

Estimating the Quadratic Covariation Matrix for Asynchronously observed high frequency stock returns corrupted by Additive Measurement Error¹

Sujin Park Seok Young Hong Oliver Linton
GSA Capital Partners *Statistical Laboratory* *Faculty of Economics*
London *University of Cambridge* *University of Cambridge*

Abstract

This paper studies the estimation problem of the covariance matrices of asset returns in the presence of microstructure noise and asynchronicity between the observations across different assets. Motivated by Malliavin and Mancino (2002, 2009) we propose a new Fourier domain based estimator of multivariate ex-post volatility, which we call the Fourier Realized Kernel (FRK). An advantage of this approach is that no explicit time alignment is required unlike the time domain based methods widely adopted in the existing literature. We derive the large sample properties and establish asymptotic normality of our estimator under some general conditions that allow for both temporal and cross-sectional correlations in the measurement error process. Our results can be viewed as Frequency domain extension of the asymptotic theories for the multivariate realized kernel estimator of Barndorff-Nielsen et al. (2011). We show in extensive simulations that our method outperforms the time domain estimators when two assets with different liquidity are traded asynchronously.

KEYWORDS: Quadratic covariation, Market microstructure noise, Asynchronous observations, Fourier Realized Kernel

1 Introduction

Over the past two decades there have been many advances in the theory and application of volatility measurement from high frequency data. The ex-post measure of volatility called the quadratic variation has been the focus of much attention. The theory has been developed in a series of papers including but not limited to: Andersen, Bollerslev, Diebold and Labys (2001), Barndorff-Nielsen and Shephard (2002, 2004), and Mykland and Zhang (2006). This literature has recently been extended to account for what is called the microstructure noise, namely the measurement error that distorts the underlying efficient price diffusion: Zhang, Mykland and Aït-Sahalia (2005), Zhang (2006), Kalnina and Linton (2008), Barndorff-Nielsen, Hansen, Lunde and Shephard (2008, 2011), Aït-Sahalia, Mykland and Zhang (2011), and Bibinger, Hautsch, Malec and Reiss (2014). Jacod, Li, Mykland, Podolskij and Vetter (2009) introduced the pre-averaging method, which involves first averaging the observed prices over a moderate number of time points to reduce the measurement error. In the multivariate case an additional issue arises in the estimation procedure; since transactions most likely occur at different time points for different assets the observations exhibit asynchronicity. Hayashi and Yoshida (2005) proposed an estimator of the integrated covariance that do not require synchronization. However, their estimator is inconsistent under the presence of microstructure noise.

¹We thank Maria Elvira Mancino, Per Mykland, Victor Todorov and two anonymous referees for many helpful comments and suggestions. Financial support from the European Research Council (ERC) is gratefully acknowledged.

In a seminal paper Malliavin and Mancino (2009) proposed a Fourier domain approach that does not require data alignment, and established consistency and asymptotic mixed normality of their estimator under a general setting (but without deriving the limiting distribution under the presence of measurement error), extending their earlier work Malliavin and Mancino (2002). Since the first version of this paper was written, Mancino and Sanfelici (2008) have shown that their univariate estimator is consistent in the presence of measurement error; they also provide mean squared error expansions for their uniform weighting estimator. Furthermore, Curato, Mancino and Sanfelici (2014) derived the bias expression due to microstructure noise, and showed consistency of their estimator. Time domain estimators addressing both asynchronicity and microstructure noise have been proposed by Zhang (2011), Barndorff-Nielsen, Hansen, Lunde and Shephard (2011), and Aït-Sahalia, Fan and Xiu (2010). The estimators are consistent with convergence rates of $O(n^{1/6})$, $O(n^{1/5})$ and $O(n^{1/4})$, respectively. The first two papers require aligning the data, although consistency of their estimator is robust to the alignment. However, the hidden cost of data alignment and non-synchronicity for these estimators is that the sample size n that appears in the convergence rate is the sample size of the aligned data. Also, the drawback of Zhang (2011) and Aït-Sahalia et al. (2010) is that the estimator cannot be generalized to dimensions higher than two unless the covariance matrix is estimated element-wise, which in turn does not guarantee the estimated covariance matrix to be positive definite. See Park and Linton (2011) for a more detailed survey.

The goal of this paper is to propose an estimator of a general multivariate volatility measure that is robust both to the microstructure noise and asynchronous data timing. The method is based on Fourier domain techniques which have been widely used in discrete time series, and is broadly similar to that of Malliavin and Mancino (2009) although we allow for more general kernel weighting in the Fourier domain. An advantage of this approach is that it does not require an explicit time alignment. The by-product of the Fourier domain based estimator is that we have a consistent estimator of the instantaneous co-volatility even under the presence of quite general dependent microstructure noise. We provide a central limit theorem for our estimator under some general conditions, and also discuss the bandwidth choice issue based on the asymptotic mean squared error expressions. Our results allow for the unbalanced case where one series may have many more observations than another, which is common in practice since stocks vary considerably in terms of their trading intensities. In Section 2 we give a setup of the model and assumptions regarding the sampling scheme. In Section 3, we propose a Fourier domain based estimator of the integrated covariance. Section 4 studies the large sample asymptotics of the proposed estimator and derives its limiting distribution under the presence of microstructure noise. The Fourier method is further extended to estimate the instantaneous covariance matrix of diffusion process. We carry out extensive simulations and empirical analysis, and report the results in Section 5.

As regards the notations, we take $a \wedge b$ and $a \vee b$ to mean the smaller and bigger value of the scalars a and b , respectively. The operator $\Delta(\cdot)$ acting on a time series process or a function thereof is the differencing operator; for example, $\Delta t_j = t_j - t_{j-1}$ and $\Delta g(t_j) = g(t_j) - g(t_{j-1})$. We denote by \implies^{stably} stable convergence in law, see Aldous and Eagleson (1978) or Jacod (1997) for rigorous definition. Also, $a_n \simeq b_n$ means a_n and b_n are the same order of magnitude. The matrix norm $\|\cdot\| = \|\cdot\|_2$ is understood as the Frobenius norm unless specified otherwise. Throughout this paper we reserve the following square bracket notation $[\cdot, \cdot]$ to denote the quadratic variation of a stochastic process. Lastly, by stationarity we mean stationarity in wide sense rather than strict stationarity, unless stated specifically otherwise.

2 Model and assumptions

2.1 Efficient Price and Parameter of Interest

The following standard assumption on the efficient price process provides the general framework that will be used throughout this paper.

ASSUMPTION 1. *The efficient price process follows a Brownian semimartingale: For a $p \times 1$ vector of logarithmic prices $P(t) = (P_1(t), \dots, P_p(t))^\top$ defined on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$, we have*

$$P(t) = \int_0^t \mu(u) du + \int_0^t \sigma(u) dW(u),$$

where $\mu(u) = [\mu_1(u), \dots, \mu_p(u)]^\top$ is a vector of predictable locally bounded drifts, $\sigma(u)$ is a symmetric $p \times p$ matrix of locally bounded càdlàg process with finite integrated quarticity: $\int_0^t \sigma(u)\sigma(u)^\top \otimes \sigma(u)\sigma(u)^\top du < \infty$ a.s., and $W(u)$ is a $p \times 1$ vector of Brownian motion that is independent of the volatility process.

The assumptions of locally bounded drift and of diffusion coefficient are needed to apply Girsanov's theorem to remove the drift term in the theoretical derivation. Consider a discrete time grid $0 = t_0 < \dots < t_n = T$ where T is fixed, and denote by $P(t_i)$ the logarithmic price observed at t_i ; $i = 1, \dots, n$. The quadratic covariation matrix of P over a time interval $[0, t]$, for some $t \leq T$ is defined by

$$[P, P]_t := \text{plim}_{n \rightarrow \infty} \sum_{i: t_i \leq t} \{P(t_i) - P(t_{i-1})\} \{P(t_i) - P(t_{i-1})\}^\top. \quad (1)$$

The limit here is finite and well-defined with probability one, and is independent of the choice of the time grid if P is a semimartingale. Furthermore, under Assumption 1 one can show that (1) is almost surely equal to the integrated covariance matrix:

$$[P, P]_t = \int_0^t \sigma(u)\sigma(u)^\top du, \quad (2)$$

where $\sigma(u)\sigma(u)^\top =: \Sigma(u)$ is the instantaneous covariance matrix. We write $[P, P] := [P, P]_T$, and its j -th diagonal element $[P_j, P_j] = \int_0^T \Sigma_{j,j}(u) du$, the integrated variance of j -th asset. We note that the integrated covariance is related to the covariance matrix of prices by

$$\text{cov}\{P(t)\} = E \left[\int_0^t \sigma(u) dW(u) \left(\int_0^t \sigma(u) dW(u) \right)^\top \right] = \int_0^t E [\sigma(u)\sigma(u)^\top] du = E[P, P]_t.$$

A natural estimator of the quadratic covariation matrix is the realized covariance, the finite sum given in the right hand side of (1), which is consistent by construction. Barndorff-Nielsen and Shephard (2002) showed in the univariate framework that the realized variance is an unbiased and \sqrt{n} -consistent estimator of the integrated variance under Assumption 1.

From a practical viewpoint, two important issues arise in estimating the integrated covariance matrix (2). First, in the multivariate setting prices of different assets may be observed at different times, leading to the issue of asynchronicity in observations. Second, observed prices are distorted by some noise due to market microstructure effects, and do not satisfy Assumption 1. The objective of this paper is to propose an estimation theory that is robust to these two problems.

2.2 Sampling scheme

In this subsection we describe the main assumptions we make on the observation times. We allow for both *unequal spacing* and *asynchronicity* in random observation times. However, since they are assumed to be strictly exogenous (see Assumption 2 below), we shall work with the conditional distributions given the observation times; all statements below should hence be interpreted as stochastic boundedness/convergence.

ASSUMPTION 2. *The time span is fixed and is scaled to vary between $[0, 2\pi]$. The logarithmic prices are observed at discrete time points: $0 = t_{0,\ell} < \dots < t_{n_\ell,\ell} = 2\pi$ for $\ell = 1, \dots, p$, where n_ℓ is the total number of observations for the ℓ -th asset. The discrete time points are allowed to be stochastic and are assumed to be independent of the price and volatility process. For asymptotics, we let the smallest number of sample sizes amongst all assets $n := \min_\ell(n_\ell) \rightarrow \infty$. For $a, b, \ell \in \{1, \dots, p\}$:*

- 1) *The discrete time points satisfy $\sup_{1 \leq i \leq n_\ell} (t_{i,\ell} - t_{i-1,\ell}) = s^*/n_\ell = O(n_\ell^{-1})$, for some finite constant $s^* > 0$.*
- 2) *Denote the intervals $I_{i,a} = [t_{i-1,a}, t_{i,a})$ and $I_{j,b} := [t_{j-1,b}, t_{j,b})$. The degree of non-synchronicity satisfies*

$$\sup_{i,j} |t_{i,a} - t_{j,b}| \mathbf{1}_{\{I_{i,a} \cap I_{j,b} \neq \emptyset\}} = O\left(\frac{1}{n_a \wedge n_b}\right).$$

Given any set of bivariate time points $\{t_{i,a}, t_{j,b}\}$ with $n_a < n_b$, we assume that

$$\sup_{0 \leq j \leq n_b} \#\{t_{j,b} \in [t_{i-1,a}, t_{i,a}) \mid \mathbf{1}_{\{I_{i,a} \cap I_{j,b} \neq \emptyset\}}\} = O\left(\frac{n_a \vee n_b}{n_a \wedge n_b}\right).$$

- 3) *Define the empirical quadratic covariation process of time: for all $a, b, c, d \in \{1, \dots, p\}$,*

$$\begin{aligned} \mathcal{Q}_{aabb}^{(n)}(t) &= (n_a \wedge n_b) \sum_{i,j:t_{i,a}, t_{j,b} < t} \Delta t_{i,a} \Delta t_{j,b} \mathbf{1}_{\{I_{i,a} \cap I_{j,b} \neq \emptyset\}} \\ \mathcal{Q}_{abab}^{(n)}(t) &= (n_a \wedge n_b) \sum_{i,j,\ell:t_{i,a}, t_{j,b}, t_{\ell,b} < t} (t_{i,a} \wedge t_{j,b} - t_{i-1,a} \vee t_{j-1,b}) \\ &\quad \times (t_{i,a} \wedge t_{\ell,b} - t_{i-1,a} \vee t_{\ell-1,b}) \mathbf{1}_{\{I_{i,a} \cap I_{j,b} \neq \emptyset\}} \mathbf{1}_{\{I_{i,a} \cap I_{\ell,b} \neq \emptyset\}}, \text{ for } n_a < n_b. \end{aligned}$$

The empirical quadratic covariation satisfies $\mathcal{Q}_{abcd}^{(n)}(t) \rightarrow \mathcal{Q}_{abcd}(t)$ as $n_a \wedge n_b \rightarrow \infty$, where the limit $\mathcal{Q}_{abcd}(t)$ is continuously differentiable in t .

The array $\mathcal{Q}_{abcd}(t)$ will appear in the asymptotic variance of our estimator for the integrated variance. In the univariate case, the expression in Assumption 2-3) specializes to $\mathcal{Q}_{aa}^{(n)}(t) = n_a \sum_{i:t_{i,a} < t} (\Delta t_{i,a})^2$. As regards Assumption 2-2) we note that $\{\mathbf{1}_{\{I_{i,a} \cap I_{j,b} \neq \emptyset\}} = 1\}$ if and only if $\{u_{ij} := (t_{i,a} \wedge t_{j,b}) > (t_{i-1,a} \vee t_{j-1,b}) =: l_{i,j}\}$. Lastly, we see that the assumption does not restrict the ratio of sample sizes of different assets to be bounded away from zero or infinity (see also Assumption 4). In summary, both unequal spacing and asynchronicity in observation are allowed in a sufficiently general way.

We introduce some notations that will be used in the sequel. Define

$$\{T_l^{(ab)}\}_{1 \leq l \leq N_T^{(ab)}} := \{t_{i,a} \cup t_{j,b}; i = 1, \dots, n_a, j = 1, \dots, n_b\}, \quad (3)$$

where $N_T^{(ab)} (\leq n_a + n_b)$ is the total number of data points for the union of time stamps. Denote the average interval size for asset ℓ by $\overline{\Delta t}_\ell := 2\pi/n_\ell$. As for two different assets say a and b that are being compared, the average interval size of the more liquid asset is denoted by $\widehat{\Delta t}_{ab} := 2\pi/(n_a \vee n_b)$.

3 Estimation

3.1 The Fourier Kernel Estimator

Motivated by the disadvantages of data alignment methods widely discussed in the literature, see for example Aït-Sahalia and Jacod (2014) for relevant discussions, we propose to use a Fourier domain approach which does not require data synchronization. Our framework follows the line of approach of Malliavin and Mancino (2002, 2009), where a nonparametric method based on Fourier analysis of returns was discussed. Frequency domain techniques are widely used in discrete time series analysis; an important application of such an approach is the estimation problem of the long-run variance of a stationary time series (which is proportional to the spectral density at zero frequency). We draw a natural link between those classical theories and the estimation of the quadratic covariation of a continuous time processes.

The Fourier basis $\{g_t(q) := e^{iqt}/\sqrt{2\pi}; q \in \mathbb{Z}\}$ where $i^2 = -1$ constitutes an orthonormal basis of $L^2([0, 2\pi])$:

$$\frac{1}{2\pi} \int_0^{2\pi} g_t(k) \overline{g_t(j)} dt = \begin{cases} 1 & \text{if } k = j \\ 0 & \text{otherwise,} \end{cases}$$

where $\overline{g_t(q)}$ refers to the complex conjugate of $g_t(q)$.

In view of completeness of the Fourier basis, we can express the continuous time signal $\Sigma(t) \in L^2([0, 2\pi])$ as a linear combination of the Fourier basis with coefficients denoted by $\mathcal{F}(\Sigma)(q)$:

$$\Sigma(t) = \frac{1}{2\pi} \sum_{q=-\infty}^{\infty} \mathcal{F}(\Sigma)(q) e^{iqt}, \quad (4)$$

where

$$\mathcal{F}(\Sigma)(q) = \int_0^{2\pi} e^{-iqt} \Sigma(t) dt; \quad q \in \mathbb{Z}. \quad (5)$$

This is the continuous time Fourier transform of the instantaneous covariance matrix. Note that (5) corresponds to the integrated covariance when $q = 0$; we will propose an estimator for the general q case. The Fourier pair above suggests that once we estimate the Fourier coefficient by a suitable estimator $\widehat{\mathcal{F}}(\Sigma)(q)$, the signal can be reconstructed via:

$$\widehat{\Sigma}(t) = \frac{1}{2\pi} \sum_{q=-n}^n \widehat{\mathcal{F}}(\Sigma)(q) e^{iqt}.$$

We now discuss the estimation procedure of (5) based on Fourier transform of the return process. Define

$$\mathcal{F}(dP_\ell)(\alpha) = \int_0^{2\pi} e^{-i\alpha t} dP_\ell(t); \quad \alpha \in \mathbb{Z}, \ell = 1, \dots, p, \quad (6)$$

where $P_\ell(t)$ refers to observed price of the ℓ -th asset satisfying Assumption 1, and let

$$\mathcal{F}_n(dP_\ell)(\alpha) = \sum_{j=1}^{n_\ell} e^{-i\alpha t_{j,\ell}} \Delta P_\ell(t_{j,\ell}), \quad (7)$$

whose vector version $\mathcal{F}_n(dP)(\alpha) = \{\mathcal{F}_n(dP_1)(\alpha), \dots, \mathcal{F}_n(dP_p)(\alpha)\}^\top$, $\alpha \in \mathbb{Z}$, can be defined to account for multiple assets altogether.

Now we consider a class of kernel called the spectral window $K_H : [-\pi/H, \pi/H] \rightarrow \mathbb{R}$, where $K_H(\lambda) \simeq H \cdot K(H\lambda)$ and $K(\cdot)$ is a function called the spectral window generator. Detailed analysis of these weighting functions is given in the next subsection (Assumption

3', (13) and their relevant discussions); see also Parzen (1967) and Rao and Gabr (1984) for earlier discussions. The estimator we propose for the integrated covariance (5) is given by

$$\widehat{\mathcal{F}}(\Sigma)(q) = \left(\widehat{\mathcal{F}}(\Sigma_{i,j})(q) \right) = \sum_{|\alpha| \leq m/2} K_H(\lambda_\alpha) \left[\mathcal{F}_n(dP)(\alpha) \right] \left[\mathcal{F}_n(dP)(q - \alpha) \right]^\top, \quad (8)$$

where $\lambda_\alpha = 2\pi\alpha/\bar{n}$ and $\bar{n} := \max_{\ell=1,\dots,p} n_\ell$. Let $m = \bar{n}/H$ where the bandwidth $H = H_n \rightarrow \infty$ but not as fast as $n = \min_\ell n_\ell \rightarrow \infty$ so that $m \rightarrow \infty$. We are smoothing over the interval $[-\pi/H, \pi/H]$, where H controls the width of the smoothing window. We shall refer to our estimator as the *Fourier Realized Kernel* (FRK). Note that the estimator considered in Malliavin and Mancino (2009) can be compared with ours with $K(x) = (1/\pi)(1 - |x|/\pi)$.

The $q = 0$ case deserves a special attention. In this case $\mathcal{F}_n(dP_1)(\alpha)\mathcal{F}_n(dP_2)(-\alpha) =: I_{12}(\alpha)$ is the realized cross periodogram between assets 1 and 2, say. Then, the $(1, 2)$ -th entry of $\widehat{\mathcal{F}}(\Sigma)(0)$ is given by kernel smoothing the realized cross periodogram around zero frequency:

$$\widehat{\mathcal{F}}(\Sigma_{12})(0) = \sum_{|\alpha| \leq m/2} K_H(\lambda_\alpha) I_{12}(\alpha). \quad (9)$$

Positive definiteness of the estimators (9) is guaranteed provided that the spectral window is non-negative.

