

# Realised Quantile-Based Estimation of the Integrated Variance

Kim Christensen

Roel Oomen

Mark Podolskij\*

January 2008

## Abstract

In this paper we propose a new jump robust measure of ex-post return variation that can be computed using potentially noisy data. The estimator exploits the link between return quantiles and volatility and is consistent for the integrated (diffusive) variance under weak conditions on the price process. We present various central limit theorems which show that the estimator converges at the best attainable rate and has excellent efficiency. Asymptotically, the estimator is immune to finite activity jumps and simulations show that also in finite sample it has superior robustness properties. In modified form, the estimator is applicable with market microstructure noise and therefore operational on high frequency data. As such, it constitutes an appealing alternative to the existing jump-robust or noise-corrected realised variance measures. An empirical application using low and high frequency data is included to further illustrate the properties of the estimator.

**Keywords:** Finite activity jumps; Integrated variance; Market microstructure noise; Order statistics; Realised variance.

**JEL Classification:** C10; C80.

---

\*Christensen is with Nordea Markets, Copenhagen, Denmark. E-mail [kim.chr@nordea.com](mailto:kim.chr@nordea.com). Oomen is with Deutsche Bank, London, UK. E-mail: [roel.oomen@db.com](mailto:roel.oomen@db.com). Podolskij is with CREATES, University of Aarhus, Denmark. E-mail: [mpodolskij@creates.au.dk](mailto:mpodolskij@creates.au.dk). Christensen is also affiliated with CREATES, University of Aarhus, Denmark and Oomen with the Department of Finance, Warwick Business School, United Kingdom and the Department of Quantitative Economics, the University of Amsterdam, The Netherlands. We would like to thank Neil Shephard, the seminar participants at London-Oxford Financial Econometrics Study Group, Imperial College, London and ECARES, Université Libre de Bruxelles, Brussels for valuable comments and discussions.

# 1 Introduction

In recent years, it has become common practice to use high-frequency data when making inference about the variation of financial asset prices. The realised variance, defined as the sum of squared intra-period returns, has been a key driver in this development (e.g. Andersen and Bollerslev, 1998; Andersen, Bollerslev, Diebold, and Labys, 2001; Barndorff-Nielsen and Shephard, 2002). However, two issues arise naturally with the use of realised variance. First, realised variance is not well-suited for handling the impact of microstructure noise, which is an important component of high-frequency data. Second, realised variance is an aggregate statistic that measures the overall variation of the price process, irrespective of whether it comes from continuous or discontinuous sample path variation. While this latter feature is desirable in some instances, in others it can be a drawback (for instance, when realised variance is used to estimate parametric models of the integrated variance, forecast future volatility, or price and hedge derivatives). Correcting for market microstructure induced biases in realised variance is a very active area of research and the three main approaches<sup>1</sup> are based on sub-sampling (see Zhang, 2006; Zhang, Mykland, and Ait-Sahalia, 2005), kernel-based autocovariance adjustments (see Barndorff-Nielsen, Hansen, Lunde, and Shephard, 2006; Zhou, 1996), and pre-filtering or pre-averaging methods (see Podolskij and Vetter, 2006, 2007; Jacod, Li, Mykland, Podolskij, and Vetter, 2007). At the same time, a largely independent literature has developed on jump-robust realised variance measures, most notably the bi-power variation of Barndorff-Nielsen and Shephard (2004a). Yet, the noise-robust realised variances are typically not robust to jumps and the jump-robust realised variances are not robust to noise (with the notable exception of Podolskij and Vetter, 2007). The contribution of this paper is to integrated both issues into a new quantile-based realised variance estimator, or “QRV”, that we show to be simultaneously robust to noise *and* jumps.

The idea behind the construction of QRV is based on the elementary relationship between quantiles and the spread of the normal distribution. In particular, for i.i.d. Gaussian returns with mean zero and variance  $\sigma^2$  we know that, for instance,  $Q_{95} = 1.645\sigma$  where  $Q_k$  denotes the  $k^{th}$  quantile of returns. Thus, with a measurement of the return quantile we can construct an estimator of the return volatility as  $\hat{Q}_{95}/1.645$ . Of course, it is also possible to use multiple quantiles, e.g.  $\hat{\sigma}_q = (\hat{Q}_{95} - \hat{Q}_5)/(2 \times 1.645)$ . Estimators of this kind can be traced back to David (1970) who studies their properties for i.i.d. normal random variables. The QRV estimator we propose in this paper is inspired by the work of David (1970), but we make considerable progress in several directions. Firstly, the proposed QRV yields consistent estimates of the integrated variance under very weak conditions on the price process, allowing for general stochastic volatility dynamics and leverage. Secondly, we explicitly consider (i.i.d) market microstructure noise and develop a modified QRV measure that remains consistent for the integrated variance of the efficient price process. Thirdly, we present various central limit

---

<sup>1</sup>Other methods include sparse sampling (e.g. Andersen, Bollerslev, Diebold, and Labys, 2000; Bandi and Russell, 2005), pre-filtering (Andersen, Bollerslev, Diebold, and Ebens, 2001; Bollen and Inder, 2002; Hansen, Large, and Lunde, 2006), model based corrections (e.g. Corsi, Zumbach, Müller, and Dacorogna, 2001), and wavelet based methods of Fan and Wang (2006). Variations and extensions of the realised kernel and subsampling approach can be found in Sun (2006) and Nolte and Voev (2007) respectively

theorems which show that the convergence rate of QRV is the best attainable, both in the absence and in the presence of microstructure noise. Thus, from a practical perspective, QRV is operational with high frequency data and – as we show in this paper – it makes an appealing alternative to the currently available realised variance measures when jump robust estimates of the integrated variance are required.

It is important to point out that the proposed QRV estimator is based on *return quantiles* and therefore distinctly different from (and not a generalization of) variance estimators based on the *price range* as studied by, e.g., Feller (1951), Garman and Klass (1980), Parkinson (1980), Beckers (1983), and Rogers and Satchell (1991). In its simplest form, these range based volatility estimators exploit the relation  $E(\max p - \min p)^2 = \sigma^2 \ln 16$  (where  $p$  is the log price process) which can be inverted to yield an estimate of the variance as  $\hat{\sigma}^2 = (\widehat{\max p} - \widehat{\min p}) / \ln 16$ . With increasingly fine sampling of  $p$ , we can obtain error-free measurements of the price range as well as the return quantiles. But, intuitively, because the relation between return quantiles and volatility is an identity, and the relation between the price range and volatility only holds in expectation, QRV is consistent whereas the above price range based estimator is not. Another important difference between the two is that QRV is robust to jumps (as it effectively ignores a fraction of the largest and smallest return observations) while the price range based estimator is clearly very sensitive to jumps. As a final remark, we point out that Christensen and Podolskij (2006) have shown that the inconsistency of the conventional price range based estimator can be overcome by computing and summing it over a growing number of sub-intervals (see also Martens and van Dijk, 2007). In this paper, we adopt a similar logic and also consider QRV on sub-intervals but the primary motivation for doing this is different: namely to account for stochastic volatility. If volatility is constant, QRV is consistent and there is no need to average over sub-intervals.

The remainder of this paper is organized as follows. In section 2, we present the quantile-based realised variance estimator and discuss its properties. By studying its asymptotic distribution, we show that it leads to efficient inference about the continuous component of the return variation. Indeed, using a double asymptotics, the efficiency of our estimator can be pushed arbitrarily close to the maximum likelihood bound of the realised variance, while retaining robustness to jumps. In section 3, we analyze the behavior of our statistic under i.i.d. microstructure noise, which we do by applying a pre-filtering method advocated by Podolskij and Vetter (2006) and then working with the QRV on the filtered data. Section 4 contains an empirical application where we implement QRV on high and low frequency data. Section 5 concludes.

## 2 Quantile-based realised variance measurement

Let  $X = (X_t)_{t \geq 0}$  denote the log-price process, defined on a filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  and adapted to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ . The theory of financial economics states that if  $X$  evolves in a frictionless market, then it has to be of semimartingale form (see Back, 1991). In this paper we assume that  $X$  is a pure Brownian semimartingale, i.e. a

continuous sample path process of the form:

$$X_t = X_0 + \int_0^t a_u du + \int_0^t \sigma_u dW_u, \quad t \geq 0, \quad (1)$$

where  $a = (a_t)_{t \geq 0}$  is a predictable locally bounded drift function,  $\sigma = (\sigma_t)_{t \geq 0}$  is a càdlàg volatility process and  $W = (W_t)_{t \geq 0}$  a standard Brownian motion.

To prove our CLTs, we will work under some stronger assumptions on  $\sigma$ .

**Assumption (V)**  $\sigma$  does not vanish ( $V_1$ ) and it satisfies the equation:

$$\sigma_t = \sigma_0 + \int_0^t a'_u du + \int_0^t \sigma'_u dW_u + \int_0^t v'_u dB'_u, \quad t \geq 0, \quad (V_2)$$

where  $a' = (a'_t)_{t \geq 0}$ ,  $\sigma' = (\sigma'_t)_{t \geq 0}$  and  $v' = (v'_t)_{t \geq 0}$  are càdlàg, with  $a'$  also being locally bounded and predictable,  $B' = (B'_t)_{t \geq 0}$  is a Brownian motion, and  $W \perp\!\!\!\perp B'$  (here  $A \perp\!\!\!\perp B$  means that  $A$  and  $B$  are stochastically independent).

This means that  $\sigma$  has its own Brownian semimartingale structure. Note the appearance of  $W$  in  $\sigma$ , which allows for leverage effects. If  $X$  is a unique strong solution of a stochastic differential equation then, under some smoothness assumptions on the volatility function  $\sigma = \sigma(t, X_t)$ , assumption (V<sub>2</sub>) (with  $v'_s = 0$  for all  $s$ ) is a consequence of Ito's formula. Thus, assumption (V<sub>2</sub>) is fulfilled for many financial models and, even though it is not a necessary condition, it simplifies the proofs considerably. A more general treatment, including the case where  $\sigma$  jumps, can be found in Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard (2006). We rule out these technical details here, as they are not important to our exposition.

In what follows, we also make use of the concept of stable convergence in law.

**Definition 1** A sequence of random variables,  $(Z_n)_{n \in \mathbb{N}}$ , converges stably in law with limit  $Z$ , defined on an appropriate extension of  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , if and only if for every  $\mathcal{F}$ -measurable, bounded random variable  $Y$  and any bounded, continuous function  $g$ , the convergence  $\lim_{n \rightarrow \infty} \mathbb{E}[Yg(Z_n)] = \mathbb{E}[Yg(Z)]$  holds. We write  $Z_n \xrightarrow{d_s} Z$ , if  $(Z_n)_{n \in \mathbb{N}}$  converges stably in law to  $Z$ .

Stable convergence implies weak convergence, or convergence in law, which can be defined as above by taking  $Y = 1$ , see Rényi (1963) or Aldous and Eagleson (1978) for more details about the properties of stably converging sequences. The extension of this concept to stable convergence of processes is discussed in Jacod and Shiryaev (2003, pp. 512-518). In our context, we need the stable convergence to transform the infeasible mixed Gaussian CLTs proved below into feasible ones that can be implemented on real data.

Central to the theory of semimartingales is the quadratic variation process, defined as:

$$[X]_t = \text{plim}_{n \rightarrow \infty} \sum_{i=1}^n (X_{t_i} - X_{t_{i-1}})^2,$$

for any sequence of partitions  $0 = t_0 < t_1 < \dots < t_n = t$  such that  $\sup_i \{t_i - t_{i-1}\} \rightarrow 0$  as  $n \rightarrow \infty$  (see, e.g. Protter, 2004). For the process in Eq. (1), the quadratic variation is equal to integrated variance, i.e.

$$[X]_t = \int_0^t \sigma_u^2 du.$$

As in Andersen and Bollerslev (1998), and many of the subsequent papers in this area, here the object of econometric interest is the integrated variance (IV hereafter).

## 2.1 The estimator and its properties

From now on, we will work on the unit time interval without loss of generality, i.e.  $t \in [0, 1]$ . We assume that  $X$  is observed at equidistant points  $t_i = i/N$ , for  $i = 0, \dots, N$ . The increments of  $X$  – the continuously compounded returns – are denoted as:

$$\Delta_i^N X = X_{i/N} - X_{(i-1)/N},$$

for  $i = 1, \dots, N$ . We further assume that  $N = nm$ , where  $m, n$  are natural numbers. Specifically, we consider  $n$  subintervals  $[(i-1)/n, i/n]$ , each containing  $m$  returns, i.e.

$$D_i^m X = (\Delta_k^N X)_{(i-1)m+1 \leq k \leq im}$$

for  $i = 1, \dots, n$ . In the asymptotic analysis below, we consider the case where (i)  $m$  is fixed and  $n \rightarrow \infty$ , (ii)  $m \rightarrow \infty$  and  $n$  is fixed, and (iii)  $n, m \rightarrow \infty$ . Next, for  $x = (x_1, \dots, x_m)$ , we define the function  $g_k : \mathbb{R}^m \rightarrow \mathbb{R}$  such that

$$g_k(x) = x_{(k)}, \quad (2)$$

where  $x_{(k)}$  is the  $k$ th order statistic of  $x$ . This allows us to define the realised (symmetric) squared  $\lambda$ -quantile as

$$q_i(m, \lambda) = g_{\lambda m}^2 \left( \sqrt{N} D_i^m X \right) + g_{m-\lambda m+1}^2 \left( \sqrt{N} D_i^m X \right), \quad (3)$$

where  $\lambda \in (1/2, 1)$  and  $\lambda m$  is a natural number. The function  $q_i(m, \lambda)$  is even in  $X$ , so its value does not change if we replace  $X$  by  $-X$ . Also note that  $\sqrt{N} D_i^m X$  has been normalized to be  $O_p(1)$ .

We are now in a position to define the quantile-based realised variance (QRV, hereafter):

$$QRV_N(m, \lambda) = \frac{1}{\nu_1(m, \lambda)} \frac{m}{N} \sum_{i=1}^{N/m} q_i(m, \lambda), \quad \text{for } \lambda \in (1/2, 1). \quad (4)$$

The scaling factor  $\nu$  in Eq. (4) is given by:

$$\nu_r(m, \lambda) = \mathbb{E} \left[ \left( |U_{(\lambda m)}|^2 + |U_{(m-\lambda m+1)}|^2 \right)^r \right], \quad (5)$$

for  $r > 0$  and where  $U_{(\lambda m)}$  is the  $(\lambda m)$ -th order statistic of an independent standard normal sample  $\{U_i\}_{i=1}^m$ .

Thus, QRV is calculated as the realised squared  $\lambda$  and  $1 - \lambda$  return quantiles, averaged across the  $n$  subintervals and scaled by the constant  $\nu_1(m, \lambda)$ . The idea of building subintervals is quite natural (see e.g. Christensen and Podolskij, 2006) and the mathematical intuition of this procedure is discussed in Mykland (2006).

As we will show below, QRV constitutes a consistent estimator of the integrated variance that, due to its reliance on return quantiles with  $\lambda \neq 1$ , is robust to finite activity jumps. In its current form, QRV uses a single pair of symmetric quantiles. As we will discuss in Section 2.2, it is possible to calculate QRV using multiple pairs of quantiles to improve its efficiency. In fact, we will see that QRV can attain the efficiency of the maximum likelihood estimator while retaining robustness to jumps. It is also possible to use asymmetric quantiles in Eq. (3), but, because this is suboptimal in the current setting (due to the symmetry of the normal distribution), we do not consider this case. Moreover, when  $q_i(m, \lambda)$  is based on asymmetric quantiles the resulting central limit theorem for QRV is unfeasible (hence, can not be applied in practice), because the function  $q_i(m, \lambda)$  is not even in  $X$  anymore (see Kinnebrock and Podolskij, 2007, for a discussion of this issue).

We now present results for QRV as defined in Eq. (4) with scalar  $\lambda$ .

**Theorem 1** *For the process in Eq. (1), and  $N = mn$  with  $m$  fixed, we have as  $N \rightarrow \infty$*

$$QRV_N(m, \lambda) \xrightarrow{p} IV, \quad (6)$$

where  $IV = \int_0^1 \sigma_u^2 du$ .

**Proof** see Appendix A.1 ■

Theorem 1 shows consistency of QRV for the integrated variance under very weak conditions on the process  $X$ . Moreover, QRV is robust to finite activity jumps.

**Proposition 1** *Theorem 1 remains valid for an extension of the process in Eq. (1) that incorporates finite activity jumps, i.e.*

$$X_t = X_0 + \int_0^t a_u du + \int_0^t \sigma_u dW_u + \sum_{i=1}^{q(t)} J_i, \quad t \geq 0,$$

where  $q = (q(t))_{t \geq 0}$  is a finite activity counting process and  $J = (J_i)_{i=1}^{q(t)}$  are non-zero random variables representing the jumps in  $X$ .

The intuition for this result is as follows: as  $N \rightarrow \infty$  each subinterval on which we compute the return quantiles shrinks so that, in the limit, it contains at most one jump with probability 1. But because  $\lambda \in (1/2, 1)$ , QRV is asymptotically immune to finite activity jumps.

**Theorem 2** For the process in Eq. (1), with condition (V) satisfied, and  $N = mn$  with  $m$  fixed, we have as  $N \rightarrow \infty$

$$\sqrt{N} (QRV_N(m, \lambda) - IV) \xrightarrow{d_s} \sqrt{\theta(m, \lambda)} \int_0^1 \sigma_u^2 dW'_u, \quad (7)$$

where

$$\theta(m, \lambda) = m \frac{\nu_2(m, \lambda) - \nu_1^2(m, \lambda)}{\nu_1^2(m, \lambda)}, \quad (8)$$

and  $W'$  is another Brownian motion defined on an extension of  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  and  $W' \perp\!\!\!\perp \mathcal{F}$ . Because  $\sigma$  is independent of the Brownian motion  $W'$ , we have stable convergence to a mixed normal:

$$\sqrt{N} (QRV_N(m, \lambda) - IV) \xrightarrow{d_s} MN(0, \theta(m, \lambda) IQ). \quad (9)$$

where  $IQ = \int_0^1 \sigma_u^4 du$ .

