Comput Stat
DOI 10.1007/s00180-014-0546-6

ORIGINAL PAPER

Estimation of realized stochastic volatility models using Hamiltonian Monte Carlo-Based methods

Didit B. Nugroho · Takayuki Morimoto

Received: 10 March 2014 / Accepted: 18 November 2014

Abstract This study develops and compares performance of Hamiltonian Monte Carlo (HMC) and Riemann manifold Hamiltonian Monte Carlo (RMHMC) samplers with that of multi-move Metropolis-Hastings sampler to estimate stochastic volatility (SV) and realized SV models with asymmetry effect. In terms of inefficiency factor, empirical results show that the RMHMC sampler give the best performance for estimating parameters, followed by multi-move Metropolis-Hastings sampler. In particular, it is also shown that RMHMC sampler offers a greater advantage in the mixing property of latent volatility chains and in the computational time than HMC sampler.

Keywords Realized stochastic volatility model · Hamiltonian Monte Carlo · Inefficiency factor

1 Introduction

Stochastic volatility model is among the most important tools for modeling volatility of a financial time series since it recognizes the time-varying return volatility found in financial markets and describes the volatility as a stochastic function. A basic and very popular SV model is perhaps the SV model introduced by Taylor (1982). In his discrete time model, the volatility process is modeled as a first-order auto-regression for the

D. B. Nugroho
Department of Mathematics, Satya Wacana Christian University, Jln. Diponegoro 52-60, Salatiga 50711, Central Java, Indonesia
e-mail: didit.budinugroho@staff.uksw.edu

T. Morimoto (✉)
Department of Mathematical Sciences, Kwansei Gakuin University, 2-1 Gakuen, Sanda, Hyogo 669-1337, Japan
e-mail: morimot@kwansei.ac.jp

Published online: 27 December 2014

log-squared volatility with innovations assumed to be independent of the innovations in the conditional return equation. When the high-frequency intra-day returns are available throughout the day, realized variance (RV) proposed by Andersen et al. (2001) has become a popular empirical measure of volatility. This measure is defined as the sum of the intra-daily squared returns. Recently, models joining SV and realized measure, known as the realized stochastic volatility (RSV) model, have been developed (e.g., Takahashi et al. 2009; Dobrev and Szerszen 2010; Koopman and Scharth 2013).

Although intuitively sound and statistically elegant, empirical applications of SV model has been limited because of difficulty to evaluate the likelihood function. The derivation of the likelihood function of model parameter involves a high dimensional integration problem where the latent volatility process is integrated out. Such computation becomes prohibitive with a large number of observations making difficult the estimation of the model parameters.

Many estimation methods have been proposed for estimating the parameters and latent volatility of the SV model from a set of observed returns [see, e.g., the survey in Broto and Ruiz (2004) and Jacquier and Polson (2011)]. One of these methods is the Bayesian approach proposed by Shephard (1993) and Jacquier et al. (1994) which use MCMC algorithms to calculate the posterior distribution of the model parameters. In this way it is possible to make exact inferences about the parameters of the model and, in particular, about the latent volatility process.

For estimating latent volatilities, Shephard and Pitt (1997), Kim et al. (1998), and Jacquier et al. (2004) considered a very simple Bayesian Metropolis–Hastings method so called "single-move Metropolis–Hastings" sampler, which is based on one-at-a-time updating. This sampler was shown to be quite inefficient because produces highly correlated samples. To solve this problem more efficient methods based on mixtures (MIX) and multi-move Metropolis–Hastings (MM-MH) samplers were suggested by, respectively, Kim et al. (1998) (MIX) and Watanabe and Omori (2004), and Asai (2005) (MM-MH). There are two another efficient methods called HMC [introduced by Duane et al. (1987) and Neal (2010)] and RMHMC [introduced by Girolami and Calderhead (2011)], which are based on the Hamiltonian dynamics. Takaishi (2009) found that the HMC method improves the mixing of the Markov chain by reducing autocorrelation of volatility variables faster than the single-move Metropolis algorithm. In turn, Girolami and Calderhead (2011) demonstrated that, among various sampling methods (including HMC), RMHMC sampling yields the best performance for parameters and latent volatilities, in terms of time-normalized effective sample size. Higher efficiency of all these methods (MIX, MM-MH, HMC, and RMHMC) is due to that they work by updating several or even entire latent volatilities in one block at once.

Furthermore, it has long been recognized that the returns of financial assets are negatively correlated with changes in the volatilities of returns. Following Black (1976), this phenomenon is known in the literature as "leverage effect" or "asymmetry effect". They note that the firm becomes riskier and the future expected volatility rises when the prices fall and the firm's financial leverage ratio (value of firm debt relative to equity) increases. In order to accommodate this effect, extensions of a simple discrete time model due to Taylor (1982) have been analyzed by, e.g., Harvey and Shephard (1996) and Jacquier et al. (2004) using a different setup. Following Omori et al. (2007), we use the term *asymmetry* in this study. In this context, Omori et al. (2007) and Omori and Watanabe (2008) extended mixture samplers proposed by Kim et al. (1998) and multi-move sampler proposed by Watanabe and Omori (2004), respectively, for estimating an SV model with asymmetry effect (ASV model). Recently, Takahashi et al. (2014) proposed multi-move sampler for estimating RSV models with asymmetry effect and generalized Student's t-distributions, which are extensions of the RSV model with asymmetry effect (RASV) proposed by Takahashi et al. (2009).

In this study, we develop both HMC and RMHMC samplers to estimate parameters in the ASV model, RASV model, and RASV model with Student's t-distribution proposed by Omori and Watanabe (2008) and Takahashi et al. (2009, 2014), respectively. We also compare their efficiency to that of MM-MH sampler proposed by these authors. So this study presents two main contributions. First, in contrast to previous studies that have proposed both HMC dan RMHMC methods to estimate a basic SV model, including Takaishi (2009) and Girolami and Calderhead (2011), we implement both methods to more general model incorporating an asymmetry effect, a Student's t-error distribution, and an RV measure since many empirical studies have shown superiority of these extensions compared to the basic SV model. Our proposed estimation method consists of two-steps that represent the joint estimation method. In the first step, we estimate the parameters of the returns and volatility process, including asymmetric parameter, using either HMC or RMHMC samplers. In the second step, we sample the parameters of the RV equation sample from its posteriors directly. Second, we investigate the performance of these methods using real data used by Omori and Watanabe (2008) and Takahashi et al. (2009, 2014). In terms of inefficiency factor, simulation results suggest that the speed of convergence of the RMHMC sampler is fastest in sampling unknown parameters and faster than that of the HMC sampler in sampling latent volatilities. Meanwhile the speed of convergence of the HMC and MM-MH samplers is found to be highly competitive in sampling unknown parameters.

This study is organized as follows. Sections 2 and 3, respectively, describe the HMC and RMHMC samplers and their application to the ASV and RASV models. Section 4 applies the model and samplers to daily real returns and then compares to the MM-MH results obtained by Omori and Watanabe (2008), Takahashi et al. (2009), and Takahashi et al. (2014). Section 5 is the conclusion of the study.

2 MCMC using Hamiltonian dynamics

2.1 HMC sampler

The HMC sampler alternately combines Gibbs updates with Metropolis updates and avoids the random walk behavior. This method proposes a new state by computing a trajectory obeying Hamiltonian dynamics Neal (2010). The trajectory is guided by first-order gradient of the log of the posterior by applying time discretization of the Hamiltonian dynamics. This gradient information encourages the HMC trajectories in the direction of high probabilities, resulting in a high-acceptance rate and ensuring that the accepted draws are not highly correlated Marwala (2012).

Let us consider position variables (parameters) $\boldsymbol{\theta} \in \mathbb{R}^D$ and introduce an independent auxiliary variable $\omega \in \mathbb{R}^D$ with density $p(\omega) = \mathcal{N}(\omega|0, \mathbf{M})$. In physical analogy, the negative logarithm of the joint probability density for the parameters of interest, $-\mathcal{L}(\boldsymbol{\theta}) \equiv -\ln p(\boldsymbol{\theta}|\mathbf{y})$, denotes a potential energy function, the auxiliary variable ω is analogous to a momentum variable, and the covariance matrix $\mathbf{M}$ denotes a mass matrix.

The Hamiltonian dynamics system is described by a function of two variables known as the Hamiltonian function, $H(\boldsymbol{\theta}, \omega)$, which is a sum of the potential energy $U(\boldsymbol{\theta})$ and kinetic energy $K(\omega)$, Neal (2010),

$$H(\boldsymbol{\theta}, \omega) = U(\boldsymbol{\theta}) + K(\omega),$$

where $U(\boldsymbol{\theta}) = -\mathcal{L}(\boldsymbol{\theta}) + \frac{1}{2}\log\left\{(2\pi)^D|\mathbf{M}|\right\}$ and $K(\omega) = \frac{1}{2}\omega'\mathbf{M}^{-1}\omega$. The second term on the $U(\boldsymbol{\theta})$ equation results from the normalization factor. The deterministic proposal for the position variable is obtained by solving the Hamiltonian equations for the momentum and position variables, respectively, given by

$$\frac{d\theta}{d\tau} = \frac{\partial H}{\partial \omega} = \mathbf{M}^{-1}\omega \quad \text{and} \quad \frac{d\omega}{d\tau} = -\frac{\partial H}{\partial \boldsymbol{\theta}} = \nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}).$$

These equations determine how $\boldsymbol{\theta}$ and ω change over a fictitious time τ. Starting with the current state $(\boldsymbol{\theta}, \omega)$, the proposed state $(\boldsymbol{\theta}^*, \omega^*)$ is then accepted as the next state of the Markov chain with probability

$$P(\boldsymbol{\theta}, \omega; \boldsymbol{\theta}^*, \omega^*) = \min\left\{1, \exp\{-H(\boldsymbol{\theta}^*, \omega^*) + H(\boldsymbol{\theta}, \omega)\}\right\}.$$

In practice, the differential equations of Hamiltonian dynamics are often simulated in a finite number of steps using the *leapfrog scheme* which will be described in the next subsection. It is possible to devise methods that have a higher order of accuracy than the leapfrog method [for example, see McLachlan and Atela (1992)].

