Lower bounds for volatility estimation in microstructure noise models

Axel Munk[1,2] and Johannes Schmidt-Hieber[1]

Institut für Mathematische Stochastik, Universität Göttingen,

Goldschmidtstr. 7, 37077 Göttingen

Email: `munk@math.uni-goettingen.de`, `schmidth@math.uni-goettingen.de`

Abstract

In this paper we derive lower bounds in minimax sense for estimation of the instantaneous volatility if the diffusion type part cannot be observed directly but under some additional Gaussian noise. Three different models are considered. Our technique is based on a general inequality for Kullback-Leibler divergence of multivariate normal random variables and spectral analysis of the processes. The derived lower bounds are indeed optimal. Upper bounds can be found in [18]. Our major finding is that the Gaussian microstructure noise introduces an additional degree of ill-posedness for each model, respectively.

AMS 2000 Subject Classification: Primary 62M10; secondary 62G08, 62G20.

Keywords: Brownian motion; Variance estimation; Kullback-Leibler divergence; Minimax rate; Microstructure noise.

1 Introduction and Discussion

Let $X_t = \int_0^t \sigma(s) \, dW_s$, where $(W_t)_{t \in [0,1]}$ denotes here and in the following a standard Brownian motion. Consider the model

$$Y_{i,n} = X_{i/n} + \tau \epsilon_{i,n}, \quad i = 1, \ldots, n, \tag{1.1}$$

where $\epsilon_{i,n} \sim \mathcal{N}(0,1)$, i.i.d. Throughout the paper $(W_t)_{t \in [0,1]}$ and $(\epsilon_{1,n}, \ldots, \epsilon_{n,n})$ are assumed to be independent. Further σ is an unknown, positive and deterministic function and $\tau > 0$ is a known constant.

---

[2]Author for correspondence, email: `munk@math.uni-goettingen.de`

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 $t = i/n$ corresponding to short time intervals $1/n$ which can be of the magnitude of seconds, nowadays. There is a vast amount of literature on volatility estimation in high-frequency models with additional microstructure noise term (see Barndorff-Nielsen et al. [1], Jacod et al. [15], Zhang [23] and Zhang et al. [24] among many others). These kinds of models have attained a lot of attention recently, since the usual quadratic variation techniques for estimation of $\int_0^1 \sigma^2(s)ds$ lead to inconsistent estimators (cf. Zhang [23]). Brown et al. [6] studied low- and high-frequency volatility models by means of asymptotic equivalence. Recently, Reiß [20] has shown that model (1.1) is asymptotically equivalent to a Gaussian shift experiment.

Closely related to model (1.1) is

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

This model can be regarded as a nonparametric extension of the model with constant σ, τ as discussed by Gloter and Jacod [11], [12] and for variogram estimation by Stein [21]. As a further natural modification of (1.1), we consider

$$\bar{Y}_{i,n} = \int_0^{i/n} X_s ds + \tau\epsilon_{i,n}, \quad i = 1, \ldots, n. \tag{1.3}$$

For constant σ, the process

$$\left(\int_0^t X_s ds\right)_{t\geq 0} \overset{\mathcal{D}}{=} \left(\int_0^t (t-s)\,\sigma(s)\,dW_s\right)_{t\geq 0}$$

is called integrated Brownian. In this case, model (1.3) has been used as a prior model for nonparametric regression (see e.g. Cox [8]).

All models have a common structure: They might be interpreted as observations coming from a particular Volterra type stochastic integral, i.e. $\int_0^t K(s,t)\,\sigma(s,t)\,dW_s$ under additional measurement noise. In this paper we derive lower bounds for the models (1.1)-(1.3). We stress that the treatment for general K is an interesting but rather difficult task.

More precisely, we derive minimax lower bounds for estimation of the instantaneous volatility, i.e. σ^2 as a function of time, with respect to L_2-loss. One of the key steps is to consider the estimation problem after taking finite differences which is typical for variance estimation (see Brown and Levine [4] or Munk et al. [16]). Usually in nonparametric regression, lower bounds are obtained under an independence assumption on the observations. In order to deal with dependend data, we introduce a new bound of the Kullback-Leibler divergence which might be of interest by its own. The lower bounds then follow from a standard multiple testing argument together with this bound and the control of the eigenvalues of the covariance operator.

In nonparametric variance estimation, $n^{-\alpha/(2\alpha+1)}$ is well-known to be the minimax rate of convergence given Hölder-smoothness α (for a definition see (2.2)) of the variance and a sufficiently smooth regression function (see Brown and Levine [4] or Munk and Ruymgaart [17]). We show that for the microstructure noise models (1.1) and (1.2) the lower bounds are $n^{-\alpha/(4\alpha+2)}$ and for model (1.3), $n^{-\alpha/(8\alpha+4)}$. If σ exceeds some minimal required smoothness these rates are shown to be also upper bounds for models (1.1) and (1.2) (Munk and Schmidt-Hieber [18]).

For constant σ, i.e. $Y_{i,n} = \sigma W_{i/n} + \tau\epsilon_{i,n}$, $8\tau\sigma^3 n^{-1/4}$ is the optimal asymptotic variance for estimation of σ^2. First this has been shown in a more general setting by Gloter and Jacod [11] but can be also derived as in Cai et *al.* [7] with respect to minimax risk via the information inequality method (see Brown and Farrell [3] and Brown and Low [5]). In order to explain the optimal rate $n^{-1/4}$ it is tempting to think that this is related to the pathwise smoothness of the Brownian motion, since optimal estimation of a Hölder continuous function with index $1/2$ results in a $n^{-1/4}$ rate of convergence. However, this reasoning is in general not true since the smoothness of integrated Brownian motion is arbitrarily close to $3/2$. By the same argument we should obtain in model (1.3) an $n^{-3/8}$ rate if σ constant. This contradicts the obtained lower bound $n^{-1/8}$.