**Proof** see Appendix A.1 ■

The above result shows that QRV converges to the true integrated variance at rate  $N^{-1/2}$ , the best attainable in this setting. Moreover, it illustrates that the efficiency of QRV depends on the choice of  $m$  and  $\lambda$  through the functional  $\theta(m, \lambda)$  multiplying IQ in Eq. (9) above. To see the magnitude and properties of  $\theta(m, \lambda)$  we report its value for different  $m$  and  $\lambda$  in Table 1. Here, we use that the joint density of  $(U_{(m-\lambda m+1)}, U_{(\lambda m)})$  is given by:

$$f_{U_{(m-\lambda m+1)}, U_{(\lambda m)}}(x, y) = 1_{\{x < y\}} \frac{m!}{(m - \lambda m)!} \frac{(\Phi(y) - \Phi(x))^{2\lambda m - m - 2} ((1 - \Phi(y))\Phi(x))^{m - \lambda m}}{(2\lambda m - m - 2)!(m - \lambda m)!} \phi(x)\phi(y) \quad (10)$$

where  $\lambda \in (1/2, 1)$ ,  $\phi(\cdot)$  and  $\Phi(\cdot)$  are the density and distribution function of a standard normal random variable. Because  $\Phi(\cdot)$  appears in the joint density function, the expectation in Eq. (5) cannot be evaluated analytically but values for  $\nu_r(m, \lambda)$  and  $\theta(m, \lambda)$  can be obtained to any desired degree of accuracy by numerical integration, which is what we have done here.

Two important patterns emerge from Table 1 (the results for  $m = \infty$  are discussed in Section 2.3 below). Firstly, there is an “optimal” choice of quantile  $\lambda$  in the range 0.93 – 0.95 that minimizes  $\theta(m, \lambda)$  and therefore maximizes the (asymptotic) efficiency of the estimator. It is quite intuitive that quantiles close to the mode of the return distribution are not informative about the variance of the process. On the other hand, realizations of quantiles far into the tail of the distribution will be informative but also erratic. The “optimal” range of quantiles balances these trade-offs and extracts the maximum amount of available information regarding the spread of the distribution. Secondly, there is an “optimal” choice of subinterval length  $m$ , namely the smaller  $m$ , the greater the efficiency of QRV. Combining these two observations, we see that the best attainable asymptotic efficiency of QRV is  $\theta(20, 0.95) \approx 2.88$ .

To place this into context, consider two natural alternatives to QRV, namely realised variance (RV) (see, e.g., Andersen and Bollerslev, 1998; Andersen, Bollerslev, Diebold, and Labys, 2001; Barndorff-Nielsen and Shephard, 2002), defined

Table 1: Asymptotic efficiency with single pair of quantiles

$\lambda$	$m$					$\lambda$	$m$				
	20	50	100	500	$\infty$		20	50	100	500	$\infty$
0.80	4.243	4.298	4.312	4.321	4.323	0.90	3.102	3.143	3.154	3.162	3.163
0.81	–	–	4.143	4.150	4.151	0.91	–	–	3.101	3.112	3.114
0.82	–	3.978	3.987	3.993	3.994	0.92	–	3.043	3.062	3.077	3.080
0.83	–	–	3.844	3.848	3.849	0.93	–	–	3.041	3.062	3.066
0.84	–	3.706	3.712	3.716	3.717	0.94	–	3.004	3.041	3.072	3.079
0.85	3.555	–	3.593	3.597	3.596	0.95	2.878	–	3.071	3.116	3.127
0.86	–	3.477	3.484	3.487	3.487	0.96	–	3.064	3.146	3.217	3.234
0.87	–	–	3.385	3.389	3.389	0.97	–	–	3.293	3.413	3.444
0.88	–	3.290	3.297	3.302	3.302	0.98	–	3.345	3.582	3.819	3.884
0.89	–	–	3.220	3.226	3.226	0.99	–	–	4.212	4.885	5.099

*Note.* This table reports  $\theta(m, \lambda)$ , as given by Eq. (8), for different values of quantiles  $\lambda$  and observation per sub-interval  $m$ .

as:

$$RV_N = \sum_{i=1}^N (\Delta_i^N X)^2, \quad (11)$$

and the jump-robust bi-power variation (BPV) of Barndorff-Nielsen and Shephard (2004b):

$$BPV_N = \frac{\pi}{2} \sum_{i=2}^N |\Delta_i^N X| |\Delta_{i-1}^N X|. \quad (12)$$

Both quantities consistently estimate the integrated variance and we know from, e.g., Jacod (1994); Jacod and Protter (1998); Barndorff-Nielsen and Shephard (2002, 2004b) that:

$$\begin{aligned} \sqrt{N} (RV_N - IV) &\xrightarrow{d_s} MN(0, 2IQ), \\ \sqrt{N} (BPV_N - IV) &\xrightarrow{d_s} MN(0, \kappa IQ), \end{aligned}$$

where  $\kappa = \pi^2/4 + \pi - 3 \simeq 2.61$ . Comparing this to Eq. (9) we observe that QRV is somewhat less efficient than RV and BPV. However, we shall see in the next section that by constructing QRV with multiple pairs of quantiles, its efficiency can be improved substantially, making it more efficient than BPV and approximate the efficiency of RV while retaining robustness to jumps.

In practice, the choice of  $\lambda$  and  $m$  can be based on asymptotic efficiency grounds as discussed above, but there are two other important considerations, namely (i) QRV's robustness to jumps in finite sample, (ii) QRV's ability to track time

varying volatility and estimate the integrated variance. For fixed  $\lambda$ , increasing  $m$  makes QRV more robust to jumps: the first and last  $(1 - \lambda)m - 1$  ordered returns (which are likely to contain the jumps if any exist) are left out of the calculation of QRV. On the other hand, for QRV to accurately measure the integrated variance the return volatility is required to be locally constant over each subinterval and lowering  $m$  is therefore desirable. Hence, in practice, the particular choice of parameters should depend on the dynamics of the variance process (with an erratic variance process, lower  $m$ ) and the arrival rate of the jumps (with frequent jumps, increase  $m$  or lower  $\lambda$ ).

## 2.2 QRV with multiple pairs of quantiles

We now extend QRV to incorporate multiple pairs of quantiles. Consider a  $(k \times 1)$  vector of quantiles  $\bar{\lambda} = (\lambda_1, \dots, \lambda_k)'$  where  $\lambda_i \in (1/2, 1)$  for  $i = 1, \dots, k$  and define

$$QRV_N(m, \bar{\lambda}, \alpha) = \alpha' QRV_N(m, \bar{\lambda}) \quad (13)$$

where  $\alpha$  is a  $(k \times 1)$  weighting vector and  $QRV_N(m, \bar{\lambda})$  is a  $(k \times 1)$  vector with  $i$ th entry equal to  $QRV_N(m, \lambda_i)$ .

**Theorem 3** *For the process in Eq. (1), with condition (V) satisfied, and  $N = mn$  with  $m$  fixed, we have as  $N \rightarrow \infty$*

$$\sqrt{N} (QRV_N(m, \bar{\lambda}, \alpha) - IV) \xrightarrow{d_s} MN(0, \theta(m, \bar{\lambda}, \alpha) IQ), \quad (14)$$

where  $\theta(m, \bar{\lambda}, \alpha) = \alpha' \Theta(m, \bar{\lambda}) \alpha$  and the  $k \times k$  matrix  $\Theta(m, \lambda) = (\Theta(m, \lambda)_{sl})_{1 \leq s, l \leq k}$  is given by

$$\Theta(m, \bar{\lambda})_{ij} = m \frac{\nu_1(m, \lambda_i, \lambda_j) - \nu_1(m, \lambda_i) \nu_1(m, \lambda_j)}{\nu_1(m, \lambda_i) \nu_1(m, \lambda_j)},$$

and

$$\nu_1(m, \lambda_i, \lambda_j) = \mathbb{E}[ (|U_{(m\lambda_i)}|^2 + |U_{(m-m\lambda_i+1)}|^2)(|U_{(m\lambda_j)}|^2 + |U_{(m-m\lambda_j+1)}|^2) ].$$

**Proof** see Appendix A.1 ■

From the above result it is clear that, for a given  $\bar{\lambda}$ , optimal weights can be assigned to the quantiles when calculating QRV, i.e.

$$\alpha^* = \frac{\Theta^{-1}(m, \bar{\lambda}) \iota}{\iota' \Theta^{-1}(m, \bar{\lambda}) \iota}. \quad (15)$$

where  $\iota$  is a  $(k \times 1)$  vector of ones. In this case, QRV attains the best efficiency with  $\theta(m, \bar{\lambda}, \alpha^*) = 1/(\iota' \Theta^{-1}(m, \bar{\lambda}) \iota)$ . As before, it is of interest to understand the magnitude of  $\theta(m, \bar{\lambda}, \alpha^*)$ , keeping in mind that the corresponding values for RV and BPV are 2 and 2.609 respectively. In Table 2 we consider a couple of different scenarios regarding the choice of quantiles (for  $m = 100$ ). For instance, using three (pairs of) “low” quantiles  $\bar{\lambda} = \{0.80; 0.85; 0.90\}$  we have an optimal weights vector  $\alpha^* = \{0.23; 0.15; 0.62\}$  and QRV attains an efficiency of 2.87. It is intuitive that a larger number of

Table 2: Optimal weights and asymptotic efficiency with multiple pairs of quantiles

$\lambda$	optimal weights $\alpha^*$					$\theta(m, \bar{\lambda}, \alpha^*)$
	0.80	0.85	0.90	0.95	0.99	
low quantiles	0.23	0.15	0.62	–	–	2.868
medium quantiles	–	0.30	0.21	0.49	–	2.471
high quantiles	–	–	0.42	0.28	0.30	2.294
all quantiles	0.17	0.11	0.18	0.26	0.28	2.139

*Note.* This table reports  $\theta(m, \lambda, \alpha)$ , as specified in Theorem 3, for different combinations of quantiles  $\lambda$ . Optimal weights  $\alpha^*$  are used and  $m = 100$ .

quantiles improves the efficiency of QRV and that the higher quantiles are more informative about IV. The most efficient estimates of the IV, however, are achieved from a mix of low to high quantiles, due to the correlation structure of the order statistics. In particular, with  $k = 5$  and quantiles as listed in Table 2, QRV attains an efficiency of 2.14. This figure can be pushed arbitrarily close to 2 by adding more quantiles, so that in the limit QRV attains the ML efficiency of RV.

### 2.3 QRV with $m \rightarrow \infty$

In the above, we kept  $m$  fixed while letting  $n \rightarrow \infty$ , i.e. as  $n$  grows, the subinterval shrinks but the number of returns contained in each subinterval stays constant (consequently, returns are measured at increasingly high frequency). Now we consider the case where  $m \rightarrow \infty$ .

**Proposition 2** *With constant volatility, i.e.  $\sigma_t = \sigma$  for  $t \in [0, 1]$ , Theorem 1, 2, and 3 remain valid for  $m \rightarrow \infty$  and  $n = 1$ , with*

$$\begin{aligned}
 \nu_1(\lambda) &\equiv \lim_{m \rightarrow \infty} \nu_1(m, \lambda) = 2c_\lambda^2, \\
 \nu_1(\lambda_i, \lambda_j) &\equiv \lim_{m \rightarrow \infty} \nu_1(m, \lambda_i, \lambda_j) = 4c_{\lambda_i}^2 c_{\lambda_j}^2, \\
 \theta(\lambda) &\equiv \lim_{m \rightarrow \infty} \theta(m, \lambda) = 2 \frac{(1 - \lambda)(2\lambda - 1)}{\phi^2(c_\lambda) c_\lambda^2}, \\
 \Theta(\bar{\lambda})_{ij} &\equiv \lim_{m \rightarrow \infty} \Theta(m, \bar{\lambda})_{ij} = 2 \frac{(1 - \lambda_j)(2\lambda_i - 1)}{\phi(c_{\lambda_i}) \phi(c_{\lambda_j}) c_{\lambda_i} c_{\lambda_j}},
 \end{aligned}$$

where  $c_\alpha$  denotes the  $\alpha$ -quantile of the standard normal distribution, and  $\phi$  is the density function of the standard normal distribution.

We emphasize that this special case where volatility is constant with  $n = 1$  and  $m \rightarrow \infty$  is analogous to David (1970), and the estimator proposed by David (1970) is equivalent to QRV with the only difference being that the latter estimates

the variance rather than the volatility. In Table 1 we report the asymptotic efficiency coefficient  $\theta(\lambda)$  of QRV with a single pair of quantiles when  $m \rightarrow \infty$ . It is interesting to note that there is only limited efficiency loss associated with the use of a single subsample, i.e.  $n = 1$  and  $m = N$ . For instance, with  $\lambda = 0.95$ , the best attainable efficiency is  $\theta(m, \lambda) \simeq 2.88$  with  $m = 20$  and  $\theta(\lambda) \simeq 3.13$  with  $m \rightarrow \infty$ . Thus, from a practical viewpoint, one may compute QRV with  $m = N$  and obtain efficient variance estimates based on merely 2 (or  $2k$  for the multi-quantile version) ordered returns!

If we want to relax the constant volatility assumption, it is still possible to have  $m \rightarrow \infty$  but we now need to let  $n$  grow for consistency of QRV and a CLT. Moreover, much stronger assumptions on the dynamics of  $X$  are required. We want to stress that the intuition behind the proofs of the results presented below differs from the “finite  $m$ ” case (and also from Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard, 2006; Christensen and Podolskij, 2006).

The crucial condition for consistency (and the CLT below) is

$$m = \delta N^\gamma, \tag{16}$$

for some  $\delta > 0$  and  $\gamma \in (0, 1/2)$ . In particular, condition (16) implies that  $m/n \rightarrow 0$  (as  $n = \delta^{-1} N^{1-\gamma}$ ).

**Theorem 4** *For the process in Eq. (1), with condition (V) and (16) satisfied, as  $N \rightarrow \infty$  we have convergence in probability*

$$QRV_N(m, \lambda) \xrightarrow{p} IV.$$

**Proof** see Appendix A.2 ■

Note that Theorem 4 requires conditions (V) and (16), whereas Theorem 1 holds under very weak assumptions on  $a$  and  $\sigma$ . To deliver a CLT, we need further smoothness conditions on the processes  $a$ ,  $\sigma'$  and  $v'$ .

**Assumption (V')**  $a$ ,  $\sigma'$  and  $v'$  satisfy condition (V)

**Theorem 5** *For the process in Eq. (1), with condition (V') and (16) satisfied, and  $\gamma \in (0, 1/4)$ ,  $\lambda \in (1/2, 1)$ , as  $N \rightarrow \infty$  we have stable convergence*

$$\sqrt{N} (QRV_N(m, \lambda) - IV) \xrightarrow{d_s} MN(0, \theta(\lambda) IQ),$$

where  $\theta(\lambda) \equiv \lim_{m \rightarrow \infty} \theta(m, \lambda)$  is as defined in Proposition 2,  $W'$  is another Brownian motion defined on an extension of  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , and  $W' \perp\!\!\!\perp \mathcal{F}$ .

**Proof** see Appendix A.2 ■

Next, we state the corresponding result for the multi-quantile extension.

**Theorem 6** For the process in Eq. (1), with condition (V') and (16) satisfied, and  $\gamma \in (0, 1/4)$ ,  $\lambda \in (1/2, 1)$ , as  $N \rightarrow \infty$  we have stable convergence

$$\sqrt{N} (QRV_N(m, \bar{\lambda}, \alpha) - IV) \xrightarrow{d_s} MN(0, \theta(\bar{\lambda}, \alpha)IQ), \quad (17)$$

where  $\theta(\bar{\lambda}, \alpha) = \alpha' \Theta(\bar{\lambda}) \alpha$  and the  $k \times k$  matrix  $\Theta(\lambda) = (\Theta(\lambda)_{sl})_{1 \leq s, l \leq k}$  is as defined in Proposition 2.

**Proof** see Appendix A.2 ■

As before, optimal weights  $\alpha^*$  can be assigned to the vector of quantiles  $\bar{\lambda}$ , with

$$\alpha^* = \frac{\Theta^{-1}(\lambda)_l}{l' \Theta^{-1}(\lambda)_l}. \quad (18)$$

## 2.4 Finite sample performance and comparison to realised variance and bi-power variation

The above results show that the asymptotic efficiency of QRV, particularly with multiple quantiles, is excellent: it compares favorably to BPV and approximates the efficiency of RV for a suitable choice of parameters. The simulations below are designed to gauge the finite sample performance of QRV. We pay particular attention to bias, efficiency, and robustness to jumps. For comparison, we include RV and BPV in the analysis.

To simulate the log price process, we adopt the following model:

$$dX_t = \sigma_t dW_t, \quad t \in [0, 1] \quad (19)$$

where  $W$  is a standard Brownian motion and the dynamics of  $\sigma_t$  are as specified below. The baseline scenario is a constant volatility Brownian motion (“BM”), i.e.

$$\sigma_t^2 = 0.0391. \quad (20)$$

To assess QRV’s ability to handle time-varying volatility, we use a Heston-type stochastic volatility (“SV”) model

$$d\sigma_t^2 = (0.3141 - 8.0369\sigma_t^2)dt + \sigma_t \sqrt{0.1827} dB_t. \quad (21)$$

where the Brownian motion  $B$  is independent of  $W$ . To gauge the impact of leverage, we also simulate from Eq. (21) with  $dW_t dB_t = -0.75dt$  (“SV-LEV”). Finally, we consider two more variance specification that are both capable of generating erratic and highly volatile sample paths, even at the intra-day frequencies considered here. The first is a model proposed by Aït-Sahalia (1996) that incorporates stochastic elasticity of variance and non-linear drift (“SEV-ND”), i.e.

$$d\sigma_t^2 = (-0.554 + 21.32\sigma_t^2 - 209.3\sigma_t^4 + 0.005\sigma_t^{-2})dt + \sqrt{0.017\sigma_t^2 + 53.97\sigma_t^{5.76}} dB_t. \quad (22)$$

with  $B$  independent of  $W$ . The second is a two-factor stochastic volatility model (“SV2F-LEV”) analyzed in Chernov, Gallant, Ghysels, and Tauchen (2003), i.e.

$$\begin{aligned}\sigma_t^2 &= s\text{-exp}(-1.2 + 0.04f_t^{(1)} + 1.5f_t^{(2)}), \\ df_t^{(1)} &= -0.000137f_t^{(1)}dt + dB_t^{(1)}, \\ df_t^{(2)} &= -1.386f_t^{(2)}dt + (1 + 0.25f_t^{(2)})dB_t^{(2)},\end{aligned}\tag{23}$$

where  $dW_t dB_t^{(1)} = dW_t dB_t^{(2)} = -0.3dt$  and  $s\text{-exp}$  denotes a “spliced” exponential function as specified and discussed in Chernov, Gallant, Ghysels, and Tauchen (2003).

The above stochastic volatility models cover a wide range of dynamic specifications and thus provide a good testing ground for QRV. The parameter values for the BM, SV, and SEV-ND models in Eqs. (20–22) are taken from the empirical study by Bakshi, Ju, and Ou-Yang (2006) and, as such, represent realistic values. The parameters for the SV2F-LEV model in Eq. (23) are taken from Huang and Tauchen (2005) and are also empirically plausible.

