# Efficient and feasible inference for the components of financial variation using blocked multipower variation

Per A. Mykland

*Department of Statistics, University of Chicago*
*Eckhart Hall, 5734 University Avenue, Chicago, IL 60637, USA*
*mykland@pascal.uchicago.edu*

Neil Shephard

*Nuffield College, New Road, OX1 1NF, UK,*
*Department of Economics, University of Oxford,*
*&*
*Oxford-Man Institute, University of Oxford*
`neil.shephard@economics.ox.ac.uk`

Kevin Sheppard

*Oxford-Man Institute, University of Oxford,*
*Eagle House, Walton Well Road, Oxford OX2 6ED, UK*
*&*
*Department of Economics, University of Oxford*
`kevin.sheppard@economics.ox.ac.uk`

21st February, 2012

**Abstract**

High frequency financial data allows us to learn more about volatility, volatility of volatility and jumps. One of the key techniques developed in the literature in recent years has been bipower variation and its multipower extension, which estimates time-varying volatility robustly to jumps. We improve the scope and efficiency of multipower variation by the use of a more sophisticated exploitation of high frequency data. This suggests very significant improvements in the power of jump tests. It also yields efficient estimates of the integrated variance of the continuous part of a semimartingale. The paper also shows how to extend the theory to the case where there is microstructure in the observations and derive the first nonparametric high frequency estimator of the volatility of volatility. A fundamental device in the paper is a new type of result showing path-by-path (strong) approximation between multipower and the (unobserved) RV based on the continuous part of the process.

## 1   Introduction

High frequency data has been demonstrated to improve our ability to understand and forecast financial volatility (e.g. Andersen, Bollerslev, Diebold, and Labys (2001) and Barndorff-Nielsen and Shephard (2002)). Nonparametric estimators like realised volatility have also been shown,

when combined with implied volatility, to powerfully explain variations in the cross sectional and temporal behaviour of risk premia (e.g. Bollerslev, Tauchen, and Zhou (2009)). Further, multipower variation and threshold tests based on high frequency data have shown convincingly that relatively frequent jumps play an important role in the evolution of price processes for commonly held assets (e.g. Barndorff-Nielsen and Shephard (2006), Huang and Tauchen (2005), Andersen, Bollerslev, and Diebold (2007) and Patton and Sheppard (2009)) while jumps play an important role in determining extreme moves in financial markets (e.g. Bollerslev and Todorov (2011)).

All of the above methods are based on the following framework. Suppose there is an underlying efficient price process $X$, which will be assumed to be a semimartingale due to the absence of arbitrage (e.g. Delbaen and Schachermeyer (2006)). We observe $n$ returns over some fixed interval of time, say time 0 to time $T$, typically representing a day. Then a natural ex-post measure of its variation is $X$'s quadratic variation (QV) (e.g. Andersen, Bollerslev, and Diebold (2009) and Jacod and Shiryaev (2003)). Some of the literature on this topic has formally allowed for the effect of market microstructure effects on the analysis (e.g. Zhang, Mykland, and Aït-Sahalia (2005), Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008) and Jacod, Li, Mykland, Podolskij, and Vetter (2009)), others use moderate frequency data such as five minute returns in order to ameliorate the effect (e.g. Andersen, Bollerslev, Diebold, and Labys (2000) and Bandi and Russell (2008)). Surveys of the mathematical work include Podolskij and Vetter (2010), Mykland and Zhang (2012) and Jacod and Protter (2012)

In this paper we focus on multipower type estimators and tests. Most of the time we will be using these statistics to estimate and make inference on the quadratic variation of $X$, $[X, X]_T$, and the quadratic variation of the continuous part of $X$, $[X^c, X^c]_T$. Multipower variation looks inside quadratic variation, splitting up the variation into that due to continuous evolution of prices and that due to jumps. It was introduced by Barndorff-Nielsen and Shephard (2004) and Barndorff-Nielsen and Shephard (2006). The choice of which power $K$ determines the configuration of the statistic, with $K = 1$ being realised variance, $K = 2$ being bipower variation, etc.

The contribution of this paper is to extend multipower variation type estimators, allowing them to be statistically efficient in the sense they achieve the non-parametrically efficiency bounds for the problem. We call this new estimator "blocked multipower variation".

An additional result from our work is the first non-parametric estimator of the volatility of volatility, or more formally the quadratic variation of the volatility process $[\sigma, \sigma]_T$. We also generalise our blocked multipower variation statistic to the case where we use miniature realised measures as the inputs into multipower variation rather than simple realised variances. This allows the estimator to be applied in the important case of market microstructure noise.

The ideas here echo Mykland and Zhang (2009) who show how to justify the use of blocks of length $M$ of high frequency returns. Econometrically we go much further however. Our results generate a path-by-path analysis of the limiting behaviour of high frequency estimators. This is new for the literature on high frequency financial econometrics and this type of novel analysis may have very wide applicability outside the scope of this paper.

To put this in context suppose log-prices have two components

$$X_t = X_t^c + J_t, \quad t \in \mathcal{R}_{\geq 0},$$

where $X_t^c$ is a continuous semimartingale $X_t^c = X_0 + \int_0^t \mu_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s$ and $J$ is a purely discontinuous finite activity jump process. Write $RV^*$ as the infeasible realised variance of the $X^c$ process. So $RV^*$ has the jumps stripped out of it and so, as long as $\sigma_t > 0$, we have that

$$n^{1/2} \left( RV - RV^* - \sum_{0 < t \leq T} |\Delta J_t|^2 \right) \overset{\mathcal{L}s}{\to} N \left( 0, 4T \sum_{0 < t \leq T} |\Delta J_t|^2 \sigma_t^2 \right), \tag{1}$$

where $\overset{\mathcal{L}s}{\to}$ denotes stable convergence as $n \to \infty$.[1] We will prove that the jump variation $JV = \sum_{0 < t \leq T} |\Delta J_t|^2$ cannot be estimated with a lower normalised variance than the scaled asymptotic variance of $RV - RV^*$

$$4T \sum_{0 < t \leq T} |\Delta J_t|^2 \sigma_t^2, \tag{2}$$

while it is well known that integrated variance $IV = \int_0^T \sigma_t^2 \mathrm{d}t$ cannot be estimated with a lower normalised variance than $2 \int_0^T \sigma_t^4 \mathrm{d}t$, a bound which is obtained by $RV^*$ (e.g. Barndorff-Nielsen and Shephard (2002)). Of course neither is immediately feasible as we do not know $RV^*$. This will be where blocked multipower variation comes in.

If the size of the block length $M$ of our new multipower variation estimator $MV_M^{(K)}$ increases as $M \approx cn^{1-\beta}$ then we report in Theorem 4 that

$$n^{1/2} \left( MV_M^{(K)} - RV^* \right) = Kc^{\frac{K-1}{K}} n^{\frac{1}{2} - \beta \frac{K-1}{K}} \sum_{0 < t \leq T} \sigma_t^{2\frac{K-1}{K}} \Delta J_t^{2/K} \tag{3}$$

$$- n^{-\beta + \frac{1}{2}} \frac{K-1}{3} c[\sigma, \sigma]_T - n^{1/2} \Upsilon_n \tag{4}$$
$$+ O_p(n^{\frac{1}{2} - \frac{3}{2}\beta} + n^{\frac{1}{2}(\beta - 1)}) + o_p(n^{\frac{1}{2} - \beta \frac{K-1}{K}}),$$

where $[\sigma, \sigma]_T$ is the volatility of volatility and $n^{1/2} \Upsilon_n$ is an edge effect which is $O_p(n^{-\beta + \frac{1}{2}})$. Thus we quantify the path-by-path difference between the multipower and $RV^*$ statistics and the impact of the jumps, edge effects and the quadratic variation of volatility. This is a significant strengthening

---

[1] For a discussion of stable convergence see Section 2 of Jacod and Protter (1998) and Section 2.2 of Mykland and Zhang (2009), which draws on Rootzén (1980). The latter is relevant for our proofs, starting with that of Theorem 4.

over a previous result in Barndorff-Nielsen, Shephard, and Winkel (2006) which showed that when $K \geq 3$ that $n^{1/2}\left(MV_1^{(K)} - MV_1^{(K)*}\right) = o_p(1)$, where $MV_1^{(K)*}$ is the infeasible multipower variation statistic applied to the $X_t^c$ process. Of course $MV_1^{(K)*}$ is not asymptotically equivalent to $RV^*$, so the results are distinct.

The theory also shows that $MV_M^{(K)}$ and $RV^*$ are asymptotically equivalent under jumps if $\beta \in (1/2, 1)$ and $\beta \frac{K-1}{K} > 1/2$. This implies, as $K$ is an integer, we must have at least $K \geq 3$. A combination which has an attractive optimality property is $\beta = 6/7$ (so the block size increases slowly with the sample size) when $K = 3$. Asymptotically the jump term has the dominant larger term, hence the rates of convergence to zero of the two terms in (3) become similar as $K$ increases. These results mean that under the presence of jumps

$$n^{1/2}\left(MV_M^{(K)} - \int_0^T \sigma_t^2 \mathrm{d}t\right) \overset{\mathcal{L}_s}{\to} N\left(0, 2T \int_0^T \sigma_t^4 \mathrm{d}t\right) \tag{5}$$

and

$$n^{1/2}\left\{\left(RV - MV_M^{(K)}\right) - \sum_{0 < t \leq T} |\Delta J_t|^2\right\} \overset{\mathcal{L}_s}{\to} N\left(0, 4T \sum_{0 < t \leq T} |\Delta J_t|^2 \sigma_t^2\right). \tag{6}$$

We thus demonstrate that this feasible approach to estimating $IV$ and $JV$ is efficient, and is helpful in generating narrow confidence intervals for the components of quadratic variation — for we can consistently estimate the "integrated quarticity" $IQ = \int_0^T \sigma_t^4 \mathrm{d}t$ and "squared jump times vol" $SJV = \sum_{0 < t \leq T} |\Delta J_t|^2 \sigma_t^2$ using block multipower statistics.

One of the interesting features of (3) is that it indicates the squared jumps appear in the asymptotics in a simple additive way, which means their effect depends roughly on the size of the quadratic variation of the jumps rather than, for example, the size of the largest jumps. This makes us speculate that this asymptotic equivalence result may also hold for infinite activity processes. One indication in this direction are the results of Veraart (2010), who studies the estimation of $JV$ by looking at the difference between $RV$ and classical $MV$ (block size 1) also for infinite activity processes, and obtains results that are similar to ours to the extent that our two setups coincide. However, her single block size means that her inference does not obtain the nonparametric efficiency bound.

One of the uses of multipower variation is to nonparametrically test for jumps over the interval $[0, T]$. Under the null of no jumps as $M \approx cn^{1-\beta}$, so long as $\beta > 1/2$, then we show that

$$J = \frac{n^{1-\beta/2}\left(RV - MV_M^{(K)}\right)}{\sqrt{T\frac{3}{4c}\int_0^T \sigma_s^4 \mathrm{d}s}} \overset{\mathcal{L}_s}{\to} N(0, 1).$$

The null is rejected when we see a large value of $J$. This result shows the asymptotic power of the Barndorff-Nielsen and Shephard (2006) jump test can be asymptotically improved by an infinite amount by using blocked multipower variation.

Alternative methods for looking at jumps include the threshold type tests instigated by Mancini (2004) and further developed by, for example, Mancini (2009), Jacod (2008), Jacod and Protter (2012), Lee and Mykland (2008) and Jacod (2012). Christensen and Podolskij (2006) have looked at bipower objects based upon ranges while Christensen, Oomen, and Podolskij (2010) focus on quantiles and Andersen, Dobrev, and Schaumburg (2009) use a nearest neighbour method.

This paper has the following structure. In Section 2 we make our notation clear and review the standard bipower variation. In Section 3 we extend bipower to allow for more efficiency and discuss the properties of the resulting statistic. In Section 3.5 we provide the technically most challenging section which allows the block size of the new efficient multipower power statistic to go to infinity with the sample size.

In Section 4 we carry out a detailed Monte Carlo analysis of blocked multipower variation, employing an empirically realistic data generating process, in order to understand the finite sample performance of our approach. In Section 5 we apply our blocked multipower variation methods to an exchange traded fund that tracks the S&P 500 Index. In Section 6 we collect extensions of our analysis which formally deal with the impact of market microstructure effects and provides the first non-parametric estimator of the volatility of volatility. In Section 7 we draw some conclusions and the Appendix has the proofs of various results we give in the main text of the paper.

## 2 Framework

### 2.1 Model and measures of variation

We will work with a univariate Brownian semimartingale (Ito process) model for log-prices defined on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$,

$$X_t = X_0 + \int_0^t \mu_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s, \tag{7}$$

where $\mu$ and $\sigma$ are predictable locally bounded drift and volatility processes, and $W$ is a standard Brownian motion adapted to $(\mathcal{F}_t)$. See, for example, Ghysels, Harvey, and Renault (1996).

Our focus will be on econometric estimators based on returns. We assume prices are recorded at times $0 = t_0, t_1, ..., t_n = T$, which we will assume are equally spaced

$$t_i = \frac{i}{n}T, \quad i = 0, 1, ..., n.$$

These returns will be written as

$$\Delta X_{t_i} = X_{t_i} - X_{t_{i-1}}.$$

The standard realised variance (RV) is

$$RV = \sum_{0 < t_i \leq T} \left( \Delta X_{t_i} \right)^2,$$

while its square root is usually labelled the realised volatility. The bipower variation statistic of Barndorff-Nielsen and Shephard (2004) is based on

$$BV = \frac{1}{k_{1,1}^2} \sum_{0 < t_{i-1}, t_i \leq T} |\Delta X_{t_i}| |\Delta X_{t_{i-1}}|,$$

where $k_{1,1}^2 = \frac{2}{\pi} \simeq 0.63661$, which is a special case of

$$k_{M,r} = \mathrm{E}\left\{(\chi_M)^r\right\} = 2^{r/2} \frac{\Gamma\left(\frac{M+r}{2}\right)}{\Gamma\left(\frac{M}{2}\right)}, \quad \text{where} \quad \chi_M \sim \left|\chi_M^2\right|^{1/2}. \tag{8}$$

Remark 2 in Section 3.2 will note some properties of $k_{M,r}$ as $M$ increases.

This bipower statistic has been extensively used as it is robust to jumps so that if the $X_t$ process also has jumps (in addition to the dynamics given in (7)),

$$BV \xrightarrow{p} \int_0^T \sigma_s^2 \mathrm{d}s$$

under very weak conditions. This compares with the celebrated result

$$RV \xrightarrow{p} \int_0^T \sigma_s^2 \mathrm{d}s + \sum_{0 \leq t \leq T} (X_t - X_{t-})^2 = IV + JV,$$

the sum of integrated variance and jump variation.

The idea of bipower variation has been generalised in a number of directions, which are collectively now called multipower variation (e.g. Barndorff-Nielsen and Shephard (2006) and Barndorff-Nielsen, Graversen, Jacod, and Shephard (2006)). Its target is the more general object $\int_0^T \sigma_t^r \mathrm{d}t$, $r > 0$. Typically focus is on $r = 2$, which we have discussed before, and $r = 4$ which is useful for carrying out inference as integrated quarticity appears in the standard error of a large number of econometrically interesting quantities (e.g. Andersen, Dobrev, and Schaumburg (2012)). The multipower extension of bipower variation is

$$MV^{(r,K)} = \frac{1}{\left(k_{1,r/K}\right)^K} \sum_{0 < t_{i-k} \leq 1} \prod_{k=0}^{K-1} |\Delta X_{t_{i-k}}|^{r/K}, \quad k_{M,s} = \mathrm{E}\left\{(\chi_M)^s\right\},$$

noting that when $r = 2$ we have that $K = 1$ delivers RV, $K = 2$ being BV and $K = 3$ being tripower variation. So long as $K > 1$ then this statistic is robust to jumps and $MV^{(r,K)} \xrightarrow{p} \int_0^T \sigma_s^r \mathrm{d}s$. In the rest of this section we will focus on RV and BV, returning to the more general case later.

## 2.2 Testing for jumps using bipower variation

It is now well known (Barndorff-Nielsen and Shephard (2006) and, e.g., Barndorff-Nielsen, Graversen, Jacod, and Shephard (2006)) that under (7)

$$\sqrt{n} \left( \begin{array}{c} RV - \int_0^T \sigma_s^2 \mathrm{d}s \\ BV - \int_0^T \sigma_s^2 \mathrm{d}s \end{array} \right) \xrightarrow{\mathcal{L}_s} N\left(0, \left( \begin{array}{cc} 2 & 2 \\ 2 & \theta_1 \end{array} \right) T \int_0^T \sigma_s^4 \mathrm{d}s \right),$$

where $\mathcal{L}s$ denotes convergence in law stably. Here $\theta_1 \simeq 2.6090$ and is given as a special case of

$$
\begin{aligned}
\theta_M &= \frac{1}{k_{M,1}^4} \left[ \text{Var} \left( \chi_{M,1} \chi_{M,2} \right) + 2\text{Cov} \left( \chi_{M,1} \chi_{M,2}, \chi_{M,3} \chi_{M,2} \right) \right] \\
&= \frac{1}{k_{M,1}^4} \left[ \left\{ \text{E} \left( \chi_M^2 \right) \right\}^2 - k_{M,1}^4 + 2 \left( k_{M,1}^2 \text{E} \left( \chi_M^2 \right) - k_{M,1}^4 \right) \right] \\
&= \frac{1}{k_{M,1}^4} \left( M^2 + 2k_{M,1}^2 M - 3k_{M,1}^4 \right),
\end{aligned} \tag{9}
$$

where $\left\{ \chi_{M,i} \right\}$ are independent with law $\chi_M$.

Barndorff-Nielsen and Shephard (2006) suggested using the difference of these statistics

$$
\sqrt{n} \left( RV - BV \right) \overset{\mathcal{L}s}{\to} N \left( 0, T \left( \theta_1 - 2 \right) \int_0^T \sigma_s^4 \mathrm{d}s \right),
$$

as a basis for a Hausman-type test for jumps as $BV$ is robust to jumps. The resulting test records

$$
J = \frac{\sqrt{n} \left( RV - BV \right)}{\sqrt{T \left( \theta_1 - 2 \right) \widehat{IQ}}} \overset{\mathcal{L}s}{\to} N(0,1),
$$

where $\widehat{IQ}$ is a jump robust consistent estimator of integrated quarticity $IQ = \int_0^T \sigma_s^4 \mathrm{d}s$. This method has been used and extended extensively empirically, e.g. Huang and Tauchen (2005), Andersen, Bollerslev, and Diebold (2007) and Lee and Mykland (2008).

The above results show that $BV$ is not as efficient as $RV$ when there are no jumps. It opens up the question of whether there is a simple generalisation of $BV$ which is significantly more efficient and can this test be made more efficient as a result?

## 2.3 A different view of bipower variation

Before we introduce our generalisation of $BV$ it is helpful to interpret $BV$ in the following way. Think of $|\Delta X_{t_i}|$ as roughly in law $\chi_1 n^{-1/2} \sigma_{t_{j-1}}$, while $\int_{\tau_{i-1}}^{\tau_i} \sigma_t^2 \mathrm{d}t$ is roughly estimated by

$$
RV_i = \sum_{\tau_{i-1} < t_j \leq \tau_i} (\Delta X_{t_j})^2, \quad i = 1, ..., n.
$$

That is the absolute value of returns is the square root of a very local realised variance estimator — computing $RV_i$ over only a single observation. Then we can write

$$
BV = \frac{1}{k_{1,1}^2} \sum_{i=2}^n |RV_i|^{\frac{1}{2}} |RV_{i-1}|^{\frac{1}{2}}.
$$

There are then two natural questions: (i) Are there gains to be made by taking longer blocks to compute a better $RV_i$, before computing $BV$? (ii) Would the block version still be robust to jumps?