Indeed, we find it more properly to look at these models from the viewpoint of statistical inverse problems. The eigenvalues of the Brownian motion and integrated Brownian motion covariance operators behave like $1/i^2$ and $1/i^4$, respectively (see Freedman [9]). So the maximal frequency i for estimation of σ is reached when $1/i^2 \sim 1/n$ (for Brownian motion) and $1/i^4 \sim 1/n$ (for integrated Brownian motion), i.e. in the first case we may use $O\left(n^{1/2}\right)$ frequencies and in the second case $O\left(n^{1/4}\right)$ resulting in the reduction of the rate of convergence by a factor of $1/2$ and $1/4$, respectively. Motivated by this heuristics, we conjecture that the optimal rate of convergence for the kernel $K\left(s,t\right) = \left(t - s\right)^q, q \in [0, \infty)$ and $\sigma\left(s,t\right) = \sigma\left(s\right)$ is $n^{-\alpha/[(2q+2)(2\alpha+1)]}$. This implies that the rate of convergence decreases as the order of the zero of K increases, or equivalently, as smoother the path of $\int_0^t K\left(s,t\right)\sigma\left(s\right) dW_s$ becomes. For a further result in this direction ($q \in (0, 1/2)$, σ constant) see Gloter and Hoffmann [10]. Note that for $q \in [1/2, 1) \cup (1, \infty)$ the spectral decomposition of the integrated Brownian motion is not known and hence our strategy of proof cannot be applied. More general techniques are required.

Note finally that the lower bounds still hold if we consider generalizations of the models. For instance consider model (1.1) and allow σ to be random itself. Assume further that $\epsilon = (\epsilon_{1,n}, \ldots, \epsilon_{n,n})$ is general microstructure noise, i.e. white noise with bounded moments (see Huang et *al.* [14] or Zhou [25]). In this model the lower bound for estimation of the instantaneous volatility is still $n^{-\alpha/(4\alpha+2)}$, of course. To show that this is indeed the upper bound is current research by Marc Hoffmann and the authors, where the estimator

is constructed by wavelet techniques and methods as in Brown et *al.* [2] are employed.

2 Results

The next Lemma might be of interest by its own and can be applied to various problems in variance estimation where we do not have independent observations. The Lemma can be viewed as a generalization of Lemma 2.1 in Golubev et *al.* [13] for matrices with non-uniformly bounded eigenvalues. See also inequality (3.8) of Reiß [19]. For our purpose it is required to allow eigenvalue sequences tending to 0 and ∞. The proof of our minimax results follows standard arguments (see Tsybakov [22]). Note, however, that the underlying dependency structure of the process causes severe technical difficulties and requires very sharp bounds for the Kullback-Leibler distance in these models as given in the next lemma. Recall that the Kullback-Leibler divergence between two probability measures P, Q is defined as

$$d_K(P,Q) := \int \log\left(\frac{dP}{dQ}\right) dP,$$

whenever $P \ll Q$ and $+\infty$ otherwise.

Lemma 2.1. *Let $X \sim \mathcal{N}(\mu, \Sigma_0)$ and $Y \sim \mathcal{N}(\mu, \Sigma_1)$ be n-variate normal r.v's with expectation μ and covariance Σ_0 and Σ_1, respectively and denote by P_X and P_Y the corresponding probability measures. Assume $0 < C\Sigma_0 \leq \Sigma_1$ for some constant $0 < C \leq 1$. Then it holds for the Kullback-Leibler divergence*

$$d_K(P_Y, P_X) \leq \frac{1}{4C^2} \left\| \Sigma_0^{-1/2}(\Sigma_1 - \Sigma_0)\Sigma_0^{-1/2} \right\|_F^2 \leq \frac{1}{4C^2} \left\| \Sigma_0^{-1}\Sigma_1 - I_n \right\|_F^2, \qquad (2.1)$$

where I_n denotes the $n \times n$ dimensional identity matrix and $\|\cdot\|_F$ is the Frobenius norm, i.e. for a matrix A, $\|A\|_F := \mathrm{tr}^{1/2}(AA^t)$.

Proof. Note that

$$d_K(P_Y, P_X) = \frac{1}{2}\left(\log\left(\det\left(\Sigma_0^{-1/2}\Sigma_1\Sigma_0^{-1/2}\right)^{-1}\right) + \mathrm{tr}\left(\Sigma_0^{-1/2}\Sigma_1\Sigma_0^{-1/2}\right) - n\right).$$

Introduce $\Sigma := \Sigma_0^{-1/2}\Sigma_1\Sigma_0^{-1/2}$ and note that Σ is positive definite. Further, by assumption $0 < C\Sigma_0 \leq \Sigma_1$ and hence $\Sigma_0^{-1/2}\Sigma_1\Sigma_0^{-1/2} \geq CI_n$. This gives $w_i := \lambda_i(\Sigma) \geq C, \quad i = 1, \ldots, n$. Recall $\det(\Sigma) = \prod_{i=1}^n w_i$ and $\mathrm{tr}(\Sigma) = \sum_{i=1}^n w_i$. Hence

$$d_K(P_Y, P_X) = \frac{1}{2}\left(-\sum_{i=1}^n \log(w_i) + \sum_{i=1}^n w_i - n\right) = \frac{1}{2}\sum_{i=1}^n \left(w_i - \log(1 + w_i - 1) - 1\right).$$

4

Assume $x \geq C - 1$. Expand $-\log(1 + x) = -x + \left(2(1 + \xi)^2\right)^{-1} x^2$ for a suitable $|\xi| \leq |x|$. For $C - 1 \leq x \leq 0$ we have $-\log(1 + x) \leq -x + x^2 / (2C^2)$ and for $x \geq 0$, $-\log(1 + x) \leq -x + x^2/2$. Therefore by Lemma A.4

$$d_K(P_Y, P_X) \leq \frac{1}{4C^2} \sum_{i=1}^{n} (w_i - 1)^2 = \frac{1}{4C^2} \left\| \Sigma_0^{-1/2} (\Sigma_1 - \Sigma_0) \Sigma_0^{-1/2} \right\|_F^2.$$

