Rounding Errors and Volatility Estimation *

Yingying Li[†] Per A. Mykland [‡]

April 2012

Abstract

Financial prices are often discretized - to the nearest cent, for example. Thus we can say that prices are observed with rounding error. Rounding errors affect the estimation of volatility. Understanding them becomes important especially when we use high frequency data. In this setting, we study the asymptotic behavior of the Realized Volatility (RV), which is commonly used as an estimator of the integrated volatility. We prove the convergence of the RV and scaled RV under different conditions on the rounding level and the number of observations. A bias-corrected volatility estimator is proposed and an associated central limit theorem is shown. Simulation and empirical results show that the improvement in statistical properties can be substantial.

Keywords: Rounding Errors, Bias-correction, Diffusion Process, Market Microstructure, Realized Volatility

*Financial support from the HKSAR RGC grants SBI09/10.BM17 and GRF-602710, and the National Science Foundation under grants DMS 06-04758 and SES 06-31605 and SES 11-24526 is gratefully acknowledged. We thank Professor Michael J. Wichura for very helpful comments related to the Lemma 4 of this paper; and we thank the Editor, an anonymous Associate Editor and two anonymous referees for their helpful comments.

[†]Department of Information Systems, Business Statistics and Operations Management, Hong Kong University of Science and Technology. yyli@ust.hk

[‡]Department of Statistics at the University of Chicago, mykland@galton.uchicago.edu

1 Introduction

In recent years, high frequency data analysis has received wide attention. One central question that people have been interested in, is the estimation of volatility. The main difficulty in estimating daily volatilities using high frequency data is the presence of market microstructure noise. Substantial developments have been seen in this area. Volatility estimators with nice convergence properties have been proposed and studied in, for example, Zhang, Mykland, and Aït-Sahalia (2005), Zhang (2006), Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008), and Xiu (2010), under the assumption that the microstructure noise is additive and i.i.d. The case when the market microstructure noise can be a combination of additive noise and rounding error has been studied in Li and Mykland (2007) and Jacod, Li, Mykland, Podolskij, and Vetter (2009). Rosenbaum (2009) proposed a novel volatility estimation approach using absolute values of the increments when rounding is the only source of the market microstructure noise.

Rounding is one important source of the market microstructure noise that should not be ignored. Stocks are traded on the grids and so the observations are effectively rounded. For some cases, especially when the stock prices are low, rounding could be the main source of the market microstructure noise. The following figure plots the second-by-second stock prices of Citigroup Inc on 01 May 2007. From which we see that the log prices of the stock don't follow a diffusion process nor a diffusion process with additive noise. Rather, they look more like diffusion rounded to the nearest tick on a grid.

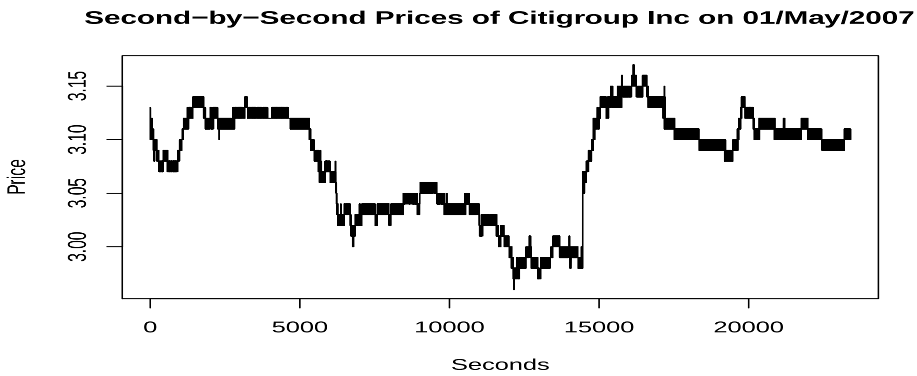

Figure 1: Second-by-second stock prices of Citigroup Inc on 01 May 2007. Rounding appears as a main feature of the data.

In this paper, we shall focus on the extreme case when there's pure rounding. We aim to study what happens to the popularly used volatility estimator, the realized volatility (RV), and how this estimator can be bias-corrected to obtain consistent volatility estimates. RV goes back to the path breaking work of Andersen and Bollerslev (1997), Andersen, Bollerslev, Diebold, and Labys (2001, 2003), Barndorff-Nielsen and Shephard (2002), Jacod and Protter (1998), and other work by the same and other authors.

We consider a security price process S, whose logarithm $X = \log S$ follows

$$dX_t = \mu_t dt + \sigma_t dW_t. \tag{1.1}$$

In other words, S is the solution to the stochastic differential equation

$$dS_t = (\mu_t + \frac{1}{2}\sigma_t^2)S_t dt + \sigma_t S_t dW_t, \quad t \in [0, 1] \tag{1.2}$$

where W_t is a standard Brownian motion. We assume that μ_t and σ_t are continuous random processes satisfying regularity conditions specified in Section 2.

It is a common practice in finance to use the sum of frequently sampled squared returns, the RV, to estimate the integrated volatility $\int_0^1 \sigma_t^2 dt$. However, empirical studies in finance have shown evidence that market microstructure noise makes this estimator upwardly biased when prices are sampled at high frequencies, while sampling sparsely gives more reasonable estimates (see, for example, the signature plots introduced by Andersen, Bollerslev, Diebold, and Labys (2000)). We investigate the case when the contamination due to market microstructure is solely due to round-off errors.

More specifically, let α_n be a sequence of positive numbers which represents the accuracy of measurement when one observes the process n times during the time period [0,1]. Suppose at time i/n ($i = 0, \cdots n$), one observes the value $k\alpha_n$ when the true value $S_{i/n}$ is in $[k\alpha_n, (k+1)\alpha_n)$ with $k \in \mathbb{Z}$. For every real s we denote by $s^{(\alpha_n)} = \alpha_n \lfloor s/\alpha_n \rfloor$ its rounded-off value at level α_n. Taking the Citigroup data as in Figure 1 for example, the rounding level is $\alpha_n = 0.01$. For that particular day, the k ranges from 296 to 317. We investigate the asymptotic behavior of the RV

$$V^n = \sum_{i=1}^{n} (\log(S_{i/n}^{(\alpha_n)}) - \log(S_{(i-1)/n}^{(\alpha_n)}))^2. \tag{1.3}$$

Jacod (1996) and Delattre and Jacod (1997) have previously studied the problem of inference for volatility based on a rounded Itô diffusion. This work is the inspiration for our work, and we seek in this paper to spell out what ensues when rounding happens on the original (dollar, euro, etc) scale and not on the log scale. As we shall see later in this paper, this more realistic type of rounding leads to a bias which requires a somewhat more complicated correction. For example, in the simple case when the volatility is constant, the bias is no longer a function of the volatility (as in section 4 of Delattre and Jacod (1997)). [1]

We shall provide the limit of V^n. This will show how RV can be problematic when the rounding errors are present; and explains why "sampling sparsely" could be a practically helpful way to estimate the volatility (however, "sampling sparsely" doesn't solve all the problems). We then propose a bias-corrected estimator, and prove an associated central limit theorem. Simulation results show that our bias-corrected

[1] We emphasize that our derivation builds on the general results of Delattre and Jacod (1997).

estimator can give substantial improvement in statistical accuracy. Empirical studies show that the bias-correction can be helpful in financial risk management. Our main bias correction applies to the case of "small rounding", as in Delattre and Jacod (1997) and Rosenbaum (2009). This kind of asymptotics is quite realistic in practice, cf. the findings for additive error in Zhang, Mykland, and Aït-Sahalia (2011). Small rounding asymptotics has also been studied in Kolassa and McCullagh (1990), where it is shown to be related to additive error. We also discuss what happens to the RV when the rounding is not "small".

