On the estimation of integrated volatility in the presence of jumps and microstructure noise

Christian Brownlees[*] Eulalia Nualart[*] Yucheng Sun[†]

September 11, 2019

Abstract

This paper is concerned with the problem of the estimation of the integrated volatility of log-prices based on high frequency data when both price jumps and market microstructure noise are present. We begin by providing a survey of the leading estimators introduced in the literature to tackle volatility estimation in this setting. We then introduce novel integrated volatility estimators based on a truncation technique and establish their properties. Finally, we carry out a simulation study to compare the performance of the different estimation techniques.

Keywords: integrated volatility, two-scales realized volatility estimator, realized kernel estimator, jumps, market microstructure noise

[*] Department of Economics and Business, Universitat Pompeu Fabra and Barcelona GSE,
[†] International School of Economics and Management, Capital University of Economics and Business,
e-mail: christian.brownlees@upf.edu, eulalia@nualart.es, sunyucheng@cueb.edu.cn
We have benefited from comments by the two anonymous referees, Nour Meddahi, George Tauchen and conference participants to the "Financial Econometrics Conference" of the Toulouse School of Economics (Toulouse, 13 and 14 of May 2016). Christian Brownlees and Eulalia Nualart acknowledge financial support from Spanish Ministry of Science and Technology (Grant MTM2015-67304-P) and from Spanish Ministry of Economy and Competitiveness through the Severo Ochoa Programme for Centres of Excellence in R&D (SEV-2011-0075). Eulalia Nualart's research is supported in part by the European Union programme FP7-PEOPLE-2012-CIG under grant agreement 333938. Yucheng Sun's research is supported by The National Social Science Fund of China No. 18VFH005.

1 Introduction

The volatility of asset prices is a fundamental ingredient for asset pricing, risk management and portfolio allocation. Over the last decade, the financial econometrics literature has developed a new generation of estimators of the daily volatility of asset prices based on intra-daily data typically referred to as realized volatility estimators. The classic realized volatility (RV) estimator of Andersen and Bollerslev (1998) for example is defined as the sum of the squares of high-frequency intra-daily returns. Under appropriate assumptions, this estimator provides a consistent estimate of the quadratic variation of asset prices when prices follow a continuous stochastic model and are directly observed (see e.g. Barndorff-Nielsen and Shephard 2002 and Andersen, Bollerslev, Diebold, and Labys 2003).

It is well acknowledged in the literature that asset prices exhibit discontinuities in their sample paths and are also contaminated by market microstructure noise (see e.g. Barndorff-Nielsen and Shephard 2006 and Hansen and Lunde 2006). Allowing for a jump component makes it more challenging to estimate the quadratic variation of the continuous part, which is commonly modeled as the integral of the spot volatility and referred to as the integrated volatility (IV). However, it is of strong economic interest to disentangle the integrated volatility from the whole quadratic variation. For example, Andersen, Bollerslev, and Diebold (2007) and Corsi, Pirino, and Reno (2010) showed that it is more accurate to predict future volatility with the integrated volatility and the jump variation than with only the whole quadratic variation. The results in Zhang, Zhou, and Zhu (2009) indicated that both integrated volatilities and the jump variations of equity prices have substantial effects on the spreads of the credit default swaps. On the other hand, the market microstructure noise also poses challenges to the estimation of the quadratic variations. In fact, in the presence of the noise the standard RV estimator is inconsistent as the sampling frequency of the data increases (see e.g. Hansen and Lunde 2006 and Bandi and Russell 2008).

These two important stylized facts of asset prices have motivated the development of a number of estimators which are consistent when there are jumps, noise or both. This paper aims to review such estimators, which are typically derived by the combination of the mechanism of the RV estimator and other techniques such as the truncation and the pre-averaging methods.

The first realized volatility estimators that appeared in the literature that are robust to price jumps are the power and multipower variation estimators. In general, the construction of this type of estimators relies on products of certain powers of adjacent intradaily-returns, since it is possible to make the effects of jumps on such products diminish as sampling frequency increases, see e.g. Barndorff-Nielsen and Shephard (2004) and Corsi, Pirino, and Reno (2010) for the case of finite activity jumps, and Barndorff-Nielsen, Shephard, and Winkel (2006b), Woerner (2006), Jacod (2008), and Jacod and Todorov (2014) for infinite activity jumps.

Mancini (2008, 2009) introduced a truncation technique that consists of excluding the intra-daily returns larger than a threshold (in absolute value) from the estimation of the quadratic variation, as these are likely to contain a realization of a jump. This leads to the truncated RV estimator which deals with both finite and infinite activity jumps. Andersen, Dobrev, and Schaumburg (2012) applied the nearest neighbor truncation method to the RV estimator by truncating the squared intra-daily returns that are larger than their adjacent ones. This technique can also make the estimator immune to jumps by removing the intervals where there are jumps. However, the truncated IV estimators in Mancini (2008, 2009) and Andersen, Dobrev, and Schaumburg (2012) are not consistent in the presence of market microstructure noise, since they are based on the RV estimator which is not robust to the noise.

The first realized volatility estimator introduced in the literature that is consistent when constructed with noisy data is the two-scales realized volatility (Zhang, Mykland, and Aït-Sahalia 2005). This estimator takes the average of many RV estimators to partially eliminate the effects of the noise, and the remaining noise effects are debiased by an estimator on the noise variance. An extension of the two-scales estimator is the multi-scales estimator proposed by Zhang (2006), which converges to the quadratic variation in probability at a faster rate than the two-scales estimator. Other estimators include e.g. the realized kernels (Barndorff-Nielsen, Hansen, Lunde, and Shephard 2008, 2011, Varneskov 2016, 2017) and realized pre-averaging estimator (Jacod, Li, Mykland, Podolskij, and Vetter 2009, Jacod, Podolskij, and Vetter 2010), using respectively kernel functions and pre-averaged data in order to smooth away the noise impact.

Finally, contributions that propose estimators that are consistent in the presence of both finite activity jumps and noise include, among others, Podolskij and Vetter (2009a,b), Christensen, Oomen, and Podolskij (2014), Christensen, Hounyo, and Podolskij (2018), Fan and Wang (2007), Barunik and Vacha (2015), Christensen, Oomen, and Podolskij (2010), Jing, Liu, and Kong (2014) and Bibinger and Winkelmann (2015). The estimators considered in Podolskij and Vetter (2009a,b) and Christensen, Oomen, and Podolskij (2014) were derived by the combination of pre-averaged data and bipower variations which belong to the family of multipower variations, and Christensen, Hounyo, and Podolskij (2018) further imposed the truncation indicators on the pre-averaged bipower variations. Fan and Wang (2007) and Barunik and Vacha (2015) employed the wavelet technique to detect jumps. Then they adjusted the data to remove jump effects, and used the adjusted data to construct the two-scales and multi-scales estimators. The estimator proposed in Christensen, Oomen, and Podolskij (2010) is based on intra-daily quantile ranges which are immune to the extreme intra-daily return values due to jumps. The truncated pre-averaging estimator considered in Jing, Liu, and Kong (2014) is derived by truncating the local average returns with large absolute values from the pre-averaging estimator, since such returns are likely to be affected by jumps. The truncated spectral estimator proposed in Bibinger and Winkelmann (2015) employs similar truncation methodology as in Mancini (2008, 2009) and Jing, Liu, and Kong (2014), but it relies heavily on spectral analysis, and in this sense, its mechanism is fundamentally different from the estimators we review in this paper. The medium blocked realized kernels in Varneskov (2017) are constructed by associating the realized kernels with the nearest neighbor truncation method, so this estimator can be robust to jumps and the noise at the same time.

Besides reviewing existing estimators, this paper introduces two novel truncated estimators, which are the truncated two-scales realized volatility and truncated flat-top realized kernel estimators. Like the truncated pre-averaging estimator, the proposed estimators here are also obtained by truncating noise-robust volatility estimators, so the truncated two-scales and kernel estimators are consistent in the presence of both jumps and the noise. Moreover, the simulation results presented in this paper suggest that the truncated estimators in general have good estimation accuracy when there are noise and jumps, which is in line with the results shown in Jing, Liu, and Kong (2014). We observe that the methodology of the wavelet-based estimators is essentially similar to the truncated estimators proposed here. In fact, the truncated two-scales estimator has the same asymptotic distribution as the wavelet-based two-scales estimator proposed in Fan and Wang (2007). The truncated estimators however are in general easier to compute than the wavelet-based ones. Moreover, the truncated pre-averaging and truncated two-scales estimators are robust to infinite activity jumps, while to the best of our knowledge, the consistency of the wavelet-based realized volatility estimators has not been theoretically justified in the presence of infinite activity jumps.

Other related contributions are Jacod and Protter (2012) and Aït-Sahalia and Jacod (2014) that proposed truncated estimators on volatilty functionals, including the truncated bipower and multipower estimators. These estimators however do not consistently estimate the quadratic variation in the presence of the noise. Besides volatility estimation, truncation techniques are widely used, for instance, to explore the relationship between jumps and spot volatility (Jacod and Todorov, 2010) and to estimate the covariation between asset returns and changes in volatility (Aït-Sahalia, Fan, Laeven, Wang, and Yang, 2017). Another strand of the literature relevant to this paper is the one that concerns testing for the presence of price jumps and cojumps (with or without noise), which includes, e.g., Jacod and Todorov (2009), Jacod, Podolskij, and Vetter (2010), Aït-Sahalia, Jacod, and Li (2012), Lee and Mykland (2012) and Li, Todorov, Tauchen, and Lin (2018). These papers have inspired the jump detection indicator based on local averages used in this work.

In the remaining of this paper, section 2 introduces basic notations. Section 3 reviews some important realized volatility estimators. Section 4 proposes the truncated two-scales and kernel estimators. Section 5 performs simulation exercises to evaluate the efficiency of some estimators reviewed or proposed. Section 6 concludes.

2 Model setup

We denote by $(y_t, t \in [0,1])$ the efficient log-price process of an asset, where 0 typically represents the opening of the trading day and 1 the closing. The process starts at an initial value $y_0 \in \mathbb{R}$ and its dynamics are given by

$$dy_t = a_t dt + \sigma_t dB_t + dJ_t, \qquad t \in]0,1] \tag{1}$$

where B is a standard Brownian motion and J is a pure jump process, both defined on a filtered probability space $(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathcal{F}, \mathrm{P})$. The coefficients a and σ are adapted and locally bounded. For some estimators, we will assume that σ follows an Itô-semimartingale process, that is:

$$\sigma_t = \sigma_0 + \int_0^t b_s ds + \int_0^t v_s dB_s + \int_0^t n_s dW_s, \tag{2}$$

where b, v and n are adapted càdlàg processes, b is predictable and locally bounded, and W is another standard Brownian motion independent of B. Concerning the jump part J, for some results we will assume it has finite activity (FA), and so it can be expressed as $J_t = \sum_{i=1}^{N_t} Y_i$, where $N_t < \infty$ for finite t. Or more generally, J is allowed to have infinite activity (IA), and can be expressed as

$$J_t = \int_0^t \int_{|x|\leq 1} x(\mu - \nu)(ds, dx) + \int_0^t \int_{|x|>1} x\mu(ds, dx),$$

where μ is the jump measure and ν is its predictable compensator.

Given m as the number of price observations in $]0,1]$, and t_i as the time point when we observe the $i-$th price, many works (e.g. Barndorff-Nielsen, Hansen, Lunde, and Shephard 2008, Christensen, Oomen, and Podolskij 2010, Varneskov 2016), including this paper unless stated otherwise, assume that the prices are equally spaced, which means $t_i = \frac{i}{m}$ for $i = 1, \ldots, m$ for the convenience of technical analysis. This assumption may be replaced by the weaker one that the sampling times are independent of y and the noise, and $\max_i(t_i - t_{i-1}) = O\left(\frac{1}{m}\right)$. This is because under this weaker assumption, typically the commonly seen estimators still maintain the same desired convergence rates towards IV, including, among others, the RV, the realized kernels, the pre-averaging, the two-scales and the multi-scales realized volatilities (see e.g. Wang and Zou 2010, Barndorff-Nielsen, Hansen, Lunde, and Shephard 2011, Fan, Li, and Yu 2012, Kim, Wang, and Zou 2016). Endogenous sampling, which can be subject to the tick-by-tick data (Robert and Rosenbaum 2010, 2012) has also been studied in the literature. For example, Li, Mykland, Renault, and Zhang (2014) showed that the estimation error by the RV estimator without jumps and the noise is still $O_{\mathrm{P}}(m^{-1/2})$, when the sampling scheme is possibly endogenous with $\max_i (t_i - t_{i-1}) = o_{\mathrm{P}}(m^{-2/3-\epsilon})$ for any fixed $\epsilon > 0$. Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008) also pointed out the robustness of the realized kernels with respect to some sampling schemes that can be possibly endogenous.