# 3    Improving the efficiency of multipower variation

## 3.1    Blocked bipower variation

Set temporal block boundaries as

$$\tau_i = M\frac{i}{n}T, \quad i = 0, 1, ..., \lfloor n/M \rfloor = n_M,$$

where $M$, the number of high frequency returns within a block, is a positive integer so $M \leq n$. Then we define the (non-overlapping) $i$-th blocked realised variance as

$$RV_i = \sum_{\tau_{i-1} < t_j \leq \tau_i} (\Delta X_{t_j})^2, \quad i = 1, 2, ..., n_M. \tag{10}$$

This computes the realised variance using all $M$ high frequency observations inside the block $(\tau_{i-1}, \tau_i]$ and estimates $\int_{\tau_{i-1}}^{\tau_i} \sigma_s^2 \mathrm{d}s$. Thus if $M = 1$ we get a single squared return, while if $M = n$ we get back to the full sample RV.

If we do this for all feasible $i$, then we have a time series, within the day, of non-overlapping *miniature realised variances* $RV_1, RV_2, ..., RV_{n_M}$. Each of these is quite noisy, but they can be averaged. If they were summed we reproduce the realised variance $RV = \sum_{i=1}^{n_M} RV_i$. Our focus is on the *block bipower variation*, which we define as

$$BV_M = \frac{M}{k_{M,1}^2} \sum_{i=2}^{n_M} (RV_{i-1} RV_i)^{1/2}, \tag{11}$$

where $k_{M,1}$ is given in (8).

In practice it is better to use an alternative definition

$$BV_M = \frac{n}{(n_M - 1) k_{M,1}^2} \sum_{i=2}^{n_M} (RV_{i-1} RV_i)^{1/2}. \tag{12}$$

$BV_M$ is exactly unbiased for constant $\sigma_t$, see Remark 1 below. For most of the paper, we shall not distinguish between (11) and (12) since $n/(n_M - 1) = M + O(M/n)$. In particular, the two statistics are asymptotically equivalent for fixed $M$.

**Remark 1** *If $X$ is scaled Brownian motion $\sigma W$, and with the definition (12), then $\mathrm{E}(BV_M) = \sigma^2 T$, $\mathrm{Var}(BV_M) = T^2 \sigma^4 d \sim T^2 \frac{\sigma^4}{n} M \theta_M$, where $d = \frac{1}{M(n_M - 1)} \left( M\theta_M - \frac{1}{(n_M - 1)} \psi_M \right)$ and $\psi_M = 2M \left( M/k_{M,1}^2 - 1 \right) = 1 + \frac{1}{4M} + O(M^{-2})$.*

**Proof.** Given in the Appendix.

## 3.2    Properties of $BV_M$ for fixed $M$

Under the more general Brownian semimartingale conditions set out in Appendix B.1, we produce:

**Theorem 1** *Under Assumption 1-2 in Appendix B.1, for fixed $M$ as $n \to \infty$:*

$$n^{1/2}\left(BV_M - \int_0^T \sigma_t^2 dt\right) \overset{\mathcal{L}_s}{\to} N\left(0, TM\theta_M \int_0^T \sigma_s^4 \mathrm{d}s\right), \tag{13}$$

$$n^{1/2}\left(RV - BV_M\right) \overset{\mathcal{L}_s}{\to} N\left(0, T\left(M\theta_M - 2\right) \int_0^T \sigma_s^4 \mathrm{d}s\right). \tag{14}$$

**Proof.** Given in the Appendix.

**Remark 2** *As $M \to \infty$ then*

$$\frac{k_{M,r}}{M^{r/2}} \sim 1 + \frac{r(r-2)}{4M} + \frac{r(r-2)(3r^2 - 14r + 8)}{4 \times 24 M^2} = 1 + \frac{a_r}{M} + \frac{b_r}{M^2},$$

$$\frac{M^{r/2}}{k_{M,r}} \sim 1 - \frac{a_r}{M} + \frac{a_r^2 - b_r}{M^2} = 1 + \frac{c_r}{M} + \frac{d_r}{M^2},$$

$$\frac{M^r}{k_{M,r}^2} \sim 1 + \frac{2c_r}{M} + \frac{c_r^2 + 2d_r}{M^2}, \quad \frac{k_{M,r}^2}{M^r} \sim 1 + \frac{2a_r}{M} + \frac{a_r^2 + 2b_r}{M^2}.$$

*Table 1 gives some useful special cases of these formula. These results mean that*

| | $\frac{k_{M,r}}{M^{r/2}} - 1$ $\simeq \frac{a_r}{M} + \frac{b_r}{M^2}$ | | $\frac{M^{r/2}}{k_{M,r}} - 1$ $\simeq \frac{c_r}{M} + \frac{d_r}{M^2}$ | | $\frac{M^r}{k_{M,r}^2} - 1$ $\simeq \frac{e_r}{M} + \frac{f_r}{M^2}$ | | $\frac{k_{M,r}^2}{M^r} - 1$ $\simeq \frac{g_r}{M} + \frac{h_r}{M^2}$ | |
|---|---|---|---|---|---|---|---|---|
| | $a_r$ | $b_r$ | $c_r$ | $d_r$ | $e_r$ | $f_r$ | $g_r$ | $h_r$ |
| $r=8$ | 12 | 44 | -12 | 100 | -24 | 344 | 24 | 232 |
| $r=4$ | 2 | 0 | $-2$ | 4 | $-4$ | 12 | 4 | 4 |
| $r=8/3$ | $\frac{4}{9}$ | $-\frac{4}{27}$ | | | | | | |
| $r=2$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $r=4/3$ | $-\frac{2}{9}$ | $\frac{4}{81}$ | $\frac{2}{9}$ | 0 | $\frac{4}{9}$ | $\frac{4}{81}$ | $-\frac{4}{9}$ | $\frac{4}{27}$ |
| $r=1$ | $-\frac{1}{4}$ | $\frac{1}{32}$ | $\frac{1}{4}$ | $\frac{1}{32}$ | $\frac{1}{2}$ | $\frac{1}{8}$ | $-\frac{1}{2}$ | $\frac{1}{8}$ |
| $r=2/3$ | $-\frac{2}{9}$ | 0 | $\frac{2}{9}$ | $\frac{4}{81}$ | $\frac{4}{9}$ | $\frac{17}{27}$ | $-\frac{4}{9}$ | $\frac{4}{81}$ |
| $r=1/2$ | $-\frac{3}{16}$ | $-\frac{7}{16\times32}$ | $\frac{3}{16}$ | $\frac{25}{16\times32}$ | $\frac{3}{8}$ | $\frac{17}{8\times16}$ | $-\frac{3}{8}$ | $\frac{1}{16\times8}$ |

Table 1: The terms in the Taylor expansion of the scaled $k_{M,r}$ function.

$$M\theta_M = M\left\{\left(\frac{M^2}{k_{M,1}^4} - 1\right) + 2\left(\frac{M}{k_{M,1}^2} - 1\right)\right\} \simeq 2 + \frac{3}{4M}.$$

*This implies that $BV_M$ is roughly efficient if $M$ is large but fixed and $n^{1/2}\left(RV - BV_M\right)$ has an asymptotic variance which is roughly $T\frac{3}{4M}\int_0^T \sigma_s^4 \mathrm{d}s$ if $M$ is large but fixed. Table 2 shows the relative efficiency of $BV_M$ as $M$ varies. By the time $M = 3$ there is a modest reduction in the asymptotic variance, but that this gain is two-thirds of all the potential gains from using the blocking. It makes realised BV nearly as efficient as RV. Under the null of no jumps $RV - BV_M$ has a much smaller variance which should improve $J_M$'s power as a test.*

**Proof.** Is straightforward but given in the Appendix for completeness.

| $M$ | 1 | 2 | 3 | 5 | 10 | 25 | 50 | $\infty$ |
|---|---|---|---|---|---|---|---|---|
| $M\theta_M$ | 2.609 | 2.335 | 2.232 | 2.143 | 2.073 | 2.029 | 2.014 | 2 |
| $M\theta_M - 2$ | .609 | .335 | .232 | .143 | .073 | .029 | .014 | 0 |
| $\frac{3}{4M}$ | .750 | .375 | .250 | .150 | .075 | .030 | .015 | 0 |

Table 2: Efficiency of the blocked bipower variation. The asymptotic efficiency is determined by $M\theta_M$, with the best number being 2. In testing for jumps the efficiency is determined by $M\theta_M - 2$.

**Remark 3** *If there are finitely many jumps in addition to the continuous part of $X$, this estimator is unaffected by these, asymptotically. Hence it is similar to bipower. This is is because if there is no jump then $\sqrt{RV_{i-1}} = O_p\left(\sqrt{\frac{M}{n}}\right)$. Hence so long as $M/n \to 0$ then under finite activity jumps $BV_M \stackrel{p}{\to} \int_0^T \sigma_s^2 \mathrm{d}s$. A more detailed analysis of jumps will appear in Section 3.5.*

The implication of this is that the block bipower test takes the form, for fixed $M$, of

$$J_M = \frac{n^{1/2}(RV - BV_M)}{\sqrt{T(M\theta_M - 2)\widehat{IQ}}} \stackrel{\mathcal{L}_s}{\to} N(0,1), \tag{15}$$

as $n \to \infty$, rejecting the null of no jump if $J_M$ is significantly positive.

**Remark 4** *When $M \sim (c/T)n^{1-\beta}$ with $\beta > 1/2$, then the asymptotic variance in (14) suggests informally that $n^{1/2}(RV - BV_M)$ has an asymptotic form $T(M\theta_M - 2) \simeq 3T^2/\{4cn^{1-\beta}\}$, and so we might predict that $n^{1-\beta/2}(RV - BV_M) \stackrel{Ls}{\to} N\left(0, \frac{3}{4c}T^2\int_0^T \sigma_s^4 \mathrm{d}s\right)$. This prediction turns out to be exactly correct when $\beta > 1/2$ where $M$ increases with $n$. So $BV_M$ is essentially efficient for large $M$. See Section 3.5 and Appendix B.2.6 for exactly matching results when $M \to \infty$ as $n \to \infty$.*

### 3.3 Blocked multipower variation

The above arguments extend to blocked multipower variation

$$MV_M^{(r,K)} = \left(\frac{n}{T}\right)^{\frac{r}{2}-1} \frac{M}{\left(k_{M,r/K}\right)^K} \sum_{i=K}^{n_M} \prod_{k=0}^{K-1} (RV_{i-k})^{r/2K}, \quad \text{where} \tag{16}$$

$$k_{M,r} = \mathrm{E}\{(\chi_M)^r\} = 2^{r/2}\frac{\Gamma\left(\frac{M+r}{2}\right)}{\Gamma\left(\frac{M}{2}\right)}.$$

When $K = 3$ this is called a blocked tripower estimator, while when $K = 4$ it is a blocked quadpower estimator. Both are used frequently empirically when $M = 1$. As with (12), when $M$ is moderately large, one is better off using a finite sample exact constant:

$$MV_M^{(r,K)} = \frac{n^{\frac{r}{2}}}{T^{\frac{r}{2}-1}} \frac{1}{(n_M - K + 1)(k_{M,r/K})^K} \sum_{i=K}^{n_M} \prod_{k=0}^{K-1} (RV_{i-k})^{r/2K}. \tag{17}$$

The two versions are equivalent when $M/n = O_p(1)$, but (17) is preferred for moderate $M$.

We now look at relative efficiency as a function of $M$ and $K$.

10

**Theorem 2** *Assume that $\sigma$ is constant. Then the $RV_i$ are i.i.d. $\chi_M^2 \sigma^2 T/n$, and so*

$$MV_M^{(r,K)} \stackrel{\mathcal{L}}{=} \sigma^r T \frac{1}{(n_M - K + 1)(k_{M,r/K})^K} \sum_{i=K}^{n_M} \prod_{k=0}^{K-1} v_{i-k}, \quad where \quad v_i \stackrel{iid}{\sim} (\chi_M)^{r/K} .$$

*Now the mean is $\sigma^r T$ and the variance is*

$$\sigma^{2r} T^2 d_{n,M,K} \sim \sigma^{2r} T^2 \frac{M}{n} \theta_M^{(r,K)}, \tag{18}$$

*as $n \to \infty$, where $d_{n,M,K}$ is given in (A.2) and*

$$\theta_M^{(r,K)} = M \left[ \left( \lambda_{M,r/K}^K - 1 \right) + 2 \left\{ \left( \lambda_{M,r/K}^{K-1} - 1 \right) + ... + \left( \lambda_{M,r/K} - 1 \right) \right\} \right],$$

*with $\lambda_{M,r} = (k_{M,2r})/(k_{M,r})^2 = \Gamma\left(\frac{M+2r}{2}\right)/\Gamma\left(\frac{M+r}{2}\right)^2 \geq 1$. Ignoring $O(M^{-2})$ terms*

$$\theta_M^{(r,K)} \quad \sim \quad \vartheta_{r/K,K,M} = \left\{ j_{r/K} + \frac{k_{r/K}}{M} \right\} \{ K + 2(K-1) + ... + 2 \}$$

$$+ \frac{j_{r/K}^2}{2M} \{ K(K-1) + 2(K-1)(K-2) + ... + 2 \},$$

*where $j_r = a_{2r} - 2a_r$, $k_r = b_{2r} + (2a_r)^2 - h_r - 2a_{2r}a_r$.*

**Proof.** Given in the Appendix.

The exact values of $\theta_M^{(r,K)}$ are presented in Table 3 for $r = 2$ and $r = 4$, the values most important in empirical work. The Table also contains the expressions for $\vartheta_{r/K,K,M}$.

| | | $r=2$ | | | | $r=4$ | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| $M$ | $K=1$ | $K=2$ | $K=3$ | $K=4$ | $K=1$ | $K=2$ | $K=3$ | $K=4$ |
| 1 | 2.000 | 2.609 | 3.061 | 3.377 | 10.66 | 12.00 | 13.65 | 14.91 |
| 2 | 2.000 | 2.335 | 2.556 | 2.697 | 10.00 | 10.00 | 10.61 | 11.09 |
| 3 | 2.000 | 2.232 | 2.376 | 2.465 | 9.600 | 9.333 | 9.702 | 9.985 |
| 5 | 2.000 | 2.143 | 2.228 | 2.277 | 9.142 | 8.800 | 9.000 | 9.156 |
| 10 | 2.000 | 2.073 | 2.114 | 2.136 | 8.666 | 8.400 | 8.492 | 8.564 |
| 25 | 2.000 | 2.029 | 2.045 | 2.054 | 8.296 | 8.160 | 8.194 | 8.222 |
| 50 | 2.000 | 2.014 | 2.022 | 2.027 | 8.153 | 8.080 | 8.097 | 8.110 |
| $\infty$ | 2.000 | 2.000 | 2.000 | 2.000 | 8.000 | 8.000 | 8.000 | 8.000 |
| $\vartheta_{r/K,K,M}$ | 2.000 | $2+\frac{3}{4M}$ | $2+\frac{92}{81M}$ | $2+\frac{43}{32M}$ | $8+\frac{8}{M}$ | $8+\frac{4}{M}$ | $8+\frac{392}{81M}$ | $8+\frac{11}{2M}$ |

Table 3: Relative efficiency is determined by $M\theta_M^{(r,K)}$. When $r = 2$ the optimal value of $M\theta_M^{(r,K)}$ is 2, when $r = 4$ the best value of $M\theta_M^{(r,K)}$ is 8. $M$ is the degree of blocking, $K$ is the degree of multipowering. $K = 2$ is a bipower variation statistic, $K = 4$ is quadpower.

The $r = 4$ case is interesting. The delta method implies that the efficiency bound for this is 8. When there are no jumps using sums of fourth power of returns is the most efficient of these estimators, but all of them are pretty inefficient. By the time $M$ reaches 5 they are all substantially improved. This is particularly the case for the $K = 3$ and $K = 4$ statistics, which are attractive as they are robust to jumps and are roughly as efficient as the $K = 1$ case as long as $M$ is 5 or more.

11

The $r = 2$ case is also important: the results are in Table 3. This shows the substantial efficiency loss of using higher values of $K$ and that increasing $M$ ameliorates the worst of these effects. By the time $M = 5$ the difference in efficiency between the different estimators is very modest.

## 3.4 Properties of $MV_M^{(r,K)}$ for fixed $M$

In the general case when $\sigma$ is nonconstant, we can proceed as in the proof of Theorem 1, to obtain

**Theorem 3** *Under Assumptions 1-2 in Appendix B.1, for fixed $M$ as $n \to \infty$:*

$$n^{1/2} \left( MV_M^{(r,K)} - \int_0^T \sigma_t^r \mathrm{d}t \right) \overset{\mathcal{L}s}{\to} N \left( 0, TM\theta_M^{(r,K)} \int_0^T \sigma_s^{2r} \mathrm{d}s \right). \tag{19}$$

**Proof.** Follows exactly the same lines as the proof of Theorem 1, so is omitted here.

Also, by the same development, we obtain for fixed $M$,

$$n^{1/2} \left( MV_M^{(2,K)} - RV \right) \overset{\mathcal{L}s}{\to} N \left( 0, T \left( M\theta_M^{(2,K)} - 2 \right) \int_0^T \sigma_s^4 \mathrm{d}s \right), \tag{20}$$

which can be used for testing.

## 3.5 Properties of $MV_M^{(2,K)}$ for $M$ increasing with $n$

We now make the assumption that

$$X_t = X_0 + \int_0^t \mu_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s + J_t, \tag{21}$$

where $J_t$ is a finite activity jump process. We allow, as $n \to \infty$,

$$n^\beta \frac{M}{n} \to \frac{c}{T},$$

that is, more informally, $M \approx \frac{c}{T} n^{1-\beta}$, where $\beta \in (0, 1)$. The scaling by $T$ means the block size is smaller when $T$ is bigger, keeping $n$ constant. This also means the number of blocks $n/M \approx \frac{T}{c} n^\beta$ grows proportionally with $T$ and less than proportionally with $n$. The boundary case $\beta = 1$ corresponds to $M$ being finite.

**Theorem 4** *Define the infeasible realised variance as*

$$RV^* = \sum_{0 < t_i \leq T} \left( \Delta X_{t_j} - \Delta J_{t_j} \right)^2, \tag{22}$$

*and let $MV_M^{(K)} = MV_M^{(2,K)}$ be given as in (17). Assume that $\sigma_t$ is a continuous semimartingale. Then, under the Assumptions 1-4 given in the Appendix B.1,*

$$n^{1/2} \left( MV_M^{(K)} - RV^* \right) = Kc^{\frac{K-1}{K}} n^{\frac{1}{2} - \beta \frac{K-1}{K}} \sum_{0 < t \leq T} \sigma_t^{2\frac{K-1}{K}} \Delta J_t^{2/K} - n^{-\beta + \frac{1}{2}} \frac{K-1}{3} c[\sigma, \sigma]_T$$

$$- n^{1/2} \Upsilon_n + O_p(n^{\frac{1}{2} - \frac{3}{2}\beta} + n^{\frac{1}{2}(\beta-1)}) + o_p(n^{\frac{1}{2} - \beta \frac{K-1}{K}}).$$

The term $o_p(n^{\frac{1}{2} - \beta \frac{K-1}{K}})$ only appears when there are jumps. $\Upsilon_n$ is an edge effect which is described exactly in Section 6.1.2. Note that $n^{1/2} \Upsilon_n = O_p(n^{-\beta + \frac{1}{2}})$.

