Model checks for the volatility under microstructure noise

Mathias Vetter and Holger Dette
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum
Germany
email: mathias.vetter@rub.de; holger.dette@rub.de
FAX: +49 2 34 32 14 559

July 1, 2009

Abstract

We consider the problem of testing the parametric form of the volatility for high frequency data. It is demonstrated that in the presence of microstructure noise commonly used tests do not keep the preassigned level and are inconsistent. The concept of pre-averaging is used to construct new tests, which do not suffer from these drawbacks. These tests are based on a Kolmogorov or Cramér-von-Mises functional of an integrated stochastic process, for which weak convergence to a (conditional) Gaussian process is established. The finite sample properties of a bootstrap version of the test are illustrated by means of a simulation study.

AMS Subject Classification: 62M02, 62G10, 62P20
Keywords and phrases: goodness-of-fit test, microstructure noise, stable convergence, parametric bootstrap, heteroscedasticity

1 Introduction

The volatility is a popular measure of risk in finance with numerous applications including the construction of optimal portfolios, hedging and pricing of options. Therefore estimating and investigating the volatility and its dynamics is of particular importance in applications and numerous models have been proposed for this purpose [see e.g. Black and Scholes [6], Vasicek [25], Cox *et al.* [8], Hull and White [15] and Heston [14] among many others]. Because the misspecification of the form of the volatility can lead to serious consequences in the subsequent data analysis numerous authors recommend to use goodness-of-fit tests for the postulated model [see e.g. Ait-Sahalia [2], Corradi and White [9], Dette *et al.* [11], Dette and Podolskij [10] among others]. The literature on statistical inference in this context can be divided into two classes depending on the type of available data. The first class of goodness-of-fit tests can be used, when the available data consists of discrete observations of the process sampled at time points Δ, 2Δ, $3\Delta, \ldots, n\Delta$, where $\Delta > 0$ is fixed and $n \to \infty$. The other class of tests addresses the situation of high frequency data, where discretely observed data of the price process is available at time $0, \Delta, 2\Delta, \ldots, n\Delta = T$, where T is fixed and $n \to \infty$ (which means that $\Delta \to 0$ for an increasing sample size).

In the present paper we consider the case of high frequency data, where - in principle - for an increasing sample size information about the whole path of the volatility would be available. However, in concrete applications the situation is much more complicated because of the presence of microstructure noise, which is usually existent in high frequency data. This additional noise is caused by many sources of the trading process such as discreteness of observations [see e.g. Harris [19], [20]], bid-ask bounces or special properties of the trading mechanism [see e.g. Black [5] or Amihud and Mendelson [4]]. While microstructure noise has been taken into account for the construction of estimators of the integrated volatility and other related quantities [see e.g. Zhang *et al.* [27], Jacod *et al.* [17] or Podolskij and Vetter [21], [22]], properties of goodness-of-fit tests in this context have not been investigated so far in the literature.

Consider for example the problem, where the process $\{Z_t\}_{t\in[0,1]}$ is observed at the n time points $1/n, 2/n, \ldots, 1$. Under the assumption that

$$Z_t = X_t \qquad \text{with} \qquad dX_t = \sigma_t \, dW_t \tag{1.1}$$

Dette and Podolskij [10] proposed to reject the hypothesis of a constant diffusion coefficient in (1.1), i.e. $H_0 : \sigma_t^2 = \sigma^2(t, X_t) = \sigma^2$, whenever

$$T_n = \sqrt{n} \sup_{t\in[0,1]} \left| \frac{\sum_{k=1}^{\lfloor nt \rfloor} |Z_{\frac{k}{n}} - Z_{\frac{k-1}{n}}|^2 - t \sum_{k=1}^{n} |Z_{\frac{k}{n}} - Z_{\frac{k-1}{n}}|^2}{\sqrt{2} \sum_{k=1}^{n} |Z_{\frac{k}{n}} - Z_{\frac{k-1}{n}}|^2} \right| > c_{1-\alpha} \,, \tag{1.2}$$

where $c_{1-\alpha}$ denotes the $(1-\alpha)$ quantile of the supremum of a Brownian Bridge. Now consider the situation, where microstructure noise is present, which is usually modeled by an additional additive component, that is

$$Z_{\frac{i}{n}} = X_{\frac{i}{n}} + U_{\frac{i}{n}}, \; i = 1, \ldots, n \tag{1.3}$$

where $\{U_{\frac{i}{n}} | \; i = 1, \ldots, n\}$ denotes a triangular array of random variables with mean 0 and variance ω^2. In Table 1 we show the finite sample behaviour of the test (1.2) for the hypothesis of a constant volatility if $\sigma_t^2 = \sigma^2(t, x) = \theta + (1-\theta)x^2$ (note that the case $\theta = 1$ corresponds to the null hypothesis). We observe that the test keeps its preassigned level only in the case where ω is rather small. In most cases the nominal level is clearly underestimated. On the other hand, the test is not able to detect any alternative. An intuitive explanation for this behaviour is that in the presence of microstructure noise the variances of the differences $Z_{\frac{k}{n}} - Z_{\frac{k-1}{n}}$ are dominated by the term ω^2. This leads to inconsistent estimates of the integrated volatility as pointed out in Zhang [27]. More precisely, a straightforward calculation shows that under microstructure noise the statistic T_n shows the same asymptotic behavior as the the statistic

$$\sqrt{n} \sup_{t\in[0,1]} \left| \frac{\sum_{k=1}^{\lfloor nt \rfloor} |U_{\frac{k}{n}} - U_{\frac{k-1}{n}}|^2 - t \sum_{k=1}^{n} |U_{\frac{k}{n}} - U_{\frac{k-1}{n}}|^2}{\sqrt{2} \sum_{k=1}^{n} |U_{\frac{k}{n}} - U_{\frac{k-1}{n}}|^2} \right|, \tag{1.4}$$

which converges weakly to

$$\sqrt{\frac{\lambda}{2}} \sup_{t\in[0,1]} |B_t|,$$

no matter if the null hypothesis is valid or not. Here B_t denotes a Brownian bridge and $\lambda = E[(U_{k/n}/\omega)^4]$. This means that in the presence of microstructure noise the test (1.2) has asymptotic level α if and only if $\lambda = 2$. In all other cases the test does not keep its preassigned level. Moreover, because the asymptotic properties under null hypothesis and alternative are the same, the test is not consistent.

[INSERT TABLE 1 HERE]

The present paper is devoted to the problem of constructing a consistent asymptotic level α test for a general parametric form of the volatility in the presence of microstructure noise. In Section 2 and 3 we present the basic model and introduce a stochastic process which can be used to test parametric hypotheses about the form of the volatility in models with microstructure noise. For this purpose we use the concept of pre-averaging, which was introduced in Podolskij and Vetter [21] and extended in several other papers [see e.g. Jacod *et al.* [17] or Podolskij and Vetter [22]] in the context of volatility estimation. Our main results are presented in Section 4, where we prove stable convergence of two stochastic processes which will form the basis of the proposed new tests for the parametric form of the volatility. The new tests can detect alternatives converging to the null hypothesis with a rate $n^{-1/4}$ and therefore achieve the optimal rate of convergence in problems of this type [see Gloter and Jacod [13]]. Section 5 deals with the problem of testing nonlinear hypotheses for the volatility. Roughly speaking, this situation can be reduced to the linear case using standard arguments from nonlinear regression models [see Seber and Wild [24]], but there appear interesting differences in the asymptotic distribution of the process, if the null hypothesis is not satisfied. In Section 6 we investigate the finite sample properties of a bootstrap version of the new tests and investigate the effect of microstructure noise in the context of goodness-of-fit testing. In particular, it is demonstrated that the new tests based on the concept of pre-averaging provide a satisfactory solution to the problem of checking model assumptions in the presence of microstructure noise. Finally, all proofs of the results and technical details are presented in an Appendix.

2 Testing parametric hypotheses for the volatility

Suppose that the process $X = (X_t)_t$ is defined on some appropriate filtered probability space $(\Omega^{(0)}, \mathcal{F}^{(0)}, (\mathcal{F}_t^{(0)})_{t\in[0,1]}, P^{(0)})$ and admits the representation

$$X_t = X_0 + \int_0^t a_s \, ds + \int_0^t \sigma_s \, dW_s, \tag{2.1}$$

where $W = (W_t)_t$ is a standard Brownian motion and the drift process a and the volatility process σ satisfy some weak regularity conditions, which will be specified later. Furthermore, we assume that the process can be observed at discrete points on a fixed time interval, say $[0, 1]$.

Various assumptions on the structure of the volatility process have been proposed in the literature, typically depending on the financial asset, whose price process is modeled by X. Among such models, a large class involves the case where σ is defined to be a local volatility process, thus merely a function of time and state [see e.g. Black and Scholes [6], Vasicek [25], Cox *et al.* [8], Chan *et al.* [7], Ait-Sahalia [2] or Ahn and Gao [3] among many others]. Because an appropriate modeling of the volatility is of particular importance for the construction of portfolios, hedging and pricing, many authors point out that the postulated model should be validated by an appropriate goodness-of-fit test [see e.g. Ait-Sahalia [2] or Corradi and White [9]]. In several cases the hypothesis for the parametric form of the volatility is linear and one has to consider the following two situations:

$$H_0 : \sigma_t^2 = \sigma^2(t, X_t) = \sum_{i=1}^d \theta_i \; \sigma_i^2(t, X_t) \quad \text{a.s.} \tag{2.2}$$

or

$$\bar{H}_0 : \sigma_t = \sigma(t, X_t) = \sum_{i=1}^d \bar{\theta}_i \; \bar{\sigma}_i(t, X_t) \quad \text{a.s.}, \tag{2.3}$$

where the functions $\sigma_1, \ldots, \sigma_d$ (or $\bar{\sigma}_1, \ldots, \bar{\sigma}_d$) are known and the parameters $\theta_1, \ldots, \theta_d$ (or $\bar{\theta}_1, \ldots, \bar{\theta}_d$) are unknown. Other models involve volatility functions, where the parameters enter nonlinearly [see Ait-Sahalia [2]] and the corresponding hypotheses will be considered later in Section 5, because the basic concepts are easier to explain in the linear context.

Let us focus on the problem raised in (2.2) for the moment, as the testing problem in (2.3) can be treated in the same way. Dette and Podolskij [10] proposed to construct a test statistic using an empirical version of the stochastic process

$$N_t := \int_0^t \left\{ \sigma_s^2 - \sum_{j=1}^d \theta_j^{min} \; \sigma_j^2(s, X_s) \right\} ds, \tag{2.4}$$

where

$$\theta^{min} = (\theta_1^{min}, \dots, \theta_d^{min})^T := \text{argmin}_{\theta \in \mathbb{R}^d} \int_0^1 \Big\{ \sigma_s^2 - \sum_{j=1}^d \theta_j \; \sigma_j^2(s, X_s) \Big\}^2 \; ds.$$

Thus, one uses the L^2 distance to determine the best approximation to the unknown volatility process σ^2 by a linear combination of the given functions $\sigma_1^2, \dots, \sigma_d^2$. It can easily be seen that the null hypothesis in (2.2) is equivalent to

$$N_t = 0 \quad \forall \; t \in [0,1] \quad \text{a.s.},$$

and a well-known result from Hilbert space theory [see Achieser [1]] implies that

$$N_t = B_t^0 - B_t^T D^{-1} C, \tag{2.5}$$

where

$$B_t^0 = \int_0^t \sigma_s^2 \; ds \qquad \text{and} \qquad B_t^i = \int_0^t \sigma_i^2(s, X_s) \; ds \quad \text{for } i = 1, \dots, d,$$

D and C denote a $d \times d$-matrix and a d-dimensional vector, respectively, with

$$D_{ij} = \int_0^1 \sigma_i^2(s, X_s) \; \sigma_j^2(s, X_s) \; ds \qquad \text{and} \qquad C_i = \int_0^1 \sigma_s^2 \; \sigma_i^2(s, X_s) \; ds.$$

Note that these quantities depend on the particular path of the process.
In practice, one does not observe the entire path of the diffusion process $X = (X_t)_t$ and it is therefore necessary to define an empirical version based on appropriate estimators for the quantities in (2.5). Let us briefly discuss the solution to the problem in the case, where the diffusion process $X = (X_t)_t$ can be observed at the discrete times $t_{n,i} = \frac{i}{n}$ $(0 \le i \le n)$ without further restrictions. Based on the decomposition above, Dette and Podolskij [10] propose to define an empirical version

$$\tilde{N}_t = \tilde{B}_t^0 - \tilde{B}_t^T \tilde{D}^{-1} \tilde{C}$$

plugging in appropriate estimators for the unknown quantities. Quite naturally, one uses a Riemann approximation of each integral, where one chooses $n|X_{\frac{k}{n}} - X_{\frac{k-1}{n}}|^2$ as a local estimate for $\sigma^2_{\frac{k-1}{n}}$. Thus,

$$\begin{aligned} \tilde{D}_{ij} &= \frac{1}{n} \sum_{k=1}^n \sigma_i^2(\frac{k}{n}, X_{\frac{k}{n}}) \; \sigma_j^2(\frac{k}{n}, X_{\frac{k}{n}}) \quad \text{for } i,j = 1, \dots, d, \\ \tilde{C}_i &= \sum_{k=1}^n \sigma_i^2(\frac{k-1}{n}, X_{\frac{k-1}{n}}) \; |X_{\frac{k}{n}} - X_{\frac{k-1}{n}}|^2 \quad \text{for } i = 1, \dots, d, \end{aligned} \tag{2.6}$$

and the quantities $\tilde{B}_t^0$ and $\tilde{B}_t = (\tilde{B}_t^1, \ldots, \tilde{B}_t^d)^T$ are given by

$$\tilde{B}_t^0 := \sum_{k=1}^{\lfloor nt \rfloor} |X_{\frac{k}{n}} - X_{\frac{k-1}{n}}|^2, \qquad \tilde{B}_t^i := \frac{1}{n} \sum_{k=1}^{\lfloor nt \rfloor} \sigma_i^2(\frac{k}{n}, X_{\frac{k}{n}}) \quad \text{for } i = 1, \ldots, d. \tag{2.7}$$

In this context one can prove a (stable) central limit theorem for the process $(\tilde{N}_t - N_t)_t$ with the optimal rate of convergence $n^{-\frac{1}{2}}$, from which one may construct test statistics of Cramér-von-Mises or Kolmogorov-Smirnov type. For example, if $d = 1$, $\sigma_1^2(t, X_t) = 1$, the hypothesis (2.2) reduces to the hypothesis of constant volatility considered in the introduction. To be precise, we have $\tilde{D} = 1$, $\tilde{C}_1 = \sum_{k=1}^n |X_{\frac{k}{n}} - X_{\frac{k-1}{n}}|^2$, $\tilde{B}_t^0 = \sum_{k=1}^{\lfloor nt \rfloor} |X_{\frac{k}{n}} - X_{\frac{k-1}{n}}|^2$, and $\tilde{B}_t^1 = \frac{\lfloor nt \rfloor}{n} \approx t$ and we obtain the process in the Kolmogorov-Smirnov statistic (1.2), which converges (stably) to the supremum of a Brownian bridge [see Dette and Podolskij [10]].
However, as pointed out in the introduction diffusion processes observed at high frequency are contaminated by microstructure effects such as rounding or bid-ask bounces. In particular, in the presence of microstructure noise the corresponding test for the hypothesis (2.2) does not keep its asymptotic level and is not consistent. Thus a modification of the corresponding test statistics is necessary, which will be discussed in the following sections.