2.2 RMHMC sampler

Recently, Girolami and Calderhead (2011) proposed a new HMC method, called RMHMC, for improving the convergence and mixing of chain. This is a sampling method derived from HMC, and provides an adaptation mechanism for HMC by exploiting the Riemannian geometry of the parameter space. RMHMC accounts for the local structure of the joint probability density by adapting the covariance matrix $\mathbf{M}$ used in HMC. In their study, $\mathbf{M}$ depends on the variable $\boldsymbol{\theta}$ and can be any positive definite matrix. $\mathbf{M}(\boldsymbol{\theta})$ is chosen to be the metric tensor, i.e.

$$\mathbf{M}(\boldsymbol{\theta}) = \text{cov}\left[\frac{\partial}{\partial \boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta})\right] = -E_{\mathbf{y}|\boldsymbol{\theta}}\left[\frac{\partial^2}{\partial \boldsymbol{\theta}^2}\mathcal{L}(\boldsymbol{\theta})\right]$$

which is the expected Fisher information matrix plus the negative Hessian of the log prior. Therefore, the Hamiltonian equations for the momentum and position variables, respectively, are now defined by

$$\frac{d\boldsymbol{\theta}}{d\tau} = \frac{\partial H}{\partial \omega} = \mathbf{M}(\boldsymbol{\theta})^{-1}\omega$$

and

$$\frac{d\omega}{d\tau} = -\frac{\partial H}{\partial \boldsymbol{\theta}} = \nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}) - \tfrac{1}{2}\mathrm{tr}\left[\mathbf{M}(\boldsymbol{\theta})^{-1}\frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right] + \tfrac{1}{2}\omega'\mathbf{M}(\boldsymbol{\theta})^{-1}\frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\mathbf{M}(\boldsymbol{\theta})^{-1}\omega.$$

The above position variable equation requires calculation of the second- and third-order derivatives of $\mathcal{L}$. This adds to the computational complexity of the algorithm and can be infeasible in many applications.

For computer implementation, the differential equations of Hamiltonian dynamics must be discretized. The leapfrog scheme that is typically used is quite simple. This is designed to generate an ergodic Markov chain, which converges to a unique stationary density (also called an equilibrium density). As Ishwaran (1999) explains, the value for ω is superfluous, although it plays a critical role in the step (1) below by introducing a stochastic transition designed to make the chain irreducible and aperiodic. Step (2) below provides the deterministic discrete time approximation to the Hamiltonian dynamics, and combined with the Metropolis acceptance rule in third step, ensures that detailed balance is satisfied.

Neal (2010) showed that the leapfrog scheme gives second order accuracy and yields better results than the standard and modified Euler's method. The generalized leapfrog algorithm, described by Leimkuhler and Reich (2004), operates as follows (for a chosen step size Δ_τ and simulation length N_L):

(1) update the momentum variable in the first half step using the equation

$$\omega_{\tau+\frac{1}{2}\Delta_\tau} = \omega_\tau - \tfrac{1}{2}\Delta_\tau \frac{\partial H\left(\boldsymbol{\theta}_\tau, \omega_{\tau+\frac{1}{2}\Delta_\tau}\right)}{\partial \boldsymbol{\theta}},$$

(2) update the parameter $\boldsymbol{\theta}$ over a full time step using the equation

$$\theta_{\tau+\Delta_\tau} = \boldsymbol{\theta}_\tau + \frac{\Delta_\tau}{2}\left\{\frac{\partial H\left(\boldsymbol{\theta}_\tau, \omega_{\tau+\frac{1}{2}\Delta_\tau}\right)}{\partial \omega} + \frac{\partial H\left(\boldsymbol{\theta}_\tau + \Delta_\tau, \omega_{\tau+\frac{1}{2}\Delta_\tau}\right)}{\partial \omega}\right\},$$

and

(3) update the momentum variable in the second half step using equation

$$\omega_{\tau+\Delta_\tau} = \omega_{\tau+\frac{1}{2}\Delta_\tau} - \tfrac{1}{2}\Delta_\tau \frac{\partial H\left(\boldsymbol{\theta}_{\tau+\Delta_\tau}, \omega_{\tau+\frac{1}{2}\Delta_\tau}\right)}{\partial \boldsymbol{\theta}}.$$

In fact, the generalized leapfrog method is the combination of the trapezoidal rule for variable $\boldsymbol{\theta}$ with a variant of the midpoint rule for variable ω. In order to ensure that the leapfrog algorithm is performing well, the step size and number of leapfrog steps must be tuned. Selection of these parameter values is particularly problematic and there is no general guidance on how these values should be chosen. So this can usually be done with some experimentation.

The full algorithm for HMC or RMHMC can then be summarized in the following three steps (for covariance matrix $\mathbf{M}$).

(1) Randomly draw a sample momentum vector $\omega \sim \mathcal{N}(\omega|0, \mathbf{M})$.
(2) Starting with the current state $(\boldsymbol{\theta}, \omega)$, run the leapfrog algorithm for N_L steps with step size Δ_τ to generate a proposal $(\boldsymbol{\theta}^*, \omega^*)$. At every leapfrog step, especially for RMHMC algorithm, the value of $\omega_{\tau+\frac{1}{2}\Delta_\tau}$ and $\boldsymbol{\theta}_{\tau+\Delta_\tau}$ are determined numerically by a fixed-point iteration method.
(3) Accept $(\boldsymbol{\theta}^*, \omega^*)$ with probability $P(\boldsymbol{\theta}, \omega; \boldsymbol{\theta}^*, \omega^*)$, otherwise retain $(\boldsymbol{\theta}, \omega)$ as the next Markov chain draw.

3 MCMC simulation in the RASV models

The RASV model with Student's t-distribution (RASVt) is expressed as

$$\left.\begin{aligned}
R_t &= e^{\frac{1}{2}\alpha_t} z_t^{\frac{1}{2}} \epsilon_t, && t = 1, \ldots, T,\\
Y_t &= \xi + h_t + \sigma_u u_t, && t = 1, \ldots, T,\\
h_{t+1} &= \mu + \phi(h_t - \mu) + \sigma_\eta \eta_t, && t = 1, \ldots, T-1,\\
h_1 &\sim \mathcal{N}\left(\mu, \sigma_\eta^2 / \left(1 - \phi^2\right)\right), &&\\
z_t &\sim \mathcal{IG}\left(\nu/2, \nu/2\right), &&\\
\begin{bmatrix} \epsilon_t \\ u_t \\ \eta_t \end{bmatrix} &\sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & \rho \\ 0 & 1 & 0 \\ \rho & 0 & 1 \end{bmatrix}\right), &&
\end{aligned}\right\} \tag{1}$$

where R_t are the returns over a unit time period from which the autocorrelations are removed, Y_t are the logarithm of RV over a unit time period, ν denotes degrees of freedom, and $\mathcal{N}$ and $\mathcal{IG}$ represent normal distribution and inverse gamma distribution, respectively. This model is similar to the RSVt model of Takahashi et al. (2014). In this case, they estimate both parameter ϕ and parameter vector (σ_η, ρ) using Metropolis–Hastings (MH) sampler separately, the latent volatilities h_t using MM-MH sampler, and the others directly from their posteriors.

In fact, the model (1) is an extension of the so called ASV-RVC models introduced by Takahashi et al. (2009). Their model (hereafter RASV model) is written as

$$\left.\begin{aligned}
R_t &= \sigma_\epsilon e^{\frac{1}{2}h_t} \epsilon_t, && t = 1, \ldots, T,\\
Y_t &= c + \alpha_t + \sigma_u u_t, && t = 1, \ldots, T,\\
\alpha_{t+1} &= \phi \alpha_t + \sigma_\eta \eta_t, && t = 1, \ldots, T-1,\\
\alpha_1 &\sim \mathcal{N}\left(0, \sigma_\eta^2 / \left(1 - \phi^2\right)\right), &&\\
\begin{bmatrix} \epsilon_t \\ u_t \\ \eta_t \end{bmatrix} &\sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & \rho \\ 0 & 1 & 0 \\ \rho & 0 & 1 \end{bmatrix}\right), &&
\end{aligned}\right\} \tag{2}$$

where $\sigma_\epsilon = e^{\frac{1}{2}\mu}$, $\alpha_t = h_t - \mu$, and $c = \xi + \mu$. This model is an extension of the so called asymmetric SV model introduced by Omori and Watanabe (2008) that do not incorporate an RV equation. Omori and Watanabe (2008) and Takahashi et al. (2009) used Metropolis–Hastings (MH) sampler to estimate both parameter ϕ and parameter vector $(\sigma_\epsilon, \sigma_\eta, \rho)$ separately and MM-MH sampler to estimate the latent variables α_t.

To adopt a Gaussian nonlinear state space form with uncorrelated return and volatility errors, we introduce a transformation given as follows:

$$\eta_t = \rho\epsilon_t + \sqrt{1-\rho^2}\zeta_t, \quad t = 1, 2, ..., T,$$

where $\zeta_t \overset{\text{iid}}{\sim} \mathcal{N}(0, 1)$ and $corr\ (\epsilon_t, \zeta_t) = 0$. By this transformation, $var\ (\eta_t) = 1$ and $corr\ (\eta_t, \epsilon_t) = \rho$, which satisfies the specifications of RASV model. A reparameterization of the models (1) and (2) yields, for $t = 1, \ldots, T-1$,

$$h_{t+1} = \mu + \phi(h_t - \mu) + \rho\sigma_\eta^{-1} R_t e^{-\frac{1}{2}h_t} z_t^{-\frac{1}{2}} + \sigma_\eta\sqrt{1-\rho^2}\zeta_t$$

and

$$\alpha_{t+1} = \phi\alpha_t + \rho\sigma_\eta\sigma_\epsilon^{-1} R_t e^{-\frac{1}{2}h_t} + \sigma_\eta\sqrt{1-\rho^2}\zeta_t,$$

respectively.

Let us denote two observation vectors $\mathbf{R} = \{R_i\}_{i=1}^T$ and $\mathbf{Y} = \{Y_i\}_{i=1}^T$, three latent vectors $\mathbf{h} = \{h_i\}_{i=1}^T$, $\boldsymbol{\alpha} = \{\alpha_i\}_{i=1}^T$, and $\mathbf{z} = \{z_i\}_{i=1}^T$, and two parameter vectors $\theta_1 = (c, \sigma_u, \sigma_\epsilon, \phi, \sigma_\eta, \rho)$ and $\theta_2 = (\xi, \sigma_u, \mu, \phi, \sigma_\eta, \rho, \nu)$. The joint posterior densities for the models (2) and (1) can be written as follows:

$$\begin{aligned} p(\theta_1, \boldsymbol{\alpha}|\mathbf{R}, \mathbf{Y}) = &\prod_{t=1}^T p(R_t|\sigma_\epsilon, \alpha_t) \times \prod_{t=1}^T p(Y_t|c, \sigma_u) \\ &\times p(\alpha_1|\phi, \sigma_\eta) \times \prod_{t=2}^T p(\alpha_t|\alpha_{t-1}, \sigma_\epsilon, \phi, \sigma_\eta, \rho) \times p(\theta_1) \end{aligned} \tag{3}$$

and

$$\begin{aligned} p(\theta_2, \mathbf{h}, \mathbf{z}|\mathbf{R}, \mathbf{Y}) = &\prod_{t=1}^T p(R_t|h_t, z_t, \nu) \times \prod_{t=1}^T p(z_t|\nu) \times \prod_{t=1}^T p(Y_t|\xi, \sigma_u) \\ &\times p(h_1|\mu, \phi, \sigma_\eta) \times \prod_{t=2}^T p(h_t|h_{t-1}, z_{t-1}, \mu, \phi, \sigma_\eta, \rho) \times p(\theta_2), \end{aligned} \tag{4}$$

respectively. Following standard practice, we assume that our models are completed by priors for the unknown structural parameters as follows:

$$\sigma_\epsilon^2 \sim \mathcal{IG}\left(\frac{a_{\sigma_\epsilon}}{2}, \frac{b_{\sigma_\epsilon}}{2}\right), \quad \sigma_u^2 \sim \mathcal{IG}\left(\frac{a_{\sigma_u}}{2}, \frac{b_{\sigma_u}}{2}\right), \quad \sigma_\eta^2 \sim \mathcal{IG}\left(\frac{a_{\sigma_\eta}}{2}, \frac{b_{\sigma_\eta}}{2}\right),$$
$$c \sim \mathcal{N}\left(m_c, v_c\right), \quad \frac{\phi+1}{2} \sim \mathcal{B}(a_\phi, b_\phi), \quad \frac{\rho+1}{2} \sim \mathcal{B}(a_\rho, b_\rho),$$
$$\xi \sim \mathcal{N}\left(m_\xi, v_\xi\right), \quad \mu \sim \mathcal{N}\left(m_\mu, v_\mu\right), \quad \nu \sim \mathcal{G}\left(a_\nu, b_\nu\right)$$

where $\mathcal{B}(\cdot,\cdot)$ and $\mathcal{G}(\cdot,\cdot)$ represent beta and gamma distributions, respectively. This choice of priors ensures that all parameters have the right support; in particular, the beta prior for ϕ and ρ ensures that $-1 < \phi, \rho < 1$. Notice that Omori and Watanabe (2008) and Takahashi et al. (2009) used the Wishart density for the prior of $\begin{bmatrix} \sigma_\epsilon^2 & \rho\sigma_\epsilon\sigma_\eta \\ \rho\sigma_\epsilon\sigma_\eta & \sigma_\eta^2 \end{bmatrix}$ and Girolami and Calderhead (2011) chose $p(\sigma_\epsilon) \propto \sigma_\epsilon^{-1}$.

By following the Omori and Watanabe (2008) and Takahashi et al. (2009) steps, the MCMC scheme for estimating the ASV and RASV model proceeds as follows:

(0) Initialize θ_1 and $\boldsymbol{\alpha}$.
(1) Draw $\boldsymbol{\alpha}^{(i+1)}$ from $p\left(\boldsymbol{\alpha} \,\middle|\, \mathbf{R}, \mathbf{Y}, \theta^{(i)}\right)$.
(2) Draw $\phi^{(i+1)}$ from $p\left(\phi \,\middle|\, \mathbf{R}, \boldsymbol{\alpha}^{(i+1)}, \sigma_\epsilon^{(i)}, \sigma_\eta^{(i)}, \rho^{(i)}\right)$.
(3) Draw $(\sigma_\epsilon, \sigma_\eta, \rho)^{(i+1)}$ from $p\left(\sigma_\epsilon, \sigma_\eta, \rho \,\middle|\, \mathbf{R}, \boldsymbol{\alpha}^{(i+1)}, \phi^{(i+1)}\right)$.
(4) Draw $\sigma_u^{(i+1)}$ from $p\left(\sigma_u \,\middle|\, \mathbf{Y}, \boldsymbol{\alpha}^{(i+1)}, c^{(i)}\right)$.
(5) Draw $c^{(i+1)}$ from $p\left(c \,\middle|\, \mathbf{Y}, \boldsymbol{\alpha}^{(i+1)}, \sigma_u^{(i+1)}\right)$.
(6) Set $i = i + 1$ and go to step (1),

where steps (4) and (5) are skipped on estimating the ASV model. In the above scheme, both the HMC and RMHMC samplers are applied in steps (1)–(3). To estimate the RASVt model, we follow the MCMC scheme proposed by Takahashi et al. (2014) as follows:

(0) Initialize θ_2, $\mathbf{h}$, and $\mathbf{z}$.
(1) Draw $\phi^{(i+1)}$ from $p\left(\phi \,\middle|\, \mathbf{R}, \mathbf{h}^{(i)}, \mathbf{z}^{(i)}, \xi^{(i)}, \sigma_u^{(i)}, \mu^{(i)}, \sigma_\eta^{(i)}, \rho^{(i)}, \nu^{(i)}\right)$.
(2) Draw $(\sigma_\eta, \rho)^{(i+1)}$ from $p\left(\sigma_\eta, \rho \,\middle|\, \mathbf{R}, \mathbf{h}^{(i)}, \mathbf{z}^{(i)}, \xi^{(i)}, \sigma_u^{(i)}, \mu^{(i)}, \phi^{(i+1)}, \nu^{(i)}\right)$.
(3) Draw $\mu^{(i+1)}$ from $p\left(\mu \,\middle|\, \mathbf{R}, \mathbf{h}^{(i)}, \mathbf{z}^{(i)}, \xi^{(i)}, \sigma_u^{(i)}, \phi^{(i+1)}, \sigma_\eta^{(i+1)}, \rho^{(i+1)}, \nu^{(i)}\right)$.
(4) Draw $\nu^{(i+1)}$ from $p\left(\nu \,\middle|\, \mathbf{z}^{(i)}\right)$.
(5) Draw $\xi^{(i+1)}$ from $p\left(\xi \,\middle|\, \mathbf{Y}, \mathbf{h}^{(i)}, \sigma_u^{(i)}\right)$.
(6) Draw $\sigma_u^{(i+1)}$ from $p\left(\sigma_u \,\middle|\, \mathbf{Y}, \mathbf{h}^{(i)}, \xi^{(i+1)}\right)$.
(7) Draw $\mathbf{z}^{(i+1)}$ from $p\left(\mathbf{z} \,\middle|\, \mathbf{R}, \mathbf{h}^{(i)}, \mu^{(i+1)}, \phi^{(i+1)}, \sigma_\eta^{(i+1)}, \rho^{(i+1)}, \nu^{(i+1)}\right)$.
(8) Draw $\mathbf{h}^{(i+1)}$ from $p\left(\mathbf{h} \,\middle|\, \mathbf{R}, \mathbf{Y}, \mathbf{z}^{(i+1)}, \theta_2^{(i+1)}\right)$.
(9) Set $i = i + 1$ and go to step (1).

In this scheme, both the HMC and RMHMC samplers are applied in steps (1), (2), (4), (7), and (8). The detail procedures to sample parameters from each conditional posterior density are given in "Appendix".

4 Empirical results on real data

In this section, properties of two HMC-based samplers are investigated on the ASV, RASV, and RASVt models using real data. The purpose of these examples is to illustrate the efficiency of the samplers, which yields fast convergence. The efficiency is based on the inefficiency factor (IF) defined as

$$\mathrm{IF} = 1 + 2\sum_{j=1}^{\infty} \rho_j,$$

where ρ_j is the autocorrelation at lag j. This quantity can be interpreted as the number of correlated samples with the same variance-reducing power as one independent sample. An IF value of one indicates that the draws are uncorrelated while large values indicate a slow convergence as well as a bad mixing.

4.1 Observed data

To analyze the performance of the HMC and RMHMC samplers, the ASV, RASV, and RASVt models, respectively, are fitted to the daily returns in percentages of TOPIX used by Omori and Watanabe (2008) from August 1, 1997 to July 31, 2002, the daily returns and RVs sampled 1-minute in percentages of TOPIX used by Takahashi et al. (2009) from April 1, 1996 to March 31, 2005, and the daily returns and realized kernels in percentages of S&P 500 used by Takahashi et al. (2014) from February 1, 2001 to July 20, 2005. The daily return and RK for the S&P 500 are particularly obtained from Oxford-Man Institute.

4.2 Setup of MCMC algorithm

Similar to Omori and Watanabe (2008), Takahashi et al. (2009), and Takahashi et al. (2014) we estimate the ASV, RASV, and RASVt models by running the corresponding MCMC algorithms 55,000, 6,000, 25,000 iterations, respectively, and taking as burn-in period 5,000, 1,000, and 5,000 iterations. The prior distributions required in the joint prior density are set as in Table 1. For initial parameter values, we set $\boldsymbol{\alpha}^{(0)} = \ln(0.8var(\mathbf{R}))$, $\sigma_\epsilon^{(0)} = 1$, $c = 0$, $\sigma_u^{(0)} = 0.1$, $\phi^{(0)} = 0.95$, $\sigma_\eta^{(0)} = 0.1$, $\rho^{(0)} = -0.3$, $\xi = 0$, $\mu = 0$, and $\nu = 20$.

To investigate the sensitivity of the IFs to the initial values, for example, we repeated the RMHMC sampling experiments 10 times for the estimation of ASV model with starting points sampled from truncated normal distributions, i.e., $\mathcal{N}_{[0,10]}(0, 10)$ for both σ_ϵ and σ_η, $\mathcal{N}_{(-1,1)}(0, 1)$ for both ϕ and ρ, and $\mathcal{N}(0, 1)$ for $\boldsymbol{\alpha}$, and fixed parameter values in the RMHMC sampler as in Table 4. The IF is computed over 10000 samples collected after a burn-in period of 5000 samples and reported in Table 2. The results confirm that the IFs are relatively not sensitive to the initial values.

4.3 Tuning parameters for HMC and RMHMC

Tuning HMC sampler has been reported by many experts to be more difficult than tuning other MCMC samplers (Ishwaran 1999; Neal 2010) and, for this reason, HMC

Table 1 Prior densities, means, and standard deviations

Parameter	Density			Mean	SD
	ASV	RASV	RASVt		
σ_ϵ^2	$\mathcal{IG}(5, 4)$		–	1.0	0.33
c, μ	$\mathcal{N}(0, 10)$			0	10
σ_u^2	$\mathcal{IG}(2.5, 0.025)$		$\mathcal{IG}(2.5, 0.1)$	0.01; 0.07	0.024; 0.094
ϕ	$\frac{\phi+1}{2} \sim \mathcal{B}(20, 1.5)$			0.86	0.11
σ_η^2	$\mathcal{IG}(2.5, 0.025)$			0.017	0.024
ρ	$\frac{\rho+1}{2} \sim \mathcal{B}(1.5, 5)$		$\frac{\rho+1}{2} \sim \mathcal{B}(1, 2)$	−0.54; −0.34	0.31; 0.47
ξ	–		$\mathcal{N}(0, 1)$	0	1
ν	–		$\mathcal{G}(5, 0.5)$	10	4.47

Table 2 Inefficiency factors of RMHMC samples on estimating the ASV model with various initial values

σ_ϵ	ϕ	σ_η	ρ	$\boldsymbol{\alpha}$ (average)
20.5	93.3	156.8	65.3	2.7
22.5	87.8	154.3	58.6	2.9
22.8	85.7	162.4	53.1	1.6
19.1	91.9	150.8	62.1	2.1
20.2	91.2	165.5	51.0	2.0
18.7	88.2	161.9	65.1	1.9
21.9	93.1	155.6	50.9	2.2
19.0	83.3	140.5	56.7	2.0
18.3	89.6	162.4	50.9	1.9
21.8	92.7	164.6	59.6	1.8

The window width 1,000 was used to compute the inefficiency factors

Table 3 Inefficiency factors of RMHMC samples on estimating the ASV model using various number of step sizes for ϕ