3.2 Comparison with some Time domain estimators

For data that is synchronized at $\{\tau_i\}$, its realized autocovariance function is defined as

$$\gamma_{12}(h) = \sum_i \Delta P_1(\tau_i) \Delta P_2(\tau_{i-h}); \quad h \in \mathbb{Z}, \quad (10)$$

where $\sum_i = \sum_{h < i \leq n}$ for $h \geq 0$, and $\sum_i = \sum_{1 \leq i \leq n+h}$ for $h < 0$. In the aligned case the realized periodogram is closely related to the realized autocovariance. In particular when τ_i are equally spaced and synchronous, i.e. $\tau_i = \tau_j + (i - j)2\pi/n$, it can be easily shown that the realized cross periodogram is the Fourier transform of the realized autocovariance; that is, $I_{12}(\alpha) = \sum_{|h| < n} e^{-i\alpha h 2\pi/n} \gamma_{12}(h)$.

Hayashi and Yoshida (2005) considered a covariation estimator defined as the realized cross periodogram at zero frequency over the interval that overlaps:

$$HY = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \Delta P_1(t_{i,1}) \Delta P_2(t_{j,2}) 1_{\{I_{i,1} \cap J_{j,2} \neq \emptyset\}}.$$

Then the estimator can be expressed in terms of the centered (i.e. frequency zero) realized cross periodogram $I_{1,2}(0)$. In particular, we have the following decomposition:

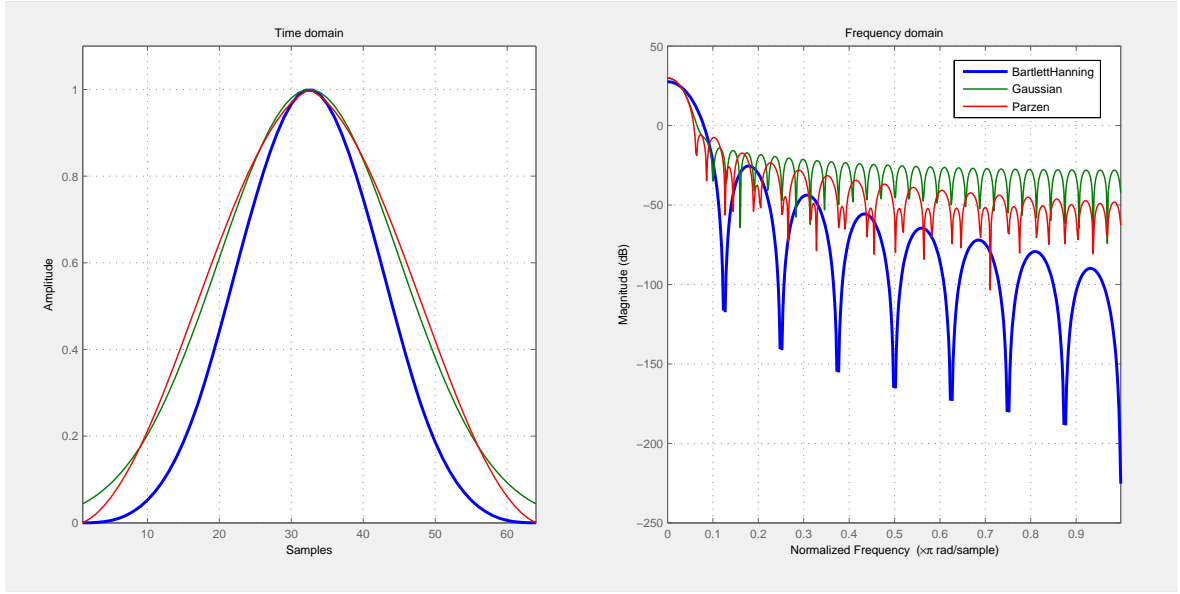
$$I_{1,2}(0) - [P_1, P_2] = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \Delta P_1(t_{i,1}) \Delta P_2(t_{j,2}) - \int_0^{2\pi} \Sigma_{12}(t) dt = M_1 + M_2,$$

where

$$M_1 = HY - \int_0^{2\pi} \Sigma_{12}(t) dt, \quad M_2 = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \Delta P_1(t_i) \Delta P_2(s_j) 1_{\{I_{i,1} \cap J_{j,2} = \emptyset\}}.$$

Hayashi and Yoshida (2008) showed that $\sqrt{n}M_1$ is asymptotically zero mean Gaussian (when data is Poisson sampled), and M_2 is mean zero with leading term of order $O_p(1)$. Hence, provided there is no microstructure noise the Hayashi and Yoshida estimator is unbiased,

Figure 1: Examples of the lag and spectral windows satisfying Assumption 3 and 3'



and achieves \sqrt{n} -consistency without requiring synchronization. The realized periodogram is also unbiased but is inconsistent due to the presence of M_2 .

We next compare our estimator (9) with the (multivariate) *realized kernel* estimator by Barndorff-Nielsen et al. (2008, 2011), denoted by $\tilde{\Sigma}$. It is given by kernel smoothing the realized autocovariances of the data aligned on $\{\tau_i\}_{i=1}^N$ using the refresh time sampling scheme. For example, its (1, 2)th entry is

$$\tilde{\Sigma}_{12} := \sum_{|h| < n} k\left(\frac{h}{H}\right) \gamma_{12}(h) = \sum_{i=1}^N \sum_{j=1}^N \Delta P_1(\tau_i) \Delta P_2(\tau_j) k\left(\frac{i-j}{H}\right), \quad (11)$$

where $k(\cdot)$ is a smoothing window in time domain. To establish the link between the time domain estimator (11) and our frequency domain estimator (9), we discuss the properties of smoothing windows; we introduce the lag window for continuous time denoted by $k(x)$, $x \in \mathbb{R}$ and compare it with the spectral kernel for continuous and bandlimited frequency. The lag window k is assumed to satisfy the following conditions (introduced as Assumption K in Barndorff-Nielsen et al. (2011)). Note that the prime notation is taken to mean differentiation with respect to the argument.

ASSUMPTION 3. *The lag window $k(\cdot)$ satisfies the following conditions: (i) k is twice continuously differentiable; (ii) $k(0) = 1$, and $k'(0) = 0$; (iii) $\|k\|^2 := \int_{-\infty}^{\infty} |k(x)|^2 dx < \infty$, $\|k^2\|^2 := \int_{-\infty}^{\infty} |k(x)|^4 dx < \infty$, $\|k'\|^2 := \int_{-\infty}^{\infty} |k'(x)|^2 dx < \infty$, $\|k''\|^2 := \int_{-\infty}^{\infty} |k''(x)|^2 dx < \infty$; (iv) $\int_{-\infty}^{\infty} k(x) \exp(-i\lambda x) dx \geq 0$, $\forall \lambda \in [-\pi, \pi]$.*

We define the spectral window generator K as a Fourier transform of the lag window k :

$$K(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} k(t) e^{-i\lambda t} dt \quad k(t) = \int_{-\pi}^{\pi} K(\lambda) e^{i\lambda t} d\lambda, \quad (12)$$

where λ denotes the angular frequency. Then it follows from simple algebra and Parseval's identity that Assumption 3 on the lag window k translates to the following conditions on the spectral window generator.

ASSUMPTION 3'. The spectral window generator $K(\cdot)$ satisfies the following conditions: (i) $\int_{-\pi}^{\pi} K(\lambda)d\lambda = 1$, $\int_{-\pi}^{\pi} \lambda K(\lambda)d\lambda = 0$; (ii) $\|K\|^2 := \int_{-\pi}^{\pi} |K(\lambda)|^2 d\lambda < \infty$, $\mu_1^2(K) := \int_{-\pi}^{\pi} |\lambda K(\lambda)|^2 d\lambda < \infty$ and $\mu_2^2(K) := \int_{-\pi}^{\pi} |\lambda^2 K(\lambda)|^2 d\lambda < \infty$; (iii) $K(\lambda) \geq 0$, $\forall \lambda \in [-\pi, \pi]$.

Figure 1 shows some examples of weighting functions k and K that satisfy Assumption 3 and 3'. In this sequel, Assumption 3 is taken to mean *both* Assumptions 3 and 3'.

As remarked by Barndorff-Nielsen et al. (2011), condition (iv) in Assumption 3 (and equivalently (iii) in Assumption 3') guarantees positive semi-definiteness of the estimators defined in (11) and (9) via Bochner's theorem. The realized periodogram is also positive semi-definite and is Hermitian as long as the spectral window is non-negative.

Consider the discrete time and discrete frequency Fourier pair:

$$K_H(\lambda_\alpha) = \frac{1}{2\pi} \sum_{|h| \leq H} k\left(\frac{h}{H}\right) e^{i\lambda_\alpha h}; \quad k\left(\frac{h}{H}\right) = \sum_{\alpha=-m/2}^{m/2-1} K_H(\lambda_\alpha) e^{-i\lambda_\alpha h}. \quad (13)$$

We refer to the weighting function K_H defined in (13) as the spectral window. To avoid the aliasing problem we assume that the signal is zero for frequencies that falls outside of the Nyquist critical frequency. As shown in Parzen (1967, page 130), the spectral window that is used to smooth the realized periodogram is related to a spectral window generator via

$$K_H(\lambda) = H \sum_{j=-\infty}^{\infty} K(H(\lambda + 2\pi j)) \simeq HK(H\lambda).$$

Now with $\lambda_\alpha = 2\pi\alpha/(n_a \vee n_b)$, $\alpha \in \mathbb{Z}$ for some $a, b = 1, \dots, p$, upon substituting $h = (n_a \vee n_b) \cdot (t_i - s_j)/(2\pi)$ into the latter expression of (13) we see that

$$\begin{aligned} \sum_{|\alpha| < m/2} K_H(\lambda_\alpha) I_{ab}(\alpha) &= \sum_{i=1}^{n_a} \sum_{j=1}^{n_b} \Delta P_a(t_i) \Delta P_b(s_j) \sum_{|\alpha| < m/2} K_H(\lambda_\alpha) e^{-i\alpha(t_i - s_j)} \\ &= \sum_{i=1}^{n_a} \sum_{j=1}^{n_b} \Delta P_a(t_i) \Delta P_b(s_j) k_H(t_i - s_j), \end{aligned} \quad (14)$$

where $k_H(t_i - s_j) := k((t_i - s_j)/(\widetilde{\Delta} t_{ab} H))$. Further, since an equally spaced and synchronized time grid satisfies $t_i = s_j + (i - j)2\pi/(n_a \vee n_b)$, the following key statement trivially holds:

REMARK 1. When trading times are synchronized and equally spaced, the Fourier realized kernel at zero frequency (9) and the multivariate realized kernel (11) are identical.

We will show later that when the data is not synchronously observed, using all the data and implementing (14) delivers a superior estimator.

It is of interest how our Fourier kernel estimator is related to other time domain estimators such as the multivariate two time scale estimator of Zhang (2011), and the modulated realized covariance (multivariate pre-averaging estimator) of Christensen, Kinnebrock and Podolskij (2010). In the univariate setting, Jacod et al. (2009) showed that their pre-averaging estimator, the univariate two time scale estimator of Zhang et al. (2005), and the flat-top realized kernel of Barndorff-Nielsen et al. (2008) can be written as a smoothed realized autocovariances, and the difference between the estimators comes from the contribution of the end points. This result holds also for the multivariate versions of the three estimators when observation points are synchronized. Our estimator can be expressed as a realized kernel only when sampling points are equally spaced and aligned. The relation between the smoothed periodogram to estimate the spectrum and data tapering (i.e. Fourier transforming the weighted return) is analogous to the relation between our estimator and the pre-averaging estimator.

4 Asymptotic Properties

4.1 Bandwidth conditions

We introduce the rate conditions on the bandwidth we require for our asymptotic theories. Our conditions allow the sample sizes of different assets to be of different order of magnitude. As aforementioned, such a situation arises often in practice as some assets may be traded much more frequently than others. As defined previously in Assumption 2, n is taken to mean the minimum amongst the individual sample sizes of the assets throughout, unless stated otherwise.

ASSUMPTION 4. *The bandwidth sequence $H = H_n$ is of order $H = O(n^\alpha)$ where $\alpha \in (0, 1)$ so that $H \rightarrow \infty$ and $n/H (=: m) \rightarrow \infty$ as $n \rightarrow \infty$. Moreover, the ratio of the bigger to smaller sample size between two assets satisfies $(n_a \vee n_b)/(n_a \wedge n_b) = o(H)$ for any $a, b \in \{1, \dots, p\}$.*

The first condition regulates the speed at which the bandwidth is allowed to increase; this is consistent with the usual setting in nonparametric literature. The next requirement $(n_a \vee n_b)/(H(n_a \wedge n_b)) = o(1)$ specifies the asymptotic behaviour of the ratio of the larger to smaller sample sizes; this will be shown to control the bias from the leading term. We define $\beta (= \max_{a,b} \beta_{ab}) \geq 1$ to be the degree of liquidity parameter where

$$\beta_{ab} = \lim_{n \rightarrow \infty} \frac{\log(n_a \vee n_b)}{\log(n_a \wedge n_b)}. \quad (15)$$

The definition suggests that if we write $n_a \wedge n_b =: n_{ab}$, then we have $n_a \vee n_b = n_{ab}^{\beta_{ab}}$; this in turn implies that $1 \leq \beta_{ab} < 2$ for all $a, b \in \{1, \dots, p\}$ because otherwise $(n_a \vee n_b)/(H(n_a \wedge n_b))$ would not converge, violating Assumption 4. As a slight abuse of notation, we will hereafter omit the subscript of β whenever it is clear from the context which assets are being considered.

4.2 Assumption on the microstructure noise

Empirical evidence from the volatility signature plot suggests that the observed price deviates from the semimartingale assumption. In fact, various studies document that the observed high frequency returns have infinite quadratic variation. The following assumption is proposed to account for this phenomena.

ASSUMPTION 5. *The observed logarithmic price of the ℓ -th asset, denoted $X_\ell(t_{j,\ell})$, is the sum of two components. The first component is a discretely observed continuous signal $P_\ell(t_{j,\ell})$ that satisfies Assumption 1, and the other one is a noise process with respect to the realization of transaction time $U_\ell(t_{j,\ell})$ that has an infinite quadratic variation: i.e. for each $\ell = 1, \dots, p$*

$$X_\ell(t_{j,\ell}) = P_\ell(t_{j,\ell}) + U_\ell(t_{j,\ell}). \quad (16)$$

This type of additive microstructure noise model has been well-studied in the literature, particularly in the univariate context; see for recent examples, Mancini and Sanfelici (2008) and Curato et al. (2014) where the noise $U_\ell(t_{j,\ell})$ is assumed to be i.i.d. There has not been much empirical work that studied cross autocorrelation of the microstructure noise for multiple assets. Amongst the few includes Ait-Sahalia et al. (2010) where the noise is set to be i.i.d in time but uncorrelated across different assets. Zhang (2011) assumed (covariance-) stationarity and an exponential alpha mixing condition with respect to the observation time.

In this paper we shall impose the following dependence condition on the microstructure noise:

ASSUMPTION 6. *The stochastic process $U_\ell(\cdot)$ is stationary, mean zero, and is independent of the efficient price process $P_\ell(\cdot)$. Furthermore, the covariance function of the noise process defined as $E[U_a(t_{i,a})U_b(t_{j,b})] =: \gamma(|t_{i,a} - t_{j,b}|/\widetilde{\Delta t}_{ab})$ satisfies*

$$\frac{1}{n_a \wedge n_b} \sum_{i=1}^{n_a-1} \sum_{j=1}^{n_b-1} \left| \gamma \left(\frac{|t_{i,a} - t_{j,b}|}{\widetilde{\Delta t}_{ab}} \right) \right| \rightarrow \Gamma_{ab} (< \infty), \quad (17)$$

where the finite limit Γ_{ab} is the (a, b) -th element of the $p \times p$ positive semi-definite covariance matrix Γ . Also, the fourth moments satisfy $|E(U_a(t_{i,a})U_b(t_{j,b})U_c(t_{r,c})U_d(t_{l,d}))| \leq \rho(M) < \infty$, where $M := \sup_{u,v,g,s} [(t_{u,g} - t_{v,s})/\widetilde{\Delta t}_{gs}]$ and $\rho(\cdot)$ is a function such that for some $\delta > 0$,

$$\sum_{\nu=1}^{\infty} \rho(\nu)(1 + \delta)^\nu < \infty. \quad (18)$$

Our assumption allows both cross-sectional and temporal correlations in the measurement error process. Note that when the data is equally spaced and balanced, (17) simplifies to $\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n |\gamma(i - j)| \rightarrow \Gamma_{ab} < \infty$. This standard absolute summability condition for the autocovariance function is often assumed in the time series literature, and is well-known to be sufficient for ergodicity and necessary for (a certain class of) mixing under stationarity, see Ibragimov and Linnik (1971). The second condition (18) implies that the supremum of the fourth moment decays sufficiently fast as the maximum cross-lag $M = \max(|i - j|, |h - l|)$ increases.

4.2.1 Asymptotic Normality

We now introduce our main theoretical results. At the end, the limiting distribution of our estimator (8) will be derived under the presence of the microstructure noise. Before we proceed, we impose a condition on the end points. This condition is in line with what is assumed in the existing literature in order to ensure consistency of the estimator at the end points, see Jacod et al. (2009) and Barndorff-Nielsen et al. (2011, section 2.2).

ASSUMPTION 7. *The observed prices at two end points, i.e. $X_\ell(t_{0,\ell})$ and $X_\ell(t_{n_\ell,\ell})$ for each $\ell = 1, \dots, p$ are respectively the arithmetic average of ξ number of distinct observations on the interval $[t_{-1,\ell}, t_{0,\ell})$ and $[t_{n_\ell,\ell}, t_{n_\ell+1,\ell})$.*

Since Assumption 5 implies $\Delta X(t_i) = \Delta P(t_i) + \Delta U(t_i)$, our estimator (8) can be decomposed as follows:

$$\begin{aligned} & \widehat{\mathcal{F}}(\Sigma)(q) - \mathcal{F}(\Sigma)(q) \\ &= \sum_{|\alpha| \leq m/2} K_H(\lambda_\alpha) \left\{ \left[\mathcal{F}_n(dP)(\alpha) \right] \left[\mathcal{F}_n(dP)(q - \alpha) \right]^\top - \left[\mathcal{F}(dP)(\alpha) \right] \left[\mathcal{F}(dP)(q - \alpha) \right]^\top \right\} \\ &+ \sum_{|\alpha| \leq m/2} K_H(\lambda_\alpha) \left[\mathcal{F}(dP)(\alpha) \right] \left[\mathcal{F}(dP)(q - \alpha) \right]^\top - \mathcal{F}(\Sigma)(q) \\ &+ \sum_{|\alpha| \leq m/2} K_H(\lambda_\alpha) \left[\mathcal{F}_n(dU)(\alpha) \right] \left[\mathcal{F}_n(dU)(q - \alpha) \right]^\top \\ &+ \sum_{|\alpha| \leq m/2} K_H(\lambda_\alpha) \left\{ \left[\mathcal{F}_n(dU)(\alpha) \right] \left[\mathcal{F}_n(dP)(q - \alpha) \right]^\top + \left[\mathcal{F}_n(dP)(\alpha) \right] \left[\mathcal{F}_n(dU)(q - \alpha) \right]^\top \right\} \\ &= (i) + (ii) + (iii) + (iv). \end{aligned} \quad (19)$$

The first term (i) is the discretization error due to sampling the continuous time signal at discrete points, and (ii) is the error due to kernel smoothing. Moreover, (iii) can be interpreted as the contribution from the smoothed realized periodogram applied to a microstructure noise. Lastly, (iv) is the sum of cross terms between the efficient price and the noise, which is of a smaller order than (iii) , see Theorem 1 of Barndorff-Nielsen et al. (2011). Below we derive the asymptotic order of the noise contribution (iii) , and establish asymptotic normality of $(i) + (ii)$. All proofs are contained in the appendix.