These main theoretical results are presented in Section 2. Simulation Studies are presented in Section 3, and empirical studies in Section 4. Section 5 concludes. The proofs are given in the Appendix.

2 THE MAIN RESULTS

We assume that the latent security price process S_t follows (1.2), where σ_t is a non-random function of S_t, of class $\mathcal{C}^5$ on $[0,\infty)$ (In the Black-Scholes model, $\sigma_t \equiv \sigma$ is a constant). Assume further that μ_t is a continuous random process (in particular, it is locally bounded).

Let $\beta_n = \alpha_n \sqrt{n}$.

Theorem 1. *Under the condition that $\alpha_n \to 0$ as $n \to \infty$ in such a way that $\beta_n \to \beta \in [0,\infty)$, we have*

$$V^n \to_P \int_0^1 \sigma_t^2 dt + \frac{\beta^2}{6}\int_0^1 \frac{1}{S_t^2} dt - \frac{\beta^2}{\pi^2}\int_0^1 \frac{1}{S_t^2}\sum_{k=1}^{\infty}\frac{1}{k^2}\exp\left(-2\pi^2 k^2 \frac{\sigma_t^2 S_t^2}{\beta^2}\right) dt.$$

One sees from this result that the bias is increasing in β, and is always positive when $\beta \neq 0$. It blows up as β grows. Also, the bias is smaller when the stock price $S_t,\ t \in [0,1]$ is larger. Figure 2 gives a visual illustration of this. Our formula is consistent with the empirical evidence that

1) subsampling helps (same α and smaller n means smaller β and therefore smaller

bias); and
2) the rounding effect is less serious for more expensive stocks (the bias is smaller for higher values of S_t).

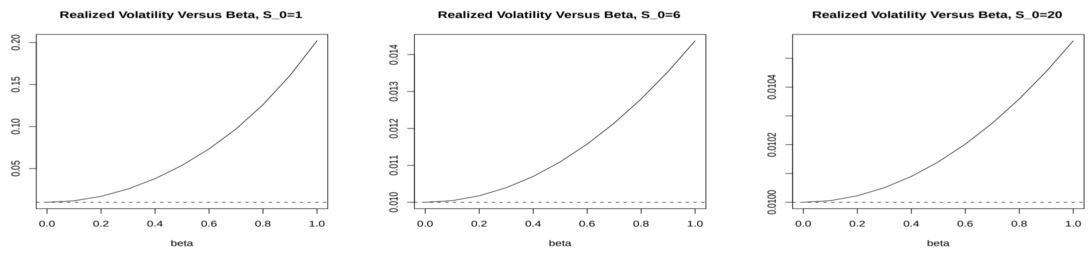

Figure 2: Realized Volatility V^n versus β based on Theorem 1 and three simulated sample paths with starting price $S_0 = 1$, $S_0 = 6$ and $S_0 = 20$ respectively. The dashed line is the true integrated volatility which is set to be 0.01; the solid curves are the limits of the realized volatility. The fact that the bias is increasing in β is illustrated by the shape of the curves, and that the bias is smaller when $S_t,\ t \in [0,1]$ is larger can be seen by comparing the ranges of the y axis of the three plots.

Theorem 1 shows that when $\beta_n \to 0$, one can have the consistency of V^n. The following theorem tells us about the convergence rate.

Theorem 2. *Under the condition that $\beta_n = O(n^{-\gamma})$ for some $\gamma > 0$, we have*

$$\sqrt{n}\left(V^n - \int_0^1 \sigma_t^2 dt - \frac{\beta_n^2}{6}\int_0^1 \frac{1}{S_t^2} dt\right) \to_{\mathcal{L}-stably} \int_0^1 \sqrt{2}\sigma_t^2 dB_t,$$

where B is a standard Brownian motion independent of W.

In this case, there is still a finite sample bias of size $\frac{\beta_n^2}{6}\int_0^1 \frac{1}{S_t^2}dt$. One can estimate the bias and find a bias-corrected estimator as the following.

Theorem 3. *Assume that $\beta_n = O(n^{-\gamma})$ for some $\gamma > 0$, and let*

$$V_0^n := V^n - \frac{\alpha_n^2}{6}\sum_{i=1}^{n} \frac{1}{(S_{i/n}^{(\alpha_n)})^2}.$$

Then as $n \to \infty$,

$$\sqrt{n}\left(V_0^n - \int_0^1 \sigma_t^2 dt\right) \to_{\mathcal{L}-stably} \int_0^1 \sqrt{2}\sigma_t^2 dB_t,$$

where B is a standard Brownian motion independent of W.

One can see from the simulation results in the next section that this bias-correction can lead to substantially improved estimates. The empirical studies in the later section further show that the bias correction can be quite helpful in risk analysis.

Remark 1. The condition of small rounding ($\alpha_n \to 0$) is necessary for the asymptotic results above. In practice, we make use of these asymptotic results via expansion – we observe only one α_n and one n for a particular price process over the time interval under consideration. If small rounding is relevant, we can make a correction as in Theorem 3, and have a better estimator (refer to the simulation studies for further illustration of the use of these results).

Remark 2. The condition that the random process σ_t is a non-random function of S_t is assumed to be able to use the framework of Delattre and Jacod (1997). In Sections 3 and 4, we see in simulation and empirical studies that even when the condition is not necessarily satisfied, the bias correction in V_0^n can still be very helpful. We conjecture that similar results hold also in stochastic volatility settings.[2]

When the small rounding condition is not satisfied, the realized volatility would blow up as the sampling frequency becomes larger. We present in the following the asymptotic result for the simple case when $\sigma_t \equiv \sigma$ to illustrate this. In this case, simple bias correction won't suffice. A correction after subsampling would help.

Theorem 4. *Let the accuracy of measurement $\alpha_n \equiv \alpha$ be independent on the number of observations n. Consider the case when $\sigma_t \equiv \sigma$ for $t \in [0, 1]$. Redefine $S_{i/n}^{(\alpha)} = \alpha$ if*

[2] A formal extension to this more general case can use a simple parametric approximation to the process, perhaps via the contiguity arguments in Mykland and Zhang (2011).

$S^{(\alpha)}_{i/n} = 0$. *As* $n \to \infty$,

$$\frac{1}{\sqrt{n}} V^n \to_P \frac{1}{\sigma} \sqrt{\frac{2}{\pi}} \sum_{k=1}^{\infty} L_1^{\log((k+1)\alpha)} (\log \frac{k+1}{k})^2,$$

where L^a_t *is the local time at the level* a *of the continuous semimartingale* $X_t = \log S_t$ *(see Revuz and Yor (1999), page 222).*

Note that by redefining $S^{(\alpha)}_{i/n} = \alpha$ if $S^{(\alpha)}_{i/n} = 0$, we rule out the possibility of taking the logarithm of zero when calculating the Realized Volatility. In practice, this simply means that the security price doesn't go below the smallest rounding grid (1 cent if $\alpha = 0.01$) during the time interval that we consider.

3 The Simulation Studies

3.1 Constant Volatility

Consider first the simplest case when $\sigma_t \equiv \sigma$ for $t \in [0, 1]$. Denote by V^n_CI and $V^n_0\_CI$ the nominal 95% confidence interval (CI) based on V^n and V^n_0 respectively, as follows.