Δ_τ for proposal ϕ	σ_ϵ	ϕ	σ_η	ρ	$\boldsymbol{\alpha}$ (average)
0.30	10.7	103.8	228.0	64.0	4.0
0.35	13.9	174.6	321.7	49.6	5.4
0.40	18.5	286.4	455.9	104.2	8.7
0.45	**10.4**	264.1	461.1	95.7	8.2
0.50	19.6	**81.8**	**152.1**	66.0	**1.8**
0.55	15.6	216.0	410.7	107.5	6.7
0.60	13.2	256.2	440.2	136.5	6.7
0.65	27.6	279.1	412.8	84.5	9.2
0.70	68.1	121.0	264.5	**39.5**	10.5

The window width 1,000 was used to compute the inefficiency factors. The bold number implies the smallest for each column

has not been widely adopted on estimating SV model. In fact, the performance of the HMC-based samplers is highly sensitive to two user-specified parameters, i.e., the step size Δ_τ and the number of leapfrog iterations N_L. Additionally, the RMHMC

Table 4 Tuning parameters for the HMC and RMHMC implementations in the ASV, RASV, and RASVt models

Sampler	Parameter of sampler	Parameter of model				
		$\boldsymbol{\alpha}, \mathbf{h}$	$\mathbf{z}$	ϕ	$(\sigma_\epsilon, \sigma_\eta, \rho)$	ν
HMC	N_L	100	100	100	100	100
	Δ_τ	0.01	0.01	0.0125	0.0125	0.03
RMHMC	N_L	50	50	6	6	6
	Δ_τ	0.1	0.1	0.5	0.5	0.5
	N_{FPI}	–	–	5	5	5

implementation requires a number of fixed point iterations, N_{FPI}. A bad choice of these parameters may result in slow mixing or incur a high computational cost in the algorithm. The selection of parameter values is particularly problematic, and there is no general guidance on how these values might best be chosen. Therefore, in order to adapt the HMC-based samplers parameters, we need to define objective functions. As Pasarica and Gelman (2008) point out, a natural objective function for adaptation is the asymptotic efficiency of an MCMC sampler, IF. Thus we tune our choices based on their sampling efficiency from preliminary MCMC runs.

To investigate the effect of HMC and RMHMC parameters on the speed of mixing to the posterior density, we repeated our experiments using different values of N_L, Δ_τ, and N_{FPI} for the hyperparameters and latent volatility. For example, we repeated the sampling experiments 10 times and averaged the IF values of MCMC samples, which are shown in Table 3 for the ASV model fitted to the daily TOPIX data from

Table 5 Posterior summary statistics of parameters in the ASV model after 50,000 iterations

Parameter	Mean (SD)	95 % Interval	p value	IF	Time (s)
Panel A: The MM-MH results in Table 5 of Omori and Watanabe (2008)					
σ_ϵ	1.259 (0.070)	[1.121, 1.398]	0.06	20.8	–
ϕ	0.945 (0.019)	[0.902, 0.974]	0.24	118.2	
σ_η	0.193 (0.033)	[0.138, 0.267]	0.32	206.7	
ρ	−0.442 (0.103)	[−0.630, −0.231]	0.89	92.7	
Panel B: Using HMC sampler					
σ_ϵ	1.241 (0.070)	[1.098, 1.377]	0.58	25.7	465.88
ϕ	0.954 (0.014)	[0.925, 0.980]	0.69	121.6	
σ_η	0.169 (0.025)	[0.120, 0.219]	0.72	218.6	
ρ	−0.461 (0.097)	[−0.648, −0.272]	0.96	28.0	
Panel C: Using RMHMC sampler					
σ_ϵ	1.238 (0.070)	[1.098, 1.378]	0.41	18.4	325.41
ϕ	0.953 (0.014)	[0.923, 0.980]	0.39	68.1	
σ_η	0.171 (0.026)	[0.122, 0.222]	0.20	129.5	
ρ	−0.456 (0.098)	[−0.636, −0.259]	0.83	40.6	

Table 6 Posterior summary statistics of parameters in the RASV model adopting RV 1-min without the Hansen and Lunde's (2005) adjustment after 5,000 iterations

Parameter	Mean (SD)	95 % Interval	p value	IF	Time (s)
Panel A: The MM-MH results in Table 9 of Takahashi et al. (2009)					
ξ	−1.2408 (0.0314)	[−1.3033, −1.1793]	0.47	1.6	–
σ_u^2	0.0971 (0.0044)	[0.08877, 0.1057]	0.79	34.9	
μ	0.2116 (0.0816)	[0.0368, 0.3632]	0.71	263.2	
ϕ	0.9590 (0.0079)	[0.9429, 0.9738]	0.19	47.9	
σ_η^2	0.0220 (0.0030)	[0.0167, 0.0282]	0.27	82.4	
ρ	−0.3086 (0.0435)	[−0.3936, −0.2230]	0.37	25.2	
Panel B: Using HMC sampler					
ξ	−1.2631 (0.0318)	[−1.3258, −1.2010]	0.74	1.6	101.24
σ_u^2	0.0983 (0.0042)	[0.0900, 0.1065]	0.47	29.4	
μ	0.2169 (0.0815)	[0.0617, 0.3723]	0.06	136.9	
ϕ	0.9609 (0.0071)	[0.9468, 0.9744]	0.10	30.3	
σ_η^2	0.0219 (0.0028)	[0.0169, 0.0276]	0.35	73.8	
ρ	−0.3137 (0.0444)	[−0.3941, −0.2222]	0.86	27.1	
Panel C: Using RMHMC sampler					
ξ	−1.2642 (0.0313)	[−1.3305, −1.2076]	0.64	0.8	62.26
σ_u^2	0.0991 (0.0044)	[0.0908, 0.1082]	0.40	18.0	
μ	0.2072 (0.0938)	[0.0344, 0.3911]	0.66	161.5	
ϕ	0.9623 (0.0076)	[0.9461, 0.9767]	0.51	25.9	
σ_η^2	0.0212 (0.0029)	[0.0150, 0.0266]	0.34	48.8	
ρ	−0.3233 (0.0425)	[−0.4049, −0.2370]	0.83	17.7	

August 1, 1997 to July 31, 2002. Results correspond to the last 10,000 of the total 15,000 iterations using the RMHMC sampler. RMHMC sampling was implemented by using different step sizes Δ_τ varying from 0.2 to 0.8 for parameter ϕ, a step size of 0.5 for the other parameters, and a step size of 0.1 for the latent volatilities. In this experiment, the number of leapfrog steps was kept fixed at 6 per parameter proposal and 50 per volatility proposal and the number of fixed point iterations was kept at 5 per parameter proposal. From Table 3, we can clearly see that the optimal step size for the proposal ϕ is 0.5 in the RMHMC implementation. Furthermore, we found that sampling generally became more efficient as those parameters was tuned around Girolami and Calderhead (2011)'s setup. These results are presented in Table 4 and will be used in our HMC and RMHMC implementations.

4.4 Comparison results

The MCMC algorithm with previous specification was coded in MATLAB 2011b (running in Microsoft Windows 7), on a desktop computer incorporating an Intel Xeon 3.47 GHz hexa-core CPU with 16GB RAM. Tables 5, 6, 7 and 8 present the

Table 7 Posterior summary statistics of parameters in the RASV model adopting RV 1-min with the Hansen and Lunde's (2005) adjustment after 5,000 iterations

Parameter	Mean (SD)	95 % Interval	p value	IF	Time (s)
Panel A: The MM-MH results in Table 10 of Takahashi et al. (2009)					
ξ	0.0587 (0.0314)	[−0.0035, 0.1202]	0.11	2.1	–
σ_u^2	0.0977 (0.0045)	[0.0891, 0.1067]	0.51	38.8	
μ	0.2018 (0.0792)	[0.0414, 0.3573]	0.35	220.0	
ϕ	0.9600 (0.0079)	[0.9436, 0.9743]	0.17	40.9	
σ_η^2	0.0215 (0.0031)	[0.0161, 0.0282]	0.47	81.1	
ρ	−0.3150 (0.0440)	[−0.4019, −0.2283]	0.12	21.1	
Panel B: Using HMC sampler					
ξ	0.0543 (0.0314)	[−0.0052, 0.1178]	0.90	2.3	101.24
σ_u^2	0.0991 (0.0043)	[0.0911, 0.1079]	0.96	24.3	
μ	0.1806 (0.0909)	[0.0125, 0.3507]	0.35	171.3	
ϕ	0.9627 (0.0074)	[0.9474, 0.9764]	0.88	41.4	
σ_η^2	0.0210 (0.0028)	[0.0161, 0.0270]	0.60	79.7	
ρ	−0.3213 (0.0436)	[−0.4089, −0.2394]	0.77	15.8	
Panel C: Using RMHMC sampler					
ξ	0.0554 (0.0315)	[−0.0085, 0.1142]	0.30	1.5	62.60
σ_u^2	0.0982 (0.0043)	[0.0898, 0.1066]	0.56	13.8	
μ	0.2220 (0.0745)	[0.0696, 0.3601]	0.80	123.9	
ϕ	0.9604 (0.0074)	[0.9461, 0.9749]	0.66	22.6	
σ_η^2	0.0221 (0.0027)	[0.0172, 0.0279]	0.44	43.1	
ρ	−0.3100 (0.0417)	[−0.3942, −0.2295]	0.75	14.1	

results obtained using the three MCMC samplers (MM-MH, HMC, and RMHMC) for estimating the ASV model, the RASV model adopting RV 1-min without adjustment, and the RASV model adopting RV 1-min with the Hansen and Lunde (2005)'s adjustment, and the RASVt model, respectively. All the reported results are based on samples which have passed the Geweke (1992)'s convergence diagnostic test for each parameter. The p value associated to the CD and the IF value is computed using a window width of 10 % of draw as in Omori and Watanabe (2008) and Takahashi et al. (2009). The computation time is only reported for two HMC-based samplers.