To study robustness of QRV to jumps we also simulate from the BM model and add jumps. In particular, we add a fixed number of  $n_J$  Gaussian jumps at random points in the sample with a combined jump variation  $v_J$  measured as a fraction of diffusive variation  $IV$ . We consider three scenarios:  $\{n_J, v_J\} = \{1, \frac{1}{4}\}$ , i.e. one large jump accounting for 20% of total variation,  $\{n_J, v_J\} = \{5, \frac{1}{4}\}$ , i.e. five small jumps accounting for 20% of total variation, and  $\{n_J, v_J\} = \{5, \frac{1}{2}\}$ , i.e. 5 medium-sized jumps accounting for one-third of total variation.

To simulate the process in Eq. (19), we use an Euler discretization scheme and set  $N = 1,000$ . RV and BPV are computed as in Eqs. (11) and (12) respectively. QRV is computed as in Eq. (13) using three pairs of quantiles  $\bar{\lambda} = \{0.85; 0.90; 0.95\}$ , optimal weights  $\alpha^*$  as in Eq. (15), and three different choices of subinterval length, namely  $m = 20, 100$ , and  $1000(\infty)$ . Over 100,000 simulation runs, we then compute a “bias” measure  $\mathbb{E}(\widehat{IV}/IV)$  and an “efficiency” measure  $\text{var}(\sqrt{N}(\widehat{IV} - IV)/\sqrt{IQ})$  where  $\widehat{IV} = \{RV_N, BPV_N, QRV_N\}$ . If the estimator is unbiased we expect the bias statistic to be one. Moreover, according to the asymptotic results presented above, the efficiency statistic should be equal to 2 for RV,  $\pi^2/4 + \pi - 3 \approx 2.609$  for BPV, and 2.456, 2.471, 2.472 for QRV with  $m = 20, 100, \infty$  respectively.

From the results in Table 3 several interesting patterns emerge. First consider the scenarios without jumps. With model BM, all estimators (including RV and BPV) perform as expected. They are unbiased and their efficiency measure is close to what the asymptotic distribution theory predicts, indicating that it affords a good approximation to finite sample performance. The efficiency of RV is best, and the efficiency of QRV is comparable to that of BPV. When introducing stochastic volatility through model SV, we find that QRV is biased downwards when few subsamples are selected (i.e.  $m$  large). However, this bias is small for  $m = 100$  and negligible for  $m = 20$ . Leverage (SV-LEV) does not have a noticeable impact on any of the results. This is not surprising since our QRV theory allows for leverage and so does RV and BPV (see Barndorff-Nielsen, Graversen, Jacod, and Shephard, 2006). With the SEV-ND and SV2F-LEV variance

Table 3: Performance of QRV with stochastic volatility and jumps

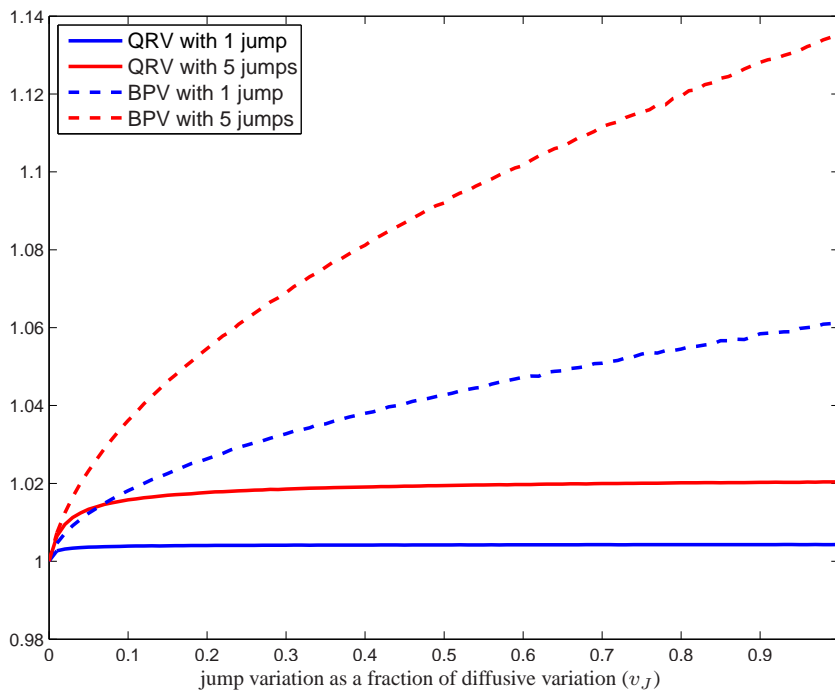
model	QRV with $\lambda = \{0.85; 0.90; 0.95\}$			RV	BPV
	$m = 20$	$m = 100$	$m = \infty$		
<i>Panel A: “bias” measure <math>\mathbb{E}(\widehat{IV}/IV)</math></i>					
BM	1.000	1.000	0.996	1.000	1.000
SV	0.997	0.985	0.950	1.000	0.999
SV-LEV	0.997	0.985	0.950	1.000	0.999
SEV-ND	0.997	0.986	0.938	1.000	0.999
SV2F-LEV	0.998	0.990	0.937	0.999	0.999
BMJ( $n_J = 1, v_J = \frac{1}{4}$ )	1.005	1.004	1.000	1.251	1.030
BMJ( $n_J = 5, v_J = \frac{1}{4}$ )	1.025	1.019	1.014	1.250	1.063
BMJ( $n_J = 5, v_J = \frac{1}{2}$ )	1.031	1.020	1.015	1.501	1.093
<i>Panel B: “efficiency” measure <math>\text{var}(\sqrt{N}(\widehat{IV} - IV)/\sqrt{IQ})</math></i>					
BM	2.465	2.484	2.469	2.006	2.617
SV	2.446	2.390	2.336	2.004	2.604
SV-LEV	2.452	2.382	2.286	2.000	2.603
SEV-ND	2.438	2.357	2.464	2.000	2.599
SV2F-LEV	2.442	2.400	2.835	1.998	2.593
BMJ( $n_J = 1, v_J = \frac{1}{4}$ )	2.497	2.505	2.488	129.674	3.660
BMJ( $n_J = 5, v_J = \frac{1}{4}$ )	3.222	2.591	2.562	28.046	3.806
BMJ( $n_J = 5, v_J = \frac{1}{2}$ )	5.189	2.591	2.566	104.135	5.227

*Note.* This table reports the bias and efficiency measure for QRV, BPV, and RV under various model specifications for  $N = 1,000$ . The bias measure in Panel A is equal to 1 for an unbiased IV estimator. The efficiency measure in Panel B takes on a minimum attainable value of 2 for the MLE.

specifications, both generating high volatility-of-volatility, the QRV estimator still performs well provided that a sufficient number of sub-intervals are selected, i.e. set  $m$  low and  $n$  high in line with discussion above. Importantly, in all of the above cases, the efficiency measure is also close to its asymptotic counterpart.

Next, consider the scenarios with jumps. It is well known that RV estimates total variation, i.e.  $(1 + v_J)IV$  in this setting, explaining the bias and low efficiency when evaluated against IV. BPV is asymptotically immune to jumps, but biased in finite samples: for the BMJ model considered here  $\mathbb{E}(BPV_N/IV) \simeq 1 + 2\sqrt{n_J v_J/N}$ . This explains the

Figure 1: Bias of QRV compared to BPV



Note. QRV is implemented using  $\bar{\lambda} = \{0.85; 0.90; 0.95\}$  with  $m = 100$  for  $N = 1000$ .

observed bias in BPV and deterioration of efficiency measure when jumps become more frequent and volatile. Turning to QRV, we observe superior robustness to jumps. As pointed out before, for given  $\bar{\lambda}$ , increasing  $m$  makes QRV more robust to jumps. This is exactly what we see in Table 3. For instance, with  $n_J = 5$  and  $v_J = \frac{1}{4}$ , QRV's bias is about 2.5% when  $m = 20$  but decreases to 1.4% with  $m = N$ .

To further illustrate the behavior of QRV with jumps, we plot its bias as a function of the jump variation parameter  $v_J$  in Figure 1. Here we use  $m = 100$  and include BPV for comparison. The observed pattern is striking: the bias of QRV grows with  $v_J$ , but only up to a point, beyond which the estimator is immune to further increases in jump variation. The intuition for this is subtle and best explained with an example. Consider a ranked sequence of diffusive returns  $\{r_{(1)}, r_{(2)}, \dots, r_{(m)}\}$  from the BM model. With  $\lambda = 0.95$  we can estimate IV unbiasedly using  $r_{(5)}$  and  $r_{(95)}$  as described above. Now suppose a large positive jump  $J$  is added that affects, say,  $r_{(60)}$ . For a sufficiently large jump, the ordered returns sequence is now  $\{r_{(1)}, r_{(2)}, \dots, r_{(59)}, r_{(61)}, r_{(m)}, r_{(60)} + J\}$  and the “realised”  $\lambda = 0.95$  quantile is now  $r_{(96)}$  and leads to a bias in QRV. Thus, when jumps are added to the diffusive price process, the original ordering of returns can be disrupted, leading to a bias in realised quantiles and QRV as a consequence. However, as the comparison to BPV illustrates, the bias has a modest magnitude.

### 3 QRV with market microstructure noise

It has long been recognized that market microstructure effects in high frequency data – such a bid-ask bounce and non-synchronous trading – distort the statistical properties of returns (e.g. Epps, 1979; Fisher, 1966; Niederhoffer and Osborne, 1966) and are detrimental to realised variance as an estimator of the integrated variance (e.g. Zhou, 1996). In this section we develop a modified version of the quantile-based realised variance that is robust to noise and delivers consistent estimates of the integrated variance.

On a filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , we consider the noisy diffusion model

$$Y_{i/N} = X_{i/N} + u_i, \quad (24)$$

for  $i = 0, 1, \dots, N$ . Here, the “efficient” price  $X$  is a Brownian semimartingale as in Eq. (1). The microstructure noise  $u$  is an i.i.d. process, independent of  $X$ , with

$$\mathbb{E}(u_i) = 0, \quad \mathbb{E}(u_i^2) = \omega^2. \quad (25)$$

The i.i.d. assumption on  $u$  is the natural starting point to analyze the noise case and is widely used in the literature. Moreover, it has some empirical support at moderate sampling frequencies (see e.g. Hansen and Lunde, 2006). See Diebold and Strasser (2007) for further discussion of this assumption. As before, the object of econometric interest is the integrated variance of  $X$ , with the additional challenge that inference is now based on discretely sampled and noisy data. The approach taken here is to compute a modified QRV measure using data that is filtered following a procedure used by Podolskij and Vetter (2006). We show that, despite the noise, QRV from filtered data is consistent, has good efficiency, and retains robustness to jumps.

#### 3.1 Noise robust data sampling

We choose a natural number  $K = K(N)$  with

$$K = cN^{1/2}, \quad (26)$$

for some constant  $c > 0$ , and define the return-like statistics

$$\bar{Y}_j^N = \frac{1}{K} \left( \sum_{i=0}^{K-1} Y_{(j+i+K)/N} - \sum_{i=0}^{K-1} Y_{(j+i)/N} \right). \quad (27)$$

An equivalent representation, which is convenient for the proof, is

$$\bar{Y}_j^N = 2 \sum_{i=1}^{2K-1} h\left(\frac{i}{2K}\right) \Delta_{j+i}^N Y, \quad h(x) = x \wedge (1-x), \quad x \in [0, 1]. \quad (28)$$

Next, we select a sub-sequence using data observed in the interval  $[i/N, (i + 2Km - 1)/N]$ :

$$\overline{D}_i^N Y = \{\overline{Y}_{i+2(j-1)K}^N\}_{j=1}^m, \quad (29)$$

and compute

$$q_i^*(m, \lambda) = g_{\lambda m}^2(N^{1/4}\overline{D}_i^N Y) + g_{m-\lambda m+1}^2(N^{1/4}\overline{D}_i^N Y).$$

The noise-adjusted quantile-based realised variance (QRV\* hereafter) is then defined as:

$$QRV_N^*(m, \lambda) = \frac{1}{\nu_1(m, \lambda)} \frac{3}{2K\sqrt{N}} \sum_{i=0}^{N-2Km+1} q_i^*(m, \lambda). \quad (30)$$

Before we present the asymptotic properties of  $QRV^*$ , we give some intuition for the outlined filtering procedure in Eqs. (26 – 29). As  $X$  is a continuous process and  $K \rightarrow \infty$ , we obtain by averaging that

$$\frac{1}{K} \sum_{i=0}^{K-1} Y_{(j+i)/N} \approx X_{j/N}. \quad (31)$$

Thus,  $N^{1/4}\overline{Y}_j^N$  “plays the role” of  $\sqrt{N}(X_{i/N} - X_{(i-1)/N})$  in the pure diffusion model. The constant  $K$  controls the stochastic order of the term  $\overline{Y}_j^N$ . In particular,

$$\overline{u}_j^{(K)} = O_p\left(\sqrt{\frac{1}{K}}\right), \quad (32)$$

$$\overline{X}_j^{(K)} = O_p\left(\sqrt{\frac{K}{N}}\right). \quad (33)$$

Thus, when  $K$  is chosen as in (26) the stochastic orders of the quantities in (32) and (33) are balanced (this implies the best rate of convergence), and under mild conditions we have

$$N^{1/4}\overline{Y}_j^N | \mathcal{F}_{j/N} \overset{asy}{\sim} N \left(0, \frac{2c}{3}\sigma_{j/N}^2 + \frac{2}{c}\omega^2\right). \quad (34)$$

This fact demonstrates again that – although it is affected by the noise process through  $\omega^2 - N^{1/4}\overline{Y}_j^N$  “behaves” like  $\sqrt{N}(X_{i/N} - X_{(i-1)/N})$ , which provides the rationale for the above filtering procedure.

### 3.2 Asymptotic properties

Our first result shows the consistency of the noise corrected QRV.

**Theorem 7** *Assume that  $m$  is a fixed number. As  $N \rightarrow \infty$ , it holds that*

$$QRV_N^*(m, \lambda) - \frac{3}{c^2}\omega^2 \xrightarrow{p} IV. \quad (35)$$

**Proof** see Appendix A.1 ■

In practice, we can construct consistent estimates of the noise variance, e.g.  $\hat{\omega}^2 = \frac{1}{2N} \sum_{i=1}^N |Y_{i/N} - Y_{(i-1)/N}|^2$  as in Bandi and Russell (2006), or a jump robust version  $\hat{\omega}^2 = -\frac{1}{N} \sum_{i=1}^{N-1} (Y_{(i+1)/N} - Y_{i/N})(Y_{i/N} - Y_{(i-1)/N})$  as in Oomen (2006a,b). Consequently, we immediately obtain the convergence

$$QRV_N^*(m, \lambda) - \frac{3}{c^2} \hat{\omega}^2 \xrightarrow{p} IV. \quad (36)$$

Before we state the CLT we introduce the following notation.

**Definition 2** For  $x \in \mathbb{R}$ ,  $u \in [0, 1]$ ,  $l = 1, \dots, m$  and  $\lambda_1, \lambda_2$  we define the quantity

$$f_{m,l,x,u}(\lambda_1, \lambda_2) = \text{cov} \left( g_{\lambda_1 m}^2(S) + g_{m-\lambda_1 m+1}^2(S), g_{\lambda_2 m}^2(T) + g_{m-\lambda_2 m+1}^2(T) \right), \quad (37)$$

where  $S = (S_1, \dots, S_m)^T$ ,  $T = (T_1, \dots, T_m)^T$  are centered and jointly normal with

- (i)  $S_i \perp S_j, T_i \perp T_j$  for all  $i \neq j$ .
- (ii)  $\text{var}(S_i) = \text{var}(T_i) = 8cw_h(0)x^2 + \frac{2}{c}w_{\nabla h}(0)\omega^2$  for all  $i$ .
- (iii)  $\text{cov}(S_{i+l-1}, T_i) = 8cw_h(u)x^2 + \frac{2}{c}w_{\nabla h}(u)\omega^2$  for all  $i$ .
- (iv)  $\text{cov}(S_{i+l}, T_i) = 8cw_h(1-u)x^2 + \frac{2}{c}w_{\nabla h}(1-u)\omega^2$  for all  $i$ .
- (v)  $\text{cov}(S_i, T_j) = 0$  for all  $|i+l-j-1| > 1$ .

Here the function  $h$  is given in (28) and  $w_h(u)$  is defined by

$$w_h(u) = \int_0^{1-u} h(y) h(y+u) dy. \quad (38)$$

When  $\lambda = \lambda_1 = \lambda_2$  we use the notation  $f_{m,l,x,u}(\lambda) = f_{m,l,x,u}(\lambda_1, \lambda_2)$ .

Notice that  $\nabla h$  (the derivative of  $h$ ) exists almost everywhere, so the quantity  $w_{\nabla h}$  makes sense.

We need an extra condition to ensure that both  $X$  and  $u$  are measurable with respect to the same type of filtration, which allows us to build on existing CLTs for high-frequency data (see, e.g., Jacod and Shiryaev, 2003). In particular, we require that  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  supports another Brownian motion  $B = (B_t)_{t \in [0,1]}$  with  $B \perp\!\!\!\perp X$ , such that the representation

$$u_i = \sqrt{N} (B_{i/N} - B_{(i-1)/N}) \omega \quad (39)$$

holds.

**Theorem 8** Assume that  $m$  is a fixed number and the conditions (V) and (39) are satisfied. As  $N \rightarrow \infty$  the statistic

$$N^{1/4} \left( QRV_N^*(m, \lambda) - \frac{3}{c^2} \hat{\omega}^2 - IV \right)$$

converges stably in law towards a random variable, defined on an extension of the original probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , which is Gaussian conditionally on  $\mathcal{F}$  with mean 0 and variance

$$\frac{9}{\nu_{1,m}^2(\lambda) c} \sum_{l=1}^m \int_0^1 \int_0^1 f_{m,l,\sigma_t,u}(\lambda) dt du. \quad (40)$$

**Proof** see Appendix A.1 ■

The rate of convergence in Theorem 8 is  $N^{-1/4}$ , which is known to be optimal in the noisy diffusion model (Gloter and Jacod, 2001a,b). Even though there is no explicit expression for the conditional variance in the above CLT, it is possible to estimate it from the data.

**Proposition 3** Assume that  $m$  is fixed and  $\mathbb{E}(u_i^4) < \infty$ . As  $N \rightarrow \infty$ , it holds that

$$\begin{aligned} & \frac{9}{4\nu_1^2(m, \lambda) K \sqrt{N}} \sum_{i=2Km-1}^{N-4Km+1} q_i^*(m, \lambda) \left( \sum_{j=i-2Km+1}^{i+2Km-1} \{q_j^*(m, \lambda) - q_{i+2Km}^*(m, \lambda)\} \right) \\ & \xrightarrow{p} \frac{9}{\nu_{1,m}^2(\lambda) c} \sum_{l=1}^m \int_0^1 \int_0^1 f_{m,l,\sigma_t,u}(\lambda) dt du. \end{aligned} \quad (41)$$

**Proof** see Appendix A.1 ■

Next, we state the corresponding result for the multi-quantile implementation.