The naive CI based on V^n relies on the classical theory of the RV, which says that when there is no observation error,

$$\sqrt{n}[V^n - \sigma^2] \to_{\mathcal{L}} N(0, 2\sigma^4).$$

The resulting nominal 95% CI is

$$V^n\_CI = \left[V^n - 1.96 * \sqrt{2(V^n)^2/n},\ V^n + 1.96 * \sqrt{2(V^n)^2/n} \right].$$

Our findings above indicate that the RV is no longer reliable when the rounding errors are present. And we have found a simple bias-corrected estimator which should work when $\alpha_n\sqrt{n}$ is reasonably small

$$V_0^n = V^n - \frac{\alpha_n^2}{6}\sum_{i=1}^{n}\frac{1}{(S_{i/n}^{(\alpha_n)})^2}.$$

By Theorem 3,

$$\sqrt{n}[V_0^n - \sigma^2] \to_{\mathcal{L}} N(0, 2\sigma^4).$$

Our adjusted nominal 95% CI is then

$$V_0^n\_CI = \left[V_0^n - 1.96 * \sqrt{2(V_0^n)^2/n},\ V_0^n + 1.96 * \sqrt{2(V_0^n)^2/n}\right].$$

To examine the performance of these volatility estimators V^n and V_0^n, we perform the following simulation study.

We simulate sample paths from (1.2) with $\mu = 0$, $\sigma = 0.1$. We run 10000 simulations of a single one-day period. The starting price of the day is taken to be $S_0 = 6$. For rounding, we use a fixed rounding level $\alpha_n \equiv \alpha = 0.01$, to be consistent with the real world where the stock prices are often rounded to the nearest cent.

Table 1 shows the simulation results. The first column of the table gives the sample size; the second column gives the corresponding sample frequencies; and the third column gives the pre-limiting β_n (so we see how our small rounding asymptotic theory works for the case of finite sample size and fixed rounding level). The last two columns contain three items each. The notation "f" stands for "actual coverage frequency", which records the frequency with which the true parameter is covered by the confidence intervals based on the corresponding volatility estimators V^n and V_0^n; "l" stands for "average length of the confidence interval", which tells how wide the confidence intervals are; "b" stands for "finite sample bias", which shows how much and to which direction the estimators are biased.

Comparing V^n to V_0^n, we see that when the sample frequency is relatively low (5 min - 1 min; see 2nd $\sim$ 3rd row), both V^n and V_0^n perform well in the sense that their nominal 95% confidence intervals are doing their jobs – these actual coverage frequencies are about 95%. This is consistent with the empirical evidence that subsampling

samp. size	samp. freq.	‘β’ $\alpha\sqrt{n}$		V^n_CI	V_0^n_CI
			f:	93.12%	92.82%
78	5 min	0.088	l:	$6.22 * 10^{-3}$	$6.19 * 10^{-3}$
			b:	$-9.44 * 10^{-5}$	$-1.31 * 10^{-4}$
			f:	95.33%	94.49%
390	1 min	0.197	l:	$3.65 * 10^{-3}$	$3.61 * 10^{-3}$
			b:	$6.97 * 10^{-5}$	$-3.98 * 10^{-5}$
			f:	78.05%	93.34%
1170	20 sec	0.342	l:	$1.71 * 10^{-3}$	$1.62 * 10^{-3}$
			b:	$5.39 * 10^{-4}$	$-9.03 * 10^{-6}$
			f:	7.83%	92.52%
2340	10 sec	0.484	l:	$1.27 * 10^{-3}$	$1.15 * 10^{-3}$
			b:	$1.09 * 10^{-3}$	$-3.79 * 10^{-6}$
			f:	0	89.07%
4680	5 sec	0.684	l:	$9.88 * 10^{-4}$	$8.10 * 10^{-4}$
			b:	$2.19 * 10^{-3}$	$-2.28 * 10^{-6}$
			f:	0	25.05%
23400	1 sec	1.53	l:	$7.46 * 10^{-4}$	$3.49 * 10^{-4}$
			b:	$1.06 * 10^{-2}$	$-3.61 * 10^{-4}$

Table 1: Performance of the nominal 95% confidence intervals based on V^n and V_0^n. “f”: actual coverage frequency of the confidence intervals; “l”: average length of the confidence intervals; “b”: finite sample bias.

helps. But since the convergence rate is square root of n, the confidence intervals are wide when the n is small. Going down to the 4th $\sim$ 6th row, we see that when the sample frequency goes a bit higher (20 sec - 5 sec), the problems with the Realized Volatility show up, the coverage frequency goes down from about 95% to 0; while the V_0^n_CI still perform quite well. Also from the biases we see that the Realized Volatility goes to something much larger than the true value, while the V_0^n stays close to the true parameter. Hence, overall, V_0^n does a better job than the uncorrected Realized Volatility V^n.

Note that for ultra high-frequency (1 sec, 7th row), the bias-corrected volatility estimator doesn’t perform as well either. This is as expected, since the bias-corrected estimator is built upon the asymptotic theory that requires the condition $\alpha_n\sqrt{n} \to 0$, which is never true in practice. For a fixed rounding level, if the sample frequency goes higher and higher, our bias correction would eventually fail. The failure at really high frequency would probably happen to all other RV-based volatility estimators, too, as a direct consequence of Theorem 4 (see Theorem 2 in Li and Mykland (2007) for a result for the Two scales Realized Volatility as an example of another RV-based volatility estimator). The above simulation suggests that for the given price level and the rounding level, when the sample frequency is lower than 5 seconds, our bias correction can be very helpful.

3.2 Stochastic Volatility

The theoretical results are established under the conditions specified in Section 2. One may wonder how the bias correction can perform if the conditions are not met. In the following, we conduct a simulation experiment. Based on a stochastic model in which the volatility process evolves by itself and is not a function of the price process. The model we adopt is the Heston Model (Heston (1993)) for the log price:

$$\begin{aligned} dX_t &= (\mu - \nu_t/2)dt + \sigma_t dB_t \\ d\nu_t &= \kappa(\eta - \nu_t)dt + \gamma \nu_t^{1/2} dW_t \end{aligned}$$

where $\nu_t = \sigma_t^2$, B and W are two standard Brownian motions with $E(dB_t dW_t) = \rho dt$, and the parameters μ, η, κ, γ and ρ are constants which are chosen to be 0.05/252, 0.1/252, 5/252, 0.5/252, -0.5 respectively in the simulation. Aït-Sahalia and Kimmel (2007) and Aït-Sahalia, Fan, and Li (2013) are taken as reference when choosing these parameter values. We use a moderate leverage effect parameter $\rho = -0.5$ to represent an individual stock. We simulate 10000 days and obtained pairs of the latent observations X_{t_i}, σ_{t_i} for $t_0 = 0, t_1 = \frac{1}{390}, \cdots, t_n = 1$ for each day (one observation per minute, $n = 390$). We compute the integrated volatility $V = n^{-1} \sum_{i=1}^{n} \sigma_{t_i}^2$ and use this as the reference measure. The observed log prices are $\log(\exp(X_{t_i})^{(\alpha)})$ with $\alpha = 0.01$ (rounded to cents). We compute the realized volatility V^n and our bias-corrected estimator V_0^n and compare their performance which is summarized in the following table. We see from the summarized results that even though under this model the

	1st Quartile	Median	3rd Quartile	Mean	Root Mean Squares
$V^n - V$	$1.28 * 10^{-4}$	$2.05 * 10^{-4}$	$3.15 * 10^{-4}$	$2.47 * 10^{-4}$	$3.02 * 10^{-4}$
$V_0^n - V$	$-4.36 * 10^{-5}$	$-1.02 * 10^{-6}$	$1.74 * 10^{-5}$	$-1.85 * 10^{-5}$	$6.98 * 10^{-5}$

Table 2: Performance of V^n and V_0^n. The estimation errors are summarized by their 1st quartile, median, 3rd quartile, mean and root mean squares.

conditions for the theoretical results are not met, the estimator V_0^n still shows clear advantage.