PROPOSITION 1. *Suppose Assumptions 2-7 hold. Then for any $a, b = 1, \dots, p$ we have*

$$E \left\{ \sum_{|\alpha| \leq m/2} K_H(\lambda_\alpha) [\mathcal{F}_n(dU_a)(\alpha)] [\mathcal{F}_n(dU_b)(q - \alpha)] \right\} \simeq \left(\frac{s^*}{2\pi} \right)^2 \frac{n_a \vee n_b}{H^2} |k''(0)| \Gamma_{ab}$$

$$E \left\{ \left| \sum_{|\alpha| \leq m/2} K_H(\lambda_\alpha) [\mathcal{F}_n(dU_a)(\alpha)] [\mathcal{F}_n(dU_b)(q - \alpha)] \right|^2 \right\} = O \left(\frac{(n_a \vee n_b)^2}{n_a \wedge n_b} H^{2\mu-3} \right),$$

where μ is some constant such that $0 < \mu < 1$, and s^* and Γ_{ab} are as defined in Assumption 2 and Assumption 6, respectively.

The proposition above implies that the leading bias term (iii) is of order $O_p((n_a \vee n_b)/H^2)$; a necessary condition for this bias contribution from the microstructure noise to be negligible in large sample is $\alpha \in (1/2, 1)$. We now move on to the main term $(i) + (ii) =: \widehat{\mathcal{F}}(\Sigma^P)(q) - \mathcal{F}(\Sigma)(q)$, and derive its asymptotic bias and variance expressions.

PROPOSITION 2. *Suppose Assumptions 1-4 hold. Then the bias and variance of $(\widehat{\mathcal{F}}(\Sigma^P)(q))_{ab}$ ($a, b = 1, \dots, p$) denoted respectively by \mathbf{B}_{ab} and \mathbf{V}_{ab} are given by:*

$$\mathbf{B}_{ab} = \left(\frac{n_a \vee n_b}{(n_a \wedge n_b)H} \right)^2 \frac{1}{2} \mathcal{A}_{ab}^2 |k''(0)| \int_0^{2\pi} e^{-itq} d[[P_a, P_b]](t)$$

$$\mathbf{V}_{ab} = \frac{H}{n_a \wedge n_b} \|k\|^2 \left[\int_0^{2\pi} e^{-i2tq} [P_a, P_a]'(t) [P_b, P_b]'(t) d\mathcal{Q}_{aabb}(t) \right. \\ \left. + \int_0^{2\pi} e^{-i2tq} ([P_a, P_b]')^2(t) d\mathcal{Q}_{abab}(t) \right],$$

with $\mathbf{B}_{aa} = 0$ (for all $a = 1, \dots, p$), where

$$\mathcal{A}_{ab} := \lim_{n_a \wedge n_b \rightarrow \infty} \frac{n_a \wedge n_b}{2\pi} \sup_{i,j} |t_i - s_j| \mathbf{1}_{\{I_{i,a} \cap I_{j,b} = \emptyset\}}.$$

This result gives the asymptotic mean squared error of our estimator in the absence of microstructure noise but with asynchronous trading. It may be compared with Barndorff-Nielsen et al. (2011) Theorem A.5. Malliavin and Mancino (2009) Theorem 4.1 gives a CLT for their estimator in the absence of microstructure noise. Our result in Proposition 2 holds under different restrictions on the bandwidth sequence, and so is not directly comparable.²

We note that \mathcal{A}_{ab} can be thought of as a measure of the degree of non-synchronicity between two assets a and b ; Assumption 2 implies that $\mathcal{A}_{ab} \in [0, \infty)$. In particular, when two series in consideration are perfectly synchronized and balanced, then $\mathcal{A}_{ab} = 0$. Otherwise it is a bounded constant i.e. $O(1)$ by Assumption 2.2.

²Note that their estimator has a specific weighting, and models the asynchronicity in a slightly different way.

From Proposition 2 we see that the order of asymptotic variance is $O(H/(n_a \wedge n_b))$. The squared bias is given by $O((n_a \vee n_b)^2/H^4)$ as this dominates that of the main term. The optimal bandwidth $H := O(n^\alpha)$ is obtained by balancing those two quantities. Denote by $n = (n_a \wedge n_b) = n_a$ say, and let $n_b = (n_a \vee n_b) = O(n^\beta)$. Then solving $n^{2\beta-4\alpha} = n^{\alpha-1}$ we have

$$H = C_0 n^{\alpha_{\text{opt}}}, \quad \alpha_{\text{opt}} = \frac{2\beta + 1}{5}, \quad (20)$$

where C_0 is some positive constant. In particular, when two sample sizes n_a and n_b are of the same order (i.e. $\beta = 1$), then the optimal order of the bandwidth is given by $\alpha_{\text{opt}} = 3/5$. The rate of convergence of the estimator is therefore $(n_a \wedge n_b)^\vartheta$, where $\vartheta = (2-\beta)/5 \in (0, 1/5]$. This suggests that the best rate $(n_a \wedge n_b)^{1/5}$ is achieved when the sample sizes are of the same order, and the rate slows down as $\beta \in [1, 2)$, the degree of relative liquidity between assets in consideration increases.

THEOREM 1. *Suppose that Assumptions 1-7 hold. Then provided that the optimal bandwidth (20) is chosen, we have for each $q \in \mathbb{Z}$*

$$\mathcal{D}_n \text{vech} \left\{ \widehat{\mathcal{F}}(\Sigma)(q) - \mathcal{F}(\Sigma)(q) \right\} \Longrightarrow^{\text{stably}} N(\mathcal{B}, \mathcal{V})$$

as $n \rightarrow \infty$, where

$$\begin{aligned} \mathcal{D}_n &= \text{diag} \{ \text{vech}(\mathcal{D}_n^*) \}, \quad \{ \mathcal{D}_n^* \}_{a,b} = (n_a \wedge n_b)^\vartheta, \quad \vartheta = \frac{2-\beta}{5}, \\ \mathcal{B} &= (2\pi)^{-2} C_0^{-2} s^{*,2} |k''(0)| \cdot \text{vech}(\Gamma), \quad \mathcal{V} = (\mathcal{V}_{ab}), \\ \mathcal{V}_{ab} &= \|k\|^2 \int_0^{2\pi} e^{-i2tq} \left\{ [P_a, P_a]'(t) [P_b, P_b]'(t) d\mathcal{Q}_{aabb}(t) + ([P_a, P_b]')^2(t) d\mathcal{Q}_{abab}(t) \right\}. \end{aligned}$$

4.3 Estimation of the Instantaneous covariance matrix

Here we extend the Fourier method discussed above to estimate the instantaneous covariance matrix of the diffusion process under the presence of microstructure noise. We can construct an estimator of the instantaneous covariation matrix by Fourier inverting (8):

$$\widehat{\Sigma}(t) = \frac{1}{2\pi} \sum_{|q| \leq m/2} K_H(\lambda_q) \exp(iqt) \widehat{\mathcal{F}}(\Sigma)(q). \quad (21)$$

Suppose that the modulus of continuity of $\Sigma(t)$ denoted by $\mathcal{C}(h)$ is given by

$$\mathcal{C}(h) := \sup_{|t-s| \leq h} \|\Sigma(t) - \Sigma(s)\|_2. \quad (22)$$

The continuity assumption is met when each element of $\Sigma(t)$ in Assumption 1 does not contain jumps, for example $\Sigma(t)$ is a Brownian semimartingale.

THEOREM 2. *Suppose that the regularity conditions of Theorem 1 and (22) hold. Then, there exists a sequence $\delta(n) \rightarrow 0$ such that*

$$\lim_{n \rightarrow \infty} \sup_{\delta(n) \leq t \leq 2\pi - \delta(n)} \|\widehat{\Sigma}(t) - \Sigma(t)\|_2 = 0.$$

Practitioners often encounter a problem of running a regression between variables that are asynchronously observed; for example, we might be interested in the effect of returns and order book information of one asset on another asset. Hannan (1975) and Robinson (1975) are the earlier literature on using frequency domain to solve such problems. Mykland and Zhang (2006) discussed a general setup of the analysis of variance for continuous time regression.

5 Numerical Study

5.1 Comparison of estimators of co-volatility

Recall that our estimator (at zero frequency $q = 0$) can be written as:

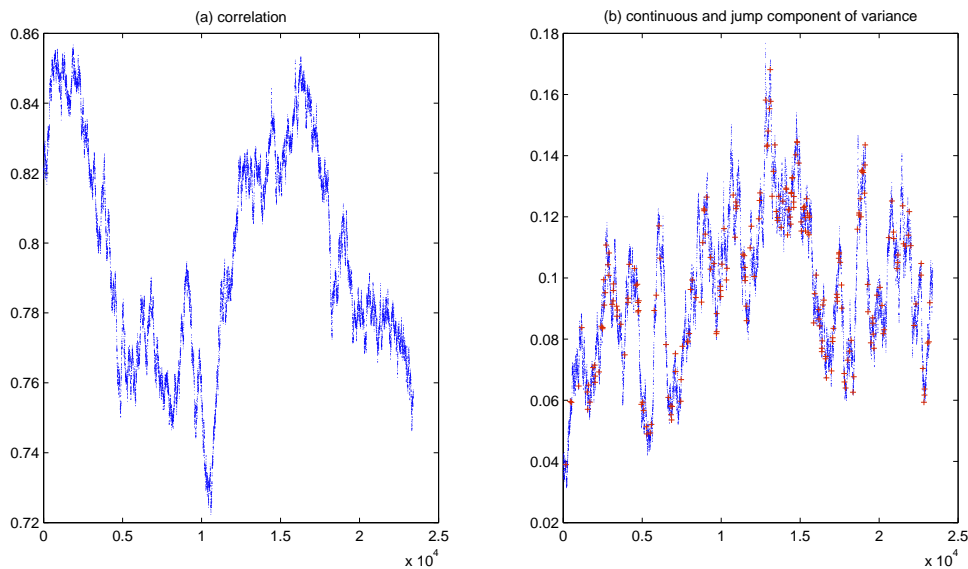
$$\begin{aligned} \sum_{\alpha} K_H(\lambda_{\alpha}) I(\lambda_{\alpha}) &:= \sum_{i,j=1}^n \Delta P_1(t_i) \Delta P_2(s_j) \sum_{\alpha} K_H(\lambda_{\alpha}) e^{-i(t_i-s_j)\alpha} \\ &\stackrel{(i)}{=} \sum_{i,j=1}^n \Delta P_1(t_i) \Delta P_2(s_j) k_H(t_i - s_j) \\ &\stackrel{(ii)}{=} \sum_{|h|<n} k\left(\frac{h}{H}\right) \gamma_h, \end{aligned}$$

where (ii) holds only when the discretization points are synchronous and equally spaced. The form of the estimator we propose to implement is (i). In the theory sections above, we assumed no leverage effect; in the simulation studies however, we relax this assumption and see if our estimator is robust to the presence of the leverage. We consider two data generating processes (DGP) for asset returns. For the first simulation, we consider the stochastic volatility model with a perfect leverage given in Barndorff-Nielsen et al. (2011). The volatility process is continuous whereas the instantaneous co-volatility is constant. For assets indexed by $j = 1, 2$,

$$\begin{aligned} dP_j(t) &= 0.03dt - 0.3\sigma_j(t)dB_j(t) + \sqrt{1 - (0.3)^2}\sigma_j(t)dW(t) \\ \sigma_j(t) &= \exp\{-5/16 + 1/8\varrho_j(t)\} \quad ; \quad d\varrho_j(t) = -1/40\varrho_j(t)dt + dB_j(t), \end{aligned} \quad (23)$$

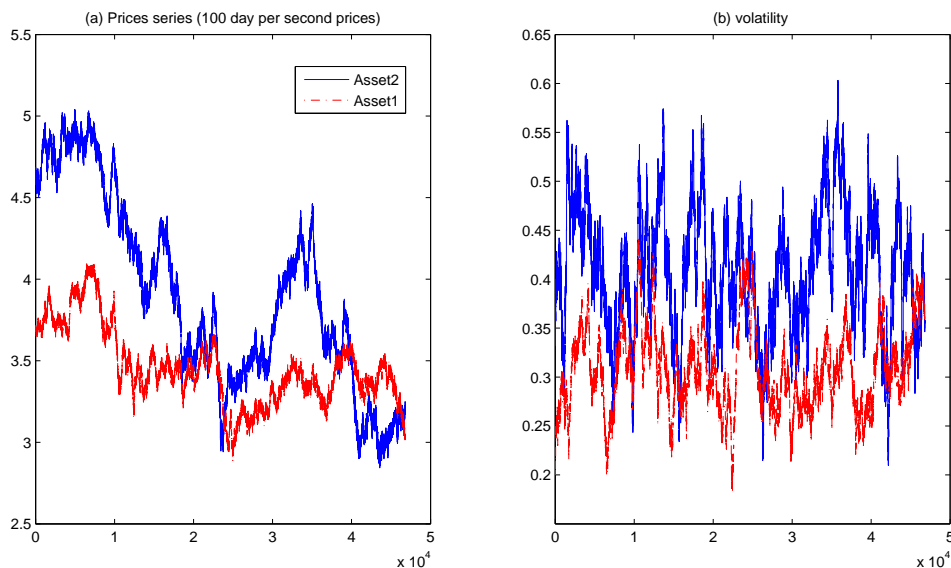
and $\varrho_j(t)$ is initialized by $\varrho_j(0) \sim N(0, 20)$. The model implies that the covariance between the returns are $EdP_1(t)dP_2(t) = 0.91\sigma_1(t)\sigma_2(t)dt$. There is a perfect statistical leverage since a single Brownian motion $B_j(t)$ which is present in the return equation, drives the volatility process.

Figure 2: Simulated intraday instantaneous co-volatility and variance



For the second simulation, the stochastic volatility is specified as a jump diffusion process and the instantaneous co-volatility coefficient follows the Cox-Ingersoll-Ross (CIR) process.

Figure 3: Simulated price and variance - per second observation



This is a modification of the DGP considered in Ait-Sahalia et al. (2010) and Barndorff-Nielsen et al. (2004). For $j = 1, 2$, we have

$$\begin{aligned} dP_j(t) &= \sigma_j(t)dW_j(t) \\ d\sigma_j^2(t) &= \kappa_j\{\bar{\sigma}_j^2 - \sigma_j^2(t)\} + a_j\sigma_j(t)dB_j(t) + \sigma_j(t_-)J_j(t)dN_j(t), \end{aligned} \quad (24)$$

where the jump size follows $J_j(t) = \exp\{z_j(t)\}$ with $z_j(t) \sim N(\mu_j, s_j)$, and $N_j(t)$ is a poisson process with intensity λ_j . The leverage effect here is $EdW_j(t)dB_j(t) := \delta_j dt$, and the covariance between the Brownian motions that are present in the price equation is given by $EdB_1(t)dB_2(t) = \rho_t dt$. The parameter values taken from Ait-Sahalia et al. (2010). We let $\rho(t) = (e^{2x(t)} - 1)/(e^{2x(t)} + 1)$, where $x(t)$ follows the CIR process

$$dx(t) = 0.03(0.64 - x(t))dt + 0.118x(t)dB_{xt}.$$

Figure 2(a) shows the time series plot of ρ_t , and (b) shows $\sigma_1^2(t)$ decomposed into a continuous and a discontinuous components. Figure 3(a) shows the time series plot of $P_j(t)$, $j = 1, 2$ and (b) shows $\sigma_j^2(t)$, $j = 1, 2$. The DGP of the microstructure noise is formed with respect to the transaction time. We consider a correlated AR(1) noise processes with smoothly decaying cross autocovariances. This can be implemented by

$$U_j(t_{i,j}) = \bar{U}_j(t_{i,j}) + \varepsilon(t_{i,j}); \quad \bar{U}_j(t_{i,j}) = \alpha\bar{U}_j(t_{i-1,j}) + \epsilon_j(t_{i,j}), \quad (25)$$

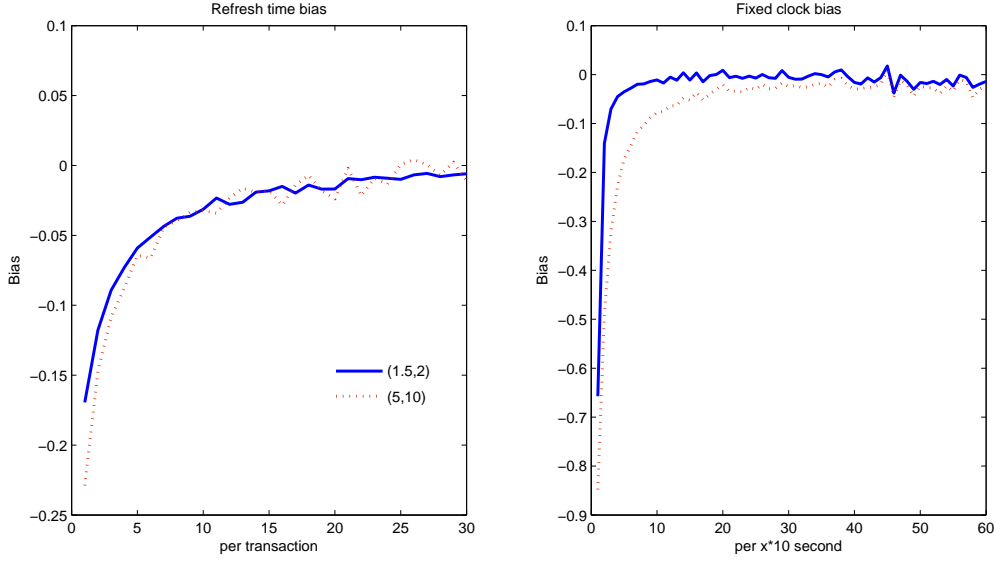
where the idiosyncratic errors are independent Gaussian, i.e. $\epsilon_j(t_{i,j}) \sim NID(0, 1)$. The common disturbance that drives the correlation between two microstructure noise is simulated by

$$\varepsilon_l = 0.5\varepsilon_{l-1} + \xi_l, \text{ for } \{T_l\}_{1 \leq l \leq N_T} = \{t_{i,1} \cup t_{j,2}, i = 1, \dots, n_1, j = 1, \dots, n_2\}, \xi \sim NID(0, 1).$$

Then we define $\{\varepsilon(t_{i,1})\}_{1 \leq i \leq n_1}$ as $\{\varepsilon_l\}_{1 \leq l \leq N_T}$ sampled at $\{T_l \cap t_{i,1}\}$ points. $\{\varepsilon(t_{j,2})\}_{1 \leq j \leq n_2}$ is defined similarly. The variance of the noise is set to be proportionate to the sample integrated quarticity; $\zeta^2 \sqrt{n_j^{-1} \sum_{i=1}^{n_j} \sigma_j^4(t_{i,j})}$, where $\zeta = \{0, 0.001, 0.01\}$ is the noise to signal ratio.

We simulate the one second data assuming 6.5 hours of daily trading time, yielding 23,400 daily data points over 100 Monte Carlo sample. The simulation is designed to assess the

Figure 4: Covariation signature plot for simulated series - bias induced by data alignment for different sampling rates



impact of the asynchronicity on the estimator: We Poisson sample the data at the rate of $\{(3/2, 30), (20, 30)\}$, where sampling at rate (a, b) means on average observation is made at every a and b seconds. To create a balanced sample for the rate $(3/2, 30)$, for the first asset, we sample on average at 1.5 second for the first half of the sample and at 30 second for the last half of the sample. For the second asset, we do this in reverse order. Then we have two assets that have the same number of transactions each day but traded very asynchronously.

