VOLATILITY AND JUMPS IN HIGH FREQUENCY FINANCIAL DATA: ESTIMATION AND TESTING

by

Nan Zhou

B.S. Mathematics, Zhejiang University, Hangzhou, China 2006

M.A. Statistics, University of Pittsburgh, Pittsburgh, PA 2007

Submitted to the Graduate Faculty of

the Arts & Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH

ARTS & SCIENCES

This dissertation was presented

by

Nan Zhou

It was defended on

April 18th 2011

and approved by

Satish Iyengar, Professor, Statistics

Jean-François Richard, Professor, Economics

Leon Gleser, Professor, Statistics

Robert Krafty, Assistant Professor, Statistics

Dissertation Director: Satish Iyengar, Professor, Statistics

ABSTRACT

VOLATILITY AND JUMPS IN HIGH FREQUENCY FINANCIAL DATA: ESTIMATION AND TESTING

Nan Zhou, PhD

University of Pittsburgh, 2011

It has been widely accepted in financial econometrics that both the microstructure noise and jumps are significantly involved in high frequency data. In some empirical situations, the noise structure is more complex than independent and identically distributed (i.i.d.) assumption. Therefore, it is important to carefully study the noise and jumps when using high frequency financial data. In this dissertation, we develop several methods related to the volatility estimation and testing for jumps.

Chapter 1 proposes a new method for volatility estimation in the case where both the noise level and noise dependence are significant. This estimator is a weighted combination of sub-sampling realized covariances, constructed from discretely observed high frequency data. It is proved to be a consistent estimator of quadratic variation in the case with either i.i.d. or dependent noise. It is also shown to have good finite-sample properties compared with existing estimators in the literature.

Chapter 2 focuses on the testing for jumps based on high frequency data. We generalize the methods in Aït-Sahalia and Jacod (2009a) and Fan and Fan (2010). The generalized method allows more flexible choices for the construction of test statistics, and has smaller asymptotic variance under both null and alternative hypotheses. However, all these methods are not effective when the microstructure noise is significant. To reduce the influence from noise, we further design a new statistical test, which is robust with the i.i.d. microstructure noise. This new method is compared with the old tests through Monte Carlo studies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I am deeply grateful to my advisor, Satish Iyengar, for his constant encouragement and guidance, for his understanding and support of my works in both academia and industry, and for sharing with me his insights and wisdom over the past five years.

I am grateful to Leon Gleser, Robert Krafty and Jean-François Richard for being on my committee and for their valuable suggestions and guidance through this process.

I also want to thank John Chadam, from whom I studied these topics in mathematical finance. Further thanks go to Yingying Li. It is her talk in the Carnegie Mellon University that inspires me to find my interest in the research of high frequency financial data.

I thank all the faculty, staff and students of the Department of Statistics at the University of Pittsburgh. I feel fortunate to study and pursue my doctoral degree in such a friendly environment.

Finally, to my parents, to whom this thesis is dedicated, goes my deepest appreciation. I want to give a special thanks to my mother for her endless love and her faith in me. I would never accomplish my goals without her standing next to me in all the good and bad times of my life. I know this is not the end but the start.

1.0 SUB-SAMPLING REALIZED VOLATILITY ESTIMATION USING HIGH-FREQUENCY DATA WITH DEPENDENT NOISE

1.1 INTRODUCTION

In financial econometrics, the modeling of volatility has been an important topic. The real-time estimates and forecasts of volatility based on discretely observed data are essential in many practical applications, like the pricing of financial instruments, portfolio allocations, performance evaluation, and risk management. While the price process of financial instruments is usually observable, the volatility is always latent, and thus brings more complexity to the study of volatility.

A classical method to deal with this fundamental latency of volatility is by building parametric models with some strong but necessary assumptions. These models include Auto Regressive Conditional Heteroskedasticity (ARCH) (e.g. Engle (1982)), Stochastic Volatility Model (e.g. Heston (1993) and Hagan et al. (2002)), and Local Volatility Model (e.g. Dupire (1994) and Derman et al. (1996)). Other related work is in Andersen et al. (2002), Chernov et al. (2003), Eraker et al. (2003), etc. An alternative approach is to derive the 'Implied Volatility' from market prices of derivative products. See the papers by Bates (1996) and Garcia et al. (2004).

In the last decade, the wide availability of reliable high frequency financial data has led to substantial improvement in the study of volatility. One popular application using high frequency data is to estimate the quadratic variation (QV), which is the integral of the squared volatility over a fixed time interval as in section 1.2.2. A classic estimator is Realized Volatility (RV), which is the sum of the frequently sampled squared returns (e.g. Andersen et al. (2001), Meddahi (2002) and Barndorff-Nielsen and Shephard (2002)). A weakness of this estimator is its high sensitivity to market microstructure noise when applied to very high frequency data such as 1 minute or less (e.g. Zhou (1996), Fang (1996) and Andersen et al. (2000)). The empirical evidence of microstructure noise is discussed in the beginning of section 1.3.

To reduce the bias introduced by microstructure noise, the classical solution uses moderate high frequency data, which is normally chosen between 5 to 30 minutes (see Bandi and Russell (2003)). However, this kind of solution uses less than one percent of available data, and thus results in very inefficient estimation. Recently, some prominent approaches are proposed to design new statistical estimators based on high frequency data, which are consistent estimators and are robust to noise in the data. Roughly, there are three main trends: Zhang et al. (2005, 2006)'s two-scales and multi-scales Realized Volatility (TSRV, MSRV), Barndorff-Nielsen et al. (2008, 2009, 2011a, 2011b)'s Realized Kernel, and Jacod et al. (2009)'s Pre-averaging approach. All of these approaches could construct consistent and efficient estimators, which converge to the true volatility at a rate of $n^{-1/4}$. This is the best attainable convergence rate even in the simplest parametric model by the maximum likelihood estimation as we show in section 1.3.1.

Most of these nonparametric approaches assume the noise is i.i.d. However, as studied in section 1.4.1, the dependence among the microstructure noise could be significant in some empirical situations. For this case, Aït-Sahalia et al. (2011) generalizes the TSRV into a sub-sampling version, which uses two sparse scales. The generalized TSRV becomes consistent in the case with dependent noise, as the number of sub-sampling interval increases to infinite.

In this paper, we develop a new estimator called SRC, which is a weighted combination of sub-sampling realized covariances with different lags, constructed from high frequency data. For lag $= 0$, the realized covariance converges in probability to quadratic variation plus the bias that depends on both noise variance and noise covariance. When the lag is greater than 0, the quadratic variation disappears in the asymptotic mean, and the mean of the realized covariance is related to the noise covariance with different lags. Therefore, choosing some specific weight function, the combination of sub-sampling realized covariances with different lags could converge in probability to the quadratic variation. The asymptotic properties of SRC and the central limit theorem are studied in section 1.5. Through the Monte Carlo simulations, this new estimator is shown to have better finite-sample performance compared with the existing methods, especially when the noise dependence is not small.

The rest of this chapter is organized as follows. Section 1.2 describes the model assumptions and necessary notations. Section 1.3 reviews the TSRV when assuming the microstructure noise is i.i.d. In section 1.4, we empirically study the noise dependence structure, based on the transactions data of 30 Dow Jones Industrials Average (DJIA) stocks. We also review the generalization of TSRV for the case with dependent noise. We develop the new estimators and study their asymptotic properties in section 1.5. The finite-sample performance of the new estimators are studied by simulations based on different noise levels and sample sizes in section 1.6. The empirical analysis is provided in the section 1.7. Section 1.8 concludes with directions for future work.

1.2 SEMIMARTINGALE AND QUADRATIC VARIATION

1.2.1 Price Process

The fundamental theory of asset prices in the frictionless arbitrage free market requires that the log-price process X_t follows a semimartingale on a filtered probability space: $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$. The most familiar semimartingale is Brownian semimartingale without jumps:

Assumption 1.1.

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s,$$

where b_t is a predictable locally bounded drifted function, σ_t is an adapted cadlag volatility process, and W_t is standard Brownian Motion.

To derive some asymptotic results, we need some further reasonable assumptions on σ:

Assumption 1.2. *σ_t does not not vanish and it satisfies:*

$$\sigma_t = \sigma_0 + \int_0^t b'_s ds + \int_0^t \sigma'_s dW'_s,$$

where b'_t and σ'_t are adapted cadlag function. W'_t is another Brownian Motion, which could be correlated with W_t.

Assumption 1.2 is fulfilled for many financial models in the literature, and it simplifies the proofs in this paper considerably. To find a more general treatment, including the case of volatility with jumps, discussions could be found in Barndorff-Nielsen et al. (2006) and Jacod (2007).

1.2.2 Quadratic Variation and Realized Volatility

Over a fixed time interval $[0, T]$, which is typically several days in practical applications, high frequency data are observed and recorded for a sequence of deterministic partitions $0 = t_0 \leq t_1 \leq \ldots \leq t_n = T$. To focus on the core issue, we suppose that the data are equally distributed: $t_i - t_{i-1} = \Delta_n = [t/n]$, which might be 1 hour, 1 minute or smaller. This equality assumption does not influence the asymptotic mean of the estimators in this paper, but only changes the asymptotic variance by a constant scale. A more natural way is to work with financial data observed in real tick time, which allows the spacing to be stochastic and endogenous. The study of stochastic transaction time could be found in section 5.3 in Barndorff-Nielsen et al.(2008). To simplify notation, we write Δ instead of Δ_n, and denote $X_i = X_{t_i}$ and $\Delta_i X = X_i - X_{i-1}$.

Quadratic Variation (QV):

A key quantitative measurement of the price process is the quadratic variation:

$$QV(X) = \int_0^T \sigma_s^2 \, ds. \tag{1.1}$$

From the probabilistic view, the QV could also be defined as

$$QV(X) = p \lim_{\overrightarrow{n \to \infty}} \sum_{i=1}^{t_i \leq T} (X_i - X_{i-1})^2, \text{ as } \max_i \{t_i - t_{i-1}\} \to 0. \tag{1.2}$$

Here,

$$A = p \lim_{\overrightarrow{n \to \infty}} A_n \text{ denotes } A_n \text{ converges in probability to } A.$$

This definition could be found in section 5.5 in Casella and Berger (2002).

Realized Volatility (RV):

A typical and intuitive method to estimate the QV is the RV:

$$[X,X]_t^n = \sum_{i=1}^{n=[t/\Delta]} \left(X_{i\Delta} - X_{(i-1)\Delta}\right)^2, \tag{1.3}$$

which has the following asymptotic properties:

$$\begin{aligned} [X,X]_t^n &= \int_0^t \sigma_s^2 \, ds + O_p(n^{-1/2}), \\ n\Big([X,X]_t^n - \int_0^t \sigma_s^2\Big) &\xrightarrow{L_s} N\Big(0, 2t^2 \int_0^t \sigma_s^4 \, ds\Big). \end{aligned} \tag{1.4}$$

The advantage of this estimator is obvious: it is model free, unbiased and consistent under mild conditions. These properties are independently discussed by Andersen and Bollerslev (1998), Comte and Renault (1998), and Barndorff-Nielsen and Shephard (2001, 2002a, 2002b). Theoretical and empirical properties of the RV have also been studied in numerous articles (see Jacod (1994), Jacod and Protter (1998), Andersen et al. (2001), Barndorff-Nielsen and Shephard (2002), and Mykland and Zhang (2006)). The multivariate generalizations to realized covariation were discussed in Andersen et al.(2003) and Barndorff-Nielsen and Shephard (2004).

1.3 TSRV WITH IID MICROSTRUCTURE NOISE

From (1.4), RV is an unbiased estimator with asymptotic variance $\frac{2t^2}{n}\int_0^t \sigma_s^4 \, ds$, which is decreasing with the sample size. Therefore, we would like to use the available data as frequently as possible to reduce the estimation error. However, empirical study shows that the RV is unacceptably sensitive to market frictions when using ultra high frequency data over time intervals such as 1 minute or less.

The existence of microstructure noise could be easily illustrated by the volatility signature plot (see Andersen (2009) and Aït-Sahalia et al. (2011)) which is the plot of RV estimator vs. different time frequencies (Δ_n). In Figure 1.3, we create the volatility signature plots based on the one year transaction data of SPY from Jan 2001 to Jan 2002, which is collected from the NYSE Trade and Quote (TAQ) database. SPY is an actively traded exchange-traded fund (ETF), and it represents an ownership in a portfolio of the equity securities that comprise the Standard & Poor's 500 Index, which usually be regarded as the overall market benchmark. It is obvious from Figure 1.3 that the RV diverges with the decreasing of sampling freuquency at a rate proportional to $1/\Delta_n$ instead of converging to a constant, which is expected to be the integrated volatility as in (1.4).

To mathematically discuss the potential influence from market microstructure noise in high frequency data, we start from a common and simple assumption that the observed log price Y_i in high frequency data is the unobservable efficient log price X_i contaminated by some noise component as another independent process E_i due to imperfections of the trading procedure:

Assumption 1.3. *X_t is the underlying unobservable log-price process, and we can observe the process*

$$Y_t = X_t + E_t,$$

where E is independent of X ($E \perp X$).

This independence assumption was questionable from a market microstructure theory viewpoint (e.g., Kalnina and Linton (2008)). However, the empirical work of Hansen and Lunde (2005) suggests that this assumption is not too damaging statistically when we analyze high frequency data.

Assumption 1.4. *We mostly work under a white noise assumption:*

$$\mathbb{E}[E] = 0, \quad Var[E], \quad Var[E^2] < \infty, \quad \text{and } E_t \perp E_s.$$

A feature of white noise is that $[E, E]_t = \infty$. Thus white noise does not belong to the semimartingale, which means the market with noise would allow arbitrage opportunities from an econometrics view.

Then instead of (1.4), we get:

$$[Y,Y]_t^n = 2n\mathbb{E}(E^2) + \int_0^t \sigma_s^2 \, ds + O_p(\sqrt{n}). \tag{1.5}$$

According to the result in (1.5), we expect to have $\ln([Y,Y]_t^n) \approx \ln(2\mathbb{E}E^2) + \ln(n)$. So a regression of $\ln([Y,Y]_t^n)$ on $\ln(n)$ should have slope coefficient close to 1, and intercept close to $\ln(2\mathbb{E}E^2)$. Figure 1.5 shows the empirical result from the transaction records of 30 DJIA stocks over the last 10 trading days in April 2004: the estimated slope is equal to 1.02, and the null value of 1 is not rejected.

The model in (1.3) and the result in (1.5) are both theoretically and empirically studied in Aït-Sahalia et al. (2005), Zhang et al. (2005), Zhang (2006), and Bandi and Russell (2004). The study of a more general noise structure is in Jacod (1996), Delattre and Jacod (1997), and Li and Mykland (2007).

Remark 1.1. Numerical facts of Microstructure Noise:
To approximately estimate and compare the true integrated volatility and microstructure noise, we use the data of "VIX", which is the Chicago Board Options Exchange Volatility Index. VIX represents a measure of the market's expectation of the (annualized) implied volatility of the S&P 500 index over the next 30-day period. The VIX Index was introduced by Whaley (1993). The simple average of the VIX over the last ten trading days in April 2004 is 16.18 as show in the Figure 1.4, which means the annualized $\sigma_s \approx \frac{16.18\%}{\sqrt{262}} = 0.01$ and the integrated volatility over one day is approximately 0.0001. The approximate estimate of the microstructure noise level is obtained from the intercept on the Figure 1.5: $\sqrt{\mathbb{E}[E^2]} \approx \sqrt{\exp(-9.2)/2} = 0.007$, which means the standard deviation is around 0.7% of original stock price, since $\ln(S) + \epsilon = \ln(Se^\epsilon) \approx \ln(S(1+\epsilon))$.

Remark 1.2. Resources of Microstructure Noise:
In the field of financial economics, it is commonly accepted that microstructure noise could be induced by some important sources such as:

1. Frictions inherent in the trading process: bid-ask spread, price discreteness (transaction price changes as multiples of ticks), price rounding, trades occurring on different markets or networks;

2. Informational effects: differences in trade sizes or informational content of price changes, gradual response of prices to a block trade, the strategic component of the order flow, inventory control effects;
3. Measurement or data recording errors: prices entered as zero, misplaced decimal points.

More details for microstructure noise is in a survey in Amihud et. al. (2006), a survey in O'Hara (2007), and an empirical analysis in Aït-Sahalia and Yu (2009).

1.3.1 Benchmark: Maximum Likelihood Estimator of QV

Before facing a more complex situation, it is helpful to have a discussion based on the simplest parametric case, which could be regarded as our benchmark.

The simplest case for a continuous process with observation noise is

$$X_t = \sigma W_t + E_t, \tag{1.6}$$

where $E_t \sim N(0, a^2)$, $X \perp E$, $E_t \perp E_s$, and σ is a constant. Then we have:

$$\begin{pmatrix} X_{1/n} - X_0 \\ X_{2/n} - X_{1/n} \\ \vdots \\ X_1 - X_{(n-1)/n} \end{pmatrix} \sim N\left(0, \frac{\sigma^2}{n} + \begin{bmatrix} 2a^2 & & & \\ -a^2 & 2a^2 & & \\ 0 & -a^2 & 2a^2 & \\ \vdots & \cdots & \cdots & \ddots \end{bmatrix} \right).$$

Let $\hat{\sigma}^2_{MLE}$ and $\hat{a}^2_{MLE}$ denote the MLEs based on results above. Their asymptotic properties are easily derived from classical results of the MA(1) process, when $a^2 > 0$,

$$\begin{pmatrix} n^{1/4}(\hat{\sigma}^2_{MLE} - \sigma^2) \\ n^{1/2}(\hat{a}^2_{MLE} - a^2) \end{pmatrix} \xrightarrow{D} N\left(0, \begin{bmatrix} 8a\sigma^3 & 0 \\ 0 & 2a^4 \end{bmatrix} \right).$$

Here, $\xrightarrow{D}$ means convergence in distribution.

The special case when there is no market microstructure noise results in a faster convergence rate:

$$n^{1/2}(\hat{\sigma}^2_{MLE} - \sigma^2) \xrightarrow{D} N(0, 2\sigma^4).$$

It shows that even with the simplest stochastic process and i.i.d. microstructure noise, the convergence rate of $\hat{\sigma}^2_{MLE}$ decreases from $n^{1/2}$ to $n^{1/4}$. It also gives us a benchmark that $n^{1/4}$ is the best achievable convergence rate when microstructure noise exists. These results have been discussed in Stein (1987), Jacod (2001), and Barndorff-Nielsen et. al. (2008).

1.3.2 Two Scales Realized Volatility

As we discussed above, using the highest frequency data contaminated with noise, the realized volatility becomes

$$[Y, Y]_t = \sum_{i=1}^{n=[t/\Delta]} \left(Y_{i\Delta} - Y_{(i-1)\Delta}\right)^2 = 2n\mathbb{E}(E^2) + \int_0^t \sigma_s^2 \, ds + O_p(\sqrt{n}).$$

It has a bias term (first term in above formula), which increases linearly with sample size n and overwhelms the effect of integrated volatility. Thus, the RV no longer approximates the integrated volatility as we expected.

To avoid the bias, a popular suggestion has long been known: do not compute RV at too high frequency. A sub-sampling interval from 5 mins to 30 mins has been suggested (e.g. Andersen et al. (2001), Barndorff-Nielsen and Shephard (2002)):

$$[Y, Y]_t^{(K-sparse)} = \sum_{i=1}^{[n/K]} \left(Y_{iK} - Y_{iK-K}\right)^2 = 2[n/K]\mathbb{E}(\epsilon^2) + \int_0^t \sigma_s^2 \, ds + O_p(\sqrt{n/K}). \qquad (1.7)$$

Zhang (2005) further generalizes it into an averaged version, which uses all available data and is thus more efficient:

$$\begin{aligned} [Y, Y]_t^{(K)} &= \sum_{i=K}^{n} \left(Y_i - Y_{i-K}\right)^2 \\ &= K \int_0^t \sigma_s^2 \, ds + 2(n - K + 1)\mathbb{E}(E^2) + O_p(\sqrt{n}); \qquad (1.8) \\ [Y, Y]_t^{(K-avg)} &= \frac{1}{K}[Y, Y]_t^{(K)} = \int_0^t \sigma_s^2 \, ds + 2\frac{n - K + 1}{K}\mathbb{E}(E^2) + O_p(\sqrt{n/K^2}). \end{aligned}$$

Compared with (1.5), the bias term is reduced in (1.8), but still exists. To completely remove the bias, first, we can construct an estimator of the noise variance:

$$\widehat{\mathbb{E}(E^2)} = \frac{1}{2n}[Y,Y]_t \xrightarrow{p} \mathbb{E}(E^2). \tag{1.9}$$

Then, combining (1.9) and (1.8), a straight bias-adjusted estimator is proposed:

$$\begin{aligned} TSRV(Y,K) &= [Y,Y]_t^{(K-avg)} - 2\frac{n-K+1}{K}\widehat{\mathbb{E}(\epsilon^2)} \\ &= [Y,Y]_t^{(K-avg)} - \frac{n-K+1}{nK}[Y,Y]_t. \end{aligned} \tag{1.10}$$

It has been proved that the number of sub-samples is optimally selected as $K = cn^{2/3}$, and $c = \left(\frac{T}{12\mathbb{E}[\epsilon^2]^2}\int_0^T \sigma_s^4 ds\right)^{-1/3}$, and we have the following theorem:

Theorem 1.1. *Under assumption 1.1, 1.2, 1.3, and 1.4, we have*

$$n^{1/6}\Big(TSRV(Y,K) - \int_0^T \sigma_s^2 ds\Big) \xrightarrow{L_s} \big[\frac{8}{c^2}(\mathbb{E}[\epsilon^2])^2 + c\frac{4T}{3}\int_0^T \sigma_s^4 ds\big]N(0,1). \tag{1.11}$$

Here, $\xrightarrow{L_s}$ means convergence stably in law, as defined below:

Definition 1.1. *Let Z_n denote a sequent a random variables defined on a probability space $(\Omega, \mathcal{F}, P)$ and taking the value in $(E, \mathcal{E})$: a complete separable metric space with Borel σ-algebra. Z_n is said to* ***converge stably in law*** *with limit Z, denoted as $Z_n \xrightarrow{L_s} Z$, if for every $\mathcal{F}-$measurable bounded random variable Y, and any bounded continuous function g, we have $\lim_{n\to\infty}\mathbb{E}[Yg(Z_n)] = \mathbb{E}[Yg(Z)]$.*

Remark 1.3. This definition is useful when we need to turn some infeasible estimation procedure into feasible one in practice. More details and the rationale were discussed in the Appendix A.1.

This estimator is originally developed in Zhang (2005). To the best of our knowledge, this is the first consistent estimator of QV when assuming the existence of microstructure noise and non-constancy of the volatility.

Motivated by the benefit of combining two scales, Zhang (2006) proposed an improved estimator (MSRV), which is a weighted average of $[Y,Y]_t^{(K-avg)}$ for multiple time scales. It has been proved that the MSRV has a convergence rate of $n^{-1/4}$, which is an improvement over the TSRV's rate of $n^{-1/6}$. This is also the best achievable convergence rate, as shown in section (1.3.1).