4 Empirical Study

To further compare the performance of V^n and V_0^n, we conduct the following analysis based on real financial data. The data we analyze are for stocks Citigroup Inc. (NYSE:C), CBS Corporation (NYSE:CBS), Dell Inc. (NYSE:DELL), Host Hotels and Resorts Inc. (NYSE:HST) and KeyCorp (NYSE:KEY) over the year of 2009. We collected the stock prices every twenty seconds (1170 observations per day), and compute V^n and V_0^n for each day. Based on the estimated volatilities, and a simple assumption that the return on each day is normally distributed with approximately zero mean and variance as estimated (as commonly assumed in risk management), we compute the 5% Value at Risk (VaR) for each day, and count the total number of days that the VaR is violated. For the 252 days considered, the following table summarizes the rate of VaR violation.

	C	CBS	DELL	HST	KEY
V^n	0.0040	0.0238	0.0159	0.0238	0.0278
V_0^n	0.0119	0.0317	0.0198	0.0317	0.0357

Table 3: 5% VaR Violation Rate. Based on 20-second stock prices of C, CBS, DELL, HST, KEY for the year of 2009.

Since we are considering the 5% VaRs, the expected rate of violation is 5%. We see that for the stocks tested, the violation rate based on V_0^n are all closer to the expected rate than the ones based on V^n. The estimate V^n tends to be over-cautious, which dramatically overestimates the daily volatilities.

5 Conclusions and Discussion

In summary, we have proved the following results:

Under the condition that $\sqrt{n}\alpha_n \to \beta \in [0, \infty)$,

$$V^n \to_P \int_0^1 \sigma_t^2 dt + \frac{\beta^2}{6}\int_0^1 \frac{1}{S_t^2}dt - \frac{\beta^2}{\pi^2}\int_0^1 \frac{1}{S_t^2}\sum_{k=1}^{\infty}\frac{1}{k^2}\exp\left\{-2\pi^2 k^2 \frac{\sigma_t^2 S_t^2}{\beta^2}\right\} dt.$$

And under the condition that $\beta_n = O(n^{-\gamma})$ for certain $\gamma > 0$, we have

$$\sqrt{n}\left(V^n - \int_0^1 \sigma_t^2 dt - \frac{\beta_n^2}{6}\int_0^1 \frac{1}{S_t^2}dt\right) \to_{\mathcal{L}-stably} \int_0^1 \sqrt{2}\sigma_t^2 dB_t,$$

and

$$\sqrt{n}\left(V^n - \frac{\alpha_n^2}{6}\sum_{i=1}^{n}\frac{1}{(S_{i/n}^{(\alpha_n)})^2} - \int_0^1 \sigma_t^2 dt\right) \to_{\mathcal{L}-stably} \int_0^1 \sqrt{2}\sigma_t^2 dB_t,$$

where B is a Brownian motion independent with the driving Brownian motion of the log price process.

We have used the later result to create a bias-correction that works for "small rounding" by defining the bias-corrected estimator to be

$$V_0^n = V^n - \frac{\alpha_n^2}{6}\sum_{i=1}^{n}\frac{1}{(S_{i/n}^{(\alpha_n)})^2}.$$

When $\alpha_n = \alpha,\ \beta_n = \alpha_n\sqrt{n} \to \infty$, V^n blows up to infinity at a rate being square root of the sample size n. We have the following result for the case when $\sigma_t = \sigma$.

$$\frac{1}{\sqrt{n}}V^n \to_P \frac{1}{\sigma}\sqrt{\frac{2}{\pi}}\sum_{k=1}^{\infty} L_1^{\log((k+1)\alpha)}(\log\frac{k+1}{k})^2,$$

where L_t^a is the local time of the continuous semimartingale $X_t = \log S_t$.

The effectiveness and practical helpfulness of the bias correction in V_0^n is shown by both simulation and empirical studies.

Note that while we work with observations on a time interval $[0, 1]$, results for the more general case of time interval $[0, T]$ is obtained by rescaling. The case of unequal observation times can be studied using the methods of Jacod and Protter (1998) and Mykland and Zhang (2006).

Appendix

A.1 Preparation

We assume without loss of generality (see section A.4 for further justification) that $\mu_t = 0$, in which case

$$d \log S_t = \sigma_t dW_t; \tag{A.1}$$

and that there exist nonrandom constants $L_\sigma,\ U_\sigma \in (0, \infty)$, such that

$$L_\sigma \leq \sigma_t \leq U_\sigma \quad \text{for } t \in [0, 1].$$

More Notation:

$$A_m := \{\omega \in \Omega : S_t(\omega)_{t\in[0,1]} \in [\frac{1}{m}, m]\};$$

$$B_n := \{\omega \in \Omega : \max_{1\leq i\leq n} \sqrt{n}\left|\frac{S_{i/n}}{S_{(i-1)/n}} - 1\right| \leq 2 \log n\};$$

$$Y_{i,n} := \sqrt{n}(S^{(\alpha_n)}_{i/n} - S^{(\alpha_n)}_{(i-1)/n});$$

$$U(n, \phi) := \frac{1}{n}\sum_{i=1}^{n} \phi(S^{(\alpha_n)}_{(i-1)/n}, Y_{i,n}) \ \text{ for } \phi : \mathbb{R}^2 \to \mathbb{R}; \tag{A.2}$$

$$h(\cdot) : \ \text{density of the standard normal law ;}$$

$$h_s(\cdot) : \ \text{density of the normal law } N(0, s^2).$$

Lemma 1. *$P(B_n) \to 1$ as $n \to \infty$.*

Proof: By (A.1),

$$S_{i/n}/S_{(i-1)/n} = \exp\left(\int_{(i-1)/n}^{i/n} \sigma_s dW_s\right).$$

Note that for any $i = 1, 2, \cdots n$,

$$
\begin{aligned}
&E(\exp\left(\sqrt{n}\int_{(i-1)/n}^{i/n}\sigma_s dW_s\right))\\
\leq &E(\exp\left(\sqrt{n}\int_{(i-1)/n}^{i/n}\sigma_s dW_s - \frac{1}{2}n\int_{(i-1)/n}^{i/n}\sigma_s^2 ds + \frac{1}{2}U_\sigma^2\right))\\
= &\exp\left(\frac{1}{2}U_\sigma^2\right).
\end{aligned}
$$

Hence for any $a > 0$

$$
\begin{aligned}
&P(\int_{(i-1)/n}^{i/n}\sigma_s dW_s > a)\\
= &P(\exp\left(\sqrt{n}\int_{(i-1)/n}^{i/n}\sigma_s dW_s\right) > \exp\left(\sqrt{n}a\right))\\
\leq &\frac{\exp\left(\frac{1}{2}U_\sigma^2\right)}{\exp\left(\sqrt{n}a\right)}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&P\left(\max_{1\leq i\leq n}(\sqrt{n}(\frac{S_{i/n}}{S_{(i-1)/n}} - 1)) > 2\log n\right)\\
= &P\left(\max_{1\leq i\leq n}(\frac{S_{i/n}}{S_{(i-1)/n}}) > \frac{2\log n}{\sqrt{n}} + 1\right)\\
= &P\left(\max_{1\leq i\leq n}(\int_{(i-1)/n}^{i/n}\sigma_s dW_s) > \log(\frac{2\log n}{\sqrt{n}} + 1)\right)\\
\leq &n\frac{\exp\left(\frac{1}{2}U_\sigma^2\right)}{\exp\left(\sqrt{n}(\log(\frac{2\log n}{\sqrt{n}} + 1))\right)}\\
\to &0 \text{ as } n \to \infty.
\end{aligned}
$$

A parallel argument gives the conclusion that

$$
P\left(\max_{1\leq i\leq n}(\sqrt{n}(1 - \frac{S_{i/n}}{S_{(i-1)/n}})) > 2\log n\right) \to 0 \text{ as } n \to \infty,
$$

hence the conclusion.

