Testing the maximal rank of the volatility process for continuous diffusions observed with noise

TOBIAS FISSLER[1] and MARK PODOLSKIJ[2]

[1]*Institute of Mathematical Statistics and Actuarial Science, University of Bern, Sidlerstrasse 5, 3012 Bern, Switzerland. E-mail: tobias.fissler@stat.unibe.ch*
[2]*Department of Mathematics, University of Aarhus, Ny Munkegade 118, 8000 Aarhus C, Denmark. E-mail: mpodolskij@math.au.dk*

In this paper, we present a test for the maximal rank of the volatility process in continuous diffusion models observed with noise. Such models are typically applied in mathematical finance, where latent price processes are corrupted by microstructure noise at ultra high frequencies. Using high frequency observations, we construct a test statistic for the maximal rank of the time varying stochastic volatility process. Our methodology is based upon a combination of a matrix perturbation approach and pre-averaging. We will show the asymptotic mixed normality of the test statistic and obtain a consistent testing procedure. We complement the paper with a simulation and an empirical study showing the performances on finite samples.

Keywords: continuous Itô semimartingales; high frequency data; microstructure noise; rank testing; stable convergence

1. Introduction

In the last twenty years, asymptotic theory for high frequency data has received a great deal of attention in probability and statistics. This is mainly motivated by financial applications, where observations of stock prices are recorded very frequently. In an ideal world, that is, under no-arbitrage conditions, price processes must follow a semimartingale, which is a celebrated result of Delbaen and Schachermayer [4]. We refer to a monograph [9] for a comprehensive study of limit theorems for Itô semimartingales and their manifold applications in statistics.

Despite the aforementioned theoretical result, at ultra high frequencies, the financial data is contaminated by *microstructure noise* such as rounding errors, bid-ask bounces and misprints. One of the standard models for the microstructure noise is an additive i.i.d. process independent of the latent price (see, e.g., [3,14] among many others; an extension of this model can be found in [7]). More formally, the model is given as

$$Y_{t_i} = X_{t_i} + \varepsilon_{t_i} \qquad \text{with } dX_t = b_t\, dt + \sigma_t\, dW_t, \tag{1.1}$$

where $(X_t)_{t \in [0,T]}$ is a d-dimensional continuous Itô semimartingale, $t_i = i\Delta_n$ and $(\varepsilon_t)_{t \in [0,T]}$ is a d-dimensional i.i.d. process independent of X with

$$\mathbb{E}[\varepsilon_t] = 0 \quad \text{and} \quad \mathbb{E}[\varepsilon_t \varepsilon_t^\star] =: \Sigma \in \mathbb{R}^{d \times d}. \tag{1.2}$$

Here and throughout the paper, we denote with $x^\star$ the transpose of x, be it a vector or a matrix. We are in the framework of infill asymptotics, that is, $\Delta_n \to 0$ while T remains fixed. This paper is devoted to the test for the maximal rank of the co-volatility matrix $c_t = \sigma_t \sigma_t^\star$ of the unobserved diffusion process X. We remark that this is an equivalent formulation of the following problem: What is the minimal amount of independent Brownian motions required for modeling the d-dimensional diffusion X? Answering this question might give a direct economical interpretation of the financial data at hand. Furthermore, testing for the full rank of c_t is connected to testing for completeness of financial markets.

In a recent paper [8], the described statistical problem has been solved in a continuous diffusion setting without noise (we also refer to an earlier article [6] for a related problem). The main idea is based upon a matrix perturbation method, which helps to identify the rank of a given matrix. The maximal rank of the stochastic co-volatility process $(c_t)_{t\in[0,T]}$ is then asymptotically identified via a certain ratio statistic, which uses the scaling property of a Brownian motion. Clearly, the test statistic becomes invalid in the framework of continuous diffusion models observed with noise. To overcome this problem, we apply the pre-averaging approach, which has been originally proposed in [7,11]. As the name suggests, weighted averages of increments of the process Y are built over a certain window in order to eliminate the influence of the noise to some extent. This in turn gives the possibility to infer the co-volatility process $(c_t)_{t\in[0,T]}$. The size of the pre-averaging window k_n is typically chosen as $k_n = O(\Delta_n^{-1/2})$ and objects as the integrated co-volatility $\int_0^T c_t \, dt$ can be estimated with the convergence rate of $\Delta_n^{-1/4}$, which is known to be optimal.

At this stage, we would like to stress that combining the pre-averaging approach and the matrix perturbation method is by far not trivial. There are mainly two problems that need to be solved. First of all, when using the optimal window size of $k_n = O(\Delta_n^{-1/2})$ in the pre-averaging approach, the diffusion and the noise parts have the same order, and it becomes virtually impossible to distinguish the rank of the co-volatility from the unknown rank of the covariance matrix Σ. Hence, we will choose a proper sub-optimal window size to still obtain a reasonable convergence rate for the test statistic. The second and more severe problem is that the ratio statistic proposed in [8] heavily relies on the scaling property of a Brownian motion. This scaling property is not shared by an i.i.d. noise process introduced in (1.2). Thus, a much deeper probabilistic analysis of the main statistic is required to come up with a valid testing procedure.

The paper is organized as follows. Section 2 gives the probabilistic description of the model, presents the main assumptions and defines the testing hypotheses. The background on matrix perturbation and pre-averaging method is demonstrated in Section 3. Section 4 presents the main results of the paper as well as an extension of the method to the usage of overlapping increments. Section 5 is concerned with a simulation study and a real data example. All proofs are collected in Section 6.

2. The setting and main assumptions

We start with a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$, satisfying the usual assumptions, on which all stochastic processes are defined. As indicated at (1.1), we observe the d-dimensional

process $Y = X + \varepsilon$ at time points $i \Delta_n$, $i = 0, 1, \ldots, [T/\Delta_n]$. The process X is given via

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

where $(b_t)_{t \in [0,T]}$ is a d-dimensional drift process, $(\sigma_t)_{t \in [0,T]}$ is a $\mathbb{R}^{d \times q}$-valued volatility process and W denotes a q-dimensional Brownian motion. We introduce the notation

$$c_t = \sigma_t \sigma_t^\star, \qquad r_t = \operatorname{rank}(c_t), \qquad R_t = \sup_{s \in [0,t)} r_s.$$

We need more structural assumptions on the processes b and σ.

Assumption (A). *The processes b and σ have the form*

$$\sigma_t = \sigma_0 + \int_0^t a_s \, ds + \int_0^t v_s \, dW_s,$$

$$b_t = b_0 + \int_0^t a_s' \, ds + \int_0^t v_s' \, dW_s, \tag{2.2}$$

$$v_t = v_0 + \int_0^t a_s'' \, ds + \int_0^t v_s'' \, dW_s,$$

where b_t and a_t' are $\mathbb{R}^d$-valued, σ_t, a_t and v_t' are $\mathbb{R}^{d \times q}$-valued, v_t and a_t'' are $\mathbb{R}^{d \times q \times q}$-valued, and v_t'' is $\mathbb{R}^{d \times q \times q \times q}$-valued, all those processes being adapted. Finally, the processes a_t, v_t', v_t'' are càdlàg and the processes a_t', a_t'' are locally bounded.

Notice that Assumption (A) is exactly the same assumption, which has been imposed in [8]. We remark that, by enlarging the dimension q of the Brownian motion W if necessary, we may assume without loss of generality that all processes X, b, σ, v are driven by the same Brownian motion. In the framework of a stochastic differential equation, that is, when $b_t = h_1(X_t)$ and $\sigma_t = h_2(X_t)$, Assumption (A) is automatically satisfied whenever $h_1 \in C^2(\mathbb{R}^d)$ and $h_2 \in C^4(\mathbb{R}^d)$ (due to Itô's formula). We also remark that Assumption (A) is rather unusual in the literature. Indeed, for classical high frequency statistics, such as for example, power variations (cf. [2]), only the first line of (2.2) is required. However, when $R_T < d$ our test statistic, which will be introduced in Section 4, turns out to be degenerate and, in contrast to classical cases, we require a higher order stochastic expansion of the increments of X. This explains the role of the second and third line of (2.2). Finally, we specify our assumptions on the noise process ε introduced at (1.2).

Assumption (E). *The i.i.d. process $(\varepsilon_t)_{t \in [0,T]}$ is $(\mathcal{F}_t)$-adapted and independent of a, a', a'', v', v'', W, hence also independent of b, σ and X. Furthermore, it is Gaussian, meaning that $\varepsilon_t \sim \mathcal{N}_d(0, \Sigma)$ and $\mathbb{E}[\varepsilon_s \varepsilon_t^\star] = 0$ for all $s, t \in [0, T]$ with $s \neq t$.*

Remark 2.1. Theoretically, we could discuss a more general structure of the noise. In particular, we could give up the assumption of the Gaussianity. What we really require is the mutual inde-

pendence of the noise at different times as well as the existence of the moments up to a certain order.

Now, for any $r \in \{0, 1, \ldots, d\}$, we introduce the following subsets of Ω:

$$\Omega_T^r := \{\omega \in \Omega : R_T(\omega) = r\}, \qquad \Omega_T^{\leq r} := \{\omega \in \Omega : R_T(\omega) \leq r\}. \tag{2.3}$$

Notice that the sets Ω_T^r and $\Omega_T^{\leq r}$ are indeed $\mathcal{F}_T$-measurable. This can be justified as follows. The rank r_t is the biggest integer $r \leq d$ such that the sum of the determinants of the matrices $(c_t^{ij})_{i,j \in J}$, where J runs through all subsets of $\{1, \ldots, d\}$ with r points, is positive; see, for example, [6], Lemma 3. Since the mapping $t \mapsto c_t$ is continuous by Assumption (A), this implies that for any r the random set $\{t : r_t(\omega) > r\}$ is open in $[0, T)$, so the mapping $t \mapsto r_t$ is lower semi-continuous. The very same argument proves that the random set

$$\{t \in [0, T) : R_T(\omega) = r_t(\omega)\}$$

is non-empty and open for each $\omega \in \Omega$. Hence, this set has a positive Lebesgue measure, which helps to statistically identify the maximal rank R_T (in contrast to lower ranks $r_t < R_T$, which might be attained at a single point on the interval $[0, T]$).

The following discussion is devoted to testing the null hypothesis $H_0 : R_T = r$ against the alternative $H_1 : R_T \neq r$ (or $H_0 : R_T \leq r$ against $H_1 : R_T > r$). Notice that this a *pathwise hypothesis*, since we test whether a given path ω belongs to Ω_T^r (or $\Omega_T^{\leq r}$) or not. It is in general impossible to know whether this hypothesis holds for another path $\omega' \in \Omega$.

3. Matrix perturbation and pre-averaging approach

3.1. Matrix perturbation method

The matrix perturbation method is a numerical approach to the computation of the rank of a given matrix. It has been introduced in [8] in the context of rank testing. To explain the main idea of our method, we need to introduce some notation. Recall that d and q are the dimensions of X and W, respectively. Let $\mathcal{M}$ denote the set of all $d \times d$ matrices and $\mathcal{M}_r$, $r \in \{0, \ldots, d\}$, the set of all matrices in $\mathcal{M}$ with rank r. Furthermore, let $\mathcal{M}'$ be the set of all $d \times q$ matrices. For any matrix A, we denote by A_i the ith column of A; for any vectors $x_1, \ldots, x_d$ in $\mathbb{R}^d$, we write $\mathrm{mat}(x_1, \ldots, x_d)$ for the matrix in $\mathcal{M}$ whose ith column is the column vector x_i. For $r \in \{0, \ldots, d\}$ and $A, B \in \mathcal{M}$ we define the quantity

$$\mathcal{M}_{A,B}^r := \{G \in \mathcal{M} : G_i = A_i \text{ or } G_i = B_i \text{ with } \#\{i : G_i = A_i\} = r\}.$$

In other words, $\mathcal{M}_{A,B}^r$ is the set of all matrices $G \in \mathcal{M}$ with r columns equal to those of A and the remaining $d - r$ ones equal to those of B (all of them being at their original places). We also define the number

$$\gamma_r(A, B) := \sum_{G \in \mathcal{M}_{A,B}^r} \det(G). \tag{3.1}$$

We demonstrate the main ideas of the matrix perturbation approach for a deterministic problem first. Let $A \in \mathcal{M}$ be an unknown matrix with unknown rank r. Assume that, although A is unknown, we have a way of computing $\det(A + \lambda B)$ for all $\lambda > 0$ and some given matrix $B \in \mathcal{M}_d$. The multilinearity property of the determinant implies the following asymptotic expansion

$$\det(A + \lambda B) = \lambda^{d-r} \gamma_r(A, B) + O(\lambda^{d-r+1}) \qquad \text{as } \lambda \downarrow 0. \tag{3.2}$$

This expansion is the key to identification of the unknown rank r. Indeed, when $\gamma_r(A, B) \neq 0$ we deduce that

$$\frac{\det(A + 2\lambda B)}{\det(A + \lambda B)} \to 2^{d-r} \qquad \text{as } \lambda \downarrow 0. \tag{3.3}$$

However, it is impossible to choose a matrix $B \in \mathcal{M}_d$ which guarantees $\gamma_r(A, B) \neq 0$ for all $A \in \mathcal{M}_r$. To solve this problem, we can use a *random* perturbation. As it has been shown in [8], for any $A \in \mathcal{M}_r$ we have $\gamma_r(A, B) \neq 0$ almost surely when B is the random matrix whose entries are independent standard normal (in fact, the random variable $\gamma_r(A, B)$ has a Lebesgue density). This is intuitively clear, because the multivariate standard normal distribution does not prefer directions. It is exactly this idea which will be the core of our testing procedure.

3.2. Pre-averaging approach

In this subsection, we briefly introduce the pre-averaging method; we refer to for example, [7,11] for a more detailed exposition.

Let $g : [0, 1] \to \mathbb{R}$ be a weight function with $g(0) = g(1) = 0$, which is continuous, piecewise C^1 with piecewise Lipschitz derivative g' and $\int_0^1 g^2(x) \, dx > 0$. A canonical choice of such a function is given by $g(x) = \min(x, 1 - x)$; see [7] for its interpretation. Now, let $(k_n)_{n \geq 1}$ be a sequence of positive integers representing the window size such that $k_n \to \infty$ and $u_n := k_n \Delta_n \to 0$. For any stochastic process V, we define the pre-averaged increments via

$$\overline{V}_i^n := \sum_{j=1}^{k_n-1} g\left(\frac{j}{k_n}\right) \Delta_{i+j}^n V = -\sum_{j=0}^{k_n-1} \left(g\left(\frac{j+1}{k_n}\right) - g\left(\frac{j}{k_n}\right)\right) V_{(i+j)\Delta_n}, \tag{3.4}$$

where $\Delta_i^n V := V_{i\Delta_n} - V_{(i-1)\Delta_n}$. Roughly speaking, this local averaging procedure reduces the influence of the noise process when we apply it to the noisy diffusion process Y defined at (1.1). Indeed, we may show that

$$\overline{X}_i^n = O_{\mathbb{P}}(\sqrt{k_n \Delta_n}) \quad \text{and} \quad \overline{\varepsilon}_i^n = O_{\mathbb{P}}(\sqrt{1/k_n}),$$

where the first approximation is essentially justified by the independence of the increments of W and the first identity of (3.4), and the second approximation follows from the i.i.d. structure of the noise process and the second identity of (3.4). We clearly see that a large k_n increases the influence of the diffusion part X and diminishes the influence of the noise part ε. However, in standard statistical problems, for example, estimation of quadratic variation, the optimal rate of

convergence is obtained when the contributions of both terms are balanced. This results in the choice of the window size k_n with $k_n\sqrt{\Delta_n} = \theta + o(\Delta_n^{1/4})$, where $\theta \in (0, \infty)$. With this window size, we deduce for instance that

$$\sqrt{\Delta_n} \sum_{i=0}^{[T/\Delta_n]-k_n+1} (\overline{Y}_i^n)(\overline{Y}_i^n)^\star \xrightarrow{\mathbb{P}} \theta\psi_2 \int_0^T c_t\, dt + \theta^{-1}\psi_1 T\Sigma,$$

where the constants ψ_1 and ψ_2 are defined by

$$\psi_1 := \int_0^1 \left(g'(x)\right)^2 dx, \qquad \psi_2 := \int_0^1 g^2(x)\, dx$$

(cf. [7]). The bias can be corrected via

$$C_t^n = \frac{\sqrt{\Delta_n}}{\theta\psi_2} \sum_{i=0}^{[T/\Delta_n]-k_n+1} (\overline{Y}_i^n)(\overline{Y}_i^n)^\star - \frac{\psi_1\Delta_n}{2\theta^2\psi_2} \sum_{i=1}^{[T/\Delta_n]} (\Delta_i^n Y)(\Delta_i^n Y)^\star \xrightarrow{\mathbb{P}} \int_0^T c_t\, dt,$$

and the statistic C_t^n becomes a consistent estimator of the quadratic covariation of X with convergence rate $\Delta_n^{-1/4}$. This rate is known to be optimal.

As explained in the Introduction, the optimal choice of the window size k_n as introduced above would not lead to a feasible testing procedure for the maximal rank R_T. Due to the complex structure of the test statistic, which will be introduced in Section 4, there is generally no de-biasing procedure as above (unless $R_T = d$; see Remark 4.12). For this reason we introduce the following window size k_n:

$$k_n\Delta_n^{2/3} = \theta + o\left(\Delta_n^{1/6}\right), \qquad \theta \in (0, \infty), \qquad u_n := k_n\Delta_n. \tag{3.5}$$

Within the framework of our test statistic, this choice of k_n leads to an optimal rate of convergence, which becomes $\Delta_n^{-1/6}$ (although better rates of convergence are theoretically possible, but they can only be derived under Ω_T^d; see again Remark 4.12). We show the intuition behind this choice in the next section. We remark that an easier choice of the window size would be $k_n = O(\Delta_n^{-3/4})$, which would completely eliminate the influence of the noise process on the central limit theorem. However, this would lead to a slower rate of convergence of $\Delta_n^{-1/8}$. For this reason, we dispense with the exact exposition of this case.

4. Main results

4.1. Test statistic

In this subsection, we introduce a random perturbation of the original data and define the main statistics. Following the basic ideas of [8] and the motivation of Section 3.1, we define a d-dimensional "perturbation" process X' by

$$X_t' = \widetilde{\sigma}\, W_t',$$

where $\widetilde{\sigma} \in \mathcal{M}_d$ is a positive definite deterministic $d \times d$ matrix and W' is a d-dimensional Brownian motion. Without loss of generality, we may assume that W' is also defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0,T]}, \mathbb{P})$. Let $\mathcal{G} \subset \mathcal{F}$ be the sub-σ-algebra, which is generated by all processes appearing in Assumption (A) and by the noise process ε. We assume that W' is independent of $\mathcal{G}$. Now, we use X' to define the perturbed process

$$Z_t^{n,\kappa} := Y_t + \sqrt{\kappa u_n} X_t', \tag{4.1}$$

where $\kappa = 1, 2$ and the sequence u_n is defined at (3.5). In some sense, the perturbation process X_t' plays the role of the random perturbation matrix B introduced after (3.3). As we will see below, our two main statistics will be constructed at two different frequencies Δ_n and $2\Delta_n$, which will be indicated by the constant $\kappa = 1, 2$.

Recall the definition of the pre-averaged quantity $\overline{V}_i^n$ introduced in (3.4) for a stochastic process V. We sometimes write $\overline{V}(g)_i^n$ instead of $\overline{V}_i^n$ if we want to stress the dependency of the term $\overline{V}_i^n$ on the weight function g. Furthermore, we use the notation $\overline{V}(g)_i^{n,\kappa}$ to indicate that the quantity $\overline{V}(g)_i^n$ is built using frequency $\kappa\Delta_n$ with $\kappa = 1, 2$, that is,

$$\overline{V}(g)_i^{n,\kappa} = \sum_{j=1}^{k_n-1} g\left(\frac{j}{k_n}\right)(V_{(i+\kappa j)\Delta_n} - V_{(i+\kappa(j-1))\Delta_n}). \tag{4.2}$$

If $V = Z^{n,\kappa}$ defined at (4.1), we will slightly abuse the notation introduced in (4.2) and use the convention $\overline{Z}(g)_i^{n,\kappa} := \overline{Z^{n,\kappa}}(g)_i^{n,\kappa}$. Now, we define our main test statistics via

$$S(g)_T^{n,\kappa} = 3du_n \sum_{i=0}^{[T/3du_n]-1} f\left(\left(\overline{Z}(g)_{((3i+\kappa-1)d+\kappa(j-1))k_n}^{n,\kappa} / \sqrt{\kappa u_n}\right)_{j=1,\ldots,d}\right), \tag{4.3}$$

for $\kappa = 1, 2$ with the test function f on $(\mathbb{R}^d)^d$ given as

$$f(x_1, \ldots, x_d) := \det\big(\mathrm{mat}(x_1, \ldots, x_d)\big)^2. \tag{4.4}$$

Note that the summands in (4.3) use non-overlapping increments of the process $Z^{n,\kappa}$, and also the statistics $S(g)_T^{n,1}$ and $S(g)_T^{n,2}$ are based on distinct increments; see also Section 4.5 for a more detailed discussion.

Remark 4.1. The statistic $S(g)_T^{n,\kappa}$ is similar in spirit to the one introduced in [8], where a d-dimensional continuous Itô semimartingale without noise has been considered. Therein, the statistics $S_T^{n,\kappa}$ defined in [8], equation (2.13), which use the raw increments instead of pre-averaged ones, satisfy the following law of large numbers

$$\frac{S_T^{n,2}}{S_T^{n,1}} \xrightarrow{\mathbb{P}} 2^{d-R_T}.$$

This should be compared with the motivation described at (3.3). The latter convergence asymptotically identifies the maximal rank R_T. The crucial difference to our framework is that this

convergence is no longer valid when we use the statistics $S(g)_T^{n,\kappa}$ introduced in (4.3). It relies on the fact that the noise process ε does not have the scaling property of the driving Brownian motion W. To overcome this issue, we will not only use different frequencies Δ_n and $2\Delta_n$, but also two different weight functions g and h, which are connected through certain identities. For this purpose, a very thorough analysis of the asymptotic behaviour of $S(g)_T^{n,\kappa}$ is required.