$\square$

Remark 2.1. *The assumption $0 < C\Sigma_0 \leq \Sigma_1$ can be relaxed at the cost of an additional symmetrization term in (2.1). More precisely, assume in Lemma 2.1 instead of $0 < C\Sigma_0 \leq \Sigma_1$ that $0 < \Sigma_0, \Sigma_1$ holds. Then we have*

$$d_K(P_Y, P_X) \leq \frac{1}{4} \left\| \Sigma_0^{-1}\Sigma_1 - I_n \right\|_F^2 + \frac{1}{4} \left\| \Sigma_1^{-1}\Sigma_0 - I_n \right\|_F^2.$$

Next we prove a lower bound for Hölder continuous functions σ^2 and τ^2. For this we need some notation. We write $[x] := \max_{z \in \mathbb{Z}} \{z \leq x\}$, $x \in \mathbb{R}$, the integer part of x. Let $0 < l < u < \infty$ some constants. The class of uniformly bounded Hölder continuous functions of index α on the interval I is defined by

$$\mathcal{C}^b(\alpha, L) := \mathcal{C}^b(\alpha, L, [l, u]) := \left\{ f : f^{(p)} \text{ exists for } p = [\alpha], \right. \tag{2.2}$$

$$\left. \left| f^{(p)}(x) - f^{(p)}(y) \right| \leq L \left| x - y \right|^{\alpha - p}, \ \forall x, y \in I, \ 0 < l \leq f \leq u < \infty \right\}.$$

For any function g we introduce the forward difference operator $\Delta_i g := g((i+1)/n) - g(i/n)$. $\log()$ is defined to be the binary logarithm and we write $\mathbb{M}_p$ and $\mathbb{D}_p$ for the space of $p \times p$ matrices and $p \times p$ diagonal matrices over $\mathbb{R}$, respectively. The i-th largest eigenvalue of a Hermitian matrix M is defined as $\lambda_i(M)$.

Theorem 2.1. *Assume model (1.1) and $\alpha > 1/2$ or model (1.2), $\alpha \geq 1$. Then there exists a $C > 0$ (depending only on α, L, l, u), such that*

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

Proof. Besides working with the observations $Y_{1,n}, \ldots, Y_{n,n}$ directly, we consider the sufficient statistics

$$\begin{aligned} \Delta Y &= (Y_{1,n}, Y_{2,n} - Y_{1,n}, \ldots, Y_{n,n} - Y_{n-1,n}), \\ \Delta \tilde{Y} &= \left(\tilde{Y}_{1,n}, \tilde{Y}_{2,n} - \tilde{Y}_{1,n}, \ldots, \tilde{Y}_{n,n} - \tilde{Y}_{n-1,n} \right). \end{aligned}$$

Let $\Delta_i W := W_{(i+1)/n} - W_{i/n}$ and $\Delta_i \epsilon_{i,n} = \epsilon_{i+1,n} - \epsilon_{i,n}$. We obtain for $i = 1, \ldots, n$

$$Y_{i,n} - Y_{i-1,n} = \int_{(i-1)/n}^{i/n} \sigma(s) dW_s + \tau \Delta_{i-1} \epsilon_{i-1,n},$$

$$\tilde{Y}_{i,n} - \tilde{Y}_{i-1,n} = \sigma\left(\frac{i}{n}\right) \Delta_{i-1} W + (\Delta_{i-1}\sigma) W_{(i-1)/n} + \tau \Delta_{i-1} \epsilon_{i-1,n}, \tag{2.3}$$

5

where $\epsilon_{0,n} := 0$, $Y_{0,n} := 0$ and $\tilde{Y}_{0,n} := 0$. We may write $\Delta Y = X_1 + X_2$ and $\Delta \tilde{Y} = X_1' + X_2 + R_1$, where X_1, X_1', X_2 have components $\int_{(i-1)/n}^{i/n} \sigma(s) dW_s$, $\sigma(i/n) \Delta_{i-1} W$ and $\tau \Delta_{i-1} \epsilon_{i-1,n}$, respectively and $(R_1)_i = (\Delta_{i-1}\sigma) W_{(i-1)/n}$.

Let us first consider model (1.2), $\alpha \geq 1$. By the ease of brevity, we simply point out in a second step the differences to model (1.1). The idea is to prove the lower bound by a multiple testing argument. Explicitly, we apply Theorem 2.5 in Tsybakov [22]. The construction of hypothesis is similar to the one given in [22], Section 2.6.1. We write $\sigma_{\min}, \sigma_{\max}$ for the lower and upper bound of σ^2, respectively, i.e. $\sigma^2 \in \mathcal{C}^b(\alpha, L, [\sigma_{\min}, \sigma_{\max}])$. Without loss of generality, we may assume that $\sigma_{\min} = 1$. Let

$$K : \mathbb{R} \to \mathbb{R}^+, \quad K(u) = a \exp\left(-\frac{1}{1 - (2u)^2}\right) \mathbb{I}_{\{|2u| \leq 1\}}, \tag{2.4}$$

where a is such that $K \in \mathcal{C}(\alpha, 1/2)$. Further for some $c > 0$, specified later on, let $m := \left[2^{-1} c n^{1/(4\alpha+2)} + 1\right]$, $h_n = (2m)^{-1}$, $t_k = h_n(k - 1/2) + 1/4$,

$$\phi_k(t) := L h_n^\alpha K\left(\frac{t - t_k}{h_n}\right), \quad k = 1, \ldots, m, \quad t \in [0, 1].$$