Lemma 2. *If $\sqrt{n}\alpha_n \to \beta \in [0,\infty)$, then for any m, there exist N large and $c_m \in (0, \frac{1}{m}]$ such that for all $n \geq N$, $i = 0, 1, 2, \cdots, n$,*

$$S_{i/n}^{(\alpha_n)} \geq c_m \text{ on } A_m.$$

Proof:

$$\forall i = 0, 1, 2, \cdots, n, \quad S_{i/n}^{(\alpha_n)} \geq S_{i/n} - \alpha_n;$$

and

$$S_{i/n} \geq \frac{1}{m} \text{ on } A_m, \text{ and } \alpha_n \to 0 \text{ as } n \to \infty,$$

hence the conclusion.

Lemma 3. *Suppose that $\beta_n = \sqrt{n}\alpha_n \to \beta \in [0,\infty)$, then for any fixed $m > 0$,*

$$\sup_{\omega \in A_m \bigcap B_n} \frac{Y_{i,n}}{\sqrt{n} S_{(i-1)/n}^{(\alpha_n)}} = O\left(\frac{\log n}{\sqrt{n}}\right).$$

Proof: On $A_m \bigcap B_n$,

$$|Y_{i,n}| = \sqrt{n}|S_{i/n}^{(\alpha_n)} - S_{(i-1)/n}^{(\alpha_n)}| \leq \sqrt{n}(|S_{i/n} - S_{(i-1)/n}| + 2\alpha_n) \leq 2m \log n + 2\beta_n.$$

By lemma 2, one can find a $c_m \in (0, \frac{1}{m}]$ such that for large n, on $A_m \bigcap B_n$,

$$\frac{|Y_{i,n}|}{\sqrt{n} S_{(i-1)/n}^{(\alpha_n)}} \leq \frac{2m \log n + 2\beta_n}{\sqrt{n} c_m}.$$

Since $\beta_n \to \beta < \infty$, the above inequality implies that for any fixed m, $\sup_{\omega \in A_m \bigcap B_n} \frac{Y_{i,n}}{\sqrt{n} S_{(i-1)/n}^{(\alpha_n)}}$ is $O\left(\frac{\log n}{\sqrt{n}}\right)$.

Lemma 4. *Let $\beta > 0$, then for all $\sigma, x > 0$,*

$$\int_0^1 \int h(y) \left(\frac{\beta \lfloor u + y\sigma x/\beta \rfloor}{x}\right)^2 dy du = \sigma^2 + \frac{1}{x^2}\left(\frac{\beta^2}{6} - \frac{\beta^2}{\pi^2} \sum_{k=1}^{\infty} \frac{1}{k^2} \exp\left(-2\pi^2 k^2 \frac{\sigma^2 x^2}{\beta^2}\right)\right).$$

Proof:

$$
\begin{aligned}
&\int_0^1 \int h(y)(\frac{\beta\lfloor u + y\sigma x/\beta\rfloor}{x})^2 dy du \\
=&E\left(\frac{\beta\lfloor U + Y\sigma x/\beta\rfloor}{x}\right)^2,\ U \sim unif[0,1],\ Y \sim N(0,1) \\
=&\frac{\beta^2}{x^2}E(\lfloor U + Y\sigma x/\beta\rfloor)^2 \\
=&\frac{\beta^2}{x^2}E(\lfloor U + Z\rfloor)^2,\ Z \sim N(0, \frac{\sigma^2 x^2}{\beta^2}) \\
=&\frac{\beta^2}{x^2}E\left(E(\lfloor U + Z\rfloor^2 \big| Z)\right) \\
=&\frac{\beta^2}{x^2}E\left((Z - \{Z\})^2(1 - \{Z\}) + (Z + 1 - \{Z\})^2\{Z\}\right) \\
=&\frac{\beta^2}{x^2}\left(EZ^2 + E(\{Z\}(1 - \{Z\}))\right) \\
=&\sigma^2 + \frac{1}{x^2}(\frac{\beta^2}{6} - \frac{\beta^2}{\pi^2}\sum_{k=1}^{\infty}\frac{1}{k^2}\exp\left(-2\pi^2 k^2 \frac{\sigma^2 x^2}{\beta^2}\right)),
\end{aligned}
$$

where $\{z\} = z - \lfloor z\rfloor$ is the fractional part of z.

The last equality above is proved by using the Fourier expansion for function $f(z) = \{z\} - \{z\}^2$.

A.2 Proof of Theorem 1

Recall that V^n is defined in (1.3). For large n,

$$\begin{aligned}
&V^n I_{A_m\cap B_n}\\
=&\sum_{i=1}^{n}(\log S_{i/n}^{(\alpha_n)}-\log S_{(i-1)/n}^{(\alpha_n)})^2 I_{A_m\cap B_n}\\
=&\frac{1}{n}\sum_{i=1}^{n}[\sqrt{n}\log(\frac{S_{i/n}^{(\alpha_n)}-S_{(i-1)/n}^{(\alpha_n)}}{S_{(i-1)/n}^{(\alpha_n)}}+1)]^2 I_{A_m\cap B_n}\\
=&\frac{1}{n}\sum_{i=1}^{n}[\sqrt{n}\log(\frac{Y_{i,n}}{\sqrt{n}S_{(i-1)/n}^{(\alpha_n)}}+1)]^2 I_{A_m\cap B_n}\\
=&\frac{1}{n}\sum_{i=1}^{n}[\sqrt{n}(\frac{Y_{i,n}}{\sqrt{n}S_{(i-1)/n}^{(\alpha_n)}}-\frac{1}{2}(\frac{Y_{i,n}}{\sqrt{n}S_{(i-1)/n}^{(\alpha_n)}})^2+\frac{1}{3}\theta^3)]^2 I_{A_m\cap B_n},\\
&\text{for } \theta\in(0,\frac{Y_{i,n}}{\sqrt{n}S_{(i-1)/n}^{(\alpha_n)}}).
\end{aligned} \tag{A.3}$$

By lemma 2, one can find

$$c_m\in(0,\frac{1}{m}] \text{ such that for large } n,\ S_{i/n}^{(\alpha_n)}\geq c_m \text{ for all } i=0,1,2,\cdots,n. \tag{A.4}$$

Define

$$\phi_{c_m}(x,y)=\begin{cases}(\frac{y}{x})^2, & \text{when } x\geq c_m;\\ (\frac{3}{{c_m}^4}x^2-\frac{8}{{c_m}^3}x+\frac{6}{{c_m}^2})y^2, & \text{when } x<c_m.\end{cases} \tag{A.5}$$

Note in particular that ϕ_{c_m} is a function satisfying Hypothesis L_r in Delattre and Jacod (1997) with $r=2$.

For n large enough, by Lemma 2 and Lemma 3, (A.3) can be rewritten as

$$\begin{aligned}
V^n I_{A_m\cap B_n}\leq&\frac{1}{n}\sum_{i=1}^{n}\phi_{c_m}(S_{(i-1)/n}^{(\alpha_n)},Y_{i,n})I_{A_m\cap B_n}+O\left(\frac{(\log n)^3}{n^{1/2}}\right)I_{A_m\cap B_n}\\
=&U(n,\phi_{c_m})I_{A_m\cap B_n}+O\left(\frac{(\log n)^3}{n^{1/2}}\right)I_{A_m\cap B_n},
\end{aligned}$$

where $U(\cdot,\cdot)$ is defined in (A.2).