Remark 4.2. Let us explain the choice of the window size k_n introduced at (3.5) and the perturbation rate $\sqrt{\kappa u_n}$. Under Assumptions (A) and (E), we will prove the following asymptotic decomposition for $i = 0, \ldots, [T/3du_n] - 1$ and $\kappa = 1, 2$

$$
\frac{1}{\sqrt{\kappa u_n}} \mathrm{mat}\big(\overline{Z}(g)_{(3i+\kappa-1)dk_n}^{n,\kappa}, \ldots, \overline{Z}(g)_{((3i+\kappa-1)d+\kappa(d-1))k_n}^{n,\kappa}\big)
$$
$$
= A(g)_i^{n,\kappa} + \sqrt{\kappa u_n}\big(B(1,g)_i^{n,\kappa} + B(2,g)_i^{n,\kappa} + B(3,g)_i^{n,\kappa}\big) \qquad (4.5)
$$
$$
+ \kappa u_n C(g)_i^{n,\kappa} + \kappa u_n D(g)_i^{n,\kappa},
$$

where the $\mathbb{R}^{d \times d}$-valued sequences $A(g)_i^{n,\kappa}, C(g)_i^{n,\kappa}, D(g)_i^{n,\kappa}$ and $B(g)_i^{n,\kappa} := B(1,g)_i^{n,\kappa} + B(2,g)_i^{n,\kappa} + B(3,g)_i^{n,\kappa}$ are tight. The matrix $A(g)_i^{n,\kappa}$, which is the dominating term in the expansion, is defined by

$$
A(g)_i^{n,\kappa} = \frac{\sigma_{(3i+\kappa-1)du_n}}{\sqrt{\kappa u_n}} \mathrm{mat}\big(\overline{W}(g)_{(3i+\kappa-1)dk_n}^{n,\kappa}, \ldots, \overline{W}(g)_{((3i+\kappa-1)d+\kappa(d-1))k_n}^{n,\kappa}\big),
$$

while $B(1,g)_i^{n,\kappa}$ depends on b, v introduced in (2.2), $B(2,g)_i^{n,\kappa}$ comes solely from the perturbation X' and $B(3,g)_i^{n,\kappa}$ is associated with the noise process ε (the third order term $C(g)_i^{n,\kappa}$ is connected to a, v', v'' and the term $D(g)_i^{n,\kappa}$ depends on a, a', a'', v', v'', defined in (2.2)). Since $\det(A(g)_i^{n,\kappa}) = 0$ whenever $R_T < d$, our statistic $S(g)_T^{n,\kappa}$ is degenerate in the sense that the second order term enters the law of large numbers. At this stage, we realize that the choice of the window size $k_n = O(\Delta_n^{-2/3})$ and the perturbation rate $\sqrt{\kappa u_n}$ creates a balance between the second order term $B(1,g)_i^{n,\kappa}$ in the stochastic expansion coming from the diffusion process, $B(3,g)_i^{n,\kappa}$ stemming from the noise process ε and $B(2,g)_i^{n,\kappa}$ induced by the perturbation process X'. The classical choice $k_n = O(\Delta_n^{-1/2})$ would make the noise part one of the dominating terms, but in this case, the estimation of the maximal rank R_T would be virtually impossible since we impose no assumptions on the covariance matrix Σ of the noise. On the other hand, when $k_n = O(\Delta_n^{-3/4})$ the noise part would enter the third order term and thus would not influence the limit theory. Although the asymptotic results become much easier in the latter case, the convergence rate gets rather low ($\Delta_n^{-1/8}$). Hence, within the framework of our test statistic, the choice $k_n = O(\Delta_n^{-2/3})$ meets the balance between feasibility of the testing procedure and the optimal rate of convergence.

Clearly, $B(g)_i^{n,\kappa}$ plays the role of the perturbation matrix B defined in Section 3.1 while $\lambda = \sqrt{\kappa u_n}$. Since it is impossible to guarantee that the matrices $B(1,g)_i^{n,\kappa}$ and $B(3,g)_i^{n,\kappa}$ have full rank, we require the presence of the matrix $B(2,g)_i^{n,\kappa}$ to insure almost sure invertibility of the sum. Thus, the perturbation process X' plays the role of regularization.

4.2. Notation

In order to state the limit theory for the statistics $S(g)_T^{n,\kappa}$, we need to introduce some notation. For any weight function g, we define the quantities

$$
\begin{aligned}
\psi_1(g) &= \int_0^1 \left(g'(x)\right)^2 dx, \qquad \psi_2(g) = \int_0^1 g^2(x)\, dx, \\
\psi_3(g) &= \int_0^1 g(x)\, dx, \qquad \psi_4(g) = \int_0^1 x g^2(x)\, dx.
\end{aligned}
\tag{4.6}
$$

For $r \in \{0, 1, \ldots, d\}$, we define the function F_r on $(\mathbb{R}^{2d})^d$ by

$$
F_r(v_1, \ldots, v_d) = \gamma_r\left(\mathrm{mat}(x_1, \ldots, x_d), \mathrm{mat}(y_1, \ldots, y_d)\right)^2, \qquad v_j = \begin{pmatrix} x_j \\ y_j \end{pmatrix} \in \mathbb{R}^{2d}, \quad (4.7)
$$

where γ_r was introduced at (3.1). Let $\overline{W}$ and $\overline{W}'$ be Brownian motions of dimension q and d, respectively, and let $\overline{\Theta} = (\overline{\Theta}_i)_{i\geq 1}$ be an i.i.d. sequences of d-dimensional standard normal random variables. $\overline{W}, \overline{W}'$, and $\overline{\Theta}$ are defined on some filtered probability space $(\overline{\Omega}, \overline{\mathcal{F}}, (\overline{\mathcal{F}}_t)_{t\geq 0}, \overline{\mathbb{P}})$ and are assumed to be independent. Let $\mathcal{M}^{\geq 0}$ be the space of all symmetric positive-semidefinite matrices $\varphi \in \mathcal{M}$. We introduce the space $\mathcal{U} = \mathcal{M}' \times \mathcal{M} \times \mathbb{R}^{dq^2} \times \mathbb{R}^d \times \mathcal{M}^{\geq 0}$, and let $\underline{u} = (\alpha, \beta, \gamma, a, \varphi) \in \mathcal{U}$. By $\varphi^{1/2} \in \mathcal{M}$ we denote the matrix root of φ.

Now, for $\kappa = 1, 2$, we define the $2d$-dimensional variables (explicitly writing the components with $l \in \{1, \ldots, d\}$)

$$
\Psi(\underline{u}, g, \kappa)_j^l = \frac{1}{\sqrt{\kappa}} \sum_{m=1}^q \alpha^{lm} \int_{\kappa(j-1)}^{\kappa j} g\big(s/\kappa - (j-1)\big)\, d\overline{W}_s^m,
\tag{4.8}
$$

$$
\begin{aligned}
\Psi(\underline{u}, g, \kappa)_j^{d+l} &= \frac{1}{\kappa} a^l \int_{\kappa(j-1)}^{\kappa j} g\big(s/\kappa - (j-1)\big)\, ds \\
&\quad + \frac{1}{\kappa} \sum_{m,k=1}^q \gamma^{lkm} \int_{\kappa(j-1)}^{\kappa j} g\big(s/\kappa - (j-1)\big) \overline{W}_s^k\, d\overline{W}_s^m \\
&\quad + \frac{1}{\sqrt{\kappa}} \sum_{m=1}^d \beta^{lm} \int_{\kappa(j-1)}^{\kappa j} g\big(s/\kappa - (j-1)\big)\, d\overline{W}_s'^m \\
&\quad + \frac{1}{\kappa} \left(\frac{\psi_1(g)}{\theta^3}\right)^{1/2} \sum_{m=1}^d (\varphi^{1/2})^{lm} \overline{\Theta}_{\kappa j}^m.
\end{aligned}
\tag{4.9}
$$

Some explanations are in order to understand these definitions.

Remark 4.3. To get an intuition for the notation, we remark that the components of $\underline{u} \in \mathcal{U}$ account for the processes in Assumption (A) that will appear in the limit. This means that α is

related to σ_t, β to $\widetilde{\sigma}$, γ to v_t and a to b_t. Finally, φ accounts for the covariance structure of the noise and is associated with Σ. As motivated above, we use different rates in our procedure. Therefore, we also have to define the limit for the two cases $\kappa = 1, 2$.

Remark 4.4. Note that the random-vectors $\Psi(\underline{u}, g, \kappa)_i$ and $\Psi(\underline{u}, g, \kappa)_j$ are uncorrelated whenever $i \neq j$.

Using the notation at (4.7), we define for a weight function g, $\underline{u} = (\alpha, \beta, \gamma, a, \varphi) \in \mathcal{U}$ and $\kappa = 1, 2$ the real-valued random variables

$$\overline{F}_r(\underline{u}, g, \kappa) = F_r\big(\Psi(\underline{u}, g, \kappa)_1, \ldots, \Psi(\underline{u}, g, \kappa)_d\big),$$

and set

$$
\begin{aligned}
\Gamma_r(\underline{u}, g, \kappa) &= \overline{\mathbb{E}}\big[\overline{F}_r(\underline{u}, g, \kappa)\big], \\
\Gamma'_r(\underline{u}, g, \kappa) &= \overline{\mathbb{E}}\big[\overline{F}_r(\underline{u}, g, \kappa)^2\big] - \Gamma_r(\underline{u}, g, \kappa)^2,
\end{aligned}
\tag{4.10}
$$

where the expectation $\overline{\mathbb{E}}$ is taken with respect to the measure $\overline{\mathbb{P}}$, introduced after equation (4.7).

Remark 4.5. Under the special assumption that $\varphi = 0$ (which corresponds to the situation without noise), the sequences $(\Psi(\underline{u}, g, 1)_j)_{j \geq 1}$ and $(\Psi(\underline{u}, g, 2)_j)_{j \geq 1}$ have the same global law which implies also that $\Gamma_r(\underline{u}, g, 1) = \Gamma_r(\underline{u}, g, 2)$ and $\Gamma'_r(\underline{u}, g, 1) = \Gamma'_r(\underline{u}, g, 2)$. This is not the case when $\varphi \neq 0$. Proposition 4.9 will demonstrate under which conditions one can find another weight function h such that $\Gamma_r(\underline{u}, g, 1) = \Gamma_r(\underline{u}, h, 2)$ and $\Gamma'_r(\underline{u}, g, 1) = \Gamma'_r(\underline{u}, h, 2)$ even in the general situation that $\varphi \neq 0$.

Remark 4.6. We have introduced the random variables $\Psi(\underline{u}, g, \kappa)_j$ only for weight functions, implying that g is continuous and piecewise C^1 with a piecewise Lipschitz derivative g'. As a matter of fact, we will often work with a discretized version g^n of g defined as

$$g^n(s) := \sum_{i=1}^{k_n - 1} g\left(\frac{i}{k_n}\right) \mathbf{1}_{\left(\frac{i-1}{k_n}, \frac{i}{k_n}\right]}(s). \tag{4.11}$$

Note that $g^n(0) = g^n(1) = 0$ and that g^n converges to g uniformly on $[0, 1]$. By definition, g^n fails to be a weight function as it is not continuous. Nevertheless, the integrals $\int_0^1 g^n(s) \, ds$, $\int_0^1 g^n(s) \, d\overline{W}_s^m$ and $\int_0^1 g^n(s) \overline{W}_s^k \, d\overline{W}_s^m$ still make sense. This corresponds to the fact that $\psi_l(g^n)$ introduced at (4.6) is well-defined for $l = 2, 3, 4$. Moreover, we have by a Riemann approximation argument that

$$\psi_l(g^n) = \psi_l(g) + O(k_n^{-1}), \qquad l = 2, 3, 4.$$

For $\psi_1(g^n)$, we must approximate the derivative and set

$$\psi_1(g^n) := \frac{1}{k_n} \sum_{i=0}^{k_n - 1} \left(\frac{g(\frac{i+1}{k_n}) - g(\frac{i}{k_n})}{1/k_n}\right)^2 = \psi_1(g) + O(k_n^{-1}), \tag{4.12}$$

where the second identity follows again by a Riemann approximation argument. With this convention, we can extend the notation and write $\Psi(\underline{u}, g^n, \kappa)_j$, $\overline{F}_r(\underline{u}, g^n, \kappa)$, $\Gamma_r(\underline{u}, g^n, \kappa)$ and $\Gamma'_r(\underline{u}, g^n, \kappa)$, respectively.

4.3. Law of large numbers

In this subsection, we present the law of large numbers for the statistic $S(g)_T^{n,\kappa}$. The quantity $\Gamma_r(\underline{u}, g, \kappa)$ defined at (4.10) will essentially determine the limit. First, we demonstrate how the terms $\Gamma_r(\underline{u}, g, \kappa)$ and $\Gamma'_r(\underline{u}, g, \kappa)$ depend on the rank of the argument α. The following lemma has been shown in [8], Lemma 3.1.

Lemma 4.7. *Let $\underline{u} = (\alpha, \beta, \gamma, a, \varphi) \in \mathcal{U}$ with $\beta \in \mathcal{M}_d$ and g be a weight function. Then, if $r \in \{0, \ldots, d\}$ and $\kappa = 1, 2$, we deduce that*

$$\operatorname{rank}(\alpha) = r \implies \Gamma_r(\underline{u}, g, \kappa) > 0, \qquad \Gamma'_r(\underline{u}, g, \kappa) > 0, \tag{4.13}$$

$$\operatorname{rank}(\alpha) < r \implies \Gamma_r(\underline{u}, g, \kappa) = \Gamma'_r(\underline{u}, g, \kappa) = 0. \tag{4.14}$$

The law of large numbers is as follows.

Theorem 4.8. *Assume that conditions (A) and (E) hold. Let $r \in \{0, \ldots, d\}$ and g be a weight function. Then, on Ω_T^r and for $\kappa = 1, 2$, we obtain the convergence*

$$(\kappa u_n)^{r-d} S(g)_T^{n,\kappa} \xrightarrow{\mathbb{P}} S(r, g)_T^\kappa := \int_0^T \Gamma_r(\sigma_s, \widetilde{\sigma}, v_s, b_s, \Sigma, g, \kappa) \, ds > 0. \tag{4.15}$$

In view of Remark 4.1, Theorem 4.8 is not directly applicable since the limit $S(r, g)_T^\kappa$ crucially depends on κ, meaning that generally $S(r, g)_T^1 \neq S(r, g)_T^2$. In particular, the ratio statistics $S(g)_T^{n,2}/S(g)_T^{n,1}$ does not contain any information about the unknown maximal rank R_T. To make use of Theorem 4.8 we need a better understanding of the structure of the functional Γ_r. The following proposition is absolutely crucial for our testing procedure.

Proposition 4.9. (i) *Fix $r \in \{0, \ldots, d\}$, $\underline{u} \in \mathcal{U}$ and $\kappa = 1, 2$. Then there exist C^∞-functions $\tau_{r,\underline{u},\kappa}, \tau'_{r,\underline{u},\kappa} : \mathbb{R}^4 \to \mathbb{R}$ such that*

$$\Gamma_r(\underline{u}, g, \kappa) = \tau_{r,\underline{u},\kappa}(\psi_1(g), \ldots, \psi_4(g)), \qquad \Gamma'_r(\underline{u}, g, \kappa) = \tau'_{r,\underline{u},\kappa}(\psi_1(g), \ldots, \psi_4(g)) \tag{4.16}$$

for any weight function g.

(ii) *Let g and h be weight functions such that $\psi_1(h) = 4\psi_1(g)$ and $\psi_l(h) = \psi_l(g)$ for $l = 2, 3, 4$. Then, for any $r \in \{0, \ldots, d\}$ and any $\underline{u} \in \mathcal{U}$, we obtain that*

$$\Gamma_r(\underline{u}, g, 1) = \Gamma_r(\underline{u}, h, 2), \qquad \Gamma'_r(\underline{u}, g, 1) = \Gamma'_r(\underline{u}, h, 2).$$

Proposition 4.9(i) says that the quantity $\Gamma_r(\underline{u}, g, \kappa)$ does not depend on the entire function g, but only on the quantities $\psi_l(g)$, $l = 1, \ldots, 4$. But most importantly, Proposition 4.9(ii) and Theorem 4.8 imply the convergence

$$\frac{S(h)_T^{n,2}}{S(g)_T^{n,1}} \xrightarrow{\mathbb{P}} 2^{d-r} \qquad \text{on } \Omega_T^r, \tag{4.17}$$

whenever the pair of weight functions g, h satisfies the conditions of Proposition 4.9(ii). This opens the door to hypothesis testing. We now give an example of a pair of weight function g, h which fulfills the conditions of Proposition 4.9(ii).

Example 4.10. We define the two auxiliary weight functions $\widetilde{g}(x) := \max(0, \min(x, 1 - x))$ and $\widetilde{h}(x) := \max(0, \min(ax, b(1 - x)))$ with $a = \frac{2}{2-\sqrt{3}}$ and $b = \frac{2}{2+\sqrt{3}}$. Then, a pair of weight functions satisfying the conditions of Proposition 4.9(ii) is given by $g_c(x) := \widetilde{g}(cx)$ and $h_c(x) := \widetilde{h}(cx - c + 1)$ where $c = \frac{8+\sqrt{3}}{8}$ (see Figure 1). Indeed, for $c \geq 1$, we obtain that

$$\psi_1(h_c) = 4\psi_1(g_c) = 4c,$$

$$\psi_2(h_c) = \psi_2(g_c) = \frac{1}{12c},$$

$$\psi_3(h_c) = \psi_3(g_c) = \frac{1}{4c},$$

$$\psi_4(h_c) = \frac{8c - 4 - \sqrt{3}}{96c^2}, \qquad \psi_4(g_c) = \frac{1}{24c^2},$$

and for $c = \frac{8+\sqrt{3}}{8}$, we have $\frac{8c-4-\sqrt{3}}{96c^2} = \frac{1}{24c^2} = \frac{536-128\sqrt{3}}{11\,163}$.

Remark 4.11. From a statistical point of view and regarding the definition of the pre-averaged increments in (3.4), we see that it is certainly not ideal to choose weight functions which are
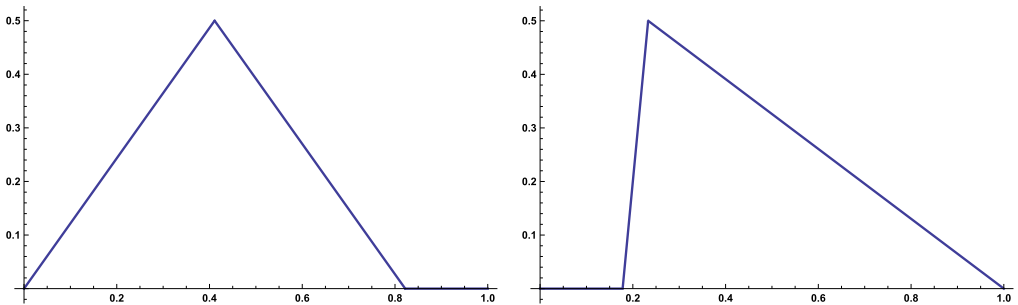

Figure 1. A particular pair of weight functions g_c (left), h_c (right) from Example 4.10 with $c = \frac{8+\sqrt{3}}{8}$ satisfying the conditions of Proposition 4.9(ii).

locally constant. Nevertheless, Example 4.10 is an attempt to reduce the parts where the weight functions are constant while still sticking to a rather simple "triangular" form.