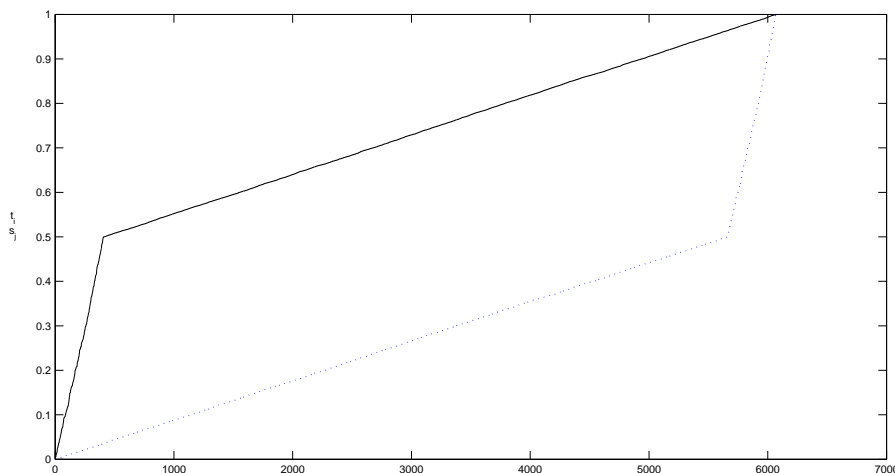
Finally, we examine the properties of the estimators in higher dimensions. We consider a simple setting where the log prices are given by $P(t) = AB(t)$; $P(t)$ is the 10×1 vector of prices, $B(t)$ is the 3×1 independent Brownian motion and A is a factor loading matrix. This is poisson sampled at rate $\{2, 2, 4, 4, 8, 8, 10, 10, 30, 30\}$ and masked by i.i.d gaussian noise. Table 2 and 3 at the end of the paper report the results for estimating the 2-dimensional covariation matrix, where the first asset is more often traded than the second asset. Table 4 reports the results for a higher dimension.

5.1.1 Realized Covariation: bias induced by data synchronization

Table 1 reports the finite sample properties of the realized covariance (RC). The efficient price follows Brownian semimartingale, given in (23). The transaction time follows a homogenous poisson process, and the microstructure noise are correlated AR(1) processes given in (25). Asynchronous data is aligned using the 5 minute fixed clock time and the refresh time. The negative bias when there is no noise is consistent with the result of Proposition 2.

When microstructure noise is present, the variance estimate has a large positive bias. The sparse sampling (5 minute aligned data) is able to reduce such bias. However, the covariance estimate has a negative bias induced by the Epps effect which dominates the positive bias induced by the microstructure noise. The degree at which Epps effect dominates the noise effect depends on the degree of non-synchronicity. Figure 4 shows the covariation signature plot for the simulated series when the price is observed without the noise. It shows that given varying degrees of non-synchronicity (rate at which assets is traded), the higher frequency we align the data (moving leftwards in x axis) the more bias it induces in estimating the integrated covariance.

Figure 5: Time stamp of two assets traded at opposite liquidity



5.1.2 Balanced Sample example

We elaborate further on the aforementioned sampling procedure: We first simulate the equally spaced data per one second for two assets. For the first asset, we sample on average at 1.5 second for the first half of the sample and at 30 second for the last half of the sample. For the second asset, this is done in reverse order - sample at 30 second for the first half and 1.5 second for later part. See Figure 5. Then we have two assets that have the same number of transactions each day but traded very asynchronously. This is like a case where two assets have opposite liquidity profile over a day.

The sample size is 607,774 over one hundred days and the refresh time aligned data reduces to a size around 750 per day. The large reduction in the sample size of the aligned data is due to severe non-synchronicity by simulation design. We compare the realized kernel and the proposed method over the range of bandwidths, $H = \{1, 5, 10, 20, 50, 100, \dots, 750, 800\}$. The Figure 6 shows that the proposed estimator is less sensitive to the choice of bandwidth - especially for large H . With large H , we can reduce the bias for the off-diagonal element more than we can do for the realized kernel. Our estimator is less sensitive to the choice of bandwidth for large values of H .

5.1.3 Unbalanced Sample example

We carried out the same exercise as above but with the unbalanced sample sizes. We Poisson sample the data at rate $\{(3/2, 30), (3/2, 2), (20, 30)\}$. For example, sampling rate $(3/2, 2)$ means that we sample the first asset on average per 1.5 second and the second asset per 2 second. The first rate is to examine the effect of different liquidity and different sample sizes. The second and third rates are to examine the effect of sparse and intense sampling of asset prices of similar liquidity.

Figure 6 shows the results for sampling rate $(3/2, 30)$. The proposed estimator has a less bias and is less sensitive to the choice of the bandwidth for large values of H . When the noise is not present, the proposed method estimates the variance of more liquid asset more precisely. When the noise is present, the bandwidth should be large for the proposed estimator to perform better. The conclusion is similar for sampling rate $(20, 30)$ as shown in Figure 8. The improvement of using the Fourier domain estimator is most evident when estimating the variance of more liquid asset when two sample sizes are very different. The

proposed covariation estimator performs better under large bandwidth. For sampling rate $(3/2, 2)$ in Figure 9, the difference of two estimator is less pronounced.

Each of these figures also show the accuracy of estimating the scalar function of the covariation matrix. We examined the maximum eigenvalues and the variance of portfolio with weight $[0.5, \sqrt{0.75}]$. Under the realistic noise to signal ratio and when two assets are of different liquidity, the proposed method delivers superior estimates. Regardless of sampling scheme, the proposed method does better in estimating these quantities when the effect of microstructure noise is not too dominant.

5.1.4 Overall Comparison and Higher Dimension Case

Table 2 and 3 shows that the proposed estimator has the best bias profile. With carefully chosen bandwidth we can achieve the best root MSE under the presence of noise. When no noise is present, the Hayashi and Yoshida estimator performs well. The refresh time aligned method often performs better in estimating the integrated variance of the less traded asset; $(2,2)$ element. This is because it effectively aligns on the time stamp of less traded asset. As shown in the analysis of asymptotic bias, when there is no noise and the number of refresh time sample is smaller in size, then the realized kernel underperforms in terms of bias. The proposed estimator overall estimates the off-diagonal elements better. We observe also that the realized covariance estimator aligned on sparsely sampled data often performs well - this is because there are two opposing effects in terms of bias: negative bias from the Epps effect and positive bias from the microstructure noise. The advantage of our estimator is most clear when we estimate the covariance matrix of higher dimension as shown in Table 4. We estimate 10 dimensional integrated covariance matrix and compare the maximum of eigenvalues and variance of the equally weighted portfolio. Under no presence of noise the refresh time based method has a large bias. We calculate the optimal bandwidth as given in Theorem 1 for each element of covariance matrix and take the minimum, maximum of these and average them. We note that our estimator seems to yield a large variance, however it performs best under the optimal bandwidth.

5.2 Empirical Application

In this section we apply the Fourier Realized Kernel to high frequency data. We analyzed five stocks of different liquidity - Microsoft (MSFT), Dell (Dell), JP Morgan (JPM) and less frequently traded Caterpillar Inc (CAT) and Banco de Chile (BCH) from WRDS TAQ database. The period of analysis is for 20 days during 05-30 March 2007. The order of least to most liquid stocks is BCH, CAT, DELL, JPM, and MSFT, whose average daily sample sizes are given by $\{48, 7526, 8337, 10337, 11451\}$, respectively. We may calculate the optimal bandwidth for individual asset by equalizing the squared bias and the variance term given in Section 4. Let $n_\ell = n^{\beta_\ell}$ where n is a minimum of all sample sizes, then it is given by $H_\ell = \{\eta \|k''(0)\|/\|k\|\}^{2/5} \zeta_\ell^{4/5} n^{\frac{1+2\beta_\ell}{5}}$, where $\eta = (s^*/2\pi)^2$. Note that ζ_ℓ^2 is the squared noise to signal ratio for each asset given by $\Gamma_{\ell\ell}/\sqrt{IQ_{\ell\ell}}$, where $\Gamma_{\ell\ell}$ is the limit defined in Assumption 6 and IQ refers to the integrated quarticity. We may estimate the variance of the microstructure noise by the Realized Variance applied on the tick data divided by $2n$, i.e. $E(U^2) \simeq RV/2n$. See Zhang et al. (2005). The square root of the integrated quarticity is estimated by the Realized Variance applied on the sparsely sampled data e.g. 10 minutes. We applied the maximum, minimum and average of the above individual bandwidths. Figure 10 and Figure 11 compare the realized covariance and the proposed method in estimating the daily covariation matrix. Since the first asset is least traded, the refresh time is effectively aligned on the trading time of the first asset. In estimating the integrated variance, the proposed method lies between the RV using pairwise refresh time (which will be dominated

by the microstructure noise) and the RV using all refresh time (which is more sparsely sampled, therefore less affected by the noise). The most interesting case is the performance in estimating the covariation for assets of different liquidity - i.e. (1, 4) and (1, 5)-th element of the estimator in our case. The daily Realized Covariance take values closer to zero due to the Epps effect whereas the proposed estimator clearly gives us non-trivial estimates.

5.3 Concluding Remarks

In this section we have shown that the Fourier domain estimator performs well even under extreme asynchronicity and the presence of microstructure noise. It may be possible to make further improvements, given the literature on discrete time estimation of the spectral density. For example, Xiao and Linton (2002) and Hirukawa (2006) developed multiplicative bias reduction methods that can improve the performance under stronger conditions without sacrificing positive semi-definiteness.

6 Proofs

We shall derive the results conditionally on the volatility matrix and the discretization time points, both of which are hence treated as deterministic throughout. Note that $C > 0$ is taken to mean some generic constant that may take different values in different places. In most cases, proofs are done under the most general framework; that is, when time stamps are asynchronous and sample sizes are unbalanced. The transaction time of the first and second assets is denoted either by $\{t_{i,1}, t_{j,2}\}$ or $\{t_i, s_j\}$.

Lemma 1. *Suppose $P(t)$ defined on some filtered probability space $(\Sigma, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ satisfies Assumption 1 with $\mu(t) = 0$. Let $f(t, s; q)$ be a bounded, measurable and square-integrable function. Then, it follows that for any $a, b, c, d = 1, \dots, p$,*

$$\begin{aligned} E \left[\int_0^{2\pi} \int_0^{2\pi} f(t, s; q) dP_a(s) dP_b(t), \int_0^{2\pi} \int_0^{2\pi} f(t, s; q') dP_c(s) dP_d(t) \right] & \quad (26) \\ &= \int_0^{2\pi} \int_0^{2\pi} f(t, s; q) f(t, s; q') d[P_a, P_c](s) d[P_b, P_d](t) \\ & \quad + \int_0^{2\pi} \int_0^{2\pi} f(t, s; q) f(s, t; q') d[P_a, P_d](s) d[P_b, P_c](t), \end{aligned}$$

where the double stochastic integral is Wiener-Itô sense.

PROOF OF LEMMA 1. From the definition of the double Wiener-Itô integral and standard properties of symmetrization operation, it is straightforward to see that

$$\begin{aligned} & \int_0^{2\pi} \int_0^{2\pi} f(t, s; q) dP_a(s) dP_b(t) \\ &= \int_0^{2\pi} \int_0^t f(t, s; q) dP_a(s) dP_b(t) + \int_0^{2\pi} \int_0^t f(s, t; q) dP_b(s) dP_a(t). \end{aligned}$$

The integrand here is measurable with respect to \mathcal{F}_t , and two terms above are martingale. Consequently, (26) can be expressed as

$$\left[\begin{aligned} & \int_0^{2\pi} \int_0^t f(t, s; q) dP_a(s) dP_b(t) + \int_0^{2\pi} \int_0^t f(s, t; q) dP_b(s) dP_a(t), \\ & \int_0^{2\pi} \int_0^t f(t, s; q') dP_c(s) dP_d(t) + \int_0^{2\pi} \int_0^t f(s, t; q') dP_d(s) dP_c(t) \end{aligned} \right].$$

Then one of the cross terms is given by

$$E \left[\int_0^{2\pi} \int_0^t f(t, s; q) dP_a(s) dP_b(t), \int_0^{2\pi} \int_0^t f(s, t; q') dP_d(s) dP_c(t) \right] \quad (27)$$

$$\begin{aligned} &= E \int_0^{2\pi} \left(\int_0^t f(t, s; q) dP_a(s) \right) \left(\int_0^t f(s, t; q') dP_d(s) \right) d[P_b, P_c](t). \\ &= \int_0^{2\pi} E \left(\int_0^t f(t, s; q) dP_a(s) \right) \left(\int_0^t f(s, t; q') dP_d(s) \right) d[P_b, P_c](t). \\ &= \int_0^{2\pi} \int_0^t f(t, s; q) f(s, t; q') d[P_a, P_d](s) d[P_b, P_c](t). \end{aligned} \quad (28)$$

Interchange of the expectation operator and the integral is justified by Fubini's Theorem. Note that $d[P_b, P_c](t) = [P_b, P_c]'(t)dt$, where the prime denotes the time derivative. The same arguments apply to the remaining terms, completing the proof. \blacksquare

Lemma 2. Consider the following kernel-weighted off-diagonal step functions:

$$\begin{aligned} f_n(t, s; q) &= \sum_{i \neq j} e^{-iqt_j} k_H(t_i - t_j) 1_{[t_{i-1}, t_i]}(t) 1_{[t_{j-1}, t_j]}(s) \\ g_n(t, s; q) &= \sum_{i=1}^n \sum_{j=1}^n e^{-iqt_j} k_H(t_i - t_j) 1_{[t_{i-1}, t_i]}(t) 1_{[t_{j-1}, t_j]}(s), \end{aligned} \quad (29)$$

where $\{t_i; i = 1, \dots, n\}$ are the discretization points that satisfy Assumption 2, and the kernel $k_H(t_i - t_j) =: k((t_i - t_j)/(H\Delta t))$ where $k(\cdot)$ is a lag window satisfying Assumption 3. Then

$$\begin{aligned} &\frac{n}{H} \int_0^{2\pi} \int_0^{2\pi} \left\{ f_n^2(t, s; q) + f_n(t, s; q) f_n(s, t; q) \right\} d[P, P](t) d[P, P](s) \\ &\quad \rightarrow 2\|k\|^2 \int_0^{2\pi} e^{-i2tq} ([P, P]'(t))^2 d\mathcal{Q}(t), \end{aligned} \quad (30)$$

where $\mathcal{Q}(t)$ is the limit of the empirical quadratic covariation $\mathcal{Q}^{(n)}(t) = n \sum_{i:t_i < t} (\Delta t_i)^2$.

PROOF OF LEMMA 2. We first show that f_n can be replaced by g_n in the integrand with an error of order $o(1)$. Given a single asset observed at discretizing points $\{t_i\}$ we have

$$\begin{aligned} &\int_0^{2\pi} \int_0^{2\pi} [g_n(t, s; q) - f_n(t, s; q)] ds dt \\ &= \int_0^{2\pi} \int_0^{2\pi} \left\{ \sum_{i=1}^n \sum_{j=1}^n - \sum_{i \neq j} \right\} \cdot e^{-it_j q} k_H(t_i - t_j) \cdot 1_{[t_{i-1}, t_i]}(t) 1_{[t_{j-1}, t_j]}(s) ds dt \\ &= \int_0^{2\pi} \int_0^{2\pi} \left\{ \sum_{j=1}^n e^{-it_j q} 1_{[t_{j-1}, t_j]}(t) 1_{[t_{j-1}, t_j]}(s) \right\} ds dt \\ &\leq C \left(\sup_{1 \leq j \leq n} \Delta t_j \right) = O\left(\frac{1}{n}\right) \end{aligned} \quad (31)$$

by Assumption 2.

Generalising this idea we can rewrite (30) as

$$\begin{aligned}
& \frac{n}{H} \int_0^{2\pi} \int_0^{2\pi} \left\{ g_n^2(t, s; q) + g_n(t, s; q)g_n(s, t; q) \right\} d[P, P](t)d[P, P](s) + o(1) \\
&= \frac{n}{H} \sum_{h=0}^{n-1} \sum_{j=1}^{n-h} \left[(e^{-it_j 2q} + e^{-it_j q} e^{-it_{j+h} q}) k_H^2(t_{j+h} - t_j) [P, P]'(t_{j+h}) [P, P]'(t_j) \Delta t_{j+h} \Delta t_j \right. \\
&\quad \left. + \sum_{h=1}^{n-1} \sum_{j=1+h}^n (e^{-it_j 2q} + e^{-it_j q} e^{-it_{j-h} q}) k_H^2(t_{j-h} - t_j) [P, P]'(t_{j-h}) [P, P]'(t_j) \Delta t_{j-h} \Delta t_j \right] + o(1) \\
&= \frac{n}{H} \left[\sum_{h=0}^{n-1} k^2 \left(\frac{t_h / \overline{\Delta t}}{H} \right) \sum_{j=1}^{n-h} (e^{-it_j 2q} + e^{-it_j q} e^{-it_{j+h} q}) [P, P]'(t_{j+h}) [P, P]'(t_j) \Delta t_{j+h} \Delta t_j \right. \\
&\quad \left. + \sum_{h=1}^{n-1} k^2 \left(\frac{-t_h / \overline{\Delta t}}{H} \right) \sum_{j=1+h}^n (e^{-it_j 2q} + e^{-it_j q} e^{-it_{j-h} q}) [P, P]'(t_{j-h}) [P, P]'(t_j) \Delta t_{j-h} \Delta t_j \right] + o(1) \tag{32}
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{H} \left[\sum_{h=0}^{n-1} k^2 \left(\frac{t_h / \overline{\Delta t}}{H} \right) + \sum_{h=1}^{n-1} k^2 \left(\frac{-t_h / \overline{\Delta t}}{H} \right) + 1 - 1 \right] \\
&\quad \times n \sum_{j=1}^n (e^{-it_j 2q} + e^{-it_j q} e^{-it_{j+h} q}) [P, P]'(t_{j+h}) [P, P]'(t_j) \Delta t_{j+h} \Delta t_j + o(1) + O \left(\frac{1}{n} \right) \tag{33}
\end{aligned}$$

$$\longrightarrow 2 \|k\|^2 \int_0^{2\pi} e^{-it 2q} ([P, P]'(t))^2 d\mathcal{Q}(t), \tag{34}$$

It is straightforward to see that the error in approximating the kernel $k_H(t_i - t_{i-h})$ by $k_H(t_h)$ in (32) is negligible. Taylor expanding the lag window we have

$$k_H(t_i - t_{i-h}) - k_H(t_h) \simeq \frac{t_i - t_{i-h} - t_h}{\Delta t H} k' \left(\frac{t_h}{\Delta t H} \right) + \frac{1}{2} \left(\frac{t_i - t_{i-h} - t_h}{\Delta t H} \right)^2 k'' \left(\frac{t_h}{\Delta t H} \right),$$

where

$$\frac{t_i - t_{i-h} - t_h}{\Delta t H} = \sum_{l=1}^h (\Delta t_{l+i-h} - \Delta t_l) \frac{1}{\Delta t H} \leq 2h \left(\frac{\sup_l \Delta t_l}{\Delta t H} \right) = Ch/H.$$

Then the approximating error in (32) is bounded by

$$C \sup_{j,h} (\Delta t_{j-h} \Delta t_j) \frac{n}{H} \sum_{|h| < n} \sum_{j=1+h}^n \left\{ \frac{h}{H} k' \left(\frac{h}{H} \right) + \left(\frac{h}{H} \right)^2 k'' \left(\frac{h}{H} \right) \right\} \simeq C \frac{1}{H} \int_{-\infty}^{\infty} x^2 k''(x) dx,$$

by Assumption 2 and 3. Further, the approximation error from (32) to (33) is given by

$$\begin{aligned}
& \frac{n}{H} \left[\sum_{h=0}^{n-1} k^2 \left(\frac{t_h / \overline{\Delta t}}{H} \right) + \sum_{h=1}^{n-1} k^2 \left(\frac{-t_h / \overline{\Delta t}}{H} \right) \right] \\
&\quad \times 2 \left(\sum_{j=1}^h (e^{-it_j 2q} + e^{-it_j q} e^{-it_{j+h} q}) [P, P]'(t_{j+h}) [P, P]'(t_j) \Delta t_{j+h} \Delta t_j \right) \\
&\leq C (\sup \Delta t_j)^2 n \sum_{|h| < n} \frac{h}{H} k^2 \left(\frac{h}{H} \right) \simeq \frac{1}{n} C \int_{-\infty}^{\infty} x k^2(x) dx.
\end{aligned}$$

Lastly, (34) is justified by the fact that both $\frac{n}{H} \sum_{|h|<n} k^2\left(\frac{h}{H}\right) \sum_{j=1}^n \Delta t_{j+h} \Delta t_j$ and $\frac{n}{H} \sum_{|h|<n} k^2\left(\frac{h}{H}\right) \sum_{j=1}^n \Delta t_j \Delta t_j$ approach to the same limit due to the presence of the kernel weights, and by Riemann approximation of the integral $\int_{-\infty}^{\infty} k^2(x) dx$. \blacksquare

Lemma 3. *Consider the following off-diagonal step functions*

$$\begin{aligned} f_n(t, s; q) &= \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} e^{-is_j q} e^{-i(t_i - s_j) \alpha} 1_{[t_{i-1}, t_i[}(t) 1_{[s_{j-1}, s_j[}(s) 1_{\{I_{i,1} \cap I_{j,2} = \emptyset\}}(t, s) \\ g_n(t, s; q) &= \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} e^{-is_j q} e^{-i(t_i - s_j) \alpha} 1_{[t_{i-1}, t_i[}(t) 1_{[s_{j-1}, s_j[}(s), \end{aligned}$$

where the discretization points $\{t_i, s_j\}$ satisfy Assumption 2. Then we have

$$\int_0^{2\pi} \int_0^{2\pi} f_n(t, s; q) ds dt = \int_0^{2\pi} \int_0^{2\pi} g_n(t, s; q) ds dt + O\left(\frac{1}{n_1 \wedge n_2}\right). \quad (35)$$

PROOF OF LEMMA 3. The arguments closely follow those in (31). The only difference is that here we consider two different assets with discretizing points $\{t_i, s_j\}$.