1.4 EXTENDED TSRV WITH DEPENDENT MICROSTRUCTURE NOISE

1.4.1 Dependence of Noise Structure

Until now, our discussion has been based on the i.i.d. assumption for the microstructure noise. We now turn to examining empirically if this assumption needs to be relaxed in practical applications.

To check whether the real data are consistent with this assumption, we collected the transactions and quotes data of 30 DJIA stocks from NYSE's TAQ database, over the first 10 trading days of January, 2010. To save the space, we list the information for six represents of those DJIA stocks: 3M Inc. (trading symbol: MMM), IBM (trading symbol: IBM), Johnson & Johnson (trading symbol: JNJ), J.P. Morgan & Co (trading symbol: JPM), General Electric (trading symbol: GE) and Intel (trading symbol: INTC). The reason to choose them is that their data have different level of time dependence. Other stocks have similar behaviors as one of them.

Figure 1.6 plots their prices over the first trading day. Table 1.5 reports the fundamental summary statistics on transaction data of these six stocks. We define the effective transactions as these leading to a price change. Averages are taken over the 10 trading days for each stock. Min and max are also computed over all the full ten days samples. First five orders of correlations are also included in the last five rows. It is interesting to find that the more liquid (more daily average effective transactions) of the stocks, the more likely to depart from the i.i.d. noise assumption.

In Figure 1.1, the top panel represents the autocorrelation plot of 3M and IBM. That part of plot corroborates with the i.i.d. noise structure assumption. The bottom panels show the corresponding autocorrelation plot of GE and Intel. However, it is clear that the i.i.d assumption does not fit these data well, and the autocorrelation is significant for some price process.

A simple generalization to capture the dependence structure is AR(1) or a mixed time series:

Assumption 1.5.

$$E_i = U_i + V_i, \tag{1.12}$$

where

- *U is white noise:* $U_i \perp U_j$*;*
- *V is AR(1):* $V_i = \rho V_{i-1} + \epsilon_i, \quad |\rho| < 1$

Under this assumption, we have the autocovariance:

$$Cov(\Delta_i Y, \Delta_j Y) = \begin{cases} \int_{t_{i-1}}^{t_{i1}} \sigma_s^2 dW_s + 2E[U^2] + 2(1-\rho)E[V^2], & \text{if } i = j; \\ -E[U^2] - (1-\rho)^2 E[V^2], & \text{if } |i-j| = 1; \\ -\rho^{j-i-1}(1-\rho)^2 E[V^2]. & \text{if } |i-j| > 1. \end{cases} \tag{1.13}$$

This model can easily be fitted by the method of moments. The estimates of $E[U^2]$, $E[V^2]$ and ρ for INTC are $3.3 * 10^{-8}$, $2.25 * 10^{-8}$ and -0.69. Figure 1.2 shows the sample ACF and the corresponding fitted ACF by the model above, illustrating the good fit of this simple generalization. It again confirms the necessity to consider the dependence microstructure noise, and to generalize the integrated volatility estimators. To estimate the quadratic variation in the case with significant noise dependence, we do not change the model assumption of the underlying stock price as in assumption 1.1. The assumption 1.3 of the noise structure is generalized as below:

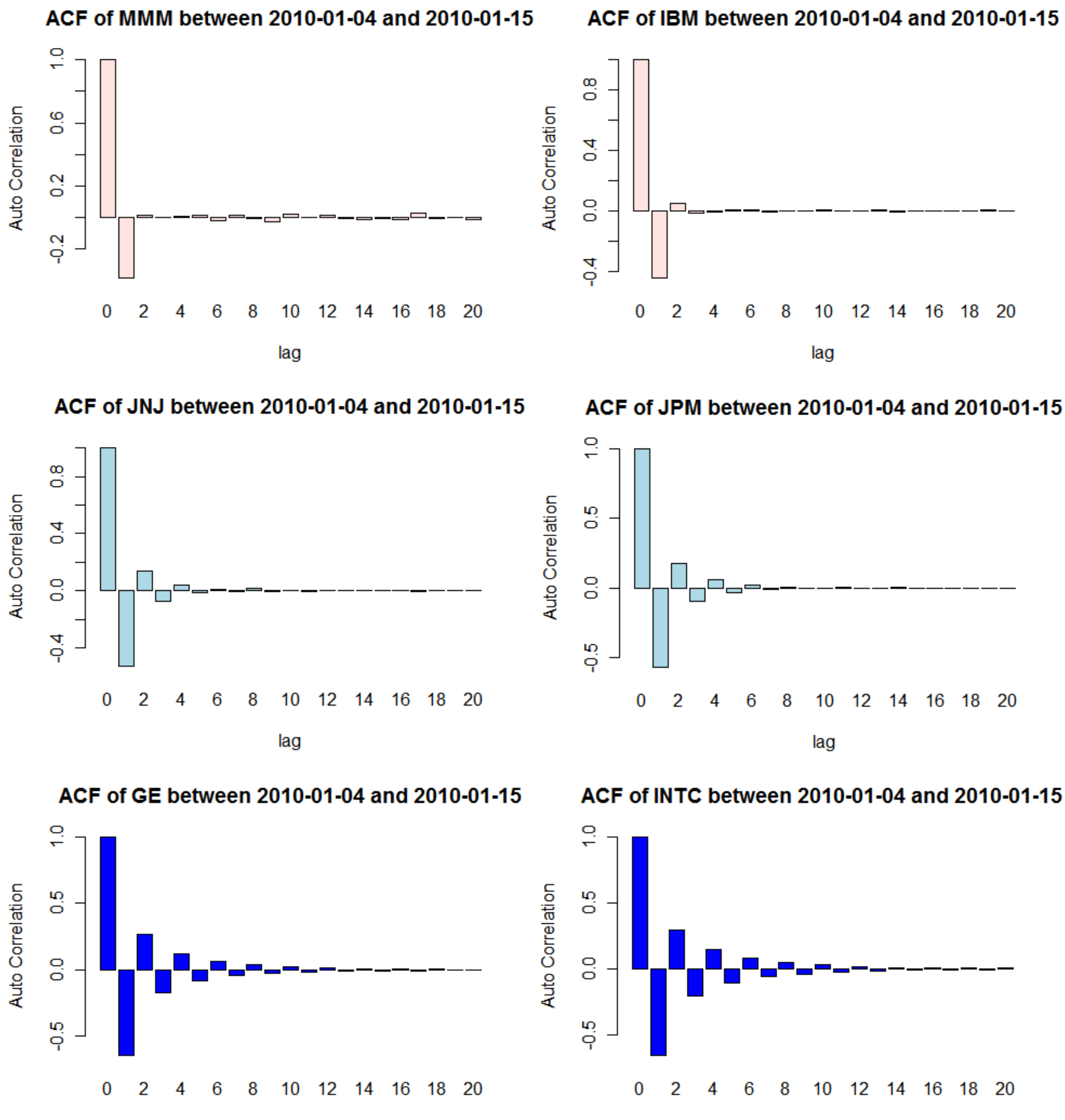

Figure 1.1: Plots of autocorrelation function of historical log price returns

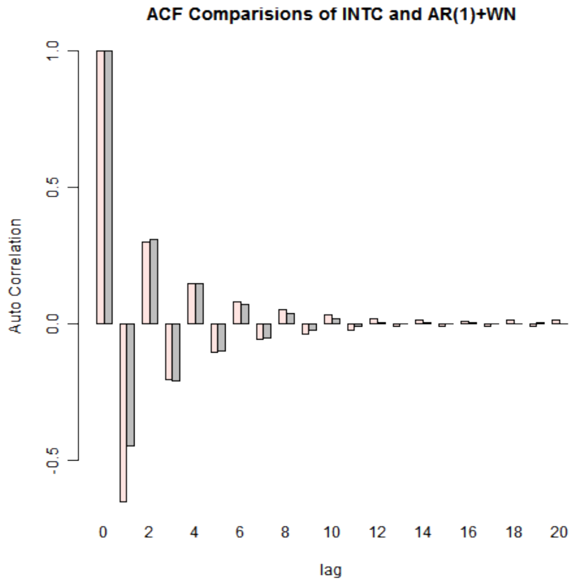

Figure 1.2: Comparison of autocorrelation function from Intel (red) and fitted value (grey)

Assumption 1.6. *The noise process E_t is a stationary process, which satisfies: $E \perp X$, and it is strong mixing with the mixing coefficients decaying exponentially (Hall and Heyde (1980)). From Theorem A.6, there exists a constant $\rho < 1$, such that for all i,*

$$|Cov(E_i, E_{i+k})| \leq \rho^k Var(E). \tag{1.14}$$

Assumption 1.7. *An alternative assumption is: the noise process E_t is a stationary process, $E \perp X$ and $|cov(E_1, E_n)| \to 0$ as $n \to \infty$. Finally, we write $V_h = Cov(E_i, E_{i+h})$.*

1.4.2 Properties of Old TSRV

To study the influence of the new noise structure to TSRV, we briefly illustrate as below:

$$
\begin{aligned}
&TSRV(Y,K)\\
&=\frac{1}{K}\Big\{[Y,Y]^{(K)}-\frac{n-K+1}{n}[Y,Y]\Big\}\\
&=\frac{1}{K}\Big\{\Big([X,X]^{(K)}+2[X,E]^{(K)}+[E,E]^{(K)}\Big)-\frac{n-K+1}{n}\Big([X,X]+2[X,E]+[E,E]\Big)\Big\}\\
&=\frac{1}{K}\Big\{\underbrace{\Big([X,X]^{(K)}-\frac{n-K+1}{n}[X,X]\Big)}_{D_{discrete}}+\underbrace{\Big([E,E]^{(K)}-\frac{n-K+1}{n}[E,E]\Big)}_{D_{noise}}\\
&\qquad+2\underbrace{\Big([X,E]^{(K)}-\frac{n-K+1}{n}[X,E]\Big)}_{D_{mix}}\Big\}.
\end{aligned}
\tag{1.15}
$$

as $n\to\infty,\ K\to\infty,$ and $K/n\to 0$,

$$
\begin{aligned}
&\mathbb{E}(\frac{1}{K}D_{discrete})\to\int_0^T\sigma_s^2ds;\\
&\mathbb{E}(\frac{1}{K}D_{mix})=0;\\
&\mathbb{E}(\frac{1}{K}D_{noise})=(n-K+1)\mathbb{E}(E_K-E_0)^2-\frac{n-K+1}{n}n\mathbb{E}(E_1-E_0)^2\\
&\qquad\qquad\quad=(n-K+1)(V_K-V_1).
\end{aligned}
\tag{1.16}
$$

Therefore,

$$\mathbb{E}\Big(TSRV(Y,K)\Big)=\int_0^T\sigma_s^2ds+(V_K-V_1)O(n). \tag{1.17}$$

Through this result, the dependence in microstructure noise introduces a bias term additionally to the integrated volatility. And this bias is linearly increasing with the sample size n. In the previous i.i.d. assumption of noise structure, both V_K and V_1 are zero, and thus the bias disappears.

1.4.3 Extended TSRV

From previous experience, sub-sampling is a common method to reduce the bias from the noise. In addition, from the assumption 1.6, the time dependence decreases exponentially. This motivates an extension of the TSRV to construct a new estimator based on sub-sampling:

$$STSRV(Y,J,K) = \frac{1}{K}\Big\{[Y,Y]^{(K)} - \frac{n-K+1}{n-J+1}[Y,Y]^{(J)}\Big\}.$$

Lemma 1.1. *Under assumptions 1.1, 1.2, 1.3 and 1.5, as $n \to 0$ and $K \to 0$, we have:*

$$[X,E]^{(K)} = \sum_{i=K}^{n}(X_i - X_{i-K})(E_i - E_{i-K}) = O_p(\sqrt{K}). \tag{1.18}$$

Proof. This is the same as lemma 1 in Aït-Sahalia et. al.(2011). □

From **Lemma** 1.1, it is easy to see that

$$[Y,Y]^{(K)} = [X,X]^{(K)} + [E,E]^{(K)} + O_p(\sqrt{K}).$$

- **Signal-Noise Decomposition:**

$$\begin{aligned}
&STSRV(Y,J,K)\\
&= \frac{1}{K}\Big\{[Y,Y]^{(K)} - \frac{n-K+1}{n-J+1}[Y,Y]^{J}\Big\}\\
&= \underbrace{\frac{1}{K}\Big([X,X]^{(K)} - \frac{n-K+1}{n-J-1}[X,X]^{J}\Big)}_{SignalTerm} + \underbrace{\frac{1}{K}\Big([E,E]^{(K)} - \frac{n-K+1}{n-J-1}[E,E]^{J}\Big)}_{NoiseTerm}\\
&\quad + O_p\Big(\sqrt{\frac{1}{J}}\Big).
\end{aligned} \tag{1.19}$$

- **Noise Term:**

$$\begin{aligned}
&\mathbb{E}[\text{NoiseTerm}]\\
&= \frac{1}{K}\mathbb{E}\Big[[E,E]^{K} - \frac{n-K+1}{n-J-1}[E,E]^{J}\Big]\\
&= \frac{n-K+1}{K}\Big(\mathbb{E}(E_i - E_{i-K})^2 - \mathbb{E}(E_i - E_{i-J})^2\Big)\\
&= 2\frac{n-K+1}{K}(V_J - V_K).
\end{aligned} \tag{1.20}$$

Lemma 1.2. *If* $\limsup \frac{J}{K} < 1$, *then as* $J, K \rightarrow \infty$

$$\frac{K}{\sqrt{n}}\Big(NoiseTerm - \mathbb{E}[NoiseTerm]\Big) \xrightarrow{D} \epsilon_{noise} Z_{noise}, \tag{1.21}$$

where $\epsilon_{noise}^2 = 8V_0^2 + 16\sum_{i=1}^{\infty} V_i^2$.

Proof. This is the same as Proposition 1 in Aït-Sahalia et. al.(2011). □

- **Signal Term:**

Lemma 1.3. *For* $1 \leq J \leq K$ *and* $\frac{K}{n} \rightarrow 0$, *as* $J, K, n \rightarrow \infty$

$$\frac{1}{\sqrt{\frac{K}{n}(1 + 2\frac{J^3}{K^3})}}\Big(SignalTerm - \int_0^T \sigma_s^2 ds\Big) \xrightarrow{D} \epsilon_{signal} Z_{signal}, \tag{1.22}$$

where $\epsilon_{signal}^2 = \frac{4}{3}T\int_0^T \sigma_s^4 ds$.

Theorem 1.2. *As* $1 \leq J \leq K$ *and* $\frac{K}{n} \rightarrow 0$,

$$\begin{aligned} STSRV(Y, J, K) = \int_0^T \sigma_s^2 ds + n^{-1/6}\Big\{&2\frac{n-K+1}{K}(V_J - V_K) + \frac{\sqrt{n}}{K}\epsilon_{noise} Z_{noise} \\ &+ \sqrt{\frac{K}{n}(1 + 2\frac{J^3}{K^3})}\epsilon_{signal} Z_{signal}\Big\}. \end{aligned} \tag{1.23}$$

Proof. This is easily proved following lemma 1.2 and lemma 1.3. □

1.5 A NEW METHOD: SUB-SAMPLING REALIZED COVARIANCE ESTIMATOR WITH DEPENDENT NOISE

1.5.1 Construction of Sub-sampling Realized Covariance

In the situation of i.i.d. microstructure noise, the number of noise terms involved in the estimator is determined by the sub-samples. The contribution of $E_{i+K} - E_i$ is similar to $E_{i+1} - E_i$, and $E_i - E_j$ is uncorrelated with $E_k - E_l$ as soon as $i, j < k, l$. However, under the time dependence noise structure, $E_i - E_j$ and $E_k - E_l$ are always correlated. The only fact we know is that their correlation decreases exponentially with the distance between them. Therefore, to reduce the correlation of noise term, we can either increase the interval size (sub-sampling), or increase the distance between these two terms. Following this logic, we define a family of estimators based on realized covariances as below:

$$\begin{aligned}
\gamma_0^{(K)}(Y,Y) &= \sum_{i=K}^{n} (Y_i - Y_{i-K})^2, \\
\gamma_1^{(K)}(Y,Y) &= \sum_{i=2K}^{n} (Y_i - Y_{i-K})(Y_{i-K} - Y_{i-2K}), \\
&\vdots \\
\gamma_h^{(K)}(Y,Y) &= \sum_{i=(h+1)K}^{n} (Y_i - Y_{i-K})(Y_{i-hK} - Y_{i-(h+1)K}), \\
&\vdots
\end{aligned} \tag{1.24}$$

The realized covariance estimators have the following asymptotic properties:

Lemma 1.4. *Under assumptions 1.1, 1.2, 1.3 and 1.5,*

$$\begin{aligned}
\mathbb{E}[\gamma_0^{(K)}(Y,Y)] &= K\int_0^T \sigma_s^2 ds + (n-K+1)(2V_0 - 2V_K) + O_p\left(\sqrt{\frac{1}{n}}\right), \\
\mathbb{E}[\gamma_1^{(K)}(Y,Y)] &= (n-2K+1)(-V_0 + 2V_K - V_{2K}), \\
&\vdots \\
\mathbb{E}[\gamma_h^{(K)}(Y,Y)] &= (n-(h+1)K+1)(-V_{(h-1)K} + 2V_{hK} - V_{(h+1)K}), \\
&\vdots
\end{aligned} \tag{1.25}$$

From these results, only the realized variance γ_0 includes the part of integrated volatility, and other realized covariances are different measurements of the covariance of the noise.

1.5.2 Sub-sampling Realized Covariance Estimator - SRC(Y,K)

We construct a family of estimators from a weighted combination of the realized covariances:

$$SRC(Y,K) = \frac{1}{K}\Big\{\gamma_0^{(K)}(Y,Y) + 2(n-K+1)\sum_{h=1}^{H}\frac{1}{n-(h+1)K+1}\gamma_h^{(K)}(Y,Y)\Big\}. \quad (1.26)$$

If we denote the vector of realized covariances as

$$\Gamma^{(K)}(X,Y) = \big(\gamma_0^{(K)}(X,Y), \gamma_1^{(K)}(X,Y), \cdots, \gamma_H^{(K)}(X,Y)\big)^T,$$

then we can rewrite the SRC(Y,K) in a matrix formula:

$$SRC(Y,K) = \frac{1}{K}W^T\Gamma^{(K)}(Y,Y),$$

where

$$W = \Big[1, K(\frac{0}{H}), \cdots, K(\frac{h-1}{H}), \cdots, K(\frac{H-1}{H})\Big].$$

There are several choices for the kernel function $K(x)$:

- **Truncated Kernel:** $W(x) = I\{x=0\}$;
- **Infinite-lag Kernel:** Bartlett, $W(x) = 1-x$; Epanechnikov, $W(x) = 1-x^2$;
- **Smooth Kernel:** Cubic, $W(x) = 1-3x^2+2x^3$, Tukey-Hanning$_n$, $W(x) = \sin^2[\pi/2(1-x)^n]$.

In this paper, we will focus on the kernel as in (1.26):

$$W = \Big(1, 2\frac{n-K+1}{n-2K+1}, \cdots, 2\frac{n-K+1}{n-(H+1)K+1}\Big) \; or \; W = \Big(1, 2, \cdots, 2\Big) + O(\frac{1}{n}).$$

Remark 1.4. This type of estimator is related to the Heteroskedastic Autocorrelation (HAC) estimators discussed by, for example, Gallant (1987), Newey and West (1987), and Andrews (1991). Its application in econometrics was first proposed in Zhou (1996), who used the first order covariance to reduce the bias from noise. Hansen and Lunde (2006) used this type of estimators with $K(x) = 1$ for general H to characterize the second-order properties of market microstructure noise. However, both of these estimators are inconsistent. The more general and consistent estimators was recently studied in Barndorff-Nielsen et. al. (2008).

Theorem 1.3. ***Asymptotic Properties of*** $\Gamma^{(K)}(Y)$***:***

Under assumptions 1.1, 1.2, 1.3 and 1.5, as $n \to \infty$*,*

(1) ***Signal Term:***

$$\sqrt{\frac{n}{T}} \left\{ \Gamma^{(K)}(X,X) - \begin{bmatrix} \int_0^T \sigma_s^2 ds \\ 0 \\ \cdots \\ 0 \end{bmatrix} \right\} \xrightarrow{L_s} N\left(0, \frac{1}{6} K(\int_0^T \sigma_s^4 ds)\Omega_X \right),$$

where

$$\Omega_X = \begin{pmatrix} 8 & & & & \\ 2 & 4 & & & \\ 0 & 1 & 4 & & \\ \vdots & 0 & 1 & \ddots & \\ 0 & \cdots & \cdots & 1 & 4 \end{pmatrix}.$$

(2) ***Mixed Term:***

as $K \to \infty$*,*

$$\Gamma^{(K)}(X,U) + \Gamma^{(K)}(U,X) = O_p(\sqrt{K}).$$

If $K=1$, we have the special case:

$$\Gamma(X,U)+\Gamma(U,X) \xrightarrow{D} N\Big(0,(\int_0^T \sigma_s^2 ds)\Omega_{XU}\Big),$$

where

$$(\Omega_{XU})_{ij} = Cov(E_i - E_{i-1}, E_j - E_{j-1}) = -V_{|i-j-1|} + 2V_{|i-j|} - V_{|i-j+1|}$$

(3) ***Noise Term:***

$$\begin{aligned}
\mathbb{E}[\Gamma(U,U)] =& n\big(2V_0 - 2V_K, -V_0 + 2V_K - V_{2K}, \cdots, -V_{(H-1)K} + 2V_{HK} - V_{(H+1)K}\big)^T \\
&+ O(1); \\
Var[\Gamma(U,U)] =& nVar(E^2)\Omega_U + O(K),
\end{aligned} \tag{1.27}$$

where

$$\Omega_U = \begin{pmatrix} 2 & & & \\ 0 & 1 & & \\ 0 & 0 & \ddots & \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

***Proof*.** See Appendix A.2. □

Based on Theorem 1.3, we can derive the large n and large K asymptotic variance of $SRC(Y,K) - \int_0^T \sigma_s^2 ds - 2\frac{n-K+1}{K}(V_{HK} - V_{(H+1)K})$ as:

$$\frac{1}{6}\frac{K}{n}T(\int_0^T \sigma_s^4 ds)W^T\Omega_X W + O(\frac{1}{K}) + \frac{n}{K^2}Var(E^2)W^T\Omega_U W.$$

To minimize the asymptotic variance above, we can select $K = cn^{2/3}$, in which case we have the following theorem:

Theorem 1.4. ***Central Limit Theorem for*** $SRC^{(K)}(Y)$***:***
Under assumptions 1.1, 1.2, 1.3 and 1.5, as $n \to 0$, $K = cn^{2/3}$, we have

$$n^{1/6}\Big\{SRC(Y,K) - \int_0^T \sigma_s^2 ds\Big\} \xrightarrow{D} N\Big(0, \frac{1}{6}T(\int_0^T \sigma_s^4 ds)W^T\Omega_X W + Var(E^2)W^T\Omega_U W\Big). \tag{1.28}$$

Proof. This follows easily from Theorem 2 and lemma .3. □

Therefore, the $SRC(Y, K)$ is a consistent estimator in the case where the noise has a time dependence structure. To compare our new estimator with other existing methods, we will present the simulation results in the next section.