Remark 4.12. Here we demonstrate why the classical choice of the local window $k_n \sqrt{\Delta_n} = \theta + o(\Delta_n^{1/4})$ does not work in the context of rank estimation. We assume for the moment that the noise covariance matrix Σ has full rank and let $A \in \mathcal{M}$ be a matrix with unknown rank r, which should represent the volatility matrix. Furthermore, let $B \in \mathcal{M}_d$ be a perturbation matrix and let $\lambda > 0$ represent the perturbation rate $\sqrt{u_n}$ as in Section 3.1. With the window size $k_n = O(\Delta_n^{-1/2})$ the noise and the diffusive parts become balanced and Proposition 4.9(i) suggests that we may estimate objects of the form

$$\Phi(\lambda, g) = \det\big((x_g A + y_g \Sigma) + \lambda B\big)^2 \qquad \text{with } x_g = \psi_2(g),\ y_g = \psi_1(g).$$

We will now expand the quantity $\Phi(\lambda, g)$ using again the multilinearity of the determinant. For this purpose, we denote with $\mathcal{P}_3$ the set of all multi-integers $\mathbf{p} = (p_1, p_2, p_3)$ with $p_j \geq 0$ and $p_1 + p_2 + p_3 = d$, and $\mathcal{I}_\mathbf{3}$ the set of all partitions $\mathbf{I} = (I_1, I_2, I_3)$ of $\{0, \ldots, d\}$ such that I_j contains exactly p_j points. For $\mathbf{p} \in \mathcal{P}_3$, $\mathbf{I} \in \mathcal{I}_\mathbf{3}$ and $A_1, A_2, A_3 \in \mathcal{M}$, we call $G^{\mathbf{I}}_{A_1, A_2, A_3}$ the matrix in $\mathcal{M}$ whose ith column is the ith column of A_j when $j \in I_j$. We then immediately deduce the identity

$$\Phi(\lambda, g) = \Bigg(\sum_{\substack{\mathbf{p}=(p_1,p_2,p_3)\in\mathcal{P}_3,\ \mathbf{I}\in\mathcal{I}_p \\ p_2\neq 0,\, p_3\leq d-r}} \lambda^{p_3} \det\big(G^{\mathbf{I}}_{x_g A, y_g \Sigma, B}\big) + \lambda^{d-r} \sum_{\mathbf{I}\in\mathcal{I}_{(r,0,d-r)}} \det\big(G^{\mathbf{I}}_{x_g A, y_g \Sigma, B}\big)$$

$$+ \sum_{\substack{\mathbf{p}=(p_1,p_2,p_3)\in\mathcal{P}_3,\ \mathbf{I}\in\mathcal{I}_p \\ p_3 > d-r}} \lambda^{p_3} \det\big(G^{\mathbf{I}}_{x_g A, y_g \Sigma, B}\big) \Bigg)^2$$

$$=: \big(\Phi_1(\lambda, g) + \Phi_2(\lambda, g) + \Phi_3(\lambda, g)\big)^2.$$

When $r < d$ and $\lambda \downarrow 0$, the first term $\Phi_1(\lambda, g)$, which in particular contains $y_g^d \det(\Sigma) \neq 0$, dominates $\Phi_2(\lambda, g)$ which has the order λ^{d-r}, while $\Phi_3(\lambda, g)$ is negligible with respect to $\Phi_2(\lambda, g)$. However, recalling the theory from Section 3.1 it is exactly the term $\Phi_2(\lambda, g)$, which contains the information about the unknown rank r. In a purely deterministic framework, we may hope to correct for the dominating bias caused by the term $\Phi_1(\lambda, g)$ by taking a suitable linear combination of a particular family $(\Phi(\lambda, g_l))_l$. Unfortunately, in a statistical setting, the objects contained in $\Phi_1(\lambda, g)$ can be only estimated with a precision rate $u_n^{-1/2}$ (cf. Theorem 4.13 below), which would still give a huge bias that dominates $\Phi_2(\lambda, g)$ whenever $r < d$. This is precisely the reason why we see little hope in using the optimal window size $k_n = O(\Delta_n^{-1/2})$.

4.4. Central limit theorem and testing procedure

In order to provide a formal testing procedure associated with the convergence in probability at (4.17), we need to show a joint stable central limit theorem for the statistics $(S(g)_T^{n,1}, S(h)_T^{n,2})$. We say that a sequence of random variables H_n converges stably in law to H ($H_n \xrightarrow{d_{\text{st}}} H$), where

H is defined on an extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ of the original probability space $(\Omega, \mathcal{F}, \mathbb{P})$, if and only if

$$\lim_{n \to \infty} \mathbb{E}\big[\phi(H_n)Z\big] = \widetilde{\mathbb{E}}\big[\phi(H)Z\big]$$

for any bounded and continuous function ϕ and any bounded $\mathcal{F}$-measurable random variable Z. We refer to [1,10] or [12] for a detailed study of stable convergence. Note that stable convergence is a stronger mode of convergence than weak convergence, but it is weaker than convergence in probability.

Now, let g and h be two weight functions satisfying the conditions of Proposition 4.9(ii). We define the statistic $U(r, g, h)^n_T = (U(r, g, h)^{n,1}_T, U(r, g, h)^{n,2}_T)$ via

$$U(r, g, h)^n_T = \frac{1}{\sqrt{u_n}}\big(u_n^{r-d} S(g)^{n,1}_T - S(r, g)^1_T, (2u_n)^{r-d} S(h)^{n,2}_T - S(r, h)^2_T\big). \tag{4.18}$$

The following theorem is one of the most important results of the paper.

Theorem 4.13. *Assume that conditions* (A) *and* (E) *are satisfied, the weight functions g, h fulfill the assumptions of Proposition* 4.9(ii) *and $R_T(\omega) \leq r$ for some $r \in \{0, \ldots d\}$. Then we obtain the stable convergence*

$$U(r, g, h)^n_T \xrightarrow{d_{\mathrm{st}}} \mathcal{MN}\big(0, V(r, g, h)_T\big), \tag{4.19}$$

where

$$
\begin{aligned}
&V(r, g, h)_T \\
&= \operatorname{diag}\left(3d \int_0^T \Gamma'_r(\sigma_s, \widetilde{\sigma}, v_s, b_s, \Sigma, g, 1)\, ds, \; 3d \int_0^T \Gamma'_r(\sigma_s, \widetilde{\sigma}, v_s, b_s, \Sigma, h, 2)\, ds\right)
\end{aligned}
\tag{4.20}
$$

is a diagonal matrix. $\mathcal{MN}(0, V(r, g, h)_T)$ denotes the two dimensional mixed normal distribution with $\mathcal{G}$-conditional mean 0 and $\mathcal{G}$-conditional covariance matrix $V(r, g, h)_T$.

Note that the rate of convergence $u_n^{-1/2}$ corresponds to $\Delta_n^{-1/6}$ for our choice of the window size k_n at (3.5). We remark that due to Proposition 4.9(ii), we know that $S(r, g)^1_T = S(r, h)^2_T$ such that the same centering term appears in both components on the right-hand side of (4.18). Again thanks to Proposition 4.9(ii) we see that the two diagonal elements of $V(r, g, h)_T$ coincide. In order to obtain a feasible version of the stable convergence in (4.19), we need to construct a consistent estimator of the $\mathcal{G}$-conditional covariance matrix $V(r, g, h)_T$. To this end, we define the following estimators for the "second moments":

$$V(g, h)^{n,11}_T = 9d^2 u_n \sum_{i=0}^{[T/3du_n]-1} f^2\big((\overline{Z}(g)^{n,1}_{(3id+(j-1))k_n}/\sqrt{u_n})_{j=1,\ldots,d}\big), \tag{4.21}$$

$$V(g, h)^{n,22}_T = 9d^2 u_n \sum_{i=0}^{[T/3du_n]-1} f^2\big((\overline{Z}(h)^{n,2}_{(3id+d+2(j-1))k_n}/\sqrt{2u_n})_{j=1,\ldots,d}\big), \tag{4.22}$$

$$V(g,h)_T^{n,12} = 9d^2 u_n \sum_{i=0}^{[T/3du_n]-1} f\left(\left(\overline{Z}(g)_{(3id+(j-1))k_n}^{n,1}/\sqrt{u_n}\right)_{j=1,\ldots,d}\right) \tag{4.23}$$

$$\times f\left(\left(\overline{Z}(h)_{(3id+d+2(j-1))k_n}^{n,2}/\sqrt{2u_n}\right)_{j=1,\ldots,d}\right),$$

where f is given at (4.4). Following the intuition from (4.17) we define an estimator $\widehat{R}(g,h)_T^n$ via

$$\widehat{R}(g,h)_T^n := d - \frac{\log(S(h)_T^{n,2}/S(g)_T^{n,1})}{\log 2}. \tag{4.24}$$

Now, we obtain the following proposition.

Proposition 4.14. *Assume that conditions* (A) *and* (E) *are satisfied and the weight functions* g, h *fulfill the assumptions of Proposition* 4.9(ii).

(i) *Let* $r \in \{0, \ldots, d\}$*. Then, on* $\Omega_T^{\leq r}$:

$$\left(u_n^2\right)^{r-d} V(g,h)_T^{n,11} \tag{4.25}$$

$$\xrightarrow{\mathbb{P}} 3d \int_0^T \Gamma_r'(\sigma_s, \widetilde{\sigma}, v_s, b_s, \Sigma, g, 1) + \Gamma_r^2(\sigma_s, \widetilde{\sigma}, v_s, b_s, \Sigma, g, 1) \, ds,$$

$$\left(4u_n^2\right)^{r-d} V(g,h)_T^{n,22} \tag{4.26}$$

$$\xrightarrow{\mathbb{P}} 3d \int_0^T \Gamma_r'(\sigma_s, \widetilde{\sigma}, v_s, b_s, \Sigma, h, 2) + \Gamma_r^2(\sigma_s, \widetilde{\sigma}, v_s, b_s, \Sigma, h, 2) \, ds,$$

$$\left(2u_n^2\right)^{r-d} V(g,h)_T^{n,12} \tag{4.27}$$

$$\xrightarrow{\mathbb{P}} 3d \int_0^T \Gamma_r(\sigma_s, \widetilde{\sigma}, v_s, b_s, \Sigma, g, 1) \Gamma_r(\sigma_s, \widetilde{\sigma}, v_s, b_s, \Sigma, h, 2) \, ds.$$

(ii) *We have the* (stable) *central limit theorem*

$$\frac{1}{\sqrt{u_n}} \frac{\widehat{R}(g,h)_T^n - R_T}{\sqrt{V(n,T,g,h)}} \xrightarrow{d_{st}} \Phi \sim \mathcal{N}(0,1), \tag{4.28}$$

where Φ *is defined on an extension* $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ *of the original probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and is independent of the* σ*-algebra* $\mathcal{G}$*. The random variable* $V(n,T,g,h)$ *is defined via*

$$V(n,T,g,h) := \frac{V(g,h)_T^{n,11} + 4^{\widehat{R}(g,h)_T^n - d} V(g,h)_T^{n,22} - 2^{1+\widehat{R}(g,h)_T^n - d} V(g,h)_T^{n,12}}{(S(g)_T^{n,1} \log 2)^2}. \tag{4.29}$$

We remark that Proposition 4.14(ii) follows directly from Theorem 4.13, Proposition 4.14(i) and the delta method for stable convergence. For this, it is essential to realize that, even though

the estimator $V(n, T, g, h)$ for the conditional variance is not $\mathcal{G}$-measurable, it converges to a $\mathcal{G}$-measurable limit due to Proposition 4.14(i) and Theorem 4.8.

Notice also that due to Proposition 4.9(ii) the right-hand side of (4.25) and (4.26) coincide and, moreover, that the right-hand side of (4.27) can be written as

$$3d \int_0^T \Gamma_r^2(\sigma_s, \widetilde{\sigma}, v_s, b_s, \Sigma, g, 1) \, ds.$$

Remark 4.15. Instead of using the estimators for the second moments given in (4.21) till (4.23), we could also use a more direct approach and consider

$$V'(g)_T^n := 3d^2 u_n \sum_{i=0}^{[T/2du_n]-1} \left\{ f\big((\overline{Z}(g)_{(2id+j-1)k_n}^{n,1} / \sqrt{u_n})_{j=1,\dots,d} \big) \right.$$

$$\left. - f\big((\overline{Z}(g)_{(2id+d+j-1)k_n}^{n,1} / \sqrt{u_n})_{j=1,\dots,d} \big) \right\}^2.$$

Notice that similar to Proposition 4.14(i), we have that

$$(u_n^2)^{r-d} V'(g)_T^n \xrightarrow{\mathbb{P}} 3d \int_0^T \Gamma_r'(\sigma_s, \widetilde{\sigma}, v_s, b_s, \Sigma, g, 1) \, ds.$$

Then (4.28) also holds upon replacing $V(n, T, g, h)$ defined at (4.29) by

$$V'(n, T, g) := \frac{2V'(g)_T^n}{(S(g)_T^{n,1} \log 2)^2}.$$

The feasible central limit theorem at (4.28) opens the door to hypothesis testing. Let us define the rejection regions via

$$\mathcal{C}_\alpha^{n,=r} := \left\{ \omega : \left| \widehat{R}(g, h)_T^n - r \right| > z_{1-\alpha/2} \sqrt{u_n V(n, T, g, h)} \right\}, \tag{4.30}$$

$$\mathcal{C}_\alpha^{n,\leq r} := \left\{ \omega : \widehat{R}(g, h)_T^n - r > z_{1-\alpha} \sqrt{u_n V(n, T, g, h)} \right\}, \tag{4.31}$$

where z_α denotes the α-quantile of the standard normal distribution. Obviously, the rejection region $\mathcal{C}_\alpha^{n,=r}$ corresponds to $H_0 : R_T = r$ vs. $H_1 : R_T \neq r$, while $\mathcal{C}_\alpha^{n,\leq r}$ corresponds to $H_0 : R_T \leq r$ vs. $H_1 : R_T > r$. The asymptotic level and consistency of the tests are demonstrated in the following corollary.

Corollary 4.16. *Assume that conditions* (A) *and* (E) *are satisfied and the weight functions* g, h *fulfill the assumptions of Proposition 4.9*(ii).

(i) *The test defined through* (4.30) *has asymptotic level* α *in the sense that*

$$A \subset \Omega_T^r, \qquad \mathbb{P}(A) > 0 \quad \Longrightarrow \quad \mathbb{P}\big(\mathcal{C}_\alpha^{n,=r} | A\big) \to \alpha. \tag{4.32}$$

Furthermore, the test is consistent, that is,

$$\mathbb{P}\big(\mathcal{C}_\alpha^{n,=r} \cap \big(\Omega_T^r\big)^c\big) \to \mathbb{P}\big(\big(\Omega_T^r\big)^c\big). \tag{4.33}$$

(ii) *The test defined through* (4.31) *has asymptotic level at most α in the sense that*

$$A \subset \Omega_T^{\leq r}, \qquad \mathbb{P}(A) > 0 \implies \limsup_{n\to\infty} \mathbb{P}\big(\mathcal{C}_\alpha^{n,\leq r}|A\big) \leq \alpha. \tag{4.34}$$

Furthermore, the test is consistent, that is,

$$\mathbb{P}\big(\mathcal{C}_\alpha^{n,\leq r} \cap \big(\Omega_T^{\leq r}\big)^c\big) \to \mathbb{P}\big(\big(\Omega_T^{\leq r}\big)^c\big). \tag{4.35}$$

4.5. Working with overlapping increments

Recall that our test statistics $S(g)_T^{n,\kappa}$, $\kappa = 1, 2$, defined at (4.3) work with "maximally non-overlapping" increments in the following sense: (i) Each single summand uses non-overlapping increments, (ii) for each $\kappa = 1, 2$ the summands in $S(g)_T^{n,\kappa}$ contain distinct increments, and finally (iii) also the statistics $S(g)_T^{n,1}$ and $S(g)_T^{n,2}$ are composed of distinct increments. The particular reasons for using non-overlapping increments on these three levels are also different. Point (i) is necessary to obtain the asymptotic decomposition explained in Remark 4.2. On the other hand, (ii) ensures the stable central limit theorem at (4.19) for each single component and (iii) implies the diagonal form of the conditional covariance matrix at (4.20), so in particular, the conditional covariance is not degenerate.

In order to reduce the asymptotic variance of the rank-estimator at (4.38) and to improve its finite sample performance, it is desirable to use overlapping increments at least at some of the different levels. Whereas (i) is crucial for the whole approximation idea, (ii) and (iii) are less compulsory. Concerning point (iii) one can allow for common increments in the two main statistics as long as the conditional covariance appearing in Theorem 4.13 is not degenerate; see, for example, [8]. A natural ansatz of tailoring new statistics using overlapping increments on the level of (ii) and (iii) is the following. We define

$$\widetilde{S}(g)_T^{n,\kappa} = \Delta_n \sum_{i=0}^{[T/\Delta_n]-\kappa dk_n-1} f\big(\big(\overline{Z}(g)_{i+\kappa(j-1)k_n}^{n,\kappa}/\sqrt{\kappa u_n}\big)_{j=1,\dots,d}\big) \tag{4.36}$$

for $\kappa = 1, 2$, a weight function g and the function f defined at (4.4). Proving a law of large numbers is a quite straightforward corollary to Theorem 4.8. The idea is to split up the sum in (4.36) into the sum of $3dk_n$ partial sums, each of them having summands with distinct increments and being of the form at (4.3) with different starting points.

Corollary 4.17. *Assume that conditions* (A) *and* (E) *hold. Let $r \in \{0, \dots, d\}$ and g be a weight function. Then, on Ω_T^r and for $\kappa = 1, 2$, we obtain the convergence*

$$(\kappa u_n)^{r-d}\widetilde{S}(g)_T^{n,\kappa} \xrightarrow{\mathbb{P}} S(r, g)_T^\kappa, \tag{4.37}$$

where $S(r, g)_T^\kappa$ is defined at (4.15).

Then, using a pair of weight functions g, h, satisfying the conditions of Proposition 4.9(ii) we obtain an analog version of the convergence at (4.17), such that

$$\widetilde{R}(g, h)_T^n := d - \frac{\log(\widetilde{S}(h)_T^{n,2}/\widetilde{S}(g)_T^{n,1})}{\log 2} \tag{4.38}$$

is a consistent estimator for the maximal rank.

In contrast to the straightforward derivation of the law of large numbers, it is far from being trivial to derive a stable central limit theorem corresponding to Theorem 4.13. In principle, one could show a one dimensional central limit theorem for each of the two components (of course with a different conditional variance) with the so called "small blocks–big blocks"-technique which has become a standard method for proving central limit theorems for dependent data; see [7,11]. However, for the joint convergence, the remaining major problem is to show that the resulting conditional covariance matrix is not degenerate. It is not clear how one could show this, and also simulation studies give evidence that the sample covariance matrix is also 'almost' degenerate. Consequently, a central limit theorem like the one at (4.28), and with that hypothesis testing, seems to be much more complicated.

5. Simulation study and real data example

5.1. Simulations

In this subsection, we want to examine how well the testing procedure for the maximal rank performs in finite samples. The main focus lies on considering the convergence results in (4.28), (4.32) and (4.33). Complementing these results, we examine how well the estimator $\widehat{R}(g, h)_T^n$ works to estimate the maximal rank R_T and we compare it with the estimator for overlapping increments $\widetilde{R}(g, h)_T^n$ (using the law of large numbers which is implicitly given by (4.28)). To this end, we also consider the truncated and integer-valued modifications of the estimators defined as

$$\widehat{R}^{\mathrm{int}}(g, h)_T^n := \begin{cases} 0, & \widehat{R}(g, h)_T^n < 1/2, \\ r, & \widehat{R}(g, h)_T^n \in [r - 1/2, r + 1/2), r \in \{1, \ldots, d - 1\}, \\ d, & \widehat{R}(g, h)_T^n \geq d - 1/2, \end{cases} \tag{5.1}$$

$$\widetilde{R}^{\mathrm{int}}(g, h)_T^n := \begin{cases} 0, & \widetilde{R}(g, h)_T^n < 1/2, \\ r, & \widetilde{R}(g, h)_T^n \in [r - 1/2, r + 1/2), r \in \{1, \ldots, d - 1\}, \\ d, & \widetilde{R}(g, h)_T^n \geq d - 1/2. \end{cases} \tag{5.2}$$

We emphasize that due to the rate of convergence of $\Delta_n^{-1/6}$ we expect a worse performance in finite samples in comparison to the simulation study in [8] (there, the rate of convergence is $\Delta_n^{-1/2}$).

5.1.1. *Models and parameters*

For all simulations, we set $T = 1$. Assuming a trading day consists of 23 400 seconds, we use the sampling frequencies $\Delta_n^{-1} = 23\,400 \times 10^k$ for $k = 0, 1, 2$. We remark that nowadays liquid assets can be observed even at the highest frequency $\Delta_n^{-1} = 23\,400 \times 10^2$. We work with a fix dimension of $d = 8$, which makes the simulation results comparable with the real data example of Section 5.2. Following the simulation study in [8], we set $\widetilde{\sigma} = I_d$ and due to [7] we use $\theta = 1/3$ for the pre-averaging procedure. Additionally, we set $k_n := [\theta / \Delta_n^{2/3}]$ such that (3.5) is satisfied. We use the weight functions g, h explicitly constructed in Example 4.10. For the noise part, we always assume a covariance structure of $\Sigma = 0.0005 I_d$ like in [7]. We perform 100 repetitions for X to uncover the finite sample properties. To reduce the influence of the random perturbation, we simulate W' 10 times for each repetition. The following quantities are reported:

- k: the implicitly given sampling frequency defined as $\Delta_n^{-1} = 23\,400 \times 10^k$;
- the first two moments of the test statistic $\dfrac{\widehat{R}(g,h)_1^n - R_1}{\sqrt{u_n V(n,1,g,h)}}$ defined at (4.28) to check for the normal approximation;
- Ω_1^r: the proportion of rejection for the possible null hypotheses Ω_1^r with $r \in \{0, \ldots, d\}$ defined at (2.3) at level $\alpha = 0.05$.

In order to compare the rank estimators for the overlapping and the non-overlapping approach, we first take the empirical means of the estimators $\widehat{R}(g,h)_1^n$ and $\widetilde{R}(g,h)_1^n$ over 10 different simulations of W', and we take the respective empirical medians of $\widehat{R}^{\text{int}}(g,h)_1^n$ and $\widetilde{R}^{\text{int}}(g,h)_1^n$. For the sake of brevity, we denote these quantities again with the same notation. We report the following quantities:

- mean: the sample mean over the 100 repetitions of $\widehat{R}(g,h)_1^n$ and $\widetilde{R}(g,h)_1^n$;
- variance: the sample variance over the 100 repetitions of $\widehat{R}(g,h)_1^n$ and $\widetilde{R}(g,h)_1^n$;
- ratio: the proportions of the event that the estimator $\widehat{R}^{\text{int}}(g,h)_1^n$ and $\widetilde{R}^{\text{int}}(g,h)_1^n$ coincides with R_1.

For the non-overlapping approach, we use again the three frequencies $\Delta_n^{-1} = 23\,400 \times 10^k$, $k = 0, 1, 2$. Due to computational restrictions, we conduct the simulation for the overlapping approach just for the lowest frequency $\Delta_n^{-1} = 23\,400$. We consider the following models for the semimartingale X defined at (2.1).

Model 1 We have a vanishing drift $b_t = 0$ and a constant volatility $\sigma_t = I_d$ of full rank.