3 Assumptions and definitions

In the case of microstructure noise it is less obvious how to estimate the unknown quantities in (2.5), basically for two reasons: One has to to find a local estimator for the unknown volatility function σ_t^2 (which has to be done in the noiseless framework as well, but becomes more complicated in this setting), and one needs an estimator for the path X_t itself, which cannot be observed directly. We solve both questions by applying the idea of pre-averaging, which was introduced in Podolskij and Vetter [21] and extended in several other papers [see e.g. Jacod *et al.* [17] or Podolskij and Vetter [22]] in the context of volatility estimation. Let us start with some basic assumptions.
Since we are dealing with microstructure noise, we have to define a second process $Z = (Z_t)_t$, which is connected to the underlying Ito semimartingale X through the equation

$$Z_t = X_t + U_t$$

for some noise process U. Even though we assume in the following that the observation times are given by $t_{n,i} = \frac{i}{n}$ for $0 \leq i \leq n$, it will be convenient to define the observed process (and thus the noise process as well, even though it will typically not be measurable in time) for any t. For this purpose we use a similar setting as in Jacod *et al.* [17].
We consider for each t in $[0, 1]$ a probability measure $Q_t(\omega^{(0)}, dz)$, which corresponds to the transition from $X_t(\omega^{(0)})$ to the observed process Z_t on $\mathbb{R}$. Thus, it is natural to define the space of observations $\Omega^{(1)} = \mathbb{R}^{[0,1]}$, equipped with its product Borel-σ-field $\mathcal{F}^{(1)}$ and the probability

measure $P^{(1)}(\omega^{(0)}, d\omega^{(1)})$, which is the product $\otimes_{t\in[0,1]} Q_t(\omega^{(0)}, \cdot)$ to ensure some sort of (conditional) independence of the noise variables. $(Z_t)_t$ is then given as the canonical process on $(\Omega^{(1)}, \mathcal{F}^{(1)}, P^{(1)})$ with the natural filtration $\mathcal{F}_t^{(1)} = \sigma(Z_s; s \leq t)$. The filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,1]}, P)$, on which both processes X and Z live, is then defined as

$$\left.\begin{array}{l} \Omega = \Omega^{(0)} \times \Omega^{(1)}, \qquad \mathcal{F} = \mathcal{F}^{(0)} \times \mathcal{F}^{(1)}, \qquad \mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s^{(0)} \times \mathcal{F}_s^{(1)}, \\ P(d\omega^{(0)}, d\omega^{(1)}) = P^{(0)}(d\omega^{(0)}) P^{(1)}(\omega^{(0)}, d\omega^{(1)}). \end{array}\right\} \tag{3.1}$$

This setting allows for quite general forms of noise; however, we restrict ourselves to the case of i.i.d. noise, thus the transition probability $Q_t(\omega^{(0)}, dz)$ depends on $\omega^{(0)}$ only through $z - X_t(\omega^{(0)})$ and has the form

$$Q_t(\omega^{(0)}, dz) = k\Big(z - X_t(\omega^{(0)})\Big)\ dz,$$

where k is a density with bounded support. Furthermore, we assume that the moment conditions

$$E[U_t] = 0, \qquad E[U_t^2] = \omega^2, \qquad E[U_t^4] < \infty \tag{3.2}$$

hold.
In order to introduce the pre-averaged statistics we have to define some further quantities. First, we choose a sequence m_n, such that

$$\frac{m_n}{\sqrt{n}} = \kappa + o(n^{-\frac{1}{4}}) \tag{3.3}$$

for some $\kappa > 0$, and a nonzero real-valued function $g : \mathbb{R} \to \mathbb{R}$, which vanishes outside of the interval $(0, 1)$, is continuous and piecewise C^1 and has a piecewise Lipschitz derivative g'. We associate with g (and n) the following real valued numbers and functions:

$$\left.\begin{array}{l} g_j^n = g(\frac{j}{m_n}), \qquad g_j'^n = g_j^n - g_{j+1}^n, \qquad \psi_1 = \int_0^1 (g'(s))^2\ ds, \qquad \psi_2 = \int_0^1 (g(s))^2\ ds \\ s \in [0,1] \ \mapsto\ \phi_1(s) = \int_s^1 g'(u) g'(u-s)\ du, \qquad \phi_2(s) = \int_s^1 g(u) g(u-s)\ du \\ i,j = 1,2: \qquad \Phi_{ij} = \int_0^1 \phi_i(s)\phi_j(s)\ ds \end{array}\right\} \tag{3.4}$$

Furthermore, we define for an arbitrary process V the random variables

$$V_j^n = V_{\frac{j}{n}}, \qquad \Delta_j^n V = V_j^n - V_{j-1}^n, \qquad \overline{V}_k^n = \sum_{j=1}^{m_n} g_j^n \Delta_{k+j}^n V. \tag{3.5}$$

Typically, we have $V = X$ or Z and for these processes $\overline{V}_k^n$ can be represented as

$$\overline{V}_k^n = \int_{\frac{k}{n}}^{\frac{k+m_n}{n}} g_n\Big(s - \frac{k}{n}\Big)\ dV_s \qquad \text{with} \qquad g_n(s) = \sum_{j=1}^{m_n} g_j^n 1_{(\frac{j-1}{n}, \frac{j}{n}]}(s), \tag{3.6}$$

where we use the convention $\int_a^b c\, dU_s = c(U_b - U_a)$ for arbitrary constants a, b and c. Finally, we set

$$\hat{X}_{\frac{k}{n}} = \frac{1}{m_n} \sum_{j=1}^{m_n} Z^n_{k+j}. \tag{3.7}$$

As pointed out before, we need additional assumptions on the process X as well as on the given basis functions in (2.2) and (2.3), respectively. Since the conditions on σ_i^2 and $\bar{\sigma}_i$ are similar, we will restrict ourselves to the first case only.
It is required that the functions $\sigma_1^2, \ldots, \sigma_d^2$ are linearly independent and that each σ_i^2 is twice continuously differentiable. Moreover, we assume that

$$E[|\det(D)|^{-\beta}] < \infty \tag{3.8}$$

for some $\beta > 0$.
Regarding the various processes in X, the assumptions are as weak as possible when testing for (2.2). We simply have to ensure that the process in (2.1) is well-defined, which follows if we assume that a is locally bounded and predictable and that σ is càdlàg. [see Jacod and Shiryaev [18] or Revuz and Yor [23]]. When working with (2.3) we propose additionally that the true volatility process σ is almost surely positive and that is has a representation of the form (2.1) as well, namely that it satisfies

$$\sigma_t = \sigma_0 + \int_0^t a'_s\, ds + \int_0^t \sigma'_s\, dW_s + \int_0^t v'_s\, dV_s,$$

where a', σ' and v' are adapted càdlàg processes, with a' also being predictable and locally bounded, and V is a second Brownian motion, independent of W.

4 Goodness-of-fit tests addressing microstructure noise

The two estimators of interest are $\overline{Z}^n_k$ and $\hat{X}_{\frac{k}{n}}$, which are both local averages of the noisy data, but with slightly different intuitions behind them. For the latter one, the filtering applies to the observations directly, and it is easy to see that such a procedure reduces the impact of the noise variables around time $\frac{k}{n}$ and still provides information about the latent price $X_{\frac{k}{n}}$, since the path of X does not fluctuate too much. For $\overline{Z}^n_k$, the averaging happens on the increments rather than on the prices, but due to the assumptions on g the interpretation is similar: one reduces the noise effects, but keeps information about the increments of X.
We start with the construction of a test for the hypothesis (2.2) again. Local estimators for σ^2 can be obtained from $|\overline{Z}^n_k|^2$, but it is well known that this quantity is not an unbiased estimate (it contains an intrinsic bias due to the noise variables U) and has a different stochastic order

than the increments $X_{\frac{k}{n}} - X_{\frac{k-1}{n}}$ in the no-noise case. Thus, we define

$$\hat{\omega}_n^2 := \frac{1}{2n} \sum_{i=1}^{n} |\Delta_i^n Z|^2,$$

which is a consistent estimator for ω^2, see Zhang *et al.* [27]. Mimicing the procedure from the no-noise case presented in Section 2, we set

$$\hat{D}_{ij} := \frac{1}{n} \sum_{k=1}^{n-m_n} \sigma_i^2(\frac{k}{n}, \hat{X}_{\frac{k}{n}})\ \sigma_j^2(\frac{k}{n}, \hat{X}_{\frac{k}{n}}) \quad \text{for } i,j = 1,\dots,d, \tag{4.1}$$

$$\hat{C}_i := \frac{1}{\kappa\psi_2} n^{-\frac{1}{2}} \sum_{k=1}^{n-m_n} \sigma_i^2(\frac{k}{n}, \hat{X}_{\frac{k}{n}})\ \Big(|\overline{Z}_k^n|^2 - n^{-\frac{1}{2}}\frac{\psi_1}{\kappa}\hat{\omega}_n^2\Big) \quad \text{for } i = 1,\dots,d, \tag{4.2}$$

as well as

$$\hat{B}_t^0 := \frac{1}{\kappa\psi_2} n^{-\frac{1}{2}} \sum_{k=1}^{\lfloor nt \rfloor - m_n} \Big(|\overline{Z}_k^n|^2 - n^{-\frac{1}{2}}\frac{\psi_1}{\kappa}\hat{\omega}_n^2\Big) \tag{4.3}$$

and

$$\hat{B}_t^i := \frac{1}{n} \sum_{k=1}^{\lfloor nt \rfloor - m_n} \sigma_i^2(\frac{k}{n}, \hat{X}_{\frac{k}{n}}) \quad \text{for } i = 1,\dots,d. \tag{4.4}$$

We define at last the process

$$\hat{N}_t = \hat{B}_t^0 - \hat{B}_t^T \hat{D}^{-1} \hat{C}, \tag{4.5}$$

which turns out to be an appropriate estimate of the process $\{N_t\}_{t\in[0,1]}$ defined in (2.4). Our first result specifies the asymptotic properties of the process $\{A_n(t)\}_{t\in[0,1]}$ with

$$A_n(t) = n^{\frac{1}{4}}(\hat{N}_t - N_t). \tag{4.6}$$

Theorem 1 *If the assumptions stated in the previous sections are satisfied, the process $(A_n(t))_{t\in[0,1]}$ defined in (4.6) converges weakly in $D[0,1]$ to a mean zero process $(A(t))_{t\in[0,1]}$. Conditionally on $\mathcal{F}$ the limiting process is Gaussian, and its finite dimensional distributions coincide with the conditional (with respect to $\mathcal{F}$) finite dimensional distributions of the process*

$$\Big\{\gamma_V\Big(I\{V \le t\} - B_t^T D^{-1} g(V, X_V)\Big) - \Big(\int_0^t \gamma_s\ ds - B_t^T D^{-1} \int_0^1 \gamma_s\ g(s, X_s)\ ds\Big)\Big\}_{t\in[0,1]}, \tag{4.7}$$

where $V \sim \mathcal{U}[0,1]$,

$$g(V, X_V) = (\sigma_1^2(V, X_V), \dots, \sigma_d^2(V, X_V))^T \tag{4.8}$$

and

$$\gamma_s^2 = \frac{4}{\psi_2^2}\Big(\Phi_{22}\kappa\sigma_s^4 + 2\Phi_{12}\frac{\sigma_s^2\omega^2}{\kappa} + \Phi_{11}\frac{\omega^4}{\kappa^3}\Big). \tag{4.9}$$

Note that the rate of convergence $n^{-\frac{1}{4}}$ is optimal for this problem, since it is already optimal for the estimation of B_t^0 even in a parametric setting [cf. Gloter and Jacod [13]].
In order to construct a test statistic based on Theorem 1 we have to define an appropriate estimator for the conditional variance of the process $\{A(t)\}_{t\in[0,1]}$, which is given by

$$s_t^2 = \int_0^t \gamma_s^2 \, ds - 2B_t^T D^{-1} \int_0^t \gamma_s^2 g(s, X_s) \, ds + B_t^T D^{-1} \int_0^1 \gamma_s^2 g(s, X_s) g^T(s, X_s) \, ds \; D^{-1} B_t.$$

Obviously, we use $\hat{B}_t$ and $\hat{D}$ as the empirical counterparts for B_t and D. In order to obtain estimates for the other random elements of s_t^2, we define

$$\begin{aligned} \Gamma_k &= \frac{4\,\Phi_{22}}{3\,\kappa\,\psi_2^4}\,|\overline{Z}_k^n|^4 + n^{-\frac{1}{2}}\,\frac{8}{\kappa^2}\Big(\frac{\Phi_{12}}{\psi_2^3} - \frac{\Phi_{22}\,\psi_1}{\psi_2^4}\Big)\,|\overline{Z}_k^n|^2\,\hat{\omega}^2 \\ &+ n^{-1}\frac{4}{\kappa^3}\Big(\frac{\Phi_{11}}{\psi_2^2} - \frac{2\,\Phi_{12}\,\psi_1}{\psi_2^3} + \frac{\Phi_{22}\,\psi_1^2}{\psi_2^4}\Big)\,\hat{\omega}^4 \end{aligned}$$

as a local estimator for the process γ^2 and observe that [see Jacod *et al.* [17]]

$$\begin{aligned} \hat{g}_0(t) &:= \sum_{k=1}^{\lfloor nt\rfloor - m_n} \Gamma_k \xrightarrow{P} \int_0^t \gamma_s^2 \, ds \\ \hat{g}_i(t) &= \sum_{k=1}^{\lfloor nt\rfloor - m_n} \Gamma_k \,\sigma_i^2(\frac{k-1}{n}, \hat{X}_{\frac{k-1}{n}}) \xrightarrow{P} \int_0^t \gamma_s^2 \,\sigma_i^2(s, X_s) \, ds \\ \hat{g}_{ij} &= \sum_{k=1}^{n} \Gamma_k \,\sigma_i^2(\frac{k-1}{n}, \hat{X}_{\frac{k-1}{n}})\,\sigma_j^2(\frac{k-1}{n}, \hat{X}_{\frac{k-1}{n}}) \xrightarrow{P} \int_0^1 \gamma_s^2 \,\sigma_i^2(s, X_s)\,\sigma_j^2(s, X_s) \, ds. \end{aligned}$$

Inserting these estimators in the corresponding elements of s_t^2 gives the consistent estimator, that is

$$\hat{s}_t^2 = \hat{g}_0(t) - 2\hat{B}_t^T \hat{D}^{-1} \hat{g}(t) + \hat{B}_t^T \hat{D}^{-1} \hat{G} \hat{D}^{-1} \hat{B}_t, \tag{4.10}$$

where $\hat{g}(t) = (\hat{g}_1(t), \ldots, \hat{g}_d(t))^T$ and $\hat{G} = (\hat{g}_{ij})_{i,j=1}^d$. A consistent test for the hypothesis (2.2) is now obtained by rejecting the null hypothesis for large values of Kolmogorov-Smirnov or Cramér-van-Mises functional of the process

$$\left\{ \frac{n^{1/4} \hat{N}_t}{\hat{s}_t} \right\}_{t\in[0,1]}.$$

In principle a similar approach can be used to construct a test for the hypothesis (2.3). However, in this case things change considerably. Quite naturally, Dette and Podolskij [10] restate this hypothesis as

$$M_t = 0 \quad \forall\, t \in [0,1] \quad \text{a.s.},$$

where

$$M_t := \int_0^t \left\{ \sigma_s - \sum_{j=1}^d \bar{\theta}_j^{min} \bar{\sigma}_j(s, X_s) \right\} ds \tag{4.11}$$

and

$$\bar{\theta}^{min} = (\bar{\theta}_1^{min}, \ldots, \bar{\theta}_d^{min})^T := \mathrm{argmin}_{\bar{\theta} \in \mathbb{R}^d} \int_0^1 \left\{ \sigma_s - \sum_{j=1}^d \bar{\theta}_j \bar{\sigma}_j(s, X_s) \right\}^2 ds.$$

Obviously, we have an analogous representation as in (2.5), namely $M_t = R_t^0 - R_t^T Q^{-1} S$, where

$$R_t^0 = \int_0^t \sigma_s \, ds \qquad \text{and} \qquad R_t^i = \int_0^t \bar{\sigma}_i(s, X_s) \, ds \quad \text{for } i = 1, \ldots, d,$$

and Q and S are a $d \times d$-matrix and a d-dimensional vector, respectively, with

$$Q_{ij} = \int_0^1 \bar{\sigma}_i(s, X_s) \, \bar{\sigma}_j(s, X_s) \, ds \qquad \text{and} \qquad S_i = \int_0^1 \sigma_s \, \bar{\sigma}_i(s, X_s) \, ds.$$

