Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data

Kim Christensen [a,*], Silja Kinnebrock [b], Mark Podolskij [a,c]

[a] CREATES, School of Economics and Management, Aarhus University, Building 1322, Bartholins Allé 10, 8000 Aarhus, Denmark
[b] Oxford-Man Institute of Quantitative Finance, University of Oxford, Eagle House, Walton Well Road, Oxford OX2 6ED, United Kingdom
[c] ETH Zürich, Department of Mathematics, Rämistrasse 101, CH-8092 Zürich, Switzerland

ARTICLE INFO

ABSTRACT

We show how pre-averaging can be applied to the problem of measuring the ex-post covariance of financial asset returns under microstructure noise and non-synchronous trading. A pre-averaged realised covariance is proposed, and we present an asymptotic theory for this new estimator, which can be configured to possess an optimal convergence rate or to ensure positive semi-definite covariance matrix estimates. We also derive a noise-robust Hayashi–Yoshida estimator that can be implemented on the original data without prior alignment of prices. We uncover the finite sample properties of our estimators with simulations and illustrate their practical use on high-frequency equity data.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

The theory of financial economics is often cast in multivariate settings, where the covariance structure of assets plays a key role to the solution of fundamental economic problems, such as optimal asset allocation and risk management. In recent years, a broader access to financial high-frequency data has improved our ability to accurately estimate and draw inference about financial covariation. The underlying idea is to use quadratic covariation, which we can estimate using realised covariance, as an ex-post measure, whose increments can be studied to learn about the properties of the true asset return covariation (e.g. Andersen et al., 2003; Barndorff-Nielsen and Shephard, 2004).

In practice, implementing realised covariance is hampered by two empirical phenomena, namely the presence of market microstructure noise (e.g., price discreteness or bid–ask spread bounce) and non-synchronous trading. The impact of microstructure noise has received much attention in the univariate setting, where its effect on the realised variance has been well documented. This builds on previous work in the noiseless case, including Andersen et al. (2001), Barndorff-Nielsen and Shephard (2002) or Mykland and Zhang (2006, 2009). A key to understanding the nature of the noise and a possible tool of how to deal with it is that microstructure noise induces autocorrelation in high-frequency returns and this leads to a bias problem (see, e.g., Zhou, 1996; Aït-Sahalia et al., 2005; Hansen and Lunde, 2006). Currently, there are three main univariate approaches, where the damage caused by the noise is explicitly fixed: the two-scale subsampler proposed by Zhang et al. (2005) or its multi-scale version of Zhang (2006), the realised kernel introduced in Barndorff-Nielsen et al. (2008a), which relies on autocovariance-based corrections, and finally the pre-averaging estimator of Podolskij and Vetter (2009) and Jacod et al. (2009).

The multivariate version of this problem is, however, more complicated in that not only does the estimator need to be robust against various types of noise, it also has to cope with non-synchronous trading (see, e.g., Fisher, 1966). Asynchronicity

* Corresponding author. Tel.: +45 8942 1550.
E-mail addresses: kchristensen@creates.au.dk (K. Christensen), silja.kinnebrock@merton.oxon.org (S. Kinnebrock), mark.podolskij@math.ethz.ch (M. Podolskij).

causes high-frequency covariance estimates to be biased towards zero as the sampling frequency increases. This feature of the data, the so-called Epps effect, was highlighted by Epps (1979). Intuitively, as the sampling frequency is increased, there are more and more zero-returns in the presence of non-synchronous trading, and this will dominate realised covariance and related statistics (e.g. realised correlation). Hayashi and Yoshida (2005) introduced an estimator, which is capable of dealing with non-synchronous data, but not with market microstructure noise. More recently, Zhang (2008) extended the two-scale RV to integrated covariance estimation in the simultaneous presence of noise and non-synchronicity, while in concurrent and independent work (Barndorff-Nielsen et al., 2008b) proposed a multivariate realised kernel. Additional work in this growing line of research includes Malliavin and Mancino (2002), Martens (2003), Renò (2003), Bandi and Russell (2005), Griffin and Oomen (2006), Large (2007), Voev and Lunde (2007) and Boudt et al. (2008), among others.

In this paper, we propose to use a "modulated" realised covariance (MRC) to estimate the ex-post integrated covariance. The econometric technique employed here for dealing with microstructure noise relies on rather simple pre-averaging of the high-frequency data, which makes the estimator both intuitive to understand and trivial to implement. It relates to previous work in the univariate case, where pre-averaging has been suggested in Podolskij and Vetter (2009) and Jacod et al. (2009). The current article draws ideas from these papers, but the multivariate extension is challenging, as it faces the additional complexity of non-synchronous trading and requires that the resulting estimator be positive semi-definite.

The pre-averaging approach depends on a bandwidth parameter, or window length, that grows with the sample and dictates the amount of averaging to be carried out. In turn, the choice of this tuning parameter controls the influence of microstructure noise on the MRC and, hence, also its asymptotic properties. In the optimal case, called balanced pre-averaging, this leads to an efficient $n^{-1/4}$ rate of convergence, which is known to be the fastest attainable (see, e.g., Gloter and Jacod, 2001a,b). This baseline MRC estimator, however, needs a bias-correction to be consistent for the integrated covariance. As a result, it is not guaranteed to be positive semi-definite in finite samples, though our empirical work indicates this shortcoming is not too much of a concern for more recent data. Nonetheless, as we show in the paper, it is straightforward to design a positive semi-definite estimator by increasing the pre-averaging window length slightly, which can also serve to make the MRC robust against more general noise processes.

The MRC is, in all its essence, a realised covariance computed on the back of pre-averaged high-frequency returns. As such, it depends on receiving synchronous observations as input, which clashes with the irregular spacing of real high-frequency data. We propose two distinct ways in which pre-averaging can be applied in the context of non-synchronous trading. First, we use traditional imputation schemes to map asynchronous data onto a common time grid, for example using previous-tick or refresh time, where the latter approach has been used in Barndorff-Nielsen et al. (2008b). An MRC computed from such returns will be asymptotically robust to non-synchronous trading. Second, we extend the Hayashi and Yoshida (2005) estimator to the case of microstructure noise by using pre-averaging and show that it is consistent. This second estimator has the property that it can be implemented directly on the irregular, non-synchronous and noisy observations without any form of imputation. It therefore omits throwing away information in the sample and further avoids potential biases arising from artificially imputed returns.

An appealing feature of pre-averaging is that it is a general statistical tool that can be applied to many estimation problems.

This proves useful in our setting, because as usual the mixed Gaussian central limit theorems feature an unknown conditional covariance matrix. In practice, this must be robustly estimated from sample data in the presence of noise and non-synchronous trading to make the distributional results feasible, such that confidence bands for elements of the integrated covariance matrix can be constructed. We outline how this can be done based on pre-averaged high-frequency data.

The paper progresses as follows. In Section 2, we formulate the theoretical setup and define the MRC estimator. In Section 3, we first show consistency of the MRC based on balanced pre-averaging and then derive its asymptotic distribution. As discussed above, this estimator needs a bias-correction, so we carry on to study a modified MRC estimator, in which the degree of pre-averaging is increased. We also discuss the application of the MRC to non-synchronous data, show its relation to the multivariate realised kernel, and finally we derive a pre-averaged version of the Hayashi–Yoshida estimator. In Section 4, we propose an estimator of the conditional covariance matrix that appears in the central limit theorem of MRC, which can be used to transform infeasible limit results into feasible ones. In Section 5, the focus is shifted towards regression and correlation analysis. A simulation study is undertaken in Section 6 to uncover the finite sample properties of our estimators, while an empirical illustration is conducted in Section 7. Section 8 draws conclusions and presents some ideas for future work. The Appendix contains the derivations of all our theoretical results.

2. Theoretical setup

We consider a vector of log-prices X defined on a probability space $(\Omega^0, \mathcal{F}^0, P^0)$ and equipped with an information filtration $(\mathcal{F}^0_t)_{t \geq 0}$. X has dimension d—the number of assets under consideration.

A standard no-arbitrage condition suggests security prices must be semimartingales (see, e.g., Back, 1991; Delbaen and Schachermayer, 1994). These processes obey the fundamental theorem of asset pricing and, as a result, are used extensively to model the evolution of asset prices through time. In accordance with this, we model X as a semimartingale that follows the equation

$$X_t = X_0 + \int_0^t a_u \mathrm{d}u + \int_0^t \sigma_u \mathrm{d}W_u, \quad t \geq 0, \tag{1}$$

where $a = (a_t)_{t \geq 0}$ is a d-dimensional predictable locally bounded drift vector, $\sigma = (\sigma_t)_{t \geq 0}$ an adapted càdlàg $d \times d$ covolatility matrix and $W = (W_t)_{t \geq 0}$ is d-dimensional Brownian motion.

This model is a Brownian semimartingale, or stochastic volatility model with drift, which permeates financial economics (cf., Ghysels et al., 1996, for a review). We think of this construct as governing an underlying efficient price process — the price that would prevail in the absence of market frictions, which we then subject to microstructure noise.

Of importance to our analysis is the quadratic covariation process of X, which is defined as

$$[X]_t = p\text{-}\lim_{n \to \infty} \sum_{i=1}^n \left(X_{t_i} - X_{t_{i-1}} \right) \left(X_{t_i} - X_{t_{i-1}} \right)' \tag{2}$$

for any sequence of deterministic partitions $0 = t_0 < t_1 < \cdots < t_n = t$ with $\sup_i \{t_i - t_{i-1}\} \to 0$ for $n \to \infty$. In our setting, the quadratic covariation of X is given by

$$[X]_t = \int_0^t \Sigma_u \mathrm{d}u, \tag{3}$$

where $\Sigma = \sigma\sigma'$. The quadratic covariation is pivotal in financial economics (see, e.g., the reviews by Barndorff-Nielsen and Shephard, 2007; Andersen et al., 2010), and we thus take Eq. (3) as defining the target that we are interested in estimating. We note that for obvious reasons the matrix in Eq. (3) is also called the integrated covariance and both terms are used interchangeably.

Throughout the remainder of the paper, and without loss of generality, we restrict the clock t to evolve in the unit interval $[0, 1]$, which we think of as representing the passing of an economic event, for example a trading day.

2.1. Microstructure noise

In practice, market microstructure noise leads to a departure from the pure semimartingale model. Microstructure noise has many forms, including price discreteness and bid–ask spread bounce, which creates spurious variation in asset prices. As a result, we do not observe X from Eq. (1) in the market but a process Y, which is the efficient price distorted by noise. More precisely, we consider the process Y, observed at time points $i/n, i = 0, 1, \ldots, n$, which is given as

$$Y_t = X_t + \epsilon_t, \tag{4}$$

where (ϵ_t) is an i.i.d. process with $X \perp\!\!\!\perp \epsilon$ (the symbol $\perp\!\!\!\perp$ is used to denote stochastic independence).

The noise process can be constructed as follows. We define a second probability space $(\Omega^1, \mathcal{F}^1, (\mathcal{F}_t^1)_{t \geq 0}, P^1)$, where Ω^1 denotes $\mathbb{R}^{[0,1]}$ and $\mathcal{F}^1$ the product Borel-σ-field on Ω^1. Next, let Q be a probability measure on $\mathbb{R}$ (Q is the marginal distribution of ϵ). For any $t \in [0, 1]$, $P_t^1 = Q$ and P^1 denotes the product $\otimes_{t \in [0,1]} P_t^1$. The filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$, on which we define the process Y, is given as

$$\left. \begin{array}{l} \Omega = \Omega^0 \times \Omega^1, \quad \mathcal{F} = \mathcal{F}^0 \times \mathcal{F}^1, \quad \mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s^0 \times \mathcal{F}_s^1, \\ P = P^0 \otimes P^1. \end{array} \right\} \tag{5}$$

The multivariate noise process ϵ is assumed to satisfy:

$$\mathbb{E}(\epsilon_t) = 0, \qquad \mathbb{E}(\epsilon_t \epsilon_t') = \Psi, \tag{6}$$

where Ψ is a positive definite $d \times d$-matrix.

Remark. The empirical results found by Hansen and Lunde (2006) show that both the i.i.d. assumption on (ϵ_t) and the independence $X \perp\!\!\!\perp \epsilon$ can be called into question when sampling the data at very high frequencies, e.g., below the 1 min mark (see also Diebold and Strasser, 2008). Jacod et al. (2009) consider more general types of (one-dimensional) noise processes. Roughly speaking, they assume that the errors ϵ_t's are, conditionally on X, centered and independent. The asymptotic theory developed in this paper still holds true for the multivariate version of such noise processes, but we restrict attention to models of the form in Eq. (4) to ease the exposition.

2.2. Pre-averaging of high-frequency data

It is intuitive that under mean zero i.i.d. microstructure noise some form of smoothing of the observed log-price Y should tend to diminish the impact of the noise. Effectively, we are going to approximate X_t, X being a continuous function of t, by an average of observations of Y in a neighborhood of t, the noise being averaged away.

Here, we describe in more detail how to conduct the pre-averaging. In particular, we consider a sequence of integers, k_n, and a number $\theta \in (0, \infty)$ such that

$$\frac{k_n}{\sqrt{n}} = \theta + o\left(n^{-1/4}\right). \tag{7}$$

An example of this would be $k_n = \lfloor \theta \sqrt{n} \rfloor$.

We also choose a function g on $[0, 1]$, which is continuous, piecewise continuously differentiable with a piecewise Lipschitz derivative g' with $g(0) = g(1) = 0$ and which satisfies $\int_0^1 g^2(s)\,ds > 0$. Furthermore, we introduce the following functions and numbers that are associated with g:

$$\phi_1(s) = \int_s^1 g'(u)\, g'(u-s)\,du,$$

$$\phi_2(s) = \int_s^1 g(u)\, g(u-s)\,du,$$

$$\psi_1 = \phi_1(0), \qquad \psi_2 = \phi_2(0),$$

$$\Phi_{11} = \int_0^1 \phi_1^2(s)\,ds, \qquad \Phi_{12} = \int_0^1 \phi_1(s)\, \phi_2(s)\,ds,$$

$$\Phi_{22} = \int_0^1 \phi_2^2(s)\,ds.$$

The functions ϕ_1 and ϕ_2 are assumed to be 0 outside the interval $[0, 1]$.

Next, with any process $V = (V_t)_{t \geq 0}$ we associate the following random variables

$$\Delta_i^n V = V_{\frac{i}{n}} - V_{\frac{i-1}{n}}, \quad \text{for } i = 1, \ldots, n$$

$$\bar{V}_i^n = \sum_{j=1}^{k_n-1} g\left(\frac{j}{k_n}\right) \Delta_{i+j}^n V, \quad \text{for } i = 0, \ldots, n - k_n + 1.$$

Applying this notation to Y, it can be seen that $\Delta_i^n Y$ represents the noisy high-frequency returns, while $\bar{Y}_i^n$ is the pre-averaged return data, using the weight function g. It follows that the stochastic order of $\bar{Y}_i^n = \bar{X}_i^n + \bar{\epsilon}_i^n$ is controlled by the sequence k_n, since

$$\bar{X}_i^n = O_p\left(\sqrt{\frac{k_n}{n}}\right), \qquad \bar{\epsilon}_i^n = O_p\left(\sqrt{\frac{1}{k_n}}\right). \tag{8}$$

Thus, taking $k_n = O(\sqrt{n})$ implies that the orders of the two terms in Eq. (8) are equal, so that $\bar{Y}_i^n = O_p\left(n^{-1/4}\right)$. This is called balanced pre-averaging and delivers the best rate of convergence. As shown below, it is also useful to look at cases in which a higher order of k_n is chosen. This results in a suboptimal rate of convergence, but it has some potentially valuable side-effects on the robustness and finite sample properties of our estimator.

The pre-averaging window length, k_n, depends on the tuning parameter θ, which needs to be chosen by the user. We will later discuss how to sensibly make this choice.

2.3. Modulated realised covariance

The core statistic of this paper is the multivariate extension of the estimator, which was introduced in Jacod et al. (2009). We call it the modulated realised covariance (MRC) and define it as

$$MRC[Y]_n = \frac{n}{n - k_n + 2} \frac{1}{\psi_2 k_n} \sum_{i=0}^{n-k_n+1} \bar{Y}_i^n \left(\bar{Y}_i^n\right)'. \tag{9}$$

The factor $n/(n - k_n + 2)$ is a finite sample correction for the true number of summands in $MRC[Y]_n$ relative to the sample size n. It is sometimes left out in the presentation below, but it is always included in implementations on data.