The efficient log-price is contaminated by market microstructure noise. That is, rather than the efficient price y_t the econometrician observes at discrete times its contaminated counterpart x_t. Specifically, we assume x_{t_i} is generated as

$$x_{t_i} = y_{t_i} + u_{t_i}, \quad i = 1, ..., m,$$

where u_{t_i} denotes the microstructure noise associated to the i-th observation with expectation 0 and variance ω^2. For notational simplicity we will use x_i, y_i and u_i to respectively denote x_{t_i}, y_{t_i} and u_{t_i}.

A common assumption in the literature is that the u_i's are mutually independent (e.g. Christensen, Oomen, and Podolskij 2014; Christensen, Hounyo, and Podolskij 2018). However, the empirical studies in Hansen and Lunde (2006) suggested that this assumption is only reasonable when the sampling frequency is moderately high, e.g. 1 minute, but not applicable to the ultra-high-frequency data such as the tick-by-tick data, in which case the noise can be time-dependent. In addition, most works assumed that u is independent of the efficient price process y, but Hansen and Lunde (2006) also pointed out that this assumption can be violated when the sampling frequency is high. Some relaxations

Table 1: Constants

Constant	Definition
K	$c_1 m^{2/3}$
H	$c_2 m^{1/2}$
K_0	$c_3 m^{1/2}$
$\overline{K}$	$c_4 m^{1/2}$
M_0	$c_5 m^{1/2}$
K_1	$c_6 m^{1/2}$
K_2	$c_7 m^{1/2}$
L	$c_8 m^a$ for $a \in (0, 3/4)$
K_3	$c_9 m^b$ for $b \in (0, 1)$

This table defines a list of constants used in the definition of the estimators. Here $c_1, c_2, \ldots, c_9$ denote positive constants that are fixed throughout the paper.

can be found in, e.g., Jacod, Li, Mykland, Podolskij, and Vetter (2009), Jing, Liu, and Kong (2014), Barndorff-Nielsen, Hansen, Lunde, and Shephard (2011), and Varneskov (2016, 2017). The first two papers assume the properties of u (zero mean and moment bound condition) conditional on y, while the others assume an endogenous component in u in the form

$$e_{t_i} = m^{1/2} \sum_{h=-\infty}^{\infty} \phi_{h,i} \left(\widetilde{W}_{t_{i-h}} - \widetilde{W}_{t_{i-h-1}} \right), \tag{3}$$

where $\widetilde{W}$ is a standard Brownian motion possibly correlated with B and the coefficients ϕ satisfy certain conditions. Finally, we will rely on the following assumption in deriving the theoretical properties of the truncated two-scales and realized kernel estimators proposed in this paper:

Assumption 1. *For any fixed positive interger n, $E(u_i^n)$ is uniformly bounded across $i \in \{1, \ldots, m\}$.*

This assumption is stronger than the common one that $E(u_i^n)$ is bounded for a fixed n, like $n = 4$ or 8, but it is still standard in the realized volatility literature. For example, Fan *et al.* (2012) and Tao, Wang, and Zhou (2013) assumed u_i as subgaussian, which makes this assumption satisfied.

An additional word on notation is in order. In Table 1 we define a number of bandwidth parameters that are required for the construction of the estimators introduced in the next section.

3 Existing integrated volatility estimators

Various estimators have been introduced in the literature to estimate the integrated volatility of an asset, which is defined as

$$\mathrm{IV} = \int_0^1 \sigma_t^2 dt \ .$$

Andersen and Bollerslev (1998) used the sum of squared intra-daily non-overlapping returns to approximate IV in the absence of jumps and noise. This leads to the classical RV estimator $\widehat{\sigma}^2_{\mathsf{RV}} = \sum_{i=1}^m (\triangle_i x)^2$, where $\triangle_i x = x_i - x_{i-1}$. Since IV is the quadratic variation of the continuous part in y, Barndorff-Nielsen and Shephard (2002) showed that when the price observations are not affected by jumps or the noise, $m^{1/2}\left(\widehat{\sigma}^2_{\mathsf{RV}} - \mathrm{IV}\right)$ is asymptotically mixed Gaussian, that is

$$m^{1/2}\left(\widehat{\sigma}^2_{\mathsf{RV}} - \mathrm{IV}\right) \xrightarrow{\mathcal{L}} \mathrm{MN}\left(0, 2\int_0^1 \sigma_s^4 ds\right),$$

as $m \to \infty$, where the convergence in law is stable. As mentioned in the introduction, this classical estimator has been modified in different directions in order to obtain robust estimators in the presence of jumps or noise. We next proceed with a review of the most important ones.

3.1 Realized range-based estimators

Christensen and Podolskij (2007) developed the realized range-based (RRV) estimator, inspired by the works of Feller (1951) and Parkinson (1980). It is defined as

$$\widehat{\sigma}^2_{\mathsf{RRV}} = \frac{1}{\upsilon_{m_1}} \sum_{i=1}^{m_2} \sup_{0 \le s,t \le m_1} \left(x_{i-1+s} - x_{i-1+t}\right)^2,$$

where m_1 and m_2 are positive integers with $m_1 m_2 = m$, and $\upsilon_{m_1} = E\Bigg(\max_{0\le s,t\le m_1}\Big(B_{\frac{t}{m_1}} - B_{\frac{s}{m_1}}\Big)^2\Bigg)$ (with s, t integers). Christensen and Podolskij (2007) showed that in the absence of jumps and noise, $\widehat{\sigma}^2_{\mathsf{RRV}}$ converges to $\int_0^1 \sigma_t^2 dt$ in probability as $m_2 \to \infty$. Moreover, they derived the asymptotic mixed Gaussian distribution for $m_2^{1/2}\left(\widehat{\sigma}^2_{\mathsf{RRV}} - \int_0^1 \sigma_t^2 dt\right)$ when σ satisfies (2). This result implies that $\widehat{\sigma}^2_{\mathsf{RRV}}$ achieves the optimal convergence rate $m^{-1/2}$ when m_1 is $O(1)$ as $m \to \infty$. Then Martens and van Dijk (2007) proposed several bias-corrected versions of $\widehat{\sigma}^2_{\mathsf{RRV}}$ in order to deal with the noise, but they did not prove the consistency of the modified estimators. Christensen, Podolskij, and Vetter (2009) showed the consistency of a type of modified RRV estimators, but the proof of the convergence relies on the assumed specific distribution for the noise terms, which is substantially stronger than the commonly seen i.i.d. assumption.

3.2 Multipower variation estimators

The realized multipower variation (MPV) estimators are defined as

$$\widehat{\sigma}_{\mathsf{MPV}}(r_1, \ldots, r_N) = m^{1-\frac{\sum_{i=1}^N r_i}{2}} \sum_{j=N}^{m} |\triangle_{j-N+1} x|^{r_1} \ldots |\triangle_j x|^{r_N}. \tag{4}$$

where $r_1, \ldots, r_N$ are constants in $(0, 1]$. The properties of this type of estimators were studied e.g. in Barndorff-Nielsen and Shephard (2004), Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard (2006a) and Barndorff-Nielsen, Shephard, and Winkel (2006b). In the absence of jumps and noise, Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard (2006a) derived the asymptotic mixed Gaussian distribution for $m^{1/2}\Big(\widehat{\sigma}_{\mathsf{MPV}} - c\int_0^1 |\sigma_s|^{\sum_{i=1}^N r_i} ds\Big)$ in the sense of stable convergence in law, where c is some positive constant depending only on $\{r_1, \ldots, r_N\}$. Thus, by setting $\sum_{i=1}^N r_i = 2$, we can make $\frac{1}{c}\widehat{\sigma}_{\mathsf{MPV}}$ a consistent estimator of IV. Barndorff-Nielsen, Shephard, and Winkel (2006b) discussed the asymptotic property of $\widehat{\sigma}_{\mathsf{MPV}}$ in the presence of FA and IA jumps. They showed that when there are FA jumps, by properly setting the values of $r_1, \ldots, r_N$, the asymptotic distribution of the multipower estimator remains the same. In the presence of IA jumps, the degree of jump activity must satisfy certain technical condition in order to yield the same asymptotic distribution for $\widehat{\sigma}_{\mathsf{MPV}}$.

When $N = 2$, $\widehat{\sigma}_{\mathsf{MPV}}$ becomes the bipower variation $\widehat{\sigma}_{\mathsf{BPV}}(r_1, r_2)$ which is an important subgroup in the family of multipower variation estimators. In particular, we note that the CLT does not in general hold for $\widehat{\sigma}_{\mathsf{BPV}}(r_1, r_2)$ when there are jumps and $r_1 + r_2 = 2$; as Barndorff-Nielsen, Shephard, and Winkel (2006b) pointed out, the CLT requires $\max\{r_1, r_2\} < 1$ in the presence of jumps. Based on the structure of $\widehat{\sigma}_{\mathsf{BPV}}$, Podolskij and Vetter (2009a,b) derived IV estimators robust to both jumps and the noise, which are the modulated bipower variations that will be reviewed later.

3.3 Two-scales realized volatility estimator

The two-scales realized volatility (TSRV) estimator is defined as

$$\widehat{\sigma}^2_{\mathsf{TSRV}} = \frac{1}{K} \sum_{i=K}^{m} (x_i - x_{i-K})^2 - \frac{m-K+1}{mK} \sum_{i=1}^{m} (\triangle_i x)^2 . \tag{5}$$

This estimator was proposed by Zhang, Mykland, and Aït-Sahalia (2005) where they showed that it consistently estimates the IV in the presence of microstructure noise. The first component on the right-side of (5) can be regarded as the average of K realized volatility estimators that are in the same form as the definition of $\widehat{\sigma}^2_{\mathsf{RV}}$, except that $x_i - x_{i-1}$ is replaced with $x_i - x_{i-K}$. Then the first component converges to IV in probability in the absence of the noise, and the second component counteracts the impact of the noise on the first component. Under the assumption that σ_t is continuous, and u_i's are i.i.d., independent of y, Zhang, Mykland, and Aït-Sahalia (2005) showed that $m^{1/6} \left(\widehat{\sigma}^2_{\mathsf{TSRV}} - \mathrm{IV}\right)$ converges stably in law to some mixed Gaussian distribution. When σ_t is not continuous, it can be checked that $\widehat{\sigma}^2_{\mathsf{TSRV}} - \mathrm{IV}$ is still $O_{\mathrm{P}}(m^{-1/6})$ (see e.g. Theorem 1 in Fan, Li, and Yu 2012).

Later we will work with the following modified version of the TSRV estimator

$$\widehat{\sigma}^2_{\mathsf{TS}} = \frac{1}{K} \sum_{i=K}^{m} (x_i - x_{i-K})^2 - \frac{1}{K} \sum_{i=K}^{m} (\triangle_i x)^2 . \tag{6}$$

It can be seen that the difference between $\widehat{\sigma}^2_{\mathsf{TSRV}}$ and $\widehat{\sigma}^2_{\mathsf{TS}}$ lies in the second component of their expressions. Specifically, the second component on the right-side of (6) is obtained by removing the first $K-1$ terms from the sum $\sum_{i=1}^{m}\left(\triangle_i x\right)^2$, and adjusting the coefficient $\frac{m-K+1}{mK}$ accordingly to maintain the balance. As $K \ll m$, it is easy to check that $\widehat{\sigma}^2_{\mathsf{TS}}$ has the same asymptotic property as the one derived in Zhang, Mykland, and Aït-Sahalia (2005) for $\widehat{\sigma}^2_{\mathsf{TSRV}}$.