**Theorem 9** Assume that the conditions of Theorem 8 are satisfied. As  $N \rightarrow \infty$ , we have the stable convergence

$$N^{1/4} \left( QRV_N^*(m, \bar{\lambda}, \alpha) - \frac{3}{c^2} \hat{\omega}^2 - IV \right) \xrightarrow{d_s} MN \left( 0, \frac{9}{c} \Sigma_m(\lambda_1, \dots, \lambda_k) \right), \quad (42)$$

where  $\Sigma_m(\lambda_1, \dots, \lambda_k) = (\Sigma_m(\lambda_1, \dots, \lambda_k)_{sl})_{1 \leq s, l \leq k}$  is given by

$$\Sigma_m(\lambda_1, \dots, \lambda_k)_{sl} = \frac{1}{\nu_{1,m}(\lambda_s) \nu_{1,m}(\lambda_l)} \sum_{l=1}^m \int_0^1 \int_0^1 f_{m,l,\sigma_t,u}(\lambda_s, \lambda_l) dt du. \quad (43)$$

**Proof** see Appendix A.1 ■

**Proposition 4** Assume that  $m$  is fixed and  $\mathbb{E}(u_i^4) < \infty$ . As  $N \rightarrow \infty$ , it holds that

$$\frac{9}{4\nu_1^2(m, \lambda) K \sqrt{N}} \sum_{i=2Km-1}^{N-4Km+1} q_i^*(m, \lambda_s) \left( \sum_{j=i-2Km+1}^{i+2Km-1} \{q_j^*(m, \lambda_l) - q_{i+2Km}^*(m, \lambda_l)\} \right) \xrightarrow{p} \frac{9}{c} \Sigma_m(\lambda_1, \dots, \lambda_k)_{sl}. \quad (44)$$

**Proof** see Appendix A.1 ■

Podolskij and Vetter (2007) show that when IV is estimated using a “sum-of-squares” estimator based on filtered data, i.e.

$$\frac{3}{2K\sqrt{N}} \sum_{j=1}^{N-2K} |\bar{Y}_j^N|^2, \quad (45)$$

the lowest attainable variance is  $8.5\sigma^3\omega$  assuming  $\sigma$  is constant (the variance of the maximum likelihood estimator is  $8\sigma^3\omega$ ). Consequently, the natural lower bound for the variance of the QRV based on filtered data is also  $8.5\sigma^3\omega$ . We note that, for suitable choice of parameters, the realised kernel approach of Barndorff-Nielsen, Hansen, Lunde, and Shephard (2006) can attain an variance of  $8.002\sigma^3\omega$ .

### 3.3 Comparison to microstructure robust realised variance measures

In the simulations below, we investigate the finite sample performance of the “noise-adjusted” QRV. In particular, we simulate the efficient price using the BM model in Eqs (19 – 20) with  $N = 10,000$  and add i.i.d. noise as in Eq. (24) with  $\omega = \xi IV/N$ . Here, the parameter  $\xi$  controls the level of noise and we consider two scenarios, namely (i) moderate noise with  $\xi = \frac{1}{4}$ , and (ii) high level of noise with  $\xi = 1$ . To implement QRV, we use the quantiles as before, i.e.  $\bar{\lambda} = \{0.85; 0.90; 0.95\}$ . The sampling parameter  $K$  that controls the averaging horizon in Eq. (27) is varied between  $K = 2$  and  $K = 15$ . Because we simulate the process with constant volatility, we set  $m = \lfloor N/(2K) \rfloor$  (or  $n = 1$ ) for simplicity and computational speed.

For comparison, we compute the multi-scale realised variance (MSRV) of Zhang (2006). Like QRV, this estimator is consistent for IV and converges at rate  $N^{-1/4}$ . We note that there are a number of closely related alternative estimators available, including two-scale realised variance of Zhang, Mykland, and Ait-Sahalia (2005) and realised kernels of Barndorff-Nielsen, Hansen, Lunde, and Shephard (2006). Because in the current setting their performance is very similar to that of MSRV we exclude these results to save space. The MSRV based on  $q$ -subsamples is defined as:

$$MSRV_N(q) = \sum_{j=1}^q \frac{a_j}{j} \sum_{h=0}^{j-1} \gamma_{h,j}(0), \quad (46)$$

where

$$\gamma_{h,q} = \sum_{i=1}^N (Y_{iq+h} - Y_{(i-1)q+h})^2,$$

and

$$a_j^* = (1 - 1/q^2)^{-1} \left( \frac{j}{q^2} h(j/q) - \frac{j}{2q^3} h'(j/q) \right),$$

for  $h(x) = 12(x - 1/2)$ . The optimal number of subsamples can be calculated as  $q_Z^* = c^* \sqrt{N}$  where

$$c^* = \arg \min_c \left\{ 2 \frac{52}{35} c IQ + \frac{48}{5} c^{-1} \omega^2 (IV + \omega^2/2) + 48 c^{-3} \omega^4 \right\}. \quad (47)$$

Table 4: Performance of QRV\* with microstructure noise ( $N = 10,000$ )

	QRV* on noise-filtered returns				raw returns		
	$K = 2$	$K = 5$	$K = 10$	$K = 15$	QRV	RV	MSRV
<i>Panel A: “bias” measure <math>\mathbb{E}(\widehat{IV}/IV)</math></i>							
$\omega^2 = \frac{1}{4}IV/N$	1.126	1.018	1.006	0.995	1.445	1.500	1.000
$\omega^2 = IV/N$	1.126	1.018	1.006	0.994	2.896	3.000	0.999
<i>Panel B: “efficiency” measure <math>\text{var}(N^{1/4}(\widehat{IV} - IV)/\sqrt{\omega\sigma^3})</math></i>							
$\omega^2 = \frac{1}{4}IV/N$	13.866	24.878	48.114	70.508	10.871	9.513	16.245
$\omega^2 = IV/N$	12.615	13.538	24.571	35.574	24.680	22.064	11.809

*Note.* The bias measure in Panel A is equal to 1 for an unbiased IV estimator. The efficiency measure in Panel B takes on a minimum attainable value of 8 for the MLE.

In the simulations, we use the optimal bandwidth to calculate MSRV (for  $\xi = \frac{1}{4}$  we have  $q_Z^* = 3$  and for  $\xi = 1$  we have  $q_Z^* = 4$ ).

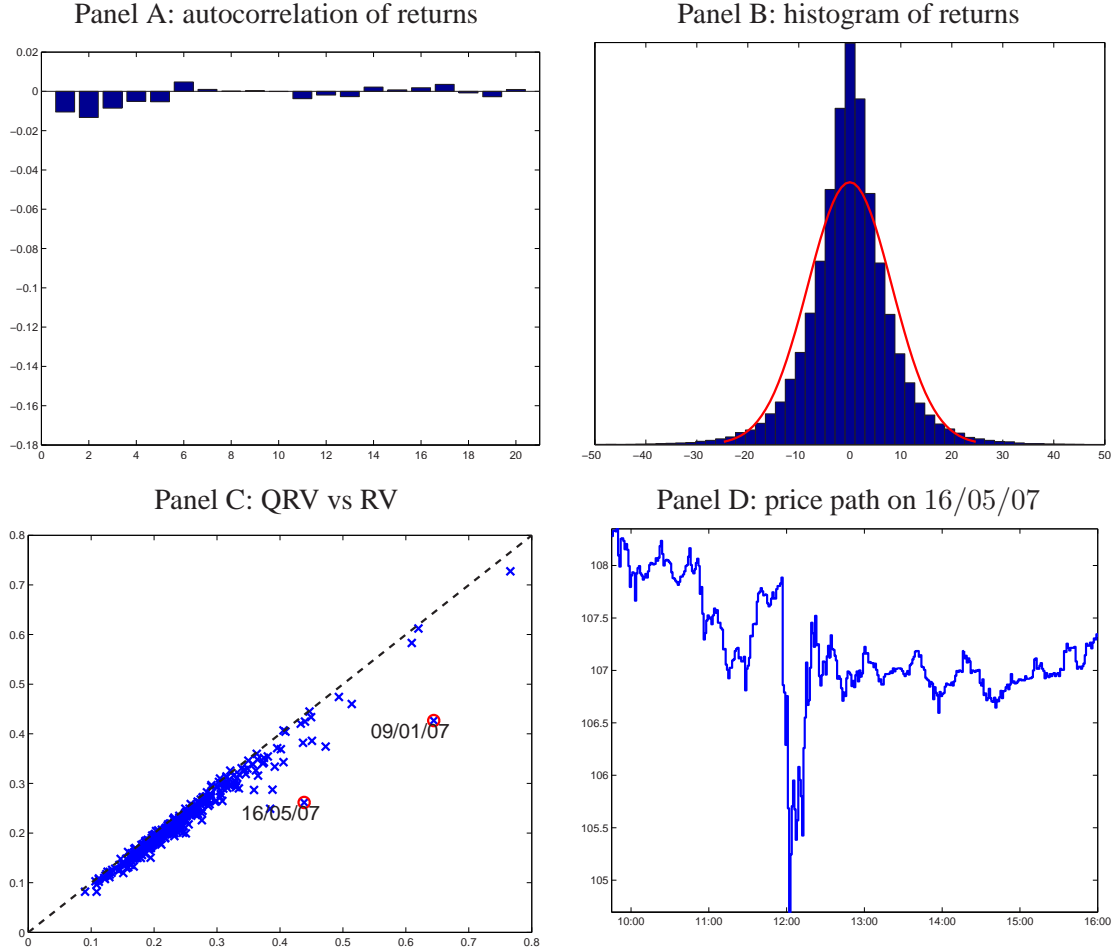
To gauge performance of the estimators we compute, as before, the bias measure  $\mathbb{E}(\widehat{IV}/IV)$  over 100,000 simulation runs where  $\widehat{IV} = \{QRV^*, MSRV\}$ . The efficiency measure is modified to reflect the noise case and now computed as  $\text{var}(N^{1/4}(\widehat{IV} - IV)/\sqrt{\omega\sigma^3})$ . Asymptotically, the best attainable efficiency of both QRV and MSRV is close to 8.

Table 4 reports the simulation results. The following observations can be made. First, both RV and QRV are heavily biased making them unsuitable for implementation with noisy data. In contrast, MSRV is unbiased irrespective of the level of noise and has excellent efficiency. Second, for low  $K$ , the noise adjusted QRV\* has comparable efficiency to MSRV but incurs a considerable bias. Increasing  $K$  to values in the range 10 – 15 largely eliminates the bias at the cost of an efficiency loss. So overall, we find that MSRV delivers superior performance in the absence of jumps, but by sacrificing some efficiency we can obtain noise *and* jump robust estimates of the integrated variance via QRV\*.

## 4 Empirical illustration

In this section, we apply the QRV estimator to a variety of data. The aim here is to illustrate the practical implementation of QRV and highlight some of its empirical properties. We use “clean” and “noisy” high frequency data over short horizons, as well as low-frequency data over long horizons, and find that in all cases the performance of the QRV compares favorably to RV and its microstructure noise adjusted counterparts.

Figure 2: QRV with “clean” AAPL 1-minute NBBO mid-quote data



#### 4.1 QRV with “clean” high-frequency data

We start with an application to high-frequency data. For the moment, we consider relatively “clean” data where microstructure noise is not a prime concern. In particular, for the highly liquid stock Apple (AAPL), we sample mid-quote data at 1 minute frequency from the national best bid and offer (nbbo) prices originating from the NYSE or NASDAQ. We sample data between 9:45 and 16:00 over the period 1 May 2006 through 31 August 2007, i.e. 336 trading days, with 376 sampled prices per day. To confirm the benign impact of noise in this data, we compute the first twenty return autocorrelations in Panel A of Figure 2 and find they are minimal (less than 1.5% in absolute value). Because QRV requires a (conditionally) normal return distribution, price discreteness can be a concern. However, Panel B of Figure 2 indicates that the 1-minute mid-quote returns are close to normal. Therefore, it seems justified to implement QRV without any noise adjustments, and make comparisons to RV. To compute QRV we use the same quantiles as before, i.e.

$\bar{\lambda} = \{0.85; 0.90; 0.96\}$ , and set  $m = N$  for simplicity (the results are again robust to reasonable choice of  $m$ ).

Panel C of Figure 2 plots the daily QRV estimates on the vertical axis versus those of RV on horizontal axis. We note that QRV estimates are closely aligned with RV estimates on most days indicating good efficiency. However, there are significant departures in specific estimates. We highlighted a couple and can identify jumps for each. For instance, on 9 January 2007, Apple announced the iPhone and Electronic News from Reed Business Information reports:

*“Confirming months of rumors and speculation, Apple Computer’s CEO Steve Jobs today announced the consumer electronics superstar’s foray into the mobile phone business, the iPhone . . . The effects of the new product launches quickly reverberated from San Francisco all the way to Wall Street. After trading at around \$85 per share for the past week and opening this morning at \$86.50, shares jumped during and after Jobs’ keynote. Apple shares were trading for as much as \$92.35 as news of the iPhone hit blogs and news sites, an increase of 6.7 percent.”*

Another large price jump occurs on 16 May 2007, see Panel D of Figure 2. The following day an article in MacNews-World, ECT News Network reports:

*“Apple hasn’t identified the prankster who circulated an e-mail at its headquarters saying that the iPhone and Leopard launches would be delayed. Nor has the identity of the person who leaked it been released. The tech blog that reported the fake news, Engadget.com, is on the hot seat for spreading the rumor that caused Apple shares to temporarily lose about \$4 billion in value.”*

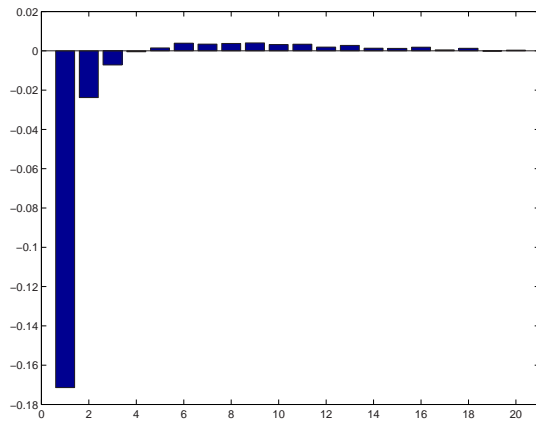
Averaging over the full sample, we find that the annualized variance estimates are 0.0681 for RV and 0.0580 for QRV, indicating that about 15% of total variation can be attributed to jumps. Interestingly, using bi-power variation technology on high frequency data for the S&P500 index over the period 1990 – 2002, Andersen, Bollerslev, and Diebold (2007) estimate a very similar figure, namely a 14.4% jump contribution to total variation. In a related study, Huang and Tauchen (2005) estimate the jump contribution at around 7% using 5 minute S&P500 index data over the period 1997–2002. While an in-depth comparison of these results is difficult (as different data, sampling frequency and horizon, and econometric techniques are used) it does illustrate that our estimates are plausible.

## 4.2 QRV with “noisy” high-frequency data

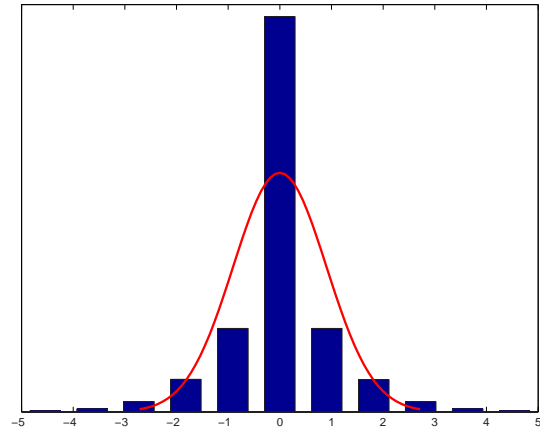
We now use AAPL *trade* data over the same period as above. We include trades from the NASDAQ primary exchange only and don’t filter the data. We have a total of 23,853,550 trades, an average of about 70,000 per day. Despite high liquidity, trade data are inherently noisy due to presence of bid-ask spread bounce. This is confirmed in Panel A of Figure 3 where we find substantial autocorrelation of returns. Also, with trade data at this frequency, price discreteness is a concern: from Panel B we see that 75.4% of return observations are zero, 20.2% are 1 tick, 3.2% are 2 ticks, and 99.9%

Figure 3: QRV with “noisy” AAPL trade data

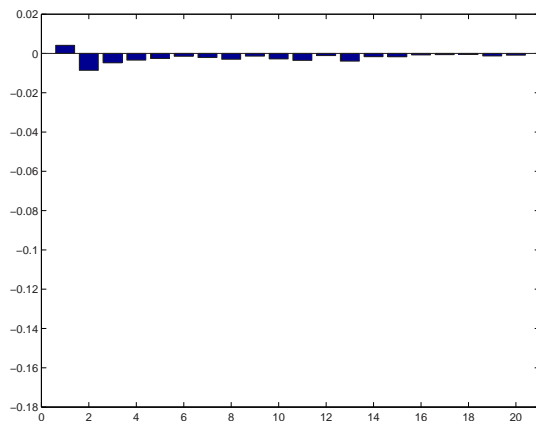
Panel A: autocorrelation of returns



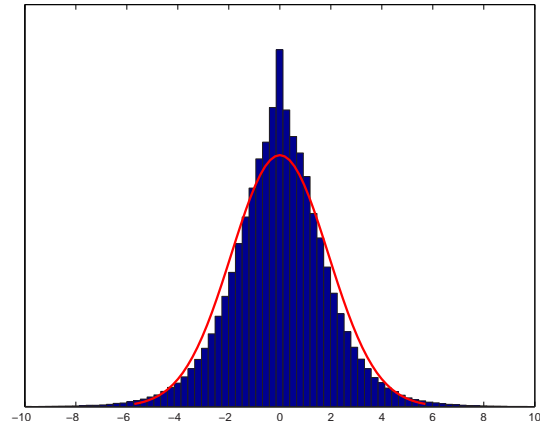
Panel B: histogram of returns



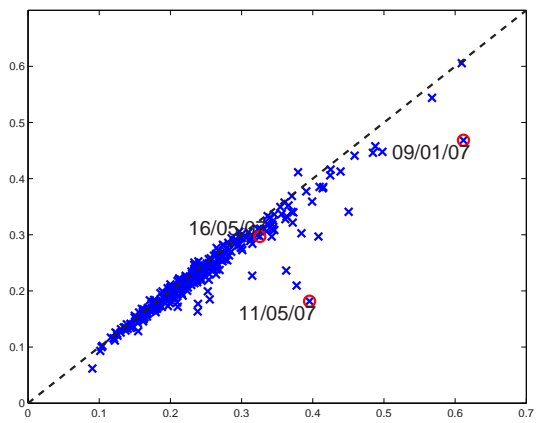
Panel C: autocorrelation of  
noise-sampled returns ( $K = 15$ )



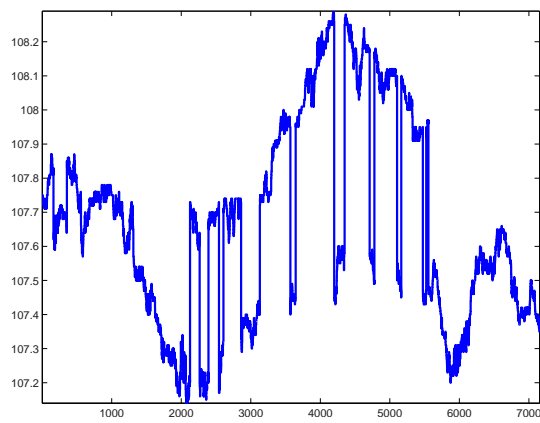
Panel D: histogram of  
noise-sampled returns ( $K = 15$ )



Panel E: QRV vs MSRV



Panel F: trades on 11/05/07 (09:55 – 10:05)



are less than or equal than 5 ticks in magnitude. For these reasons, it is clearly inappropriate to apply the standard QRV. Instead, we first filter the data using the procedure outlined above. Panel C and D of Figure 3 report the autocorrelation and histogram of noise filtered returns with  $K = 15$ . The results are striking: not only is the autocorrelation virtually eliminated, the return distribution is also much closer to normal so that  $QRV^*$  can be readily applied. Panel E draws the scatter plot of  $QRV^*$  versus  $MSRV^2$  and, as before, we find a close alignment of the two estimates. It is interesting to note that for the previously classified “jump day” 16/05/07 based on 1 minute data, the  $QRV^*$  estimate from trade data is no longer much different from  $MSRV$ . Closer inspection of the data reveals that during the 6 minutes between 11:55 and 12:01 where the AAPL price dropped nearly than 3%, a total of 8720 trades took place (more than 10% of average daily trade volume) with a maximum absolute trade-by-trade price change of only 8 cents! This is a good example where an extreme burst of volatility over a short horizon is mistaken for a jump in prices at lower frequency. This logic is also what underlies the family of jump tests developed by Aït-Sahalia and Jacod (2006), namely true jumps can only be identified by increasing the sampling frequency to the limit. An example of a day where  $QRV^*$  deviates substantially from  $MSRV$  is 11/05/2007. Panel F of Figure 3 plots the price path between 09:55 – 10:05 and we observe a very erratic behavior. Closer inspection of the data reveals that many of the trades here are “out-of-sequence” (as indicated by condition code “Z” in the TAQ data). Obviously, such observations can (and should) be filtered out beforehand, but in practice filtering may be imperfect and this example further illustrates the robustness of QRV.

