Statistics and Its Interface Volume 8 (2015) 305–319

Spectral analysis of quadratic variation in the presence of market microstructure noise*

Fangfang Wang

We analyze the ex-post variation of equity prices in the frequency domain. A realized periodogram-based estimator is proposed, which consistently estimates the quadratic variation of the log equilibrium price process. For prices which are contaminated by market microstructure noise, the proposed estimator behaves like a filter: it removes the noise by filtering out high frequency periodograms. In other words, the proposed estimator converts high frequency data into low frequency periodograms. We show, through a simulation study and an application to the General Electric transaction prices, that the proposed estimator is insensitive to the choice of sampling frequency and it is competitive with other existing volatility measures.

Keywords and phrases: Jump diffusion, Quadratic variation, Periodogram, Discrete Fourier transform, Spectral density, Market microstructure noise.

1. INTRODUCTION

Suppose that $X(t)$ is the log equilibrium price of an equity at time t and it follows a continuous-time semimartingale. Denote by $[X,X](T)$ its quadratic variation (henceforth denoted by QV) over a fixed interval of time $[0,T]$. It follows from probability theory that a sum of squared increments over fine intervals consistently estimates $[X,X](T)$. To be specific, supposing that X is discretely sampled at $\Pi_n = \{0 = t_0 < t_1 < \ldots < t_n = T\}$, $RV_n^X \doteq \sum_{j=1}^n (X(t_j) - X(t_{j-1}))^2$ converges to $[X,X](T)$ in probability as the mesh of the partition Π_n goes to zero, and RV_n^X is known as realized variance in financial econometrics. See for instance [4], [3], [29], [8], [1], [10], [25], [31], among many others.

It has been widely recognized that ultra high frequency prices are contaminated by market microstructure(MS) noise. In other words, when the mesh of Π_n is small enough (less than 1 minute), the observed prices are not true prices. Let Y_j be noisy observation of $X(t)$ at t_j for $j = 0,\ldots,n$. The realized variance computed from the noisy prices, $RV_n^Y = \sum_{j=1}^n (Y_j - Y_{j-1})^2$, can not identify the true price variation – see for instance [2], [22], [7], [11], among others. Because RV_n^Y can be decomposed as a sum of periodograms that are attributable to different Fourier frequencies and RV_n^Y contains both signal X and noise, one can think of extracting signal from RV_n^Y by filtering out periodograms due to noise.

In this paper, we propose a realized periodogram-based estimator of QV which is the weighted average of periodograms over non-zero Fourier frequencies. This is motivated by the fact that the estimation of QV is analogue to estimating spectral density of a stationary process at zero frequency. The proposed estimator transfers high frequency data into low frequency periodograms. The estimator is determined by three factors: sample size, a so-called cut-off frequency which controls the number of periodograms to be included, and spectral window. We will study the periodogram-based estimator with and without the MS contamination, and derive the optimal cut-off frequency which minimizes the effect of the noise on the periodogram-based estimator.

The proposed estimator generalizes the Fourier estimator of [27], which is defined as $\hat{\sigma}^{2(X)}_{N,n} = (2N+1)^{-1}\sum_{k=-N}^{N} |c_k|^2$ with $c_k = \sum_{j=1}^n e^{-2\pi\sqrt{-1}kt_{j-1}/T}(X(t_j) - X(t_{j-1}))$. The proposed estimator features a weighting structure and it does not include zero frequency periodogram. Assigning different weight (or a spectral window) to a different periodogram can potentially improve the finite sample performance. We will show, through a simulation study, that the weight does matter. Periodogram at zero-frequency is excluded so that the estimator is invariant to the mean shift in the data. In other words, one does not need to demean raw data before computing the estimator.

The new estimator has a natural connection with the auto-covariance-based estimators of QV in the time domain, for instance, the kernel-based estimator of [36] and [22], the realized kernel of [11]. Therefore this work fills the gap between time-domain and frequency-domain analysis of QV, and it enriches the literature of the ex-post measure of price variation. Moreover, compared with the auto-covariance-based estimators, the proposed estimator is easy to implement and the fast Fourier transform makes it computationally efficient.

The rest of paper is organized as follows. In section 2, we lay out the model and introduce the periodogram-based estimator. Its properties are derived in Section 3 and Section 4. The first focuses on the true price process, while the

*The author would like to thank the Editor, Associate Editor and the Referee for their valuable comments.

latter deals with the noisy observations. Section 5 discusses the choice of the cut-off frequency in a finite sample. In Section 6, we perform a simulation experiment to evaluate the accuracy of the proposed estimator, and an empirical exercise is presented in Section 7. Section 8 concludes the paper. All the proofs are collected in an Appendix.

2. A REALIZED PERIODOGRAM-BASED ESTIMATOR OF QV

Suppose that the log equilibrium price $X(t)$ is defined on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ and has the dynamics

(1)
$$X(t) = X_0 + \int_0^t \mu(s)ds + \int_0^t \sigma(s)dB(s) + \int_{[0,t]\times\mathbb{R}} x\tilde{\mathcal{J}}_X(ds \times dx),$$

where B is a standard Brownian motion, $\mu(t)$ and $\sigma^2(t)$ are adapted processes satisfying $E[\int_0^T \mu^2(t)dt] < \infty$ and $E[\int_0^T \sigma^4(t)dt] < \infty$. The volatility $\sigma(t)$ is almost surely positive. The jump measure $\mathcal{J}_X$ is a Poisson random measure with Lévy measure ν satisfying $\int_{|x|\leq 1} x\nu(dx) < \infty$, $\int_{|x|>1} x^2\nu(dx) < \infty$, and $\tilde{\mathcal{J}}_X$ is the compensated jump measure.

The MS noise $\epsilon_j = Y_j - X(t_j)$ satisfies the following assumption

Assumption 2.1. *(1) $\{\epsilon_j\}$ is white noise with variance σ_ϵ^2, and $E(\epsilon_j^4) = \eta\sigma_\epsilon^4 < \infty$. (2) $\{\epsilon_j\}$ is uncorrelated with X at all lags and leads.*

Let $\Delta X_j = X(t_j) - X(t_{j-1})$ and $\Delta Y_j = Y_j - Y_{j-1}$. If $\{\Delta X_j\}$ is covariance stationary, the spectral density of ΔY_j is given by

$$f_{\Delta Y}(\omega) = (2\pi n)^{-1}E[X, X](T) + \sigma_\epsilon^2(1 - \cos\omega)/\pi, \tag{2}$$

for $\omega \in (-\pi, \pi]$. At $\omega = 0$, $2\pi n f_{\Delta Y}(0)$ is the expected value of QV, $E[X, X](T)$. The noise affects spectral density at non-zero frequencies. The spectral density increases as ω moves towards π or $-\pi$. This is consistent with the empirical feature of tick-by-tick data. Figure 1 plots the power spectrum of log-returns for General Electric transaction prices on Feb 1, 2007.[1] The periodogram is small around zero frequency and tends to increase with the frequency. So the problem of estimating QV in the presence of noise would be analogue to the estimation of $f_{\Delta Y}(0)$, i.e., using periodograms close to zero frequency.

Define the discrete Fourier transform of an arbitrary sequence $\{Z_j\}_{j=1}^n$ at frequency $\omega \in (-\pi, \pi]$:

$$J_n^Z(\omega) = \sum_{j=1}^n e^{-i\omega j} Z_j, \tag{3}$$

[1]See Section 7 for details.

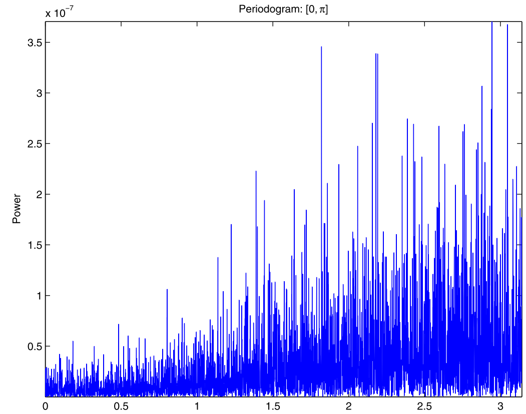


Figure 1. Power Spectrum for GE transaction prices on Feb 1, 2007.

where $i = \sqrt{-1}$ is the imaginary unit. The sample periodogram of $\{Z_j\}_{j=1}^n$ at ω is then

$$I_n^Z(\omega) = n^{-1}|J_n^Z(\omega)|^2. \tag{4}$$

We propose a realized periodogram-based estimator of QV:

$$\widehat{F}^{\Delta Y}(N, n) = n \sum_{1\leq k\leq N} W(k) I_n^{\Delta Y}(\omega_k), \tag{5}$$

where $\omega_k = 2\pi k/n$ is the Fourier frequency, and the weight $W(k)$ satisfies:

Assumption 2.2. *$W(k) \geq 0$ for $k \in \mathbb{Z}$. $\sum_{k=1}^N W(k) = 1$, $W(-k) = W(k)$ for $1 \leq k \leq N$ and 0 otherwise. $\lim_{N\to\infty} \sum_{k=1}^N W^2(k) = 0$.*

As a result $\widehat{F}^{\Delta Y}(N, n) = \sum_{1\leq|k|\leq N} 2^{-1}W(k)|J_n^{\Delta Y}(\omega_k)|^2$. We exclude periodogram at zero frequency, in that periodograms are invariant to location shift at the non-zero Fourier frequencies.[2] Thus we don't need to demean raw data before computing the estimator.

The proposed estimator is determined by two types of frequencies: the sampling frequency T/n and the cut-off Fourier frequency $2\pi N/n$ (for irregularly sampled data, T/n measures the average distance between successive observations). The first determines how often the data should be sampled, while the latter controls the number of Fourier frequencies to be included in estimation. Note that $RV_n^Y = \sum_{1\leq k\leq n} |J_n^{\Delta Y}(\omega_k)|^2$ due to the Parseval's theorem. When n is even and $W(k) = 1/N$, we have $\widehat{F}^{\Delta Y}(n/2, n) = RV_n^Y$. So the periodogram-based estimator includes realized variance as a special case.

[2]This is common practice in estimating spectral density at zero-frequency, see for instance [21], [18].

Remark 2.1. *The periodogram-based estimator has the following time-domain representation*

$$\widehat{F}^{\Delta Y}(N,n) = \sum_{h=-(n-1)}^{n-1} W^*(h)\gamma_Y(h), \tag{6}$$

where $\gamma_Y(h) = \sum_{j=1}^{n-|h|} \Delta Y_j \Delta Y_{j+|h|}$ *and* $W^*(h) = \sum_{k=1}^{N} W(k)\cos(\omega_k h)$ *is lag window. The right hand side of (6) is similar in spirit to the Realized Kernel of [11] and [14]. In particular, the non-negative Realized Kernel of [14] is defined as*

$$K(Y) = \sum_{h=-H}^{H} \Lambda(\frac{h}{H})\gamma_Y(h), \tag{7}$$

where the kernel weight $\Lambda(\cdot)$ *has the property of* $\Lambda(0) = 1$, $\Lambda(1) = 0$, *and* $\Lambda'(0) = \Lambda'(1) = 0$. *Moreover, considering* $\gamma_Y(h) = \sum_{1\le k\le n} I_n^{\Delta Y}(\omega_k)e^{ih\omega_k}$, *we can rewrite* $K(Y)$ *as*

$$K(Y) = \sum_{1\le k\le n} \Lambda^*(k)|J_n^{\Delta Y}(\omega_k)|^2, \tag{8}$$

where $\Lambda^*(k) = n^{-1}\sum_{h=-H}^{H} \Lambda(\frac{h}{H})e^{ih\omega_k}$ *and* $\sum_{k=1}^{n} \Lambda^*(k) = 1$. *The two estimators,* $\widehat{F}^{\Delta Y}(N,n)$ *and* $K(Y)$, *are linked with each other. Yet, they work differently:* $\widehat{F}^{\Delta Y}(N,n)$ *eliminates the noise by removing high frequency periodograms – in the frequency domain, while the realized kernel estimator eliminates the noise by removing the autocovariances at higher lags – in the time domain.*

3. FREQUENCY-DOMAIN REPRESENTATION OF QV

We first study the periodogram-based estimator in the absence of the MS noise, namely, $\widehat{F}^{\Delta X}(N,n) = \sum_{1\le k\le N} W(k)|J_n^{\Delta X}(\omega_k)|^2$.