3.4 Realized kernel estimator

The realized kernel (RK) estimator is defined as

$$\widehat{\sigma}^2_{\mathsf{RK}} = \gamma_0(X) + \sum_{h=1}^{H} k\left(\frac{h-1}{H}\right)\left(\gamma_h(X) + \gamma_{-h}(-X)\right), \tag{7}$$

where for each $h \in \{-H, \dots, H\}$, $\gamma_h(X) = \sum_{i=1}^{m} \triangle_i x \triangle_{i-h} x$, and the kernel function k satisfies (i) $k(0) = 1$; (ii) $k(x)$ is twice differentiable with bounded derivatives on $[0,1]$; (iii) $k(1) = k'(0) = k'(1) = 0$. This estimator was introduced in Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008) and uses the kernel function to smooth away the impact of the noise, so it consistently estimates IV in the presence of the u_i's that are mutually independent and also independent of y. Notice that the construction of $\widehat{\sigma}^2_{\mathsf{RK}}$ requires some data outside the period $[0,1]$. This issue can be addressed by changing the start and end points of the process X, which, as pointed out by Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008), does not cause additional technical problems, but can ease the exposition. Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008) showed that without jumps, $m^{1/4}\left(\widehat{\sigma}^2_{\mathsf{RK}} - \mathrm{IV}\right)$ converges stably in law to some mixed Gaussian distribution.

Barndorff-Nielsen, Hansen, Lunde, and Shephard (2011) made some modification on $\widehat{\sigma}^2_{\mathsf{RK}}$, and the modified kernel estimator allows some dependency in the noise series, but it converges to IV in probability at a slower rate, $m^{-1/5}$. In order to deal with dependent noise terms while keeping the convergence rate $m^{-1/4}$, Varneskov (2016, 2017) employed flat-top kernel functions in constructing $\widehat{\sigma}^2_{\mathsf{RK}}$, and the corresponding flat-top realized kernel

(FTRK) estimator is defined as

$$\widehat{\sigma}^2_{\mathsf{FRK}} = \gamma_0(X) + \sum_{h=1}^{H(1+C)} \overline{k}\left(\frac{h}{H}\right)(\gamma_h(X) + \gamma_{-h}(-X)), \text{ where} \tag{8}$$

$$\overline{k}(x) = \mathbf{1}_{\{|x|\le C\}} + k(|x| - C)\mathbf{1}_{\{|x|>C\}},$$

and $C = H^{-c}$ for some $c \in (0,1)$. Compared to the kernel function in Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008, 2011), a property of the flat-top kernel function is that $k(x) = 1$ in a neighbourhood of zero. Varneskov (2017) demonstrated that when there are no jumps, $m^{1/4}\left(\widehat{\sigma}^2_{\mathsf{FRK}} - \mathrm{IV}\right)$ is asymptotically mixed Gaussian in the sense of stable convergence in law. Moreover, the CLT is derived in the presence of a noise that can be decomposed as

$$u_i = e_i + v_i, \tag{9}$$

where e_i is the endogenous noise that can be described by (3), and v_i is independent of y with the $\alpha-$mixing property and a polynomial decaying mixing coefficient. Additionally, it can be checked that the order of the difference between $\widehat{\sigma}^2_{\mathsf{FRK}}$ and the IV is $m^{-1/4}$ when the u_i's are independent of y and $M-$dependent with a fixed integer $M > 0$ (see the proof of Theorem 3). In this case the assumption on the dynamics of σ_t like (2) is not needed. Recall that M-dependent noise means that u_i is independent of u_j if $|i - j| > M$.

3.5 Pre-averaging estimator

Jacod, Li, Mykland, Podolskij, and Vetter (2009) used the pre-averaging approach to construct a noise-robust IV estimator. The idea of the pre-averaging (PA) estimator is to first compute the locally weighted averages of price observations, and then use the averaged observations to approximate the IV, since the impact of noise on the averaged observations is much smaller than on the original ones. The definition of the PA estimator depends on q weight function g satisfying: (i) g is continuous, piecewise C^1 with a piecewise Lipschitz derivative g'; (ii) $g(0) = g(1) = 0$, $\int_0^1 g(s)^2 ds > 0$. For a generic i, a locally weighted average return used in the construction of the PA estimator is defined as

$$\triangle_{i,K_0}X(g) = \sum_{j=1}^{K_0-1} g\left(\frac{j}{K_0}\right)\triangle_{i+j}x. \tag{10}$$

Then taking the sum of the rescaled squares of $\triangle_{i,K_0}X(g)$ across i leads to the first component of the PA estimator, and the second component eliminates the bias caused by the noise. Specifically, the PA estimator is defined as

$$\widehat{\sigma}^2_{\mathsf{PA}} = \frac{1}{K_0\varphi_2}\sum_{i=0}^{m-K_0+1}\left(\triangle_{i,K_0}X(g)\right)^2 - \frac{\varphi_1}{2K_0^2\varphi_2}\sum_{i=1}^{m}\left(\triangle_i x\right)^2, \tag{11}$$

where $\varphi_1 = \int_0^1 g'(u)^2du$ and $\varphi_2 = \int_0^1 g(u)^2du$. Furthermore, Jacod, Li, Mykland, Podolskij, and Vetter (2009) showed that in the absence of jumps, $m^{1/4}\left(\widehat{\sigma}_{\mathsf{PA}} - \mathrm{IV}\right)$ is asymptotically mixed Gaussian in the sense of stable convergence in law, assuming the zero mean and bounded moment properties of u conditional on x. Jing, Liu, and Kong (2014) and Koike (2016) adopted similar assumptions on u when showing the properties of the truncated pre-averaging estimators that will be reviewed later in the text.

3.6 Truncated realized volatility

Mancini (2008, 2009) and Cont and Mancini (2011) applied the truncation technique to $\widehat{\sigma}^2_{\mathsf{RV}}$, and the truncated $\widehat{\sigma}^2_{\mathsf{RV}}$, denoted as $\widehat{\sigma}^2_{\mathsf{TRV}}$, consistently estimates the IV in the presence of jumps. In the absence of the noise, when x is continuous on $]t_{i-1}, t_i]$, the increment of x over this interval is small in absolute value for large m, otherwise the value of the increment is close to the jump size. Then the idea of truncation is to detect the presence of jumps by checking whether $|\triangle_i x|$ is larger than some threshold, and if it is larger, $\triangle_i x$ is removed from the realized volatility estimator. Thus $\widehat{\sigma}^2_{\mathsf{TRV}}$ is defined as

$$\widehat{\sigma}^2_{\mathsf{TRV}} = \sum_{i=1}^{m}\left(\triangle_i x\right)^2 \mathbf{1}_{\{|\triangle_i x|\le T(m)\}},$$

where the threshold $T(m)$ converges to zero at some rate slower than the convergence rate of the continuous part in $\triangle_i x$ as $m \to \infty$.

Mancini (2009) pointed out that in the presence of FA jumps, $\widehat{\sigma}^2_{\mathsf{TRV}}$ has the same asymptotic distribution as $\widehat{\sigma}^2_{\mathsf{RV}}$ constructed in the setting where there are no jumps, since truncation removes the effects of jumps, and the magnitude of information loss is $o_{\mathrm{P}}(m^{-1/2})$. Assuming jumps follow a Lévy process, Mancini (2009) also showed the consistency of $\widehat{\sigma}^2_{\mathsf{TRV}}$ while allowing jumps to have infinite activity. Cont and Mancini (2011) further demonstrated that under the condition where jumps have finite variation, the IA Lévy jump process does not affect the asymptotic distribution of $\widehat{\sigma}^2_{\mathsf{TRV}}$, which is the same as that of $\widehat{\sigma}^2_{\mathsf{RV}}$ in the absence of jumps.

We have seen that when estimating the IV, the mechanisms of the BPV and TRV have the potential to deal with jumps, while the ideas of the TSRV, RK and PA estimators lead to noise-robust estimators. Therefore, it is natural to think of combining the techniques of these estimators in hopes of deriving new estimators robust to both jumps and the noise. Accordingly, the bipower-type estimators considered in Podolskij and Vetter (2009a,b) and Christensen, Oomen, and Podolskij (2014) were obtained by introducing the pre-averaging procedure in the bipower variations, and truncating the PA gives rise to the truncated pre-averaging estimator proposed in Jing, Liu, and Kong (2014). Besides the aforementioned estimators, we will also introduce the truncated TSRV and truncated FTRK estimators. Moreover, another estimator reviewed later is the quantile-based realized volatility (QRV), which also consistently estimates the IV in the presence of FA jumps and the noise.

3.7 Modulated bipower variation

The earliest modulated bipower variation (MBV) estimator introduced by Podolskij and Vetter (2009a,b) is similar to the BPV, except that $\triangle x$'s are replaced with local (weighted) average returns. Taking the pre-averaged returns makes the MBV robust to the noise. Generally, this type of estimators are analyzed with the u_i's that are mutually independent and also independent of y, and equation (2) is a necessary condition to derive the corresponding CLT results (Podolskij and Vetter 2009a,b, Christensen, Oomen, and Podolskij 2014; Christensen, Hounyo, and Podolskij 2018). Specifically, the MBV in Podolskij and

Vetter (2009b) is given by

$$\widehat{\sigma}_{\mathsf{MBV}}(r,\ell) = m^{\frac{r+\ell}{4}-\frac{1}{2}} \sum_{i=1}^{M_0} \left|\overline{X}_i\right|^r \left|\overline{X}_{i+1}\right|^\ell, \text{ where } \overline{X}_i = \frac{1}{\frac{m}{M_0} - \overline{K} + 1} \sum_{j=\frac{(i-1)m}{M_0}}^{\frac{im}{M_0}-\overline{K}} \left(x_{j+\overline{K}} - x_j\right), \quad (12)$$

for some $r, \ell > 0$. Podolskij and Vetter (2009b) showed that with FA jumps and the noise, when $r + \ell = 2$, $\widehat{\sigma}_{\mathsf{MBV}}(r,\ell)$ converges in probability to some quantity depending on ω^2 and IV. Given that ω^2 can be conveniently estimated by $\widehat{\omega}^2 = \frac{1}{2m}\sum_{i=1}^m \triangle_i x^2$, a consistent estimator on IV can be derived from $\widehat{\sigma}_{\mathsf{MBV}}(r, 2-r)$, but the convergence rate is unknown.

Podolskij and Vetter (2009a) modified $\widehat{\sigma}_{\mathsf{MBV}}$ with a weight function g, and the modified MBV is defined as

$$\widetilde{\sigma}_{\mathsf{MBV}}(r,\ell) = m^{\frac{r+\ell}{4}-1} \sum_{i=0}^{m-2\overline{K}+1} |\widetilde{X}_i|^\ell |\widetilde{X}_{i+\bar{k}}|^r, \text{ where } \widetilde{X}_i = \sum_{j=1}^{\overline{K}} g\left(j/\overline{K}\right) \triangle_{i+j} x,$$

for $r, \ell > 0$. In the presence of IA jumps that satisfy some conditions, Podolskij and Vetter (2009a) computed the limit of $\widetilde{\sigma}_{\mathsf{MBV}}(r,\ell)$ in probability when $r + \ell = 2$. This limit is also determined by the IV and ω^2. Then a consistent IV estimator can be naturally derived from $\widetilde{\sigma}_{\mathsf{MBV}}(r, 2-r)$ and $\widehat{\omega}^2$ for $r \in (0,2)$. In the presence of IA jumps, Podolskij and Vetter (2009a) showed a CLT for $\widetilde{\sigma}_{\mathsf{MBV}}(r,\ell)$ when $r, \ell < 1$. They also provided a CLT for $\widetilde{\sigma}_{\mathsf{MBV}}(r,\ell)$ when $r + \ell = 2$ in the absence of jumps. Based on this result, Christensen, Hounyo, and Podolskij (2018) established a CLT for the truncated MBV defined as

$$\widetilde{\sigma}_{\mathsf{TMBV}}(r,\ell) = m^{\frac{r+\ell}{4}-1} \sum_{i=0}^{m-2\overline{K}+1} |\widetilde{X}_i|^\ell \mathbf{1}_{\left\{|\widetilde{X}_i|<v_m\right\}} |\widetilde{X}_{i+\bar{k}}|^r \mathbf{1}_{\left\{|\widetilde{X}_{i+\bar{k}}|<v_m\right\}}, \text{ where}$$

$$r, \ell > 0, v_m = c u_m^{\overline{\omega}} \text{ for some } \overline{\omega} \in (0, 1/2), \text{ and } u_m = \frac{\overline{K}}{m},$$

with a certain restriction on the degree of activity of the IA jumps. In particular, the CLT indicates that when $r + \ell = 2$ the limit in probability of $\widetilde{\sigma}_{\mathsf{TMBV}}(r,\ell)$ is the same as that of $\widetilde{\sigma}_{\mathsf{MBV}}(r,\ell)$. Thus a consistent IV estimator can be derived from $\widetilde{\sigma}_{\mathsf{TMBV}}(r, 2-r)$, with the estimation error multiplied by $m^{1/4}$, being asymptotically mixed Gaussian in the sense of stable convergence in law. Moreover, Christensen, Oomen, and Podolskij (2014) showed that in the presence of FA jumps, $m^{1/4}\left(\widehat{\sigma}_{\mathsf{BV}} - \mathrm{IV}\right)$ is asymptotically mixed Gaussian in the sense of stable convergence in law, in which $\widehat{\sigma}_{\mathsf{BV}}$ is also an MBV-type estimator given by

$$\widehat{\sigma}_{\mathsf{BV}} = \frac{m\pi}{4(m-2K_1+2)K_1\phi_{K_1}} \sum_{i=0}^{m-2K_1+1} |X_{i,K_1}|\,|X_{i+K_1,K_1}| - \frac{\widehat{\omega}^2}{c_6^2\phi_{K_1}}, \text{ where}$$

$$X_{i,K_1} = \frac{1}{K_1} \sum_{j=K_1/2}^{K_1-1} \left(x_{i+j} - x_{i+j-\frac{K_1}{2}}\right), \ \ \phi_{K_1} = \frac{2+K_1^2}{12K_1^2},$$

and c_6 is the constant from Table 1.