The difference between the integrals in (35) is given by

$$\begin{aligned} & \int_0^{2\pi} \int_0^{2\pi} \left(\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \left[1 - 1_{\{I_{i,1} \cap I_{j,2} = \emptyset\}}(t, s) \right] e^{-iqs_j} e^{-i\alpha(t_i - s_j)} 1_{[t_{i-1}, t_i[}(t) 1_{[s_{j-1}, s_j[}(s) \right) ds dt \\ &= \int_0^{2\pi} \int_0^{2\pi} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} e^{-iqs_j} e^{-i\alpha(t_i - s_j)} 1_{[t_{i-1}, t_i[}(t) 1_{[s_{j-1}, s_j[}(s) 1_{\{I_{i,1} \cap I_{j,2} \neq \emptyset\}}(t, s) ds dt \\ &\leq \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \Delta t_i \Delta s_j 1_{\{I_{i,1} \cap I_{j,2} \neq \emptyset\}} \leq \sup_{1 \leq i \leq n_1} (\Delta t_i) \sup_{1 \leq j \leq n_2} (\Delta s_j) \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} 1_{\{I_{i,1} \cap I_{j,2} \neq \emptyset\}} \\ &= \frac{(s^*)^2}{n_1 n_2} \#\{t_i \cup s_j, 0 \leq t_i, s_j \leq 2\pi\} = O\left(\frac{1}{n_1 \wedge n_2}\right), \end{aligned}$$

because the total number of union of time two stamps is of order $O(n_1 \vee n_2)$. \blacksquare

Lemma 4. *Suppose that $P(t)$ is defined on some probability space $(\Sigma, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ and satisfies Assumption 1, and denote by $\mathcal{G} = \sigma(P)$ a sub- σ -field of \mathcal{F} . Let Z be a standard Gaussian variable on a suitable extension of the probability space, and \mathcal{V} be the \mathcal{G} -measurable stochastic variance. Then for $f_n(\cdot)$ given in Lemma 2,*

$$\sqrt{\frac{n}{H}} \int_0^{2\pi} \int_0^{2\pi} f_n(t, s; q) dP_1(s) dP_2(t) \Longrightarrow^{\text{stably}} \sqrt{\mathcal{V}} Z,$$

provided that the conditions of Theorem 3.2 of Jacod (1997) are met.

PROOF OF LEMMA 4. Given the discretized filtration \mathcal{F}_i , $i = \max_j \{t_j \leq t\}$ and the discretized sequence adapted to \mathcal{F}_i :

$$\chi_i^n = \sqrt{n/H} \Delta P_1(t_i) \sum_{j: s_j < t_i} \Delta P_2(s_j) k_H(t_i - s_j) e^{-is_j q},$$

we show stable convergence of $Z_t^n := \sum_{\max_j \{t_j \leq t\}} \chi_i^n$ to $Z_t = \int_0^t v_s dW_s$, a \mathcal{F}_t -conditional Gaussian martingale. Then by Theorem 3.2 of Jacod (1997) we have $Z^n \Longrightarrow Z$ stably. For future reference we note that a sufficient condition for the conditional Lindeberg condition in (4) is the Lyapunov condition $\sum_i E(\{\chi_i^n\}^{2+\varepsilon} | \mathcal{F}_{i-1}) \rightarrow_p 0$, for some $\varepsilon > 0$. We will show this for $\varepsilon = 2$ in the proof of Theorem 1. \blacksquare

6.1 Proof of Proposition 1

PROOF. For the sake of simplicity of notation take $a = 1$ and $b = 2$, w.l.o.g.

Using (14) we may write

$$\begin{aligned}
& \sum_{|\alpha| \leq m/2} K_H(\lambda_\alpha) [\mathcal{F}_n(dU_1)(\alpha)] [\mathcal{F}_n(dU_2)(q - \alpha)] \\
&= \sum_{|\alpha| \leq m/2} K_H(\lambda_\alpha) \left[\sum_{i=1}^{n_1} e^{-i\alpha t_i} \Delta U_1(t_i) \right] \left[\sum_{j=1}^{n_2} e^{-i(q-\alpha)s_j} \Delta U_2(s_j) \right] \\
&= \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \Delta U_1(t_i) \Delta U_2(s_j) e^{-iqs_j} \left[\sum_{|\alpha| \leq m/2} K_H(\lambda_\alpha) e^{-i\alpha(t_i - s_j)} \right] \\
&= \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \Delta U_1(t_i) \Delta U_2(s_j) e^{-iqs_j} k_H(t_i - s_j). \tag{36}
\end{aligned}$$

We first consider the terms that are not affected by the end points (i.e. t_{n_1} , t_0 , s_{n_2} and s_0). Expanding the difference operators we see that those terms are given by

$$\begin{aligned}
& \sum_{i=1}^{n_1-1} \sum_{j=1}^{n_2-1} U_1(t_i) U_2(s_j) \left\{ e^{-iqs_j} \left[-k_H(t_{i+1} - s_j) + k_H(t_i - s_j) \right] \right. \\
& \quad \left. + e^{-iqs_{j+1}} \left[k_H(t_{i+1} - s_{j+1}) - k_H(t_i - s_{j+1}) \right] \right\}. \tag{37}
\end{aligned}$$

By repeated use of Taylor expansion the terms in the curly bracket of (37) become

$$\begin{aligned}
& \left\{ e^{-is_j q} \left[-k \left(\frac{t_i - s_j + \Delta t_{i+1}}{\bar{H}} \right) + k \left(\frac{t_i - s_j}{\bar{H}} \right) \right] \right. \\
& \quad \left. + e^{-is_{j+1} q} \left[k \left(\frac{t_i - s_j + \Delta t_{i+1} - \Delta s_{j+1}}{\bar{H}} \right) - k \left(\frac{t_i - s_j - \Delta s_{j+1}}{\bar{H}} \right) \right] \right\} \\
&= \left\{ -e^{-is_j q} \left[\frac{\Delta t_{i+1}}{\bar{H}} k' \left(\frac{t_i - s_j}{\bar{H}} \right) \right] + e^{-is_{j+1} q} \left[\frac{\Delta t_{i+1}}{\bar{H}} k' \left(\frac{t_i - s_j - \Delta s_{j+1}}{\bar{H}} \right) \right] + o(1) \right\} \\
&= \left\{ -e^{-is_j q} \left[\frac{\Delta t_{i+1}}{\bar{H}} k' \left(\frac{t_i - s_j}{\bar{H}} \right) - \frac{\Delta t_{i+1}}{\bar{H}} k' \left(\frac{t_i - s_j - \Delta s_{j+1}}{\bar{H}} \right) \right] \right. \\
& \quad \left. + e^{-is_{j+1} q} \left[\frac{\Delta t_{i+1}}{\bar{H}} k' \left(\frac{t_i - s_j - \Delta s_{j+1}}{\bar{H}} \right) \right] - e^{-is_j q} \frac{\Delta t_{i+1}}{\bar{H}} k' \left(\frac{t_i - s_j - \Delta s_{j+1}}{\bar{H}} \right) + o(1) \right\} \\
&= \left\{ e^{-is_j q} \left[\frac{\Delta t_{i+1} \Delta s_{j+1}}{\bar{H}^2} k'' \left(\frac{t_i - s_j}{\bar{H}} \right) \right] + e^{-is_j q} \left[(e^{-i\Delta s_{j+1} q} - 1) \frac{\Delta t_{i+1}}{\bar{H}} k' \left(\frac{t_i - s_j}{\bar{H}} \right) \right] + o(1) \right\},
\end{aligned}$$

where $\bar{H} = H \widetilde{\Delta t}_{12}$. In this sequel $\widetilde{\Delta t}$ is taken to mean $\widetilde{\Delta t}_{12}$ for the sake of simplicity. Note that under Assumptions 2 and 4 $\Delta t_{i+1} \Delta s_{j+1} \frac{1}{\bar{H}^2} = o(1)$ and $\Delta t_{i+1} \frac{1}{\bar{H}} = O(H^{-1}) = o(1)$.

When samples are equally spaced and balanced, the curly bracket terms can be simplified further as follows:

$$\left\{ \cdot \right\} = -e^{-is_j q} \frac{1}{H^2} k'' \left(\frac{i-j}{H} \right) + e^{-is_j q} (e^{-i\Delta s_{j+1} q} - 1) \frac{1}{H} k'_H \left(\frac{i-j-1}{H} \right).$$

It now follows that (37) can be written as a sum of the following two terms (up to additional terms of negligible order $o(1)$):

$$A_1 + A_2 = \frac{1}{\bar{H}^2} \sum_{i=1}^{n_1-1} \sum_{j=1}^{n_2-1} e^{-is_j q} \left[U_1(t_i) U_2(s_j) k'' \left(\frac{t_i - s_j}{\bar{H}} \right) \right] \Delta t_{i+1} \Delta s_{j+1} \quad (38)$$

$$+ \frac{1}{\bar{H}} \sum_{i=1}^{n_1-1} \sum_{j=1}^{n_2-1} e^{-is_j q} \cdot (e^{-i\Delta s_{j+1} q} - 1) \left[U_1(t_i) U_2(s_j) k' \left(\frac{t_i - s_j}{\bar{H}} \right) \right] \Delta t_{i+1}. \quad (39)$$

We show that A_2 is of negligible order of magnitude, and that A_1 is the leading term that dominates in large sample. The expectation of A_2 is bounded above by:

$$\begin{aligned} E(A_2) &\leq \frac{1}{\widetilde{\Delta t H}} \sup_j |1 - e^{-iq\Delta s_{(j+1)}}| \sup_i (\Delta t_{i+1}) \left| \sum_{i=1}^{n_1-1} \sum_{j=1}^{n_2-1} E[U_1(t_i) U_2(s_j)] \cdot k' \left(\frac{t_i - s_j}{\bar{H}} \right) \right| \\ &= C \frac{n_1 \vee n_2}{H} \frac{1}{n_1 n_2} \left| \left[\sum_{|t_i - s_j|/\widetilde{\Delta t} \leq \sqrt{H}} + \sum_{|t_i - s_j|/\widetilde{\Delta t} > \sqrt{H}} \right] E[U_1(t_i) U_2(s_j)] \cdot k' \left(\frac{t_i - s_j}{\bar{H}} \right) \right| \\ &\leq \frac{C}{H(n_1 \wedge n_2)} \left\{ \sup_{|t_i - s_j|/\widetilde{\Delta t} \leq \sqrt{H}} \left| k' \left(\frac{t_i - s_j}{\bar{H}} \right) \right| \cdot \left| \sum_{|t_i - s_j|/\widetilde{\Delta t} \leq \sqrt{H}} \gamma \left(\frac{t_i - s_j}{\widetilde{\Delta t}} \right) \right| \right. \\ &\quad \left. + \sup_{|t_i - s_j|/\widetilde{\Delta t} > \sqrt{H}} \left| \gamma \left(\frac{t_i - s_j}{\widetilde{\Delta t}} \right) \right| \cdot \left| \sum_{|t_i - s_j|/\widetilde{\Delta t} > \sqrt{H}} k' \left(\frac{t_i - s_j}{\bar{H}} \right) \right| \right\} \\ &= \frac{C}{H(n_1 \wedge n_2)} \left\{ O(1) O(n_1 \wedge n_2) + o(1) \right\} = O\left(\frac{1}{H}\right) = o(1), \end{aligned}$$

where in the second and the last equality we used Taylor approximation and the fact that $k'(0) = 1$ along with Assumption 6, respectively.

Now consider a set $\mathcal{S} := \{i, j, r, l; (t_i - t_r)/\bar{\Delta t} < H^\mu, (s_j - s_l)/\bar{\Delta s} < H^\mu\}$ where μ is some positive constant between $(0, 1)$. Note that $\max_{i,r} (t_i - t_r)/\bar{\Delta t} = n$, and recall $H = n^\alpha$, $\alpha \in (0, 1)$. We then see that $E(A_2^2)$ is bounded above by

$$\begin{aligned} &\frac{CH^{-2}}{(n_1 \wedge n_2)^2} \left\{ \sum_{i,j,r,l} E[U_1(t_i) U_2(s_j) U_1(t_r) U_2(s_l)] k' \left(\frac{t_i - s_j}{\bar{H}} \right) k' \left(\frac{t_r - s_l}{\bar{H}} \right) \right\} \quad (40) \\ &= \frac{CH^{-2}}{(n_1 \wedge n_2)^2} \left\{ \sum_{i,j,r,l \in \mathcal{S}} + \sum_{i,j,r,l \in \mathcal{S}^c} \right\} E[U_1(t_i) U_2(s_j) U_1(t_r) U_2(s_l)] k' \left(\frac{t_i - s_j}{\bar{H}} \right) k' \left(\frac{t_r - s_l}{\bar{H}} \right) \\ &\leq \frac{CH^{-2}}{(n_1 \wedge n_2)^2} \sup_{i,j,r,l \in \mathcal{S}} |EU_1(t_i) U_2(s_j) U_1(t_r) U_2(s_l)| \cdot \left| \sum_{i,j} \sum_{|h|,|v| < H^\mu} k' \left(\frac{t_i - s_j}{\bar{H}} \right) k' \left(\frac{t_{i-h} - s_{j-v}}{\bar{H}} \right) \right| \\ &\quad + \frac{CH^{-2}}{(n_1 \wedge n_2)^2} n_1^2 n_2^2 \sup_{i,j,r,l \in \mathcal{S}^c} |EU_1(t_i) U_2(s_{i-h}) U_1(t_r) U_2(s_{r-v})| = B_1 + B_2. \end{aligned}$$

When the data is balanced and equally spaced, B_1 simplifies to

$$\begin{aligned} &\frac{C}{(n_1 \wedge n_2)^2 H^2} \left[\sup_{|i-r| < H^\mu, |h-v| < H^\mu} |EU_1(t_i) U_2(s_{i-h}) U_1(t_r) U_2(s_{r-v})| \right. \\ &\quad \left. \times \left| \sum_{|i-r| < H^\mu, |h-v| < H^\mu} k' \left(\frac{t_i - s_{i-h}}{\bar{H}} \right) k' \left(\frac{t_r - s_{r-v}}{\bar{H}} \right) \right| \right]. \quad (41) \end{aligned}$$

In particular, when the sample size of two assets is equal, it holds that $(t_i - s_{i-h})/\bar{H} = h/H + o(1)$ under Assumption 2. Hence the squared bracket term in (41) is given by

$$\begin{aligned} & \sum_{|i-r| < H^\mu} \sum_{|h-v| < H^\mu} k' \left(\frac{h}{H} \right) k' \left(\frac{v}{H} \right) = 2H^\mu n \sum_{|h-v| < H^\mu} k' \left(\frac{h}{H} \right) k' \left(\frac{v}{H} \right) \\ & = 2H^\mu n \left\{ \sum_{0 < l < H^\mu} \sum_{h=1+l}^n k' \left(\frac{h}{H} \right) k' \left(\frac{h-l}{H} \right) + \sum_{0 < l < H^\mu} \sum_{h=1}^{n-l} k' \left(\frac{h}{H} \right) k' \left(\frac{h+l}{H} \right) \right\} \\ & \leq 4H^{2\mu} n \sum_{h=1}^n \left\{ k' \left(\frac{h}{H} \right) \right\}^2. \end{aligned}$$

In the unbalanced case, the Riemann approximation gives

$$\sum_{i=1}^{n_1-1} \sum_{j=1}^{n_2-1} \left\{ k' \left(\frac{t_i - s_j}{\bar{H}} \right) \right\}^2 = (n_1 \wedge n_2) H \int_{-\infty}^{\infty} \{k'(x)\}^2 dx + o(1),$$

and the order of $\#\{0 \leq i, r \leq n_1; \frac{t_i - t_r}{\Delta t} < H^\mu\}$ is the same as when the data is equally spaced under Assumption 2; therefore, it follows that

$$B_1 = \frac{CH^{-2}}{(n_1 \wedge n_2)^2} \times \left[4\rho(0)(n_1 \wedge n_2)H^{2\mu+1} \int_{-\infty}^{\infty} (k'(x))^2 dx \right] = O\left(\frac{H^{2\mu-1}}{n_1 \wedge n_2}\right) = o(1).$$

Also, we have $B_2 = CH^{-2}(n_1 \wedge n_2)^{-2}n_1^2n_2^2 \cdot \sup_{|\tau| > H^\mu} \rho(\tau)$ which vanishes in large sample as a consequence of (17) in Assumption 6.

