightarrow \mathcal{C}(\eta(\alpha,\delta))$$

is continuous and since it is linear also bounded. This yields

$$\sup_{f \in \Theta_s^b(\alpha,Q)} \|f\|_{\mathcal{C}(\eta(\alpha,\delta))} \leq \|\iota\| \sup_{f \in \Theta_s^b(\alpha,Q)} \|f\|_{W^{\alpha,2}[-1,1]|_{[0,1]}} < \infty.$$

□

Acknowledgments

We are grateful to T. Tony Cai, Marc Hoffmann, Mark Podolskij and Ingo Witt for helpful comments and discussions. Moreover, we thank an associate editor and a referee for carefully examining our paper and providing us a number of important comments.

References

[1] Alvarez, A., Panloup, F., Pontier, M., and Savy, N. Estimation of the instantaneous volatility. 2010. arxiv:0812.3538, Math arXiv Preprint.

[2] Bandi, F. and Russell, J. Microstructure noise, realized variance, and optimal sampling. *Rev. Econom. Stud.*, 75:339–369, 2008. MR2398721

[3] Barndorff-Nielsen, O., Hansen, P., Lunde, A., and Stephard, N. Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica*, 76(6):1481–1536, 2008. MR2468558

[4] Barucci, E., Malliavin, P., and Mancino, M.E. Harmonic analysis methods for nonparametric estimation of volatility: theory and applications. In *Stochastic processes and applications to mathematical finance. Proceedings of the 5th Ritsumeikan international symposium, Kyoto, Japan, March 3-6, 2005*, pages 1–34, 2006. MR2279147

[5] Bissantz, N., Hohage, T., Munk, A., and Ruymgaart, F. Convergence rates of general regularization methods for statistical inverse problems and applications. *SIAM J. Numerical Analysis*, 45:2610–2636, 2007. MR2361904

[6] Britanak, V., Yip, P.C., and Yao, K.R. *Discrete Cosine and Sine Transforms: General Properties, Fast Algorithms and Integer Approximations.* Academic Press, 2006. MR2293207

[7] Butucea, C. and Tsybakov, A.B. Sharp optimality for density deconvolution with dominating bias, I. *Theory Probab. Appl.*, 52(1):111–128, 2007. MR2354572

[8] Butucea, C. and Tsybakov, A.B. Sharp optimality for density deconvolution with dominating bias, II. *Theory Probab. Appl.*, 52(2):336–349, 2007. MR2354572

[9] Cai, T., Munk, A., and Schmidt-Hieber, J. Sharp minimax estimation of the variance of Brownian motion corrupted with Gaussian noise. *Statist. Sinica*, 2009. Forthcoming.

[10] Delaigle, A. and Gijbels, I. Practical bandwidth selection in deconvolution kernel density estimation. *Comput. Stat. Data Anal.*, 45(2):249–267, 2004. MR2045631

[11] Dey, A.K., Mair, B.A., and Ruymgaart, F.H. Cross-validation for parameter selection in inverse estimation problems. *Scand. J. Statist.*, 23:609–620, 1996. MR1439715

[12] Fan, J. On the optimal rates of convergence for nonparametric deconvolution problem. *Ann. Statist.*, 19(3):1257–1272, 1991. MR1126324

[13] Fan, J., Jiang, J., Zhang, C., and Zhou, Z. Time-dependent diffusion models for term structure dynamics. *Statist. Sinica*, 13:965–992, 2003. MR2026058

[14] Gloter, A. and Jacod, J. Diffusions with measurement errors. I. Local asymptotic normality. *ESAIM Probab. Stat.*, 5:225–242, 2001. MR1875672

[15] Gloter, A. and Jacod, J.. Diffusions with measurement errors. II. Optimal estimators. *ESAIM Probab. Stat.*, 5:243–260, 2001. MR1875673

[16] Jacod, J., Li, Y., Mykland, P.A., Podolskij, M., and Vetter, M.

Microstructure noise in the continuous case: The pre-averaging approach. *Stochastic Process. Appl.*, 119(7):2249–2276, 2009. MR2531091

[17] Jacod, J., Podolskij, M., and Vetter, M. Limit theorems for moving averages of discretized processes plus noise. *Ann. Statist.*, 38(3):1478–1545, 2010. MR2662350

[18] Kinnebrock, S. Asymptotics for semimartingales and related processes with applications in econometrics. 2008. PhD thesis.

[19] Madahavan, A. Market microstructure: A survey. *Journal of Financial Markets*, 3:205–258, 2000.

[20] Munk, A. and Schmidt-Hieber, J. Lower bounds for volatility estimation in microstructure noise models. *Borrowing Strength: Theory Powering Applications - A Festschrift for Lawrence D. Brown, IMS Collection*, 6:43–55, 2010.

[21] Neubauer, A. The convergence of a new heuristic parameter selection criterion for general regularization methods. *Inverse Problems*, 24:055005, 2008. MR2438940

[22] Pereverzev, S. and Schock, E. On the adaptive selection of the parameter in regularization of ill-posed problems. *SIAM J. Numer. Anal.*, 43(5):2060–2076, 2005. MR2192331

[23] Podolskij, M. and Vetter, M. Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. *Bernoulli*, 15(3):634–658, 2009. MR2555193

[24] Reiss, M. Asymptotic equivalence and sufficiency for volatility estimation under microstructure noise. 2010. arxiv:1001.3006, Math arXiv Preprint.

[25] Rosenbaum, M. Integrated volatility and round-off error. *Bernoulli*, 15(3):687–720, 2009. MR2555195

[26] Steele, M. *Stochastic Calculus and Financial Applications*. Springer, New York, 2001. MR1783083

[27] Steele, M. Minimum norm quadratic estimation of spatial variograms. *J. Amer. Statist. Assoc.*, 82(399):765–772, 1987. MR0909981

[28] Taylor, M. *Partial Differential Equations III: Nonlinear Equations*. Springer, New York, 1996. MR1477408

[29] Tsybakov, A.B. *Introduction to Nonparametric Estimation (Springer Series in Statistics XII)*. Springer-Verlag, New York, 2009. MR2013911

[30] Zhang, L. Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. *Bernoulli*, 12:1019–1043, 2006. MR2274854

[31] Zhang, L., Mykland, P., and Ait-Sahalia, Y. A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *J. Amer. Statist. Assoc.*, 472:1394–1411, 2005. MR2236450