ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 4, Issue 6, November 2015

Bayesian Inference of Realized Stochastic Volatility Model and Bias of Realized Volatility

Tetsuya Takaishi
Hiroshima University of Economics, Hiroshima, Japan

***Abstract*— We analyze time series data of a stock price return on the Tokyo Stock Exchange by the realized stochastic volatility (RSV) model, which combines the standard stochastic volatility model and the dynamics of the realized volatility. Bayesian inference is used for the parameter estimation of the RSV model, obtained by the Markov chain Monte Carlo method. Volatility updates, being the most time-consuming part of the calculation, are achieved by the hybrid Monte Carlo method. In the RSV model, bias of the realized volatility is obtained as a model parameter. We obtain the bias correction parameter to the realized volatility as a function of the sampling frequency and find that the bias correction parameter agrees with the Hansen and Lunde adjustment factor, although a slight difference is seen at high sampling frequencies where the microstructure noise dominates**

***Index Terms*—Bayesian inference, Hybrid Monte Carlo method, Microstructure noise, Realized stochastic volatility model, Realized volatility.**

I. INTRODUCTION

In financial risk management, it is crucial to measure precisely the volatility of returns, defined as a variance. Since volatility is not directly observed in the financial markets, we need to estimate it through appropriate estimation techniques such as volatility modeling. One widely used approach is the stochastic volatility (SV) model [1], [2] which allows the volatility to be a stochastic process. To determine the parameters of the SV model, one usually uses Bayesian inference performed by the Markov Chain Monte Carlo (MCMC) method. Recently, an extended SV model, called the realized stochastic volatility (RSV) model, has been proposed by Takahashi *et al*. [3], which utilizes the daily realized volatility [4] as additional information on the volatility process. The realized volatility is a model-free estimate of the integrated volatility and is calculated from high-frequency asset price data. Under ideal circumstances, the realized volatility converges with the integrated volatility. However, such ideal conditions are usually violated in the real financial markets, due to microstructure noise and non-trading hours on the financial markets. Thus, the realized volatility actually calculated from the high-frequency data is biased. Hansen and Lunde [5] introduced an adjustment factor to modify the bias of the realized volatility. By contrast, the RSV model does not require such an adjustment factor, but introduces a bias correction parameter that is determined through the parameter estimation of the RSV model through Bayesian inference.

The most time-consuming part in the Bayesian inference process of the RSV model is the MCMC method for the volatility variables. In this study, we use the hybrid Monte Carlo (HMC) method [6] for the volatility updates. Whereas the HMC method was first invented for lattice simulations in quantum chromo dynamics [7], it can also be used in other fields as an MCMC method. The HMC method has been applied for the SV model [8], [9] and the RSV model [10], and it is shown that the HMC method can de-correlate Monte Carlo samples rapidly enough. Therefore, in our analysis we also use the HMC method for the volatility updates.

In this paper, we analyze return time series of Nomura Holdings, Inc., traded on the Tokyo Stock Exchange, using the RSV model. We use realized volatility calculated at various sampling frequencies and determine the bias correction parameter of the RSV model as a function of sampling frequency. We then compare the bias correction parameter with the adjustment factor introduced by Hansen and Lunde.

II. REALIZED VOLATILITY

The realized volatility [4] is constructed as a sum of squared returns calculated from high-frequency asset price data. The returns sampled at sampling frequency Δ are given by

$$r_{t+i\Delta} = \ln P_{t+i\Delta} - \ln P_{t+(i-1)\Delta}, \quad i = 1, \cdots, n, \tag{1}$$

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 4, Issue 6, November 2015

where P_t is an asset price at time t and n is the number of returns sampled at Δ. Since in empirical finance we usually use a daily realized volatility, n is given by $n = 1\, day / \Delta$. The realized volatility RV_t is given by

$$RV_t = \sum_{i=1}^{n} r_{t+i\Delta}^2 . \tag{2}$$

Equation (2) goes to the integrated volatility, i.e., the true volatility, in the limit of $n \to \infty$. However, the returns we observe in the financial market are not the true returns; they are biased by microstructure noise. Let $\ln P_t^*$ be a log-price observed in the markets and contaminated by microstructure noise. Here, following [11], we assume that the microstructure noise is given by an independently distributed noise distribution $\xi_t \sim N(0, \omega^2)$. Then $\ln P_t^*$ is written as

$$\ln P_t^* = \ln P_t + \xi_t . \tag{3}$$

The return observed in the markets, r_t^* is given by

$$r_t^* = r_t + \rho_t , \tag{4}$$

where $\rho_t = \xi_t - \xi_{t-\Delta}$.