We next move on to A_1 . Its expectation is given by

$$\begin{aligned} & \bar{H}^{-2} \sum_{|t_i - s_j|/\widetilde{\Delta t} \leq \sqrt{H}} E[U_1(t_i)U_2(s_j)] e^{-is_j q} k'' \left(\frac{t_i - s_j}{\bar{H}} \right) \Delta t_{i+1} \Delta s_{j+1} \\ & + \bar{H}^{-2} \sum_{|t_i - s_j|/\widetilde{\Delta t} > \sqrt{H}} E[U_1(t_i)U_2(s_j)] e^{-is_j q} k'' \left(\frac{t_i - s_j}{\bar{H}} \right) \Delta t_{i+1} \Delta s_{j+1} = D_1 + D_2, \end{aligned}$$

where D_2 is bounded above by

$$\begin{aligned} & \bar{H}^{-2} \sup_i (\Delta t_{i+1}) \sup_j (\Delta s_{j+1}) \sup_{|t_i - s_j|/\widetilde{\Delta t} > \sqrt{H}} |EU_1(t_i)U_2(s_j)| \sum_{|t_i - s_j|/\widetilde{\Delta t} > \sqrt{H}} \left| k'' \left(\frac{t_i - s_j}{\bar{H}} \right) \right| \\ & \leq C \left(\frac{n_1 \vee n_2}{H^2} \right) \sup_{|t_i - s_j|/\widetilde{\Delta t} > \sqrt{H}} \left| \gamma \left(\frac{|t_i - s_j|}{\widetilde{\Delta t}} \right) \right| \cdot H \int_{-\infty}^{\infty} |k''(x)| dx = o(1), \end{aligned}$$

and moreover

$$\begin{aligned} D_1 & \leq \frac{(n_1 \vee n_2)^2 s^* s^*}{(2\pi)^2 H^2} \frac{1}{n_1 n_2} \sum_{|t_i - s_j|/\widetilde{\Delta t} \leq \sqrt{H}} E[U_1(t_i)U_2(s_j)] k'' \left(\frac{t_i - s_j}{\bar{H}} \right) e^{-is_j q} \\ & = \left(\frac{s^*}{2\pi} \right)^2 \frac{n_1 \vee n_2}{H^2 (n_1 \wedge n_2)} |k''(0)| \sum_{|t_i - s_j|/\widetilde{\Delta t} \leq \sqrt{H}} E[U_1(t_i)U_2(s_j)] e^{-is_j q} + o(1) \\ & \leq \left(\frac{s^*}{2\pi} \right)^2 \frac{n_1 \vee n_2}{H^2} |k''(0)| \Gamma_{12} + o(1) = O\left(\frac{n_1 \vee n_2}{H^2}\right) \end{aligned}$$

by Assumption 6.

The expectation of the square of (38) is bounded above by

$$\begin{aligned}
& C \left(\frac{n_1 \vee n_2}{H^2(n_1 \wedge n_2)} \right)^2 E \left\{ \sum_{i=1}^{n_1-1} \sum_{j=1}^{n_2-1} [U_1(t_i)U_2(s_j)] k'' \left(\frac{t_i - s_j}{\bar{H}} \right) \right\}^2 \\
&= C \left(\frac{n_1 \vee n_2}{H^2(n_1 \wedge n_2)} \right)^2 \rho(0) \cdot 4(n_1 \wedge n_2) H^{2\mu+1} \int_{-\infty}^{\infty} (k''(x))^2 dx + o(1) \\
&= O \left(\frac{(n_1 \vee n_2)^2}{n_1 \wedge n_2} H^{2\mu-3} \right). \tag{42}
\end{aligned}$$

All other terms involving the end points are of smaller order by similar arguments given in Barndorff-Nielsen et al. (2011) along with Assumption 7. The proof is now complete. \blacksquare

6.2 Proof of Proposition 2 and Theorem 1

PROOF. The proof consists of three parts; we show stable convergence of the diagonal and off-diagonal terms of the centered estimator in the first and second parts, respectively, followed by a brief justification for joint convergence via the Cramér-Wold device.

Without loss of generality consider the first element of the centred estimator (without noise); that is, the term involving the price process of the first asset:

$$\begin{aligned}
\mathcal{E}_1 &= \left[\widehat{\mathcal{F}}(\Sigma^P)(q) - \mathcal{F}(\Sigma^P)(q) \right]_{11} = \left(\text{vech} \left(\widehat{\mathcal{F}}(\Sigma^P)(q) - \mathcal{F}(\Sigma^P)(q) \right) \right)_1 \\
&= \sum_{|\alpha| \leq m/2} K_H(\lambda_\alpha) \left[\mathcal{F}_n(dP_1)(\alpha) \right] \left[\mathcal{F}_n(dP_1)(q - \alpha) \right] - \mathcal{F}(\Sigma_{11})(q).
\end{aligned}$$

For the sake of simplicity we drop the subscript denoting asset; for example $P(t_j)$ and n are hereafter taken to mean $P_1(t_{j,1})$ and n_1 , respectively. In view of the relationship between the spectral window K and the lag window k (13), we have the following ‘bias-variance’ decomposition:

$$\begin{aligned}
\mathcal{E}_1 &= \sum_{|\alpha| \leq m/2} K_H(\lambda_\alpha) \left[\sum_{j=1}^n e^{-i\alpha t_j} \Delta P(t_j) \right] \left[\sum_{k=1}^n e^{-i(q-\alpha)t_k} \Delta P(t_k) \right] - \int_0^{2\pi} e^{-iqt} d[P, P](t) \\
&= \sum_{j=1}^n \sum_{k=1}^n \left[\sum_{|\alpha| \leq \frac{m}{2}} K_H(\lambda_\alpha) e^{-i\alpha(t_j - t_k)} \right] \Delta P(t_j) \Delta P(t_k) e^{-iqt_k} - \int_0^{2\pi} e^{-iqt} d[P, P](t) \\
&= \sum_{j=1}^n \sum_{k=1}^n k_H(t_j - t_k) \Delta P(t_j) \Delta P(t_k) e^{-iqt_k} - \int_0^{2\pi} e^{-iqt} d[P, P](t) \\
&= \left[\sum_{j=1}^n (\Delta P(t_j))^2 e^{-iqt_j} - \int_0^{2\pi} e^{-iqt} d[P, P](t) \right] + \left[\sum_{j \neq k} k_H(t_j - t_k) \Delta P(t_j) \Delta P(t_k) e^{-iqt_k} \right] \\
&=: M_1 + M_2.
\end{aligned}$$

We first show that $\sqrt{n/\bar{H}}M_1$ converges in probability to zero.

Since Itô’s lemma gives

$$[\Delta P(t_j)]^2 e^{-iqt_j} = 2 \int_{t_{j-1}}^{t_j} \{P(t) - P(t_{j-1})\} \cdot e^{-iqt_j} dP(t) + \int_{t_{j-1}}^{t_j} e^{-iqt_j} d[P, P](t),$$

we can further decompose M_1 into a martingale M_{11} and a predictable finite variation component A , where

$$M_{11} = 2 \sum_{j=1}^n \int_{t_{j-1}}^{t_j} \{P(t) - P(t_{j-1})\} e^{-iqt_j} dP(t) = O_p(n^{-1/2}),$$

$$A = \sum_{j=1}^n \int_{t_{j-1}}^{t_j} (e^{-itq} - e^{-it_jq}) d[P, P](t) = O(n^{-1}).$$

This is the Euler discretization error whose distribution is given by the Theorem 5.5 of Jacod and Protter (1998). Therefore, it follows that $\sqrt{n/H}M_1 = O_p(1/\sqrt{H}) = o_p(1)$.

Now we show stable convergence of M_2 . Using the off-diagonal step function $f_n(t, s; q)$ defined in Lemma 3:

$$f_n(t, s; q) = \sum_{j \neq k} e^{-iqt_k} k_H(t_j - t_k) 1_{[t_{j-1}, t_j]}(t) 1_{[t_{k-1}, t_k]}(s),$$

we write

$$M_2 = \sum_{j \neq k} k_H(t_j - t_k) \Delta P(t_j) \Delta P(t_k) e^{-iqt_k} = \int_0^{2\pi} \int_0^{2\pi} f_n(t, s; q) dP(s) dP(t).$$

Clearly, the expectation of M_2 is zero. Furthermore, using Lemma 1 we see that

$$E[M_2, M_2] = 2 \int_0^{2\pi} \int_0^{2\pi} f_n^2(t, s; q) d[P, P](s) d[P, P](t),$$

whose limiting behaviour is given by Lemma 2. Consequently the distribution of \mathcal{E}_1 follows by Lemma 4, provided the conditional Lindeberg condition holds. It suffices to verify the Lyapunov condition: For some $\varepsilon > 0$, $\sum_j E(\{\chi_j^n\}^{2+\varepsilon} | \mathcal{F}_{j-1}) \rightarrow^p 0$, as we shall do here for $\varepsilon = 2$.

For later reference we note that the fourth moment of the return is given by

$$\begin{aligned} E([\Delta P(t_j)]^4) &= E\left(2 \int_{t_{j-1}}^{t_j} \{P(t) - P(t_{j-1})\} dP(t) + \int_{t_{j-1}}^{t_j} d[P, P](t)\right)^2 \\ &= 4E\left(\int_{t_{j-1}}^{t_j} \{\Delta P(t_j)\} dP(t)\right)^2 + E\left(\int_{t_{j-1}}^{t_j} d[P, P](t)\right)^2 \\ &= 2E \int_{t_{j-1}}^{t_j} \int_{t_{j-1}}^t d[P, P](s) d[P, P](t) + E\left(\int_{t_{j-1}}^{t_j} d[P, P](t)\right)^2 \\ &= 3E\left(\int_{t_{j-1}}^{t_j} d[P, P]_t\right)^2 \end{aligned}$$

due to Itô's lemma and the isometry property.

Now letting

$$\chi_j^n := \sum_{k < j} \left[\sqrt{\frac{n}{H}} \left\{ \Delta P(t_j) \Delta P(t_k) k_H(t_j - t_k) \cdot (e^{-it_kq} + e^{-it_jq}) \right\} \right],$$

we can construct an upper bound of $E|\chi_j^n|^4$ for $j = n$:

$$\begin{aligned}
E|\chi_j^n|^4 &\leq 2^4 \left(\frac{n}{H}\right)^2 E \left\{ \sum_{h=1}^n \Delta P(t_j) \Delta P(t_{j-h}) \cdot k \left(\frac{t_h}{\Delta t H} \right) \right\}^4 \\
&= 2^4 \left(\frac{n}{H}\right)^2 \sum_{h=1}^n E \left([\Delta P(t_j)]^4 \right) E \left([\Delta P(t_{j-h})]^4 \right) k^4 \left(\frac{t_h}{\Delta t H} \right) \\
&\quad + 2^4 \cdot 6 \left(\frac{n}{H}\right)^2 \sum_{h,m=1}^n E \left([\Delta P(t_j)]^4 \right) E \left([\Delta P(t_{j-h})]^2 \right) E \left([\Delta P(t_{j-m})]^2 \right) k_H^2(t_h) k_H^2(t_m) \\
&= 2^4 \cdot 3^2 \left(\frac{n}{H}\right)^2 \sum_{h=1}^n E \left(\int_{t_{j-1}}^{t_j} [P, P]'(t) dt \right)^2 \left(\int_{t_{j-h-1}}^{t_{j-h}} [P, P]'(t) dt \right)^2 k^4 \left(\frac{t_h}{\Delta t H} \right) \\
&\quad + 2^4 \cdot 6 \cdot 3 \left(\frac{n}{H}\right)^2 \left[\sum_{h,m=1}^n E \left(\int_{t_{j-1}}^{t_j} [P, P]'(t) dt \right)^2 E \left(\int_{t_{j-h-1}}^{t_{j-h}} [P, P]'(t) dt \right) \right. \\
&\quad \quad \left. \times E \left(\int_{t_{j-m-1}}^{t_{j-m}} [P, P]'(t) dt \right) k^2 \left(\frac{t_h}{\Delta t H} \right) k^2 \left(\frac{t_m}{\Delta t H} \right) \right] \\
&\leq 144 \cdot n^2 \times \sup_t \{ [P, P]'(t) \}^4 \times \sup_{1 \leq j \leq n} (\Delta t_i^4) \\
&\quad \times \left[H^{-1} \frac{1}{H} \sum_{h=1}^n k^4 \left(\frac{t_h}{\Delta t H} \right) + \frac{2}{H^2} \sum_{h=1}^n k^2 \left(\frac{t_h}{\Delta t H} \right) \sum_{m=1}^n k^2 \left(\frac{t_m}{\Delta t H} \right) \right] \\
&= n^{-2} H^{-1} C_1 \left(\int_0^\infty k^4(x) dx \right) + 2n^{-2} C_2 \left(\int_0^\infty k^2(x) dx \right)^2 = O(n^{-2}) = o(1), \quad (43)
\end{aligned}$$

provided $\int_0^\infty k^4(x) dx < \infty$, where the last line is due to Assumption 2.1. Consequently the distributional result in Lemma 4 follows.

We now give a result for the off-diagonal element of the estimator. When time stamps are synchronous and sample sizes are balanced, the proof is same as the univariate case. We will give a proof for the most general case; that is, when time stamps are asynchronous and sample sizes are unbalanced. We first show for the bivariate case and will extend the result to general $p \times p$ dimension. Denote the transaction time of the first asset $t_{i,1} = t_i$ and the second asset $t_{j,2} = s_j$ for the sake of simplicity. The centered estimator in (8) can be decomposed into, $\mathcal{E} = M_1 + M_2$, where

$$\begin{aligned}
M_1 &= \sum_{i,j} e^{-is_j q} k_H(t_i - s_j) \Delta P_1(t_i) \Delta P_2(s_j) 1_{\{I_{i,1} \cap I_{j,2} \neq \emptyset\}} - \int_0^{2\pi} e^{-iqt} d[P_1, P_2](t) \\
M_2 &= \sum_{i,j} e^{-is_j q} k_H(t_i - s_j) \Delta P_1(t_i) \Delta P_2(s_j) 1_{\{I_{i,1} \cap I_{j,2} = \emptyset\}}.
\end{aligned}$$

We first derive the asymptotic bias. Let $u_{ij} = t_i \wedge s_j$ and $l_{ij} = t_{i-1} \vee s_{j-1}$. Then,

$$\begin{aligned}
E(M_1) &= E \left(\sum_{i,j} e^{-is_j q} \int_{l_{i,j}}^{u_{i,j}} dP_1(t) \int_{l_{i,j}}^{u_{i,j}} dP_2(s) 1_{\{I_{i,1} \cap I_{j,2} \neq \emptyset\}} - \int_0^{2\pi} e^{-iqt} d[P_1, P_2](t) \right) \\
&\quad - E \left(\sum_{i,j} e^{-is_j q} \int_{l_{i,j}}^{u_{i,j}} dP_1(t) \int_{l_{i,j}}^{u_{i,j}} dP_2(s) \{1 - k_H(t_i - s_j)\} 1_{\{I_{i,1} \cap I_{j,2} \neq \emptyset\}} \right).
\end{aligned}$$

By multivariate Itô calculus, conditionally on $1_{\{I_{i,1} \cap I_{j,2} = \emptyset\}}$, $E(M_1)$ can be written as the expectation of following sums

$$\sum_{i,j} e^{-is_j q} \int_{l_{i,j}}^{u_{i,j}} \{P_1(t) - P_1(l_{i,j})\} dP_2(t) + e^{-is_j q} \int_{l_{i,j}}^{u_{i,j}} \{P_2(t) - P_2(l_{i,j})\} dP_1(t) \quad (44)$$

$$+ \sum_{i,j} \int_{l_{i,j}}^{u_{i,j}} (e^{-is_j q} - e^{-itq}) d[P_1, P_2](t) \quad (45)$$

$$- \sum_{i,j} e^{-is_j q} \int_{l_{i,j}}^{u_{i,j}} dP_1(t) \int_{l_{i,j}}^{u_{i,j}} dP_2(s) \{1 - k_H(t_i - s_j)\}. \quad (46)$$

Using the notation of (3), we see that the order of magnitude of the first term in (44) is

$$\begin{aligned} & \sum_{1 \leq i \leq n_1, 1 \leq j \leq n_2} e^{-is_j q} \int_{l_{i,j}}^{u_{i,j}} \{P_1(t) - P_1(l_{i,j})\} dP_2(t) \\ &= \sum_{l=1}^{N_T} \int_{T_{l-1}}^{T_l} \{P_1(t) - P_1(T_{l-1})\} dP_2(t) - \sum_{i,j} (1 - e^{-is_j q}) \int_{l_{i,j}}^{u_{i,j}} \{P_1(t) - P_1(l_{i,j})\} dP_2(t) \\ &= O_p(N_T^{-1/2}) + O_p(n_2^{-1} N_T^{-1/2}). \end{aligned}$$

The order of the magnitude for the second term in (44) is derived in a similar way. The change of discretization points to the union of the time points are without error and holds analytically. As for the second term we discretize the deterministic function e^{-itq} over the time stamp of s_j ; the order of (45) is therefore $O_p(n_2^{-1})$ since we can replace the summation $\sum_{i,j} \int_{l_{i,j}}^{u_{i,j}}$ by $\sum_{1 \leq j \leq n_2} \int_{s_{j-1}}^{s_j}$. This term is zero for the integrated (co)variance estimator, $q = 0$. Now the asymptotic bias term conditional on the volatility path is given by

$$\begin{aligned} & E \left(\sum_{i,j} e^{-is_j q} \int_{l_{i,j}}^{u_{i,j}} dP_1(t) \int_{l_{i,j}}^{u_{i,j}} dP_2(s) \{1 - k_H(t_i - s_j)\} 1_{\{I_{i,1} \cap I_{j,2} \neq \emptyset\}} \right) \\ & \simeq \sum_{i,j} e^{-is_j q} \int_{l_{i,j}}^{u_{i,j}} d[P_1, P_2](t) \left\{ -\frac{1}{2} k''(0) \left(\frac{t_i - s_j}{\widetilde{\Delta t H}} \right)^2 \right\} 1_{\{I_{i,1} \cap I_{j,2} \neq \emptyset\}} \\ & \leq \left(\frac{n_1 \vee n_2}{(n_1 \wedge n_2) H} \right)^2 \frac{1}{2} \left\{ \frac{n_1 \wedge n_2}{2\pi} \sup_{i,j} |t_i - s_j| 1_{\{I_{i,1} \cap I_{j,2} \neq \emptyset\}} \right\}^2 |k''(0)| \sum_{i,j} e^{-is_j q} \int_{l_{i,j}}^{u_{i,j}} d|[P_1, P_2]|(t) \\ & = \left(\frac{n_1 \vee n_2}{(n_1 \wedge n_2) H} \right)^2 \frac{1}{2} \mathcal{A}^2 |k''(0)| \int_0^{2\pi} e^{-itq} d|[P_1, P_2]|(t). \quad (47) \end{aligned}$$

The first approximation holds by Taylor expansion of $\{k(0) - k_H(t_i - s_j)\}$, Assumption 2.3 and $k'(0) = 0$. Then the order of the stochastic bias M_1 is given by $O_p(N_T^{-1/2}) + O_p(n_2^{-1}) + O_p(\{\frac{n_1 \vee n_2}{(n_1 \wedge n_2) H}\}^2)$ for the estimator at non-zero frequency and $O_p(N_T^{-1/2}) + O_p(\{\frac{n_1 \vee n_2}{(n_1 \wedge n_2) H}\}^2)$ for the integrated (co)variance estimator. In both cases, the leading order term for the bias is the last term under the optimal bandwidth.