Remark. The sum of outer products in Eq. (9) is a realised covariance based on pre-averaged data. To build some intuition for

our approach, we explain the usage of $\bar{Y}_i^n$ in more detail. Suppose k_n is an even number and write

$$\hat{Y}_i^n = \frac{2}{k_n} \sum_{j=0}^{k_n/2-1} Y_{\frac{i+j}{n}},$$

which is a simple average of Y over $k_n/2$ terms. Because of this pre-averaging, $\hat{Y}_i^n$ will be closer to the efficient price $X_{\frac{i}{n}}$. Next, we compute the realised covariation estimator based on these filtered increments by setting

$$\bar{Y}_i^n = \frac{1}{2}\left(\hat{Y}_{i+\frac{k_n}{2}}^n - \hat{Y}_i^n\right) = \frac{1}{k_n}\left(\sum_{j=k_n/2}^{k_n-1} Y_{\frac{i+j}{n}} - \sum_{j=0}^{k_n/2-1} Y_{\frac{i+j}{n}}\right).$$

(However, as we shall see this induces a bias, which is a function of Ψ.) This method was originally proposed by Podolskij and Vetter (2009) and using the above definition of $\bar{Y}_i^n$ corresponds to choosing the weight function

$$g(x) = \min(x, 1-x), \tag{10}$$

which is the most intuitive example. Here we explicitly give the numerical values of the asymptotic constants for this choice of function g, as it is the one used for all our simulations and empirical work:

$$\psi_1 = 1, \qquad \psi_2 = \frac{1}{12}, \qquad \Phi_{11} = \frac{1}{6},$$
$$\Phi_{12} = \frac{1}{96}, \qquad \Phi_{22} = \frac{151}{80,640}.$$

Remark. As noted in Jacod et al. (2009), to avoid biases in small samples we should replace the asymptotic constants and functions $\psi_1, \psi_2, \phi_1, \phi_2, \Phi_{11}, \Phi_{12}$, and Φ_{22} by their Riemann approximations:

$$\psi_1^{k_n} = k_n \sum_{i=1}^{k_n}\left(g\left(\frac{i}{k_n}\right) - g\left(\frac{i-1}{k_n}\right)\right)^2,$$

$$\psi_2^{k_n} = \frac{1}{k_n}\sum_{i=1}^{k_n-1} g^2\left(\frac{i}{k_n}\right),$$

$$\phi_1^{k_n}(j) = \sum_{i=j+1}^{k_n-1}\left(g\left(\frac{i-1}{k_n}\right) - g\left(\frac{i}{k_n}\right)\right)$$
$$\times \left(g\left(\frac{i-j-1}{k_n}\right) - g\left(\frac{i-j}{k_n}\right)\right),$$

$$\phi_2^{k_n}(j) = \sum_{i=j+1}^{k_n-1} g\left(\frac{i}{k_n}\right) - g\left(\frac{i-j}{k_n}\right),$$

$$\Phi_{11}^{k_n} = k_n\left(\sum_{j=0}^{k_n-1}\left(\phi_1^{k_n}(j)\right)^2 - \frac{1}{2}\left(\phi_1^{k_n}(0)\right)^2\right),$$

$$\Phi_{12}^{k_n} = \frac{1}{k_n}\left(\sum_{j=0}^{k_n-1}\phi_1^{k_n}(j)\,\phi_2^{k_n}(j) - \frac{1}{2}\phi_1^{k_n}(0)\,\phi_2^{k_n}(0)\right),$$

$$\Phi_{22}^{k_n} = \frac{1}{k_n^3}\left(\sum_{j=0}^{k_n-1}\left(\phi_2^{k_n}(j)\right)^2 - \frac{1}{2}\left(\phi_2^{k_n}(0)\right)^2\right).$$

These are the actual terms that appear in the computations of the MRC. Note that for all appropriate indices of i and j, $\psi_i^{k_n} \to \psi_i$, $\phi_i^{k_n} \to \phi_i$, $\Phi_{ij}^{k_n} \to \Phi_{ij}$ as $n \to \infty$ at smaller order than $n^{-1/4}$, so the finite sample versions can be replaced with the asymptotic constants in all the limit theorems given below, including the central limit theorems.

3. Asymptotic properties of MRC

3.1. Consistency

Our first result inspects the probability limit of $MRC\,[Y]_n$.

Theorem 1. *Assume that $\mathbb{E}\left(|\epsilon^j|^4\right) < \infty$ for all $j = 1, \ldots, d$ and (k_n, θ) satisfy Eq. (7). As $n \to \infty$, it holds that*

$$MRC\,[Y]_n \xrightarrow{p} \int_0^1 \Sigma_s ds + \frac{\psi_1}{\theta^2 \psi_2}\Psi. \tag{11}$$

Proof. See Appendix. □

A couple of points are worth highlighting. First, as Theorem 1 shows $MRC\,[Y]_n$ is consistent for $\int_0^1 \Sigma_s ds$ up to a bias-correction. The bias term depends on the unknown Ψ, which must be estimated from the data.

We set

$$\hat{\Psi}_n = \frac{1}{2n}\sum_{i=1}^{n} \Delta_i^n Y\left(\Delta_i^n Y\right)', \tag{12}$$

which is linked to the univariate estimator proposed by Aït-Sahalia et al. (2005) and Bandi and Russell (2006, 2008). Then, we obtain the convergence $\hat{\Psi}_n \xrightarrow{p} \Psi$, such that

$$MRC\,[Y]_n - \frac{\psi_1^{k_n}}{\theta^2 \psi_2^{k_n}}\hat{\Psi}_n \xrightarrow{p} \int_0^1 \Sigma_s ds.$$

Hence, for the remainder of the paper, we shall incorporate the bias-correction term into the definition of $MRC\,[Y]_n$.[1] In doing so, we are no longer ensured that $MRC\,[Y]_n$ is positive semi-definite in finite samples. To deal with this problem, we propose an alternative formulation of the MRC below, which uses a longer pre-averaging window k_n to avoid this step. Second, since $\hat{\Psi}_n$ is a $\sqrt{n}$-estimator of Ψ, it will not affect the CLT of $MRC\,[Y]_n$, as the latter converges at a slower rate. Third, our initial MRC estimator is based on synchronous data, assuming all components of the log-price vector Y are observed contemporaneously. We will later extend the MRC to the non-synchronous setting.

3.2. The central limit theorem

We proceed with the central limit theorem for $MRC\,[Y]_n$. As in Jacod et al. (2009), we only require a moment condition on ϵ to prove this result.

The notion of stable convergence is used, which we describe here. A sequence of random variables V^n is said to converge stably in law towards V, where V is defined on an appropriate extension $(\Omega', \mathcal{F}', P')$ of the probability space $(\Omega, \mathcal{F}, P)$, if and only if for any $\mathcal{F}$-measurable, bounded random variable W and any bounded, continuous function f, the convergence

$$\lim_{n\to\infty} \mathbb{E}\left[Wf(V^n)\right] = \mathbb{E}'\left[Wf(V)\right]$$

holds.

We write this as $V^n \xrightarrow{d_s} V$ and note that stable convergence is a slightly stronger mode of convergence than weak convergence, or

---

[1] Since $\sum_{i=1}^{n} \Delta_i^n Y\left(\Delta_i^n Y\right)' = 2n\Psi + \int_0^1 \Sigma_s ds + o_p(n^{-1})$, where the error of this approximation has expectation zero, the bias-corrected $MRC\,[Y]_n$ actually estimates $\left(1 - \frac{\psi_1^{k_n}}{\theta^2 \psi_2^{k_n}}\frac{1}{2n}\right)\int_0^1 \Sigma_s ds$ and thus needs to be rescaled by $1/\left(1 - \frac{\psi_1^{k_n}}{\theta^2 \psi_2^{k_n}}\frac{1}{2n}\right)$. We include this rescaling in our simulations and empirical work but omit it throughout the remainder of the text to simplify notation.

convergence in law, which is the special case obtained by taking $W = 1$ (see, e.g., Rényi, 1963; Aldous and Eagleson, 1978, for further details on stable convergence). Jacod and Shiryaev (2003) discuss the extension of this concept to stable convergence of processes. The key reason we require the convergence in law stably is that the conditional covariance matrix in the CLT of $MRC[Y]_n$, $avar_{MRC_n}$, is a function of σ and therefore random, and the usual convergence in law is insufficient to ensure joint convergence of the bivariate vector $(MRC[Y]_n, avar_{MRC_n})$, which we need to apply the delta method to the joint asymptotic distribution, for example in order to construct confidence intervals.

Theorem 2. *Assume that $\mathbb{E}\left(|\epsilon^j|^8\right) < \infty$ for all $j = 1, \ldots, d$ and (k_n, θ) satisfy Eq. (7). As $n \to \infty$, it holds that*

$$n^{1/4}\left(MRC\,[Y]_n - \int_0^1 \Sigma_s \mathrm{d}s\right) \xrightarrow{d_s} \sum_{j',k'=1}^d \int_0^1 \gamma_s^{jk,j'k'} \mathrm{d}B_s^{j'k'}, \tag{13}$$

where B is a standard d^2-dimensional Brownian motion defined on an extension of $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq0}, P)$ with $B \perp\!\!\!\perp \mathcal{F}$,

$$\sum_{j,m=1}^d \gamma_s^{kl,jm}\gamma_s^{k'l',jm}$$
$$= \frac{2}{\psi_2^2}\left(\Phi_{22}\theta \Lambda_s^{kl,k'l'} + \frac{\Phi_{12}}{\theta}\Theta_s^{kl,k'l'} + \frac{\Phi_{11}}{\theta^3}\Upsilon^{kl,k'l'}\right),$$

and where Λ, Θ and Υ are $d \times d \times d \times d$ arrays with elements

$$\Lambda_s = \left\{\Sigma_s^{kk'}\Sigma_s^{ll'} + \Sigma_s^{kl'}\Sigma_s^{lk'}\right\}_{k,k',l,l'=1,\ldots,d},$$
$$\Theta_s = \left\{\Sigma_s^{kk'}\Psi^{ll'} + \Sigma_s^{kl'}\Psi^{k'l} + \Sigma_s^{k'l}\Psi^{kl'}\right.$$
$$\left. + \Sigma_s^{ll'}\Psi^{kk'}\right\}_{k,k',l,l'=1,\ldots,d}, \tag{14}$$
$$\Upsilon = \left\{\Psi^{kk'}\Psi^{ll'} + \Psi^{kl'}\Psi^{lk'}\right\}_{k,k',l,l'=1,\ldots,d}.$$

Proof. See Appendix. □

Because $B \perp\!\!\!\perp \mathcal{F}$, we can write the convergence statement in Theorem 2 as follows:

$$n^{1/4}\left(MRC\,[Y]_n - \int_0^1 \Sigma_s \mathrm{d}s\right) \xrightarrow{d_s} MN\left(0, avar_{MRC}\right),$$

where

$$avar_{MRC} = \frac{2}{\psi_2^2}\left(\Phi_{22}\theta \int_0^1 \Lambda_s \mathrm{d}s + \frac{\Phi_{12}}{\theta}\int_0^1 \Theta_s \mathrm{d}s + \frac{\Phi_{11}}{\theta^3}\Upsilon\right) \tag{15}$$

is the conditional covariance matrix. This means that the asymptotic distribution of $MRC\,[Y]_n$ is mixed normal. To make use of this result to construct confidence intervals for elements of $\int_0^1 \Sigma_u \mathrm{d}u$ in practice, we need to estimate $avar_{MRC}$, which is addressed in Section 4.[2]

---

[2] The assumption that data be equidistant is not required for the consistency to hold true. It would also apply under identical observation times (i.e. synchronous but non-equidistant data). Here, $MRC\,[Y]_n$ is consistent with $\Delta_i^n V$ redefined as $\Delta_i^n V = V_{t_i} - V_{t_{i-1}}$ for any process V and $k_n = \theta\sqrt{n} + o(n^{1/4})$. This is not surprising: the realised covariance also remains consistent for irregular observations (by definition). The CLT also holds, but here the variable of integration ds needs to be replaced by $\mathrm{d}H_s$, where H_s is the so-called "quadratic variation of time", see, e.g., Barndorff-Nielsen et al. (2008b).

3.3. Choosing θ in practice

The $avar_{MRC}$ matrix in Theorem 2 depends on the parameter θ, or in other words the window size k_n. If the purpose is to estimate some one-dimensional parameters (such as realised covariance, regression or correlation) by real-valued functions of the MRC, it is in principle possible to minimize $avar_{MRC}$ by choosing the "best" θ for a fixed function g.[3]

To illustrate this point, we focus on the estimation problem in the univariate case, $d = 1$. In this situation, the expressions reduce to

$$avar_{MRC} = \frac{4}{\psi_2^2}\left(\Phi_{22}\theta \int_0^1 \sigma_s^4 \mathrm{d}s + \frac{2\Phi_{12}}{\theta}\Psi^2 \int_0^1 \sigma_s^2 \mathrm{d}s + \frac{\Phi_{11}}{\theta^3}\Psi^4\right),$$

where $IV = \int_0^1 \sigma_s^2 \mathrm{d}s$ and $IQ = \int_0^1 \sigma_s^4 \mathrm{d}s$ are called the integrated variance and integrated quarticity, respectively. Minimizing this term with respect to θ results in solving a quadratic equation. Thus, for the optimal choice of θ, say θ^*, we get $\theta^* = f(IV, IQ, \Psi)$. Then, we can use an iterative procedure to find an approximation of θ^*: (i) Choose a "reasonable" value for θ and compute $\hat{IV}$, $\hat{IQ}$ and $\hat{\Psi}$, (ii) from these compute $\hat{\theta}^*$ by setting $\hat{\theta}^* = f(\hat{IV}, \hat{IQ}, \hat{\Psi})$ and (iii) then, assuming the values of $\hat{\theta}^*$ converge, repeat this process until $\hat{\theta}^*$ has stabilized.

In practice, a simple guide that informs us about how to select θ for small values of n would certainly be valuable. Unfortunately, θ comes from asymptotic statistics and therefore it does not give any precise instructions on this issue. Nonetheless, some plausible range of values of θ can be inferred from previous work in this area. For example, Jacod et al. (2009) report that the univariate pre-averaged realised variance measure is fairly robust to the choice of k_n, and they suggest to take $\theta = 1/3$. Christensen et al. (forthcoming) use simulations to gauge the influence of sample size and noise on the optimal choice of θ. Interestingly, they show that the MSE curve of their pre-averaged quantile-based realised variance is highly asymmetric in k_n and they generally prefer to use a slightly higher value of k_n than what would be optimal. In their empirical analysis, they also show that conservative choices of k_n helps to heavily reduce the detrimental effects of price discreteness. In the simulation section and empirical illustration below, we therefore pick conservative values of k_n by choosing $\theta = 1$, which seems to work well.

3.4. Positive semi-definite estimators

In the previous section, we used an optimal pre-averaging window length to construct the MRC, which balances the impact of the noise with the estimation of the integrated covariance matrix. This choice leads to an optimal rate of convergence – $n^{-1/4}$ – but requires that we subtract an estimate of Ψ to eliminate the bias induced by noise, and the final estimator is then not positive semi-definite in general. Here, we demonstrate how positive semi-definite estimates of $\int_0^1 \Sigma_s \mathrm{d}s$ can be formed by increasing the bandwidth parameter k_n as to kill the influence of the noise, rather than balancing it. This comes at the cost of slowing down the speed at which MRC converges to the true integrated covariation.

Now, we take:

$$\frac{k_n}{n^{1/2+\delta}} = \theta + o\left(n^{-1/4+\delta/2}\right) \tag{16}$$

---

[3] The optimal choice of θ will depend on the original real-valued functions of the MRC. In this sense, there is no universal optimal θ.

for some $0 < \delta < 1/2$, and set

$$MRC\,[Y]_n^\delta = \frac{n}{n-k_n+2}\frac{1}{\psi_2 k_n}\sum_{i=0}^{n-k_n+1}\bar{Y}_i^n\left(\bar{Y}_i^n\right)'. \qquad (17)$$

The following result shows that $MRC\,[Y]_n^\delta$ is consistent without a bias-correction.

Theorem 3. *Assume that* $\mathbb{E}\left(|\epsilon^j|^4\right) < \infty$ *for all* $j = 1,\ldots,d$ *and* (k_n,θ) *satisfy Eq.* (16). *As* $n \to \infty$, *it holds that*

$$MRC\,[Y]_n^\delta \xrightarrow{p} \int_0^1 \Sigma_s ds. \qquad (18)$$

Proof. See Appendix. □

In Theorem 3, the properties of the noise process do not show up in the stochastic limit of Eq. (18), because the influence of the noise is negligible by the choice of order for k_n made in Eq. (16) (refer back to Eq. (8)). This has some appealing advantages. First, $MRC\,[Y]_n^\delta$ is positive semi-definite by construction. Second, although we state and prove this result in the i.i.d. noise case, Theorem 3 does in fact allow for more general noise dynamics than in Theorem 1. In particular, so long as $\mathbb{E}(\epsilon_t \mid X) = 0$ and $\bar{\epsilon}_i^n$ admits asymptotic normality at the usual rate $k_n^{-1/2}$, the theorem will hold (so we do not require any assumptions on the dependence between X and ϵ). Of course, the rate $k_n^{-1/2}$ is achieved in the i.i.d. case, but there are more general cases where it also holds (e.g., for q-dependent and mixing processes).