Model 2 We observe pure noise, so $b_t = 0$ and $\sigma_t = 0$.

Model 3 We have a constant drift $b_t = 0.5$ and a constant volatility structure with rank $R_1 = 6$ which is given by

$$\sigma_t = \begin{pmatrix} 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Model 4 We have a vanishing drift $b_t = 0$ and a time varying volatility process given by $\sigma_t = \max(\sin(2\pi t), 0) I_d$, such that $R_1 = 8$.

5.1.2. *Summary*

According to our theoretical results, the empirical counterparts of the first two moments, level and power reported in Table 1 seem to converge to their theoretical analogues as the frequency increases. However, the speed of convergence depends on the particular model, and generally, even at the highest frequency $\Delta_n^{-1} = 23\,400 \times 10^2$, the empirical results are still quite far away from the theoretical asymptotic results, which is due to the rather slow convergence rate of $\Delta_n^{-1/6}$. More specifically, the numerical results at the lowest frequency $\Delta_n^{-1} = 23\,400$ are less reliable for all four models.

In particular, for models 1, 2 and 3, we observe reasonable results for the convergence of the empirical moments. Only at the lowest frequency $\Delta_n^{-1} = 23\,400$, model 3 performs considerably worse than models 1 and 2. An explanation is that there are several small order terms in the expansion of the main statistic, which seem to influence the finite sample performance at relatively low frequencies. Recall that there is a drift term in model 3, whereas models 1 and 2 have vanishing drift and constant volatility, such that these lower order terms do not appear. Also the relatively bad performance of the empirical moments in model 4 is in line with that explanation. It is the only model with time varying volatility – and the range of the rank is quite extreme: one has full rank in $[0, 1/2)$ and pure noise in $[1/2, 0]$. Astonishingly, in model 4 one obtains the best results for the empirical moments at the lowest frequency.

The approximation of power again depends on the complexity of the time-varying coefficients of the model. Quite intuitively, we observe a better power performance for alternative hypotheses, which are more distant to the true one. In models 1 and 3, at the lowest frequency, the underestimation of the true maximal rank (7.29 and 5.87; see Table 2) induces a bias in the corresponding

Table 1. The results for the statistic $\widehat{R}(g, h)_1^n$ in four models are summarized according to their order

k	1st mt	2nd mt	Ω_1^0	Ω_1^1	Ω_1^2	Ω_1^3	Ω_1^4	Ω_1^5	Ω_1^6	Ω_1^7	Ω_1^8
0	-1.84	715.27	0.933	0.903	0.843	0.781	0.693	0.609	0.553	0.527	0.535
1	0.19	15.71	0.990	0.962	0.908	0.836	0.734	0.637	0.539	0.471	0.444
2	-0.22	3.99	1.000	0.981	0.954	0.922	0.800	0.632	0.435	0.313	0.285
0	2.30	6472.00	0.567	0.597	0.622	0.676	0.743	0.814	0.866	0.898	0.935
1	-0.11	13.51	0.360	0.400	0.491	0.602	0.738	0.841	0.917	0.952	0.980
2	0.11	3.72	0.253	0.281	0.448	0.647	0.791	0.916	0.969	0.988	0.996
0	-23.09	2.29×10^5	0.830	0.764	0.714	0.657	0.607	0.565	0.580	0.590	0.617
1	-0.15	17.40	0.903	0.825	0.721	0.594	0.494	0.433	0.400	0.451	0.534
2	-0.02	3.84	0.975	0.927	0.811	0.684	0.499	0.346	0.278	0.330	0.515
0	0.54	10.44	0.818	0.780	0.723	0.679	0.631	0.591	0.560	0.540	0.521
1	17.20	2.76×10^5	0.905	0.883	0.856	0.825	0.772	0.717	0.681	0.637	0.606
2	-1.79	1281.66	0.933	0.915	0.889	0.824	0.725	0.642	0.553	0.505	0.517

Table 2. The results for the rank estimators $\widehat{R}(g,h)_T^n$ and $\widetilde{R}(g,h)_T^n$ of the overlapping and non-overlapping approach for the four models are summarized according to their order

Overlapping	k	Mean	Variance	Ratio
No	0	7.29	8.86	0.48
No	1	8.12	7.01	0.64
No	2	7.80	3.40	0.61
Yes	0	7.88	0.55	0.72
No	0	0.02	1.41	0.55
No	1	−0.04	0.70	0.59
No	2	0.11	0.39	0.58
Yes	0	0.08	0.04	0.93
No	0	5.87	5.33	0.15
No	1	5.82	2.81	0.27
No	2	6.00	1.34	0.29
Yes	0	5.96	0.24	0.67
No	0	9.10	27.82	0.60
No	1	8.63	20.77	0.68
No	2	7.99	10.62	0.56
Yes	0	8.93	1.49	0.90

tests: in model 1, the proportion of rejection for the null hypothesis Ω_1^7 is 0.527 which is smaller than 0.535, the proportion of rejection for the true hypothesis Ω_1^8 (also compare the values for Ω_1^5 and Ω_1^6 for model 3). In model 4, we have the same effect at the highest frequency where the proportion of rejection is 0.505 for Ω_1^7 and 0.517 for Ω_1^8. This cannot merely be explained by the relatively small bias of -0.01 of $\widehat{R}(g,h)_1^n$; see Table 2. However, the rounded but not truncated version of $\widehat{R}(g,h)_1^n$ takes 17 times the value 7, but only 11 times the value 8 in our simulated sample, which can serve as an explanation of the biased test.

If we first focus on the results for the non-overlapping statistics in Table 2 we can see that the bias and the variance of the corresponding rank-estimator seem to become smaller as $\Delta_n \to 0$. Especially for the variance we can see this in all models. In line with this convergence, one would expect that the ratio of events that $\widehat{R}^{\text{int}}(g,h)_1^n$ coincides with the true maximal rank increases as $\Delta_n \to 0$. However, we can assert this increase only for model 3. This is the only model where the true maximal rank, which is 6, is an "inner point" of the possible values $\{0,\ldots,8\}$. In the other models, truncation effects in the definition at (5.1) take place. For example, in model 4, at the frequencies $\Delta_n^{-1} = 23\,400 \times 10^k$, $k = 0, 1$, there is actually an upwards bias ($+1.10$ and $+0.63$), whereas for the highest frequency, there is a negligible bias (-0.01).

Comparing the results for the non-overlapping approach with the overlapping approach outlined in Section 4.5 yields qualitatively expectable, but quantitatively impressive results. First of all, the bias of the estimator using the overlapping approach is of comparable size as the one for the non-overlapping approach at the same frequency $\Delta_n^{-1} = 23\,400$. On the other hand, the sample variances of the two different approaches is dramatically different. Comparing the esti-

mators at the same frequency, the overlapping approach reduces the variance by a factor of 16 (model 1) up to 35 (model 2). Even for the unfair comparison at different frequencies for the different approaches ($\Delta_n^{-1} = 23\,400 \times 10^2$ for the non-overlapping approach and $\Delta_n^{-1} = 23\,400$ for the overlapping approach), the variance in the overlapping approach is still smaller than the one in the non-overlapping approach by a factor of 5 (model 3) up to 10 (model 2). As a result of the variance reduction, the proportion of the event that the estimators defined at (5.1) and (5.2) coincide with the true maximal rank increases in all models when switching to the overlapping approach. In our simulation study, this proportion of correct estimation in the overlapping approach is considerably higher than that proportion for the non-overlapping approach uniformly over all frequencies. Only the comparison on model 4 seems to be slightly unfair due to the positive bias of $\widetilde{R}(g, h)_1^n$ and the truncation effects when calculating $\widetilde{R}^{\text{int}}(g, h)_1^n$.

We remark that we have conducted several pre-simulations for lower dimensions. In theses cases, the speed of convergence both of the empirical moments and of the empirical powers is the higher the lower the dimension is. This can be explained by the fact that the true rate of convergence is $[T/3du_n]^{1/2}$ rather than $\Delta_n^{-1/6}$, which decreases when d is growing. On the other hand, for constant volatility models the speed of convergence seems to increase as the time horizon increases, which is also in line with the above statement. For more complicated models, the relation is not so easy. The explanation is that one should rather replace T with the Lebesgue measure of the set $\{t \in [0, T]: r_t = R_T\}$ in the rate of convergence given above.

5.2. Real data example

5.2.1. *Data description*

In order to test our theoretical results on real data, we consider a homogeneous market of eight stocks from big American banks: American International, Allstate, Bank of America, Citigroup, Goldman Sachs, J. P. Morgan, MetLife, and Prudential Financial. The sample period is 2006–2009, both years included, which results in a total number of 1007 trading days. We performed a pre-cleaning of the data to exclude misprints. We have applied the previous-tick method to obtain 1-second equidistant observations in each trading day, that is 23 401 prices in total for each day.

5.2.2. *Data cleaning*

As our asymptotic theory is not robust to jumps in the price process X, we need to eliminate possible jumps first. To remove possible intraday jumps, we apply the following algorithm:

1. For each component Y^m of the observed price process, we calculate the mean daily squared volatility, denoted by $(\widehat{\sigma}_i^m)^2$ for day i. To this end, we use the jump-robust pre-averaged estimator defined at [5], equation (4.35). As parameters we use the standard weight function $g(x) = \min(x, 1 - x)$, $\theta = 1/3$ and $k_n' = [\theta/\Delta_n^{1/2}]$, resulting in $\psi_2 = \psi_2(g) = 1/12$; see (4.6).
2. For each day i and each component Y^m, we define a threshold

$$\text{Threshold}_i^m := 3\widehat{\sigma}_i^m \sqrt{\psi_2 k_n' \Delta_n}.$$

3. We delete a pre-averaged increment $\overline{Y}_j^n$ within day i of the multi-dimensional price process whenever there is at least one component $m \in \{1, \dots, d\}$ such that

$$\left| \overline{Y^m}_j^n \right| > \text{Threshold}_i^m.$$

Since we have to cancel the increments in all components whenever there occurs a jump in a single component, the algorithm removes quite a number of increments. The number of remaining increments is about 71.9% of the original number before the data cleaning.

5.2.3. *Results*

In line with the simulation study, we set $\widetilde{\sigma} = I_d$, $\theta = 1/3$ and $k_n = [\theta/\Delta_n^{2/3}]$ and use the weight functions g, h which are constructed in Example 4.10. Over each time window, we use 10 independent realizations of the random perturbation W'. Again, we report $\widehat{R}(g, h)_T^n$ ($\widetilde{R}(g, h)_T^n$, respectively) as the mean over that sample of the respective estimators, and $\widehat{R}^{\text{int}}(g, h)_T^n$ ($\widetilde{R}^{\text{int}}(g, h)_T^n$, respectively) as the respective median.

   We report the results for the non-overlapping increments approach for a time window of one day in Figure 2 and over a rolling time window of 10 days in Figure 3. One can find the results for the overlapping increments approach for a time window of one day in Figure 4. For a first analysis of the results, we ignore the time varying structure of the data and compute the sample mean and variance of the estimators $\widehat{R}(g, h)_1^n$, $\widehat{R}(g, h)_{10}^n$, and $\widetilde{R}(g, h)_1^n$; see Table 3.

   For the non-overlapping situation with a one-day time window, we clearly see a positive bias. The sample variance reduces both when switching to a larger time horizon and when switching
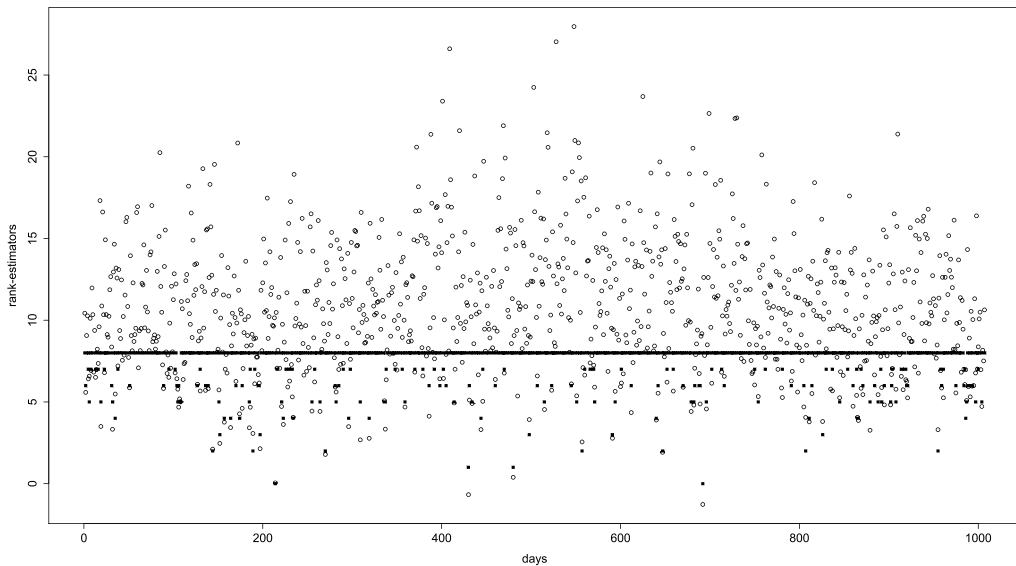

Figure 2. Estimators $\widehat{R}(g, h)_1^n$ (circles) and $\widehat{R}^{\text{int}}(g, h)_1^n$ (squares) over a one-day time window.
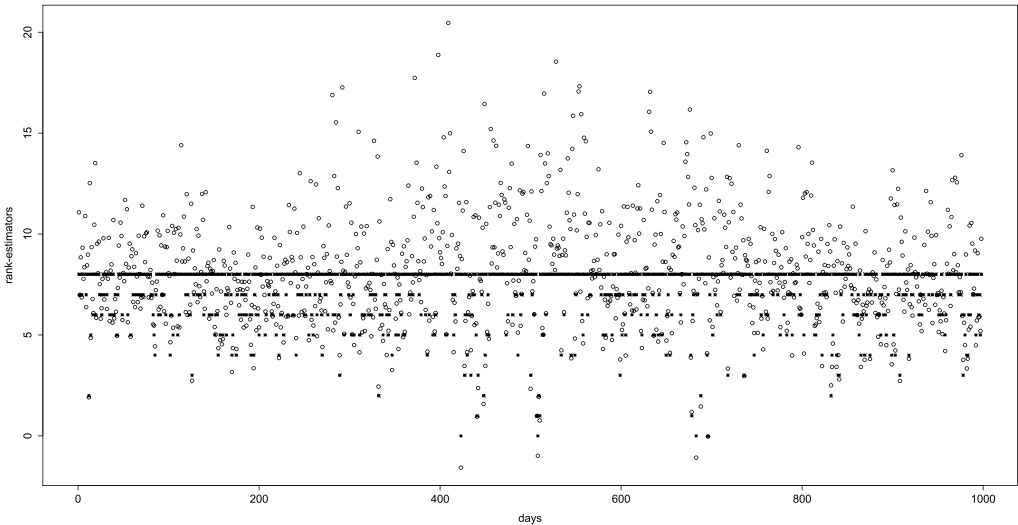
Figure 3. Estimators $\widehat{R}(g,h)_{10}^{n}$ (circles) and $\widehat{R}^{\mathrm{int}}(g,h)_{10}^{n}$ (squares) over a 10-days rolling time window.

to the overlapping increments approach. Consequently, it is clearly preferable to work with the results of the overlapping increments approach. An obvious and interesting question is whether one could reject the null hypothesis that the market at hand is complete. This amounts to asking
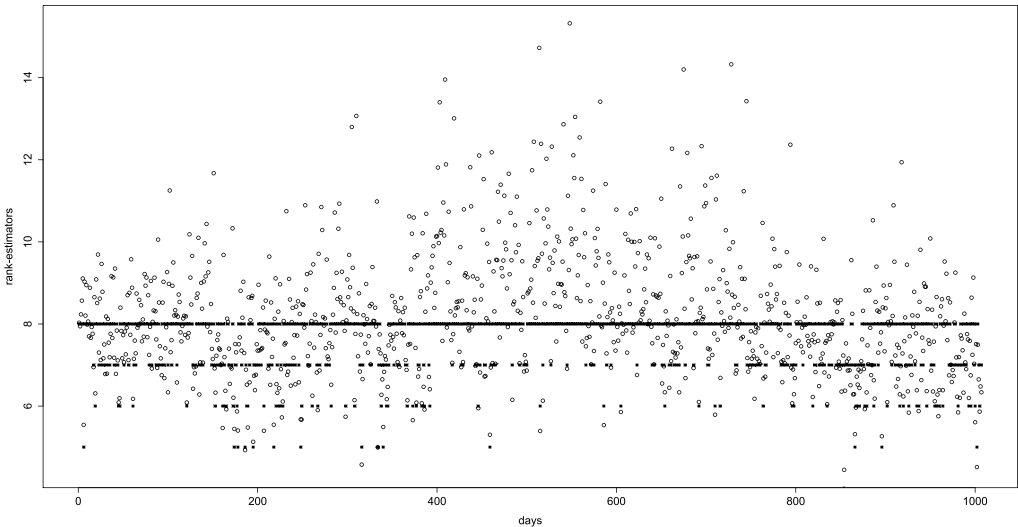

Figure 4. Estimators $\widetilde{R}(g,h)_{1}^{n}$ (circles) and $\widetilde{R}^{\mathrm{int}}(g,h)_{1}^{n}$ (squares) over a one-day time window.

Table 3. Sample mean and variance of $\widehat{R}(g,h)_1^n$, $\widehat{R}(g,h)_{10}^n$, and $\widetilde{R}(g,h)_1^n$

Overlapping	T	Mean	Variance
No	1	10.85	16.11
No	10	8.20	8.12
Yes	1	8.22	2.31

whether the maximal rank of the co-volatility is 8 for all days in the sample period and the days where $\widetilde{R}^{\mathrm{int}}(g,h)_1^n$ yields another result is just due to the variance of the estimator (and the low frequency). To this end, we count the events that $\widetilde{R}^{\mathrm{int}}(g,h)_1^n = 8$ which is for 700 out of the 1007 trading days in the sample period. For 223 days, the estimator yield 7, and for the rest of the days, it yields smaller values. If we compare these numbers with the results of model 1 in the simulation study for the overlapping approach (see also Table 2), those numbers are of comparable size: 72 out of 100 simulations, the estimator yields 8, 25 times the value 7, and 3 times the value 6. Of course, with 2.31 the sample variance of the estimator in the data example is much bigger than the variance in the simulation study with constant and full rank, which is 0.55. On the other hand, it is plausible that for the real data, there is both a non-vanishing drift term and a more complex volatility structure than a merely constant volatility. Hence, probably due to the lower order terms in the expansion of the main statistic, the variance is bigger; see also the discussion of the simulation study.

Finally, we see different rank regimes in the data. It seems that the rank estimator yields higher values in the mid-third of the sample period in comparison to the first and last third.

6. Proofs

Before presenting the proofs in detail, let us briefly outline the roadmap of this section. In Section 6.1, we introduce some technical results about expansions of determinants. We justify the asymptotic expansion at (3.2) and also show some more involved results.

In Section 6.2, we show that – using a standard localization procedure – we obtain the stochastic decomposition explained in Remark 4.2. Moreover, we show how the law of the dominating term in the expansion can be expressed in terms of the notation introduced in Section 4.2.

Section 6.3 is especially concerned with the proof of Proposition 4.9. To this end, we perform a very detailed analysis of the terms Γ_r and Γ_r' introduced at (4.10) and their dependency on the weight function g. This mainly relies on an application of the Leibniz rule for the calculation of the determinants and a repeated use of the Itô isometry to calculate the expectations.

Section 6.4 deals with the proofs of the main Theorem 4.13 and of Proposition 4.14. First, we show that – thanks to the stochastic expansion established in Section 6.2 – the main approximation idea motivated in Section 3.1 works in the stochastic setting. The second main step in the proof of Theorem 4.13 is the application of a stable central limit theorem for semimartingales (see, e.g., [10], Theorem IX.7.28). Proposition 4.14(i) follows along the lines of parts of the proof of Theorem 4.13. Proposition 4.14(ii) follows by Theorem 4.13, Theorem 4.8 – which in turn is

a direct consequence of Theorem 4.13 – and Proposition 4.14(ii) by applying the delta method for stable convergence. Note that this procedure does only work under a proper choice of the pair of weight functions which fulfills Proposition 4.9(ii).

The proof of Corollary 4.16 is essentially a consequence of the stable convergence at (4.28) and is referred to Section 6.5.

6.1. Expansion of determinants

Due to Section 3.1, the key to identifying the unknown rank of a matrix $A \in \mathcal{M}$ is the matrix perturbation method which results in the expansion at (3.2). While we could show the law of large numbers at (4.15) with an expansion like the one at (3.2), we need a higher order expansion of the determinant to derive the central limit theorem at (4.19). Therefore, we shall introduce some additional notation to the one in Section 3.1 which is similar to the one introduced in [8].

In the sequel, $\|A\|$ denotes the Euclidean norm of a matrix $A \in \mathcal{M}$. For any positive integer $m \geq 1$, we denote with $\mathcal{P}_m$ the set of all multi-integers $\mathbf{p} = (p_1, \ldots, p_m)$ with $p_j \geq 0$ and $p_1 + \cdots + p_m = d$, and $\mathcal{I}_\mathbf{p}$ the set of all partitions $\mathbf{I} = (I_1, \ldots, I_m)$ of $\{0, \ldots, d\}$ such that I_j contains exactly p_j points. For $\mathbf{p} \in \mathcal{P}_m$, $\mathbf{I} \in \mathcal{I}_\mathbf{p}$ and $A_1, \ldots, A_m \in \mathcal{M}$, we call $G^\mathbf{I}_{A_1, \ldots, A_m}$ the matrix in $\mathcal{M}$ whose ith column is the ith column of A_j when $j \in I_j$.

