Power Variations and Testing for Co-Jumps: The Small Noise Approach

DAISUKE KURISU
Graduate School of Economics, University of Tokyo

ABSTRACT. In this paper, we study the effects of noise on bipower variation, realized volatility (RV) and testing for co-jumps in high-frequency data under the small noise framework. We first establish asymptotic properties of bipower variation in this framework. In the presence of the small noise, RV is asymptotically biased, and the additional asymptotic conditional variance term appears in its limit distribution. We also propose consistent estimators for the asymptotic variances of RV. Second, we derive the asymptotic distribution of the test statistic proposed in (Ann. Stat. 37, 1792-1838) under the presence of small noise for testing the presence of co-jumps in a two-dimensional Itô semimartingale. In contrast to the setting in (Ann. Stat. 37, 1792-1838), we show that the additional asymptotic variance terms appear and propose consistent estimators for the asymptotic variances in order to make the test feasible. Simulation experiments show that our asymptotic results give reasonable approximations in the finite sample cases.

Key words: bipower variation, co-jump test, high-frequency financial data, market microstructure noise, small noise asymptotics, realized volatility

1. Introduction

Recently, considerable attention has been paid to estimation and testing for underlying continuous time stochastic processes based on high-frequency financial data. In the analysis of high-frequency data, it is important to take into account the influence of market microstructure noise. The market microstructure noise captures a variety of frictions inherent in trading processes such as bid-ask bounce and discreteness of price changes Aït-Sahalia & Yu (2009). There are a large number of papers on the high-frequency data analysis in the presence of noise. For example, Zhang et al. (2005), Bandi & Russel (2006) and Bibinger & Reiß (2014) investigate the case when the log-price follows a diffusion process observed with an additive noise. They assume that the size of noise does not depend on the observation frequency. To be precise, they assume that observed log-prices are of the form

\[ Y_{t_i}^{(m)} = X_{t_i}^{(m)} + v_i^{(m)}, \quad i = 1, \ldots, n, \quad m = 1, \ldots, d, \]  

where \( X_t = (X_t^{(1)}, \ldots, X_t^{(d)})^\top \) is the underlying \( d \)-dimensional continuous time log-price process and \( v_i = (v_i^{(1)}, \ldots, v_i^{(d)})^\top \) are \( d \)-dimensional i.i.d. random noise vectors of which each component has mean 0 and constant variance independent of the process \( X_t \). Zhang et al. (2005) and Bandi & Russel (2006) study the one-dimensional \((d = 1)\) case, and Bibinger & Reiß (2014) study the multi-dimensional case. Intuitively, the assumption of the constant noise variance means that the noise is dominant to the log-price when the observation frequency increases. In this paper, we instead assume that the effect of noise depends on the frequency of the observation, and the observed log-prices are of the form

\[ Y_{t_i}^{(m)} = X_{t_i}^{(m)} + \epsilon_{n,m} v_i^{(m)}, \quad i = 1, \ldots, n, \quad m = 1, \ldots, d, \]
where $\epsilon_n = (\epsilon_{n,1}, \ldots, \epsilon_{n,d})^\top$ is a $d$-dimensional non-stochastic sequence satisfying $\epsilon_{n,m} \downarrow 0$ as $n \to \infty$ for each $m$. We call this assumption small noise. Under the small noise assumption, the noise is vanishing as the observation frequency increases. Hence, the small noise assumption is interpreted as an intermediate assumption between the no noise assumption and the constant noise variance assumption. Related literature that considers small noise includes Gloter & Jacod (2001a, 2001b), Barndorff-Nielsen et al. (2008), Li et al. (2016), Li et al. (2015) and among others. Hansen & Lunde (2006) give an empirical evidence that the market microstructure noise is small. The analysis in the present paper is also related to those of Rosenbaum (2011), Bacry et al. (2013a, 2013b) and Bacry & Muzy (2014). Under the constant noise assumption, the signature plot that is defined by the realized volatility (RV) explodes as $\Delta \downarrow 0$. This is not completely consistent with the behavior of the RV as they observed in a fine scale of a few seconds. On the other hand, under the small noise assumption, the signature plot could increase as the observation frequency increases, however, it does not explode because the order of the variance of the noise is the same as (or smaller than) that of quadratic variation of a latent process. Those papers as we mentioned earlier consider the situation that the signature plot increases as $\Delta \downarrow 0$ but not explode. This is very reasonable in practice as discussed in those papers. In particular, Rosenbaum (2011) study the effect of market microstructure noise by using a noise index proposed in the paper. The simulation results in the paper shows that under the assumption of the additive small noise the signature plot reproduce the results of empirical data well. Bacry et al. (2013a, 2013b) and Bacry & Muzy (2014) use marked Hawkes process for modeling microstructure noise and other stylized facts of market prices microstructure such as Epps effect and lead–lag effect. Our approach can also be seen as an alternative approach of their point process modeling of asset prices.

The first purpose of the paper is to investigate the effect of small noise on bipower variation (BPV) proposed in Barndorff-Nielsen & Shephard (2004, 2006) and the estimation of the integrated volatility. We establish the asymptotic properties of BPV when the latent process $(X_t)_{0 \leq t \leq 1}$ is a one-dimensional Itô semimartingale. We also propose procedures to estimate integrated volatility using RV and the asymptotic conditional variances which appear in the limit distribution of RV under the small noise assumption. In contrast to the no noise model, RV is asymptotically biased and an additional asymptotic conditional variance term appears in its limit distribution (Bandi & Russel (2006), Hansen & Lunde (2006), Li et al. (2015), Kunitomo & Kurisu (2017) and among others). In the recent related literature, Li et al. (2015) proposed the unified approach for estimating the integrated volatility of a diffusion process when both small noise and asymptotically vanishing rounding error are present. In this paper, we only consider the additive noise, but we assume that the log-price process $(X_t)_{0 \leq t \leq 1}$ is a $d$-dimensional Itô semimartingale that includes a diffusion process as a special case.

The second purpose of this paper is to propose a procedure to test the existence of co-jumps in two log-prices when the small noise exists. Examining whether two asset prices have contemporaneous jumps (co-jumps) or not is one of the effective approaches toward distinguishing between systematic and idiosyncratic jumps of asset prices and also important in option pricing and risk management. From the empirical side, Glider et al. (2014) investigate co-jumps and give a strong evidence for the presence of co-jumps. Bollerslev et al. (2013) provide another empirical evidence for the dependence in the extreme tails of the distribution governing jumps of two stock prices. In spite of the importance of this problem, a testing procedure for co-jumps is not sufficiently studied. Jacod & Todorov (2009) is the seminal paper in this literature and other important contributions include Mancini & Gobbi (2012) and Bibinger & Winkelmann (2015). Mancini & Gobbi (2012) study the
no noise model. Bibinger & Winkelmann (2015) is a recent important contribution to co-jump test for the model (1). Their co-jump test is based on the wild bootstrap-type approach and for testing the null hypothesis that observed two log-prices have no co-jumps. On the other hand, a great variety of testing methods for detecting the presence of jumps in the one-dimensional case have been developed. For example, Barndorff-Nielsen & Shephard (2006), Fan & Wang (2007), Jiang & Oomen (2008), Bollerslev et al. (2008), Jacod (2008), Mancini (2009) and Aït-Sahalia & Jacod (2009) for the no noise model, and Aït-Sahalia et al. (2012) and Li (2013) for the model (1) with, conditionally on $X$, mutually independent noise.

Our idea of estimating integrated volatility and the asymptotic conditional variance of $RV$ is based on the separated information maximum likelihood (SIML) method developed in Kunitomo & Sato (2013) for correcting the bias of $RV$ and the truncation method developed in Mancini (2009). For a construction of a co-jump test, we assume that the process $(X_t)_{0 \leq t \leq 1}$ in the model (2) is a two-dimensional Itô semimartingale and investigate the asymptotic properties of the test statistic proposed in Jacod & Todorov (2009). We show that, because of the presence of the small noise, the asymptotic distribution of the test statistic is different from their result. In fact, the additional asymptotic conditional variance appears in its limit distribution. We develop a fully data-driven procedure to estimate the asymptotic variance of the test statistics based on similar technique used in the estimation of integrated volatility and asymptotic variance of $RV$.

From a technical point of view, our results in the present paper are non-trivial applications of developed in the literature of high-frequency data analysis. The analysis of the present paper has some connections to that of Kunitomo & Kurisu (2017) that study the effects of small noise on the asymptotic distribution of the $RV$ and some existing test statistics proposed for the statistical test of the presence of jumps in no noise case. They investigate one-dimensional case and do not study the consistent estimation of the asymptotic conditional variances that appear in the asymptotic distributions of the test statistics. For the consistent estimation of the asymptotic conditional variance of the $RV$ or co-jump test statistics proposed in the present paper, we have to estimate the local volatility for some $t \geq 0$ or some functionals of volatility process. For this, we can use truncated functionals in no noise case for example. However, in the presence of the small noise, the truncation method is not sufficient to distinguish the continuous part of $X$ with the noise because the size of noise gets smaller as the observation frequency increases. Therefore, we have to evaluate and correct the bias which comes from the presence of noise.

The numerical experiments show that our proposed method gives a good approximation of the limit distribution of $RV$ and reasonable result for the estimation of integrated volatility. Our proposed testing procedure of co-jumps also improves the empirical size in the presence of noise compared with the test in Jacod & Todorov (2009).

This paper is organized as follows. In section 2, we describe the theoretical settings of the underlying Itô semimartingale and market microstructure noise. In section 3, we investigate the effects of noise on the asymptotic properties of $BPV$, and give some comments on the stable limit theorems of $RV$. We also propose an estimation method of the integrated volatility in the small noise framework. In section 4, we study statistics related to the detection of co-jumps in the two-dimensional setting when the noise satisfy the small noise assumption. Then, we propose a testing procedure for the presence of co-jumps. In section 5, we give estimation methods of asymptotic conditional variances which appear in the limit theorems of $RV$ and co-jump test statistic studied in the previous sections. We report some simulation results in section 6, and we give some concluding remarks in section 7. Proofs are collected in appendix A.
2. Setting

We consider a continuous-time financial market in a fixed terminal time \( T \). We set \( T = 1 \) without loss of generality. The underlying log-price is a \( d \)-dimensional Itô semimartingale. We observe the log-price process in high-frequency contaminated by the market microstructure noise.

Let a first filtered probability space be \( (\Omega^{(0)}, \mathcal{F}^{(0)}, (\mathcal{F}_t^{(0)})_{t \geq 0}, P^{(0)}) \) on which the \( d \)-dimensional Itô semimartingale \( X = (X_t)_{0 \leq t \leq 1} \) is defined. We adopt the construction of the whole filtered probability space \( (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0,1]}, P) \) where both the process \( X \) and the noise is defined as in Christensen et al. (2013). Let \( \mathcal{B}^d \) be the Borel \( \sigma \)-field of \( \mathbb{R}^d \) and \( Q \) be a probability measure on \((\mathbb{R}^d, \mathcal{B}^d)\). We consider a second filtered probability space \((\Omega^{(1)}, \mathcal{F}^{(1)}, (\mathcal{F}_t^{(1)})_{t \in [0,1]}, P^{(1)})\), where \( \Omega^{(1)} \) is the set of functions from \([0,1]\) to \( \mathbb{R}^d \), \( \mathcal{F}^{(1)} \) is the Borel \( \sigma \)-field on \( \Omega^{(1)} \), and \( P^{(1)} = \otimes_{t \in [0,1]} P_t \) with \( P_t = Q \). Define a process \( \tilde{v} = (\tilde{v}_t)_{t \in [0,1]} \) as the canonical process on \( \Omega^{(1)} \), and \( V = (\mathcal{F}_t^{(1)})_{t \in [0,1]} \) where \( \mathcal{F}_t^{(1)} \) is the marginal distribution of \( \tilde{v}_t \). We set \( t_i = \tilde{v}^{(1)}_{n,i} \) for the discrete time points \( 0 = t_0^n < t_1^n < \cdots < t_n^n \leq 1 \) and consider the filtered space \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0,1]}, P)\) where \( \Omega = \Omega^{(0)} \times \Omega^{(1)}, \mathcal{F} = \mathcal{F}^{(0)} \times \mathcal{F}^{(1)}, \mathcal{F}_t = \mathcal{F}_t^{(0)} \otimes \mathcal{F}_t^{(1)} \) and \( P = P^{(0)} \times P^{(1)} \).

We consider the following model for the observed log-price at \( t_i^n \in [0,1] \)

\[
Y^{(m)}_{t_i^n} = X^{(m)}_{t_i^n} + \epsilon_{n,m} v_i^{(m)}, \quad i = 1, \ldots, n, \quad m = 1, \ldots, d,
\]

where \( v_i = (v_i^{(1)}, \ldots, v_i^{(d)})^\top \) are \( d \)-dimensional i.i.d. random noise and noise coefficient \( \epsilon_{n,m} = (\epsilon_{1,n,\ldots,d,n})^\top \) is a sequence of \( d \)-dimensional vector that depends on sample size \( n \) and for each \( m, \epsilon_{m,n} \rightarrow 0 \). We assume that these terms satisfy assumptions 2 and 3 described in the succeeding text. Moreover, let \( X = (X_t)_{0 \leq t \leq 1} \) be an Itô semimartingale of the form

\[
X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + \int_0^t \int_{\mathbb{R}^d} \kappa \circ \delta(s,x)(\mu - \nu)(ds,dx) + \int_0^t \int_{\mathbb{R}^d} \kappa' \circ \delta(s,x)\mu(ds,dx),
\]

where \( (W_s) \) is a \( d' \)-dimensional standard Brownian motion, \( (b_s) \) is a \( d \)-dimensional adapted process, \( \sigma_s \) is a \((d \times d')\)-(instantaneous) predictable volatility process, and we define the process \( c = \sigma \sigma^\top \). Furthermore, \( \delta(s,x) \) is a predictable function on \( \Omega \times \mathbb{R}_+ \times \mathbb{R}^d \), \( \kappa : \mathbb{R}^d \rightarrow \mathbb{R}^d \) is a continuous truncation function with compact support and \( \kappa'(x) = x - \kappa(x), \mu(\cdot) \) is a Poisson random measure on \( \mathbb{R}_+ \times \mathbb{R}^d \), and \( v(ds,dz) = ds \otimes \lambda(dz) \) is a predictable compensator or intensity measure of \( \mu \) with a \( \sigma \)-finite measure \( \lambda \). We partially follow the notation used in Jacod & Protter (2012). We assume that the observed times \( 0 = t_0^n < t_1^n < \cdots < t_n^n = 1 \) are such that \( t_i^n - t_{i-1}^n = 1/n = \Delta_n \). When \( d = 2 \) (bivariate case), let

\[
\Delta X_t = X_t - X_{t-}, \quad \tau = \inf\{t : \Delta X_t^{(1)} \Delta X_t^{(2)} \neq 0\},
\]

\[
\tilde{\Gamma} = \{(\omega, t, x) : \delta^1(\omega, t, x) \delta^2(\omega, t, x) \neq 0\},
\]

and for \( i = 1, 2 \), define

\[
\delta_{\tau^i}(\omega) = \begin{cases} \int_{\mathbb{R}^d} \kappa(\omega, t, x) \lambda(dx) & \text{if the integral is well-defined}, \\ +\infty & \text{otherwise}. \end{cases}
\]

We also make the following assumptions.
Assumption 1.

