

Pricing Options under Stochastic Volatility Models Using Bayesian Estimation via Hamiltonian Monte Carlo Method

Hidetoshi Mitsui*

1 Introduction

In recent years, the Hamiltonian Monte Carlo (HMC) method, a type of Markov chain Monte Carlo (MCMC) method based on Bayesian statistics, has attracted significant attention.¹⁾ The HMC method employs molecular dynamics techniques developed in physics for updating parameters in statistical models²⁾, and employs the Metropolis-Hastings (M-H) method for accepting or rejecting updated parameters. Because HMC combines two distinct methods for molecular dynamics and M-H, it is also referred to as hybrid Monte Carlo. Unlike the M-H method and Gibbs sampling method commonly used in MCMC methods, the HMC method has the characteristic that the adoption probability does not decrease even when many parameters are sampled at once. In the molecular dynamics method used to update the parameters of the HMC method, Hamilton's equations of motion are used. Since Hamilton's equations of motion cannot be solved analytically, they must be solved approximately using numerical integration. In order for the HMC method to satisfy the detailed balance condition, one of the basic conditions of the MCMC method, this numerical integration must satisfy time reversal and volume conservation. A numerical integration method that satisfies these two conditions is called a symplectic integrator. In the HMC method, leapfrog integration, which is a second-order symplectic integrator, is used. In leapfrog integration, there are two parameters that represent the time advanced in a single update: step size and number of steps. These parameters affect the stability of Hamilton's equations of motion and the efficiency of sampling, respectively.

Time-series data on risk asset returns such as stocks, stock indices, and foreign exchange rates are characterized by fluctuations in the second moment, known as volatility. In analyzing such risk asset return data, the Stochastic Volatility (SV) model is widely used. However, having latent variables in the SV model makes maximum likelihood estimation of the model parameters difficult, necessitating alternative estimation methods. To address this issue, many previous studies have employed Bayesian estimation using MCMC methods.³⁾ Additionally, when using the SV model, it is necessary to simultaneously sample not only the parameters describing the model but also the latent variable volatility from the posterior distribution. Therefore, the volatility becomes the same number as the number of observations, making efficient sampling crucial. In the aforementioned prior studies using SV models with MCMC methods, these MCMC included the M-H method and Gibbs sampling. However, such MCMC methods suffer from reduced rejection rates, among others, when estimating many random variables

* College of Economics, Nihon University, e-mail: mitsui.hidetoshi@nihon-u.ac.jp

simultaneously. Therefore, Takaishi (2008, 2009, 2013) proposed Bayesian estimation using hybrid Monte Carlo methods for SV model estimation. Nugroho and Morimoto (2015) conducted empirical analyses on the TOPIX and Standard & Poor's 500 Stock Index using the SV model, employing the Riemann manifold Hamiltonian Monte Carlo method, an improved version of the HMC method⁴⁾. Mitsui and Totsuka (2022) conducted empirical research by estimating the SV model using Bayesian estimation via the HMC method with Nikkei Stock Average data.

Furthermore, in option pricing theory, the volatility of the underlying asset's return plays a crucial role and is a parameter highly sensitive to option prices. Therefore, the SV model can be applied to option theory. In option theory based on the SV model, it is often assumed that volatility fluctuations exhibit mean reversion and that the logarithm of volatility undergoes stochastic fluctuations. Heston (1993) provided a closed-form solution for the SV option, but because this solution is a continuous-type model, parameter estimation is not possible, making empirical research impossible. Much of the empirical research to date has estimated the parameters of the SV model by some method and then performed option valuation via Monte Carlo simulation using the estimated parameters. The problem with such an approach is that it does not account for parameter estimation error. Therefore, the present paper proposes Bayesian estimation using HMC for option pricing based on SV models.

The remainder of this paper is organized as follows: Section 2 describes option pricing using a continuous-time SV model and its closed-form solution. Section 3 explains discrete-time approximation of continuous-time SV models. Section 4 introduces an SV model with asymmetry and an SV model with a fat-tails distribution. Section 5 describes Bayesian inference with the HMC method. Section 6 proposes estimation of SV options using the HMC method. Section 7 presents conclusions.