The sample average of annualized variance estimates is 0.0669 for  $MSRV$  and 0.0595 for  $QRV^*$ .<sup>3</sup> Both figures are close to that obtained from 1 minute data, indicating the absence of any serious biases, while the contribution of jumps to total variation is now somewhat lower at about 10%. As the above illustrates, when moving to the highest trade frequency, some of the previously identified jump days vanish while others may appear due to the emergence of spurious “noise” jumps.

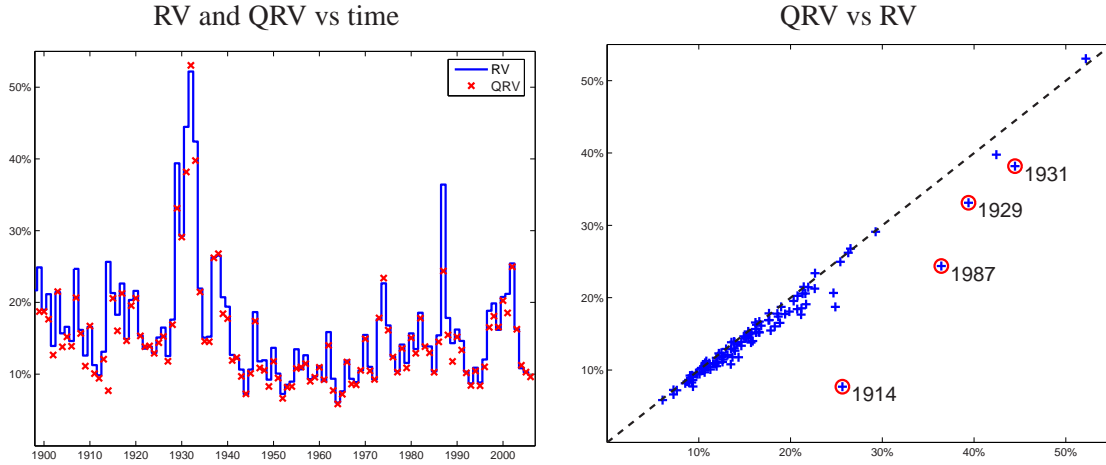
### 4.3 QRV with low frequency data

The use of QRV, like any other realised variance measure, is not only limited to “high” frequency data over short horizons, but can also be applied to “low” frequency data over longer horizons. In the latter case the impact of market microstructure noise is benign and can be ignored for all practical purposes. See, for instance, Schwert (1989) who calculates monthly realised volatilities using daily data or, more recently, Andersen, Bollerslev, Diebold, and Wu (2006) who study quarterly realised variances and realised betas calculated from daily data.

<sup>2</sup>The  $MSRV$  is implemented with the optimal bandwidth as in Eq. (47), estimated for each day in the sample separately. The noise variance is estimated from the first order autocovariance of the trade data, whereas IV and IQ estimates are calculated from subsampling the equivalent of 5 minute data in trade time (i.e. each sub-sample consists of 79 price observations per day). The average bandwidth computed is 3.58, with  $q^* = 3$  for three out of every four days, a minimum of 2, and a maximum of 9.

<sup>3</sup>Noise adjustment is key here. The average variance is estimated as 0.0881 using the noise-uncorrected QRV, 0.1909 using RV, and 0.1201 using BPV

Figure 4: QRV with daily data



For our final illustration, we have available daily data<sup>4</sup> for the Dow Jones Industrial Average (DJIA) stock index over the period January 1897 - December 2006, i.e. 27,553 observations covering 110 years. For each year in the sample, we estimate the return variance using QRV and RV from daily data (i.e.  $N \simeq 250$ ). To implement QRV, we use  $\bar{\lambda} = \{0.86, 0.90, 0.94\}$  with  $m = 50$ , but we find that the results are robust to reasonable alternative choices of quantiles and subsamples (even setting  $m = N$  leads to only minor differences).

In Figure 4 we plot the time series of variance estimates in the first panel and a cross plot of QRV (on the vertical axis) versus RV (on the horizontal axis) in the second. Two observations can be made. Firstly, the years where QRV differs substantially from RV are all years with extreme market movements. For instance, in 1914 the DJIA closed on July 30, 1914 at 71.42 and reopened more than 4 months later on December 14, 1914 at 56.76, reflecting a 20% drop in value. In 1929, the start of the great depressions, the DJIA index fell 13.5% October 28, and another 11.7% the following day, only to rebound on October 30 by 12.3%. Similarly, in 1987 the stock market crashed again, experiencing a daily return of  $-22.6\%$  on October 19 which, even with the RV estimate of 38% for that year, constitutes a nine-standard deviation event. The only year that is seemingly different is 1931. Here, no stock market crash occurred, but considering the list of “20 all-time greatest DJIA positive daily percentage returns” compiled by Dow Jones Indexes, we see that 5 of these occur in 1931. All this illustrates the robustness of QRV to jumps. Secondly, in all remaining years, the QRV estimates are close to those of RV (with a correlation exceeding 99%) indicating good efficiency.

Over the full 110 year sample period, we obtain average annual variance estimates of 0.0321 for RV and 0.0274 for QRV. This indicates that about 15% of total variation is due to jumps. If we leave out the four years discussed above, these figures drop to 0.0281 and 0.0254 respectively, indicating that about 9% of total variation is due to jumps.

<sup>4</sup>Source: Dow Jones Indexes (<http://www.djindexes.com/>)

Interestingly, these figures are broadly in line with the results from the intra-day AAPL data and the above mentioned studies by Andersen, Bollerslev, and Diebold (2007) and Huang and Tauchen (2005). A more in-depth study into the jump contribution to total variation, and its dependence on the observation frequency, is clearly of interest but beyond the scope of the current paper.

## 5 Concluding remarks

In this paper we propose a new quantile-based realised variance measure, QRV, that delivers jump-robust and consistent estimates of the integrated variance. We present an asymptotic distribution theory for QRV under weak assumptions on the price process and show that the estimator has excellent properties. In the presence of i.i.d. market microstructure noise, QRV can be implemented in modified form on filtered data, retaining consistency and attaining the best possible convergence rate of  $N^{-1/4}$ . Extensive simulations, and a brief empirical application, further illustrate the finite sample properties of our estimator.

The methodology outlined in this paper can be extended into various directions. For instance, it is possible to develop a joint distribution theory for RV and QRV allowing the construction of a formal jump test in the spirit of Barndorff-Nielsen and Shephard (2006). Also, it is possible to modify QRV to produce jump and noise robust estimates of the integrated quarticity, a key quantity when making inference about integrated variance and testing for jumps. With these tools available, it may then be interesting to revisit some of the empirical work on non-parametric jump tests (e.g. Aït-Sahalia and Jacod, 2006; Barndorff-Nielsen and Shephard, 2006; Fan and Wang, 2006; Jiang and Oomen, 2007; Lee and Mykland, 2007) and the measurement and modeling of the jump contribution to total variation (e.g. Huang and Tauchen, 2005; Andersen, Bollerslev, and Diebold, 2007). Recent work by Aït-Sahalia and Jacod (2006), Barndorff-Nielsen, Shephard, and Winkel (2006) and Woerner (2006) has shown that bi-power variation is not only robust to finite activity jumps (as considered in this paper) but also to certain infinite activity jump specifications. An investigation of the properties of QRV in such a scenario might be of interest and allow for further comparison to alternative jump robust estimators. Finally, the data filtering method of Podolskij and Vetter (2006) underlying the construction of the noise adjusted QRV may be modified to allow for certain types of dependent noise (as studied by for instance Aït-Sahalia, Mykland, and Zhang, 2006; Barndorff-Nielsen, Hansen, Lunde, and Shephard, 2006; Jacod, Li, Mykland, Podolskij, and Vetter, 2007) and this warrants further study. All the above is well beyond the scope of the current paper, and will be left for future research.

## A Proofs

In this part of the paper, we state the proofs of the theorems given in the main text. Throughout, we use the approximation

$$\Delta_i^N X \approx \sigma_{\frac{i-1}{N}} \Delta_i^N W.$$

Thus, to prove our asymptotic results we first replace  $\Delta_i^N X$  with  $\sigma_{\frac{i-1}{N}} \Delta_i^N W$  and then show that the error caused by this approximation is asymptotically negligible.

Let us fix some notations. We set

$$\beta_i^n = \sqrt{N} \left( \sigma_{\frac{i-1}{n}} \Delta_k^N W \right)_{(i-1)m+1 \leq k \leq im},$$

and define

$$w_i^{(n,m)}(\lambda) = g_{\lambda m}^2(\beta_i^n) + g_{m-\lambda m+1}^2(\beta_i^n).$$

Before we start to prove the main results, we state a simple Lemma.

**Lemma 1** *The function  $g_k$  defined in (2) has the following properties:*

1.  $g_k$  is continuous.
2.  $g_k$  is differentiable on the set  $\{x \in \mathbb{R}^m \mid x_i \neq x_j, 1 \leq i < j \leq m\}$ , that is

$$\frac{1}{\epsilon} [g_k(x + \epsilon y) - g_k(x)] \rightarrow y_{k*} \quad \epsilon \searrow 0,$$

where  $y \in \mathbb{R}^m$  and

$$k* = i \text{ with } x_i = x_{(k)}.$$

In the following we assume without loss of generality that  $a, \sigma, a', \sigma'$  and  $v'$  are bounded (for details see e.g. Section 3 in Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard, 2006). Moreover, the constants used in the proofs will all be denoted by  $C$ .

### A.1 Results with $N \rightarrow \infty$ and $m$ fixed

**Proof of Theorem 1** First, we define:

$$\xi_i^n = \nu_{1,m}^{-1}(\lambda) w_i^{(n,m)}(\lambda),$$

$$U_n = \frac{1}{n} \sum_{i=1}^n \xi_i^n.$$

Note that:

$$\mathbb{E} \left[ \xi_i^n \mid \mathcal{F}_{\frac{i-1}{n}} \right] = \sigma_{\frac{i-1}{n}}^2,$$

so

$$\frac{1}{n} \sum_{i=1}^n \mathbb{E} \left[ \xi_i^n \mid \mathcal{F}_{\frac{i-1}{n}} \right] \xrightarrow{p} \int_0^1 \sigma_u^2 du. \quad (48)$$

Now, by setting

$$\eta_i^n = \xi_i^n - \mathbb{E} \left[ \xi_i^n \mid \mathcal{F}_{\frac{i-1}{n}} \right], \quad (49)$$

we get:

$$\mathbb{E} \left[ |\eta_i^n|^2 \mid \mathcal{F}_{\frac{i-1}{n}} \right] = \frac{\nu_{2,m}(\lambda) - \nu_{1,m}^2(\lambda)}{\nu_{1,m}^2(\lambda)} \sigma_{\frac{i-1}{n}}^4.$$

Therefore,

$$\frac{1}{n^2} \sum_{i=1}^n \mathbb{E} \left[ |\eta_i^n|^2 \mid \mathcal{F}_{\frac{i-1}{n}} \right] \xrightarrow{p} 0.$$

Hence, the assertion  $U_n \xrightarrow{p} \int_0^1 \sigma_u^2 du$  follows directly from (48). Now, we are left to prove that

$$QRV_N(m, \lambda) - U_n \xrightarrow{p} 0. \quad (50)$$

Note that

$$QRV_N(m, \lambda) - U_n = \frac{\nu_{1,m}^{-1}(\lambda)}{n} \sum_{i=1}^n \zeta_i^n,$$

where

$$\zeta_i^n = q_i^{(n,m)}(\lambda) - w_i^{(n,m)}(\lambda).$$

Now, we use the decomposition

$$\zeta_i^n = \zeta_i^n(1) + \zeta_i^n(2),$$

where  $\zeta_i^n(k)$ ,  $k = 1, 2$ , are given by

$$\zeta_i^n(1) = g_{\lambda m}^2 \left( \sqrt{N} D_i^m X \right) - g_{\lambda m}^2(\beta_i^n), \quad (51)$$

$$\zeta_i^n(2) = g_{m-\lambda m+1}^2 \left( \sqrt{N} D_i^m X \right) - g_{m-\lambda m+1}^2(\beta_i^n). \quad (52)$$

In the following, we show that

$$\frac{\nu_{1,m}^{-1}(\lambda)}{n} \sum_{i=1}^n \zeta_i^n(1) \xrightarrow{p} 0. \quad (53)$$

The corresponding result for  $\zeta_i^n(2)$  can be proven similarly. We begin with the following Lemma.

**Lemma 2** For  $x \in \mathbb{R}^m$ , we define a norm  $\|x\| = \sum_{k=1}^m |x_k|$ . Then we have

$$\frac{1}{n} \sum_{i=1}^n \mathbb{E} \left[ \|\sqrt{N} D_i^m X - \beta_i^n\|^2 \right] \rightarrow 0. \quad (54)$$

**Proof of Lemma 2** The boundedness of the drift function  $a$  and  $\|x\|^2 \leq m \sum_{k=1}^m |x_k|^2$  yield

$$\begin{aligned} \mathbb{E} \left[ \|\sqrt{N} D_i^m X - \beta_i^n\|^2 \right] &\leq mC \left( \frac{1}{n} + N \sum_{k=1}^m \mathbb{E} \left[ \int_{\frac{i-1}{n} + \frac{k-1}{N}}^{\frac{i-1}{n} + \frac{k}{N}} |\sigma_u - \sigma_{\frac{i-1}{n}}|^2 du \right] \right) \\ &= mC \left( \frac{1}{n} + N \mathbb{E} \left[ \int_{\frac{i-1}{n}}^{\frac{i}{n}} |\sigma_u - \sigma_{\frac{i-1}{n}}|^2 du \right] \right). \end{aligned}$$

Hence, the left side of (54) is smaller than

$$mC \left( \frac{1}{n} + m \int_0^1 \mathbb{E} [|\sigma_u - \sigma_{[nu]/n}|^2] du \right).$$

Because  $\sigma$  is bounded and càdlàg, the assertion of Lemma 2 is proved by Lebesgue's theorem. ■

Next, we set

$$m_A(\epsilon) = \sup\{|g_{\lambda m}^2(x) - g_{\lambda m}^2(y)| : \|x - y\| \leq \epsilon, \|x\| \leq A\}.$$

For all  $\epsilon \in (0, 1]$  and  $A > 1$ , we obtain the estimate

$$\begin{aligned} \zeta_i^n(1) &\leq C \left( m_A(\epsilon) + A^2 1_{\{\|\sqrt{N}D_i^m X - \beta_i^n\| > \epsilon\}} \right. \\ &\quad \left. + \left( g_{\lambda m}^2(\sqrt{N}D_i^m X) + g_{\lambda m}^2(\beta_i^n) \right) \left( 1_{\{\|\sqrt{N}D_i^m X\| > A\}} + 1_{\{\|\beta_i^n\| > A\}} \right) \right) \\ &\leq C \left( m_A(\epsilon) + \frac{A^2 \|\sqrt{N}D_i^m X - \beta_i^n\|^2}{\epsilon^2} + \frac{(\|\sqrt{N}D_i^m X\| + \|\beta_i^n\|)^3}{A} \right). \end{aligned}$$

The boundedness of  $a$  and  $\sigma$  and Burkholder's inequality imply that

$$\mathbb{E} [\|\sqrt{N}D_i^m X\|^p] + \mathbb{E} [\|\beta_i^n\|^p] \leq C_p,$$

for all  $p \geq 0$ . This means that

$$\frac{1}{n} \sum_{i=1}^n \mathbb{E} [\|\zeta_i^n(1)\|] \leq C \left( m_A(\epsilon) + \frac{1}{A} + \frac{A^2}{n\epsilon^2} \sum_{i=1}^n \mathbb{E} [\|\sqrt{N}D_i^m X - \beta_i^n\|^2] \right).$$

Because  $m_A(\epsilon) \rightarrow 0$  as  $\epsilon \rightarrow 0$  for every  $A$ , we obtain by Lemma 2

$$\frac{1}{n} \sum_{i=1}^n \mathbb{E} [\|\zeta_i^n(1)\|] \rightarrow 0, \tag{55}$$

by first choosing  $A$  large, then  $\epsilon$  small and finally  $n$  large. Then (53) holds and the proof of Theorem 1 is complete. ■