To show the CLT, we require a restriction on δ. This is because the bias caused by the noise, which is negligible in Theorem 3, becomes more substantial when multiplying with the rate of convergence.

Theorem 4. *Assume that* $\mathbb{E}\left(|\epsilon^j|^8\right) < \infty$ *for all* $j = 1,\ldots,d$ *and* (k_n,θ) *satisfy Eq.* (16). *As* $n \to \infty$, *it holds that*

(i) *If* $\delta > 0.1$

$$n^{1/4-\delta/2}\left(MRC\,[Y]_n^\delta - \int_0^1 \Sigma_s ds\right)$$
$$\xrightarrow{d_s} MN\left(0, \frac{2\Phi_{22}\theta}{\psi_2^2}\int_0^1 \Lambda_s ds\right), \qquad (19)$$

where (Λ_s) *is defined in* (14).

(ii) *If* $\delta = 0.1$

$$n^{1/5}\left(MRC\,[Y]_n^\delta - \int_0^1 \Sigma_s ds\right)$$
$$\xrightarrow{d_s} MN\left(\frac{\psi_1}{\theta^2\psi_2}\Psi, \frac{2\Phi_{22}\theta}{\psi_2^2}\int_0^1 \Lambda_s ds\right). \qquad (20)$$

Proof. See Appendix. □

Theorem 4 amounts to a classical bias-variance trade-off. It shows the expected result that using a longer pre-averaging window $k_n = O(n^{1/2+\delta})$ averages enough to make $MRC\,[Y]_n^\delta$ consistent without a bias-correction, but it also slows down its rate of convergence. The larger is δ, the harder is this effect. Note that the asymptotic variance depends solely on the volatility process σ, while the two noise-dependent terms (a cross-term and a pure noise term) appearing in Eq. (15) are wiped out.

The optimal choice $\delta = 0.1$ results in a $n^{-1/5}$ rate of convergence, and a large sample bias of the order $n^{-1/5}\frac{\psi_1}{\theta^2\psi_2}\Psi$. If

desired, the bias can be estimated using $\hat{\Psi}_n$, but it should be small for more recent data and could be ignored.[4]

This result bears some resemblance to that of the multivariate realised kernel of Barndorff-Nielsen et al. (2008b). A flat-top kernel estimator can be designed to converge at rate $n^{-1/4}$, but the resulting estimator may go negative. To guarantee positive semi-definiteness, the authors propose kernel corrections that are not entirely flat-top and they show this results in $n^{-1/5}$ rate of convergence and a non-zero asymptotic mean as produced here as well.

3.5. Mapping MRC into a realised kernel estimator

We can indeed make a stronger link between the multi-scale RV, the kernel approach and our pre-averaging method, when estimating quadratic variation. Here, we provide some more details about this relationship. To fix ideas, we take $d = 1$ (i.e., the univariate setting) and compare only to the kernel approach. Barndorff-Nielsen et al. (2008a) then show how the multi-scale RV fits into the realised kernel setting.

When we explicitly include the bias-correction and ignore finite sample issues, the $MRC\,[Y]_n$ estimator is given by the formula

$$MRC\,[Y]_n = \frac{1}{\psi_2 k_n}\sum_{i=0}^{n-k_n+1}|\bar{Y}_i^n|^2 - \frac{\psi_1}{\theta^2\psi_2}\hat{\Psi}_n. \qquad (21)$$

Consider a flat-top kernel-based estimator with kernel weights

$$k(s) = \frac{\phi_2(s)}{\psi_2},$$

where $\phi_2(s)$ and ψ_2 are defined as in Section 2.2. We call the function k a flat-top kernel, if (i) $k(0) = 1$, (ii) $k(1) = 0$ and (iii) $k'(0) = k'(1) = 0$. We note that $k(s) = \phi_2(s)/\psi_2$ satisfies all three conditions: (i) and (ii) are trivial, while the third condition follows from the identity:

$$k'(s) = -\frac{1}{\psi_2}\int_s^1 g(x)g'(x-s)dx,$$

and integration by parts (recall that $g(0) = g(1) = 0$).

In the next step, we map the MRC statistic into a kernel-like one as follows

$$MRC\,[Y]_n = \sum_{i=1}^n \delta_{0i}|\Delta_i^n Y|^2 + 2\sum_{h=1}^{k_n-2}\sum_{i=1}^{n-h}\delta_{hi}\Delta_i^n Y\Delta_{i+h}^n Y \qquad (22)$$

with

$$\delta_{0i} = \begin{cases} \frac{1}{\psi_2 k_n}\sum_{j=1}^i g^2\left(\frac{j}{k_n}\right) - \frac{\psi_1}{\theta^2\psi_2 2n} : \\ \quad 1 \le i \le k_n - 2 \\ \frac{1}{\psi_2 k_n}\sum_{j=1}^{k_n-1} g^2\left(\frac{j}{k_n}\right) - \frac{\psi_1}{\theta^2\psi_2 2n} : \\ \quad k_n - 1 \le i \le n - k_n + 2 \\ \frac{1}{\psi_2 k_n}\sum_{j=1}^{n-i+1} g^2\left(\frac{k_n-j}{k_n}\right) - \frac{\psi_1}{\theta^2\psi_2 2n} : \\ \quad n - k_n + 3 \le i \le n \end{cases} \qquad (23)$$

---

[4] If we do bias-correct $MRC\,[Y]_n^\delta$, the resulting estimator is then again not ensured to be positive semi-definite.

and

$$\delta_{hi} = \begin{cases} \dfrac{1}{\psi_2 k_n} \displaystyle\sum_{j=1}^{i} g\left(\dfrac{j}{k_n}\right) g\left(\dfrac{j+h}{k_n}\right): \\ \quad 1 \le i \le k_n - h - 2 \\ \dfrac{1}{\psi_2 k_n} \displaystyle\sum_{j=1}^{k_n-h-1} g\left(\dfrac{j}{k_n}\right) g\left(\dfrac{j+h}{k_n}\right): \\ \quad k_n - h - 1 \le i \le n - k_n + 2 \\ \dfrac{1}{\psi_2 k_n} \displaystyle\sum_{j=1}^{n-i+1} g\left(\dfrac{k_n-j}{k_n}\right) g\left(\dfrac{k_n-j+h}{k_n}\right): \\ \quad n - k_n + 3 \le i \le n \end{cases} \tag{24}$$

for $1 \le h \le k_n - 2$.

An example and some remarks are now in order.

Example. Take $g(x) = \min(x, 1 - x)$ for $x \in [0, 1]$, which is our canonical weight function. Then, the corresponding kernel $k(s) = \phi_2(s)/\psi_2$ is the Parzen kernel.

Remark. Apart from border terms, i.e. terms close to 0 and 1, the pre-averaging estimator coincides with the kernel-based estimator using the flat-top kernel function $k(s) = \phi_2(s)/\psi_2$. Both estimators possess the optimal rate of convergence $n^{-1/4}$, but the border terms affect their asymptotic distribution. In fact, to arrive at their central limit theorem for the realised kernel, Barndorff-Nielsen et al. (2008b) need to apply some averaging (or "jittering") to edge terms, while the pre-averaging estimator is asymptotically mixed normal "by construction".

Remark. Using the definition $k(s) = \phi_2(s)/\psi_2$, we learn that for every weight function g there exists a flat-top kernel k. However, the converse statement is not true in general. Thus, the class of flat-top kernels is larger than our class of pre-averaging functions, but this does not appear to be a noticeable disadvantage for practical applications (Barndorff-Nielsen et al., 2008b, for example, advocate using the Parzen kernel in practice, which corresponds to our g function by the example).

It is worthwhile to underscore that pre-averaging is a general approach, which can be used to estimate various characteristics of semimartingales, when they are cloaked with noise. Jacod et al. (2009), for example, pre-average realised variance, but it has also been used in jump-robust inference in conjunction with the bipower variation statistic (e.g., Podolskij and Vetter, 2009) or the quantile-based realised variance (e.g., Christensen et al., forthcoming). This is also the reason, why we can estimate the asymptotic conditional variance in the CLT using the same type of estimator to deliver a feasible result. Other estimation methods are typically designed to estimate quadratic variation and cannot, in general, be used to solve other estimation problems.

3.6. Applying the MRC to non-synchronous data

Non-synchronous trading has long been a recognized feature of multivariate high-frequency data (see, e.g., Fisher, 1966; Lo and MacKinlay, 1990). This causes spurious cross-autocorrelation amongst asset price returns sampled at regular intervals in calendar time, as new information gets built into prices at varying intensities. It is well known that high-frequency realised covariance estimates, using for example the previous-tick method to align prices, are biased towards zero in this setting (e.g., Epps, 1979).

Motivated by these shortcomings of realised covariance, a number of alternative procedures have been proposed in the literature. Scholes and Williams (1977) and Dimson (1979) suggested to include leads and lags of sample autocovariances

of high-frequency returns into the realised covariance to correct for stale prices. This results in a bias reduction but also increases the variance of the estimator (e.g., Griffin and Oomen, 2006). Importantly, the lead–lag realised covariance is generally inconsistent in the presence of noise. More recently, (Zhang, 2008) extends the two-scale RV to the multivariate setting and shows that it is consistent under non-synchronous prices and noise. An analytic derivation of the impact of Epps effect on realised covariance under previous-tick sampling is also given. In independent and concurrent work to ours, Barndorff-Nielsen et al. (2008b) propose a multivariate realised kernel. They rely on the concept of refresh time to match prices in calendar time and show that stale pricing errors under refresh time sampling do not change the asymptotic distribution of the multivariate realised kernel.

A related approach can be taken to make the MRC robust to non-synchronous observation times. In particular, under suitable regularity conditions on the sampling times, non-synchronous trading does not alter the asymptotic distribution of the MRC, when constructing synchronous time series of prices. A multivariate time series of high-frequency returns obtained in this fashion can therefore be plugged into the asymptotic theory developed above without concern. The proof of these results are highly technical, but their validity should be clear in the light of the comparison with the multivariate realised kernel given above. Therefore, we omit the detailed proofs of these results. Instead, we conduct some simulations in Section 6 to verify the correctness of these conjectures.

3.6.1. A pre-averaged Hayashi–Yoshida estimator

Hayashi and Yoshida (2005, 2008) develop an alternative procedure for covariance measurement in the noise-free case, which is based on the original non-synchronous data (see, e.g., de Jong and Nijman, 1997; Martens, 2003; Palandri, 2006; Corsi and Audrino, 2007, for related work). This estimator has the profound advantage that it does not throw away information that is typically lost using a synchronization procedure. Here, we show how this estimator can also be made robust to noise by using pre-averaging.

Given the vector of log-prices $Y = (Y^1, \ldots, Y^d)'$, which is defined by the noisy diffusion model in Eq. (4), we now assume that the component processes (Y^k) are observed at non-random time points $t_i^{(k)}$, for $i = 1, \ldots, n_k$, with $(t_i^{(k)})$ being a partition of the interval $[0, 1]$ and $k = 1, \ldots, d$. In addition, we need some regularity conditions on the sampling such that all time schemes are comparable in the following way: $\max |t_i^{(k)} - t_{i-1}^{(k)}| \to 0$ as $n_k \to \infty$, for $k = 1, \ldots, d$ and

$$\max_{1 \le i \le n_k} \# \left\{ t_j^{(k)} \mid t_j^{(k)} \in \left[t_i^{(l)}, t_i^{(l)}\right] \right\} \le K, \tag{25}$$

for $1 \le k, l \le d$ and some $K > 0$, where K is independent of n_k, for $k = 1, \ldots, d$. The latter condition says that data from one process do not cluster in any single interval of the others. Finally, the following condition is needed

$$\frac{\max |t_i^{(k)} - t_{i-1}^{(k)}|}{\min |t_i^{(k)} - t_{i-1}^{(k)}|} \le c, \tag{26}$$

where $c > 0$ is a constant independent of n_k for all $1 \le k \le d$. This condition implies that $n_k \max |t_i^{(k)} - t_{i-1}^{(k)}| \le c$, which could be assumed instead.

A consistent estimator of the integrated covariance $\int_0^1 \Sigma_s^{kl} ds$ between assets Y^k and Y^l can then be constructed as follows. We set $n = \sum_{k=1}^{d} n_k$ and define the statistic

$$HY[Y]_n^{(k,l)}$$

$$= \frac{1}{(\psi_{HY} k_n)^2} \sum_{i=0}^{n_k - k_n + 1} \sum_{j=0}^{n_l - k_n + 1} \bar{Y}_i^{k^n} \bar{Y}_j^{l^n} \mathbb{I}_{\{(t_i^{(k)}, t_{i+k_n}^{(k)}] \cap (t_j^{(l)}, t_{j+k_n}^{(l)}] \neq \emptyset\}}, \tag{27}$$

where $\psi_{HY} = \int_0^1 g(x)dx$ and $\mathbb{I}_{\{\bullet\}}$ is the indicator function discarding pre-averaged returns that do not overlap in time. Hence, $HY[Y]_n^{(k,l)}$ is a pre-averaged version of the Hayashi and Yoshida (2005) estimator. Note that under the previous conditions, n, n_k and n_l are of the same order and that n controls the universal pre-averaging window k_n.[5]

Theorem 5. *Assume that* $\mathbb{E}\left(|\epsilon^j|^4\right) < \infty$ *for all* $j = 1, \ldots, d$, (k_n, θ) *satisfy Eq.* (7) *and* $g(x) > 0$ *for* $x \in (0, 1)$. *As* $n \to \infty$, *it holds that*

$$HY[Y]_n^{(k,l)} \xrightarrow{p} \int_0^1 \Sigma_s^{kl} ds, \tag{28}$$

for $1 \leq k, l \leq d$.

Proof. See Appendix. □

Interestingly, Theorem 5 shows the somewhat surprising result that there is no asymptotic noise-induced bias in $HY[Y]_n^{(k,l)}$, not even when the spacings $(t_i^{(k)})$ and $(t_i^{(l)})$ are identical. This can be seen as follows. First, under the i.i.d. structure on the noise only products of the form $\epsilon_{t_i^{(k)}}^k \epsilon_{t_j^{(l)}}^l$ with $t_i^{(k)} = t_j^{(l)} = t$ contribute to potential bias. We consider that set of points and assume that $k_n \leq i \leq n_1 - k_n$ and $k_n \leq j \leq n_2 - k_n$ (this is innocent, for the summands which do not fulfill this are negligible). Then, an inspection of $HY[Y]_n^{(k,l)}$ shows that all products $\epsilon_t^k \epsilon_t^l$ appear with the factor

$$\left(\sum_{j=0}^{k_n-1} g\left(\frac{j+1}{k_n}\right) - g\left(\frac{j}{k_n}\right)\right)^2$$

in front. But $\sum_{j=0}^{k_n-1} g\left(\frac{j+1}{k_n}\right) - g\left(\frac{j}{k_n}\right) = 0$, because $g(0) = g(1) = 0$, so these terms drop out of the summation.

While $HY[Y]_n^{(k,l)}$ has the advantage that it is free of prior alignment of log-prices and hence does not throw away information in the sample, it does suffer from being a pairwise estimator, which means that once we assemble all the single variance/covariance estimates into a full covariance matrix, it is not guaranteed to be positive semi-definite. Still, there are some problems in financial economics, in which it is only the accuracy of the estimator that matters and positive semi-definiteness is less important, for example in asset allocation and risk management under gross exposure constraints (e.g. Fan et al., 2009). Moreover, our empirical results show that $(HY[Y]_n^{(k,l)})_{1 \leq k,l \leq d}$ does not fail to be positive semi-definite on a single day for the $d = 5$-dimensional vector of asset prices considered there.

Proposition 1. *Assume that* $\mathbb{E}\left(|\epsilon^j|^4\right) < \infty$ *for all* $j = 1, \ldots, d$. *If* (k_n, θ) *satisfy Eq.* (7) *and* $g(x) > 0$ *for* $x \in (0, 1)$, *then*

$$\text{var}\left(HY[Y]_n^{(k,l)}\right) = O\left(n^{-1/2}\right),$$

that is, $HY[Y]_n^{(k,l)}$ *has the optimal rate of convergence.*

---

[5] Our choice here implies that k_n is identical across all pairs of asset combinations. In general, the best way of choosing k_n depends on what is being estimated (e.g., covariance, correlation or beta). Thus, if one only cares about efficiency in pairwise covariance estimation, a better $HY[Y]_n^{(k,l)}$ estimator might be based on a local k_n, for example $k_n^{k,l} = \theta_{k,l}\sqrt{n} + o(n^{1/2})$. This would also serve to make the estimated covariances universe independent (i.e., they will not change by adding or removing assets). To make a more qualified, statistical argument on this issue, however, we need to compare an expression for the asymptotic variance of $HY[Y]_n^{(k,l)}$ under different ways of choosing k_n (e.g., from a CLT). This is beyond the scope of this paper and will be left for future research, where we hope to shed more light on the subject.