We next move on to M_2 , and write

$$M_2 = \int_0^{2\pi} \int_{s < t} f_n(t, s; q) dP_2(s) dP_1(t) + \int_0^{2\pi} \int_{s < t} f_n(s, t; q) dP_1(s) dP_2(t),$$

where $f_n(t, s; q)$ is given in Lemma 3. Its expectation is zero, and moreover by Lemma 1 the

expectation of the quadratic variation $E[M_2, M_2]$ is given by

$$E[M_2, M_2] = \int_0^{2\pi} \int_0^{2\pi} f_n^2(t, s; q) d[P_2, P_2](s) d[P_1, P_1](t) \quad (48)$$

$$+ \int_0^{2\pi} \int_0^{2\pi} f_n(t, s; q) f_n(s, t; q) d[P_2, P_1](s) d[P_1, P_2](t). \quad (49)$$

Now using similar arguments of the proof of Lemma 2, we can show that

$$\begin{aligned} \frac{n_1 \wedge n_2}{H} \int_0^{2\pi} \int_0^{2\pi} f_n^2(t, s; q) d[P_1, P_1](t) d[P_2, P_2](s) \\ \longrightarrow \|k\|^2 \int_0^{2\pi} e^{-i2tq} [P_1, P_1]'(t) [P_2, P_2]'(t) d\mathcal{Q}_{1122}(t) \end{aligned} \quad (50)$$

$$\begin{aligned} \frac{n_1 \wedge n_2}{H} \int_0^{2\pi} \int_0^{2\pi} f_n(t, s; q) f_n(s, t; q) d[P_1, P_2](t) d[P_2, P_1](s) \\ \longrightarrow \|k\|^2 \int_0^{2\pi} e^{-i2tq} ([P_1, P_2]')^2(t) d\mathcal{Q}_{1212}(t). \end{aligned} \quad (51)$$

Hence the proof of Proposition 2 is complete in view of (47), (50) and (51).

Now similar arguments as in the univariate case yields $\sup_i E|\chi_i^n|^4 = O((n_1 \wedge n_2)^{-2})$, where

$$\chi_i^n = \sum_{j: s_j < t_i} \sqrt{m} \Delta P_1(t_i) \Delta P_2(s_j) k_H(t_i - s_j) e^{-is_j q} \mathbf{1}_{\{I_{i,1} \cap I_{j,2} = \emptyset\}};$$

so stable convergence can be established using Lemma 4 and Proposition 1. Higher-dimensional extension of the asymptotic results involves the use of the Cramér-Wold device; it is sufficient to show that any linear combination of the elements of the matrix estimator converges to the corresponding univariate Gaussian random variable. Denote by $\mathcal{R}(q) := \widehat{\mathcal{F}}(\Sigma)(q) - \mathcal{F}(\Sigma)(q)$ and consider the linear combinations $a^\top \mathcal{R}(q) b$ and $c^\top \mathcal{R}(q) d$, where a, b, c, d are some arbitrary constant vectors of conformable dimension. Note that $a^\top \mathcal{R}(q) c b^\top \mathcal{R}(q) d = \text{vech}(ab^\top)^\top (\mathcal{R}(q) \otimes \mathcal{R}(q)) \text{vech}(dc^\top)$, and that its expectation depends on $E\{\mathcal{R}(q) \otimes \mathcal{R}(q)\}$. So the same arguments as above using lemmas in Section 6.1 complete the proof. \blacksquare

6.3 Proof of Theorem 2

By the triangle inequality we have

$$\begin{aligned} & \left\| \Sigma(t) - \frac{1}{2\pi} \sum_{|q| \leq m/2} K_H(\lambda_q) \exp(iqt) \widehat{\mathcal{F}}(\Sigma)(q) \right\|_2 \\ & \leq \left\| \Sigma(t) - \frac{1}{2\pi} \sum_{|q| \leq m/2} K_H(\lambda_q) \exp(iqt) \mathcal{F}(\Sigma)(q) \right\|_2 \\ & \quad + \left(\frac{1}{2\pi} \right)^2 \sum_{|q|, |\ell| \leq m/2} K_H(\lambda_q) K_H(\lambda_\ell) \exp(i(q + \ell)t) \left\| \mathcal{F}(\Sigma)(q) - \widehat{\mathcal{F}}(\Sigma)(q) \right\|_2. \end{aligned}$$

where $m = n/H$. Now Theorem 1 implies that (8) converges in probability to $\mathcal{F}(\Sigma)(q)$ for each q . So if we assume the modulus of continuity of $\Sigma(t)$ is available and is given by (22) then there exists a sequence $\delta(n) \rightarrow 0$ such that

$$\sup_{\delta(n) \leq t \leq 2\pi - \delta(n)} \left\| \Sigma(t) - \sum_{|q| \leq m/2} K_H(\lambda_q) \exp(iqt) \widehat{\mathcal{F}}(\Sigma)(q) \right\|_2 \leq \mathcal{C} \left(\frac{4}{m} \right),$$

as required. \blacksquare

References

- [1] Aït-Sahalia, Y., Fan, J. and Xiu, D. (2010). High-frequency covariance estimates with noisy and asynchronous financial data. *Journal of the American Statistical Association*, 105(492), 1504-1517.
- [2] Aït-Sahalia, Y., Mykland, P. A. and Zhang, L. (2011). Ultra high frequency volatility estimation with dependent microstructure noise. *Journal of Econometrics*, 160(1), 160-175.
- [3] Aldous, D. J. and Eagleson, G. K. (1978). On mixing and stability of limit theorems. *Annals of Probability*, 6(2), 325-331.
- [4] Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, 96(453), 42-55.
- [5] Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. and Shephard, N. (2008). Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica*, 76(6), 1481-1536.
- [6] Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. and Shephard, N. (2011). Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. *Journal of Econometrics*, 162(2), 149-169.
- [7] Barndorff-Nielsen, O. E. and Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(2), 253-280.
- [8] Barndorff-Nielsen, O. E. and Shephard, N. (2004). Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica*, 72(3), 885-925.
- [9] Bibinger, M., Hautsch, N., Malec, P. and Reiss, M. (2014). Estimating the Quadratic Covariation Matrix from Noisy Observations: Local Method of Moments and Efficiency. *Annals of Statistics*, 42(4), 1312-1346.
- [10] Christensen, K., Kinnebrock, S. and Podolskij, M. (2010). Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data. *Journal of Econometrics*, 159(1), 116-133.
- [11] Curato, I. V., Mancino, M. E. and Sanfelici, S. (2014). High Frequency Volatility of Volatility Estimation Free from Spot Volatility Estimates. SSRN Working Paper available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2437989
- [12] Hannan, E. J. (1975). Linear regression in continuous time. *Journal of the Australian Mathematical Society: Series A*, 19(02), 146-159.
- [13] Hayashi, T. and Yoshida, N. (2005). On covariance estimation of non-synchronously observed diffusion processes. *Bernoulli*, 11(2), 359-379.
- [14] Hayashi, T. and 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.

- [15] Hirukawa, M. (2006). A modified nonparametric prewhitened covariance estimator. *Journal of Time Series Analysis*, 27(3), 441-476.
- [16] Ibragimov, I. A. and Linnik, Y. V. (1971). Independent and stationary sequences of random variables. Wolters-Noordhoff, Groningen.
- [17] Jacod, J. (1997). On continuous conditional Gaussian martingales and stable convergence in law. *Séminaire de probabilités de Strasbourg*, 31, 232-246.
- [18] Jacod, J., Li, Y., Mykland, P. A., Podolskij, M. and Vetter, M. (2009). Microstructure noise in the continuous case: the pre-averaging approach. *Stochastic processes and their applications*, 119(7), 2249-2276.
- [19] Jacod, J. and Protter, P. (1998). Asymptotic error distributions for the Euler method for stochastic differential equations. *Annals of Probability*, 26(1), 267-307.
- [20] Kalnina, I. and Linton, O. (2008). Estimating quadratic variation consistently in the presence of endogenous and diurnal measurement error. *Journal of Econometrics*, 147(1), 47-59.
- [21] Malliavin, P. and Mancino, M. E. (2002). Fourier series method for measurement of multivariate volatilities. *Finance and Stochastics*, 6(1), 49-61.
- [22] Malliavin, P. and Mancino, M. E. (2009). A Fourier transform method for nonparametric estimation of multivariate volatility. *Annals of Statistics*, 37(4), 1983-2010.
- [23] Mancino, M. E. and Sanfelici, S. (2008). Robustness of Fourier estimator of integrated volatility in the presence of microstructure noise. *Computational Statistics and Data Analysis*, 52(6), 2966-2989.
- [24] Mykland, P. A. and Zhang, L. (2006). ANOVA for diffusions and Itô processes. *Annals of Statistics*, 34(4), 1931-1963.
- [25] Park, S. and Linton, O. (2012). Realized Volatility: Theory and Applications. in *Handbook of Volatility Models and Their Applications*, Wiley, eds. Bauwens, L., Hafner, C. and Sebastien, L. London, 319-345.
- [26] Parzen, E. (1967). The role of spectral analysis in time series analysis. *Review of the International Statistical Institute*, 35(2), 125-141.
- [27] Rao, T. S. and Gabr, M. M. (1984). An introduction to bispectral analysis and bilinear time series models (Lecture Notes in Statistics). Springer-Verlag, New York.
- [28] Robinson, P. M. (1975). Continuous time regressions with discrete data. *Annals of Statistics*, 3(3), 688-697.
- [29] Xiao, Z. and Linton, O. (2002). A nonparametric prewhitened covariance estimator. *Journal of Time Series Analysis*, 23(2), 215-250.
- [30] Zhang, L. (2006). Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. *Bernoulli*, 12(6), 1019-1043.
- [31] Zhang, L. (2011). Estimating covariation: Epps effect, microstructure noise. *Journal of Econometrics*, 160(1), 33-47.
- [32] Zhang, L., Mykland, P. A. and Aït-Sahalia, Y. (2005). A tale of two time scales. *Journal of the American Statistical Association*, 100(472), 1394-1411.

Table 1: Realized Covariance

		Realized Covariance 5 min aligned					
		BIAS			rMSE		
Sampling	NoiseSignal	(1,1)	(2,2)	(1,2)	(1,1)	(2,2)	(1,2)
Equal	0	(0.01)	0.00	(0.00)	0.17	0.17	0.16
	0.001	0.02	0.03	0.01	0.17	0.18	0.16
	0.01	0.24	0.26	0.13	0.31	0.34	0.23
(3/2,30)	0	(0.00)	0.01	(0.08)	0.17	0.17	0.17
	0.001	0.02	0.03	(0.08)	0.17	0.18	0.17
	0.01	0.25	0.26	(0.07)	0.32	0.33	0.18
(3/2,2)	0	(0.00)	0.01	(0.00)	0.17	0.17	0.16
	0.001	0.03	0.03	0.01	0.18	0.18	0.16
	0.01	0.27	0.26	0.09	0.34	0.35	0.21
(20,30)	0	(0.00)	0.01	(0.07)	0.17	0.17	0.17
	0.001	0.02	0.02	(0.07)	0.17	0.17	0.18
	0.01	0.24	0.22	(0.04)	0.31	0.30	0.19

		Realized Covariance Refresh Time aligned					
		BIAS			rMSE		
Sampling	NoiseSignal	(1,1)	(2,2)	(1,2)	(1,1)	(2,2)	(1,2)
Equal	0	(0.00)	0.00	(0.00)	0.01	0.01	0.01
	0.001	4.52	4.54	2.10	4.52	4.54	2.10
	0.01	45.24	45.41	21.07	45.24	45.41	21.07
(3/2,30)	0	0.01	0.01	(0.02)	0.08	0.08	0.08
	0.001	0.23	0.23	0.07	0.25	0.24	0.10
	0.01	2.24	2.25	0.92	2.25	2.26	0.93
(3/2,2)	0	(0.00)	0.00	(0.17)	0.02	0.02	0.17
	0.001	1.78	1.80	0.53	1.78	1.80	0.53
	0.01	17.83	17.98	6.77	17.84	17.99	6.78
(20,30)	0	0.01	0.00	(0.27)	0.09	0.07	0.28
	0.001	0.16	0.15	(0.24)	0.18	0.17	0.25
	0.01	1.55	1.55	0.08	1.56	1.56	0.15

Table 2: 2 dimensional covariation matrix - continuous SV ($\cdot/100$)

NRS	Realized Cov Refresh Time		Realized Cov 5 min		HY		Realized Kernel Refresh Time		FRK average H		FRK opt H		
	(1,1)	(1,2)	(1,2)	(2,2)	(1,2)	(2,2)	(1,1)	(1,2)	(2,2)	(1,1)	(1,2)	(2,2)	
Sampling: (3/2,30) Balanced													
0	bias 1.0	(1.9)	0.8	(0.1)	(7.6)	0.9	1.0	1.0	(1.3)	0.8	0.9	(0.6)	1.4
	rmse 8.0	7.6	8.1	17.1	17.1	16.8	7.6	8.2	7.5	8.0	20.9	19.6	20.8
0.001	bias 23.2	7.3	22.9	2.2	(7.7)	3.3	10.3	1.6	0.9	1.6	1.0	0.8	1.7
	rmse 24.7	10.6	24.6	17.2	17.2	17.6	12.9	16.7	15.2	15.4	30.2	28.9	32.3
0.01	bias 225	92.3	225	24.5	(6.6)	25.5	96.4	3.5	2.3	4.0	0.6	(1.2)	(1.5)
	rmse 226	93.6	226	31.6	17.9	33.0	97.6	25.3	23.7	25.5	43.8	40.1	46.7
Sampling: (3/2,30) Unbalanced													
0	bias 0.1	(2.6)	0.5	(0.9)	(8.2)	0.9	0.4	0.0	(1.9)	0.6	1.0	(3.1)	1.4
	rmse 6.7	6.9	7.0	17.4	18.0	17.8	6.6	7.0	7.0	7.3	15.9	14.8	15.9
0.001	bias 23.5	6.6	21.3	1.4	(8.0)	3.8	9.5	2.0	1.8	3.1	1.5	(12.5)	5.0
	rmse 24.7	9.6	22.6	17.9	18.5	18.7	11.7	15.7	14.6	15.6	10.5	15.5	12.2
0.01	bias 239	92.5	210	23.5	(6.9)	27.3	95.7	2.6	2.3	4.6	1.9	(1.9)	11.6
	rmse 240	93.9	211	31.7	20.3	35.2	96.8	24.0	23.3	25.9	18.1	16.7	21.7
Sampling: (20,30) Unbalanced													
0	bias 0.1	(27.1)	0.7	0.6	(6.7)	1.3	0.8	0.2	(23.7)	0.7	0.9	(0.4)	1.8
	rmse 8.3	28.0	8.2	17.2	17.5	17.8	9.3	8.4	24.8	8.2	20.2	19.1	20.9
0.001	bias 16.8	(23.1)	16.2	3.1	(6.1)	3.7	4.9	2.0	0.1	2.8	1.1	0.1	1.9
	rmse 19.0	24.4	18.6	18.2	18.0	18.8	11.0	16.3	15.0	16.2	23.3	22.1	24.2
0.01	bias 162	10.4	155	26.1	(1.7)	26.0	41.1	3.8	2.2	4.7	1.7	0.6	1.9
	rmse 164	16.5	156	34.0	20.4	34.3	43.6	24.3	23.4	26.7	32.2	30.8	34.5
													35.8
													34.1
													39.4

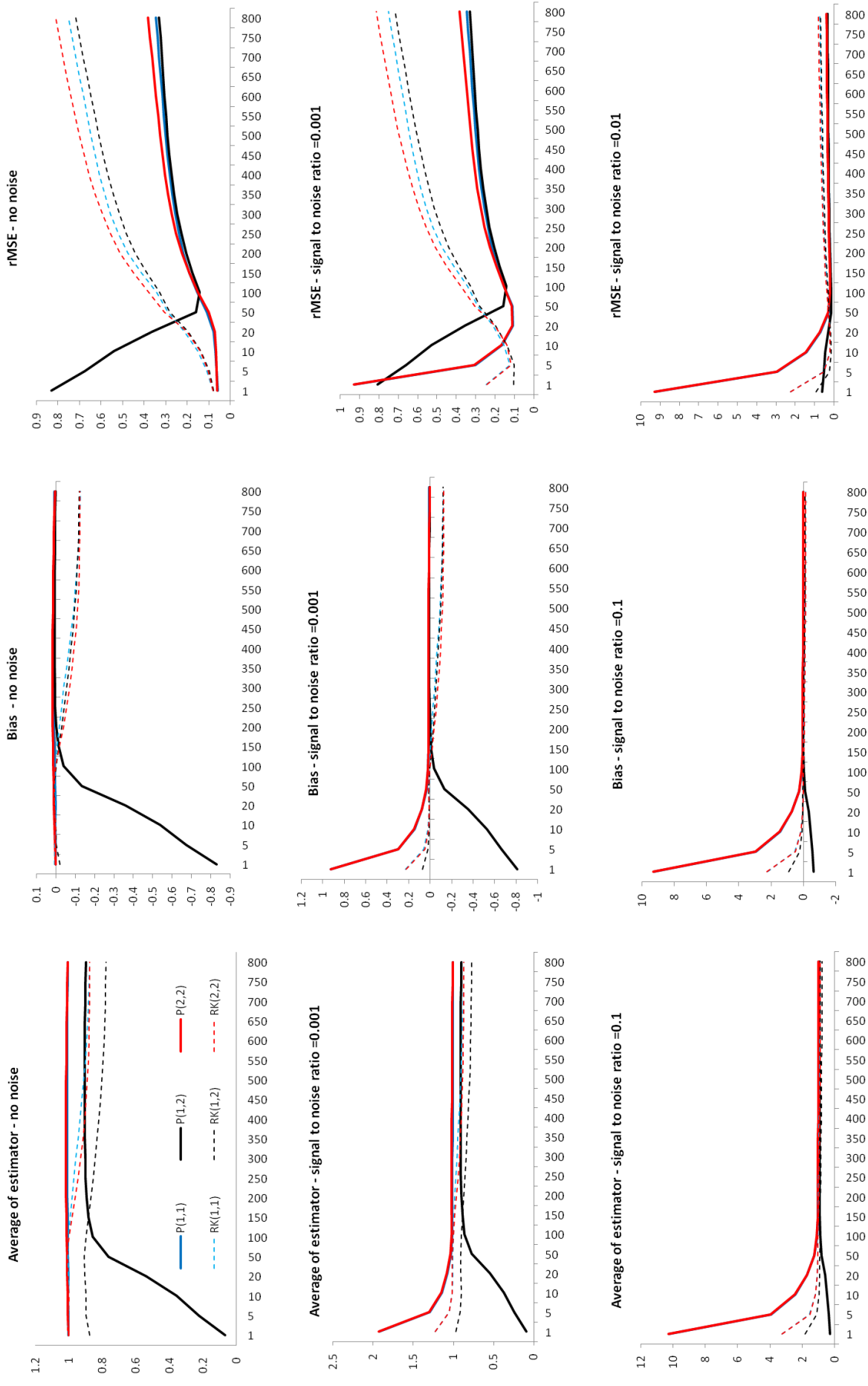
Table 3: 2 dimensional covariation matrix - jump diffusion SV ($\cdot/100$)