Define $\Omega := \{\omega = (\omega_1, \ldots, \omega_m), \omega_i \in \{0, 1\}\}$ and consider the set $\mathcal{E} := \{\sigma_\omega^2(t) : \sigma_\omega^2(t) = 1 + \sum_{k=1}^m \omega_k \phi_k(t), \omega \in \Omega\}$. Then it holds $\forall \omega, \omega' \in \Omega$

$$\left\| \sigma_\omega^2 - \sigma_{\omega'}^2 \right\|_2^2 = \int_0^1 \left(\sigma_\omega^2(t) - \sigma_{\omega'}^2(t)\right)^2 dt = L^2 h_n^{2\alpha+1} \|K\|_2^2 \rho(\omega, \omega'), \tag{2.5}$$

where $\rho(\omega, \omega') = \sum_{k=1}^m \mathbb{I}_{\{\omega_k \neq \omega_k'\}}$ is the Hamming distance. By the Varshamov-Gilbert bound (cf. Tsybakov [22]) there exists for all $m \geq 8$ a subset $\{\omega_0, \ldots, \omega_M\}$ of Ω such that $\omega_0 = (0, \ldots, 0)$, $\rho(\omega_i, \omega_j) \geq m/8$, $\forall 0 \leq j < k \leq M$ and $M \geq 2^{m/8}$. Define the hypothesis H_i for $i = 0, \ldots, M$ by the probability measure P_i induced by $\sigma_{i,n}^2 := \sigma_{\omega_i,n}^2$. By Theorem 2.5 in Tsybakov [22] the proof is finished once we have established

$$\begin{aligned}
&(i) \quad \sigma_{i,n}^2 \in \mathcal{C}(\alpha, L), \ 1 \leq \sigma_{i,n}^2 \leq \sigma_{\max} \\
&(ii) \quad \left\| \sigma_{i,n}^2 - \sigma_{j,n}^2 \right\|_2 \geq 2s \geq c n^{-\alpha/(4\alpha+2)}, \quad i \neq j, \ c > 0, \\
&(iii) \quad \frac{1}{M} \sum_{j=1}^M d_K(P_j, P_0) \leq \kappa \log M, \quad j = 1, \ldots, M, \quad \kappa < 1/10.
\end{aligned} \tag{2.6}$$

(i) is obviously fulfilled for sufficiently large n. By (2.5) it follows for $m > 8$, $\left\| \sigma_{i,n}^2 - \sigma_{j,n}^2 \right\|_2 \geq 16^{-1} L h_n^\alpha \|K\|_2$ and hence (ii). (iii) We apply Lemma 2.1 in combination with Lemma A.2. Let $\Pi_k \in \mathbb{D}_n$, $k = 1, \ldots, M$ with entries $(\Pi_k)_{i,j} = \sigma_{k,n}(i/n) \delta_{i,j}$. For the observation vector

6

$\tilde{Y} = \left(\tilde{Y}_{1,n}, \ldots, \tilde{Y}_{n,n} \right)$, we have $\tilde{Y} \sim \mathcal{N}\left(0, \Sigma'_k\right)$ under H_k, $k = 0, \ldots, M$, where

$$\Sigma'_0 = \left(\frac{i \wedge j}{n} \right)_{i,j=1,\ldots,n} + \tau^2 I_n,$$

$$\Sigma'_k = \Pi_k \left(\frac{i \wedge j}{n} \right)_{i,j=1,\ldots,n} \Pi_k + \tau^2 I_n, \quad k = 1, \ldots, M.$$

Because of $\sigma_{k,n} \geq 1$ it follows $|\sigma_{k,n}(x) - \sigma_{k,n}(y)| \leq \left| \sigma^2_{k,n}(x) - \sigma^2_{k,n}(y) \right| \leq L\,|x - y|$, i.e. $\sigma \in \mathcal{C}(1, L)$. By Lemma A.2, $0 < \left(2 + 12L^2\right)^{-1} \Sigma'_0 < \Sigma'_k$. These inequalities remain valid under any invertible linear transformations of $\tilde{Y}$. Hence if we denote by Σ_k the covariance of $\Delta \tilde{Y}$ under H_k $(k = 0, \ldots, M)$ it follows $0 < \left(2 + 12L^2\right)^{-1} \Sigma_0 < \Sigma_k$ and we may apply Lemma 2.1 with $C = \left(2 + 12L^2\right)^{-1}$. Hence

$$d_K(P_k, P_0) \leq \frac{\left(2 + 12L^2\right)^2}{4} \left\| \Sigma_0^{-1} \Sigma_k - I_n \right\|_F^2 \quad k = 1, \ldots, M.$$

Let for $1 \leq i, j \leq n$, the matrix A be defined by

$$(A)_{i,j} := \begin{cases} 2 & \text{for} \quad i = j \quad \text{and} \quad i > 1 \\ -1 & \text{for} \quad |i - j| = 1 \\ 1 & \text{for} \quad i = j = 1 \\ 0 & \text{else} \end{cases} \tag{2.7}$$

and let $\Gamma_k := \Pi_k - I_n$, $k = 1, \ldots, M$. Clearly, $\Gamma_k \leq L h_n^\alpha \|K\|_\infty$. We abbreviate the covariance of two column vectors X and Y as the matrix with covariances of XY^t. Then we have the explicit representations

$$\Sigma_0 = \frac{1}{n} I_n + \tau^2 A,$$

$$\Sigma_k = \Sigma_0 + \frac{2}{n} \Gamma_k + \frac{1}{n} \Gamma_k^2 + \text{Cov}_{H_k}\left(X_1', R_1\right) + \text{Cov}_{H_k}\left(R_1, X_1'\right) + \text{Cov}_{H_k}\left(R_1\right),$$

where the subscript H_k means that these covariances are taken with respect to the probability measure induced by H_k. We remark that due to $\Sigma_0 \geq I_n/n$ it holds $\lambda_1\left(\Sigma_0^{-1}\right) = \lambda_n^{-1}\left(\Sigma_0\right) \leq \lambda_n^{-1}\left(I_n/n\right) = n$. This yields using Lemma A.3, (i)

$$d_K(P_k, P_0) \leq 5 \frac{\left(2 + 12L^2\right)^2}{4} \left(\frac{4}{n^2} \left\| \Sigma_0^{-1} \Gamma_k \right\|_F^2 + \frac{1}{n^2} \left\| \Sigma_0^{-1} \Gamma_k^2 \right\|_F^2 + \left\| \Sigma_0^{-1} \text{Cov}_{H_k}\left(X_1', R_1\right) \right\|_F^2 \right.$$

$$\left. + \left\| \Sigma_0^{-1} \text{Cov}_{H_k}\left(R_1, X_1'\right) \right\|_F^2 + \left\| \Sigma_0^{-1} \text{Cov}_{H_k}\left(R_1\right) \right\|_F^2 \right)$$

$$\leq 5 \left(1 + 6L^2\right)^2 \left(\left(4L^2 \|K\|_\infty^2 + L^4 \|K\|_\infty^4 h_n^{2\alpha} \right) h_n^{2\alpha} n^{-2} \left\| \Sigma_0^{-1} \right\|_F^2 \right.$$

$$\left. + n^2 \left\| \text{Cov}_{H_k}\left(X_1', R_1\right) \right\|_F^2 + n^2 \left\| \text{Cov}_{H_k}\left(R_1, X_1'\right) \right\|_F^2 + n^2 \left\| \text{Cov}_{H_k}\left(R_1\right) \right\|_F^2 \right).$$