Due to the multilinearity property of the determinant we have the following identity for all $A_1, \ldots, A_m \in \mathcal{M}$

$$\det(A_1 + \cdots + A_m) = \sum_{\mathbf{p} \in \mathcal{P}_m} \sum_{\mathbf{I} \in \mathcal{I}_p} \det\big(G^\mathbf{I}_{A_1, \ldots, A_m}\big). \tag{6.1}$$

For $A, B, C \in \mathcal{M}$ and $r \in \{0, \ldots, d\}$, recalling (3.1), we recover the identity

$$\gamma_r(A, B) = \sum_{\mathbf{I} \in \mathcal{I}_{(r, d-r)}} \det\big(G^\mathbf{I}_{A, B}\big) \tag{6.2}$$

and set

$$\gamma'_r(A, B, C) := \sum_{\mathbf{I} \in \mathcal{I}_{(r, d-r-1, 1)}} \det\big(G^\mathbf{I}_{A, B, C}\big), \tag{6.3}$$

with the convention that $\gamma_{-1}(A, B) = 0$ and $\gamma'_d(A, B, C) = 0$. Let $A, B, C, D \in \mathcal{M}$ and rank$(A) \leq r$. Using (6.1), we obtain the asymptotic expansion

$$\begin{aligned}
\det&\big(A + \lambda B + \lambda^2 C + \lambda^2 D\big) \\
&= \lambda^{d-r} \gamma_r(A, B) \\
&\quad + \lambda^{d-r+1}\big(\gamma_{r-1}(A, B) + \gamma'_r(A, B, C) + \gamma'_r(A, B, D)\big) + O\big(\lambda^{d-r+2}\big) \qquad \text{as } \lambda \downarrow 0.
\end{aligned} \tag{6.4}$$

This observation gives rise to the following lemma (see also [8], Lemma 6.2).

Lemma 6.1. *There is a constant $K > 0$ such that for all $r \in \{0, \ldots, d\}$, all $\lambda \in (0, 1]$ and all $A, B, C, D \in \mathcal{M}$ with $\mathrm{rank}(A) \leq r$ we have with $\Lambda = \|A\| + \|B\| + \|B\| + \|D\|$*

$$\left|\det\left(A + \lambda B + \lambda^2 C + \lambda^2 D\right) - \lambda^{d-r}\gamma_r(A, B) - \lambda^{d-r+1}\left(\gamma_{r-1}(A, B) + \gamma_r'(A, B, C)\right)\right| \tag{6.5}$$
$$\leq K\lambda^{r-d+1}\Lambda^{d-1}\left(\lambda\Lambda + \|D\|\right),$$

$$\left|\frac{1}{\lambda^{2d-2r}}\det\left(A + \lambda B + \lambda^2 C + \lambda^2 D\right)^2 - \gamma_r(A, B)^2\right.$$
$$\left. - 2\lambda\gamma_r(A, B)\left(\gamma_{r-1}(A, B) + \gamma_r'(A, B, C)\right)\right| \tag{6.6}$$
$$\leq K\lambda\Lambda^{2d-1}\left(\lambda\Lambda + \|D\|\right).$$

If further $\lambda' \in (0, 1]$, $A', B', C', D' \in \mathcal{M}$ with $\mathrm{rank}(A') \leq r$ and $\Lambda' = \|A'\| + \|B'\| + \|C'\| + \|D'\|$, then

$$\left|\frac{1}{(\lambda\lambda')^{2d-2r}}\det\left(A + \lambda B + \lambda^2 C + \lambda^2 D\right)^2\det\left(A' + \lambda' B' + \lambda'^2 C' + \lambda'^2 D'\right)^2\right.$$
$$\left. - \gamma_r(A, B)^2\gamma_r\left(A', B'\right)^2\right| \leq K\left(\lambda + \lambda'\right)\Lambda\Lambda'. \tag{6.7}$$

Proof. The inequalities at (6.5) and (6.7) essentially follow from the asymptotic expansion at (6.4) and the fact that there is a $K > 0$ such that for any $\mathbf{p} \in \mathcal{P}_4$, $\mathbf{I} \in \mathcal{I}_{\mathbf{p}}$ and $\lambda \in (0, 1]$ we have

$$\left|\det\left(G^{\mathbf{I}}_{A, \lambda B, \lambda^2 C, \lambda^2 D}\right)\right| = \lambda^{p_2+2p_3+2p_4}\left|\det\left(G^{\mathbf{I}}_{A, B, C, D}\right)\right| \leq K\lambda^{p_2+2p_3+2p_4}\Lambda^{d-p_4}\|D\|^{p_4}.$$

The inequality at (6.6) follows from (6.5) by taking squares. □

6.2. The stochastic decomposition

Under Assumption (A) and by a standard localization procedure (see, e.g., [2], Section 3), it is no restriction to make the following technical assumption.

Assumption (A1). *Assumption (A) holds and the processes X_t, b_t, σ_t, a_t, v_t, a_t', v_t', a_t'', v_t'' defined at (2.1) and (2.2) are uniformly bounded in (ω, t).*

We make the convention that all constants are denoted by K, or K_p if they depend on an additional parameter p. The constants never depend on T, t, n, i, j. To ease notation, we use generic constants that may change from line to line. We introduce the filtration $(\mathcal{H}_t)_{t \in [0, T]}$ defined as

$$\mathcal{H}_t := \mathcal{F}_t \vee \sigma(\varepsilon),$$

where $\sigma(\varepsilon)$ is the σ-field generated by the whole process $(\varepsilon)_{t\in[0,T]}$. For any process V and for the filtrations $(\mathcal{F}_t)_{t\in[0,T]}$, $(\mathcal{H}_t)_{t\in[0,T]}$ and $\kappa = 1, 2$, we will use the simplifying notation

$$V_i^{n,\kappa} = V_{(3i+\kappa-1)du_n}, \qquad \mathcal{F}_i^{n,\kappa} = \mathcal{F}_{(3i+\kappa-1)du_n}, \qquad \mathcal{H}_i^{n,\kappa} = \mathcal{H}_{(3i+\kappa-1)du_n}. \qquad (6.8)$$

Note that we have the "nesting property" $\mathcal{F}_i^{n,1} \subset \mathcal{F}_i^{n,2}$ and $\mathcal{H}_i^{n,1} \subset \mathcal{H}_i^{n,2}$, respectively. Now, we show that under Assumption (A) we can obtain the stochastic decomposition at (4.5) explained in Remark 4.2. To do so, we notice (see [8], Section 6) that under Assumption (A), and for any $z \leq t \leq s$, we have the following expansion for the increment $X_s - X_t = \int_t^s b_u\, ds + \int_t^s \sigma_s\, dW_s$ (using vector notation):

$$\int_t^s b_u\, du = \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4,$$

$$\int_t^s \sigma_u\, dW_u = \alpha_5 + \alpha_6 + \alpha_7 + \alpha_8 + \alpha_9 + \alpha_{10} + \alpha_{11},$$

where

$$\alpha_1 = b_z(s-t), \qquad \alpha_2 = \int_t^s \left(\int_z^u a_w'\, dw \right) du,$$

$$\alpha_3 = v_z' \int_t^s (W_u - W_z)\, du, \qquad \alpha_4 = \int_t^s \left(\int_z^u (v_w' - v_z')\, dW_w \right) du,$$

$$\alpha_5 = \sigma_z(W_s - W_t), \qquad \alpha_6 = a_z \int_t^s (u-z)\, dW_u,$$

$$\alpha_7 = \int_t^s \left(\int_z^u (a_w - a_z)\, dw \right) dW_u, \qquad \alpha_8 = v_z \int_t^s (W_u - W_z)\, dW_u, \qquad (6.9)$$

$$\alpha_9 = \int_t^s \left(\int_z^u \left(\int_z^w a_r''\, dr \right) dW_w \right) dW_u,$$

$$\alpha_{10} = v_z'' \int_t^s \left(\int_z^u (W_w - W_z)\, dW_w \right) dW_u,$$

$$\alpha_{11} = \int_t^s \left(\int_z^u \left(\int_z^w (v_r'' - v_z'')\, dW_r \right) dW_w \right) dW_u.$$

By the Burkhölder–Gundy inequality (see, e.g., [13]), we have under Assumption (A1) for all $p, t, s > 0$ and for $V = X, \sigma, b, v$ that

$$\mathbb{E}\left[\sup_{u\in[0,s]} \|V_{t+u} - V_t\|^p \Big| \mathcal{F}_t \right] \leq K_p s^{p/2}. \qquad (6.10)$$

We set

$$\eta_{t,s} = \sup_{u\in[0,s], V=a,v',v''} \|V_{t+u} - V_t\|^2, \qquad n_i^{n,\kappa} = \sqrt{\mathbb{E}\big[\eta_{(3i+\kappa-1)du_n, \kappa du_n} | \mathcal{F}_i^{n,\kappa}\big]}.$$

Using the Burkhölder–Gundy inequality and Hölder inequality leads to (recall that $z \leq s$)

$$
\mathbb{E}\big[\|\alpha_j\|^p|\mathcal{F}_z\big] \leq
\begin{cases}
K_p(s-z)^{p/2} & \text{if } j = 5, \\
K_p(s-z)^p & \text{if } j = 1, 8, \\
K_p(s-z)^{3p/2} & \text{if } j = 3, 6, 10, \\
K_p(s-z)^{2p} & \text{if } j = 2, 9, \\
K_p(s-z)^{3p/2}\mathbb{E}\big[\eta_{z,s-z}^p|\mathcal{F}_z\big] & \text{if } j = 4, 7, 11.
\end{cases}
\tag{6.11}
$$

Let g be a weight function (see Section 3.2) and g^n its discretization introduced at (4.11). For $\kappa = 1, 2$, we define the function

$$
g_{n,\kappa}(x) := g^n(\kappa u_n x) = \sum_{j=1}^{k_n-1} g\left(\frac{j}{k_n}\right)\mathbf{1}_{(\kappa(j-1)\Delta_n, \kappa j \Delta_n]}(x).
$$

Using (6.9) with $z = (3i + \kappa - 1)du_n$, $t = ((3i + \kappa - 1)d + \kappa(j-1))u_n$, $s = ((3i + \kappa - 1)d + \kappa j)u_n$ with $i \in \{0, \ldots, [T/3du_n] - 1\}$, $j \in \{1, \ldots, d\}$ and $\kappa = 1, 2$ we then obtain the stochastic decomposition at (4.5), namely

$$
\frac{1}{\sqrt{\kappa u_n}}\, \text{mat}\big(\overline{Z}(g)_{(3i+\kappa-1)dk_n}^{n,\kappa}, \ldots, \overline{Z}(g)_{((3i+\kappa-1)d+\kappa(d-1))k_n}^{n,\kappa}\big)
$$
$$
= A(g)_i^{n,\kappa} + \sqrt{\kappa u_n}\big(B(1, g)_i^{n,\kappa} + B(2, g)_i^{n,\kappa} + B(3, g)_i^{n,\kappa}\big)
$$
$$
+ \kappa u_n C(g)_i^{n,\kappa} + \kappa u_n D(g)_i^{n,\kappa},
$$

where (using vector notation)

$$
A(g)_{i,j}^{n,\kappa} = \frac{\sigma_i^{n,\kappa}}{\sqrt{\kappa u_n}}\int_{((3i+\kappa-1)d+\kappa(j-1))u_n}^{((3i+\kappa-1)d+\kappa j)u_n} g_{n,\kappa}\big(s - ((3i + \kappa - 1)d + \kappa(j-1))u_n\big)\,dW_s,
$$

$$
B(1, g)_{i,j}^{n,\kappa} = \frac{b_i^{n,\kappa}}{\kappa u_n}\int_{((3i+\kappa-1)d+\kappa(j-1))u_n}^{((3i+\kappa-1)d+\kappa j)u_n} g_{n,\kappa}\big(s - ((3i + \kappa - 1)d + \kappa(j-1))u_n\big)\,ds
$$
$$
+ \frac{v_i^{n,\kappa}}{\kappa u_n}\int_{((3i+\kappa-1)d+\kappa(j-1))u_n}^{((3i+\kappa-1)d+\kappa j)u_n} g_{n,\kappa}\big(s - ((3i + \kappa - 1)d + \kappa(j-1))u_n\big)
$$
$$
\times (W_s - W_{(3i+\kappa-1)du_n})\,dW_s,
$$

$$
B(2, g)_{i,j}^{n,\kappa} = \frac{\widetilde{\sigma}}{\sqrt{\kappa u_n}}\int_{((3i+\kappa-1)d+\kappa(j-1))u_n}^{((3i+\kappa-1)d+\kappa j)u_n} g_{n,\kappa}\big(s - ((3i + \kappa - 1)d + \kappa(j-1))u_n\big)\,dW_s',
$$

$$
B(3, g)_{i,j}^{n,\kappa} = \frac{1}{\kappa u_n}\overline{\varepsilon}(g)_{((3i+\kappa-1)d+\kappa(j-1))k_n}^{n,\kappa},
$$

$$
C(g)_{i,j}^{n,\kappa} = \frac{a_i^{n,\kappa}}{(\kappa u_n)^{3/2}}\int_{((3i+\kappa-1)d+\kappa(j-1))u_n}^{((3i+\kappa-1)d+\kappa j)u_n} g_{n,\kappa}\big(s - ((3i + \kappa - 1)d + \kappa(j-1))u_n\big)
$$

$$\times \left(s - (3i + \kappa - 1)du_n \right) dW_s$$

$$+ \frac{v'^{n,\kappa}_i}{(\kappa u_n)^{3/2}} \int_{((3i+\kappa-1)d+\kappa(j-1))u_n}^{((3i+\kappa-1)d+\kappa j)u_n} g_{n,\kappa}\big(s - ((3i + \kappa - 1)d + \kappa(j - 1))u_n \big)$$

$$\times (W_s - W_{(3i+\kappa-1)du_n}) \, ds$$

$$+ \frac{v''^{n,\kappa}_i}{(\kappa u_n)^{3/2}} \int_{((3i+\kappa-1)d+\kappa(j-1))u_n}^{((3i+\kappa-1)d+\kappa j)u_n} g_{n,\kappa}\big(s - ((3i + \kappa - 1)d + \kappa(j - 1))u_n \big)$$

$$\times \left(\int_{(3i+\kappa-1)du_n}^{s} (W_u - W_{(3i+\kappa-1)du_n}) \, dW_u \right) dW_s,$$

and $D(g)^{n,\kappa}_{i,j}$ is the remainder term. In the sequel, we will make the convention that $B(g)^{n,\kappa}_i :=$ $B(1, g)^{n,\kappa}_i + B(2, g)^{n,\kappa}_i + B(3, g)^{n,\kappa}_i$. With the following lemma, we can deduce that under Assumption (A1) the $\mathbb{R}^{d \times d}$-valued sequences $A(g)^{n,\kappa}_i, B(g)^{n,\kappa}_i, C(g)^{n,\kappa}_i, D(g)^{n,\kappa}_i$ are tight (see also equation (6.15) in [8]).

Lemma 6.2. *Let the Assumptions* (A1) *and* (E) *be satisfied. For $p \geq 1$, there is a $K_p > 0$ such that we have the following estimates*

$$\mathbb{E}\big[\big\| A(g)^{n,\kappa}_{i,j} \big\|^p + \big\| B(g)^{n,\kappa}_{i,j} \big\|^p + \big\| C(g)^{n,\kappa}_{i,j} \big\|^p | \mathcal{F}^{n,\kappa}_i \big] \leq K_p,$$

$$\mathbb{E}\big[\big\| D(g)^{n,\kappa}_{i,j} \big\|^p | \mathcal{F}^{n,\kappa}_i \big] \leq K_p \big(u_n^{p/2} + \big(\eta^{n,\kappa}_i \big)^{p \wedge 2} \big) \leq K_p.$$

Proof. To show the estimate for the term $B(3, g)^{n,\kappa}_{i,j}$, we refer to [9], equation (16.2.3), which implies that $\mathbb{E}[\|\overline{\varepsilon}(g)^{n,\kappa}_i\|^p] \leq K_p k_n^{-p/2}$, such that the claim follows by recalling (3.5). For the remaining terms, we use (6.11) with $z = (3i + \kappa - 1)du_n, t = ((3i + \kappa - 1)d + \kappa(j - 1))u_n, s = ((3i + \kappa - 1)d + \kappa j)u_n$ plus the fact that $g_{n,\kappa}$ is uniformly bounded in n. □

Lemma 6.3. *Assume Assumptions* (A1) *and* (E). *Then* $u_n \mathbb{E}[\sum_{i=0}^{[T/3du_n]-1} \eta^{n,\kappa}_i] \to 0$.

Proof. The proof follows along the lines of the proof of Lemma 6.3 in [8]. □

Lemma 6.4. *Let the Assumptions* (A1) *and* (E) *be satisfied. Fix a weight function g. Then, for any $r \in \{0, \ldots, d\}$, $i = 0, \ldots, [T/3du_n] - 1$ and $\kappa = 1, 2$ the $\mathcal{F}^{n,\kappa}_i$-conditional law of*

$$\gamma_r \big(A(g)^{n,\kappa}_i, B(g)^{n,\kappa}_i \big)^2$$

coincides with the $\mathcal{F}^{n,\kappa}_i$-conditional law of

$$\overline{F}_r \big(\sigma^{n,\kappa}_i, \widetilde{\sigma}, v^{n,\kappa}_i, b^{n,\kappa}_i, \Sigma^n, g^n, \kappa \big),$$

where

$$\Sigma^n := \frac{\theta^3}{k_n^3 \Delta_n^2} \Sigma. \tag{6.12}$$

Proof. The quantity $\overline{F}_r(\sigma_i^{n,\kappa}, \tilde{\sigma}, v_i^{n,\kappa}, b_i^{n,\kappa}, \Sigma^n, g^n, \kappa)$ can be realized on the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0,T]}, \mathbb{P})$ by taking

$$\overline{W}_t = \frac{W_{((3id+\kappa-1)d+t)u_n} - W_{(3id+\kappa-1)du_n}}{\sqrt{u_n}}, \qquad \overline{W}'_t = \frac{W'_{((3id+\kappa-1)d+t)u_n} - W'_{(3id+\kappa-1)du_n}}{\sqrt{u_n}}.$$

Then define $\overline{\Theta}$ implicitly, such that

$$\frac{1}{\kappa}\left(\frac{\psi_1(g)}{\theta^3}\right)^{1/2} \sum_{m=1}^{d} ((\Sigma^n)^{1/2})^{lm} \overline{\Theta}_{\kappa j}^m = B(3, g)_{i,j}^{n,\kappa}.$$

Indeed, we know that $B(3, g)_{i,j}^{n,\kappa}$ is a centered Gaussian random variable with covariance matrix

$$\frac{1}{\kappa^2 u_n^2} \sum_{\mu=0}^{k_n-1} \left(g\left(\frac{\mu+1}{k_n}\right) - g\left(\frac{\mu}{k_n}\right)\right)^2 \Sigma = \frac{\psi_1(g^n)}{\kappa^2 k_n^3 \Delta_n^2} \Sigma = \frac{1}{\kappa^2} \frac{\psi_1(g^n)}{\theta^3} \Sigma^n,$$

where $\psi_1(g^n)$ is defined at (4.12). We also remark that due to (3.5) we obtain that $\Sigma^n = (1 + o(\Delta_n^{1/6}))\Sigma$. $\qquad\square$

As a direct consequence of Lemma 6.4, we can deduce that

$$\mathbb{E}\big[\gamma_r\big(A(g)_i^{n,\kappa}, B(g)_i^{n,\kappa}\big)^2 | \mathcal{F}_i^{n,\kappa}\big] = \Gamma_r\big(\sigma_i^{n,\kappa}, \tilde{\sigma}, v_i^{n,\kappa}, b_i^{n,\kappa}, \Sigma^n, g^n, \kappa\big), \tag{6.13}$$

$$\mathrm{Var}\big[\gamma_r\big(A(g)_i^{n,\kappa}, B(g)_i^{n,\kappa}\big)^2 | \mathcal{F}_i^{n,\kappa}\big] = \Gamma'_r\big(\sigma_i^{n,\kappa}, \tilde{\sigma}, v_i^{n,\kappa}, b_i^{n,\kappa}, \Sigma^n, g^n, \kappa\big). \tag{6.14}$$

6.3. Proof of Lemma 4.7 and Proposition 4.9

Proof of Lemma 4.7. The proof of Lemma 4.7 follows along the lines of the proof of [8], Lemma 3.1. (Notice that we can incorporate the additional terms with $\overline{\Theta}_j$ appearing in (4.9) in the terms Θ_j at [8], equation (6.8).) $\qquad\square$

Proof of Proposition 4.9. We start with the proof of part (i). Let $r \in \{0, \ldots, d\}$, $\underline{u} = (\alpha, \beta, \gamma, a, \varphi) \in \mathcal{U}$, $\kappa = 1, 2$ and g be any weight function. Using the notation at (4.8) and (4.9), we define the matrices

$$A(\underline{u}, g, \kappa) := \big(\Psi(\underline{u}, g, \kappa)_i^j\big)_{i,j=1,\ldots,d}, \qquad B(\underline{u}, g, \kappa) := \big(\Psi(\underline{u}, g, \kappa)_i^{d+j}\big)_{i,j=1,\ldots,d},$$

being elements of $\mathcal{M}$. Furthermore, for $\mathbf{I} \in \mathcal{I}_{(r,d-r)}$ we will use the notation

$$G_{A(\underline{u},g,\kappa),B(\underline{u},g,\kappa)}^{\mathbf{I}} = \big((G_{A(\underline{u},g,\kappa),B(\underline{u},g,\kappa)}^{\mathbf{I}})_i^j\big)_{i,j=1,\ldots,d}.$$