(i) The path \( t \mapsto b_t(\omega) \) is locally bounded.

(ii) The process \( \sigma \) is continuous.

(iii) We have \( \sup_{\omega,x} \| \delta(\omega,t,x) \| / \gamma(x) \) is locally bounded for a deterministic non-negative function satisfying \( \int_{\mathbb{R}^d} (\gamma(x)^h \wedge 1) \lambda(dx) < \infty. \) for some \( h \in (0,2) \) \((\| \cdot \| \) denotes the Euclidean norm on \( \mathbb{R}^d \)).

(iv) For each \( \omega \) and \( i = 1, 2 \), the path \( t \mapsto \delta_t^i(\omega) \) is locally bounded on the interval \( [0, \tau(\omega)) \).

(v) We have \( \int_t^{t+u} \| \sigma_s \| ds > 0 \) a.s. for all \( t, u > 0 \).

We call the second and third terms in the right hand side as drift part and continuous part of \( X \), respectively. If \( X \) does not have the last two terms of the right hand side of (3) (these are jump parts of \( X \)), then we say that \( X \) is continuous. Otherwise, we say that \( X \) is discontinuous.

For the noise term, we assume the following conditions.

Assumption 2. There exist some \( q \geq 0 \) and \( \zeta_m > 0, 1 \leq m \leq d \) such that

\[ n \epsilon_{m,n}^2 = \zeta_m n^{-2q} + O \left( n^{-(1+2q)} \right). \]

Assumption 3. \( \{v_i\}_{i=1}^{\infty} \) is a sequence of i.i.d. \( d \)-dimensional standard normal random variables.

When \( q = 0 \), Assumption 2 coincides with the small noise assumption in Kunitomo & Kurisu (2017). If the noise coefficient does not depend on the sampling scheme, that is, for each component there exist some positive constants \( \epsilon_m \) such that \( \text{Var}(v_{m,n}^{(m)}) = \epsilon_m^2 \) in the model (2), then the effect of noise is asymptotically dominant. This case corresponds to the assumption that the variance of noise is constant. Assumption 2 means that the effect of noise depends on a sample number \( n \). Hence, the effect of noise gets smaller if the observation frequency increases.

2.1. Discussion on assumption 2

We discuss the relevance of the assumption 2 and refer to the papers which are related to our small noise assumption and mentioned in section 1. The small noise assumption has been considered in not a few number of papers. For example, Aït-Sahalia et al. (2005) consider in section 9.2 the case when the variance contributed by noise to a observed log-return is of the same order as the latent return. In Zhang et al. (2005), the size of additive noise is considered to be \( 0.0005 \). This means that if \( n = 23,400 \), which corresponds to second by second prices of stocks traded between 9:30:00 and 16:00:00, \( n \epsilon_n^2 \) (around 0.00585) is still quite small and hence, assumption 2 seems not unrealistic. Hansen & Lunde (2006) investigate the properties of market microstructure noise and its effects on the estimation of integrated volatility based on the RV. As reported in the summary and concluding remarks in the paper, their empirical study of the stock returns for 30 equities of the Dow Jones Industrial Average shows that the noise is very small. Li et al. (2015) and Li et al. (2016), which are considering the small noise as a market microstructure noise, are recent contribution in the literature of high-frequency financial econometrics in both theoretical and empirical point of view. Li et al. (2015) consider the estimation of integrated volatility when the latent price is contaminated by i.i.d. normal small noise with asymptotically vanishing rounding error, and they analyzed New York Stock Exchange trade and quote data for four stocks: Citigroup Inc., Intel Corp., Bank of America and Microsoft Corp. which are all actively traded. Li et al. (2016) consider the efficient estimation of integrated volatility based on the RV-type estimator for the one-dimensional model.
\[ Y_{tn}^i = X_{tn}^i + g(Z_{tn}^i; \theta_0) + v_t^i, \quad i = 1, \ldots, n. \]  

where \( g \) is a parametric function with parameters \( \theta_0 \) which partially models the market microstructure, \( Z_{tn}^i \) is the information set which can include trade type, trading volume and bid-ask-bounds, and \((v_t^i)_{1 \leq t \leq n}\) are i.i.d. noise with mean 0 and variance \( \epsilon_n^2 = O(1/n) \) (small noise), and these noise can be considered as the noise which cannot be explained by the past trading information. If we estimate \( \theta_0 \) by \( \hat{\theta}_0 \) that proposed in their paper, the model (4) reduced to the following model:

\[ Y_{tn}^i - g(Z_{tn}^i; \hat{\theta}_0) = X_{tn}^i + v_t^i, \quad i = 1, \ldots, n. \]

Regarding \( Y_{tn}^i - g(Z_{tn}^i; \hat{\theta}_0) \) as observed log-prices, this model corresponds to our model (2). In this model, we can perform our proposed methods for the estimation of integrated volatility and for testing the presence of co-jumps explained in the following sections. As an empirical study, they analyzed New York Stock Exchange trade and quote data for four stocks: Arch Coal Inc., Dell Inc., EMC Corp and General Electric Co. Because empirical studies in these papers, we referred here shows the relevance of the small noise assumption, our assumption on the noise also appear to be reasonable on empirical side.

3. The effects of small noise on bipower variation and realized volatility

In this section, we assume that the process \((X_t)_{0 \leq t \leq 1}\) is one-dimensional \((d = 1)\), and give the asymptotic properties of BPV and give some remarks on the problem of an estimation of integrated volatility and the asymptotic conditional variance of RV under the presence of small noise.

3.1. Asymptotic properties of bipower variation and realized volatility

Bipower variation and RV are often used for estimating integrated volatility. We give some results on asymptotic properties of BPV and RV. Let \( \Delta_t^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n} \) and define the following statistics:

\[
V_p^{(n)}(X) = \sum_{i=1}^{n} |\Delta_t^n X|^p, \\
V_{r,s}^{(n)}(X) = \sum_{i=1}^{n-1} |\Delta_t^n X|^r |\Delta_{i+1}^n X|^s.
\]

According to the above definition, \( \tilde{V}_2^{(n)}(X) \) is the realized volatility (RV(X)) and \( \tilde{V}_{1,1}^{(n)}(X) \) is the bipower variation (BPV(X)). First, we give asymptotic properties of above statistics. The following result describes the effect of small noise.

**Proposition 1.** Suppose Assumptions 1, 2 and 3 are satisfied. Let \( r, s \) and \( k \) be positive integers, then

\[
V_{2r}^{(n)}(Y) - \tilde{V}_{2r}^{(n)}(X) = O_P(n^{1-(r+q)}), \\
V_{r,s}^{(n)}(Y) - \tilde{V}_{r,s}^{(n)}(X) = O_P(n^{1-(r+s)/2-q}).
\]

In the following results, we freely use the stable convergence arguments and \( \mathcal{F}^{(0)} \)-conditionally Gaussianity, which have been developed and explained by Jacod (2008) and
Jacro & Protter (2012), and use the notation $\xrightarrow{\mathcal{L}}$ as stable convergence in law. For the general reference on stable convergence, we also refer to Häusler & Luschgy (2015). The following proposition on the asymptotic properties of BPV describes the case when the effect of noise is asymptotically negligible. The result follows immediately from proposition 1.

**Proposition 2.** Suppose Assumptions 1, 2, and 3 are satisfied. Let $X$ be continuous, and $r$ and $s$ be positive integers such that $(r+s)/2$ is integral. If $q > 0$, then we have the following convergence in probability:

$$n^{(r+s)/2-1}\tilde{V}_{r,s}^{(n)}(Y) \xrightarrow{p} m_rm_s \int_0^1 \sigma_u^{r+s} du,$$

and if $q > 1/2$, then we have the following stable convergence in law:

$$\sqrt{n} \left(n^{(r+s)/2-1}\tilde{V}_{r,s}^{(n)}(Y) - m_rm_s \int_0^1 \sigma_u^{r+s} du\right) \xrightarrow{\mathcal{L}} U,$$

where $m_r = 2^{r/2}\Gamma((r+1)/2)\Gamma(2^{-1})/\Gamma(r)$ and $U$ is $\mathcal{F}^{(0)}$-conditionally Gaussian with mean 0 and $\mathcal{F}^{(0)}$-conditional variance $E[U^2 | \mathcal{F}^{(0)}] = (m_2m_{2s} + 2m_rm_sm_{r+s} - 3m_2^2m_s^2)\int_0^1 \sigma_u^{2(r+s)} du$.

From the remark of theorem 2.5 in Barndorff-Nielsen et al. (2006), if $X$ is continuous, then

$$\sqrt{n} \left(BPV(X) - m_1^2 \int_0^1 \sigma_s^2 ds\right) \xrightarrow{\mathcal{L}} \tilde{U},$$

where $\tilde{U}$ has the same distribution as $U$ in proposition 2 (replacing $r,s = 1$). The latter part of proposition 2 implies that if $q > 1/2$, we can replace $\tilde{V}_{1,1}^{(n)}(X)$ as $\tilde{V}_{1,1}^{(n)}(Y)$. In such a case, we can use BPV as the consistent estimator of integrated volatility.

Next, we consider the asymptotic properties of RV. When the underlying process $X$ is continuous, the RV($X$) is often used for estimating integrated volatility in the no noise case. In this case, to construct a confidence interval of integrated volatility or construct a jump test proposed in Barndorff-Nielsen & Shephard (2006) for example, we must consistently estimate the asymptotic conditional variance of the limit distribution of RV and jump test statistics. In place of BPV, multipower variation (MPV) is often used for estimating volatility functionals of RV, multipower variation (MPV) is often used for estimating volatility functionals. In place of BPV, multipower variation (MPV) is often used for estimating volatility functionals.

$$\Delta_n^{-1} \sum_{i=1}^{n-3} |\Delta_{i+3}^n Y| |\Delta_{i+2}^n Y| |\Delta_{i+1}^n Y| |\Delta_i^n Y|$$

$$= \Delta_n^{-1} \sum_{i=1}^{n-3} |\Delta_{i+3}^n X||\Delta_{i+2}^n X||\Delta_{i+1}^n X||\Delta_i^n X| + O_P(n^{-q}).$$

Therefore, when the asymptotic order of noise is sufficiently higher ($q > 0$), MPV is a consistent estimator of $\int_0^1 \sigma_s^4 ds$. In this case, we can use the same procedure as that of the no noise case.

If small noise satisfy the condition $q = 0$ in assumption 2, the effect of noise cannot be ignored. From the first part of proposition 1, if $q = 0$ and $r = 1$, then $\text{RV}(Y) - \text{RV}(X) =$
OP (1). Kunitomo & Kurisu (2017) proved that RV is asymptotically biased and derived the following two stable convergence results under assumptions 1, 2 and 3 with \( q = 0 \) in Assumption 2. If \( X \) is continuous, then

\[
RV(Y) \overset{p}{\longrightarrow} \int_0^1 \sigma_s^2 ds + 2\zeta_1 \equiv U_{0,1},
\]

\[
\sqrt{n}(RV(Y) - U_{0,1}) \xrightarrow{D} U_1 + U_2 + U_3,
\]

where \( \zeta_1 > 0 \) is a constant which appears in the assumption 2, \( U_j \) for \( j = 1, 2, 3 \) are \( \mathcal{F}^{(0)} \)-conditionally mutually independent Gaussian random variables with mean 0 and \( \mathcal{F}^{(0)} \)-conditional variances \( E(U_1^2 | \mathcal{F}^{(0)}) = 2 \int_0^1 \sigma_s^4 ds \), \( E(U_2^2 | \mathcal{F}^{(0)}) = 8\zeta_1 \int_0^1 \sigma_s^2 ds \) and \( E(U_3^2 | \mathcal{F}^{(0)}) = 12\zeta_1^2 \).

If \( X \) is the Itô semimartingale of the form (3), then

\[
RV(Y) \overset{p}{\longrightarrow} \int_0^1 \sigma_s^2 ds + \sum_{0 \leq s \leq 1} (\Delta X_s)^2 + 2\zeta_1 \equiv U_{0,2},
\]

\[
\sqrt{n}(RV(Y) - U_{0,2}) \xrightarrow{D} U_1 + U_2,
\]

where \( U_j \) for \( j = 1, 2 \) are \( \mathcal{F}^{(0)} \)-conditionally mutually independent Gaussian random variables with mean 0 and \( \mathcal{F}^{(0)} \)-conditional variances

\[
E(U_1^2 | \mathcal{F}^{(0)}) = 2 \int_0^1 \sigma_s^4 ds + 4 \sum_{0 \leq s \leq 1} \sigma_s^2 (\Delta X_s)^2 \quad \text{and} \quad E(U_2^2 | \mathcal{F}^{(0)}) = 8\zeta_1 \left[ \int_0^1 \sigma_s^2 ds + \sum_{0 \leq s \leq 1} (\Delta X_s)^2 \right].
\]

To construct a feasible procedure to estimate RV, we need to estimate the noise parameter \( \zeta_1 \) and the asymptotic conditional variance of its limit distribution. In section 5, we construct estimators of the asymptotic conditional variances of RV.