It can be seen that the produced parameter posteriors are almost identical for all cases studied. Regarding the estimation efficiency, the comparison of IFs for parameters shows that the fastest convergence has been achieved by RMHMC sampler, indicating that the RMHMC sampler generally outperforms both the HMC and MM-MH samplers. All IFs of the RMHMC sampler are smaller than the MM-MH sampler. Tables 5 and 6 indicate that the RMHMC sampler is less efficient than the HMC sampler only for sampling ρ in the ASV model and μ in the RASV model adopting RV 1-min without adjustment. Comparing performance between the HMC sampler and the MM-MH sampler, Table 5 indicates the HMC sampler on estimating the ASV model

Table 8 Posterior summary statistics of parameters in the RASVt model adopting RK after 5,000 iterations

Parameter	Mean (SD)	95 % Interval	p value	IF	Time (s)
Panel A: The MM-MH results in Table 2 of Takahashi et al. (2014)					
ξ	−0.2135 (0.0435)	[−0.3006, −0.1307]	0.343	29.67	–
σ_u	0.3590 (0.0100)	[0.3403, 0.3792]	0.865	6.96	
μ	−0.1867 (0.1772)	[−0.5346, 0.1674]	0.109	2.17	
ϕ	0.9698 (0.0069)	[0.9561, 0.9699]	0.094	5.95	
σ_η	0.1791 (0.0090)	[0.1624, 0.1978]	0.474	22.08	
ρ	−0.3904 (0.0496)	[−0.4832, −0.2905]	0.357	14.45	
ν	24.7185 (5.4374)	[15.9969, 37.4976]	0.712	65.55	
Panel B: Using HMC sampler to draw $(\phi, \sigma_\eta, \rho, \nu)$					
ξ	−0.2096 (0.0487)	[−0.3055, −0.1153]	0.132	53.56	555.02
σ_u	0.3703 (0.0102)	[0.3499, 0.3898]	0.488	4.56	
μ	−0.2031 (0.1896)	[−0.5845, 0.1677]	0.382	1.93	
ϕ	0.9725 (0.0044)	[0.9640, 0.9810]	0.252	9.30	
σ_η	0.1637 (0.0100)	[0.1441, 0.1833]	0.184	12.82	
ρ	−0.3719 (0.0473)	[−0.4168, −0.3239]	0.658	14.26	
ν	22.7302 (5.0587)	[13.1347, 32.3821]	0.136	54.46	
Panel C: Using RMHMC sampler to draw $(\phi, \sigma_\eta, \rho, \nu)$					
ξ	−0.2084 (0.0471)	[−0.3036, −0.1194]	0.921	30.13	226.67
σ_u	0.3711 (0.0102)	[0.3509, 0.3913]	0.770	4.27	
μ	−0.2048 (0.1865)	[−0.5863, 0.1491]	0.784	1.92	
ϕ	0.9728 (0.0042)	[0.9645, 0.9808]	0.651	3.27	
σ_η	0.1621 (0.0096)	[0.1441, 0.1807]	0.943	6.93	
ρ	−0.3747 (0.0466)	[−0.4180, −0.3271]	0.728	13.80	
ν	23.2007 (5.0526)	[13.6528, 33.1834]	0.821	44.60	

is more efficient for sampling ρ only. On estimating the RASV model, Tables 6 and 7 indicate that the HMC sampler always reduces inefficiencies for sampling $(\mu, \sigma_u^2, \sigma_\eta^2)$ using the MM-MH sampler. On estimating the RASVt model, Table 8 indicate that the MM-MH sampler is more efficient than the HMC sampler only for sampling ϕ. These results show that the HMC and MM-MH samplers are very competitive.

The autocorrelation and trace plots for posterior samples of RMHMC sampling are displayed in Figs. 1 and 2, respectively. The autocorrelation plots show quick decay of the autocorrelation as time lag between samples increases, indicating that the process is stationary. Trace plots of samples indicate that the chains fluctuate to be around their means, indicating that chains could have reached the right distribution. We can conclude that the mixing of chains is quite good here.

Unfortunately, when we consider on sampling the latent variables, we can only compare the efficiency of RMHMC and HMC samplers since Omori and Watanabe (2008), Takahashi et al. (2009), and Takahashi et al. (2014) did not present the efficiency of MM-MH sampler. Figure 3 displays the IF plots of latent variables $\boldsymbol{\alpha}$ and

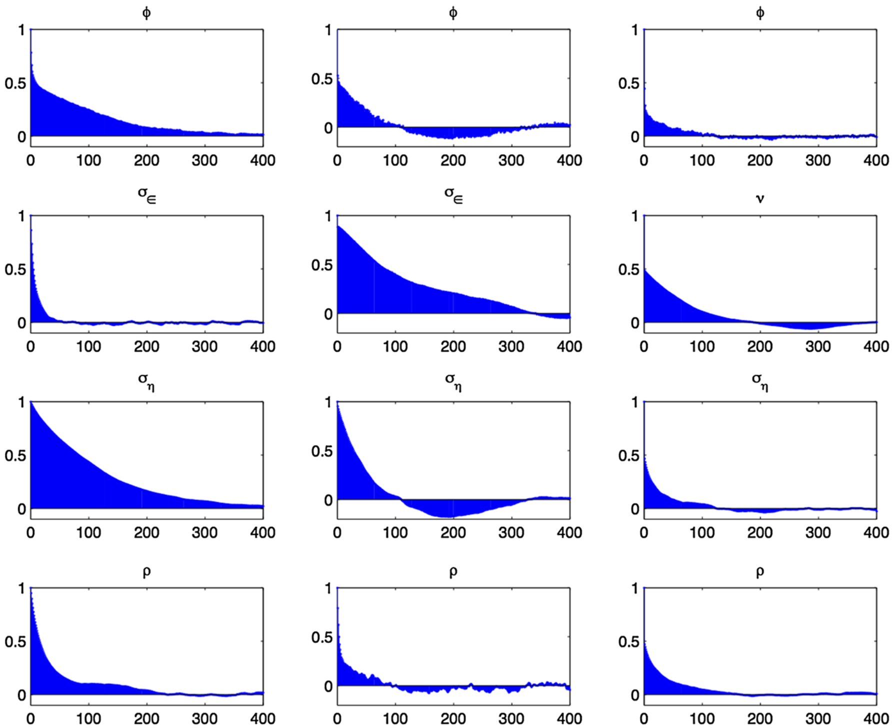

Fig. 1 Autocorrelation functions of RMHMC samples for the ASV model (*left*), the RASV model adopting RV 1-min without adjustment (*center*), and the RASV model adopting RK (*right*)

h sampled by the HMC and RMHMC samplers. Clearly, it shows that all IFs of the RMHMC sampler for sampling latent variable are always smaller than the HMC sampler, indicating that the RMHMC sampler offers a greater advantage in the mixing property of MCMC chains than the HMC sampler. In these cases, the average IFs of latent variable in the RMHMC and HMC samplers, respectively, are 1.8 and 6.1 on estimating the ASV model, 5.6 and 23.4 on estimating the RASV model without adjustment, 8.9 and 22.6 on estimating the RASV model with adjustment, and 1.8 and 2.9 on estimating the RASVt model. In addition, the average IFs of z_t in the RMHMC and HMC samplers on estimating the RASVt model are 3.6 and 4.7, respectively. These smaller IF values are results of higher acceptance rates in the RMHMC sampler for the latent variables as shown in Table 9. Girolami and Calderhead (2011) point out that the use of the metric tensor allows the RMHMC sampler to converge much more quickly to the target density.

We employed the RMHMC scheme, such that every 100 iterations we achieve the acceptance rates of between 92 and 100 % for latent variables $\boldsymbol{\alpha}$ and **h**, between 80 and 90 % for latent variable **z**, between 80 and 98 % for ϕ, and between 89 and 100 % for $(\sigma_\epsilon, \sigma_\eta, \rho)$, and between 67 and 100 % for ν. In the HMC scheme, the acceptance rates are between 67 and 99 % for latent variables $\boldsymbol{\alpha}$ and **h**, between 77 and 90 % for latent variable **z**, between 80 and 99 % for ϕ, between 81 and 100 % for $(\sigma_\epsilon, \sigma_\eta, \rho)$,

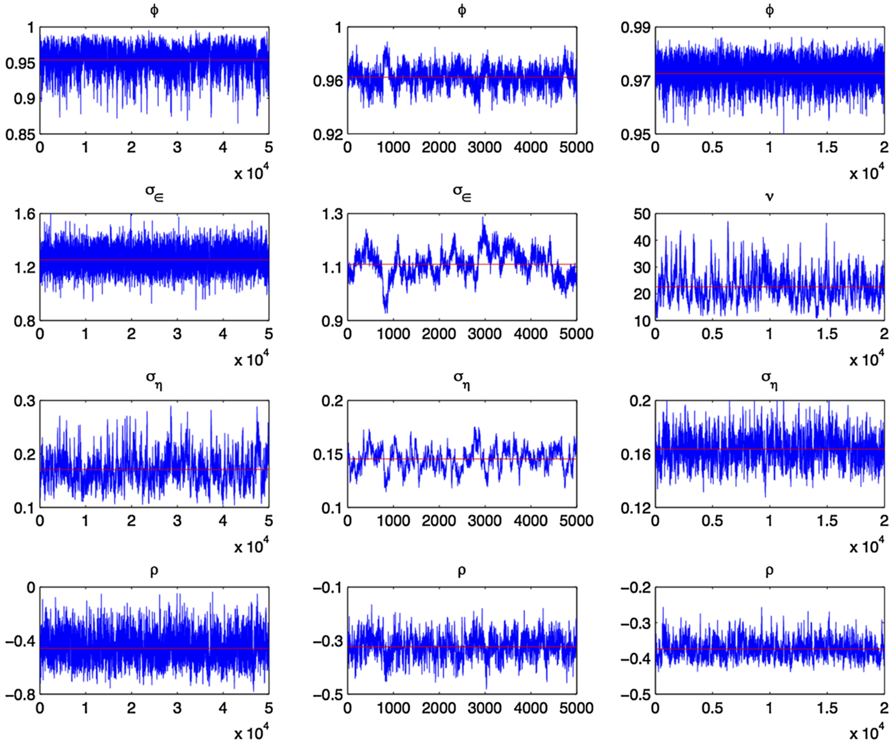

Fig. 2 Trace plots of RMHMC samples for the ASV model (*left*), the RASV model adopting RV 1-min without adjustment (*center*), and the RASVt model adopting RK (*right*)

and between 60 and 100 % for ν. Although the acceptance rates of ϕ and $(\sigma_\epsilon, \sigma_\eta, \rho)$ in the RMHMC sampler are relatively smaller than those in the MM-MH sampler as shown in Table 9, IFs of the RMHMC sampler for sampling those parameters are found to be smaller. The acceptance rates of ϕ on estimating the RASV model by the RMHMC sampler are also smaller than those by the HMC sampler, but these yield smaller IF values.

Finally, regarding computational time, the RMHMC sampler runs faster than the HMC sampler. This is partly caused by the computationally efficient tridiagonal structure of the metric tensor for the latent variables. This metric tensor is fixed and so the generalized leapfrog integration scheme is not implemented in RMHMC sampling, during each of which the metric tensor and its inverse is not recalculated, similar to the situation with the HMC sampling. Moreover, the RMHMC sampler requires significantly fewer leapfrog iterations than the HMC sampler to explore the target density.