Suppose that there is a continuous record on the sample path of X over $[0,T]$. We define a continuous-time counterpart of $\widehat{F}^{\Delta X}(N,n)$:

$$\widehat{F}^{dX}(N) = T^2 \sum_{1\le k\le N} W(k)|\mathcal{F}(dX)(k)|^2,$$

where $\mathcal{F}(dX)(k) = T^{-1}\int_0^T e^{-2\pi ikt/T}dX(t)$ is the continuous-time Fourier transform of X at $k \in \mathbb{Z}$. Proportion 3.1 below shows that the drift has a negligible effect on $\widehat{F}^{dX}(N)$.

Proposition 3.1. *Let* $X^c(t) = X(t) - \int_0^t \mu(s)ds$. *Under Assumption 2.2, we have*

$$\lim_{N\to\infty} \sum_{1\le k\le N} W(k)\left(|\mathcal{F}(dX)(k)|^2 - |\mathcal{F}(dX^c)(k)|^2\right) = 0 \text{ in } L^2.$$

Moreover,

Proposition 3.2. *Let* $X^c(t) = X(t) - \int_0^t \mu(s)ds$, *and define* $\widehat{F}^{dX^c}(N)$ *accordingly. Under Assumption 2.2,*

$$plim_{N\to\infty}\widehat{F}^{dX^c}(N) = \int_0^T \sigma(t)^2dt + \sum_{t\le T}(X(t) - X(t-))^2, \tag{9}$$

where 'plim' means convergence in probability.

Therefore $\widehat{F}^{dX}(N)$ is consistent for $[X,X](T)$.

Consider the discretely sampled process $\{X(t_j), j = 0,1,\ldots,n\}$. The next theorem derives the consistency of $\widehat{F}^{\Delta X}(N,n)$.

Theorem 3.1. *Suppose that the conditions in Proposition 3.2 hold. Let* $\rho(n) = \max_{j=1,2,\ldots,n}|t_j - Tj/n| \vee |t_{j-1} - Tj/n|$. *Assume that* $\rho(n) = O(n^{-1})$ *and* $N = o(n)$. *Then*

$$plim_{N,n\to\infty}\widehat{F}^{\Delta X}(N,n) = \int_0^T \sigma(t)^2dt + \sum_{t\le T}(X(t) - X(t-))^2. \tag{10}$$

Remark 3.1. *As a corollary to Propositions 3.1 and 3.2, we have*

$$\mathcal{F}(dX) \star_B \mathcal{F}(dX) = \frac{1}{T}\mathcal{F}([dX,dX]), \tag{11}$$

where $\mathcal{F}(d[X,X])(k) = T^{-1}\int_0^T e^{-2\pi ikt/T}d[X,X](t)$, *and* $\star_B$ *is the Bohr convolution product.*[3] *The convergence of convolution product in (11) is held in probability. Equality (11) extends Theorem 2.1 of [27] to jump diffusion processes. More generally, consider two jump diffusion processes* $X_1(t)$ *and* $X_2(t)$ *satisfying (1). Denote by* $[X_1,X_2](t)$ *the quadratic covariation between* $X_1(t)$ *and* $X_2(t)$. *Note that* $[X_1,X_2](t) = 2^{-1}\{[X_1+X_2,X_1+X_2](t) - [X_1,X_1](t) - [X_2,X_2](t)\}$. *We have*

$$\begin{aligned}&T^{-1}\mathcal{F}(d[X_1,X_2])\\&= 2^{-1}\{\mathcal{F}(dX_1)\star_B\mathcal{F}(dX_2) + \mathcal{F}(dX_2)\star_B\mathcal{F}(dX_1)\}.\end{aligned}$$

If the true price process were observable, the peridogram-based estimator consistently estimates QV. For sufficiently large n, $\widehat{F}^{\Delta X}(N,n)$ provides a spectral decomposition of $[X,X](T)$. Since the drift $\mu(t)$ does not influence the asymptotic behavior of $\widehat{F}^{dX}(N)$ and periodogram is invariant to location shift at the non-zero Fourier frequencies, we will assume $\mu(t) = 0$ in the rest of the paper.

4. ESTIMATE QV USING NOISY OBSERVATIONS

In this section, we will study the impact of market MS noise on the periodogram-based estimator with a focus on the finite sample properties.

[3] For f and g: $\mathbb{Z} \mapsto \mathbb{C}$, the Bohr convolution product between f and g, denoted by $f \star_B g$, defines a map from $\mathbb{Z}$ to $\mathbb{C}$ such that $(f\star_B g)(k) = \lim_{N\to\infty}(2N+1)^{-1}\sum_{s=-N}^{N} f(s)g(k-s)$ provided that $\lim_{N\to\infty}(2N+1)^{-1}\sum_{s=-N}^{N} f(s)g(k-s)$ exists.

Denote by $I_n^{\Delta X,\Delta\epsilon}(\omega)$ the cross periodogram between $\{\Delta X_j\}$ and $\{\Delta\epsilon_j\}$:

$$I_n^{\Delta X,\Delta\epsilon}(\omega) = n^{-1} J_n^{\Delta X}(\omega) J_n^{\Delta\epsilon}(-\omega).$$

For any $\omega \in (-\pi,\pi]$, $I_n^{\Delta Y}(\omega) = I_n^{\Delta X}(\omega) + I_n^{\Delta\epsilon}(\omega) + I_n^{\Delta X,\Delta\epsilon}(\omega) + I_n^{\Delta X,\Delta\epsilon}(-\omega)$. Consequently, we can rewrite $\widehat{F}^{\Delta Y}(N,n)$ as the sum of three components:

(12)
$$\widehat{F}^{\Delta Y}(N,n) = \widehat{F}^{\Delta X}(N,n) + \widehat{F}^{\Delta\epsilon}(N,n) + \hat{F}^{\Delta X,\Delta\epsilon}(N,n),$$

where $\widehat{F}^{\Delta\epsilon}(N,n) = \sum_{k=1}^{N} W(k)|J_n^{\Delta\epsilon}(\omega_k)|^2$, and $\hat{F}^{\Delta X,\Delta\epsilon}(N,n) = \sum_{k=1}^{N} W(k)(nI_n^{\Delta X,\Delta\epsilon}(\omega_k) + nI_n^{\Delta X,\Delta\epsilon}(-\omega_k))$. The two additional terms are due to the presence of the noise. We measure the accuracy of $\widehat{F}^{\Delta Y}(N,n)$ as an estimator of QV by its mean squared error: $E(\widehat{F}^{\Delta Y}(N,n) - [X,X](T))^2$. To understand the noise-induced bias, we derive the mean and variance of $\widehat{F}^{\Delta X}(N,n)$, $\widehat{F}^{\Delta\epsilon}(N,n)$, and $\widehat{F}^{\Delta X,\Delta\epsilon}(N,n)$ as functions of N and n.

We consider $\widehat{F}^{\Delta\epsilon}(N,n)$ first. If $\widehat{F}^{\Delta\epsilon}(N,n)$ is divided by the sample size n, $n^{-1}\widehat{F}^{\Delta\epsilon}(N,n)$ estimates the spectral density of the MS noise at 0 frequency and hence $E(n^{-1}\widehat{F}^{\Delta\epsilon}(N,n)) = o(1)$ and $(\sum_k W(k)^2)^{-1} Var(n^{-1}\widehat{F}^{\Delta\epsilon}(N,n)) = o(1)$ if $N \to \infty$ and $N/n \to 0$ as $n \to \infty$.[4] We show below a refined result for $\widehat{F}^{\Delta\epsilon}(N,n)$.

Proposition 4.1. *Suppose that $N \to \infty$, $n\sum_k W(k)^2 \to \infty$, and $N^2/n \to 0$ as $n \to \infty$, and assumptions 2.1, 2.2 hold. Then*

$$\text{(13)}\quad Var(|J_n^{\Delta\epsilon}(\omega_s)|^2) = \sigma_\epsilon^4\left[2(\eta+1) + 2(\eta-1)n\omega_s^2 + O(N^4/n^2)\right],\quad 0 < s \le N,$$

$$\text{(14)}\quad Cov(|J_n^{\Delta\epsilon}(\omega_s)|^2, |J_n^{\Delta\epsilon}(\omega_l)|^2) = \sigma_\epsilon^4\left[2(\eta+1) - 2\omega_s^2 - 2\omega_l^2 + O(N^4/n^3)\right],\quad 0 \le s \ne l \le N.$$

As a result,

$$E\hat{F}^{\Delta\epsilon}(N,n) = 2\sigma_\epsilon^2[1 + 2^{-1}(n-1)\sum_{1\le k\le N} W(k)\omega_k^2] + O(N^4/n^3),$$
$$Var\hat{F}^{\Delta\epsilon}(N,n) = \sigma_\epsilon^4[2(\eta+1) + 4n\sum_{1\le k\le N} W(k)^2\omega_k^2] + o(N^4/n^2).$$

We turn to the cross periodogram next which characterizes the interaction between noise and the true price. Assume that

Assumption 4.1. *The volatility $\sigma^2(u)$ is stationary, and Π_n is an equally spaced partition of $[0,T]$.*

Proposition 4.2. *Suppose that $N \to \infty$, $n\sum_k W(k)^2 \to \infty$, and $N^2/n \to 0$ as $n \to \infty$, and assumptions 2.1, 2.2 and 4.1 hold. We have $E\hat{F}^{\Delta X,\Delta\epsilon}(N,n) = 0$, and*

$$Var\hat{F}^{\Delta X,\Delta\epsilon}(N,n) = 4\sigma_\epsilon^2\sigma_X^2\sum_k W(k)^2(1 + n\omega_k^2/2) + o(N^4/n^2),$$

where $\sigma_X^2 = E[X,X](T)$.

Last but not least, we look at $\hat{F}^{\Delta X}(N,n)$. Because $I_n^{\Delta X}(\omega_s) = n^{-1}\sum_{|h|<n}\gamma_X(h)e^{-i\omega_s h}$ where $\gamma_X(h) = \sum_{j=1}^{n-|h|}\Delta X_j \Delta X_{j+|h|}$, we have $E|J_n^{\Delta X}(\omega_s)|^2 = \sigma_X^2$. Let $E\Delta X_j^4 = \xi_n\sigma_X^4/n^2$. Note that ξ_n is $O(1)$ due to the Burkholder-Davis-Gundy Inequality. Then $Var(|J_n^{\Delta X}(\omega_s)|^2) = (n^{-1}(\xi_n - 3) + 1)\sigma_X^4$ for $0 < \omega_s < \pi$, and $Cov(|J_n^{\Delta X}(\omega_s)|^2, |J_n^{\Delta X}(\omega_l)|^2) = n^{-1}(\xi_n - 3)\sigma_X^4$ for $\omega_s \ne \omega_l$. Therefore

Proposition 4.3. *Under assumption 4.1, $E\widehat{F}^{\Delta X}(N,n) = \sigma_X^2$ and $Var\widehat{F}^{\Delta X}(N,n) = n^{-1}(\xi_n - 3)\sigma_X^4 + \sum_{k=1}^{N} W(k)^2\sigma_X^4$, where $\sigma_X^2 = E[X,X](T)$.*

By virtue of Propositions 4.1, 4.2 and 4.3, we have the following theorem which presents the main result.

Theorem 4.1. *Suppose that $N = o(n^{1/2})$, $\lim_{n,N\to\infty} n\sum_k W(k)^2 = \infty$, and Assumptions 2.1, 2.2 and 4.1 hold. We have*

$$\text{(15)}\quad \lim_{N,n\to\infty} E(\widehat{F}^{\Delta Y}(N,n) - [X,X](T)) = 2E(\epsilon_j^2),$$

and

$$\text{(16)}\quad \lim_{N,n\to\infty} E(\widehat{F}^{\Delta Y}(N,n) - [X,X](T))^2 = 2E(\epsilon_j^4) + 6(E(\epsilon_j^2))^2.$$

Remark 4.1. *In the presence of the MS noise, the MSE does not vanish when both n and N go to infinity. This is due to end-point effects, and it can be eliminated asymptotically by jittering, namely, average the first m and last m observations (see for instance [24] and [14]). The revised estimator is asymptotically unbiased when $m \to \infty$, and it is consistent for QV.*[5] *Yet, jittering will complicate the computation of the mean and variance of $\widehat{F}^{\Delta X}(N,n)$, $\widehat{F}^{\Delta\epsilon}(N,n)$, and $\widehat{F}^{\Delta X,\Delta\epsilon}(N,n)$ in finite samples. So we will not pursue it in this paper.*

Remark 4.2. *The asymptotic bias of $\widehat{F}^{\Delta Y}(N,n)$, $2E(\epsilon_j^2)$, does not hinge on the spectral window. This contradicts Theorem 2 of [28], which states that the Fourier estimator, in the presence of the MS noise, is asymptotically unbiased. This is due to an error in the proof of Theorem 2 of [28]: j' should be 2 on line 27 and line 28.*

The discussion so far assumes that the MS noise is uncorrelated with the true price process at all frequencies. This assumption is reasonable when intraday returns are sampled

[4] See for instance [15], [32], [20], [17], [16], [18].