However, an appropriate definition of an empirical version of the form

$$\hat{M}_t = \hat{R}_t^0 - \hat{R}_t^T \hat{Q}^{-1} \hat{S}$$

requires some less obvious modifications, because local estimators for σ_s are more difficult to obtain in this setting. Using a pre-averaged estimator of the form $|\overline{Z}_k^n|$ again causes an intrinsic bias, but due to the absolute value (instead of the square as in the previous setting) its correction turns out to be impossible at the optimal rate. However, it has been argued in Podolskij and Vetter [21] that using in (3.3) a sequence of a larger magnitude than $n^{\frac{1}{2}}$ reduces the impact of the noise terms in $\overline{Z}_k^n$. This modification makes inference about σ_s possible, though resulting in a worse rate of convergence. To be precise, we fix some $\delta > \frac{1}{6}$ and choose l_n such that

$$\frac{l_n}{n^{\frac{1}{2}+\delta}} = \rho + o(n^{-(\frac{1}{4}+\frac{\delta}{2})})$$

for some $\rho > 0$. Using the sequence l_n instead of m_n, we define all quantities from (3.4) to (3.7) in the straightforward way. Next we set

$$\hat{S}_i = \frac{1}{\mu_1 \sqrt{\rho \psi_2}} n^{-(\frac{3}{4}+\frac{\delta}{2})} \sum_{k=1}^{n-l_n} \bar{\sigma}_i(\frac{k}{n}, \hat{X}_{\frac{k}{n}}) \, |\overline{Z}_k^n| \quad \text{for } i = 1, \ldots, d,$$

and

$$\hat{R}_t^0 \;\; = \;\; \frac{1}{\mu_1 \sqrt{\rho \psi_2}} n^{-(\frac{3}{4}+\frac{\delta}{2})} \sum_{k=1}^{\lfloor nt \rfloor - l_n} |\overline{Z}_k^n|,$$

where μ_1 denotes the first absolute moment of a standard normal distribution. Moreover, it is natural to use the following estimators $\hat{R}_t = (\hat{R}_{\hat{t}}, \dots, \hat{R}_t^d)^T$ and $\hat{Q} = (\hat{Q}_{ij})_{i,j=1}^d$, for the quantities R_t and Q:

$$\hat{R}_t^i := \frac{1}{n} \sum_{k=1}^{\lfloor nt \rfloor - l_n} \bar{\sigma}_i(\frac{k}{n}, \hat{X}_{\frac{k}{n}}) \quad \text{for } i = 1, \dots, d$$

and

$$\hat{Q}_{ij} = \frac{1}{n} \sum_{k=1}^{n-l_n} \bar{\sigma}_i(\frac{k}{n}, \hat{X}_{\frac{k}{n}})\ \bar{\sigma}_j(\frac{k}{n}, \hat{X}_{\frac{k}{n}}) \quad \text{for } i = 1, \dots, d.$$

Finally, we define

$$B_n(t) = n^{\frac{1}{4} - \frac{\delta}{2}}(\hat{M}_t - M_t) \tag{4.12}$$

for any $t \in [0,1]$ and obtain the following result.

Theorem 2 *If the assumptions stated in the previous sections are satisfied, the process $(B_n(t))_{t \in [0,1]}$ defined in (4.12) converges weakly in $D[0,1]$ to a mean zero process $(B(t))_{t \in [0,1]}$. Conditionally on $\mathcal{F}$ the limiting process is Gaussian, and its finite dimensional distributions coincide with the conditional (with respect to $\mathcal{F}$) finite dimensional distributions of the process*

$$\Big\{\bar{\gamma}_V\Big(I\{V \le t\} - R_t^T Q^{-1} \bar{g}(V, X_V)\Big) - \Big(\int_0^t \bar{\gamma}_s\ ds - R_t^T Q^{-1} \int_0^1 \bar{\gamma}_s\ \bar{g}(s, X_s)\ ds\Big)\Big\}_{t \in [0,1]}, \tag{4.13}$$

where $V \sim \mathcal{U}[0,1]$, $\bar{g}(V, X_V) = (\bar{\sigma}_1(V, X_V), \dots, \bar{\sigma}_d(V, X_V))^T$ and

$$\begin{aligned} \bar{\gamma}_s^2 &= \frac{2\rho\Xi}{\mu_1^2}\sigma_s^2, \\ \Xi &= \int_0^1 \xi(s)\ ds, \quad \xi(s) = f\Big(\frac{\phi_2(s)}{\psi_2}\Big), \\ f(u) &= \frac{2}{\pi}\Big(u \arcsin(u) + \sqrt{1-u^2} - 1\Big). \end{aligned}$$

The estimation of the conditional variance of the process $\{B(t)\}_{t \in [0,1]}$

$$r_t^2 = \int_0^t \bar{\gamma}_s^2\ ds - 2R_t^T Q^{-1} \int_0^t \bar{\gamma}_s^2 \bar{g}(s, X_s)\ ds + R_t^T D^{-1} \int_0^1 \bar{\gamma}_s^2 \bar{g}(s, X_s) \bar{g}^T(s, X_s)\ ds\ Q^{-1} R_t.$$

becomes easier in this context. With the notation

$$\bar{\Gamma}_k = n^{-(\frac{1}{2}+\delta)}\ \frac{2\ \Xi}{\psi_2\ \mu_1^2}\ |\overline{Z}_k^n|^2,$$

we have

$$\begin{aligned}
\hat{h}_0(t) &= \sum_{k=1}^{\lfloor nt\rfloor - l_n} \bar{\Gamma}_k \xrightarrow{P} \int_0^t \bar{\gamma}_s^2 \, ds \\
\hat{h}_i(t) &= \sum_{k=1}^{\lfloor nt\rfloor - l_n} \bar{\Gamma}_k \; \bar{\sigma}_i(\frac{k-1}{n}, \hat{X}_{\frac{k-1}{n}}) \xrightarrow{P} \int_0^t \bar{\gamma}_s^2 \; \bar{\sigma}_i(s, X_s) \, ds \\
\hat{h}_{ij} &= \sum_{k=1}^{n} \bar{\Gamma}_k \; \bar{\sigma}_i(\frac{k-1}{n}, \hat{X}_{\frac{k-1}{n}}) \; \bar{\sigma}_j(\frac{k-1}{n}, \hat{X}_{\frac{k-1}{n}}) \xrightarrow{P} \int_0^1 \bar{\gamma}_s^2 \; \bar{\sigma}_i(s, X_s) \; \bar{\sigma}_j(s, X_s) \, ds
\end{aligned}$$

and consequently a consistent estimator $\hat{r}_t^2$ for the conditional variance is given by

$$\hat{r}_t^2 = \hat{h}_0(t) - 2\hat{R}_t^T \hat{Q}^{-1} \hat{h}(t) + \hat{R}_t^T \hat{Q}^{-1} \hat{H} \hat{Q}^{-1} \hat{R}_t, \tag{4.14}$$

where $\hat{h}(t) = (\hat{h}_1(t), \dots, \hat{h}_d(t))^T$ and $\hat{H} = (\hat{h}_{ij})_{i,j=1}^d$. A consistent test for the hypothesis (2.3) is now obtained by rejecting the null hypothesis for large values of the Kolmogorov-Smirnov or Cramér-van-Mises functional of the process

$$\left\{ \frac{n^{1/4-\delta/2} \hat{M}_t}{\hat{r}_t} \right\}_{t\in[0,1]} .$$

Note that one knows from previous work that it is neither necessary to define X to be an Ito semimartingale with continuous paths as in (2.1) nor to model the noise terms U as being independent and identically distributed to obtain similar results as in Theorem 1 and 2. In fact, for an underlying Ito semimartingale exhibiting jumps one can use bipower-type estimators as discussed in Podolskij and Vetter [22] in order to define an estimator closely related to $\hat{B}_t^0$. Moreover, it has been argued in Jacod *et al.* [17] that even for a noise process with càdlàg variance (depending on $\omega^{(0)}$) a similar theory as presented in this paper applies.

5 Nonlinear hypotheses

In this section we briefly discuss the case of a nonlinear hypothesis

$$H_0 : \sigma_t^2 = \sigma^2(t, X_t) = \sigma^2(t, X_t, \theta), \tag{5.1}$$

where $\theta \in \Theta \subset \mathbb{R}^d$ denotes the unknown parameter. Under suitable conditions on the parameter space Θ, H_0 can be restated as $N_t = 0 \;\forall\, t \in [0,1]$ a.s., where the process $\{N_t\}_{t\in[0,1]}$ is defined by

$$N_t = B_t^0 - B_t(\theta_0) := \int_0^t \left\{ \sigma_s^2 - \sigma^2(s, X_s, \theta_0) \right\} ds.$$

Here, θ_0 is the parameter corresponding to the best L^2-approximation of σ_s^2 by the parametric class, that is

$$\theta_0 = \text{argmin}_{\theta\in\Theta}\ g(\theta), \quad \text{where} \quad g(\theta) = \int_0^t \Big\{\sigma_s^2 - \sigma^2(s, X_s, \theta)\Big\}^2\, ds.$$

An analogue of the process $\hat{N}_t$ introduced in (4.5) is given by

$$\hat{N}_t = \hat{B}_t^0 - \hat{B}_t(\hat{\theta}), \tag{5.2}$$

where $\hat{B}_t^0$ is defined by (4.3),

$$\hat{B}_t(\hat{\theta}) = \frac{1}{n} \sum_{k=1}^{\lfloor nt\rfloor - m_n} \sigma^2(\frac{k}{n}, \hat{X}_{\frac{k}{n}}, \hat{\theta}), \tag{5.3}$$

$$\hat{\theta} = \text{argmin}_{\theta\in\Theta}\ g_n(\theta), \quad \text{where} \quad g_n(\theta) = \sum_{k=1}^{n-m_n} \Big\{s_k^2 - \frac{1}{n}\sigma^2(\frac{k}{n}, \hat{X}_{\frac{k}{n}}, \theta)\Big\}^2 \tag{5.4}$$

and

$$s_k^2 = \frac{n^{-\frac{1}{2}}}{\kappa\psi_2}\left(|\overline{Z}_k^n|^2 - n^{-\frac{1}{2}}\frac{\psi_1}{\kappa}\hat{\omega}_n^2\right). \tag{5.5}$$

From similar arguments as in the proof of Theorem 3 in the Appendix we see that

$$B_t(\theta_0) - \hat{B}_t(\hat{\theta}) = \int_0^t \Big\{\sigma^2(t, X_t, \theta_0) - \sigma^2(t, X_t, \hat{\theta})\Big\}\, ds + o_p(n^{-\frac{1}{4}}).$$

Assuming the common regularity conditions for nonlinear regression [see Gallant [12] or Seber and Wild [24]] θ_0 is the unique minimum of g and attained at an interior point of Θ. It is easy to see that $\hat{\theta} \to \theta_0$ in probability in this case, and thus we can assume that $\hat{\theta}$ satisfies $g_n'(\hat{\theta}) = 0$. This implies that

$$0 = g_n'(\hat{\theta}) = g_n'(\theta_0) + g_n''(\tilde{\theta})(\hat{\theta} - \theta_0) \quad \Leftrightarrow \quad \hat{\theta} - \theta_0 = -(g_n''(\tilde{\theta}))^{-1}\, g_n'(\theta_0)$$

for a suitable choice of $\tilde{\theta}$. Moreover,

$$\begin{aligned}
-g_n'(\theta_0) &= 2\sum_{k=1}^{n-m_n} \Big\{s_k^2 - \frac{1}{n}\sigma^2(\frac{k}{n}, \hat{X}_{\frac{k}{n}}, \theta)\Big\}\, \frac{\partial}{\partial\theta}\sigma^2(\frac{k}{n}, \hat{X}_{\frac{k}{n}}, \theta_0)\big|_{\theta=\theta_0} \\
&= 2\Big(\sum_{k=1}^{n-m_n} s_k^2\, \frac{\partial}{\partial\theta}\sigma^2(\frac{k}{n}, \hat{X}_{\frac{k}{n}}, \theta)\big|_{\theta=\theta_0} - \int_0^1 \sigma^2(s, X_s, \theta_0)\, \frac{\partial}{\partial\theta}\sigma^2(s, X_s, \theta)\big|_{\theta=\theta_0}\, ds\Big) + o_p(n^{-\frac{1}{4}}) \\
&= 2\Big(\sum_{k=1}^{n-m_n} s_k^2\, \frac{\partial}{\partial\theta}\sigma^2(\frac{k}{n}, \hat{X}_{\frac{k}{n}}, \theta)\big|_{\theta=\theta_0} - \int_0^1 \sigma_s^2\, \frac{\partial}{\partial\theta}\sigma^2(s, X_s, \theta)\big|_{\theta=\theta_0}\, ds\Big) + o_p(n^{-\frac{1}{4}}),
\end{aligned}$$

where the last equality follows from the definition of θ_0. Thus, the quantity $-g_n'(\theta_0)$ has a similar structure as the term $\hat{C}-C$ in the linear case, and in particular it is of order $O_p(n^{-\frac{1}{4}})$ as well. Furthermore, we have $\tilde{\theta} \to \theta_0$ in probability, and thus it can be assumed that $g_n''(\tilde{\theta})$ is positive definite and that the difference $||g_n''(\tilde{\theta})-g_n''(\theta_0)||$ is small. We conclude that $(\hat{\theta}-\theta_0) = O_p(n^{-\frac{1}{4}})$, and thus

$$\begin{aligned} B_t(\theta_0) - \hat{B}_t(\hat{\theta}) &= \int_0^t \Big(\frac{\partial}{\partial\theta}\sigma^2(s,X_s,\theta)\big|_{\theta=\theta_0}\Big)^T ds\cdot(\hat{\theta}-\theta_0) + o_p(n^{-\frac{1}{4}}) \\ &= -\int_0^t \Big(\frac{\partial}{\partial\theta}\sigma^2(s,X_s,\theta)\big|_{\theta=\theta_0}\Big)^T ds\ (g_n''(\theta_0))^{-1}\ g_n'(\theta_0) + o_p(n^{-\frac{1}{4}}). \end{aligned}$$

Furthermore, the $d\times d-$dimensional matrix $g_n''(\theta_0)$ takes the form

$$g_n''(\theta_0) = 2\Big(\frac{1}{n}S^TS - \sum_{k=1}^{n-m_n}\Big\{s_k^2 - \frac{1}{n}\sigma^2(\frac{k}{n},\hat{X}_{\frac{k}{n}},\theta_0)\Big\}H_k\Big),$$

where the $(n-m_n)\times d$ matrix S is given by

$$S = \Big(\frac{\partial}{\partial\theta}\sigma^2(\frac{k}{n},\hat{X}_{\frac{k}{n}},\theta)\big|_{\theta=\theta_0}\Big)_{k=1,\dots,n-m_n}$$

and H_k denotes the Hessian

$$H_k = \frac{\partial^2}{\partial\theta^2}\sigma^2(\frac{k}{n},X_{\frac{k}{n}},\theta)\big|_{\theta=\theta_0}.$$

Again, a similar calculation as given in the Appendix shows that

$$g_n''(\theta_0) = g''(\theta_0) + O_p(n^{-\frac{1}{4}}),$$

where the $d\times d$ matrix

$$\begin{aligned} g''(\theta_0) &= 2\int_0^1 \Big(\Big(\frac{\partial}{\partial\theta}\sigma^2(s,X_s,\theta)\big|_{\theta=\theta_0}\Big)^T\Big(\frac{\partial}{\partial\theta}\sigma^2(s,X_s,\theta)\big|_{\theta=\theta_0}\Big)ds \\ &- 2\int_0^1 \Big\{\sigma_s^2 - \sigma^2(s,X_s,\theta_0)\Big\}\frac{\partial^2}{\partial\theta^2}\sigma^2(s,X_s,\theta)\big|_{\theta=\theta_0}ds \end{aligned}$$

is positive definite. Note that the second term in this sum vanishes, when either the hypothesis is linear (since the Hessian is zero) or the null hypothesis is valid (since σ_s^2 equals $\sigma^2(s,X_s,\theta_0)$). In these cases the matrix $g''(\theta_0)$ takes precisely the same form as D in the linear setting. In any case, $g''(\theta_0)$ is of order $O_p(1)$, and thus we end up with the representation

$$B_t(\theta_0) - \hat{B}_t(\hat{\theta}) = -\int_0^t \Big(\frac{\partial}{\partial\theta}\sigma^2(s,X_s,\theta)\big|_{\theta=\theta_0}\Big)^T ds\ (g''(\theta_0))^{-1}\ g_n'(\theta_0) + o_p(n^{-\frac{1}{4}}),$$

and the asymptotics are driven by $g_n'(\theta_0)$. Consequently, it follows from the proof of Theorem 1 in the Appendix that in the case of testing a nonlinear hypothesis of the form (5.1), the process $\{\sqrt{n}(\hat{N}_t - N_t)\}_{t\in[0,1]}$ exhibits a similar asymptotic behavior as in the linear case, that is

$$\{n^{\frac{1}{4}}(\hat{N}_t - N_t)\}_{t\in[0,1]} \Longrightarrow \{A(t)\}_{t\in[0,1]},$$

where conditionally on $\mathcal{F}$ the limiting process is Gaussian, and its finite dimensional distributions coincide (conditionally on $\mathcal{F}$) with the finite dimensional distributions of the process

$$\Big\{\gamma_V\Big(I\{V \le t\} - \int_0^t \Big(\frac{\partial}{\partial\theta}\sigma^2(s,X_s,\theta)\big|_{\theta=\theta_0}\Big)^T ds\ (g''(\theta_0))^{-1}\ \Big(\frac{\partial}{\partial\theta}\sigma^2(V,X_V,\theta)\big|_{\theta=\theta_0}\Big)^T\Big)$$
$$-\ \ \Big(\int_0^t \gamma_s\ ds - \int_0^t \Big(\frac{\partial}{\partial\theta}\sigma^2(s,X_s,\theta)\big|_{\theta=\theta_0}\Big)^T ds\ (g''(\theta_0))^{-1}\ \int_0^1 \gamma_s \Big(\frac{\partial}{\partial\theta}\sigma^2(s,X_s,\theta)\big|_{\theta=\theta_0}\Big)^T ds\Big)\Big\}_{t\in[0,1]},$$

where the constant γ_u is defined in (4.9). We finally note again that, in the case of a fixed alternative and a nonlinear null hypothesis, this expression has a different structure than the corresponding term in Theorem 1.