2 Option Pricing Using Continuous-Time SV Model

2.1 SV Options

This section explains option valuation using the continuous-time stochastic volatility model developed by Heston (1993). First, we consider option pricing using the continuous-time SV model and review the characteristics of its parameters. The underlying asset price S is assumed to obey the following stochastic differential equation⁵⁾:

$$dS = \mu S dt + \sigma S dz_1. \quad (2.1)$$

Here, μ is the drift term, dt is the infinitesimal change in time, σ is the volatility⁶⁾, and dz_1 represents the Wiener process. Furthermore, suppose a stochastic process with volatility σ follows the following stochastic differential equation, which is an Ornstein-Uhlenbeck process:

$$d\sigma = -\beta\sigma dt + \delta dz_2. \quad (2.2)$$

Here, β and δ represent the coefficients of variation. By Ito's lemma, the variance σ^2 is assumed to obey the following stochastic differential equation:

$$d\sigma^2 = [\delta^2 - 2\beta\sigma^2]dt + 2\delta\sigma dz_2. \quad (2.3)$$

Equation (2.3) can be approximated by the following square root process⁷⁾:

$$d\sigma^2 = \kappa[\theta - \sigma^2]dt + \delta\sigma dz_2. \quad (2.4)$$

Here, θ represents the long-term mean of the reversion⁸⁾, and κ represents the adjustment speed of the reversion. Put another way, volatility fluctuates toward the long-term mean θ , whose rate of mean reversion is determined by κ . Therefore, when the mean volatility θ increases, option prices also rise. Furthermore, let dz_2 be correlated with dz_1 by a factor of ρ . To assume a constant interest rate r , we define the price at time t of a unit of a discount bond maturing at time $t + \tau$ as follows:

$$B(t, t + \tau) = e^{-r\tau}. \quad (2.5)$$

The assumptions in Eqs. (2.1)–(2.5) are insufficient for option pricing; volatility risk must be considered. From the asset equilibrium model of Cox *et al.* (1985), the value of the option $U(S, \sigma^2, t)$ must satisfy the following stochastic differential equation:

$$\begin{aligned} \frac{1}{2}\sigma^2 S^2 + \frac{\partial^2 U}{\partial S^2} + \rho\delta\sigma^2 S \frac{\partial^2 U}{\partial S \partial \sigma^2} + \frac{1}{2}\delta^2\sigma^2 \frac{\partial^2 U}{\partial \sigma^4} + rS \frac{\partial U}{\partial S} \\ + \{\kappa[\theta - \sigma^2] - \lambda(S, \sigma^2, t)\} \frac{\partial U}{\partial \sigma^2} - rS + \frac{\partial U}{\partial t} = 0. \end{aligned} \quad (2.6)$$

Here, $\lambda(S, \sigma^2, t)$ represents the risk premium associated with volatility. From Breeden's (1979) consumption-based model, it satisfies the following relationship:

$$\lambda(S, \sigma^2, t)dt = \gamma Cov \left[d\sigma^2, \frac{dc}{c} \right] \quad (2.7)$$

Here, C represents the consumption rate, and γ denotes the investor's relative risk aversion. The consumption process is assumed to obey the following stochastic differential equation:

$$dC = \mu_c C dt + \delta_c \sigma C dz_3. \quad (2.8)$$

Consumption growth exhibits a certain correlation with the returns on real assets. Since the risk premium is proportional to σ^2 , the following holds:

$$\lambda(S, \sigma^2, t) = \lambda\sigma^2. \quad (2.9)$$

The risk premium shall be expressed in the form described above.

2.2 Closed-Form Solution for SV Options

Heston (1993) solved the partial differential equation (2.6) by imposing the following appropriate boundary conditions when the underlying asset price and volatility follow Eqs. (2.1) and (2.4), thereby providing a closed-form solution for the call option price $C(S, \sigma^2, t)$:

$$\begin{aligned} C(S, \sigma^2, T) &= \text{Max}(0, S - K), \\ C(0, \sigma^2, t) &= 0, \\ \frac{\partial C}{\partial S}(\infty, \sigma^2, t) &= 1, \\ rS \frac{\partial C}{\partial S}(S, 0, t) + \kappa\theta \frac{\partial C}{\partial \sigma^2}(S, 0, t) - r(S, 0, t) + \frac{\partial C}{\partial t}(S, 0, t) &= 0, \\ C(S, \infty, t) &= S. \end{aligned}$$

By analogy with the Black-Scholes formula⁹⁾, we can infer the following closed-form solution:

$$C(S, \sigma^2, t) = SP_1 - KP(t, T)P_2. \quad (2.10)$$

Here, SP_1 represents the present value of the underlying asset upon optimal exercise, and $KP(t, T)P_2$ denotes the present value of the strike price. Taking the logarithm of the current price,

$$x = \ln[S] \quad (2.11)$$

Using this definition, P_1 and P_2 satisfy the following partial differential equations:

$$\begin{aligned} \frac{1}{2}\sigma^2 \frac{\partial^2 P_j}{\partial x^2} + \rho\delta\sigma^2 \frac{\partial^2 P_j}{\partial x \partial \sigma^2} + \frac{1}{2}\delta^2\sigma^2 \frac{\partial^2 P_j}{\partial \sigma^4} + (r + u_j\sigma^2) \frac{\partial P_j}{\partial x} \\ + (a_j - b_j\sigma^2) \frac{\partial P_j}{\partial \sigma^2} + \frac{\partial P_j}{\partial t} = 0, j = 1, 2. \end{aligned} \quad (2.12)$$