Furthermore,

$$\begin{aligned}
V^n I_{A_m} &= V^n I_{A_m\cap B_n} + V^n I_{A_m\cap B_n^c}\\
&\le U(n,\phi_{c_m})I_{A_m} + (V^n - U(n,\phi_{c_m}))I_{A_m\cap B_n^c} + O(\frac{(\log n)^3}{n^{1/2}})I_{A_m\cap B_n}\\
&= U(n,\phi_{c_m})I_{A_m} + o_p(1) \qquad \text{(by Lemma 1).}
\end{aligned}$$

By Theorem 3.1 of Delattre and Jacod (1997),

$$U(n,\phi_{c_m}) \to_P \begin{cases} \int_0^1\int_0^1\int h(y)\phi_{c_m}(S_t, \beta\lfloor u + y\sigma_t S_t/\beta\rfloor)dydudt, \text{ if } \beta>0;\\ \int_0^1\int h(y)\phi_{c_m}(S_t, y\sigma_t S_t)dydt, \text{ if } \beta = 0.\end{cases}$$

Note that since $c_m \le 1/m$, we have

$$\phi_{c_m}(S^{(\alpha_n)}_{(i-1)/n}, Y) = \left(\frac{Y}{S^{(\alpha_n)}_{(i-1)/n}}\right)^2 I_{A_m} + \phi_{c_m}(S^{(\alpha_n)}_{(i-1)/n}, Y)I_{A_m^c}.$$

Lemma 4 gives, when $\beta > 0$,

$$U(n,\phi_{c_m})I_{A_m} \to_P \int_0^1 \frac{1}{S_t^2}(\sigma_t^2 S_t^2 + \frac{\beta^2}{6} - \frac{\beta^2}{\pi^2}\sum_{k=1}^\infty \frac{1}{k^2}\exp\left(-2\pi^2 k^2\frac{\sigma_t^2 S_t^2}{\beta^2}\right))dtI_{A_m}.$$

It is easy to check that the above convergence is also true when $\beta = 0$.

Therefore, for $\beta \in [0,\infty)$,

$$\begin{aligned}
&V^n I_{A_m}\\
=&U(n,\phi_{c_m})I_{A_m} + o_p(1)\\
\to_P&\int_0^1 \frac{1}{S_t^2}(\sigma_t^2 S_t^2 + \frac{\beta^2}{6} - \frac{\beta^2}{\pi^2}\sum_{k=1}^\infty \frac{1}{k^2}\exp\left(-2\pi^2 k^2\frac{\sigma_t^2 S_t^2}{\beta^2}\right))dtI_{A_m}.
\end{aligned}$$

That is to say, for any $\delta > 0$, $\epsilon > 0$, there exists N, such that for all $n > N$,

$$P(|V^n I_{A_m} - \int_0^1 \frac{1}{S_t^2}(\sigma_t^2 S_t^2 + \frac{\beta^2}{6} - \frac{\beta^2}{\pi^2}\sum_{k=1}^{\infty}\frac{1}{k^2}\exp\left(-2\pi^2 k^2 \frac{\sigma_t^2 S_t^2}{\beta^2}\right))dt I_{A_m}| > \delta) < \epsilon.$$

On the other hand, since $A_m \nearrow \Omega$ as $m \to \infty$, there exists M large, such that

$$P(A_M^c) < \epsilon.$$

Therefore, for $n > N$,

$$\begin{aligned}
&P(|V^n - \int_0^1 \frac{1}{S_t^2}(\sigma_t^2 S_t^2 + \frac{\beta^2}{6} - \frac{\beta^2}{\pi^2}\sum_{k=1}^{\infty}\frac{1}{k^2}\exp\left(-2\pi^2 k^2 \frac{\sigma_t^2 S_t^2}{\beta^2}\right))dt| > \delta) \\
\le& P(A_M^c) + \\
&P\left(|V^n I_{A_M} - \int_0^1 \frac{1}{S_t^2}(\sigma_t^2 S_t^2 + \frac{\beta^2}{6} - \frac{\beta^2}{\pi^2}\sum_{k=1}^{\infty}\frac{1}{k^2}\exp\left(-2\pi^2 k^2 \frac{\sigma_t^2 S_t^2}{\beta^2}\right))dt I_{A_M}| > \delta\right) \\
<& 2\epsilon.
\end{aligned}$$

This proves Theorem 1.

A.3 Proof of Theorem 2 and Theorem 3

By (A.3), for large n,

$$\begin{aligned}
&\sqrt{n} V^n I_{A_m \cap B_n} \\
=&\sqrt{n}\frac{1}{n}\sum_{i=1}^{n}[\sqrt{n}(\frac{Y_{i,n}}{\sqrt{n}S_{(i-1)/n}^{(\alpha_n)}} - \frac{1}{2}(\frac{Y_{i,n}}{\sqrt{n}S_{(i-1)/n}^{(\alpha_n)}})^2 + \frac{1}{3}\theta^3)]^2 I_{A_m \cap B_n}, \\
&\text{for } \theta \in (0, \frac{Y_{i,n}}{\sqrt{n}S_{(i-1)/n}^{(\alpha_n)}}).
\end{aligned} \tag{A.6}$$

Using the $c_m \in (0, \frac{1}{m}]$ as in (A.4), we define

$$\psi_{c_m}(x,y) = \begin{cases} (\dfrac{y}{x})^3, \text{ when } x \geq c_m; \\ (\dfrac{4}{{c_m}^3} - \dfrac{3x}{{c_m}^4})y^3, \text{ when } x < c_m. \end{cases} \tag{A.7}$$

(A.6) can be further written as

$$\begin{aligned} &\sqrt{n}V^n I_{A_m \cap B_n} \\ \leq &\sqrt{n}U(n,\phi_{c_m})I_{A_m\cap B_n} - U(n,\psi_{c_m})I_{A_m\cap B_n} + O(\frac{(\log n)^4}{n^{1/2}})I_{A_m\cap B_n}; \end{aligned}$$

and

$$\begin{aligned} &\sqrt{n}V^n I_{A_m} \\ =&\sqrt{n}V^n I_{A_m\cap B_n} + \sqrt{n}V^n I_{A_m\cap B_n^c} \\ \leq&(\sqrt{n}U(n,\phi_{c_m}) - U(n,\psi_{c_m}))I_{A_m} \\ &+ (\sqrt{n}V^n - \sqrt{n}U(n,\phi_{c_m}) + U(n,\psi_{c_m}))I_{A_m\cap B_n^c} + O(\frac{(\log n)^4}{n^{1/2}})I_{A_m\cap B_n} \\ =&\sqrt{n}U(n,\phi_{c_m})I_{A_m} - U(n,\psi_{c_m})I_{A_m} + o_p(1), \end{aligned}$$

where ϕ_{c_m} is defined in (A.5), ψ_{c_m} in (A.7) and $U(\cdot,\cdot)$ in (A.2), and we have used Lemma 3 in the above.