Proof. See Appendix. □

Proposition 1 shows that the rate of convergence associated with $HY[Y]_n^{(k,l)}$ is $n^{-1/4}$. A proof of the much stronger result, the CLT for $HY[Y]_n^{(k,l)}$, will be given in a companion paper to this one (see Christensen et al., 2010). In the empirical section, we gauge the properties of this estimator on actual data and find that the pre-averaged version performs very well.

4. An estimator of the asymptotic covariance matrix

In order to make Theorems 2 and 4 feasible, we need to estimate the asymptotic covariance matrix $avar_{MRC}$, as it appears in Eq. (15). Here, we give an explicit estimator of $avar_{MRC}$. More precisely, we present an estimator of the asymptotic covariance matrix of the vectorized statistic in Eq. (13).

First, we set

$$\chi_i^n = \text{vec}\left(\bar{Y}_i^n (\bar{Y}_i^n)'\right), \tag{29}$$

where $\text{vec}(\cdot)$ is the vectorization operator that stacks columns of a matrix below one another. Next, we define the statistic

$$V_n(g) = \sum_{i=0}^{n-k_n+1} \chi_i^n (\chi_i^n)' - \frac{1}{2} \sum_{i=0}^{n-2k_n+1} \left(\chi_i^n (\chi_{i+k_n}^n)' + \chi_{i+k_n}^n (\chi_i^n)'\right).$$

We should note that $V_n(g)$ depends on both the bandwidth parameter θ and the pre-averaging function g, and that it is positive semi-definite by construction. Moreover, for any $1 \leq k, k', l, l' \leq d$, we get the following convergence

$$V_n^{(k-1)d+k',(l-1)d+l'}(g) \xrightarrow{p} a_B(g, \theta) \int_0^1 \Lambda_u^{kk',ll'} du$$
$$+ a_M(g, \theta) \int_0^1 \Theta_u^{kk',ll'} du + a_N(g, \theta) \Upsilon^{kk',ll'},$$

where $a_B(g, \theta) = \theta^2 \psi_2^2$, $a_M(g, \theta) = \psi_1 \psi_2$ and $a_N(g, \theta) = \frac{\psi_1^2}{\theta^2}$ (the proof of this result is achieved by using arguments alike the ones presented in the proof of Theorem 1; see Kinnebrock and Podolskij, 2008, for further details).

What needs to be estimated is

$$avar_{MRC} = \frac{2}{\psi_2^2}\left(\Phi_{22}\theta \int_0^1 \Lambda_u du + \frac{\Phi_{12}}{\theta}\int_0^1 \Theta_u du + \frac{\Phi_{11}}{\theta^3}\Upsilon\right),$$

with all constants referring to a given function g_0. Suppose that $g_0(x) = \min(x, 1 - x)$ as above. Now, we take three different functions g_1, g_2 and g_3 that satisfy all the conditions of g_0, and which are chosen such that the matrix

$$A(g_1, g_2, g_3) = \begin{pmatrix} a_B(g_1, \theta) & a_M(g_1, \theta) & a_N(g_1, \theta) \\ a_B(g_2, \theta) & a_M(g_2, \theta) & a_N(g_2, \theta) \\ a_B(g_3, \theta) & a_M(g_3, \theta) & a_N(g_3, \theta) \end{pmatrix}$$

is invertible. We can then construct a weight vector

$$C(g_1, g_2, g_3) = \left(\frac{2\Phi_{22}\theta}{\psi_2^2}, \frac{2\Phi_{12}}{\psi_2^2\theta}, \frac{2\Phi_{11}}{\psi_2^2\theta^3}\right) A^{-1}(g_1, g_2, g_3). \tag{30}$$

Finally, we form the statistic

$$\widehat{avar}_{MRC,n} = C^{(1)}(g_1, g_2, g_3)V_n(g_1) + C^{(2)}(g_1, g_2, g_3)V_n(g_2)$$
$$+ C^{(3)}(g_1, g_2, g_3)V_n(g_3), \tag{31}$$

where $V_n(g_k)$ are the above estimators associated with the functions g_k, $k = 1, 2, 3$. Then, it holds that

$$\widehat{avar}_{MRC,n}^{(k-1)d+k',(l-1)d+l'} \xrightarrow{p} avar_{MRC}^{kk',ll'},$$

for any $1 \leq k, k', l, l' \leq d$.[6]

---

[6] The weights $\left(\frac{2\Phi_{22}\theta}{\psi_2^2}, \frac{2\Phi_{12}}{\psi_2^2\theta}, \frac{2\Phi_{11}}{\psi_2^2\theta^3}\right)$ in Eq. (30) are the coefficients in front of $(\int_0^1 \Lambda_u du, \int_0^1 \Theta_u du, \Upsilon)$ in the expression for the asymptotic variance of MRC.

There are various classes of functions from which the g_k can be selected, for example $g(x) = x^a (1 - x)^b$ with $a, b \geq 1$ or $g(x) = \sin(c\pi x)$ for integer c. If one manages to find a set of functions g_k for which the coefficients $C^{(k)}(g_1, g_2, g_3)$ are all positive, $k = 1, 2$ and 3, then the statistic in Eq. (31) is guaranteed to be positive semi-definite, because it is a linear combination of positive semi-definite statistics $V_n(g_k)$ using positive weights $C^{(k)}(g_1, g_2, g_3)$. It appears to be quite hard to find such a combination in practice, and so far we did not succeed at this. Nonetheless, $\widehat{avar}_{MRC,n}$ remains a consistent estimator.

5. Asymptotic theory for covariance, regression and correlation

The results in Sections 3 and 4 can be applied in order to compute confidence intervals for some functionals of $\int_0^1 \Sigma_u du$ that are important in practice, such as covariance, regression and correlation. For the ith and jth asset, these quantities are given by

$$\int_0^1 \Sigma_u^{ij} du, \quad \beta^{(ji)} = \frac{\int_0^1 \Sigma_u^{ij} du}{\int_0^1 \Sigma_u^{ii} du},$$

$$\rho^{(ji)} = \frac{\int_0^1 \Sigma_u^{ij} du}{\sqrt{\int_0^1 \Sigma_u^{ii} du \int_0^1 \Sigma_u^{jj} du}}. \tag{32}$$

Theorems 1, 3 or 5 can be invoked to provide consistent estimates of $\int_0^1 \Sigma_u^{ij} du$, $\beta^{(ji)}$ and $\rho^{(ji)}$, e.g. for $MRC[Y]_n$

$$MRC[Y]_n^{i,j} \xrightarrow{p} \int_0^1 \Sigma_u^{ij} du, \quad \hat{\beta}_n^{(ji)} = \frac{MRC[Y]_n^{i,j}}{MRC[Y]_n^{i,i}} \xrightarrow{p} \beta^{(ji)},$$

$$\hat{\rho}_n^{(ji)} = \frac{MRC[Y]_n^{i,j}}{\sqrt{MRC[Y]_n^{i,i} MRC[Y]_n^{j,j}}} \xrightarrow{p} \rho^{(ji)} \tag{33}$$

for any $1 \leq i, j \leq d$. The estimators in (33) are called modulated realised covariance, regression and correlation, respectively. In the next theorem we present the associated feasible central limit theorems, which follow from Theorem 2 and the delta method for stable convergence.

Theorem 6. Assume that $\mathbb{E}\left(|\epsilon^j|^8\right) < \infty$ for all $j = 1, \ldots, d$ and (k_n, θ) satisfy Eq. (7). As $n \to \infty$, it holds that

$$\frac{n^{1/4}\left(MRC[Y]_n^{i,j} - \int_0^1 \Sigma_u^{ij} du\right)}{\widehat{avar}_{MRC,n}^{(i-1)d+j,(i-1)d+j}} \xrightarrow{d} N(0, 1), \tag{34}$$

$$\frac{n^{1/4}(\hat{\beta}_n^{(ji)} - \beta^{(ji)})}{\sqrt{\left(MRC[Y]_n^{i,i}\right)^{-2} g_n^{(ji)}}} \xrightarrow{d} N(0, 1), \tag{35}$$

$$\frac{n^{1/4}(\hat{\rho}_n^{(ji)} - \rho^{(ji)})}{\sqrt{\left(MRC[Y]_n^{i,i} MRC[Y]_n^{j,j}\right)^{-1} h_n^{(ji)}}} \xrightarrow{d} N(0, 1), \tag{36}$$

where $\widehat{avar}_{MRC,n}$ is given in (31) and $g_n^{(ji)}$ and $h_n^{(ji)}$ are defined by

$$g_n^{(ji)} = \left(1, -\hat{\beta}_n^{(ji)}\right) \Gamma_n \left(1, -\hat{\beta}_n^{(ji)}\right)',$$

$$h_n^{(ji)} = \left(-\frac{1}{2}\hat{\beta}_n^{(ji)}, 1, -\frac{1}{2}\hat{\beta}_n^{(ji)}\right) \overline{\Gamma}_n \left(-\frac{1}{2}\hat{\beta}_n^{(ji)}, 1, -\frac{1}{2}\hat{\beta}_n^{(ji)}\right)'$$

with

$$\Gamma_n = \begin{pmatrix} \widehat{avar}_{MRC,n}^{(i-1)d+j,(i-1)d+j} & \widehat{avar}_{MRC,n}^{(i-1)d+j,(i-1)d+i} \\ \bullet & \widehat{avar}_{MRC,n}^{(i-1)d+i,(i-1)d+i} \end{pmatrix},$$

$$\overline{\Gamma}_n = \begin{pmatrix} \widehat{avar}_{MRC,n}^{(i-1)d+i,(i-1)d+i} & \widehat{avar}_{MRC,n}^{(i-1)d+j,(i-1)d+i} & \widehat{avar}_{MRC,n}^{(i-1)d+i,(j-1)d+j} \\ \bullet & \widehat{avar}_{MRC,n}^{(i-1)d+j,(i-1)d+j} & \widehat{avar}_{MRC,n}^{(i-1)d+j,(j-1)d+j} \\ \bullet & \bullet & \widehat{avar}_{MRC,n}^{(j-1)d+j,(j-1)d+j} \end{pmatrix}.$$

All the required terms are easy to compute, so it is rather simple to implement the estimators.

6. Simulation study

We now demonstrate the finite sample accuracy of some of the asymptotic results developed above using simulations. The design of our Monte Carlo study, which we briefly describe here, is identical to the analysis conducted in Barndorff-Nielsen et al. (2008b).

Specifically, to simulate log-prices we consider the following bivariate stochastic volatility model

$$dX_t^{(i)} = a^{(i)} dt + \rho^{(i)} \sigma_t^{(i)} dB_t^{(i)}$$
$$+ \sqrt{1 - [\rho^{(i)}]^2} \sigma_t^{(i)} dW_t, \quad \text{for } i = 1, 2, \tag{37}$$

where $B^{(i)}$ and W are independent Brownian motions. In this model, the term $\rho^{(i)} \sigma_t^{(i)} dB_t^{(i)}$ is an idiosyncratic component, while $\sqrt{1 - [\rho^{(i)}]^2} \sigma_t^{(i)} dW_t$ is a common factor.

The spot volatility is modeled as $\sigma_t^{(i)} = \exp(\beta_0^{(i)} + \beta_1^{(i)} \varrho_t^{(i)})$ with an Ornstein–Uhlenbeck specification for $\varrho_t^{(i)}$: $d\varrho_t^{(i)} = \alpha^{(i)} \varrho_t^{(i)} dt + d\varrho_t^{(i)}$. This implies that there is perfect correlation between the innovations of $\rho^{(i)} \sigma_t^{(i)} dB_t^{(i)}$ and $\sigma_t^{(i)}$, while it is $\rho^{(i)}$ between the increments of $X_t^{(i)}$ and $\varrho_t^{(i)}$. Finally, the magnitude of correlation between the two underlying price processes $X_t^{(1)}$ and $X_t^{(2)}$ is $\sqrt{1 - [\rho^{(1)}]^2}\sqrt{1 - [\rho^{(2)}]^2}$. The reported results are based on the following configuration of parameters for both processes: $(a^{(i)}, \beta_0^{(i)}, \beta_1^{(i)}, \alpha^{(i)}, \rho^{(i)}) = (0.03, -5/16, 1/8, -1/40, -0.3)$, so that $\beta_0^{(i)} = [\beta_1^{(i)}]^2/[2\alpha^{(i)}]$. We note that this particular choice of parameters also means that the volatility process has been normalized, in the sense that $\mathbb{E}(\int_0^1 [\sigma_s^{(i)}]^2 ds) = 1$.

We simulate 1000 paths of this model over the interval [0, 1]. Motivated by our empirical data, we let [0, 1] represent 6.5 hours worth of trading, which is then further decomposed into $N = 23{,}400$ subintervals of equal length $1/N$ (N denoting the number of seconds in 6.5 hours). In constructing noisy prices $Y^{(i)}$, we first generate a complete high-frequency record of N equidistant observations of the efficient price $X^{(i)}$ using a standard Euler scheme. The initial values for the $\varrho_t^{(i)}$ processes at each simulation run is drawn randomly from the stationary distribution of $\varrho_t^{(i)}$, which is $\varrho_t^{(i)} \sim N(0, [-2\alpha^{(i)}]^{-1})$.[7]

We add simulated microstructure noise $Y^{(i)} = X^{(i)} + \epsilon^{(i)}$ by taking

$$\epsilon^{(i)} \mid \{\sigma, X\} \stackrel{\text{i.i.d}}{\sim} N(0, \omega^2) \quad \text{with } \omega^2 = \gamma^2 \sqrt{\frac{1}{N} \sum_{j=1}^N \sigma_{j/N}^{(i)4}}. \tag{38}$$

This choice again follows Barndorff-Nielsen et al. (2008b) and means that the variance of the noise process increases with the

---

This choice results in a consistent estimator of $avar_{MRC}$. More generally, one can estimate any linear combination of $(\int_0^1 \Lambda_u du, \int_0^1 \Theta_u du, \Upsilon)$ by choosing an appropriate weight vector in Eq. (30). For example, the multivariate version of the integrated quarticity, $\int_0^1 \Lambda_u du$, can be estimated by plugging the linear combination $(1, 0, 0)A^{-1}(g_1, g_2, g_3)$ into Eq. (31).

[7] Note that the Ornstein–Uhlenbeck process admits an exact discretization (see, e.g., Glasserman, 2004). We use that result here to avoid discretization errors in approximating the continuous time distribution of $d\varrho^{(i)}$ over discrete time steps of size $\Delta t = 1/N$.

Table 1
Simulation results.