**Proof.** From Theorem 13 but then add the impact of jumps and edge effects.

In other words: for multipower variation of order $K \geq 3$, there is a range of $\beta$ for which the integrated volatility ($IV$) is estimated efficiently. The asymptotically equivalence of $MV_M^{(K)}$ and $RV^*$ is remarkable and theoretically extremely helpful. On the other hand, bipower variation is a boundary case which is consistent, but converges to $IV$ only at rate $n^{-\beta/2}$. Hence bipower is not quite asymptotically equivalent to $RV^*$. This result is due to the presence of jumps.

This asymptotic equivalence result immediately implies a number of properties.

**Theorem 5** *Assume the condition of Theorem 4. Also assume that $M \approx \frac{c}{T} n^{1-\beta}$, $K \geq 3$ and*

$$\frac{1}{2} \frac{K}{K-1} < \beta < 1. \tag{23}$$

*Then*

$$n^{1/2} \left( MV_M^{(K)} - \int_0^T \sigma_t^2 \mathrm{d}t \right) \xrightarrow{\mathcal{L}_s} N \left( 0, 2T \int_0^T \sigma_t^4 \mathrm{d}t \right), \tag{24}$$

$$n^{1/2} \left\{ \left( RV - MV_M^{(K)} \right) - \sum_{0 < t \leq T} |\Delta J_t|^2 \right\} \xrightarrow{\mathcal{L}_s} N \left( 0, 4T \sum_{0 < t \leq T} |\Delta J_t|^2 \sigma_t^2 \right). \tag{25}$$

*When there are no jumps when $K = 2$ and $\beta > 1/2$*

$$J_M = \frac{n^{1-\beta/2} \left( RV - MV_M^{(2,K)} \right)}{\sqrt{\frac{3}{4c} T^2 \int_0^T \sigma_s^4 \mathrm{d}s}} \xrightarrow{\mathcal{L}_s} N(0, 1). \tag{26}$$

**Proof.** The result (24) follows from Theorem 4. The second result (25) follows from (24) and from (1) in the Introduction. Equation (26) is a special case of Theorem 7.

When there are jumps the average value of $J_M$ should be $n^{1-\beta/2} \sum_{0 < t \leq T} |\Delta J_t|^2 / \sqrt{\frac{3}{4c} T \int_0^T \sigma_s^4 \mathrm{d}s}$. So the power of the test should increase with the size of $JV$ and $n$ as long as $K \geq 3$ and $\frac{K}{K-1} < 2\beta$. However, when $\int_0^T \sigma_s^4 \mathrm{d}s$ is high the power of the test will be low, keeping constant $\sum_{0 < t \leq T} |\Delta J_t|^2$.

**Remark 5** *(Optimal choice of $\beta$ for IV). When both jumps and volatility of volatility is present, it follows from Theorem 4 that the jumps part will always dominate the $[\sigma, \sigma]_T$ part. Thus there is no trade-off between the two biases. The bias due to jumps can, however, be traded off against the $O_p(n^{\frac{1}{2}(\beta-1)})$ stochastic term. A bias-variance trade-off to minimize the error in $MV_M^{(K)} - RV^*$ will thus seek to set these two orders to be equal. This yields the optimal choice of $\beta$*

$$\beta^* = \frac{2K}{3K-2}, \tag{27}$$

which means that as $K$ increases so $\beta^*$ converges from above to $2/3$. This choice of $\beta$ satisfies condition (23) in Theorem 5. For tripower, $\beta^* = 6/7$.

## 3.6 Optimality for jump estimation of $RV - MV_M^{(2,K)}$

In addition to estimating $IV$ efficiently, we also obtain an efficient estimate of $JV$.

**Theorem 6** *Assume that $\widehat{JV}_n$ is an estimator of $JV$, and that $n^{1/2}(\widehat{JV}_n - JV) \xrightarrow{\mathcal{L}_s} N\left(0, \gamma^2\right)$, where $\gamma^2$ is measurable with respect to the underlying filtration. Also assume that $J_t$ is compound Poisson, independent of $X_t^c$, with jump sizes that are independent of the jump times and of each other, and that the distributions of jumps are absolutely continuous with respect to Lebesgue measure. Then, under Assumptions 1-2 of Appendix B.1, with probability one*

$$\gamma^2 \;\geq\; 4T \sum_{0 < t \leq T} |\Delta J_t|^2 \, \sigma_t^2. \tag{28}$$

**Proof.** Given in the Appendix.

The additional assumptions on $J_t$ are made to invoke results on superefficiency, see LeCam (1953) and Lehmann (1983, Ch. 6). There is no minimal set of conditions, and for other relevant results on superefficiency, see Bahadur (1964) and Bahadur (1980), as well as the literature on Hájek-LeCam convolution, as in Hájek (1969). There are circumstances where the bound (28) can be improved on, in particular when the jump sizes are known in advance.

## 3.7 Estimating the asymptotic variance

To carry out inference robustly to jumps we need to consistently estimate the "integrated quarticity" $IQ = \int_0^T \sigma_t^4 \mathrm{d}t$ and the "squared jump times vol" $SJV = \sum_{0 < t \leq T} |\Delta J_t|^2 \, \sigma_t^2$. Barndorff-Nielsen and Shephard (2006) solved the former using multipower variation. We have already improved upon this by using a blocked version to estimate $IQ$. Here our focus will be on estimating $\sum_{0 < t \leq T} |\Delta J_t|^2 \, \sigma_t^2$.

Using a blocked approach this problem is simple. Rearranging terms

$$RV - BV_M = \sum_{i=1}^{n_M} \left[ RV_i - \frac{M}{k_M} \left( RV_i RV_{i-1} \right)^{1/2} \right] \xrightarrow{p} \sum_{0 < t \leq T} |\Delta J_t|^2,$$

as $n \to \infty$ if $M$ is fixed or slowly increasing. We now extend this using multipower ideas

$$\sum_{i=1}^{n_M} \left\{ RV_i - \frac{M}{k_M} \left( RV_i RV_{i-1} \right)^{1/2} \right\} \left\{ \frac{1}{k_M} \left( RV_{i-2} RV_{i-3} \right)^{1/2} \right\} \xrightarrow{p} \sum_{0 < t \leq T} |\Delta J_t|^2 \, \sigma_t^2.$$

Hence all our new distribution theory is feasible. Veraart (2010) provided an alternative estimator of $SJV$, which is somewhat more complicated.

## 3.8 Finite sample improvement via transformation

The jump test (15) can be transformed using the delta method so we look at

$$J_M = \frac{n^{1/2} \left( \log RV - \log BV_M \right)}{\sqrt{T \left( M\theta_M - 2 \right) \frac{\widehat{IQ}}{(BV_M)^2}}} \overset{\mathcal{L}s}{\to} N(0,1). \tag{29}$$

This log-transform was used by Barndorff-Nielsen and Shephard (2002) to produce better size properties. Further noting that by Jensen's inequality $R = \int_0^T \sigma_s^4 \mathrm{d}s / \left( \int_0^T \sigma_s^2 \mathrm{d}s \right)^2 \geq \frac{1}{T}$, it makes sense to replace (29) by

$$J_M = \frac{\log RV - \log BV_M}{\sqrt{T d_{n,M,K} \max \left( \frac{\widehat{IQ}}{(BV_M)^2}, \frac{1}{T} \right)}} \overset{\mathcal{L}s}{\to} N(0,1), \tag{30}$$

where $d_{n,M,K}$ is the exact constant in Theorem 1. This makes a large difference if $M$ is large. Of course further refinements are possible, such as the bootstrap of Goncalves and Meddahi (2004).

## 3.9 Finite sample improvement via time-change and diurnality

Quite a substantial amount of the variation of the volatility process within a day is caused by the average diurnal pattern playing out. This is one of the causes of $R > 1/T$ in practice. Can we perform a time-change to bring it down closer to one, making the test more efficient?

Consider the absolutely continuous time-change, and the $T = 1$ case for ease of exposition. Then assume and define $h_t^2 > 0$, $\int_0^1 h_t^2 \mathrm{d}t = 1$, $H_t = \int_0^t h_s^2 \mathrm{d}s$ then $X_{H_t}$ is a Brownian semimartingale with spot variance at time $t$ of $\widetilde{\sigma}_t^2 = \sigma_{H_t}^2 h_t^2$, while $[X \circ H]_1 = [X]_1$. Hence if we know the path $\sigma_t$ then we can design the time-change $h_t$ so that $\widetilde{\sigma}_t^2$ is time-invariant and so $R = 1$. Of course in general this is impossible, but we can time-change to take out the expected diurnal pattern. We can do this safely in our model, for if the expected diurnality is modelled incorrectly then our inference procedure will still be valid, for the time-change process is always a Brownian semimartingale whatever absolutely continuous time-change we use.

Consider the diurnal model $\sigma_t^2 = \psi_t d_t$ then $\widetilde{\sigma}_t^2 = \psi_{H_t}^2 d_{H_t} h_t^2 = \psi_{H_t}^2$ if we design $h_t^{-2} = d \circ \int_0^t h_u^2 \mathrm{d}u$.

How do we implement this idea? Here we estimate the diurnal shape non-parametrically using block bipower-type statistics. Define the $n - 2M + 1$ overlapping unnormalised "bipower blocks" as $\widetilde{BV}_{M,i} = \sqrt{RV_{M,i} RV_{M,i+M}}$, $i = 1, 2, \ldots, n-2M+1$. The diurnal pattern can the be estimated using $S$ days of historical data, $\hat{\delta}_i^2 = S^{-1} \sum_{t=t-S}^{t-1} \left( \frac{1}{u_i - l_i} \sum_{i=l_i}^{u_i} \widetilde{BV}_{M,i,t} \right)$, where $u_i = \min \left( n - 2M + 1, i \right)$ and $l_i = \max \left( 1, i - 2M + 1 \right)$. This estimator uses all $M$-sample blocks which contain return $i$ to estimate the diurnal pattern in the data. Except for the period near the beginning or end of the day, the diurnal effect is estimated by averaging the $2M$ bipower blocks that contain return $i$. Periods at the beginning or end use between 1 and $2M - 1$ blocks.

| Parameter | $\mu$ | $\beta_0$ | $\beta_1$ | $\alpha_v$ | $\rho$ |
|-----------|-------|-----------|-----------|------------|--------|
| Value | 0.030 | 0 | 0.125 | -0.100 | -0.620 |

Table 4: Parameters which index the single factor SV simulation model.

# 4 Monte Carlo study of efficient bipower variation

For purposes of simulation, we take $T = 1$ and always use 50,000 replications, representing 50,000 independent days. When constructing estimators and t tests we use versions which are exactly unbiased in the Brownian motion case, rather than the simpler asymptotic equivalent versions. This does make a great deal of difference in terms of Monte Carlo performance.

## 4.1 Single factor SV model

To assess the usefulness of the asymptotic results we have conducted two Monte Carlo experiments. The first, and simplest, uses a single factor SV model. It has the following log-normal structure $dX_t = \mu dt + \exp(\beta_0 + \beta_1 v_t) dW_{p,t} + dJ_t$, where $dv_t = \alpha_v v_t dt + dW_{v,t}$. Here $W_p$ and $W_v$ are correlated standard Brownian motions, writing $\rho$ to denote the correlation. This process was implemented using the values given in Table 4, which follow from the experiments reported in Huang and Tauchen (2005). This was designed to have relatively fast mean reversion with a half-life of 10 days, and so there is some variation in the volatility during each day.

### 4.1.1 Finite activity jumps

Within each day the high frequency data was simulated using $23,400$ steps per day, and then "1-minute" returns were sampled by skipping 60 steps. The result is $n = 390$. Jumps, when present, have a random location uniformly distributed on $1, \ldots, 23,400$ and were a deterministic percentage of the day's integrated variance (e.g. a jump, when squared, represents 10% of IV, or $JV/IV = 0.1$). Throughout our simulations a single jump occurred on each day.

### 4.1.2 One factor SV model with diurnal effects

The simulated model was augmented to include a diurnal effect $\delta_t$ which produces variances which differ deterministically by a factor of 3 between the middle of the day and the beginning or end. $dX_t = \mu dt + \exp(\beta_0 + \beta_1 v_t)\sqrt{\delta_t} dW_{p,t} + dJ_t$, $dv_t = \alpha_v v_t dt + dW_{v,t}$, $\delta_t = 0.6 + 4.8(t - 0.5)^2$, where the average effect $\int_0^1 \delta_t dt = 1$ and the start and the end of the day $\delta_0 = \delta_1 = 3\delta_{0.5}$. This choice is motivated by our empirical results in Section 5.3.

## 4.2 Bias, variance, size and power of jump tests

Size was computed by setting the jump size to 0. The test statistics we study are

$$J_M = \frac{RV - BV_M}{\sqrt{d_{n,M,K}\widehat{IQ}}}, \quad \text{and} \quad J_{M,\log} = \frac{\log RV - \log BV_M}{\sqrt{d_{n,M,K}\max\left(\frac{\widehat{IQ}}{(BV_M)^2}, 1\right)}},$$

where $d_{n,M,K}$ is given by (18). We estimate integrated quarticity using the unbiased blocked tripower quarticity estimator (17) with $r = 4$, $K = 3$ and the same value of $M$ as used for $BV_M$. All tests were conducted using 1-sided upper tailed critical values. We also carried out the Monte Carlo using a unblocked tripower variation statistic to estimate $IQ$ and there was little difference in the corresponding size.

Table 5 reports the empirical size both the standard 1-factor model and the model that includes deterministic diurnal effects. We only report the size of a 5% one-sided test since both 10% and 1% sized tests performed analogously.

In the Table we use the following important notation. **FS**. This reports results using the finite sample variance for a standard Brownian motion, using equation (A.1). The results using the large sample values for $M\theta_M$ were poor and so we do not recommend their use and so we are not reporting them here. The distortions to the size of the test statistics computed using the large sample $M\theta_M$ is caused by the overconfidence in the precision, and block sizes larger than 5 lead to unacceptable large deviations from 1. **FS-DC**. This results using the finite sample variance and applying the diurnal correction developed in Section 3.9. Throughout this uses the previous $S = 100$ days of simulated data. The test statistics were computed by first constructing a modified set of returns, $\Delta X_i/\hat{\delta}_i$, where $\hat{\delta}_i$ is the diurnal correction for the relevant block. This is exactly the same as having standardised the local realised variance estimator by the estimated diurnal feature of the data. We then compute $RV$, $BV_M$ and $TPQ$ using the transformed data.

In the non-diurnal simulation results, using the large sample variance of the jump test leads to large distortions when the block size is moderately large. The finite sample results show little size distortion for any block size. When the deterministic diurnality is introduced a large distortion, driven by the difference in volatility across blocks, appears which increases with the block length.

Table 6 reports the empirical size-adjusted power of the of the test statistics when using a 5% test. The power peaks with a block size of 5 which is consistent with the minimum of finite sample variance for $n = 390$. The size-adjusted power falls off for large blocks due to both the increase in variance, and in the simulations with the diurnal component, the substantial size distortions.

Table 5 contains size results for the jump test using the finite sample variance and the diurnality correction (FS-DC). When there is no diurnality present the modification has no effect on size. When the diurnality is present the transformation produces an empirical size which is virtually

| | No Diurnality | | | | Diurnality | | | |
| | FS | | FS-DC | | FS | | FS-DC | |
| $M$ | $J_M$ | $J_{M,\log}$ | $J_M$ | $J_{M,\log}$ | $J_M$ | $J_{M,\log}$ | $J_M$ | $J_{M,\log}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 6.4 | 5.5 | 6.5 | 5.6 | 7.3 | 6.5 | 6.7 | 5.7 |
| 2 | 6.1 | 5.5 | 6.2 | 5.6 | 8.8 | 8.3 | 6.7 | 6.0 |
| 3 | 6.4 | 5.9 | 6.4 | 5.8 | 10.9 | 10.5 | 6.1 | 5.6 |
| 5 | 7.0 | 6.6 | 6.4 | 6.0 | 14.7 | 14.2 | 6.4 | 5.8 |
| 10 | 6.7 | 6.4 | 6.9 | 6.5 | 33.6 | 33.0 | 7.0 | 6.8 |
| 15 | 7.3 | 7.0 | 6.9 | 6.5 | 56.1 | 55.4 | 7.5 | 7.1 |
| 30 | 8.1 | 7.6 | 8.3 | 7.9 | 90.2 | 89.5 | 8.6 | 8.1 |

Table 5: Empirical size of SV1F with and without (left panel) deterministic diurnality using one minute simulated returns. Values under $J_M$ correspond to the standard version of $BV_M$ and $J_{M,\log}$ correspond to the log version of the jump test statistic. Values under FS use the finite sample version of $d_{n,M,K}$ in (A.1). Values below FS-DC use finite sample variance and an estimated diurnality correction.

| | 10% Jump Size | | | | | | | |
| | No Diurnality | | | | Diurnality | | | |
| | FS | | FS-DC | | FS | | FS-DC | |
| $M$ | $J_M$ | $J_{M,\log}$ | $J_M$ | $J_{M,\log}$ | $J_M$ | $J_{M,\log}$ | $J_M$ | $J_{M,\log}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 41.4 | 41.9 | 41.4 | 41.9 | 37.7 | 37.7 | 48.2 | 48.7 |
| 2 | 50.6 | 50.7 | 50.6 | 50.7 | 46.9 | 46.9 | 56.9 | 57.0 |
| 3 | 54.9 | 55.4 | 54.9 | 55.4 | 50.8 | 50.7 | 60.3 | 60.5 |
| 5 | 57.9 | 58.1 | 57.9 | 58.1 | 52.3 | 52.5 | 63.4 | 63.6 |
| 10 | 52.9 | 53.0 | 52.9 | 53.0 | 41.1 | 41.3 | 59.1 | 59.3 |
| 15 | 44.1 | 44.2 | 44.1 | 44.2 | 31.4 | 31.3 | 51.7 | 51.9 |
| 30 | 26.4 | 26.2 | 26.4 | 26.2 | 15.7 | 15.6 | 29.0 | 29.0 |
| | 20% Jump Size | | | | | | | |
| 1 | 85.8 | 86.0 | 85.8 | 86.0 | 82.3 | 82.2 | 83.5 | 83.9 |
| 2 | 93.4 | 93.6 | 93.4 | 93.6 | 91.0 | 91.0 | 90.1 | 90.2 |
| 3 | 95.9 | 96.1 | 95.9 | 96.1 | 93.4 | 93.4 | 92.3 | 92.4 |
| 5 | 96.8 | 96.9 | 96.8 | 96.9 | 93.9 | 93.9 | 93.4 | 93.5 |
| 10 | 95.8 | 95.8 | 95.8 | 95.8 | 89.6 | 89.7 | 92.7 | 92.8 |
| 15 | 91.8 | 92.0 | 91.8 | 92.0 | 80.5 | 80.5 | 90.0 | 90.1 |
| 30 | 68.5 | 68.4 | 68.5 | 68.4 | 43.5 | 43.2 | 75.4 | 75.3 |

Table 6: Empirical size-adjusted power for jump sizes representing 10% (top panel) and 20% (bottom panel) of the integrated variance. The left 4 columns correspond to the SV1F model and the right columns add a deterministic diurnal effect. FS indicates that the finite sample variance was used and FS-DC indicates that both the finite sample variance and the diurnal correction were used to compute the statistic.

identical to the size of the FS estimator in the simulation without the diurnality. Finally, Table 6 shows that the use of the diurnal correction has no effect on size-adjusted power when the data are not diurnal, and improves the size-adjusted power when the data is diurnal.