1.6 SIMULATIONS AND COMPARISONS

For practical applications, it is important to consider these estimators' finite sample performance. It is also useful to check their sensitivity to different noise levels and different dependence levels. Therefore, in this section, we conduct an extensive Monte Carlo study to examine the performance of our new estimator $SRC(Y, K)$, and compare it with other estimators: RV, sparse realized volatility (SRV), TSRV, and RTSRV.

1.6.1 Monte Carlo Setup

To generate the simulated data, we use the stochastic volatility model of Heston:

$$\begin{aligned} dX_t &= (\mu - v_t/2)dt + \sigma_t dW_t \Leftrightarrow dS_t = \mu S_t d_t + \sigma_t S_t dW_t, \\ dv_t &= a(\bar{v} - v_t)dt + r\sqrt{v_t}dW_t. \end{aligned} \tag{1.29}$$

We used the following parameters: $a = 5$, $\bar{v} = .05/262$, $r = 0.5$, $\rho = -0.5$ as in Zhang (2011). For each experiment, 5000 sample paths are generated using the Euler scheme with time interval Δ =1 second. Figure 1.7 is an example of a simulated path over one day without observation noise, along with the underlying, but unobservable volatility process. The plot is created as 3-mins OHLC (Open/High/Low/Close) candlestick charts. It is easy to see the mean-reversing of the volatility process, and the negative correlation between the log-price process and volatility process.

For the case of i.i.d. microstructure noise, we generate the observation process by the underlying process X_t plus a white noise process: $Y_i = X_i + E_i$, where $E_i \perp X_i$, and E_i are $i.i.d.$ $\sim$ $N(0, \sigma_E^2)$.

For the case of dependent microstructure noise, we generate the noise process following an AR(1) setup:

$$E_i = U_i + V_i, \tag{1.30}$$

where

- U is white noise: $U_i \perp U_j$;
- V is AR(1): $V_i = \rho V_{i-1} + \epsilon_i, \quad |\rho| < 1$.

1.6.2 Results: No Noise

Figure 1.8 compares the different estimators, using the simulated data without observation noise. In this ideal situation, all these estimators are converging to the real integrated volatility, as the sub-sampling interval decreases. It is obvious that TSRV and SRC do not improve the estimation of RV, and the RV is the best choice here.

1.6.3 Results: i.i.d. Noise

Figure 1.9 compares the different estimators, using the simulated data with i.i.d noise. To consider different situations, we compare the results under different noise levels: 0.0001, 0.0005 and 0.001, which is around 1, 2 and 3 multiples (noise-signal-ratio) of the volatility in each sub-sampling.

From the left panel, RV diverges as the sub-sampling interval decreases from 500 seconds to 1 seconds. The right panel shows the comparisons of TSRV and SRC. The adjTSRV is an adjusted version of TSRV with the same asymptotic properties (Zhang (2005)). They all converge to the true QV as expected, when the sub-sampling decreases. It is interesting that although our new estimator SRC is designed for the dependent noise, SRC works better with smaller finite-sample bias in the case with high noise level.

1.6.4 Results: Dependent Noise

To evaluate the performance of these estimators, we compare their relative bias and relative MSE separately for each stock, with different sample sizes. The relative bias is calculated as an approximation of $\mathbb{E}\Big(\frac{\text{estimator}-\int_0^T \sigma_s^2 ds}{\int_0^T \sigma_s^2 ds}\Big)$ over these 5000 sample paths; the relative MSE is calculated as an approximation of $Var\Big(\frac{\text{estimator}-\int_0^T \sigma_s^2 ds}{\int_0^T \sigma_s^2 ds}\Big)$.

To compare their small sample properties, we did experiments with with different sample sizes (1 day with $n = 23,400$, 4 hours with $n = 14,400$, 2 hours with $n = 7,200$, 1 hour with $n = 3,600$, 30 mins with $n = 1,800$, 15 mins with $n = 900$, and 10 mins with $n = 600$). We use three levels of microstructure noise: low ($\mathbb{E}(E^2) = 0.00005$), medium ($\mathbb{E}(E^2) = 0.0005$), and high ($\mathbb{E}(E^2) = 0.002$) to evaluate their sensitivity to noise level.

Figure 1.10 shows how the relative MSE changes with different sub-sampling choice, using the 1 day simulated data with medium level noise. The optimal choice of sub-sampling size could be theoretically derived, but it is not the focus of this chapter. From the figure we can see that the new estimator SRC has smaller relative MSE compared with the revised TSRV, and that it favors more frequent sub-sampling.

Table 1.1 shows the Monte Carlo results in the case of medium level noise. The volatility used in the Stochastic Volatility Model is on average 0.05 annually, which is $\sqrt{\frac{0.05}{262*23400}} \approx 0.0001$ for every second. The autocorrelation of the noise dependence is assumed as -0.6, which is similar to the one from our empirical estimation.

It is obvious that the new estimator SRC(K,1) has smaller relative MSE compared with other estimators. And we observe that the relative bias of the new estimator is much smaller. Actually, this observation is consistent with the our logic for constructing this new estimator: reduce the bias of noise by combining different realized covariances, while the revised TSRV mostly relies on the sub-samplings.

Table 1.2 shows their performance with ultra high frequency data. $\Delta = 5$ *secs* means, on average, we can observe 1 data point per 5 seconds.

Table 1.3 and 1.4 show the results separately for the low noise level ($\sqrt{\mathbb{E}[E^2]} = 0.00005$) and high level of noise ($\sqrt{\mathbb{E}[E^2]} = 0.002$). The new method SRC consistently has the smallest relative bias and relative MSE. Additionally we observe that when the noise level is very low, the simple sub-sampling RV (Sparse RV) is comparable with TSRV and SRC, especially when the sample size is not large.

1.7 EMPIRICAL ANALYSIS

Based on the theoretical studies in this and the previous chapters, we now turn to the comparisons of the empirical performance of the RV, TSRV and our new SRC estimators. We collect the transaction data of SPY from the first eight trading days in 2001 from NYSE's TAQ database. The reason that we analyze this data is that SPY is an actively traded exchange-traded fund (ETF), and it represents an ownership in a portfolio of the equity securities that comprise the Standard & Poor's 500 Index, which usually be regarded as the overall market benchmark. We also collect the transaction data of 30 DJIA stocks from NYSE's TAQ database, over the first ten trading days of January, 2010.

- **Marketwise: SPY**

  Figure 1.11 and 1.12 are results of different estimators based on SPY data on the first eight days in 2001, which represents the marketwise averaged noise and dependence level. We can see the divergence of RV with the decrease of sample interval. Also from Figure 12, TSRV and SRC are stable with respect to the sub-sampling choices, while the RV is quite jagged.

- **High Noise Dependence: INTC**

  Figure 1.13 and 1.14 show results of different estimators based on Intel, over the first eight days in 2010. As discussed before and shown in Figure 1.1, the autocorrelation is very strong among the log-return price of Intel. In this situation, the TSRV becomes worse, and its bias increases with the sample sizes, but is smaller than the RV estimator. Our new method SRC estimator is robust for this case with high noise dependence.

- **Low Noise Dependence: MMM**

  Figure 1.15 is for the MMM's stock. We already discussed and showed in Figure 1.1 that MMM does not have significant time dependence structure. In this case, the figure shows that the TSRV and SRC estimators are very close.

1.8 CONCLUSION AND FUTURE WORK

In this chapter, we have reviewed different approaches to estimate the quadratic variation using high frequency data. The presence and significant influence of the microstructure noise has also been empirically studied.

To reduce the bias introduced by the noise in the estimator of QV, Zhang (2005) proposed the fist consistent estimator TSRV based on high frequency data with the assumption of i.i.d. noise. TSRV has been generalized to a sparse version in Zhang (2011) to make it is still consistent in the case with dependence noise structure.

We propose a new estimator SRC, which is constructed by a weighted combination of sub-sampling realized covariances. The advantage of bringing in the covariance is that the realized covariances introduce more information of high order noise dependence, which is significantly nonzero for some stocks like INTC.

Here, we only focus on a special case of the new Sub-sampling Realized Covariance Estimator, which uses the truncated kernel. Similar to the discussion in Barndorff-Nielsen et al. (2008), different kernel functions will give different results, and some might increase the convergence rate from $n^{1/6}$ to $n^{1/4}$. Further discussion in this direction will be a part of our future work.

1.9 TABLES AND FIGURES

Table 1.1: Performance with different T: Medium Noise Level = 0.0005, $\rho_{AR} = -0.6$

T		RV	SRV	TSRV(Y,K)	STSRV(Y,J,5J)	SRC(Y,K)
15 mins	Relative Bias	216	0.0209	0.2111	0.2609	0.0370
	Relative MSE	223	0.5214	0.3887	0.4231	0.5366
1 hours	Relative Bias	240	0.1251	0.2858	0.1423	0.0164
	Relative MSE	259	0.4544	0.4448	0.2654	0.2927
2 hour	Relative Bias	247	0.0764	0.2128	0.1340	0.0035
	Relative MSE	274	0.3731	0.3648	0.2127	0.2064
4 hours	Relative Bias	233	0.4609	0.1647	0.0887	0.0018
	Relative MSE	249	0.7808	0.2876	0.1720	0.1459
1 day	Relative Bias	23.39	0.1252	0.0863	0.0640	0.0015
	Relative MSE	249.82	0.2889	0.1832	0.1400	0.1186

Table 1.2: Performance with different Δ_n: Medium Noise Level = 0.0005, $\rho_{AR} = -0.8$

Δ_n		RV	SRV	TSRV(Y,K)	STSRV(Y,J,5J)	SRC(Y,K)
30 secs	Relative Bias	8	0.0418	0.1139	0.2146	0.0095
	Relative MSE	9	0.3367	0.1961	0.3207	0.2291
5 secs	Relative Bias	63	0.1455	0.1411	0.1014	0.0111
	Relative MSE	67	0.3187	0.2553	0.1808	0.1750
1 sec	Relative Bias	23.39	0.1252	0.0863	0.0640	0.0015
	Relative MSE	249.82	0.2889	0.1832	0.1400	0.1186

Table 1.3: Performance with different T: Heston Model, Low Noise Level = 0.00005

T		RV	SRV	TSRV(Y,K)	STSRV(Y,J,5J)	SRC(Y,K)
15 mins	Relative Bias	2.17	0.0236	0.0883	0.1834	0.0826
	Relative MSE	2.26	0.2950	0.1919	0.2745	0.1690
1 hours	Relative Bias	2.40	0.0158	0.0835	0.0964	0.0152
	Relative MSE	2.62	0.1591	0.1250	0.1832	0.0940
2 hour	Relative Bias	2.49	0.0209	0.0591	0.0932	0.0179
	Relative MSE	2.75	0.1158	0.1017	0.1453	0.0722
4 hours	Relative Bias	2.38	0.0195	0.0623	0.0937	0.0212
	Relative MSE	2.60	0.0862	0.0861	0.1250	0.0538
1 day	Relative Bias	2.32	0.0721	0.0529	0.0462	0.0144
	Relative MSE	2.46	0.0277	0.0711	0.1049	0.0431

Table 1.4: Performance with different T: Heston Model, High Noise Level = 0.002

T		RV	SRV	TSRV(Y,K)	STSRV(Y,J,5J)	SRC(Y,K)
2 hour	Relative Bias	247	0.0764	0.3917	0.1928	0.0200
	Relative MSE	274	0.3731	0.5612	0.3117	0.3400
4 hours	Relative Bias	212	0.0726	0.4643	0.1747	0.0089
	Relative MSE	245	0.4342	0.5660	0.2594	0.2213
1 day	Relative Bias	928	0.2771	0.2433	0.1093	0.0048
	Relative MSE	993	0.4545	0.3891	0.2070	0.1951

Table 1.5: Descriptive Statistics for DJIA stocks in first 10 days of 2010

Descriptive Statistics	MMM	IBM	JNJ	JPM	GE	INTC
Avg. Effective Transaction	19026	33043	44671	143447	134031	154835
Avg. time between Transaction	1.30	0.73	0.55	0.17	0.19	0.17
Min log-return	-0.01	-0.01	-0.01	-0.02	-0.02	-0.02
Max log-return	0.01	0.01	0.01	0.02	0.02	0.02
Avg. daily 1st order Corr.	-0.38	-0.44	-0.53	-0.57	-0.64	-0.65
Avg. daily 2nd order Corr.	0.01	0.05	0.13	0.18	0.27	0.30
Avg. daily 3nd order Corr.	-0.00	-0.01	-0.07	-0.10	-0.18	-0.20
Avg. daily 4nd order Corr.	0.00	-0.00	0.04	0.06	0.12	0.14
Avg. daily 2nd order Corr.	0.01	0.01	-0.02	-0.03	-0.09	-0.11

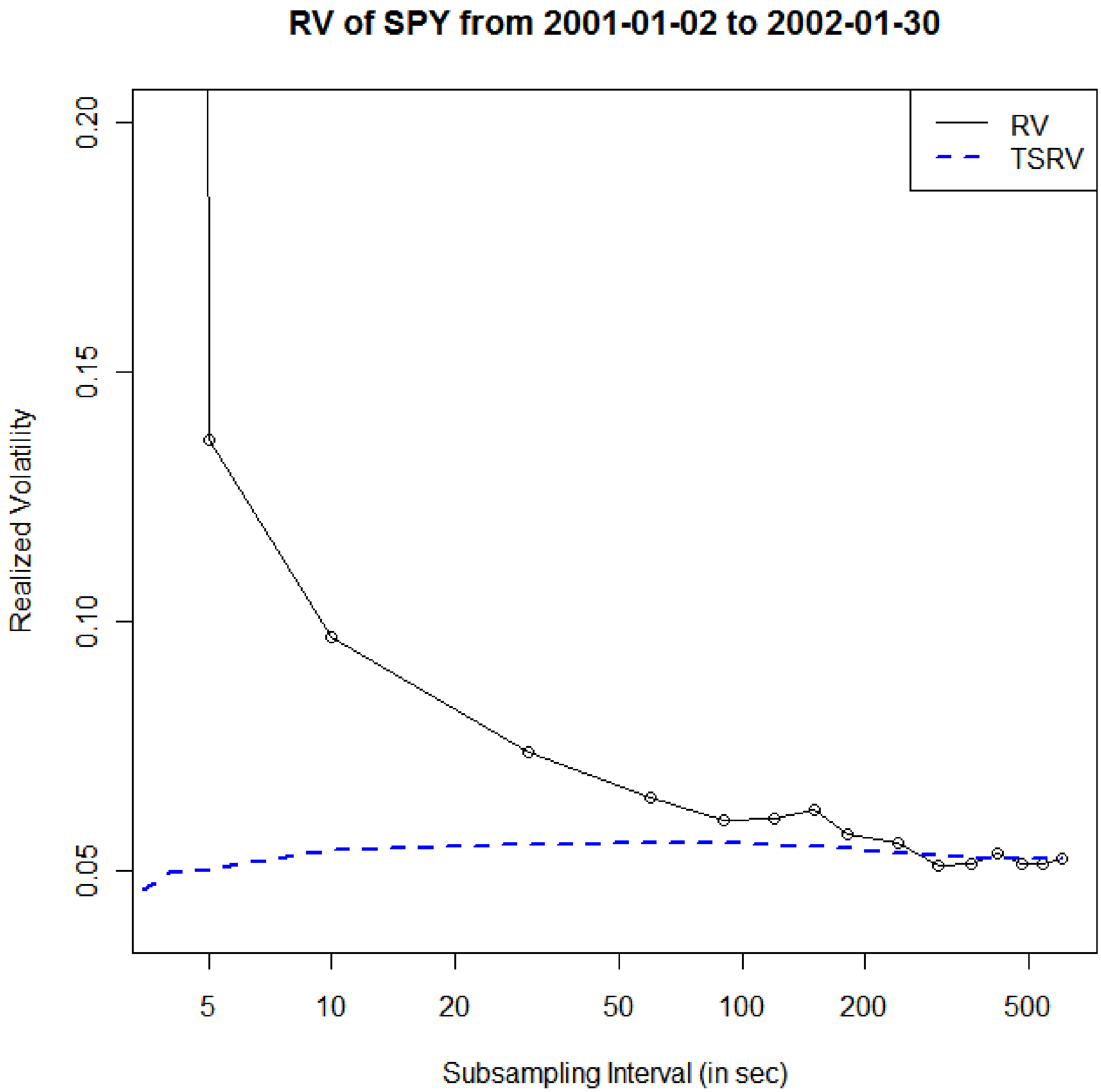

Figure 1.3: Volatility Signature Plot: RV vs. Sub-sampling Δ

This plot shows the RV estimator $[Y, Y]_t^n$ *plotted against the sub-sampling interval* Δ*. The RV estimator is computed based on SPY transaction price from Jan 2001 to Jan 2002. The plot illustrates the divergence of RV as* $\Delta \to 0$*, which is also very common for many other financial data.*

Figure 1.4: Historical data of VIX from the year of 2004 to 2009

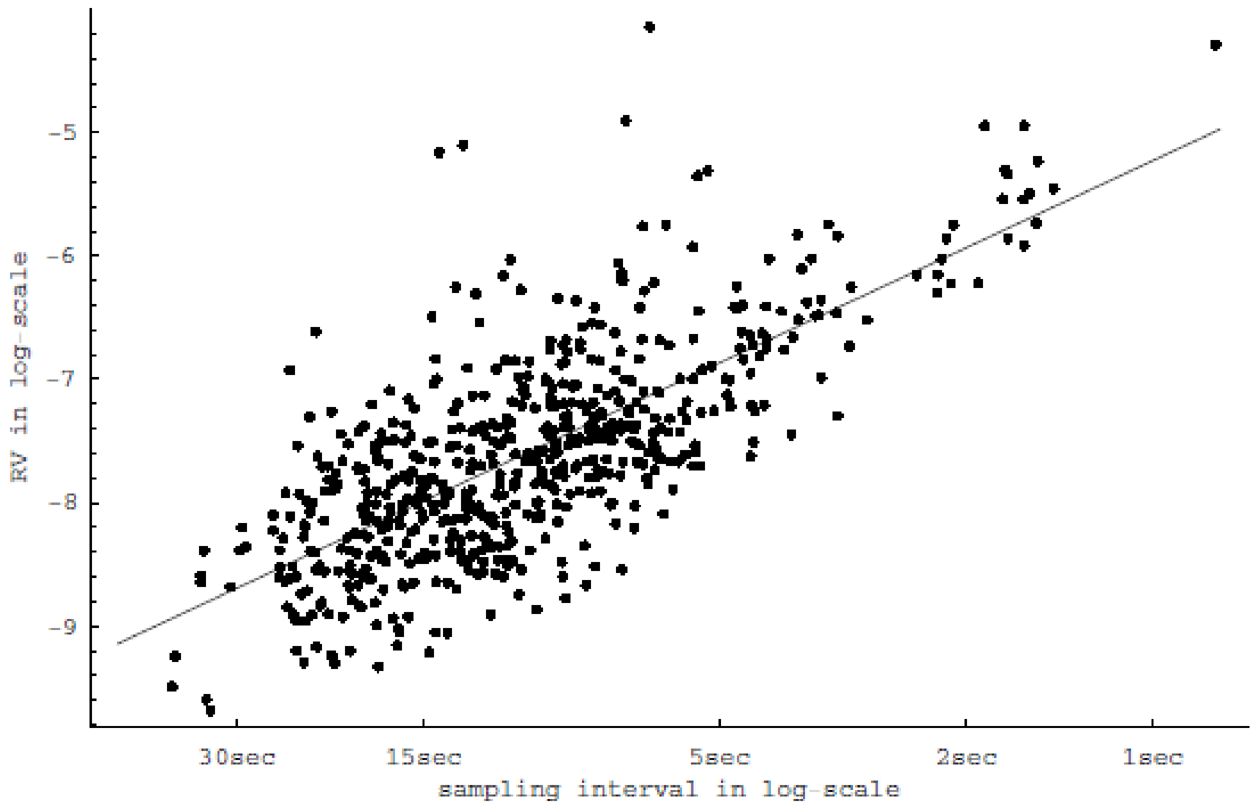

Figure 1.5: Plot of ln(RV) vs. ln(sample size)

This plot is from Aït-Sahalia, Mykland, and Zhang (2011). It shows a regression of $\ln([Y,Y]_t^n)$ *against* $\ln\ n$.