### 3.2. Estimation of integrated volatility under small noise

In the model (1) with constant noise variance, it is well known that the variance of noise can be estimated by \( (2n)^{-1} \sum_{i=1}^n (\Delta_i^q Y)^2 \overset{p}{\longrightarrow} \text{Var}(v_1) \). However, under the small noise assumption, for example \( q = 0 \), this estimation does not work. In fact for the small noise case, \( (2n)^{-1} \sum_{i=1}^n (\Delta_i^q Y)^2 \overset{p}{\longrightarrow} 0 \) regardless of the value of \( \zeta_1 \). Thus, we must consider another procedure for estimating \( \zeta_1 \). The separated information maximum likelihood (SIML) method investigated in Kunitomo & Sato (2013) can be used to estimate \( \zeta_1 \). The SIML estimator is a consistent estimator of quadratic variation \( [X, X] \) of the process \( X \) under both models (1) and (2). Therefore, under assumptions 1, 2 and 3, we have

\[
[X, X]_{\text{SIML}} \overset{p}{\longrightarrow} [X, X].
\]

where \([X, X]_{\text{SIML}}\) is the SIML estimator discussed in section 5. From (5), (6) and (7), we obtain

\[
\hat{\zeta}_1 = \frac{1}{2} \left( \sum_{i=1}^n (\Delta_i^q Y)^2 - [X, X]_{\text{SIML}} \right) \overset{p}{\longrightarrow} \zeta_1.
\]

We consider two types of the truncated version RV:
where $\alpha > 0$ and $\theta \in (0, 1/2)$. When $X$ have jumps, we can estimate the jump part of quadratic variation by TRVJ when assumptions 1, 2 and 3 are satisfied:

$$\text{TRVJ} \xrightarrow{p} \sum_{0 \leq s \leq 1} (\Delta X_s)^2.$$

Then, integrated volatility (IV) can be estimated consistently:

**Proposition 3.** Suppose assumptions 1, 2, and 3 are satisfied. Then

$$\hat{\text{IV}} = [\hat{X}, \hat{X}]_{\text{SIML}} - \text{TRVJ} \xrightarrow{p} \int_0^1 \sigma_s^2 ds.$$

Proposition 3 implies that $\hat{\text{IV}}$ is robust to jumps and small noise. In particular, if $q = 0$ in assumption 2, then we can rewrite $\hat{\text{IV}} = [\hat{X}, \hat{X}]_{\text{SIML}} - \text{TRVJ} = \text{TRVC} - 2\hat{\gamma}_1$. Therefore, from the remark in section 3.1 and (8), we can estimate integrated volatility by the bias correction of RV.

In this section, we considered only one-dimensional case, however, an extension to the multivariate case is straightforward for the estimation of the covolatility,

$$c^{(p, q)}_{s} \sigma_s^2 ds$$

where $c^{(p, q)}_{s}$ is the $(p, q)$ component of the volatility process $c_s$ defined in section 2.

### 4. Co-jump test under small noise

One of the interesting problems in high-frequency financial econometrics is whether two asset prices have co-jumps or not. To the best of our knowledge, none of the existing literature has so far proposed a co-jump test in the small noise framework. In this section, we consider two dimensional case $X_t = (X^{(1)}_t, X^{(2)}_t)^T_{0 < t \leq 1}$ and propose a testing procedure to detect the existence of co-jumps for discretely observed processes contaminated by small noise. For this purpose, we study the asymptotic property of the following test statistic proposed in Jacod & Todorov (2009):

$$T^{(n)} = \frac{S^{(n)}_{2, 2, 2}(Y)}{S^{(n)}_{1, 2, 2}(Y)}, \quad S^{(n)}_{k, r, s}(Y) = \sum_{i = 1}^{[n/k]} (\Delta_i^n Y^{(k, 1)})^r (\Delta_i^n Y^{(k, 2)})^s,$$

where $\Delta_i^n Y^{(k, l)} = Y^{(l)}_{ik\Delta_n} - Y^{(l)}_{(i-1)k\Delta_n}$ for $k \geq 2, l = 1, 2$, and $[x]$ is the integer part of $x \in \mathbb{R}$. To describe our result, we first decompose the sample space $\Omega$ into three disjoint sets

$$\begin{align*}
\Omega^{(j)} &= \{ \omega : \text{on } [0, 1] \text{ the process } \Delta X^{(1)}_s \Delta X^{(2)}_s \text{ is not identically } 0 \}, \\
\Omega^{(d)} &= \{ \omega : \text{on } [0, 1] \text{ the processes } \Delta X^{(1)}_s \text{ and } \Delta X^{(2)}_s \text{ are not identically } 0, \\
&\quad \text{but the process } \Delta X^{(1)}_s \Delta X^{(2)}_s \text{ is} \}, \\
\Omega^{(c)} &= \{ \omega : \text{on } [0, 1] \text{ } X^{(1)} \text{ and } X^{(2)} \text{ is continuous} \}.
\end{align*}$$

Kunitomo & Kurisu (2017) investigate the effects of small noise on the asymptotic properties of the testing procedure for the presence of jumps developed in Aït-Sahalia & Jacod (2009).
Although Kunitomo & Kurisu (2017) only derived the limiting distribution of the test statistics in the presence of small noise, we can estimate the asymptotic conditional variance in the same way as the estimation of $\zeta_m$, $D_{p,q}^{(l,m)}(r,s)$ and $J^{(l,m)}(r,s)$ which are discussed in section 5. Therefore, we can extend the testing procedure in the case when the underlying log-price process $(X_t)_{0 \leq t \leq 1}$ is contaminated by small noise. By using the procedure, we can decide in principle whether we are on $\Omega^{(c)}$ or not as a first step. If the conclusion is that we are not on $\Omega^{(c)}$, we need to decide that we are on $\Omega^{(j)}$ or $\Omega^{(d)}$ as a second step. The second step is the main purpose of this section.

We test the null hypothesis $H_0$ that observed two log-prices have co-jumps, that is, we are on $\Omega^{(j)}$ against the alternative hypothesis $H_1$ that observed two log-prices have no co-jumps but each log-price have jumps, that is, we are on $\Omega^{(d)}$.

We first provide the asymptotic property of $T(n)$ under the null hypothesis. For this purpose, we consider the asymptotic property of $S_{k,2,2}^{(n)}(X)$. Evaluating the discretization error of the process, in restriction to the set $\Omega^{(j)}$, for $k = 1, 2$, we have

$$
S_{k,2,2}^{(n)}(Y) = S_{k,2,2}^{(n)}(X) + 2R_{k,2,1}^{(n)}(X) + 2R_{k,1,2}^{(n)}(X) + o_P(1/\sqrt{n}).
$$

where

$$
R_{k,r,s}^{(n)}(X) = \epsilon_{r \wedge s,n} \sum_{i=1}^{[n/k]} (\Delta_i^{n} X^{(k,1)})^r (\Delta_i^{n} X^{(k,2)})^s (\Delta v_{i}^{(k,r \wedge s)}),
$$

$$
\Delta v_{i}^{(k,l)} = \sum_{j=k(i-1)+1}^{ki} (v_{ij}^{(l)} - v_{j-1}^{(l)}) = v_{k(i)}^{(l)} - v_{k(i-1)}^{(l)},
$$

where $r \wedge s = \min(r,s)$. Refer to the proof of theorem 1 in appendix A for details. Then it suffices to evaluate the three terms $S_{k,2,2}^{(n)}(X)$, $R_{k,2,1}^{(n)}(X)$ and $R_{k,1,2}^{(n)}(X)$. We also have, in restriction to the set $\Omega^{(j)}$,

$$
S_{k,2,2}^{(n)}(X) \overset{p}{\rightarrow} \sum_{0 \leq s \leq 1} (\Delta X_s^{(1)})^2 (\Delta X_s^{(2)})^2 \equiv S_0,
$$

and $R_{k,2,1}^{(n)}(X) = R_{k,1,2}^{(n)}(X) = O_P(1/\sqrt{n})$. Finally, we obtain the joint stable convergence of these three terms:

$$
\sqrt{n}(S_{k,2,2}^{(n)}(X) - S_0, R_{k,2,1}^{(n)}(X), R_{k,1,2}^{(n)}(X)) \overset{\mathcal{L}-s}{\longrightarrow} (U_{1,1}^{(1,2)}, U_{2,1}^{(1,2)}, U_{3,1}^{(1,2)}),
$$

where $U_{1,1}^{(l,2)}$, $U_{l}^{(1,2)}$ for $l = 2, 3$ are $\mathcal{F}^{(0)}$-conditionally mutually independent Gaussian random variables with mean 0 and the following variances:

$$
E[(U_{1,1}^{(1,2)})^2 | \mathcal{F}^{(0)}] = kF_{2,2},
$$

$$
E[(U_{2,1}^{(1,2)})^2 | \mathcal{F}^{(0)}] = 8\xi_2 \sum_{0 \leq s \leq 1} (\Delta X_s^{(1)})^4 (\Delta X_s^{(2)})^2,
$$

$$
E[(U_{3,1}^{(1,2)})^2 | \mathcal{F}^{(0)}] = 8\xi_1 \sum_{0 \leq s \leq 1} (\Delta X_s^{(1)})^2 (\Delta X_s^{(2)})^4,
$$

and where

© 2017 Board of the Foundation of the Scandinavian Journal of Statistics.
\[
F_{r,s} = \sum_{0 \leq u \leq 1} (r^2 c_{u}^{(1,1)} (\Delta X_{u}^{(1)})^{2(r - 1)} (\Delta X_{u}^{(2)})^{2s} + 2 r s c_{u}^{(1,2)} (\Delta X_{u}^{(1)})^{2r - 1} (\Delta X_{u}^{(2)})^{2s - 1} \\
+ s^2 c_{u}^{(2,2)} (\Delta X_{u}^{(1)})^{2r} (\Delta X_{u}^{(2)})^{2(s - 1)}).
\]

Refer to the proof of theorem 1 in appendix A for details. Therefore, we obtain the following theorem which plays an important role in the construction of our co-jump test.

**Theorem 1.** Suppose Assumptions 1, 2 and 3 are satisfied with \( q = 0 \) in assumption 2. Let \( S_0 = \sum_{0 \leq s \leq 1} (\Delta X_s^{(1)})^2 (\Delta X_s^{(2)})^2 \) and \( k = 1 \) or 2. Then, in restriction to \( \Omega^{(f)} \), we have

\[
\sqrt{n}(S_{k,2,2}^{(n)}(Y) - S_0) \xrightarrow{\mathcal{L}-s} \hat{U}^{(1,2)} = U_{1,k}^{(1,2)} + U_2^{(1,2)} + U_3^{(1,2)}.
\]

Now we propose a co-jump test for two-dimensional high-frequency data under the presence of small noise. From theorem 1, we have \( T^{(n)} \overset{P}{\rightarrow} 1 \) under the null hypothesis. Because

\[
T^{(n)} - 1 = \frac{S_{2,2,2}^{(n)}(Y) - S_{1,2,2}^{(n)}(Y)}{S_{1,2,2}^{(n)}(Y)},
\]

we must have the asymptotic distribution of \( S_{2,2,2}^{(n)}(Y) - S_{1,2,2}^{(n)}(Y) \). Considering the decomposition of \( S_{k,2,2}^{(n)} \) in (9), the asymptotic distribution can be derived from the joint limit distribution of \( (S^{(n)}, R_{1}^{(n)}, R_{2}^{(n)}) \), where \( S^{(n)} = S_{2,2,2}^{(n)}(Y) - S_{1,2,2}^{(n)}(Y) \), \( R_{1}^{(n)} = R_{2,2,1}^{(n)}(X) - R_{1,2,1}^{(n)}(X) \) and \( R_{2}^{(n)} = R_{1,2,2}^{(n)}(X) - R_{1,1,2}^{(n)}(X) \). In restriction to the set \( \Omega^{(j)} \), we have

\[
\sqrt{n}(S^{(n)}, R_{1}^{(n)}, R_{2}^{(n)}) \xrightarrow{\mathcal{L}-s} (U_{1,1}^{(1,2)}, U_2^{(1,2)}, U_3^{(1,2)}).
\]

Refer to the proof of theorem 2 in appendix A for details. Hence, we obtain the asymptotic property of the test statistics under the null hypothesis.

**Theorem 2.** Suppose Assumptions 1, 2 and 3 are satisfied with \( q = 0 \) in assumption 2. Then under the null hypothesis \( \mathbb{H}_0 \), we have

\[
\sqrt{n}(T^{(n)} - 1) \xrightarrow{\mathcal{L}-s} U = \frac{U_{1,1}^{(1,2)} + U_2^{(1,2)} + U_3^{(1,2)}}{(U_0^{(1,2)})^2},
\]

where \( U_0^{(1,2)} = \sum_{0 \leq s \leq 1} (\Delta X_s^{(1)})^2 (\Delta X_s^{(2)})^2 \), and \( U_{1,1}^{(1,2)}, U_l^{(1,2)} \) for \( l = 2, 3 \) are \( \mathcal{F}^{(0)} \)-conditionally mutually independent Gaussian random variables appearing in theorem 1.

We notice that in contrast to theorem 4.1 in JACOD & Todorov (2009), additional asymptotic variance terms appear in the limit distribution of the test statistics that must be estimated for the construction of a co-jump test under the small noise assumption. If \( \xi_1 = \xi_2 = 0 \) (no noise case), the result in theorem 2 corresponds to their result. Hence, theorem 2 includes their result as a special case. We propose methods to estimate the asymptotic variance of the test statistic in section 5.

Next, we consider the asymptotic property of the test statistic under the alternative hypothesis \( \mathbb{H}_1 \). In restriction to the set \( \Omega^{(d)} \), because of the presence of small noise, we have

\[
S_{k,2,2}^{(n)}(Y) = S_{k,2,2}^{(n)}(X) + 2R_{k,2,1}^{(n)}(X) + 2R_{k,1,2}^{(n)}(X) + o_P(1/n).
\]

© 2017 Board of the Foundation of the Scandinavian Journal of Statistics.
It is possible to obtain a joint limit theorem of \( n \times (S_{2,2,2}^{(n)}(X), S_{1,2,2}^{(n)}(X), \tilde{R}_1^{(n)}, \tilde{R}_2^{(n)}) \), where 
\[
\tilde{R}_1^{(n)} = R_{2,2,1}^{(n)}(X) + R_{2,1,2}^{(n)}(X) \quad \text{and} \quad \tilde{R}_2^{(n)} = R_{1,2,1}^{(n)}(X) + R_{1,1,2}^{(n)}(X)
\]
in restriction to the set \( \Omega^{(d)} \). This yields
\[
T^{(n)} \xrightarrow{\mathcal{L}} \Phi,
\]
where \( \Phi \) is a random variable which is almost surely different from 1. The detailed description of \( \Phi \) and proof is given in appendix A. Then, we obtain the following proposition.

**Proposition 4.** Suppose assumptions 1, 2 and 3 are satisfied with \( q = 0 \) in assumption 2. Then the test statistic \( T^{(n)} \) has asymptotic size \( \alpha \) under the null hypothesis \( \mathbb{H}_0 \) and is consistent under the alternative hypothesis \( \mathbb{H}_1 \), that is,
\[
\begin{align*}
&P\left( \frac{\sqrt{n}(T^{(n)} - 1)}{V^{(j)}_{1,2}} \geq q_{1-\alpha/2} \middle| \mathbb{H}_0 \right) \rightarrow \alpha, \quad \text{if } P(\Omega^{(j)}) > 0, \\
&P\left( \frac{\sqrt{n}(T^{(n)} - 1)}{V^{(j)}_{1,2}} \geq q_{1-\alpha/2} \middle| \mathbb{H}_1 \right) \rightarrow 1, \quad \text{if } P(\Omega^{(d)}) > 0,
\end{align*}
\]
where \( V^{(j)}_{1,2} \) is the asymptotic conditional variance of the random variable given in theorem 2, \( P(\cdot|\mathbb{H}_0) \) and \( P(\cdot|\mathbb{H}_1) \) are conditional probabilities with respect to the sets \( \Omega^{(j)} \) and \( \Omega^{(d)} \) and \( q_\alpha \) be the \( \alpha \)-quantile of standard normal distribution.