5 Conclusion

In this study, we have developed RMHMC and HMC samplers for estimating parameters and latent variables in the ASV, RASV, and RASVt models and compared their efficiencies to the MM-MH sampler proposed by Omori and Watanabe (2008),

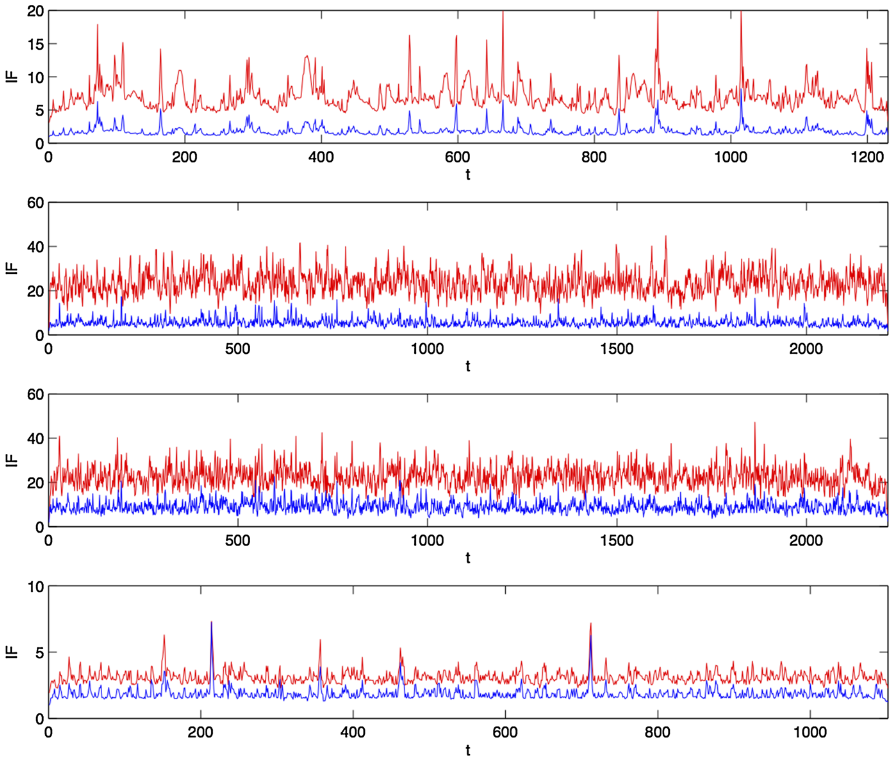

Fig. 3 IF plots for latent variable α_t of the ASV model (*first row*), the RASV model adopting RV 1-min without adjustment (*second row*), the RASV model adopting RV 1-min with adjustment (*third row*), and for latent volatility h_t of the RASVt model (*fourth row*) sampled by HMC sampler (*red line*) and RMHMC sampler (*blue line*)

Table 9 Acceptance rates of the MCMC samplers

Parameter	Acceptance rates						
	ASV model			RASV model without adjustment		RASVt model	
	MM-MH	HMC	RMHMC	HMC	RMHMC	HMC	RMHMC
h	0.856	0.889	0.975	0.955	0.967	0.940	0.978
ϕ	0.955	0.910	0.901	0.924	0.917	0.880	0.894
Σ	0.990	0.952	0.966	0.902	0.949	–	–
(σ_η, ρ)	–	–	–	–	–	0.950	0.967
ν	–	–	–	–	–	0.788	0.799
z	–	–	–	–	–	0.838	0.874

Takahashi et al. (2009), and Takahashi et al. (2014). The empirical results on real data show that the RMHMC sampler generally give the best performance in terms of inefficiency factor for all cases studied. These are results of high acceptance rates in the

RMHMC sampler for latent variables. On estimating the ASV model, the HMC sampler exhibits faster convergence than the MM-HMC sampler for sampling ρ only. On estimating the RASV models, the HMC and MM-MH samplers are very competitive, where the HMC sampler always outperforms the MM-MH sampler for sampling σ_η and ν. Therefore, this study recommends the RMHMC and HMC samplers to estimate SV and RSV models with asymmetry effect.

Appendix 1: Estimation of RASV model

Let us take the logarithm of the full conditional posterior density defined in Eq. (3):

$$\begin{aligned}\ln p(\theta_1, \boldsymbol{\alpha}|\mathbf{R}, \mathbf{Y}) = constant &+ \sum_{t=1}^{T}\left(-\frac{1}{2}\alpha_t - \frac{1}{2}\widetilde{R}_t^2\right) - T\ln(\sigma_\epsilon\sigma_u)\\ &-\frac{1}{2\sigma_u^2}\sum_{t=1}^{T}(Y_t - c - \alpha_t)^2 - T\ln(\sigma_\eta) + \frac{1}{2}\ln(1-\phi^2)\\ &-\frac{T-1}{2}\ln(1-\rho^2) - \frac{1-\phi^2}{2\sigma_\eta^2}\alpha_1^2 - \frac{1}{2\sigma_\eta^2(1-\rho^2)}\sum_{t=2}^{T}\widetilde{\alpha}_t^2\\ &-\frac{1}{2}\ln(v_c) - \frac{(c-m_c)^2}{2v_c} - \left(\frac{a_{\sigma_u}}{2}+1\right)\ln\left(\sigma_u^2\right) - \frac{b_{\sigma_u}}{2\sigma_u^2}\\ &-\left(\frac{a_{\sigma_\epsilon}}{2}+1\right)\ln\left(\sigma_\epsilon^2\right) - \frac{b_{\sigma_\epsilon}}{2\sigma_\epsilon^2} + (a_\phi - 1)\ln(1+\phi)\\ &+(b_\phi - 1)\ln(1-\phi) - \left(\frac{a_{\sigma_\eta}}{2}+1\right)\ln\left(\sigma_\eta^2\right) - \frac{b_{\sigma_\eta}}{2\sigma_\eta^2}\\ &+(a_\rho - 1)\ln(1+\rho) + (b_\rho - 1)\ln(1-\rho) \end{aligned} \tag{5}$$

where

$$\begin{aligned}\widetilde{R}_t &= R_t\sigma_\epsilon^{-1}e^{-\frac{1}{2}h_t}, \quad \text{for } t = 1, \ldots, T, \text{ and}\\ \widetilde{\alpha}_t &= \alpha_t - \phi\alpha_{t-1} - \rho\sigma_\eta\widetilde{R}_{t-1}, \quad \text{for } t = 2, \ldots, T.\end{aligned}$$

The following subsections describe how to sample latent variable $\boldsymbol{\alpha}$ and parameters in $\theta_1 = (c, \sigma_u, \sigma_\epsilon, \phi, \sigma_\eta, \rho)$.

Updating latent variable $\boldsymbol{\alpha}$

Sampling from the conditional posterior density $p(\boldsymbol{\alpha}|\mathbf{R}, \mathbf{Y}, \theta_1)$, is a difficult work because of the high-dimensional log-normal structure. Therefore, we sample latent variable $\boldsymbol{\alpha}$ using the HMC and RMHMC algorithms, where the expressions required by both algorithms to sample the entire latent volatility at once are evaluated as follows.

On the basis of Eq. (5), the logarithm of the full conditional posterior density of $\boldsymbol{\alpha}$ has the following expression:

$$\mathcal{L}(\boldsymbol{\alpha}) = constant + \sum\nolimits_{t=1}^{T} \left(-\frac{1}{2}\alpha_t - \frac{1}{2}\widetilde{R}_t^2\right) - \frac{1}{2\sigma_u^2}\sum_{t=1}^{T}(Y_t - c - \alpha_t)^2$$
$$-\frac{1-\phi^2}{2\sigma_h^2}\alpha_1^2 - \frac{1}{2\sigma_h^2(1-\rho^2)}\sum_{t=2}^{T}\widetilde{\alpha}_t^2.$$

The important aspect of the HMC and RMHMC algorithms is that both algorithms require the calculation of a gradient vector $\nabla_{\alpha_t}\mathcal{L}(\alpha_t)$. These are expressed by

$$\nabla_{\alpha_1}\mathcal{L}(\boldsymbol{\alpha}) = \nabla_{\alpha_1}\mathcal{L}_{\mathbf{R},\mathbf{Y}} - \frac{1-\phi^2}{\sigma_\eta^2}\alpha_1 + \frac{1}{\sigma_\eta^2(1-\rho^2)}\widetilde{\alpha}_2\left(\phi - \frac{1}{2}\rho\sigma_\eta\widetilde{R}_1\right),$$
$$\nabla_{\alpha_T}\mathcal{L}(\boldsymbol{\alpha}) = \nabla_{\alpha_T}\mathcal{L}_{\mathbf{R},\mathbf{Y}} - \frac{1}{\sigma_\eta^2(1-\rho^2)}\widetilde{\alpha}_T,$$
$$\nabla_{\alpha_i}\mathcal{L}(\boldsymbol{\alpha}) = \nabla_{\alpha_i}\mathcal{L}_{\mathbf{R},\mathbf{Y}} - \frac{1}{\sigma_\eta^2(1-\rho^2)}\left[\widetilde{\alpha}_i - \widetilde{\alpha}_{i+1}\left(\phi - \frac{1}{2}\rho\sigma_\eta\widetilde{R}_i\right)\right],$$

for $i = 2, \ldots, T-1$, and

$$\nabla_{\alpha_t}\mathcal{L}_{\mathbf{R},\mathbf{Y}} = -\frac{1}{2} + \frac{1}{2}\widetilde{R}_t^2 + \frac{1}{\sigma_u^2}(Y_t - c - \alpha_t), \quad t = 1, \ldots, T$$

To apply the leapfrog-RMHMC algorithm, then we also need expressions for the metric tensor and its partial derivatives. The metric tensor $\mathbf{M}(\boldsymbol{\alpha})$ is a symmetric tridiagonal matrix, where elements of main diagonal are calculated as follows:

$$\mathbf{M}(1,1) = \frac{1}{2} + \frac{1}{\sigma_u^2} + \frac{1-\phi^2}{\sigma_\eta^2} + \frac{1}{\sigma_\eta^2(1-\rho^2)}\left(\phi^2 + \frac{1}{4}\rho^2\sigma_\eta^2\right),$$
$$\mathbf{M}(i,i) = \frac{1}{2} + \frac{1}{\sigma_u^2} + \frac{1}{\sigma_\eta^2(1-\rho^2)}\left(1 + \phi^2 + \frac{1}{4}\rho^2\sigma_\eta^2\right),$$
$$\mathbf{M}(T,T) = \frac{1}{2} + \frac{1}{\sigma_u^2} + \frac{1}{\sigma_\eta^2(1-\rho^2)}$$

for $i = 2, \ldots, T-1$, and elements of superdiagonal and subdiagonal are calculated as

$$\mathbf{M}(i-1,i) = \mathbf{M}(i,i-1) = -\frac{\phi}{\sigma_\eta^2(1-\rho^2)},$$

for $i = 2, \ldots, T$. Since the above metric tensor is fixed, that is not a function of the latent volatility $\boldsymbol{\alpha}$, the associated partial derivatives with respect to the latent volatility are zero. This implies that both metric tensor and its inverse are not recalculated in the generalized leapfrog integration scheme.

Updating parameter ϕ

As far as ϕ is only concerned to be sampled, the logarithm of conditional posterior density for ϕ on the basis of Eq. (5) has the following expression:

$$\mathcal{L}(\phi) = constant + \frac{1}{2}\ln(1-\phi^2) - \frac{\alpha_1^2}{2\sigma_\eta^2}(1-\phi^2) - \frac{1}{2\sigma_\eta^2(1-\rho^2)}\sum_{t=2}^{T}\widetilde{\alpha}_t^2 \\ + (a_\phi - 1)\ln(1+\phi) + (b_\phi - 1)\ln(1-\phi),$$

which is in non-standard form, and therefore we cannot sample from it directly. In the following, we evaluate the expressions required by the HMC and RMHMC algorithms to sample ϕ.