Then, developing the determinant with the Leibniz rule, we obtain the identity

$$
\begin{aligned}
\Gamma_r(\underline{u}, g, \kappa) &= \overline{\mathbb{E}}\big[\gamma_r\big(A(\underline{u}, g, \kappa), B(\underline{u}, g, \kappa)\big)^2\big] \\
&= \overline{\mathbb{E}}\Bigg[\sum_{\mathbf{I}, \mathbf{I}' \in \mathcal{I}_{(r, d-r)}} \det\big(G^{\mathbf{I}}_{A(\underline{u}, g, \kappa), B(\underline{u}, g, \kappa)}\big) \det\big(G^{\mathbf{I}'}_{A(\underline{u}, g, \kappa), B(\underline{u}, g, \kappa)}\big)\Bigg] \\
&= \overline{\mathbb{E}}\Bigg[\sum_{\mathbf{I}, \mathbf{I}' \in \mathcal{I}_{(r, d-r)}} \sum_{\pi, \pi' \in \mathfrak{S}_d} \operatorname{sgn}(\pi) \operatorname{sgn}(\pi') \\
&\qquad \times \prod_{i=1}^{d} \big(G^{\mathbf{I}}_{A(\underline{u}, g, \kappa), B(\underline{u}, g, \kappa)}\big)^{\pi(i)}_i \prod_{j=1}^{d} \big(G^{\mathbf{I}'}_{A(\underline{u}, g, \kappa), B(\underline{u}, g, \kappa)}\big)^{\pi'(j)}_j\Bigg] \\
&= \sum_{\mathbf{I}, \mathbf{I}' \in \mathcal{I}_{(r, d-r)}} \sum_{\pi, \pi' \in \mathfrak{S}_d} \operatorname{sgn}(\pi) \operatorname{sgn}(\pi') \\
&\qquad \times \prod_{i=1}^{d} \overline{\mathbb{E}}\big[\big(G^{\mathbf{I}}_{A(\underline{u}, g, \kappa), B(\underline{u}, g, \kappa)}\big)^{\pi(i)}_i \big(G^{\mathbf{I}'}_{A(\underline{u}, g, \kappa), B(\underline{u}, g, \kappa)}\big)^{\pi'(i)}_i\big],
\end{aligned} \tag{6.15}
$$

where $\mathfrak{S}_d$ denotes the group of all permutations of the set $\{1, \ldots, d\}$ and $\operatorname{sgn}(\pi) \in \{-1, 1\}$ is the sign of the permutation $\pi \in \mathfrak{S}_d$. The last step in the computation is due to the fact that the vectors $(G^{\mathbf{I}}_{A(\underline{u}, g, \kappa), B(\underline{u}, g, \kappa)})_i$ and $(G^{\mathbf{I}'}_{A(\underline{u}, g, \kappa), B(\underline{u}, g, \kappa)})_j$ are uncorrelated if $i \neq j$. Thus, for fixed $r \in \{0, \ldots, d\}$ and $\kappa = 1, 2$, the mapping $(\underline{u}, g) \mapsto \Gamma_r(\underline{u}, g, \kappa)$ can be considered as a polynomial in $\binom{d}{r}^2 \times (d!)^2 \times d$ variables of the form

$$
\overline{\mathbb{E}}\big[\big(G^{\mathbf{I}}_{A(\underline{u}, g, \kappa), B(\underline{u}, g, \kappa)}\big)^{\pi(i)}_i \big(G^{\mathbf{I}'}_{A(\underline{u}, g, \kappa), B(\underline{u}, g, \kappa)}\big)^{\pi'(i)}_i\big], \tag{6.16}
$$

where $\mathbf{I}, \mathbf{I}' \in \mathcal{I}_{(r, d-r)}$, $\pi, \pi' \in \mathfrak{S}_d$ and $i = 1, \ldots, d$. Using Itô's isometry, (6.16) takes one of the following three forms with $l, l' \in \{1, \ldots, d\}$:

$$
\overline{\mathbb{E}}\big[\Psi(\underline{u}, g, \kappa)^l_i \Psi(\underline{u}, g, \kappa)^{l'}_i\big] = \psi_2(g) \sum_{m=1}^{q} \alpha^{lm}\alpha^{l'm},
$$

$$
\overline{\mathbb{E}}\big[\Psi(\underline{u}, g, \kappa)^l_i \Psi(\underline{u}, g, \kappa)^{d+l'}_i\big] = 0,
$$

$$
\overline{\mathbb{E}}\big[\Psi(\underline{u}, g, \kappa)^{d+l}_i \Psi(\underline{u}, g, \kappa)^{d+l'}_i\big] = \psi_3(g)^2 a^l a^{l'} + \big(\psi_4(g) + (i-1)\psi_2(g)\big) \sum_{m,k=1}^{q} \gamma^{lkm}\gamma^{l'km}
$$

$$
+ \psi_2(g) \sum_{m=1}^{d} \beta^{lm}\beta^{l'm} + \frac{\psi_1(g)}{\kappa^2\theta^3} \sum_{m=1}^{d} \big(\varphi^{1/2}\big)^{lm}\big(\varphi^{1/2}\big)^{l'm}.
$$

Hence, if we additionally fix $\underline{u} \in \mathcal{U}$, then there is a polynomial $\tau_{r,\underline{u},\kappa} : \mathbb{R}^4 \to \mathbb{R}$ such that the mapping $g \mapsto \Gamma_r(\underline{u}, g, \kappa)$ can be written as

$$g \mapsto \tau_{r,\underline{u},\kappa}\big(\psi_1(g), \psi_2(g), \psi_3(g), \psi_4(g)\big).$$

This shows the first part of (4.16). To show the second part, we use the relationship

$$\Gamma_r'(\underline{u}, g, \kappa) = \overline{\mathbb{E}}\big[\gamma_r\big(A(\underline{u}, g, \kappa), B(\underline{u}, g, \kappa)\big)^4\big] - \Gamma_r(\underline{u}, g, \kappa)^2.$$

By a similar calculation as in (6.15), we obtain that

$$\overline{\mathbb{E}}\big[\gamma_r\big(A(\underline{u}, g, \kappa), B(\underline{u}, g, \kappa)\big)^4\big]$$

$$= \sum_{\substack{\mathbf{I},\mathbf{I'},\mathbf{I''},\mathbf{I'''} \\ \in \mathcal{I}_{(r,d-r)}} \sum_{\substack{\pi,\pi',\pi'',\pi''' \\ \in \mathfrak{S}_d}} \operatorname{sgn}(\pi)\operatorname{sgn}(\pi')\operatorname{sgn}(\pi'')\operatorname{sgn}(\pi''') \prod_{i=1}^{d} \overline{\mathbb{E}}\big[\big(G^{\mathbf{I}}_{A(\underline{u},g,\kappa), B(\underline{u},g,\kappa)}\big)_i^{\pi(i)}$$

$$\times \big(G^{\mathbf{I'}}_{A(\underline{u},g,\kappa), B(\underline{u},g,\kappa)}\big)_i^{\pi'(i)} \big(G^{\mathbf{I''}}_{A(\underline{u},g,\kappa), B(\underline{u},g,\kappa)}\big)_i^{\pi''(i)} \big(G^{\mathbf{I'''}}_{A(\underline{u},g,\kappa), B(\underline{u},g,\kappa)}\big)_i^{\pi'''(i)}\big].$$

If we fix again $r \in \{0, \ldots, d\}$, $\kappa = 1, 2$, the mapping $(\underline{u}, g) \mapsto \overline{\mathbb{E}}[\gamma_r(A(\underline{u}, g, \kappa), B(\underline{u}, g, \kappa))^4]$ can be considered as a polynomial in $\binom{d}{r}^4 \times (d!)^4 \times d$ variables of the form

$$\overline{\mathbb{E}}\big[\big(G^{\mathbf{I}}_{A(\underline{u},g,\kappa), B(\underline{u},g,\kappa)}\big)_i^{\pi(i)} \big(G^{\mathbf{I'}}_{A(\underline{u},g,\kappa), B(\underline{u},g,\kappa)}\big)_i^{\pi'(i)}$$

$$\times \big(G^{\mathbf{I''}}_{A(\underline{u},g,\kappa), B(\underline{u},g,\kappa)}\big)_i^{\pi''(i)} \big(G^{\mathbf{I'''}}_{A(\underline{u},g,\kappa), B(\underline{u},g,\kappa)}\big)_i^{\pi'''(i)}\big], \tag{6.17}$$

where $\mathbf{I}, \mathbf{I'}, \mathbf{I''}, \mathbf{I'''} \in \mathcal{I}_{(r,d-r)}$, $\pi, \pi', \pi'', \pi''' \in \mathfrak{S}_d$ and $i = 1, \ldots, d$. By a careful calculation, we can see that (6.17) takes one of the following five forms with $l, l', l'', l''' \in \{1, \ldots, d\}$:

$$\overline{\mathbb{E}}\big[\Psi(\underline{u}, g, \kappa)_i^l \Psi(\underline{u}, g, \kappa)_i^{l'} \Psi(\underline{u}, g, \kappa)_i^{l''} \Psi(\underline{u}, g, \kappa)_i^{l'''}\big] = \psi_2(g)^2 K_\alpha,$$

$$\overline{\mathbb{E}}\big[\Psi(\underline{u}, g, \kappa)_i^l \Psi(\underline{u}, g, \kappa)_i^{l'} \Psi(\underline{u}, g, \kappa)_i^{l''} \Psi(\underline{u}, g, \kappa)_i^{d+l'''}\big] = 0,$$

$$\overline{\mathbb{E}}\big[\Psi(\underline{u}, g, \kappa)_i^l \Psi(\underline{u}, g, \kappa)_i^{l'} \Psi(\underline{u}, g, \kappa)_i^{d+l''} \Psi(\underline{u}, g, \kappa)_i^{d+l'''}\big]$$

$$= \psi_2(g) \sum_{m=1}^{q} \alpha^{lm} \alpha^{l'm} \left(\psi_3(g)^2 a^{l''} a^{l'''} + \big(\psi_4(g) + (i-1)\psi_2(g)\big) \sum_{m,k=1}^{q} \gamma^{l''km} \gamma^{l'''km} \right.$$

$$\left. + \psi_2(g) \sum_{m=1}^{d} \beta^{l''m} \beta^{l'''m} + \frac{\psi_1(g)}{\kappa^2 \theta^3} \sum_{m=1}^{d} (\varphi^{1/2})^{l''m} (\varphi^{1/2})^{l'''m} \right),$$

$$\overline{\mathbb{E}}\big[\Psi(\underline{u}, g, \kappa)_i^l \Psi(\underline{u}, g, \kappa)_i^{d+l'} \Psi(\underline{u}, g, \kappa)_i^{d+l''} \Psi(\underline{u}, g, \kappa)_i^{d+l'''}\big] = 0,$$

$$\overline{\mathbb{E}}\big[\Psi(\underline{u}, g, \kappa)_i^l \Psi(\underline{u}, g, \kappa)_i^{l'} \Psi(\underline{u}, g, \kappa)_i^{d+l''} \Psi(\underline{u}, g, \kappa)_i^{d+l'''}\big]$$

$$= \psi_3(g)^4 a^l a^{l'} a^{l''} a^{l'''} + \big(\psi_4(g) + (i-1)\psi_2(g)\big)^2 K_\gamma + \psi_2(g)^2 K_\beta + \left(\frac{\psi_1(g)}{\kappa^2\theta^3}\right)^2 K_\varphi$$

$$+ \psi_3(g)^2\big(\psi_4(g) + (i-1)\psi_2(g)\big)K_{a,\gamma}$$

$$+ \psi_3(g)^2\psi_2(g)K_{a,\beta} + \psi_3(g)^2\frac{\psi_1(g)}{\kappa^2\theta^3}K_{a,\varphi} + \big(\psi_4(g) + (i-1)\psi_2(g)\big)\psi_2(g)K_{\gamma,\beta}$$

$$+ \big(\psi_4(g) + (i-1)\psi_2(g)\big)\frac{\psi_1(g)}{\kappa^2\theta^3}K_{\gamma,\varphi} + \psi_2(g)\frac{\psi_1(g)}{\kappa^2\theta^3}K_{\beta,\varphi}.$$

We remark that the constants do not depend on κ. Consequently, if we additionally fix $\underline{u} \in \mathcal{U}$, there is a polynomial $\tau'_{r,\underline{u},\kappa} : \mathbb{R}^4 \to \mathbb{R}$ such that the mapping $g \mapsto \Gamma'_r(\underline{u}, g, \kappa)$ can be written as

$$g \mapsto \tau'_{r,\underline{u},\kappa}\big(\psi_1(g), \psi_2(g), \psi_3(g), \psi_4(g)\big),$$

which proves part (i) of Proposition 4.9. By an inspection of the previous calculations, we see that the only term where κ appears is the term $\frac{\psi_1(g)}{\kappa^2\theta^3}$. Hence, for any $r \in \{0, \dots, d\}$, $\underline{u} \in \mathcal{U}$, we have

$$\tau_{r,\underline{u},1}(x_1, x_2, x_3, x_4) = \tau_{r,\underline{u},2}(4x_1, x_2, x_3, x_4),$$

$$\tau'_{r,\underline{u},1}(x_1, x_2, x_3, x_4) = \tau'_{r,\underline{u},2}(4x_1, x_2, x_3, x_4), \qquad (x_1, x_2, x_3, x_4) \in \mathbb{R}^4.$$

This shows part (ii) of Proposition 4.9. $\square$

6.4. Proof of Theorem 4.13 and Proposition 4.14

Let g be a weight function. We begin by constructing approximations for the main test statistics $S(g)_T^{n,\kappa}$ defined at (4.3) and $V(g, h)_T^{n,\kappa\kappa'}$ given at (4.21), (4.22) and (4.23). The approximations are discretized versions of $S(r, g)_T^\kappa$ (see (4.15)) and the right-hand sides of (4.25) to (4.27)

$$S(r, g)_T^{n,\kappa} := 3du_n \sum_{i=0}^{[T/3du_n]} \gamma_r\big(A(g)_i^{n,\kappa}, B(g)_i^{n,\kappa}\big)^2,$$

$$V(r, g, h)_T^{n,\kappa\kappa'} := \begin{cases} 9d^2u_n \displaystyle\sum_{i=0}^{[T/3du_n]-1} \gamma_r\big(A(g)_i^{n,1}, B(g)_i^{n,1}\big)^4, & \text{if } \kappa = \kappa' = 1, \\[2.5em] 9d^2u_n \displaystyle\sum_{i=0}^{[T/3du_n]-1} \gamma_r\big(A(h)_i^{n,2}, B(h)_i^{n,2}\big)^4, & \text{if } \kappa = \kappa' = 2, \\[2.5em] 9d^2u_n \displaystyle\sum_{i=0}^{[T/3du_n]-1} \gamma_r\big(A(g)_i^{n,1}, B(g)_i^{n,1}\big)^2 \\[0.5em] \quad \times \gamma_r\big(A(h)_i^{n,2}, B(h)_i^{n,2}\big)^2, & \text{if } \kappa = 1, \kappa' = 2. \end{cases} \tag{6.18}$$

The lemma is based on the asymptotic expansion at (6.4).

Lemma 6.5. *Assume Assumptions* (A1), (E), *let* $r \in \{0, \dots, d\}$, $\kappa, \kappa' = 1, 2$ *and* g, h *be two weight functions* (*not necessarily satisfying the conditions of Proposition* 4.9(ii)). *Then, on* $\Omega_T^{\leq r}$, *we have that*

$$\frac{1}{\sqrt{u_n}} \left(\frac{1}{(\kappa u_n)^{d-r}} S(g)_T^{n,\kappa} - S(r,g)_T^{n,\kappa} \right) \xrightarrow{\mathbb{P}} 0, \tag{6.19}$$

$$\frac{1}{(\kappa\kappa' u_n^2)^{d-r}} V(g,h)_T^{n,\kappa\kappa'} - V(r,g,h)_T^{n,\kappa\kappa'} \xrightarrow{\mathbb{P}} 0. \tag{6.20}$$

Proof. The proof is an adaption of the proof of [8], Lemma 6.4. Let $\xi(g)_i^{n,\kappa}$ denote the ith summand on the right-hand side of (4.3). We start by showing (6.19). To this end, we use the fact that $\mathrm{rank}(A(g)_i^{n,\kappa}) \leq r$ for all i to apply the inequality at (6.6) with $\lambda = \sqrt{\kappa u_n}$ to obtain

$$\frac{1}{(\kappa u_n)^{d-r}} \xi(g)_i^{n,\kappa} = \gamma_r \left(A(g)_i^{n,\kappa}, B(g)_i^{n,\kappa} \right)^2 + 2\sqrt{\kappa u_n} \zeta(g)_i^{n,\kappa} + \widetilde{\zeta}(g)_i^{n,\kappa},$$

where with the Cauchy–Schwarz inequality (using the conventions after (6.3))

$$\begin{aligned} \zeta(g)_i^{n,\kappa} &:= \gamma_r \left(A(g)_i^{n,\kappa}, B(g)_i^{n,\kappa} \right) \\ &\quad \times \left(\gamma_{r-1} \left(A(g)_i^{n,\kappa}, B(g)_i^{n,\kappa} \right) + \gamma_r' \left(A(g)_i^{n,\kappa}, B(g)_i^{n,\kappa}, C(g)_i^{n,\kappa} \right) \right), \end{aligned}$$
$$\mathbb{E}\left[|\widetilde{\zeta}(g)_i^{n,\kappa}| \right] \leq K u_n + K\sqrt{u_n} \mathbb{E}[\eta_i^{n,\kappa}].$$

Applying Lemma 6.3, we deduce that $\sqrt{u_n} \sum_{i=0}^{[T/3du_n]-1} \widetilde{\zeta}(g)_i^{n,\kappa} \xrightarrow{\mathbb{P}} 0$. Regarding the structure of $S(r,g)_T^{n,\kappa}$ we need to prove that $u_n \sum_{i=0}^{[T/3du_n]-1} \zeta(g)_i^{n,\kappa} \xrightarrow{\mathbb{P}} 0$. To this end, we consider the decomposition $\zeta(g)_i^{n,\kappa} = \zeta'(g)_i^{n,\kappa} + \zeta''(g)_i^{n,\kappa}$, where $\zeta''(g)_i^{n,\kappa} = \mathbb{E}[\zeta(g)_i^{n,\kappa}|\mathcal{F}_i^{n,\kappa}]$. We obtain

$$\begin{aligned} \mathbb{E}\left[\left(u_n \sum_{i=0}^{[T/3du_n]-1} \zeta''(g)_i^{n,\kappa} \right)^2 \right] &= u_n^2 \sum_{i,j=0}^{[T/3du_n]-1} \mathbb{E}\left[\zeta''(g)_i^{n,\kappa} \zeta''(g)_j^{n,\kappa} \right] \\ &= u_n^2 \sum_{i=0}^{[T/3du_n]-1} \mathbb{E}\left[|\zeta''(g)_i^{n,\kappa}|^2 \right] \\ &\leq u_n K T \to 0, \end{aligned} \tag{6.21}$$

where the second identity follows from the fact that $\zeta''(g)_i^{n,\kappa}$ is $\mathcal{F}_{i+1}^{n,\kappa}$-measurable and the last estimate is a consequence of Lemma 6.2 and the fact that γ_r and γ_r' are continuous functions. Hence, we know that $u_n \sum_{i=0}^{[T/3du_n]-1} \zeta''(g)_i^{n,\kappa} \xrightarrow{\mathbb{P}} 0$. So it is sufficient to show that $\zeta'(g)_i^{n,\kappa} = 0$, or the even stronger result that

$$\mathbb{E}\left[\zeta(g)_i^{n,\kappa} | \mathcal{H}_i^{n,\kappa} \vee \sigma(W') \right] = 0, \tag{6.22}$$

where $\sigma(W')$ is the σ-field generated by the whole process W' and $\mathcal{H}_i^{n,\kappa}$ was introduced before and in (6.8). Recalling the definitions at (6.2) and (6.3), equation (6.22) follows by the

implication

$$\mathbf{I} \in \mathcal{I}_{(r,d-r)}, \qquad \mathbf{I}' \in \mathcal{I}_{(r-1,d-r+1)}, \qquad \mathbf{I}'' \in \mathcal{I}_{(r,d-r-1,1)}$$

$$\implies \quad \mathbb{E}\big[\det\big(G^{\mathbf{I}}_{A(g)^{n,\kappa}_i, B(g)^{n,\kappa}_i}\big)\det\big(G^{\mathbf{I}'}_{A(g)^{n,\kappa}_i, B(g)^{n,\kappa}_i}\big)|\mathcal{H}^{n,\kappa}_i \vee \sigma(W')\big] = 0, \qquad (6.23)$$

$$\mathbb{E}\big[\det\big(G^{\mathbf{I}}_{A(g)^{n,\kappa}_i, B(g)^{n,\kappa}_i}\big)\det\big(G^{\mathbf{I}''}_{A(g)^{n,\kappa}_i, B(g)^{n,\kappa}_i, C(g)^{n,\kappa}_i}\big)|\mathcal{H}^{n,\kappa}_i \vee \sigma(W')\big] = 0. \quad (6.24)$$

Note that due to the conventions after (6.3) the left-hand side of (6.23) is 0 if $r = 0$, and the left-hand side of (6.24) is 0 if $r = d$. The d-dimensional variables $A(g)^{n,\kappa}_{i,j}$, $B(g)^{n,\kappa}_{i,j}$ and $C(g)^{n,\kappa}_{i,j}$ can be written in the form

$$\Phi\big(\omega, \big(W(\omega)_{(3i+\kappa-1)du_n+t} - W(\omega)_{(3i+\kappa-1)du_n}\big)_{t \geq 0}\big),$$

where Φ is a $(\mathcal{H}^{n,\kappa}_i \vee \sigma(W')) \otimes \mathcal{C}^q$-measurable function on $\Omega \times C(\mathbb{R}_+, \mathbb{R}^q)$. Here, $C(\mathbb{R}_+, \mathbb{R}^q)$ is the set of all continuous functions on $\mathbb{R}_+$ with values in $\mathbb{R}^q$ and $\mathcal{C}^q$ is its Borel σ-field for the local uniform topology. Notice that for $\Phi = A(g)^{n,\kappa}_{i,j}$ or $\Phi = C(g)^{n,\kappa}_{i,j}$, the mapping $x \mapsto \Phi(\omega, x)$ is odd, meaning that $\Phi(\omega, -x) = -\Phi(\omega, x)$, and for $\Phi = B(g)^{n,\kappa}_{i,j}$, it is even, meaning that $\Phi(\omega, -x) = \Phi(\omega, x)$. We set

$$\Psi = \det\big(G^{\mathbf{I}}_{A(g)^{n,\kappa}_i, B(g)^{n,\kappa}_i}\big),$$

$$\Psi' = \det\big(G^{\mathbf{I}'}_{A(g)^{n,\kappa}_i, B(g)^{n,\kappa}_i}\big),$$

$$\Psi'' = \det\big(G^{\mathbf{I}''}_{A(g)^{n,\kappa}_i, B(g)^{n,\kappa}_i, C(g)^{n,\kappa}_i}\big),$$

where Ψ, Ψ', Ψ'' are functions similar to Φ. Due to the multilinearity of the determinant we can deduce that if r is even, then Ψ is even and Ψ', Ψ'' are odd. If r is odd, Ψ is odd and Ψ', Ψ'' are even. Thus, in all cases, the products $\Psi\Psi'$ and $\Psi\Psi''$ are odd. Now, the $(\mathcal{H}^{n,\kappa}_i \vee \sigma(W'))$-conditional law of $(W_{(3i+\kappa-1)du_n+t} - W_{(3i+\kappa-1)du_n})_{t \geq 0}$ is invariant under the map $x \mapsto -x$ on $C(\mathbb{R}_+, \mathbb{R}^q)$, which implies (6.23), and hence (6.22).