6 Simulation study

We have indicated in the introduction that the original test for a constant volatility from the noise-free model loses its asymptotic properties in the presence of noise. Unsurprisingly, for a smaller variance of the noise variables, the data look more like observations from a continuous semimartingale and thus the test statistics behaves roughly in the same way as before, provided that the sample size is not too large. On the other hand, for a large variance of the error terms these are dominating, and thus the whole procedure breaks down even for small sample sizes. The same problem arises if the variance of the error is small but the sample size is large (see the discussion in the introduction). We start with a further example simulating the level of the bootstrap test proposed by Dette and Podolskij [10] for a parametric hypothesis, assessing its quality for various sample sizes n and different variances ω^2.

[INSERT TABLE 2 HERE]

Precisely, we have used the bootstrap test in Dette and Podolskij [10] for testing the hypothesis $H_0 : \sigma^2(t,x) = \theta x^2$, where $b(t,x) = 0.1x$. The results are obtained from 1000 simulation runs and 500 bootstrap replications and displayed in Table 2 for various sample sizes and standard deviations ω of the noise process. We observe that for $n = 256$ and a (small) standard deviation of $\omega = 0.001$ the test does roughly keep its asymptotic level, whereas it cannot be used at all when the variance becomes larger. Moreover, even if the variance is small but the sample size is increased, the test does not keep its pre-assigned level (see the results for $\omega = 0.001$ and $n = 1024$ in Table 2). Roughly speaking, we observe from these and similar simulation

results that there is no need for using tests, which address the problem of microstructure noise, if both the variance of the noise terms and the sample sizes (in our example $n \leq 256$) are small. On the other hand, it is known from empirical research that it is not realistic to assume extremely large values of ω, but the sample size for high frequency data is usually much larger than 256. Consequently, in many applications tests ignoring the presence of microstructure will neither keep their pre-assigned level nor be consistent, and the application of testing procedures addressing the problem of microstructure noise is strictly recommended.
In the following section we illustrate the finite sample properties of a bootstrap version of the Kolmogorov-Smirnov test based on the processes investigated in Section 4 and 5. Since the stochastic order of $|\Delta_i^n Z|$ is basically determined by the maximum of $n^{-\frac{1}{2}}$ and ω (which are the orders of $|\Delta_i^n X|$ and $|\Delta_i^n U|$, respectively), we kept $n\omega^2 = 0.1024$ fixed in order to have comparable results for different sample sizes n. The regularisation parameters κ and ρ were set to be 1/2 each. All simulation results presented in the following paragraphs are based on 1000 simulation runs and 500 bootstrap replications (if the bootstrap is applied to estimate critical values).
For all testing problems discussed below we have not used exactly the statistics $\hat{N}_t$ and $\hat{M}_t$, but related versions accounting for finite sample adjustments. Following Jacod *et al.* [17], where it has been shown that finite sample corrections improve the behaviour of the estimate $\hat{B}_t^0$ (and presumably of $\hat{C}$ as well) substantially, we have replaced the quantities ψ_i and Φ_{ij} in (3.4) by certain numbers ψ_i^n and Φ_{ij}^n, which constitute the "true" quantities for finite samples, but are replaced by their limits ψ_i and Φ_{ij} in the asymptotics. See Jacod *et al.* [17] for details.

6.1 Testing for homoscedasticity

In the problem of testing for homoscedasticity the limiting process $(A(t))_{t\in[0,1]}$ in (4.7) has an extremely simple form, when the null hypothesis of a constant process $(\sigma_t)_{t\in[0,1]}$ holds. In fact, the finite dimensional distributions of the process $(A(t))_{t\in[0,1]}$ coincide with the finite dimensional distributions of the process

$$\{\gamma(I\{V \leq t\} - t)\}_{t\in[0,1]}$$

for $V \sim \mathcal{U}[0,1]$, which means that $(A(t))_{t\in[0,1]}$ is a rescaled Brownian bridge. Thus we obtain the weak convergence

$$\Big(\frac{A_n(t)}{\hat{s}_t}\Big)_{t\in[0,1]} \xrightarrow{\mathcal{D}} (B_t)_{t\in[0,1]}, \tag{6.1}$$

where $(B_t)_{t\in[0,1]}$ is a standard Brownian bridge. Of course, this result can be used to construct a Kolmogorov-Smirnov or a Cramér-von-Mises test, and we have investigated the properties of the Kolmogorov-Smirnov test for different sample sizes n, where the noise satisfies $U \sim \mathcal{N}(0,\omega^2)$ and the drift function is again given by $b(t,x) = 0.1x$. A similar test can be constructed using

Theorem 2, but the corresponding results are omitted for the sake of brevity as the rate of convergence in this case becomes worse.

[INSERT TABLE 3 HERE]

In Table 3 we present the simulated level of the Kolmogorov-Smirnov test using the critical values from the asymptotic distribution. It can be seen that the asymptotic level of the test is slightly underestimated. This effect becomes less visible for a larger sample size, but even then it is still apparent. Note that these findings are in line with previous simulations on noisy observations and it is likely that they are due to the fact the rate of convergence for most testing problems is only $n^{-\frac{1}{4}}$, just as in our case.

6.2 Testing general hypotheses

For a general null hypothesis in (2.2), the distribution of the limiting process $(A(t))_{t\in[0,1]}$ depends on the path of the underlying semimartingale $(X_t)_{t\in[0,1]}$ and on the volatility $(\sigma_t)_{t\in[0,1]}$, and thus we cannot use it directly for the calculation of critical values. For this reason we propose the application of the parametric bootstrap in order to obtain simulated critical values. First we compute the global estimators $\hat{\omega}^2$ and $\hat{\theta} = \hat{D}^{-1}\hat{C}$ as well as each $n^{\frac{1}{4}}\hat{N}_t$ and $\hat{s}_t^2$ from the observed data. Under the null hypothesis N_t equals zero, and thus it is intuitively clear that the null hypothesis has to be rejected for large values of the standardised Kolmogorov-Smirnov statistic

$$Y_n = \sup_{t\in[0,1]} \left|\frac{n^{\frac{1}{4}}\hat{N}_t}{\hat{s}_t}\right|.$$

In a second step, we generate bootstrap data

$$(Z^{*(j)}_{\frac{i}{n}}, i = 1, \dots n, j = 1, \dots \beta),$$

where $Z^{*(j)}_{\frac{1}{n}} = X^{*(j)}_{\frac{1}{n}} + U^{*(j)}_{\frac{1}{n}}$, the $X^{*(j)}_{\frac{i}{n}}$ are realisations of the process in (2.1) with $b_s \equiv 0$ and $\sigma_s^2 = \sigma^2(s, X_s) = \sum_{k=1}^{d} \hat{\theta}_k \sigma_k^2(s, X_s)$ (corresponding to the null hypothesis) and each $U^{*(j)}_{\frac{i}{n}}$ is normally distributed with mean zero and variance $\hat{\omega}^2$. Using these data, we calculate the corresponding bootstrap statistics $Y_n^{*(j)}$ and use these to compute the quantiles of the bootstrap distribution. Finally, the null hypothesis is rejected if Y_n is larger than the $(1-\alpha)$-quantile of the bootstrap distribution.

[INSERT TABLE 4 HERE]

In order to investigate the approximation of the nominal level we consider the hypothesis of constant volatility and the hypothesis $H_0 : \sigma^2(t, x) = \theta x^2$. The data is generated under the null hypothesis with drift function $b(t, x) = 0.1x$ and the rejection probabilities are depicted in Table 4. These results show that the bootstrap approximation works well even for a small n. In particular, we see that in the case of homoscedasticity the exact asymptotic test using

the weak convergence of Y_n to the supremum of a standard Brownian bridge is outperformed (compare with Table 3). In the case of testing the parametric hypothesis $H_0 : \sigma^2(t,x) = x^2$ we observe a slight overestimation of the nominal level by the bootstrap test.
As an example for testing the hypothesis $\bar{H}_0$ defined in (2.3) we have chosen $\sigma(t,x) = \theta|x|$ and investigated the properties of the analogues of Y_n and $Y_n^{*(j)}$ from above, where we have replaced $n^{\frac{1}{4}}\hat{N}_t$ and $\hat{s}_t$ by $n^{\frac{1}{4}-\frac{\delta}{2}}\hat{M}_t$ and $\hat{r}_t$, respectively. In this case we chose $\delta = \frac{1}{4}$, corresponding to $l_n = O(n^{-\frac{3}{4}})$ and a rate of convergence $n^{-\frac{1}{8}}$. Note that in this particular situation there is no need for stating the hypothesis in terms of $\bar{H}_0$ as it is equivalent to $\sigma^2(t,x) = \theta|x|^2$, but nevertheless it gives a reasonable impression on how well the bootstrap approximation works for testing hypotheses of the form (2.3).

[INSERT TABLE 5 HERE]

We observe from the results in Table 5 that even though the rate of convergence in Theorem 2 is worse than in Theorem 1, there is no substantial difference in the approximation of the nominal level by the bootstrap test for both types of hypotheses: The nominal level is slightly overestimated, but in general the parametric bootstrap yields to a satisfactory and reliable approximation of the nominal level.
Finally, Table 6 contains the rejection probabilities of the bootstrap test under the alternative. The null hypothesis is given by $H_0 : \sigma^2(t,x) = \theta|x|^2$ and two alternatives, namely $\sigma^2(t,x) = 1$ and $\sigma^2(t,x) = 1 + |x|$, and one alternative coming from a stochastic volatility model is considered. For this case we chose the Heston model, i.e.

$$X_t = X_0 + \int_0^t (\mu - \nu_s/2)\, ds + \int_0^t \sigma_s\, dW_t \quad \text{with} \quad \nu_t = \nu_0 + \delta \int_0^t (\alpha - \nu_s)\, ds + \gamma \int_0^1 \nu_s^{1/2} dB_s,$$

where $\nu_t = \sigma_t^2$ and $\mathrm{Corr}(W,B) = \eta$ and the parameters were chosen as $\mu = 0.05/252, \delta = 5/252, \alpha = 0.04/252, \gamma = 0.05/252$ and $\rho = -0.5$.

[INSERT TABLE 6 HERE]

We observe from the results depicted in Table 6 that the bootstrap test indicates in all cases that the null hypothesis is not satisfied. It is also remarkable that it is more difficult to detect the alternatives $\sigma^2(t,x) = 1$ and $\sigma^2(t,x) = 1 + |x|$ than the one coming from the Heston model. In the latter case, the rejection probabilities are extremely large even for a small sample size, in contrary to the first two situations.

7 Appendix: Proof of Theorem 1 and 2

Before we come to the proof of the two theorems, we start with a typical localisation argument, which allows us to assume that several of the quantities and processes involved are bounded. Recall first that a and σ are locally bounded by assumption, from which is follows that X is locally bounded as well. Thus we can conclude along the lines of Jacod [16] that we may assume without loss of generality that each of these processes is actually bounded. Since further each σ_i^2 is continuous and because U has a compact support, we may conclude that both (s, X_t) and $(s, \hat{X}_{\frac{k}{n}})$ (for arbitrary s, t, k and n) are living on a compact set, and thus $\sigma_i^2(s, X_t)$ and $\sigma_i^2(s, \hat{X}_{\frac{k}{n}})$ are also bounded, the latter one uniformly in n. Similar results hold for the first two derivatives of σ_i^2 as well as for any of the functions $\bar{\sigma}_i$. Constants are denoted by K throughout this section.

7.1 Some preparations

The proofs of Theorem 1 and 2 are based on several preliminary results, which will be presented and proved in this subsection. We start with two results determining the rate of convergence of the quantities $\hat{B}_t^i - B_t^i$ and $\hat{D}_{ij} - D_{ij}$ defined in (2.7) and (2.6), respectively. The following result ensures that the (conditional) variance in a limit theorem for $\hat{N}_t - N_t$ will not depend on $\hat{B}_t^i$ and $\hat{D}_{ij}$, since the rate of convergence will be $n^{-\frac{1}{4}}$. Thus, we will focus in the following on the behavior of $\hat{C}_i$ and $\hat{B}_t^0$.

Theorem 3 *Under the assumptions from Section 3 we have*

$$\hat{B}_t^i - B_t^i = o_p(n^{-\frac{1}{4}}), \quad \textit{for } i = 1, \ldots, d, \tag{7.1}$$

$$\hat{D}_{ij} - D_{ij} = o_p(n^{-\frac{1}{4}}), \quad \textit{for } i, j = 1, \ldots, d, \tag{7.2}$$

where the first result holds uniformly with respect to $t \in [0, 1]$.