[5] This is currently studied by [35].

every 15 ticks or so (see for instance [22]). We will next consider endogenous noise, namely, noise correlated with the true price. As in [22] and [11], we assume that the noise ϵ_j has the following decomposition

Assumption 4.2. *$\epsilon_j = \alpha\Delta X_j + \tilde{\epsilon}_j$ where $\tilde{\epsilon}_j$ is white noise with variance $\sigma^2_{\tilde{\epsilon}}$ and $E(\tilde{\epsilon}^4_j) = \eta\sigma^4_{\tilde{\epsilon}} < \infty$, and $\{\tilde{\epsilon}_j\}$ is uncorrelated with X at all lags and leads.*

The next theorem shows that the asymptotic bias (15) and the asymptotic MSE (16) are invariant of the interaction between noise and the true price process.

Theorem 4.2. *Theorem 4.1 still holds if Assumption 2.1 is replaced with Assumption 4.2.*

5. THE CHOICE OF CUT-OFF FREQUENCY IN A FINITE SAMPLE

A practical issue arises: how to choose cut-off frequency N for a given data. In this section, we discuss optimal cut-off frequency by virtue of MSE minimization. Because N should be less than $n/2$ to avoid aliasing, we define the optimal cut-off frequency as follows

$$N_{opt} = arg \min_{N \le n/2, N\in\mathbb{Z}} E(\widehat{F}^{\Delta Y}(N,n) - [X,X](T))^2. \tag{17}$$

Theorem 5.1 below provides an approximation to choose the optimal cut-off frequency N_{opt}.

Theorem 5.1. *Suppose that $N = o(n^{1/2})$, $\lim_{n,N\to\infty} n\sum_k W(k)^2 = \infty$, and Assumptions 4.2, 2.2 and 4.1 hold. To leading order,*

$$N_{opt} = arg \min_{0<N \le n/2, N\in\mathbb{Z}} \{n \sum_{1\le k\le N} W(k)\omega_k^2 + \sum_{1\le k\le N} W(k)^2 \left[\rho^2 + 2\rho + (1+\rho)n\omega_k^2\right]\}, \tag{18}$$

where $\omega_k = 2\pi k/n$, $\rho = (2\sigma^2_\epsilon)^{-1}\sigma^2_X$, and $\sigma^2_\epsilon = E(\epsilon^2_j)$, $\sigma^2_X = E\int_0^T \sigma(t)^2 dt + T\int_{\mathbb{R}} x^2\nu(dx)$.

In the absence of the noise, i.e., $\sigma^2_\epsilon = 0$, we have $N_{opt} = [n/2]$ the integer part of $n/2$. This is consistent with Theorem 3.1.

As an application of Theorem 5.1, we discuss two special cases when one can find N_{opt} in closed-form.

Define the rectangular window:

$$W(k) = 1/N \text{ for } |k| = 1,\dots,N \text{ and 0 otherwise.} \tag{19}$$

It satisfies Assumption 2.2. We have $\hat{F}^{\Delta Y}(N,n) = N^{-1}\sum_{k=1}^N |J_n^{\Delta Y}(\omega_k)|^2$. This is the Fourier estimator of [26]. By virtue of Theorem 5.1, the optimal cut-off frequency is $N_{opt} = \min([N_*],[n/2])$, where

$$N_* = -b + (-b^3 - d + \sqrt{d(d+2b^3)})^{1/3} + (-b^3 - d - \sqrt{d(d+2b^3)})^{1/3}, \tag{20}$$

and $b = 12^{-1}(5+2\rho)$, $d = -\frac{3n\rho(\rho+2)}{16\pi^2}$. See also [34].

Next we consider a triangular window which is defined in the following corollary.

Corollary 5.1. *Let $H(k) = (N-|k|)/N^2$ for $|k| = 0,1,\dots,N-1$. The proposed estimator, with $W(1) = H(0)$ and $W(k) = 2H(|k|-1)$ for $|k| = 2,\dots,N$ and 0 otherwise, is given by*

$$\widehat{F}^{\Delta Y}(N,n) = nN^{-1}I_n^{\Delta Y}(\omega_1) + 2n\sum_{1\le k\le N-1}(N-|k|)N^{-2}I_n^{\Delta Y}(\omega_{k+1}). \tag{21}$$

The optimal cut-off frequency is then $N_{opt} = \min([N_*],[n/2])$, where N_* has the form of (20) with $b = \frac{2(6+\rho)}{15}$, and $d = -\frac{n\rho(\rho+2)}{2\pi^2}$.

The following is a variant of the triangular window defined in Corollary 5.1.

Corollary 5.2. *With $W(k) = 2(N+1-|k|)(N(N+1))^{-1}$ for $|k| = 1,\dots,N$ and 0 otherwise, we have $\hat{F}^{\Delta Y}(N,n) = 2(N(N+1))^{-1}\sum_{1\le k\le N}(N+1-|k|)|J_n^{\Delta Y}(\omega_k)|^2$. The optimal cut-off frequency is $N_{opt} = \min([N_*],[n/2])$, where N_* approximately has the form of (20) with $b = \frac{19+4\rho}{30}$, and $d = -\frac{n\rho(\rho+2)}{2\pi^2}$.*

Theorem 5.1 is practically important. It provides a simple approximation to compute the optimal cut-off frequency for a given data set. It should be noted that the results hold when the volatility $\sigma^2(t)$ is stationary and the prices are sampled at regular intervals (i.e., Assumption 4.1). Though the stationarity of $\sigma^2(t)$ is required, it does not rule out conditional heteroskedasticity in the volatility process. Stationary volatility processes have been widely discussed in literature. Examples include, but are not limited to, the constant elasticity of variance (CEV) process, the non-Gaussian Ornstein-Uhlenbeck (OU) process, and the sum or superpositions of independent CEV or OU processes. See also the eigenfunction stochastic volatility(SV) models of [6]. It has been shown that these SV models are empirically reasonable. See [13], [8], [19], [30], [6], [33], among others.

Theorem 5.1 applies to intraday returns that are sampled at evenly spaced intervals, i.e., calendar time sampling. Calendar time sampling is common for exchange rates data, for instance, the data from Olsen & Associates. When dealing with trade and quote data, it is more natural to sample prices in tick time (number of trades between observations). We will consider irregularly spaced data in a simulation study, and assess numerically how sampling frequency affects the optimal cut-off frequency given in equation (18).

6. SIMULATION STUDY

In this section a simulation experiment is performed to evaluate the accuracy of $\widehat{F}^{\Delta Y}(N,n)$ as an estimator of QV in finite samples, especially at $N = N_{opt}$.

Two data generating processes are considered. The first is a GARCH diffusion process

$$\begin{aligned} dX(t) &= \sigma(t)dB_1(t), \\ d\sigma(t)^2 &= \theta(\varpi - \sigma(t)^2)dt + \sqrt{2\lambda\theta}\sigma(t)^2 dB_2(t), \end{aligned} \tag{22}$$

where $B_1(t)$ and $B_2(t)$ are Brownian motions, with the leverage correlation $corr(dB_1, dB_2) = \varphi$. We use the parameter setting in [5] and [23]: $\theta = 0.0350$, $\varpi = 0.6365$, $\lambda = 0.2962$, $\varphi = -0.62$. The second model is a stochastic volatility model with jumps (SVJ)

$$\begin{aligned} dX(t) &= \sigma(t)dB_1(t) + J(t)dN(t), \\ d\sigma(t)^2 &= \theta(\varpi - \sigma(t)^2)dt + \lambda\sigma(t)dB_2(t), \end{aligned} \tag{23}$$

where $N(t)$ is a compound Poisson process with intensity λ_J. We use values reported in [9]: $\theta = 0.01$, $\varpi = 0.5$, $\lambda = 0.05$, $\lambda_J = 2$ and the jump size is Gaussian with mean 0 and variance 0.64ϖ, and $corr(dB_1, dB_2) = -0.62$.

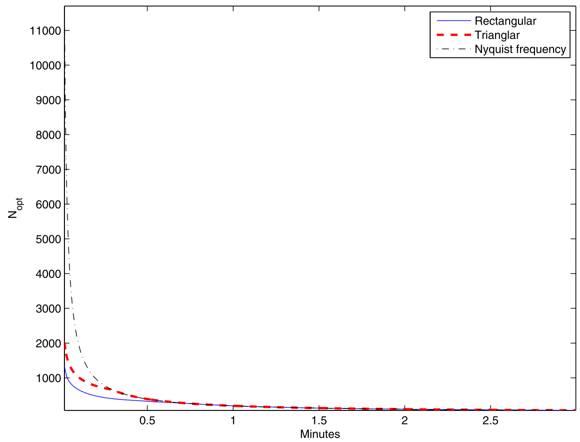
The true log price process is generated via the Euler discretization: take one day as a reference measure, and simulate 6.5 hours of trading with $dt = 1/23{,}400$, i.e., a total of 23,400 second-by-second returns per day. We simulate 500 daily replications, with $X(0) = \log(100)$ and $\sigma(0)^2 = \varpi$. The noise ϵ_t is specified as Gaussian with mean 0 and variance σ_ϵ^2. Two variance levels are considered: $\sigma_\epsilon^2 = 0.000142$ and 0.00142.

Denote by Nopt_R and Nopt_T the optimal cut-off frequencies for the rectangular window and the triangular window respectively, which are calculated from (20) and Corollary 5.1. Figure 2 displays Nopt_R and Nopt_T as functions of sampling interval ranging from 1 second to 3 minutes for $\sigma_\epsilon^2 = 0.000142$, and from 20 seconds to 9 minutes for $\sigma_\epsilon^2 = 0.00142$, where $X(t)$ follows GARCH diffusion (22). The dashed curve corresponds to the Nyquist frequency $n/2$. When the sampling interval is smaller than 30 seconds for $\sigma_\epsilon^2 = 0.000142$ and 3 minutes for $\sigma_\epsilon^2 = 0.00142$, Nopt_R and Nopt_T are much smaller than the Nyquist frequency. This observation shows that the effect of the MS noise is pronounced at ultra high frequencies. The plots for SVJ model (23) are similar and hence are omitted.

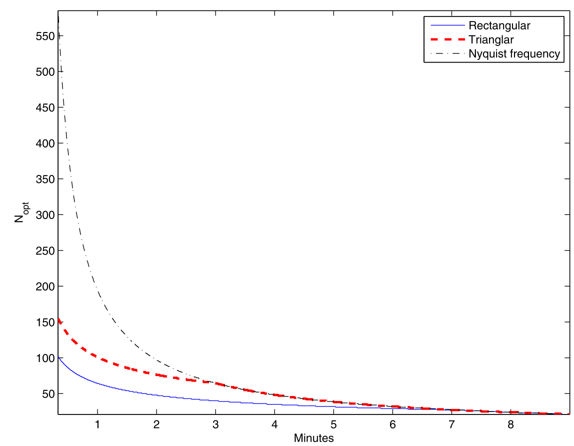
We then compute the periodogram-based estimators with rectangular and triangular windows. For comparison purposes, we also use the Realized Kernel to estimate daily QV. As suggested in [14], we compute the Realized Kernel with the Parzen weight

$$\Lambda(x) = \begin{cases} 1 - 6x^2 + 6x^3 & \text{if } 0 \le x \le 1/2, \\ 2(1-x)^3 & \text{if } 1/2 \le x \le 1, \\ 0 & \text{if } x > 1. \end{cases} \tag{24}$$

The preferred bandwidth is $H^* = 3.51\xi^{4/5}n^{3/5}$ with $\xi^2 = \frac{\sigma_\epsilon^2}{\sqrt{\int_0^1 \sigma(u)^4 du}}$. Prior to computing the realized kernel, we average the first two prices and the last two prices to remove end-effects as suggested in [14]. Since the MS effect is pronounced at ultra high frequencies, we focus on prices at frequencies of 1 second, 5 seconds, 10 seconds, 20 seconds and 30 seconds. Meanwhile, we also consider returns that are sampled at randomly spaced time points. Noisy prices are picked up at times $\{\tilde{t}_j\}$ where $\tilde{t}_j - \tilde{t}_{j-1}$ is Exponential distributed with mean duration τ. We consider the following values of τ: 5 seconds, 10 seconds, 20 seconds and 30 seconds.



(a) σ_ϵ^2 =0.000142



(b) σ_ϵ^2 =0.00142

Figure 2. N_{opt} versus sampling interval, GARCH diffusion (22).