The proof of (6.20) is more direct. We apply the estimate at (6.7) with $\lambda = \sqrt{\kappa u_n}$, $\lambda' = \sqrt{\kappa' u_n}$. With the previous notation and Lemma 6.2 we obtain

$$\mathbb{E}\bigg[\bigg|\frac{1}{(\kappa\kappa' u_n^2)^{d-r}} V(g,h)^{n,\kappa\kappa'}_T - V(r,g,h)^{n,\kappa\kappa'}_T\bigg|\bigg]$$

$$\leq 9d^2 u_n \sum_{i=0}^{[T/3du_n]-1} \mathbb{E}\bigg[\bigg|\frac{1}{(\kappa\kappa' u_n^2)^{d-r}} \xi(g)^{n,\kappa}_i \xi(h)^{n,\kappa'}_i$$

$$- \gamma'_r\big(A(g)^{n,\kappa}_i, B(g)^{n,\kappa}_i\big)^2 \gamma'_r\big(A(h)^{n,\kappa'}_i, B(h)^{n,\kappa'}_i\big)^2\bigg|\bigg]$$

$$\leq 9d^2 KT\sqrt{u_n} \to 0,$$

which implies (6.20). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

With respect to Lemma 6.5, Theorem 4.13 follows by showing the following lemma.

Lemma 6.6. *Assume Assumptions* (A1), (E). *Let* $r \in \{0, \ldots, d\}$ *and* g, h *be two weight function satisfying the conditions of Proposition* 4.9(ii). *Then, on* $\Omega_T^{\leq r}$, *we have the stable convergence*

$$U'(r, g, h)_T^n \xrightarrow{d_{\mathrm{st}}} \mathcal{MN}\big(0, V(r, g, h)_T\big),$$

where $\mathcal{MN}(0, V(r, g, h)_T)$ *is defined after equation* (4.20). *The two-dimensional statistic* $U'(r, g, h)_T^n = (U'(r, g, h)_T^{n,1}, U'(r, g, h)_T^{n,2})$ *is given via*

$$U'(r, g, h)_T^n := \frac{1}{\sqrt{u_n}}\big(S(r, g)_T^{n,1} - S(r, g)_T^1, \, S(r, h)_T^{n,2} - S(r, h)_T^2\big).$$

We will do the proof of Lemma 6.6 in three steps:

(i) Recall that due to Proposition 4.9(ii) we have that $S(r, g)_T^1 = S(r, h)_T^2$. By a Riemann approximation argument, one can show that

$$\begin{aligned}
&\frac{1}{\sqrt{u_n}}\Bigg(3du_n \sum_{i=0}^{[T/3du_n]-1} \Gamma_r\big(\sigma_i^{n,1}, \widetilde{\sigma}, v_i^{n,1}, b_i^{n,1}, \Sigma, g, 1\big) \\
&\qquad - \int_0^T \Gamma_r(\sigma_s, \widetilde{\sigma}, v_s, b_s, \Sigma, g, 1)\, ds\Bigg) \xrightarrow{\mathbb{P}} 0, \\
&\frac{1}{\sqrt{u_n}}\Bigg(3du_n \sum_{i=0}^{[T/3du_n]-1} \Gamma_r\big(\sigma_i^{n,2}, \widetilde{\sigma}, v_i^{n,2}, b_i^{n,2}, \Sigma, h, 2\big) \\
&\qquad - \int_0^T \Gamma_r(\sigma_s, \widetilde{\sigma}, v_s, b_s, \Sigma, h, 2)\, ds\Bigg) \xrightarrow{\mathbb{P}} 0.
\end{aligned} \tag{6.25}$$

More precisely, we use the fact that for a fixed weight function g and $\kappa = 1, 2$, the map $\mathcal{U} \ni \underline{u} \mapsto \Gamma_r(\underline{u}, g, \kappa)$ is a polynomial (and hence C^∞) as well as the fact that thanks to Assumption (A) the processes σ, v and b are Itô semimartingales and hence càdlàg (see Section 8 in [2] for more details).

(ii) We identify the limit by proving that

$$\begin{aligned}
3d\sqrt{u_n} \sum_{i=0}^{[T/3du_n]-1} \big(\Gamma_r\big(\sigma_i^{n,1}, \widetilde{\sigma}, v_i^{n,1}, b_i^{n,1}, \Sigma^n, g^n, 1\big) \\
- \Gamma_r\big(\sigma_i^{n,1}, \widetilde{\sigma}, v_i^{n,1}, b_i^{n,1}, \Sigma, g, 1\big)\big) \xrightarrow{\mathbb{P}} 0,
\end{aligned} \tag{6.26}$$

$$\begin{aligned}
3d\sqrt{u_n} \sum_{i=0}^{[T/3du_n]-1} \big(\Gamma_r\big(\sigma_i^{n,2}, \widetilde{\sigma}, v_i^{n,2}, b_i^{n,2}, \Sigma^n, h^n, 2\big) \\
- \Gamma_r\big(\sigma_i^{n,2}, \widetilde{\sigma}, v_i^{n,2}, b_i^{n,2}, \Sigma, h, 2\big)\big) \xrightarrow{\mathbb{P}} 0.
\end{aligned} \tag{6.27}$$

(iii) We prove the stable convergence

$$U''(r, g, h)_T^n \xrightarrow{d_{\mathrm{st}}} \mathcal{MN}\big(0, V(r, g, h)_T\big), \tag{6.28}$$

for the two-dimensional statistic $U''(r, g, h)_T^n = (U''(r, g, h)_T^{n,1}, U''(r, g, h)_T^{n,2})$ with components

$$U''(r, g, h)_T^{n,1} = 3d\sqrt{u_n} \sum_{i=0}^{[T/3du_n]-1} \big(\gamma_r\big(A(g)_i^{n,1}, B(g)_i^{n,1}\big)^2 - \Gamma_r\big(\sigma_i^{n,1}, \widetilde{\sigma}, v_i^{n,1}, b_i^{n,1}, \Sigma^n, g^n, 1\big)\big),$$

$$U''(r, g, h)_T^{n,2} = 3d\sqrt{u_n} \sum_{i=0}^{[T/3du_n]-1} \big(\gamma_r\big(A(h)_i^{n,2}, B(h)_i^{n,2}\big)^2 - \Gamma_r\big(\sigma_i^{n,2}, \widetilde{\sigma}, v_i^{n,2}, b_i^{n,2}, \Sigma^n, h^n, 2\big)\big).$$

The following lemma is concerned with the convergence at (6.26) and (6.27), respectively.

Lemma 6.7. *Assume Assumptions* (A1), (E). *Let* $r \in \{0, \ldots, d\}$, $\kappa = 1, 2$ *and* g *be a weight function. Then, on* $\Omega_T^{\leq r}$, *it holds that*

$$\sqrt{u_n} \sum_{i=0}^{[T/3du_n]-1} \big(\Gamma_r\big(\sigma_i^{n,\kappa}, \widetilde{\sigma}, v_i^{n,\kappa}, b_i^{n,\kappa}, \Sigma^n, g^n, \kappa\big)$$

$$- \Gamma_r\big(\sigma_i^{n,\kappa}, \widetilde{\sigma}, v_i^{n,\kappa}, b_i^{n,\kappa}, \Sigma, g, \kappa\big)\big) \xrightarrow{\mathbb{P}} 0. \tag{6.29}$$

Proof. Fix $r \in \{0, \ldots, d\}$, $\kappa = 1, 2$ and a weight function g. Recall that by Proposition 4.9(i) for any $\underline{u} \in \mathcal{U}$ there is a polynomial $\tau_{r,\underline{u},\kappa}$ such that we have $\Gamma_r(\underline{u}, g, \kappa) = \tau_{r,\underline{u},\kappa}(\psi_1(g), \ldots, \psi_4(g))$. An inspection of the proof of Proposition 4.9(i) yields that the map

$$\mathcal{U} \times \mathbb{R}^4 \to \mathbb{R}, \qquad \big(\alpha, \beta, \gamma, a, \Sigma, \psi_1(g), \ldots, \psi_4(g)\big) \mapsto \tau_{r,(\alpha,\beta,\gamma,a,\Sigma),\kappa}\big(\psi_1(g), \ldots, \psi_4(g)\big)$$

is a C^∞-function. Consider the first order partial derivatives in $(\Sigma, \psi_1(g), \ldots, \psi_4(g))$. For fixed Σ, g, they are continuous in $(\alpha, \beta, \gamma, a)$. Therefore, by a first order Taylor expansion, we obtain that for any compact set $A \subset \mathcal{M}' \times \mathcal{M} \times \mathbb{R}^{dq^2} \times \mathbb{R}^d$

$$\sup_{(\alpha,\beta,\gamma,a)\in A} \big|\Gamma_r\big(\alpha, \beta, \gamma, a, \Sigma^n, g^n, \kappa\big) - \Gamma_r\big(\alpha, \beta, \gamma, a, \Sigma, g, \kappa\big)\big|$$

$$\leq K_A \big\|\big(\Sigma^n, \psi_1\big(g^n\big), \ldots, \psi_4\big(g^n\big)\big) - \big(\Sigma, \psi_1(g), \ldots, \psi_4(g)\big)\big\|_{\mathbb{R}^{d^2} \times \mathbb{R}^4},$$

where $\| \cdot \|_{\mathbb{R}^{d^2} \times \mathbb{R}^4}$ is the Euclidean norm on $\mathbb{R}^{d^2} \times \mathbb{R}^4$. Combining (6.12) and (3.5) we get that $(\Sigma^n)_{ij} - \Sigma_{ij} = o(\Delta_n^{1/6})$, $i, j = 1, \ldots, d$, and with (4.11), (4.12), we have that $\psi_l(g^n) - \psi_l(g) = O(k_n^{-1})$, $l = 1, \ldots, 4$. Again using (3.5) this implies that

$$\big\|\big(\Sigma^n, \psi_1\big(g^n\big), \ldots, \psi_4\big(g^n\big)\big) - \big(\Sigma, \psi_1(g), \ldots, \psi_4(g)\big)\big\|_{\mathbb{R}^{d^2} \times \mathbb{R}^4} = o\big(\Delta_n^{1/6}\big).$$

Now, we apply Assumption (A1) to deduce that

$$\sup_{s\in[0,T]} \mathbb{E}\big[\big|\Gamma_r(\sigma_s, \tilde{\sigma}, v_s, b_s, \Sigma^n, g^n, \kappa) - \Gamma_r(\sigma_s, \tilde{\sigma}, v_s, b_s, \Sigma, g, \kappa)\big|\,|\mathcal{F}_s\big] = o\big(\Delta_n^{1/6}\big),$$

and hence

$$\sqrt{u_n} \sum_{i=0}^{[T/3du_n]-1} \mathbb{E}\big[\big|\Gamma_r\big(\sigma_i^{n,\kappa}, \tilde{\sigma}, v_i^{n,\kappa}, b_i^{n,\kappa}, \Sigma^n, g^n, \kappa\big) - \Gamma_r\big(\sigma_i^{n,\kappa}, \tilde{\sigma}, v_i^{n,\kappa}, b_i^{n,\kappa}, \Sigma, g, \kappa\big)\big|\,|\mathcal{F}_i^{n,\kappa}\big]$$

$$= o\bigg(\frac{\Delta_n^{1/6}}{\sqrt{u_n}}\bigg) = o(1),$$

which implies (6.29). □

The next lemma deals with the stable convergence at (6.28).

Lemma 6.8. *Assume Assumptions (A1), (E). Let $r \in \{0, \ldots, d\}$ and g, h be two weight function satisfying the conditions of Proposition 4.9(ii). Then, on $\Omega_T^{\leq r}$, we have the stable convergence*

$$U''(r, g, h)_T^n \xrightarrow{d_{\mathrm{st}}} \mathcal{MN}\big(0, V(r, g, h)_T\big),$$

where $\mathcal{MN}(0, V(r, g, h)_T)$ is defined after equation (4.20). The two-dimensional statistic $U''(r, g, h)_T^n = (U''(r, g, h)_T^{n,1}, U''(r, g, h)_T^{n,2})$ is given after (6.28).

Proof. We apply a simplified version of Theorem IX.7.28 in [10]. To this end, we introduce the two-dimensional variables ξ_i^n with components

$$\xi_i^{n,1} = 3d\sqrt{u_n}\big(\gamma_r\big(A(g)_i^{n,1}, B(g)_i^{n,1}\big)^2 - \Gamma_r\big(\sigma_i^{n,1}, \tilde{\sigma}, v_i^{n,1}, b_i^{n,1}, \Sigma^n, g^n, 1\big)\big),$$

$$\xi_i^{n,2} = 3d\sqrt{u_n}\big(\gamma_r\big(A(h)_i^{n,2}, B(h)_i^{n,2}\big)^2 - \Gamma_r\big(\sigma_i^{n,2}, \tilde{\sigma}, v_i^{n,2}, b_i^{n,2}, \Sigma^n, h^n, 2\big)\big).$$

We must prove the following five statements where $\kappa, \kappa' = 1, 2$:

$$\sum_{i=0}^{[T/3du_n]-1} \mathbb{E}\big[\xi_i^{n,\kappa}|\mathcal{F}_i^{n,1}\big] \xrightarrow{\mathbb{P}} 0, \tag{6.30}$$

$$\sum_{i=0}^{[T/3du_n]-1} \mathbb{E}\big[\xi_i^{n,\kappa}\xi_i^{n,\kappa'}|\mathcal{F}_i^{n,1}\big] \xrightarrow{\mathbb{P}} V(r, g, h)_T^{\kappa,\kappa'}, \tag{6.31}$$

$$\sum_{i=0}^{[T/3du_n]-1} \mathbb{E}\big[\xi_i^{n,\kappa}\big(W_{3(i+1)du_n}^m - W_{3idu_n}^m\big)|\mathcal{F}_i^{n,1}\big] \xrightarrow{\mathbb{P}} 0, \tag{6.32}$$

$$\sum_{i=0}^{[T/3du_n]-1} \mathbb{E}\big[\|\xi_i^n\|^2 \mathbf{1}_{\{\|\xi_i^n\|>\delta\}}|\mathcal{F}_i^{n,1}\big] \xrightarrow{\mathbb{P}} 0 \qquad \forall \delta > 0, \tag{6.33}$$

$$\sum_{i=0}^{[T/3du_n]-1} \mathbb{E}\big[\xi_i^{n,\kappa}(N_{3(i+1)du_n} - N_{3idu_n})|\mathcal{F}_i^{n,1}\big] \xrightarrow{\mathbb{P}} 0, \tag{6.34}$$

where W^m is any of the components of W and N is a one-dimensional bounded martingale, orthogonal to (W, W') in the sense that the covariation between N and W^m, as well as the covariation between N and W'^m vanishes. We will later specify the conditions on N. If (6.30) to (6.34) hold, then Theorem IX.7.28 in [10] yields that

$$U''(r, g, h)_T^n \xrightarrow{d_{st}} \mathcal{U}''(r, g, h)_T,$$

where the random variable $\mathcal{U}''(r, g, h)_T$ is defined on an extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ of the original probability space $(\Omega, \mathcal{F}, \mathbb{P})$. It can be realized as

$$\mathcal{U}''(r, g, h)_T = \int_0^T y_s \, dW'_s + \int_0^T z_s \, d\widetilde{W}_s, \tag{6.35}$$

where $\widetilde{W}$ is a d-dimensional Brownian motion independent of $\mathcal{F}$, and – for fixed $\widetilde{\sigma}, \Sigma, g, h - y$ and z are càdlàg processes with values in $\mathbb{R}^{d \times d}$ which are adapted to the filtration generated by σ, b, v. Moreover, y and z can be characterized by

$$\sum_{i=0}^{[T/3du_n]-1} \mathbb{E}\big[\xi_i^{n,\kappa}\big(W'^m_{3(i+1)du_n} - W'^m_{3idu_n}\big)|\mathcal{F}_i^{n,1}\big] \xrightarrow{\mathbb{P}} \int_0^T y_s^m \, ds,$$

and

$$V(r, g, h)_T = \int_0^T \big(y_s y_s^\star + z_s z_s^\star\big) \, ds.$$

Since $\widetilde{W}$ and W' are independent of $\mathcal{G}$ and y, z are $\mathcal{G}$-measurable, (6.35) yields that $\mathcal{U}''(r, g, h)_T$ is mixed normal with $\mathcal{G}$-conditional mean 0 and $\mathcal{G}$-conditional covariance $V(r, g, h)_T$. Now, we turn to the proof of (6.30) to (6.34).

(i) We use equation (6.13) to derive that $\mathbb{E}[\xi_i^{n,\kappa}|\mathcal{F}_i^{n,\kappa}] = 0$ for $\kappa = 1, 2$. Using the nesting property $\mathcal{F}_i^{n,1} \subseteq \mathcal{F}_i^{n,2}$ and the tower property, we immediately obtain (6.30).

(ii) With equation (6.14) one can show that

$$\mathbb{E}\big[\xi_i^{n,1}\xi_i^{n,1}|\mathcal{F}_i^{n,1}\big] = 9d^2 u_n \Gamma_r'\big(\sigma_i^{n,1}, \widetilde{\sigma}, v_i^{n,1}, b_i^{n,1}, \Sigma^n, g^n, 1\big),$$

$$\mathbb{E}\big[\xi_i^{n,2}\xi_i^{n,2}|\mathcal{F}_i^{n,2}\big] = 9d^2 u_n \Gamma_r'\big(\sigma_i^{n,2}, \widetilde{\sigma}, v_i^{n,2}, b_i^{n,2}, \Sigma^n, h^n, 2\big).$$

Now, we have to carefully evaluate the term $\mathbb{E}[\Gamma_r'(\sigma_i^{n,2}, \tilde{\sigma}, v_i^{n,2}, b_i^{n,2}, \Sigma^n, h^n, 2)|\mathcal{F}_i^{n,1}]$. Recall (6.10) which implies that

$$\sup_{V=\sigma,v,b} \mathbb{E}\big[\|V_i^{n,2} - V_i^{n,1}\| |\mathcal{F}_i^{n,1}\big] \le K\sqrt{u_n}.$$

Using the multilinearity property of the determinant and the fact that Γ_r' consists of determinants to the power four, we end up with

$$\mathbb{E}\big[\Gamma_r'\big(\sigma_i^{n,2}, \tilde{\sigma}, v_i^{n,2}, b_i^{n,2}, \Sigma^n, h^n, 2\big)|\mathcal{F}_i^{n,1}\big]$$
$$= \mathbb{E}\big[\Gamma_r'\big(\sigma_i^{n,1} + \big(\sigma_i^{n,2} - \sigma_i^{n,1}\big), \tilde{\sigma}, v_i^{n,1} + \big(v_i^{n,2} - v_i^{n,1}\big), b_i^{n,1} + \big(b_i^{n,2} - b_i^{n,1}\big), \Sigma^n, h^n, 2\big)|\mathcal{F}_i^{n,1}\big]$$
$$= \Gamma_r'\big(\sigma_i^{n,1}, \tilde{\sigma}, v_i^{n,1}, b_i^{n,1}, \Sigma^n, h^n, 2\big) + O\big(u_n^2\big).$$

Hence,

$$\sum_{i=0}^{[T/3du_n]-1} \big(\mathbb{E}\big[\xi_i^{n,2}\xi_i^{n,2}|\mathcal{F}_i^{n,1}\big] - \Gamma_r'\big(\sigma_i^{n,1}, \tilde{\sigma}, v_i^{n,1}, b_i^{n,1}, \Sigma^n, h^n, 2\big)\big) \xrightarrow{\mathbb{P}} 0.$$

Since $\xi_i^{n,1}$ is $\mathcal{F}_i^{n,2}$-measurable, $\mathcal{F}_i^{n,1} \subseteq \mathcal{F}_i^{n,2}$ and $\mathbb{E}[\xi_i^{n,2}|\mathcal{F}_i^{n,2}] = 0$, we can deduce that $\mathbb{E}[\xi_i^{n,1}\xi_i^{n,2}|\mathcal{F}_i^{n,1}] = 0$. It follows along the lines of the proof of Lemma 6.7 that

$$9d^2 u_n \sum_{i=0}^{[T/3du_n]-1} \big(\Gamma_r'\big(\sigma_i^{n,1}, \tilde{\sigma}, v_i^{n,1}, b_i^{n,1}, \Sigma^n, g^n, 1\big) - \Gamma_r'\big(\sigma_i^{n,1}, \tilde{\sigma}, v_i^{n,1}, b_i^{n,1}, \Sigma, g, 1\big)\big) \xrightarrow{\mathbb{P}} 0,$$

$$9d^2 u_n \sum_{i=0}^{[T/3du_n]-1} \big(\Gamma_r'\big(\sigma_i^{n,1}, \tilde{\sigma}, v_i^{n,1}, b_i^{n,1}, \Sigma^n, h^n, 2\big) - \Gamma_r'\big(\sigma_i^{n,1}, \tilde{\sigma}, v_i^{n,1}, b_i^{n,1}, \Sigma, h, 2\big)\big) \xrightarrow{\mathbb{P}} 0.$$

By a Riemann approximation argument similar to the one used to show (6.25), one can deduce that

$$9d^2 u_n \sum_{i=0}^{[T/3du_n]-1} \Gamma_r'\big(\sigma_i^{n,1}, \tilde{\sigma}, v_i^{n,1}, b_i^{n,1}, \Sigma, g, 1\big) - 3d \int_0^T \Gamma_r'(\sigma_s, \tilde{\sigma}, v_s, b_s, \Sigma, g, 1)\, ds \xrightarrow{\mathbb{P}} 0,$$

$$9d^2 u_n \sum_{i=0}^{[T/3du_n]-1} \Gamma_r'\big(\sigma_i^{n,1}, \tilde{\sigma}, v_i^{n,1}, b_i^{n,1}, \Sigma, h, 2\big) - 3d \int_0^T \Gamma_r'(\sigma_s, \tilde{\sigma}, v_s, b_s, \Sigma, h, 2)\, ds \xrightarrow{\mathbb{P}} 0$$

which gives (6.31).