**Proof of Theorem 2** We proceed with a three-stage proof of Theorem 2. First, we prove a CLT for the sequence

$$\bar{U}_n = \sqrt{\frac{m}{n}} \sum_{i=1}^n \eta_i^n,$$

with  $\eta_i^n$  defined by (49). The second step is to define a new sequence:

$$U'_n = \nu_{1,m}^{-1}(\lambda) \sqrt{\frac{m}{n}} \sum_{i=1}^n \left( q_i^{(n,m)}(\lambda) - \mathbb{E} \left[ q_i^{(n,m)}(\lambda) \mid \mathcal{F}_{\frac{i-1}{n}} \right] \right),$$

and show the result

$$U'_n - \bar{U}_n \xrightarrow{p} 0.$$

Finally, in part III, the theorem follows from the convergences:

$$\begin{aligned} &\sqrt{\frac{m}{n}} \sum_{i=1}^n \left( \nu_{1,m}^{-1}(\lambda) \mathbb{E} \left[ q_i^{(n,m)}(\lambda) \mid \mathcal{F}_{\frac{i-1}{n}} \right] - \mathbb{E} \left[ \xi_i^n \mid \mathcal{F}_{\frac{i-1}{n}} \right] \right) \xrightarrow{p} 0, \\ &\sqrt{\frac{m}{n}} \left( \sum_{i=1}^n \mathbb{E} \left[ \xi_i^n \mid \mathcal{F}_{\frac{i-1}{n}} \right] - \int_0^1 \sigma_u^2 du \right) \xrightarrow{p} 0. \end{aligned}$$

**Proof of part I** Notice that:

$$\frac{m}{n} \sum_{i=1}^n \mathbb{E} \left[ (\eta_i^n)^2 \mid \mathcal{F}_{\frac{i-1}{n}} \right] \xrightarrow{p} m \frac{\nu_{2,m}(\lambda) - \nu_{1,m}^2(\lambda)}{\nu_{1,m}^2(\lambda)} \int_0^1 \sigma_u^4 du.$$

Moreover, since  $W \stackrel{d}{=} -W$  and  $w_i^{(n,m)}(\lambda)$  is an even functional in  $W$ , we have

$$\mathbb{E} \left[ \eta_i^n \Delta_i^n W \mid \mathcal{F}_{\frac{i-1}{n}} \right] = 0.$$

Next, let  $N = (N_t)_{t \in [0,1]}$  be a bounded martingale on  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , which is orthogonal to  $W$  (i.e., with quadratic covariation  $[W, N] = 0$ , almost surely). Then,

$$\sqrt{\frac{m}{n}} \sum_{i=1}^n \mathbb{E} \left[ \eta_i^n \Delta_i^n N \mid \mathcal{F}_{\frac{i-1}{n}} \right] = 0.$$

This result is shown in Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard (2006) (see the proof of Proposition 4.1) and in Christensen and Podolskij (2006). Finally, stable convergence in law follows by Theorem IX 7.28 in Jacod and Shiryaev (2003):

$$\bar{U}_n \xrightarrow{d_s} \sqrt{m \frac{\nu_{2,m}(\lambda) - \nu_{1,m}^2(\lambda)}{\nu_{1,m}^2(\lambda)}} \int_0^1 \sigma_u^2 dW'_u.$$

which completes the proof of part I. □

**Proof of part II** We begin by setting

$$\delta_i^n = \nu_{1,m}^{-1}(\lambda) \sqrt{\frac{m}{n}} \left( q_i^{(n,m)}(\lambda) - w_i^{(n,m)}(\lambda) \right),$$

and obtain the identity:

$$U'_n - \bar{U}_n = \sum_{i=1}^n \left( \delta_i^n - \mathbb{E} \left[ \delta_i^n \mid \mathcal{F}_{\frac{i-1}{n}} \right] \right).$$

To complete the second step, it suffices that

$$\sum_{i=1}^n \mathbb{E} [|\delta_i^n|^2] \rightarrow 0.$$

We omit the proof of this result, as it be shown by using exactly the same methods behind the proof of the convergence in (55) in Theorem 1. □

**Proof of part III** It holds that:

$$\sqrt{\frac{m}{n}} \left( \sum_{i=1}^n \mathbb{E} \left[ \xi_i^n \mid \mathcal{F}_{\frac{i-1}{n}} \right] - \int_0^1 \sigma_u^2 du \right) = \sqrt{mn} \sum_{i=1}^n \int_{\frac{i-1}{n}}^{\frac{i}{n}} \left( \sigma_{\frac{i-1}{n}}^2 - \sigma_u^2 \right) du.$$

Exploiting the results of Section 8 in Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard (2006) (recall that  $m$  is a fixed number), we find that, under condition (V), the convergence

$$\sqrt{\frac{m}{n}} \left( \sum_{i=1}^n \mathbb{E} \left[ \xi_i^n \mid \mathcal{F}_{\frac{i-1}{n}} \right] - \int_0^1 \sigma_u^2 du \right) \xrightarrow{p} 0,$$

holds. Now, we prove the first convergence of part III stated above. Under condition (V), we introduce the decomposition

$$\sqrt{N} D_i^m X - \beta_i^n = \mu_i^n(1) + \mu_i^n(2),$$

where  $\mu_i^n(1)$  and  $\mu_i^n(2)$  are  $m$ -dimensional vectors with components defined by

$$\begin{aligned}
(\mu_i^n(1))_k &= \sqrt{N} \int_{\frac{i-1}{n} + \frac{k-1}{N}}^{\frac{i-1}{n} + \frac{k}{N}} \left( a_u - a_{\frac{i-1}{n}} \right) du + \sqrt{N} \int_{\frac{i-1}{n} + \frac{k-1}{N}}^{\frac{i-1}{n} + \frac{k}{N}} \left( \int_{\frac{i-1}{n} + \frac{k-1}{N}}^u a'_s ds \right. \\
&\quad \left. + \int_{\frac{i-1}{n} + \frac{k-1}{N}}^u \left( \sigma'_s - \sigma'_{\frac{i-1}{n}} \right) dW_s + \int_{\frac{i-1}{n} + \frac{k-1}{N}}^u \left( v'_s - v'_{\frac{i-1}{n}} \right) dB'_s \right) dW_u, \\
(\mu_i^n(2))_k &= \sqrt{N} \left( \frac{1}{N} a_{\frac{i-1}{n}} + \sigma'_{\frac{i-1}{n}} \int_{\frac{i-1}{n} + \frac{k-1}{N}}^{\frac{i-1}{n} + \frac{k}{N}} \left( W_u - W_{\frac{i-1}{n}} \right) dW_u \right. \\
&\quad \left. + v'_{\frac{i-1}{n}} \int_{\frac{i-1}{n} + \frac{k-1}{N}}^{\frac{i-1}{n} + \frac{k}{N}} \left( B'_u - B'_{\frac{i-1}{n}} \right) dW_u \right), \tag{56}
\end{aligned}$$

for  $k = 1, \dots, m$ . Moreover, we decompose

$$\nu_{1,m}^{-1}(\lambda) q_i^{(n,m)}(\lambda) - \xi_i^n = \theta_i^n(1) + \theta_i^n(2),$$

where

$$\begin{aligned}
\theta_i^n(1) &= \nu_{1,m}^{-1}(\lambda) \left( g_{\lambda m}^2 \left( \sqrt{N} D_i^m X \right) - g_{\lambda m}^2(\beta_i^n + \mu_i^n(2)) \right. \\
&\quad \left. + g_{m-\lambda m+1}^2 \left( \sqrt{N} D_i^m X \right) - g_{m-\lambda m+1}^2(\beta_i^n + \mu_i^n(2)) \right), \\
\theta_i^n(2) &= \nu_{1,m}^{-1}(\lambda) \left( g_{\lambda m}^2(\beta_i^n + \mu_i^n(2)) - g_{\lambda m}^2(\beta_i^n) \right. \\
&\quad \left. + g_{m-\lambda m+1}^2(\beta_i^n + \mu_i^n(2)) - g_{m-\lambda m+1}^2(\beta_i^n) \right). \tag{57}
\end{aligned}$$

Using the same methods as for the proof of (55) in Theorem 1, we obtain

$$\sqrt{\frac{m}{n}} \sum_{i=1}^n \mathbb{E} [|\theta_i^n(1)|] \rightarrow 0,$$

which implies

$$\sqrt{\frac{m}{n}} \sum_{i=1}^n \mathbb{E} \left[ \theta_i^n(1) \mid \mathcal{F}_{\frac{i-1}{n}} \right] \xrightarrow{p} 0.$$

Thus, we are left to prove that

$$\sqrt{\frac{m}{n}} \sum_{i=1}^n \mathbb{E} \left[ \theta_i^n(2) \mid \mathcal{F}_{\frac{i-1}{n}} \right] \xrightarrow{p} 0.$$

Now we apply Lemma 1 to the term  $\theta_i^n(2)$  with  $\epsilon = N^{-1/2}$  and

$$x = \beta_i^n,$$

$$y = \sqrt{N} \mu_i^n(2).$$

Notice that as  $\sigma$  does not vanish, we have  $(\beta_i^n)_k \neq (\beta_i^n)_l$  for all  $1 \leq k < l \leq m$  almost surely, and consequently the assumptions of Lemma 1 are satisfied. Finally, we obtain the stochastic expansion

$$\begin{aligned} \sqrt{\frac{m}{n}} \sum_{i=1}^n \mathbb{E} \left[ \theta_i^n(2) \mid \mathcal{F}_{\frac{i-1}{n}} \right] &= 2\nu_{1,m}^{-1}(\lambda) \sqrt{\frac{m}{n}} \sum_{i=1}^n \mathbb{E} \left[ g_{\lambda m}(\beta_i^n) \left( \mu_i^n(2) \right)_{(\lambda m)*} \right. \\ &\quad \left. + g_{m-\lambda m+1}(\beta_i^n) \left( \mu_i^n(2) \right)_{(m-\lambda m+1)*} \mid \mathcal{F}_{\frac{i-1}{n}} \right] + o_p(1), \end{aligned}$$

where we recall that  $(\lambda m)*$  is defined by

$$(\lambda m)* = k \quad \text{with} \quad (\beta_i^n)_k = (\beta_i^n)_{(\lambda m)}.$$

Now  $(W, V) \stackrel{d}{=} - (W, V)$  and

$$g_{\lambda m}(\beta_i^n) \left( \mu_i^n(2) \right)_{(\lambda m)*} + g_{m-\lambda m+1}(\beta_i^n) \left( \mu_i^n(2) \right)_{(m-\lambda m+1)*}$$

is odd in  $(W, V)$ , which implies that

$$\mathbb{E} \left[ g_{\lambda m}(\beta_i^n) \left( \mu_i^n(2) \right)_{(\lambda m)*} + g_{m-\lambda m+1}(\beta_i^n) \left( \mu_i^n(2) \right)_{(m-\lambda m+1)*} \mid \mathcal{F}_{\frac{i-1}{n}} \right] = 0.$$

Consequently,

$$\sqrt{\frac{m}{n}} \sum_{i=1}^n \mathbb{E} \left[ \theta_i^n(2) \mid \mathcal{F}_{\frac{i-1}{n}} \right] \xrightarrow{p} 0,$$

which completes the proof of part III and, hence, Theorem 2. ■

**Proof of Theorem 3** This result is shown in the same way as the proof of Theorem 2. ■

## A.2 Results with $N \rightarrow \infty$ and $m \rightarrow \infty$

Here we use the same notations as in the previous subsection (although the quantities now essentially depend on  $m$  we suppress this dependency in our notations).

Just as above we need to prove an analogue of Lemma 1. However, things become more complicated, because  $m \rightarrow \infty$ . In the following, we use the notation

$$\|x\|_\infty = \max_i |x_i|,$$

to denote the maximum norm.

**Lemma 3** *We consider the vectors  $x^{(m)} \in \{x \in \mathbb{R}^m \mid x_i \neq x_j, 1 \leq i < j \leq m\}$ ,  $y^{(m)} \in \mathbb{R}^m$ . Let  $\epsilon_m \rightarrow 0$  be a positive sequence with*

$$\epsilon_m \|y^{(m)}\|_\infty = O(m^{-\rho}),$$

*for some  $\rho > 0$  (as  $m \rightarrow \infty$ ). Moreover, we assume that*

$$\|x^{(m)}\|_\infty = O(m^\delta) \quad \max_{1 \leq i \neq j \leq m} |x_i^{(m)} - x_j^{(m)}| \geq K > 0,$$

for some  $\delta \geq 0$  and  $K > 0$ . Let  $a_m = 1$  or  $a_m = \sqrt{m}$ . Then we have for all  $\lambda \in (0, 1)$

$$H_m := \frac{a_m}{\epsilon_m} \left[ g_{\lambda m}^2 \left( x^{(m)} + \epsilon_m y^{(m)} \right) - g_{\lambda m}^2 \left( x^{(m)} \right) \right] = 2a_m x_{(\lambda m)^*}^{(m)} y_{(\lambda m)^*}^{(m)} + a_m \epsilon_m \left( y_{(\lambda m)^*}^{(m)} \right)^2 + o(1),$$

where

$$(\lambda m)^* = i \quad \text{with} \quad x_i^{(m)} = x_{(\lambda m)^*}^{(m)}.$$

**Proof of Lemma 3** We use the decomposition

$$\begin{aligned} H_m &= \left( 1 - 1_{\{\epsilon_m \|y^{(m)}\|_\infty \geq \max_{1 \leq i \neq j \leq m} |x_i^{(m)} - x_j^{(m)}|\}} \right) H_m + 1_{\{\epsilon_m \|y^{(m)}\|_\infty \geq \max_{1 \leq i \neq j \leq m} |x_i^{(m)} - x_j^{(m)}|\}} H_m \\ &= H_m(1) + H_m(2). \end{aligned}$$

Note that

$$|H_m| \leq \frac{3a_m}{\epsilon_m} \left( \|x^{(m)}\|_\infty^2 + \epsilon_m^2 \|y^{(m)}\|_\infty^2 \right),$$

and we obtain

$$|H_m(2)| \leq \frac{3a_m \epsilon_m^t \|y^{(m)}\|_\infty^t}{\epsilon_m \left( \max_{1 \leq i \neq j \leq m} |x_i^{(m)} - x_j^{(m)}| \right)^t} \left( \|x^{(m)}\|_\infty^2 + \epsilon_m^2 \|y^{(m)}\|_\infty^2 \right) \rightarrow 0,$$

for some  $t$  with  $t\rho > \frac{1}{2} + \rho + 2\delta$ . Next, note that on  $\{\epsilon_m \|y^{(m)}\|_\infty < \max_{1 \leq i \neq j \leq m} |x_i^{(m)} - x_j^{(m)}|\}$

$$H_m = 2a_m x_{(\lambda m)^*}^{(m)} y_{(\lambda m)^*}^{(m)} + a_m \epsilon_m \left( y_{(\lambda m)^*}^{(m)} \right)^2.$$

It follows that

$$H_m(1) = 2a_m x_{(\lambda m)^*}^{(m)} y_{(\lambda m)^*}^{(m)} + a_m \epsilon_m \left( y_{(\lambda m)^*}^{(m)} \right)^2 + o(1),$$

and we are done. ■

**Proof of Theorem 4** We define:

$$\xi_i^{n,m} = (c_\lambda^2 + c_{1-\lambda}^2)^{-1} w_i^{(n,m)}(\lambda),$$

$$U_{n,m} = \frac{1}{n} \sum_{i=1}^n \xi_i^{n,m}.$$

Exactly as in the proof of Theorem 1, we can conclude that

$$U_{n,m} \xrightarrow{p} \int_0^1 \sigma_u^2 du.$$

Now we are left to prove that ( $k = 1, 2$ )

$$\frac{1}{n} \sum_{i=1}^n \zeta_i^n(k) \xrightarrow{p} 0,$$

where  $\zeta_i^n(1), \zeta_i^n(2)$  are defined by (51) and (52). In the following, we show that

$$\frac{1}{n} \sum_{i=1}^n \zeta_i^n(1) \xrightarrow{p} 0,$$

but we omit the proof for  $\zeta_i^n(2)$  as it is identical. We prove the result by applying Lemma 3. First, note that the boundedness of  $a$  and Burkholder's inequality imply

$$\begin{aligned}\mathbb{E} \|\sqrt{N}D_i^m X - \beta_i^n\|_\infty^2 &\leq N \mathbb{E} \left( \sup_{s,t \in [\frac{i-1}{n}, \frac{i}{n}]} |(X_t - X_s) - \sigma_{\frac{i-1}{n}}(W_t - W_s)|^2 \right) \\ &\leq CN \left( \frac{1}{n^2} + \mathbb{E} \left[ \int_{\frac{i-1}{n}}^{\frac{i}{n}} (\sigma_u - \sigma_{\frac{i-1}{n}})^2 du \right] \right).\end{aligned}$$

Next, observe that under assumption (V) we have that

$$\mathbb{E} [|\sigma_t - \sigma_s|^p] \leq C_p |t - s|^{\frac{p}{2}},$$

for all  $p > 0$ . This implies that

$$\mathbb{E} \|\sqrt{N}D_i^m X - \beta_i^n\|_\infty^2 \leq \frac{CN}{n^2} = CN^{2\gamma-1},$$

where  $\gamma \in (0, 1/2)$  is defined by (16). Finally, we find that

$$\mathbb{E} \left[ \left( \max_{1 \leq k \neq l \leq m} |(\beta_i^n)_k - (\beta_i^n)_l| \right)^{-p} \right] < \infty,$$

for fixed  $p > 0$ , if  $m$  is large enough. This property follows from Lemma 1 in Christensen and Podolskij (2006). Next, we apply Lemma 3 with

$$\epsilon_m = m^{\frac{\gamma-1/2}{\gamma}} = O(N^{\gamma-1/2}),$$

$$x^{(m)} = \beta_i^n,$$

$$y^{(m)} = \frac{1}{\epsilon_m} (\sqrt{N}D_i^m X - \beta_i^n),$$

and obtain

$$\begin{aligned}\frac{1}{n} \sum_{i=1}^n \mathbb{E} (|\zeta_i^n(1)|) &= \frac{1}{n} \sum_{i=1}^n \mathbb{E} \left( |2(\beta_i^n)_{(\lambda m)} (\sqrt{N}D_i^m X - \beta_i^n)_{(\lambda m)^*} + \left( (\sqrt{N}D_i^m X - \beta_i^n)_{(\lambda m)^*} \right)^2| \right) + o_p(1) \\ &\leq \frac{1}{n} \sum_{i=1}^n \mathbb{E} \left( 2|(\beta_i^n)_{(\lambda m)}| \|\sqrt{N}D_i^m X - \beta_i^n\|_\infty + \|\sqrt{N}D_i^m X - \beta_i^n\|_\infty^2 \right) + o_p(1).\end{aligned}$$