	Cov^{15m}		Cov^{1m}		$MRC[Y]_{n=RT}$		$MRC[Y]_{n=RT}^{\delta=0.1}$		$HY[Y]_n^{(k,l)}$	
	bias	rmse	bias	rmse	bias	rmse	bias	rmse	bias	rmse
Panel A: Integrated covariance										
$\xi^2 = 0$										
$\lambda = (3, 6)$	−0.015	0.159	−0.052	0.103	−0.007	0.156	−0.018	0.222	−0.044	0.278
$\lambda = (5, 10)$	−0.025	0.170	−0.090	0.147	−0.010	0.167	−0.023	0.233	−0.055	0.296
$\lambda = (10, 20)$	−0.044	0.174	−0.178	0.266	−0.017	0.207	−0.027	0.270	−0.069	0.338
$\lambda = (30, 60)$	−0.128	0.255	−0.387	0.547	−0.028	0.259	−0.038	0.326	−0.101	0.426
$\lambda = (60, 120)$	−0.228	0.367	−0.505	0.702	−0.038	0.300	−0.047	0.370	−0.123	0.513
$\xi^2 = 0.001$										
$\lambda = (3, 6)$	−0.019	0.166	−0.054	0.122	−0.008	0.148	−0.018	0.220	−0.044	0.275
$\lambda = (5, 10)$	−0.025	0.174	−0.089	0.161	−0.011	0.176	−0.022	0.239	−0.053	0.298
$\lambda = (10, 20)$	−0.046	0.188	−0.183	0.292	−0.016	0.198	−0.028	0.267	−0.071	0.345
$\lambda = (30, 60)$	−0.125	0.246	−0.383	0.538	−0.030	0.263	−0.036	0.329	−0.100	0.432
$\lambda = (60, 120)$	−0.226	0.368	−0.506	0.705	−0.039	0.295	−0.043	0.366	−0.123	0.518
$\xi^2 = 0.01$										
$\lambda = (3, 6)$	−0.014	0.350	−0.048	0.519	−0.007	0.147	−0.019	0.216	−0.044	0.276
$\lambda = (5, 10)$	−0.019	0.350	−0.070	0.485	−0.010	0.174	−0.021	0.237	−0.054	0.295
$\lambda = (10, 20)$	−0.056	0.346	−0.192	0.604	−0.017	0.202	−0.029	0.264	−0.069	0.341
$\lambda = (30, 60)$	−0.116	0.365	−0.392	0.682	−0.029	0.243	−0.040	0.323	−0.099	0.430
$\lambda = (60, 120)$	−0.225	0.494	−0.508	0.767	−0.039	0.324	−0.042	0.386	−0.121	0.525
Panel B: Integrated correlation										
$\xi^2 = 0$										
$\lambda = (3, 6)$	−0.016	0.028	−0.069	0.070	−0.001	0.017	−0.002	0.026	−0.012	0.051
$\lambda = (5, 10)$	−0.026	0.037	−0.119	0.121	−0.002	0.020	−0.003	0.030	−0.014	0.059
$\lambda = (10, 20)$	−0.051	0.060	−0.237	0.239	−0.002	0.024	−0.004	0.036	−0.020	0.076
$\lambda = (30, 60)$	−0.157	0.167	−0.514	0.517	−0.003	0.034	−0.008	0.047	−0.032	0.112
$\lambda = (60, 120)$	−0.287	0.299	−0.672	0.674	−0.006	0.044	−0.013	0.061	−0.044	0.159
$\xi^2 = 0.001$										
$\lambda = (3, 6)$	−0.141	0.149	−0.438	0.440	−0.001	0.018	−0.005	0.027	−0.012	0.050
$\lambda = (5, 10)$	−0.146	0.154	−0.463	0.465	−0.002	0.020	−0.005	0.030	−0.014	0.060
$\lambda = (10, 20)$	−0.170	0.179	−0.528	0.530	−0.003	0.025	−0.007	0.037	−0.019	0.077
$\lambda = (30, 60)$	−0.257	0.266	−0.657	0.659	−0.005	0.035	−0.012	0.050	−0.032	0.117
$\lambda = (60, 120)$	−0.368	0.378	−0.738	0.739	−0.007	0.046	−0.017	0.062	−0.045	0.159
$\xi^2 = 0.01$										
$\lambda = (3, 6)$	−0.559	0.571	−0.813	0.815	−0.002	0.026	−0.024	0.039	−0.011	0.051
$\lambda = (5, 10)$	−0.564	0.574	−0.817	0.819	−0.003	0.030	−0.027	0.044	−0.014	0.063
$\lambda = (10, 20)$	−0.579	0.590	−0.831	0.833	−0.005	0.037	−0.034	0.055	−0.020	0.079
$\lambda = (30, 60)$	−0.620	0.632	−0.851	0.853	−0.007	0.053	−0.044	0.074	−0.035	0.121
$\lambda = (60, 120)$	−0.657	0.667	−0.859	0.861	−0.011	0.067	−0.054	0.091	−0.043	0.166
Panel C: Integrated beta										
$\xi^2 = 0$										
$\lambda = (3, 6)$	−0.020	0.092	−0.090	0.128	−0.000	0.065	−0.002	0.098	−0.026	0.179
$\lambda = (5, 10)$	−0.034	0.100	−0.159	0.219	−0.001	0.076	−0.002	0.112	−0.030	0.214
$\lambda = (10, 20)$	−0.069	0.143	−0.316	0.427	−0.003	0.092	−0.003	0.135	−0.043	0.266
$\lambda = (30, 60)$	−0.212	0.316	−0.693	0.926	−0.001	0.129	−0.005	0.178	−0.069	0.385
$\lambda = (60, 120)$	−0.384	0.538	−0.902	1.190	−0.001	0.170	−0.008	0.222	−0.106	0.504
$\xi^2 = 0.001$										
$\lambda = (3, 6)$	−0.188	0.277	−0.586	0.777	0.000	0.067	−0.004	0.099	−0.026	0.177
$\lambda = (5, 10)$	−0.195	0.297	−0.623	0.832	−0.002	0.076	−0.004	0.112	−0.028	0.217
$\lambda = (10, 20)$	−0.226	0.325	−0.716	0.956	−0.003	0.093	−0.007	0.136	−0.039	0.263
$\lambda = (30, 60)$	−0.354	0.511	−0.903	1.200	−0.004	0.131	−0.010	0.183	−0.070	0.388
$\lambda = (60, 120)$	−0.495	0.680	−1.005	1.330	−0.006	0.169	−0.014	0.215	−0.089	0.496
$\xi^2 = 0.01$										
$\lambda = (3, 6)$	−0.747	1.010	−1.093	1.450	0.000	0.091	−0.031	0.116	−0.025	0.185
$\lambda = (5, 10)$	−0.758	1.033	−1.096	1.442	−0.002	0.114	−0.035	0.134	−0.032	0.220
$\lambda = (10, 20)$	−0.772	1.047	−1.123	1.501	−0.009	0.137	−0.045	0.159	−0.045	0.289
$\lambda = (30, 60)$	−0.839	1.149	−1.154	1.530	−0.006	0.181	−0.059	0.215	−0.076	0.380
$\lambda = (60, 120)$	−0.888	1.203	−1.169	1.559	−0.005	0.222	−0.065	0.239	−0.094	0.486

Note. This table shows the results of the simulation analysis.

level of volatility of $X^{(i)}$, as documented by Bandi and Russell (2006). γ^2 takes the values 0, 0.001, 0.01, which covers scenarios with no noise through low-to-high levels of noise.

Finally, we extract irregular, non-synchronous data from the complete high-frequency record using Poisson process sampling to generate actual observation times, $\{t_j^{(i)}\}$. In particular, we consider two independent Poisson processes with intensity parameter $\lambda = (\lambda_1, \lambda_2)$. Here λ_i denotes the average waiting time (in seconds) for

new data from process $Y^{(i)}$, so that an average day will have N/λ_i observations of $Y^{(i)}$, $i = 1, 2$. We vary λ_1 through (3, 5, 10, 30, 60) to capture the influence of liquidity on the performance of our estimators, and we set $\lambda_2 = 2\lambda_1$ such that on average $Y^{(2)}$ refreshes at half the pace of $Y^{(1)}$.

In Table 1, we report the results of the simulations. As the results for estimating the variance components of the 2×2 covariance matrix are as expected compared to prior work, we

Table 2
Descriptive statistics and number of data before and after filtering.

Stock	BMY	LMT	ORCL	SLE	SPY
Exchange	N	N	Q	N	P
Panel A: Transaction data					
Raw trades	1085,420	811,607	8153,413	533,517	7685,215
Corrected/Abnormal/Zeros	111	61	13806	85	2584
Time aggregation	231,826	139,181	6599,286	78,039	6000,898
#Trades	853,483	672,365	1540,321	455,393	1681,733
Intensity	3400	2679	6137	1814	6700
Noise ratio, γ	0.363	0.336	0.484	0.656	0.202
Panel B: Quotation data					
Raw quotes	5,402,607	3,245,315	23,411,495	3,208,830	17,536,447
Negative/Wide/Zeros	643	3917	2623	604	256
Time aggregation	2,547,054	1,075,914	20,050,224	979,299	12,851,464
#Quotes	2,854,910	2,165,484	3,358,648	2,228,927	4,684,727
Intensity	11,374	8627	13,381	8880	18,664
Avg. spread (in cents)	1.273	2.389	1.017	1.215	1.575
Noise ratio, γ	0.205	0.219	0.203	0.310	0.109

Note. This table reports some descriptive statistics and liquidity measures for the selection of stocks included in our empirical application. We show the exchange from which data are extracted. The exchange code is: N = NYSE, Q = NASDAQ and P = Pacific. Raw trades/quotes is the total number of data available from these exchanges during the trading session, while # trades/quotes is the total sample remaining after filtering the data. Intensity is the average number of data per day, while the noise ratio is defined in Oomen (2006).

give focus here towards estimating the integrated covariance, correlation and beta. Also, because the refresh time sampled MRC estimators perform somewhat better than the previous-tick based MRC estimators, we only report the results for the MRC estimators based on refresh time sampling ($n = RT$).[8] In the three panels of the table, we therefore provide the bias and root mean squared error (rmse) of the various estimators in terms of estimating the integrated covariance, correlation and beta. We compare our results to a standard realised covariance sampled at either a 1 min or 15 min frequency.

Looking at the table, we see that the MRC estimators are very efficient across all scenarios of noise and non-synchronous trading considered here, matching or outperforming the standard realised covariance. The $HY[Y]^{(k,l)}$ estimator is less efficient, owing largely to a larger finite sample bias in this estimator. We will study the possibilities of making finite sample adjustments to $HY[Y]^{(k,l)}$, elsewhere.

Above, we conjectured without a proof that the MRC was robust to stale prices and retained its rate of convergence under non-synchronous trading. If this is to be true, and ignoring finite sample biases, we should then expect the rmse of the two MRC estimators to decrease at rate $n^{-1/4}$ and $n^{-1/5}$. As can be seen from the table, this is exactly what we find. For example, when estimating the integrated covariance and going from $\lambda = (60, 120)$ to $\lambda = (3, 6)$, i.e. an approximate twenty-fold increase in sample size, the rmse of $MRC[Y]_n$ decreases roughly by half and the rmse of $MRC[Y]_n^\delta$ by slightly less than a half, which is consistent with the rates given above.

All in all, the simulation results show that the estimators proposed in this paper are very good at estimating the integrated covariance, correlation and beta across a wide range of noise and liquidity scenarios, and that the asymptotic predictions given above are also reasonable guides to their finite sample behavior.

7. Empirical illustration

To illustrate some empirical features of the pre-averaging theory developed above, we retrieved high-frequency data for a five-dimensional vector of assets from Wharton Research Data Services (WRDS). We picked four equities at random from the

S&P 500 constituents list as of July 1, 2009. We then added a 5th element, namely the S&P 500 Depository Receipt (ticker symbol SPY), which is an exchange-traded fund that tracks the large-cap segment of the US stock market. As such, it can be viewed as generating market-wide index returns. The four remaining stocks are the following (with ticker symbol and industry classification in parenthesis): Bristol-Myers Squibb (BMY, health care), Lockheed Martin (LMT, industrials), Oracle (ORCL, information technology) and Sara Lee (SLE, consumer staples), thus representing a broad category of industries. We use both trades and quotes data for the sample period that covers the whole of 2006, which results in 251 trading days.

Table 2 reports some descriptive statistics for our universe of stocks and sample period. As can be seen, these equities display varying degrees of liquidity with ORCL and SPY being the most liquid, while LMT and SLE are the least liquid. Also reported in the table is the univariate noise ratio statistic, γ, which is a noise-to-signal measure that describes the level of microstructure noise to integrated variance (see, e.g., Oomen, 2006, for further details on the noise ratio). Generally speaking, there is a tendency for more frequently traded companies to contain less microstructure noise, the notable exception being the transaction data for ORCL.

7.1. Filtering procedures

As a preliminary step, we subdued the sample data to some cleaning procedures. Pre-cleaning high-frequency data is necessary, because the raw data has many invalid observations (e.g., data with misplaced decimal points, or trades that are reported out-of-sequence). Our filter is roughly identical to that used by Barndorff-Nielsen et al. (2008b) with some minor differences. Here, we briefly describe the filtering rules we employ.

Trades and quotes: The following rules are applied to both trades and quotes data. (a) We keep data from a single exchange: Pacific for SPY and primary exchange for the 4 remaining equities, see Table 2, (b) we delete data with time stamps outside the regular exchange opening hours from 9:30am to 4:00pm, (c) we delete rows with a transaction price, bid or ask quote of zero, and (d) we aggregate data with identical time stamp using volume-weighted average prices (using total transaction volume or quoted bid and ask volume, respectively).

Trades only: We delete entries with a correction indicator $\neq 0$ or with abnormal sales condition.

---

[8] All unreported results are available upon request.

Table 3
Average of high-frequency covariance matrix estimates.

	$MRC[Y]_{n=CT(390)}$					$MRC[Y]^{\delta=0.1}_{n=CT(390)}$				
	BMY	LMT	ORCL	SLE	SPY	BMY	LMT	ORCL	SLE	SPY
BMY	1.465	0.203	0.259	0.154	0.233	1.396	0.195	0.259	0.152	0.231
LMT	0.201	0.975	0.232	0.149	0.247	0.194	0.905	0.222	0.158	0.235
ORCL	0.255	0.228	1.804	0.143	0.316	0.256	0.218	1.718	0.151	0.316
SLE	0.144	0.149	0.136	0.955	0.162	0.144	0.156	0.143	0.911	0.167
SPY	0.228	0.249	0.307	0.154	0.317	0.228	0.237	0.309	0.162	0.310
	$MRC[Y]_{n=RT}$					$MRC[Y]^{\delta=0.1}_{n=RT}$				
	BMY	LMT	ORCL	SLE	SPY	BMY	LMT	ORCL	SLE	SPY
BMY	1.394	0.197	0.226	0.134	0.220	1.397	0.198	0.249	0.141	0.227
LMT	0.193	0.955	0.224	0.134	0.242	0.201	0.924	0.222	0.147	0.241
ORCL	0.196	0.198	1.756	0.128	0.297	0.228	0.220	1.726	0.139	0.310
SLE	0.117	0.114	0.105	0.878	0.147	0.131	0.134	0.122	0.898	0.157
SPY	0.200	0.237	0.243	0.120	0.310	0.216	0.249	0.285	0.138	0.311
	HY					$HY[Y]^{(k,l)}_n$				
	BMY	LMT	ORCL	SLE	SPY	BMY	LMT	ORCL	SLE	SPY
BMY	1.160	0.093	0.097	0.103	0.111	1.417	0.192	0.255	0.144	0.227
LMT	0.049	0.761	0.075	0.072	0.097	0.202	0.902	0.223	0.147	0.238
ORCL	0.032	0.036	1.922	0.077	0.112	0.240	0.227	1.903	0.135	0.309
SLE	0.029	0.026	0.013	1.053	0.095	0.134	0.137	0.127	0.900	0.159
SPY	0.041	0.051	0.040	0.023	0.253	0.219	0.250	0.292	0.143	0.309
	Cov^{15s}					$Cov^{avg(5m,15m)}$				
	BMY	LMT	ORCL	SLE	SPY	BMY	LMT	ORCL	SLE	SPY
BMY	1.823	0.088	0.118	0.091	0.120	1.499	0.192	0.288	0.149	0.221
LMT	0.095	0.854	0.084	0.059	0.094	0.191	0.953	0.209	0.152	0.221
ORCL	0.095	0.086	4.043	0.069	0.162	0.261	0.206	1.936	0.153	0.306
SLE	0.056	0.049	0.043	1.947	0.072	0.146	0.148	0.142	1.008	0.161
SPY	0.111	0.118	0.117	0.057	0.293	0.219	0.233	0.290	0.152	0.303

Note. This table reports average covariance matrix estimates. In all subpanels, the numbers in the upper diagonal (including diagonal elements) are based on transaction prices, while the lower diagonal is based on mid-quote data. $Cov^{avg(5m,15m)}$ is a simple time series average of the realised covariance computed from 5 and 15 min returns.

Quotes only: We delete quotes with negative spreads and rows where the quoted spread exceeds 10 times the median spread for that day.
Table 2 also reports how many observations that are lost by passing these filters through the data. It should be noted that the "Trades Only" and "Quotes Only" filters generally tend to reduce the sample by only a very small fraction.

*7.2. High-frequency covariance analysis*

Here, we inspect the outcome of applying the estimators introduced above, after which we look at transforms of the covariance matrix. As a comparison, we also compute the standard realised covariance from 15 min, 5 min, and 15 s previous-tick data.

We implement both the $MRC[Y]_n$ and $MRC[Y]^{\delta}_n$ estimators of Sections 2 and 3.4, respectively. Recall that $MRC[Y]_n$ converges at rate $n^{-1/4}$, it needs to be corrected for bias, and as a result is not guaranteed to be positive semi-definite. We base $MRC[Y]^{\delta}_n$ on $\delta = 0.1$, among many plausible choices, which results in a $n^{-1/5}$ rate of convergence, a small finite sample bias that we omit correcting and, hence, positive semi-definiteness by construction. Two sampling schemes are used, calendar time and refresh time sampling, which yields a total of four combinations. We use pre-averaging windows found as $k_n = \lfloor \theta n^{\delta'} \rfloor$, where $\theta = 1$ and $\delta' = (0.5, 0.6)$ for our two choices. The selection of θ follows the conservative rule discussed above. We set $n = 390$ for the calendar time-based estimators, which is 1 min sampling, while n is determined automatically by the data for the refresh time sampling scheme.