The realized volatility observed in the market RV_t^* is obtained as a sum of the squared returns r_t^*,

$$RV_t^* = \sum_{i=1}^{n} (r_{t+i\Delta}^*)^2 = RV_t + 2\sum_{i=1}^{n} r_{t+i\Delta}\rho_{t+i\Delta} + \sum_{i=1}^{n} \rho_{t+i\Delta}^2 . \tag{5}$$

After averaging RV_t^*, we find that the bias term in RV_t^* is given by $\sum_{i=1}^{n} \rho_{t+i\Delta}^2$, which corresponds to $2n\omega$. Thus, due to the bias, RV_t^* diverges as $n \to \infty$.

There exists another bias, created by non-trading hours on the financial markets. At the Tokyo Stock Exchange, domestic stocks are traded in two trading sessions from 9:00 to 11:00 and from 12:30 to 15:00. Thus, no high-frequency data are available outside of the trading sessions and the daily realized volatility calculated without including returns during non-trading hours can be underestimated.

To circumvent the bias problems, Hansen and Lunde introduced an adjustment factor that modifies the original realized volatility, so that the average of the realized volatility matches the variance of the daily returns [5]. The Hansen and Lunde (HL) adjustment factor c is given by

$$c = \frac{\sum_{t=1}^{N} (R_t - \bar{R})^2}{\sum_{t=1}^{N} RV_t^*} , \tag{6}$$

where R_t and $\bar{R}$ are the daily return and the average daily return respectively. N is the length of the time series. Using c, the modified realized volatility is then obtained as cRV_t^*.

III. REALIZED STOCHASTIC VOLATILITY MODEL

The realized stochastic volatility (RSV) model is an extended version of the stochastic volatility model and is formulated by Takahashi *et al.* [3] as

$$R_t = \exp(h_t)\varepsilon_t , \quad \varepsilon_t \sim N(0,1) \tag{7}$$

$$h_t = \mu + \phi(h_{t-1} - \mu) + \eta_t , \quad \eta_t \sim N(0, \sigma_\eta^2) \tag{8}$$

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 4, Issue 6, November 2015

$$\ln RV_t^* = \zeta + h_t + u_t, \quad u_t \sim N(0, \sigma_u^2) \tag{9}$$

where R_t is the daily return and RV_t^* is the daily realized volatility; h_t is a latent volatility defined by $h_t = \ln \sigma_t^2$. Equation (9) incorporates additional information on daily volatility into the standard SV model and thus it can be expected that the RSV model will have better predictive power for the daily volatility. The parameter ζ stands for a bias correction parameter and captures the bias that appears on the RV. Due to the fact that in the RSV model the bias correction on the RV is made through this parameter, we do not need to adjust the RV by means such as the HL adjustment factor of (6). The RSV model has five parameters $\theta = (\theta_1, \cdots, \theta_5) \equiv (\phi, \mu, \varsigma, \sigma_u^2, \sigma_\eta^2)$ that have to be determined so that the model matches the observed time series data.

IV. BAYESIAN INFERENCE BY HYBRID MONTE CARLO

In order to determine the parameters of the RSV model, we use Bayesian inference. From Bayes' theorem, we obtain a posterior probability density of parameters and the values of parameters are obtained as expectation values. Let $f(\theta, h)$ be the conditional posterior density of the RSV model [3]. The expectation value of θ_i for $i = 1, \cdots, 5$ is calculated as

$$\mathrm{E}[\theta_i] = \int \theta_i f(\theta, h) d\theta_1 \cdots d\theta_5 dh_1 \cdots dh_T . \tag{10}$$

The expectation values in (10) are obtained as average values through the MCMC method. The most time-consuming part of the MCMC method is the Monte Carlo update for the volatility variables h_t. We update volatility variables by the HMC method [6]. The HMC method was first developed for lattice simulations in quantum chromo dynamics [7]. A beneficial property of the HMC method is that it is a global algorithm that can update all variables simultaneously. This global property can accelerate the decorrelation between Monte Carlo samples. In practice, the HMC method has been applied to the SV model [8], [9] and the RSV model [10], and it is shown that the HMC method can generate well-decorrelated Monte Carlo samples of volatility variables. In this study we also use the HMC method for the volatility update.