As

$$U_{(\lambda m)} \xrightarrow{p} c_\lambda,$$

where  $U_i \sim N(0, 1)$ , for  $i = 1, \dots, m$ , it follows by the Cauchy-Schwarz inequality that

$$\begin{aligned}\frac{2}{n} \sum_{i=1}^n \mathbb{E} \left( |(\beta_i^n)_{(\lambda m)}| \|\sqrt{N}D_i^m X - \beta_i^n\|_\infty \right) &\leq \frac{2}{n} \sum_{i=1}^n \mathbb{E} \left( |(\beta_i^n)_{(\lambda m)}|^2 \right)^{1/2} \mathbb{E} \left( \|\sqrt{N}D_i^m X - \beta_i^n\|_\infty^2 \right)^{1/2} \\ &= o_p(1).\end{aligned}$$

Similarly

$$\frac{1}{n} \sum_{i=1}^n \mathbb{E} \|\sqrt{N} D_i^m X - \beta_i^n\|_\infty^2 = o_p(1).$$

Consequently

$$\frac{1}{n} \sum_{i=1}^n \zeta_i^n(1) \xrightarrow{p} 0,$$

which completes the proof of Theorem 4. ■

**Proof of Theorem 5** As in the previous subsection, we show Theorem 5 in three steps.

**Proof of part I** Recall that we need to prove the stable convergence

$$\bar{U}_n \xrightarrow{d_s} U(\lambda),$$

with

$$\bar{U}_n = \sqrt{\frac{m}{n}} \sum_{i=1}^n \eta_i^n,$$

where  $\eta_i^n$  is given by (49). By simple calculations, it follows that

$$\frac{m}{n} \sum_{i=1}^n \mathbb{E} \left[ |\eta_i^n|^2 \mid \mathcal{F}_{\frac{i-1}{n}} \right] = m \frac{\nu_{2,m}(\lambda) - \nu_{1,m}^2(\lambda)}{\nu_{1,m}^2(\lambda)} \int_0^1 \sigma_u^4 \mathbf{d}u + o_p(1).$$

By standard results for order statistics (see, e.g., David, 1970), we have

$$m \frac{\nu_{2,m}(\lambda) - \nu_{1,m}^2(\lambda)}{\nu_{1,m}^2(\lambda)} \rightarrow \frac{8c_\lambda^2(1-\lambda)(2\lambda-1)}{\phi^2(c_\lambda)(c_\lambda^2 + c_{1-\lambda}^2)^2},$$

as  $m \rightarrow \infty$ . The rest is shown as in the proof of part I of the previous subsection. □

**Proof of part II** As before, we need to prove that

$$U'_n - \bar{U}_n = \sum_{i=1}^n \left( \delta_i^n - \mathbb{E} \left[ \delta_i^n \mid \mathcal{F}_{\frac{i-1}{n}} \right] \right) \xrightarrow{p} 0,$$

where  $\delta_i^n$  is defined by

$$\delta_i^n = \nu_{1,m}^{-1}(\lambda) \sqrt{\frac{m}{n}} \left( q_i^{(n,m)}(\lambda) - w_i^{(n,m)}(\lambda) \right).$$

Again, it suffices to show that

$$\sum_{i=1}^n \mathbb{E} [|\delta_i^n|^2] \rightarrow 0.$$

Next, note that

$$\nu_{1,m}(\lambda) \rightarrow c_\lambda^2 + c_{1-\lambda}^2.$$

By applying Lemma 3, we obtain

$$\begin{aligned}
\sum_{i=1}^n \mathbb{E} [|\delta_i^n|^2] &= \frac{\nu_{1,m}^{-2}(\lambda)}{n} \sum_{i=1}^n \mathbb{E} \left[ |\sqrt{m} \left( q_i^{(n,m)}(\lambda) - w_i^{(n,m)}(\lambda) \right)|^2 \right] \\
&= \frac{\nu_{1,m}^{-2}(\lambda)}{n} \sum_{i=1}^n \mathbb{E} \left[ |\sqrt{m} \left( 2(\beta_i^n)_{(\lambda m)} \left( \sqrt{N} D_i^m X - \beta_i^n \right)_{(\lambda m)*} + \left( \left( \sqrt{N} D_i^m X - \beta_i^n \right)_{(\lambda m)*} \right)^2 \right. \right. \\
&\quad \left. \left. + 2(\beta_i^n)_{(m-\lambda m+1)} \left( \sqrt{N} D_i^m X - \beta_i^n \right)_{(m-\lambda m+1)*} + \left( \left( \sqrt{N} D_i^m X - \beta_i^n \right)_{(m-\lambda m+1)*} \right)^2 \right)|^2 \right] + o_p(1).
\end{aligned}$$

Using similar methods as in the proof of Theorem 4, we deduce that

$$\sum_{i=1}^n \mathbb{E} [|\delta_i^n|^2] = O\left(\frac{m^2}{n}\right) = O(N^{3\gamma-1}),$$

and because  $\gamma \in (0, 1/4)$ , we also get

$$\sum_{i=1}^n \mathbb{E} [|\delta_i^n|^2] \rightarrow 0,$$

which completes the proof of part II. □

**Proof of part III** To derive the assertion of Theorem 5, we must show that

$$\begin{aligned}
&\sqrt{\frac{m}{n}} \sum_{i=1}^n \left( \nu_{1,m}^{-1}(\lambda) \mathbb{E} \left[ q_i^{(n,m)}(\lambda) \mid \mathcal{F}_{\frac{i-1}{n}} \right] - \mathbb{E} \left[ \xi_i^n \mid \mathcal{F}_{\frac{i-1}{n}} \right] \right) \xrightarrow{p} 0, \\
&\sqrt{\frac{m}{n}} \left( \sum_{i=1}^n \mathbb{E} \left[ \xi_i^n \mid \mathcal{F}_{\frac{i-1}{n}} \right] - \int_0^1 \sigma_u^2 du \right) \xrightarrow{p} 0.
\end{aligned}$$

First, note that

$$\sqrt{\frac{m}{n}} \left( \sum_{i=1}^n \mathbb{E} \left[ \xi_i^n \mid \mathcal{F}_{\frac{i-1}{n}} \right] - \int_0^1 \sigma_u^2 du \right) = \sqrt{mn} \sum_{i=1}^n \int_{\frac{i-1}{n}}^{\frac{i}{n}} (\sigma_{\frac{i-1}{n}}^2 - \sigma_u^2) du = O_p \left( \sqrt{\frac{m}{n}} \right),$$

as was shown in Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard (2006) (Section 8) under assumption (V). As  $m/n \rightarrow 0$ , we deduce that

$$\sqrt{\frac{m}{n}} \left( \sum_{i=1}^n \mathbb{E} \left[ \xi_i^n \mid \mathcal{F}_{\frac{i-1}{n}} \right] - \int_0^1 \sigma_u^2 du \right) \xrightarrow{p} 0.$$

Next, recall the definitions of  $\mu_i^n(1)$ ,  $\mu_i^n(2)$  and  $\theta_i^n(1)$ ,  $\theta_i^n(2)$  in (56) - (57). As in the previous subsection, it suffices to prove that

$$\sqrt{\frac{m}{n}} \sum_{i=1}^n \mathbb{E} \left[ \theta_i^n(k) \mid \mathcal{F}_{\frac{i-1}{n}} \right] \xrightarrow{p} 0.$$

for  $k = 1, 2$ . Applying Lemma 3 (as in the proof of part II), we obtain the approximation

$$\begin{aligned}
\sqrt{\frac{m}{n}} \sum_{i=1}^n \mathbb{E} [|\theta_i^n(1)|] &= \nu_{1,m}^{-1}(\lambda) \sqrt{\frac{m}{n}} \sum_{i=1}^n \mathbb{E} \left[ 2(\beta_i^n + \mu_i^n(2))_{(\lambda m)} (\mu_i^n(1))_{(\lambda m)*} + \left( (\mu_i^n(1))_{(\lambda m)*} \right)^2 \right. \\
&\quad \left. + 2(\beta_i^n + \mu_i^n(2))_{(m-\lambda m+1)} (\mu_i^n(1))_{(m-\lambda m+1)*} + \left( (\mu_i^n(1))_{(m-\lambda m+1)*} \right)^2 \right] + o(1).
\end{aligned}$$

Next, note that under assumption (V') we have

$$\mathbb{E}[|a_t - a_s|^p] + \mathbb{E}[|\sigma'_t - \sigma'_s|^p] + \mathbb{E}[|v'_t - v'_s|^p] \leq C_p |t - s|^{\frac{p}{2}}$$

for all  $p > 0$ . Consequently, the Cauchy-Schwarz and Burkholder inequalities imply that

$$\sqrt{\frac{m}{n}} \sum_{i=1}^n \mathbb{E}[|\theta_i^n(1)|] = O\left(\sqrt{\frac{m^2}{n}}\right) + o(1) = O\left(N^{\frac{3\gamma-1}{2}}\right) + o(1).$$

Since  $\gamma \in (0, 1/4)$ , we get

$$\sqrt{\frac{m}{n}} \sum_{i=1}^n \mathbb{E}\left[\theta_i^n(1) \mid \mathcal{F}_{\frac{i-1}{n}}\right] \xrightarrow{p} 0.$$

Similarly, we obtain

$$\begin{aligned} \sqrt{\frac{m}{n}} \sum_{i=1}^n \mathbb{E}\left[\theta_i^n(2) \mid \mathcal{F}_{\frac{i-1}{n}}\right] &= \nu_{1,m}^{-1}(\lambda) \sqrt{\frac{m}{n}} \sum_{i=1}^n \mathbb{E}\left[2(\beta_i^n)_{(\lambda m)} (\mu_i^n(2))_{(\lambda m)*} + \left((\mu_i^n(2))_{(\lambda m)*}\right)^2\right. \\ &\quad \left.+ 2(\beta_i^n)_{(m-\lambda m+1)} (\mu_i^n(2))_{(m-\lambda m+1)*} + \left((\mu_i^n(2))_{(m-\lambda m+1)*}\right)^2 \mid \mathcal{F}_{\frac{i-1}{n}}\right] + o(1). \end{aligned}$$

As was shown in the previous subsection, it holds that

$$\mathbb{E}\left[(\beta_i^n)_{(\lambda m)} (\mu_i^n(2))_{(\lambda m)*} + (\beta_i^n)_{(m-\lambda m+1)} (\mu_i^n(2))_{(m-\lambda m+1)*} \mid \mathcal{F}_{\frac{i-1}{n}}\right] = 0.$$

Moreover, we have

$$\mathbb{E}[|\mu_i^n(2)|_\infty^2] \leq C \frac{m}{n}.$$

Consequently, it follows that

$$\sqrt{\frac{m}{n}} \sum_{i=1}^n \mathbb{E}\left[\theta_i^n(2) \mid \mathcal{F}_{\frac{i-1}{n}}\right] = O_p\left(\sqrt{\frac{m^3}{n}}\right) + o_p(1) = O_p\left(N^{\frac{4\gamma-1}{2}}\right) + o_p(1).$$

Finally,

$$\sqrt{\frac{m}{n}} \sum_{i=1}^n \mathbb{E}\left[\theta_i^n(2) \mid \mathcal{F}_{\frac{i-1}{n}}\right] \xrightarrow{p} 0,$$

because  $\gamma \in (0, 1/4)$ . This completes the proof of part III and Theorem 5. ■

**Proof of Theorem 6** The result is shown in the same manner as the proof of Theorem 5. ■

**Proof of Theorem 7** First, we state the following Lemma, which is a simple consequence of the representation (28).

**Lemma 4** Recall that  $h(x) = x \wedge (1 - x)$ ,  $x \in [0, 1]$ . The following identities hold:

$$\begin{aligned} \mathbb{E}[|N^{1/4} \overline{W}_j^N|^2] &= 8c \int_0^1 h^2(s) ds + O(K^{-1}) = \frac{2c}{3} + O(K^{-1}), \\ \mathbb{E}[|N^{1/4} \overline{u}_j^N|^2] &= \frac{2\omega^2}{c} \int_0^1 \nabla h^2(s) ds + O(K^{-1}) = \frac{2}{c} \omega^2 + O(K^{-1}). \end{aligned}$$

Next, we set

$$\beta_i^{*N} = N^{1/4} \{ \sigma_{\frac{i}{N}} \overline{W}_{i+2(j-1)K}^N + \overline{u}_{i+2(j-1)K}^N \}_{j=1}^m ,$$

and define

$$w_i^{*(n,m)} = g_{\lambda m}^2(\beta_i^{*N}) + g_{m-\lambda m+1}^2(\beta_i^{*N}). \quad (58)$$

By the same methods as presented in Podolskij and Vetter (2007) we conclude that

$$Q_{RV_N^*}(m, \lambda) - \frac{1}{\nu_1(m, \lambda)} \frac{3}{2K\sqrt{N}} \sum_{i=0}^{N-2Km+1} w_i^{*(n,m)} = o_p(1). \quad (59)$$

By representation (28) and Lemma 4 we obtain the (conditional) convergence in distribution

$$\beta_i^{*N} | \mathcal{F}_{\frac{i}{N}} \xrightarrow{d} N_m \left( 0, \text{diag} \left( \frac{2c}{3} \sigma_{\frac{i}{N}}^2 + \frac{2}{c} \omega^2, \dots, \frac{2c}{3} \sigma_{\frac{i}{N}}^2 + \frac{2}{c} \omega^2 \right) \right).$$

From this we deduce that  $\mathbb{E}[w_i^{*(n,m)} | \mathcal{F}_{\frac{i}{N}}] = \nu_1(m, \lambda) (\frac{2c}{3} \sigma_{\frac{i}{N}}^2 + \frac{2}{c} \omega^2) + o_p(1)$  (uniformly in  $i$ ), and consequently the convergence

$$\frac{1}{\nu_1(m, \lambda)} \frac{3}{2K\sqrt{N}} \sum_{i=0}^{N-2Km+1} \mathbb{E}[w_i^{*(n,m)} | \mathcal{F}_{\frac{i}{N}}] \xrightarrow{p} IV + \frac{3\omega^2}{c^2} \quad (60)$$

holds. In view of (59) and (60) we are left to proving

$$\frac{1}{\nu_1(m, \lambda)} \frac{3}{2K\sqrt{N}} \sum_{i=0}^{N-2Km+1} \eta_i^{*N} \xrightarrow{p} 0, \quad \eta_i^{*N} = w_i^{*(n,m)} - \mathbb{E}[w_i^{*(n,m)} | \mathcal{F}_{\frac{i}{N}}]. \quad (61)$$

Observe that due to the construction of  $\beta_i^{*N}$  and the boundedness of  $\sigma$  we obtain the estimate

$$\mathbb{E}[\eta_i^{*N} \eta_j^{*N}] \leq C, \quad |i - j| < 2mK, \quad (62)$$

whereas  $\mathbb{E}[\eta_i^{*N} \eta_j^{*N}] = 0$  for  $|i - j| \geq 2mK$ . Since  $m$  is fixed, we deduce the estimate

$$\mathbb{E} \left[ \left| \frac{1}{\nu_1(m, \lambda)} \frac{3}{2K\sqrt{N}} \sum_{i=0}^{N-2Km+1} \eta_i^{*N} \right|^2 \right] \leq \frac{C}{K}, \quad (63)$$

which completes the proof of Theorem 7. ■

**Proof of Theorem 8** Here we will apply “big blocks & small blocks”-technique. More precisely, we will build big blocks of size  $p$ , which will be separated by a small block of size 1. This procedure ensures the (conditional) independence of big blocks. Later, we let  $p$  converge to infinity.

For this purpose we require some additional notations. First, set

$$a_i(p) = 2i(p+1)mK, \quad b_i(p) = 2i(p+1)mK + 2pmK,$$

and let  $A_i(p)$  denote the set of integers  $l$  with  $a_i(p) \leq l < b_i(p)$  and  $B_i(p)$  the set of integers  $l$  with  $b_i(p) \leq l < a_{i+1}(p)$ . Furthermore, let  $j_N(p)$  denote the largest integer  $j$  with  $b_j(p) \leq N$ . Notice that  $j_N(p) = O(\sqrt{N}/p)$ . Finally, we set  $i_N(p) = 2(j_N(p) + 1)(p+1)mK$ .

Next, we define an approximation of  $\overline{D}_i^N Y$  by

$$\overline{D}_{i,l}^N Y = \{\sigma_{\frac{l}{N}} \overline{W}_{i+2(j-1)K}^N + \overline{u}_{i+2(j-1)K}^N\}_{j=1}^m$$

with  $l \leq i$ , and we set

$$q_{i,l}^*(m, \lambda) = g_{\lambda m}^2(N^{1/4} \overline{D}_{i,l}^N Y) + g_{m-\lambda m+1}^2(N^{1/4} \overline{D}_{i,l}^N Y).$$

We further set

$$\Upsilon_{i,l}^N = q_{i,l}^*(m, \lambda) - \mathbb{E}[q_{i,l}^*(m, \lambda) | \mathcal{F}_{\frac{l}{N}}],$$

and

$$\tilde{\Upsilon}_j^N = \begin{cases} \frac{1}{\nu_1(m, \lambda)} \frac{3}{2K\sqrt{N}} \Upsilon_{j, a_i(p)}^N, & j \in A_i(p) \\ \frac{1}{\nu_1(m, \lambda)} \frac{3}{2K\sqrt{N}} \Upsilon_{j, b_i(p)}^N, & j \in B_i(p) \\ \frac{1}{\nu_1(m, \lambda)} \frac{3}{2K\sqrt{N}} \Upsilon_{j, i_N(p)}^N, & j \geq i_N(p) \end{cases}$$

Finally, we define

$$\zeta(p, 1)_j^N = \sum_{l=a_j(p)}^{b_j(p)-1} \tilde{\Upsilon}_l^N, \quad \zeta(p, 2)_j^N = \sum_{l=b_j(p)}^{a_{j+1}(p)-1} \tilde{\Upsilon}_l^N,$$

and

$$M(p)^N = \sum_{j=0}^{j_N(p)} \zeta(p, 1)_j^N, \quad N(p)^N = \sum_{j=0}^{j_N(p)} \zeta(p, 2)_j^N, \quad C(p)^N = \sum_{j=i_N(p)}^N \tilde{\Upsilon}_j^N$$

Notice that the big blocks are collected in  $M(p)^N$ , the small blocks are contained in  $N(p)^N$  and  $C(p)^N$  is the sum of the border terms.