Tables 1 and 2 report mean squared errors for regularly-sampled and irregularly-sampled prices, respectively. For the periodogram-based estimators, we report not only the mean squared error at the optimal cut-off frequency – see the number in the parentheses, but also the minimum mean squared error (min MSE) with the associated cut-off frequency – see the number in the parentheses. Evidently, the approximation in equation (18) offers a good representation of the optimal cutting frequency, and it is not sensitive to the sampling frequency. Moreover, the periodogram-based estima-

Table 1. Mean squared error, evenly sampled data

	n	Triangular window		Rectangular window		Realized Kernel
		min MSE	MSE at N_{opt}	min MSE	MSE at N_{opt}	MSE
			$\sigma^2_\epsilon = 0.000142$, GARCH diffusion (22)			
1 sec	23400	4.25 (1528)	4.82 (2017)	4.45 (1169)	4.83 (1343)	9.01
5 sec	4680	7.69 (860)	8.70 (1091)	7.90 (577)	8.77 (699)	18.96
10 sec	2340	10.19 (671)	10.95 (828)	10.53 (439)	11.26 (520)	25.07
20 sec	1170	12.76 (585)	12.76 (585)	13.49 (386)	13.58 (384)	32.94
30 sec	780	15.99 (390)	15.99 (390)	14.39 (389)	15.99 (320)	38.71
			$\sigma^2_\epsilon = 0.00142$, GARCH diffusion (22)			
1 sec	23400	10.17 (645)	12.01 (463)	10.38 (408)	12.97 (321)	25.00
5 sec	4680	21.80 (343)	24.03 (260)	22.76 (231)	27.93 (176)	47.10
10 sec	2340	28.48 (260)	31.17 (201)	30.97 (186)	34.33 (134)	62.52
20 sec	1170	38.38 (196)	41.14 (155)	40.73 (142)	44.88 (102)	84.21
30 sec	780	43.71 (171)	47.63 (133)	47.93 (106)	52.49 (86)	97.12
			$\sigma^2_\epsilon = 0.000142$, SVJ model (23)			
1 sec	23400	11.59 (1960)	14.46 (2799)	11.63 (1105)	13.96 (1843)	23.71
5 sec	4680	17.01 (978)	22.61 (1498)	17.25 (707)	20.34 (945)	46.00
10 sec	2340	21.56 (846)	24.23 (1130)	23.71 (675)	24.46 (699)	60.93
20 sec	1170	26.83 (585)	26.83 (585)	27.29 (577)	29.82 (513)	80.08
30 sec	780	35.67 (390)	35.67 (390)	31.40 (373)	32.38 (390)	96.87
			$\sigma^2_\epsilon = 0.00142$, SVJ model (23)			
1 sec	23400	22.65 (778)	23.45 (675)	23.14 (543)	25.28 (463)	66.75
5 sec	4680	43.19 (384)	43.20 (376)	46.36 (322)	52.21 (251)	163.29
10 sec	2340	57.68 (302)	57.91 (290)	62.35 (199)	65.96 (191)	229.08
20 sec	1170	81.86 (249)	83.89 (222)	89.48 (163)	98.31 (143)	302.53
30 sec	780	92.41 (227)	98.12 (190)	99.76 (160)	113.88 (121)	352.59

Note: The mean squared errors have been multiplied by 10^4.

Table 2. Mean squared error, unevenly sampled data

τ	Triangular window		Rectangular window		Realized Kernel
	min MSE	MSE at N_{opt}	min MSE	MSE at N_{opt}	MSE
		$\sigma^2_\epsilon = 0.000142$, GARCH diffusion (22)			
5 sec	9.40 (834)	10.40 (1095)	9.39 (583)	10.46 (701)	20.92
10 sec	13.11 (666)	14.04 (828)	13.33 (452)	13.69 (520)	28.24
20 sec	18.04 (614)	18.04 (614)	18.79 (389)	19.01 (392)	38.78
30 sec	22.87 (402)	22.87 (402)	21.65 (379)	22.65 (325)	47.37
		$\sigma^2_\epsilon = 0.00142$, GARCH diffusion (22)			
5 sec	23.97 (335)	25.73 (260)	23.76 (235)	26.72 (176)	50.34
10 sec	34.16 (255)	36.34 (200)	36.89 (163)	40.23 (133)	63.27
20 sec	44.01 (198)	47.67 (155)	45.76 (152)	52.02 (102)	89.83
30 sec	54.47 (176)	59.35 (131)	57.91 (115)	66.49 (85)	113.29
		$\sigma^2_\epsilon = 0.000142$, SVJ (23)			
5 sec	20.93 (883)	27.13 (1511)	21.23 (637)	27.07 (954)	45.84
10 sec	28.03 (858)	30.45 (1140)	30.29 (424)	31.03 (706)	71.77
20 sec	43.23 (610)	43.23 (610)	41.48 (514)	42.76 (523)	104.94
30 sec	70.64 (389)	70.64 (389)	66.06 (384)	67.42 (389)	113.59
		$\sigma^2_\epsilon = 0.00142$, SVJ (23)			
5 sec	46.12 (386)	46.32 (373)	56.08 (302)	59.59 (249)	257.97
10 sec	67.69 (323)	68.72 (294)	74.85 (192)	75.12 (194)	312.55
20 sec	85.97 (250)	88.37 (219)	94.44 (168)	107.37 (141)	409.86
30 sec	110.59 (184)	110.63 (187)	115.59 (108)	121.02 (120)	555.99

Note: The mean squared errors have been multiplied by 10^4.

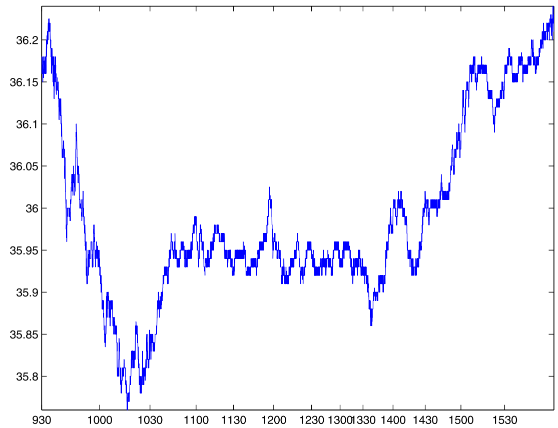


Figure 3. GE transaction prices on Feb 1, 2007.

tors are competitive with the Realized Kernel. When the sample path does not feature jumps, the triangular window is preferred over the rectangular window. When jumps are present, the rectangular window achieves a smaller mean squared error for $\sigma_\epsilon^2 = 0.000142$. The simulation results also suggest using returns sampled at the highest possible frequency to construct the periodogram-based estimators.

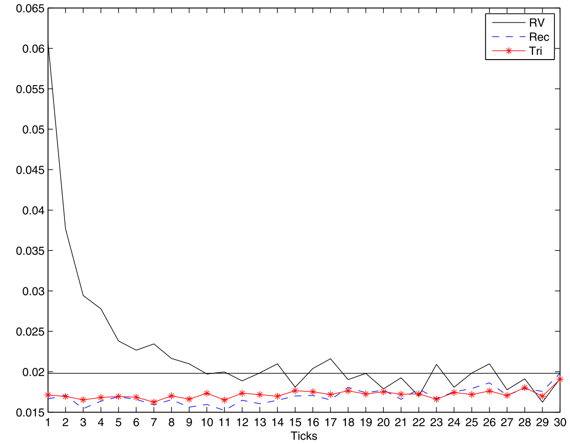
7. REAL DATA ANALYSIS

In this section, we consider an application on the transaction prices of General Electric on Feb 1, 2007. The intraday data are extracted from the Wharton Research Data Services, TAQ database. We focus on the prices from the New York Stock Exchange. Transaction prices before 9:30 am and after 4:00 pm were removed. We deleted entries with zero prices. For multiple transactions with the same time stamp, we used the median price. We also filtered out prices that are more than one spread away from the bid and ask quotes. There are 8,052 observations left in the sample, so the prices are recorded every 3 seconds on average. The price series are plotted in Figure 3.

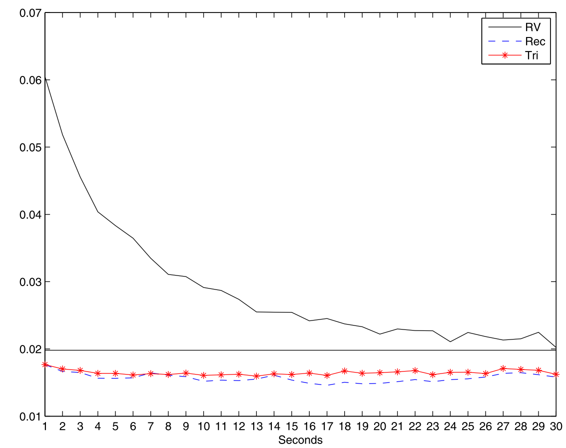
As suggested by the simulation study, we will use all the available data to construct the periodogram-based estimator. To compute the optimal cut-off frequency, we need to estimate σ_X^2 and σ_ϵ^2. Denote by RV_{sparse} the subsampled RV based on 20-minute returns, which is the average of 1,200 RV's by shifting the time of the first observation in 1 second increment. Define

$$\hat{\sigma}_e^2 = q^{-1} \sum_{j=1}^{q} (2n_{(j)})^{-1} RV^{(j)},$$

where $RV^{(1)}, RV^{(2)}, \ldots, RV^{(q)}$ are q distinct realized variances based on every qth trade by varying the starting point, and $n_{(j)}$ is the number of non-zero returns in the calculation of $RV^{(j)}$. We use RV_{sparse} and $\hat{\sigma}_e^2$ as estimators of σ_X^2 and σ_ϵ^2 respectively, with $q = 50$ (see [12] for details). We obtained $\hat{\sigma}_e^2 = 0.0065\%$ and $\hat{\sigma}_X^2 = 1.98\%$. The optimal cut-off frequencies are 169 for the rectangular window and 247 for the triangular window. The periodogram-based estimators at the optimal cut-off frequency are 1.66% and 1.71% for the rectangular window and the triangular window respectively.



(a) Tick time sampling



(b) Calendar time sampling

Figure 4. Volatility Signature Plot.

Figure 4 presents volatility signature plot for the realized variance, periodogram-based estimators using rectangular window and triangular window. We consider both tick time sampling and calendar time sampling. The horizontal line is the subsampled 20-minute RV. The signature plots for the periodogram-based estimators are almost level, especially for the triangular window, which shows that the periodogram-based estimators are robust to market microstructure noise.

8. CONCLUSION

In this paper, we proposed a new frequency-domain estimator of quadratic variation. It is based on the weighted average of periodograms at non-zero Fourier frequencies. It is nonparametric in nature. The performance of the estimator depends on both sample size n and cut-off frequency N. We provide a simple approximation to calculate the optimal cut-off frequency for a given data set. This work fills the gap between the time-domain and the frequency-domain analysis of the ex-post measure of price variation. The appeal of the proposed estimator is its flexibility and simplicity. It is easy to implement and is computationally efficient as well. We show, through a simulation study, that the proposed estimators with rectangular window and triangular window are competitive with the realized kernel of [14], and they are insensitive to the choice of the sampling frequency.

APPENDIX A. TECHNICAL APPENDICES

A.1 Proof of Proposition 3.1

Let $A(t) = \int_0^t \mu(s)ds$. Note that $\mathcal{F}(dX)(k) = \mathcal{F}(dA)(k) + \mathcal{F}(dX^c)(k)$. We obtained

$$\begin{aligned} &|\mathcal{F}(dX)(k)|^2 \\ &= |\mathcal{F}(dA)(k)|^2 + |\mathcal{F}(dX^c)(k)|^2 \\ &\quad + \mathcal{F}(dA)(k)\mathcal{F}(dX^c)(-k) + \mathcal{F}(dA)(-k)\mathcal{F}(dX^c)(k). \end{aligned}$$

It suffices to show that

$$\lim_{N\to\infty} \sum_{1\le k\le N} W(k)|\mathcal{F}(dA)(k)|^2 = 0 \quad \text{in } L^2. \tag{A.1}$$

Note that $\sum_{1\le k\le N} W(k)|\mathcal{F}(dA)(k)|^2 = T^{-2}\int_0^T b(s)\mu(s)ds$ where $b(s) = \int_0^T D_N((s-t)/T)\mu(t)dt$, and $D_N(v) = \sum_{1\le k\le N} W(k)e^{2\pi ikv}$. Because

$$\begin{aligned} &E\left(\int_0^T b(s)\mu(s)ds\right)^2 \\ &\le E\int_0^T b(s)^2 ds E\int_0^T \mu(s)^2 ds \\ &\le T^2\int_0^1 |D_N(u)|^2 du \left(E\int_0^T \mu(s)^2 ds\right)^2, \end{aligned}$$

and $\int_0^1 D_N(v)e^{-2\pi ikv}dv = W(k)$ for $1 \le k \le N$ and 0 otherwise, it follows from the Parseval's theorem that $\int_0^1 |D_N(v)|^2 dv = \sum_{1\le k\le N} W(k)^2$ and hence (A.1) holds.