7

The remaining part of the proof is concerned with bounding these terms. We make use of the properties on Frobenius norms collected in Lemma A.4 and obtain

$$\left\|\mathrm{Cov}_{H_k}(R_1)\right\|_F^2 \leq n^{-4}L^4 \left\|\left(\frac{(i\wedge j)-1}{n}\right)_{i,j=1,\ldots,n}\right\|_F^2 \leq L^4 n^{-2}.$$

Let $E \in \mathbb{M}_n$ given by

$$(E)_{i,j} := \begin{cases} 1 & \text{if } j > i \\ 0 & \text{otherwise} \end{cases}.$$

Also let $\Delta\Pi_k \in \mathbb{D}_n$, $k=1,\ldots,M$ with entries $(\Delta\Pi_k)_{i,j} = (\Delta_{i-1}\sigma_{k,n})\,\delta_{i,j}$, where $\sigma_{k,n}(0)=0$. Then $\mathrm{Cov}_{H_k}(X_1',R_1) = n^{-1}\Pi_k E(\Delta\Pi_k)$ and for n large enough

$$\left\|\mathrm{Cov}_{H_k}(X_1',R_1)\right\|_F^2 \leq 4n^{-2}\,\mathrm{tr}\left(E(\Delta\Pi_k)^2 E^t\right) \leq 4L^2 n^{-4}\|E\|_F^2 \leq 4L^2 n^{-2}.$$

With the same arguments we can bound $\left\|\Sigma_0^{-1}\mathrm{Cov}_{H_k}(R_1,X_1')\right\|_F^2$. Further we have for $j=1,\ldots,M$ using Lemma A.1

$$\left\|\Sigma_0^{-1}\right\|_F^2 \leq \sum_{i=1}^n \left(\frac{1}{n} + \tau^2 i^2/\left(4n^2\right)\right)^{-2} \leq C_\tau n^{5/2},$$

for a constant C_τ only depending on τ. This gives

$$\frac{1}{M}\sum_{j=1}^M d_K(P_j,P_0)$$
$$\leq 5\left(1+6L^2\right)^2\left(\left(4L^2\|K\|_\infty^2 + L^4\|K\|_\infty^4 h_n^{2\alpha}\right)n^{1/2}h_n^{2\alpha}C_\tau + L^4 + 8L^2\right)$$
$$\leq C_{\tau,L,\|K\|_\infty}n^{1/2}h_n^{2\alpha} \leq C_{\tau,L,\|K\|_\infty}c^{-2\alpha-1}2m \leq \kappa\log M,$$

where $C_{\tau,L,\|K\|_\infty}$ is independent of n and the last inequality holds if

$$c > \left(16C_{\tau,L,\|K\|_\infty}\kappa^{-1}\right)^{1/(2\alpha+1)},$$

using the Varshamov-Gilbert bound.

The proof for model (1.1) and $\alpha > 1/2$ is almost the same as for Theorem 2.1. So we only sketch it here. Note that Lemma 2.1 can be applied directly without use of Lemma A.2. The construction of hypothesis is the same. Let $\Pi_k \in \mathbb{D}_n$, $k=1,\ldots,M$ defined by $(\Pi_k)_{i,i} := \int_{(i-1)/n}^{i/n}\sigma_{k,n}^2(s)\,ds$ and $\Gamma_k := \Pi_k - I_n$. With the notation as in Theorem 2.1 the problem can be reduced to a testing problem where we have to test the hypothesis that a centered random vector has covariance matrix $\Sigma_0 = I_n + A$ against the M alternative covariance matrices $\Sigma_k = \Sigma_0 + \Gamma_k$, $k=1,\ldots,M$. With $\max_{1\leq i\leq n}\max_{1\leq k\leq M}(\Gamma_k)_{i,i} = O(h_n^\alpha)$ and Lemma 2.1 the Theorem follows. $\square$

The proof of the lower bound for estimation of σ^2 in model (1.3) differs from the previous one. Instead of using first order differences we transform the data by taking second order differences. The key step is to show that for constant σ this is "close" to the model where we observe $Z_{i,n} = \sigma n^{-3/2} \eta_{i,n} + \tau \xi_{i,n}, i = 1, \ldots, n, \ \eta_{i,n} \sim \mathcal{N}(0,1)$, i.i.d. Here, $\xi = (\xi_{1,n}, \ldots, \xi_{n,n})$ is a particular MA(2)-process, independent of $\eta = (\eta_{1,n}, \ldots, \eta_{n,n})$.

Theorem 2.2. *Assume model (1.3) and $\alpha > 1/2$. Then there exists a $C > 0$ (depending only on α, L, l, u), such that*

$$\liminf_{n \to \infty} \inf_{\hat{\sigma}_n^2} \sup_{\sigma^2 \in \mathcal{C}^b(\alpha, L)} \mathrm{E}\left(n^{\frac{\alpha}{4\alpha+2}} \left\| \hat{\sigma}^2 - \sigma^2 \right\|_2^2 \right) \geq C.$$

Furthermore, there exists a $\tilde{C} > 0$ (depending on $\sigma_{\min}, \sigma_{\max}$) such that for constant σ and $0 < \sigma_{\min} < \sigma_{\max} < \infty$

$$\liminf_{n \to \infty} \inf_{\hat{\sigma}_n^2} \sup_{\sigma^2 \in [\sigma_{\min}, \sigma_{\max}]} \mathrm{E}\left(n^{1/4} \left(\hat{\sigma}^2 - \sigma^2 \right)^2 \right) \geq \tilde{C}.$$