The HMC combines the molecular dynamics (MD) simulations and the Metropolis test [6]. In the first step of the HMC method, candidates of the volatility variables are obtained by the MD simulations, which solve Hamilton's equations of motion,

$$\frac{dh_t}{d\tau} = \frac{\partial H}{\partial p_t}, \tag{11}$$

$$\frac{dp_t}{d\tau} = -\frac{\partial H}{\partial h_t}, \tag{12}$$

where p_t for $t = 1, \cdots, T$ is a conjugate momentum to h_t and the Hamiltonian H is defined by

$$H(p, h) = \frac{1}{2}\sum_{t=1}^{T} p_t^2 - \ln f(h, \theta). \tag{13}$$

In general, (11)-(12) are not solved analytically. Usually, therefore, (11)-(12) are integrated numerically by MD simulations. The standard and popular integrator for MD simulations in the HMC method is the second-order leapfrog integrator. We could also use higher-order integrators [12], [13] or other type of integrators [14] for the HMC method, although the performance of integrators depends on the model used. The minimum norm (MN) integrator [14] was tested for the HMC method in the RSV model [9] and it is found that the MN integrator performs better than the leapfrog integrator. Therefore, in this study we also use the MN integrator for the HMC method. The final step of the HMC method is the Metropolis test, in which the candidate volatility variables obtained by the MD simulations are chosen according to a probability $\sim \min\{1, \exp(-\Delta H)\}$, where ΔH is the error of the Hamiltonian caused by the numerical integration.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 4, Issue 6, November 2015

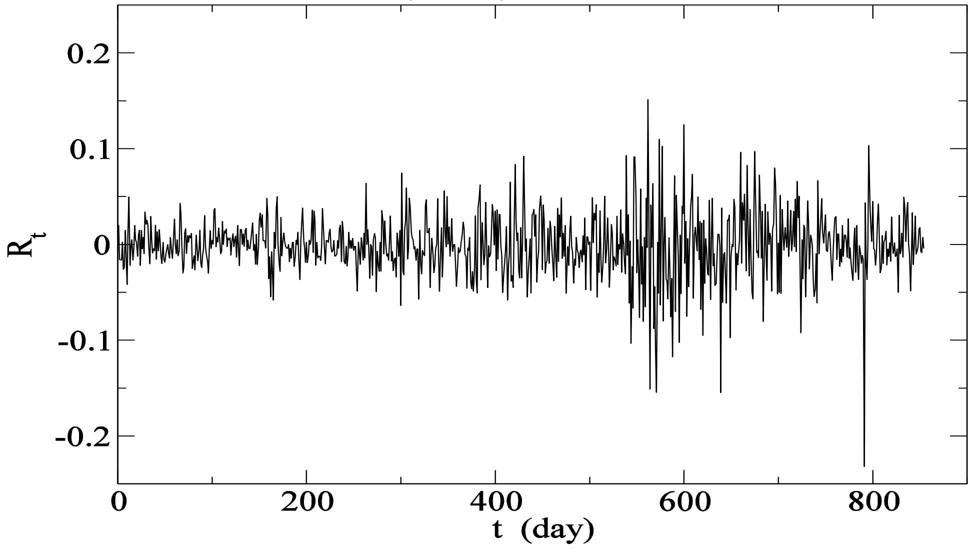

Fig. 1 Return time series of Nomura Holdings, Inc., from June 3, 2006 to December 30, 2009.

V. EMPIRICAL RESULTS

We analyze high-frequency stock price data of Nomura Holdings, Inc., from June 3, 2006 to December 30, 2009, as traded on the Tokyo Stock Exchange. Fig. 1 shows return time series of Nomura Holdings, Inc. We calculate realized volatility at various sampling frequencies, $\Delta = (1, \cdots, 35)$ min. Fig. 2 shows the realized volatilities calculated at $\Delta =$ 1 and 35 as representative examples. In order to see the bias from the microstructure noise, we plot the average values of realized volatility as a function of sampling frequency in Fig. 3. Such a graph is called the volatility signature plot [15]. We find that the average realized volatility increases as $\Delta \to 0$, which indicates that the bias from the microstructure noise increases with decreasing Δ, as expected from (5).