In Table 3, we report the sample covariance matrix estimates averaged across the 251 days. This table is constructed in the usual way, displaying the results based on transaction prices in the upper diagonal (including the main diagonal), while the strict lower diagonal elements are the corresponding results based on quotation data. Consistent with prior literature, we see from the table that the standard realised covariance is suffering from Epps effect, when sampling runs quickly. All estimated covariance terms lie in the positive region, but for Cov^{15s} they are heavily compressed towards zero. This is less of a concern for $Cov^{avg(5m,15m)}$, which should tend to capture the average level of the covariance structure well, while not being seriously influenced by microstructure frictions and Epps effect. Turning next to the estimators proposed in this paper, we note that the time series average of both MRC versions and the noise-robust HY estimator are in line with that produced by $Cov^{avg(5m,15m)}$, showing that they appear free of any systematic bias. The Hayashi and Yoshida (2005) estimator produces a strong downwards bias in the covariance estimates, when it is applied directly to noisy and irregular high-frequency data. This reaffirms previous empirical work (see, e.g., Barndorff-Nielsen et al., 2008b; Griffin and Oomen, 2006; Voev and Lunde, 2007), so these results are not surprising or novel. The pre-averaged version $HY[Y]^{(k,l)}_n$, however, does a much better job and tends to agree with the average level of other noise-robust estimators.

Finally, we turn to the issue of positive semi-definiteness. As noted above, $MRC[Y]_n$ and $HY[Y]^{(k,l)}_n$ are not guaranteed to possess this property. Nonetheless, they do not fail to be positive semi-definite on a single instance across our sample period. This is true for both the transaction and quotation data, and both combinations of the bias-corrected MRC estimator. Thus, while theoretically
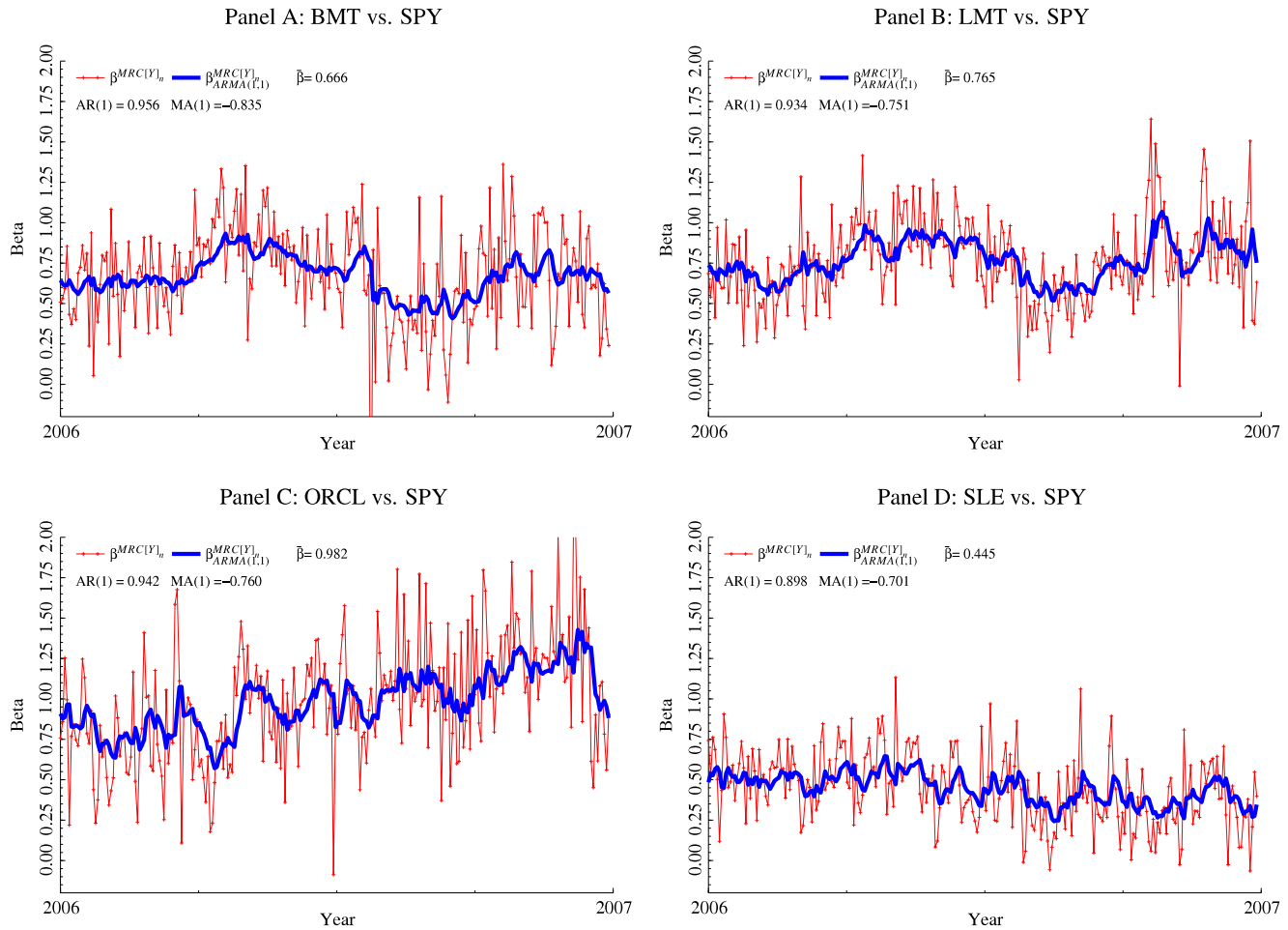
Note. $\beta^{MRC[Y]_n}$ is the time series of daily $MRC\,[Y]_n$-based beta estimates, using transaction prices and refresh time sampling for various asset combinations (in subpanels). $\beta^{MRC[Y]_n}_{ARMA(1,1)}$ are fitted values from an ARMA(1,1) filter, with estimated autoregressive parameter AR(1) and moving average parameter MA(1). The sample mean MRC beta is reported as $\bar{\beta}$.

Fig. 1. MRC-based beta.

a concern, this problem does not appear to occur frequently in practice, although the conclusion might change for other data sets.

7.3. Analysing realised beta

We now focus on estimating $\beta^{(ji)}$ by $\hat{\beta}_n^{(ji)} = MRC\,[Y]_n^{i,j} / MRC\,[Y]_n^{i,i}$, where we take $i = $ SPY and form regressions by using $j = $ BMY, LMT, ORCL, SLE. This type of regression, where individual equity covariances with the market are regressed onto a market-wide realised variance measure, is important in financial economics, for example within the conditional CAPM (see, e.g., Jagannathan and Wang, 1996; Lettau and Ludvigson, 2001), since only systematic risk should be rewarded with expected excess returns.

In Fig. 1, we plot the MRC-based betas from transaction prices and refresh time sampling. The corresponding plots from the other estimators proposed in this paper are qualitatively similar. As in Barndorff-Nielsen et al. (2008b), we smooth the daily beta estimates by passing them through an ARMA(1, 1) filter. The figure shows that beta is time-varying and predictable, and that it tends to fluctuate around its mean level. Evidently, the estimated processes exhibit substantial memory with autoregressive roots at 0.90 or higher (see also Andersen et al., 2006, who study the persistence of quarterly realised beta estimated from daily asset returns).

8. Concluding remarks

In this paper, we present a simple solution, based on applying pre-averaging to financial high-frequency data, to the problem of how to estimate the multivariate ex-post integrated covariance matrix, possibly in the simultaneous presence of market microstructure noise and non-synchronous trading.

A modulated realised covariance (MRC) estimator is introduced. The MRC bears close resemblance to a standard realised covariance, being a sum of outer products of high-frequency returns, but it relies instead on pre-averaging to reduce the harmful impact of microstructure noise. We study the properties of this new estimator by showing its consistency and asymptotic mixed normality under mild conditions on the dynamics of the price process. As shown in the paper, the MRC can be configured to possess an optimal rate of convergence or to guarantee positive semi-definite covariance matrix estimates. In the presence of non-synchronous trading, we also outline how to modify the MRC by using an imputation scheme (for example the previous-tick rule or refresh time sampling) to match high-frequency prices in time. An MRC constructed on the back of such artificial returns will again be consistent for the integrated covariance.

Another novelty developed in this paper is a pre-averaged version of the Hayashi and Yoshida (2005) estimator that can

be implemented directly on the raw noisy *and* non-synchronous observations, without any prior alignment of prices. We also show the consistency of this estimator and derive a rate for its variance, but otherwise we defer further theoretical analysis of finite sample improvements and asymptotic properties of the pre-averaged Hayashi–Yoshida estimator to future research (see Christensen et al., 2010).

We demonstrate with a set of simulations that these newly proposed estimators can bring substantial efficiency gains with them compared to a standard realised covariance in the realistic setting with both microstructure noise and non-synchronous trading. Furthermore, an empirical illustration highlights their applicability to real high-frequency data. We therefore look forward to future applications of these estimators, including an investigation of their informational content about future volatility. Being able to produce good forecasts of future volatility is paramount in financial economics. Thus, as an example, it could be interesting to apply our estimators in the context of portfolio choice to calculate their economic value, akin to Fleming et al. (2001, 2002), Bandi and Russell (2006) and others. However, generally they should also be useful in many other areas, including asset and option pricing or risk management.

Acknowledgements

Appendix. Proofs

In the following, we assume that the processes a and σ are bounded. This is without loss of generality and can be justified by a standard localization procedure (see, e.g., Barndorff-Nielsen et al., 2006). Moreover, we denote constants by C, or C_p if they depend on an additional parameter p. The main parts of the proofs are based upon Podolskij and Vetter (2009) and Jacod et al. (2009).

Proof of Theorem 1. Due to the triangular equality $[V, W] = \frac{1}{4}([V + W, V + W] - [V - W, V - W])$, it suffices to prove the univariate case $d = 1$ (i.e. all processes are one-dimensional). We use the decomposition

$$MRC\,[Y]_n = \frac{1}{k_n} \sum_{l=0}^{k_n-1} MRC\,[Y]_n^l - \frac{\psi_1}{\theta^2 \psi_2} \hat{\Psi}_n, \tag{39}$$

with

$$MRC\,[Y]_n^l = \frac{1}{\theta \psi_2 \sqrt{n}} \sum_{j=0}^{[n/k_n]-1} |\bar{Y}_{l+jk_n}^n|^2 .$$

Notice that, for any $l = 0, \ldots, k_n - 1$, the summands in the definition of $MRC\,[Y]_n^l$ are asymptotically uncorrelated. This type of estimators have been discussed in Podolskij and Vetter (2009) and we can deduce by the methods presented therein (see the proof of Theorem 1) that

$$MRC\,[Y]_n^l \xrightarrow{p} \int_0^1 \sigma_s^2 ds + \frac{\psi_1}{\psi_2 \theta^2} \Psi,$$

where the convergence holds uniformly in l (due to the boundedness of the processes a and σ). On the other hand we have that

$$\hat{\Psi}_n = \frac{1}{2n} \sum_{i=1}^n |\Delta_i^n Y|^2 \xrightarrow{p} \Psi.$$

This implies the convergence

$$MRC\,[Y]_n \xrightarrow{p} \int_0^1 \sigma_s^2 ds,$$

which completes the proof. □

Proof of Theorem 2. Here we apply the "big blocks & small blocks"-technique used in Jacod et al. (2009). The role of the small blocks (which will be asymptotically negligible) is to ensure the asymptotic independence of the big blocks. More precisely, we choose an integer p, set

$$a_i(p) = i(p + 1)k_n \quad \text{and} \quad b_i(p) = i(p + 1)k_n + pk_n,$$

and let $A_i(p)$ denote the set of integers l satisfying $a_i(p) \le l < b_i(p)$ and $B_i(p)$ the integers satisfying $b_i(p) \le l < a_{i+1}(p)$. We further define $j_n(p)$ to be the largest integer j such that $b_j(p) \le n$ holds, which gives the identity

$$j_n(p) = \left\lfloor \frac{n}{k_n(p + 1)} \right\rfloor - 1. \tag{40}$$

Moreover, we use the notation $i_n(p) = (j_n(p) + 1)(p + 1)k_n$.

Next, we introduce the random variable

$$\bar{Y}_{i,m}^n = \sum_{j=1}^{k_n-1} g\left(\frac{j}{k_n}\right) \left(\sigma_{\frac{m}{n}} \Delta_{i+j}^n W + \Delta_{i+j}^n \epsilon\right), \tag{41}$$

which can be interpreted as an approximation of some $\bar{Y}_j^n$. Moreover, we set

$$\Upsilon_{j,m}^n = \bar{Y}_{i,m}^n \left(\bar{Y}_{i,m}^n\right)' - \mathbb{E}\left[\bar{Y}_{i,m}^n \left(\bar{Y}_{i,m}^n\right)' \mid \mathcal{F}_{\frac{m}{n}}\right], \tag{42}$$

and define

$$\tilde{Y}_j^n = \begin{cases} \Upsilon_{j,a_i(p)}^n, & j \in A_i(p) \\ \Upsilon_{j,b_i(p)}^n, & j \in B_i(p) \\ \Upsilon_{j,i_n(p)}^n, & j \ge i_n(p) \end{cases}$$

as well as

$$\zeta(p, 1)_j^n = \sum_{l=a_j(p)}^{b_j(p)-1} \tilde{Y}_l^n, \qquad \zeta(p, 2)_j^n = \sum_{l=b_j(p)}^{a_{j+1}(p)-1} \tilde{Y}_l^n.$$

Notice that $\zeta(p, 1)_j^n$ contains pk_n summands ("big block") whereas $\zeta(p, 2)_j^n$ contains k_n summands ("small block"). Finally, we set

$$M(p)^n = n^{-\frac{1}{2}} \sum_{j=0}^{j_n(p)} \zeta(p, 1)_j^n, \qquad N(p)^n = n^{-\frac{1}{2}} \sum_{j=0}^{j_n(p)} \zeta(p, 2)_j^n,$$

$$C(p)^n = n^{-\frac{1}{2}} \sum_{j=i_n(p)}^n \tilde{Y}_j^n$$

and note that

$$\mathbb{E}\left[\zeta(p, 1)_j^n \mid \mathcal{F}_{\frac{a_j(p)}{n}}\right] = 0 = \mathbb{E}\left[\zeta(p, 2)_j^n \mid \mathcal{F}_{\frac{b_j(p)}{n}}\right] \tag{43}$$

by construction.

Now, by the same approximations as presented in Jacod et al. (2009) (see the identity (5.14), Lemmas 5.5 and 5.6 therein) we get that

$$n^{1/4}\left(MRC\,[Y]_n - \int_0^1 \Sigma_s ds\right)$$

$$= \frac{n^{\frac{1}{4}}}{\theta \psi_2}(M(p)^n + N(p)^n + C(p)^n) + R(p)_n \tag{44}$$

where the last three summands satisfy the convergence

$$\lim_{p\to\infty} \limsup_{n\to\infty} P(\|n^{\frac{1}{4}}N(p)^n\|$$

$$+ \|n^{\frac{1}{4}}C(p)^n\| + \|R(p)^n\| > \delta) = 0 \tag{45}$$

for any $\delta > 0$. Notice that the term $R(p)_n$ stands for the approximation error in Eq. (41).

In the next lemma we show the stable convergence $\frac{n^{\frac{1}{4}}}{\theta \psi_2} M(p)^n \xrightarrow{d_s} U(p)$ (for any fixed p). On the other hand, we will see that, as $p \to \infty$, $U(p) \xrightarrow{p} U$, where U is the limiting variable defined in Theorem 2. By combining this with Eqs. (44) and (45) we arrive at the assertion of Theorem 2.

Lemma 1. *If the assumptions of Theorem 2 are satisfied we obtain (for any fixed p)*

$$\frac{n^{\frac{1}{4}}}{\theta \psi_2} M(p)^n \xrightarrow{d_s} U(p) = \sum_{j',k'=1}^{d} \int_0^1 \gamma_s^{jk,j'k'}(p) dB_s^{j'k'},$$

and

$$\sum_{j,m=1}^{d} \gamma_s^{kl,jm}(p) \gamma_s^{k'l',jm}(p)$$

$$= A_s^{kl,k'l'} = \frac{2}{\psi_2^2} \left(\frac{\theta p}{p+1} \Lambda_s^{kl,k'l'} \int_0^1 \left(1 - \frac{u}{p}\right) \phi_2^2(u) du \right.$$

$$+ \frac{p}{\theta(p+1)} \Theta_s^{kl,k'l'} \int_0^1 \left(1 - \frac{u}{p}\right) \phi_1(u)\phi_2(u) du$$

$$\left. + \frac{p}{\theta^3(p+1)} \Upsilon^{kl,k'l'} \int_0^1 \left(1 - \frac{u}{p}\right) \phi_1^2(u) du \right),$$

where the processes Λ_s, Θ_s and Υ are given in Theorem 2.