Proposition 4 implies that if we have a consistent estimator of \( V^{(j)}_{1,2} \), then we can carry out the co-jump test. The procedures to estimate the asymptotic variance of test statistics are discussed in section 5.

**5. Consistent estimation of the asymptotic conditional variances**

In this section, we first construct an estimator of the noise parameter \( \xi_m \). Then, we propose estimation procedures of the asymptotic conditional variance of RV and co-jump test statistics.

**5.1. Estimation of the noise variances**

The most important characteristic of the SIML is its simplicity compared with the pre-averaging method in Jacod et al. (2009) and spectral method in Bibinger & Winkelmann (2015) for estimating the quadratic variation consistently, and have asymptotic robustness for the rounding-error Kunitomo & Sato (2010, 2011). It is also quite easy to deal with the multivariate high-frequency data in this approach as demonstrated in Kunitomo & Sato (2011).

Let \( W_n = (Y_{\Delta n}^1, Y_{\Delta n}^2, \ldots, Y_{\Delta n}^d)^\top \) be a \( n \times d \) matrix, where \( Y_{j\Delta n}^j \) is the \( j \) th observation of the process \( Y \) and \( C_n \) be \( n \times n \) matrix,
\[
C_n = \begin{pmatrix}
1 & 0 & \cdots & 0 \\
1 & 0 & \cdots & 0 \\
1 & 1 & \cdots & 0 \\
1 & \cdots & 1 & 0
\end{pmatrix}, \quad C_n^{-1} = \begin{pmatrix}
1 & 0 & \cdots & 0 \\
-1 & 1 & \cdots & 0 \\
0 & -1 & 1 & \cdots \\
0 & \cdots & -1 & 1
\end{pmatrix}.
\]
We consider the spectral decomposition of \( C_n^{-1}(C_n^{-1})^\top \), that is,
\[
C_n^{-1}(C_n^{-1})^\top = P_nD_nP_n^\top = 2I_n - 2A_n.
\]
where $D_n$ is a diagonal matrix with the $k$th element

$$d_k = 2 \left[ 1 - \cos \left( \pi \left( \frac{2k - 1}{2n + 1} \right) \right) \right], \quad k = 1, \ldots, n,$$

and

$$p_{nk} = \sqrt{\frac{2}{n + (1/2)}} \cos \left[ \frac{2\pi}{2n + 1} \left( k - \frac{1}{2} \right) \left( j - \frac{1}{2} \right) \right], \quad 1 \leq j, k \leq n,$$

$$A_n = \frac{1}{2} \begin{pmatrix}
1 & 1 & \cdots & 0 & 0 \\
1 & 0 & 1 & \cdots & 0 \\
0 & 1 & 0 & \cdots & 0 \\
0 & \cdots & 1 & 0 & 1 \\
0 & \cdots & 0 & 1 & 0
\end{pmatrix}.$$

We transform $W_n$ to $Z_n$ by

$$Z_n = \Delta_n^{-1/2} P_n C_n^{-1} (W_n - \overline{W}_0),$$

where $\overline{W}_0 = (Y_0^T, \ldots, Y_0^T)^T$. The SIML estimator of the quadratic variation $[X, X]$ is defined by

$$[\hat{X}, \hat{X}]_{\text{SIML}} = \frac{1}{m_n} \sum_{k=1}^{m_n} Z_n^T Z_n, \quad m_n = n^p, \quad 0 < p < \frac{1}{2}. \quad (12)$$

From the straightforward extension of theorem 1 of Kunitomo & Sato (2013), we obtain

$$[\hat{X}, \hat{X}]_{\text{SIML}} \to_p [X, X] = \int_0^1 c_s ds + \sum_{0 \leq s \leq 1} (\Delta X_s)(\Delta X_s)^T. \quad (13)$$

From (5), (6) and (13), we obtain the next proposition.

**Proposition 5.** Suppose Assumptions 1, 2, and 3 with $q = 0$ in assumption 2 are satisfied. Then

$$\frac{1}{2} \left( \text{RV}(Y^{(m)}) - [\hat{X}, \hat{X}]_{\text{SIML}}^{(m,m)} \right) \to_p \zeta_m, \quad m = 1, \ldots, d,$$

where $[\hat{X}, \hat{X}]_{\text{SIML}}^{(m,m)}$ is the $(m, m)$ component of the SIML estimator.

The SIML estimator can be regarded as a modification of the standard maximum likelihood method under the Gaussian process and an extension for the multivariate case of the maximum likelihood estimation of the univariate diffusion process with market microstructure noise by Gloter & Jacod (2001b) (Kunitomo & Sato, 2013). In stead of the SIML estimator, we can also use the quasi maximum likelihood estimator studied in Xiu (2010) for the noise robust estimation of quadratic variation.
5.2. Estimation of the asymptotic conditional variances of realized volatility and the test statistic

We propose procedures to estimate the the asymptotic conditional variances of co-jump test statistics studied in section 4, and RV when the process \((X_t)_{0 \leq t \leq 1}\) has continuous path or discontinuous path. We introduce some notations. For \(l, m, p, q = 1, \ldots, d\),

\[
D^{(l,m)}_{p,q}(r, s) = \sum_{0 < u \leq 1} \left[ (\Delta X^{(l)}_u)^r (\Delta X^{(m)}_u)^s \right], \quad r, s \geq 1,
\]

\[
F^{(l,m)}(r, s) = r^2 D^{(l,m)}_{1,1} + 2rs D^{(l,m)}_{1,m} (2r-1, 2s-1) + s^2 D^{(l,m)}_{m,m} (2r, 2(s-1)).
\]

\[
A^{(l,m)}(r) = \int_0^1 (\epsilon_s^{(l,m)})^r ds, \quad r \geq 1,
\]

\[
J^{(l,m)}(r, s) = \sum_{0 < u \leq 1} (\Delta X^{(l)}_u)^r (\Delta X^{(m)}_u)^s, \quad r, s \geq 2.
\]

To estimate the asymptotic conditional variance of the co-jump test statistic, we consider following statistics. For \(n, m, p, q = 1, \ldots, d\) and \(r, s \geq 1\),

\[
\hat{D}^{(l,m)}_{p,q}(r, s) = \sum_{i=1}^n (\Delta^n Y^{(l)})^r (\Delta^n Y^{(m)})^s
\]

\[
\times \left[ \frac{1}{2kn \Delta_n} \sum_{j \in I_n(i)} (\Delta^n_j Y^{(p)}) (\Delta^n_j Y^{(q)})(\Delta^n_j Y^{(r)}) \right]_{\|\Delta^n_j Y\| \leq \alpha \Delta_n} - 2\delta_{pq} \sqrt{\hat{\xi}_p \hat{\xi}_q},
\]

and for \(r, s \geq 2\),

\[
\hat{F}^{(l,m)}(r, s) = r^2 \hat{D}^{(l,m)}_{1,1} (2r-1, 2s) + 2rs \hat{D}^{(l,m)}_{1,m} (2r-1, 2s-1) + s^2 \hat{D}^{(l,m)}_{m,m} (2r, 2(s-1)),
\]

\[
\hat{J}^{(l,m)}(r, s) = \sum_{i=1}^n (\Delta^n Y^{(l)})^r (\Delta^n Y^{(m)})^s,
\]

where \(I_n(i) = \{ j \in \mathbb{N} : j \neq i, 1 \leq j \leq n, |i-j| \leq k_n \}, k_n \to \infty, k_n \Delta_n \to 0\) as \(n \to \infty\), \(\alpha > 0\) and \(\theta \in (0, 1/2)\).

Although we can construct the consistent estimators of the asymptotic variances based on the truncated functionals (refer to (Mancini, 2009) and (Jacod & Todorov, 2009)) for the no noise case, the truncation method gives us biased estimators because we now assume the presence of the small noise. In fact, because the size of noise gets smaller as the observation frequency increases, truncation is not sufficient to distinguish the continuous part of process \(X\) with the noise. To be precise, from the similar argument of theorem 9.3.2 in Jacod & Protter (2012), we have

\[
\frac{1}{2kn \Delta_n} \sum_{j \in I_n(i)} (\Delta^n_j Y^{(p)}) (\Delta^n_j Y^{(q)})(\Delta^n_j Y^{(r)}) \to 0,
\]

\[
\frac{1}{2kn \Delta_n} \sum_{j \in I_n(i)} (\Delta^n_j Y^{(p)}) + \epsilon_{p,n} \Delta_j W^{(p)}(\Delta^n_j Y^{(q)}) + \epsilon_{q,n} \Delta_j W^{(q)} \to 0.
\]

where \(C_t = \int_0^t \sigma_s dW_s\) and \(\Delta^n Y^{(p)}\) is the \(p\)th component of \(C_j \Delta_n - C_{(j-1)\Delta_n}\).
Because $E[\Delta_i^n C^{(p)} \epsilon_{i,n} \Delta_i v^{(q)}] = 0$ and $\epsilon_{p,n} \epsilon_{q,n} E[\Delta_i v^{(p)} \Delta_i v^{(q)}] = 2\delta_{pq} \sqrt{\xi_p \xi_q} \Delta_n + o(\Delta_n)$ for all $1 \leq p, q \leq d$, $\#I_n(i) = 2k_n$ where $\delta_{pq}$ is a Dirac’s delta function, it follows that

$$\frac{1}{2k_n \Delta_n} \sum_{j \in I_n(i)} \Delta_j^n C^{(p)} \Delta_j^n C^{(q)} - c_i^{(p,q)} \rightarrow 0.$$

Therefore, we obtain

$$\frac{1}{2k_n \Delta_n} \sum_{j \in I_n(i)} (\Delta_j^n Y^{(p)}) (\Delta_j^n Y^{(q)}) 1_{\{|Y_j^n| \leq \alpha \Delta_n^a\}} - [c_i^{(p,q)} + 2\delta_{pq} \sqrt{\xi_p \xi_q}] \rightarrow 0.$$

Then, for $\hat{D}_{p,q}^{(l,m)}(r,s)$,

$$\frac{1}{2k_n \Delta_n} \sum_{j \in I_n(i)} (\Delta_i^n Y^{(p)}) (\Delta_i^n Y^{(q)}) 1_{\{|Y_i^n| \leq \alpha \Delta_n^a\}} - 2\delta_{pq} \sqrt{\xi_p \xi_q}$$

is an estimator of the spot volatility $c_i^{(p,q)}$. By the simple extension of theorem 2 in Kunitomo & Kurisu (2017), it can be obtained that $(\Delta_i^n Y^{(l)})^T (\Delta_i^n Y^{(m)})^t$ is an unbiased estimator of $(\Delta X_i^{(l)})^T (\Delta X_i^{(m)})^t$. Hence, for estimating the asymptotic conditional variance of RV, we consider following statistics:

$$\hat{c}_i^{(l,m)} = \frac{1}{2k_n \Delta_n} \sum_{j \in I_n(i)} (\Delta_i^n Y^{(l)}) (\Delta_i^n Y^{(m)}) 1_{\{|Y_i^n| \leq \alpha \Delta_n^a\}} - 2\delta_{lm} \sqrt{\xi_l \xi_m},$$

$$\hat{A}^{(l,m)}(r) = \Delta_n \sum_{i=1}^{n-k_n+1} (\hat{c}_i^{(l,m)})^r,$$

$$\hat{D}_{p,q}^{(l,m)}(1,1) = \sum_{i=1}^{n} c_i^{(p,q)} (\Delta_i^n Y^{(l)}) (\Delta_i^n Y^{(m)}) 1_{\{|Y_i^n| > \alpha \Delta_n^a\}}.$$

Then, we obtain the next result.

**Proposition 6.** Suppose assumptions 1, 2 and 3 are satisfied with $q = 0$ in assumption 2. Then, for $l, m, p, q = 1, \ldots, d$, we have for $r, s \geq 1$

$$\hat{D}_{p,q}^{(l,m)}(r,s) \xrightarrow{p} D_{p,q}^{(l,m)}(r,s), \quad \hat{A}^{(l,m)}(r) \xrightarrow{p} A^{(l,m)}(r),$$

and for $r, s \geq 2$,

$$\hat{F}^{(l,m)}(r,s) \xrightarrow{p} F^{(l,m)}(r,s), \quad \hat{J}^{(l,m)}(r,s) \xrightarrow{p} J^{(l,m)}(r,s).$$

From propositions 5 and 6, we can construct the consistent estimator of the asymptotic variance of co-jump test statistics

$$\hat{V}_{1,2} = \frac{\hat{F}^{(1,2)}(2,2) + 8(\hat{\xi}_2 \hat{J}^{(1,2)}(2,4) + \hat{\xi}_1 \hat{J}^{(1,2)}(4,2))}{n \times \hat{J}^{(1,2)}(2,2)},$$

and for consistent estimators of the asymptotic variances of RV.

© 2017 Board of the Foundation of the Scandinavian Journal of Statistics.
\[ \hat{V}^{(c)} = (2\widehat{A}^{(1,1)}(4) + 8\widehat{\zeta}_1\widehat{A}^{(1,1)}(2) + 12\zeta_1^2)/n, \quad \text{if } X \text{ is continuous,} \]
\[ \hat{V}^{(j)} = (2\widehat{A}^{(1,1)}(4) + 4\widehat{\zeta}_1^{(1,1)}(1,1) + 8\zeta_1^2[X, X]_{(1,1)}^{SIML})/n, \quad \text{if } X \text{ is an Itô semimartingale of the form (3).} \]

Therefore, we obtain the next two central limit theorems.