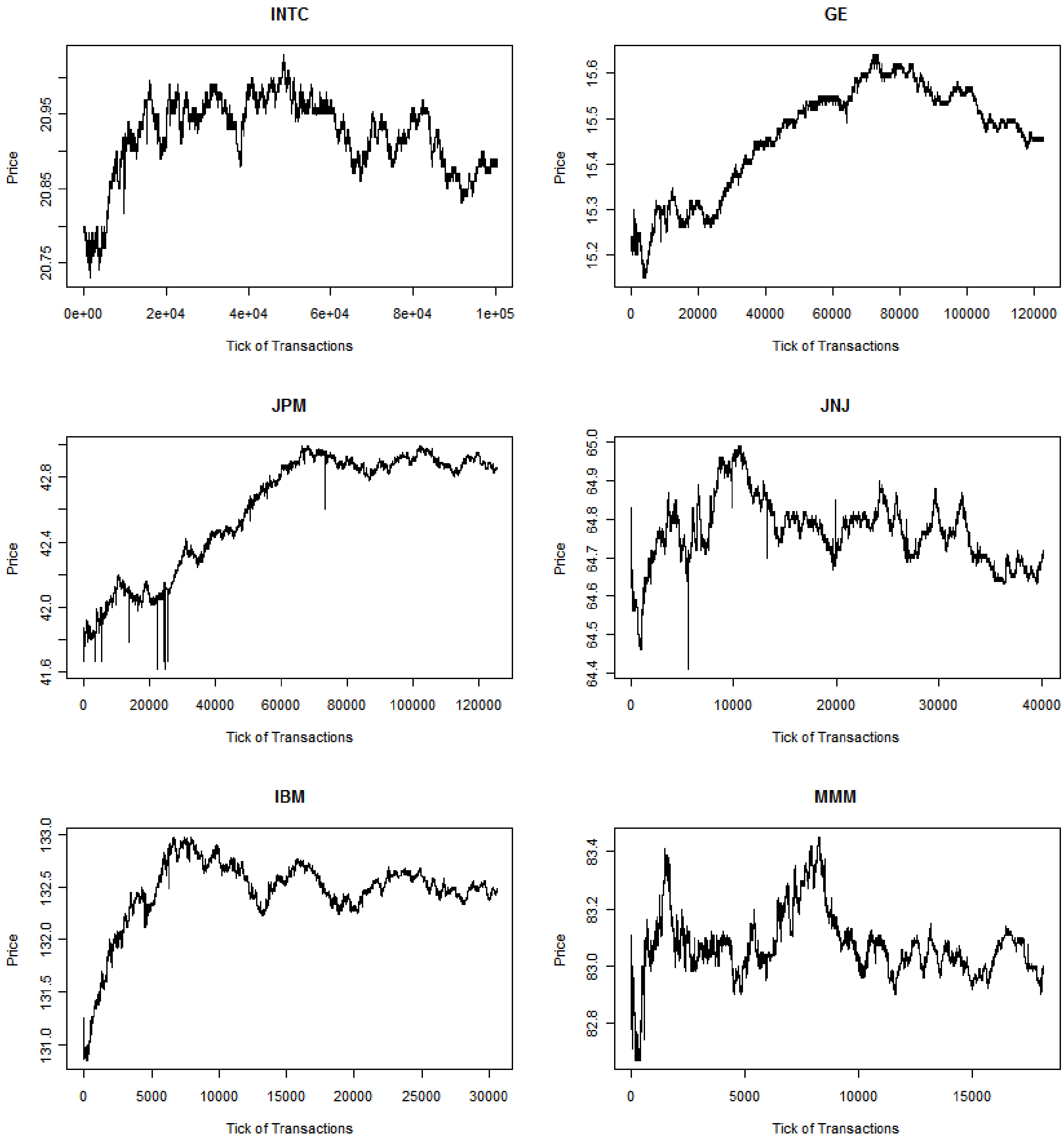

Figure 1.6: Plots of Six DJIA Stock Prices on the first trading day in 2010

Figure 1.7: One path example of Stochastic Volatility Models

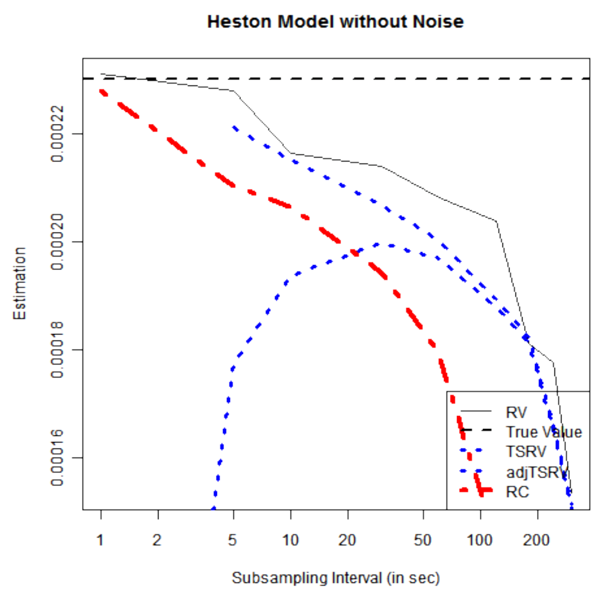

Figure 1.8: Comparisons of RV, TSRV, adjTSRV, and SRC

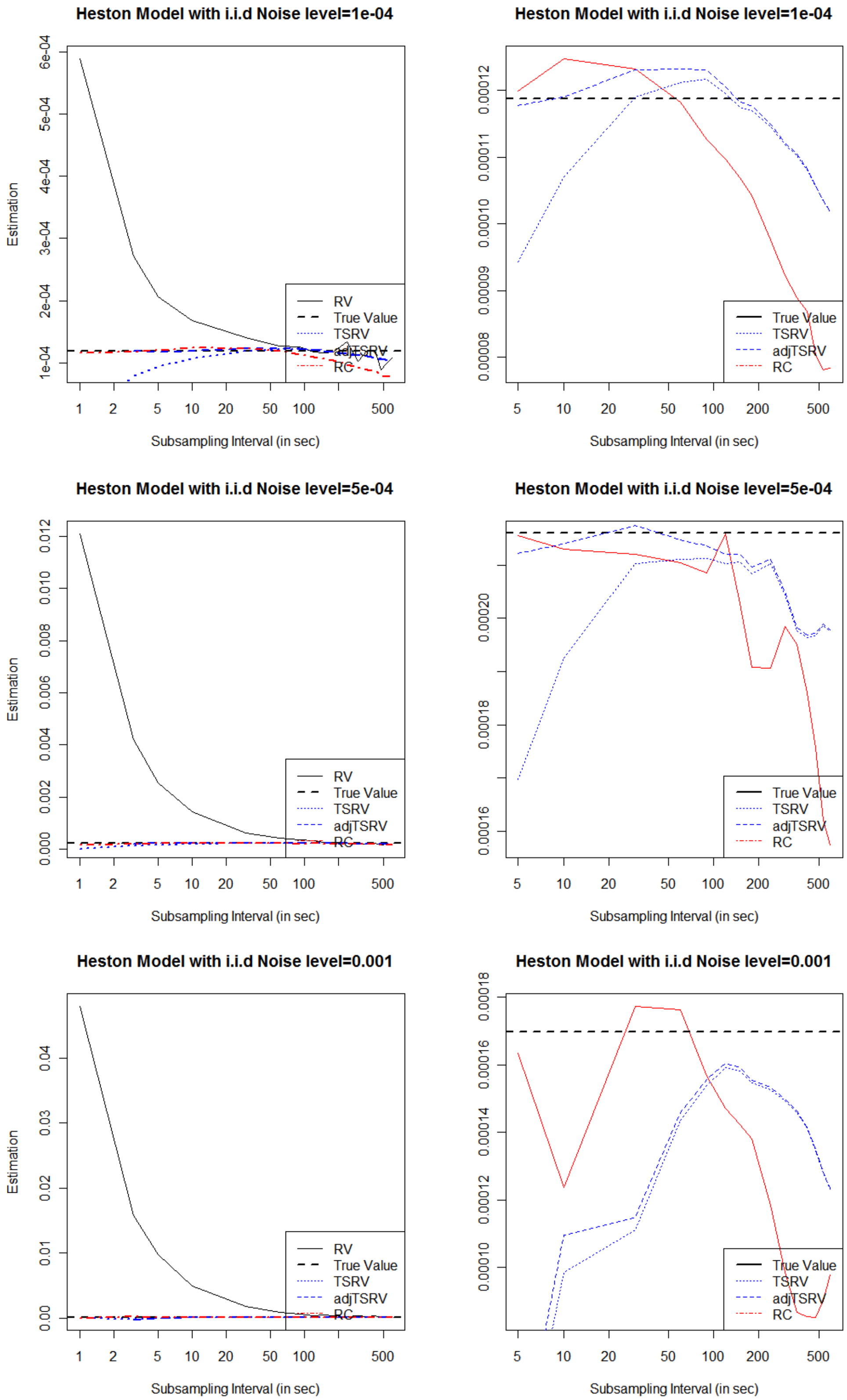

Figure 1.9: Comparisons of RV, TSRV, adjTSRV, and SRC under i.i.d noise

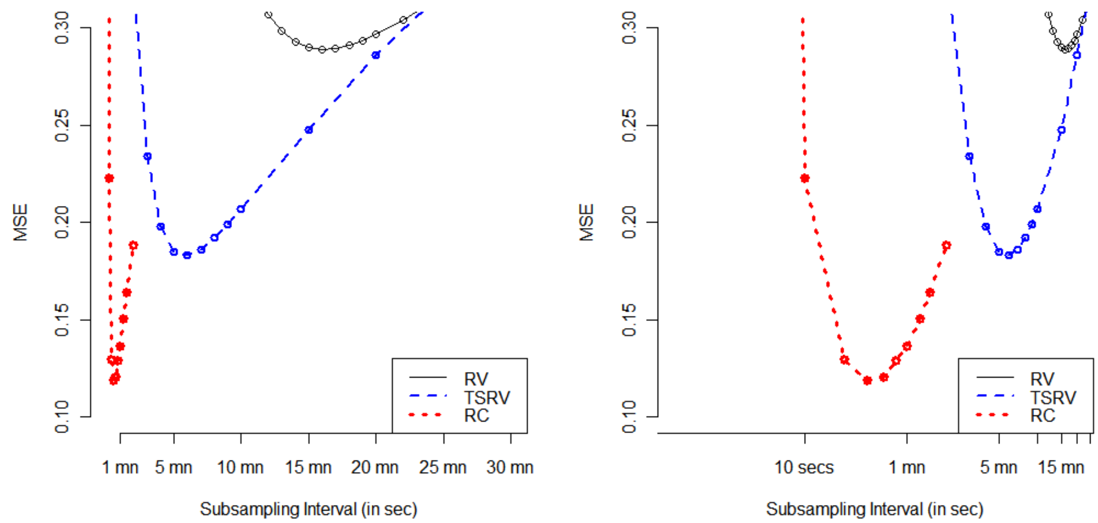

Figure 1.10: Comparisons of RV, TSRV, adjTSRV, and SRC under time dependence noise

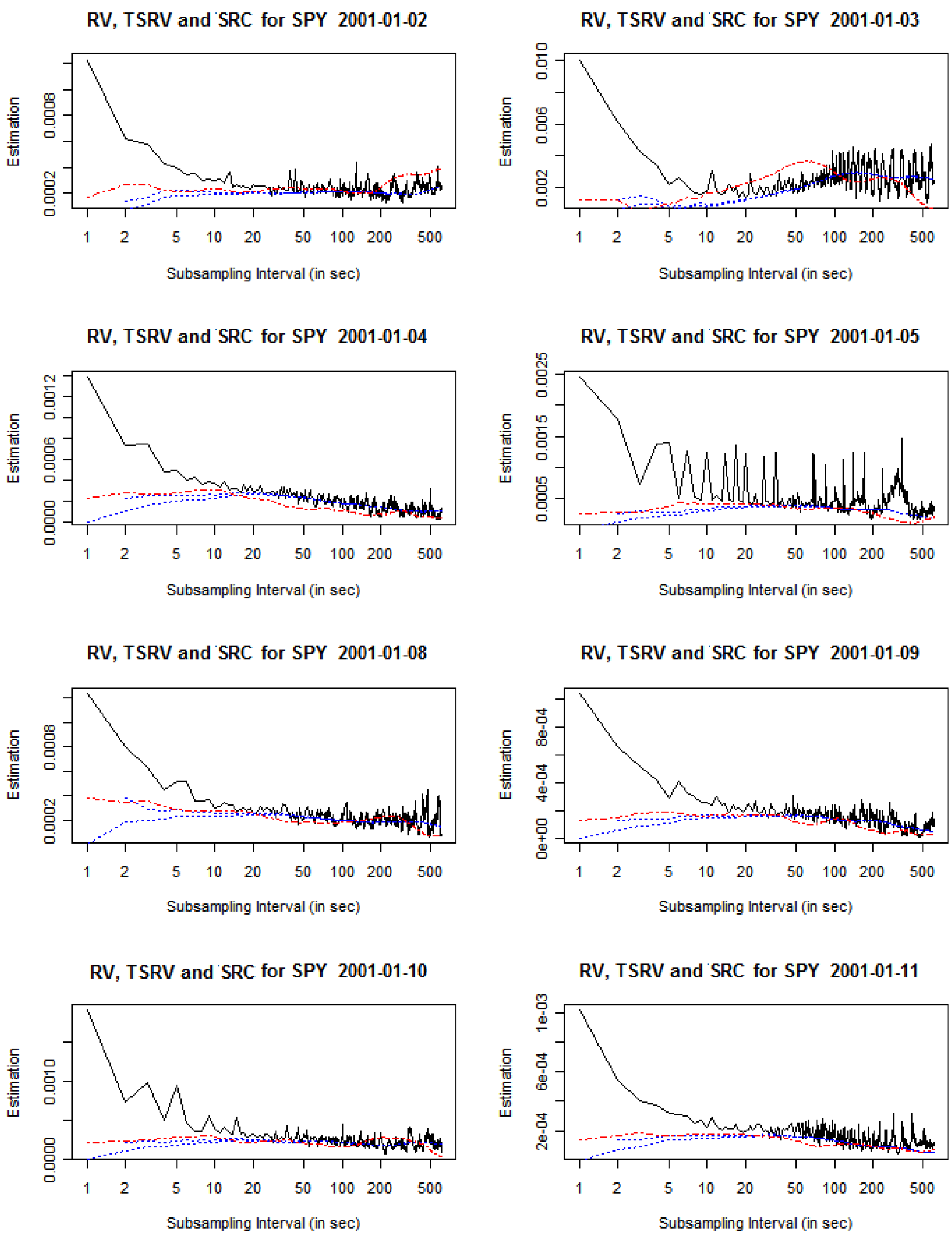

Figure 1.11: Comparisons of RV, adjusted TSRV, and SRC for SPY, computed on a daily basis

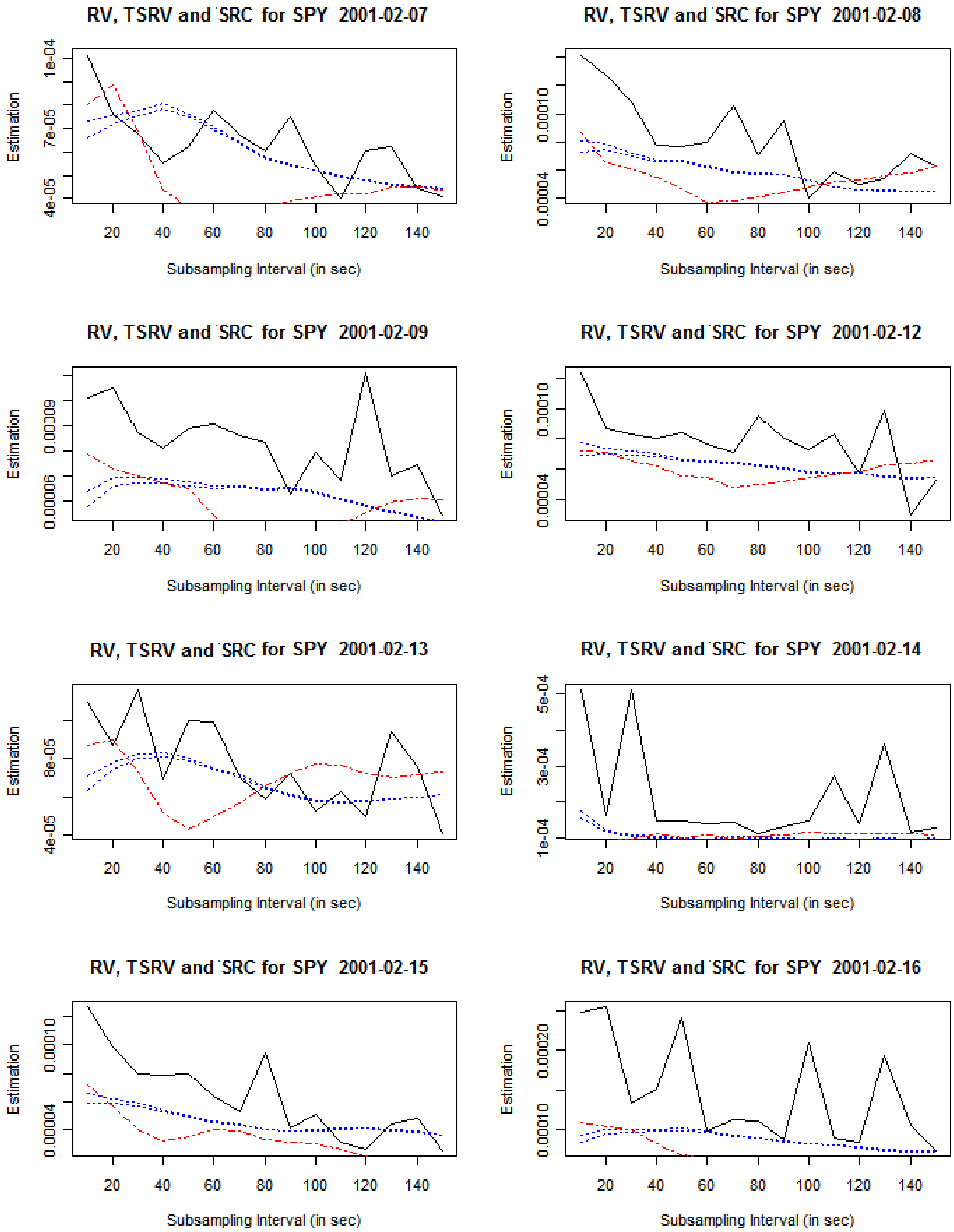

Figure 1.12: Robustness of RV, adjusted TSRV, and SRC for SPY, computed on a daily basis

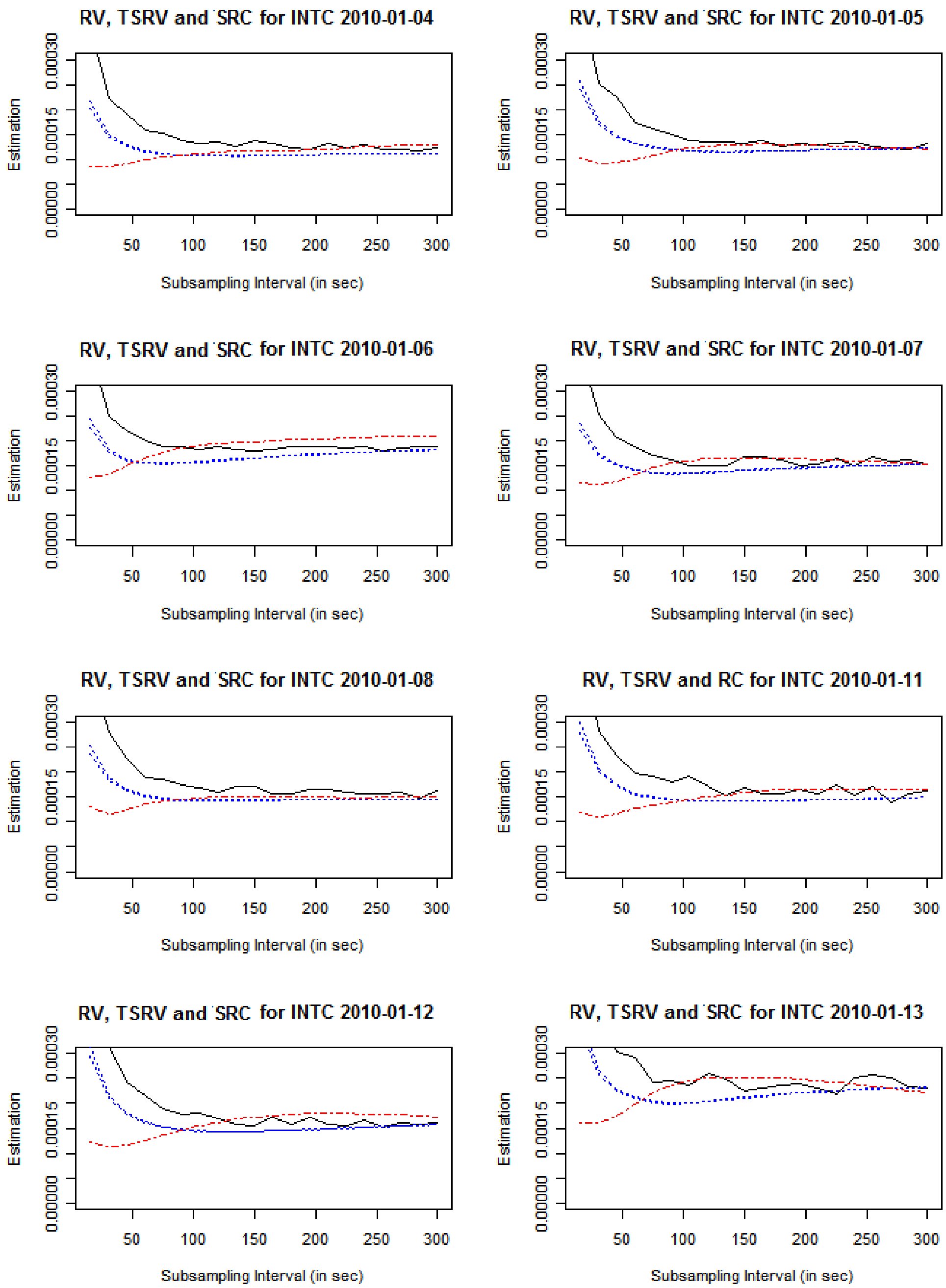

Figure 1.13: Comparisons of RV, adjusted TSRV, and SRC for Intel, computed on a daily basis for 2010 data

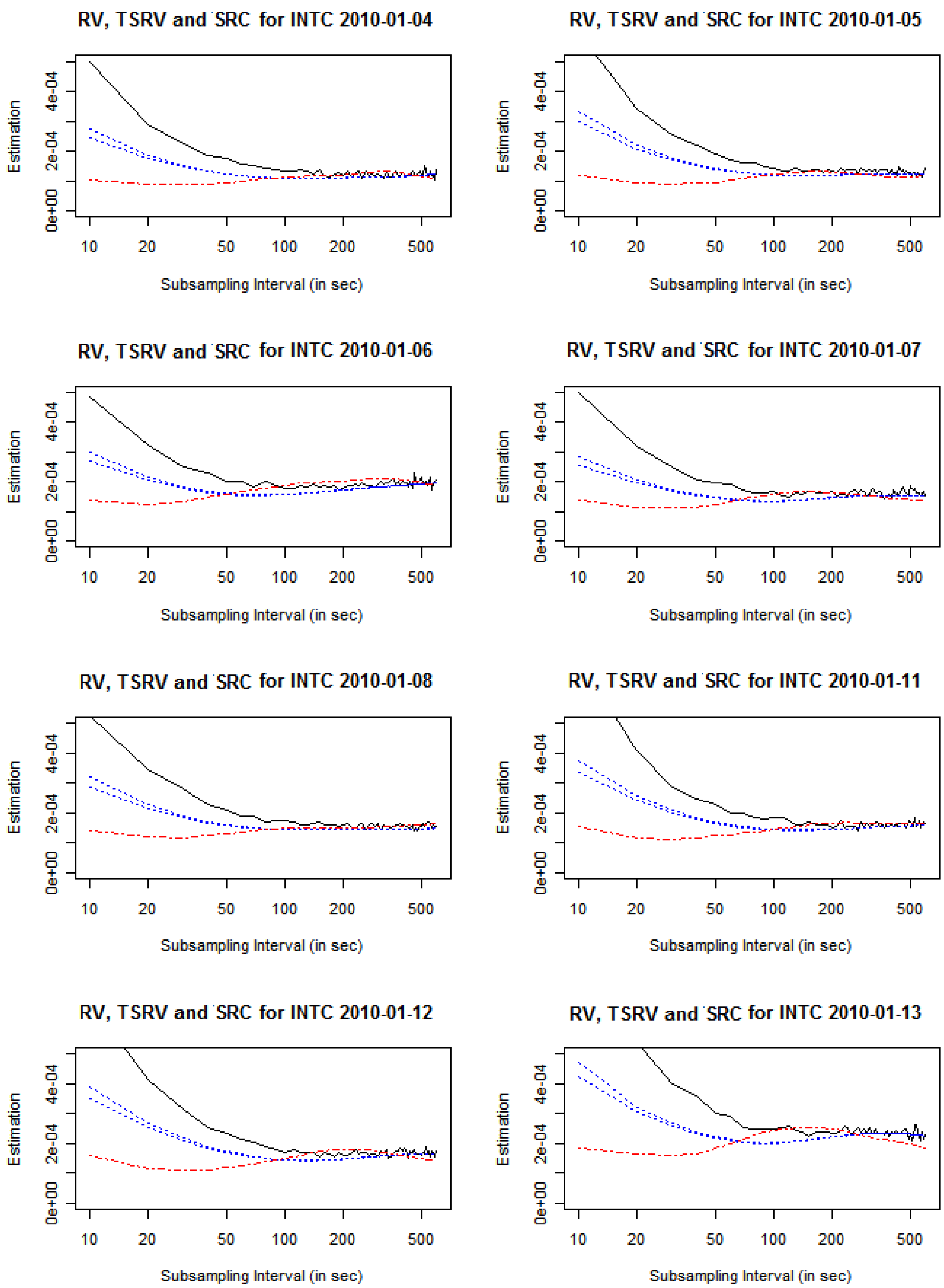

Figure 1.14: Log scale Comparisons of RV, adjusted TSRV, and SRC for Intel, computed on a daily basis for 2010 data