A.2 Proof of Proposition 3.2

Note that $|\mathcal{F}(dX^c)(k)|^2 = \mathcal{F}(dX^c)(k)\mathcal{F}(dX^c)(-k)$. It follows from the Ito's lemma that

$$\begin{aligned} &T^2|\mathcal{F}(dX^c)(k)|^2 \\ &= [X,X](T) + \int_0^T I_k(h-)dI_{-k}(h) + \int_0^T I_{-k}(h-)dI_k(h), \end{aligned}$$

where $I_k(h) = \int_0^h e^{-2\pi ikt/T}dX^c(t)$. Let $R(k) = \int_0^T I_k(h-)dI_{-k}(h) + \int_0^T I_{-k}(h-)dI_k(h)$. Next we will show that

$$\lim_{N\to\infty} \sum_{1\le k\le N} W(k)R(k) = 0 \quad \text{in } L^2. \tag{A.2}$$

It suffices to prove that $(E|A_N|^2)^2 \le C\int_0^1 |D_N(v)|^4 dv$ for some constant $C > 0$, where $A_N = \int_0^T\int_0^{h-} D_N((h-t)/T)dX^c(t)dX^c(h)$ and $D_N(v) = \sum_{1\le k\le N} W(k)e^{2\pi ikv}$.

Let $Z(t) = \int_0^t \sigma(s)dB(s)$ and $G(h) = \int_0^{h-} D_N((h-t)/T)dX^c(t)$. Then

$$A_N = \underbrace{\int_0^T G(h)dZ(h)}_{\mathbb{T}_1} + \underbrace{\int_0^T\int_{-\infty}^{\infty} G(h)x\tilde{\mathcal{J}}_X(dh\times dx)}_{\mathbb{T}_2}. \tag{A.3}$$

Note that

$$\begin{aligned} E|\mathbb{T}_1|^2 &= E\int_0^T |G(h)|^2\sigma(h)^2 dh \\ &\le \sqrt{E\int_0^T |G(h)|^4 dh}\sqrt{E\int_0^T \sigma(h)^4 dh}, \end{aligned} \tag{A.4}$$

$$\begin{aligned} E|\mathbb{T}_2|^2 &= E\int_0^T\int_{-\infty}^{\infty} |G(h)|^2 x^2 dh\nu(dx) \\ &= E\int_0^T |G(h)|^2 dh\int_{-\infty}^{\infty} x^2\nu(dx). \end{aligned} \tag{A.5}$$

Let $G^c(h) = \int_0^h D_N((h-t)/T)dZ(t)$, and $G^J(h) = \int_0^{h-}\int_{-\infty}^{\infty} D_N((h-t)/T)x\tilde{\mathcal{J}}_X(dt\times dx)$. Then

$$E\int_0^T |G(h)|^4 dh \le 27\left(E\int_0^T |G^c(h)|^4 dh + E\int_0^T |G^J(h)|^4 dh\right).$$

It follows from the Burkholder-Davis-Gundy inequality that there exists a constant $C_1 > 0$ such that

$$\begin{aligned} &E\left(\int_0^T |G^c(h)|^4 dh\right) \\ &\le C_1 E\int_0^T\int_0^h |D_N((h-t)/T)|^4\sigma(t)^4 dtdh \\ &= C_1\int_0^T |D_N(u/T)|^4 du E\int_0^T \sigma(t)^4 dt, \end{aligned} \tag{A.6}$$

and

$$\begin{aligned} &E|G^J(h)|^4 \\ &\le C_1 E\left(\int_0^{h-}\int_{-\infty}^{\infty} D_N((h-t)/T)^2 x^2 J_X(dt\times dx)\right)^2 \\ &\le C_1 T\int_0^h D_N((h-t)/T)^4 dt\left(\int_{-\infty}^{\infty} x^2\nu(dx)\right)^2. \end{aligned}$$

Therefore,

$$\begin{aligned} &E\int_0^T |G^J(h)|^4 dh \\ &\le C_1 T^2\int_0^T D_N(u/T)^4 du\left(\int_{-\infty}^{\infty} x^2\nu(dx)\right)^2. \end{aligned} \tag{A.7}$$

In view of (A.6) and (A.7), we have

$$(E|A_N|^2)^2 \le C \int_0^1 |D_N(v)|^4\, dv$$

for some constant $C > 0$.

A.3 Proof of Theorem 3.1

It suffices to prove that

$$\text{(A.8)} \qquad \left\| |J_n^X(\omega_k)|^2 - T^2 |\mathcal{F}(dX)(k)|^2 \right\|_2 \le Ck\rho(n),$$

for some $C > 0$.

It follows from Lemma 3.1 and Proposition 3.2 that $E|\mathcal{F}(dX)(k)|^4 < \infty$. Let $H(t) = \sum_{j=1}^n e^{-i\omega j} 1_{(t_{j-1},t_j]}(t)$. Then

$$\text{(A.9)} \qquad J_n^X(\omega) = \underbrace{\int_0^T H(t)\mu(t)dt}_{\mathbb{T}_0} + \underbrace{\int_0^T H(t)\sigma(t)dB(t)}_{\mathbb{T}_1} + \underbrace{\int_0^T \int_{-\infty}^{\infty} H(t) x \tilde{\mathcal{J}}_X(dt \times dx)}_{\mathbb{T}_2}.$$

Since $|H(t)| \le 1$, there exists a constant $C_1 > 0$ such that $E|\mathbb{T}_0|^4 \le C_1 (E\int_0^T \mu(s)^2 ds)^2$, $E|\mathbb{T}_1|^4 \le C_1 E\int_0^T \sigma(t)^4 dt$, $E|\mathbb{T}_2|^4 \le C_1 (\int_{-\infty}^{\infty} x^2 \nu(dx))^2$ due to the Burkholder-Davis-Gundy inequality. Hence $E|J_n^X(\omega)|^4$ is bounded uniformly in n.

Let $M(t) = \sum_{j=1}^n (e^{-2\pi ijk/n} - e^{-2\pi ikt/T}) 1_{(t_{j-1},t_j]}(t)$. We have

$$\begin{aligned} &\left\| |J_n^X(\omega_k)|^2 - T^2 |\mathcal{F}(dX)(k)|^2 \right\|_2 \\ &\le \left(\|J_n^X(\omega_k)\|_4 + \|T\mathcal{F}(dX)(k)\|_4 \right) \|\mathbb{T}\|_4, \end{aligned}$$

where $\mathbb{T} = \int_0^T M(t) dX(t)$. Decompose $\mathbb{T}$ as $\mathbb{T} = \mathbb{T}_0 + \mathbb{T}_1 + \mathbb{T}_2$, where $\mathbb{T}_k$ is defined in (A.9) with $H(\cdot)$ replaced by $M(\cdot)$. Note that $|M(t)| \le 2\pi k \rho(n)/T$. For some constant $C_2 > 0$,

$$\text{(A.10)} \qquad \begin{aligned} E|\mathbb{T}_0|^4 &\le C_2 k^4 \rho(n)^4 \left(E \int_0^T \mu(s)^2 ds \right)^2, \\ E|\mathbb{T}_1|^4 &\le C_2 k^4 \rho(n)^4 (E \int_0^T \sigma(t)^2 dt)^2, \\ E|\mathbb{T}_2|^4 &\le C_2 k^4 \rho(n)^4 (\int_{-\infty}^{\infty} x^2 \nu(dx))^2, \end{aligned}$$

where (A.10) is due to equation (35) of [27]. Therefore there exists a constant $C_3 > 0$ which is independent of n such that $\|\mathbb{T}\|_4 \le C_3 k \rho(n)$. The proof is complete.

A.4 Proof of Equality (11)

By definition,

$$\begin{aligned} &(\mathcal{F}(dX) \star_B \mathcal{F}(dX))(k) \\ &= \lim_{N\to\infty} \frac{1}{2N+1} \sum_{s=-N}^{N} \mathcal{F}(dX)(s)\mathcal{F}(dX)(k-s). \end{aligned}$$

Let $I_k(h) = \int_0^h e^{-2\pi ikt/T} dX(t)$. Because $T^2 \mathcal{F}(dX)(s_1)\mathcal{F}(dX)(s_2) = \int_0^T e^{-2\pi i(s_1+s_2)t/T} d[X,X](t) + \int_0^T I_{s_1}(h-) dI_{s_2}(h) + \int_0^T I_{s_2}(h-) dI_{s_1}(h)$ due to Ito's Lemma, we have

$$\begin{aligned} &\frac{1}{2N+1} \sum_{s=-N}^{N} \mathcal{F}(dX)(s)\mathcal{F}(dX)(k-s) \\ &= \frac{1}{T} \mathcal{F}(d[X,X])(k) \\ &\quad + \frac{1}{(2N+1)T^2} \sum_{s=-N}^{N} \Big(\int_0^T I_s(h-) dI_{k-s}(h) \\ &\quad + \int_0^T I_{k-s}(h-) dI_s(h) \Big). \end{aligned}$$

We need to show that $\frac{1}{2N+1} \sum_{s=-N}^{N} (\int_0^T I_s(h-) dI_{k-s}(h) + \int_0^T I_{k-s}(h-) dI_s(h)) = o_p(1)$ for any $k \in \mathbb{Z}$. It suffices to prove

$$\text{(A.11)} \qquad \frac{1}{2N+1} \sum_{s=-N}^{N} (\int_0^T I_s(h-) dI_{k-s}(h)) = o_p(1).$$

Note that

$$\begin{aligned} &\frac{1}{2N+1} \sum_{s=-N}^{N} \left(\int_0^T I_s(h-) dI_{k-s}(h) \right) \\ &= \int_0^T \int_0^{h-} D_N((h-t)/T) dX(t) dX(h), \end{aligned}$$

where $D_N(v) = e^{-2\pi ikh/T} \frac{1}{2N+1} \sum_{s=-N}^{N} e^{2\pi isv}$. Similar to the proofs of Lemma 3.1 and Proposition 3.2, we have

$$\begin{aligned} &\left(E \left| \int_0^T \int_0^{h-} D_N((h-t)/T) dX(t) dX(h) \right|^2 \right)^2 \\ &\le C \int_0^1 |D_N(v)|^4\, dv, \end{aligned}$$

for some constant $C > 0$. Since $\int_0^1 |D_N(v)|^4 dv \le \int_0^1 |D_N(v)|^2 dv = (2N+1)^{-1}$, (A.11) is true and this completes the proof.

A.5 Proof of Proposition 4.1

Equalities (13) and (14) follow from Lemma A.1. Therefore

$$\begin{aligned} &Var \hat{F}^{\Delta\epsilon}(N,n) \\ &= \sum_k W(k)^2 Var|J_n^{\Delta\epsilon}(\omega_k)|^2 \\ &\quad + \sum_{k\ne k'} W(k)W(k') Cov(|J_n^{\Delta\epsilon}(\omega_k)|^2, |J_n^{\Delta\epsilon}(\omega_{k'})|^2) \\ &= \sigma_\epsilon^4 [2(\eta-1) + 4(\eta-3)(n-1)(1 - \sum_k W(k)\cos\omega_k)^2 \\ &\quad + 4(\sum_k W(k)\cos\omega_k)^2 \\ &\quad + 4n \sum_k W(k)^2 (1-\cos\omega_k)(n-(n-2)\cos\omega_k)] \end{aligned}$$

$$= \sigma_\epsilon^4[2(\eta+1) + 4n\sum_k W(k)^2\omega_k^2] + o(N^4/n^2).$$

Note that $I_n^{\Delta\epsilon}(\omega) = n^{-1}\sum_{|h|<n}\gamma_\epsilon(h)e^{-i\omega h}$ where $\gamma_\epsilon(h) = \sum_{j=1}^{n-|h|}\Delta\epsilon_j\Delta\epsilon_{j+|h|}$. We have

$$\begin{aligned} &E\hat{F}^{\Delta\epsilon}(N,n) \\ &= 2\sigma_\epsilon^2[n-(n-1)\sum_{1\le k\le N}W(k)\cos\omega_k] \\ &= 2\sigma_\epsilon^2[1+(n-1)\sum_{1\le k\le N}W(k)\omega_k^2/2] + O(N^4/n^3). \end{aligned}$$

The proof of Proposition 4.1 needs the following lemma.