**Corollary 1.** Suppose assumptions 1, 2 and 3 are satisfied with \( q = 0 \) in assumption 2. Then we have the following convergence in law:

(i) If \( X \) is continuous, then
\[ (\hat{V}^{(c)})^{-1/2}(RV(Y) - U_0^{(c)}) \to N(0, 1), \]
where \( U_0^{(c)} = \int_0^1 \sigma^2_s ds + 2\zeta_1. \)

(ii) If \( X \) is the Itô semimartingale of the form (3), then
\[ (\hat{V}^{(j)})^{-1/2}(RV(Y) - U_0^{(j)}) \to N(0, 1), \]
where \( U_0^{(j)} = \int_0^1 \sigma^2_s ds + \sum_{0 \leq s \leq 1} (\Delta X_s)^2 + 2\zeta_1. \)

**Corollary 2.** Suppose assumptions 1, 2 and 3 are satisfied with \( q = 0 \) in assumption 2. Then, under the null hypothesis \( H_0 \), we have
\[ (\hat{V}_{1,2}^{(j)})^{-1/2}(T^{(n)} - 1) \to N(0, 1). \]

6. Numerical experiments

In this section, we report several results of numerical experiment. We simulate a data generating process according to the procedure in Cont & Tankov (2004). First, we consider the estimation of integrated volatility. We used following data generating processes:

\[ dX_t = \sigma_t dW_t + dJ_t^\mathcal{P}, \quad \text{ (14)} \]
\[ dX_t = \sigma_t dW_t + dJ_t^\mathcal{B}, \quad \text{ (15)} \]
where \( W \) is a standard Brownian motion. \( J_t^\mathcal{P} \) is a compound Poisson process in \( t \in [0, 1] \), where the intensity of the Poisson process is \( \lambda = 10 \), and the jump size is a uniform distribution on \([-0.3, -0.05] \cup [0.05, 0.3] \). \( J_t^\mathcal{B} \) is a \( \beta \)-stable process with \( \beta = 1.5 \). We also set the truncate level \( \alpha \Delta_\theta^\alpha \) as \( \alpha = 2 \) and \( \theta = 0.48 \). These values are also used in the simulation of the co-jump test. As a stochastic volatility model, we use
\[ d\sigma_t^2 = \alpha(\beta - \sigma_t^2)dt + \sigma_t dW_t^\sigma, \quad \text{ (16)} \]
where \( \alpha = 5, \beta = 0.2 \) and \( \rho = E[dW_t dW_t^\sigma] = -0.5 \). For the market microstructure noise, we have adopted the three types of Gaussian noise \( \mathcal{N}(0, \xi \Delta_n) \) with \( \xi = 0, 10^{-4}, \) and \( 10^{-2} \) (we call these cases as (i), (ii) and (iii)). To estimate \( \xi \), we used SIML with \( \rho = 0.49 \) in (12) and we also use this value in the following simulations. We also consider following cases:

(i) CJ1: \( 1/\Delta_n = 20,000, \) \( X \) follows the process (14).
(ii) CJ2: \( 1/\Delta_n = 30,000, \) \( X \) follows the process (14).
(iii) SJ1: $1/\Delta_n = 20,000$, $X$ follows the process (15).
(iv) SJ2: $1/\Delta_n = 30,000$, $X$ follows the process (15).

CJ and SJ correspond to a finite activity jump case and an infinite activity jump case, respectively. We present the root mean square error for each case in Table 1. Compared with CJ, the root mean square errors in SJ tends to be large. This is because of the difficulty in a finite sample case to distinguish infinite activity jumps and the other part of the observed process.

Next, we check the performance of the proposed feasible central limit theorem (CLT) of $RV$. We use following data generating processes:

$$
\begin{align*}
    dX_t &= \sigma_t dW_t, \quad \text{for continuous case}, \\
    dX_t &= \sigma_t dW_t + dJ^P_t, \quad \text{for jump case}, \\
\end{align*}
$$

(17)

where the process (17) and $\sigma$ are the same processes as (16) and (14). Figures 1 and 2 give the standardized empirical densities when the small noise is $N(0, 10^{-2}\Delta_n)$. The simulation size is $N = 1,000$ and the number of the observations is $n = 20,000$. The red line is the density of standard normal distribution. The left figure corresponds to the bias-variance corrected case implied by corollary 1 and the right figure corresponds to the no correction case. We found that contamination of small noise still has the significant effect on the distribution of the limiting random variables when $\zeta_m \neq 0$ in the continuous case. We can see this in Fig 1. However, we have a good approximation for the finite sample distribution of statistics, if we correct the effect of noise by using the small noise asymptotics.

Finally, we give simulation results of the co-jump test. We simulate a two-dimensional Itô semimartingale by the following model:

Table 1. \textit{RMSEs of $\hat{\mathbf{V}}$. Values are reported as multiples of $10^{-3}$}

| Case    | CJ1-(i) | CJ1-(ii) | CJ1-(iii) | CJ2-(i) | CJ2-(ii) | CJ2-(iii) |
|---------|---------|----------|-----------|---------|----------|-----------|
| RMSE    | 2.137   | 2.192    | 2.349     | 1.701   | 1.735    | 1.878     |
| Case    | SJ1-(i) | SJ1-(ii) | SJ1-(iii) | SJ2-(i) | SJ2-(ii) | SJ2-(iii) |
| RMSE    | 11.45   | 11.47    | 11.40     | 9.501   | 9.496    | 9.470     |

RMSE, root mean square error

Fig 1. Empirical distributions of $RV$ when $dX_t = \sigma_t dW_t$, $\xi = 10^{-2}$. The left figure corresponds to the bias-variance corrected case implied by corollary 1(i). The right figure corresponds to the no correction case (statistics are standardized as $\xi_1 = 0$ in Corollary 1(i)). The red line is the density of a standard normal distribution. [Colour figure can be viewed at wileyonlinelibrary.com]
Fig. 2. Empirical distributions of realized volatility (RV) when \( dX_t = \alpha_t dW_t + dJ_t^\rho, \xi = 10^{-2} \). We plot the left and right figures in the same way as Figure 2. [Colour figure can be viewed at wileyonlinelibrary.com]

\[
\begin{align*}
    dX_t^{(1)} &= \alpha_t^{(1)} dW_t^{(1)} + Z_t^{(1)} dN_t^{(1)} + Z_t^{(3)} dN_t^{(3)}, \\
    dX_t^{(2)} &= \alpha_t^{(2)} dW_t^{(2)} + Z_t^{(2)} dN_t^{(2)} + Z_t^{(4)} dN_t^{(3)},
\end{align*}
\]

where \( W = (W^{(1)}, W^{(2)}) \) is a two-dimensional Brownian motion and \( N^{(j)} \) for \( j = 1, 2, 3 \) are Poisson processes with intensities \( \lambda_j \) which are mutually independent and also independent of \( W \). \( Z = (Z^{(1)}, Z^{(2)}, Z^{(3)}, Z^{(4)}) \) is the vector of jump sizes, cross-sectionally and serially independently distributed with laws \( F_{Z^{(j)}} \). In this simulation, we set \( \lambda_1 = \lambda_2 = 5 \) and \( \lambda_3 = 10 \), and jump size distributions are \( Z_t^{(1)}, Z_t^{(2)} \sim N(0.5^{-2}), Z_t^{(3)}, Z_t^{(4)} \sim N(0, 10^{-2}) \). For the volatility process \( \sigma \), we set

\[
    d(\sigma_t^{(j)})^2 = \alpha_j (\beta_j - (\sigma_t^{(j)})^2) dt + \sigma_t^{(j)} dW_t^{\sigma^{(j)}}, \quad j = 1, 2,
\]

where \( \sigma^{(1)} \) and \( \sigma^{(2)} \) are independent and \( \alpha_1 = \alpha_2 = 5 \), \( \beta_1 = 0.2 \), \( \beta_2 = 0.15 \), \( E[dW_t^{\sigma^{(j)}} dW_t^{\sigma^{(j)}}] = \rho_j dt, \rho_1 = -0.5 \), and \( \rho_2 = -0.4 \). For the market microstructure noise, we use the four types of Gaussian noise for each component with mean 0 and the same variance \( \sigma_{1,n} = \sigma_{2,n} = \zeta/n \) where \( \zeta = 10^{-4}, 10^{-2}, 10^{-1} \), and 1 (we call these cases as I, II, III, and IV). We consider the 5% significant level for following cases:

(i) C1 : \( 1/\Delta_n = 20,000 \), co-jump.
(ii) C2 : \( 1/\Delta_n = 30,000 \), co-jump.
(iii) D1 : \( 1/\Delta_n = 20,000 \), no co-jump.
(iv) D2 : \( 1/\Delta_n = 30,000 \), no co-jump.

In Fig 3, we plot the empirical distribution obtained from corollary 2 in the case C1–IV. It is interesting to see that our proposed testing procedure gives a good approximation of the limit distribution of test statistics even in large noise case \( (\xi = 1) \). In Table 2, we also give the empirical size and power of the proposed co-jump test and the test proposed in Jacod & Todorov (2009) (we call this JT test). We find that JT test is sensitive to the small noise. Theorem 2 implies that in the presence of small noise, the critical value of the test is larger than that of no noise case. Hence, the critical value of JT test is small compared with that of the proposed test, and the empirical size of JT test tends to be larger than the significant level. In particular, when the effect of noise is large (C1-III, C1-IV, C2-III and C2-IV), we can see the size distortion of the JT test. On the other hand, the empirical size of the proposed test is close to 0.05. This result shows that we need to correct the asymptotic variance of JT test in the presence of small noise. Additionally, the empirical power of our proposed test is very close to 1 in each case.
Fig. 3. Empirical distribution of the co-jump test implied by corollary 2 in C1–IV. The red line is the density of a standard normal distribution. [Colour figure can be viewed at wileyonlinelibrary.com]

Table 2. Empirical size and power of the co-jump test (5% significant level) are reported for the proposed test (top) and test proposed in Jacod & Todorov (2009) (bottom)

| Case | C1-I | C1-II | C1-III | C1-IV | C2-I | C2-II | C2-III | C2-IV |
|------|------|-------|--------|-------|------|-------|--------|-------|
| Size | 0.068 | 0.064 | 0.065 | 0.071 | 0.058 | 0.053 | 0.052  | 0.046 |
| Power| 0.082 | 0.078 | 0.139 | 0.253 | 0.069 | 0.070 | 0.128  | 0.237 |

| Case | D1-I | D1-II | D1-III | D1-IV | D2-I | D2-II | D2-III | D2-IV |
|------|------|-------|--------|-------|------|-------|--------|-------|
| Power| 0.986 | 0.994 | 0.989 | 0.985 | 0.978 | 0.984 | 0.986  | 0.990 |

7. Extensions

In this section, we consider some extension of our results given in the previous sections. First, we discuss about the assumption on the volatility process \( \sigma_t \), and then we discuss about the assumption on the noise. In assumption 1, we assume that the volatility process \( \sigma_t \) is continuous (condition (b)) for the simplicity of our analysis. On the other hand, there are some empirical evidence that \( \sigma_t \) and the original process \( X_t \) have common jumps, refer to Jacod & Todorov (2009, 2010) and Bibinger & Winkelmann (2016) for examples. Instead of (b) in assumption 1, we consider the following condition:

(b') The process \( \sigma_t \) is càdlàg.

This condition implies that \( \sigma_t \) and \( X_t \) could have common-jumps. Under assumptions 1 with (b') in place of (b), 2 with \( q = 0 \) and 3, every results in our paper holds and our proofs are still valid for every results with minor and straightforward changes. However, the limit distributions of proposed test statistics and the RV are somewhat complicated in that case. Let

\[
\tilde{F}_{r,s} = \sum_{0 \leq u \leq 1} \left( r^2(c_{u-}^{(1,1)} + c_u^{(1,1)}) (\Delta X_u^{(1)})^{2r-1} (\Delta X_u^{(2)})^{2s} 
+ 2rs(c_{u-}^{(1,2)} + c_u^{(1,2)}) (\Delta X_u^{(1)})^{2r-1} (\Delta X_u^{(2)})^{2s-1} s^2(c_{u-}^{(2,2)})
+ c_u^{(2,2)} (\Delta X_u^{(1)})^{2r} (\Delta X_u^{(2)})^{2(s-1)} \right),
\]

where \( c_{u-}^{(p,q)} = \lim_{t \uparrow u} c_t^{(p,q)} \). Then theorem 1 still holds by replacing the \( \mathcal{F}^{(0)} \)-conditional variance of \( U_{1,k}^{(1,2)} \) with \( k \tilde{F}_{2,2} \). Theorem 2 and proposition 4 also hold under this general
setting. If the process $X$ is one dimension and have jumps, the asymptotic distribution of the RV also changes and we have the following stable convergence:

$$
\sqrt{n}(RV(Y) - U_{0,2}) \xrightarrow{d} \tilde{U}_1 + U_2,
$$

where $U_{0,2} = \int_0^1 \sigma_s^2 ds + \sum_{0 \leq s \leq 1} (\Delta X_s)^2 + 2\xi_1$, $\tilde{U}_1$ and $U_2$ are $\mathcal{F}(0)$-conditionally independent random variables with mean 0 and $\mathcal{F}(0)$-conditional variances

$$
E(\tilde{U}_1^2 | \mathcal{F}(0)) = 2 \int_0^1 \sigma_s^4 ds + 2 \sum_{0 \leq s \leq 1} (\sigma_{s-}^2 + \sigma_s^2)(\Delta X_s)^2,
$$

$$
E(U_2^2 | \mathcal{F}(0)) = 8\xi_1 \left[ \int_0^1 \sigma_s^2 ds + \sum_{0 \leq s \leq 1} (\Delta X_s)^2 \right].
$$

Moreover, it should be noted that estimators $\tilde{D}_{p,q}^{(l,m)}(r,s)$, $\tilde{J}_{(l,m)}(r,s)$, $\tilde{F}^{(l,m)}(r,s)$ and $\tilde{A}^{(l,m)}(r)$ proposed in our paper are still consistent estimators of $D_{p,q}^{(l,m)}(r,s)$, $J_{(l,m)}(r,s)$, $F^{(l,m)}(r,s)$ and $A^{(l,m)}(r)$, where

$$
\tilde{D}_{p,q}^{(l,m)}(r,s) = \frac{1}{2} \sum_{0 \leq u \leq 1} (c_{u}^{(p,q)} + c_{u}^{(p,q)}) (\Delta X_u)^r (\Delta X_u^m)^s,
$$

$$
\tilde{F}^{(l,m)}(r,s) = r^2 \tilde{D}_{1,l}^{(l,m)}(2(r - 1), 2s) + 2rs \tilde{D}_{l,m}^{(l,m)}(2r - 1, 2s - 1) + s^2 \tilde{D}_{m,m}^{(l,m)}(2r, 2(s - 1)),
$$

and $J_{(l,m)}(r,s)$ and $A^{(l,m)}(r)$ are the same values defined in section 5.2 in our paper. Therefore, proposition 5, corollaries 1 and 2 hold. We call the set of conditions (a), (c), (d) and (e) in assumption 1, and (b') as assumption 1'.