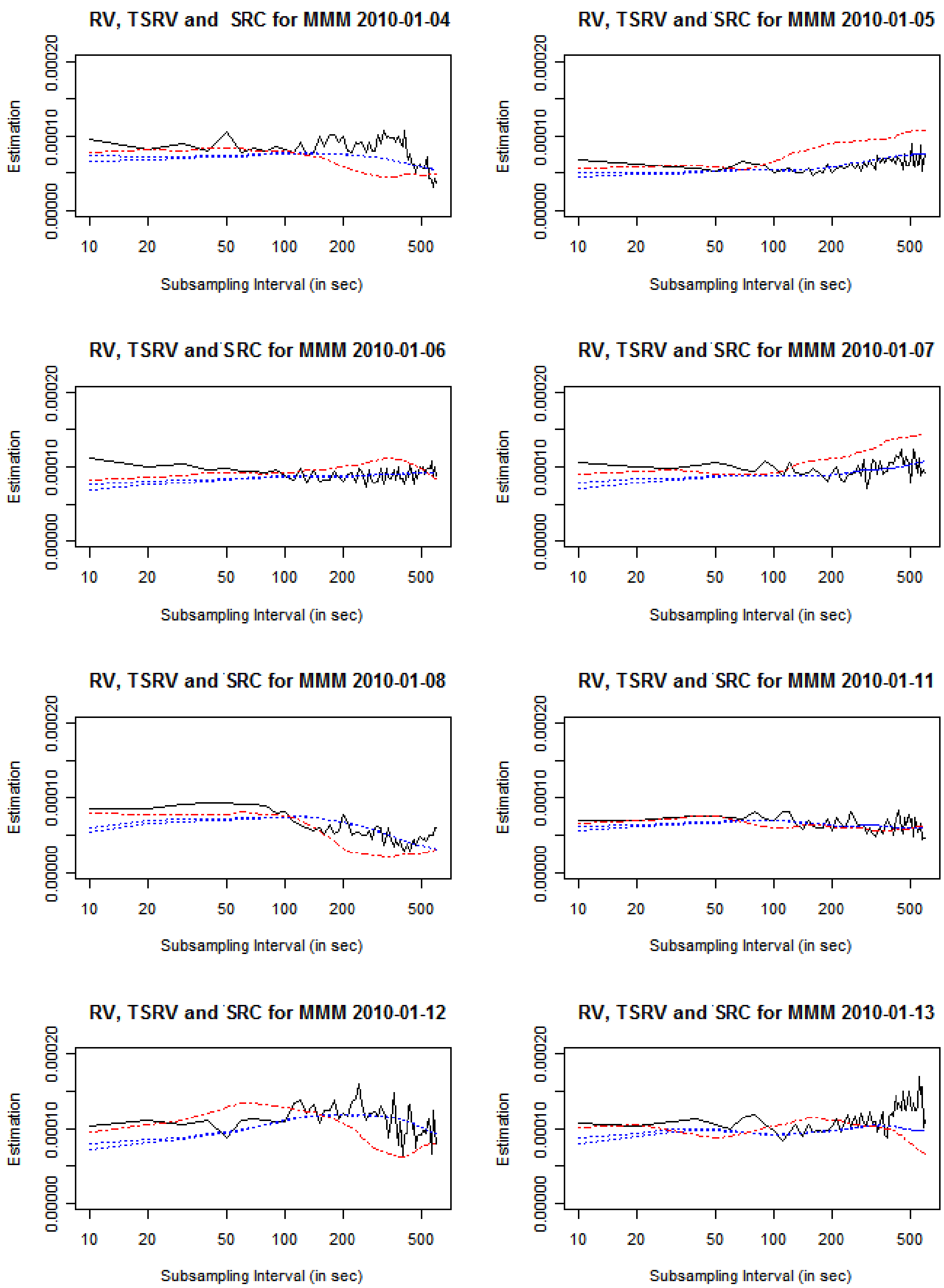

Figure 1.15: Log Scale Comparisons of RV, adjusted TSRV and SRC for MMM, computed on a daily basis for 2010 data

2.0 TESTING FOR JUMPS USING HIGH FREQUENCY DATA WITH NOISE

2.1 INTRODUCTION

2.1.1 Motivations: Nontrivial Jumps

Efforts to prove the existence of jumps and study their implications have a long history, going back to Merton (1976). The studies of Barndorff-Nielsen and Shephard (2006), Andersen et al. (2007) and Huang and Tauchen (2005) have given nonparametric evidence for the presence of nontrivial jumps.

Sometimes, the jump is large enough to be detected by a simple glance as the plots on the top panel of Figure 2.1, and these large jumps could be easily associated with macroeconomic news. For example, the timing of the jump in the DM/$ exchange rate, as in the first plot of Figure 2.1, as evidenced by the apparent discontinuity at 13:30, corresponds exactly to the release of the U.S. trade deficit for the month of October. Quoting from the Wall Street Journal: "The trade gap swelled to a record $17.63 billion in October, sending the dollar and bonds plunging.". The timing of the jump in the stock market on June 30, 1999 corresponds exactly to the time of the 0.25% increase in the Fed funds rate at 13:15. The timing of the jump in the Bond market corresponds to the release of the National Association of Purchasing Managers (NAPM) index at 9:00.

However, most of the time, a visual inspection can not give clear evidence for whether a small or medium size jump belongs to a jump component, as shown in these plots on the bottom panel of Figure 2.1. Thus, it is important to provide the formal statistical testing of jumps.

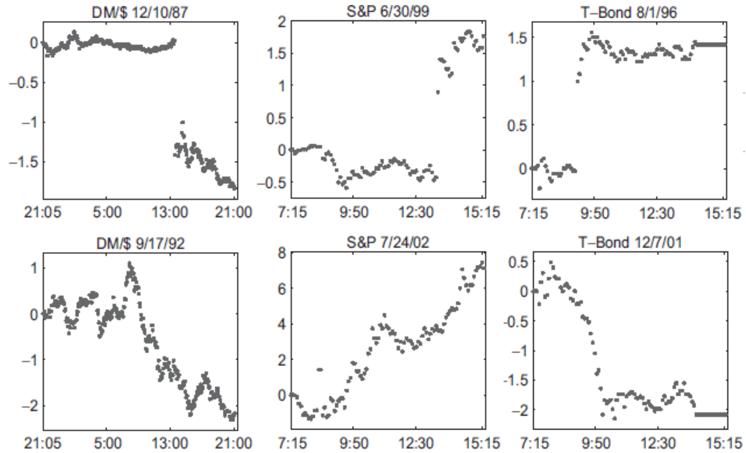

Figure 2.1: Evidence of Jumps in Real High Frequency Financial Data

This plot is from Anderson (2006). It shows the five-minute increments in the log prices for FX, equity and bond markets. For ease of comparison, the log price has been normalized to zero at the beginning of each day, so that a unit increment on the plots corresponds to a 1% return in the log prices.

High frequency data enables researchers to develop nonparametric approaches to accurately test and estimate jumps: Barndorff-Nielsen and Shephard (2006, 2006b) designed realized bipower variation (RBV) and realized multipower variation (RMV) which could separate the continuous part of the total realized variation. Further, they constructed statistical tests using the ratio or difference of RBV and total quadratic variation, and studied their asymptotic distributions under the null hypothesis (no jump). Aït-Sahalia and Jacod (2009a, 2009b, 2009c) provided a series of studies about jumps based on an Itô semimartingale: in Aït-Sahalia and Jacod(2009a), they constructed a nonparametric test statistic for the presence of both finite large or infinite many small jumps by the ratio of realized p-power variation ($p > 3$) with two different time scales (Δ_n and $k\Delta_n$); in Aït-Sahalia and Jacod(2009b), they defined a generalized index of jump activity to study the behavior of infinite but small jumps; in Aït-Sahalia and Jacod(2009c), they construct a new statistical test for the presence of Brownian Motion, in favor of the pure jump process, which is meaningful, because the mathematical treatment of pure jump models is quite different from the models combining Brownian Motion and jumps.

The statistical test in Aït-Sahalia and Jacod(2009a) is very powerful because it works as soon as the price process follows an Itô semimartingale, and it depends neither on the law of the process nor on the coefficients of the equation which it solves. Also, the availability of asymptotic distributions under both alternatives enables us to construct tests with a given significance level, and to calculate the corresponding test power.

However, there is a trade-off between asymptotic means and asymptotic variances for this test statistic that the difference of asymptotic means and the asymptotic variances are increasing at the same time with p and k, some of the parameters in the test statistic. To make this hypothesis test more powerful in the application, Fan and Fan (2009) proposed a new test statistic based on the idea of variance reduction. The principles and details for the classical variance reduction method - 'control variable' can be found in Glasserman (2004). This method consistently smaller asymptotic variances compared with the old one. They further developed an approach to detect the jump locations, using a multiple comparison method.

Other related works include Carr and Wu (2003), Mancini (2004), and Johannes et. al. (2004). These works have given much insight into the effect of jumps with different behaviors, but few of the resulting procedures is robust with respect to microstructure noise. As we know, the only systematic study to estimate jumps from noisy data is by Fan and Wang (2007), who developed wavelet methods for jump testing and estimating based on a Compound Poisson process.

2.1.2 Contributions of My Work and Structure of This Paper

Our purpose in this chapter is to develop a general study considering both microstructure noise and jumps based on high frequency data. There are mainly two contributions:

1. This paper generalizes the statistical tests of jumps in Aït-Sahalia and Jacod (2009a) and Fan and Fan (2010), based on discretely observed high frequency data, without consid-

ering microstructure noise. Compared with the previous work, our approach gives more flexible choices for different sampling frequencies and has smaller asymptotic variance under both null and alternative hypotheses (thus smaller type II error).

2. This paper further designs a new statistical test of jumps. The power of this new test is its robustness with the i.i.d. microstructure noise, which is very common in practical applications. This test considers both the jumps and microstructure noise, and thus is more robust and powerful compared with the old test.

The rest of this paper is organized as follows. Section 2.2 describes the model assumptions of *Itô* semimartingale and related notations. Section 2.3 discusses the two-scales subsampling methods. The construction of Realized Multi-Power Covariances(RMPC) and their asymptotic properties are studied in section 2.4. In section 2.5.1, we re-examine the properties of RMPC in the case of i.i.d. microstructure noise, and show that the old jump test is invalid. A new test method is proposed in section 2.5.2. We also study its asymptotic properties. In section 2.6, we describe the Monte Carlo simulations to compare the new method RPMC with the old one. section 2.7 concludes with a summary and points to future work.

2.2 NOTATION, DEFINITION, AND BACKGROUND

2.2.1 Itô semimartingales

In this paper, the underlying process X is assumed to be a 1-dimensional Itô semimartingale defined on a filtered probability space $(\Omega, F, (F_t)_{t\geq 0}, P)$. Mathematically, it is written as:

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + \int_0^t \int_{\mathbb{R}} x\mathbb{1}_{(|x|<a)}(\mu-\nu)(ds,dx) + \int_0^t \int_{\mathbb{R}} x\mathbb{1}_{(|x|>a)}\mu(ds,dx), \tag{2.1}$$

where

- W_s is the standard Wiener process or Brownian Motion;

- b_s and σ_s are optional (for ex., cadlag function) processes;
- μ is the random measure defined on $\Omega \otimes (\mathbb{R}^+ \times \mathbb{R})$: if we denote the size of the jump of X at time t as $\Delta X_t = X_t - X_{t-}$, then $\mu(\omega, ds, dx) = \sum_{s>0, \Delta X_s(\omega)\neq 0} \mathbb{1}(s, \Delta X_s(\omega))(ds, dx)$
- ν is the predictable compensator of μ, which is the unique measure on $\mathbb{R}^+ \times \mathbb{R}$ which can be written as $\nu(dt, dx) = dt \times \lambda(dx)$, where λ is σ-finite or infinite measure without atoms, and, for any Borel set A of $\mathbb{R}$ and a positive time t, the difference $\mu((0,t]\times A) - \nu((0,t]\times A)$ is a martingale on (Ω, F, P).
- a could be any deterministic value. It is used to distinguish "small jumps" and "big jumps", which are represented respectively by the last two integrations in (2.1).

  There are finitely many large jumps to ensure that the large jump integral is finite, but there may be infinitely many small jumps.

This is a standard setup and more details are in Jacod and Shiryaev (2003). Before we continue our discussion, we need to present some basic assumptions which are similar with those in Jacod (2007).

Assumption 2.1. *1. The process X_t has the form (2.1), and the volatility process σ_t also follows another Itô semimartingale of the form:*

$$\sigma_t = \sigma_0 + \int_0^t \tilde{b}_s ds + \int_0^t \tilde{\sigma}_s d\tilde{W}_s + \int_0^t \int_{\mathbb{R}} x\mathbb{1}_{(|x|<a)}(\tilde{\mu} - \tilde{\nu})(ds, dx) + \int_0^t \int_{\mathbb{R}} x\mathbb{1}_{(|x|>a)}\tilde{\mu}(ds, dx). \tag{2.2}$$

2. The process $\tilde{b}_t$ is locally bounded, which means there exists an increasing sequence of stopping times (τ_n) with $\tau_n \to \infty$, and $(\tilde{b}_{t\wedge\tau_n})$ is bounded by a constant for $\forall n$. So is the process b_t;

3. All paths of $b_t, \tilde{b}_s, \sigma_s, \tilde{\sigma}$ are left continuous with right limits;

4. There exist deterministic nonnegative function $f(x)$ and $\tilde{f}(x)$ satisfying $\int_{\mathbb{R}}(f(x) \vee a)\lambda(dx) < \infty$ and $\int_{\mathbb{R}}(\tilde{f}(x) \vee a)\tilde{\lambda}(dx) < \infty$;

5. $\int_0^t |\sigma_s| ds > 0$ a.s. for any $t > 0$

Assumption 2.2. *The processes of X and σ_t have no common jumps.*

2.2.2 Measurements of Volatility and Jumps

The common measurements of volatility is the integrated volatility $\int_0^t \sigma_s^2\, ds$ as discussed in 1.2. Here, we introduce a number of processes which are similar to the integrated volatility, and all measure different aspects of the variability of X, focusing on continuous and jump components (if jump indeed exists) separately:

$$A(p)_t = \int_0^t |\sigma_s|^p ds, \quad B(p)_t = \sum_{s \le t} |X_s - X_{s-}|^p, \text{for } \forall\ p > 0. \tag{2.3}$$

Under **Assumption** 2.1,

- $A(p)$ measures the integrated p-th absolute power volatility for the continuous component in the semimartingale. It is finite-valued as soon as $p > 0$
- $B(p)$ measures the summation of p-th absolute power jumps for the jump component. If the jump component is trivial (μ is a.s. zero on $\Omega \otimes (\mathbb{R}^+ \times \mathbb{R})$), then $B(p) = 0$.

  When $p = 2$, we have:

$$p \lim_{\overrightarrow{n \to \infty}} [X, X]_t^n = A(2)_t + B(2)_t.$$

It has an additional jump component, compared with the results of RV in (1.4).

2.3 JUMP TESTING BY RATIO OF REALIZED ABSOLUTE POWER USING DIFFERENT SCALES

2.3.1 Realized Absolute P-th Power

To test the existence of jumps, Aït-Sahalia (2007) constructs a nonparametric method, using the ratio of the realized absolute p-th power at two different sample scales. The estimator is:

$$\hat{B}(p, \Delta_n)_t = \sum_{i=1}^{[n=T/\Delta]} |X_i - X_{i-1}|^p.$$

To see the underlying logic, we can roughly write this estimator as:

$$\hat{B}(p,\Delta_n)_t \approx \sum_{i=1}^{[n=T/\Delta]} |\sigma_i\sqrt{\Delta}Z_i + J_i|^p \approx \sum_i |\sigma_i\sqrt{\Delta}Z_i|^p + \sum_j |J_j|^p,$$

$$\hat{B}(p,\Delta_n)_t = \Delta^{p/2-1}\mathbb{E}(|Z|^p)A(p)_t + B(p)_t + O_p(n^{-1/2}).$$

We have the following convergences in probability, locally uniform in t:

$$\text{as } n \to \infty, \begin{cases} p > 2, & \hat{B}(p,\Delta_n)_t \xrightarrow{p} B(p)_t; \\ p = 2, & \hat{B}(p,\Delta_n)_t \xrightarrow{p} A(2); \\ p < 2, & \hat{B}(p,\Delta_n)_t \xrightarrow{p} \infty, \\ & \text{and } \frac{\Delta_n^{1-p/2}}{m_p}\hat{B}(p,\Delta_n)_t \xrightarrow{p} A(p)_t; \\ X \ is \ \text{ continuous}, & \frac{\Delta_n^{1-p/2}}{m_p}\hat{B}(p,\Delta_n)_t \xrightarrow{p} A(p)_t. \end{cases} \tag{2.4}$$

where $m_p = \mathbb{E}(|Z|^p) = \frac{2^{p/2}}{\sqrt{\pi}}\Gamma(\frac{p+1}{2})$ is the pth absolute moment of a standard Gaussian random variable. These properties could also be found in Lepingle (1976) for all semimartingales.

2.3.2 Test Statistics in Aït-Sahalia and Jacod (2009a)

For testing the existence of jumps, they use the ratio of volatility estimates from two different time scales (Δ_n $vs.$ $k\Delta_n$):

$$\hat{S}(p,k,\Delta_n) = \frac{\hat{B}(p,k\Delta_n)}{\hat{B}(p,\Delta_n)},$$

where k is a positive number.

Corollary 2.1.

$$\hat{S}(p,k,\Delta_n) \longrightarrow_p \begin{cases} k^{p/2-1}, & under \ \ H_0; \\ 1, & under \ \ H_\alpha. \end{cases}$$

Proof. This is the same as theorem 1 in Aït-Sahalia et. al.(2009a). □

If we choose $k = 2$ and $p = 4$, from the result above, we know that the test statistic converges to 2 for the paths with jumps; and converges to 1 for the paths without jump.

2.4 JUMP TESTING FROM REALIZED MULTI-POWER COVARIANCES (RPMC)

2.4.1 Construction of RMPC

To retrieve useful information from these measurements defined in (2.3), we define a general family of estimators as:

$$\widehat{RMPC}(X, m, \vec{k}, \vec{p}, \vec{d}) = \sum_{i=|\vec{k}|+|\vec{d}|}^{n} \big(\prod_{j=1}^{m} |X_{r_{i,j}} - X_{l_{i,j}}|^{p_j}\big), \tag{2.5}$$

where

- m is the number of terms in each cross products;
- $\vec{k} = [k_1, \ldots, k_m]'$ are the sampling intervals for the cross terms, k_i is the positive integer, and $|\vec{k}| = \sum_{i=1}^{m} k_i$;
- $\vec{p} = [p_1, \ldots, p_m]'$ are the powers for the cross terms and $p_i > 0$;
- $\vec{d} = [d_1, \ldots, d_{m-1}]'$ are the distances between the adjacent cross terms, d_i is the positive integer, and $|\vec{d}| = \sum_{i=1}^{m-1} k_i$;
- $r_{i,1} = i\Delta,\ l_{i,1} = (i - k_1)\Delta$, and Δ is the smallest time interval;
- $r_{i,j} = r_{i,j-1} - k_{j-1} - d_{j-1}$ and $l_{i,j} = l_{i,j-1} - d_{j-1} - k_j$, for $j = 2, \ldots, m$.

This is the most general framework of estimators similar as the RV, and its construction is showed in Figure 2.2.

2.4.2 Specific Examples

To illustrate its generality, we list several specific examples here:

- When $m = 1, k = 1, \vec{p} = p, d = 0$, $\widehat{RMPC}(X, 1, 1, p, 0)$ reduces to be

$$\hat{B}(p, \Delta_n)_t = \sum_{i=1}^{n} \big(|X_{i\Delta} - X_{(i-1)\Delta}|^p\big),$$

which is the denominator in the test statistics constructed by Aït-Sahalia and Jacod (2009a); that paper proved the following asymptotic properties.

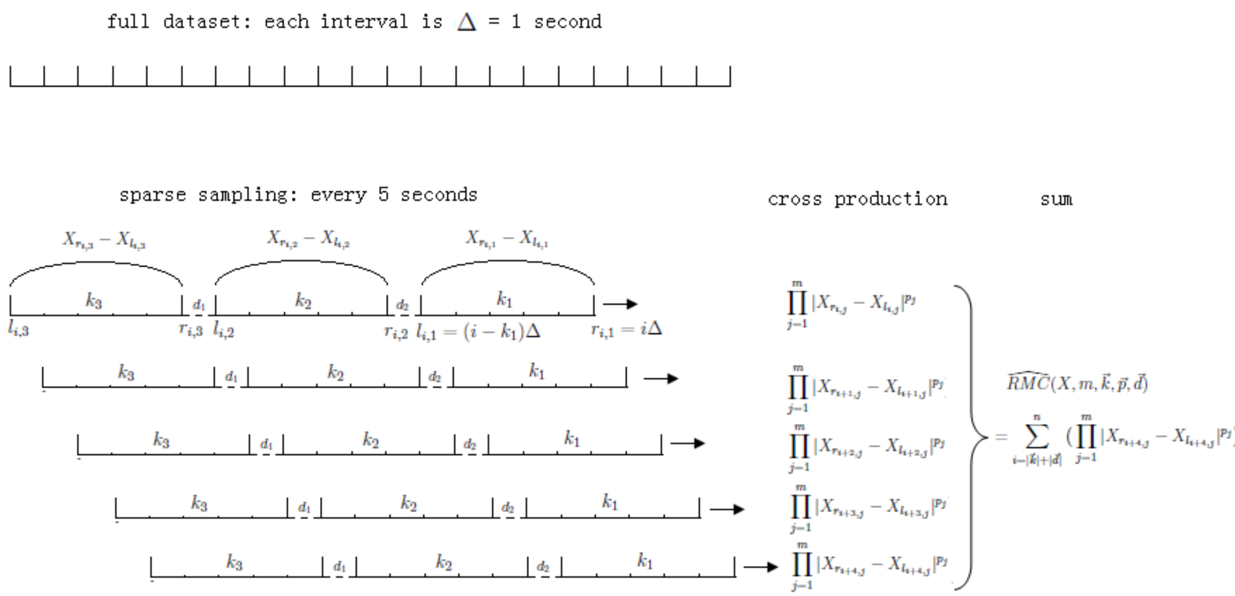

Figure 2.2: Illustration of the construction of Realized Multi-Power Covariances

This example is constructed by the summation of the cross products. Here,

$$m = 3,\ k_1 = k_2 = k_3 = 5,\ d_1 = d_2 = 1.$$

- When $m = 1, k > 1, \vec{p} = p, d = 0$, $\frac{1}{k}\widehat{RMPC}_n(X, 1, k, p, 0)$ reduces to be the most optimal choice in the weighted estimator $\sum_{l=1}^{K} a_j \hat{B}(p, K\Delta)_l$ constructed by Fan and Fan (2010), and it has similar asymptotic result as above, just replacing Δ by $K\Delta$ in (2.4);
- When $m > 1, k = 1, \vec{p} = \vec{r}, \vec{d} = 0$, $\widehat{RMPC}_n(X, m, 1, \vec{r}, 0)$ reduces to be the multipower variation in BNS (2006).