## 4.3 $MV_M^{(2,K)}$ for $M$ increasing with $n$

We will now repeat the above experiments but allow $n$ to vary, driving some movements in $M$. Throughout this subsection we will take $K = 3$ and the corresponding optimal $\beta = 6/7$ following Remark 5, where $M = cn^{1-\beta}$. As $cn^{1-\beta}$ will not be an integer we take it to be the nearest integer but requiring it also to be greater than or equal to one. We do not have a theory driven value for $c$, so we take it from the above Monte Carlos which suggest for $n = 390$ then $M = 3$ provides

|  |  | No Diurnality | | | | Diurnality | | | |
|  |  | FS | | FS-DC | | FS | | FS-DC | |
| $n$ | $M$ | $J_M$ | $J_{M,\log}$ | $J_M$ | $J_{M,\log}$ | $J_M$ | $J_{M,\log}$ | $J_M$ | $J_{M,\log}$ |
| | | | | | Empirical size | | | | |
| 780 | 3 | 6.1 | 5.8 | 6.1 | 5.8 | 8.7 | 8.3 | 6.1 | 5.8 |
| 1,560 | 3 | 5.5 | 5.3 | 5.5 | 5.3 | 7.4 | 7.1 | 5.5 | 5.3 |
| 1,560 | 4 | 5.6 | 5.5 | 5.6 | 5.5 | 7.9 | 7.7 | 5.6 | 5.5 |
| | | | | | Power: 10% Jump Size | | | | |
| 780 | 3 | 87.3 | 87.4 | 87.3 | 87.4 | 85.2 | 85.2 | 85.2 | 85.3 |
| 1,560 | 3 | 99.6 | 99.7 | 99.6 | 99.7 | 99.0 | 99.0 | 97.2 | 97.3 |
| 1,560 | 4 | 99.8 | 99.8 | 99.8 | 99.8 | 99.3 | 99.3 | 98.0 | 98.0 |

Table 7: Impact of increasing $M$ with $n$. Empirical size and size-adjusted power for jump sizes representing 10% of the integrated variance. The left 4 columns correspond to the SV1F model and the right columns add a deterministic diurnal effect. FS indicates that the finite sample variance was used and FS-DC indicates that both the finite sample variance and the diurnal correction were used to compute the statistic.

satisfactory results. This suggests taking $c = M/n^{1/7} \simeq 1.28$ as we vary $n$.

In our Monte Carlo we will double $n$ to 780 and double it again to $1,560$. The first change leaves $M$ unaltered at 3, the second change drives $M$ up to 4. Throughout we use the finite $M$ form of the t-statistic in implementing the test.

The results are in Table 7. They suggest an improving size performance as $n$ and $M$ increase together, while the power naturally increases very significantly with an increase in the sample size. Hence these results are encouraging for the blocking and its asymptotic analysis. The Table also shows the results from having $M$ is fixed as $n$ increases in the $n = 1,560$ and $M = 3$ case (the theory suggests $M$ should have nudged up to 4 due to the increase in $n$). The results suggest a loss in power in that case, although the differences are mild.

# 5  Empirical results

## 5.1  Database

Our application will be based around trade data for the SPDR S&P 500 ETF (SPY), which is a exchange traded fund (ETF) that tracks the S&P 500 Index. This is the most liquid equity in U.S. markets and typically has a spread of 0.01 which represents approximately 0.01% of the price of the instrument. The small spread allows for frequent sampling without substantial market microstructure noise, and so we employ returns computed every 60 seconds. As a result $n = 390$.

Our database contains transactions from January 3, 2005 until December 31, 2009. We will not compute statistics on days which are short (i.e. open or close early) or days where there are sequences of no trades for more than 60 consecutive seconds. These rules leaves us with 1,226 full days of high frequency data. Prices were cleaned to remove outliers in a similar way to Barndorff-Nielsen, Hansen, Lunde, and Shephard (2009). The top panel of Figure 1 contains a plot of the
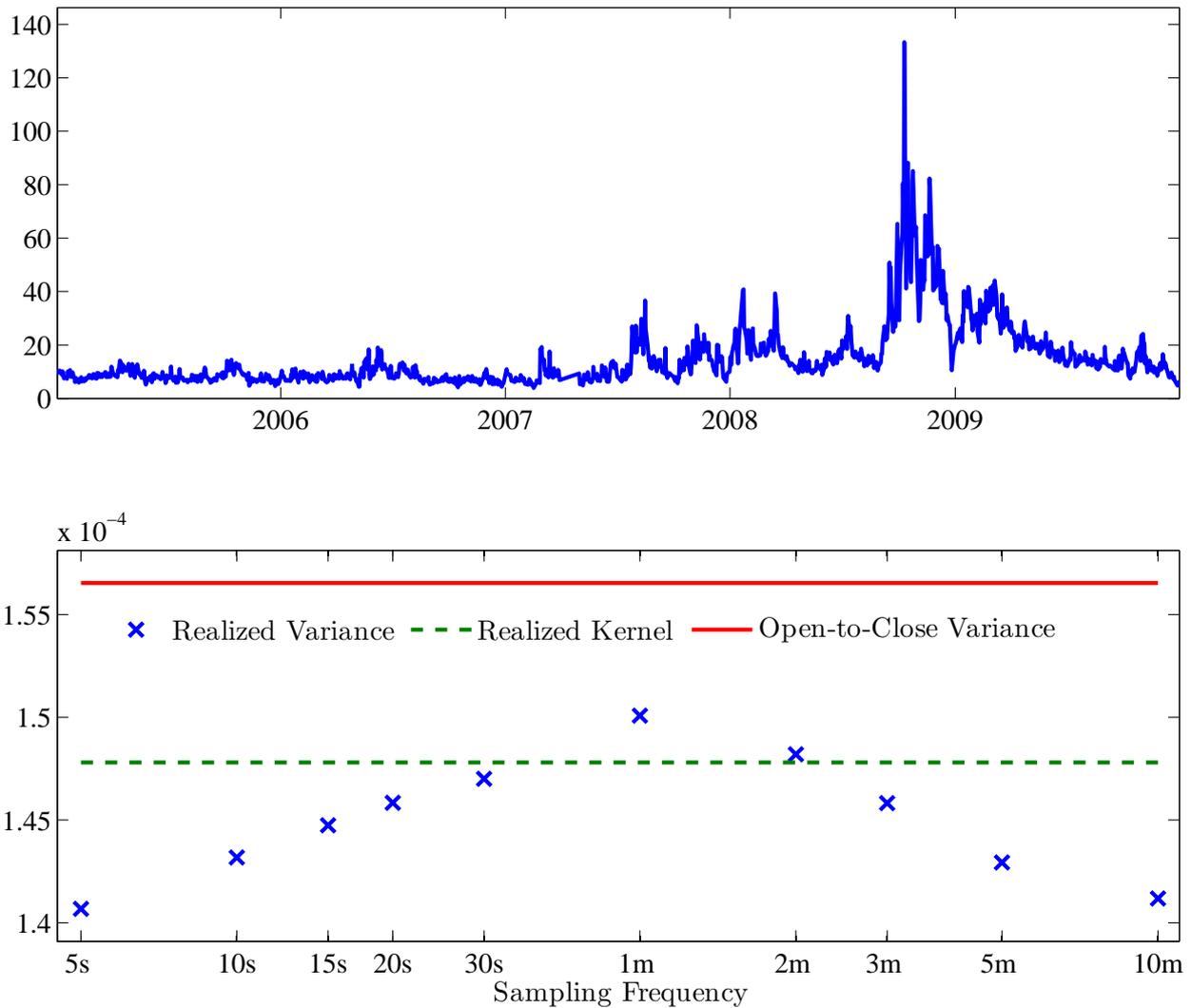
Figure 1: The top panel contains the annualized realised volatility for the S&P500 SPDR computed using one-minute returns from January 2005 to December 2009. The bottom panel contains the volatility signature plot for realized variance using sample frequencies from 5-seconds to 10-minutes, the average value of realized kernels computed daily, and the variance of daily open-to-close returns.

annualized volatility (%) computed from the daily realised variance estimates, and the bottom panel contains the volatility signature plot for sampling times between 5 seconds and 10 minutes, as well as the average value of daily realized kernels and the open-to-close return variance.[2] Sampling using 1-minute prices does not appear to be affected by market microstructure noise.

## 5.2 Empirical rejection rates

Both the jump test and the log-version FS-DC were computed daily with $n = 390$. All tests were implemented as one-sided upper-tail tests, rejecting for large values of the test statistic. Throughout we estimate integrated quarticity using blocked TPQ. The empirical rejection rates are given in

---

[2]Realized Kernels computed using all returns following the method detailed in Barndorff-Nielsen, Hansen, Lunde, and Shephard (2009).

|   | $J_M$ | | | $J_{M,\log}$ | | |
|---|---|---|---|---|---|---|
| $M$ | 5% | 1% | .1% | 5% | 1% | .1% |
| 1 | 48.2 | 31.2 | 19.9 | 49.3 | 33.6 | 21.4 |
| 2 | 58.0 | 41.3 | 24.5 | 68.3 | 56.2 | 43.4 |
| 3 | 59.7 | 42.7 | 26.0 | 74.5 | 65.0 | 53.4 |
| 5 | 60.2 | 44.0 | 30.0 | 79.6 | 72.6 | 63.1 |
| 10 | 63.1 | 46.6 | 30.8 | 85.4 | 80.8 | 74.6 |
| 15 | 68.2 | 52.2 | 36.9 | 88.7 | 84.4 | 79.7 |

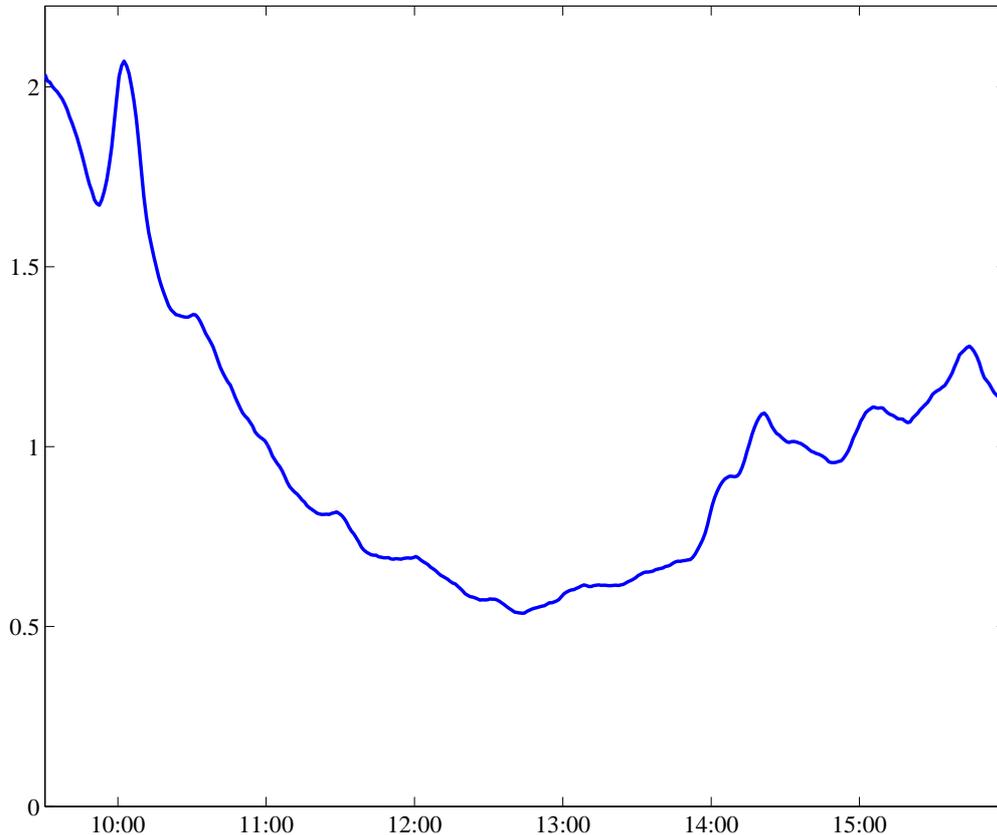Table 8: Empirical size using the blocked bipower jump tests for different levels of blocking $M$.



Figure 2: *Full sample estimated dirunal pattern of volatility $(\hat{\delta}^2)$ for the S&P500 SPDR.*

Table 8. The table contains results from the test statistics applied directly to returns, and show that the rejection rate is monotonically increasing with the block size for both test statistics.

## 5.3 Diurnality Correction

The top panel of figure 2 shows an estimated averaged diurnality in volatility. This figure was computed using the estimator described in the Monte Carlo with two changes: (i) High frequency returns are standardized by the daily $\sqrt{BV_5}$ prior to estimating the diurnality. This mitigates the effect of changing volatility throughout the sample, and is similar to weighting in a generalised least squares regression. (ii) The entire sample was used to compute the diurnal volatility pattern.

The Figure indicates that volatility is substantially higher at the start of the day, even when

| | No Correction | | | | | | Diurnal Correction | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $J_M$ | | | $J_{M,\log}$ | | | $J_M$ | | | $J_{M,\log}$ | | |
| $M$ | 5% | 1% | .1% | 5% | 1% | .1% | 5% | 1% | .1% | 5% | 1% | .1% |
| 1 | 48.2 | 31.2 | 19.9 | 49.3 | 33.6 | 21.4 | 55.4 | 38.7 | 25.4 | 58.9 | 42.3 | 28.4 |
| 2 | 58.0 | 41.3 | 24.5 | 68.3 | 56.2 | 43.4 | 62.9 | 46.1 | 31.1 | 73.6 | 63.5 | 51.1 |
| 3 | 59.7 | 42.7 | 26.0 | 74.5 | 65.0 | 53.4 | 63.1 | 47.1 | 30.4 | 78.4 | 69.1 | 59.2 |
| 5 | 60.2 | 44.0 | 30.0 | 79.6 | 72.6 | 63.1 | 59.6 | 44.9 | 30.7 | 78.0 | 72.8 | 64.3 |
| 10 | 63.1 | 46.6 | 30.8 | 85.4 | 80.8 | 74.6 | 51.3 | 35.6 | 23.1 | 75.4 | 69.7 | 63.9 |
| 15 | 68.2 | 52.2 | 36.9 | 88.7 | 84.4 | 79.7 | 44.9 | 32.7 | 20.6 | 70.2 | 65.1 | 58.8 |

Table 9: Empirical size using the blocked bipower jump tests for different levels of blocking $M$. The left two panels contains rejection rates for the usual version of the test statistic. The right panels contains rejection rates for the test statistic based on returns with a dirunal correction.

| $M$ | 1 | 2 | 3 | 5 | 6 | 10 | 13 | 15 |
|---|---|---|---|---|---|---|---|---|
| Calendar Time | 11.7 | 2.0 | 0.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |

Table 10: Percentage of "bipower blocks" which are 0 for each block size.

controlling for the transaction rate, and that volatility rises at the end of the day.

Table 9 contains test statistics computed using returns transformed by an estimate of the diurnal effect. This diurnal correction was estimated using the previous returns from the previous 100 days. This correction removes the increase in the rejection rate as the block size grows, and the rejection rate peaks with $M = 3$ or 5, consistent with the Monte Carlo study. In both panels the log version of the test rejects substantially more often than the standard version.

## 5.4 Validation

Various methods were used to provide a check on these results. Here we detail two. (i) We computed the time-series average of $BV_M$ divided by the time-series average of RV. If the bias is important one would expect this to decline sharply as $M$ increases. Figure 3 shows this ratio for all three versions of the test statistics. The diurnality correction appear to reverse the decline in the average BV as the block size decreases. (ii) Table 10 contains the percentage of the "bipower blocks" which are 0. In a BSM 0 returns should not occur, although due to market microstructure noise they do arise approximately 10% of the time when sampling using 1 minute returns. Larger blocks have substantially fewer zeros which may be an additional advantage.

## 6 Some additional points

### 6.1 A general theory for realised measures

#### 6.1.1 Background

The same kind of theory for blocked multipower variation can be established using miniature realised measures, rather than only miniature realised variances. The virtue of using this approach is that it provides a unified theory to dealing with the asymptotics when $M$ increases with $n$. Ex-
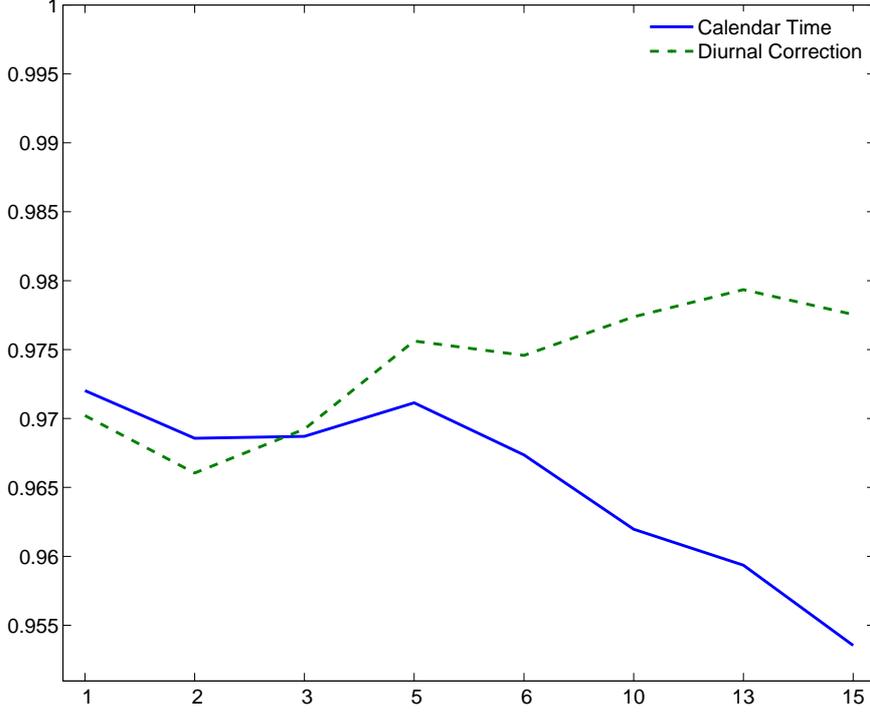
Figure 3: *Full sample estimated ratio of average $BM_M$ to average $RV$ for prices sampled in calendar time (solid), diurnally corrected returns (dashed) and prices sampled in business time (dashed).*

amples of potential miniature measures include using noise robust measures such as the multiscale estimators, realised kernels and preaveraging estimators (e.g. Zhang, Mykland, and Aït-Sahalia (2005), Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008) and Jacod, Li, Mykland, Podolskij, and Vetter (2009)). But the scope of realised measures is wider than that, it includes for example multipower variation and realised range statistics (e.g. Martens and van Dijk (2007) and Christensen and Podolskij (2007)).

### 6.1.2 Edge corrected definitions

Denote $RM_i$ the miniature realised measure which estimates $\int_{\tau_{i-1}}^{\tau_i} \sigma_t^2 \mathrm{d}t$, computed using the $i$-th block of data. The "edge corrected RV" based on miniature realised measures is

$$ECRV_n^{(K)} = \frac{1}{K} \sum_{m=1}^{K} RV_n^{(K,m)} = RV_n - \Upsilon_n, \quad RV_n^{(K,m)} = \sum_{i=m}^{m-1+K\lfloor (n_M-m+1)/K \rfloor} RM_i.$$

The leading case of this is where $K = 2$, then this estimator is $ECRV_n^{(2)} = \frac{1}{2}RM_1 + \sum_{i=2}^{n_M-1} RM_i + \frac{1}{2}RM_{n_M}$. The corresponding "edge corrected MV" based on miniature realised measures is

$$UMV_n^{(K)} = \frac{1}{K} \sum_{m=1}^{K} SMV_n^{(K,m)}, \quad SMV_n^{(K,m)} = K \sum_{1 \le i; Ki+m-1 \le n_M} (RM_{Ki-K+m}...RM_{Ki+m-1})^{1/K}.$$

The impact of edge effects is sometimes important and sometimes it does not. The following characterise the differences and will drive the application of some of the results below.