NRS	Realized Cov		Realized Cov		HY	Realized Kernel		FRK		FRK		
	Refresh Time	(1,1) (1,2) (2,2)	5 min	(1,1) (1,2) (2,2)		Refresh Time	average H	opt H				
Sampling: (3/2,30) Balanced												
0	bias	(0.3) (2.0) (18.7)	(2.2) (6.0) (20.4)	(0.3) (0.2) (1.5) (18.6)	(0.1) (0.7) (18.8)	(0.1) (0.7) (18.8)	(0.7) (18.8)					
	rmse	6.7 18.1 20.3	17.2 21.9 25.4	18.6 7.1 18.3 20.3	18.0 23.4 24.6	18.0 23.4 24.6	18.0 23.4 24.6					
0.001	bias	23.3 8.2 4.9	0.5 (5.6) (17.2)	9.8 1.5 1.3 (16.9)	0.0 0.1 (19.3)	0.0 0.1 (19.3)	0.1 (19.3)					
	rmse	24.7 19.8 9.9	16.8 22.0 23.0	21.1 13.4 22.0 21.8	29.1 29.2 31.9	29.1 29.2 31.9	29.1 29.2 31.9					
0.01	bias	233 96.7 211	24.2 (3.6) 7.0	98.8 2.6 2.1 (15.8)	1.4 1.3 (18.0)	10.2 (0.7) (10.0)	(0.7) (10.0)					
	rmse	234 99.6 212	30.8 22.6 19.9	101.8 19.8 24.5 25.3	42.6 38.9 42.8	17.2 21.2 17.1	17.2 21.2 17.1					
Sampling: (3/2,30) Unbalanced												
0	bias	(0.3) (1.9) (18.8)	(1.6) (6.6) (21.2)	(0.1) (0.2) (1.6) (18.8)	0.2 (2.5) (19.1)	0.0 (0.5) (18.9)						
	rmse	7.1 17.7 20.4	16.9 22.0 26.1	18.2 7.1 17.9 20.5	13.7 21.1 23.0	20.1 24.4 25.3						
0.001	bias	24.0 8.0 2.6	0.8 (6.6) (19.0)	10.0 0.5 (17.5)	1.5 (8.3) (14.6)	0.7 (3.2) (17.5)						
	rmse	25.5 19.7 9.3	17.6 22.4 24.7	21.1 14.0 20.7 21.5	9.2 19.1 18.0	12.8 20.5 21.6						
0.01	bias	243 94.1 192	23.1 (6.1) 3.4	97.8 1.4 1.0 (15.7)	1.8 (1.6) (8.6)	3.8 (3.3) 2.6						
	rmse	244 97.1 193	32.1 24.2 20.8	100.7 22.1 24.7 23.8	15.2 21.6 17.3	12.6 19.6 14.8						
Sampling: (20,30) Unbalanced												
0	bias	(0.1) (16.7) (18.0)	(1.4) (5.4) (20.1)	(0.0) (0.1) (14.7) (18.0)	(0.0) (0.9) (18.9)	(0.2) (0.1) (18.9)						
	rmse	6.8 21.5 19.5	16.8 21.7 26.3	19.2 6.8 20.4 19.6	17.6 23.3 24.2	24.2 27.0 28.3						
0.001	bias	15.3 (13.2) (2.7)	1.4 (4.7) (17.6)	3.7 1.6 (0.4) (17.9)	0.1 (0.4) (18.5)	0.1 (0.0) (18.8)						
	rmse	17.1 19.1 8.9	17.1 21.6 24.7	19.6 14.5 21.2 21.7	19.7 24.5 25.1	23.9 26.7 28.0						
0.01	bias	159 20.6 136	24.9 0.1 4.9	40.1 3.0 1.2 (15.9)	0.8 0.9 (17.8)	6.2 0.7 (10.5)						
	rmse	160 26.8 137	32.2 22.3 20.9	22.7 25.8 24.1	30.2 30.6 32.6	17.9 22.6 17.7						

Table 4: Scalar function of 10 dimensional covariation matrix

		max (eigenvalue)		portfolio		
Noise to Signal Ratio=0		Bias	rMSE	Bias	rMSE	
	RV_refresh	(2.34)	2.75	1.76	2.74	
	RV_fixed	(0.85)	3.18	0.14	4.09	
	Realized Kernel	(2.21)	2.65	1.66	2.67	
	Fourier RK	(1.18)	2.17	0.26	2.51	
Noise to Signal Ratio=0.001						
	RV_refresh	7.22	7.50	27.20	27.47	
	RV_fixed	0.40	3.28	4.14	6.31	
	Realized Kernel	0.38	3.00	1.67	4.16	
	Fourier RK	(0.28)	1.95	3.88	4.87	
		avgH	(0.47)	2.64	0.73	3.67
		maxH	(0.52)	3.13	(0.20)	3.99
Noise to Signal Ratio=0.01						
	RV_refresh	127.24	127.81	256.02	257.05	
	RV_fixed	15.42	16.54	40.37	41.99	
	Realized Kernel	1.29	4.82	3.23	6.86	
	Fourier RK	1.22	3.27	6.88	8.55	
		avgH	0.13	4.02	1.76	5.58
		maxH	(0.03)	4.92	0.67	5.96

Figure 6: Simulation Result : Balanced, Sampled at $= \{3/2, 30\}$.^a



^aComparison of 2×2 covariance matrix and its scalar function estimated by Realized Kernel and Fourier Realized Kernel for range of bandwidths. Sample size of two assets are unbalanced and time stamps are poisson sampled at rate $= \{3/2, 30\}$.

Figure 7: Simulation Result : Unbalanced, Sampled at $= \{3/2, 30\}$

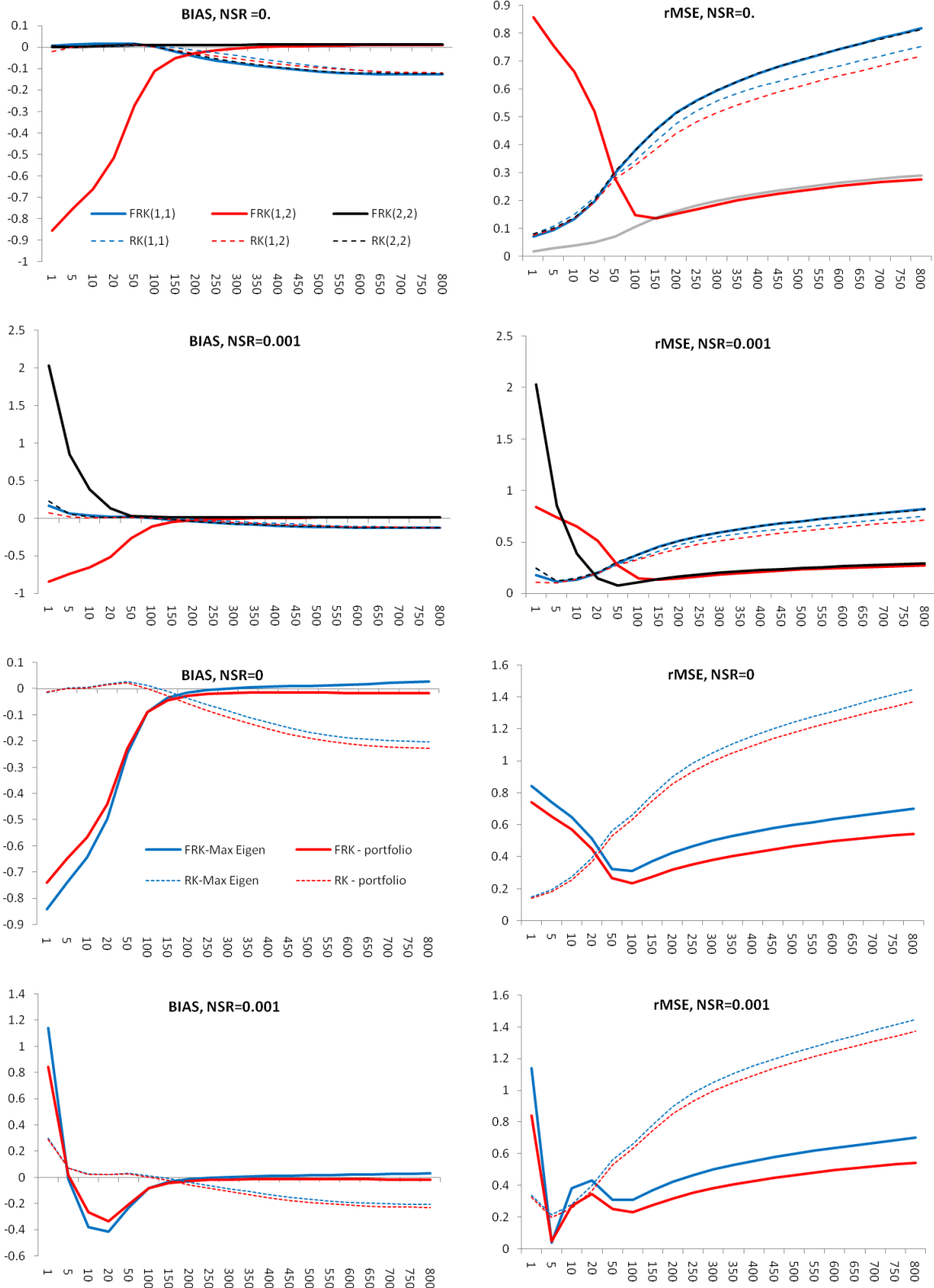


Figure 8: Simulation Result : Unbalanced, Sampled at = {20, 30}

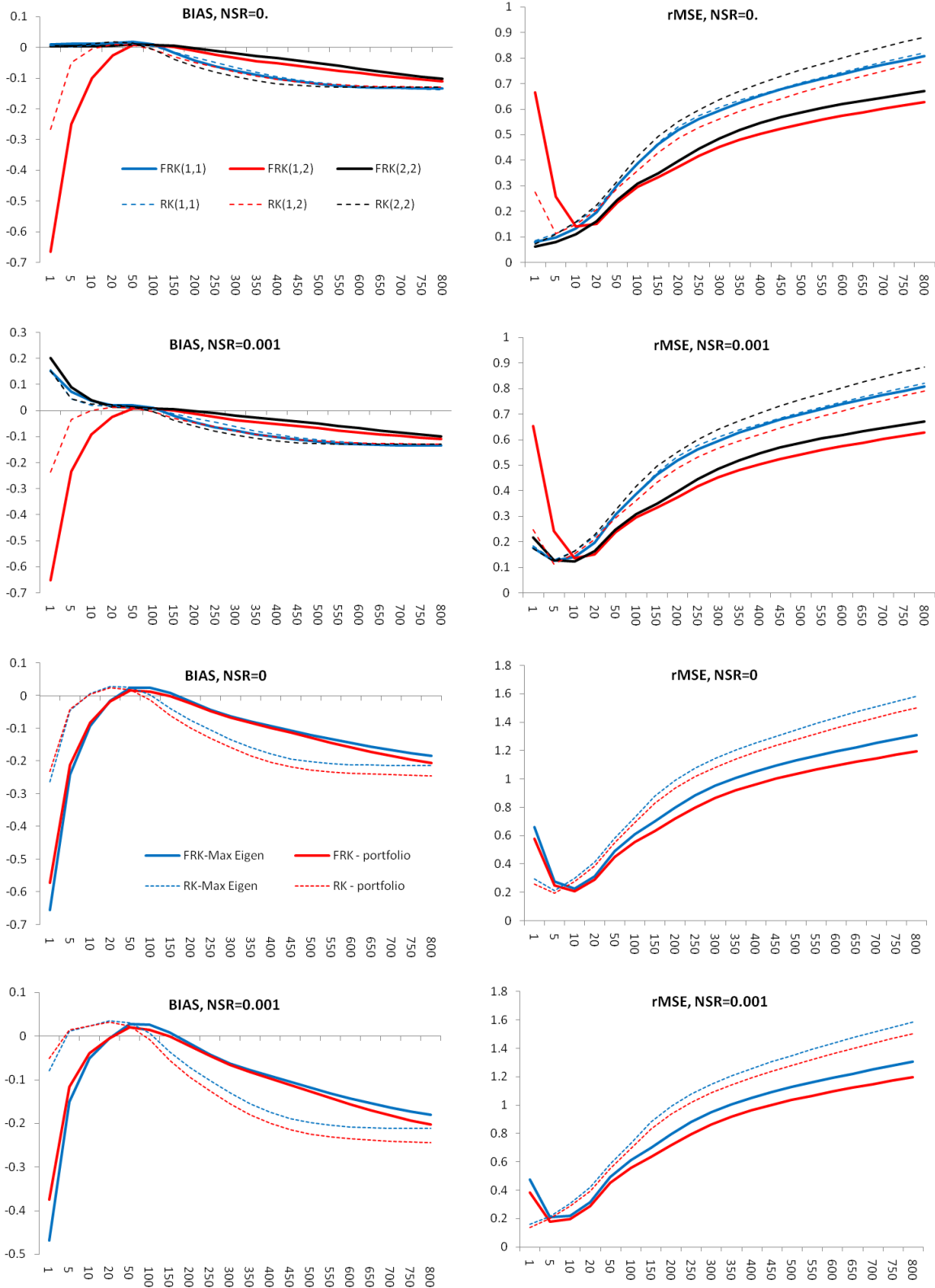


Figure 9: Simulation Result : Unbalanced, Sampled at $= \{3/2, 2\}$

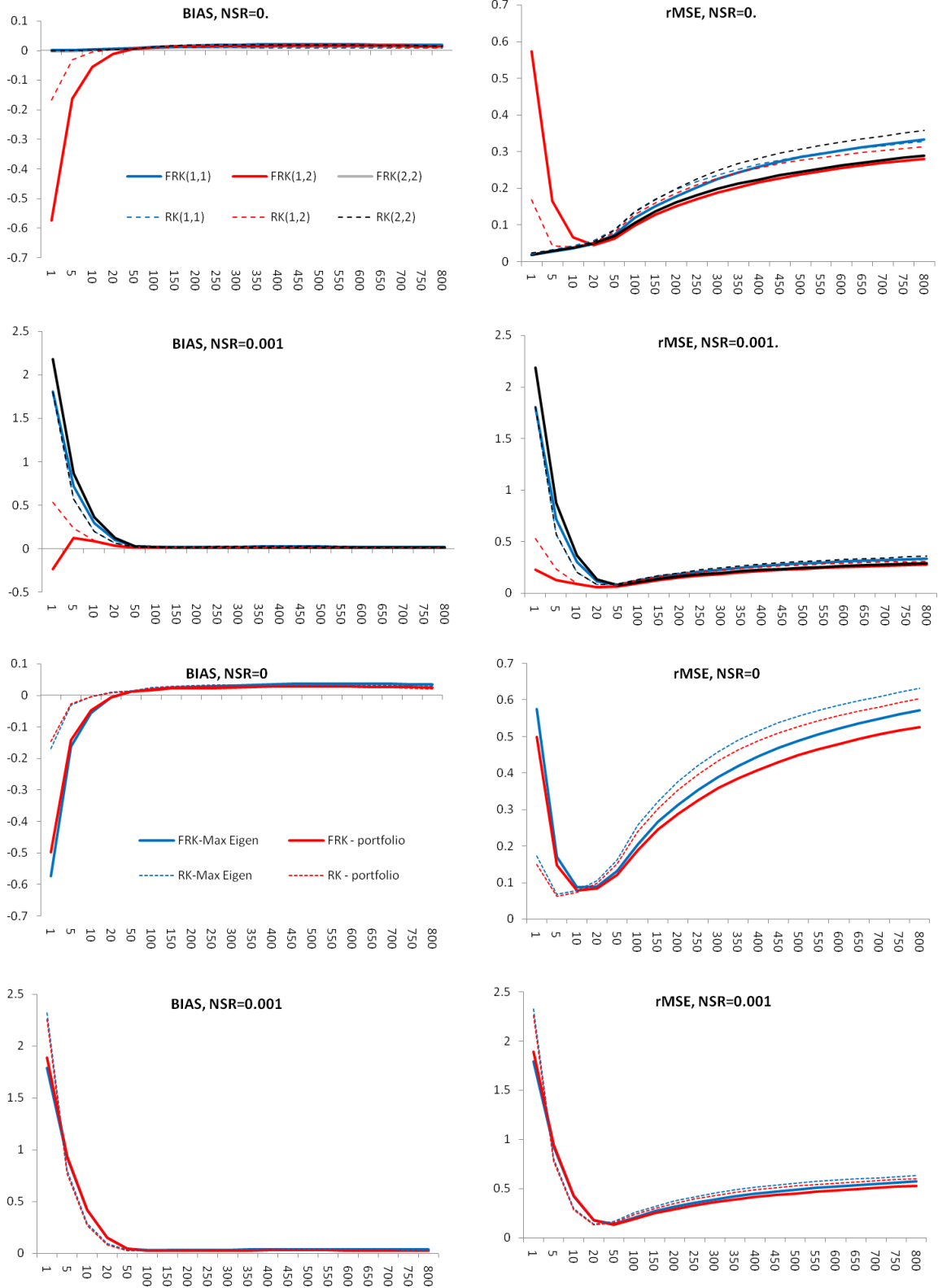


Figure 10: Covariation matrix estimates : element (1, 1) to (2, 5)

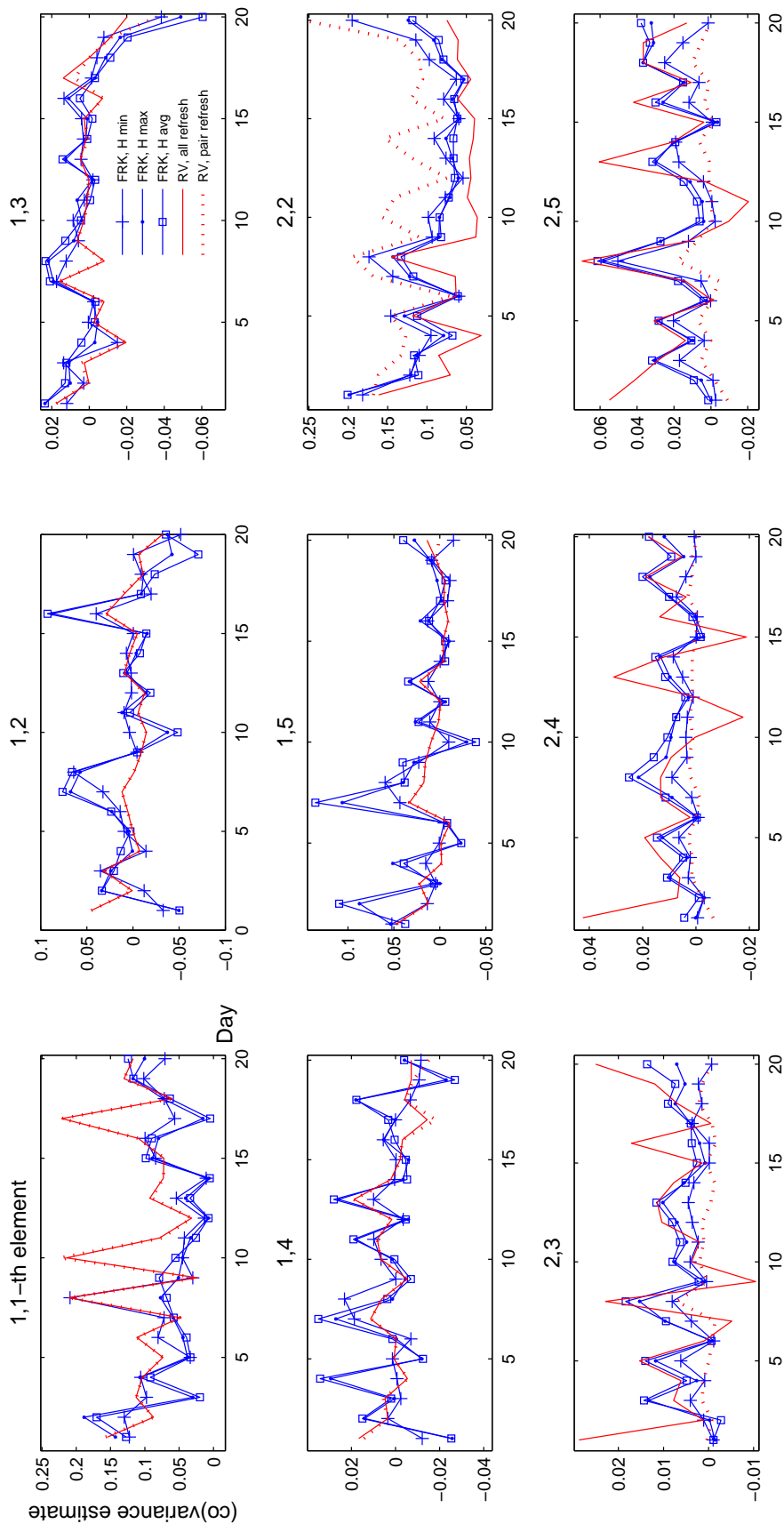


Figure 11: Covariation matrix estimates : element (3, 3) to (5, 5)