Proof of Theorem 3: For a proof of the first estimate (7.1) we use for a fixed index i the decomposition

$$\hat{B}_t^i - B_t^i = \frac{1}{n} \sum_{k=1}^{\lfloor nt \rfloor - m_n} \Big(\sigma_i^2(\frac{k}{n}, \hat{X}_{\frac{k}{n}}) - \sigma_i^2(\frac{k}{n}, X_{\frac{k}{n}}) \Big) + \Big(\frac{1}{n} \sum_{k=1}^{\lfloor nt \rfloor - m_n} \sigma_i^2(\frac{k}{n}, X_{\frac{k}{n}}) - \int_0^t \sigma_i^2(s, X_s) \, ds \Big).$$

Regarding the first term in this sum, note that

$$\begin{aligned} \hat{X}_{\frac{k}{n}} - X_{\frac{k}{n}} &= \frac{1}{m_n} \sum_{j=1}^{m_n} U_{\frac{k+j}{n}} + \frac{1}{m_n} \sum_{j=1}^{m_n} (X_{\frac{k+j}{n}} - X_{\frac{k}{n}}) \\ &= \frac{1}{m_n} \sum_{j=1}^{m_n} \Big(U_{\frac{k+j}{n}} + \int_{\frac{k}{n}}^{\frac{k+j}{n}} \sigma_s \, dW_s \Big) + O_p(n^{-\frac{1}{2}}), \end{aligned}$$

and thus $\hat{X}_{\frac{k}{n}} - X_{\frac{k}{n}} = O_p(n^{-\frac{1}{4}})$. Hence a Taylor expansion gives

$$\sigma_i^2(\frac{k}{n}, \hat{X}_{\frac{k}{n}}) - \sigma_i^2(\frac{k}{n}, X_{\frac{k}{n}}) = \frac{\partial}{\partial y} \sigma_i^2(\frac{k}{n}, X_{\frac{k}{n}}) \; (\hat{X}_{\frac{k}{n}} - X_{\frac{k}{n}}) + \frac{\partial^2}{\partial y^2} \sigma_i^2(\frac{k}{n}, \xi_{k,n}) \; (\hat{X}_{\frac{k}{n}} - X_{\frac{k}{n}})^2$$

for some random variables $\xi_{k,n}$ with $|\xi_{k,n} - X_{\frac{k}{n}}| \le |\hat{X}_{\frac{k}{n}} - X_{\frac{k}{n}}|$. As noted before, we have that $\frac{\partial^2}{\partial y^2}\sigma_i^2(\frac{k}{n}, \xi_{k,n})$ is bounded for all k and n which yields

$$\frac{1}{n}\sum_{k=1}^{\lfloor nt\rfloor - m_n}\Big(\sigma_i^2(\frac{k}{n}, \hat{X}_{\frac{k}{n}}) - \sigma_i^2(\frac{k}{n}, X_{\frac{k}{n}})\Big) = \frac{1}{n}\sum_{k=1}^{\lfloor nt\rfloor - m_n}\frac{\partial}{\partial y}\sigma_i^2(\frac{k}{n}, X_{\frac{k}{n}})\ (\hat{X}_{\frac{k}{n}} - X_{\frac{k}{n}}) + O_p(n^{-\frac{1}{2}}).$$

Hence, with

$$A_{k,n} = \frac{1}{m_n}\sum_{j=1}^{m_n}\frac{\partial}{\partial y}\sigma_i^2(\frac{k}{n}, X_{\frac{k}{n}})\ \Big(U_{\frac{k+j}{n}} + \int_{\frac{k}{n}}^{\frac{k+j}{n}}\sigma_s\ dW_s\Big)$$

it suffices to prove that

$$\frac{1}{n}\sum_{k=1}^{\lfloor nt\rfloor - m_n} A_{k,n} = o_p(n^{-\frac{1}{4}}). \tag{7.3}$$

However, we have $E[A_{k,n}A_{l,n}] = O(n^{-\frac{1}{2}})$ for arbitrary k and l as well as $E[A_{k,n}A_{k+l,n}] = 0$ for $l \ge m_n$ by conditioning on $\mathcal{F}_{\frac{k+l}{n}}$. This yields

$$E\Big[(\frac{1}{n}\sum_{k=1}^{\lfloor nt\rfloor - m_n} A_{k,n})^2\Big] = \frac{1}{n^2}\sum_{k=m_n}^{\lfloor nt\rfloor - 2m_n}\sum_{l=-m_n}^{m_n} E[A_{k,n}A_{k+l,n}] + O(\frac{m_n}{n^2}) = O(\frac{1}{n}),$$

and (7.3) follows. For the second term in the decomposition of $\hat{B}_t^i - B_t^i$ it holds that

$$\begin{aligned}
&\frac{1}{n}\sum_{k=1}^{\lfloor nt\rfloor - m_n}\sigma_i^2(\frac{k}{n}, X_{\frac{k}{n}}) - \int_0^t \sigma_i^2(s, X_s)\ ds \\
=\ & \sum_{k=1}^{\lfloor nt\rfloor}\int_{\frac{k-1}{n}}^{\frac{k}{n}}\Big(\sigma_i^2(\frac{k-1}{n}, X_{\frac{k-1}{n}}) - \sigma_i^2(s, X_s)\Big)\ ds + O_p(n^{-\frac{1}{2}}) \\
=\ & \sum_{k=1}^{\lfloor nt\rfloor}\int_{\frac{k-1}{n}}^{\frac{k}{n}}\Big(\sigma_i^2(\frac{k-1}{n}, X_{\frac{k-1}{n}}) - \sigma_i^2(s, X_{\frac{k-1}{n}}) + \sigma_i^2(s, X_{\frac{k-1}{n}}) - \sigma_i^2(s, X_s)\Big)\ ds + O_p(n^{-\frac{1}{2}}).
\end{aligned}$$

Since by assumption

$$\Big|\sigma_i^2(\frac{k-1}{n}, X_{\frac{k-1}{n}}) - \sigma_i^2(s, X_{\frac{k-1}{n}})\Big| < Kn^{-1}$$

for $\frac{k-1}{n} \le s \le \frac{k}{n}$ and by a similar expansion as above the claim follows. The result on $\hat{D}_{ij} - D_{ij}$ can be shown in the same way. □

The following result specifies the convergence of the finite dimensional distributions of the processes, which are used for the construction of $\{\hat{N}_t\}_{t\in[0,1]}$. Below we use the notation $G_n \xrightarrow{\mathcal{D}_{st}} G$ to indicate stable convergence of a sequence of random variables (G_n) to a limiting variable G, which is defined on an appropriate extension $(\Omega', \mathcal{F}', (\mathcal{F}'_t)_{t\in[0,1]}, P')$ of the original probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,1]}, P)$. For details on stable convergence see Jacod and Shiryaev [18].

Theorem 4 *Define for any fixed $t_1, \ldots, t_k \in [0,1]$ the $(k+d) \times (k+d)$ matrix*

$$\Sigma_{t_1,\ldots,t_k}(s, X_s) = \gamma_s^2 \ \ell(s, X_s)\ell^T(s, X_s)$$

where $\ell(s, X_s) = (1_{[0,t_1]}(s), \ldots, 1_{[0,t_k]}(s), g^T(s, X_s))^T$ and the vector $g(s, X_s)$ and γ_s^2 are defined by (4.8) and (4.9), respectively. Then we have

$$n^{\frac{1}{4}}\Big(\hat{B}^0_{t_1} - B^0_{t_1}, \ldots, \hat{B}^0_{t_k} - B^0_{t_k}, \hat{C}_1 - C_1, \ldots, \hat{C}_d - C_d\Big)^T \xrightarrow{\mathcal{D}_{st}} \int_0^1 \Sigma^{\frac{1}{2}}_{t_1,\ldots,t_k}(s, X_s)\ dW'_s,$$

where W' is another Brownian motion, which is independent of the σ-algebra $\mathcal{F}$.

Proof of Theorem 4: Observe first that $\hat{C}_i$ can be decomposed as follows:

$$\begin{aligned}
\hat{C}_i &= \frac{1}{\kappa\psi_2} n^{-\frac{1}{2}} \sum_{k=1}^{n-m_n} \sigma_i^2(\frac{k}{n}, X_{\frac{k}{n}})\ \Big(|\overline{Z}^n_k|^2 - n^{-\frac{1}{2}}\frac{\psi_1}{\kappa}\omega^2\Big) \\
&+ \frac{\psi_1}{\kappa^2\psi_2} n^{-1} \sum_{k=1}^{n-m_n} \sigma_i^2(\frac{k}{n}, X_{\frac{k}{n}})\ (\omega^2 - \hat{\omega}_n^2) \\
&+ \frac{1}{\kappa\psi_2} n^{-\frac{1}{2}} \sum_{k=1}^{n-m_n} \Big(\sigma_i^2(\frac{k}{n}, \hat{X}_{\frac{k}{n}}) - \sigma_i^2(\frac{k}{n}, X_{\frac{k}{n}})\Big)\ \Big(|\overline{Z}^n_k|^2 - n^{-\frac{1}{2}}\frac{\psi_1}{\kappa}\omega^2\Big) \\
&+ \frac{\psi_1}{\kappa^2\psi_2} n^{-1} \sum_{k=1}^{n-m_n} \Big(\sigma_i^2(\frac{k}{n}, \hat{X}_{\frac{k}{n}}) - \sigma_i^2(\frac{k}{n}, X_{\frac{k}{n}})\Big)\ (\omega^2 - \hat{\omega}_n^2).
\end{aligned}$$

Since $\omega^2 - \hat{\omega}_n^2 = O_p(n^{-\frac{1}{2}})$, the second and the fourth term in this sum are of the same order. Moreover, we find from similar arguments as given in the proof of Theorem 3 that the third term is of order $o_p(n^{-\frac{1}{4}})$ and thus asymptotically negligible as well. Therefore we are left to focus on

$$F_{in} = \frac{1}{\kappa\psi_2} n^{-\frac{1}{2}} \sum_{k=1}^{n-m_n} \sigma_i^2(\frac{k}{n}, X_{\frac{k}{n}})\Big(|\overline{Z}^n_k|^2 - n^{-\frac{1}{2}}\frac{\psi_1}{\kappa}\omega^2\Big).$$

Due to the dependence structure of the summands in F_{in} it will be convenient to use a "small-blocks-big-blocks"-technique as in Jacod *et al.* [17] in order to prove Theorem 4. To this end we choose an integer p, which later will go to infinity, and partition the n observations into several subsets: We define

$$b_k(p) = k(p+1)m_n \qquad \text{and} \qquad c_k(p) = k(p+1)m_n + pm_n$$

and define $j_n(p)$ to be the largest integer k such that $c_k(p) \le n - m_n$ holds, which gives the identity

$$j_n(p) = \Big\lfloor \frac{n}{(p+1)m_n} \Big\rfloor - 1. \tag{7.4}$$

Moreover, we use the notation $i_n(p) = (j_n(p)+1)pm_n$, and introduce for each $0 \le k \le j_n(p)$ and any p the following random variables:

$$G(k,p)_1^n = \frac{1}{\kappa\psi_2} n^{-\frac{1}{2}} \sigma_i^2\Big(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}}\Big) \sum_{j=b_k(p)}^{c_k(p)-1} \Big(|\overline{Z}_j^n|^2 - n^{-\frac{1}{2}} \frac{\psi_1}{\kappa} \omega^2 \Big),$$

$$G(k,p)_2^n = \frac{1}{\kappa\psi_2} n^{-\frac{1}{2}} \sigma_i^2\Big(\frac{c_k(p)}{n}, X_{\frac{c_k(p)}{n}}\Big) \sum_{j=c_k(p)}^{b_{k+1}(p)-1} \Big(|\overline{Z}_j^n|^2 - n^{-\frac{1}{2}} \frac{\psi_1}{\kappa} \omega^2 \Big).$$

The remainder terms are gathered in

$$G(p)_3^n = n^{-\frac{1}{2}} \frac{1}{\kappa\psi_2} \sum_{j=i_n(p)}^{n-m_n} \sigma_i^2\Big(\frac{i_n(p)}{n}, X_{\frac{i_n(p)}{n}}\Big) \Big(|\overline{Z}_j^n|^2 - n^{-\frac{1}{2}} \frac{\psi_1}{\kappa} \omega^2 \Big).$$

Note that each of these quantities depends on i, although it does not appear in the notation. The main intuition behind these quantities is that the terms $G(k,p)_1^n$ are defined on non-overlapping intervals, which means that the intervals on which each $\overline{Z}_j^n$ within $G(k,p)_1^n$ lives are disjoint from any $\overline{Z}_j^n$ within any other $G(l,p)_1^n$. This is sufficient to ensure some type of conditional independence, which will be used in order to prove Theorem 4. The variables $G(k,p)_2^n$ and $G(p)_3^n$ are filling the gaps between $G(k,p)_1^n$ and $G(l,p)_1^n$ and can be shown to be asymptotically negligible.
An important tool will be the following decomposition of $|\overline{Z}_j^n|^2$. We set

$$V_s^j = \int_{\frac{j}{n}}^{\frac{j}{n}+s} g_n\Big(u - \frac{j}{n}\Big)\, a_u \, du + \int_{\frac{j}{n}}^{\frac{j}{n}+s} g_n\Big(u - \frac{j}{n}\Big)\, \sigma_u \, dW_u,$$

and obtain from the representation of $\overline{X}_j^n$ as in (3.6) and by an application of Ito's formula

$$|\overline{X}_j^n|^2 = \int_{\frac{j}{n}}^{\frac{j+m_n}{n}} 2V_s^j \, g_n\Big(s-\frac{j}{n}\Big)\, a_s \; + \; g_n^2\Big(s-\frac{j}{n}\Big)\, \sigma_s^2 \, ds + 2\int_{\frac{j}{n}}^{\frac{j+m_n}{n}} V_s^j \, g_n\Big(s-\frac{j}{n}\Big)\, \sigma_s \, dW_s.$$

Thus,

$$\begin{aligned} |\overline{Z}_j^n|^2 &= |\overline{X}_j^n|^2 + |\overline{U}_j^n|^2 + 2\overline{X}_j^n \, \overline{U}_j^n \\ &= 2\int_{\frac{j}{n}}^{\frac{j+m_n}{n}} V_s^j \, g_n\Big(s-\frac{j}{n}\Big)\, a_s \, ds + 2\int_{\frac{j}{n}}^{\frac{j+m_n}{n}} V_s^j \, g_n\Big(s-\frac{j}{n}\Big)\, \sigma_s \, dW_s \\ &+ \int_{\frac{j}{n}}^{\frac{j+m_n}{n}} g_n^2\Big(s-\frac{j}{n}\Big)\, \sigma_s^2 \, ds + |\overline{U}_j^n|^2 + 2\overline{U}_j^n \int_{\frac{j}{n}}^{\frac{j+m_n}{n}} g_n\Big(s-\frac{j}{n}\Big)\, a_s \, ds \\ &+ 2\overline{U}_j^n \int_{\frac{j}{n}}^{\frac{j+m_n}{n}} g_n\Big(s-\frac{j}{n}\Big)\, \sigma_s \, dW_s =: \sum_{l=1}^{6} D(j)_l^n, \end{aligned} \qquad (7.5)$$

where the last identity defines the quantities $D(j)_l^n$ in an obvious manner.
For $b_k(p) \leq j < c_k(p)$ we introduce further

$$\tilde{D}(k,j,p)_2^n = 2\sigma^2_{\frac{b_k(p)}{n}} \int_{\frac{j}{n}}^{\frac{j+m_n}{n}} \Big(\int_{\frac{j}{n}}^{\frac{j}{n}+s} g_n(u-\frac{j}{n})\ dW_u \Big)\ g_n(s-\frac{j}{n})\ dW_s,$$

$$\tilde{D}(k,j,p)_6^n = 2\sigma_{\frac{b_k(p)}{n}}\ \overline{U}_j^n \int_{\frac{j}{n}}^{\frac{j+m_n}{n}} g_n(s-\frac{j}{n})\ dW_s$$

as approximations for the quantities $D(j)_2^n$ and $D(j)_6^n$. Additionally, we set

$$H(k,p)^n = \sigma_i^2(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}})\ Y(k,p)^n,$$

where

$$Y(k,p)^n = \frac{1}{\kappa\psi_2} n^{-\frac{1}{2}} \sum_{j=b_k(p)}^{c_k(p)-1} \Big\{ \tilde{D}(k,j,p)_2^n + \tilde{D}(k,j,p)_6^n + \Big(D(j)_4^n - n^{-\frac{1}{2}} \frac{\psi_1}{\kappa}\omega^2 \Big) \Big\}. \tag{7.6}$$

Finally, we define

$$\chi(p)_k^n = \Big(E\Big[\Big(\sup_{s,t\in[\frac{b_k(p)}{n},\frac{c_k(p)}{n}]} |a_s - a_t| + |\sigma_s - \sigma_t| \Big)^2 \Big| \mathcal{F}_{\frac{b_k(p)}{n}} \Big] \Big)^{\frac{1}{2}}.$$

We start with two auxiliary results which specify the asymptotic properties of F_{in} and prove the first assertion in detail.