**Remark 6** $ECRV_n^{(K)} - RV_n = \Upsilon_n = O_p(n^{-\beta})$.

**Proof.** Special case of Remark 8 in the Appendix.

**Remark 7** $UMV_n^{(k)} = MV_M^{(2,K)}\left\{1 + O_p(n^{-\beta})\right\}$.

**Proof.** Follows from (B.20).

### 6.1.3 Core result

The following result drives all the paper's analysis of the case where $M$ increases with $n$.

**Theorem 7** *Assume that under the statistical risk neutral distribution (Mykland and Zhang (2009, Sect 2.2)), there is a sequence of continuous martingales $(M_t^{(n)})_{0 \leq t \leq T}$ so that*

$$RM_i = \int_{\tau_{i-1}}^{\tau_i} \sigma_t^2 \mathrm{d}t + M_{\tau_i}^{(n)} - M_{\tau_{i-1}}^{(n)} = \int_{\tau_{i-1}}^{\tau_i} \sigma_t^2 \mathrm{d}t + \Delta M_{\tau_i}^{(n)}.$$

*Suppose that as $n \to \infty$, $n^{2\alpha}[M^{(n)}, M^{(n)}]_t \xrightarrow{p} \int_0^t f_s^2 \mathrm{d}s$ and $n^{\alpha}[M^{(n)}, W]_t \xrightarrow{p} 0$. Also assume the conditions of Appendix B.1, except that $J \equiv 0$. In the case when $\beta = \alpha$, set $b_1 = \frac{1}{4}c^{-1}\int_0^T \sigma_t^{-2} f_t^2 \mathrm{d}t$ and $b_2 = \frac{1}{3}c\,[\sigma,\sigma]_T$, while define $a^2 = \frac{3}{16}\int_0^T c^{-1}\left(\sigma_t^{-2} f_t^2 + c^2[\sigma,\sigma]_t'\right)^2 \mathrm{d}t$. Then*

$$n^{\frac{3}{2}\alpha}\left(ECRV_n^{(2)} - UMV_n^{(2)} - (n^{-2\alpha+\beta}b_1 + n^{-\beta}b_2)\right) \xrightarrow{\mathcal{L}s} a \times Z \tag{31}$$

*where $Z$ is standard normal and independent of the underlying filtration. When $\beta > \alpha$, only keep the first term inside the brackets in $a^2$, and replace $n^{\frac{3}{2}\alpha}$ by $n^{2\alpha - \frac{1}{2}\beta}$ in (31). When $\beta < \alpha$, only keep the second term inside the brackets in $a^2$, and replace $n^{\frac{3}{2}\alpha}$ by $n^{\frac{3}{2}\beta}$ in (31).*

**Proof.** Given in the Appendix.

Theorem 11 extends the result to the general multipower case[3] with

$$ECRV_n^{(2)} - UMV_n^{(K)} = n^{-2\alpha+\beta}b_1 + n^{-\beta}b_2 + O_p(n^{-2\alpha + \frac{1}{2}\beta} + n^{-\frac{3}{2}\beta}).$$

Theorem 14 specialises the result to the case of no microstructure effects.

The following Corollary thus holds, containing most of the cases of practical importance.

---

[3]Using the same method of proof as Theorem 7 in the case where $\beta > \alpha$ then

$$n^{2\alpha - \frac{1}{2}\beta}\left(ECRV_n^{(2)} - UMV_n^{(K)} - (n^{-2\alpha+\beta}b_1 + n^{-\beta}b_2)\right) \xrightarrow{\mathcal{L}s} \sqrt{\frac{\vartheta_K}{4}\int_0^T c^{-1}\left(\sigma_t^{-2} f_t^2\right)^2 \mathrm{d}t} \times Z.$$

Here, using the constants in Table 3,

$$\vartheta_K = k_{2/K}\left\{K + 2(K-1) + ... + 2\right\} + \frac{j_{2/K}^2}{2}\left\{K(K-1) + 2(K-1)(K-2) + ... + 2\right\}.$$

In particular, $\vartheta_2 = 4/3$, $\vartheta_3 = 92/81$ and $\vartheta_3 = 43/32$.

**Corollary 1** *Assume $0 < \alpha < \beta$, $2\alpha > \beta$ and $f_t^2 = dT\sigma_t^4$ where $d > 0$ is a constant, then*

$$n^{2\alpha - \beta/2} \left( RV_n - \left\{ 1 + n^{\beta - 2\alpha} \frac{Td}{4c} \right\} UMV_M^{(K)} \right) \overset{\mathcal{L}s}{\to} \sqrt{\frac{d^2 T^2}{2c} \int_0^T \sigma_t^4 dt} \times Z.$$

**Example 1** *In the realised variance case then $\alpha = 1/2$, $d = 2$. This result drives Theorem 5.*

**Example 2** *The realised kernels of Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008, equation 19) and Barndorff-Nielsen, Hansen, Lunde, and Shephard (2011) obey this structure when the bandwidth parameter $H = bM^{2/3}$ and the weight function is very smooth. In the Parzen case, $d = 4bk_\bullet^{0,0}$, where $k_\bullet^{0,0} = 0.269$, so within a fixed block*

$$M^{1/6} \left( RM_i - \int_{\tau_{i-1}}^{\tau_i} \sigma_t^2 dt \right) \overset{\mathcal{L}s}{\to} N \left( 0, 4bk_\bullet^{0,0} (\tau_i - \tau_{i-1}) \int_{\tau_{i-1}}^{\tau_i} \sigma_t^4 dt \right).$$

*This implies $\alpha = 1/6$ so we must have $\beta < 1/3$, which means $M = (c/T) n^{1-\beta}$ must grow quite quickly with $n$, indicating very large blocks are needed. In this case $f_t^2 = 4bk_\bullet^{0,0} T\sigma_t^4$. Most noise robust estimators do not apply as they do not have the feature that $f_t^2 = dT\sigma_t^4$, instead usually the variance of the noise complicates the distribution. The Barndorff-Nielsen, Hansen, Lunde, and Shephard (2011) kernel has the virtue in this context of being non-negative.*

### 6.1.4   A two scale estimator

In order to produce more efficient results we can use two multipower variations $UMV_M^{(K,c_1)}$ and $UMV_M^{(K,c_2)}$, taking $c_1 > c_2$. Define a "two scale" estimator

$$UMV_M^{(K,c_1,c_2)} = (c_1 - c_2)^{-1}(c_1 UMV_M^{(K,c_1)} - c_2 UMV_M^{(K,c_2)}).$$

Again following Section B.2.3, we obtain

$$
\begin{aligned}
n^\alpha \left( UMV_M^{(K,c_1,c_2)} - RV_n \right) = {} & K \frac{c_1^{\frac{-1}{K}} - c_2^{\frac{-1}{K}}}{c_1 - c_2} n^{\alpha - \beta \frac{K-1}{K}} \sum_{0 < t \leq T} \sigma_t^{2\frac{K-1}{K}} \Delta J_t^{2/K} \\
& + n^{-\beta + \alpha} \frac{K-1}{3} \frac{1}{c_1 c_2} [\sigma, \sigma]_T + \text{ differential edge effect} \\
& + O_p(n^{-\alpha + \frac{1}{2}\beta} + n^{\alpha - \frac{3}{2}\beta}) + o_p(n^{\frac{1}{2} - \beta \frac{K-1}{K}}).
\end{aligned}
$$

With this expression, the orders become analogous to those in Theorem 4, with $\alpha$ replacing $1/2$. In particular, the other bias terms disappear to order $o_p(1)$ provided

$$\alpha \frac{K}{K-1} < \beta < 1, \tag{32}$$

generalising (23). Notice that as $\alpha \leq 1/4$, the permissible range for $\beta$ is much bigger. The reason for this is that the bias terms scale with $n$ as $n^{\alpha - \beta}$ so low values of $\alpha$ allow $\beta$ to be lower.

A version of Theorem 5 for microstructure thus follows.

**Theorem 8** *Assume the setting of this section, that $K \geq 2$ and that (32) holds. Then*

$$n^\alpha \left( UMV_M^{(K)} - \int_0^T \sigma_t^2 \mathrm{d}t \right) \overset{\mathcal{L}s}{\to} N \left( 0, \int_0^T f_s^2 \mathrm{d}s \right), \tag{33}$$

*and*

$$n^{1/2} \left\{ \left( RV_n - UMV_M^{(K)} \right) - \sum_{0<t\leq T} |\Delta J_t|^2 \right\} \overset{\mathcal{L}s}{\to} N \left( 0, 4T \sum_{0<t\leq T} |\Delta J_t|^2 \sigma_t^2 \right). \tag{34}$$

The optimal $\beta$ is for estimating $IV$ and $JV$ is, similarly to (27), given by

$$\beta^* = 2\alpha \frac{2K}{3K-2}. \tag{35}$$

If $K = 3$ and $\alpha = 1/4$, then $\beta^* = 3/7$. Hence, when there is microstructure it makes sense to have much bigger blocks due to the slow rate of convergence.

## 6.2   Estimating the integrated variance of volatility

How quickly volatility moves is importantly economically. Mathematically this is measured by the quadratic variation of the volatility process, $[\sigma, \sigma]_T$. We call this the integrated variance of volatility $[\sigma, \sigma]_T = \int_0^T [\sigma, \sigma]_t' \mathrm{d}t$. It would be helpful to estimate $[\sigma, \sigma]_T$ nonparametrically.

Now following Theorem 4 we set $\beta = 1/2$ and define an estimator of $[\sigma, \sigma]_T$ of the form

$$\widehat{[\sigma, \sigma]}_T^{(K,n)} = c^{-1} \frac{3}{K-1} n^{1/2} \left( ECRV^{(K)} - UMV_M^{(K)} \right). \tag{36}$$

This suggests $\widehat{[\sigma, \sigma]}_T^{(K,n)} = [\sigma, \sigma]_T + O_p(n^{-\frac{1}{4}})$ where the error is conjectured to be mixed normal.

We now provide the precise result for the case of bipower variation, as follows.

**Corollary 2** *Let $a^2 = \frac{27}{16} \int_0^T c^{-3} \left( 2T\sigma_t^2 + c^2[\sigma, \sigma]_t' \right)^2 \mathrm{d}t$, the conditions of Theorem 4 hold, and $X$ have no jumps. As $n \to \infty$ when $\beta = 1/2$ then*

$$n^{\frac{1}{4}} \left( \widehat{[\sigma, \sigma]}_T^{(2,n)} - [\sigma, \sigma]_T \right) \overset{\mathcal{L}s}{\to} a \times Z,$$

*where $Z$ is standard normal and independent of the underlying filtration.*

**Proof.** This is a direct implication of Theorem 7.

# 7   Conclusion

Jumps are important in financial economics. Recent development of econometric methods to analyse jumps using high frequency data has quantified the importance of jumps. The recent empirical work has suggested they are quite common empirically.

One of the leading methods for detecting jumps is bi- and multi-power power variation. In this paper we have suggested ways of improving the efficiency of these estimators, making them efficient, and increasing the power of jump tests. We also show how to robustify multipower variation to noise and develop a central limit theory for the resulting estimator. Our results are extended to allow the first nonparametric estimator of the volatility of volatility.

We carry out detailed Monte Carlo studies of the finite sample behaviour of our blocked version of MPV and apply it to data from U.S. equity markets.

# References

Abramowitz, M. and I. A. Stegun (1970). *Handbook of Mathematical Functions*. New York: Dover Publications Inc.

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*, 707–720.

Andersen, T. G., T. Bollerslev, and F. X. Diebold (2009). Parametric and nonparametric measurement of volatility. In Y. Aït-Sahalia and L. P. Hansen (Eds.), *Handbook of Financial Econometrics*, pp. 67–138. Amsterdam: North Holland.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2000). Great realizations. *Risk 13*, 105–108.

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

Andersen, T. G., D. Dobrev, and E. Schaumburg (2009). Jump-robust volatility estimation using nearest neighbor truncation. Unpublished paper.

Andersen, T. G., D. Dobrev, and E. Schaumburg (2012). A functional filtering and truncation approach to integrated return quarticity estimation. *Journal of Econometrics*. Forthcoming.

Bahadur, R. (1964). On Fisher's bound for asymptotic variances. *Ann. Math. Statist. 35*, 1545–1552.

Bahadur, R. (1980). Hodges superefficiency. In *Encyclopedia of Statistical Sciences*, Volume 3 (F-H). New York: Wiley.

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

Barndorff-Nielsen, O. E., S. E. Graversen, J. Jacod, and N. Shephard (2006). Limit theorems for realised bipower variation in econometrics. *Econometric Theory 22*, 677–719.

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

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2009). Realised kernels in practice: trades and quotes. *Econometrics Journal 12*, C1–C32.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2011). Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. *Journal of Econometrics 162*, 149–169.

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

Barndorff-Nielsen, O. E. and N. Shephard (2004). Power and bipower variation with stochastic volatility and jumps (with discussion). *Journal of Financial Econometrics 2*, 1–48.

Barndorff-Nielsen, O. E. and N. Shephard (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics 4*, 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*, 796–806.

Bollerslev, T., G. Tauchen, and H. Zhou (2009). Expected stock returns and variance risk premia. *Review of Financial Studies 22*, 4463–4492.

Bollerslev, T. and V. Todorov (2011). Estimation of jump tails. *Econometrica 79*, 1727–1783.

Christensen, K., R. Oomen, and M. Podolskij (2010). Realised quantile-based estimation of the integrated variance. *Journal of Econometrics 159*, 74–98.

Christensen, K. and M. Podolskij (2006). Asymptotic theory for range-based estimation of quadratic variation of a discontinuous semi-martingale. Unpublished paper: Aarhus School of Business.

Christensen, K. and M. Podolskij (2007). Asymptotic theory for range-based estimation of integrated volatility of a continuous semi-martingale. *Journal of Econometrics 141*, 323–349.

Delbaen, F. and W. Schachermeyer (2006). *The Mathematics of Arbitrage.* Springer.

Dellacherie, C. and P. Meyer (1982). *Probabilities and Potential B.* Amsterdam: North-Holland.

Ghysels, E., A. C. Harvey, and E. Renault (1996). Stochastic volatility. In C. R. Rao and G. S. Maddala (Eds.), *Statistical Methods in Finance*, pp. 119–191. Amsterdam: North-Holland.

Goncalves, S. and N. Meddahi (2004). Bootstrapping realized volatility. *Econometrica 77*, 283–306.

Hájek, J. (1969). A characterization of limiting distributions of regular estimates. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 14*, 323–330.

Huang, X. and G. Tauchen (2005). The relative contribution of jumps to total price variation. *Journal of Financial Econometrics 3*, 456–499.

Jacod, J. (2008). Asymptotic properties of realized power variations and related functionals of semimartingales. *Stochastic Processes and Their Applications 118*, 517–559.

Jacod, J. (2012). Statistics and high frequency data. In M. Kessler, A. Lindner, and M. Sorensen (Eds.), *Statistical methods for stochastic differential equations*, pp. 191–310. Chapman and Hall.

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

Jacod, J. and P. Protter (1998). Asymptotic error distributions for the Euler method for stochastic differential equations. *Annals of Probability 26*, 267–307.

Jacod, J. and P. Protter (2012). *Discretization of Processes.* Heidelberg: Springer.

Jacod, J. and A. N. Shiryaev (2003). *Limit Theorems for Stochastic Processes* (2 ed.). Springer: Berlin.

LeCam, L. (1953). On some asymptotic properties of maximum likelihood estimates and related Bayes' estimates. *Univ. of California Publ. in Statistics 1*, 277–330.

Lee, S. and P. A. Mykland (2008). Jumps in financial markets: A new nonparametric test and jump dynamics. *Review of Financial Studies 21*, 2535–2563.

Lehmann, E. (1983). *Theory of Point Estimation.* New York: Wiley.

Mancini, C. (2004). Estimation of the characteristics of jump of a general Poisson-diffusion process. *Scandinavian Actuarial Journal 1*, 42–52.

Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics 36*, 270–296.

Martens, M. and D. van Dijk (2007). Measuring volatility with the realized range. *Journal of Econometrics 138*, 181–207.

Mykland, P. A. (1994). Bartlett type identities for martingales. *Annals of Statistics 22*, 21–38.

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

Mykland, P. A. and L. Zhang (2009). Inference for continuous semimartingales observed at high frequency. *Econometrica 77*, 1403–1455.

Mykland, P. A. and L. Zhang (2010). The econometrics of high frequency data. (to appear in *Statistical Methods for Stochastic Differential Equations*, M. Kessler, A. Lindner, and M. Sørensen, eds., Chapman and Hall/CRC Press).

Mykland, P. A. and L. Zhang (2012). The econometrics of high frequency data. In M. Kessler, A. Lindner, and M. Sørensen (Eds.), *Statistical Methods for Stochastic Differential Equations*, pp. 109–190. Chapman & Hall/CRC Press.

Patton, A. J. and K. K. Sheppard (2009). Good volatility, bad volatility: signed jumps and persistence of volatility. Unpublished paper: Oxford-Man Institute, University of Oxford.

Podolskij, M. and M. Vetter (2010). Understanding limit theorems for semimartingales: a short survey. *Statistica Nederlandica 64*, 329–351.

Protter, P. (2004). *Stochastic Integration and Differential Equations*. New York: Springer-Verlag.

Qi, F. (2010). Bounds for the ratio of two gamma functions. *Journal of Inequalities and Applications 2010*, 1–84.

Rootzén, H. (1980). Limit distributions for the error in approximations of stochastic integrals. *Annals of Probability 8*, 241–251.

Veraart, A. E. D. (2010). Inference for the jump part of quadratic variation of ito semimartingales. *Econometric Theory 26*, 331–368.

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*, 1394–1411.

# APPENDIX

The results to prove from the main text are list, together with the sections of this appendix where they are given.

- Section A: Remarks 1 and 2; Theorems 2, 6 and 1.

- Section B: Theorems 4 and 7.