3.8 Quantile-based realized volatility

Christensen, Oomen, and Podolskij (2010) proposed two kinds of QRV estimators, respectively for the no-noise and noisy data conditions, and here, we only focus on the latter. Like the PA estimator and the MBV in Podolskij and Vetter (2009a), the QRV employs a weight function g to compute the weighted averages of intra-daily returns within overlapping intervals in order to relieve the impact of the noise. For a generic interval within which we perform the pre-averaging procedure, if it contains a jump, the jump will cause extreme values in the averaged returns. For the sake of excluding such extreme values and so the effects of jumps, Christensen, Oomen, and Podolskij (2010) selected the averaged returns based on their empirical quantiles across each overlapping interval, and used the selected data to construct the QRV.

Specifically, the first step to construct the QRV estimator is to compute the weighted average returns:

$$X_i = \sum_{j=1}^{K_2-1} g\left(\frac{i}{K_2}\right) \triangle_{i+j} x, \ \text{for } i = 0, \ldots, m-K_2+1,$$

where the function g also satisfies conditions (i) and (ii) above (10). Then for each interval $\left[\frac{i}{m}, \frac{i+c(K_2-1)}{m}\right]$, where c is some positive constant, we consider the sequence $C_i =$

$\left\{X_{i+(j-1)(K_2-1)}\right\}_{j=1}^{c}$ that contains the averaged returns in it, and compute

$$q_i^*(\lambda_j) = h^2_{\lambda_j c}\left(m^{1/4}C_i\right) + h^2_{\lambda_j(1-c)+1}\left(m^{1/4}C_i\right),$$

where $h_j(F) = F_{(j)}$ is the j-th order statistic of the sequence $F = (f_1, \ldots, f_k)$, and $\lambda_j \in \left(\frac{1}{2}, 1\right)$. Define

$$\widehat{\sigma}^2_{\mathsf{QRV},j} = \frac{1}{c_7\varphi_2(m - c(K_2-1)+1)} \sum_{i=0}^{m-c(K_2-1)} \frac{q_i(\lambda_j)}{v(c,\lambda_j)}, \tag{13}$$

where c_7 is defined in Table 1, φ_2 is defined in (11), $v(c,\lambda_j) = E\left(|U_{c\lambda_j}|^2 + |U_{c-c\lambda_j+1}|^2\right)$, and $U_{c\lambda_j}$ is the $(c\lambda_j)$-th order statistic of an independent standard normal sample $\{U_i\}_{i=1}^{c}$. Then the QRV estimator $\widehat{\sigma}^2_{\mathsf{QRV}}$ is obtained by taking the weighted average of $\widehat{\sigma}^2_{\mathsf{QRV},j}$, and correcting the bias caused by the noise:

$$\widehat{\sigma}^2_{\mathsf{QRV}} = \sum_{i=1}^{\tilde{c}} b_i \widehat{\sigma}^2_{\mathsf{QRV},i} - \frac{\varphi_1}{c^2\varphi_2}\widehat{\omega}^2,$$

where $\tilde{c}$ is a positive constant, φ_1 is defined in (11), $\widehat{\omega}^2 = \frac{1}{2m}\sum_{i=1}^{m} \triangle_i x^2$, $b_i \geq 0$, and $\sum_{i=1}^{\tilde{c}} b_i = 1$. Since the increments of the Brownian motion are normal, the rescaling factor $v(c,\lambda_j)$ in (13) is from the quantiles of the standard normal distribution. (2) is needed in order to obtain the desired property for $\widehat{\sigma}^2_{\mathsf{QRV}}$, as σ_t is required to be roughly unchanged within small intervals. Then Christensen, Oomen, and Podolskij (2010) showed that in the presence of FA jumps and the noise, $m^{1/4}\left(\widehat{\sigma}^2_{\mathsf{QRV}} - \mathrm{IV}\right)$ is asymptotically mixed Gaussian in the sense of stable convergence in law.

3.9 Truncated pre-averaging estimator

Truncated pre-averaging estimators that consistently estimate the IV in the presence of IA jumps and the noise can be found, among others, in Wang, Liu, and Liu (2013), Jing, Liu, and Kong (2014) and Koike (2016). The first paper shows the consistency without specifying the convergence rate, while the other two show a CLT for their estimators. Jing, Liu, and Kong (2014) used the absolute value of $\triangle_{i,K_0}X(g)$ defined in (10) to detect the presence of jumps. If $|\triangle_{i,K_0}X(g)|$ is abnormally large, it is likely that there is a jump(s) in $]t_i, t_{i+K_0-1}]$, and this jump can change the asymptotic distribution of the PA estimator. Then Jing, Liu, and Kong (2014) obtained the truncated pre-averaging (TPA) estimator by subtracting $(\triangle_{i,K_0}X(g))^2$ from the PA estimator when $|\triangle_{i,K_0}X(g)|$ is larger than some threshold, so that the TPA estimator is robust to both jumps and the noise. Specifically, the TPA estimator $\widehat{\sigma}^2_{\mathsf{TPA}}$ is defined as

$$\widehat{\sigma}^2_{\mathsf{TPA}} = \frac{1}{K_0\varphi_2} \sum_{i=0}^{m-K_0+1} (\triangle_{i,K_0}X(g))^2 \, \mathbf{1}_{\{|\triangle_{i,K_0}X(g)| \le u_m\}} - \frac{\varphi_1}{2K_0^2\varphi_2} \sum_{i=1}^{m} (\triangle_i x)^2, \qquad (14)$$

where the threshold u_m is set so that $u_m m^{-\omega_1} \to 0$ and $u_m m^{-\omega_2} \to \infty$ for some $0 \le \omega_1 < \omega_2 < \frac{1}{4}$, and φ_1, φ_2 are the same as in (11). Notice that there is no truncation on the second component of the right-side of (14), since the impact of jumps on that component is $o_{\mathrm{P}}(m^{-1/4})$, and the downward bias due to truncations on the first component is small. Jing, Liu, and Kong (2014) allowed jumps to have infinite activity. They showed that $\widehat{\sigma}^2_{\mathsf{TPA}}$ has the same asymptotic distribution as the one derived for $\widehat{\sigma}^2_{\mathsf{PA}}$ by Jacod, Li, Mykland, Podolskij, and Vetter (2009) that did not consider jumps, when the IA jump process satisfies certain properties, the dynamics of σ can be described by (2), and the noise shares the same properties as in Jacod, Li, Mykland, Podolskij, and Vetter (2009). In addition, Koike (2016) applied the Hayashi-Yoshida technique to the truncated pre-averaging estimator. As Hayashi and Yoshida (2005) designed this technique to deal with the asynchronicity problem for multivariate analysis, the truncated pre-averaging Hayashi-Yoshida estimator by Koike (2016) is robust to some irregular sampling schemes.

3.10 Nearest neighbor truncation estimators

Developed by Andersen, Dobrev, and Schaumburg (2012), the nearest neighbor truncation mechanism eliminates the jump impact by comparing each absolute intraday return with its adjacent absolute return(s), and throwing away the bigger returns in constructing the estimators. Specifically, Andersen, Dobrev, and Schaumburg (2012) provided the following estimators by applying this technique to $\widehat{\sigma}^2_{\mathsf{RV}}$:

$$\widehat{\sigma}^2_{\mathsf{minRV}} = \frac{\pi}{\pi - 2}\frac{m}{m-1}\sum_{i=1}^{m-1} \min\left(\left|\triangle_i x\right|, \left|\triangle_{i+1} x\right|\right)^2, \text{ and}$$

$$\widehat{\sigma}^2_{\mathsf{medRV}} = \frac{\pi}{6 - 4\sqrt{3} + \pi}\frac{m}{m-2}\sum_{i=2}^{m-1} \operatorname{med}\left(\left|\triangle_{i-1} x\right|, \left|\triangle_i x\right|, \left|\triangle_{i+1} x\right|\right)^2,$$

where $\frac{m}{m-1}$ and $\frac{m}{m-2}$ are for finite sample bias corrections, and the rescaled factors $\frac{\pi}{\pi-2}$, $\frac{\pi}{6-4\sqrt{3}+\pi}$ are obtained following the Brownian motion properties, as their values, e.g., can be easier to pin down when σ is a constant. When there is no noise and jumps have finite activity, Andersen, Dobrev, and Schaumburg (2012) demonstrated the consistency of $\widehat{\sigma}^2_{\mathsf{minRV}}$ and $\widehat{\sigma}^2_{\mathsf{medRV}}$ as IV estimators, and the asymptotic centered mixed Gaussian distributions for $m^{1/2}\left(\widehat{\sigma}^2_{\mathsf{minRV}} - \mathrm{IV}\right)$ and $m^{1/2}\left(\widehat{\sigma}^2_{\mathsf{medRV}} - \mathrm{IV}\right)$ in the sense of stable convergence in law when σ follows (2). The estimators $\widehat{\sigma}^2_{\mathsf{minRV}}$ and $\widehat{\sigma}^2_{\mathsf{medRV}}$ are robust to FA jumps because when m is large enough, for two adjacent intervals of the form $(t_i, t_{i+1}]$, at most one of them contains a jump, which makes the corresponding absolute return the bigger one and truncated from the estimators. Andersen, Dobrev, and Schaumburg (2012) further applied this nearest neighbor truncation method to realized variations of higher powers, and thus obtained consistent estimators of $\int_0^1 \sigma_t^p dt$ for any positive and even p, when there are FA jumps and no noise.

To deal with the time-dependent noise and FA jumps, Varneskov (2017) combines the above technique with the realized kernel estimators. Specifically, he considered the partitions $\left\{0, \frac{L}{m}, \ldots, \frac{m_L L}{m}\right\}$, where $m_L = \lfloor m/L \rfloor$, and for each $i \in \{1, \ldots, m_L\}$, constructs the flat-top realized kernels $\widehat{\sigma}^2_{\mathsf{FRK},i}$ based on the data in $\left[\frac{(i-1)L}{m}, \frac{iL}{m}\right]$. When σ satisfies (2), $\widehat{\sigma}^2_{\mathsf{FRK},i} \approx \sigma^2_{iL}\frac{L}{m}$, if there is no jump in $\left[\frac{(i-1)L}{m}, \frac{iL}{m}\right]$. Then in the spirit of $\widehat{\sigma}^2_{\mathsf{medRV}}$, the medium block realized kernels are defined as

$$\widehat{\sigma}^2_{\mathsf{MBRK}} = \sum_{i=2}^{m_L - 1} \operatorname{med}\left(\widehat{\sigma}^2_{\mathsf{FRK},i-1}, \widehat{\sigma}^2_{\mathsf{FRK},i}, \widehat{\sigma}^2_{\mathsf{FRK},i+1}\right),$$

and Varneskov (2017) showed under (2) that $\widehat{\sigma}^2_{\mathsf{MBRK}}$ converges to IV in probability at the rate of $m^{-1/4}$ in the presence of FA jumps and the same type of noise as displayed in (9).

4 Two novel truncated estimators

In line with the methodology of truncated volatility estimators, we propose the truncated two-scales and kernel estimators respectively in this and the next subsections. The general idea is to employ the β_i's defined below as a jump indicator, and then truncate the intervals that contain jumps from the two-scales and kernel estimators.