Note that $\psi_{c_m}(S_t, \sigma_t S_t y)$ is an odd function of y, and $\beta = 0$; by Theorem 3.1 of Delattre and Jacod (1997),

$$U(n,\psi_{c_m}) \to_P \int_0^1 \int h(y)\psi_{c_m}(S_t,\sigma_t S_t y)dydt = 0.$$

Therefore,

$$U(n,\psi_{c_m})I_{A_m} \to_P 0.$$

As a consequence,

$$\sqrt{n}V^n I_{A_m} = \sqrt{n}U(n,\phi_{c_m})I_{A_m} + o_p(1). \tag{A.8}$$

Also by Corollary 3.3 of Delattre and Jacod (1997), since $\phi_{c_m}(x,y)$ is even in y,

$$\begin{aligned}&\sqrt{n}[U(n,\phi_{c_m}) - \int_0^1 \Gamma\phi_{c_m}(S_t,\beta_n)dt]\\ \to_{\text{stably in law}} &\int_0^1 \Delta(\phi_{c_m},\phi_{c_m})(S_t,0)^{1/2}dB_s,\end{aligned} \tag{A.9}$$

where $B \perp\!\!\!\perp W$, and

$$\begin{aligned}&\Gamma\phi_{c_m}(S_t,\beta_n)\\ =&\int_0^1\int h(y)\phi_{c_m}(S_t,\beta_n\lfloor u + y\sigma_t S_t/\beta_n\rfloor)dydu\\ =&\int_0^1\int h(y)\left((\frac{\beta_n\lfloor u + y\sigma_t S_t/\beta_n\rfloor}{S_t})^2 I_{A_m} + \phi_{c_m}(S_t,\beta_n\lfloor u + y\sigma_t S_t/\beta_n\rfloor)I_{A_m^c}\right)dydu\\ =&(\sigma_t^2 + \frac{\beta_n^2}{6}\frac{1}{S_t^2} - \frac{\beta_n^2}{\pi^2}\frac{1}{S_t^2}\sum_{k=1}^{\infty}\frac{1}{k^2}\exp\left(-2\pi^2k^2\frac{\sigma_t^2S_t^2}{\beta_n^2}\right))I_{A_m}+\\ &\int_0^1\int h(y)\phi_{c_m}(S_t,\beta_n\lfloor u + y\sigma_t S_t/\beta_n\rfloor)dyduI_{A_m^c}\quad\text{(by Lemma 4)}\ ;\end{aligned} \tag{A.10}$$

and

$$\begin{aligned}&\Delta(\phi_{c_m},\phi_{c_m})(S_t,0)\\ =&\int h_{\sigma_tS_t}(y)\phi_{c_m}^2(S_t,y)dy - (\int h_{\sigma_tS_t}(y)\phi_{c_m}(S_t,y)dy)^2\\ =&\int h_{\sigma_tS_t}(y)[(\frac{y}{S_t})^4I_{A_m} + \phi_{c_m}^2(S_t,y)I_{A_m^c}]dy\\ &-(\int h_{\sigma_tS_t}(y)[(\frac{y}{S_t})^2I_{A_m} + \phi_{c_m}(S_t,y)I_{A_m^c}]dy)^2\\ =&[\int h_{\sigma_tS_t}(y)(\frac{y}{S_t})^4dy - (\int h_{\sigma_tS_t}(y)(\frac{y}{S_t})^2dy)^2]I_{A_m}+\\ &[\int h_{\sigma_tS_t}(y)\phi_{c_m}(S_t,y)^2dy - (\int h_{\sigma_tS_t}(y)\phi_{c_m}(S_t,y)dy)^2]I_{A_m^c};\end{aligned}$$

hence

$$
\begin{aligned}
&\Delta(\phi_{c_m},\phi_{c_m})(S_t,0)^{1/2}\\
=&[\int h_{\sigma_t S_t}(y)(\frac{y}{S_t})^4 dy-(\int h_{\sigma_t S_t}(y)(\frac{y}{S_t})^2 dy)^2]^{1/2} I_{A_m}+\\
&[\int h_{\sigma_t S_t}(y)\phi_{c_m}(S_t,y)^2 dy-(\int h_{\sigma_t S_t}(y)\phi_{c_m}(S_t,y)dy)^2]^{1/2} I_{A_m^c}\\
=&(2\sigma_t^4)^{1/2} I_{A_m}+[\int h_{\sigma_t S_t}(y)\phi_{c_m}(S_t,y)^2 dy-(\int h_{\sigma_t S_t}(y)\phi_{c_m}(S_t,y)dy)^2]^{1/2} I_{A_m^c}.
\end{aligned}
\tag{A.11}
$$

Plug (A.10) and (A.11) into (A.9), and note that by the assumption that $\beta_n = O(n^{-\gamma})$,

$$
\sqrt{n}\frac{\beta_n^2}{\pi^2}\int_0^1 \frac{1}{S_t^2}\sum_{k=1}^{\infty}\frac{1}{k^2}\exp\left(-2\pi^2 k^2\frac{\sigma_t^2 S_t^2}{\beta_n^2}\right)dt \to 0 \text{ a.s. on } A_m \text{ as } n\to\infty.
$$

One has,

$$
\begin{aligned}
&\sqrt{n}[U(n,\phi_{c_m})-(\int_0^1 \sigma_t^2 dt+\frac{\beta_n^2}{6}\int_0^1 \frac{1}{S_t^2}dt)]I_{A_m}\\
&+\sqrt{n}[U(n,\phi_{c_m})-\int_0^1 \Gamma\phi_{c_m}(S_t,\beta_n)dt]I_{A_m^c}\\
\to_{stably\ in\ law}&\\
&ZI_{A_m}+\int_0^1[\int h_{\sigma_t S_t}(y)\phi_{c_m}(S_t,y)^2 dy-(\int h_{\sigma_t S_t}(y)\phi_{c_m}(S_t,y)dy)^2]^{1/2} dB_s I_{A_m^c},
\end{aligned}
$$

where $Z\sim\int_0^1(2\sigma_t^4)^{1/2}dB_s,\ B \perp\!\!\!\perp W$.

For any continuous function g that vanishes outside a compact set, the above stable convergence implies that $\forall E\in\mathcal{F}$,

$$
\begin{aligned}
&E[g(\sqrt{n}[U(n,\phi_{c_m})-(\int_0^1\sigma_t^2 dt+\frac{\beta_n^2}{6}\int_0^1\frac{1}{S_t^2}dt)]I_{A_m})I_{A_m}I_E]\\
\to &E[g(\int_0^1(2\sigma_t^4)^{1/2}dB_s I_{A_m})I_{A_m}I_E].
\end{aligned}
\tag{A.12}
$$

And by defining $\eta_{c_m}(\cdot,\cdot)$ to be

$$\eta_{c_m}(x,y) = \begin{cases} (\frac{1}{x})^2, \text{ when } x \geq c_m; \\ (\frac{3}{{c_m}^4}x^2 - \frac{8}{{c_m}^3}x + \frac{6}{{c_m}^2}), \text{ when } x < c_m, \end{cases}$$

one has,

$$V_0^n I_{A_m} = V^n I_{A_m} - \frac{\beta_n^2}{6} U(n,\eta_{c_m}) I_{A_m}. \tag{A.13}$$

Again, by Theorem 3.1 of Delattre and Jacod (1997),

$$U(n,\eta_{c_m})I_{A_m} = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{(S_{i/n}^{(\alpha_n)})^2} I_{A_m} \to_P \int_0^1 \frac{1}{S_t^2} dt I_{A_m}.$$

and

$$\sqrt{n}(\frac{\beta_n^2}{6}U(n,\eta_{c_m}) - \frac{\beta_n^2}{6}\int_0^1 \frac{1}{S_t^2}dt)I_{A_m} = O_P(\beta_n^2) = o_P(1). \tag{A.14}$$

By (A.8), (A.13) and (A.14),

$$\sqrt{n}V_0^n I_{A_m} = \sqrt{n}(U(n,\phi_{c_m}) - \frac{\beta_n^2}{6}\int_0^1 \frac{1}{S_t^2}dt)I_{A_m} + o_p(1).$$