Proof. Except for changing the definition of $m := \left[2^{-1} c n^{1/(4\alpha+2)} + 1 \right]$ to $m := \left[2^{-1} c n^{1/(8\alpha+4)} + 1 \right]$ we construct the same hypothesis as in the proof of Theorem 2.1. In order to prove the first part of the statement it remains to show (iii) in (2.6). We consider second order differences, i.e.

$$Y_i^* := \Delta_{i-1}^2 \bar{Y}_i, \ \bar{Y}_0 := 0, \ i = 2, 3, \ldots, \ Y_1^* := \sqrt{2} \bar{Y}_1,$$

where $\Delta_i^2 = \Delta_i \circ \Delta_i$. Note that for $i \geq 2$

$$Y_i^* = \int_{(i-1)/n}^{i/n} \left(\frac{i}{n} - s \right) \sigma(s) \, dW_s + \int_{(i-2)/n}^{(i-1)/n} \left(s - \frac{i-2}{n} \right) \sigma(s) \, dW_s + \tau \Delta_i^2 \epsilon_i$$

Let $Y^* := (Y_1^*, \ldots, Y_n^*)$. Obviously, this is equivalent to observing $\bar{Y}$. The covariance of Y^* under hypothesis H_k is denoted by $\Sigma_k, k = 0, \ldots, M$. Since by construction $\sigma_{0,n}^2 \leq \sigma_{k,n}^2$, it follows from elementary computations that $\Sigma_k - \Sigma_0$ is the covariance of Y^* under $\sigma = \left(\sigma_{k,n}^2 - \sigma_{0,n}^2 \right)^{1/2}$ and $\tau = 0$. From this we conclude that

$$\Sigma_k - \Sigma_0 \geq 0, \tag{2.8}$$

i.e. $\Sigma_0 \leq \Sigma_k$. Note that $\mathrm{Var}(X + Y) \leq 2 \mathrm{Var} X + 2 \mathrm{Var} Y$ (in the sense of Loewner ordering, see also Lemma A.3, (iii)). This yields together with (2.8)

$$\Sigma_k - \Sigma_0 \leq \Gamma,$$

where Γ is diagonal with entries

$$(\Gamma)_{i,j} = \frac{4 L h_n^\alpha \|K\|_\infty}{3 n^3} \delta_{i,j},$$

and $\delta_{i,j}$ denotes the Kronecker delta. Let P_k denote the probability measure of Y_k^* under H_k. Due to $\Sigma_0 \leq \Sigma_k$ we may apply Lemma 2.1 and obtain

$$d_K\left(P_k, P_0\right) \leq \frac{1}{4}\left\|\Sigma_0^{-1/2}\left(\Sigma_k - \Sigma_0\right)\Sigma_0^{-1/2}\right\|_F^2 \leq \frac{16L^2 h_n^{2\alpha}\left\|K\right\|_\infty^2}{9n^6}\left\|\Sigma_0^{-1}\right\|_F^2. \qquad (2.9)$$

Direct computations give

$$\Sigma_0 = \frac{1}{n^3}I_n - \frac{1}{6n^3}A + \frac{\sqrt{2}-1}{6n^3}V_1 + \tau^2\left(A^2 + V_2\right),$$

where A is as defined in (2.7),

$$(V_1)_{i,j} := \begin{cases} 1 & \text{if } i = 1, j = 2 \text{ or } i = 2, j = 1, \\ 0 & \text{otherwise.} \end{cases}$$

and V_2 is symmetric and $(V_2)_{i,j} \neq 0$ only if $i, j \leq 3$. Obviously, the smallest eigenvalue of V_1 is -1. Hence we can estimate by Lemma A.1,

$$\Sigma_0 \geq \frac{1}{6n^3}I_n + \tau^2\left(A^2 + V_2\right).$$

Since V_2 has only non-zero entries in the first three rows and columns, it has only three non-zero eigenvalues. By standard bounds on eigenvalues (see Lemma A.3, (ii)) this allows to estimate for $i \geq 3$

$$\lambda_{n-i}\left(A^2 + V_2\right) \geq \lambda_{n-i+3}\left(A^2\right) + \lambda_{n-3}\left(V_2\right) = \lambda_{n-i+3}^2\left(A\right).$$

Let $r_n := \left[n^{1/4}\right]$. Then for sufficiently large n by Lemma A.1

$$\left\|\Sigma_0^{-1}\right\|_F^2 = \sum_{i=1}^n \lambda_i^2\left(\Sigma_0^{-1}\right) \leq \sum_{i=1}^{r_n} \lambda_{n-i+1}^{-2}\left(\frac{1}{6n^3}I_n\right) + \tau^4 \sum_{i=r_n+1}^n \lambda_{n-i+1}^{-2}\left(A^2 + V_2\right)$$

$$\leq 36n^{25/4} + \tau^4 \sum_{i=r_n+1}^n \lambda_{n-i+4}^{-4}\left(A\right) \leq 36n^{25/4} + 4^4\tau^4 \sum_{i=r_n+1}^n \frac{n^8}{(i-3)^8}$$

$$\leq 36n^{25/4} + 2^{24}n^{25/4}\frac{1}{n^{1/4}} \sum_{i=r_n+1}^n \frac{1}{\left(i/n^{1/4}\right)^8}.$$

Because the last part converges as a Riemann sum to a finite integral, we may find a constant C_τ depending only on τ such that $\left\|\Sigma_0^{-1}\right\|_F^2 \leq C_\tau n^{25/4}$. This gives with (2.9),

$$\frac{1}{M}\sum_{k=1}^M d_K\left(P_k, P_0\right) \leq \frac{16}{9}C_\tau L^2\left\|K\right\|_\infty^2 h_n^{2\alpha}n^{1/4} \leq \frac{2^8}{9}C_\tau L^2\left\|K\right\|_\infty^2 c^{-2\alpha-1}m \leq \kappa \log M$$

whenever $c \geq \left(2^8/9C_\tau L^2\left\|K\right\|_\infty^2 \kappa^{-1}\right)^{1/(2\alpha+1)}$.