For dealing with the constraints $-1 < \phi < 1$, it is necessary to implement the transformations $\phi = \tanh(\tilde{\phi})$ and then introduce a Jacobian factor into the acceptance ratio. The derivative of the transformation is $\mathrm{d}\phi/\mathrm{d}\tilde{\phi} = 1 - \tanh^2(\tilde{\phi}) = 1 - \phi^2$. Then, the partial derivative of the above log joint posterior with respect to the transformed parameter is as follows:

$$\nabla_{\tilde{\phi}}\mathcal{L}(\phi) = -\phi + \frac{\alpha_1^2}{\sigma_\eta^2}(\phi - \phi^3) + \frac{1-\phi^2}{\sigma_\eta^2(1-\rho^2)}\sum_{t=2}^{T}\widetilde{\alpha}_t\alpha_{t-1} \\ + (a_\phi - 1)(1-\phi) - (b_\phi - 1)(1+\phi).$$

To apply the leapfrog RMHMC algorithm, then we also need expressions for the metric tensor and its partial derivative with respect to the transformed parameter. The metric tensor $\mathbf{M}(\phi)$ for the log joint posterior density is calculated as follows:

$$\mathbf{M}(\phi) = 2\phi^2 + \frac{(T-1)(1-\phi^2)}{1-\rho^2} + (a_\phi + b_\phi - 2)(1-\phi^2).$$

Thus the partial derivative of metric tensor $\mathbf{M}(\phi)$ follows as

$$\frac{\partial\mathbf{M}(\phi)}{\partial\tilde{\phi}} = 2\phi(1-\phi^2)\left(4 - \frac{T-1}{1-\rho^2} - a_\phi - b_\phi\right).$$

Updating parameters $(\sigma_\epsilon, \sigma_\eta, \rho)$

On the basis of Eq. (5), the log joint posterior density of $\Sigma = (\sigma_\epsilon, \sigma_\eta, \rho)$ has the following expression:

$$
\begin{aligned}
\mathcal{L}(\Sigma) &= constant - T\ln(\sigma_\epsilon\sigma_\eta) - \frac{1}{2}\sum_{t=1}^{T}\widetilde{R}_t^2 - \frac{(1-\phi^2)\alpha_1^2}{2\sigma_\eta^2} \\
&\quad -\frac{T-1}{2}\ln(1-\rho^2) - \frac{1}{2\sigma_\eta^2(1-\rho^2)}\sum_{t=2}^{T}\widetilde{\alpha}_t^2 \\
&\quad - \left(a_{\sigma_\epsilon}+2\right)\ln\left(\sigma_\epsilon\right) - \frac{b_{\sigma_\epsilon}}{2\sigma_\epsilon^2} - \left(a_{\sigma_\eta}+2\right)\ln\left(\sigma_\eta\right) - \frac{b_{\sigma_\eta}}{2\sigma_\eta^2} \\
&\quad + (a_\rho - 1)\ln(1+\rho) + (b_\rho - 1)\ln(1-\rho),
\end{aligned}
$$

which is in non-standard form. The expressions required by the HMC and RMHMC algorithms to sample Σ are evaluated as follows.

For dealing with the constraints $-1 < \rho < 1$ and $\sigma_\epsilon, \sigma_\eta > 0$, it is necessary to implement the transformations $\rho = \tanh(\tilde{\rho})$, $\sigma_\epsilon = \exp(\tilde{\sigma}_\epsilon)$, and $\sigma_\eta = \exp(\tilde{\sigma}_\eta)$ and then introduce a Jacobian factor into the acceptance ratio. Notice that Girolami and Calderhead (2011) did not take a transformation for the parameter σ_ϵ. The derivatives of the transformations are $\mathrm{d}\rho/\mathrm{d}\tilde{\rho} = 1-\rho^2$, $\mathrm{d}\sigma_\epsilon/\mathrm{d}\tilde{\sigma}_\epsilon = \sigma_\epsilon$, and $\mathrm{d}\sigma_\eta/\mathrm{d}\tilde{\sigma}_\eta = \sigma_\eta$. Then, the partial derivatives of the above log joint posterior with respect to the transformed parameters are as follows:

$$
\begin{aligned}
\nabla_{\tilde{\sigma}_\epsilon}\mathcal{L}(\Sigma) &= -T + \sum_{t=1}^{T}\widetilde{R}_t^2 - \frac{\rho}{\sigma_\eta(1-\rho^2)}\sum_{t=2}^{T}\widetilde{\alpha}_t\widetilde{R}_{t-1} - \left(a_{\sigma_\epsilon}+2\right) + \frac{b_{\sigma_\epsilon}}{\sigma_\epsilon^2} \\
\nabla_{\tilde{\sigma}_\eta}\mathcal{L}(\Sigma) &= -T + \frac{(1-\phi^2)\alpha_1^2}{\sigma_\eta^2} + \frac{1}{\sigma_\eta^2(1-\rho^2)}\sum_{t=2}^{T}\widetilde{\alpha}_t^2 \\
&\quad + \frac{\rho}{\sigma_\eta(1-\rho^2)}\sum_{t=2}^{T}\widetilde{\alpha}_t\widetilde{R}_{t-1} - \left(a_{\sigma_\eta}+2\right) + \frac{b_{\sigma_\eta}}{\sigma_\eta^2} \\
\nabla_{\tilde{\rho}}\mathcal{L}(\Sigma) &= (T-1)\rho - \frac{\rho}{\sigma_\eta^2(1-\rho^2)}\sum_{t=2}^{T}\widetilde{\alpha}_t^2 + \frac{1}{\sigma_\eta}\sum_{t=2}^{T}\widetilde{\alpha}_t\widetilde{R}_{t-1} \\
&\quad + (a_\rho - 1)(1-\rho) - (b_\rho - 1)(1+\rho)
\end{aligned}
$$

To apply the leapfrog RMHMC algorithm, then we also need expressions for the metric tensor and its partial derivatives with respect to the transformed parameter. The metric tensor $\mathbf{M}(\sigma_\epsilon, \sigma_\eta, \rho)$ is a symmetric tridiagonal matrix, where the nonzero elements are calculated as follows:

$$
\begin{aligned}
\mathbf{M}(1,1) &= 2T + \frac{(T-1)\rho^2}{1-\rho^2} + \frac{2b_{\sigma_\epsilon}}{\sigma_\epsilon^2}, \\
\mathbf{M}(2,2) &= 2T + \frac{(T-1)\rho^2}{1-\rho^2} + \frac{2b_{\sigma_\eta}}{\sigma_\eta^2}, \\
\mathbf{M}(3,3) &= (T-1)(1+\rho^2) + (a_\rho + b_\rho - 2)(1-\rho^2),
\end{aligned}
$$

$$\mathbf{M}(1,2) = \mathbf{M}(2,1) = -(T-1)\frac{\rho^2}{1-\rho^2},$$
$$\mathbf{M}(1,3) = \mathbf{M}(3,1) = -(T-1)\rho,$$
$$\mathbf{M}(2,3) = \mathbf{M}(3,2) = -(T-1)\rho,$$

and its partial derivatives follow as

$$\frac{\partial \mathbf{M}(\Sigma)}{\partial \tilde{\sigma}_\epsilon} = \begin{bmatrix} -\frac{4b_{\sigma_\epsilon}}{\sigma_\epsilon^2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \frac{\partial \mathbf{M}(\Sigma)}{\partial \tilde{\sigma}_\eta} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\frac{4b_{\sigma_\eta}}{\sigma_\eta^2} & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\frac{\partial \mathbf{M}(\Sigma)}{\partial \tilde{\rho}} = -(T-1)\begin{bmatrix} -\frac{2\rho}{1-\rho^2} & \frac{2\rho}{1-\rho^2} & 1-\rho^2 \\ \frac{2\rho}{1-\rho^2} & -\frac{2\rho}{1-\rho^2} & 1-\rho^2 \\ 1-\rho^2 & 1-\rho^2 & \frac{2(a_\rho + b_\rho - T - 1)\rho(1-\rho^2)}{T-1} \end{bmatrix}.$$

Updating parameters σ_u^2 and c

In the steps 4 and 5, we sample from the conditional posterior density of σ_u^2 and c directly:

$$\sigma_u^2 | c, \boldsymbol{\alpha}, \mathbf{Y} \sim \mathcal{IG}\left(d_{\sigma_u}/2, D_{\sigma_u}/2\right),$$
$$c | \sigma_u^2, \boldsymbol{\alpha}, \mathbf{Y} \sim \mathcal{N}\left(M_c, V_c\right),$$

where

$$d_{\sigma_u} = a_{\sigma_u} + T, \quad D_{\sigma_u} = b_{\sigma_u} + \sum_{t=1}^{T} (Y_t - c - \alpha_t)^2,$$

$$V_c = \left[\frac{1}{v_c} + \frac{T}{\sigma_u^2}\right]^{-1}, \quad \text{and} \quad M_c = V_c\left[\frac{m_c}{v_c} + \frac{1}{\sigma_u^2}\sum_{t=1}^{T}(Y_t - \alpha_t)\right].$$

Appendix 2: Estimation of RASVt model

Let us take the logarithm of the full conditional posterior density defined in Eq. (4):

$$\ln p(\theta_2, \mathbf{h}|\mathbf{R}, \mathbf{Y}) = constant + \sum_{t=1}^{T}\left(-\frac{1}{2}h_t - \frac{1}{2}\ln(z_t) - \frac{1}{2}\widetilde{R}_t^2\right) - T\ln(\sigma_\epsilon \sigma_u)$$
$$-\frac{1}{2\sigma_u^2}\sum_{t=1}^{T}(Y_t - \xi - h_t)^2 - T\ln(\sigma_\eta) + \frac{1}{2}\ln(1-\phi^2)$$

$$
\begin{aligned}
&-\frac{T-1}{2}\ln(1-\rho^2)-\frac{1-\phi^2}{2\sigma_\eta^2}\widetilde{h}_1^2-\frac{1}{2\sigma_\eta^2(1-\rho^2)}\sum_{t=2}^{T}\widetilde{h}_t^2\\
&+\frac{\nu}{2}T\ln\left(\frac{\nu}{2}\right)-T\ln\Gamma\left(\frac{\nu}{2}\right)-\sum\nolimits_{t=1}^{T}\left[\left(\frac{\nu}{2}+1\right)\ln(z_t)+\frac{\nu}{2}z_t^{-1}\right]\\
&-\frac{1}{2}\ln(v_\xi)-\frac{(\xi-m_\xi)^2}{2v_\xi}-\left(\frac{a_{\sigma_u}}{2}+1\right)\ln\left(\sigma_u^2\right)-\frac{b_{\sigma_u}}{2\sigma_u^2}\\
&-\frac{1}{2}\ln(v_\mu)-\frac{(\mu-m_\mu)^2}{2v_\mu}+(a_\phi-1)\ln(1+\phi)+(b_\phi-1)\ln(1-\phi)\\
&-\left(\frac{a_{\sigma_\eta}}{2}+1\right)\ln\left(\sigma_\eta^2\right)-\frac{b_{\sigma_\eta}}{2\sigma_\eta^2}+(a_\rho-1)\ln(1+\rho)\\
&+(b_\rho-1)\ln(1-\rho)+(a_\nu-1)\ln(\nu)-b_\nu\nu
\end{aligned}
\tag{6}
$$

where

$$
\widetilde{R}_t=R_t\sigma_\epsilon^{-1}e^{-\frac{1}{2}h_t}z_t^{-\frac{1}{2}},\quad \text{for } t=1,\ldots,T, \text{ and} \tag{7}
$$

$$
\widetilde{h}_t=\begin{cases} h_1-\mu, & \text{for } t=1;\\ h_t-\mu-\phi(h_{t-1}-\mu)-\rho\sigma_\eta\widetilde{R}_{t-1}, & \text{for } t=2,\ldots,T.\end{cases} \tag{8}
$$

On the basis of the above density, the full conditional densities of $\mathbf{h}$, ξ, σ_u, ϕ, and σ_η are similar to $\boldsymbol{\alpha}$, c, σ_u, ϕ, and σ_η, respectively, for the RASV model. Thus, we follow the same sampling procedure described in the previous section with different specifications of $\widetilde{R}_t$ and $\widetilde{h}_t$ defined in Eqs. (7) and (8), respectively. To sample parameters μ and ν and latent variable $\mathbf{z}$, we describe these sampling procedures as follows.