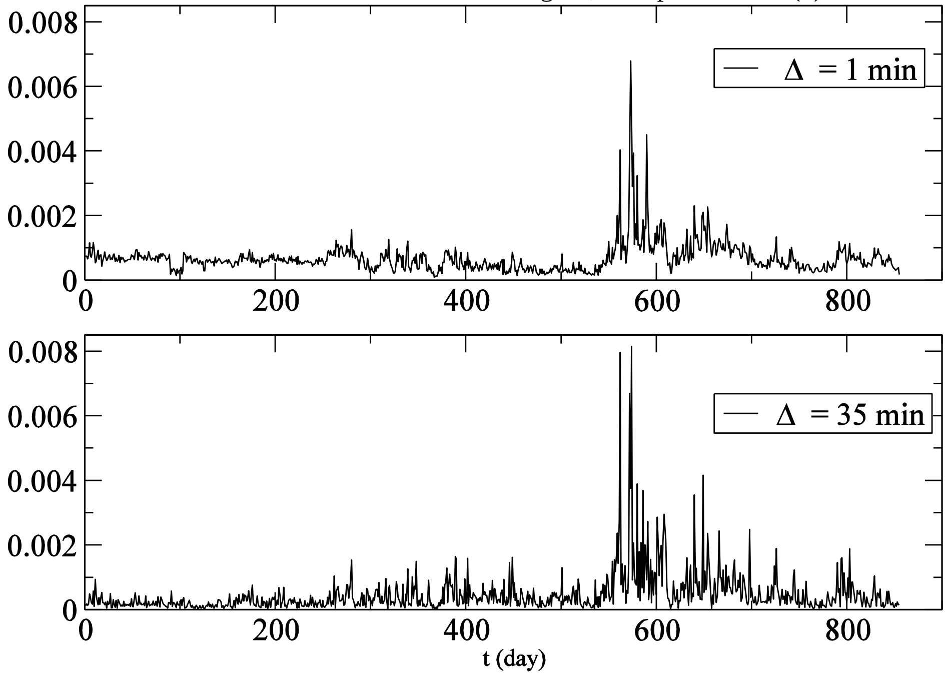

Fig. 2 Representative time series of realized volatility from June 3, 2006 to December 30, 2009.

We perform Bayesian inference of the RSV model by the MCMC method for each realized volatility data sample, i.e., each sample for $\Delta = (1, \cdots, 35)$. We discard the first 5000 MCMC samples and after that collect 50,000

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 4, Issue 6, November 2015

MCMC samples for analysis. Table 1 shows the results of model parameters obtained by the MCMC method. All the parameters ϕ are obtained to be very close to one, which indicates that the volatility time series is very persistent and long-correlated. It is also found that the parameter σ_u^2,which captures the variance of realized volatility, increases with the sampling frequency Δ. This finding is consistent with the fact that the realized volatility calculated at large sampling frequency usually has a large variance, since the number of returns to construct the realized volatility is small at large sampling frequencies.

The parameter ς corresponds to the bias correction parameter. Since the bias correction parameter explains the bias in $\ln RV_t^*$, the relationship between ς and the HL adjustment factor c is given by $\varsigma = -\ln(c)$. Fig. 4 shows the bias correction parameter ς and HL adjustment factor translated into $-\ln(c)$ as a function of the sampling frequency. It is found that both ς and $-\ln(c)$ show similar behavior, i.e., they decrease as the sampling frequency Δ increases. However, it is also found that the bias correction parameter ς is slightly larger than the HL adjustment factor $-\ln(c)$ at small sampling frequencies, where the effect of the microstructure noise is strong.