Similarly to the TPA estimator, we remove the effect of the jumps based on the absolute values of the locally averaged returns β_i which are defined as:

$$\beta_i(m) = \frac{1}{K_3} \sum_{j=i}^{i+K_3-1} (x_j - x_{j-K_3}) \,, \quad \text{for} \quad i = 1, \ldots, m.$$

Notice that like $\widehat{\sigma}^2_{\mathsf{RK}}$ defined in (7), obtaining the values of the β_i's for $i < K_3$ and $i > m - K_3 + 1$ requires observations outside the period $[0, 1]$. This is for the simplicity of explanation, and can be achieved by extending the whole period without additional technical problems. The properties of the β_i's related to the jump detection when jumps have finite activity are summarized in Lemma 1 of the appendix. Then in Lemma 2 of the appendix we analyze the behavior of the β_i's in the presence of IA jumps, which is a more involved case.

4.1 Truncated two-scales estimator

The truncated two-scales realized volatility (TTSRV) estimator is obtained by applying the jump indictor β to $\widehat{\sigma}^2_{\mathsf{TS}}$, and performing the truncation accordingly as follows:

$$\widehat{\sigma}^2_{\mathsf{TTS}} = \frac{1}{K} \sum_{j=K}^{m} (x_j - x_{j-K})^2 \mathbf{1}_{E_j} - \frac{1}{K} \sum_{j=K}^{m} (x_j - x_{j-1})^2 \mathbf{1}_{E_j}, \tag{15}$$

where $E_j = \{|\beta_i| \le r(m), \text{ for all } i = j - K + 1, \ldots, j\}$, and $r(m)$ is the threshold.

Notice that on the right-side of (15), the scheme can truncate many terms in the second component that are not affected by jumps in the presence of FA jumps. The reason is that due to the noise effects, substantial downward bias which can make the truncated two-scales estimator inconsistent could occur, if we truncate $(x_j - x_{j-1})^2$ from the second component only when there is a jump in $[t_{j-1}, t_j]$. This truncation scheme is different from that of $\widehat{\sigma}^2_{\mathsf{TPA}}$, since there is no truncation on the second component of the right-side of (14). This is because for the $\widehat{\sigma}^2_{\mathsf{TPA}}$, the impact of the noise on the truncated terms from the first component is $o_{\mathrm{P}}(m^{-1/4})$, which is negligible, so there is no need to truncate the second component to cancel the noise effects. This difference is natural since the noise impact on the first component of $\widehat{\sigma}^2_{\mathsf{TS}}$ is $O_{\mathrm{P}}(m^{1/3})$, while it is only $O_{\mathrm{P}}(1)$ for the first component of $\widehat{\sigma}^2_{\mathsf{PA}}$.

The estimator $\widehat{\sigma}^2_{\mathsf{TTS}}$ is similar to the truncated two-scales estimators proposed by Fan and Wang (2007) and Boudt and Zhang (2015), since they both use some jump indicators to truncate the intervals that may contain jumps. However, there are notable differences in the truncation schemes used. Fan and Wang (2007) adopted the wavelet technique to locate jumps, and the technical analysis relies on the assumption that the jump process is independent of the continuous process in y, which is not needed for $\widehat{\sigma}^2_{\mathsf{TTS}}$. Boudt and Zhang (2015) detected the presence of jumps in $(t_i, t_j]$ by checking whether $|x_j - x_i|$ is larger than some threshold, but the validity of this indicator can be disturbed by the noise. Thus Boudt and Zhang (2015) did not deliver the consistency of their estimator. Based on the theoretical property of $\widehat{\sigma}^2_{\mathsf{TSRV}}$ derived in Zhang, Mykland, and Aït-Sahalia (2005), we can obtain the asymptotic distribution of $\widehat{\sigma}^2_{\mathsf{TTS}}$ under FA jumps as follows:

THEOREM 1. *Consider the assumptions of Lemma 1 in the appendix, and that* $\lim_{m\to\infty} \frac{K_3}{K} = 0$. *If* σ_t *is continuous and the* u_i*'s are mutually independent, then as* $m \to \infty$,

$$m^{1/6}\left(\widehat{\sigma}^2_{\mathsf{TTS}} - \int_0^1 \sigma_t^2 dt\right) \xrightarrow{\mathcal{L}} \mathrm{MN}\left(0, 8c_1^{-2}\omega^4 + \frac{4}{3}c_1 \int_0^1 \sigma_t^4 dt\right),$$

where c_1 *is the constant from Table 1, the convergence is stable in law, and "*MN*" means a mixed Gaussian distribution.*

Observe that by Theorem 1, $\widehat{\sigma}^2_{\mathsf{TTS}}$ has the same asymptotic distribution as the TSRV estimator in Zhang, Mykland, and Aït-Sahalia (2005) that did not consider jumps. This is because in the proof we show that the difference between $\widehat{\sigma}^2_{\mathsf{TTS}}$ and the TSRV built with data not affected by jumps, which is defined as $\overline{\sigma}^2_{\mathsf{TS}}$ in the proof, is $o_{\mathrm{P}}(m^{-1/6})$.

The assumption that σ_t is continuous is needed for proving the asymptotic distribution (Zhang, Mykland, and Aït-Sahalia 2005). Without this assumption, it can be still checked that $\overline{\sigma}^2_{\mathsf{TS}} - \text{IV}$ is $O_{\text{P}}(m^{-1/6})$ (see e.g. Fan, Li, and Yu 2012). In addition, when proving $\widehat{\sigma}^2_{\mathsf{TTS}} - \overline{\sigma}^2_{\mathsf{TS}} = o_{\text{P}}(m^{-1/6})$, σ_t does not need to be continuous, so the difference between $\widehat{\sigma}^2_{\mathsf{TTS}}$ and IV is still $O_{\text{P}}(m^{-1/6})$ when σ_t is not continuous.

$\widehat{\sigma}^2_{\mathsf{TTS}}$ consistently estimates the IV also in the presence of infinite activity jumps. Specifically, we assume $J_t = J_{1,t} + J_{2,t}$, where $J_{1,t}$ is a general FA jump process with $J_{1,t} = \sum_{i=1}^{N_t} Y_i$, and $J_{2,t}$ is an IA Lévy pure jump process. Under these conditions, we show the consistency of TTSRV in the following theorem:

Theorem 2. *Assume the hypothesis of Lemma 1 in the appendix except the assumption that J_t has finite activity. Assume also that J_2 is independent of N, and u_i's are mutually independent. If there exists $\alpha > 0$ such that*

$$\lim_{m\to\infty} \frac{K_3^3}{r^2(m)m^{1/3-\alpha}} = 0,$$

then as $m \to \infty$, $\widehat{\sigma}^2_{\mathsf{TTS}} \xrightarrow{\text{P}} \int_0^1 \sigma_t^2 dt$.

The structure of the proof of this theorem is similar to that of (Mancini, 2009, Theorem 4). On one hand, we use the measure β_i and the threshold $r(m)$ in order to cut off the jumps from J_1. On the other hand, we truncate the jumps in J_2 with absolute values larger than $\sqrt{\delta + 16r^2(m)}$, where $\delta > 0$ is arbitrary, and show that the information loss caused by the truncation is negligible.

4.2 Truncated flat-top realized kernel estimator

Another novel truncated estimator introduced in this paper is the truncated flat-top realized kernel (TFTRK) estimator, which consistently estimates the IV in the presence of jumps and M-dependent noise with the optimal convergence rate $m^{-1/4}$. Here $M-$dependence means that u_i is independent of u_j when $|i - j| > M$. The TFTRK estimator is defined as

$$\widehat{\sigma}^2_{\mathsf{TRK}} = \sum_{i=1}^{m} \left(\sum_{h=1}^{H(1+C)} \overline{k}\left(\frac{h}{H}\right) (\triangle_{i-h}x + \triangle_{i+h}x) + \triangle_i x \right) \triangle_i x \mathbf{1}_{\overline{E}_i}, \tag{16}$$

where $\overline{E}_i = \{|\beta_j| \leq r(m), \text{ for all } j = i - H(1+C), \ldots, i + H(1+C)\}$. Without the indicator function, the right-side of (16) is the same as $\widehat{\sigma}^2_{\mathsf{FRK}}$ defined in (8). Notice that the scheme can truncate many terms of the form $\overline{k}\left(\frac{i-j}{H}\right) \triangle_i x \triangle_j x$ even though there are no jumps in $[t_{i-1}, t_i]$ or $[t_{j-1}, t_j]$. On the contrary, if we truncate $\overline{k}\left(\frac{i-j}{H}\right) \triangle_i x \triangle_j x$ only when it is affected by a jump, the truncated kernel estimator would be inconsistent due to the noise effects.

Moreover, in computing $\widehat{\sigma}^2_{\mathsf{TRK}}$ we adopt the similar jittering procedure as in Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008, 2011) and Varneskov (2017) in order to smooth away the boundary effects which can increase the asymptotic variance of the estimator. The difference is that for $\widehat{\sigma}^2_{\mathsf{TRK}}$, the boundary effects arise when we truncate each component due to one jump, so the boundary points of each truncated component should be jittered, and the corresponding mechanism and procedure are described in the appendix for the case of infinite activity jumps. In addition, we use the jittered points to compute the value of the right-side of (16) except $\mathbf{1}_{\overline{E}_i}$, that is, the β_i's and so the $\overline{E}_i$'s are computed still based on the original observations.

Given relevant analysis in the literature (e.g. Kim, Wang, and Zou 2016 and Varneskov 2017), one can show that when there are no jumps, the difference $\widehat{\sigma}^2_{\mathsf{FRK}} - \mathrm{IV}$ is $O_{\mathrm{P}}(m^{-1/4})$. As the jump indicator $\mathbf{1}_{\overline{E}_i}$ can remove finite activity jumps from the estimator, and the magnitude of information loss due to the truncation is $o_{\mathrm{P}}(m^{-1/4})$, $\widehat{\sigma}^2_{\mathsf{TRK}}$ converges to IV in probability also at the rate of $m^{-1/4}$, which gives the following theorem:

Theorem 3. *Under the assumptions of Lemma 1 in the appendix and that* $\lim_{m\to\infty} \frac{K_3}{m^{1/2}} = 0$, *we have*

$$\widehat{\sigma}^2_{\mathsf{TRK}} - \int_0^1 \sigma_t^2 dt = O_{\mathrm{P}}(m^{-\frac{1}{4}}), \qquad as \quad m \to \infty.$$

In the setting of Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008) where the u_i's are mutually independent and (2) is true, it can be easily checked that $\widehat{\sigma}^2_{\mathsf{FRK}}$ defined in

(8) has the same asymptotic distribution as $\widehat{\sigma}^2_{\mathsf{RK}}$ defined in (7), which can be also verified by Theorem 1 in Varneskov (2017). Then the proof of Theorem 3 implies that $\widehat{\sigma}^2_{\mathsf{TRK}}$ has the same asymptotic distribution holds when FA jumps are present, which leads to the following corollary:

Corollary 1. *Consider the assumptions of Lemma 1 in the appendix, that* $\lim_{m\to\infty}\frac{K_3}{m^{1/2}}=0$, *$\sigma$ is subject to* (2), *and the u_i's are mutually independent. Then, as $m\to\infty$,*

$$m^{1/4}\left(\widehat{\sigma}^2_{\mathsf{TRK}}-\int_0^1\sigma_u^2du\right)\xrightarrow{\mathcal{L}_{\mathcal{H}_1}}\mathrm{MN}\left(0,4\int_0^1\sigma_u^4du\left(c_2k_1+2c_2^{-1}k_2\rho_1\rho_2^2+c_2^{-3}k_3\rho_2^4\right)\right),\quad(17)$$

where c_2 is defined in Table 1, $\xrightarrow{\mathcal{L}_{\mathcal{H}_1}}$ means $\mathcal{H}_1-$stable convergence in law with $\mathcal{H}_1$ being the σ-algebra generated by (y_s,σ_s) for $s\in[0,1]$, $k_1=\int_0^1k(x)^2dx, k_2=\int_0^1k'(x)^2dx, k_3=\int_0^1k''(x)^2dx, \rho_1=\int_0^1\sigma_u^2du/\sqrt{\int_0^1\sigma_u^4du}$, and $\rho_2=\omega^2/\sqrt{\int_0^1\sigma_u^4du}$.