In order to prove the second statement of the theorem we consider two hypothesis $\sigma_0^2 := \sigma_{\min}$ and $\sigma_1^2 := \sigma_{\min} + cn^{-1/8}$ and apply Theorem 2.5 in Tsybakov [22] for $M = 2$. Using the bounds from the first part, the remaining part of the proof is straightforward and thus omitted. $\qquad\square$

Appendix A   Technical tools

Lemma A.1. *Let* A, Q^{-1} *be as in (2.7) and (A.1), respectively. Then*

$$\lambda_{n-i+1}(A) = \lambda_{n-i+1}(Q^{-1}) = 4\sin^2\left(\frac{(2i-1)\pi}{4n+2}\right) \geq \frac{i^2}{4n^2}, \quad i = 1,\dots,n$$

Proof. The i-th eigenvector v_i of Q^{-1} is given by $v_i = (\sin(x_i), \sin(2x_i), \dots, \sin(nx_i))$, where $x_i := (2i-1)\pi/(2n+1), i = 1,\dots,n$. The corresponding eigenvalues are

$$\lambda_{n-i+1}(Q^{-1}) = 4\sin^2\left(\frac{(2i-1)\pi}{4n+2}\right).$$

If $v_i = (v_{i,1},\dots,v_{i,n})$ is an eigenvector of Q^{-1} so is $\tilde{v}_i = (v_{i,n},\dots,v_{1,1})$ an eigenvector of A with the same eigenvalue. Using $x\pi/2 \leq \sin(x\pi)$ whenever $x \in [0, 1/2]$, we obtain

$$4\sin^2(x_i\pi/2) \geq x_i^2\pi^2/4 \geq \frac{i^2}{4(2n+1)^2}\pi^2 \geq \frac{i^2}{4n^2}.$$

$\square$

Lemma A.2. *Let* $\sigma \geq 1$ *be a Hölder* $\mathcal{C}(1, L)$ *function and let* $\Sigma \in \mathbb{D}_n$ *be a diagonal matrix with* $\Sigma_{i,i} = \sigma(i/n)$. *Further introduce* $Q := (i \wedge j)_{i,j=1,\dots,n}$. *Then*

$$\left(2 + 12L^2\right)^{-1} Q \leq \Sigma Q \Sigma$$

in the sense of partial Loewner ordering of symmetric matrices.

Proof. Obviously

$$Q_{i,j}^{-1} = \begin{cases} 2 & \text{for} \quad i = j \quad \text{and} \quad i < n, \\ -1 & \text{for} \quad |i - j| = 1, \\ 1 & \text{for} \quad i = j = n, \\ 0 & \text{else.} \end{cases} \tag{A.1}$$

Let $O \in \mathbb{M}_n$ given by

$$O_{i,j} = \begin{cases} 1 & \text{for} \quad i = j, \\ -1 & \text{for} \quad i = j - 1, \\ 0 & \text{else.} \end{cases}$$

Note that $Q^{-1} = OO^t$. We have with

$$\left(\tilde{\Sigma}\right)_{i,j} := \begin{cases} \left(\Delta_i\sigma^{-1}\right)^2 & \text{for} \quad i = j - 1, \\ \left(\Delta_j\sigma^{-1}\right)^2 & \text{for} \quad i = j + 1, \\ 0 & \text{else} \end{cases}$$

11

that

$$\Sigma^{-1}Q^{-1}\Sigma^{-1} = \frac{1}{2}\Sigma^{-2}Q^{-1} + \frac{1}{2}Q^{-1}\Sigma^{-2} + \frac{1}{2}\tilde{\Sigma}$$

$$\leq Q^{-1} + \frac{1}{2}\left(\Sigma^{-2}O - O\Sigma^{-2}\right)O^t + \frac{1}{2}O\left(O^t\Sigma^{-2} - \Sigma^{-2}O^t\right) + \frac{1}{2}\tilde{\Sigma}. \qquad \text{(A.2)}$$

As can be seen by direct calculations

$$\left(\Sigma^{-2}O - O\Sigma^{-2}\right)O^t + O\left(O^t\Sigma^{-2} - \Sigma^{-2}O^t\right)$$

$$= -\begin{pmatrix} 2\Delta_1\sigma^{-2} & -\Delta_1\sigma^{-2} & & \\ -\Delta_1\sigma^{-2} & \ddots & & \ddots \\ & \ddots & 2\Delta_{n-1}\sigma^{-2} & -\Delta_{n-1}\sigma^{-2} \\ & & -\Delta_{n-1}\sigma^{-2} & 0 \end{pmatrix}. \qquad \text{(A.3)}$$

Define $M \in \mathbb{M}_n$ by

$$M_{i,j} = \begin{cases} 1 & \text{for} \quad i = j, \\ \Delta_i\sigma^{-2} - 1 & \text{for} \quad i = j - 1, \\ 0 & \text{else.} \end{cases}$$

Due to $MM^t \geq 0$ and (A.3) it follows

$$\left(\Sigma^{-2}O - O\Sigma^{-2}\right)O^t + O\left(O^t\Sigma^{-2} - \Sigma^{-2}O^t\right)$$

$$= Q^{-1} + \begin{pmatrix} \left(\Delta_1\sigma^{-2}\right)^2 & & & \\ & \ddots & & \\ & & \left(\Delta_{n-1}\sigma^{-2}\right)^2 & \\ & & & 0 \end{pmatrix} \leq Q^{-1} + \frac{4L^2}{n^2}I_n,$$

where we used in the last step that

$$\left|\frac{1}{\sigma^2(x)} - \frac{1}{\sigma^2(y)}\right| = |\sigma(x) - \sigma(y)|\left|\frac{1}{\sigma^2(x)\sigma(y)} + \frac{1}{\sigma(x)\sigma^2(y)}\right| \leq 2L|x-y|,$$

i.e. $\sigma \in \mathcal{C}(1, L)$ and $\sigma \geq 1$ implies $\sigma^{-2} \in \mathcal{C}(1, 2L)$.