Lemma 1 *We have*

$$\lim_{p\to\infty} \limsup_{n\to\infty} n^{\frac{1}{4}} \Big\{ \Big(\sum_{k=0}^{j_n(p)} (G(k,p)_1^n + G(k,p)_2^n) + G(p)_3^n - C_i \Big) - \sum_{k=0}^{j_n(p)} H(k,p) \Big\} = 0.$$

Proof of Lemma 1: The proof goes through a rather large number of steps and makes extensive use of the decomposition in (7.5). We will show first that the influence of the random variables $D(j)_1^n$ and $D(j)_5^n$ within $G(k,p)_1^n$ is asymptotically negligible, that is

$$\lim_{p\to\infty} \limsup_{n\to\infty} n^{-\frac{1}{4}} \sum_{k=0}^{j_n(p)} \sigma_i^2(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}}) \sum_{j=b_k(p)}^{c_k(p)-1} (D(j)_1^n + D(j)_5^n) = 0. \tag{7.7}$$

Completely analogous results hold for the corresponding results on $G(k,p)_2^n$ and $G(p)_3^n$ as well. For a proof of (7.7), assume without loss of generality that $b_k(p) \leq j < c_k(p)$, and thus we have the decomposition $D(j)_1^n = D'(j)_1^n + D''(j)_1^n$ with

$$D'(j)_1^n = 2\int_{\frac{j}{n}}^{\frac{j+m_n}{n}} \Big(\int_{\frac{j}{n}}^{\frac{j}{n}+s} g_n(u-\frac{j}{n})\ a_u\ du \Big)\ g_n(s-\frac{j}{n})\ a_s\ ds,$$

$$D''(j)_1^n = 2\int_{\frac{j}{n}}^{\frac{j+m_n}{n}} \Big(\int_{\frac{j}{n}}^{\frac{j}{n}+s} g_n(u-\frac{j}{n})\ \sigma_u\ dW_u \Big)\ g_n(s-\frac{j}{n})\ a_s\ ds.$$

Obviously, we have $E\Big[|D'(j)^n_1| \;\Big|\mathcal{F}_{\frac{b_k(p)}{n}}\Big] \le Kn^{-1}$, which allows us to focus on the second term only. Using the decomposition

$$\begin{aligned} D''(j)^n_1 &= 2a_{\frac{b_k(p)}{n}} \int_{\frac{j}{n}}^{\frac{j+m_n}{n}} \Big(\int_{\frac{j}{n}}^{\frac{j}{n}+s} g_n(u-\frac{j}{n})\ \sigma_u\ dW_u\Big)\ g_n(s-\frac{j}{n})\ ds \\ &+\ 2\int_{\frac{j}{n}}^{\frac{j+m_n}{n}} \Big(\int_{\frac{j}{n}}^{\frac{j}{n}+s} g_n(u-\frac{j}{n})\ \sigma_u\ dW_u\Big)\ g_n(s-\frac{j}{n})(a_s - a_{\frac{b_k(p)}{n}})\ ds, \end{aligned}$$

the martingale property of a stochastic integral with respect to Brownian motion and the Cauchy-Schwarz inequality we derive that

$$\Big|E\Big[D''(j)^n_1 \;\Big|\mathcal{F}_{\frac{b_k(p)}{n}}\Big]\Big| \le K\ n^{-\frac{3}{4}}\ \chi(p)^n_k.$$

Thus with the notation $\delta(k,p)^n_1 = \sum_{j=b_k(p)}^{c_k(p)-1} D''(j)^n_1$, we conclude that

$$\Big|E\Big[\delta(k,p)^n_1 \;\Big|\mathcal{F}_{\frac{b_k(p)}{n}}\Big]\Big| \le K\ p\ n^{-\frac{1}{4}}\ \chi(p)^n_k.$$

For the same reasons we have

$$E\Big[\Big(\delta(k,p)^n_1\Big)^2 \;\Big|\mathcal{F}_{\frac{b_k(p)}{n}}\Big] = \sum_{j,l=b_k(p)}^{c_k(p)-1} E\Big[D''(j)^n_1\ D''(l)^n_1 \;\Big|\mathcal{F}_{\frac{b_k(p)}{n}}\Big] \le K\ p^2\ n^{-\frac{1}{2}},$$

and with $k > l$ it follows

$$\Big|E\Big\{\sigma_i^2(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}})\ \sigma_i^2(\frac{b_l(p)}{n}, X_{\frac{b_l(p)}{n}})\ \delta(l,p)^n_1\ E\Big[\delta(k,p)^n_1 \;\Big|\mathcal{F}_{\frac{b_k(p)}{n}}\Big]\Big\}\Big| \le K\ p^2\ n^{-\frac{1}{2}}\ \chi(p)^n_k.$$

Finally, we obtain

$$\begin{aligned} & E\Big[\Big(n^{-\frac{1}{4}} \sum_{k=0}^{j_n(p)} \sigma_i^2(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}}) \sum_{j=b_k(p)}^{c_k(p)-1} D''(j)^n_1\Big)^2\Big] \\ =&\ \ n^{-\frac{1}{2}} \sum_{k=0}^{j_n(p)} E\Big[\sigma_i^4(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}})\ (\delta(k,p)^n_1)^2\Big] \\ +&\ \ 2n^{-\frac{1}{2}} \sum_{k>l}^{j_n(p)} E\Big[\sigma_i^2(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}})\ \sigma_i^2(\frac{b_l(p)}{n}, X_{\frac{b_l(p)}{n}})\ \delta(k,p)^n_1\ \delta(l,p)^n_1\Big] \\ \le&\ \ K\Big(p\ n^{-\frac{1}{2}} + \sum_{k>l}^{j_n(p)} p^2\ n^{-1}\ E[\chi(p)^n_k]\Big). \end{aligned}$$

From Lemma 5.4. in Jacod *et al.* [17] it follows that $\lim_{n\to\infty} n^{-\frac{1}{2}} \sum_{k=1}^{j_n(p)} E[\chi(p)^n_k] = 0$ for any p, which gives that the first term in the sum (7.7) converges to 0. For a proof of a corresponding statement for the second term, we define

$$\delta(k,p)_5^n = \sum_{j=b_k(p)}^{c_k(p)-1} D(j)_5^n$$

and obtain from the independence of X and U that

$$E\Big[\delta(k,p)_5^n \,\Big|\mathcal{F}_{\frac{b_k(p)}{n}}\Big] = 0 \qquad \text{and} \qquad E\Big[(\delta(k,p)_5^n)^2 \,\Big|\mathcal{F}_{\frac{b_k(p)}{n}}\Big] \leq K\ p^2\ n^{-\frac{1}{2}}.$$

Hence, a standard martingale argument gives

$$E\Big[\Big(n^{-\frac{1}{4}} \sum_{k=0}^{j_n(p)} \sigma_i^2(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}})\ \delta(k,p)_5^n\Big)^2\Big] \ \leq\ K\ p\ n^{-\frac{1}{2}},$$

which finishes the proof of (7.7).
The next step is devoted to the analysis of the term $D(j)_2^n$. We prove

$$\lim_{p\to\infty} \limsup_{n\to\infty} n^{-\frac{1}{4}} \sum_{k=0}^{j_n(p)} \sigma_i^2(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}}) \sum_{j=b_k(p)}^{c_k(p)-1} \Big(D(j)_2^n - \tilde{D}(k,j,p)_2^n\Big) = 0 \tag{7.8}$$

as well as

$$\lim_{p\to\infty} \limsup_{n\to\infty} n^{-\frac{1}{4}} \sum_{k=0}^{j_n(p)} \sigma_i^2(\frac{c_k(p)}{n}, X_{\frac{c_k(p)}{n}}) \sum_{j=c_k(p)}^{b_{k+1}(p)-1} D(j)_2^n = 0, \tag{7.9}$$

$$\lim_{p\to\infty} \limsup_{n\to\infty} n^{-\frac{1}{4}}\ \sigma_i^2(\frac{i_n(p)}{n}, X_{\frac{i_n(p)}{n}}) \sum_{j=i_n(p)}^{n-m_n} D(j)_2^n = 0. \tag{7.10}$$

Set $b_k(p) \leq j < c_k(p)$ again and observe the decomposition $D(j)_2^n = D'(j)_2^n + D''(j)_2^n$, where

$$\begin{aligned} D'(j)_2^n &= 2\int_{\frac{j}{n}}^{\frac{j+m_n}{n}} \Big(\int_{\frac{j}{n}}^{\frac{j}{n}+s} g_n(u-\frac{j}{n})\ a_u\ du\Big)\ g_n(s-\frac{j}{n})\ \sigma_s\ dW_s, \\ D''(j)_2^n &= 2\int_{\frac{j}{n}}^{\frac{j+m_n}{n}} \Big(\int_{\frac{j}{n}}^{\frac{j}{n}+s} g_n(u-\frac{j}{n})\ \sigma_u\ dW_u\Big)\ g_n(s-\frac{j}{n})\ \sigma_s\ dW_s. \end{aligned}$$

From

$$E\Big[D'(j)_2^n \,\Big|\mathcal{F}_{\frac{b_k(p)}{n}}\Big] = 0 \qquad \text{and} \qquad E\Big[|D'(j)_2^n\ D'(l)_2^n| \,\Big|\mathcal{F}_{\frac{b_k(p)}{n}}\Big] \leq K\ n^{-\frac{3}{2}}$$

we conclude

$$\lim_{p\to\infty} \limsup_{n\to\infty} E\Big[\Big(n^{-\frac{1}{4}} \sum_{k=0}^{j_n(p)} \sigma_i^2(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}}) \sum_{j=b_k(p)}^{c_k(p)-1} D'(j)_2^n\Big)^2\Big] = 0$$

from a similar martingale argument as in the previous paragraph and may thus focus on $D''(j)_2^n$. We have

$$E\Big[D''(j)_2^n \,\Big|\mathcal{F}_{\frac{b_k(p)}{n}}\Big] = 0 \qquad \text{and} \qquad E\Big[|D''(j)_2^n D''(l)_2^n| \,\Big|\mathcal{F}_{\frac{b_k(p)}{n}}\Big] \leq K\ n^{-1},$$

thus (7.10) follows easily. For (7.9), note that $E[(\sum_{j=c_k(p)}^{b_{k+1}(p)-1} D''(j)_2^n)^2] \leq K$, which gives (recall the definition of $j_n(p), b_k(p)$ and $c_k(p)$)

$$n^{-\frac{1}{2}} \sum_{k=0}^{j_n(p)} E\Big[\sigma_i^4(\frac{c_k(p)}{n}, X_{\frac{c_k(p)}{n}})\Big(\sum_{j=c_k(p)}^{b_{k+1}(p)-1} D''(j)_2^n\Big)^2\Big] \leq K\ n^{-\frac{1}{2}}\ \frac{n^{\frac{1}{2}}}{p} = K\ \frac{1}{p},$$

which converges to zero as p tends to infinity. We are thus left to prove

$$\lim_{p\to\infty} \limsup_{n\to\infty} n^{-\frac{1}{4}} \sum_{k=0}^{j_n(p)} \sigma_i^2(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}}) \sum_{j=b_k(p)}^{c_k(p)-1} \Big(D''(j)_2^n - \tilde{D}(k,j,p)_2^n\Big) = 0.$$

This time, we have $E[D''(j)_2^n - \tilde{D}(k,j,p)_2^n\ |\mathcal{F}_{\frac{b_k(p)}{n}}] = 0$ and

$$E\Big[\Big|\Big(D''(j)_2^n - \tilde{D}(k,j,p)_2^n\Big)\ \Big(D''(l)_2^n - \tilde{D}(k,l,p)_2^n\Big)\Big|\ \Big|\mathcal{F}_{\frac{b_k(p)}{n}}\Big] \leq K\ n^{-1}\ (\chi(p)_k^n)^2.$$

Thus

$$\begin{aligned}
& E\Big[\Big\{n^{-\frac{1}{4}} \sum_{k=0}^{j_n(p)} \sigma_i^2(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}}) \sum_{j=b_k(p)}^{c_k(p)-1} \Big(D(j)_2^n - \tilde{D}(k,j,p)_2^n\Big)\Big\}^2\Big] \\
= \;& n^{-\frac{1}{2}} \sum_{k=0}^{j_n(p)} E\Big[\sigma_i^4(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}})\Big(\sum_{j=b_k(p)}^{c_k(p)-1} \Big(D(j)_2^n - \tilde{D}(k,j,p)_2^n\Big)\Big)^2\Big] \\
\leq \;& K\ n^{-\frac{1}{2}} \sum_{k=0}^{j_n(p)} \sum_{j,l=b_k(p)}^{c_k(p)-1} E\Big[\Big(D''(j)_2^n - \tilde{D}(k,j,p)_2^n\Big)\ \Big(D''(l)_2^n - \tilde{D}(k,l,p)_2^n\Big)\Big] \\
\leq \;& K\ n^{-\frac{3}{2}} \sum_{k=0}^{j_n(p)} \sum_{j,l=b_k(p)}^{c_k(p)-1} (\chi(p)_k^n)^2 \leq K\ p^2\ n^{-\frac{1}{2}} \sum_{k=0}^{j_n(p)} E\Big[(\chi(p)_k^n)^2\Big].
\end{aligned}$$

With a similar argument as in the proof of (7.7) we are done. Proving that $D(j)_6^n$ can be replaced by $\tilde{D}(k,j,p)_6^n$ works analogously, thus we finish the proof of Lemma 1 showing

$$\begin{aligned}
\lim_{p\to\infty} \limsup_{n\to\infty} n^{\frac{1}{4}} \Big\{ \frac{1}{\kappa\psi_2} n^{-\frac{1}{2}} \Big(\sum_{k=0}^{j_n(p)} \Big(\sigma_i^2(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}}) \sum_{j=b_k(p)}^{c_k(p)-1} D(j)_3^n \qquad\qquad (7.11) \\
+ \;\; \sigma_i^2(\frac{c_k(p)}{n}, X_{\frac{c_k(p)}{n}}) \sum_{j=c_k(p)}^{b_{k+1}(p)-1} D(j)_3^n\Big) + \sum_{j=i_n(p)}^{n-m_n} D(j)_3^n\Big) - C_i \Big\} = 0.
\end{aligned}$$

We start with the following proposition:

$$\lim_{p\to\infty}\limsup_{n\to\infty} n^{\frac{1}{4}}\Big\{\Big(\sum_{k=0}^{j_n(p)}\Big(\int_{\frac{b_k(p)}{n}}^{\frac{c_k(p)}{n}} \sigma_i^2(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}})\ \sigma_s^2\ ds \tag{7.12}$$
$$+ \int_{\frac{c_k(p)}{n}}^{\frac{b_{k+1}(p)}{n}} \sigma_i^2(\frac{c_k(p)}{n}, X_{\frac{c_k(p)}{n}})\ \sigma_s^2\ ds\Big) + \int_{\frac{i_n(p)}{n}}^{1} \sigma_i^2(\frac{i_n(p)}{n}, X_{\frac{i_n(p)}{n}})\ \sigma_s^2\ ds\Big) - C_i\Big\} = 0.$$

As in the proof of Theorem 3 we obtain

$$\begin{aligned}
&\int_{\frac{b_k(p)}{n}}^{\frac{c_k(p)}{n}} \Big(\sigma_i^2(s,X_s) - \sigma_i^2(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}})\Big)\ \sigma_s^2\ ds \\
= \ &\int_{\frac{b_k(p)}{n}}^{\frac{c_k(p)}{n}} \Big(\sigma_i^2(s,X_s) - \sigma_i^2(s, X_{\frac{b_k(p)}{n}}) + \sigma_i^2(s, X_{\frac{b_k(p)}{n}}) - \sigma_i^2(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}})\Big)\ \sigma_s^2\ ds \\
= \ &\int_{\frac{b_k(p)}{n}}^{\frac{c_k(p)}{n}} \frac{\partial}{\partial y}\sigma_i^2(s, X_{\frac{b_k(p)}{n}})\Big(\int_{\frac{b_k(p)}{n}}^{s} \sigma_u\ dW_u\Big)\ \sigma_s^2\ ds + O_p\Big(\frac{p^2 m_n^2}{n^2}\Big) \\
=: \ &\delta'(k,p)_3^n + \delta''(k,p)_3^n + O_p\Big(\frac{p^2 m_n^2}{n^2}\Big),
\end{aligned} \tag{7.13}$$

where

$$\delta'(k,p)_3^n = \sigma^3_{\frac{b_k(p)}{n}} \int_{\frac{b_k(p)}{n}}^{\frac{c_k(p)}{n}} \frac{\partial}{\partial y}\sigma_i^2(s, X_{\frac{b_k(p)}{n}})\Big(\int_{\frac{b_k(p)}{n}}^{s} dW_u\Big)\ ds$$

and $\delta''(k,p)_3^n$ is defined implicitly by equation (7.13). From

$$E\Big[\delta'(k,p)_3^n\ \Big|\mathcal{F}_{\frac{b_k(p)}{n}}\Big] = 0 \qquad \text{and} \qquad E\Big[(\delta'(k,p)_3^n)^2\ \Big|\mathcal{F}_{\frac{b_k(p)}{n}}\Big] \le K\ p^3\ n^{-\frac{3}{2}}$$

we conclude

$$\lim_{p\to\infty}\limsup_{n\to\infty} n^{\frac{1}{2}} E\Big[\Big(\sum_{k=0}^{j_n(p)} \delta'(k,p)_3^n\Big)^2\Big] = 0.$$

For $\delta''(k,p)_3^n$ we have $E[|\delta''(k,p)_3^n|\ |\mathcal{F}_{\frac{b_k(p)}{n}}] \le K\ p^{\frac{3}{2}}\ n^{-\frac{3}{4}}\ \chi(p)_k^n$ as usual, thus

$$\lim_{p\to\infty}\limsup_{n\to\infty} n^{\frac{1}{4}} \sum_{k=0}^{j_n(p)} E\Big[|\delta''(k,p)_3^n|\Big] \le \lim_{p\to\infty}\limsup_{n\to\infty} K\ p^{\frac{3}{2}}\ n^{-\frac{1}{2}} \sum_{k=0}^{j_n(p)} E\Big[\chi(p)_k^n\Big] = 0.$$

The corresponding results for the other summands in (7.12) can be shown analogously.