2.4.3 Construction of Test Statistics

Based on the estimators constructed and studied in the last two sections, we construct a new test statistic:

$$\hat{S}(X, K_1, K_2, p) = \frac{\frac{1}{K_2}\widehat{RMPC}_n(X, 1, K_2, p, 0)}{\frac{1}{K_1}\widehat{RMPC}_n(X, 1, K_1, p, 0)}, \;\; K_2 > K_1. \tag{2.6}$$

Let us compare the new test statistic with those of Aït-Sahalia and Jacod (2009a):

$$\hat{S}(X, K, p)_{AJ} = \frac{\sum_{i=1}^{n=[t/\Delta_{sparse}]} |X_{i\Delta_{sparse}} - X_{(i-1)\Delta_{sparse}}|^p}{\sum_{i=1}^{n=[t/K\Delta_{sparse}]} |X_{iK\Delta_{sparse}} - X_{(i-1)K\Delta_{sparse}}|^p}, \;\; K = K_2/K_1, \;\; \Delta_{sparse} = K_1\Delta. \tag{2.7}$$

and those of Fan and Fan (2010):

$$\hat{S}(X, K, p)_{FF} = \frac{\sum_{i=1}^{n=[t/\Delta_{sparse}]} |X_{i\Delta_{sparse}} - X_{(i-1)\Delta_{sparse}}|^p}{\frac{1}{K}\sum_{i=K}^{n=[t/\Delta_{sparse}]} |X_{i\Delta_{sparse}} - X_{(i-K)\Delta_{sparse}}|^p}, \;\; K = K_2/K_1, \;\; \Delta_{sparse} = K_1\Delta. \tag{2.8}$$

Both of the numerator and denominator in our new test statistic in equation (2.6) utilize all available high frequency data observed at every time interval of Δ, while the numerator and denominator in equation (2.7) separately use only $\frac{1}{K_1}$ and $\frac{1}{K_1 K_2}$ proportion of all available data, and both of them in equation (2.8) use $\frac{1}{K_1}$ proportion of all available data.

The asymptotic properties of RMPC and the new test statistics are given below:

2.4.4 Central Limit Theorem on Paths with Jumps

Lemma 2.1. ***First Convergence in Probability for RMPC:***
If Assumption 2.1 holds, $p > 0$ and $\Delta \to 0$. Then we have the following conditional convergence in probability:

(a) For $p > 2$, $\frac{1}{K}\widehat{RMPC}_n(X, 1, K, p, 0)|\Omega_j \xrightarrow{p} B(p) = \sum_{s\leq t} |X_s - X_{s-}|^p$;

(b) For $p = 2$, $\frac{1}{K}\widehat{RMPC}_n(X, 1, K, p, 0)|\Omega_j \xrightarrow{p} A(p)+B(p) = \int_0^t |\sigma_s|^p ds+\sum_{s\leq t} |X_s - X_{s-}|^p$;

(c) For $p < 2$, $\frac{1}{K}\widehat{RMPC}_n(X, 1, K, p, 0)|\Omega_j \xrightarrow{p} \infty$, *and*
$\Delta^{1-p/2}\frac{1}{K}\widehat{RMPC}_n(X, 1, K, p, 0)|\Omega_j \xrightarrow{p} m_p A(p)$,

where Ω_j denotes the collection of all paths with nontrivial jumps.

Proof. This is the same as (2.4). □

Theorem 2.2. ***First Central Limit Theorem for RMPC:***
Under same assumptions as in lemma 2.1, and additionally $p > 2$, we have the following central limit theorem:

$$\frac{1}{\sqrt{\Delta}}\Big[\frac{1}{K}\widehat{RMPC}_n(X, 1, K, p, 0) - B(p)\Big]|\Omega_j \overset{L}{\Rightarrow} \sum_{s\leq t:|X_s-X_{s-}|>0} \Big[p|J_s|^{p-1}\sum_{d=0}^{K-1}\big(U_s^d + U_s\big)\Big] \triangleq Y_{p,K}. \tag{2.9}$$

where U_s and U_s^d are defined on an extension probability space $(\tilde{\Omega}, \tilde{F}, (\tilde{F}_t)_{t>0}, \tilde{P})$:

- $U_s \overset{L}{=} \sqrt{\kappa}\sigma_{s-}Z_s^L + \sqrt{1-\kappa}\sigma_s Z_s^R$; $\quad U_s^d \overset{L}{=} \sqrt{d}\sigma_{s-}\bar{Z}_s^L + \sqrt{K-d-1}\sigma_s\bar{Z}_s^R$
- $\{\bar{Z}_s^L, Z_s^L, Z_s^R, \bar{Z}_s^R\}_s \overset{i.i.d}{\sim} N(0, 1) \perp \{\kappa_s\}_s \overset{i.i.d}{\sim} U(0, 1)$;
- *Jump locations $\{s : s \leq t, |X_s - X_{s-}| > 0\}$ might be finite or infinite depending on the jump properties.*

Proof. The proof is an extension of the proof of Theorem 2.12 (i) of jacod (2006). It could be easily proved step by step following Theorem 8 in Aït-Sahalia and Jacod (2009a) or Theorem 4 in Fan and Fan (2010). □

Corollary 2.3. *Conditional on Ω_j,*

(a) $Y_{p,K}$ is independent with $Y_{p,K} - Y_{p,1}$ for $\forall K > 1$;

(b) $COV\big[(Y_{p,K} - Y_{p,1}), (Y_{p,K} - Y_{p,1})\big] = \frac{(K_1-1)(3K_2-K_1-1)}{6K_2} D(p)$,

where $D(p) = p^2 \sum_{s \le t: |X_s - X_{s-}| > 0} |J_s|^{2p-2}(\sigma_{s-}^2 + \sigma_s^2)$.

Theorem 2.4. ***First Central Limit Theorem for Test Statistics:***

Conditional on Ω_j*:*

$$\frac{1}{\sqrt{\Delta}}\big[\hat{S}(X, K_1, K_2, p) - 1\big] \overset{L_s}{\Rightarrow} S^j_{p,K_1,K_2}, \tag{2.10}$$

where

$$\tilde{\mathbb{E}}\big(S^j_{p,K_1,K_2}\big) = 0, \tag{2.11}$$

$$\tilde{\mathbb{E}}\big((S^j_{p,K_1,K_2})^2\big) = \frac{2K_1(K_2 - K_1)^2 + (K_2 - K_1)}{6K_1^2 K_2} \frac{D(p)}{B(p)^2} \triangleq V^j. \tag{2.12}$$

Proof. See Appendix A.2. □

2.4.5 Central Limit Theorem on Continuous Paths

Theorem 2.5. ***Second Convergence in Probability for RMPC:***

Under same assumptions as in lemma 2.1, we have the following conditional convergence in probability with different value of p:

$$\Delta^{1-p/2}\frac{1}{K}\widehat{RMPC}_n(X, 1, K, p, 0)|\Omega_c \xrightarrow{p} K^{p/2-1} m_p A_p. \tag{2.13}$$

where Ω_c *is the collection of all continuous paths.*

Corollary 2.6. ***Second Central Limit Theorem for RMPC:***

Under same assumptions as in lemma 2.1, and additionally $p > 2$*, we have the following central limit theorem conditional on* Ω_c*:*

$$\frac{1}{\Delta}\begin{pmatrix} \Delta^{1-p/2}\frac{1}{K_1}\widehat{RMPC}_n(X, 1, K_1, p, 0) - K_1^{p/2-1} m_p A_p \\ \Delta^{1-p/2}\frac{1}{K_2}\widehat{RMPC}_n(X, 1, K_2, p, 0) - K_2^{p/2-1} m_p A_p \end{pmatrix} \overset{L}{\Rightarrow} MVN\Big(0, \Sigma_c\Big), \tag{2.14}$$

where

$$\Sigma_c = \begin{bmatrix} K_1^{-2}\Sigma(K_1) & K_1^{-1}K_2^{-1}\Sigma(K_1, K_2) \\ K_1^{-1}K_2^{-1}\Sigma(K_1, K_2) & K_2^{-2}\Sigma(K_2) \end{bmatrix};$$

$$\Sigma(K) = 2\sum_{d=1}^{K-1} m_p(d, K-d, d) + K^p(m_{2p} - m_p^2);$$

$$\Sigma(K_1, K_2) = 2\sum_{d=1}^{K_1-1} m_p(d, K_1-d, K_2-K_1+d) + (K_2-K_1+1)m_p(0, K_1, K_2-K_1).$$

m_p is the p-th absolute moment of standard Gaussian random variable as before; $m_p(a,b,c)$ is a generalized version defined by $m_p(a,b,c) = \mathbb{E}(|aZ_1+bZ_2|^p|bZ_2+cZ_3|^p)$, where $Z_1,\ Z_2,\ Z_3$ are i.i.d. standard Gaussian random variables.

Proof. The proof is similar to Theorem 2 in Fan and Fan (2010). □

Theorem 2.7. ***Second Central Limit Theorem for Test Statistics:***
Conditional on Ω_c:

$$\frac{1}{\sqrt{\Delta}}\big[\hat{S}(X, K_1, K_2, p) - (\frac{K_2}{K_1})^{p/2-1}\big] \overset{L}{\Rightarrow} S^c_{p,K_1,K_2}, \tag{2.15}$$

where

$$\tilde{\mathbb{E}}\big(S^c_{p,K_1,K_2}\big) = 0, \tag{2.16}$$

$$\tilde{\mathbb{E}}\big((S^c_{p,K_1,K_2})^2\big) = \frac{A_{2p}}{m_p^2 A_p^2}(\frac{1}{K_1})^p(\frac{K_1}{K_2})^2\big[\Sigma(K_2) + (\frac{K_1}{K_2})^p\Sigma(K_1) - 2(\frac{K_1}{K_2})^{p/2}\Sigma(K_1, K_2)\big] \triangleq V^c. \tag{2.17}$$

Proof. See Appendix A.2. □

Considering the asymptotic properties and practical applications, our new test statistic designed here has at least two important advantages:

First, it is more flexible to choose different time scales of any integer $K_2 > K_1 > 0$. For example, we can use the ratio of our estimators under the scale of 45 seconds and 30 seconds to construct the test statistic, while when choosing 30 seconds for the denominator, only 60 seconds, 90 seconds and etc. are available to be used for the numerator in the old test statistics.

Second, it could be proved both theocratically and empirically that the new test statistic has smaller asymptotic variance under both null and alternative hypotheses:

- If we take $K_1 = \Delta, K_2 = K\Delta$, where $\frac{1}{\Delta}$ is the highest available frequency (ex. $\Delta = 1/23400 = 1$ second), then our new test statistic reduces to version using the highest available frequency in Fan and Fan (2010);
- However, both Aït-Sahalia and Jacod (2009a) and Fan and Fan (2010) discussed that we can not use such high frequency data in the application, because of the presence of microstructure noise (which we will discuss in section 2.5.1). 1 min to 3 mins time interval is more common. In this situation, if we take $K_1 = \Delta_{\text{sparse}}, K_2 = K * K_1$, where the K is the same one as that in Aït-Sahalia and Jacod (2009a) and Fan and Fan (2010), then the asymptotic variance of our test statistic under null hypothesis is $\left(\frac{(K-1)^2}{3K} + \frac{K-1}{6K\Delta_{\text{sparse}}^2}\right)\frac{D(p)}{B(p)^2} \approx \frac{(K-1)^2}{3K}\frac{D(p)}{B(p)^2}$. Compared with $\frac{K-1}{2}\frac{D(p)}{B(p)^2}$ and $\frac{(2K-1)(K-1)}{6K}\frac{D(p)}{B(p)^2}$, our new test statistic reduces the variance under null hypothesis by factors of $\frac{2K-2}{3K}$ and $\frac{2K-2}{2K-1}$ respectively.
- Theoretical proof of the variance reduction under alternative hypothesis is more complex, so I will give the numerical results instead in section 2.6.

2.4.6 Testing for Jumps

Finally, it is time to design our tests for jumps:

$$\begin{aligned} H_0: &\quad X_{(\omega)} \textit{ has no jump} \Leftrightarrow B_{(\omega)}(p)_t = 0; \\ H_\alpha: &\quad X_{(\omega)} \textit{ has jumps} \Leftrightarrow B_{(\omega)}(p)_t > 0. \end{aligned} \tag{2.18}$$

The test statistics:

$$\hat{S}(X, K_1, K_2, p) = \frac{\frac{1}{K_2}\widehat{RMPC}_n(X, 1, K_2, p, 0)}{\frac{1}{K_1}\widehat{RMPC}_n(X, 1, K_1, p, 0)} \xrightarrow{L} \begin{cases} N(1, V^j), & \text{under} \;\; H_0; \\ N(k^{p/2-1}, V^c), & \text{under} \;\; H_\alpha. \end{cases}$$

Decision Rule-Rejection Region:

$$RR = \{\hat{S}(p, k, \Delta_n) > x\},$$

where $x \in (1, (\frac{K_2}{K_1})^{p/2-1})$.

Here, V^c and V^j are functions of $D(p), A(p),$ and $B(p)$, which are unknown and need to be estimated in practice. We use similar estimators as those constructed by Aït-Sahalia and

Jacod (2009a):

$$\hat{D}(p) = \sum_{i=1}^{n} \big[|\Delta_i x|^{2p-2} \frac{1}{\parallel I_i \parallel \Delta} \big(\sum_{j \in I_i} (\Delta_i X)^2 \mathbb{1}\{|\Delta_i X| \leq \alpha \Delta^{\gamma}\} \big) \big]. \tag{2.19}$$

where α is a deterministic positive value; $\gamma \in (0, \frac{1}{2})$; I_i is the local window around $i\Delta$, with its window size $\parallel I_i \parallel$ satisfying $\parallel I_i \parallel \to 0$ as $\Delta \to 0$. An estimator of $A(p)$ could be

$$\hat{A}(p) = \frac{\Delta^{1-p/2}}{m_p} \sum_{1}^{n} |\Delta_i X|^p \mathbb{1}\{|\Delta_i X| \leq \alpha \Delta^{\gamma}\}. \tag{2.20}$$

An alternative estimator of $A(p)$ could be found from the family of estimators as discussed in section 2.2.2, for example, $\hat{A}'_p = c\widehat{RMPC}_n(X, m > 1, 1, \vec{p} = [\frac{p}{m}, \ldots, \frac{p}{m}]')$, where c is a normalizing constant. When $m = 2$ or 3, this estimator is almost the same as Realized Bipower or Realized Multipower as designed in BNS(2006a,2006b).

$$\hat{B}(p) = \widehat{RMPC}_n(X, 1, 1, p, 0), \ p > 2. \tag{2.21}$$

Let

$$\hat{V}^j = \frac{2K_1(K_2 - K_1)^2 + (K_2 - K_1)}{6K_1^2 K_2} \frac{\hat{D}(p)}{\hat{B}(p)^2},$$

$$\hat{V}^c = \frac{\hat{A}(2p)}{m_p^2 \hat{A}(p)^2} (\frac{1}{K_1})^p (\frac{K_1}{K_2})^2 \big[\Sigma(K_2) + (\frac{K_1}{K_2})^p \Sigma(K_1) - 2(\frac{K_1}{K_2})^{p/2} \Sigma(K_1, K_2)\big].$$

then, we have the following corollary,

Corollary 2.8. *Under same assumptions as in lemma 2.1, and additonally* $p > 2$, $(\hat{V}^j)^{-1/2}\big(\hat{S}(X, K_1, K_2, p) - 1\big)|\Omega_j$ *and* $(\hat{V}^c)^{-1/2}\big(\hat{S}(X, K_1, K_2, p) - (\frac{K_2}{K_1})^{p/2-1}\big)|\Omega_c$ *both converge stably in law to a standard normal distribution.*

Proof. This corollary is immediately from **theorem** 2.4 and **theorem** 2.7, combined with the properties for stable convergence. □

Type I & Type II Errors:

The two error functions of this jump testing are:

Type I error: $\alpha_n(x) = \mathbb{P}(\hat{S}(X, K_1, K_2, p) \leq x | H_0)$;

Power Function: $\beta_n(x) = \mathbb{P}(\hat{S}(X, K_1, K_2, p) \leq x | H_1)$.

Theorem 2.9. *Assume that* ***Assumption 1*** *holds and the critical value* $c \in (1, (\frac{K_2}{K_1})^{p/2-1})$*. Then we have:*

(a) If $\mathbb{P}(\Omega_j) > 0$*,*$\alpha_n(x) \to 0$*, that is, the rejection region has an asymptotic size 0 if there are jumps in the path;*

(b) $\beta_n(x) \to 1$*, as* $n \to \infty$

Similarly, these results hold if the null and alternative hypotheses are switched.

Proof. The proof is similar to Theorem 6 in Aït-Sahalia and Jacod (2009a) and Theorem 3 in Fan and Fan (2010). □

2.5 A NEW TEST BASED ON RMPC WITH IID NOISE

2.5.1 Influence of Microstructure Noise

As we already discussed in section 1.3, the real world is not as ideal as we expect, the observation noise is very common in high frequency financial data as in (1.3): $Y_t = X_t + \epsilon_t$. Instead of the underlying X_t, we can only observe the noisy Y_t. Therefore, all these results for the estimators constructed based on X_t in section 2.3 and section 2.4 should be revised for the real world based on Y_t. To simplify our explanations, we fix $p = 4$.

Compared with the asymptotic properties for the estimators based on the ideal world (X_t) in lemma 2.1 and Corollary 2.2:

$$\frac{1}{K}\widehat{RMPC}_n(X, 1, K, 4, 0)|\Omega_j \xrightarrow{p} B_4 = \sum_{s \le t} |X_s - X_{s-}|^4,$$
$$\text{and} \tag{2.22}$$
$$\frac{1}{\Delta}\frac{1}{K}\widehat{RMPC}_n(X, 1, K, 4, 0)|\Omega_c \xrightarrow{p} 3KA_4 = 3K\int_0^t \sigma_s^4 ds.$$

The estimators for the real world (Y_t) become:

$$
\begin{aligned}
&\frac{1}{K}\widehat{RMPC}_n(Y,1,K,4,0) \\
&= \frac{1}{K}\sum_{i=K}^{n}(\Delta Y_i)^4 = \frac{1}{K}\sum_{i=K}^{n}(\Delta X_i + \Delta\epsilon_i)^4 \\
&= \frac{1}{K}\sum_{i=K}^{n}\left[(\Delta X_i)^4 + 4(\Delta X_i)^3\Delta\epsilon_i + 6(\Delta X_i)^2(\Delta\epsilon_i)^2 + 4\Delta X_i(\Delta\epsilon_i)^3 + (\Delta\epsilon_i)^4\right]
\end{aligned}
\tag{2.23}
$$

$$
\begin{aligned}
\Longrightarrow \mathbb{E}\Big[\frac{1}{K}\widehat{RMPC}_n(Y,1,K,4,0)\Big] =& [3K\Delta A_4 + B_4] + \frac{n-K+1}{K}\mathbb{E}((\Delta\epsilon)^4) \\
&+ [6\mathbb{E}((\Delta\epsilon)^2)(A_2+B_2)] + O_p(n^{-1/2}).
\end{aligned}
\tag{2.24}
$$

where the first and third terms correspond with the results in (2.22), which behaves differently for continuous paths and jump paths. However, the last two terms are coming from microstructure noise (ϵ) regardless of the existence of jump and they overwhelm the first two terms when sample size is large and noise is not small. In this case, the old test statistic: $\hat{S}(Y,K_1,K_2,4) = \frac{\frac{1}{K_2}\widehat{RMPC}_n(Y,1,K_2,4,0)}{\frac{1}{K_1}\widehat{RMPC}_n(Y,1,K_1,4,0)}$ always converges in probability to $\frac{K_1}{K_2}$ as $n \to \infty$ and $n/\max\{K_1,K_2\} \to \infty$. Thus, it loses the power to distinguish continuous and jumps paths.

This influence of microstructure noise to our jump testing is obvious when we apply the old test to real data as in Figure 2.3, where many test statistics cluster at $\frac{K1}{K2} = 0.5$.

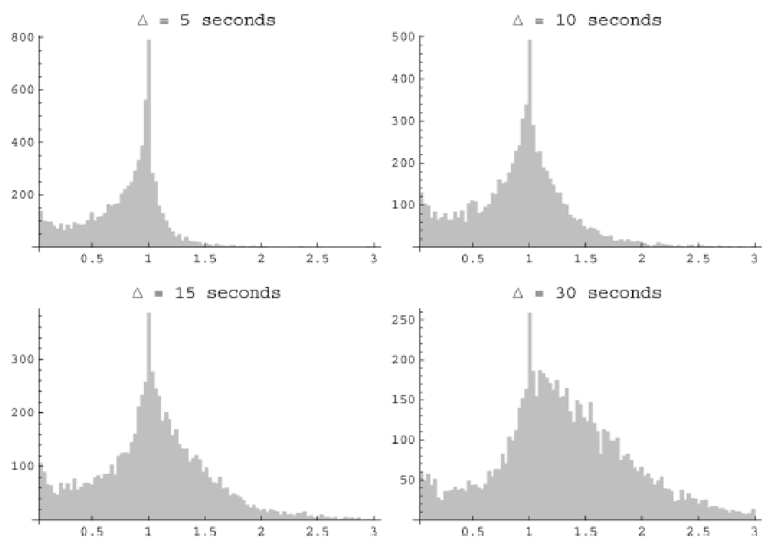

Figure 2.3: Evidence of Noise in Jump Test

This plot is constructed in Aït-Sahalia and Jacod (2009a). It shows the empirical distribution of the old test statistic $\hat{S}(Y, K_1, K_2, 4)$ *for different values of the sampling interval* Δ_n*, based on 2005 DJIA data.*

2.5.2 New Test Statistics based on RMPC

An obvious approach to solve the problem is to counteract the biasing effect from microstructure noise through the linear combination of different estimators. To achieve this purpose, we need to introduce another specific construction from our general framework of estimators:

$$
\begin{aligned}
\frac{1}{K}\widehat{RMPC}_n(Y,2,K,4,d) &= \frac{1}{K}\sum_{i=K+d}^{n}(\Delta Y_i)^2(\Delta Y_{i-d})^2 \\
&= \frac{1}{K}\sum_{i=K+d}^{n}(\Delta X_i+\Delta\epsilon_i)^2(\Delta X_{i-d}+\Delta\epsilon_{i-d})^2 \\
\approx [K\Delta A_4] + 2\mathbb{E}[(\Delta\epsilon)^2](A_2+B_2) &+ \frac{n-K-d+1}{K}\mathbb{E}[(\Delta_i\epsilon)^2(\Delta_{i-d}\epsilon)^2].
\end{aligned} \tag{2.25}
$$

In addition, it is not hard to see that $\frac{\widehat{RMPC}_n(Y,1,1,4,0)}{n}$ is an efficient estimator of $\mathbb{E}[\Delta\epsilon^4]$ and $\frac{\widehat{RMPC}_n(Y,2,1,2,d)}{n-1-d}$ is an efficient estimator of $\mathbb{E}[(\Delta\epsilon^2)^2]$. Therefore, our new estimator considering the bias adjustment as:

$$
\begin{aligned}
NB(Y,K) = &[\frac{1}{2K}\widehat{RMPC}_n(Y,1,2K,4,0) - 3\frac{1}{K}\widehat{RMPC}_n(Y,2,K,4,K)] \\
&\approx 3K\Delta A_4 + B_4,
\end{aligned} \tag{2.26}
$$

or

$$
\begin{aligned}
&NB^{adj}(Y,D_1,D_2,d) \\
=&[\frac{1}{D_1}\widehat{RMPC}_n(Y,1,D_1,4,0) - 3\frac{1}{D_2}\widehat{RMPC}_n(Y,2,D_2,4,d)] \\
&-[\frac{n-D_1+1}{D_1}\frac{\widehat{RMPC}_n(Y,1,1,4,0)}{n} - 3\frac{n-D_2-d+1}{D_2}\frac{\widehat{RMPC}_n(Y,2,1,2,d)}{n-1-d}] \\
\approx&3(D_1-D_2)\Delta A_4 + B_4,
\end{aligned} \tag{2.27}
$$

which is a small sample adjusted version.