The notion of $\mathcal{H}_1-$stable convergence in law follows from Theorem 1 in Varneskov (2017), and its definition can be found in Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008). Here it amounts to the statement that $(\mathcal{Y}_m,Z)\xrightarrow{\mathcal{L}}(\mathcal{Y},Z)$ for any random variable Z which is $\mathcal{H}_1$-measurable, where $\mathcal{Y}_m$ is the left-side of (17) and $\mathcal{Y}$ is the right-side. Lemmas 1-3 and Proposition 5 by Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008) revealed some properties of this type of convergence.

After reviewing and introducing the above IV estimators, we summarize their properties in Table 2 concerning their convergence rates and robustness to jumps and the noise.

5 Simulation study

In this section we first perform simulation exercises to evaluate the efficiency of the truncated two-scales and kernel estimators in finite samples, as these two estimators are newly introduced in this paper. Then by simulations we compare the performance of $\widehat{\sigma}_{\mathsf{MBV}}(r,\ell),\widehat{\sigma}^2_{\mathsf{QRV}},\widehat{\sigma}^2_{\mathsf{TPA}},\widehat{\sigma}^2_{\mathsf{TTS}}$, $\widehat{\sigma}^2_{\mathsf{TRK}}$ and $\widehat{\sigma}^2_{\mathsf{MBRK}}$. We simulate the dynamics of y_t according

Table 2: Property of the realized volatility estimators

	Robust to jumps	Robust to the noise	Convergence rate	Dependence of u_i's
$\widehat{\sigma}^2_{\mathsf{RV}}$	No	No	$m^{-1/2}$	/
$\widehat{\sigma}^2_{\mathsf{RRV}}$	No	No	$m^{-1/2}$	/
$\widehat{\sigma}_{\mathsf{MPV}}$	Yes, robust to IA jumps	No	$m^{-1/2}$	/
$\widehat{\sigma}_{\mathsf{BPV}}$	Yes, robust to IA jumps	No	Unknown	/
$\widehat{\sigma}^2_{\mathsf{TS}}$	No	Yes	$m^{-1/6}$	independent
$\widehat{\sigma}^2_{\mathsf{RK}}$	No	Yes	$m^{-1/4}$	independent
$\widehat{\sigma}^2_{\mathsf{FRK}}$	No	Yes	$m^{-1/4}$	see equation (9)
$\widehat{\sigma}^2_{\mathsf{PA}}$	No	Yes	$m^{-1/4}$	independent conditional on y
$\widehat{\sigma}^2_{\mathsf{TRV}}$	Yes, robust to IA jumps	No	$m^{-1/2}$	/
$\widehat{\sigma}_{\mathsf{MBV}}$	Yes, robust to FA jumps	Yes	Unknown	independent
$\widetilde{\sigma}_{\mathsf{MBV}}$	Yes, robust to IA jumps	Yes	Unknown	independent
$\widetilde{\sigma}_{\mathsf{TMBV}}$	Yes, robust to IA jumps	Yes	$m^{-1/4}$	independent
$\widehat{\sigma}_{\mathsf{BV}}$	Yes, robust to FA jumps	Yes	$m^{-1/4}$	independent
$\widehat{\sigma}^2_{\mathsf{QRV}}$	Yes, robust to FA jumps	Yes	$m^{-1/4}$	independent
$\widehat{\sigma}^2_{\mathsf{TPA}}$	Yes, robust to IA jumps	Yes	$m^{-1/4}$	independent conditional on y
$\widehat{\sigma}^2_{\mathsf{minRV}}$	Yes, robust to FA jumps	No	$m^{-1/2}$	/
$\widehat{\sigma}^2_{\mathsf{medRV}}$	Yes, robust to FA jumps	No	$m^{-1/2}$	/
$\widehat{\sigma}^2_{\mathsf{MBRK}}$	Yes, robust to FA jumps	Yes	$m^{-1/4}$	see equation (9)
$\widehat{\sigma}^2_{\mathsf{TTS}}$	Yes, robust to IA jumps	Yes	$m^{-1/6}$ (FA jumps)	independent
$\widehat{\sigma}^2_{\mathsf{TRK}}$	Yes, robust to FA jumps	Yes	$m^{-1/4}$	$M-$dependent

This table shows the convergence rates and robustness to the jumps and noise of the reviewed estimators. The last column indicates the type of dependency among the noise terms that can be allowed by the estimator.

to (1), and

$$\sigma_t = \exp(\beta_0 + \beta_1 \tau_t),$$

where $d\tau_t = \alpha\tau_t dt + dW_t$ with W_t being a standard Brownian motion and $\mathrm{Corr}(dW_t, dB_t) = \rho$. We let $\mu = 0.03, \beta_0 = 0.3125, \beta_1 = 0.125, \alpha = -0.025, \rho = -0.3$, and these coefficients are selected from Jing, Liu, and Kong (2014) and Podolskij and Vetter (2009b). The noise u_i is a discrete i.i.d. $\mathrm{N}(0, \omega^2)$. The jump process J is a compound Poisson process with a constant intensity $\lambda = 2$, and the jump sizes are i.i.d. $\mathrm{N}(0, \xi^2)$.

We assume that a trading day is 6.5 hours long and the observed price x_t is measured for every 3 seconds (that is, $m = 7800$). The simulation is carried out using the Euler simulation scheme. Throughout this section we set $K_3 = \lfloor m^{\frac{1}{3}} \rfloor = 19$. For the TFTRK, we use the Parzen kernel, and choose $C = H^{-\frac{3}{5}}$ following Varneskov (2017), $H = \lfloor 0.5 m^{\frac{1}{2}} \rfloor = 44$, and $m_0 = 3$ which, as described in the appendix, is the number of jittered points when jittering occurs; and for the TTSRV we set $K = \lfloor 0.1 m^{\frac{2}{3}} \rfloor$. For each setting the simulation is replicated 1000 times.

Figure 1 shows the plots of the MSE of the TTSRV and TFTRK estimators as a function of $r(m)$ for different magnitudes of ξ (0.25, 0.50, 1), while ω is fixed at 0.10.

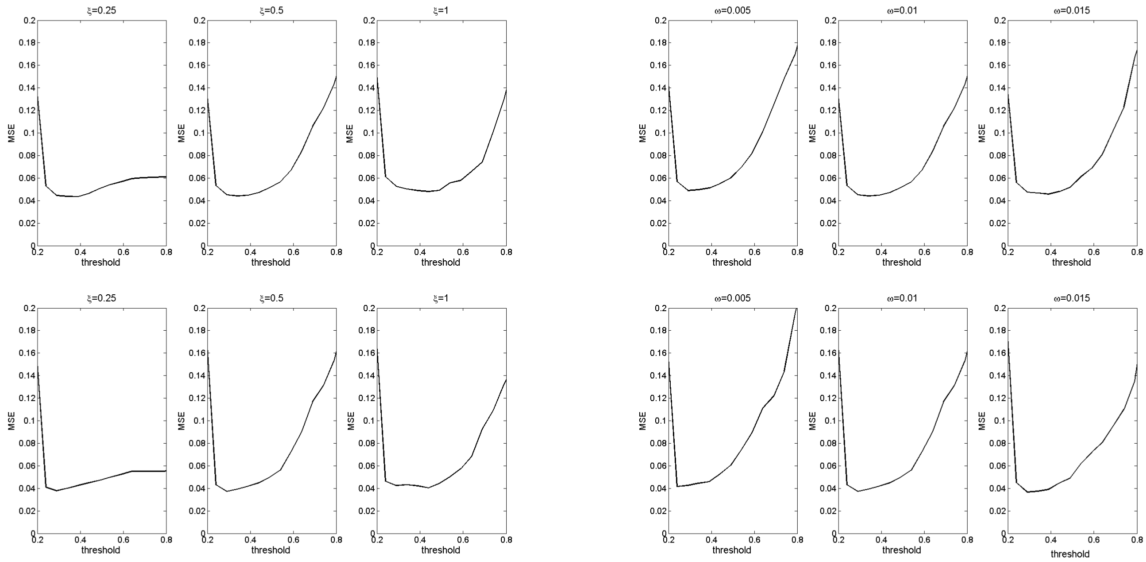

Figure 1: The figure shows the plot of the MSE of the TTSRV and TFTRK estimators as a function of the threshold $r(m)$. The top left panel shows the MSE of TTSRV for different values of the jump size standard deviation ξ ($\xi = 0.25, 0.5, 1$). The top right panel shows the MSE of TTSRV for different values of the market microstructure noise standard deviation ω ($\omega = 0.005, 0.010, 0.015$). The bottom left panel shows the MSE of TFTRK for different values of the jump size standard deviation ξ ($\xi = 0.25, 0.5, 1$). The bottom right panel shows the MSE of TFTRK for different values of the market microstructure noise standard deviation ω ($\omega = 0.005, 0.010, 0.015$).

The plots show that in both cases the MSE is a decreasing function of $r(m)$ when $r(m)$ is small. This is because when the threshold is too small, many intervals that do not contain jumps are truncated, and this can cause severe downward bias for the TTSRV and TFTRK estimators. On the other hand, the MSE is an increasing function of $r(m)$ when $r(m)$ is large. This is because when the threshold is too large, the intervals that contain price jumps are not truncated, which leads to upward bias for the TTSRV and TFTRK estimators. Figure 1 also displays the plots of the MSE of the TTSRV and TFTRK as a function of $r(m)$ for different magnitudes of ω (0.005, 0.01, 0.015), while ξ is fixed at 0.50. The MSE curves have the same convex shape documented in previous cases.

Next we investigate the finite sample distributions of the TTSRV and TFTRK estimators. We define the standardized estimation errors of the TTSRV and TFTRK estimators

Table 3: Distribution of the TTSRV estimator.

	$\xi = 0.50$	$\xi = 0.75$	$\xi = 1.0$
Mean	-0.204	-0.269	-0.311
Std Dev	0.968	1.061	1.00
Skewness	0.106	0.286	0.202
Kurtosis	2.68	2.98	2.86
Pct	0.035	0.080	0.063

Statistics of the standardized estimation errors of the TTSRV estimator. Pct means the percentage of the standardized errors whose absolute values are larger than 1.96 (asymptotically this percentage is 0.05).

Table 4: Distribution of the TFTRK estimator.

	$\xi = 0.50$	$\xi = 0.75$	$\xi = 1.0$
Mean	-0.059	-0.065	-0.166
Std Dev	1.079	1.029	1.022
Skewness	0.12	0.14	0.12
Kurtosis	3.28	3.14	2.76
Pct	0.058	0.053	0.053

Statistics of the standardized estimation errors of the TFTRK estimator. Pct means the percentage of the standardized errors whose absolute values are larger than 1.96 (asymptotically this percentage is 0.05).

respectively as

$$z_1 = \frac{m^{1/6}\left(\widehat{\sigma}^2_{\mathsf{TTS}} - \int_0^1 \sigma_s^2 ds\right)}{\left(8c_1^{-2}\omega^4 + \frac{4}{3}c_1\int_0^1 \sigma_s^4 ds\right)^{1/2}},$$

and

$$z_2 = \frac{m^{1/4}\left(\widehat{\sigma}^2_{\mathsf{TRK}} - \int_0^1 \sigma_s^2 ds\right)}{\sqrt{4\int_0^1 \sigma_u^4 du(c_2k_1 + 2c_2^{-1}k_2\rho_1\rho_2^2 + c_2^{-3}k_3\rho_2^4)}},$$

Theorem 1 and Corollary 1 imply that if the sample size is sufficiently large, z_1 and z_2 should be approximately standard normally distributed. We compute z_1 and z_2 keeping the value of the threshold $r(m)$ fixed at 0.35 and $\eta = 0.01$. From the last set of simulations we can see that MSE is relatively small when $r(m)$ is around 0.35.

Tables 3 and 4 report summary statistics of z_1 and z_2. We can see that the truncation of intervals causes a downward bias in estimation for both estimators, and this bias increases as the standard deviation of the jump size grows. Overall, z_2 is more unbiased against the standard normal distribution, and the inspection of the histogram and the normal qqplot of z_2 (nor reported in the paper) conveys that the approximation provided by the asymptotic theory is adequate.