Notice that $\sum_{j,m=1}^{d} \gamma_s^{kl,jm}(p) \gamma_s^{k'l',jm}(p) \xrightarrow{p} \sum_{j,m=1}^{d} \gamma_s^{kl,jm} \gamma_s^{k'l',jm}$ ($1 \le k, k', l, l' \le d$), where γ_s is defined in Theorem 2. From this we deduce the convergence $U(p) \xrightarrow{p} U$.

Proof of Lemma 1. Due to Theorem IX 7.28 in Jacod and Shiryaev (2003) the following conditions need to be shown (for all $1 \le k, k', l, l' \le d$)

$$\frac{n^{-1/2}}{\theta^2 \psi_2^2} \sum_{j=0}^{j_n(p)} \mathbb{E}\left[\zeta(p,1)_j^{n,kl} \zeta(p,1)_j^{n,k'l'} \mid \mathcal{F}_{\frac{a_j(p)}{n}} \right] \xrightarrow{p} \int_0^1 A_u^{kl,k'l'} du, \tag{46}$$

$$n^{-1} \sum_{j=0}^{j_n(p)} \mathbb{E}\left[\|\zeta(p,1)_j^n\|^4 \mid \mathcal{F}_{\frac{a_j(p)}{n}} \right] \xrightarrow{p} 0, \tag{47}$$

$$n^{-1/4} \sum_{j=0}^{j_n(p)} \mathbb{E}\left[\zeta(p,1)_j^{n,kl} \Delta W(p)_j^{n,k'} \mid \mathcal{F}_{\frac{a_j(p)}{n}} \right] \xrightarrow{p} 0, \tag{48}$$

$$n^{-1/4} \sum_{j=0}^{j_n(p)} \mathbb{E}\left[\zeta(p,1)_j^{n,kl} \Delta N(p)_j^n \mid \mathcal{F}_{\frac{a_j(p)}{n}} \right] \xrightarrow{p} 0, \tag{49}$$

where $\Delta V(p)_j^n = V_{n/b_j(p)} - V_{n/a_j(p)}$ for any process V and Eq. (49) holding for any one-dimensional bounded martingale N being orthogonal to W. For proving Eqs. (47) and (49), it is no restriction to assume that $d = 1$. Then these conditions are already shown in Jacod et al. (2009) (Lemma 5.7). On the other hand, the functional $\zeta(p,1)_j^n$ is even in W. Since W and ϵ are independent, we readily deduce that

$$\mathbb{E}\left[\zeta(p,1)_j^{n,kl} \Delta W(p)_j^{n,k'} \mid \mathcal{F}_{\frac{a_j(p)}{n}} \right] = 0,$$

which implies the condition in Eq. (48). Hence, we are left to proving Eq. (46).

First, notice the identity

$$\bar{V}_i^n = \sum_{j=1}^{k_n} g\left(\frac{j}{k_n}\right) \Delta_{i+j}^n V = -\sum_{j=0}^{k_n-1} \left(g\left(\frac{j+1}{k_n}\right) - g\left(\frac{j}{k_n}\right) \right) V_{\frac{i+j}{n}}.$$

The second equality is useful for the computation of the moments of $\bar{\epsilon}_i^n$. By the smoothness assumption on the function g and the above identity we obtain the approximations ($1 \le k, l \le d$)

$$\mathbb{E}[\overline{W}_j^{n,k} \overline{W}_{j'}^{n,l}] = \delta_{kl} \frac{k_n}{n} \psi_2 \left(\frac{|j-j'|}{k_n} \right) + O(n^{-1}),$$

$$\mathbb{E}[\bar{\epsilon}_j^{n,k} \bar{\epsilon}_{j'}^{n,l}] = \frac{\Psi^{kl}}{k_n} \psi_1 \left(\frac{|j-j'|}{k_n} \right) + O(n^{-1}) \tag{50}$$

for $|j - j'| < k_n$, whereas the above expectations vanish when $|j - j'| \ge k_n$ (here δ_{kl} denotes the Kronecker symbol). Next, we introduce the decomposition

$$\zeta(p,1)_j^n = \upsilon(p,1)_j^n + \upsilon(p,2)_j^n + \upsilon(p,3)_j^n,$$

where the terms $\upsilon(p,1)_j^n$, $\upsilon(p,2)_j^n$ and $\upsilon(p,3)_j^n$ are given by

$$\upsilon(p,1)_j^n = \sum_{l=a_j(p)}^{b_j(p)-1} \sigma_{\frac{a_j(p)}{n}} \overline{W}_l^n \left(\sigma_{\frac{a_j(p)}{n}} \overline{W}_l^n \right)'$$

$$- \mathbb{E}\left[\sigma_{\frac{a_j(p)}{n}} \overline{W}_l^n \left(\sigma_{\frac{a_j(p)}{n}} \overline{W}_l^n \right)' \mid \mathcal{F}_{\frac{a_j(p)}{n}} \right],$$

$$\upsilon(p,2)_j^n = \sum_{l=a_j(p)}^{b_j(p)-1} \bar{\epsilon}_l^n \left(\bar{\epsilon}_l^n \right)' - \mathbb{E}\left[\bar{\epsilon}_l^n \left(\bar{\epsilon}_l^n \right)' \right],$$

$$\upsilon(p,3)_j^n = \sum_{l=a_j(p)}^{b_j(p)-1} \sigma_{\frac{a_j(p)}{n}} \overline{W}_l^n \left(\bar{\epsilon}_l^n \right)' + \bar{\epsilon}_l^n \left(\sigma_{\frac{a_j(p)}{n}} \overline{W}_l^n \right)'.$$

By a straightforward calculation (and Eq. (50)) we obtain for all $1 \le k, l, k'l' \le d$

$$\mathbb{E}\left[\upsilon(p,1)_j^{n,kl} \upsilon(p,1)_j^{n,k'l'} \mid \mathcal{F}_{\frac{a_j(p)}{n}} \right]$$

$$= \frac{2pk_n^4}{n^2} \Lambda_{\frac{a_j(p)}{n}}^{kl,k'l'} \int_0^1 \left(1 - \frac{u}{p}\right) \phi_2^2(u) du + o_p(1),$$

$$\mathbb{E}\left[\upsilon(p,2)_j^{n,kl} \upsilon(p,2)_j^{n,k'l'} \mid \mathcal{F}_{\frac{a_j(p)}{n}} \right]$$

$$= 2p \Upsilon^{kl,k'l'} \int_0^1 \left(1 - \frac{u}{p}\right) \phi_1^2(u) du + o_p(1),$$

$$\mathbb{E}\left[\upsilon(p,3)_j^{n,kl} \upsilon(p,3)_j^{n,k'l'} \mid \mathcal{F}_{\frac{a_j(p)}{n}} \right]$$

$$= \frac{2pk_n^2}{n} \Theta_{\frac{a_j(p)}{n}}^{kl,k'l'} \int_0^1 \left(1 - \frac{u}{p}\right) \phi_1(u)\phi_2(u) du + o_p(1),$$

where the approximation holds uniformly in j. Now recall that $j_n(p) = \left\lfloor \frac{n}{k_n(p+1)} \right\rfloor - 1$. Consequently, by Riemann integrability we deduce that

$$\frac{n^{-\frac{1}{2}}}{\theta^2 \psi_2^2} \sum_{j=0}^{j_n(p)} \mathbb{E}\left[\zeta(p,1)_j^{n,kl} \zeta(p,1)_j^{n,k'l'} \mid \mathcal{F}_{\frac{a_j(p)}{n}} \right] \xrightarrow{p} \int_0^1 A_u^{kl,k'l'} du,$$

which completes the proof of Lemma 1. $\quad\square$

Proof of Theorem 3 and 4. Recall that $\frac{k_n}{n^{1/2+\delta}} = \theta + o(n^{-1/4+\delta/2})$. A straightforward calculation shows that

$$\mathbb{E}\left[MRC\left[\epsilon\right]_n^\delta\right] = \frac{\psi_1}{\theta^2\psi_2 n^{2\delta}}\Psi + o(n^{-1/4+\delta/2}).$$

By similar methods as presented in the proof of Theorem 1 we deduce that

$$MRC\left[Y\right]_n^\delta - \left(\int_0^1 \Sigma_s ds + \frac{\psi_1}{\theta^2\psi_2 n^{2\delta}}\Psi\right) \xrightarrow{p} 0.$$

Hence, we obtain the convergence in probability

$$MRC\left[Y\right]_n^\delta \xrightarrow{p} \int_0^1 \Sigma_s ds,$$

which implies the assertion of Theorem 3. Following the same scheme demonstrated in the proof of Theorem 2 we get

$$n^{1/4-\delta/2}\left(MRC\left[Y\right]_n^\delta - \left(\int_0^1 \Sigma_s ds + \frac{\psi_1}{\theta^2\psi_2 n^{2\delta}}\Psi\right)\right)$$

$$\xrightarrow{d_s} MN\left(0, \frac{2\Phi_{22}\theta}{\psi_2^2}\int_0^1 \Lambda_s ds\right),$$

since $\frac{\psi_1}{\theta^2\psi_2 n^{2\delta}}\Psi$ is an appropriate centering for the noise term in this case. Now

$$\frac{\psi_1 n^{1/4-\delta/2}}{\theta^2\psi_2 n^{2\delta}}\Psi \xrightarrow{p} 0$$

for $\delta > 1/10$, whereas

$$\frac{\psi_1 n^{1/4-\delta/2}}{\theta^2\psi_2 n^{2\delta}}\Psi = \frac{\psi_1}{\theta^2\psi_2}\Psi$$

for $\delta = 1/10$. Hence, Theorem 4 follows.

Proof of Theorem 5. First, we start with the decomposition

$$HY[Y]_n^{(k,l)} = \frac{1}{(\psi_{HY}k_n)^2}\left(\sum_{i,j}\bar{X}^k{}_i^n\bar{X}^l{}_j^n\mathbb{I}_{\{(t_i^{(k)},t_{i+k_n}^{(k)}]\cap(t_j^{(l)},t_{j+k_n}^{(l)}]\neq\emptyset\}}\right.$$

$$+ \sum_{i,j}\left(\bar{X}^k{}_i^n\bar{\epsilon}^l{}_j^n + \bar{\epsilon}^k{}_i^n\bar{X}^l{}_j^n\right)\mathbb{I}_{\{(t_i^{(k)},t_{i+k_n}^{(k)}]\cap(t_j^{(l)},t_{j+k_n}^{(l)}]\neq\emptyset\}}$$

$$\left.+ \sum_{i,j}\bar{\epsilon}^k{}_i^n\bar{\epsilon}^l{}_j^n\mathbb{I}_{\{(t_i^{(k)},t_{i+k_n}^{(k)}]\cap(t_j^{(l)},t_{j+k_n}^{(l)}]\neq\emptyset\}}\right)$$

$$=: HY[Y]_n^1 + HY[Y]_n^2 + HY[Y]_n^3.$$

As X and ϵ are independent, it follows that $\mathbb{E}\left(HY[Y]_n^2\right) = 0$, and a simple computation shows that

$$\text{var}(HY[Y]_n^2) \to 0.$$

Thus, $HY[Y]_n^2 \xrightarrow{p} 0$. Next, we consider the term $HY[Y]_n^3$. This expression can be further decomposed as

$$HY[Y]_n^3 = \frac{1}{(\psi_{HY}k_n)^2}\left(\sum_{t_i\in J_{k,l}}a_i^n(k,l)\epsilon_{t_i}^k\epsilon_{t_i}^l + \sum_{t_i\in J_{k,l}^c}b_i^n(k,l)\epsilon_{t_i}^k\epsilon_{t_i}^l\right.$$

$$\left.+ \sum_{(i,j)\in F_{k,l}}c_{ij}^n(k,l)\epsilon_{t_i^{(k)}}^k\epsilon_{t_j^{(l)}}^l\mathbb{I}_{\{t_i^{(k)}\neq t_j^{(l)}\}}\right),$$

for certain numbers $a_i^n(k,l)$, $b_i^n(k,l)$ and $c_{ij}^n(k,l)$. Here $J_{k,l}$ denotes the set of common points of $(t_i^{(k)})_{k_n\leq i\leq n_k-k_n}$ and $(t_i^{(l)})_{k_n\leq i\leq n_l-k_n}$,

and $J_{k,l}^c$ denotes the set of all common points of $(t_i^{(k)})_{1\leq i\leq n_k}$ and $(t_i^{(l)})_{1\leq i\leq n_l}$ excluded $J_{k,l}$. The set $F_{k,l}$ is given by

$$F_{k,l} = \left\{(i,j)\mid \exists r,s \text{ with } r\leq i\leq r+k_n, s\leq j\leq s+k_n,\right.$$

$$\left.(t_r^{(k)},t_{r+k_n}^{(k)}]\cap(t_s^{(l)},t_{s+k_n}^{(l)}]\neq\emptyset\right\}.$$

Since $\bar{\epsilon}^k{}_i^n = -\sum_{j=1}^{k_n}\left(g\left(\frac{j}{k_n}\right)-g\left(\frac{j-1}{k_n}\right)\right)\epsilon_{t_{i+j}^{(k)}}^k$ and g is piecewise differentiable it holds that

$$|b_i^n(k,l)| + |c_{ij}^n(k,l)| \leq C.$$

A straightforward computation shows that

$$a_i^n(k,l) = \left(\sum_{j=1}^{k_n}g\left(\frac{j}{k_n}\right)-\left(\frac{j-1}{k_n}\right)\right)^2 = (g(1)-g(0))^2 = 0,$$

because $g(0) = g(1) = 0$. Thus, the first summand in the decomposition of $HY[Y]_n^3$ disappears, which is absolutely crucial for the proof. On the other hand, $\sharp J_{k,l}^c \leq Ck_n$ which means that

$$\frac{1}{(\psi_{HY}k_n)^2}\sum_{t_i\in J_{k,l}^c}b_i^n(k,l)\epsilon_{t_i}^k\epsilon_{t_i}^l \xrightarrow{p} 0.$$

Finally, note that the summands $\epsilon_{t_i^{(k)}}^k\epsilon_{t_j^{(l)}}^l\mathbb{I}_{\{t_i^{(k)}\neq t_j^{(l)}\}}$ have expectation 0 and are mutually uncorrelated. Since $\sharp F_{k,l} \leq Cnk_n$ this implies that

$$\frac{1}{(\psi_{HY}k_n)^2}\sum_{(i,j)\in F_{k,l}}c_{ij}^n(k,l)\epsilon_{t_i^{(k)}}^k\epsilon_{t_j^{(l)}}^l\mathbb{I}_{\{t_i^{(k)}\neq t_j^{(l)}\}} \xrightarrow{p} 0.$$

Hence,

$$HY[Y]_n^3 \xrightarrow{p} 0.$$

Now, we consider the term $HY[Y]_n^1$. We decompose

$$HY[Y]_n^1 = \frac{1}{(\psi_{HY}k_n)^2}$$

$$\times\left(\sum_{(i,j)\in I_{k,l}}\bar{a}_{ij}^n(k,l)\Delta_i^{n_k}X^k\Delta_j^{n_l}X^l\mathbb{I}_{\{(t_{i-1}^{(k)},t_i^{(k)}]\cap(t_{j-1}^{(l)},t_j^{(l)}]\neq\emptyset\}}\right.$$

$$+ \sum_{(i,j)\in I_{k,l}^c}\bar{b}_{ij}^n(k,l)\Delta_i^{n_k}X^k\Delta_j^{n_l}X^l\mathbb{I}_{\{(t_{i-1}^{(k)},t_i^{(k)}]\cap(t_{j-1}^{(l)},t_j^{(l)}]\neq\emptyset\}}$$

$$\left.+ \sum_{(i,j)\in F_{k,l}}\bar{c}_{ij}^n(k,l)\Delta_i^{n_k}X^k\Delta_j^{n_l}X^l\mathbb{I}_{\{(t_{i-1}^{(k)},t_i^{(k)}]\cap(t_{j-1}^{(l)},t_j^{(l)}]=\emptyset\}}\right), \quad (51)$$

for some constants $\bar{a}_{ij}^n(k,l)$, $\bar{b}_{ij}^n(k,l)$, $\bar{c}_{ij}^n(k,l)$, $I_{k,l} = \{(i,j) : k_n \leq i \leq n_k - k_n, k_n \leq j \leq n_l - k_n\}$ and $I_{k,l}^c = \{(i,j) : 1 \leq i \leq n_k, 1 \leq j \leq n_l\} - I_{k,l}$. Notice that all "border terms" are collected in the second summand whereas all terms with empty intersection of the intervals are in the third summand (in fact, we will see that both are negligible).