Lemma A.1. *Consider an MA(1) process $P_j = (1-\theta B)e_j$ where B is the backshift operator and e_t is white noise with variance σ_e^2 and $Ee_t^4 = \eta_e\sigma_e^4 < \infty$. Then for $0 < \omega_s = 2\pi s/n < \pi$,*

$$\begin{aligned} &Var(|J_n^P(\omega_s)|^2) \\ &= \sigma_e^4(\eta_e-3)(1+\theta^4+(n-1)(1+\theta^2-2\theta\cos\omega_s)^2) \\ &\quad + 4\theta^2\sigma_e^4 + \sigma_e^4((1+\theta^2)n - 2\theta(n-1)\cos\omega_s)^2, \end{aligned}$$

and for $0\le\omega_s\ne\omega_l\le\pi$,

$$\begin{aligned} &Cov(|J_n^P(\omega_s)|^2, |J_n^P(\omega_l)|^2) \\ &= \sigma_e^4(\eta_e-3)(n-1)(1+\theta^2-2\theta\cos\omega_s)(1+\theta^2-2\theta\cos\omega_l) \\ &\quad + \sigma_e^4(\eta_e-3)[1+\theta^4] + 4\theta^2\sigma_e^4(1+\cos\omega_s\cos\omega_l). \end{aligned}$$

Proof of Lemma A.1: Let $\omega = \omega_j$, $\lambda = \omega_k$. According to the proof of Theorem 10.3.2 of [18],

$$\begin{aligned} &Cov(I^P(\omega), I^P(\lambda)) \\ &= n^{-2}(\eta_e-3)\sigma_e^4\sum_{s,t,u,v=1}^n \psi^*_{s,t,u,v}e^{-i\omega(t-s)}e^{-i\lambda(v-u)} \\ &\quad + \left(n^{-1}\sum_{s,u=1}^n\gamma(u-s)e^{-i\omega s-i\lambda u}\right)\left(n^{-1}\sum_{t,v=1}^n\gamma(v-t)e^{i\omega t+i\lambda v}\right) \\ &\quad + \left(n^{-1}\sum_{s,v=1}^n\gamma(v-s)e^{-i\omega s+i\lambda v}\right)\left(n^{-1}\sum_{t,u=1}^n\gamma(u-t)e^{i\omega t-i\lambda u}\right) \end{aligned}$$

where $\psi^*_{s,t,u,v} = \psi_{t-s}\psi_{u-s}\psi_{v-s} + \psi_1\psi_{t-s+1}\psi_{u-s+1}\psi_{v-s+1}$, and $\psi_0 = 1$, $\psi_1 = -\theta$ and $\psi_k = 0$ for $k\ne 0,1$; and $\gamma(h) = (1+\theta^2)\sigma_e^2$ if $h=0$, $-\theta\sigma_e^2$ if $h=\pm1$ and 0 for the others.

Note that for $0\le\omega,\lambda\le\pi$,

$$\begin{aligned} &\sum_{s,t,u,v=1}^n \psi_{t-s}\psi_{u-s}\psi_{v-s}e^{-i\omega(t-s)}e^{-i\lambda(v-u)} \\ &= \sum_{s=1}^n(\sum_{t=1-s}^{n-s}\psi_t e^{-i\omega t})(\sum_{v=1-s}^{n-s}\psi_v e^{-i\lambda v})(\sum_{u=1-s}^{n-s}\psi_u e^{i\lambda u}) \\ &= 1 + (n-1)(1-\theta e^{-i\omega})|1-\theta e^{-i\lambda}|^2, \end{aligned}$$

and

$$\begin{aligned} &\sum_{s,t,u,v=1}^n \psi_{t-s+1}\psi_{u-s+1}\psi_{v-s+1}e^{-i\omega(t-s)}e^{-i\lambda(v-u)} \\ &= -\theta^3 + (n-1)e^{i\omega}(1-\theta e^{-i\omega})|1-\theta e^{-i\lambda}|^2. \end{aligned}$$

Moreover,

$$\text{(A.12)}\quad \begin{aligned} &\sum_{s,u=1}^n\gamma(u-s)e^{-i\omega s-i\lambda u} \\ &= \begin{cases} \theta\sigma_e^2(e^{-i\omega}+e^{-i\lambda}) & \text{for } 0<\omega,\lambda<\pi, \\ (n+n\theta^2-2(n-1)\theta)\sigma_e^2 & \text{for } \omega=\lambda=0, \\ (n+n\theta^2+2(n-1)\theta)\sigma_e^2 & \text{for } \omega=\lambda=\pi, \end{cases} \end{aligned}$$

and

$$\text{(A.13)}\quad \begin{aligned} &\sum_{s,v=1}^n\gamma(v-s)e^{-i\omega s+i\lambda v} \\ &= \begin{cases} \theta\sigma_e^2(e^{-i\omega}+e^{i\lambda}) & \text{for } 0\le\omega\ne\lambda\le\pi, \\ \sigma_e^2((1+\theta^2)n-2\theta(n-1)\cos\lambda) & \text{for } 0\le\omega=\lambda\le\pi. \end{cases} \end{aligned}$$

As a result, for $0\le\omega\ne\lambda\le\pi$,

$$\begin{aligned} &Cov(nI^P(\omega), nI^P(\lambda)) \\ &= \sigma_e^4(\xi_n-3)(1+\theta^4+(n-1)(1+\theta^2-2\theta\cos\omega) \\ &\quad \times(1+\theta^2-2\theta\cos\lambda)) \\ &\quad + \theta^2\sigma_e^4[4+4\cos\omega\cos\lambda], \end{aligned}$$

and for $0<\omega<\pi$

$$\begin{aligned} &Var(nI^P(\omega)) \\ &= \sigma_e^4(\xi_n-3)(1+\theta^4+(n-1)(1+\theta^2-2\theta\cos\omega)^2) \\ &\quad + 4\theta^2\sigma_e^4 + \sigma_e^4((1+\theta^2)n-2\theta(n-1)\cos\omega)^2. \end{aligned}$$

A.6 Proof of Proposition 4.2

By the definition of $\hat{F}^{\Delta X,\Delta\epsilon}(N,n)$, $E\hat{F}^{\Delta X,\Delta\epsilon}(N,n) = 0$.

Let $\hat{F}_1^{X,\epsilon}(N,n) = \sum_{k=1}^N W(k)J_n^{\Delta X}(\omega_k)J_n^{\Delta\epsilon}(-\omega_k)$. We have

$$\begin{aligned} &Var\left(\hat{F}_1^{X,\epsilon}(N,n)\right) = E\left|\hat{F}_1^{X,\epsilon}(N,n)\right|^2 \\ &= \sum_k W(k)^2E|J_n^{\Delta X}(\omega_k)|^2E|J_n^{\Delta\epsilon}(\omega_k)|^2 \\ &\quad + \sum_{k\ne k'}W(k)W(k')E\left[J_n^{\Delta X}(\omega_k)J_n^{\Delta X}(-\omega_{k'})\right] \\ &\quad \times E\left[J_n^{\Delta\epsilon}(-\omega_k)J_n^{\Delta\epsilon}(\omega_{k'})\right]. \end{aligned}$$

Note that $E(J_n^{\Delta\epsilon}(\omega_s)J_n^{\Delta\epsilon}(\omega_k)) = \sum_{j,l=1}^n e^{-i\omega_s j-i\omega_k l}\gamma(j-l)$. In view of (A.12) and (A.13), we have

$$\begin{aligned} &E(J_n^{\Delta\epsilon}(\omega_s)J_n^{\Delta\epsilon}(\omega_k)) \\ &= \sigma_\epsilon^2(e^{-i\omega_s}+e^{-i\omega_k}), \quad \text{for } 0\le\omega_s,\omega_k\le\pi, \\ &E(J_n^{\Delta\epsilon}(\omega_s)J_n^{\Delta\epsilon}(-\omega_k)) \\ &= \begin{cases} \sigma_\epsilon^2(e^{-i\omega_s}+e^{i\omega_k}) & \text{for } 0\le\omega_s\ne\omega_k\le\pi, \\ 2\sigma_\epsilon^2(n-(n-1)\cos(\omega_s)) & \text{for } 0\le\omega_s=\omega_k\le\pi. \end{cases} \end{aligned}$$

Note also that $E(J_n^{\Delta X}(\omega_s)J_n^{\Delta X}(\omega_k)) = \sum_{j=1}^n e^{-i\omega_{s+k}j}E\Delta X_j^2 = 0$ for $\omega_s+\omega_k\ne 0 \pmod{2\pi}$, otherwise, $EJ_n^{\Delta X}(\omega_s)J_n^{\Delta X}(\omega_k) = \sigma_X^2$. As a result,

$$Var\left(\hat{F}_1^{X,\epsilon}(N,n)\right) = 2\sigma_\epsilon^2\sigma_X^2\sum_k W(k)^2(n-(n-1)\cos(\omega_k)).$$

Let $\hat{F}_2^{\epsilon,X}(N,n) = \sum_{k=1}^N W(k)[J_n^{\Delta\epsilon}(\omega_k)J_n^{\Delta X}(-\omega_k)]$. We have for $N < n/2$,

$$
\begin{aligned}
&Cov\left(\hat{F}_1^{X,\epsilon}(N,n), \hat{F}_2^{\epsilon,X}(N,n)\right) \\
&= \sum_{k,k'} W(k)W(k')\left(\sum_{j=1}^n e^{-i\omega_{k+k'}j}E\Delta X_j^2\right)\left(\sigma_\epsilon^2(e^{i\omega_k}+e^{i\omega_{k'}})\right) \\
&= 0.
\end{aligned}
$$

It follows that $Var(\hat{F}^{\Delta X,\Delta\epsilon}(N,n)) = 4\sigma_\epsilon^2\sigma_X^2\sum_k W(k)^2(n-(n-1)\cos\omega_k) = 4\sigma_\epsilon^2\sigma_X^2\sum_k W(k)^2(1+n\omega_k^2/2)+o(N^4/n^2)$.

A.7 Proof of Theorem 4.1

Let $V = [X,X](T)$. Because $Cov(\hat{F}^{\Delta X,\Delta\epsilon}(N,n),\widehat{F}^{\Delta\epsilon}(N,n)) = 0$ and $Cov(\hat{F}^{\Delta X,\Delta\epsilon}(N,n),\widehat{F}^{\Delta X}(N,n)) = 0$, we have

$$
\begin{aligned}
&E(\widehat{F}^{\Delta Y}(N,n)-V)^2 \\
&= \left(E\widehat{F}^{\Delta\epsilon}(N,n)\right)^2 + Var\widehat{F}^{\Delta X}(N,n) \\
&\quad + Var\widehat{F}^{\Delta\epsilon}(N,n) + Var\hat{F}^{\Delta X,\Delta\epsilon}(N,n).
\end{aligned}
$$

Note that

$$
\begin{aligned}
\left(E\widehat{F}^{\Delta\epsilon}(N,n)\right)^2 &= 4\sigma_\epsilon^4 + 4\sigma_\epsilon^4 n\sum_k W(k)\omega_k^2 + O(N^4/n^2), \\
Var\widehat{F}^{\Delta X}(N,n) &= n^{-1}(\xi_n-3)\sigma_X^4 + \sum_k W(k)^2\sigma_X^4, \\
Var\widehat{F}^{\Delta\epsilon}(N,n) &= \sigma_\epsilon^4[2(\eta+1)+4n\sum_k W(k)^2\omega_k^2] + o(N^4/n^2), \\
Var\hat{F}^{\Delta X,\Delta\epsilon}(N,n) &= 4\sigma_\epsilon^2\sigma_X^2\sum_k W(k)^2(1+n\omega_k^2/2)+o(N^4/n^2).
\end{aligned}
$$

As a result,

$$
\begin{aligned}
\text{(A.14)}\quad &E(\widehat{F}^{\Delta Y}(N,n)-V)^2 \\
&= (2\eta+6)\sigma_\epsilon^4 + n^{-1}(\xi_n-3)\sigma_X^4 + 4\sigma_\epsilon^4 n\sum_k W(k)\omega_k^2 \\
&\quad + \sum_k W(k)^2\left[\sigma_X^4 + 4\sigma_\epsilon^4 n\omega_k^2 + 4\sigma_\epsilon^2\sigma_X^2(1+n\omega_k^2/2)\right] \\
&\quad + O(N^4/n^2).
\end{aligned}
$$

Therefore with $N^2 = o(n)$, $E(\widehat{F}^{\Delta Y}(N,n)-V)^2 = (2\eta+6)\sigma_\epsilon^4 + o(1)$.