Now we proceed as follows. By the same methods as presented in Podolskij and Vetter (2007) we can show that

$$N^{1/4} \left( QRV_N^*(m, \bar{\lambda}, \alpha) - \frac{3}{c^2} \hat{\omega}^2 - IV \right) = N^{1/4} (M(p)^N + N(p)^N + C(p)^N) + \gamma_N(p), \quad (64)$$

where  $\gamma_N(p)$  is a sequence of random variables with  $\lim_{p \rightarrow \infty} \limsup_{N \rightarrow \infty} P(|\gamma_N(p)| > \epsilon) = 0$  for any  $\epsilon > 0$ . They also showed that the convergence

$$\lim_{p \rightarrow \infty} \limsup_{N \rightarrow \infty} P(N^{1/4} |N(p)^N + C(p)^N| > \epsilon) = 0 \quad (65)$$

holds. Now we prove that, as  $N \rightarrow \infty$ , the sequence  $N^{1/4} M(p)^N$  converges stably in law to a random variable, defined on an extension of  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , which is Gaussian conditionally on  $\mathcal{F}$  with mean 0 and variance  $\Gamma(p)$ , and

$$\Gamma(p) \xrightarrow{p} \frac{9}{\nu_{1,m}^2(\lambda) c} \sum_{l=1}^m \int_0^1 \int_0^1 f_{m,l,\sigma_t,u}(\lambda) dt du$$

as  $p \rightarrow \infty$ . In view of (64) and (65) the above stable convergence implies the assertion of Theorem 8.

In order to prove the stable convergence for  $N^{1/4} M(p)^N$  we need to show the following conditions (see Theorem IX 7.28 in

Jacod and Shiryaev, 2003) when  $p$  is fixed and  $N \rightarrow \infty$

$$N^{1/2} \sum_{j=0}^{j_N(p)} \mathbb{E}[|\zeta(p, 1)_j^N|^2 | \mathcal{F}_{\frac{a_j(p)}{N}}] \xrightarrow{p} \Gamma(p) \quad (66)$$

$$N \sum_{j=0}^{j_N(p)} \mathbb{E}[|\zeta(p, 1)_j^N|^4 | \mathcal{F}_{\frac{a_j(p)}{N}}] \xrightarrow{p} 0 \quad (67)$$

$$N^{1/2} \sum_{j=0}^{j_N(p)} \mathbb{E}[\zeta(p, 1)_j^N \Delta Z(p)_j^N | \mathcal{F}_{\frac{a_j(p)}{N}}] \xrightarrow{p} 0 \quad (68)$$

$$N^{1/2} \sum_{j=0}^{j_N(p)} \mathbb{E}[\zeta(p, 1)_j^N \Delta V(p)_j^N | \mathcal{F}_{\frac{a_j(p)}{N}}] \xrightarrow{p} 0, \quad (69)$$

where  $\Delta Y(p)_j^N = Y_{\frac{b_j(p)}{N}} - Y_{\frac{a_j(p)}{N}}$  for any process  $Y$ ,  $Z = W$  or  $B$  and (69) holding for any bounded martingale  $V$  that is orthogonal to  $(W, B)$ . Conditions (68) and (69) have already been shown in Podolskij and Vetter (2006) (see Lemma 4 therein), so we concentrate on proving (66) and (67). A straightforward calculation shows that

$$\mathbb{E}[|\zeta(p, 1)_j^N|^4 | \mathcal{F}_{\frac{a_j(p)}{N}}] \leq \frac{Cp^2}{N^2}$$

since  $\sigma$  is bounded. This implies (67) because  $j_N(p) = O(\sqrt{N}/p)$ .

Finally, let us show (66). First, notice that (28) implies the identities

$$N^{1/2} \mathbb{E}[\overline{W}_j^N \overline{W}_i^N] = 8cw_h\left(\frac{|j-i|}{2K}\right) + O(K^{-1}), \quad N^{1/2} \mathbb{E}[\overline{u}_j^N \overline{u}_i^N] = \frac{2}{c} w_{\nabla h}\left(\frac{|j-i|}{2K}\right) \omega^2 + O(K^{-1}), \quad |j-i| < 2K, \quad (70)$$

where the functional  $w_f(u)$  is given in Definition 2. When  $|j-i| \geq 2K$  the above covariances are 0. Assume now that  $j > i$  with  $j = i + 2(l-1)K + d$  for some  $1 \leq l \leq m$  and  $0 \leq d \leq 2K-1$ , and  $a_z(p) \leq i, j \leq b_z(p) - 1$ . Due to the identities in (70) we obtain

$$\mathbb{E}[\Upsilon_{j,a_z(p)}^N \Upsilon_{i,a_z(p)}^N | \mathcal{F}_{\frac{a_z(p)}{N}}] = f_{m,l,\sigma_{a_z(p)}, \frac{d}{2K}}(\lambda) + O_p(K^{-1}),$$

where  $f_{m,l,x,u}(\lambda)$  is given in Definition 2. This implies the identity

$$\begin{aligned} N^{1/2} \sum_{j=0}^{j_N(p)} \mathbb{E}[|\zeta(p, 1)_j^N|^2 | \mathcal{F}_{\frac{a_j(p)}{N}}] &= o_p(1) + \frac{9}{2\nu_{1,m}^2(\lambda)K^2\sqrt{N}} \sum_{j=0}^{j_N(p)} \sum_{l=1}^m \sum_{d=0}^{2K-1} (2pmK - 2(l-1)K - d) f_{m,l,\sigma_{a_j(p)}, \frac{d}{2K}}(\lambda) \\ &\xrightarrow{p} \frac{p}{p+1} \frac{9}{\nu_{1,m}^2(\lambda)c} \sum_{l=1}^m \left(1 - \frac{l-1}{pm}\right) \int_0^1 \int_0^1 f_{m,l,\sigma_t,u}(\lambda) dt du = \Gamma(p). \end{aligned}$$

Clearly, we have

$$\Gamma(p) \xrightarrow{p} \frac{9}{\nu_{1,m}^2(\lambda)c} \sum_{l=1}^m \int_0^1 \int_0^1 f_{m,l,\sigma_t,u}(\lambda) dt du$$

as  $p \rightarrow \infty$ , and the proof is complete. ■

**Proof of Proposition 3** Recall the definition of  $w_i^{*(n,m)}$  in (58). As in the proof of Theorem 7 we have that

$$\begin{aligned} & \frac{9}{4\nu_1^2(m, \lambda)K\sqrt{N}} \sum_{i=2Km-1}^{N-4Km+1} \left( q_i^*(m, \lambda) \left( \sum_{j=i-2Km+1}^{i+2Km-1} \{q_j^*(m, \lambda) - q_{i+2Km}^*(m, \lambda)\} \right) \right. \\ & \left. - 2w_i^{*(n,m)} \left( \sum_{j=i+1}^{i+2Km-1} \{w_j^{*(n,m)} - w_{i+2Km}^{*(n,m)}\} \right) \right) = o_p(1) \end{aligned}$$

Next, notice the identity

$$\begin{aligned} & \frac{9}{2\nu_1^2(m, \lambda)K\sqrt{N}} \sum_{i=2Km-1}^{N-4Km+1} \mathbb{E} \left[ w_i^{*(n,m)} \left( \sum_{j=i+1}^{i+2Km-1} \{w_j^{*(n,m)} - w_{i+2Km}^{*(n,m)}\} \right) \middle| \mathcal{F}_{\frac{i}{N}} \right] \\ &= \frac{9}{2\nu_1^2(m, \lambda)K\sqrt{N}} \sum_{i=2Km-1}^{N-4Km+1} \sum_{l=1}^m \sum_{d=0}^{2K-1} f_{m,l,\sigma_{\frac{i}{N}}, \frac{d}{2K}}(\lambda) + o_p(1) \\ &\xrightarrow{p} \frac{9}{\nu_{1,m}^2(\lambda)c} \sum_{l=1}^m \int_0^1 \int_0^1 f_{m,l,\sigma_t,u}(\lambda) dt du. \end{aligned}$$

On the other hand, we deduce that

$$\begin{aligned} & \frac{9}{2\nu_1^2(m, \lambda)K\sqrt{N}} \sum_{i=2Km-1}^{N-4Km+1} \left( w_i^{*(n,m)} \left( \sum_{j=i+1}^{i+2Km-1} \{w_j^{*(n,m)} - w_{i+2Km}^{*(n,m)}\} \right) \right. \\ & \left. - \mathbb{E} \left[ w_i^{*(n,m)} \left( \sum_{j=i+1}^{i+2Km-1} \{w_j^{*(n,m)} - w_{i+2Km}^{*(n,m)}\} \right) \middle| \mathcal{F}_{\frac{i}{N}} \right] \right) \xrightarrow{p} 0 \end{aligned}$$

as in (61). This completes the proof of Proposition 3.

**Proof of Theorem 9** This result is shown in the same way as the proof of Theorem 8. ■

**Proof of Proposition 4** This result is shown in the same way as the proof of Proposition 3. ■

## References

- Aït-Sahalia, Y., 1996, "Testing continuous-time models of the spot interest rate," *Review of Financial Studies*, 9(2), 385–426.
- Aït-Sahalia, Y., and J. Jacod, 2006, "Testing for Jumps in a Discretely Observed Process," *Annals of Statistics*, forthcoming.
- Aït-Sahalia, Y., P. A. Mykland, and L. Zhang, 2006, "Ultra high frequency volatility estimation with dependent microstructure noise," Working paper, Princeton University.
- Aldous, D. J., and G. K. Eagleson, 1978, "On mixing and stability of limit theorems," *Annals of Probability*, 6(2), 325–331.
- Andersen, T. G., and T. Bollerslev, 1998, "Answering the skeptics: Yes, standard volatility models do provide accurate forecasts," *International Economic Review*, 39(4), 885–905.
- Andersen, T. G., T. Bollerslev, and F. X. Diebold, 2007, "Roughing It Up: Including Jump Components in the Measurement, Modeling, and Forecasting of Return Volatility," *Review of Economics and Statistics*, 89(4), 701–720.
- Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens, 2001, "The distribution of realized stock return volatility," *Journal of Financial Economics*, 61(1), 43–76.
- Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys, 2000, "Great Realizations," *Risk*, 13 (3), 105–108.
- Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys, 2001, "The distribution of realized exchange rate volatility," *Journal of the American Statistical Association*, 96(453), 42–55.
- Andersen, T. G., T. Bollerslev, F. X. Diebold, and J. Wu, 2006, "Realized Beta: Persistence and Predictability," in *Advances in Econometrics: Econometric Analysis of Economic and Financial Time Series in Honor of R.F. Engle, and C.W.J. Granger*, ed. by T. Fomby, and D. Terrell. pp. 1 – 40, Volume B.
- Back, K., 1991, "Asset prices for general processes," *Journal of Mathematical Economics*, 20(4), 317–395.
- Bakshi, G. S., N. Ju, and H. Ou-Yang, 2006, "Estimation of continuous-time models with an application to equity volatility," *Journal of Financial Economics*, 82(1), 227–249.
- Bandi, F. M., and J. R. Russell, 2005, "Microstructure noise, realized variance, and optimal sampling," Working paper, Graduate School of Business, University of Chicago.
- , 2006, "Separating microstructure noise from volatility," *Journal of Financial Economics*, 79(3), 655–692.

- Barndorff-Nielsen, O. E., S. E. Graversen, J. Jacod, M. Podolskij, and N. Shephard, 2006, “A central limit theorem for realized power and bipower variations of continuous semimartingales,” in *From Stochastic Calculus to Mathematical Finance: The Shiryaev Festschrift*, ed. by Y. Kabanov, R. Lipster, and J. Stoyanov. Springer-Verlag, pp. 33–68.
- Barndorff-Nielsen, O. E., S. E. Graversen, J. Jacod, and N. Shephard, 2006, “Limit theorems for bipower variation in financial econometrics,” *Econometric Theory*, 22, 677–719.
- Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard, 2006, “Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise,” Working paper, Nuffield College, University of Oxford.
- Barndorff-Nielsen, O. E., and N. Shephard, 2002, “Econometric analysis of realized volatility and its use in estimating stochastic volatility models,” *Journal of the Royal Statistical Society: Series B*, 64(2), 253–280.
- , 2004a, “Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics,” *Econometrica*, 72(3), 885–925.
- , 2004b, “Power and bipower variation with stochastic volatility and jumps,” *Journal of Financial Econometrics*, 2(1), 1–48.
- , 2006, “Econometrics of testing for jumps in financial economics using bipower variation,” *Journal of Financial Econometrics*, 4(1), 1–30.
- Barndorff-Nielsen, O. E., N. Shephard, and M. Winkel, 2006, “Limit theorems for multipower variation in the presence of jumps,” *Stochastic Processes and Their Applications*, 116(5), 796–806.
- Beckers, S., 1983, “Variances of security price returns based on high, low, and closing prices,” *Journal of Business*, 56(1), 97–112.
- Bollen, B., and B. Inder, 2002, “Estimating daily volatility in financial markets utilizing intraday data,” *Journal of Empirical Finance*, 9(5), 551–562.
- Chernov, M., A. R. Gallant, E. Ghysels, and G. Tauchen, 2003, “Alternative models for stock price dynamics,” *Journal of Econometrics*, 116(1-2), 225–257.
- Christensen, K., and M. Podolskij, 2006, “Range-based estimation of quadratic variation,” Working paper, Aarhus School of Business.
- Corsi, F., G. Zumbach, U. A. Müller, and M. M. Dacorogna, 2001, “Consistent high-precision volatility from high-frequency data,” *Economic Notes*, 30(2), 183–204.

- David, H., 1970, *Order Statistics*. John Wiley & Sons, Inc., New York, NY.
- Diebold, F. X., and G. H. Strasser, 2007, “On the Correlation Structure of Microstructure Noise in Theory and Practice,” manuscript University of Pennsylvania.
- Epps, T. W., 1979, “Comovements in Stock Prices in the Very Short Run,” *Journal of the American Statistical Association*, 74 (366), 291–298.
- Fan, J., and Y. Wang, 2006, “Multi-scale Jump and Volatility Analysis for High-Frequency Financial Data,” manuscript Princeton University.
- Feller, W., 1951, “The asymptotic distribution of the range of sums of independent random variables,” *Annals of Mathematical Statistics*, 22(3), 427–432.
- Fisher, L., 1966, “Some New Stock-Market Indexes,” *Journal of Business*, 39 (1-2), 191–225.
- Garman, M. B., and M. J. Klass, 1980, “On the estimation of security price volatilities from historical data,” *Journal of Business*, 53(1), 67–78.
- Gloter, A., and J. Jacod, 2001a, “Diffusions with measurement errors. I - local asymptotic normality,” *ESAIM: Probability and Statistics*, 5, 225–242.
- , 2001b, “Diffusions with measurement errors. II - measurement errors,” *ESAIM: Probability and Statistics*, 5, 243–260.
- Hansen, P. R., J. Large, and A. Lunde, 2006, “Moving Average-Based Estimators of Integrated Variance,” forthcoming *Econometric Reviews*.
- Hansen, P. R., and A. Lunde, 2006, “Realized variance and market microstructure noise,” *Journal of Business and Economic Statistics*, 24(2), 127–161.
- Huang, X., and G. Tauchen, 2005, “The relative contribution of jumps to total price variance,” *Journal of Financial Econometrics*, 3(4), 456–499.
- Jacod, J., 1994, “Limit of random measures associated with the increments of a Brownian semimartingale,” Preprint number 120, Laboratoire de Probabilités, Université Pierre et Marie Curie, Paris.
- Jacod, J., Y. Li, P. A. Mykland, M. Podolskij, and M. Vetter, 2007, “Microstructure Noise in the Continuous Case: The Pre-Averaging Approach,” manuscript University of Chicago.

- Jacod, J., and P. E. Protter, 1998, "Asymptotic error distributions for the Euler method for stochastic differential equations," *Annals of Probability*, 26(1), 267–307.
- Jacod, J., and A. N. Shiryaev, 2003, *Limit Theorems for Stochastic Processes*. Springer-Verlag, 2 edn.
- Jiang, G. J., and R. C. A. Oomen, 2007, "Testing for Jumps When Asset Prices are Observed with Noise - A "Swap Variance" Approach," manuscript, Warwick Business School.
- Kinnebrock, and M. Podolskij, 2007, "A Note on the Central Limit Theorem for Bipower Variation of General Functions," *Stochastic Processes and Their Applications*, forthcoming.
- Lee, S. S., and P. A. Mykland, 2007, "Jumps in Financial Markets: A New Nonparametric Test and Jump Dynamics," *Review of Financial Studies*, 20, forthcoming.
- Martens, M., and D. van Dijk, 2007, "Measuring volatility with the realized range," *Journal of Econometrics*, 138(1), 181–207.
- Mykland, P. A., 2006, "A Gaussian calculus for inference from high frequency data," manuscript University of Chicago.
- Niederhoffer, V., and M. F. M. Osborne, 1966, "Market Making and Reversal on the Stock Exchange," *Journal of the American Statistical Association*, 61 (316), 897–916.
- Nolte, I., and V. Voev, 2007, "Estimating High-Frequency Based (Co-) Variances: A Unified Approach," manuscript University of Konstanz.
- Oomen, R. C., 2006a, "Comment on 2005 JBES invited address "Realized variance and market microstructure noise" by Peter R. Hansen and Asger Lunde," *Journal of Business & Economic Statistics*, 24 (2), 195 – 202.
- , 2006b, "Properties of Realized Variance under Alternative Sampling Schemes," *Journal of Business & Economic Statistics*, 24 (2), 219 – 237.
- Parkinson, M., 1980, "The extreme value method for estimating the variance of the rate of return," *Journal of Business*, 53(1), 61–65.
- Podolskij, M., and M. Vetter, 2006, "Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps," Working paper, CREATES, University of Aarhus and Ruhr University of Bochum.
- , 2007, "Bipower-type estimation in a noisy diffusion setting," Working paper, CREATES, University of Aarhus and Ruhr University of Bochum.
- Protter, P. E., 2004, *Stochastic Integration and Differential Equations*. Springer-Verlag, 1 edn.

- Rényi, A., 1963, “On stable sequences of events,” *Sankhya: The Indian Journal of Statistics; Series A*, 25(3), 293–302.
- Rogers, L. C. G., and S. E. Satchell, 1991, “Estimating variances from high, low, and closing prices,” *Annals of Applied Probability*, 1(4), 504–512.
- Schwert, G. W., 1989, “Why does stock market volatility change over time,” *Journal of Finance*, 44(5), 1115–1153.
- Sun, Y., 2006, “Best Quadratic Unbiased Estimators of Integrated Variance in the Presence of Market Microstructure Noise,” manuscript University of California, San Diego.
- Woerner, J. H. C., 2006, “Power and multipower variation: inference for high frequency data,” in *Proceedings of the International Conference on Stochastic Finance 2004*, ed. by A. N. Shiryaev, M. do R. Grossihno, P. Oliviera, and M. Esquivel, pp. 343–364, Berlin. Springer-Verlag.
- Zhang, L., 2006, “Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach,” *Bernoulli*, 12(6), 1019–1043.
- Zhang, L., P. A. Mykland, and Y. Aït-Sahalia, 2005, “A tale of two time scales: determining integrated volatility with noisy high-frequency data,” *Journal of the American Statistical Association*, 100(472), 1394–1411.
- Zhou, B., 1996, “High-frequency data and volatility in foreign-exchange rates,” *Journal of Business and Economic Statistics*, 14(1), 45–52.