Also since that g is uniformly continuous, $\forall E \in \mathcal{F}$,

$$\begin{aligned} &\lim_{n\to\infty} E[g(\sqrt{n}(V_0^n - \int_0^1 \sigma_t^2 dt)I_{A_m})I_{A_m}I_E] \\ =&\lim_{n\to\infty} E[g(\sqrt{n}[U(n,\phi_{c_m}) - (\int_0^1 \sigma_t^2 dt + \frac{\beta_n^2}{6}\int_0^1 \frac{1}{S_t^2}dt)]I_{A_m})I_{A_m}I_E] \\ =&E[g(\int_0^1 (2\sigma_t^4)^{1/2} dB_t I_{A_m})I_{A_m}I_E] \quad (by\ (A.12)), \end{aligned}$$

which implies, for any $\epsilon > 0$, there exists N, such that $\forall n \geq N$,

$$|E[g(\sqrt{n}[V_0^n - \sigma_t^2]I_{A_M})I_{A_M}I_E] - E[g(\int_0^1 (2\sigma_t^4)^{1/2} dB_t I_{A_M})I_{A_M}I_E]| < \epsilon.$$

Note also that g is bounded, suppose $|g| \leq M_g$. Recall that $P(A_M^c) \to 0$, one can choose M such that $P(A_M^c) < \epsilon/M_g$.

So for $n \geq N$,

$$\begin{aligned}
&|E[g(\sqrt{n}[V_0^n - \int_0^1 \sigma_t^2 dt])I_E] - E[g(\int_0^1 (2\sigma_t^4)^{1/2} dB_t)I_E]| \\
\leq &|E[g(\sqrt{n}[V_0^n - \int_0^1 \sigma_t^2 dt]I_{A_M})I_{A_M}I_E] - E[g(\int_0^1 (2\sigma_t^4)^{1/2} dB_t I_{A_M})I_{A_M}I_E]| + 2M_g * P(A_M^c) \\
\leq &3\epsilon
\end{aligned}$$

Hence we've proved that for all continuous function g that vanishes outside a compact set, $\forall E \in \mathcal{F}$,

$$\lim_{n\to\infty} E[g(\sqrt{n}[V_0^n - \int_0^1 \sigma_t^2 dt])I_E] = E[g(\int_0^1 (2\sigma_t^4)^{1/2} dB_t)I_E],$$

i.e.,

$$\sqrt{n}[V_0^n - \int_0^1 \sigma_t^2 dt] \to_{\mathcal{L}-stably} \int_0^1 (2\sigma_t^4)^{1/2} dB_t.$$

This finishes the proof of Theorem 3. The proof of Theorem 2 is basically contained in the proof above.

A.4 The Case of General μ_t and σ_t

Step 1: For general cases when $\mu_t \neq 0$, if there exists $L_\sigma,\ U_\sigma,\ C_\mu \in (0, \infty)$, such that $L_\sigma \leq \sigma_t \leq U_\sigma$ and $|\mu_t| \leq C_\mu$ for $t \in [0, 1]$, the previous results all hold.

For the simplicity of notation, we consider the log scale. Let P be the probability measure corresponding to the system

$$dX_t = \sigma_t dW_t$$

and Q the probability measure corresponding to the system

$$dX_t = \mu_t dt + \sigma_t dW_t^Q,$$

where W_t and W_t^Q are standard Brownian motions under P and Q respectively.

Note that by the Girsanov Theorem (see, for example, page 164 of Øksendal (2003)), for bounded σ_t and μ_t (as stated in the conditions of "Step 1"), P and Q are mutually absolutely continuous.

The following proposition justifies the conclusion of "Step 1".

Proposition (Mykland and Zhang (2009)) *Suppose that ζ_n is a sequence of random variables which converges stably to $N(b, a^2)$ under P (meaning that $N(b, a^2) = b + aN(0, 1)$, where $N(0, 1)$ is a standard normal variable independent of $\mathcal{F}$, also a and b are $\mathcal{F}$ measurable). Then ζ_n converges stably in law to $b + aN(0, 1)$ under Q, where $N(0, 1)$ remains independent of $\mathcal{F}$ under Q.*

Step 2: for locally bounded σ_t and μ_t, the stable convergence and the convergence in probability stay valid.

This can be proved by a localization argument which uses essentially the same techniques as in the derivation in the last part of section A.3. For example, to unbound σ_t, one considers a sequence of stopping times τ_m corresponding to a sequence of positive constants σ_m which increases to infinity as $m \to \infty$: $\tau_m = min\{t : \sigma_t^2 \geq \sigma_m^2\}$, and note the fact that the sets $\{\tau_m > T\} \nearrow \Omega$.

In particular, the locally bounded assumption is automatically satisfied when σ_t and μ_t are continuous.

A.5 Proof of Theorem 4

Similar argument as the Proof of Theorem 3 in Li and Mykland (2007) gives the result.

REFERENCES

Aït-Sahalia, Y., Fan, J., and Li, Y. (2013), "The Leverage Effect Puzzle: Disentangling Sources of Bias at High Frequency," *Journal of Financial Economics*, forthcoming.

Aït-Sahalia, Y. and Kimmel, R. (2007), "Maximum Likelihood Estimation of Stochastic Volatility Models," *Journal of Financial Economics*, 83, 413–452.

Andersen, T. G. and Bollerslev, T. (1997), "Intraday Periodicity and Volatility Persistence in Financial Markets," *Journal of Empirical Finance*, 4, 115–158.

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

— (2001), "The Distribution of Exchange Rate Realized Volatility," *Journal of the American Statistical Association*, 96, 42–55.

— (2003), "Modeling and Forecasting Realized Volatility," *Econometrica*, 71, 579–625.

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

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

Delattre, S. and Jacod, J. (1997), "A Central Limit Theorem for Normalized Functions of the Increments of a Diffusion Process, in the Presence of Round-Off Errors," *Bernoulli*, 3, 1–28.

Heston, S. (1993), "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bonds and Currency Options," *Review of Financial Studies*, 6, 327–343.

Jacod, J. (1996), "La Variation Quadratique du Brownien en Présence d'Erreurs d'Arrondi," *Astérisque*, 236, 155–161.

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

Jacod, J. and Protter, P. (1998), "Asymptotic Error Distributions for the Euler Method for Stochastic Differential Equations," *Annals of Probability*, 26, 267–307.

Kolassa, J. and McCullagh, P. (1990), "Edgeworth series for lattice distributions," *Annals of Statistics*, 18, 981–985.

Li, Y. and Mykland, P. A. (2007), "Are Volatility Estimators Robust with Respect to Modeling Assumptions?" *Bernoulli*, 13, 601–622.

Mykland, P. A. and Zhang, L. (2006), "ANOVA for Diffusions," *The Annals of Statistics*, 34, # 4.

— (2009), "Inference for continuous semimartingales observed at high frequency: A general approach," *Econometrica*, 77, 1403–1445.

— (2011), "The Double Gaussian Approximation for High Frequency Data," *Scandinavian Journal of Statistics*, 38, 215–236.

Øksendal, B. (2003), *Stochastic differential equations*, Universitext, Berlin: Springer-Verlag, sixth ed., an introduction with applications.

Revuz, D. and Yor, M. (1999), *Continuous Martingales and Brownian Motion*, Berlin, Germany: Springer-Verlag, 3rd ed.

Rosenbaum, M. (2009), "Integrated Volatility and Round-off Errors," *Bernoulli*, 15, 687–720.

Xiu, D. (2010), "Quasi-Maximum Likelihood Estimation of Volatility with High Frequency Data," *Journal of Econometrics*, 159, 235–250.

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

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

— (2011), "Edgeworth Expansions for Realized Rolatility and Related Estimators," *Journal of Econometrics*, 160, 160–175.