Its asymptotic properties are as below:

Lemma 2.2. ***Convergence in Probability of Estimator:***
As $n\to\infty$, $NB^{adj}(Y,D_1,D_2,d) \xrightarrow{p} B_4$;
If $B_4 = 0$, *then* $\frac{1}{\Delta(D_1-D_2)}NB^{adj}(Y,D_1,D_2,d) \xrightarrow{p} 3A_4$.

Proof. See Appendix A.2. □

Now we can construct our new test statistic:

$$S^{New}(Y, K_1, K_2, 4) = \frac{NB^{adj}(Y, K + D_2, K, 1)}{NB^{adj}(Y, K + D_1, K, 1)}. \tag{2.28}$$

This test statistic has the asymptotic property as we expect to distinguish continuous paths and paths with jumps:

Theorem 2.10. ***Convergence in Probability of Test Statistics:***

$$As\ n \to 0,\ S^{New}(Y, K_1, K_2, 4) \xrightarrow{p} \begin{cases} 1, & \textit{on continuous paths;} \\ \frac{D_2}{D_1}, & \textit{on paths with jumps.} \end{cases} \tag{2.29}$$

Proof. This result immediately follows Lemma 2.2, using Theorem 5.5.4 in Casella and Berger (2002). □

2.6 SIMULATIONS AND COMPARISONS

2.6.1 Continuous Stochastic Volatility Models without Noise

$$\begin{aligned} \frac{dX_t}{X_t} &= \sigma_t\ dW_t^1, \\ v_t = \sigma_t^2,\ dv_t &= \kappa(\bar{v} - v_t)dt + \gamma\sqrt{v_t}dW_t^2, \\ \mathbb{E}[dW_t^1 dW_t^2] &= \rho dt. \end{aligned} \tag{2.30}$$

We simulate 100 sample paths of prices over a one-day period with parameters $\bar{v} = 0.4^2,\ \gamma = 0.5,\ \kappa = 5$ and $\rho = -0.5$. This setup was similar to that in Aït-Sahalia and Jacod (2009a) and Fan and Fan (2010), and all parameters are realistic for a stock studied in Aït-Sahalia and Kimmel (2007). The sampling frequencies are taken as $K_1\Delta_n$ =5 seconds, 30 seconds, 1 minutes and 2 minutes. In each of these simulations, we study our test statistic $\hat{S}(X, K_1, K_2, p)$ for $\frac{K_2}{K_1} = 2$ and 3.

Figure 2.4 shows one continuous path generated from the stochastic volatility model in (2.30).

Figure 2.5 shows the Monte Carlo simulation of our new test statistics based on these continuous paths. It is consistent with the results in Theorem 2.7, which says the test statistics is asymptotically normally distributed around $(\frac{K_2}{K_1})^{p/2-1}$.

Table 2.1 compares our new test statistics (RMPC) with the previous methods in Aït-Sahalia and Jacod (2009a) (AJ) and Fan and Fan(2009) (FF) under the assumption that the path does not have a jump component. While all of them have the asymptotic mean close to $\frac{K_2}{K_1}$, the asymptotic standard deviation for RMPC is consistently smaller than that of AJ and FF. The problem of optimal sampling frequency to achieve the most powerful test statistic is also another interesting topic, but not our focus in this chapter.

2.6.2 Stochastic Volatility Models with Compound Poisson Processes without Noise

To conduct the Monte Carlo simulations for comparisons of methods conditional of paths with jumps, we consider the following model:

$$\begin{aligned} \frac{dX_t}{X_t} &= \sigma_t \; dW_t^1 + J_t dN_t, \\ v_t = \sigma_t^2, \; dv_t &= \kappa(\bar{v} - v_t)dt + \gamma\sqrt{v_t}dW_t^2, \\ \mathbb{E}[dW_t^1 dW_t^2] &= \rho dt. \end{aligned} \tag{2.31}$$

where N_t is an independent Poisson process with intensity $\lambda = 3$, J_t measures the jump size, and $W_t^1$and W_t^2 are both Brownian Motions.

Figure 2.6 shows one continuous path generated from the stochastic volatility model in (2.31).

Figure 2.7 shows the Monte Carlo simulation of our new test statistics based on these continuous paths. It is consistent with the results in Theorem 2.4 that the test statistics is asymptotically normally distributed around 1.

Table 2.2 compares our new test statistics (RMPC) with the methods of AJ and FF under the assumption that the path has nontrivial jump component. While all of them have the asymptotic mean close to 1, the asymptotic standard deviation of RMPC is consistently smaller than that of AJ and FF again.

2.6.3 Jump test for High Frequency Data with i.i.d. Microstructure Noise

To compare the old with the new test statistics, we use the same setup of stochastic volatility model with Poisson Processes as in (2.30) to generate the underlying log price X_t. As for the microstructure noise, we simply generate the noise from the normal distribution with different noise levels ($\mathbb{E}\epsilon^2 = 0.01^2, 0.005^2,$ and 0.0005^2).

Table 2.3 and Table 2.4 summarize the comparison results under different noise levels and different choices of sampling frequencies. It is obvious that the new test statistic is indeed robust with observation noise, and is consistently much better than the old test statistic.

2.7 CONCLUSION AND FUTURE WORK

2.7.1 Asymptotic Results, Optimal Sampling Size and Convergence Rates

1. In this chapter, we have studied the asymptotic means of the proposed estimators and the new test statistic $S^{New}(Y, K_1, K_2, 4)$, and we also compared the asymptotic variance of the new test statistics with the old one through Monte Carlo sfimulations. For the next step, we plan to do an analytic study on the asymptotic distributions, and compare the old and new test statistics when applied to the real high frequency data on the stock market, foreign exchange market, and bond market.

2. We have given a general framework, which provides a whole family of estimators with flexible choices of different sampling frequencies. On the other hand, it is interesting to study how to optimally choose the sampling frequencies. The common method is to minimize the mean square error (MSE) or minimize the Kullback-Leibler divergence of the asymptotic distributions under null and alternative hypotheses.
3. As studied in section 1.3.1, under the influence of microstructure noise, the best attainable convergence rate of RV is $n^{1/4}$, which is also the same even for the simplest available parametric model: $dX_t = \sigma_1 dW_t$ and $Y_t = X_t + \sigma_2 Z_t$. Is $n^{1/4}$ the best attainable convergence rate in the testing of jumps? If so, how should we construct our new test statistic? These questions will be studied in my future work.

2.7.2 Empirical Study of Microstructure Noise

Until now, we only study the simplest case of microstructure noise: it is independent with stock price, and itself is i.i.d distributed. However, the real world is much more complex. It is reasonable to empirically study two possible dependent structures:

1. As in the paper of Aït-Sahalia, Mykland, and Zhang (2011), there is some evidence of serially dependent structure for the microstructure noise.
2. Li and Mykland (2007) argue that it is reasonable to believe that there might be some dependent relationship between microstructure noise and the underlying stock prices. Following the notation in Li and Mykland (2007), the law of Y_t could be:

$$P(Y_t \leq y | X_t) = Q(X_t, y) \tag{2.32}$$

Therefore, our next step is to study whether our new test statistic is robust with respect to these dependent structure. If not, how to adjust our test statistic under different dependent structure?

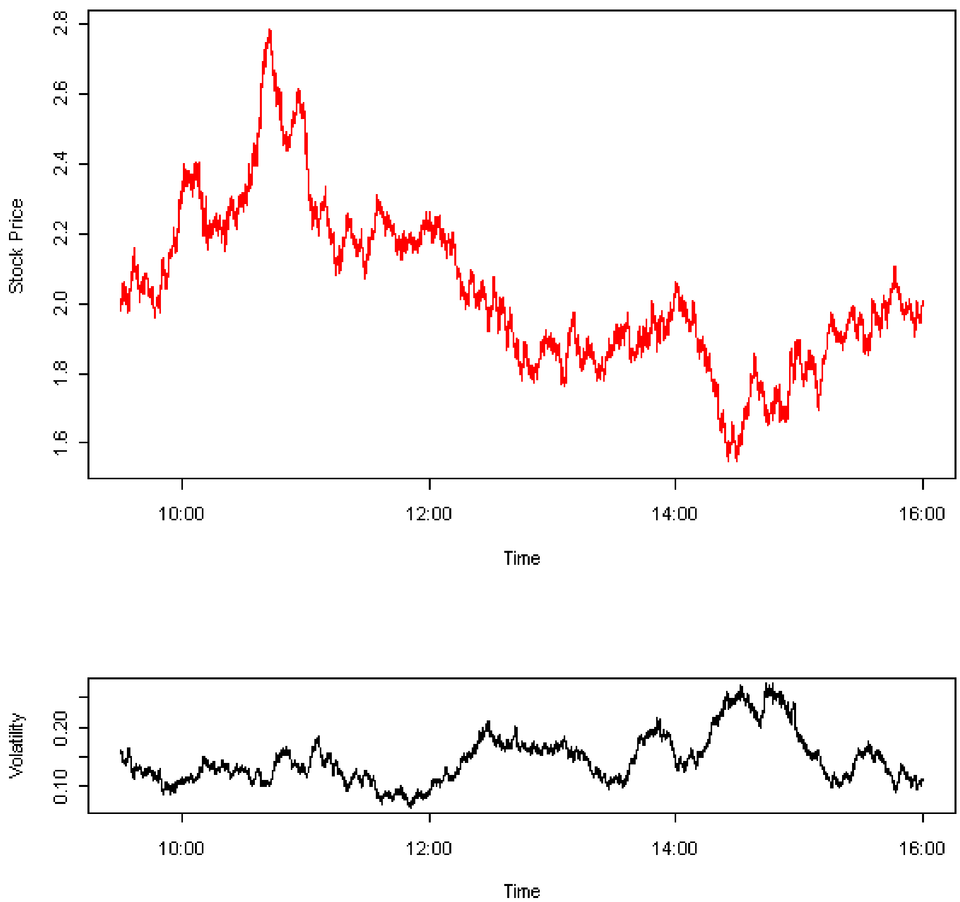

Figure 2.4: One Continuous Path from Our Simulations

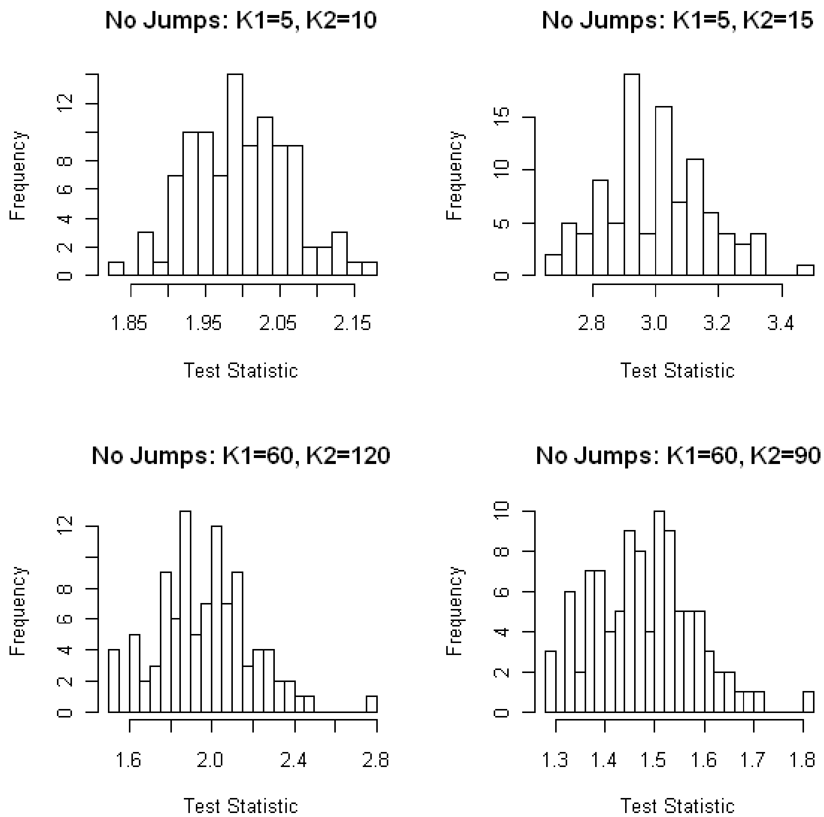

Figure 2.5: Monte Carlo asymptotic distribution of our new test statistics for continuous paths

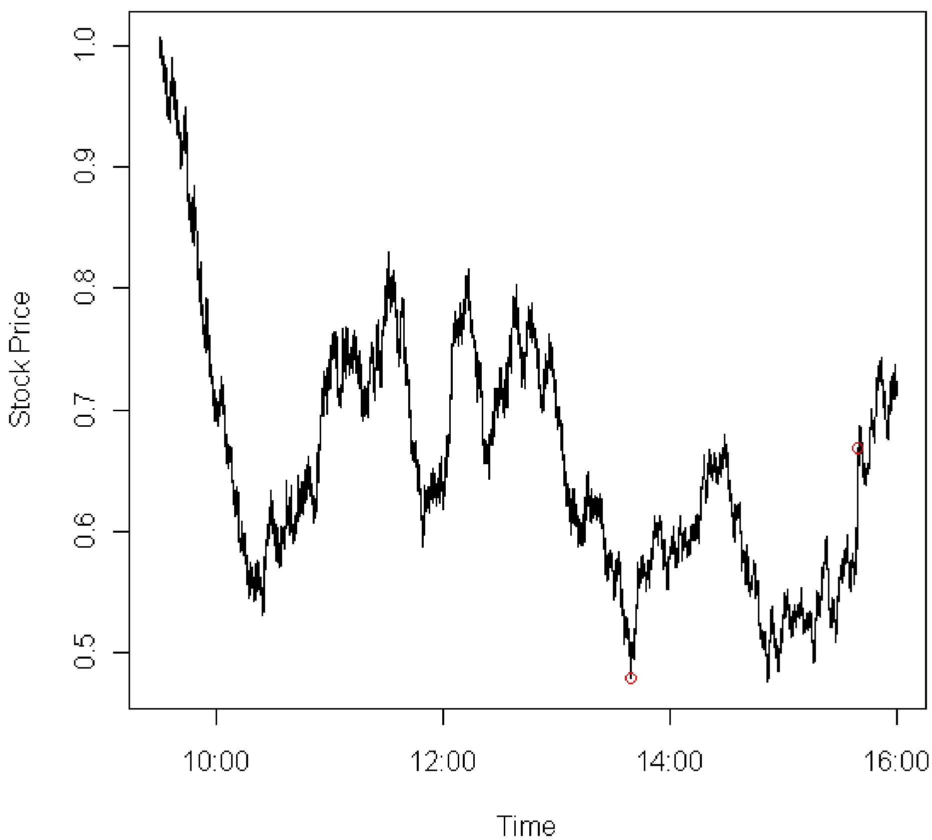

Figure 2.6: One Path with Jumps from Our Simulations

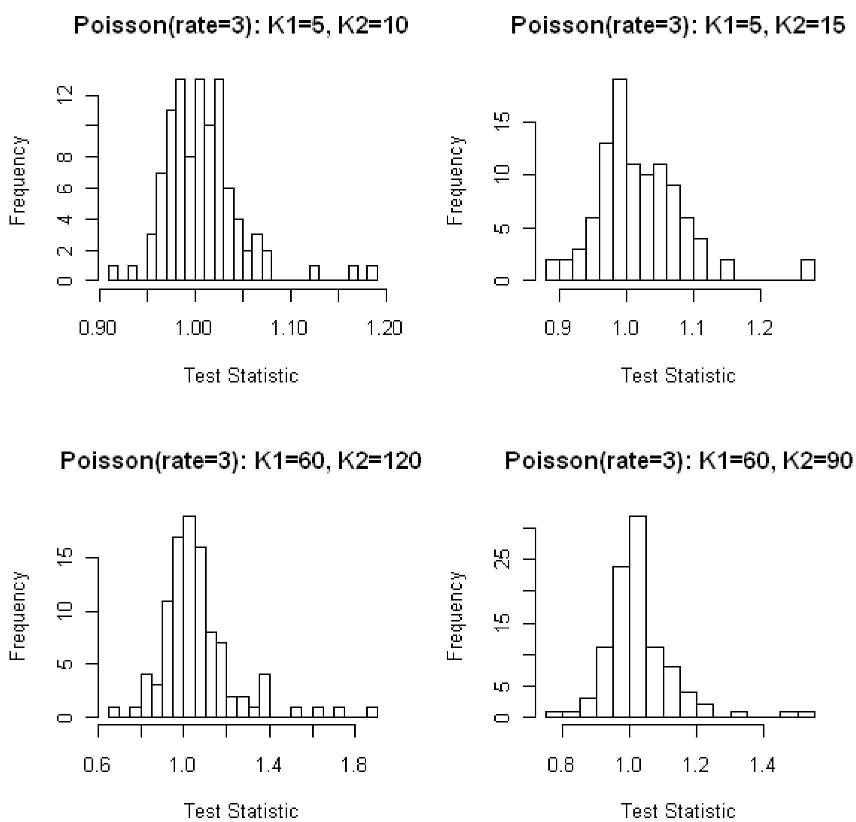

Figure 2.7: Monte Carlo asymptotic distribution of our new test statistics for paths with jumps

Table 2.1: Monte Carlo Mean and Standard Deviation for continuous paths

Frequency	Method	$\frac{K_2}{K_1} = 2$	$\frac{K_2}{K_1} = 3$
5 sec	AJ	2.0078(0.1472)	3.0345(0.244)
5 sec	FF	2.0028(0.1020)	3.0074(0.1962)
5 sec	RMPC	1.9976(0.0673)	2.9999(0.164)
30 sec	AJ	2.0404(0.3308)	2.9211(0.615)
30 sec	FF	1.9668(0.2388)	2.9228(0.4762)
30 sec	RMPC	1.9790(0.1757)	2.9396(0.4057)
1 min	AJ	1.947(0.4146)	2.9524(0.8095)
1 min	FF	1.9286(0.3376)	2.933(0.6317)
1 min	RMPC	1.9712(0.2291)	2.9988(0.5251)
2 min	AJ	2.1359(0.5702)	3.1174(1.0488)
2 min	FF	2.0587(0.3941)	3.1189(0.8084)
2 min	RMPC	2.0493(0.2987)	3.1114(0.745)

Table 2.2: Monte Carlo Mean and Standard Deviation for paths with jumps

Frequency		K=2	K=3
5 sec	AJ	1.0070(0.0715)	1.0151(0.1121)
5 sec	FF	1.0090(0.0604)	1.0194(0.0811)
5 sec	RMPC	1.0076(0.0407)	1.018(0.066)
30 sec	AJ	1.0163(0.1804)	1.0994(0.3473)
30 sec	FF	1.0424(0.2218)	1.0892(0.3756)
30 sec	RMPC	1.0354(0.1840)	1.0799(0.3395)
1 min	AJ	1.0937(0.3039)	1.2568(0.766)
1 min	FF	1.1241(0.4369)	1.2377(0.6124)
1 min	RMPC	1.0676(0.183)	1.168(0.3288)
2 min	AJ	1.1614(0.5306)	1.4684(0.8733)
2 min	FF	1.2497(0.563)	1.4544(0.7383)
2 min	RMPC	1.178(0.292)	1.3737(0.5236)

Table 2.3: Monte Carlo Comparisons of Old Test Statistic and New Test Statistic, $\mathbb{E}\epsilon^2 = 0.01^2$

sampling frequency				Expectation		New Test		Old Test	
$K+K_2$	K	$K+K_1$	K	$\frac{K_2}{K_1}$	Path	Mean	Std.	Mean	Std.
30	10	20	10	2	Cont	1.94	0.17	1.23	0.15
				1	Jump	1.03	0.06	1.02	0.05
20	10	15	10	2	Cont	1.99	0.27	0.98	0.13
				1	Jump	1.03	0.04	1.00	0.04
40	10	20	10	3	Cont	2.86	0.33	1.53	0.28
				1	Jump	1.05	0.11	1.02	0.10
50	10	20	10	4	Cont	3.78	0.52	1.84	0.40
				1	Jump	1.07	0.17	1.03	0.13

Table 2.4: Monte Carlo Comparisons of Old Test Statistic and New Test Statistic, $\mathbb{E}\epsilon^2 = 0.005^2$

sampling frequency				Expectation		New Test		Old Test	
$K+K_2$	K	$K+K_1$	K	$\frac{K_2}{K_1}$	Path	Mean	Std.	Mean	Std.
6	2	4	2	2	Cont	1.76	0.08	1.15	0.14
				1	Jump	1.12	0.03	1.00	0.04
20	10	15	10	2	Cont	2.01	0.14	1.49	0.12
				1	Jump	1.03	0.03	1.01	0.04
25	10	15	10	3	Cont	2.99	0.28	2.00	0.25
				1	Jump	1.05	0.05	1.01	0.07
30	10	15	10	4	Cont	3.96	0.47	2.53	0.36
				1	Jump	1.06	0.09	1.03	0.08

APPENDIX

ADDITIONAL RESULTS AND PROOFS

A.1 STABLE CONVERGENCE IN LAW

Let X_n denote a sequence a random variables defined on a probability space $(\Omega, \mathcal{F}, P)$ and taking the value in $(E, \mathcal{E})$, a complete separable metric space with Borel σ-algebra.