To finish the proof of Lemma 1 we have to show

$$\begin{aligned}\lim_{p\to\infty}\limsup_{n\to\infty} n^{\frac14}\Big\{\sum_{k=0}^{j_n(p)}\Big(\sigma_i^2(\frac{b_k(p)}{n},X_{\frac{b_k(p)}{n}})\Big(\frac{1}{\kappa\psi_2}n^{-\frac12}\sum_{j=b_k(p)}^{c_k(p)-1}D(j)_3^n-\int_{\frac{b_k(p)}{n}}^{\frac{c_k(p)}{n}}\sigma_s^2\ ds\Big)\\ +\ \sigma_i^2(\frac{c_k(p)}{n},X_{\frac{c_k(p)}{n}})\Big(\frac{1}{\kappa\psi_2}n^{-\frac12}\sum_{j=c_k(p)}^{b_{k+1}(p)-1}D(j)_3^n-\int_{\frac{c_k(p)}{n}}^{\frac{b_{k+1}(p)}{n}}\sigma_s^2\ ds\Big)\Big)\\ +\ \sigma_i^2(\frac{i_n(p)}{n},X_{\frac{i_n(p)}{n}})\Big(\frac{1}{\kappa\psi_2}n^{-\frac12}\sum_{j=i_n(p)}^{n-m_n}D(j)_3^n-\int_{\frac{i_n(p)}{n}}^{1}\sigma_s^2\ ds\Big)\Big\}=0.\end{aligned}$$

The last term in the sum is negligible. For the other terms we fix k for a moment and observe that

$$\frac{1}{\kappa\psi_2}n^{-\frac12}\sum_{j=b_k(p)}^{c_k(p)-1}D(j)_3^n = \int_{\frac{b_k(p)}{n}}^{\frac{b_{k+1}(p)}{n}} h_{n,p}\Big(s-\frac{b_k(p)}{n}\Big)\ \sigma_s^2\ ds, \tag{7.14}$$

$$\frac{1}{\kappa\psi_2}n^{-\frac12}\sum_{j=c_k(p)}^{b_{k+1}(p)-1}D(j)_3^n = \int_{\frac{c_k(p)}{n}}^{\frac{b_{k+1}(p)+m_n}{n}} \bar h_{n,p}\Big(s-\frac{c_k(p)}{n}\Big)\ \sigma_s^2\ ds \tag{7.15}$$

with

$$\begin{aligned}h_{n,p}(s) &= h^1_{n,p}(s)\ 1_{[0,\frac{m_n}{n})}(s)+h^2_{n,p}\ 1_{[\frac{m_n}{n},\frac{pm_n}{n})}(s)+h^3_{n,p}(s)\ 1_{[\frac{pm_n}{n},\frac{(p+1)m_n}{n})}(s),\\ \bar h_{n,p}(s) &= h^1_{n,p}(s)\ 1_{[0,\frac{m_n}{n})}(s)+h^3_{n,p}(s)\ 1_{[\frac{m_n}{n},\frac{2m_n}{n})}(s) \qquad \text{and}\\ h^1_{n,p}(s) &= \frac{1}{\kappa\psi_2}n^{-\frac12}\sum_{j=0}^{m_n-1}\sum_{i=1}^{j+1}(g_i^n)^2 1_{[\frac{j}{n},\frac{j+1}{n})}(s),\\ h^2_{n,p} &= \frac{1}{\kappa\psi_2}n^{-\frac12}\sum_{i=1}^{m_n}(g_i^n)^2 = 1+O(n^{-\frac12}),\\ h^3_{n,p}(s) &= \frac{1}{\kappa\psi_2}n^{-\frac12}\sum_{j=0}^{m_n-1}\sum_{i=j+2}^{m_n}(g_i^n)^2 1_{[\frac{j}{n},\frac{j+1}{n})}(s).\end{aligned}$$

Thus,

$$\Big|n^{\frac14}\sum_{k=0}^{j_n(p)}\sigma_i^2(\frac{b_k(p)}{n},X_{\frac{b_k(p)}{n}})\int_{\frac{b_k(p)+m_n}{n}}^{\frac{c_k(p)}{n}}\Big(h_{n,p}\Big(s-\frac{b_k(p)}{n}\Big)-1\Big)\ \sigma_s^2\ ds\Big|\le K\ n^{-\frac14}.$$

Other integrals than those between $\frac{b_k(p)+m_n}{n}$ and $\frac{c_k(p)}{n}$ occur in the following way:

$$\begin{aligned}\int_{\frac{b_k(p)}{n}}^{\frac{b_k(p)+m_n}{n}}&\Big\{\sigma_i^2(\frac{b_k(p)}{n},X_{\frac{b_k(p)}{n}})\ \Big(h^1_{n,p}\Big(s-\frac{b_k(p)}{n}\Big)-1\Big)\\ &+\ \sigma_i^2(\frac{c_{k-1}(p)}{n},X_{\frac{c_{k-1}(p)}{n}})\ h^3_{n,p}\Big(s-\frac{c_{k-1}(p)}{n}\Big)\Big\}\ \sigma_s^2\ ds,\end{aligned}$$

where the first term and the second term in the integrand come from (7.14) and (7.15), respectively. A similar result holds for the integral from $\frac{b_k(p)+m_n}{n}$ to $\frac{c_k(p)}{n}$. By definition, we have

$$h^1_{n,p}\Big(s - \frac{b_k(p)}{n}\Big) + h^3_{n,p}\Big(s - \frac{c_{k-1}(p)}{n}\Big) = h^2_{n,p}$$

for $\frac{b_k(p)}{n} \le s \le \frac{b_k(p)+m_n}{n}$, and hence it is enough to prove that

$$n^{\frac{1}{4}} \sum_{k=0}^{j_n(p)} \int_{\frac{b_k(p)}{n}}^{\frac{b_k(p)+m_n}{n}} h^3_{n,p}\Big(s - \frac{b_k(p)}{n}\Big) \ \Big(\sigma_i^2(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}}) - \sigma_i^2(\frac{c_k(p)}{n}, X_{\frac{c_k(p)}{n}})\Big) \ \sigma_s^2 \ ds$$

converges to zero in the usual way. Again, this follows from a Taylor expansion and a similar argument as in the first part of the proof of (7.11). □

Lemma 2 *We have*

$$\lim_{p\to\infty} \limsup_{n\to\infty} n^{\frac{1}{4}} \Big\{ F_{in} - \Big(\sum_{k=0}^{j_n(p)} (G(k,p)^n_1 + G(k,p)^n_2) + G(p)^n_3 \Big) \Big\} = 0.$$

Proof of Lemma 2: Without loss of generality is suffices to show

$$\lim_{p\to\infty} \limsup_{n\to\infty} n^{-\frac{1}{4}} \sum_{k=0}^{j_n(p)} \sum_{j=b_k(p)}^{c_k(p)-1} \Big(\sigma_i^2(s, X_s) - \sigma_i^2(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}})\Big) \ \Big(|\overline{Z}^n_j|^2 - n^{-\frac{1}{2}} \frac{\psi_1}{\kappa}\omega^2\Big) = 0.$$

From another Taylor expansion we have

$$\sigma_i^2(s, X_s) - \sigma_i^2(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}}) = \frac{\partial}{\partial y}\sigma_i^2(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}}) \int_{\frac{b_k(p)}{n}}^{s} \sigma_u \ dW_u + O_p\Big(\frac{pm_n}{n}\Big),$$

thus we are left to prove

$$\lim_{p\to\infty} \limsup_{n\to\infty} n^{-\frac{1}{4}} \sum_{k=0}^{j_n(p)} \sum_{j=b_k(p)}^{c_k(p)-1} \frac{\partial}{\partial y}\sigma_i^2(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}})\Big(\int_{\frac{b_k(p)}{n}}^{s} \sigma_u \ dW_u\Big) \ \Big(|\overline{Z}^n_j|^2 - n^{-\frac{1}{2}} \frac{\psi_1}{\kappa}\omega^2\Big) = 0.$$

However, this result follows from similar arguments as in the proof of Lemma 1. □

Note that we have completely analogous results for a decomposition of $\hat{B}'^0_t - B^0_t$. Thus, we end up with

$$\begin{aligned} \lim_{p\to\infty} \limsup_{n\to\infty} n^{\frac{1}{4}} \Big\{ (\hat{B}'^0_t - B^0_t) - \sum_{k=0}^{j_n(p)} Y(k,p) 1_{\{\frac{c_k(p)}{n} \le t\}} \Big\} &= 0, \qquad (7.16) \\ \lim_{p\to\infty} \limsup_{n\to\infty} n^{\frac{1}{4}} \Big\{ (\hat{C}'_i - C_i) - \sum_{k=0}^{j_n(p)} \sigma_i^2(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}}) \ Y(k,p) \Big\} &= 0, \end{aligned}$$

where $Y(k,p)$ was defined in (7.6). Since $n\ E[(Y(k,p))^2|\mathcal{F}_{\frac{b_k(p)}{n}}] = p\ \kappa\ \gamma^2_{\frac{b_k(p)}{n}} + o_p(1)$ and $E[Y(k,p)|\mathcal{F}_{\frac{b_k(p)}{n}}] = 0$ as in Jacod *et al.* [17], we conclude

$$\lim_{p\to\infty}\lim_{n\to\infty} n^{\frac{1}{2}} \sum_{k=0}^{j_n(p)} E\Big[Y(k,p)^2 1_{\{\frac{c_k(p)}{n}\le t_i\wedge t_j\}}\Big|\mathcal{F}_{\frac{b_k(p)}{n}}\Big] = \int_0^1 \gamma_s^2\ 1_{[0,t_i\wedge t_j]}(s)\ ds$$

$$\lim_{p\to\infty}\lim_{n\to\infty} n^{\frac{1}{2}} \sum_{k=0}^{j_n(p)} E\Big[Y(k,p)^2 1_{\{\frac{c_k(p)}{n}\le t_i\}}\sigma_i^2(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}})\Big|\mathcal{F}_{\frac{b_k(p)}{n}}\Big] = \int_0^1 \gamma_s^2\ 1_{[0,t_i]}(s)\ \sigma_j^2(s,X_s)\ ds$$

$$\lim_{p\to\infty}\lim_{n\to\infty} n^{\frac{1}{2}} \sum_{k=0}^{j_n(p)} E\Big[Y(k,p)^2 \sigma_i^2(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}})\sigma_j^2(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}})\Big|\mathcal{F}_{\frac{b_k(p)}{n}}\Big]$$
$$= \int_0^1 \gamma_s^2\ \sigma_i^2(s,X_s)\ \sigma_j^2(s,X_s)\ ds$$

Theorem 4 follows now from Theorem IX 7.28 in Jacod and Shiryaev [18], since the missing conditions can be shown in the same way as in Jacod *et al.* [17]. □

7.2 Proof of Theorem 1

The convergence of the finite dimensional distributions follows from the delta method for stably converging sequences, since we have

$$n^{\frac{1}{4}}(\hat{N}_{t_1} - N_{t_1}, \dots, N_{t_k} - N_{t_k})^T \xrightarrow{\mathcal{D}_{st}} Y \int_0^1 \Sigma^{\frac{1}{2}}_{t_1,\dots,t_k}(s,X_s)\ dW_s,$$

where the $k\times(d+k)$-dimensional matrix Y has the form

$$Y = \begin{pmatrix} I_{k\times k} & -Y^* \end{pmatrix}, \qquad Y^* = \begin{pmatrix} B_{t_1}^T D^{-1} \\ \vdots \\ B_{t_k}^T D^{-1} \end{pmatrix}.$$

A straightforward calculation shows that the conditional covariance coincides with the conditional covariance of the finite dimensional distributions of the process defined in (4.1). Thus we are left to prove the tightness of the process $n^{\frac{1}{4}}(\hat{N}_t - N_t)$. We have the uniform decomposition

$$n^{\frac{1}{4}}(\hat{N}_t - N_t) = n^{\frac{1}{4}}(\hat{B}_t^0 - B_t^0) + n^{\frac{1}{4}} B_t^T D^{-1}(\hat{C}' - C) + o_p(1)$$

and will prove the tightness of each of the two sequences on the right hand side separately. To this end, we use Theorem VI. 4.5 in Jacod and Shiryaev [18], which says (in a special case) that a family of processes $(X_t^n)_{t\le 1}$ living on the same probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, P)$ is tight, as long as the following two conditions are satisfied:

(i) For all $\epsilon > 0$ there exists some $n_0 \in \mathbb{N}$ and $K > 0$ such that

$$P\Big(\sup_{t\le 1}|X_t^n| > K\Big) < \epsilon \tag{7.17}$$

for all $n > n_0$.

(ii) For all $\epsilon > 0$ we have

$$\lim_{\eta\downarrow 0}\limsup_{n\to\infty}\sup_{R,S\in\mathcal{T};R\le S\le R+\eta} P(|X_R^n - X_S^n| > \epsilon) = 0, \tag{7.18}$$

where $\mathcal{T}$ denotes the set of all stopping times bounded by 1.