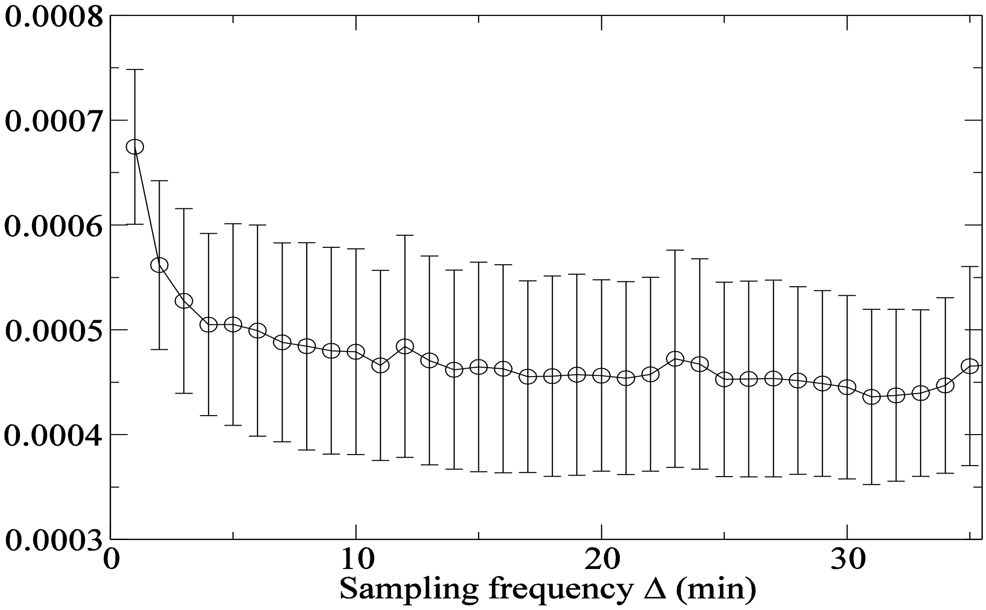

Fig. 3 Volatility signature plot (average realized volatility).

Table 1. Results obtained by the Bayesian inference performed by the MCMC method. The values in parentheses show statistical errors from the MCMC method.

Δ	ϕ	μ	ς	σ_η^2	σ_u^2
1 min	0.9354(4)	-7.01(2)	-0.458(3)	0.0320(2)	0.0549(2)
2 min	0.9385(4)	-7.10(2)	-0.600(3)	0.0328(2)	0.0834(2)
5 min	0.9526(4)	-7.20(4)	-0.721(3)	0.0376(3)	0.1457(3)
10 min	0.9539(6)	-7.24(5)	-0.797(3)	0.0440(6)	0.2194(5)
15 min	0.9550(2)	-7.26(6)	-0.864(3)	0.0463(2)	0.2925(4)
20 min	0.9572(2)	-7.26(6)	-0.885(3)	0.0453(5)	0.3500(5)
25 min	0.9532(7)	-7.26(5)	-0.921(3)	0.0499(8)	0.3987(8)
30 min	0.9549(5)	-7.26(15)	-0.956(3)	0.0488(6)	0.4738(8)
35 min	0.9609(5)	-7.27(67)	-0.932(3)	0.0435(6)	0.5002(9)

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 4, Issue 6, November 2015

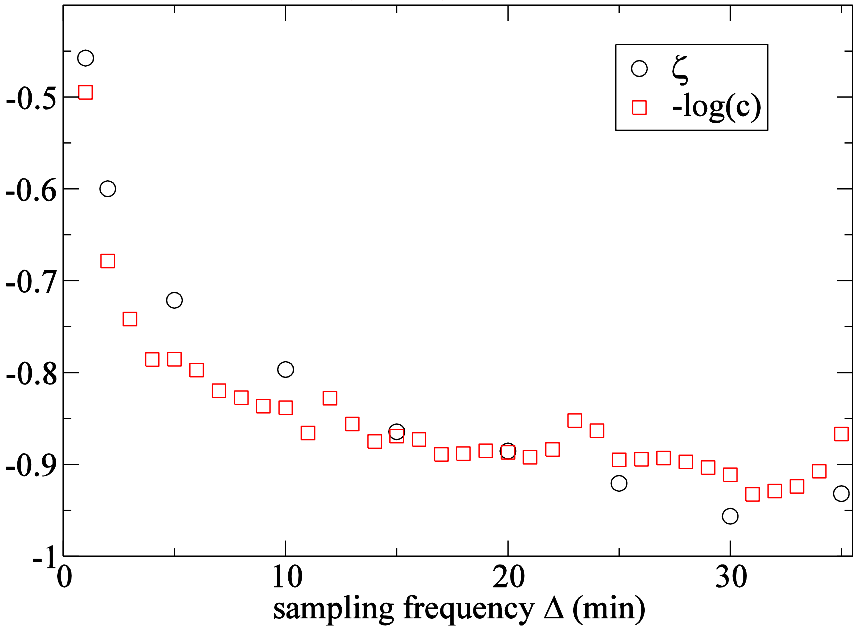

Fig. 4 The bias correction parameter ζ and HL adjustment factor $-\ln(c)$ against sampling frequency Δ.