Definition .1. *X_n is said to* ***converge stably in law****, denoted as $X_n \xrightarrow{L_s} X$ if there exists a probability measure μ on $(\Omega \times E, \mathcal{F} \times \mathcal{E})$, such that $\mu(A \times E) = P(A)$ for all $A \in \mathcal{F}$, and for $n \to \infty$,*

$$\mathbb{E}(Yf(X_n)) = \int Y(\omega)f(x)\mu(d\omega, dx).$$

for all bounded continuous function f on E and bounded random variable Y on $(\Omega, \mathcal{F})$.

Lemma .1. *Stable convergence implies weakly convergence: If $X_n \xrightarrow{L_s} X \Longrightarrow X_n \xrightarrow{D} X$.*

Proof. Let $Y \equiv 1$ *a.s.*, from the definition and Theorem 25.8 in Billingsley, it is easy to prove. □

Why do we need this definition? From Slutsky's Theorem, we know that if $X_n \overset{D}{\to} \sigma Z$, and $S_n \overset{p}{\to} \sigma$, then $X_n/S_n \overset{D}{\to} Z$. However, in many cases, we have the conditional convergence, $X_n|\Sigma \overset{D}{\to} \Sigma Z$, but Σ is a random variable with an unknown law. But we can find a sequence of statistics Σ_n such that $(X_n, \Sigma_n) \overset{D}{\to} (X, \Sigma)$. So we want the new statistics $Z_n = X_n\Sigma_n^{-1}$ could converges in law to $N(0, 1)$. This is the why we need the stable convergence.

This concept was first introduced by Renyi (1963), for the same reasons as ours. For more details, see Jacod and Shiryaev (2003, p 512 - 518); and for an early use in econometrics, see Phillips and Ouliaris (1990).

Most important is that the analog of Slutsky's theorem for ordinary convergence in law holds for stable convergence in law:

Lemma .2. *If* $X_n \xrightarrow{L_s} X$ *and* $Y_n/Y \xrightarrow{D} 1$, *then* $X_n Y_n \xrightarrow{D} XY$.

Lemma .3. *If* $X_n \xrightarrow{L_s} X$ *and* $Y_n \xrightarrow{D} Y$, *then* $X_n + Y_n \xrightarrow{L_s} X + Y$.

A.2 PROOFS

A.2.1 Proof of Lemma 1.3

$$\begin{aligned}
\gamma_0^{(K)}(Y,Y) &= \sum_{i=K}^{n} (Y_i - Y_{i-K})^2, \\
\gamma_1^{(K)}(Y,Y) &= \sum_{i=2K}^{n} (Y_i - Y_{i-K})(Y_{i-K} - Y_{i-2K}), \\
&\vdots \\
\gamma_h^{(K)}(Y,Y) &= \sum_{i=(h+1)K}^{n} (Y_i - Y_{i-K})(Y_{i-hK} - Y_{i-(h+1)K}), \\
&\vdots
\end{aligned} \tag{.1}$$

It is easy to prove some basic results:

$$\begin{aligned}
\mathbb{E}[\gamma_0^{(K)}(Y,Y)] &= \mathbb{E}\big[\gamma_0^{(K)}(X,X) + \gamma_0^{(K)}(E,E) + 2\gamma_0^{(K)}(X,E)\big] \\
&= K\int_0^T \sigma_s^2 ds + O_p\left(\sqrt{\frac{1}{n}}\right) + (n-K+1)\mathbb{E}[(E_i - E_{i-K})^2] \\
&= K\int_0^T \sigma_s^2 ds + (n-K+1)(2V_0 - 2V_K) + O_p\left(\sqrt{\frac{1}{n}}\right), \\
&\vdots
\end{aligned} \tag{.2}$$

$$
\begin{aligned}
\mathbb{E}[\gamma_1^{(K)}(Y,Y)] &= \mathbb{E}\big[\gamma_1^{(K)}(X,X) + \gamma_1^{(K)}(E,E) + 2\gamma_1^{(K)}(X,E)\big] \\
&= (n-2K+1)\mathbb{E}[(E_i - E_{i-K})(E_{i-K} - E_{i-2K})] \\
&= (n-2K+1)(-V_0 + 2V_K - V_{2K}), \\
&\vdots \\
\mathbb{E}[\gamma_h^{(K)}(Y,Y)] &= (n-(h+1)K+1)(-V_{(h-1)K} + 2V_{hK} - V_{(h+1)K}), \\
&\vdots
\end{aligned}
\tag{.3}
$$

A.2.2 Proof of Theorem 1.3

Under assumption 1.6, as $n \to \infty$,

(1) **Signal Term:**

$$
\sqrt{\frac{n}{T}}\left\{\frac{1}{K}\begin{bmatrix} \gamma_0^{(K)}(X) \\ \gamma_1^{(K)}(X) \\ \cdots \\ \gamma_H^{(K)}(X) \end{bmatrix} - \begin{bmatrix} \int_0^T \sigma_s^2 ds \\ 0 \\ \cdots \\ 0 \end{bmatrix}\right\} \xrightarrow{L_s} N\left(0, \frac{1}{6}K(\int_0^T \sigma_s^4 ds)\Omega_X\right).
$$

where

$$
\Omega_X = \begin{pmatrix} 8 & & & & \\ 2 & 4 & & & \\ 0 & 1 & 4 & & \\ \vdots & 0 & 1 & \ddots & \\ 0 & \cdots & \cdots & 1 & 4 \end{pmatrix}.
$$

Consider a H-dimensional function $f = (f_1, \cdots, f_H):\ \mathbb{R}^H \to \mathbb{R}^H$ with $f_1(X_1, \cdots, X_H) = (X_1)^2$, $f_2(X_1, \cdots, X_H) = X_1X_2$, $f_3(X_1, \cdots, X_H) = X_1X_3$, $\cdots$, $f_H(X_1, \cdots, X_H) = X_1X_H$. Denote $\rho_\sigma^{\otimes K}(f) = \int f(x)\rho_\sigma^{\otimes K}(dx)$ with $\rho_\sigma^{\otimes K}$ the K-fold tensor product of the law

$N(0,\sigma^2)$. Define

$$\begin{aligned} V'(f,K,\Delta)_t &= \sum_{i=1}^{[t/\Delta]-H+1} f(\Delta_i X/\sqrt{\Delta},\cdots,\Delta_{i+H-1}X/\sqrt{\Delta}) \\ &\text{and} \\ V(f,K,\Delta)_t &= \sum_{i=1}^{n-K+1} C_i^T f(\Delta_i X/\sqrt{\Delta},\cdots,\Delta_{i+H-1}X/\sqrt{\Delta}). \end{aligned} \tag{.4}$$

where

$$C_i = (C_i^1,\cdots,C_i^H)^T, \text{ and } C_i^h = 1 \text{ if } 0 \le i\Delta \le \cdots \le (i+h)\Delta \le T, \ = 0 \text{ otherwise.}$$

From Theorem 7.1 of Jacod (2007), the H-dimensional processes

$$\frac{1}{\sqrt{\Delta}}\left(\Delta V'(f,K,\Delta)_t - \int_0^t \rho_{\sigma_s}^{\otimes K}(f)ds\right) \tag{.5}$$

converges stably in law to a continuous process $V'(f,K)$ defined on an extension $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{P})$ of the original space $(\Omega,\mathcal{F},P)$, which conditionally on the σ-field $\mathcal{F}$ is a centered Gaussian $\mathbb{R}^H$- valued process with independent increments, satisfying

$$\tilde{\mathbb{E}}[V'(f_i,K)V'(f_j,K)] = \int_0^t R_{\sigma_s}^{ij}(f,K)ds, \tag{.6}$$

where $R_\sigma^{ij}(f,K)$ is defined as

$$\begin{aligned} R_\sigma^{ij}(f,K) = &\sum_{d=-H+1}^{H-1} \tilde{\mathbb{E}}[f_i(\sigma Z_H,\cdots,\sigma Z_{2H-1})f_j(\sigma Z_{H+d},\cdots,\sigma Z_{2H-1+d})] \\ &- (2H-1)\tilde{\mathbb{E}}[f_i(\sigma Z_H,\cdots,\sigma Z_{2H-1})]\tilde{\mathbb{E}}[f_j(\sigma Z_{H+d},\cdots,\sigma Z_{2H-1+d})], \end{aligned} \tag{.7}$$

where (Z_i) are independent standard Gaussian random variables. By the definition of f, we can derive that

$$\begin{aligned} \rho_{\sigma_s}^{\otimes K}(f) &= \left(\mathbb{E}(\sigma_s^2 Z_1^2),\mathbb{E}(\sigma_s^2 Z_1 Z_2),\cdots,\mathbb{E}(\sigma_s^2 Z_1 Z_H)\right)^T \\ &= (\sigma_s^2,0,\cdots,0)^T, \\ \int_0^t \rho_{\sigma_s}^{\otimes K}(f)ds &= \left(\int_0^t \sigma_s^2 ds, 0,\cdots,0\right)^T, \end{aligned} \tag{.8}$$

and

$$
\begin{aligned}
&R^{11}_{\sigma}(f,K)\\
&= \frac{1}{K^2} Cov\Big((\Delta^K_i X)^2, (\Delta^K_{i-K+1}X)^2 + \ldots + (\Delta^K_{i+K-1}X)^2 \Big)\\
&= \frac{1}{K^2}\sigma^4 Var(Z^2) \left(\sum_{d=1-K}^{K-1} d^2 \right)\\
&= \frac{1}{K^2}\sigma^4 Var(Z^2) \left(\frac{(K-1)K(2K-1)}{6} + K^2 \right)\\
&= \frac{1}{K^2}\sigma^4 Var(Z^2) \int_{-K}^{K} x^2 dx + O(K^2)\\
&= \frac{4}{3} K A_4 + O(K^2),
\end{aligned} \tag{.9}
$$

$$
\begin{aligned}
&R^{hh}_{\sigma}(f,K)\\
&= \frac{1}{K^2} Cov\Big(\Delta^K_i X \Delta^K_{i-K} X \quad , \quad \Delta^K_{i-K+1} X \Delta^K_{i-2K+1} X + \ldots + \Delta^K_{i+K-1} X \Delta^K_{i-1} X \Big)\\
&= \frac{1}{K^2}\sigma^4 Var(Z_1 Z_2) \int_{-K}^{K} x^2 dx + O(K^2))\\
&= \frac{2}{3} K A_4 + O(K^2),
\end{aligned}
$$

$$
\begin{aligned}
&R^{12}_{\sigma}(f,K)\\
&= \frac{1}{K^2} Cov\Big(\Delta^K_i X \Delta^K_{i-hK} X \quad , \quad (\Delta^K_{i-K+1}X)^2 + \ldots + (\Delta^K_{i+K-1}X)^2 \Big)\\
&= \frac{1}{K^2}\sigma_i^4 Var(Z_1 Z_2) \int_{0}^{K} 2x(K-x) dx + O(K^2)\\
&= \frac{1}{3} K A_4 + O(K^2),\\
&R^{(h-1)h}_{\sigma}(f,K)\\
&= \frac{1}{K^2} Cov\Big(\Delta^K_i X \Delta^K_{i-hK} X \quad , \quad \Delta^K_{i-K+1} X \Delta^K_{i-(h+1)K+1} X + \ldots + \Delta^K_{i-1} X \Delta^K_{i-hK-1} X \Big)\\
&= \frac{1}{K^2}\sigma^4 Var(Z_1 Z_2) \int_{0}^{K} x(K-x) dx + O(K^2)\\
&= \frac{1}{6} K A_4 + O(K^2).
\end{aligned} \tag{.10}
$$

Therefore,

$$\frac{1}{\sqrt{\Delta}}\left(\frac{1}{K}V'(f,K,\Delta)_t-\int_0^t \rho_{\sigma_s}^{\otimes K}(f)ds\right)\xrightarrow{L_s} N\left(0,\frac{1}{6}K(\int_0^T\sigma_s^4 ds)\Omega_X\right). \tag{.11}$$

Finally, it is easy to prove $\Gamma_H^{(K)}(X)=V(f,K,\Delta)_t=V'(f,K,\Delta)_t+O_p(K\Delta)$, and the theorem could be proved from lemma .3.

(2) **Mixed Term:**

We prove the result for the $\gamma_0^{(K)}(X,E)$. Others could be proved similarly.

$$\gamma_0^{(K)}(X,E)=\sum_{i=K}^{n}(X_i-X_{i-K})(E_i-E_{i-K})=\sum_{i=0}^{n}(C_i-C_{i-K})E_i,$$

where $C_i=\Delta_i^K X=X_i-X_{i-K}$ if $0\le i-K<i\le n$, and $=0$ otherwise. Then we have:

$$\begin{aligned}
&\mathbb{E}\left(\gamma_0^{(K)}(X,E)|X\right)=\mathbb{E}\left(\left(\sum_{i=0}^{n}(-C_{i-K}+C_i)E_i\right)^2|X\right)\\
&\le V_0\left(\sum_{i=0}^{n}(-C_{i-K}+C_i)^2+2\sum_{d=1}^{n}\rho^d|\sum_i(-C_{i-K}+C_i)(-C_{i-K+d}+C_{i+d})|\right)\\
&\le V_0\sum_{i=0}^{n}(-C_{i-K}+C_i)^2(1+4\sum_{d=1}^{n}\rho^d)\\
&\le V_0K\frac{[X,X]^{(K)}}{K}(1+4\rho/(1-\rho)).
\end{aligned} \tag{.12}$$

The last two steps use the Cauchy-Schwarz inequality. Then using Markov's Inequality, we can prove the result.

A.2.3 Proof of Theorem 2.4

Write $U_n=(\Delta_n)^{-1/2}\left(\frac{\widehat{RMPC}_n(X,1,K_2,p,0)}{K_2}\right)$ and $V_n=(\Delta_n)^{-1/2}\left(\frac{\widehat{RMPC}_n(X,1,K_1,p,0)}{K_1}\right)$. Then

$$\begin{aligned}
&\hat{S}(X,K_1,K_2,p)-1\\
&=\frac{\frac{1}{K_2}\widehat{RMPC}_n(X,1,K_1,p,0)}{\frac{1}{K_1}\widehat{RMPC}_n(X,1,K_1,p,0)}-1\\
&=(\Delta_n)^{1/2}\frac{U_n-V_n}{\frac{1}{K_1}\widehat{RMPC}_n(X,1,K_1,p,0)}.
\end{aligned} \tag{.13}$$

Then U_n-V_n converges stably in law to $Y_{p,K}$ as in (2.9). Theorem 2.4 follows from corollary 2.3.

A.2.4 Proof of Theorem 2.7

Write $U'_n = (\Delta_n)^{-1/2}\left(\Delta^{1-p/2}\frac{1}{K_2}\widehat{RMPC}_n(X,1,K_2,p,0) - K_2^{p/2-1}m_pA_p\right)$ and $V'_n = (\Delta_n)^{-1/2}\left(\Delta^{1-p/2}\frac{1}{K_1}\widehat{RMPC}_n(X,1,K_1,p,0) - K_1^{p/2-1}m_pA_p\right)$. Then

$$\begin{aligned}
&\hat{S}(X,K_1,K_2,p) - (\frac{K_2}{K_1})^{p/2-1} \\
&= \frac{\frac{1}{K_2}\widehat{RMPC}_n(X,1,K_1,p,0)}{\frac{1}{K_1}\widehat{RMPC}_n(X,1,K_1,p,0)} - (\frac{K_2}{K_1})^{p/2-1} \\
&= (\Delta_n)^{1/2}\frac{U'_n - \frac{K_2}{K_1}V'_n}{\frac{1}{K_1}\widehat{RMPC}_n(X,1,K_1,p,0)}.
\end{aligned} \tag{.14}$$

Then Theorem 2.7 follows from corollary 2.6.

BIBLIOGRAPHY

[1] Y. Aït-Sahalia, J. Jacod. Testing for Jumps in a Discretely Observed Process. *Annals of Statistics*, 37(1):184–222, 2009a.

[2] Y. Aït-Sahalia, J. Jacod. Estimating the degree of activity of jumps in high frequency data. *The Annals of Statistics*, 37(5A):2202–2244, 2009b.

[3] Y. Aït-Sahalia, J. Jacod. Is Brownian Motion Necessary to Model High Frequency Data? *Accepted, The Annals of Statistics*, 2009c.

[4] Y. Aït-Sahalia, R. Kimmel. Maximum likelihood estimation of stochastic volatility models. *Journal of Financial Economics*, 83(2):413–452, 2007.

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

[6] Y. Aït-Sahalia, P. Mykland, L. Zhang. Ultra high frequency volatility estimation with dependent microstructure noise. *Unpublished paper: Department of Economics, Princeton University*, 2005.

[7] Y. Aït-Sahalia, P. Mykland, L. Zhang. Ultra high frequency volatility estimation with dependent microstructure noise. *Journal of Econometrics*, 160:160–175, 2011.

[8] Y. Aït-Sahalia, J. Yu. High frequency market microstructure noise estimates and liquidity measures. *Annals of Applied Statistics*, 3(1):422–457, 2009.

[9] Y. Amihud, H. Mendelson, L. Pedersen. *Liquidity and asset prices.* Now Pub, 2006.

[10] T. Andersen, L. Benzoni, J. Lund. An empirical investigation of continuous-time equity return models. *The Journal of Finance*, 57(3):1239–1284, 2002.

[11] T. Andersen, T. Bollerslev. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39(4):885–905, 1998.

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

[13] T. Andersen, T. Bollerslev, F. Diebold, P. Labys. The Distribution of Realized Exchange Rate Volatility. *Journal of the American Statistical Association*, 96(453), 2001.

[14] T. Andersen, T. Bollerslev, F. Diebold, P. Labys. Modeling and forecasting realized volatility. *Econometrica*, 71(2):579–625, 2003.

[15] T. Andersen, T. Terasvirta. Realized volatility. *Handbook of Financial Time Series*, strony 555–575, 2009.

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

[17] F. Bandi, J. Russell. Microstructure noise, realized variance, and optimal sampling. *Review of Economic Studies*, 75(2):339–369, 2008.

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

[19] O. Barndorff-Nielsen, P. Hansen, A. Lunde, N. Shephard. Realized Kernels in Practice: Trades and Quotes. *Econometrics Journal*, 12(3):1–32, 2009.

[20] O. Barndorff-Nielsen, P. Hansen, A. Lunde, N. Shephard. Subsampling realised kernels. *Journal of Econometrics*, 160(1):204–219, 2011a.

[21] O. Barndorff-Nielsen, P. Hansen, A. Lunde, N. Shephard. Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. *Journal of Econometrics*, 2011b.

[22] O. Barndorff-Nielsen, N. Shephard. Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 63(2):167–241, 2001.

[23] O. Barndorff-Nielsen, N. Shephard. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, strony 253–280, 2002a.

[24] O. Barndorff-Nielsen, N. Shephard. Estimating quadratic variation using realized variance. *Journal of Applied Econometrics*, 17(5):457–477, 2002b.

[25] O. Barndorff-Nielsen, N. Shephard. Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica*, 72(3):885–925, 2004.

[26] O. Barndorff-Nielsen, N. Shephard. Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics*, 4(1):1, 2006.

[27] O. Barndorff-Nielsen, N. Shephard, M. Winkel. Limit theorems for multipower variation in the presence of jumps. *Stochastic processes and their applications*, 116(5):796–806, 2006.

[28] D. Bates. Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche Mark options. *The Review of Financial Studies*, 9(1):69–107, 1996.

[29] P. Carr, L. Wu. What type of process underlies options? A simple robust test. *Journal of Finance*, 58(6):2581–2610, 2003.

[30] G. Casella, R. Berger. *Statistical inference*. Duxbury Pr, 2002.

[31] M. Chernov, A. Ronald Gallant, E. Ghysels, G. Tauchen. Alternative models for stock price dynamics. *Journal of Econometrics*, 116(1-2):225–257, 2003.

[32] F. Comte, E. Renault. Long memory in continuous-time stochastic volatility models. *Mathematical Finance*, 8(4):291–323, 1998.

[33] S. Delattre, J. Jacod. A central limit theorem for normalized functions of the increments of a diffusion process, in the presence of round-off errors. *Bernoulli*, 3(1):1–28, 1997.

[34] R. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, strony 987–1007, 1982.

[35] B. Eraker, M. Johannes, N. Polson. The impact of jumps in volatility and returns. *Journal of Finance*, strony 1269–1300, 2003.

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

[37] Y. Fan, J. Fan. Testing and detecting jumps based on a discretely observed process. *Manuscript*, 2010.

[38] R. Garcia, E. Ghysels, E. Renault. The econometrics of option pricing. *CIRANO Working Papers*, 2004.

[39] P. Glasserman. *Monte Carlo methods in financial engineering*, wolumen 53. Springer Verlag, 2004.

[40] P. Hansen, A. Lunde. A realized variance for the whole day based on intermittent high-frequency data. *Journal of Financial Econometrics*, 3(4):525, 2005.

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

[42] J. Jacod. Limit of random measures associated with the increments of a Brownian semimartingale. *Preprint*, 120:155–162, 1994.

[43] J. Jacod. La variation quadratique du brownien en presence d'erreurs d'arrondi. *Astérisque*, 236:155–162, 1996.

[44] J. Jacod. Statistics and high frequency data. *Lecture notes of SEMSTAT course in La Manga*, 2007.

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

[46] J. Jacod, P. Protter. Asymptotic error distributions for the Euler method for stochastic differential equations. *The Annals of Probability*, 26(1):267–307, 1998.

[47] J. Jacod, A. Shiryaev. Limit theorems for stochastic processes, volume 288 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 2003.

[48] M. Johannes, N. Polson, J. Stroud. Sequential parameter estimation in stochastic volatility models with jumps. Raport instytutowy, Working paper, Columbia GSB, 2005.

[49] I. Kalnina, O. Linton. Estimating quadratic variation consistently in the presence of endogenous and diurnal measurement error. *Journal of econometrics*, 147(1):47–59, 2008.

[50] Y. Li, P. Mykland. Are volatility estimators robust with respect to modeling assumptions? *Bernoulli*, 13(3):601, 2007.

[51] C. Mancini. Estimation of the characteristics of the jumps of a general Poisson-diffusion model. *Scandinavian Actuarial Journal*, 2004(1):42–52, 2004.

[52] R. Merton. Option pricing when underlying stock returns are discontinuous* 1. *Journal of Financial Economics*, 3(1-2):125–144, 1976.

[53] P. Mykland, L. Zhang. ANOVA for diffusions and Ito processes. *Annals of statistics*, 34(4):1931, 2006.

[54] M. O'hara. *Market microstructure theory.* Blackwell, 2007.

[55] R. Whaley. Derivatives on Market Volatility. *The journal of Derivatives*, 1(1):71–84, 1993.

[56] L. Zhang. Efficient estimation of stochastic volatility using noisy observations: a multiscale approach. *Bernoulli*, 12(6):1019–1043, 2006.

[57] L. Zhang. Efficient estimation of stochastic volatility using noisy observations: A multiscale approach. *Bernoulli*, 12(6):1019, 2006.

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