## A   Proof of Theorems 2, 6 and 1 and Remarks 1 and 2

### A.1   Proof of Remark 1

For simplicity, take $T = 1$. Now

$$BV_M \overset{\mathcal{L}}{=} \frac{\sigma^2}{n} \frac{n}{(n_M - 1) k_{M,1}^2} \sum_{i=2}^{n_M} \chi_{M,i-1} \chi_{M,i} \overset{\mathcal{L}}{=} \frac{\sigma^2}{(n_M - 1) k_{M,1}^2} \sum_{i=2}^{n_M} \chi_{M,i-1} \chi_{M,i},$$

which has a expectation and variance of

$$\frac{\sigma^2}{(n_M - 1) \{\mathrm{E}(\chi_M)\}^2} (n_M - 1) \{\mathrm{E}(\chi_M)\}^2 = \sigma^2$$

and noting that

$$\mathrm{Var}\left(\chi_{M,1}\chi_{M,2}\right) = \mathrm{E}\left(\chi_{M,1}^2 \chi_{M,2}^2\right) - \left\{\mathrm{E}\left(\chi_{M,1}\right)\right\}^4 = M^2 - k_{M,1}^4$$

$$\mathrm{Cov}\left(\chi_{M,1}\chi_{M,2}, \chi_{M,3}\chi_{M,2}\right) = M k_{M,1}^2 - k_{M,1}^4$$

we have

$$\frac{\sigma^4}{(n_M - 1)^2 k_{M,1}^4} \left[(n_M - 1)\mathrm{Var}\left(\chi_{M,1}\chi_{M,2}\right) + 2(n_M - 2)\mathrm{Cov}\left(\chi_{M,1}\chi_{M,2}, \chi_{M,3}\chi_{M,2}\right)\right]$$

$$= \frac{\sigma^4}{(n_M - 1)^2 k_{M,1}^4} \left\{ (n_M - 1) \left( M^2 - k_{M,1}^4 \right) + 2(n_M - 2) \left( Mk_{M,1}^2 - k_{M,1}^4 \right) \right\}$$

$$= \frac{\sigma^4}{(n_M - 1)^2} \left\{ (n_M - 1) \frac{\left( M^2 + 2Mk_{M,1}^2 - 3k_{M,1}^4 \right)}{k_{M,1}^4} - \frac{2 \left( Mk_{M,1}^2 - k_{M,1}^4 \right)}{k_{M,1}^4} \right\}$$

$$= \frac{\sigma^4}{M(n_M - 1)^2} \left\{ (n_M - 1) M\theta_M - \psi_M \right\}, \quad \psi_M = 2M \left( M/k_{M,1}^2 - 1 \right) = 1 + o(1),$$

$$= \frac{\sigma^4}{M(n_M - 1)} \left( M\theta_M - \frac{1}{(n_M - 1)} \psi_M \right) \sim \frac{\sigma^4}{n} (M\theta_M), \tag{A.1}$$

the latter transition for large $n$ so long as $M/n = o(1)$. In practice it makes sense to use the finite sample constant

$$d = \frac{1}{M(n_M - 1)} \left( M\theta_M - \frac{1}{(n_M - 1)} \psi_M \right).$$

## A.2 Proof of Theorem 2

Here

$$d_{n,M,K} = \frac{1}{M(n_M - K + 1)^2} \frac{1}{\left( k_{M,r/K} \right)^{2K}} \tag{A.2}$$

$$\times \left\{ (n_M - K + 1)\mathrm{Var} \left( \prod_{i=1}^{K} v_i \right) + 2 \sum_{k=1}^{K-1} (n_M - K + 1 - k)\mathrm{Cov} \left( \prod_{i=1}^{K} v_i, \prod_{i=1}^{K} v_{i-k} \right) \right\},$$

where

$$\mathrm{Cov} \left( \prod_{i=1}^{K} v_i, \prod_{i=1}^{K} v_{i-k} \right) = \left[ \prod_{i=k+1}^{K} \mathrm{E} \left( v_i^2 \right) - \left\{ \prod_{i=k+1}^{K} \mathrm{E} \left( v_i \right)^2 \right\} \right] \prod_{i=1}^{k} \mathrm{E} \left( v_i \right)^2$$

$$= \left\{ \left( k_{M,2r/K} \right)^{K-k} - \left( k_{M,r/K} \right)^{2(K-k)} \right\} \left( k_{M,r/K} \right)^{2k}$$

$$= \left( \lambda_{M,r,K}^{K-k} - 1 \right) \left( k_{M,r/K} \right)^{2K}.$$

Now

$$\theta_M^{(r,K)} = \frac{1}{\left( k_{M,r/K} \right)^{2K}} \left\{ \mathrm{Var} \left( \prod_{i=1}^{K} v_i \right) + 2 \sum_{k=1}^{K-1} \mathrm{Cov} \left( \prod_{i=1}^{K} v_i, \prod_{i=1}^{K} v_{i-k} \right) \right\}.$$

This means that

$$\theta_M^{(r,K)} = \left( \lambda_{M,r,K}^{K} - 1 \right) + 2 \sum_{k=1}^{K-1} \left( \lambda_{M,r,K}^{K-k} - 1 \right) = 2\lambda_{M,r,K} \sum_{k=0}^{K-1} \lambda_{M,r,K}^{k} - \lambda_{M,r,K}^{K} - 2K + 1.$$

Returning to the notation in Remark 2, we have

$$\lambda_r = \frac{\left( k_{M,2r} \right)}{\left( k_{M,r} \right)^2} = \frac{\left( k_{M,2r}/M^r \right)}{\left( k_{M,r} \right)^2 /M^r} \simeq \frac{1 + \frac{a_{2r}}{M} + \frac{b_{2r}}{M^2}}{1 + \frac{g_r}{M} + \frac{h_r}{M^2}}$$

$$\sim \left( 1 + \frac{a_{2r}}{M} + \frac{b_{2r}}{M^2} \right) \left( 1 - \frac{g_r}{M} + \frac{g_r^2 - h_r}{M^2} \right) \sim 1 + \frac{a_{2r} - g_r}{M} + \frac{b_{2r} + g_r^2 - h_r - a_{2r}g_r}{M^2}$$

$$= 1 + \frac{j_r}{M} + \frac{k_r}{M^2}.$$

Of course, by Taylor expansion,

$$\lambda_r^K - 1 \sim K\frac{j_r}{M} + \frac{1}{2}\frac{K(K-1)j_r^2 + 2Kk_r}{M^2}.$$

Now

$$M\left\{\left(\lambda_{M,r/K}^K - 1\right) + 2\left(\lambda_{M,r/K}^{K-1} - 1\right) + 2\left(\lambda_{M,r/K}^{K-2} - 1\right) + ... + 2\left(\lambda_{M,r/K} - 1\right)\right\}$$

$$\sim \ j_{r/K}\left\{K + 2(K-1) + ... + 2\right\} + \frac{j_{r/K}^2}{2M}\left\{K(K-1) + 2(K-1)(K-2) + ... + 2\right\}$$

$$+ \frac{k_{r/K}}{M}\left\{K + 2(K-1) + ... + 2\right\},$$

$$= \ \vartheta_{r/K,K,M}$$

$$j_r = \ a_{2r} - 2a_r,$$

$$k_r = \ b_{2r} + (2a_r)^2 - h_r - 2a_{2r}a_r$$

The important cases are given in Tables 11 and 12.

| $r$ | 4 | 2 | $\frac{4}{3}$ | 1 | $\frac{2}{3}$ | $\frac{1}{2}$ |
|-----|---|---|---------------|---|---------------|---------------|
| $j_r$ | 8 | 2 | $\frac{8}{9}$ | $\frac{1}{2}$ | $\frac{2}{9}$ | $\frac{1}{8}$ |
| $k_r$ | 8 | 0 | $\frac{8}{81}$ | $\frac{1}{8}$ | $\frac{8}{81}$ | $\frac{9}{16\times8}$ |

Table 11: Coefficients $j_r$ and $k_r$ as a function of $r$.

| $K$ | $K + 2(K-1)$ $+...+2$ | $K(K-1) + 2(K-1)(K-2)$ $+...+2$ | $\vartheta_{2/K,K,M}$ | $\vartheta_{4/K,K,M}$ |
|-----|-----------------------|----------------------------------|-----------------------|-----------------------|
| 1 | 1 | 0 | 2 | $8 + 8\frac{1}{M}$ |
| 2 | 4 | 2 | $2 + \frac{3}{4}\frac{1}{M}$ | $8 + 4\frac{1}{M}$ |
| 3 | 9 | 10 | $2 + \frac{92}{81}\frac{1}{M}$ | $8 + \frac{392}{81}\frac{1}{M}$ |
| 4 | 16 | 28 | $2 + \frac{43}{32}\frac{1}{M}$ | $8 + \frac{11}{2}\frac{1}{M}$ |

Table 12: Coefficients $\vartheta_{2/K,K,M}$ and $\vartheta_{4/K,K,M}$ as a function of $K$.

## A.3 Proof of Remark 2

Note that (e.g. Abramowitz and Stegun (1970, 6.1.47) and Qi (2010, equation (1.11))) for large $z$

$$z^{b-a}\frac{\Gamma(z+a)}{\Gamma(z+b)} \sim 1 + \frac{(a-b)(a+b-1)}{2z} + \frac{(a-b)(a+b-1)}{24z^2}\left\{3(a+b-1)^2 - a + b - 1\right\}...$$

Thus as $M \to \infty$ then

$$\frac{k_{M,r}}{M^{r/2}} = \left(\frac{M}{2}\right)^{-r/2}\frac{\Gamma\left(\frac{M+r}{2}\right)}{\Gamma\left(\frac{M}{2}\right)}$$

$$\sim \quad 1 + \frac{\frac{r}{2}\left(\frac{r}{2}-1\right)}{2\left(\frac{M}{2}\right)} + \frac{\frac{r}{2}\left(\frac{r}{2}-1\right)}{24\left(\frac{M}{2}\right)^2}\left\{3\left(\frac{r}{2}-1\right)^2 - \frac{r}{2} - 1\right\}$$

$$= \quad 1 + \frac{r\,(r-2)}{4M} + \frac{r\,(r-2)}{24M^2}\left\{\frac{3}{4}(r-2)^2 - \frac{r}{2} - 1\right\}$$

$$= \quad 1 + \frac{r\,(r-2)}{4M} + \frac{r\,(r-2)}{24M^2}\left\{\frac{3r^2}{4} - \frac{7r}{2} + 2\right\}$$

$$= \quad 1 + \frac{r\,(r-2)}{4M} + \frac{r\,(r-2)\left(3r^2 - 14r + 8\right)}{4 \times 24M^2}$$

An important special case is when $r = 1$, then

$$\frac{k_{M,1}}{M^{1/2}} \sim 1 - \frac{1}{4M} + \frac{3}{4}\frac{1}{24M^2}...$$

Of course as $x \downarrow 0$ so

$$g(x) = (1 + ax + bx^2)^{-1} = g(0) + xg'(0) + \frac{x^2}{2}g''(0) + O(x^3) = 1 - ax + x^2(a^2 - b) + ...$$

So

$$\frac{M^{1/2}}{k_{M,1}} \sim 1 + \frac{1}{4M} + \left(\frac{1}{16} - \frac{3}{4}\frac{1}{24}\right)\frac{1}{M^2} = 1 + \frac{1}{4M} + \frac{1}{32}\frac{1}{M^2}.$$

More generally

$$\frac{M^{r/2}}{k_{M,r}} \sim 1 - \frac{r\,(r-2)}{4M} + \left\{\frac{r^2\,(r-2)^2}{16} - \frac{r\,(r-2)\left(3r^2 - 14r + 8\right)}{4 \times 24}\right\}\frac{1}{M^2}.$$

## A.4  Proof of Theorem 6

Since $J$ is independent of $X^c$, we can condition on the $J$ process. Also, it is enough to show the result when it is known that the jumps are in intervals $(t_{i_1}, t_{i_1+1}], ..., (t_{i_\nu}, t_{i_\nu+1}]$, where $\nu$ is nonrandom (this adds information, and so can only reduce any lower bound). By the absolute continuity condition in the statement of the theorem, we thus only need to show the result when the jumps are of nonrandom size almost everywhere (with respect to Lebesgue measure). The problem then reduces to estimating the jump sizes (as parameters) when the observations are $X_{t_{i_1+1}} - X_{t_{i_1}}, ..., X_{t_{i_\nu+1}} - X_{t_{i_\nu}}$. If these increments had normal distributions around jumps, the result would follow from the superefficiency results in LeCam (1953). The generalization to the non-normal case follows from Mykland and Zhang (2009, Thm 1, p. 1411), since the contiguity correction does not alter the asymptotic variance.

## A.5  Proof of Theorem 1

For compactness of notation, set $\theta = \int_0^T \sigma_t^2 dt$, and let $\hat{\theta}_n$ be $BV_M$ based on $n$ observations. (Not to be confused with $\theta_M$ in (9).) $RV_i$ is based on blocks of size $M$. Let $\mathcal{M}$ be a large integer.

Decompose

$$\hat{\theta}_n = \hat{\theta}_n^{(1)} + \hat{\theta}_n^{(2)}, \text{ where}$$

$$\hat{\theta}_n^{(1)} = \frac{M}{k_M} \sum_k \sum_{i=k\mathcal{M}+1}^{(k+1)\mathcal{M}-1} (RV_i RV_{i+1})^{1/2}, \quad \text{and} \quad \hat{\theta}_n^{(2)} = \frac{M}{k_M} \sum_k (RV_{k\mathcal{M}} RV_{k\mathcal{M}+1})^{1/2}.$$

For the first term, note that,

$$n^{1/2} \sum_k \left( \frac{M}{k_M} \sum_{i=k\mathcal{M}+1}^{(k+1)\mathcal{M}-1} (RV_i RV_{i+1})^{1/2} - \frac{\mathcal{M}-1}{\mathcal{M}} \sigma_{\tau_{n,k}}^2 \Delta\tau \right)$$

is a sum of $(\mathcal{Y}_{n,k}, Q_n)$ martingale increments in the sense of the definitions in Mykland and Zhang (2009, p. 1417 & p. 1421). The block size is $M\mathcal{M}$ in lieu of $M$. $\tau_{n,k}$ is also as defined in this paper.

The martingale has quadratic variation

$$n \sum_k \text{Var}_{Q_n} \left( \frac{M}{k_M} \sum_{i=k\mathcal{M}+1}^{(k+1)\mathcal{M}-1} (RV_i RV_{i+1})^{1/2} - \frac{\mathcal{M}-1}{\mathcal{M}} \sigma_{\tau_{n,k}}^2 \Delta\tau \mid \mathcal{Y}_{n,k} \right)$$

$$= n \sum_k \sigma_{\tau_{n,k}}^4 \left( \frac{M}{k_M} \right)^2 \Delta t_n^2 \text{Var} \left( \sum_{i=1}^{\mathcal{M}-1} \chi_{M,i} \chi_{Mi+1} \right)$$

$$\to \frac{T}{M\mathcal{M}} \left( \frac{M}{k_M} \right)^2 \text{Var} \left( \sum_{i=1}^{\mathcal{M}-1} \chi_{M,i} \chi_{M,i+1} \right) \int_0^T \sigma_t^4 \mathrm{d}t, \tag{A.3}$$

in probability as $n \to \infty$, where the $\chi_{M,i}$ are i.i.d. $\chi_M$.

The same arguments as Mykland and Zhang (2009) imply $n^{1/2}(\hat{\theta}_n^{(1)} - \theta(\mathcal{M}-1)/\mathcal{M})$ converges stably under $Q_n$ to a mixed normal distribution with mean 0 and variance (A.3). It is easy to see that there is no adjustment to $P_n^*$, $P^*$, or $P$, as these measures are defined in the referenced paper.

Similar arguments, using a first block of size $\mathcal{M}$, and then blocks of size $M\mathcal{M}$, yield that $n^{1/2}(\hat{\theta}_n^{(2)} - \theta/\mathcal{M})$ converges stably, to a variance which is of $o_p(1)$ as $\mathcal{M} \to \infty$ ($\mathcal{M}$ is sent to $\infty$ *after* $n$). Standard weak convergence arguments involving tightness and subsequences of subsequences thus yield that $n^{1/2}(\hat{\theta}_n - \theta)$ converges stably under $P$ to a mixed normal limit with mean zero and a random variance which is the limit of (A.3) as $\mathcal{M} \to \infty$, in other words, $TM\theta_M \int_0^T \sigma_t^4 \mathrm{d}t$, where this $\theta_M$ is (9). This shows the first part of the theorem. The second part follows similarly.

# B  Multipower variation as $M \to \infty$ and $n \to \infty$

## B.1  Assumptions

**Assumption 1** . *The continuous part of the process $X_t$ is adapted to a filtration $(\mathcal{F}_t)$ which is generated by Brownian motions $W_t$. That*

$$X_t = X_0 + \int_0^t \mu_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s + J_t, \tag{B.4}$$

*where $\mu$ is a predictable locally bounded drift, $W$ is a standard Brownian motion and $J$ is a finite activity jump process. $\sigma_t$ is also locally bounded away from zero, and is itself a continuous semimartingale (Itô process).*

**Assumption 2** . *Observations $t_i = \frac{i}{n}$, $i = 0, 1, 2, ..., n$. With blocks of size $M$ we have block boundaries $\tau_i = \frac{Mi}{n}$, $i = 0, 1, ..., \lfloor n/M \rfloor = n_M$.*

**Assumption 3** . *As $n \to \infty$ with $\beta \in (0, 1)$,*

$$n^\beta \frac{M}{n} \to \frac{c}{T}. \tag{B.5}$$

**Assumption 4**

$$0 < \beta < 1. \tag{B.6}$$

In the following section, the observed process is assumed continuous, i.e., $J_t \equiv 0$. The task is to prove the Theorems 4 and 7 given in the main text. To obtain those results we go through various preliminary results and definitions. Proofs of each of the stated theorems are given at the end of this section.

## B.2 Some building blocks

### B.2.1 Exact representation of multipower variation

To transparently study the case where $M \to \infty$ as $n \to \infty$, we use the simplified estimator "semi-multipower variation" for the exponent $r = 2$:

$$SMV_n^{(K)} = K \sum_{1 \leq i; KiM \leq n} (RV_{Ki-K+1} \times ... \times RV_{Ki})^{1/K}, \tag{B.7}$$

with "semi-bipower variation" given by:

$$SBV_n = SMV_n^{(2)} = 2 \sum_{1 \leq i; 2iM \leq n} (RV_{2i-1} RV_{2i})^{1/2}. \tag{B.8}$$

We have here used the limiting value 1 of $(k_{M,2/K})^K/M$ (for $K = 2$, this is the same as $k_M/M$). For simplicity we avoid overlapping intervals. They are studied in Appendix B.2.4 and beyond.

To avoid edge effects, we modify the realised volatility as follows:

$$RV_n^{(K)} = \sum_{i=1}^{K \lfloor n/KM \rfloor} RV_i. \tag{B.9}$$

Now set

$$f(x_1, ..., x_K) = \sum_{k=1}^{K} x_k - K \prod_{k=1}^{K} x_k^{1/K}. \tag{B.10}$$

34

The key *exact* representation is now the following:

$$RV_n^{(K)} - SMV_n^{(K)} = \sum_{1 \le i; KiM \le n} f(RV_{Ki-K+1}, ..., RV_{Ki})$$

$$= \sum_{1 \le i; KiM \le n} f(RV_{Ki-K+1}/\Delta\tau, ..., RV_{Ki}/\Delta\tau)\Delta\tau \tag{B.11}$$

with $\Delta\tau = M\Delta t = MT/n$. In the bipower case, $f(x, y) = (y^{1/2} - x^{1/2})^2$, which is non-negative.

### B.2.2 An approximation theorem for bipower variation

The results here and the next subsection are stronger than those of Theorem 4 in that they provide stable convergence to a normal limit in place of an $O_p$ term. This gives reasonable certainty that a similar stable convergence holds in the earlier theorem, and thus that the $O_p$ term cited there does not vanish. We show in Appendix B.2.6 that the stronger result holds for bipower variation.

Based on the representation (B.11), we obtain the following result under model (7).

**Theorem 9** *Under the assumptions of Theorem 4, setting* $a^2 = \frac{1}{2} \int_0^T c^{-1} \left(2T\sigma_t^2 + \frac{1}{3}c^2[\sigma, \sigma]_t'\right)^2 dt$, *and if X has no jumps,* $b_1 = c^{-1}\frac{1}{2}T \int_0^T \sigma_t^2 dt$ *and* $b_2 = \frac{1}{3}c[\sigma, \sigma]_T$ . *Then if* $\beta = 1/2$,

$$n^{\frac{3}{4}} \left\{ RV_n^{(2)} - SBV_n - (n^{-1+\beta}b_1 + n^{-\beta}b_2) \right\} \xrightarrow{\mathcal{L}s} a \times Z, \tag{B.12}$$

*where Z is standard normal and independent of the underlying filtration. When* $\beta > 1/2$, *only keep the first term inside the brackets in* $a^2$, *and replace* $n^{\frac{3}{4}}$ *by* $n^{1-\frac{1}{2}\beta}$ *in (B.12). When* $\beta < 1/2$, *only keep the second term inside the brackets in* $a^2$, *and replace* $n^{\frac{3}{4}}$ *by* $n^{\frac{3}{2}\beta}$.

**Proof.** Given in the Appendix.

### B.2.3 Microstructure and multipower

To show that it is possible to pursue a similar theory in the more general case of estimation under microstructure, we provide a theorem where "$RV_i$" is a general estimator of the integrated volatility in interval # $i$.