(iii) We will show (6.32) by proving that

$$\mathbb{E}\big[\xi_i^{n,1}\big(W_{3(i+1)du_n}^m - W_{3idu_n}^m\big)|\mathcal{H}_i^{n,1} \vee \sigma(W')\big] = 0, \tag{6.36}$$

$$\mathbb{E}\big[\xi_i^{n,2}\big(W_{3(i+1)du_n}^m - W_{(3i+1)du_n}^m\big)|\mathcal{H}_i^{n,2} \vee \sigma(W')\big] = 0. \tag{6.37}$$

Indeed, for $\kappa = 1$, (6.36) directly implies (6.32). For $\kappa = 2$, we use the relationship

$$\mathbb{E}\big[\xi_i^{n,1}\big(W_{3(i+1)du_n}^m - W_{3idu_n}^m\big)|\mathcal{F}_i^{n,1}\big]$$

$$= \mathbb{E}\big[\big(W_{(3i+1)du_n}^m - W_{3idu_n}^m\big)\mathbb{E}\big[\xi_i^{n,2}|\mathcal{F}_i^{n,2}\big]|\mathcal{F}_i^{n,1}\big]$$

$$+ \mathbb{E}\big[\mathbb{E}\big[\big(W_{3(i+1)du_n}^m - W_{(3i+1)du_n}^m\big)\xi_i^{n,2}|\mathcal{H}_i^{n,2} \vee \sigma(W')\big]|\mathcal{F}_i^{n,1}\big].$$

Since $\mathbb{E}[\xi_i^{n,2}|\mathcal{F}_i^{n,2}] = 0$, showing (6.37) implies (6.32) in this case. Similar to the proof of Lemma 6.5 one can write $\xi_i^{n,\kappa}$ as function of the form

$$\Phi\big(\omega, \big(W(\omega)_{(3i+\kappa-1)du_n+t} - W(\omega)_{(3i+\kappa-1)du_n}\big)_{t\geq 0}\big),$$

where Φ is a $(\mathcal{H}_i^{n,\kappa} \vee \sigma(W')) \otimes \mathcal{C}^q$-measurable function on $\Omega \times C(\mathbb{R}_+, \mathbb{R}^q)$. We have already seen that $A(g)_{i,j}^{n,\kappa}$ and $B(g)_{i,\kappa}^{n,1}$ can also be considered as function of the form (6.4) where $A(g)_{i,j}^{n,\kappa}$ is an odd function and $B(g)_{i,j}^{n,\kappa}$ is an even function. Since $\xi_i^{n,\kappa}$ consists of squared determinants, the function Φ in (6.4) is always even in the sense that $\Phi(\omega, (x_1, \ldots, x_q)) = \Phi(\omega, -(x_1, \ldots, x_q))$, no matter if r is even or odd. Consequently the map

$$(x_1, \ldots, x_q) \mapsto x_m \Phi\big(\omega, (x_1, \ldots, x_q)\big)$$

is odd such that (6.36), (6.37) follow by a standard argument.

(iv) Lemma 6.2 implies that $\mathbb{E}[\|\xi_i^n\|^4|\mathcal{F}_i^{n,1}] \leq Ku_n^2$, such that (6.33) follows by a standard argument.

(v) The proof of (6.34) is somewhat more involved than the previous steps. First, we introduce two filtrations: $(\mathcal{F}_t^{(0)})_{t\in[0,T]}$ which is generated by all processes appearing in Assumption (A) plus the Brownian motion W'. In contrast, the filtration $(\mathcal{F}_t^{(1)})_{t\in[0,T]}$ is generated by the noise process ε only. Note that due to Assumption (E), $\mathcal{F}_t^{(0)}$ and $\mathcal{F}_t^{(1)}$ are independent. Following the proof of [7], Lemma 5.7, it is sufficient to show (6.34) for all one-dimensional bounded martingales in a set $\mathcal{N} = \mathcal{N}^0 \cup \mathcal{N}^1$. Here, $\mathcal{N}^0$ consists of all $(\mathcal{F}_t^{(0)})$-martingales which are orthogonal to (W, W'). The set $\mathcal{N}^1$ comprises all $(\mathcal{F}_t^{(1)})$-Lévy-martingales N, such that there exists an integer $m \geq 1$, time points $0 \leq t_1 < \cdots < t_m \leq T$ and a bounded Borel-function $\tilde{f} : (\mathbb{R}^d)^m \to \mathbb{R}$ with the relation

$$N_t = \mathbb{E}\big[N_\infty|\mathcal{F}_t^{(1)}\big], \qquad N_\infty = \tilde{f}(\varepsilon_{t_1}, \ldots, \varepsilon_{t_m}). \tag{6.38}$$

Let $N \in \mathcal{N}^0$. With a similar argumentation like in point (iii), (6.34) follows by proving that

$$\mathbb{E}\big[\xi_i^{n,\kappa}(N_{3(i+1)du_n} - N_{(3i+\kappa-1)du_n})|\mathcal{H}_i^{n,\kappa}\big] = 0. \tag{6.39}$$

By assumption, N is independent of ε so N is also orthogonal to (W, W') conditionally on $\mathcal{H}_i^{n,\kappa}$. The variable $\xi_i^{n,\kappa}$ can be considered as a $\mathcal{H}_i^{n,\kappa} \otimes \mathcal{C}^q \otimes \mathcal{C}^d$-measurable function on $\Omega \times C(\mathbb{R}_+, \mathbb{R}^q) \times C(\mathbb{R}_+, \mathbb{R}^d)$ of the form

$$\Phi\big(\omega, \big(W(\omega)_{(3i+\kappa-1)du_n+t} - W(\omega)_{(3i+\kappa-1)du_n}\big)_{t\geq 0},$$

$$\big(W'(\omega)_{(3i+\kappa-1)du_n+t} - W'(\omega)_{(3i+\kappa-1)du_n}\big)_{t\geq 0}\big).$$

By virtue of the representation theorem (see [13], Proposition V.3.2), we can – conditionally on $\mathcal{H}_i^{n,\kappa}$ – write Φ as the sum of a constant and a stochastic integral over the interval $((3i + \kappa - 1)du_n, 3(i+1)du_n]$ with respect to (W, W') for a suitable $(q + d)$-dimensional predictable integrand. Then, thanks to the Itô-isometry and the fact that the covariation of N and any component of (W, W') vanishes, one ends up with (6.39).

Now, let $N \in \mathcal{N}^1$ with the representation (6.38). If $\{t_1, \ldots, t_m\} \cap (3idu_n, 3(i+1)du_n] = \varnothing$, then $\xi_i^{n,\kappa}$ and $(N_{3(i+1)du_n} - N_{3idu_n})$ are independent conditionally on $\mathcal{F}_i^{n,1}$, so we obtain that $\mathbb{E}[\xi_i^{n,\kappa}(N_{3(i+1)du_n} - N_{3idu_n})|\mathcal{F}_i^{n,1}] = 0$. If $\{t_1, \ldots, t_m\} \cap (3idu_n, 3(i+1)du_n] \neq \varnothing$, the fact that $\tilde{f}$ is bounded plus Lemma 6.2 imply that

$$\mathbb{E}\big[\big|\xi_i^{n,\kappa}(N_{3(i+1)du_n} - N_{3idu_n})\big|\big|\mathcal{F}_i^{n,1}\big] \leq K\sqrt{u_n}.$$

Since the intervals $(3idu_n, 3(i+1)du_n]$ are disjoint for different i, the number of such intervals having a non-empty intersection with $\{t_1, \ldots, t_m\}$ is bounded by m. Consequently, we end up with

$$\sum_{i=0}^{[T/3du_n]-1} \mathbb{E}\big[\big|\xi_i^{n,\kappa}(N_{3(i+1)du_n} - N_{3idu_n})\big|\big|\mathcal{F}_i^{n,1}\big] \leq mK\sqrt{u_n},$$

which gives us (6.34). This completes the proof of Lemma 6.8 and therefore the proof of Theorem 4.13. $\qquad\square$

The proof of Proposition 4.14 is somewhat simpler in comparison to the proof of Theorem 4.13. Regarding Lemma 6.5, part (i) of Proposition 4.14 follows by showing the following lemma.

Lemma 6.9. *Assume Assumptions* (A1), (E). *Let* $r \in \{0, \ldots, d\}$, $\kappa, \kappa' = 1, 2$ *and* g, h *be any weight functions. Then, on* $\Omega_T^{\leq r}$, *we have that*

$$
V(r, g, h)_T^{n,\kappa\kappa'}
$$

$$
\xrightarrow{\mathbb{P}}
\begin{cases}
3d \displaystyle\int_0^T \Gamma_r'(\sigma_s, \tilde{\sigma}, v_s, b_s, \Sigma, g, 1) + \Gamma_r(\sigma_s, \tilde{\sigma}, v_s, b_s, \Sigma, g, 1)^2 \, ds, \\
\quad \text{if } \kappa = \kappa' = 1, \\
3d \displaystyle\int_0^T \Gamma_r'(\sigma_s, \tilde{\sigma}, v_s, b_s, \Sigma, h, 2) + \Gamma_r(\sigma_s, \tilde{\sigma}, v_s, b_s, \Sigma, h, 2)^2 \, ds, \\
\quad \text{if } \kappa = \kappa' = 2, \\
3d \displaystyle\int_0^T \Gamma_r(\sigma_s, \tilde{\sigma}, v_s, b_s, \Sigma, g, 1)\Gamma_r(\sigma_s, \tilde{\sigma}, v_s, b_s, \Sigma, h, 2) \, ds, \\
\quad \text{if } \kappa = 1, \kappa' = 2.
\end{cases}
\tag{6.40}
$$

Proof. Define the variables

$$\rho(g,h)_i^{n,\kappa\kappa'} = \begin{cases} \gamma_r\big(A(g)_i^{n,1}, B(g)_i^{n,1}\big)^4, & \text{if } \kappa = \kappa' = 1, \\ \gamma_r\big(A(h)_i^{n,2}, B(h)_i^{n,2}\big)^4, & \text{if } \kappa = \kappa' = 2, \\ \gamma_r\big(A(g)_i^{n,1}, B(g)_i^{n,1}\big)^2 \gamma_r\big(A(h)_i^{n,2}, B(h)_i^{n,2}\big)^2, & \text{if } \kappa = 1, \kappa' = 2, \end{cases}$$

which is the ith summand in the right-hand side of (6.18). Define the variables

$$\rho'(g,h)_i^{n,\kappa\kappa'} = \mathbb{E}\big[\rho(g,h)_i^{n,\kappa\kappa'}|\mathcal{F}_i^{n,1}\big], \qquad \rho''(g,h)_i^{n,\kappa\kappa'} = \rho(g,h)_i^{n,\kappa\kappa'} - \rho'(g,h)_i^{n,\kappa\kappa'}.$$

Using Lemma 6.4, we get that

$$\rho'(g)_i^{n,\kappa} = \Gamma_r'\big(\sigma_i^{n,\kappa}, \widetilde{\sigma}, v_i^{n,\kappa}, b_i^{n,\kappa}, \Sigma^n, g^n, \kappa\big) + \Gamma_r\big(\sigma_i^{n,\kappa}, \widetilde{\sigma}, v_i^{n,\kappa}, b_i^{n,\kappa}, \Sigma^n, g^n, \kappa\big)^2.$$

Just as in the proof of (6.31) we can deduce that $9d^2 u_n \sum_{i=0}^{[T/3du_n]-1} \rho'(g,h)_i^{n,\kappa\kappa'}$ converges in probability to the right-hand side of (6.40). By construction, the sequence $(\rho''(g,h)_i^{n,\kappa\kappa'})_{i\geq 0}$ is a $(\mathcal{F}_i^{n,1})$-martingale. Hence, we can use Doob's inequality and a calculation similar to the one in (6.21) to end up with

$$9d^2 u_n \sum_{i=0}^{[T/3du_n]-1} \rho''(g,h)_i^{n,\kappa\kappa'} \xrightarrow{\mathbb{P}} 0,$$

which completes the proof of (6.40). $\square$

Part (ii) of Proposition 4.14 essentially follows by the next lemma.

Lemma 6.10. *Assume Assumptions* (A1), (E). *Let* $r \in \{0,\ldots,d\}$ *and* g, h *be two weight function satisfying the conditions of Proposition* 4.9(ii). *Then, on* Ω_T^r, *we have that*

$$\frac{\widehat{R}(g,h)_T^n - r}{\sqrt{u_n}} - \frac{U(r,g,h)_T^{n,1} - U(r,g,h)_T^{n,2}}{\log 2 S(r,g)_T^1} \xrightarrow{\mathbb{P}} 0. \qquad (6.41)$$

Proof. Using the fact that $S(r,g)_T^1 = S(r,h)_T^2$, we obtain by an elementary calculation that, on Ω_T^r,

$$\widehat{R}(g,h)_T^n - r = \frac{\log(1 + \sqrt{u_n}U(r,g,h)_T^{n,1}/S(r,g)_T^1) - \log(1 + \sqrt{u_n}U(r,g,h)_T^{n,2}/S(r,g)_T^1)}{\log 2}.$$

Due to (4.19) the sequence $U(r,g,h)_T^n$ is tight. By a Taylor expansion, one obtains that $\log(1 + x) = x + O(x^2)$ for $|x| < 1$, so we get (for n sufficiently large)

$$\log\big(1 + \sqrt{u_n}U(r,g,h)_T^{n,\kappa}/S(r,g)_T^1\big) = \sqrt{u_n}U(r,g,h)_T^{n,\kappa}/S(r,g)_T^1 + O_{\mathbb{P}}(u_n).$$

This readily implies (6.41). $\square$

The continuous mapping theorem for stable convergence then implies that, on Ω_T^r,

$$\frac{U(r, g, h)_T^{n,1} - U(r, g, h)_T^{n,2}}{\log 2 S(r, g)_T^1} \xrightarrow{d_{\mathrm{st}}} \frac{\mathcal{U}''(r, g, h)_T^1 - \mathcal{U}''(r, g, h)_T^2}{\log 2 S(r, g)_T^1}, \tag{6.42}$$

where $\mathcal{U}''(r, g, h)_T$ is the limit in (4.19) (see also equation (6.35)). The right-hand side of (6.42) is mixed normal with $\mathcal{G}$-conditional mean 0 and $\mathcal{G}$-conditional variance

$$\frac{3d \int_0^T \Gamma_r'(\sigma_s, \widetilde{\sigma}, v_s, b_s, \Sigma, g, 1)\, ds + 3d \int_0^T \Gamma_r'(\sigma_s, \widetilde{\sigma}, v_s, b_s, \Sigma, h, 2)\, ds}{(S(r, g)_T^1 \log 2)^2} > 0.$$

The positivity of the variance is a consequence of (4.13) in Lemma 4.7. At this stage, (4.28) follows by part (i) of Proposition 4.14, Theorem 4.8 and the delta method for stable convergence.

6.5. Proof of Corollary 4.16

The implication at (4.32) is a direct consequence of the stable convergence at (4.28). To prove the consistency at (4.33), it is sufficient to show that for any $r' \neq r$ we have that

$$\mathbb{P}\big(C_\alpha^{n,=r} \cap \Omega_T^{r'}\big) \to \mathbb{P}\big(\Omega_T^{r'}\big).$$

Let Φ be the right-hand side of (4.28). Then we have by Proposition 4.14(ii) that

$$\mathbb{P}\big(C_\alpha^{n,=r} \cap \Omega_T^{r'}\big) - \widetilde{\mathbb{P}}\left(\left\{\left|\Phi + \frac{r'-r}{\sqrt{u_n V(n, T, g, h)}}\right| > z_{1-\alpha/2}\right\} \cap \Omega_T^{r'}\right) \to 0.$$

By Proposition 4.14, Theorem 4.8 and Lemma 4.7, $V(n, T, g, h)$ converges in probability to a positive-valued limit, such that $u_n V(n, T, g, h) \xrightarrow{\mathbb{P}} 0$ and hence

$$\widetilde{\mathbb{P}}\left(\left\{\left|\Phi + \frac{r'-r}{\sqrt{u_n V(n, T, g, h)}}\right| > z_{1-\alpha/2}\right\} \cap \Omega_T^{r'}\right) \to \mathbb{P}\big(\Omega_T^{r'}\big),$$

which shows (4.33). To show (4.34), let $A \subset \Omega_T^{\leq r}$ with $\mathbb{P}(A) > 0$. Then we obtain

$$\mathbb{P}\big(C_\alpha^{n,\leq r} | A\big) = \sum_{r' \leq r} \mathbb{P}\big(C_\alpha^{n,\leq r} \cap \Omega_T^{r'} | A\big) \leq \sum_{r' \leq r} \mathbb{P}\big(C_\alpha^{n,\leq r'} \cap \Omega_T^{r'} | A\big)$$

$$\to \sum_{r' \leq r} \widetilde{\mathbb{P}}\big(\{\Phi > z_{a-\alpha}\} \cap \Omega_T^{r'} | A\big) = \alpha \mathbb{P}\big(\Omega_T^{\leq r} | A\big) = \alpha.$$

We essentially used the convergence at (4.28) as well as the fact that Φ is independent of $\mathcal{G}$ and $\Omega_T^{r'} \in \mathcal{G}$. The consistency result at (4.35) follows in the same manner as (4.33).

Acknowledgement

References

[1] Aldous, D.J. and Eagleson, G.K. (1978). On mixing and stability of limit theorems. *Ann. Probab.* **6** 325–331. MR0517416

[2] Barndorff-Nielsen, O.E., Graversen, S.E., Jacod, J., Podolskij, M. and Shephard, N. (2006). A central limit theorem for realised power and bipower variations of continuous semimartingales. In *From Stochastic Calculus to Mathematical Finance* 33–68. Berlin: Springer. MR2233534

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

[4] Delbaen, F. and Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing. *Math. Ann.* **300** 463–520. MR1304434

[5] Hautsch, N. and Podolskij, M. (2013). Preaveraging-based estimation of quadratic variation in the presence of noise and jumps: Theory, implementation, and empirical evidence. *J. Bus. Econom. Statist.* **31** 165–183. MR3055330

[6] Jacod, J., Lejay, A. and Talay, D. (2008). Estimation of the Brownian dimension of a continuous Itô process. *Bernoulli* **14** 469–498. MR2544098

[7] Jacod, J., Li, Y., Mykland, P.A., Podolskij, M. and Vetter, M. (2009). Microstructure noise in the continuous case: The pre-averaging approach. *Stochastic Process. Appl.* **119** 2249–2276. MR2531091

[8] Jacod, J. and Podolskij, M. (2013). A test for the rank of the volatility process: The random perturbation approach. *Ann. Statist.* **41** 2391–2427. MR3127870

[9] Jacod, J. and Protter, P. (2012). *Discretization of Processes. Stochastic Modelling and Applied Probability* **67**. Heidelberg: Springer. MR2859096

[10] Jacod, J. and Shiryaev, A.N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **288**. Berlin: Springer. MR1943877

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

[12] Rényi, A. (1963). On stable sequences of events. *Sankhyā Ser. A* **25** 293–302. MR0170385

[13] Revuz, D. and Yor, M. (2005). *Continuous Martingales and Brownian Motion*, 3rd corrected ed. Berlin: Springer.

[14] Zhang, L., Mykland, P.A. and Aït-Sahalia, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *J. Amer. Statist. Assoc.* **100** 1394–1411. MR2236450