VI. CONCLUSION

The Bayesian inference of the realized stochastic volatility (RSV) model was performed by the Markov Chain Monte Carlo method to analyze time series data of Nomura Holdings, Inc., from June 3, 2006 to December 30, 2009, as traded on the Tokyo Stock Exchange. We used the hybrid Monte Carlo method for updates of the volatility variables. We obtained the bias correction parameter ζ, which serves to modify the bias of the realized volatility as a function of the sampling frequency, and found that the bias correction parameter ζ shows similar behavior to the Hansen and Lunde adjustment factor, although there is a slight difference at high sampling frequencies, i.e., small Δ, where the microstructure noise dominates.

To test the efficiency of the RSV model further, it might be interesting to see whether the mixture of distributions hypothesis (MDH) can be applied to the observed return dynamics. In previous studies [16], [17] it is shown that the MDH holds for stock return dynamics on the Tokyo Stock Exchange when using the realized volatility as a proxy of the true volatility. If we obtain accurate volatilities from the RSV model, we can verify that the MDH also holds by using the RSV volatility.

ACKNOWLEDGMENT

Numerical calculations in this work were carried out at the Yukawa Institute Computer Facility and the facilities of the Institute of Statistical Mathematics. This work was supported by JSPS KAKENHI Grant Number 25330047.

REFERENCES

[1] S.J. Taylor, "Modeling Financial Time Series," John Wiley & Sons, New Jersey, 1986.

[2] S. Kim, N. Shephard, and S. Chib, "Stochastic volatility: Likelihood inference and comparison with ARCH models," Review of Economic Studies, vol. 65, pp. 361-393, 1998.

[3] M. Takahashi, Y. Omori, and T. Watanabe, "Estimating stochastic volatility models using daily returns and realized volatility simultaneously," Compt. Stat. & Data Anal., vol. 53, pp. 2404-2426, 2009.

[4] T.G. Andersen and T. Bollerslev, "Answering the Skeptics: Yes, Standard Volatility Models Provide Accurate Forecasts," International Economic Review, vol. 39, pp. 885-905, 1998.

[5] P.R. Hansen and A. Lunde, "A forecast comparison of volatility models: does anything beat a GARCH (1, 1)?" Journal of Applied Econometrics, vol. 20, pp. 873-889, 2005.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 4, Issue 6, November 2015
[6] S. Duane, A.D. Kennedy, B.J. Pendleton, and D. Roweth "Hybrid Monte Carlo," Phys. Lett. B, vol. 195, pp. 216-222, 1987.

[7] A. Ukawa, "Lattice QCD Simulations Beyond the Quenched Approximation," Nucl. Phys. B (Proc. Suppl.), vol. 10, pp. 66-145, 1989.

[8] T. Takaishi, "Financial Time Series Analysis of SV Model by Hybrid Monte Carlo," Lecture Notes in Computer Science, vol. 5226, pp. 929-936, 2008.

[9] T. Takaishi, "Bayesian Inference of Stochastic Volatility Model by Hybrid Monte Carlo," Journal of Circuits, Systems, and Computers, vol. 18, 1381-1396, 2009.

[10] T. Takaishi, "Bayesian estimation of realized stochastic volatility model by Hybrid Monte Carlo algorithm," Journal of Physics: Conference Series, vol. 490, pp. 1742-6596, 2014.

[11] B. Zhou, "High-frequency data and volatility in foreign-exchange rates," Journal of Business & Economics Statistics, vol. 14, pp. 45-52, 1996.

[12] T. Takaishi, "Choice of Integrators in the Hybrid Monte Carlo Algorithm," Comput. Phys. Commun., vol. 133, pp. 6-17, 2000.

[13] T. Takaishi, "Higher Order Hybrid Monte Carlo at Finite Temperature," Phys. Lett. B, vol. 540, pp. 159-165, 2002.

[14] T. Takaishi and P. de Forcrand, "Testing and Tuning Symplectic Integrators for Hybrid Monte Carlo Algorithm in Lattice QCD," Phys. Rev. E, vol. 73, 036706, 2006.

[15] T.G. Andersen, T. Bollerslev, F.X. Diebold, and P. Labys, "Great Realization," Risk, pp. 105-108, March 2000.

[16] T. Takaishi, T.T. Chen, and Z. Zheng, "Analysis of Realized Volatility in Two Trading Sessions of the Japanese Stock Market," Progress of Theoretical Physics Suppl. 194, pp. 43-54, 2012

[17] T. Takaishi, "Finite-Sample Effects on the Standardized Returns of the Tokyo Stock Exchange," Procedia—Social and Behavioral Sciences, vol. 65, pp. 968-973, 2012.