**Theorem 10** *Assume that under the statistical risk neutral distribution (Mykland and Zhang (2009, Sect 2.2)), there is a sequence of continuous martingales* $(M_t^{(n)})_{0 \le t \le T}$ *so that*

$$RV_i = \int_{\tau_{i-1}}^{\tau_i} \sigma_t^2 dt + M_{\tau_i}^{(n)} - M_{\tau_{i-1}}^{(n)} = \int_{\tau_{i-1}}^{\tau_i} \sigma_t^2 dt + \Delta M_{\tau_i}^{(n)}. \tag{B.13}$$

*Suppose that as* $n \to \infty$,

$$n^{2\alpha}[M^{(n)}, M^{(n)}]_t \xrightarrow{p} \int_0^t f_s^2 ds \text{ and } n^{\alpha}[M^{(n)}, W^{(i)}]_t \xrightarrow{p} 0 \text{ for } i = 1, ..., p. \tag{B.14}$$

*Also assume the conditions of Appendix B.1, except that $J \equiv 0$. In the case when $\beta = \alpha$, set*

$$a^2 = \frac{1}{2} \int_0^T c^{-1}(\sigma_t^{-2} f_t^2 + \frac{1}{3} c^2 [\sigma, \sigma]_t')^2 \mathrm{d}t \text{ and } b_1 = \frac{1}{4} c^{-1} \int_0^T \sigma_t^{-2} f_t^2 \mathrm{d}t \text{ and } b_2 = \frac{1}{3} c [\sigma, \sigma]_T . \quad \text{(B.15)}$$

*Then, under Assumption 1,*

$$n^{\frac{3}{2}\alpha} \left\{ RV_n - SBV_n - (n^{-2\alpha+\beta} b_1 + n^{-\beta} b_2) \right\} \overset{\mathcal{L}s}{\to} a \times Z \quad \text{(B.16)}$$

*where $Z$ is standard normal and independent of the underlying filtration. When $\beta > \alpha$, only keep the first term inside the brackets in $a^2$, and replace $n^{\frac{3}{2}\alpha}$ by $n^{2\alpha-\frac{1}{2}\beta}$ in (B.16). When $\beta < \alpha$, only keep the second term inside the brackets in $a^2$, and replace $n^{\frac{3}{2}\alpha}$ by $n^{\frac{3}{2}\beta}$.*

**Proof.** Given in the Appendix.

Likewise, for multipower variation, we have

**Theorem 11** *Assume the conditions of Theorem 10. Set*

$$b_1 = \frac{1}{2} c^{-1} \frac{K-1}{K} \int_0^T \sigma_t^{-2} f_t^2 \mathrm{d}t \text{ and } b_2 = \frac{K-1}{3} c [\sigma, \sigma]_T . \quad \text{(B.17)}$$

*Then*

$$RV_n^{(K)} - SMV_n^{(K)} = n^{-2\alpha+\beta} b_1 + n^{-\beta} b_2 + O_p(n^{-2\alpha+\frac{1}{2}\beta} + n^{-\frac{3}{2}\beta}). \quad \text{(B.18)}$$

### B.2.4 Passage to overlapping blocks, and the edge effect

Define unnormalised multipower variation:

$$UMV_n^{(K)} = \sum_{i=K}^{n_M} \prod_{k=0}^{K-1} (RV_{i-k})^{1/K}. \quad \text{(B.19)}$$

As in the previous section, we ignore (for the purposes of notation only) the difference between $\widehat{RV}_i$ and $RV_i$ (microstructure and no microstructure). For comparison, for $r = 2$,

$$MV_M^{(2,K)} \text{ in eq.(16)} = \frac{M}{\left(k_{M,2/K}\right)^K} UMV_n^{(K)} \text{ and}$$

$$MV_M^{(2,K)} \text{ in eq.(17)} = \frac{n}{(n_M - K + 1)(k_{M,2/K})^K} UMV_M^{(K)}.$$

For these normalizations, note for later reference that

$$1 - \frac{(k_{M,2/K})^K}{M} = \frac{1}{M} \frac{K-1}{K} + O(M^{-2})) \text{ and}$$

$$\left\{ \frac{(n_M - K + 1)}{n} - \frac{1}{M} \right\} (k_{M,2/K})^K = O(M/n) = O(n^{-\beta}), \text{ so that}$$

$$1 - \frac{(n_M - K + 1)(k_{M,2/K})^K}{n} = \frac{1}{M} \frac{K-1}{K} + O(M^{-2}) + O(M/n). \quad \text{(B.20)}$$

Replace (B.7)-(B.9) by

$$SMV_n^{(K,m)} = K \sum_{1 \leq i; Ki+m-1 \leq n_M} (RV_{Ki-K+m} \times ... \times RV_{Ki+m-1})^{1/K} \text{ and}$$

$$RV_n^{(K,m)} = \sum_{i=m}^{m-1+K\lfloor (n_M-m+1)/K \rfloor} RV_i, \tag{B.21}$$

where $1 \leq m \leq K$. The same representation (B.11), and theorems, will apply. Now observe that

$$UMV_n^{(K)} = \frac{1}{K} \sum_{m=1}^{K} SMV_n^{(K,m)}. \tag{B.22}$$

To verify this note that no term in $UMV_n^{(K)}$ occurs in more than once $SMV_n^{(K,m)}$ for $1 \leq m \leq K$.

The comparable exact representation is therefore based on an edge effect adjusted RV, and on using equation (B.11) on each component of an average of $RV_n^{(K,m)} - SMV_n^{(K,m)}$:

$$ECRV_n^{(K)} - UMV_n^{(K)} = \frac{1}{K} \sum_{m=1}^{K} \left( RV_n^{(K,m)} - SMV_n^{(K,m)} \right), \tag{B.23}$$

where the Edge Effect Corrected RV is given by

$$ECRV_n^{(K)} = \frac{1}{K} \sum_{m=1}^{K} RV_n^{(K,m)}. \tag{B.24}$$

**Remark 8** *(Largest common index.)* *The index $m = n_M - K + 1$ is the largest one appearing in all the terms of the above sum (B.24). This motivates (B.28) below. To see this, set $m_0 = n_M - K\lfloor n_M/K \rfloor$ and $u = m_0 - m + 1$, and note that the largest common index has value*

$$\min_{1 \leq m \leq K} (m - 1 + K\lfloor (n_M - m + 1)/K \rfloor) = K\lfloor n_M/K \rfloor + \min_{1 \leq m \leq K} (m - 1 + K\lfloor (m_0 - m + 1)/K \rfloor)$$

$$= K\lfloor n_M/K \rfloor + \min_{m_0-K+1 \leq u \leq m_0} (m_0 - u + K\lfloor u/K \rfloor)$$

$$= n_M + \min_{m_0-K+1 \leq u \leq m_0} (-u + K\lfloor u/K \rfloor)$$

$$= n_M - K + 1, \quad \text{since } 0 = m_0 < K.$$

By averaging and with very minor modifications in the proof, Theorem 11 applies identically:

**Theorem 12** *(Extension of Theorem 11). Under the assumptions of Theorem 11,*

$$ECRV_n^{(K)} - UMV_n^{(K)} = n^{-2\alpha+\beta}b_1 + n^{-\beta}b_2 + O_p(n^{-2\alpha+\frac{1}{2}\beta} + n^{-\frac{3}{2}\beta}), \tag{B.25}$$

$$RV_n - UMV_n^{(K)} = n^{-2\alpha+\beta}b_1 + n^{-\beta}b_2 + \Upsilon_n + O_p(n^{-2\alpha+\frac{1}{2}\beta} + n^{-\frac{3}{2}\beta}). \tag{B.26}$$

**Remark 9** *First consider the ordinary RV up to observation $Mn_M$. Denote this by $\tilde{RV}$:*

$$\tilde{RV}_n = \sum_{i=1}^{m_n} RV_i. \tag{B.27}$$

37

*The difference between ECRV and ordinary RV is given by*

$$\tilde{RV}_n - ECRV_n^{(K)} = \sum_{m=1}^{K} \left(1 - \frac{m}{K}\right) RV_m \;+\; \text{a corresponding term on the right edge}$$

$$= \left\{ \sum_{m=1}^{K} \left(1 - \frac{m}{K}\right) \sigma_0^2 \Delta\tau + \sum_{m=1}^{K} \left(1 - \frac{m}{K}\right) \sigma_T^2 \Delta\tau \right\} \left\{1 + O_p(n^{-\alpha})\right\}$$

$$= \frac{1}{2}\Delta\tau(K-1)(\sigma_0^2 + \sigma_T^2)\Delta\tau \left\{1 + O_p(n^{-\alpha}) + O_p(\Delta\tau^{1/2})\right\}$$

$$= n^{-\beta}\frac{1}{2}c(K-1)(\sigma_0^2 + \sigma_T^2)\left\{1 + O_p(n^{-\alpha}) + O_p(n^{-\frac{1}{2}\beta})\right\}. \tag{B.28}$$

*On the other hand, if $RV_n$ denotes the estimator based on all observations,*

$$RV_n = \tilde{RV}_n = T\frac{n - Mn_M}{n}\sigma_T^2 \left\{1 + O_p(n^{-\alpha}) + O_p(n^{-\frac{1}{2}\beta})\right\}. \tag{B.29}$$

*The full edge effect in Theorems 4 is thus given by the sum of these two:*

$$\Upsilon_n = RV_n - ECRV_n^{(K)} = O_p(n^{-\beta}).$$

### B.2.5 The case with no microstructure

**Theorem 13** *Under the assumptions of Theorem 4, and if $X$ has no jumps,*

$$n^{1/2}\left(MV_M^{(K)} - ECRV^{(K)}\right) = -n^{-\beta+\frac{1}{2}}\frac{K-1}{3}c[\sigma,\sigma]_T + O_p(n^{-\frac{1}{2}+\frac{1}{2}\beta} + n^{\frac{1}{2}-\frac{3}{2}\beta}). \tag{B.30}$$

**Proof.** Given below.

### B.2.6 Overlapping blocks: $[\sigma,\sigma]$, and reconciling the fixed $M$ case

A similar result to Theorem 9 covers overlapping blocks. This is declared as Theorem 7 in the main text. In the no microstructure special case, we obtain.

**Theorem 14** *(Bipower, overlapping, no noise). Under the assumptions of Theorem 4, if $X$ has no jumps, let $b_1$ & $b_2$ be from Theorem 9, while $a^2 = \frac{3}{16}\int_0^T c^{-1}\left(2T\sigma_t^2 + c^2[\sigma,\sigma]_t'\right)^2 dt$. Then*

$$n^{\frac{3}{4}}\left(ECRV_n^{(2)} - UMV_n^{(2)} - (n^{-1+\beta}b_1 + n^{-\beta}b_2))\right) \overset{\mathcal{L}s}{\to} a \times Z \text{ and}$$

$$n^{\frac{3}{4}}\left(ECRV_n^{(2)} - MV_n^{(2)} - n^{-\beta}b_2\right) \overset{\mathcal{L}s}{\to} a \times Z, \tag{B.31}$$

*where $Z$ is standard normal & independent of the underlying filtration. When $\beta > 1/2$, only keep the first term inside the brackets in $a^2$, and replace $n^{\frac{3}{4}}$ by $n^{1-\frac{1}{2}\beta}$ in (B.31). When $\beta < 1/2$, only keep the second term inside the brackets in $a^2$, and replace $n^{\frac{3}{4}}$ by $n^{\frac{3}{2}\beta}$ in (B.31).*

## B.3 Proofs

### B.3.1 Proof of Theorem 10

We Taylor expand $f(RV_{2i-1}/\Delta\tau, RV_{2i}/\Delta\tau)$ around $f(\sigma^2_{\tau_{2i-2}}, \sigma^2_{\tau_{2i-2}})$. Note that $f(z,z) = f_x(z,z) = f_y(z,z) = 0$, while $f_{x,x}(z,z) = f_{y,y}(z,z) = -f_{x,y}(z,z) = 1/2z$. Hence, up to second order,

$$
\begin{aligned}
RV^{(2)} - SBV &\approx \frac{1}{2}\Delta\tau \sum_{1\leq i;2iM\leq n} \left\{ \left(\frac{RV_{2i-1}}{\Delta\tau} - \sigma^2_{\tau_{2i-2}}\right)^2 f_{x,x}(\sigma^2_{\tau_{2i-2}}, \sigma^2_{\tau_{2i-2}}) \right. \\
&\qquad\qquad + \left(\frac{RV_{2i}}{\Delta\tau} - \sigma^2_{\tau_{2i-2}}\right)^2 f_{y,y}(\sigma^2_{\tau_{2i-2}}, \sigma^2_{\tau_{2i-2}}) \\
&\qquad\qquad \left. + 2\left(\frac{RV_{2i-1}}{\Delta\tau} - \sigma^2_{\tau_{2i-2}}\right)\left(\frac{RV_{2i}}{\Delta\tau} - \sigma^2_{\tau_{2i-2}}\right) f_{x,y}(\sigma^2_{\tau_{2i-2}}, \sigma^2_{\tau_{2i-2}}) \right\} \\
&= \frac{1}{4}\Delta\tau \sum_{1\leq i;2iM\leq n} \sigma^{-2}_{\tau_{2i-2}} \left(\frac{RV_{2i}}{\Delta\tau} - \frac{RV_{2i-1}}{\Delta\tau}\right)^2 \\
&= \frac{1}{4}\frac{1}{\Delta\tau} \sum_{1\leq i;2iM\leq n} \sigma^{-2}_{\tau_{2i-2}} (RV_{2i} - RV_{2i-1})^2 .
\end{aligned}
\tag{B.32}
$$

The remainder term in the above is of order $o_p(n^{-\frac{3}{2}\alpha})$, as explained at the end of the proof. Since both the bias and stochastic terms are of order $O_p(n^{-\frac{3}{2}\alpha})$, the remainder can be ignored.

To get a further handle on this quantity recall (B.13). By Itô's formula,

$$
\int_{\tau_{2i-1}}^{\tau_{2i}} \sigma^2_t dt = \sigma^2_{\tau_{2i-1}}\Delta\tau + \int_{\tau_{2i-1}}^{\tau_{2i}} (\tau_{2i} - t)d\sigma^2_t \text{ and } \int_{\tau_{2i-2}}^{\tau_{2i-1}} \sigma^2_t dt = \sigma^2_{\tau_{2i-1}}\Delta\tau - \int_{\tau_{2i-2}}^{\tau_{2i-1}} (t - \tau_{2i-2})d\sigma^2_t.
\tag{B.33}
$$

Thus, $RV_{2i} - RV_{2i-1} = L^{(n)}_{\tau_{2i}} - L^{(n)}_{\tau_{2i-2}}$, where $dL^{(n)}_t = -dM^{(n)}_t - (t - \tau_{2i-2})d\sigma^2_t$ for $\tau_{2i-2} < t < \tau_{2i-1}$, and $dL^{(n)}_t = dM^{(n)}_t + (\tau_{2i} - t)d\sigma^2_t$ for $\tau_{2i-1} < t < \tau_{2i}$. In particular,

$$
d[L^{(n)}, L^{(n)}]_t = \begin{cases} d[M^{(n)}, M^{(n)}]_t + (t - \tau_{2i-2})^2 d[\sigma^2, \sigma^2]_t + 2(t - \tau_{2i-2})d[M^{(n)}, \sigma^2]_t, & \tau_{2i-2} < t < \tau_{2i-1}, \\ d[M^{(n)}, M^{(n)}]_t + (\tau_{2i} - t)^2 d[\sigma^2, \sigma^2]_t + 2(\tau_{2i} - t)d[M^{(n)}, \sigma^2]_t, & \tau_{2i-1} < t < \tau_{2i}. \end{cases}
$$

The cross term involving $[M^{(n)}, \sigma^2]_t$ will disappear to relevant order, from Condition (B.14).

Recall that $n^\alpha M^{(n)}_t$ converges in law to $M_t$ as a process, so that in particular, $n^{2\alpha}[M^{(n)}, M^{(n)}]_t$ converges to $[M, M]_t$. Note that

$$
\begin{aligned}
[L^{(n)}, L^{(n)}]_{\tau_{2i}} - [L^{(n)}, L^{(n)}]_{\tau_{2i-2}} &\approx n^{-2\alpha}([M, M]_{\tau_{2i}} - [M, M]_{\tau_{2i-2}}) \\
&\qquad + \int_{\tau_{2i-2}}^{\tau_{2i-1}} (t - \tau_{2i-2})^2 d[\sigma^2, \sigma^2]_t + \int_{\tau_{2i-1}}^{\tau_{2i}} (\tau_{2i} - t)^2 d[\sigma^2, \sigma^2]_t \\
&\approx n^{-2\alpha}([M, M]_{\tau_{2i}} - [M, M]_{\tau_{2i-2}}) + \frac{2}{3}(\Delta\tau)^3 [\sigma^2, \sigma^2]'_{\tau_{2i-2}},
\end{aligned}
\tag{B.34}
$$

where "prime" denotes differentiation with respect to time.

39

We therefore get

r.h.s. of $(B.32)$

$$\approx \frac{1}{4}\frac{1}{\Delta\tau}\sum_{1\leq i;2iM\leq n}\sigma_{\tau_{2i-2}}^{-2}\left([M^{(n)},M^{(n)}]_{\tau_{2i}} - [M^{(n)},M^{(n)}]_{\tau_{2i-2}} + \frac{2}{3}(\Delta\tau)^3[\sigma^2,\sigma^2]'_{\tau_{2i-2}}\right) \quad (B.35)$$

$$\approx \frac{1}{4}\frac{1}{\Delta\tau}\int_0^T \sigma_t^{-2}\mathrm{d}[M^{(n)},M^{(n)}]_t + \frac{1}{12}(\Delta\tau)\int_0^T \sigma_t^{-2}\mathrm{d}[\sigma^2,\sigma^2]_t$$

$$\approx \frac{1}{4}\frac{1}{n^{2\alpha}\Delta\tau}\int_0^T \sigma_t^{-2}\mathrm{d}[M,M]_t + \frac{1}{12}\Delta\tau\int_0^T \sigma_t^{-2}\mathrm{d}[\sigma^2,\sigma^2]_t. \quad (B.36)$$

From (B.6) we obtain

$$\text{r.h.s. of } (B.32) \approx n^{-2\alpha+\beta}\frac{1}{4}c^{-1}\int_0^T \sigma_t^{-2}d[M,M]_t + n^{-\beta}\frac{1}{12}c\int_0^T \sigma_t^{-2}\mathrm{d}[\sigma^2,\sigma^2]_t$$

$$= n^{-2\alpha+\beta}\frac{1}{4}c^{-1}\int_0^T \sigma_t^{-2}d[M,M]_t + n^{-\beta}\frac{1}{3}c\,[\sigma,\sigma]_T \quad (B.37)$$

since $\mathrm{d}\sigma_t^2 = 2\sigma_t\mathrm{d}\sigma_t + \mathrm{d}[\sigma,\sigma]_t$.