For the first sequence note that (7.17) and (7.18) follows easily from Theorem 4, since it yields the stable convergence $n^{\frac{1}{4}}(\hat{B}_t^0 - B_t^0) \xrightarrow{\mathcal{D}_{st}} \int_0^t \gamma_s \, dW_s'$, and the process (γ_t) is bounded.
The proof of the tightness of the second sequence is slightly more involved. Note first that Cramér's rule gives $D^{-1} = \operatorname{adj}(D)/\det(D)$, where $\operatorname{adj}(D)$ denotes the adjoint matrix of D. From the boundedness of the functions σ_i^2 we conclude that each entry of $\operatorname{adj}(D)$ is bounded as well, and thus (3.8) yields $E[|D_{ij}^{-1}|^\beta] < K$ for all i and j and some $\beta > 0$. Moreover, we have $|B_t^i| < K$ uniformly in t, and using Markov's and Hölder's inequality we conclude for any $\epsilon > 0$ and arbitrary $1 \le i, j \le d$:

$$\begin{aligned} P\Big(\Big|n^{\frac{1}{4}}D_{ij}^{-1}(\hat{C}_j' - C_j)\Big| > K\Big) &\le K^{-\frac{\beta}{2}}\, n^{\frac{\beta}{8}}\, E\Big[|D_{ij}^{-1}|^{\frac{\beta}{2}}\, |\hat{C}' - C|^{\frac{\beta}{2}}\Big] \\ &\le K^{-\frac{\beta}{2}}\, E\Big[|D_{ij}^{-1}|^\beta\Big]^{\frac{1}{2}}\, E\Big[|n^{\frac{1}{4}}\,(\hat{C}' - C)|^\beta\Big]^{\frac{1}{2}}. \end{aligned}$$

From the proof of the previous theorem we know that the latter expectation is bounded (uniformly in n) as well. Thus, for all $\epsilon > 0$ there exists some $K > 0$ such that

$$P\Big(\sup_{t\le 1}\Big|n^{\frac{1}{4}}B_t^T D^{-1}(\hat{C}' - C)\Big| > K\Big) < \epsilon,$$

for all $n > n_0$. This gives (7.17). Note for the same reasons that

$$\lim_{\eta\downarrow 0}\limsup_{n\to\infty} P(|n^{\frac{1}{4}}\ \eta\ D_{ij}^{-1}(\hat{C}_j' - C_j)| > \epsilon) = 0$$

for all $\epsilon > 0$, and since we have $|B_R^i - B_S^i| \le K \cdot \eta$ for all such stopping times R, S with $R \le S \le R + \eta$ (7.18) follows and we are done. □

7.3 Proof of Theorem 2

For most parts the proof works in the same way as the ones for the preceding results. However, since $x \mapsto |x|$ is not differentiable, we cannot use Ito's formula to obtain a decomposition of $|\overline{Z}_k^n|$

and have to proceed in a slightly different manner in order to prove a result, which is similar to Theorem 4. To this end, we define the sequences $b_k(p), c_k(p), i_n(p)$ and $j_n(p)$ completely analogous, but with m_n replaced by l_n.
Note first that we have $E\Big[|\overline{U}_k^n|\Big] \le K \frac{1}{\sqrt{ln}} = O_p(n^{-(\frac{1}{4}+\frac{\delta}{2})})$, from which we conclude

$$\hat{S}_i = \frac{1}{\mu_1\sqrt{\rho\psi_2}} n^{-(\frac{3}{4}+\frac{\delta}{2})} \sum_{k=1}^{n-l_n} \bar{\sigma}_i(\frac{k}{n}, \hat{X}_{\frac{k}{n}})|\overline{X}_k^n| + o_p(n^{-(\frac{1}{4}-\frac{\delta}{2})}),$$

due to the constraint $\delta > \frac{1}{6}$. Furthermore, we have

$$n^{-(\frac{3}{4}+\frac{\delta}{2})} \sum_{k=1}^{n-l_n} \bar{\sigma}_i(\frac{k}{n}, \hat{X}_{\frac{k}{n}})|\overline{X}_k^n| = n^{-(\frac{3}{4}+\frac{\delta}{2})} \sum_{k=1}^{n-l_n} \bar{\sigma}_i(\frac{k}{n}, X_{\frac{k}{n}})|\overline{X}_k^n| + O_p(\frac{l_n}{n})$$

as in Theorem 3. Suppose $b_k(p) \le j < c_k(p)$. This yields

$$\bar{\sigma}_i(\frac{j}{n}, X_{\frac{j}{n}}) - \bar{\sigma}_i(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}}) = \frac{\partial}{\partial y}\bar{\sigma}_i(\frac{j}{n}, X_{\frac{b_k(p)}{n}}) \int_{\frac{b_k(p)}{n}}^{\frac{j}{n}} \sigma_s \, dW_s + O_p\Big(\frac{pl_n}{n}\Big)$$

once again and we end up with

$$\lim_{p\to\infty} \limsup_{n\to\infty} n^{-(\frac{1}{2}+\delta)} \sum_{k=0}^{j_n(p)} \sum_{j=b_k(p)}^{c_k(p)-1} \Big(\bar{\sigma}_i(\frac{j}{n}, X_{\frac{j}{n}}) - \bar{\sigma}_i(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}})\Big) \; |\overline{X}_j^n| = 0$$

for the same reasons as in the proof of Theorem 4. Moreover, we obtain from similar arguments as in the proof of Theorem 3 in Podolskij and Vetter [22]

$$\lim_{p\to\infty} \limsup_{n\to\infty} n^{-(\frac{1}{2}+\delta)} \sum_{k=0}^{j_n(p)} \sum_{j=b_k(p)}^{c_k(p)-1} \bar{\sigma}_i(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}}) \; \Big(|\overline{X}_j^n| - \sigma_{\frac{b_k(p)}{n}} \; |\overline{W}_j^n|\Big) = 0.$$

On the other hand we have

$$\lim_{p\to\infty} \limsup_{n\to\infty} n^{\frac{1}{4}-\frac{\delta}{2}} \Big\{ S_i - \sum_{k=0}^{j_n(p)} \Big(\int_{\frac{b_k(p)}{n}}^{\frac{c_k(p)}{n}} \sigma_{\frac{b_k(p)}{n}} \; \bar{\sigma}_i(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}}) \; ds$$
$$- \int_{\frac{c_k(p)}{n}}^{\frac{b_{k+1}(p)}{n}} \sigma_{\frac{c_k(p)}{n}} \; \bar{\sigma}_i(\frac{c_k(p)}{n}, X_{\frac{c_k(p)}{n}}) \; ds \Big) - \int_{\frac{i_n(p)}{n}}^{1} \sigma_{\frac{i_n(p)}{n}} \; \bar{\sigma}_i(\frac{i_n(p)}{n}, X_{\frac{i_n(p)}{n}}) \; ds \Big\} = 0$$

as before. Mimicing the proof of Theorem 4 we conclude

$$\lim_{p\to\infty} \limsup_{n\to\infty} n^{\frac{1}{4}-\frac{\delta}{2}} \Big\{ (\hat{R}_t^0 - R_t^0) - \sum_{k=0}^{j_n(p)} \tilde{Y}(k,p) 1_{\{\frac{c_k(p)}{n} \le t\}} \Big\} \;\; = \;\; 0,$$
$$\lim_{p\to\infty} \limsup_{n\to\infty} n^{\frac{1}{4}-\frac{\delta}{2}} \Big\{ (\hat{S}_i - S_i) - \sum_{k=0}^{j_n(p)} \bar{\sigma}_i(\frac{b_k(p)}{n}, X_{\frac{b_k(p)}{n}}) \; \tilde{Y}(k,p) \Big\} \;\; = \;\; 0,$$

where

$$\tilde{Y}(k,p) = \frac{1}{\mu_1\sqrt{\rho\psi_2}} n^{-(\frac{3}{4}+\frac{\delta}{2})}\sigma_{\frac{b_k(p)}{n}} \sum_{j=b_k(p)}^{c_k(p)-1} (|\overline{W}_j^n| - E[|\overline{W}_j^n|]).$$

We have

$$\begin{aligned} E\Big[\Big(\sum_{j=b_k(p)}^{c_k(p)-1}(|\overline{W}_j^n| - E[|\overline{W}_j^n|])\Big)^2\Big] &= 2\sum_{j=b_k(p)}^{c_k(p)-1}\sum_{i=0}^{l_n}(E[|\overline{W}_j^n\ \overline{W}_{i+j}^n|] - E[|\overline{W}_j^n|]^2) + O_p(\frac{l_n^3}{n}) \\ &= 2\ p\ l_n\sum_{i=0}^{l_n}(E[|\overline{W}_0^n\ \overline{W}_i^n|] - E[|\overline{W}_0^n|]^2) + O_p(\frac{l_n^3}{n}) \end{aligned}$$

and

$$E[|\overline{W}_0^n\ \overline{W}_i^n|] - E[|\overline{W}_0^n|]^2 = \rho\psi_2\ n^{-\frac{1}{2}+\delta}(E[|N_0\ N_i|] - \mu_1^2) + o_p(n^{-\frac{1}{2}+\delta})$$

with $N_0, N_i \sim \mathcal{N}(0,1)$ and $E[N_0\ N_i] = \frac{\phi_2(\frac{i}{l_n})}{\psi_2}$. From $\mu_1^2 = \frac{2}{\pi}$ and Wellner and Smythe [26] we conclude $E[|N_0\ N_i|] - \mu_1^2 = f\Big(\frac{\phi_2(\frac{i}{l_n})}{\psi_2}\Big)$ with f as defined in Theorem 2, thus a Riemann sum argument gives

$$E[(\tilde{Y}(k,p))^2|\mathcal{F}_{\frac{b_k(p)}{n}}] = \frac{2\ \Xi}{\mu_1^2}\sigma^2_{\frac{b_k(p)}{n}}\frac{pl_n^2}{n^2} + o_p(\frac{pl_n^2}{n^2}).$$

Theorem 2 can now be derived easily and the details are omitted for the sake of brevity. □

Acknowledgements The authors would like to thank Martina Stein, who typed parts of this manuscript with considerable technical expertise. This work has been supported in part by the Collaborative Research Center "Statistical modeling of nonlinear dynamic processes" (SFB 823) of the German Research Foundation (DFG), the BMBF Project SKAVOE and the NIH grant award IR01GM072876:01A1.

References

[1] N. J. Achieser (1956). Theory of approximation. Dover Publications Inc., New York.

[2] Y. Ait-Sahalia (1996). Testing Continuous-Time Models of the Spot Interest Rate, Review of Financial Studies 9, 385-426.

[3] D. Ahn, B. Gao (1999). A Parametric Nonlinear Model of Term Structure Dynamics. Review of Financial Studies 12, 721-762.

[4] Y. Amihud, H. Mendelson (1987). Trading mechanisms and stock returns: An empirical investigation. Journal of Finance 42(3), 533-553.

[5] F. Black (1986). Noise. Journal of Finance 41, 529-543.

[6] F. Black, M. Scholes (1973). The Pricing of Options and Corporate Liabilities. Journal of Political Economy, 81, 637-659.

[7] K. C. Chan, G.A. Karolyi, F.A. Longstaff, and A.B. Sanders (1992). An Empirical Comparison of Alternative Models of the Short-Term Interest Rate. Journal of Finance 47, 1209-1227.

[8] J. C. Cox, J.E. Ingersoll, S.A. Ross (1985). A Theory of the Term Structure of Interest Rates, Econometrica 53, 385-407.

[9] V. Corradi, H. White (1999). Specification tests for the variance of a diffusion process. Journal of Time Series Analysis 20, 253-270.

[10] H. Dette, M. Podolskij (2008). Testing the parametric form of the volatility in continuous time diffusion models–a stochastic process approach. Journal of Econometrics 143, 56-73.

[11] H. Dette, M. Podolskij, M. Vetter (2006). Estimation of integrated volatility in continuous time financial models with applications to goodness-of-fit testing. Scandinavian Journal of Statistics 33, 259-278.

[12] A. R. Gallant (1987). Nonlinear Statistical Models. Wiley, N.Y.

[13] A. Gloter, J. Jacod (2001). Diffusions with measurement errors. II. Optimal estimators. ESAIM Probability and Statistics 5, 243-260.

[14] S. L. Heston (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. Review of Financial Studies 6(2), 327-343.

[15] J. Hull, A. White (1987). The Pricing of Options on Assets with Stochastic Volatilities. Journal of Finance, 42(2), 281-300.

[16] J. Jacod (2008). Asymptotic properties of realized power variations and related functionals of semimartingales. Stochastic Process. Appl. 118 (4), 517-559.

[17] J. Jacod, Y. Li, P. A. Mykland, M. Podolskij, M. Vetter (2007). Microstructure Noise in the Continuous Case: The Pre-Averaging Approach. Stochastic Process. Appl., forthcoming.

[18] J. Jacod, A. N. Shiryaev (2003). Limit theorems for stochastic processes. Springer, Berlin.

[19] L. Harris (1990). Estimation of stock variance and serial covariance from discrete observations. Journal of Financial and Quantitative Analysis 25, 291-306.

[20] L. Harris (1991). Stock price clustering and discreteness. Review of Financial Studies 4(3), 389-415.

[21] M. Podolskij, M. Vetter (2009). Estimation of Volatility Functionals in the Simultaneous Presence of Microstructure Noise and Jumps. Bernoulli, forthcoming.

[22] M. Podolskij, M. Vetter (2009). Bipower-type estimation in a noisy diffusion setting. Stoch. Proc. Appl., forthcoming.

[23] D. Revuz, M. Yor (1999). Continuous Martingales and Brownian Motion. Springer, Berlin.

[24] G. A. F. Seber, C. J. Wild (1989). Nonlinear Regression. Wiley, N.Y.

[25] O. Vasicek (1977). An Equilibrium Characterization of the Term Structure. Journal of Financial Economics 5, 177-188.

[26] J. A. Wellner, R. T. Smythe (2002). Computing the Covariance of two Brownian Area Integrals. Statist. Neerlandica 56(1), 101-109.

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

ω	0.01			0.0025			0.000625		
θ \ α	.025	.05	.1	.025	.05	.1	.025	.05	.1
1	.01	.02	.038	.023	.058	.104	.024	.047	.101
.75	.004	.01	.02	.004	.009	.022	.003	.007	.015
.5	.003	.006	.013	.002	.004	.014	.000	.000	.002
.25	.002	.004	.015	.001	.002	.003	.001	.003	.004
0	.000	.005	.019	.003	.006	.015	.004	.007	.016

Table 1: *Simulated level of the test (1.2) for various choices of* ω *and* θ*, where the true volatility function is* $\sigma^2(t,x) = \theta + (1-\theta)x^2$ *and the noise terms* U *are normally distributed with mean zero and variance* ω^2*. In all cases the sample size is given by* $n = 16384$*.*

n	256			1024		
ω \ α	.025	.05	.1	.025	.05	.1
.001	.033	.062	.111	.333	.415	.512
.002	.158	.243	.324	.810	.862	.907
.004	.392	.518	.650	.993	.996	.998
.005	.497	.628	.742	.991	.994	.998
.01	.596	.754	.873	.987	.998	.999

Table 2: *Simulated level of the bootstrap test proposed by Dette and Podolskij [10], where the volatility function equals* $H_0 : \sigma^2(t,x) = \theta x^2$*, but the observations are corrupted with normally distributed noise having variance* ω^2*.*

n \ α	.025	.05	.1
256	.008	.022	.058
1024	.007	.023	.062
4096	.013	.029	.079
16384	.017	.038	.077

Table 3: *Simulated nominal level of the test, which rejects the null hypothesis of homoscedasticity for a large value of* $\sup \left| \frac{A_n(t)}{\hat{s}_t} \right|$*, using the critical values from the asymptotic theory [see (6.1)]. The variance of the noise process is defined by* $n\omega^2 = 0.1024$*.*

$\sigma_1^2(t,x)$	1			x^2		
n \ α	.025	.05	.1	.025	.05	.1
256	.019	.046	.113	.03	.066	.118
1024	.02	.049	.099	.034	.07	.119

Table 4: *Simulated level of the bootstrap test based on the standardised Kolmogorov-Smirnov functional of* $(\hat{N}_t)$ *for various hypotheses. The variance of the noise process is defined by* $n\omega^2 = 0.1024$.

n \ α	.025	.05	.1
256	.040	.076	.136
1024	.032	.057	.119

Table 5: *Simulated level of the bootstrap test based on the standardised Kolmogorov-Smirnov functional of* $(\hat{M}_t)$ *for* $\sigma(t,x) = \theta|x|$. *The variance of the noise process is defined by* $n\omega^2 = 0.1024$.

alt	1			$1+|x|$			Heston		
n \ α	.025	.05	.1	.025	.05	.1	.025	.05	.1
256	.057	.128	.237	.073	.152	.263	.722	.870	.941
1024	.170	.230	.329	.224	.326	.465	.975	.980	.985

Table 6: *Simulated rejection probabilities of the bootstrap test based on the standardised Kolmogorov-Smirnov functional of* $(\hat{N}_t)$ *for various alternatives. The data is simulated with* $\sigma^2(t,x) = \theta|x|^2$ *and the variance of the noise process is defined by* $n\omega^2 = 0.1024$.