For the noise terms, assumption 3 is used to derive the asymptotic distribution of the proposed test statistics. In particular, we need assumption 3 for theorems 1, and 2, propositions 4 and 5, and corollaries 1 and 2. However, to prove other results, we only used the finiteness of moments of noise. So, the Gaussianity of noise can be removed in those results. We consider the following assumption in place of assumption 3 in the present paper:

**Assumption 4.** \(\{v_{i,j}\}_{i=1}^{\infty}\) is a sequence of i.i.d. $d$-dimensional random variables such that $v_{i,j}$, \(j = 1, \ldots, d\) are independent and $E[|v_{i,j}|^\eta] < \infty$ for some $\eta > 0$.

Concretely, proposition 1 holds under assumptions 1', 2 and 4 with $\eta = 4r$ for the first result and $\eta = 2(r + s)$ for the second result, and Proposition 2 and 3 hold under Assumptions 1', 2, and 4 with $\eta = 2(r + s)$ and $\eta = 4$, respectively. Furthermore, Proposition 5 holds under assumptions 1' and 2 with $q = 0$ and 4 with $\eta = 4$.

8. Concluding remarks

In this paper, we developed the small noise asymptotic analysis when the size of the market microstructure noise depends on the frequency of the observation. By using this approach, we can identify the effects of jumps and noise in high-frequency data analysis. We investigated the asymptotic properties of BPV and RV in one-dimensional case in the presence of small noise. We proposed methods to estimate integrated volatility and the asymptotic conditional
variance of RV. As a result, feasible central limit theorems of RV is established under the small noise assumption when the latent process \( X \) is an Itô semimartingale. Our method gives a good approximation of the limiting distributions of the sequence of random variables.

We also proposed a testing procedure of the presence of co-jumps when the two-dimensional latent process is observed with small noise for testing the null hypothesis that the observed two latent processes have co-jumps. Our proposed co-jump test is an extension the co-jump test in Jacod & Todorov (2009) for the noisy observation setting. Estimators of the asymptotic conditional variance of the test statistics can be constructed in a simple way. We found that the empirical size of the proposed test works well in finite sample. In particular, proposed co-jump test has a good performance even when the effect of noise is large.

Acknowledgements

I would like to thank the editor, associate editor and anonymous referees for carefully reading the manuscript and for their constructive comments that helped improve on the quality of the manuscript. One referee kindly pointed out relevant references Bacry et al. (2013a), Bacry & Muzy (2014) and Rosenbaum (2011). I also would like to thank Kengo Kato, Naoto Kunitomo and Yasuhiro Omori for their helpful suggestions and encouragements. This work is partially supported by Grant-in-Aid for JSPS Research Fellow (16J06454) from the JSPS.

References

Aït-Sahalia, Y. & Jacod, J. (2009). Testing for jumps in a discretely observed process. *Ann. Stat.* 37, 184–222.
Aït-Sahalia, Y., Jacod, J. & Li, J. (2012). Testing for jumps in noisy high-frequency data. *J. Econometrics* 168, 207–222.
Aït-Sahalia, Y., Mykland, P. & Zhang, L. (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *Rev. Fin. Stud.* 18, 351–416.
Aït-Sahalia, Y. & Yu, J. (2009). High-frequency market microstructure noise estimated and liquidity measures. *Ann. Appl. Statist.* 3, 422–457.
Bacry, E., Delatte, S., Hoffmann, M. & Muzy, J.F. (2013a). Modelling microstructure noise with mutually exciting point processes. *Quant. Finance* 13, 65–77.
Bacry, E., Delatte, S., Hoffmann, M. & Muzy, J.F. (2013b). Scaling limits for Hawkes processes and application to financial statistics. *Stochastic Process. Appl.* 123, 2475–2499.
Bacry, E. & Muzy, J.F. (2014). Hawkes model for price and trades high-frequency dynamics. *Quant. Finance* 14, 1147–1166.
Bandi, F. M. & Russel, J. R. (2006). Separating microstructure noise from volatility. *J. Financial Economics* 79, 655–692.
Barndorff-Nielsen, O. E., Graversen, S. E., Jacod, J., Podolskij, M. & Shephard, N. (2006). A central limit theorem for realised power and bipower variations of continuous semimartingales. In *From Stochastic Analysis to Mathematical Finance* (eds Y. Kabanov, R. Liptser & J. Stoyanov), 33–68. Springer, Berlin Heidelberg.
Barndorff-Nielsen, O. E. & Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *J. Financial Econometrics* 2, 1–48.
Barndorff-Nielsen, O. E. & Shephard, N. (2006). Econometrics of testing for jumps in financial economics using bipower variation. *J. Financial Econometrics* 4, 1–30.
Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. & Shephard, N. (2008). Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica* 76, 1481–1536.
Bibinger, M. & Reiß, M. (2014). Spectral estimation of covolatility from noisy observations using local weights. *Scand. J. Statist.* 41, 23–50.
Bibinger, M. & Winkelmann, L. (2015). Econometrics of co-jumps in high-frequency data with noise. *J. Econometrics* 184, 361–378.
Bibinger, M. & Winkelmann, L. (2016). Common price and volatility jumps in noisy high-frequency data. arXiv:1407.4376.
Bollerslev, T., Law, T. H. & Tauchen, G. (2008). Risk, jumps, and diversification. *J. Econometrics* 144, 234–256.
Bollerslev, T., Todorov, V. & Li, S. Z. (2013). Jump tails, extreme dependencies, and the distribution of stock returns. *J. Econometrics* **172**, 307–324.

Christensen, K., Podolskij, M. & Vetter, M. (2013). On covariation estimation for multivariate continuous Itô semimartingales with noise in non-synchronous observation schemes. *J. Multivariate Anal.* **120**, 59–84.

Cont, R. & Tankov, P. (2004). *Financial modeling with jump process*, Chapman & Hall.

Fan, J. & Wang, Y. (2007). Multi-scale jump and volatility analysis for high-frequency data. *J. Amer. Statist. Assoc.* **102**, 1349–1362.

Glider, D., Shackleton, M. B. & Taylor, S. J. (2014). Cojumps in stock prices: Empirical evidence. *J. Banking. Finance* **40**, 443–459.

Gloter, A. & Jacod, J. (2001a). Diffusions with measurement errors. I-Local asymptotic normality. *ESAIM: Probab. Stat.* **5**, 225–242.

Gloter, A. & Jacod, J. (2001b). Diffusions with measurement errors. II-Optimal estimators. *ESAIM: Probab. Stat.* **5**, 243–260.

Hansen, P. R. & Lunde, A. (2006). Realized variance and market microstructure noise. *J. Business. Econ. Statist.* **24**, 127–161.

Häusler, E. & Luschgy, H. (2015). *Stable convergence and stable limit theorems*, Springer International Publishing, Switzerland.

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

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

Jacod, J. & Protter, P. (2012). *Discretization of processes*, Springer, Berlin Heidelberg.

Kunitomo, N. & Kurisu, D. (2017). Effects of jumps and small noise in high-frequency financial econometrics. *Asia-Pac. Financial Markets* **24**, 39–73.

Kunitomo, N. & Sato, S. (2010). Robustness of the separating information maximum likelihood method estimation of realized volatility with micro-market noise. CIRJE Discussion Paper F-733, Graduate School of Economics, University of Tokyo, Tokyo.

Kunitomo, N. & Sato, S. (2011). The SIML estimation of realized volatility of Nikkei-225 futures and hedging coefficient with micro-market noise. *Math. Comput. Simul.* **81**, 1272–1289.

Kunitomo, N. & Sato, S. (2013). Separating information maximum likelihood estimation of the integrated volatility and covariance with micro-market noise. *The North Amer. J. Econ. Finance* **26**, 282–309.

Li, J. (2013). Robust estimation and inference for jumps in noisy high-frequency data: A local-to-continuity theory for the pre-averaging method. *Econometrica* **81**, 1673–1693.

Li, Y., Xie, S. & Zheng, X. (2016). Efficient estimation of integrated volatility incorporating trading information. *J. Econometrics* **195**, 33–50.

Li, Y., Zhang, Z. & Li, Y. (2015). A unified approach to volatility estimation in the presence of both rounding and random market microstructure noise. working paper.

Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scand. J. Statist.* **36**, 270–296.

Mancini, C. & Gobbi, F. (2012). Identifying the Brownian covariation from the co-jumps given discrete observations. *Econom. Theor.* **28**, 249–273.

Rosenbaum, M. (2011). A new microstructure index. *Quant. Finance* **11**, 883–899.

Xiu, D. (2010). Quasi-maximum likelihood estimation of volatility with high frequency data. *J. Econometrics* **159**, 235–250.

Zhang, L., Mykland, P. A. & Aït-Sahalia, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *J. Amer. Statist. Assoc.* **100**, 1394–1411.

Received March 2017, in final form October 2017

Daisuke Kurisu, Graduate School of Economics, University of Tokyo 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan.

E-mail: dkurisu.mathstat@gmail.com
Appendix A: Proofs

A.1 Proofs for section 3

Throughout appendix, \( K \) denotes a generic constant which may change from line to line. We use the techniques developed in Jacod & Protter (2012). We can replace assumption 1 to the local boundedness assumption below and such a replacement can be established by the localizing procedure provided in Jacod & Protter (2012).

Assumption 5. We have assumption 1 and there are a constant \( A \) and a non-negative function \( \Gamma \) on \( \mathbb{R}^d \) for all \( \omega, t, x \in \Omega \times \mathbb{R}_+ \times \mathbb{R}^d \) such that

\[
\max\{||b_t(\omega)||, ||\sigma_t(\omega)||, ||X_t(\omega)||, ||\Gamma(x)||\} \leq A.
\]

Under assumption 5, we can obtain following inequalities for Itô semimartingale. First, we consider the decomposition of \( X_t \),

\[
X_t = X_0 + B_t + C_t + J_t,
\]

where

\[
B_t = \int_0^t \left( b_s + \int_{\mathbb{R}^d} \kappa \circ \delta(s, x) \lambda(dx) \right) ds,
\]

\[
C_t = \int_0^t \sigma_s dW_s, \quad J_t = \int_0^t \int_{\mathbb{R}^d} \delta(s, x) (\mu - \nu)(ds, dx).
\]

For all \( p \geq 1, s, t \geq 0 \), we obtain the following inequalities:

\[
E \left( \sup_{0 \leq u \leq t} ||B_{s+u} - B_s||^p \big| \mathcal{F}_s^{(0)} \right) \leq K t^p, \tag{18}
\]

\[
E \left( \sup_{0 \leq u \leq t} ||C_{s+u} - C_s||^p \big| \mathcal{F}_s^{(0)} \right) \leq K t^{p/2}, \tag{19}
\]

\[
E \left( \sup_{0 \leq u \leq t} ||J_{s+u} - J_s||^p \big| \mathcal{F}_s^{(0)} \right) \leq K t^{p/\gamma + 1}. \tag{20}
\]

Refer to section 2.1.5 of Jacod & Protter (2012) for details of the derivation of these inequalities.

Proof of Proposition 1. First, we have

\[
\tilde{V}_{2r}^{(n)}(Y) - \tilde{V}_{2r}^{(n)}(X) = \sum_{i=1}^n \sum_{j=1}^{2r} \left( \frac{2r}{p} \right) (\Delta_i^n X)^{2r-p} (\epsilon_n \Delta v_i)^p.
\]

By using the inequalities (18), (19) and (20), it follows that

\[
E \left[ (\Delta_i^n X)^{2r-p} (\epsilon_n \Delta v_i)^p \right] \leq K \Delta_n^{\left(\frac{r}{2} + q\right)p} \times \Delta_n^{(r - \frac{p}{2})} = K \Delta_n^q + r, \quad i = 1, \ldots, n.
\]

Therefore, we obtain \( \tilde{V}_{2r}^{(n)}(Y) - \tilde{V}_{2r}^{(n)}(X) = O_P(\Delta_n^{q+r-1}) \). Second, we have
\[ \hat{V}_{r,s}(n) - \hat{V}_{r,s}(X) = \sum_{i=1}^{n-1} (|\Delta_i^n X + \epsilon_n \Delta v_i|^{r} |\Delta_{i+1}^n X + \epsilon_n \Delta v_{i+1}|^{s} - |\Delta_i^n X|^{r} |\Delta_{i+1}^n X|^{s}) . \]

By using the inequalities (18), (19) and (20) repeatedly, we have \( E[|\Delta_i^n X|^{r} |\Delta_{i+1}^n X|^{s}] \) for \( p \geq 1 \). Then, we have

\[
E \left[ \left( \frac{|\Delta_i^n X + \epsilon_n \Delta v_i|^{r} |\Delta_{i+1}^n X + \epsilon_n \Delta v_{i+1}|^{s} - |\Delta_i^n X|^{r} |\Delta_{i+1}^n X|^{s}}{\sqrt{n}^q} \right) \right] \\
= E \left[ \left( \frac{|\Delta_i^n X + \epsilon_n \Delta v_i|^{r} |\Delta_{i+1}^n X + \epsilon_n \Delta v_{i+1}|^{s} - |\Delta_i^n X|^{r} |\Delta_{i+1}^n X|^{s}}{\sqrt{n}^q} \right) \right] \\
\leq K \epsilon_n n^{-1/2} = O(\Delta_i^n), \quad i = 1, \ldots, n.
\]

Therefore, we obtain \( \hat{V}_{r,s}(n) - \hat{V}_{r,s}(X) = O_P(\Delta_i^q / (r+s)/2 + 1) \).

**Proof of proposition 2.** By proposition 1,

\[
n^{(r+s)/2-1} \hat{V}_{r,s}(n) - m_r m_s \int_0^{1/k} \sigma_s^{r+s} ds = n^{(r+s)/2-1} \left( \hat{V}_{r,s}(n) - \hat{V}_{r,s}(X) \right) \\
+ \left( n^{(r+s)/2-1} \hat{V}_{r,s}(X) - m_r m_s \int_0^{1/k} \sigma_s^{r+s} ds \right) \\
= O_P \left( \left( \frac{1}{\sqrt{n}} \right) \right).
\]

Therefore, if \( q > 0 \), then \( n^{(r+s)/2-1} \hat{V}_{r,s}(n) \) converges in probability to \( m_r m_s \int_0^{1/k} \sigma_s^{r+s} ds \). If \( q > 1/2 \), then from theorem 2.3 of Barndorff-Nielsen et al. (2006), the first part of above decomposition converges in probability to 0, and the second part converges stably to \( U \) defined in proposition 2. Then we obtain the desired result.