Notice that

$$|\bar{a}_{ij}^n(k,l)| + |\bar{b}_{ij}^n(k,l)| + |\bar{c}_{ij}^n(k,l)| \leq Cn.$$

Furthermore, we have

$$\mathbb{E}\left[|\Delta_i^{n_k}X^k\Delta_j^{n_l}X^l|\right] \leq C\sqrt{\left(t_i^{(k)}-t_{i-1}^{(k)}\right)\left(t_j^{(l)}-t_{j-1}^{(l)}\right)}$$

and $\sharp(I^c_{k,l} \cap \{(i,j) : (t^{(k)}_{i-1}, t^{(k)}_i] \cap (t^{(l)}_{j-1}, t^{(l)}_j] \neq \emptyset\}) \leq Ck_n$ by Eq. (25). This implies by Eq. (26) that

$$\frac{1}{(\psi_{HY} k_n)^2} \sum_{(i,j) \in I^c_{k,l}} \bar{b}^n_{ij}(k,l) \Delta^{n_k}_i Y^k \Delta^{n_l}_j Y^l \mathbb{I}_{\{(t^{(k)}_{i-1}, t^{(k)}_i] \cap (t^{(l)}_{j-1}, t^{(l)}_j] \neq \emptyset\}}$$

$$= O_p(n^{-1/2}),$$

and thus the second summand in Eq. (51) is negligible.

Now, recall that $\Delta^{n_k}_i X^k$ can be replaced with $(\sigma_s \Delta^{n_k}_i W)^k$ for any s with $t^{(k)}_i - s = O(k_n/n)$ ($1 \leq k \leq d$) without changing the first-order asymptotics (see, e.g., Podolskij and Vetter, 2009). Notice also that the terms $\Delta^{n_k}_i W^k \Delta^{n_l}_j W^l \mathbb{I}_{\{(t^{(k)}_{i-1}, t^{(k)}_i] \cap (t^{(l)}_{j-1}, t^{(l)}_j] = \emptyset\}}$ are mutually uncorrelated. Hence,

$$\frac{1}{(\psi_{HY} k_n)^2} \sum_{(i,j) \in F_{k,l}} \bar{c}^n_{ij}(k,l) \Delta^{n_k}_i X^k \Delta^{n_l}_j X^l \mathbb{I}_{\{(t^{(k)}_{i-1}, t^{(k)}_i] \cap (t^{(l)}_{j-1}, t^{(l)}_j] = \emptyset\}}$$

$$= o_p(1).$$

Finally, consider the first summand in Eq. (51). As above, we approximate $\Delta^{n_k}_i X^k \Delta^{n_l}_j X^l$ by

$$X^n_{i,j}(k,l) = \left(\sigma_{t^{(k)}_{i-1} \wedge t^{(l)}_{j-1}} \Delta^{n_k}_i W \right)^k \left(\sigma_{t^{(k)}_{i-1} \wedge t^{(l)}_{j-1}} \Delta^{n_l}_j W \right)^l$$

whenever $(t^{(k)}_{i-1}, t^{(k)}_i] \cap (t^{(l)}_{j-1}, t^{(l)}_j] \neq \emptyset$, and set

$$\overline{HY}[X]^{(k,l)}_n$$

$$= \frac{1}{(\psi_{HY} k_n)^2} \sum_{(i,j) \in I_{k,l}} \bar{a}^n_{ij}(k,l) X^n_{i,j}(k,l) \mathbb{I}_{\{(t^{(k)}_{i-1}, t^{(k)}_i] \cap (t^{(l)}_{j-1}, t^{(l)}_j] \neq \emptyset\}}.$$

Note that the quantities $HY[Y]^{(k,l)}_n$ and $\overline{HY}[X]^{(k,l)}_n$ have the same first-order asymptotics. Then, it can be shown that

$$\bar{a}^n_{ij}(k,l) = \left(\sum_{h=1}^{k_n-1} g\left(\frac{h}{k_n}\right) \right)^2 = k_n^2 \left(\int_0^1 g(x)\mathrm{d}x \right)^2 + o\left(k_n^2\right).$$

Next, note that

$$\mathbb{E}\left[\Delta^{n_k}_i W^a \Delta^{n_l}_j W^b \mathbb{I}_{\{(t^{(k)}_{i-1}, t^{(k)}_i] \cap (t^{(l)}_{j-1}, t^{(l)}_j] \neq \emptyset\}} \mid \mathcal{F}_{t^{(k)}_{i-1} \wedge t^{(l)}_{j-1}} \right]$$

$$= \delta_{ab} \left[\left(t^{(k)}_i \wedge t^{(l)}_j \right) - \left(t^{(k)}_{i-1} \vee t^{(l)}_{j-1} \right) \right]. \tag{52}$$

Eq. (52) and Riemann integrability then delivers the convergence

$$\frac{1}{(\psi_{HY} k_n)^2} \sum_{(i,j) \in I_{k,l}} \bar{a}^n_{ij}(k,l)$$

$$\times \mathbb{E}\left[X^n_{i,j}(k,l) \mathbb{I}_{\{(t^{(k)}_{i-1}, t^{(k)}_i] \cap (t^{(l)}_{j-1}, t^{(l)}_j] \neq \emptyset\}} \mid \mathcal{F}_{t^{(k)}_{i-1} \wedge t^{(l)}_{j-1}} \right] \xrightarrow{p} \int_0^1 \Sigma^{kl}_s \mathrm{d}s.$$

By usual martingale arguments we have that

$$\overline{HY}[X]^{(k,l)}_n - \frac{1}{(\psi_{HY} k_n)^2} \sum_{(i,j) \in I_{k,l}} \bar{a}^n_{ij}(k,l)$$

$$\times \mathbb{E}\left[X^n_{i,j}(k,l) \mathbb{I}_{\{(t^{(k)}_{i-1}, t^{(k)}_i] \cap (t^{(l)}_{j-1}, t^{(l)}_j] \neq \emptyset\}} \mid \mathcal{F}_{t^{(k)}_{i-1} \wedge t^{(l)}_{j-1}} \right] \xrightarrow{p} 0.$$

On the other hand, $HY[Y]^{(k,l)}_n - \overline{HY}[X]^{(k,l)}_n \xrightarrow{p} 0$. Thus, collecting terms produces

$$HY[Y]^{(k,l)}_n \xrightarrow{p} \int_0^1 \Sigma^{kl}_s \mathrm{d}s,$$

which completes the proof. □

Proof of Proposition 1. The terms $\bar{Y}^{k^n}_i \bar{Y}^{l^n}_j \mathbb{I}_{\{(t^{(k)}_i, t^{(k)}_{i+k_n}] \cap (t^{(l)}_j, t^{(l)}_{j+k_n}]\}}$ and $\bar{Y}^{k^n}_r \bar{Y}^{l^n}_s \mathbb{I}_{\{(t^{(k)}_r, t^{(k)}_{r+k_n}] \cap (t^{(l)}_s, t^{(l)}_{s+k_n}]\}}$ are (asymptotically) uncorrelated when the intervals $(t^{(k)}_i, t^{(k)}_{i+k_n}]$ and $(t^{(l)}_j, t^{(l)}_{j+k_n}]$ do not intersect with $(t^{(k)}_r, t^{(k)}_{r+k_n}]$ and $(t^{(l)}_s, t^{(l)}_{s+k_n}]$. Thus, due to the assumption in Eq. (25), there are $O(nk_n^3)$ correlated terms, and each covariance has order $O(n^{-1})$ (cf. Eq. (8)). This implies that $\mathrm{var}(HY[Y]^{(k,l)}_n) = O(n^{-1/2})$. □

References

Aït-Sahalia, Y., Mykland, P.A., Zhang, L., 2005. How often to sample a continuous-time process in the presence of market microstructure noise. Review of Financial Studies 18 (2), 351–416.

Aldous, D.J., Eagleson, G.K., 1978. On mixing and stability of limit theorems. Annals of Probability 6 (2), 325–331.

Andersen, T.G., Bollerslev, T., Diebold, F.X., 2010. Parametric and nonparametric volatility measurement. In: Hansen, L.P., Ait-Sahalia, Y. (Eds.), Handbook of Financial Econometrics. North-Holland, pp. 67–138.

Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2001. The distribution of realized exchange rate volatility. Journal of the American Statistical Association 96 (453), 42–55.

Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2003. Modeling and forecasting realized volatility. Econometrica 71 (2), 579–625.

Andersen, T.G., Bollerslev, T., Diebold, F.X., Wu, J., 2006. Realized beta: persistence and predictability. In: Fomby, T., Terrell, D. (Eds.), Advances in Econometrics: Econometric Analysis of Economic and Financial Time Series in Honor of R.F. Engle, and C.W.J. Granger, vol. B. pp. 1–40.

Back, K., 1991. Asset prices for general processes. Journal of Mathematical Economics 20 (4), 371–395.

Bandi, F.M., Russell, J.R., 2005. Realized covariation, realized beta and microstructure noise. Working Paper. Graduate School of Business, University of Chicago.

Bandi, F.M., Russell, J.R., 2006. Separating microstructure noise from volatility. Journal of Financial Economics 79 (3), 655–692.

Bandi, F.M., Russell, J.R., 2008. Microstructure noise, realized variance, and optimal sampling. Review of Economic Studies 75 (2), 339–369.

Barndorff-Nielsen, O.E., Graversen, S.E., Jacod, J., Podolskij, M., Shephard, N., 2006. A central limit theorem for realized power and bipower variations of continuous semimartingales. In: Kabanov, Y., Lipster, R., Stoyanov, J. (Eds.), From Stochastic Calculus to Mathematical Finance: The Shiryaev Festschrift. Springer-Verlag, pp. 33–68.

Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N., 2008a. Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. Econometrica 76 (6), 1481–1536.

Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N., 2008b. Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. Working Paper. University of Aarhus.

Barndorff-Nielsen, O.E., Shephard, N., 2002. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. Journal of the Royal Statistical Society: Series B 64 (2), 253–280.

Barndorff-Nielsen, O.E., Shephard, N., 2004. Econometric analysis of realized covariation: high frequency based covariance, regression, and correlation in financial economics. Econometrica 72 (3), 885–925.

Barndorff-Nielsen, O.E., Shephard, N., 2007. Variation, jumps, market frictions and high frequency data in financial econometrics. In: Blundell, R., Torsten, P., Newey, W.K. (Eds.), Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress. Cambridge University Press.

Boudt, K., Croux, C., Laurent, S., 2008. Outlyingness weighted quadratic covariation. Working Paper. Leuven, K.U.

Christensen, K., Oomen, R.C.A., Podolskij, M., 2010. Realised quantile-based estimation of the integrated variance. Journal of Econometrics, forthcoming (doi:10.1016/j.jeconom.2010.04.008).

Christensen, K., Podolskij, M., Vetter, M., 2010. A central limit theorem for the Hayashi–Yoshida estimator under microstructure noise and non-synchronous trading. Working Paper. University of Aarhus.

Corsi, F., Audrino, F., 2007. Realised correlation tick-by-tick. Working Paper. Available at: http://ssrn.com/abstract=957997.

de Jong, F., Nijman, T., 1997. High frequency analysis of lead–lag relationships between financial markets. Journal of Empirical Finance 4 (2–3), 259–277.

Delbaen, F., Schachermayer, W., 1994. A general version of the fundamental theorem of asset pricing. Mathematische Annalen 300, 463–520.

Diebold, F.X., Strasser, G.H., 2008. On the correlation structure of microstructure noise in theory and practice. Working Paper. University of Pennsylvania.

Dimson, E., 1979. Risk measurement when shares are subject to infrequent trading. Journal of Financial Economics 7 (2), 197–226.

Epps, T.W., 1979. Comovements in stock prices in the very short run. Journal of the American Statistical Association 74 (366), 291–298.

Fan, J., Zhang, J., Yu, K., 2009. Asset allocation and risk assessment with gross exposure constraints for vast portfolios. Working Paper. Princeton University.

Fisher, L., 1966. Some new stock-market indexes. Journal of Business 39 (1–2), 191–225.

Fleming, J., Kirby, C., Ostdiek, B., 2001. The economic value of volatility timing. Journal of Finance 56 (1), 329–352.

Fleming, J., Kirby, C., Ostdiek, B., 2002. The economic value of volatility timing using "realized" volatility. Journal of Financial Economics 67 (3), 473–509.

Ghysels, E., Harvey, A.C., Renault, E., 1996. Stochastic volatility. In: Maddala, G.S., Rao, C.R. (Eds.), Handbook of Statistics, vol. 14. North-Holland, pp. 119–191.

Glassermann, P., 2004. Monte Carlo Methods in Financial Engineering. Springer-Verlag.

Gloter, A., Jacod, J., 2001a. Diffusions with measurement errors. I — local asymptotic normality. ESAIM: Probability and Statistics 5, 225–242.

Gloter, A., Jacod, J., 2001b. Diffusions with measurement errors. II — measurement errors. ESAIM: Probability and Statistics 5, 243–260.

Griffin, J.E., Oomen, R.C.A., 2006. Covariance measurement in the presence of non-synchronous trading and market microstructure noise. Working Paper. University of Warwick.

Hansen, P.R., Lunde, A., 2006. Realized variance and market microstructure noise. Journal of Business and Economic Statistics 24 (2), 127–161.

Hayashi, T., Yoshida, N., 2005. On covariance estimation of non-synchronously observed diffusion processes. Bernoulli 11 (2), 359–379.

Hayashi, T., Yoshida, N., 2008. Asymptotic normality of a covariance estimator for nonsynchronously observed diffusion processes. Annals of the Institute of Statistical Mathematics 60 (2), 367–406.

Jacod, J., Li, Y., Mykland, P.A., Podolskij, M., Vetter, M., 2009. Microstructure noise in the continuous case: the pre-averaging approach. Stochastic Processes and their Applications 119 (7), 2249–2276.

Jacod, J., Shiryaev, A.N., 2003. Limit Theorems for Stochastic Processes, 2 ed. Springer-Verlag.

Jagannathan, R., Wang, Z., 1996. The conditional CAPM and the cross-section of expected returns. Journal of Finance 51 (1), 3–53.

Kinnebrock, S., Podolskij, M., 2008. An econometric analysis of modulated realised covariance, regression and correlation in noisy diffusion models. Working Paper. CREATES, University of Aarhus.

Large, J., 2007. Accounting for the Epps effect: realized covariation, cointegration and common factors. Working Paper. Oxford-Man Institute, University of Oxford.

Lettau, M., Ludvigson, S., 2001. Resurrecting the (C)CAPM: a cross-sectional test when risk premia are time-varying. Journal of Political Economy 109 (6), 1238–1287.

Lo, A., MacKinlay, C., 1990. An econometric analysis of nonsynchronous trading. Journal of Econometrics 45 (1–2), 181–211.

Malliavin, P., Mancino, M.E., 2002. Fourier series method for measurement of multivariate volatilities. Finance and Stochastics 6 (1), 49–61.

Martens, M., 2003. Estimating unbiased and precise realized covariances. Working Paper. Erasmus School of Economics, Rotterdam.

Mykland, P.A., Zhang, L., 2006. ANOVA for diffusions and Itô processes. Annals of Statistics 34 (4), 1931–1963.

Mykland, P.A., Zhang, L., 2009. Inference for continuous semimartingales observed at high frequency: A general approach. Econometrica 77 (5), 1403–1445.

Oomen, R.C.A., 2006. Properties of realized variance under alternative sampling schemes. Journal of Business and Economic Statistics 24 (2), 219–237.

Palandri, A., 2006. Consistent realized covariance for asynchronous observations contaminated by market microstructure noise. Working Paper. University of Copenhagen.

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

Renò, R., 2003. A closer look at the Epps effect. International Journal of Theoretical and Applied Finance 6 (1), 87–102.

Rényi, A., 1963. On stable sequences of events. Sankhya: The Indian Journal of Statistics; Series A 25 (3), 293–302.

Scholes, M., Williams, J., 1977. Estimating betas from nonsynchronous data. Journal of Financial Economics 5 (3), 309–327.

Voev, V., Lunde, A., 2007. Integrated covariance estimation using high-frequency data in the presence of noise. Journal of Financial Econometrics 5 (1), 68–104.

Zhang, L., 2006. Efficient estimation of stochastic volatility using noisy observations: a multi-scale approach. Bernoulli 12 (6), 1019–1043.

Zhang, L., 2008. Estimating covariation: Epps effect, microstructure noise. Working Paper. University of Illinois at Chicago.

Zhang, L., Mykland, P.A., Aït-Sahalia, Y., 2005. A tale of two time scales: determining integrated volatility with noisy high-frequency data. Journal of the American Statistical Association 100 (472), 1394–1411.

Zhou, B., 1996. High-frequency data and volatility in foreign-exchange rates. Journal of Business and Economic Statistics 14 (1), 45–52.