Here,

$$u_1 = \frac{1}{2}, u_2 = -\frac{1}{2}, a = \kappa\theta, b_1 = \kappa + \lambda - \rho\delta, b_2 = \kappa + \lambda.$$

The initial conditions are as follows:

$$P_j(x, \sigma^2, T; \ln K) = 1_{\{x \geq \ln K\}}. \quad (2.13)$$

Here, x obeys the following process:

$$dx_t = [r + u_j\sigma^2]dt + \sigma dz_1, \quad (2.14)$$

$$d\sigma^2 = (a_j - b_j\sigma^2)dt + \delta\sigma dz_2. \quad (2.15)$$

Let P_j denote the conditional probability that the option is exercised in-the-money,

$$P_j(x, \sigma^2, T; \ln [K]) = Pr[x(T) \geq \ln [K] | x(t) = x, \sigma^2(t) = \sigma^2]. \quad (2.16)$$

However, it is not useful in closed-form. Therefore, by utilizing¹⁰⁾ $f_1(x, \sigma^2, T; \phi)$ and $f_2(x, \sigma^2, T; \phi)$ ¹¹⁾, it satisfies partial differential equation (2.12) under the following terminal condition constraint:

$$f_j(x, \sigma^2, T; \phi) = e^{i\phi x}. \quad (2.17)$$

Using the above definitions, the solution to the characteristic function is given as follows¹²⁾:

$$f_j(x, \sigma^2, T; \phi) = e^{C(T-t; \phi) + D(T-t; \phi)\sigma^2 + i\phi x}. \quad (2.18)$$

Here,

$$\begin{aligned} C(\tau; \phi) &= r\phi i\tau + \frac{a}{\delta^2} \left\{ (b_j - \rho\delta\phi i + d)\tau - 2 \ln \left[\frac{1 - ge^{d\tau}}{1 - g} \right] \right\}, \\ D(\tau; \phi) &= \frac{b_j - \rho\delta\phi i + d}{\delta^2} \left[\frac{1 - e^{d\tau}}{1 - ge^{d\tau}} \right], \\ g &= \frac{b_j - \rho\delta\phi i + d}{b_j - \rho\delta\phi i - d}, \end{aligned}$$

$$d = \sqrt{(\rho\delta\phi i - b_j)^2 - \delta^2(2\mu_j\phi i - \phi^2)}.$$

Therefore, by inverting the characteristic function as follows, we can obtain the original probability (Fourier inversion):

$$P_j(x, \sigma^2, T; \ln [K]) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \operatorname{Re} \left[\frac{e^{-i\phi \ln [K]} f_j(x, \sigma^2, T; \phi)}{i\phi} \right] d\phi. \quad (2.19)$$

Here, $\operatorname{Re}[\cdot]$ denotes the real part of a complex number. The solution for a European call option is obtained from Eqs. (2.10), (2.18), and (2.19).

Heston (1993) obtained the following simulation results for the given parameters. (i) The correlation coefficient ρ influences the skewness of the underlying asset's return. When $\rho > 0$, a high value of σ widens the right tail of the underlying asset's return probability density, but a low value of σ does not widen the left tail. When $\rho < 0$, the opposite result occurs compared to when $\rho > 0$. (ii) δ controls the volatility of volatility. (iii) When $\delta = 0$, volatility is constant. Fluctuations in σ increase the kurtosis of the underlying asset's return. (iv) Fluctuations in σ increase the price of far-in-the-money options and decrease the price of near-the-money options. They have a greater impact on the price of in-the-money options compared to out-of-the-money options.

The results above demonstrate that the correlation coefficient between the underlying asset return and volatility σ , as well as fluctuations in volatility σ , are critical factors in option pricing. However, as stated above, Heston's (1993) model is continuous in time, making parameter estimation impossible and preventing its direct use in empirical research. Therefore, the next section explains the approximation of the continuous-time model to a discrete-time model.