Updating parameter μ

On the basis of Eq. (6), parameter μ can be sampled directly from its conditional posterior density:

$$
\mu|\phi,\sigma_\eta,\rho,\mathbf{h},\mathbf{z},\mathbf{R}\sim\mathcal{N}\left(M_\mu,V_\mu\right),
$$

where

$$
\begin{aligned}
V_\mu&=\left[\frac{1}{v_\mu}+\frac{(1-\rho^2)(1-\phi^2)+(T-1)(1-\phi)^2}{\sigma_\eta^2(1-\rho^2)}\right]^{-1},\\
M_\mu&=V_\mu\left[\frac{m_\mu}{v_\mu}+\frac{(1-\rho^2)(1-\phi^2)h_1+(1-\phi)\sum_{t=2}^{T}(h_t-\phi h_{t-1}-\rho\sigma_\eta\widetilde{R}_{t-1})}{\sigma_\eta^2(1-\rho^2)}\right].
\end{aligned}
$$

Updating parameter ν

The log posterior of ν and its gradient vector required in both leapfrog-HMC and -RMHMC algorithms are given by

$$\mathcal{L}(\nu) = constant + \frac{\nu}{2} T \ln\left(\frac{\nu}{2}\right) - T \ln \Gamma\left(\frac{\nu}{2}\right) - \frac{\nu}{2} \sum\nolimits_{t=1}^{T} \left[\ln(z_t) + z_t^{-1}\right] \\ +(a_\nu - 1)\ln(\nu) - b_\nu \nu$$

and

$$\nabla_\nu \mathcal{L}(\nu) = \frac{1}{2} T \left[\ln\left(\frac{\nu}{2}\right) + 1\right] - \frac{1}{2} T \Psi\left(\frac{\nu}{2}\right) - \frac{1}{2} \sum_{t=1}^{T} \left[\ln(z_t) + z_t^{-1}\right] + \frac{a_\nu - 1}{\nu} - b_\nu.$$

respectively, where $\Psi(x)$ is a digamma function defined by $\Psi(x) = d \ln \Gamma(x)/dx$. In the implementation of leapfrog-RMHMC algorithm, the metric tensor and its partial derivatives are calculated as follows:

$$\mathbf{M}(\nu) = -\frac{T}{2\nu} + \frac{T}{4} \Psi'\left(\frac{\nu}{2}\right) + \frac{a_\nu - 1}{\nu^2}$$

and

$$\frac{\partial \mathbf{M}}{\partial \nu} = \frac{T}{2\nu^2} + \frac{T}{8} \Psi''\left(\frac{\nu}{2}\right) - \frac{2(a_\nu - 1)}{\nu^3},$$

respectively, where $\Psi'(x)$ is a trigamma function defined by $\Psi'(x) = d\Psi(x)/dx$ and $\Psi''(x)$ is a tetragamma function defined by $\Psi''(x) = d\Psi'(x)/dx$.

Updating the latent variable $\mathbf{z}$

The logarithm of the full conditional posterior density of the latent variable $\mathbf{z}$ is

$$\mathcal{L}(\mathbf{z}) = constant + \sum_{t=1}^{T} \left[-\frac{1}{2} \widetilde{R}_t^2 - \frac{\nu + 3}{2} \ln(z_t) - \frac{\nu}{2z_t}\right] - \frac{1}{2\sigma_\eta^2 (1 - \rho^2)} \sum_{t=2}^{T} \widetilde{h}_t^2.$$

From this conditional density, draws can be obtained by implementing the transformation $z_t = \exp(\hat{z}_t)$ to ensure constrained sampling for $z_t > 0$. Gradient vector of the log posterior required in both leapfrog-HMC and -RMHMC algorithms is then evaluated as follows:

$$\nabla_{\hat{z}_t} \mathcal{L}(z_t) = -\widetilde{R}_t \frac{\partial \widetilde{R}_t}{\partial \hat{z}_t} - \frac{\nu + 3}{2} + \frac{\nu}{2z_t} + \frac{\rho}{\sigma_\eta (1 - \rho^2)} \widetilde{h}_{t+1} \frac{\partial \widetilde{R}_t}{\partial \hat{z}_t} I_{t<T},$$

where I is an indicator function and

$$\frac{\partial \widetilde{R}_t}{\partial \hat{z}_t} = -\frac{1}{2} R_t \sigma_\epsilon^{-1} e^{-\frac{1}{2}h_t} z_t^{-\frac{1}{2}}.$$

Furthermore, the metric tensor $\mathbf{M}(\mathbf{z})$ required to apply the leapfrog-RMHMC algorithm is a diagonal matrix whose diagonal entries are given by

$$\mathbf{M}(t, t) = \frac{\nu + 1}{2} + \frac{\rho^2}{4(1 - \rho^2)} I_{t<T}.$$

Since the above metric tensor is fixed, that is not a function of the latent variable $\mathbf{z}$, the associated partial derivatives with respect to the transformed latent variable are zero. This implies that both metric tensor and its inverse are not recalculated in the generalized leapfrog integration scheme.

References

Andersen TG, Bollerslev T, Diebold FX, Labys P (2001) The distribution of realized exchange rate volatility. J Am Stat Assoc 96:42–55

Asai M (2005) Comparison of MCMC methods for estimating stochastic volatility models. Comput Econ 25:281–301

Black F (1976) Studies of stock market volatility changes. Am Stat Assoc, Bus Econ Stat Sect 1:177–181

Broto C, Ruiz E (2004) Estimation methods for stochastic volatility models: a survey. J Econ Surv 18: 613–649

Dobrev D, Szerszen P (2010) The information content of high-frequency data for estimating equity return models and forecasting risk. Working Paper, Finance and Economics Discussion Series. Retrieved from http://www.federalreserve.gov/pubs/feds/2010/201045/

Duane S, Kennedy AD, Pendleton B, Roweth D (1987) Hybrid Monte Carlo. Phys Lett B 195:216–222

Geweke J (1992) Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: Bernardo JM, Berger JO, Dawid AP, Smith AFM (eds) Bayesian statistics 4. Clarendon Press, Oxford, pp 169–194 with discussion

Girolami M, Calderhead B (2011) Riemann manifold Langevin and Hamiltonian Monte Carlo methods. J R Stat Soc, Ser B 73:1–37

Hansen PR, Lunde A (2005) A forecast comparison of volatility models: does anything beat a GARCH(1,1). J Appl Econ 20:873–889

Harvey AC, Shephard N (1996) The estimation of an asymmetric stochastic volatility model for asset returns. J Bus Econ Stat 14:429–434

Ishwaran H (1999) Applications of Hybrid Monte Carlo to Bayesian generalized linear models: quasicomplete separation and neural networks. J Comput Graph Stat 8:779–799

Jacquier E, Polson NG (2011) Bayesian methods in finance. In: Geweke J, Koop G, van Dijk H (eds) Handbook of Bayesian econometrics. Oxford University Press, Oxford, pp 439–512

Jacquier E, Polson NG, Rossi PE (1994) Bayesian analysis of stochastic volatility models. In: Shephard N (ed) Stochastic volatility: selected readings. Oxford University Press, New York, pp 247–282

Jacquier E, Polson NG, Rossi PE (2004) Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. J Econ 122:185–212

Kim S, Shephard N, Chib S (1998) Stochastic volatility: likelihood inference and comparison with ARCH models. In: Shephard N (ed) Stochastic volatility: selected readings. Oxford University Press, New York

Koopman SJ, Scharth M (2013) The analysis of stochastic volatility in the presence of daily realized measures. J Financ Econ 11:76–115

Leimkuhler B, Reich S (2004) Simulating Hamiltonian dynamics. Cambridge University Press, Cambridge

Marwala T (2012) Condition monitoring using computational intelligence methods. Springer, Berlin

McLachlan RI, Atela P (1992) The accuracy of symplectic integrators. Nonlinearity 5:541–562

Neal RM (2010) MCMC using Hamiltonian dynamics. In: Brooks S, Gelman A, Jones G, Meng X-L (eds) Handbook of Markov Chain Monte Carlo. Chapman & Hall/CRC Press, Boca Raton, pp 113–162

Omori Y, Chib S, Shephard N, Nakajima J (2007) Stochastic volatility model with leverage: fast likelihood inference. J Econ 140:425–449

Omori Y, Watanabe T (2008) Block sampler and posterior mode estimation for asymmetric stochastic volatility models. Comput Stat Data Anal 52:2892–2910

Pasarica C, Gelman A (2008) Adaptive scaling the metropolis algorithm using expected squared jumped distance. Stat Sin 20:343–364

Shephard N (1993) Fitting non-linear time series models, with applications to stochastic variance models. J Appl Econ 8:135–152

Shephard N, Pitt MK (1997) Likelihood analysis of non-Gaussian measurement timeseries. Biometrika 84:653–667

Takahashi M, Omori Y, Watanabe T (2009) Estimating stochastic volatility models using daily returns and realized volatility simultaneously. Comput Stat Data Anal 53:2404–2426

Takahashi M, Omori Y, Watanabe T (2014) Volatility and quantile forecasts by realized stochastic volatility models with generalized hyperbolic distribution. Working Paper Series CIRJE-F-921, CIRJE, Faculty of Economics, University of Tokyo. Retrieved from http://www.cirje.e.u-tokyo.ac.jp/research/dp/2014/2014cf921

Takaishi T (2009) Bayesian inference of stochastic volatility by Hybrid Monte Carlo. J Circuits Syst Comput 18:1381–1396

Taylor SJ (1982) Financial returns modelled by the product of two stochastic processes–a study of the daily sugar prices 1961–75. In: Shephard N (ed) Stochastic volatility: selected readings. Oxford University Press, New York, pp 60–82

Watanabe T, Omori Y (2004) A multi-move sampler for estimating non-Gaussian times series models: comments on Shephard and Pitt. Biometrika 91:246–248