A.8 Proof of Theorem 4.2

Let $Z_j = (1-(1+\alpha)^{-1}\alpha B)\Delta X_j$ where B is the backshift operator. We have

$$
\text{(A.15)}\quad \widehat{F}^{\Delta Y}(N,n) = (1+\alpha)^2\widehat{F}^Z(N,n) + \widehat{F}^{\Delta\tilde{\epsilon}}(N,n) + (1+\alpha)\hat{F}^{Z,\Delta\tilde{\epsilon}}(N,n),
$$

where $\hat{F}^{Z,\Delta\tilde{\epsilon}}(N,n) = \sum_{1\le k\le N} W(k)(J_n^Z(\omega_k)J_n^{\Delta\tilde{\epsilon}}(-\omega_k) + J_n^Z(-\omega_k)J_n^{\Delta\tilde{\epsilon}}(\omega_k)$. Because Proposition 4.1 applies also to $\tilde{\epsilon}$, we obtain the following

$$
\begin{aligned}
E\hat{F}^{\Delta\tilde{\epsilon}}(N,n) &= 2\sigma_{\tilde{\epsilon}}^2[1+(n-1)\sum_k W(k)\omega_k^2/2] + O(N^4/n^3), \\
Var\hat{F}^{\Delta\tilde{\epsilon}}(N,n) &= \sigma_{\tilde{\epsilon}}^4[2(\eta+1)+4n\sum_k W(k)^2\omega_k^2] + o(N^4/n^2).
\end{aligned}
$$

Moreover,

$$
\text{(A.16)}\quad E\hat{F}^Z(N,n) = \sigma_X^2(1-\theta)^2 + O(N^2/n^2),
$$

$$
\begin{aligned}
\text{(A.17)}\quad Var\hat{F}^Z(N,n) &= \sigma_X^4\sum_k W(k)^2\left[(1-\theta)^4 + 4\theta(1-\theta)^2 n^{-1}\right] \\
&\quad + n^{-1}(\xi_n-3)\sigma_X^4(1-\theta)^4 + o(N^2/n^2),
\end{aligned}
$$

and

$$
E\hat{F}^{Z,\Delta\tilde{\epsilon}}(N,n) = 0,
$$

$$
\begin{aligned}
\text{(A.18)}\quad Var(\hat{F}^{Z,\Delta\tilde{\epsilon}}(N,n)) &= 4(1-\theta)^2\sigma_{\tilde{\epsilon}}^2\sigma_X^2\sum_k W(k)^2(1+n\omega_k^2/2) \\
&\quad + 16\theta\sigma_X^2\sigma_{\tilde{\epsilon}}^2 n^{-1} + o(N^2/n^2),
\end{aligned}
$$

where $\theta = (1+\alpha)^{-1}\alpha$. (They will be verified later.)

Because the variance of ϵ_j is $\sigma_\epsilon^2 = \alpha^2\sigma_X^2/n + \sigma_{\tilde{\epsilon}}^2$, we have

$$
\begin{aligned}
&E\widehat{F}^{\Delta Y}(N,n) - \sigma_X^2 \\
&= (1+\alpha)^2 E\widehat{F}^Z(N,n) + E\widehat{F}^{\Delta\tilde{\epsilon}}(N,n) - \sigma_X^2 \\
&= 2E(\epsilon_j^2) + o(N^2/n).
\end{aligned}
$$

Note that $Cov(\hat{F}^{Z,\Delta\tilde{\epsilon}}(N,n),\widehat{F}^{\Delta\tilde{\epsilon}}(N,n)) = 0$, $Cov(\hat{F}^{Z,\Delta\tilde{\epsilon}}(N,n),\widehat{F}^Z(N,n)) = 0$ and $E\epsilon_j^4 = \eta\sigma_{\tilde{\epsilon}}^4 + 6\alpha^2\sigma_{\tilde{\epsilon}}^2\sigma_X^2 n^{-1} + \alpha^4\xi_n\sigma_X^4 n^{-2}$. Consequently,

$$
\begin{aligned}
\text{(A.19)}\quad &E(\widehat{F}^{\Delta Y}(N,n) - [X,X](T))^2 \\
&= \left(E\widehat{F}^{\Delta\tilde{\epsilon}}(N,n)\right)^2 + Var\widehat{F}^{\Delta\tilde{\epsilon}}(N,n) + (1+\alpha)^4 Var\widehat{F}^Z(N,n) \\
&\quad + (1+\alpha)^2 Var\hat{F}^{Z,\Delta\tilde{\epsilon}}(N,n) \\
&= 2E\epsilon_j^4 + 6\sigma_\epsilon^4 + \left[(\xi_n-3)\sigma_X^4 + 8\alpha(2-\alpha)\sigma_\epsilon^2\sigma_X^2\right]n^{-1} \\
&\quad + 4\sigma_\epsilon^4 n\sum_k W(k)\omega_k^2 + (\sigma_X^4 + 4\sigma_\epsilon^2\sigma_X^2)\sum_k W(k)^2 \\
&\quad + (4\sigma_\epsilon^4 + 2\sigma_\epsilon^2\sigma_X^2)n\sum_k W(k)^2\omega_k^2 + O(N^4/n^2) \\
&\quad + o(N^2/n\sum W(k)^2) \\
&= 2E\epsilon_j^4 + 6\sigma_\epsilon^4 + o(1).
\end{aligned}
$$

Proof of (A.16): Note that $I_n^Z(\omega) = n^{-1}\sum_{|h|<n}\gamma_Z(h)e^{-i\omega h}$ where $\gamma_Z(h) = \sum_{j=1}^{n-|h|} Z_j Z_{j+|h|}$, and $\cos\omega = 1-\omega^2/2+O(\omega^4)$. We have

$$
\begin{aligned}
&E\hat{F}^Z(N,n) \\
&= \sigma_{\Delta X}^2[n(1+\theta^2) - 2\theta(n-1)\sum_{1\le k\le N} W(k)\cos\omega_k] \\
&= \sigma_X^2(1-\theta)^2 + O(N^2/n^2).
\end{aligned}
$$

Proof of (A.17): It follows from Lemma A.1 that, with $E\Delta X_j^4 = \xi_n\sigma_X^4/n^2$,

$$
\begin{aligned}
&Var(nI_n^Z(\omega_s)) \\
&= n^{-2}\sigma_X^4(\xi_n-3)(1+\theta^4+(n-1)(1+\theta^2-2\theta\cos\omega_s)^2) \\
&\quad + 4\theta^2 n^{-2}\sigma_X^4 + n^{-2}\sigma_X^4((1+\theta^2)n - 2\theta(n-1)\cos\omega_s)^2,
\end{aligned}
$$

for $0 < \omega_s = 2\pi s/n < \pi$, and

$$\begin{aligned} &Cov(nI_n^Z(\omega_s), nI_n^Z(\omega_l)) \\ &= n^{-2}\sigma_X^4(\xi_n - 3)(n-1)(1+\theta^2 - 2\theta\cos\omega_s)(1+\theta^2-2\theta\cos\omega_l) \\ &\quad + n^{-2}\sigma_X^4(\xi_n-3)(1+\theta^4) + 4\theta^2 n^{-2}\sigma_X^4\left(1+\cos\omega_s\cos\omega_l\right), \end{aligned}$$

for $0 \le \omega_s \ne \omega_l \le \pi$. Note also that $\xi_n = O(1)$. We have

$$\begin{aligned} &Var\hat{F}^Z(N,n) \\ &= \sum_s W(s)^2 Var|J_n^Z(\omega_s)|^2 \\ &\quad + \sum_{s\ne l} W(s)W(l)Cov(|J_n^Z(\omega_s)|^2, |J_n^Z(\omega_l)|^2) \\ &= n^{-2}\sigma_X^4(\xi_n-3)(1+\theta^4) + 4\theta^2 n^{-2}\sigma_X^4 \\ &\quad + n^{-2}\sigma_X^4(\xi_n-3)(n-1)\left[1+\theta^2\right]^2 + (1+\theta^2)^2\sigma_X^4\sum_s W(s)^2 \\ &\quad - 4\theta(1+\theta^2)n^{-2}\sigma_X^4(\xi_n-3)(n-1)\left[\sum_s W(s)(\cos\omega_s)\right] \\ &\quad + 4\theta^2 n^{-2}\sigma_X^4[(\xi_n-3)(n-1)+1]\left[\sum_s W(s)(\cos\omega_s)\right]^2 \\ &\quad + 4\theta^2\sigma_X^4(1-2n^{-1})\sum_s W(s)^2(\cos\omega_s)^2 \\ &\quad - 4(1+\theta^2)\theta(1-n^{-1})\sigma_X^4\sum_s W(s)^2\cos\omega_s \\ &= \sigma_X^4\sum_s W(s)^2\left[(1-\theta)^4 + 4\theta(1-\theta)^2 n^{-1}\right] \\ &\quad + n^{-1}(\xi_n-3)\sigma_X^4(1-\theta)^4 + o(N^2/n^2). \end{aligned}$$

Proof of (A.18): Let $\hat{F}_1^{Z,\Delta\tilde{\epsilon}}(N,n) = \sum_{k=1}^N W(k)J_n^Z(\omega_k)J_n^{\Delta\tilde{\epsilon}}(-\omega_k)$. We have

$$\begin{aligned} &Var\left(\hat{F}_1^{Z,\Delta\tilde{\epsilon}}(N,n)\right) \\ &= \sum_{k\ne k'} W(k)W(k')E\left[J_n^Z(\omega_k)J_n^Z(-\omega_{k'})\right]E\left[J_n^{\Delta\tilde{\epsilon}}(-\omega_k)J_n^{\Delta\tilde{\epsilon}}(\omega_{k'})\right] \\ &\quad + \sum_k W(k)^2 E|J_n^Z(\omega_k)|^2 E|J_n^{\Delta\tilde{\epsilon}}(\omega_k)|^2. \end{aligned}$$

In view of (A.12) and (A.13), we obtain

$$\begin{aligned} &E(J_n^{\Delta\tilde{\epsilon}}(\omega_s)J_n^{\Delta\tilde{\epsilon}}(\omega_k)) \\ &= \sigma_{\tilde{\epsilon}}^2(e^{-i\omega_s}+e^{-i\omega_k}), \quad \text{for } 0<\omega_s,\omega_k<\pi, \\ &E(J_n^{\Delta\tilde{\epsilon}}(\omega_s)J_n^{\Delta\tilde{\epsilon}}(-\omega_k)) \\ &= \begin{cases} \sigma_{\tilde{\epsilon}}^2(e^{-i\omega_s}+e^{i\omega_k}) & \text{for } 0<\omega_s\ne\omega_k<\pi, \\ 2\sigma_{\tilde{\epsilon}}^2(n-(n-1)\cos\omega_s) & \text{for } 0<\omega_s=\omega_k<\pi, \end{cases} \end{aligned}$$

and

$$\begin{aligned} &E(J_n^Z(\omega_s)J_n^Z(\omega_k)) \\ &= \theta\sigma_{\Delta X}^2(e^{-i\omega_s}+e^{-i\omega_k}), \quad \text{for } 0<\omega_s,\omega_k<\pi, \\ &E(J_n^Z(\omega_s)J_n^Z(-\omega_k)) \\ &= \begin{cases} \theta\sigma_{\Delta X}^2(e^{-i\omega_s}+e^{i\omega_k}) & \text{for } 0<\omega_s\ne\omega_k<\pi, \\ \sigma_{\Delta X}^2((1+\theta^2)n-2\theta(n-1)\cos\omega_s) & \text{for } 0<\omega_s=\omega_k<\pi, \end{cases} \end{aligned}$$

where $\sigma_{\Delta X}^2 = n^{-1}\sigma_X^2$. Consequently,

$$\begin{aligned} &Var\left(\hat{F}_1^{Z,\Delta\tilde{\epsilon}}(N,n)\right) \\ &= \underbrace{2\sigma_{\tilde{\epsilon}}^2\sigma_{\Delta X}^2\sum_k W(k)^2(n-(n-1)\cos\omega_k)}_{\mathbb{T}_1} \\ &\quad \times \underbrace{((1+\theta^2)n-2\theta(n-1)\cos\omega_k)}_{\mathbb{T}_1} \\ &\quad \underbrace{+\theta\sigma_{\Delta X}^2\sigma_{\tilde{\epsilon}}^2\sum_{k\ne s}W(k)W(s)(e^{-i\omega_s}+e^{i\omega_k})(e^{i\omega_s}+e^{-i\omega_k})}_{\mathbb{T}_2}. \end{aligned}$$