Next we bound the eigenvalues of the symmetric matrix $\tilde{\Sigma}$ by showing that the corresponding characteristic polynomial $\chi_{\tilde{\Sigma}}(t)$ does not have any zero in $[2\omega, \infty)$, where $\omega = \max_{i,j}\left|\left(\tilde{\Sigma}\right)_{i,j}\right|$. In order to see this introduce the notation $s(i) := \left(\tilde{\Sigma}\right)_{i,i+1}/\omega$ and note for

12

$t \geq 2\omega$

$$
\begin{pmatrix}
t\omega^{-1} & -s\,(1) & & \\
-s\,(1) & \ddots & & \ddots \\
& \ddots & \ddots & -s\,(n-1) \\
& & -s\,(n-1) & t\omega^{-1}
\end{pmatrix}
$$

$$
\geq
\begin{pmatrix}
1+s\,(1)^2 & -s\,(1) & & \\
-s\,(1) & \ddots & & \ddots \\
& \ddots & 1+s\,(n-1)^2 & -s\,(n-1) \\
& & -s\,(n-1)^2 & 1
\end{pmatrix}
$$

$$
=
\begin{pmatrix}
1 & -s\,(1) & & \\
& \ddots & \ddots & \\
& & \ddots & -s\,(n-1) \\
& & & 1
\end{pmatrix}
\begin{pmatrix}
1 & -s\,(1) & & \\
& \ddots & \ddots & \\
& & \ddots & -s\,(n-1) \\
& & & 1
\end{pmatrix}^t
> 0
$$

and therefore $\chi_{\tilde{\Sigma}}(t) > 0$ for $t \geq 2\omega$. Because of $w \leq L^2/n^2$ this shows that

$$
\Sigma^{-1}Q^{-1}\Sigma^{-1} \leq 3/2\,Q^{-1} + \frac{3L^2}{n^2}I_n. \tag{A.4}
$$

From Lemma A.1 follows $1/(4n^2) \leq \lambda_1^{-1}(Q) = \lambda_n\left(Q^{-1}\right)$ and hence

$$
\frac{L^2}{n^2}I_n \leq 4L^2\lambda_n\left(Q^{-1}\right)I_n \leq 4L^2 Q^{-1}.
$$

This gives with (A.4) finally

$$
\Sigma^{-1}Q^{-1}\Sigma^{-1} \leq 2Q^{-1} + 12L^2 Q^{-1}.
$$

$\square$

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

Lemma A.3.   *(i) Let $A, B \in \mathbb{M}_n$ are positive semidefinite matrices. Denote by $\lambda_1(A)$ the largest eigenvalue of A. Then $\operatorname{tr}(AB) \leq \lambda_1(A)\operatorname{tr}(B)$.*

(ii) Let $A, B \in \mathbb{M}_n$ are Hermitian. Then

$$
\lambda_{n-r-s}(A+B) \geq \lambda_{n-r}(A) + \lambda_{n-s}(B).
$$

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

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

13

Lemma A.4 (Frobenius norm)**.** *Let* $A \in \mathbb{M}_n$. *Then*

(i)

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

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

(ii) It holds

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

(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) is well known and omitted. (iii) By assumption it holds $0 \le X^t A X \le X^t B X$. Hence $\lambda_i^2\left(X^t A X\right) \le \lambda_i^2\left(X^t B X\right)$ and the result follows. $\square$

Acknowledgments

References

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

[2] L. Brown, T. Cai, and H. Zhou. Robust nonparametric estimation via wavelet median regression. *Ann. Statist.*, 36(5):2055–2084, 2008.

[3] L. Brown and R. H. Farrell. A lower bound for the risk in estimating the value of a probability density. *J. Amer. Stat. Assoc.*, 85(412):1147–1153, 1990.

[4] L. Brown and M. Levine. Variance estimation in nonparametric regression via the difference sequence method. *Ann. Statist.*, 35(5):2219–2232, 2007.

[5] L. Brown and M. G. Low. Information inequality bounds on the minimax risk (with an application to nonparametric regression). *Ann. Statist.*, 19(1):329–337, 2007.

[6] L. Brown, Y. Wang, and L. H. Zhao. On the statistical equivalence at suitable frequencies of garch and stochastic volatility models with the corresponding diffusion model. *Statist. Sinica*, 13:993–1013, 2003.

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

[8] D. D. Cox. An analysis of bayesian inference for nonparametric regression. *Ann. Statist.*, 21(2):903–923, 1993.

[9] D. Freedman. On the Bernstein-von Mises theorem with infinite-dimensional parameters. *Ann. Statist.*, 27(4):1119–1140, 1999.

[10] A. Gloter and M. Hoffmann. Estimation of the Hurst parameter from discrete noisy data. *Ann. Statist.*, 35:1947–1974, 2007.

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

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

[13] G. Golubev, M. Nussbaum, and H. Zhou. Asymptotic equivalence of spectral density estimation and Gaussian white noise. *Ann. Statist.*, 38:181–214, 2010.

[14] S. J. Huang, Q. Liu, and J. Yu. Realized Daily Variance of S& P 500 Cash Index: A Revaluation of Stylized Facts. *Annals of Economics and Finance*, 1:33–56, 2007.

[15] J. Jacod, Y. Li, P. A. Mykland, M. Podolskij, and M. Vetter. Microstructure noise in the continuous case: The pre-averaging approach. *Stochastic Process. Appl.*, 119(7):2249–2276, 2009.

[16] A. Munk, N. Bissantz, T. Wagner, and G. Freitag. On difference based variance estimation in nonparametric regression when the covariate is high dimensional. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 67:19–41, 2005.

[17] A. Munk and F. Ruymgaart. Minimax rates for estimating the variance and its derivatives in nonparametric regression. *Austr. New Zeal. Journ. Statist.*, 44:479–488, 2002.

[18] A. Munk and J. Schmidt-Hieber. Nonparametric estimation of the volatility function in a high-frequency model corrupted by noise. 2009. arxiv:0908.3163, Math arXiv Preprint.

[19] M. Reiß. Asymptotic equivalence for nonparametric regression with multivariate and random design. *Ann. Statist.*, 36:1957–1982, 2008.

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

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

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

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

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

[25] B. Zhou. High-frequency data and volatility in foreign-exchange rates. *J. Business Econom. Statist.*, 14:45–52, 1996.