Meanwhile, the error in approximation in (B.36) has quadratic variation

stoch term qv

$$= \left(\frac{1}{4}\frac{1}{\Delta\tau}\right)^2\sum_{1\leq i;2iM\leq n}\sigma_{\tau_{2i-2}}^{-4}4\int_{\tau_{2i-2}}^{\tau_{2i}}([L^{(n)},L^{(n)}]_t - [L^{(n)},L^{(n)}]_{\tau_{2i-2}})\mathrm{d}[L^{(n)},L^{(n)}]_t \quad (B.38)$$

$$= \left(\frac{1}{4}\frac{1}{\Delta\tau}\right)^2\sum_{1\leq i;2iM\leq n}\sigma_{\tau_{2i-2}}^{-4}2([L^{(n)},L^{(n)}]_{\tau_{2i}} - [L^{(n)},L^{(n)}]_{\tau_{2i-2}})^2$$

$$\approx \left(\frac{1}{4}\frac{1}{\Delta\tau}\right)^2\sum_{1\leq i;2iM\leq n}\sigma_{\tau_{2i-2}}^{-4}2(n^{-2\alpha}([M,M]_{\tau_{2i}} - [M,M]_{\tau_{2i-2}}) + \frac{2}{3}(\Delta\tau)^3[\sigma^2,\sigma^2]'_{\tau_{2i-2}})^2$$

$$\approx \left(\frac{1}{4}\frac{1}{\Delta\tau}\right)^2\sum_{1\leq i;2i)M\leq n}\sigma_{\tau_{2i-2}}^{-4}2(n^{-2\alpha}2\Delta\tau[M,M]'_{\tau_{2i-2}} + \frac{2}{3}(\Delta\tau)^3[\sigma^2,\sigma^2]'_{\tau_{2i-2}})^2, \quad (B.39)$$

where the relative error is of smaller order in probability, from (B.34). Under (B.6), consider first the case where $\beta = \alpha$. Then

$$\text{stoch term qv } = \frac{1}{16}\sum_{1\leq i;2iM\leq n}\sigma_{\tau_{2i-2}}^{-4}2n^{-3\alpha}c^{-1}\Delta\tau(2[M,M]'_{\tau_{2i-2}} + \frac{2}{3}c^2[\sigma^2,\sigma^2]'_{\tau_{2i-2}})^2 + o_p(n^{-3\alpha})$$

$$= n^{-3\alpha}\frac{1}{8}\int_0^T c^{-1}\sigma_t^{-4}(2[M,M]'_t + \frac{2}{3}c^2[\sigma^2,\sigma^2]'_t)^2dt + o_p(n^{-3\alpha})$$

$$= n^{-3\alpha}\frac{1}{2}\int_0^T c^{-1}(\sigma_t^{-2}[M,M]'_t + \frac{1}{3}c^2[\sigma,\sigma]'_t)^2dt + o_p(n^{-3\alpha}). \quad (B.40)$$

Meanwhile, by similar arguments, it is easy to see that the covariation of the stochastic term with underlying processes is of order $o_p(n^{-\frac{3\alpha}{2}})$. Since these derivations remain valid over intervals $[0,t]$ (replacing $T$ with $t$), we have shown that the right hand side of (B.32) converges as specified in Theorem 10. This follows from Theorem 6 in Mykland and Zhang (2010).

40

Consider separately the case where $\beta > \alpha$ and $\beta < \alpha$. In the former case, the first term inside the brackets in (B.39)-(B.40) dominate, and you replace $c^{-1}n^{-3\alpha}$ by $n^{-4\alpha}\Delta\tau^{-1} = c^{-1}n^{-4\alpha+\beta}$. In the latter case, the second term dominates, and you replace $c^3 n^{-3\alpha}$ by $\Delta\tau^3 = c^3 n^{-3\beta}$.

To complete the proof, we need to show that the remainder terms in (B.32) are of lower order. For simplicity of exposition, we just show that when $\beta = \alpha$, the remainder is of order $o_p(n^{-\frac{3}{2}\alpha})$. For this, note that for $p + q \geq 2$, $\partial^{p+q}f(x,y)/\partial x^p \partial y^q = -2a_p a_q x^{\frac{1}{2}-p} y^{\frac{1}{2}-q}$, where $a_0 = 1$, and $a_k = \frac{1}{2}...\left(\frac{1}{2} - k + 1\right)$ ($k$ factors) for $k \geq 1$. Thus, if we truncate at the forth derivative, by the usual truncation and stopping arguments, the expectation of this term is of order proportional to

$$\Delta\tau E\left\{\sum_{1\leq i;2iM\leq n}\left(\frac{RV_{2i}}{\Delta\tau} - \frac{RV_{2i-1}}{\Delta\tau}\right)^4\right\} \leq c\Delta\tau^{-3}\sum_{1\leq i;2iM\leq n} E([M^{(n)}, M^{(n)}]_{\tau_{2i}} - [M^{(n)}, M^{(n)}]_{\tau_{2i-2}})^2$$
$$= O(\Delta\tau^{-4}(n^{-3\alpha})^2) = O(n^{-2\alpha})$$

from the Burkholder-Davis-Gundy inequality (see Dellacherie and Meyer (1982, Sect 3 of Ch. VII), or Protter (2004, p. 193 & 222)) from (B.6) and (B.34). This leaves the third order term, which is

$$-\frac{1}{3!}\Delta\tau\sum_{1\leq i;2iM\leq n}\sigma_{\tau_{2i-2}}^{-4}\left(\frac{RV_{2i}}{\Delta\tau} - \frac{RV_{2i-1}}{\Delta\tau}\right)^3.$$

The Burkholder-Davis-Gundy bound is in this case of exact order $O(n^{-\frac{3\alpha}{2}})$. A long but tedious calculation shows, however, that the term goes away at this order. (Heuristically, this is because the third cumulant of an asymptotically normal sequence will vanish to first order. An exact derivation involves the Bartlett-type identities for martingales (Mykland (1994)).

### B.3.2 Proof of Theorem 9

To verify Theorem 9 from Theorem 10, we note that in this case, $\alpha = 1/2$. $M_t^{(n)}$ is the usual martingale associated with realised volatility, and stable convergence follows from Jacod and Protter (1998), Barndorff-Nielsen and Shephard (2002), and Mykland and Zhang (2006). The conditions (B.14) is one path to verifying the stable convergence, and otherwise (B.14) follows from the stable convergence via Jacod and Shiryaev (2003, Corollary VI.6.30, p. 385). From the same literature, $f_t^2 = 2T\sigma_t^4$.

### B.3.3 Proof of Theorem 11

We use the representation (B.11). As in the bipower case, note that $f(z,...,z) = f_{x_i}(z,...,z) = 0$, while $f_{x_i x_i}(z,...,z) = z^{-1}(K-1)/K$ and $f_{x_i x_j}(z,...,z) = z^{-1}(-1)/K$ for $i \neq j$. It follows that

$$f(x_1,...,x_K) = \frac{1}{2}z^{-1}\left(\sum_{k=1}^{K}(x_k - z)^2 - \frac{1}{K}(\sum_{k=1}^{K}(x_k - z))^2\right)$$

41

$$+ \text{ terms involving derivatives of order greater than two .}$$

Now set $x_k = RV_{Ki-k+1}/\Delta\tau$ and $z = \sigma^2_{\tau_{K(i-1)}}$. For the same reasons as those discussed in detail in the proof of Theorem 10, we thus obtain that up to error of order $O_p(n^{-2\alpha+\frac{1}{2}\beta} + n^{-\frac{3}{2}\beta})$,

$$RV_n^{(K)} - SMV_n^{(K)} \approx$$
$$\frac{1}{2\Delta\tau} \sum_{1 \leq i; KiM \leq n} \sigma^{-2}_{\tau_{K(i-1)}} \left( \sum_{k=1}^{K} (RV_{Ki-k+1} - \sigma^2_{\tau_{K(i-1)}}\Delta\tau)^2 - \frac{1}{K}(\sum_{k=1}^{K}(RV_{Ki-k+1} - \sigma^2_{\tau_{K(i-1)}}\Delta\tau))^2 \right).$$
(B.41)

As in equation (B.33), we note that

$$\int_{\tau_a}^{\tau_b} \sigma_t^2 \mathrm{d}t = \sigma^2_{\tau_a}\Delta\tau + \int_{\tau_a}^{\tau_b} (\tau_b - t)\mathrm{d}\sigma_t^2.$$

Thus, to sufficient approximation (when aggregated in (B.41))

$$\sum_{k=1}^{K} (RV_{Ki-k+1} - \sigma^2_{\tau_{K(i-1)}}\Delta\tau)^2$$

$$= \sum_{k=1}^{K} \left( \Delta M^{(n)}_{\tau_{Ki-k+1}} + \int_{\tau_{Ki-k}}^{\tau_{Ki-k+1}} (\tau_{Ki-k} - t)\mathrm{d}\sigma_t^2 + \sigma^2_{\tau_{Ki-k}}\Delta\tau - \sigma^2_{\tau_{K(i-1)}}\Delta\tau \right)^2$$

$$\approx \sum_{k=1}^{K} \left( \Delta[M^{(n)}, M^{(n)}]_{\tau_{Ki-k+1}} + \int_{\tau_{Ki-k}}^{\tau_{Ki-k+1}} (\tau_{Ki-k+1} - t)^2 \mathrm{d}[\sigma^2, \sigma^2]_t \right.$$
$$\left. + ([\sigma^2, \sigma^2]_{\tau_{Ki-k}} - [\sigma^2, \sigma^2]_{\tau_{K(i-1)}})\Delta\tau^2 \right)$$

$$\approx ([M^{(n)}, M^{(n)}]_{\tau_{Ki}} - [M^{(n)}, M^{(n)}]_{\tau_{K(i-1)}}) + \frac{1}{3}\Delta\tau^2([\sigma^2, \sigma^2]_{\tau_{Ki}} - [\sigma^2, \sigma^2]_{\tau_{K(i-1)}})$$
$$+ \frac{K-1}{2}\Delta\tau^2([\sigma^2, \sigma^2]_{\tau_{Ki}} - [\sigma^2, \sigma^2]_{\tau_{K(i-1)}})$$

$$\approx ([M^{(n)}, M^{(n)}]_{\tau_{Ki}} - [M^{(n)}, M^{(n)}]_{\tau_{K(i-1)}}) + \left( \frac{K}{2} - \frac{1}{6} \right) \Delta\tau^2([\sigma^2, \sigma^2]_{\tau_{Ki}} - [\sigma^2, \sigma^2]_{\tau_{K(i-1)}}),$$
(B.42)

while

$$-\frac{1}{K} \left( \sum_{k=1}^{K} (RV_{Ki-k+1} - \sigma^2_{\tau_{K(i-1)}}\Delta\tau) \right)^2$$

$$= -\frac{1}{K} \left( M^{(n)}_{\tau_{Ki}} - M^{(n)}_{\tau_{K(i-1)}} + \int_{\tau_{K(i-1)}}^{\tau_{Ki}} (\tau_{Ki} - t)\mathrm{d}\sigma_t^2 \right)^2$$

$$\approx -\frac{1}{K} \left( [M^{(n)}, M^{(n)}]_{\tau_{Ki}} - [M^{(n)}, M^{(n)}]_{\tau_{K(i-1)}} + \int_{\tau_{K(i-1)}}^{\tau_{Ki}} (\tau_{Ki} - t)^2 \mathrm{d}[\sigma^2, \sigma^2]_t \right)$$

$$-\frac{1}{K} \left( [M^{(n)}, M^{(n)}]_{\tau_{Ki}} - [M^{(n)}, M^{(n)}]_{\tau_{K(i-1)}} \right) - \frac{1}{3}K\Delta\tau^2 \left( [\sigma^2, \sigma^2]_{\tau_{Ki}} - [\sigma^2, \sigma^2]_{\tau_{K(i-1)}} \right).$$
(B.43)

Combining (B.42)-(B.43) in (B.41), we obtain, again up to error of order $O_p(n^{-2\alpha+\frac{1}{2}\beta} + n^{-\frac{3}{2}\beta})$,

$$RV_n^{(K)} - SMV_n^{(K)}$$

$$\approx \frac{1}{2\Delta\tau} \sum_{1 \le i; KiM \le n} \sigma_{\tau_{K(i-1)}}^{-2} \frac{K-1}{K} \left([M^{(n)}, M^{(n)}]_{\tau_{Ki}} - [M^{(n)}, M^{(n)}]_{\tau_{K(i-1)}}\right)$$

$$+ \frac{1}{2\Delta\tau} \sum_{1 \le i; KiM \le n} \sigma_{\tau_{K(i-1)}}^{-2} \frac{1}{6}(K-1)\Delta\tau^2 \left([\sigma^2, \sigma^2]_{\tau_{Ki}} - [\sigma^2, \sigma^2]_{\tau_{K(i-1)}}\right)$$

$$\approx \frac{1}{2}\frac{K-1}{K}\Delta\tau^{-1}n^{-2\alpha}\int_0^T \sigma_t^{-2}f_t^2 dt + \frac{K-1}{3}\Delta\tau[\sigma,\sigma]_T. \tag{B.44}$$

Substituting by (B.6) gives the result of the theorem.

### B.3.4   Proof of Theorem 13

Under no noise $f_t^2 = 2T\sigma_t^4$ and $\alpha = 1/2$, so that $b_1 = c^{-1}\frac{K-1}{K}T\int_0^T \sigma_t^2 dt$ and Theorem 12 yields

$$n^{1/2}\left(UMV_M^{(K)} - ECRV^{(K)}\right) = -n^{-\beta+\frac{1}{2}}\frac{K-1}{3}c[\sigma,\sigma]_T + n^{-\frac{1}{2}+\beta}b_1 + O_p(n^{-\frac{1}{2}+\frac{1}{2}\beta} + n^{\frac{1}{2}-\frac{3}{2}\beta}).$$

Meanwhile, from the discussion in Remark 9, $ECRV_n^{(K)} = RV_n + O_p(n^{-\beta}) = \int_0^T \sigma_t^2 dt + O_p(n^{-\beta} + n^{-1/2})$. Since also $n^{1/2}\frac{v}{M} \sim n^{\beta-\frac{1}{2}}\left(\frac{c}{T}\right)^{-1}v$, it follows that

$$n^{1/2}\frac{v}{M}ECRV_n^{(K)} = n^{\beta-\frac{1}{2}}\left(\frac{c}{T}\right)^{-1}v\int_0^T \sigma_t^2 dt + O_p(n^{-1/2} + n^{\beta-1})$$

$$= n^{\beta-\frac{1}{2}}b_1 + O_p(n^{-1/2} + n^{\beta-1}) \tag{B.45}$$

if $v = (K-1)/K$. Thus

$$n^{1/2}\left(UMV_M^{(K)} - (1 + \frac{v}{M})ECRV^{(K)}\right) = -n^{-\beta+\frac{1}{2}}\frac{K-1}{3}c^{-1}[\sigma,\sigma]_T + O_p(n^{-\frac{1}{2}+\frac{1}{2}\beta} + n^{\frac{1}{2}-\frac{3}{2}\beta}).$$

Divide this expression by $(1 + \frac{v}{M})$, and use equation (B.20), to finish the result.

### B.3.5   Proof of Theorem 7

Refer to the proof of Theorem 10 in Section B.3.1. The only change (apart from notational complexity) when passing from $RV_n^{(2)} - SBV_n$ to $ECRV_n^{(2)} - UMV_n^{(2)}$ is that the right hand side of equation B.32 gets replaced by

$$\frac{1}{8}\frac{1}{\Delta\tau}\sum_{2 \le i \le n} \sigma_{\tau_{i-1}}^{-2}(RV_i - RV_{i-1})^2. \tag{B.46}$$

The bias term arising from this is unchanged from Theorem 10, by the same argument as in the proof of Theorem 4. For the variance term, we note that

$$RV_i - RV_{i-1} = \Delta M_{\tau_{i+1}}^{(n)} - \Delta M_{\tau_i}^{(n)} + \int_{\tau_i}^{\tau_{i+1}} \sigma_t^2 dt - \int_{\tau_{i-1}}^{\tau_i} \sigma_t^2 dt = \Delta V_{\tau_{i+1}} - \Delta V_{\tau_i} + \Delta\sigma_{\tau_{i+1}}^2 \Delta\tau$$

where

$$\Delta V_{\tau_{i+1}} = \Delta M_{\tau_{i+1}}^{(n)} + \int_{\tau_i}^{\tau_{i+1}} (t - \tau_i)d\sigma_t^2. \tag{B.47}$$

Extending Mykland and Zhang (2009, Sect 2.2), we can without much loss of generality take both $X$ and $\sigma_t^2$ to be martingales.[4] To sufficient approximation, we can thus write (B.46) as

$$\frac{1}{8}\frac{1}{\Delta\tau}\sum_{2\leq i\leq n}\sigma_{\tau_{i-1}}^{-2}\left(2(\Delta V_{\tau_{i+1}})^2 - 2\Delta V_{\tau_{i+1}}\Delta V_{\tau_i} + (\Delta\sigma_{\tau_{i+1}}^2)^2\Delta\tau^2 + 2\Delta V_{\tau_{i+1}}\Delta\sigma_{\tau_{i+1}}^2\Delta\tau - 2\Delta V_{\tau_i}\Delta\sigma_{\tau_{i+1}}^2\Delta\tau\right).$$
(B.48)

To get rid of the terms arising from $\Delta\sigma_{\tau_{i+1}}^2\Delta\tau$, note that taking quadratic variation,

$$(\Delta\tau)^{-1}\sum_{2\leq i\leq n}\{(\Delta\sigma_{\tau_{i+1}}^2)^2 - \Delta[\sigma^2,\sigma^2])\Delta\tau^2\} = O_p(\tau^{-3/2}).$$

Similarly,

$$(\Delta\tau)^{-1}\sum_{2\leq i\leq n}\{\Delta V_{\tau_{i+1}}\Delta\sigma_{\tau_{i+1}}^2 - \Delta[V,\sigma^2]_{\tau_{i+1}}\}\Delta\tau$$

and $(\Delta\tau)^{-1}\sum_{2\leq i\leq n}\Delta V_{\tau_i}\Delta\sigma_{\tau_{i+1}}^2\Delta\tau$ are of order $O_p(n^{-\alpha}\Delta\tau^{-1/2}) + O_p(\Delta\tau^{-3/2})$. Thus (B.46) is

$$\frac{1}{4}\frac{1}{\Delta\tau}\sum_{2\leq i\leq n}\sigma_{\tau_{i-1}}^{-2}\left((\Delta V_{\tau_{i+1}})^2 - \Delta[V,V]_{\tau_{i+1}} - \Delta V_{\tau_{i+1}}\Delta V_{\tau_i}\right)$$
$$+\text{ stochastic drift }+ O_p(n^{-\alpha}\Delta\tau^{-1/2}) + O_p(\Delta\tau^{-3/2}).$$
(B.49)

The quadratic variation is thus, up to relative error of $1 + o_p(1)$,

$$\frac{3}{16}\frac{1}{\Delta\tau^2}\sum_{2\leq i\leq n}\sigma_{\tau_{i-1}}^{-4}(\Delta[V,V]_{\tau_{i+1}})^2$$
$$\sim \frac{3}{16}\frac{1}{\Delta\tau^2}\sum_{2\leq i\leq n}\sigma_{\tau_{i-1}}^{-4}(n^{-2\alpha}\Delta[M,M]_{\tau_{i+1}} + \frac{1}{3}\Delta\tau^3[\sigma^2,\sigma^2]_{\tau_i}')^2$$
$$\sim \frac{3}{16}\frac{1}{\sum_{2\leq i\leq n}}\sigma_{\tau_{i-1}}^{-4}(n^{-2\alpha}[M,M]_{\tau_i}' + \frac{1}{3}\Delta\tau^2[\sigma^2,\sigma^2]_{\tau_i}')^2$$
$$\sim \frac{3}{16}\int_0^T\sigma_t^{-4}(n^{-2\alpha}\Delta\tau^{-1/2}[M,M]_t' + \frac{1}{3}\Delta\tau^{3/2}[\sigma^2,\sigma^2]_t')^2dt$$
$$\sim 3\int_0^T(\frac{1}{4}c^{-\frac{1}{2}}n^{-2\alpha+\frac{1}{2}\beta}\sigma^{-2}[M,M]_t' + \frac{1}{3}c^{\frac{3}{2}}n^{-\frac{3}{2}\beta}[\sigma,\sigma]_t')^2dt,$$
(B.50)

since $\Delta\tau \sim cn^{-\beta}$. The result then follows in the same way as in Appendix B.3.1.

---

[4]Tedious exact calculation will give the same result without the regularity conditions.