**A.2 Proofs for section 4**

**Proof of theorem 1.** We decompose \( S_{1,2,2}^{(n)}(Y) \) as follows:

\[
S_{1,2,2}^{(n)}(Y) = U_{1,n} + \sum_{j=1}^{2} U_{2,j,n} + \sum_{j=1}^{3} U_{3,j,n} + \sum_{j=1}^{3} U_{4,j,n},
\]

where

\[
U_{1,n} = \sum_{i=1}^{[n/k]} \left( \frac{|\Delta_i^n X|^{(k,1)} |\Delta_{i+1}^n X|^{(k,2)} |\Delta_{i+1}^n X|^{(k,1)} }{\sqrt{n}^q} \right), \quad U_{2,1,n} = 2R_{1,2,1}^{(n)}(X), \\
U_{2,2,n} = 2R_{1,2,2}^{(n)}(X), \\
U_{3,1,n} = \epsilon_{2,n}^{2} \sum_{i=1}^{[n/k]} \left( \frac{|\Delta_i^n X|^{(k,1)} |\Delta_{i+1}^n X|^{(k,2)} |\Delta_{i+1}^n X|^{(k,1)} }{\sqrt{n}^q} \right), \\
U_{3,2,n} = 4\epsilon_{1,n} \epsilon_{2,n} \sum_{i=1}^{[n/k]} \left( \frac{|\Delta_i^n X|^{(k,1)} |\Delta_{i+1}^n X|^{(k,2)} |\Delta_{i+1}^n X|^{(k,1)} }{\sqrt{n}^q} \right), \\
U_{3,3,n} = \epsilon_{1,n}^{2} \sum_{i=1}^{[n/k]} \left( \frac{|\Delta_i^n X|^{(k,2)} |\Delta_{i+1}^n X|^{(k,1)} }{\sqrt{n}^q} \right),
\]

© 2017 Board of the Foundation of the Scandinavian Journal of Statistics.


\[ U_{4,1,n} = 2\epsilon_{1,n}\epsilon_{2,n} \left[ n/k \right] \sum_{i=1}^{\left[ n/k \right]} (\Delta_i^{(k,1)} X_{t_i}^{(k,1)} (\Delta_i^{(k,1)} X_{t_i}^{(k,2)})^2, \]

\[ U_{4,2,n} = 2\epsilon_{1,n}^2 \epsilon_{2,n} \left[ n/k \right] \sum_{i=1}^{\left[ n/k \right]} (\Delta_i^{(k,2)} X_{t_i}^{(k,1)} (\Delta_i^{(k,1)} X_{t_i}^{(k,2)})^2, \]

\[ U_{4,3,n} = \epsilon_{1,n}^2 \epsilon_{2,n} \left[ n/k \right] \sum_{i=1}^{\left[ n/k \right]} (\Delta_i^{(k,1)} (\Delta_i^{(k,2)})^2, \]

and where \( R_{k,1,1}^{(n)}(X), R_{k,1,2}^{(n)}(X), \Delta_i^{(k,1)} \) and \( \Delta_i^{(k,2)} \) are defined in (10) and (11). For the proof of theorem 1, we first prove the stable convergence of \( U_{1,n}, U_{2,1,n} \) and \( U_{2,2,n} \). Then we prove the joint convergence of these terms. Finally, we will prove that \( \sum_{j=1}^{3} U_{3,j,n} \) and \( \sum_{j=1}^{3} U_{4,j,n} \) in the decomposition of \( S_{k,1,2}(Y) \) in (21) is asymptotically negligible.

**Evaluation of \( U_{1,n} \):** By theorem 3.3.1 and proposition 15.3.2 in Jacod & Protter (2012), we have \( U_{1,n} \xrightarrow{p} \sum_{0\leq u\leq 1} (\Delta X_u^{(1)})^2 (\Delta X_u^{(2)})^2 \equiv U_0 \), and

\[ \sqrt{n}(U_{1,n} - U_0) \xrightarrow{d} N(0, k F_{2,2}). \]

**Evaluation of \( U_{2,1,n} \) and \( U_{2,2,n} \):** Let \( (Z_p^{(1)}) \) and \( (Z_p^{(2)}) \) be the mutually independent sequences of i.i.d. standard normal random variables defined on the second filtered probability space \( (\Omega^{(1)}, \mathcal{F}^{(1)}), (\mathcal{F}_t^{(1)})_{t\geq 0}, P^{(1)} \) and \( (\tau_p) \) be the co-jump times of the first and second component of the process \( (X_t = (X_t^{(1)}, X_t^{(2)}))_{0\leq t\leq 1} \). We will prove the following result:

\[ (\sqrt{n}U_{2,1,n}, \sqrt{n}U_{2,2,n}) \xrightarrow{d} (U_2^{(1,2)}, U_3^{(1,2)}), \]

\[ U_2^{(1,2)} = \sqrt{8}\epsilon_2 \sum_{p=1}^{\infty} (\Delta X_{\tau_p}^{(1)})^2 (\Delta X_{\tau_p}^{(2)}) Z_p^{(2)} 1(\tau_p \leq 1), \]

\[ U_3^{(1,2)} = \sqrt{8}\epsilon_1 \sum_{p=1}^{\infty} (\Delta X_{\tau_p}^{(1)}) (\Delta X_{\tau_p}^{(2)})^2 Z_p^{(1)} 1(\tau_p \leq 1). \]

For the first step of the proof, we prove our result in a simple case when the process \( X \) has at most finite jumps in the interval \([0, 1]\). Then we prove the general case when \( X \) may have infinite jumps in the interval \([0, 1]\).

(Step1): In this step, we introduce some notations. Let \( (T_p) \) be the reordering of the double sequence \( (T(m, j) : m, j \geq 1) \). A random variable \( T(m, j) \) is the successive jump time of the Poisson process \( \{ A_m \}_{m \in \mathbb{N}} \). Where \( A_m = \{ z : \Gamma(z) > 1/m \} \). Let \( P_m \) denote the set of all indices \( p \) such that \( T_p = T(m', j) \) for some \( j \geq i \) and some \( m' \leq m \). For \( (i - 1)k\Delta_n < T_p \leq ik\Delta_n \), we define following random variables:

\[ W_{-}^{(k,m)}(n, p) = \frac{1}{\sqrt{k\Delta_n}} (X^{(m)}_{T_p -} - X^{(m)}_{(i-1)k\Delta_n}), \]

\[ W_{+}^{(k,m)}(n, p) = \frac{1}{\sqrt{k\Delta_n}} (X^{(m)}_{ik\Delta_n} - X^{(m)}_{T_p}), \]

\[ W^{(k,m)}(n, p) = W_{-}^{(k,m)}(n, p) + W_{+}^{(k,m)}(n, p). \]

We also set following stochastic processes:
\[ b(m)_t = b_t - \int_{A_m \cap \{ z : \| \delta(t, z) \| \leq 1 \}} \delta(t, z) \lambda(z), \]
\[ X(m)_t = X_0 + \int_0^t b(m)_s ds + \int_0^t \sigma_s dW_s + (\delta_1 A_m) * (\mu - v)_t, \]
\[ X'(m) = X - X(m) = (\delta_1 A_m) * \mu. \]

Let \( \Omega_n(m) \) denote the set of all \( \omega \) such that each interval \([0, 1] \cap ((-1)k \Delta_n, k \Delta_n]\) contains at most one jump of \( X'(m) \), and that \( ||X(m)(\omega)_{t+s} - X'(m)(\omega)_t|| \leq 2/m \) for all \( t \in [0, 1] \) and \( s \in [0, \Delta_n] \). Moreover, let
\[
\eta^{n,k}_p = \left( \Delta X^{(1)}_{T_p} + \sqrt{k \Delta_n} W^{(k,1)}(n, p) \right)^2 \left( \Delta X^{(2)}_{T_p} + \sqrt{k \Delta_n} W^{(k,2)}(n, p) \right),
\]
\[
\tilde{\eta}^{n,k}_p = \sqrt{\xi_2} \eta^{n,k}_p (\Delta v^{(k,2)}_i).
\]

By the above notations, on the set \( \Omega_n(m) \), we have
\[
\sqrt{n} U_{2,1,n} = 2 \sqrt{n} R^{(n)}_{2,2,1}(X(m)) + 2 \sqrt{n} R^{(n)}_{2,2,1}(X'(m)). \quad (22)
\]

(Step 2) : In this step, we will prove the stable convergence of \( U_{2,1,n} \) and \( U_{2,2,n} \). A Taylor expansion of \( f(x_1, x_2) = x_1^2 x_2 \) yields \( \eta^{n,k}_p - (\Delta X^{(1)}_{T_p})^2 (\Delta X^{(2)}_{T_p})^p \rightarrow 0 \). By Proposition 4.4.10 in Jacod & Protter (2012), we have
\[
\left( \eta^{n,k}_p \right)_{p \geq 1} \xrightarrow{\mathcal{L}} \left( \Delta X^{(1)}_{T_p} \right)^2 \Delta X^{(2)}_{T_p}. \quad (23)
\]

The sequence \( (\Delta v^{(k,2)}_i) \) for \( k = 1, 2 \) consists of correlated Gaussian random variables with mean 0 and has the covariance structure
\[
\text{Cov}(\Delta v^{(k,2)}_i, \Delta v^{(k,2)}_j) = \begin{cases} 2 & \text{if } i = j, \\ -1 & \text{if } |i - j| = 1, \\ 0 & \text{if } |i - j| > 1. \end{cases}
\]

Using the inequalities (18), (19) and (20), if \( |i - j| \geq 1 \), then we have \( E[|\xi^n_i \xi^n_j|] \leq K \Delta_n^{3/2} \) where \( \xi^n_i = (\Delta_i X^{(k,1)}_n)^2 ((\Delta_i X^{(k,2)}_n)(\Delta v^{(k,2)}_i)) \). Therefore, the correlation between \( \xi^n_i \) and \( \xi^n_j \) when \( |i - j| = 1 \) is asymptotically negligible. Because the set \( \{ T_p : p \in \mathcal{P}_m \} \cap [0, 1] \) is finite, it follows that
\[
2 \sqrt{n} R^{(n)}_{2,2,1}(X'(m)) \xrightarrow{\mathcal{L}} \sqrt{8 \xi_2} \sum_{p \in \mathcal{P}_m : T_p \leq 1} (\Delta X^{(1)}_{T_p})^2 (\Delta X^{(2)}_{T_p}) Z^{(2)}_p \equiv R_{2,2,1}(X'(m)), \quad (24)
\]

where \( Z^{(2)}_p \) is the sequence of i.i.d. standard normal random variables introduced before (Step 1).

(Step 3) : We will prove the joint stable convergence of \( (U_{1,n} - U_0, U_{2,1,n}, U_{2,2,n}) \) in this step. From a slight modification of the proof of Theorem 5.1.2 in Jacod & Protter (2012), it is possible to prove
\[ R_{2,2,1}(X'(m)) \xrightarrow{p} U_2^{(1,2)}, \]  

(25)

\[
\lim_{m \to \infty} \limsup_{n \to \infty} \mathbb{P} \left( \Omega_n(m) \cap \left\{ |\sqrt{n} R_{2,2,1}^{(n)}(X(m))| > \eta \right\} \right) = 0, \quad \text{for all } \eta > 0. \]  

(26)

Combining these with (22) and (24), and by proposition 2.2.4 in Jacod & Protter (2012), we obtain \( \sqrt{n} U_{2,1,n} \xrightarrow{\mathbb{L}} U_3^{(2,1)} \). We also have \( \sqrt{n} U_{2,2,n} \xrightarrow{\mathbb{L}} U_3^{(1,2)} \) in the same way. Since \( U_{1,n}, U_{2,1,n} \) and \( U_{2,2,n} \) are \( \mathcal{F}^{(0)} \)-conditionally mutually independent by the definition of small noise, we obtain the joint convergence

\[
\sqrt{n}(U_{1,n} - U_0, U_{2,1,n}, U_{2,2,n}) \xrightarrow{\mathbb{D}} (U_{1,k}^{(1,2)}, U_2^{(1,2)}, U_3^{(1,2)}).
\]

Therefore, we also obtain \( \sqrt{n}(U_{1,n} + U_{2,1,n} + U_{2,2,n} - U_0) \xrightarrow{\mathbb{D}} U_{1,k}^{(1,2)} + U_2^{(1,2)} + U_3^{(1,2)} \).

**Evaluation of the remaining terms**: We shall prove that the remaining terms \( \sum_{j=1}^{3} U_{3,j,n} \) and \( \sum_{j=1}^{3} U_{4,j,n} \) in the decomposition of \( S_{k,2,2}^{(n)}(Y) \) in (21) are asymptotically negligible. Let \( \xi_i^n = e_2^n (\Delta_i X^{(k,1)})^2 (\Delta_i v_i^{(k,2)})^2 \). Then \( U_{3,1,n} = \sum_{i=1}^{n/k} \xi_i^n \). Using the inequalities (18), (19) and (20), we have

\[
E[|\xi_i^n|] = e_2^n E \left[ (\Delta_i v_i^{(k,2)})^2 \right] E[ (\Delta_i X^{(k,1)})^2 ] \leq K \Delta_n^2.
\]

Therefore, we obtain \( U_{3,1,n} = O_P(1/n) \). Similarly, we have \( U_{3,2,n} = U_{3,3,n} = U_{4,1,n} = U_{4,2,n} = O_P(1/n) \). Because \( n^{-1} \sum_{i=1}^{n/k} (\Delta_i v_i^{(k,1)})^2 (\Delta_i v_i^{(k,2)})^2 \xrightarrow{p} 4\xi_1\xi_2 \), we have \( U_{4,3,n} = O_P(1/n\sqrt{n}) \). Consequently, we obtain \( \sqrt{n}(S_{k,2,2}^{(n)}(Y) - U_0) \xrightarrow{\mathbb{D}} U_{1,k}^{(1,2)} + U_2^{(1,2)} + U_3^{(1,2)} \).

\[ \square \]

**Proof of theorem 2**: We first prove a technical tool for the proof of theorem 2. The result of theorem 2 follows immediately from the following proposition.