Last, we compare the efficiency of the TTSRV and TFTRK with the other estimators mentioned at the beginning of this section. In this exercise the threshold $r(m)$ of the TTSRV and TFTRK is fixed at 0.35, and besides the i.i.d. assumption, we also consider

Table 5: MSE comparisons for different values of the noise variance (i.i.d. noise)

ω	0	0.002	0.004	0.006	0.008	0.010	0.012	0.014
TPV	0.116	0.128	0.110	0.116	0.114	0.119	0.116	0.117
MBV	0.362	0.464	0.582	0.363	0.386	0.463	0.388	0.426
QRV	0.793	0.553	0.488	0.173	0.144	0.026	0.137	0.587
MBRK	0.085	0.071	0.093	0.079	0.131	0.078	0.069	0.065
TFTRK	0.038	0.043	0.042	0.038	0.037	0.045	0.039	0.045
TTSRV	0.044	0.050	0.045	0.044	0.041	0.047	0.043	0.049

Table 6: MSE comparisons for different values of the noise variance (MA(1) noise)

ω	0	0.002	0.004	0.006	0.008	0.010	0.012	0.014
TPV	0.133	0.108	0.090	0.098	0.092	0.124	0.091	0.093
MBV	0.373	0.578	0.460	0.393	0.442	0.352	0.361	0.408
QRV	0.793	0.618	0.555	0.245	0.192	0.099	0.167	0.675
MBRK	0.083	0.108	0.090	0.080	0.089	0.090	0.065	0.078
TFTRK	0.041	0.037	0.037	0.039	0.032	0.036	0.033	0.038
TTSRV	0.051	0.050	0.048	0.049	0.039	0.056	0.043	0.045

the following MA(1) process to generate u_i, which was also employed by Kim, Wang, and Zou (2016) and Varneskov (2017) in their numerical studies:

$$u_i = \theta_{i+1} - 0.5\theta_i, \tag{18}$$

where θ_i is a discrete i.i.d. $\mathrm{N}(0, \omega^2/1.25)$.

To compute $\widehat{\sigma}_{\mathsf{MBV}}(r, \ell)$ described in (12), we set $r = \ell = 1$, $\overline{K} = \lfloor m^{\frac{1}{2}} \rfloor = 88$, and $M_0 = \lfloor 0.625 m^{\frac{1}{2}} \rfloor = 55$, following Podolskij and Vetter (2009b). Recall the definition of $\widehat{\sigma}^2_{\mathsf{TPA}}$ in (14). In computing $\widehat{\sigma}^2_{\mathsf{TPA}}$, following Jing, Liu, and Kong (2014), we set $K_0 = \lfloor m^{\frac{1}{2}} \rfloor = 88$, $g(s) = \min(s, 1-s)$ for $s \in (0,1)$, and the threshold $u_m = m^{-0.23}$.

Given the process of constructing $\widehat{\sigma}^2_{\mathsf{QVR}}$ described in section 3.8, following Christensen, Oomen, and Podolskij (2010), we set $c = 40, \tilde{c} = 4$, $(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (0.8, 0.85, 0.9, 0.95)$, $g(x) = \min(x, 1-x)$ for $x \in (0,1)$ and the value of K_2 is determined through simulations. That is, we set $K_2 = 23$ after many experiments, which roughly yields smallest estimation errors on average in our setting. Moreover, the weight vector (b_1, b_2, b_3, b_4) is determined by equation (11) in Christensen, Oomen, and Podolskij (2010).

Analogous to the choice in Varneskov (2017), we choose $m_L = 20$ in computing $\widehat{\sigma}^2_{\mathsf{MBRK}}$, and so $L = \frac{m}{m_L} = 390$. We also adopt the Parzen kernel, and set $H_L = \lfloor 0.5 L^{\frac{1}{2}} \rfloor = 9$ and $C_L = H_L^{-\frac{3}{5}}$ when constructing $\widehat{\sigma}_{\mathsf{FRK},i}$, where H_L and C_L for the FTRK over $\left[\frac{(i-1)L}{m}, \frac{iL}{m}\right]$ are the counterparts of H and C for the FTRK over the whole period.

Figures 2 and 3 show the plots of the MSE of the estimators as a function of ω, and the

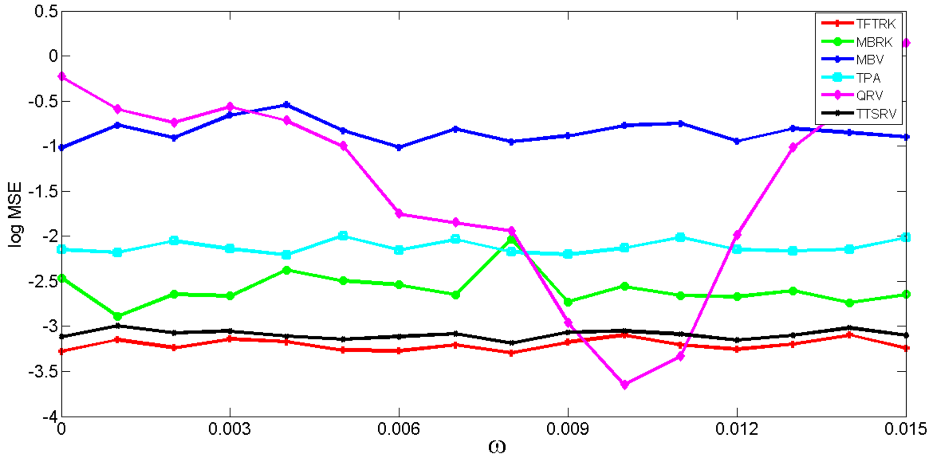

Figure 2: The figure shows the plot of log values of the MSE of the MBV, MBRK, TFTRK, TPA, QRV and TTSRV estimators as a function of standard deviation ω of the i.i.d. noise.

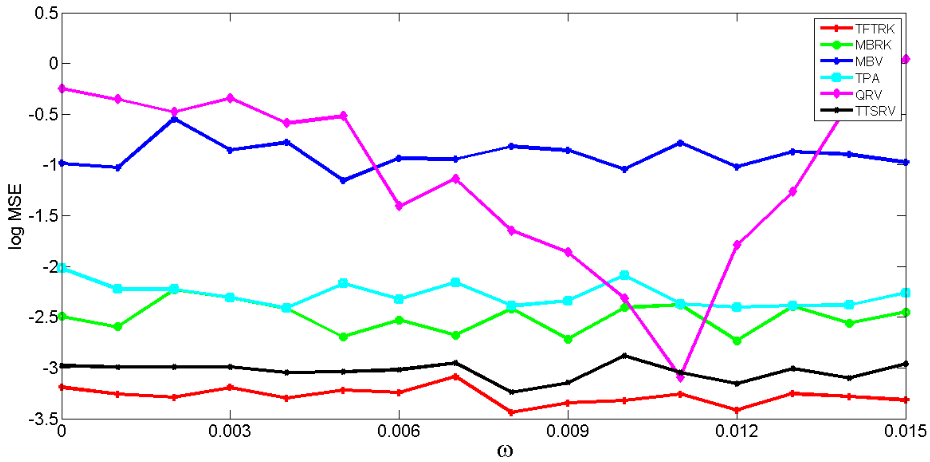

Figure 3: The figure shows the plot of log values of the MSE of the MBV, MBRK, TFTRK, TPA, QRV and TTSRV estimators as a function of standard deviation ω of the MA(1) noise.

value of ξ is fixed to 0.5, respectively under the condition where u_i's are i.i.d. or MA(1) as described by (18). Tables 5 and 6 report the corresponding MSE of the estimators for selected values of ω. The plots show that the MBRK, TPV, TTSRV and TFTRK estimators can steadily yield small estimation erros over the range of ω considered, and the TTSRV and TFTRK dominate the other estimators most of the time. Under i.i.d. noise, the QRV estimator yields the smallest MSE when ω is around 0.01. However, the efficiency of this estimator is quite sensitive to the value of ω, and it can lead to large

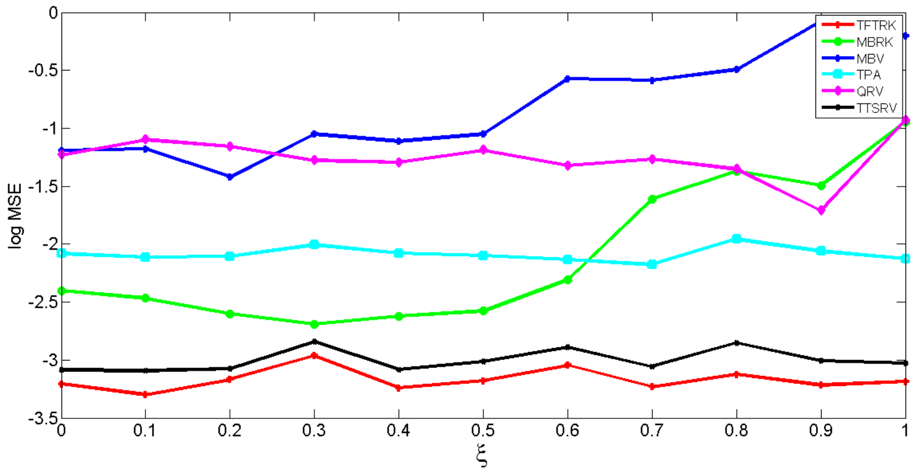

Figure 4: The figure shows the plot of log values of the MSE of the MBV, MBRK, TFTRK, TPA, QRV and TTSRV estimators as a function of the jump size standard deviation ξ in the presence of the i.i.d. noise.

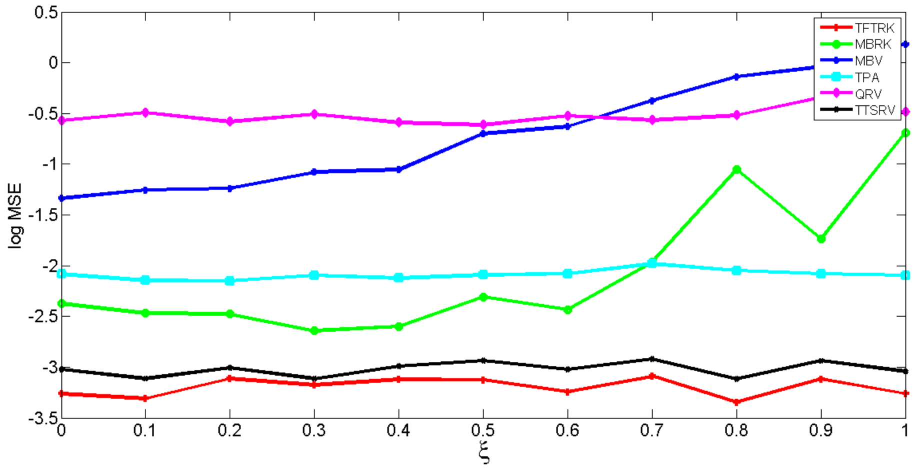

Figure 5: The figure shows the plot of log values of the MSE of the MBV, MBRK, TFTRK, TPA, QRV and TTSRV estimators as a function of the jump size standard deviation ξ in the presence of the MA(1) noise.

estimation error when ω is substantially different from 0.01. Moreover, it can be noticed that compared to the condition with i.i.d. noise, time-dependency in the noise can always increase the estimation error of the QRV.

Figures 4 and 5 show the plots of the MSE of the estimators as a function of ξ, respectively under the condition where u_i's are i.i.d. or MA(1), with ω fixed at 0.005. Tables 7 and 8 report the MSE of the estimators for selected values of ξ. The TTSRV

Table 7: MSE comparisons for different jump size variance (i.i.d. noise)

ξ	0	0.2	0.4	0.6	0.8	1
TPV	0.125	0.122	0.125	0.119	0.141	0.119
MBV	0.304	0.241	0.329	0.565	0.611	0.817
QRV	0.293	0.315	0.274	0.267	0.258	0.395
MBRK	0.091	0.074	0.073	0.100	0.254	0.388
TFTRK	0.041	0.042	0.039	0.048	0.044	0.041
TTSRV	0.046	0.046	0.046	0.055	0.058	0.048

Table 8: MSE comparisons for different jump size variance (MA(1) noise)

ξ	0	0.2	0.4	0.6	0.8	1
TPV	0.125	0.116	0.120	0.125	0.129	0.123
MBV	0.263	0.290	0.349	0.532	0.869	1.198
QRV	0.565	0.558	0.553	0.591	0.595	0.613
MBRK	0.093	0.084	0.074	0.088	0.349	0.502
TFTRK	0.038	0.045	0.044	0.039	0.035	0.038
TTSRV	0.049	0.050	0.050	0.049	0.044	0.048

and TFTRK achieve the best performance overall, and the estimation error of the TPA is also steadily small. The estimation precision of the MBRK is good when ξ is small, but when ξ is larger than 0.6, the MSE of the MBRK, as well as that of the MBV, increases significantly as ξ grows.