Because $(n-(n-1)\cos\omega_k)((1+\theta^2)n-2\theta(n-1)\cos\omega_k) = (1-\theta)^2 n + 2\theta + (1-\theta)^2 n^2\omega_k^2/2 + O(n\omega_k^2)$, we have

$$\begin{aligned} \mathbb{T}_1 &= 2\sigma_{\tilde{\epsilon}}^2\sigma_X^2\sum_k W(k)^2((1-\theta)^2 + 2\theta n^{-1} + (1-\theta)^2 n\omega_k^2/2) \\ &\quad + O(N^2/n^2\sum W(k)^2). \end{aligned}$$

Note also that $(e^{-i\omega_s}+e^{i\omega_k})(e^{i\omega_s}+e^{-i\omega_k}) = 2(1+\cos\omega_k\cos\omega_s - \sin\omega_k\sin\omega_s)$. With some algebra, we have

$$\mathbb{T}_2 = 4\theta\sigma_X^2\sigma_{\tilde{\epsilon}}^2 n^{-1}\left[1-\sum W(k)^2\right] + O(N^2/n^3).$$

It follows that

$$\begin{aligned} &Var\left(\hat{F}_1^{Z,\Delta\tilde{\epsilon}}(N,n)\right) \\ &= 2(1-\theta)^2\sigma_{\tilde{\epsilon}}^2\sigma_X^2\sum_k W(k)^2(1+n\omega_k^2/2) \\ &\quad + 4\theta\sigma_X^2\sigma_{\tilde{\epsilon}}^2 n^{-1} + O(N^2/n^2\sum W(k)^2). \end{aligned}$$

Let $\hat{F}_2^{\Delta\tilde{\epsilon},Z}(N,n) = \sum_{k=1}^N W(k)[J_n^{\Delta\tilde{\epsilon}}(\omega_k)J_n^Z(-\omega_k)]$. Then

$$Var\left(\hat{F}_2^{\Delta\tilde{\epsilon},Z}(N,n)\right) = Var\left(\hat{F}_1^{Z,\Delta\tilde{\epsilon}}(N,n)\right),$$

and

$$\begin{aligned} &Cov\left(\hat{F}_1^{Z,\Delta\tilde{\epsilon}}(N,n), \hat{F}_2^{\Delta\tilde{\epsilon},Z}(N,n)\right) \\ &= n^{-1}\theta\sigma_X^2\sigma_{\tilde{\epsilon}}^2\sum_{k,k'}W(k)W(k')(e^{-i\omega_k}+e^{-i\omega_{k'}})(e^{i\omega_k}+e^{i\omega_{k'}}). \end{aligned}$$

Because $\sum_{k,k'}W(k)W(k')(e^{-i\omega_k}+e^{-i\omega_{k'}})(e^{i\omega_k}+e^{i\omega_{k'}}) = 4 + O(N^2/n^2)$, we have

$$Cov\left(\hat{F}_1^{Z,\Delta\tilde{\epsilon}}(N,n), \hat{F}_2^{\Delta\tilde{\epsilon},Z}(N,n)\right) = 4n^{-1}\theta\sigma_X^2\sigma_{\tilde{\epsilon}}^2 + O(N^2/n^3).$$

Consequently,

$$\begin{aligned} &Var\left(\hat{F}_1^{Z,\Delta\tilde{\epsilon}}(N,n) + \hat{F}_2^{\Delta\tilde{\epsilon},Z}(N,n)\right) \\ &= 16\theta\sigma_X^2\sigma_{\tilde{\epsilon}}^2 n^{-1} + 4(1-\theta)^2\sigma_{\tilde{\epsilon}}^2\sigma_X^2\sum_k W(k)^2(1+n\omega_k^2/2) \\ &\quad + o(N^2/n^2). \end{aligned}$$

A.9 Proof of Theorem 5.1

It follows from (A.19) that for a fixed n, the optimal cut-off frequency should be such that minimizes to leading order

$$4\sigma_\epsilon^4 n \sum_k W(k)\omega_k^2 + (\sigma_X^4 + 4\sigma_\epsilon^2\sigma_X^2)\sum_k W(k)^2 + (4\sigma_\epsilon^4 + 2\sigma_\epsilon^2\sigma_X^2) n \sum_k W(k)^2\omega_k^2$$

(A.20)

$$= 4\sigma_\epsilon^4 \left(n\sum_k W(k)\omega_k^2 + \sum_k W(k)^2 \left[\rho^2 + 2\rho + (1+\rho) n\omega_k^2\right]\right)$$

where $\rho = (2\sigma_\epsilon^2)^{-1}\sigma_X^2$.

A.10 Proof of Corollary 5.1

Note that $\sum_{1\le k\le N} W(k)^2 = \frac{1}{N^2} + \frac{2(N-1)(2N-1)}{3N^3}$, $\sum_{1\le k\le N} W(k)k^2 = [\frac{(N+1)^2(N+2)}{6N} - \frac{1}{N}]$, and $\sum_{1\le k\le N} W(k)^2k^2 = \frac{2(N+1)(N^3+4N^2+6N+4)}{15N^3} - \frac{3}{N^2}$. Equation (A.20) becomes

$$\frac{2\pi^2}{3n}N^2 + \frac{8\pi^2(6+\rho)}{15n}N + \frac{4(\rho^2+2\rho)}{3}\frac{1}{N}.$$

For $N \in (0,\infty)$, $N \mapsto \frac{2\pi^2}{3n}N^2 + \frac{8\pi^2(6+\rho)}{15n}N + \frac{4(\rho^2+2\rho)}{3}\frac{1}{N}$ is minimized at $N_* = -b + (-b^3 - d + \sqrt{d(d+2b^3)})^{1/3} + (-b^3 - d - \sqrt{d(d+2b^3)})^{1/3}$ where $b = \frac{2(6+\rho)}{15}$, and $d = -\frac{n\rho(\rho+2)}{2\pi^2}$, and $\rho = \sigma_X^2/(2\sigma_\epsilon^2)$. The proof is complete.

A.11 Proof of Corollary 5.2

Because $\sum_k W(k)k^2 = \frac{(N+1)(N+2)}{6}$ and $\sum_k W(k)^2k^2 = \frac{2}{15}[N+3+\frac{4}{N} - \frac{1}{N+1}]$, (A.20) is equal to

$$\frac{2(\rho^2+2\rho)}{3}\left[\frac{1}{N} + \frac{1}{N+1}\right] + \frac{4\pi^2}{n}\frac{(N+1)(N+2)}{6} + \frac{8\pi^2(1+\rho)}{15n}\left[N+3+\frac{4}{N} - \frac{1}{N+1}\right] \approx \frac{2\pi^2}{3n}N^2 + \frac{2\pi^2(19+4\rho)}{15n}N + \frac{4(\rho^2+2\rho)}{3}\frac{1}{N}.$$

For $N \in (0,\infty)$, $N \mapsto \frac{2\pi^2}{3n}N^2 + \frac{2\pi^2(19+4\rho)}{15n}N + \frac{4(\rho^2+2\rho)}{3}\frac{1}{N}$ is minimized at $N_* = -b + (-b^3 - d + \sqrt{d(d+2b^3)})^{1/3} + (-b^3 - d - \sqrt{d(d+2b^3)})^{1/3}$ $b = \frac{19+4\rho}{30}$, and $d = -\frac{n\rho(\rho+2)}{2\pi^2}$. The proof is complete.

Received 17 January 2014

REFERENCES

[1] Aït-Sahalia, Y. and Mykland, P. A. Estimators of diffusions with randomly spaced discrete observations: A general theory. *The Annals of Statistics*, 32(5):2186–2222, 2004. MR2102508

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

[3] Andersen, T., Bollerslev, T., Diebold, F. X., and Labys, P. The distribution of exchange rate volatility. *Journal of the American statistical Association*, 96:42–55, 2001. MR1952727

[4] Andersen, T. G. and Bollerslev, T. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 885–905, 1998.

[5] Andersen, T. G., Bollerslev, T., and Lange, S. Forecasting financial market volatility: Sample frequency vis-à-vis forecast horizon. *Journal of Empirical Finance*, 6(5):457–477, 1999.

[6] Andersen, T. G., Bollerslev, T., and Meddahi, N. Realized volatility forecasting and market microstructure noise. *Journal of Econometrics*, 160(1):220–234, 2011. MR2745879

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

[8] Barndorff-Nielsen, O. E. and Shephard, N. 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, 2002. MR1904704

[9] Barndorff-Nielsen, O. E. and Shephard, N. Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2(1):1, 2004.

[10] Barndorff-Nielsen, O. E., Graversen, S. E., Jacod, J., and Shephard, N. Limit theorems for bipower variation in financial econometrics. *Econometric Theory*, 22(4):677, 2006. MR2283032

[11] Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica*, 76(6):1481–1536, 2008. MR2468558

[12] Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. Realized kernels in practice: Trades and quotes. *The Econometrics Journal*, 12(3):C1–C32, 2009. MR2597033

[13] Barndorff-Nielsen, O. E. and Shephard, N. Non-gaussian ornstein–uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2):167–241, 2001. MR1841412

[14] Barndorff-Nielsen, O. E, Hansen, P. R., Lunde, A., and Shephard, N. 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, 2011. MR2795610

[15] Bartlett, M. S. *An Introduction to Stochastic Processes: With Special Reference to Methods and Applications*. Cambridge University Press, 1977.

[16] Bloomfield, P. *Fourier Analysis of Time Series: An Introduction*. Wiley-Interscience, 2004. MR1884963

[17] Brillinger, D. R. *Time Series: Data Analysis and Theory*, volume 36. Society for Industrial Mathematics, 2001. MR1853554

[18] Brockwell, P. J. and Davis, R. A. *Time Series: Theory and Methods*. Springer Verlag, 2009.

[19] Chernov, M., Gallant, A. R., Ghysels, E., and Tauchen, G. Alternative models for stock price dynamics. *Journal of Econometrics*, 116(1):225–257, 2003. MR2011152

[20] Dahlhaus, R. Asymptotic normality of spectral estimates. *Journal of Multivariate Analysis*, 16(3):412–431, 1985. MR0793500

[21] Grenander, U. and Rosenblatt, M. *Statistical Analysis of Stationary Time Series*. Chelsea Pub Co, 2008. MR0890514

[22] Hansen, P. R. and Lunde, A. Realized variance and market microstructure noise. *Journal of Business & Economic Statistics*, 24(2):127–161, 2006. MR2234447

[23] Huang, X. and Tauchen, G. The relative contribution of jumps to total price variance. *Journal of Financial Econometrics*, 3(4): 456–499, 2005.

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

[25] Kanatani, T. Integrated volatility measuring from unevenly sampled observations. *Economics Bulletin*, 3(36):1–8, 2004.

[26] Malliavin, P. and Mancino, M. E. Fourier series method for measurement of multivariate volatilities. *Finance and Stochastics*, 6(1):49–61, 2002. MR1885583

[27] Malliavin, P. and Mancino, M. E. A fourier transform method for nonparametric estimation of multivariate volatility. *The Annals of Statistics*, 37(4):1983–2010, 2009. MR2533477

[28] Mancino, M. E. and Sanfelici, S. Robustness of fourier estimator of integrated volatility in the presence of microstructure noise. *Computational Statistics & Data Analysis*, 52(6):2966–2989, 2008. MR2424771

[29] Meddahi, N. A theoretical comparison between integrated and realized volatility. *Journal of Applied Econometrics*, 17(5):479–508, 2002.

[30] Meddahi, N. and Renault, E. Temporal aggregation of volatility models. *Journal of Econometrics*, 119(2):355–379, 2004. MR2057104

[31] Mykland, P. A. and Zhang, L. Anova for diffusions and ito processes. *The Annals of Statistics*, 34(4):1931–1963, 2006. MR2283722

[32] Priestley, M. B. *Spectral Analysis and Time Series*. Academic Press, 1981.

[33] Todorov, V. and Tauchen, G. The realized laplace transform of volatility. *Econometrica*, 80(3):1105–1127, 2012. MR2963883

[34] Wang, F. Optimal design of fourier estimator in the presence of microstructure noise. Computational Statistics & Data Analysis, forthcoming, 2013.

[35] Wang, F. An unbiased measure of integrated volatility in the frequency domain. Working paper, UIC, 2014.

[36] Zhou, B. High-frequency data and volatility in foreign-exchange rates. *Journal of Business & Economic Statistics*, 45–52, 1996.

Fangfang Wang
Department of Information and Decision Sciences
University of Illinois at Chicago
Chicago, IL 60607-7124
USA
E-mail address: ffwang@uic.edu