**Proposition 7**: Suppose assumptions 1, 2 and 3 are satisfied with \( q = 0 \) in assumption 2. Let

\[
S_0 = \sum_{0 \leq s \leq 1} (\Delta X_s^{(1)})^2 (\Delta X_s^{(2)})^2,
\]

\[
W^{(n)} = \sqrt{n} \left[ (S_{2,2,2}^{(n)}(Y) - S_0) - (S_{1,2,2}^{(n)}(Y) - S_0) \right].
\]

Then, in restriction to the set \( \Omega^{(j)} \), we have

\[
W^{(n)} \xrightarrow{\mathbb{L}} U_{1,1}^{(1,2)} + U_2^{(1,2)} + U_3^{(1,2)}.
\]

**Proof of proposition 7**: From the result of theorem 1, for \( k = 1, 2 \), we have

\[
S_{k,2,2}^{(n)}(Y) = S_{k,2,2}^{(n)}(X) + 2R_{k,1,2}^{(n)}(X) + 2R_{k,2,1}^{(n)}(X) + o_P(1/\sqrt{n}),
\]

where \( R_{k,r,s}^{(n)}(X) \) are defined in (10). We decompose \( W^{(n)} \) into three leading terms.
\[ W^{(n)} = \sqrt{n} (S_{2.2.2}^{(n)}(X) - S_{1.2.2}^{(n)}(X)) + 2\sqrt{n} (R_{2.1.2}^{(n)}(X) - R_{1.1.2}^{(n)}(X)) \\
+ 2\sqrt{n} (R_{2.2.1}^{(n)}(X) - R_{1.2.1}^{(n)}(X)) + o_P(1) \\
= I_{n.1} + II_{n.2} + III_{n.3} + o_P(1). \]

Because of the independence of the noise \( v_i^{(1)} \) and \( v_i^{(2)} \), the three terms \( I_{n.1}, II_{n.1} \) and \( III_{n.1} \) are asymptotically mutually independent. Therefore, it suffices to evaluate each term. We can rewrite \( II_{n.2} \) by using the estimation inequalities (18), (19) and (20),

\[
II_{n.2} = 2\sqrt{\frac{n}{k}} \sum_{i=1}^{[n/2]} \sum_{j=2(i-1)+1}^{2i} (\Delta v_i^{(1)}(X))(\Delta v_j^{(2)}X^2)^2 \left( \sum_{l=2(i-1)+1}^{2i} \Delta v_i^{(1)} - \Delta v_j^{(1)} \right) \\
+ O_P(1/\sqrt{n}) = \tilde{II}_{n.2} + O_P(1/\sqrt{n}).
\]

By the application of proposition 15.3.2 in Jacod & Protter (2012) to \( I_{n.1} \) and the similar argument of the evaluation of \( U_{2,1,n} \) in the proof of theorem 1, we obtain that \( W^{(n)} \) converge stably to \( U_4^{(1,2)} + U_2^{(1,2)} + U_3^{(1,2)} \).

Finally, we prove the stable convergence of \( T^{(n)} \). From the definition of \( T^{(n)} \), we have

\[
T^{(n)} - 1 = \frac{S_{2.2.2}^{(n)}(Y) - S_{1.2.2}^{(n)}(Y)}{S_{1.2.2}^{(n)}(Y)}.
\]

Then by proposition 7, we obtain the desired result.

**Proof of proposition 7.** Because the asymptotic size of the test statistic \( T^{(n)} \) follows from theorem 2, we prove the consistency of the test here. To describe the limit variable of \( T^{(n)} \) under the alternative hypothesis \( \mathbb{H}_1 \), we introduce some notations. We use some notations in Jacod & Todorov (2009).

(i) a sequence \( (\kappa_p) \) of uniform variables on \([0, 1]\).

(ii) a sequence \( (L_p) \) of uniform variables on \([0, 1]\), that is, \( P(L_p = 0) = P(L_p = 1) = 1/2 \).

(iii) four sequences \( (U_p), (U'_p), (\tilde{U}_p), (\tilde{U}'_p) \) of two-dimensional \( N(0, I_2) \) variables.

(iv) two sequences \( (\tilde{Z}_p), (\tilde{Z}'_p), (\tilde{Z}''_p) \) of two-dimensional \( N(0, I_2) \) variables.

The variables introduced above are defined on \( (\Omega^{(1)}, \mathcal{F}^{(1)}, (\mathcal{F}^{(1)}_t)_{t \geq 0}, P^{(1)}) \). Then we define two dimensional variables:

\[
R_p = \sigma_T \left( \sqrt{\kappa_q} U_p + \sqrt{1 - \kappa_p} U'_p \right), \\
R'_p = \sigma_T \left( \sqrt{L_q} \tilde{U}_p + \sqrt{1 - L_p} \tilde{U}'_p \right), \\
R''_p = R_p + R'_p.
\]
We also define following variables:

\[
\begin{align*}
D'' &= \sum_{p : T_p \leq 1} \left( (\Delta X^{(1)}_{T_p} R^{(2)}_p) + (\Delta X^{(2)}_{T_p} R''_p) \right), \\
D &= \sum_{p : T_p \leq 1} \left( (\Delta X^{(1)}_{T_p} R^{(2)}_p) + (\Delta X^{(2)}_{T_p} R''_p) \right), \\
L'' &= \sum_{p : T_p \leq 1} \left( \sqrt{\xi_2 (\Delta X^{(1)}_{T_p} R^{(2)}_p) \tilde{Z}_p} + \sqrt{\xi_1 (\Delta X^{(2)}_{T_p} R''_p) \tilde{Z}_p} \right), \\
L &= \sum_{p : T_p \leq 1} \left( \sqrt{\xi_2 (\Delta X^{(1)}_{T_p} R^{(2)}_p) \tilde{Z}_p} + \sqrt{\xi_1 (\Delta X^{(2)}_{T_p} R''_p) \tilde{Z}_p} \right), \\
H &= \int_0^1 \left( c_s^{(1,1)} c_s^{(2,2)} + 2(c_s^{(1,2)})^2 \right) ds.
\end{align*}
\]

By using (18), (19) and (20), in restriction to the set \( \Omega^{(d)} \), we have

\[
S_{k, 2, 2}^{(n)}(Y) = S_{k, 2, 2}^{(n)}(X) + 2R_{k, 2, 1}^{(n)}(X) + 2R_{k, 1, 2}^{(n)}(X) + o_P(1/n),
\]

where \( R_{k, r, s}^{(n)}(X) \) is defined in (10). Moreover, from similar argument of the proof of theorem 3.1 in Jacod & Todorov (2009), in restriction to the set \( \Omega^{(d)} \), it is possible to have

\[
n \times (S_{k, 2, 2}^{(n)}(X), S_{1, 2, 2}^{(n)}(X), R_{2, 2, 1}^{(n)}(X) + R_{2, 1, 2}^{(n)}(X), R_{1, 2, 1}^{(n)}(X) + R_{1, 1, 2}^{(n)}(X)) \xrightarrow{\mathcal{L}} (D'' + 2H, \hat{D} + H, L'', L).
\]

Therefore, we obtain

\[
T(n) \xrightarrow{\mathcal{L}} \Phi = \frac{\hat{D}'' + 2H + 2L''}{\hat{D} + H + 2L}.
\]

Because \((\hat{D}'', \hat{D}, L'', L)\) admits a density conditionally on \( \mathcal{F} \) and on being in \( \Omega^{(d)} \), \( T(n) \neq 1 \) a.s. on \( \Omega^{(d)} \).

\[\Box\]

A.3 Proofs for section 5

\textit{Proof of proposition 5.} For the consistency of the SIML, refer to the proof of theorem 1 in Kunitomo & Sato (2013).

\[\Box\]

\textit{Proof of proposition 6.} We only give the proof of a consistency of \( \hat{D}_{p,q}^{(l,m)}(1, 1) \) and \( \hat{A}^{(l,m)}(r) \). The proofs of consistency of the other estimators are similar. Under the assumptions of Proposition 6, for any \( \eta > 0, \alpha > 0, \theta \in (0, 1/2) \), we have

\[
P \left( \sup_{1 \leq t \leq n} \frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbb{1}_{\{\epsilon_{j,n} M_i^{(j)} > \alpha \Delta_n^j \} > \eta} \right) \leq P \left( \sum_{i=1}^n \mathbb{1}_{\{\epsilon_{j,n} M_i^{(j)} > \alpha \Delta_n^j \} > \eta} \right)
\]

\[
\leq P \left( \bigcup_{i=1}^n \left\{ \epsilon_{j,n} |M_i^{(j)}| > \alpha \Delta_n^j \right\} \right) \leq \sum_{i=1}^n P \left( \epsilon_{j,n} |M_i^{(j)}| > \alpha \Delta_n^j \right)
\]

\[
\leq 4 \sum_{i=1}^n \max \left\{ P \left( \epsilon_{j,n} M_i^{(j)} > \frac{\alpha \Delta_n^j}{2} \right), P \left( \epsilon_{j,n} M_i^{(j)} < -\frac{\alpha \Delta_n^j}{2} \right) \right\} = o(1).
\]
Then \( \sup_{1 \leq i \leq n} \left| 1_{\{|\Delta_i^n X| > \alpha \Delta_i^n\}} - 1_{\{|\Delta_i^n X| > \alpha \Delta_i^n/2\}} \right| \stackrel{p}{\rightarrow} 0. \) Therefore, we obtain

\[
\sum_{i=1}^{n} (\Delta_i^n Y^{(1)})(\Delta_i^n Y^{(2)}) 1_{\{|\Delta_i^n Y| > \alpha \Delta_i^n\}} - \sum_{i=1}^{n} (\Delta_i^n Y^{(1)})(\Delta_i^n Y^{(2)}) 1_{\{|\Delta_i^n X| > \alpha \Delta_i^n/2\}} \stackrel{p}{\rightarrow} 0.
\]

Moreover, for \( \eta > 0, \)

\[
P \left( \epsilon_{1,n} \epsilon_{2,n} \sum_{i=1}^{n} (\Delta_i^n X^{(1)})(\Delta_i^n X^{(2)})(\Delta_i^n v^{(1)})(\Delta_i^n v^{(2)}) 1_{\{|\Delta_i^n X| > \alpha \Delta_i^n\}} > \eta \right)
\]

\[
\leq \left( \frac{\epsilon_{1,n} \epsilon_{2,n}}{\eta^2} \right)^2 E \left[ \left( \sum_{i=1}^{n} (\Delta_i^n X^{(1)})(\Delta_i^n X^{(2)})(\Delta_i^n v^{(1)})(\Delta_i^n v^{(2)} \right)^{2} \right]
\]

\[
\leq \left( \frac{\epsilon_{1,n} \epsilon_{2,n}}{\eta^2} \right)^2 \left( \sum_{i=1}^{n} A_{i,1}^{1,n} + \sum_{i=2}^{n} A_{i,2}^{2,n} + \sum_{i=1}^{n} A_{i,3}^{3,n} \right),
\]

where

\[
A_{i,1}^{1,n} = E \left[ (\Delta_i^n X^{(1)})^2 (\Delta_i^n X^{(2)})^2 (\Delta_i^n v^{(1)})^2 (\Delta_i^n v^{(2)})^2 \right],
\]

\[
A_{i,2}^{2,n} = E \left[ (\Delta_i^n X^{(1)})(\Delta_{i-1}^n X^{(1)})(\Delta_i^n X^{(2)})(\Delta_{i-1}^n X^{(2)})(\Delta_i^n v^{(1)})(\Delta_{i-1}^n v^{(1)})(\Delta_i^n v^{(2)})(\Delta_{i-1}^n v^{(2)}) \right],
\]

\[
A_{i,3}^{3,n} = E \left[ (\Delta_{i+1}^n X^{(1)})(\Delta_i^n X^{(1)})(\Delta_{i+1}^n X^{(2)})(\Delta_i^n X^{(2)})(\Delta_{i+1}^n v^{(1)})(\Delta_i^n v^{(1)})(\Delta_{i+1}^n v^{(2)})(\Delta_i^n v^{(2)}) \right].
\]

Now we evaluate the last three terms in the above inequality. By Hölder’s inequality and inequalities (18), (19) and (20), we have \( A_{i,1}^{1,n} = A_{i,2}^{2,n} = A_{i,3}^{3,n} = O(n^{-2}). \) Therefore,

\[
\epsilon_{1,n} \epsilon_{2,n} \sum_{i=1}^{n} (\Delta_i^n X^{(1)})(\Delta_i^n X^{(2)})(\Delta_i^n v^{(1)})(\Delta_i^n v^{(2)}) 1_{\{|\Delta_i^n X| > \alpha \Delta_i^n\}} \stackrel{p}{\rightarrow} 0.
\]

From the similar argument, for \( 1 \leq l, m \leq 2, \) we have

\[
\epsilon_{m,n}^2 \sum_{i=1}^{n} (\Delta_i^n X^{(l)})(\Delta_i^n v^{(m)})^2 1_{\{|\Delta_i^n X| > \alpha \Delta_i^n\}} \stackrel{p}{\rightarrow} 0.
\]

\[
\epsilon_{m,n} \sum_{i=1}^{n} (\Delta_i^n X^{(l)})(\Delta_i^n X^{(m)})(\Delta_i^n v^{(m)}) 1_{\{|\Delta_i^n X| > \alpha \Delta_i^n\}} \stackrel{p}{\rightarrow} 0.
\]

\[
\epsilon_{1,n} \epsilon_{m,n}^2 \sum_{i=1}^{n} (\Delta_i^n X^{(l)})(\Delta_i^n v^{(l)})(\Delta_i^n v^{(m)}) 1_{\{|\Delta_i^n X| > \alpha \Delta_i^n\}} \stackrel{p}{\rightarrow} 0.
\]

Hence, we have

\[
\sum_{i=1}^{n} (\Delta_i^n Y^{(1)})(\Delta_i^n Y^{(2)}) 1_{\{|\Delta_i^n Y| > \alpha \Delta_i^n\}} - \sum_{i=1}^{n} (\Delta_i^n Y^{(1)})(\Delta_i^n Y^{(2)}) 1_{\{|\Delta_i^n X| > \alpha \Delta_i^n/2\}} \stackrel{p}{\rightarrow} 0.
\]

Then \( \sum_{i=1}^{n} (\Delta_i^n Y^{(1)})(\Delta_i^n Y^{(2)}) 1_{\{|\Delta_i^n Y| > \alpha \Delta_i^n\}} \stackrel{p}{\rightarrow} \sum_{0 \leq s \leq 1} (\Delta X_s^{(1)})(\Delta X_s^{(2)}). \) Therefore, from theorems 9.4.1 and 9.5.1 in Jacod & Protter (2012), we have
\[
\sum_{i=1}^{n} c_{i}^{(1,2)} (\Delta_{i}^{n} Y^{(1)})(\Delta_{i}^{n} Y^{(2)}) 1_{i,|\Delta_{i}^{n} Y^{(1)}|>\alpha \Delta_{i}^{n}} \xrightarrow{p} D_{1,2}^{(1,2)} (1, 1), \\
\Delta_n \sum_{i=1}^{n-k_n+1} (c_{i}^{(1,2)})^{r} \xrightarrow{p} A^{(1,2)}(r).
\]

Therefore, we obtain the desired results.

**Proofs of corollaries 1 and 2.** The result immediately follows from the remarks in section 3, theorem 2, and propositions 5 and 6.