6 Conclusion

Estimating the integrated volatility with high frequency price data is a rapidly growing field of research in financial econometrics. In this paper, we have reviewed a number of IV estimators that are robust to either jumps or the market microstructure noise, or both. We can see that stemming from the classic RV estimator, methods including the use of two time scales, the kernel functions and the pre-averaging procedure have been developed to deal with the noise, and the mechanisms of the multipower variations, the quantile-based technique and the (nearest neighbor) truncations can be adopted to remove the disturbance from jumps. Then consistent IV estimators can be derived by the combinations of these methodologies, like the modulated bipower variations, the truncated pre-averaging and the medium block realized kernels. We also fill the gap in the literature by introducing the truncated versions of the two-scales and realized kernels estimators, and justify their consistency in the presence of jumps and the noise. Finally, we perform numerical studies to evaluate the efficiency of the different estimation approaches.

References

Aït-Sahalia, Y. and Jacod, J. (2014). High-Frequency Financial Econometrics. *Princeton, N.J.: Princeton University Press, first edition.*

Aït-Sahalia, Y., Jacod, J., and Li, J. (2012). Testing for Jumps in Noisy High-Frequency Data. *Journal of Econometrics*, **168**(2), 207–222.

Aït-Sahalia, Y., Fan, J., Laeven, R. J., Wang, C. D., and Yang, X. (2017). Estimation of the Continuous and Discontinuous Leverage Effects. *Journal of the American Statistical Association*, **112**(520), 1744–1758.

Andersen, T. G. and Bollerslev, T. (1998). Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts. *International Economic Review*, **39**(4), 885–905.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modeling and Forecasting Realized Volatility. *Econometrica*, **71**, 579–625.

Andersen, T. G., Bollerslev, T., and Diebold, F. X. (2007). Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *The Review of Economics and Statistics*, **89**(4), 701–720.

Andersen, T. G., Dobrev, D., and Schaumburg, E. (2012). Jump-Robust Volatility Estimation Using Nearest Neighbor Truncation. *Journal of Econometrics*, **169**(1), 75–93.

Bandi, F. M. and Russell, J. R. (2008). Microstructure Noise, Realized Variance, and Optimal Sampling. *The Review of Economic Studies*, **75**(2), 339–369.

Barndorff-Nielsen, O. E. and Shephard, N. (2002). Econometric Analysis of Realized Volatility and its Use in Estimating Stochastic Volatility Models. *Journal of the Royal Statistical Society, Series B*, **64**(2), 253–280.

Barndorff-Nielsen, O. E. and Shephard, N. (2004). Power and Bipower Variation with Stochastic Volatility and Jumps. *Journal of Financial Econometrics*, **2**, 1–48.

Barndorff-Nielsen, O. E. and Shephard, N. (2006). Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation. *Journal of Financial Econometrics*, **4**(1), 1–30.

Barndorff-Nielsen, O. E., Graversen, S. E., Jacod, J., Podolskij, M., and Shephard, N. (2006a). A Central Limit Theorem for Realised Power and Bipower Variations of Continuous Semimartingales. *From Stochastic Calculus to Mathematical Finance, Springer*, pages 33–68.

Barndorff-Nielsen, O. E., Shephard, N., and Winkel, M. (2006b). Limit Theorems for Multipower Variation in the Presence of Jumps. *Stochastic Processes and their Applications*, **116**(5), 796–806.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2008). Designing Realized Kernels to Measure the ex−post Variation of Equity Prices in the Presence of Noise. *Econometrica*, **76**, 1481–1536.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2011). Multivariate Realised Kernels: Consistent Positive Semi-Definite Estimators of the Covariation of Equity Prices with Noise and Non-Synchronous Trading. *Journal of Econometrics*, **162**(2), 149–169.

Barunik, J. and Vacha, L. (2015). Realized Wavelet-Based Estimation of Integrated Variance and Jumps in the Presence of Noise. *Quantitative Finance*, **15**(8), 1347–1364.

Bibinger, M. and Winkelmann, L. (2015). Econometrics of Co-Jumps in High-Frequency Data with Noise. *Journal of Econometrics*, **184**(2), 361–378.

Boudt, K. and Zhang, J. (2015). Jump Robust Two Time Scale Covariance Estimation and Realized Volatility Budgets. *Quantitative Finance*, **15**(6), 1041–1054.

Christensen, K. and Podolskij, M. (2007). Realized Range-Based Estimation of Integrated Variance. *Journal of Econometrics*, **141**, 323–349.

Christensen, K., Podolskij, M., and Vetter, M. (2009). Bias-Correcting the Realized Range-Based Variance in the Presence of Market Microstructure Noise. *Finance and Stochastics*, **13**(2), 239–268.

Christensen, K., Oomen, R., and Podolskij, M. (2010). Realised Quantile-Based Estimation of the Integrated Variance. *Journal of Econometrics*, **159**(1), 74–98.

Christensen, K., Oomen, R. C., and Podolskij, M. (2014). Fact or Friction: Jumps at Ultra High Frequency. *Journal of Financial Economics*, **114**(3), 576–599.

Christensen, K., Hounyo, U., and Podolskij, M. (2018). Is the Diurnal Pattern Sufficient to Explain Intraday Variation in Volatility? A Nonparametric Assessment. *Journal of Econometrics*, **205**(2), 336–362.

Cont, R. and Mancini, C. (2011). Nonparametric tests for pathwise properties of semimartingales. *Bernoulli*, **17**(2), 781–813.

Corsi, F., Pirino, D., and Reno, R. (2010). Threshold Bipower Variation and the Impact of Jumps on Volatility Forecasting. *Journal of Econometrics*, **159**(2), 276–288.

Fan, J. and Wang, Y. (2007). Multi-Scale Jump and Volatility Analysis for High-Frequency Financial Data. *Journal of the American Statistical Association*, **102**(480), 1349–1362.

Fan, J., Li, Y., and Yu, K. (2012). Vast Volatility Matrix Estimation Using High-Frequency Data for Portfolio Selection. *Journal of the American Statistical Association*, **107**(497), 412–428.

Feller, W. (1951). The Asymptotic Distribution of the Range of Sums of Independent Random Variables. *The Annals of Mathematical Statistics*, **22**(3), 427–432.

Hansen, P. R. and Lunde, A. (2006). Realized Variance and Market Microstructure Noise. *Journal of Business and Economic Statistics*, **24**(2), 127–161.

Hayashi, T. and Yoshida, N. (2005). On Covariance Estimation of Non-Synchronously Observed Diffusion Processes. *Bernoulli*, **11**, 359–379.

Jacod, J. (2008). Asymptotic Properties of Realized Power Variations and Related Functionals of Semimartingales. *Stochastic Processes and their Applications*, **118**(4), 517–559.

Jacod, J. and Protter, P. (2012). Discretization of Processes. *New York: Springer-Verlag, first edition.*

Jacod, J. and Todorov, V. (2009). Testing for Common Arrivals of Jumps for Discretely Observed Multidimensional Processes. *The Annals of Statistics*, **37**(4), 1792–1838.

Jacod, J. and Todorov, V. (2010). Do Price and Volatility Jump Together? *The Annals of Applied Probability*, **20**(4), 1425–1469.

Jacod, J. and Todorov, V. (2014). Efficient Estimation of Integrated Volatility in Presence of Infinite Variation Jumps. *The Annals of Statistics*, **42**(3), 1029–1069.

Jacod, J., Li, Y., Mykland, P. A., Podolskij, M., and Vetter, M. (2009). Microstructure Noise in the Continuous Case: The Pre-averaging Approach. *Stochastic Processes and their Applications*, **119**(7), 2249–2276.

Jacod, J., Podolskij, M., and Vetter, M. (2010). Limit Theorems for Moving Averages of Discretized Processes plus Noise. *The Annals of Statistics*, **38**(3), 1478–1545.

Jing, B.-Y., Liu, Z., and Kong, X.-B. (2014). On the Estimation of Integrated Volatility With Jumps and Microstructure Noise. *Journal of Business and Economic Statistics*, **32**(3), 457–467.

Kim, D., Wang, Y., and Zou, J. (2016). Asymptotic Theory for Large Volatility Matrix Estimation Based on High-frequency Financial Data. *Stochastic Processes and their Applications*, **126**, 3527–3577.

Koike, Y. (2016). Estimation of Integrated Covariances in the Simultaneous Presence of Nonsynchronicity, Microstructure Noise and Jumps. *Econometric Theory*, **32**(3), 533–611.

Lee, S. S. and Mykland, P. A. (2012). Jumps in Equilibrium Prices and Market Microstructure Noise. *Journal of Econometrics*, **168**(2), 396–406.

Li, J., Todorov, V., Tauchen, G., and Lin, H. (2018). Rank Tests at Jump Events. *Journal of Business and Economic Statistics, preprint*.

Li, Y., Mykland, P. A., Renault, E., and Zhang, L. (2014). Realized Volatility When Sampling Time Are Possibly Endogenous. *Econometric Theory*, **30**, 580–605.

Mancini, C. (2008). Large Deviation Principle for an Estimator of the Diffusion Coefficient in a Jump-diffusion Process. *Statistics and Probability Letters*, **78**(7), 869–879.

Mancini, C. (2009). Non-parametric Threshold Estimation for Models with Stochastic Diffusion Coefficient and Jumps. *Scandinavian Journal of Statistics*, **36**(2), 270–296.

Martens, M. and van Dijk, D. (2007). Meausring Volatility with the realized Range. *Journal of Econometrics*, **138**, 181–207.

Parkinson, M. (1980). The Extreme Value Method for Estimating the Variance of the Rate of Return. *The Journal of Business*, **53**(1), 61–65.

Podolskij, M. and Vetter, M. (2009a). Bipower-type Estimation in a Noisy Diffusion Setting. *Stochastic Processes and their Applications*, **119**(9), 2803–2831.

Podolskij, M. and Vetter, M. (2009b). Estimation of Volatility Functionals in the Simultaneous Presence of Microstructure Noise and Jumps. *Bernoulli*, **15**(3), 634–658.

Robert, C. Y. and Rosenbaum, M. (2010). On the Microstructural Hedging Error. *SIAM Journal on Financial Mathematics*, **1**(1), 427–453.

Robert, C. Y. and Rosenbaum, M. (2012). Volatility and Covariation Estimation When Microstructure Noise and Trading Times are Endogenous. *Mathematical Finance*, **22**(1), 133–164.

Tao, M., Wang, Y., and Zhou, H. H. (2013). Optimal Sparse Volatility Matrix Estimation for High-dimensional Itô Processes with Measurement Errors. *The Annals of Statistics*, **41**(1), 1816–1864.

Varneskov, R. T. (2016). Flat-Top Realized Kernel Estimation of Quadratic Covariation with Nonsynchronous and Noisy Asset Prices. *Journal of Business and Economic Statistics*, **34**, 1–22.

Varneskov, R. T. (2017). Estimating the Quadratic Variation Spectrum of Noisy Asset Prices Using Generalized Flat-top Realized Kernels. *Econometric Theory*, **33**(6), 1457–1501.

Wang, K., Liu, J., and Liu, Z. (2013). Disentangling the Effect of Jumps on Systematic Risk Using a New Estimator of Integrated Co-Volatility. *Journal of Banking and Finance*, **37**(5), 1777–1786.

Wang, Y. and Zou, J. (2010). Vast Volatility Matrix Estimation for High-Frequency Financial Data. *The Annals of Statistics*, **38**(2), 943–978.

Woerner, W. (2006). Power and Multipower Variation: Inference for High Frequency Data. *In Stochastic Finance, A.N. Shiryaev, M. do Rosário Grosshino, P. Oliviera, M. Esquivel, eds., 343-354, Springer*, pages 343–354.

Zhang, B. Y., Zhou, H., and Zhu, H. (2009). Explaining Credit Default Swap Spreads with the Equity Volatility and Jump Risks of Individual Firms. *The Review of Financial Studies*, **22**(12), 5099–5131.

Zhang, L. (2006). Efficient Estimation of Stochastic Volatility Using Noisy Observations: a Multi-Scale Approach. *Bernoulli*, **12**(6), 1019–1043.

Zhang, L., Mykland, P. A., and Aït-Sahalia, Y. (2005). A Tale of Two Time Scales: Determining Integrated Volatility With Noisy High-Frequency Data. *Journal of the American Statistical Association*, **100**(472), 1394–1411.