3 Discrete-Time Approximation of Continuous-Time SV Models

3.1 Discrete Approximation of Simple SV Models

First, consider a simple continuous-time SV model. The underlying asset price follows the process described by Eq. (2.1). Furthermore, the variation of σ^2 is formulated as a continuous-time process as follows:

$$d\sigma^2 = \psi\sigma^2 dt + \delta\sigma^2 dz_2. \quad (3.1)$$

Here, the Wiener process dz_2 may or may not be correlated with dz_1 in Eq. (2.1), corresponding to $\rho \neq 0$ and $\rho = 0$, respectively. In order to conduct an empirical study, we need to convert the continuous-time model in Eqs. (2.1) and (3.1) into a discrete-time model. The discrete approximations of Eqs. (2.1) and (3.1) are as follows:

$$y_t = \psi \exp\left(\frac{h_t}{2}\right) u_t, \quad t = 1, \dots, n, \quad (3.2)$$

$$h_t = \beta h_{t-1} + \eta_t, \quad t = 1, \dots, n, \quad (3.3)$$

$$\begin{pmatrix} u_t \\ \eta_t \end{pmatrix} \sim i.i.d. \mathcal{N} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & \sigma_\eta^2 \end{bmatrix} \right).$$

Here, y_t is the rate of return, volatility $\sigma_t^2 = \exp(h_t/2)$ with $h_t = \ln \sigma_t^2$ is the variance of rate of return y_t , and ψ is a scale parameter. u_t has mean 0 and variance 1 and η_t has mean 0 and variance σ_η^2 , and these error terms each follow a normal distribution. Prefix “*i.i.d.*” denotes independent and identically distributed. As shown, u_t and η_t are assumed to be uncorrelated or correlated. In the following, we simply refer to the discrete approximation SV model as the SV model. When conducting empirical research, for example, let S_t denote the Dow or Nikkei Stock Average at time t . The return y_t at time t is given by

$$y_t = (\ln S_t - \ln S_{t-1}) \times 100. \quad (3.4)$$

Similarly, if Index_t is the level of the S&P 500 or TOPIX at time t , the return y_t at time t is given by

$$y_t = (\ln \text{Index}_t - \ln \text{Index}_{t-1}) \times 100. \quad (3.5)$$

In addition, most studies of continuous-time SV models for risky asset prices have formulated the variation of σ^2 as the following continuous-time stochastic process with the mean-reverting property like the Ornstein-Uhlenbeck process:

$$y_t = \exp\left(\frac{h_t}{2}\right) u_t, \quad t = 1, \dots, n, \quad (3.6)$$

$$h_{t+1} = \mu + \phi(h_t - \mu) + \eta_t, \quad t = 0, \dots, n-1, \quad (3.7)$$

$$\begin{pmatrix} u_t \\ \eta_t \end{pmatrix} \sim i.i.d. \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & \sigma_\eta^2 \end{bmatrix}\right). \quad (3.8)$$

Here, h_t is a latent variable indicating the variation of volatility, μ is the mean volatility, and ϕ is a parameter indicating the persistence of shocks in volatility. For stationarity, we assume that $|\phi| < 1$. Equation (3.7) shows that the latent variable h_t follows a first-order autoregressive process.

Well-known estimation methods for nonlinear models such as SV models include the moment method (MM), generalized moment method (GMM), maximum likelihood estimation (ML), and Bayesian estimation.

3.2 Option Pricing Using Monte Carlo Simulation

Here, we explain a method for valuing option prices using Monte Carlo simulation under the assumption that investors are risk neutral. Taking the utility function of a typical investor as logarithmic, in Eq. (2.9), $\lambda\sigma^2 = 0$ holds, and the call option price can be obtained by applying appropriate boundary conditions to the partial differential equation (2.6). However, as explained in Section 2, even when closed-form solutions are obtained, such as in Heston (1993), the parameters of the return process and the volatility process cannot be estimated. Consequently, option prices cannot be estimated empirically. Therefore, assuming that representative investors have risk-neutral preferences, we utilize the fact that, since there is no risk premium, the expected return on an option equals the expected return on the risk-free asset. As a result, when investors are risk-neutral, the price of a European option is the discounted present value of the expected value of the option price at maturity, discounted at the risk-free interest rate r^{13} .

Let t denote the current time, T denote the option's expiration date, and C_t denote the price of a call option with strike price K at time t . Then C_t is expressed as follows:

$$\begin{aligned}
C_t &= e^{-r(T-t)} \int_K^\infty [\text{Max}(S_T - K, 0)] p(S_T | S, \sigma^2) dS_T \\
&= e^{-r(T-t)} \hat{E} [\text{Max}(S_T - K, 0)].
\end{aligned} \tag{3.9}$$

Here, \hat{E} denotes the expected value in a risk-neutral world, S_t represents the underlying asset price at maturity, and $p(\cdot)$ denotes the conditional probability density function. Similarly, the put option price P_t is expressed by the following equation:

$$P_t = e^{-(T-t)r} \hat{E} [\text{Max}(K - S_T, 0)]. \tag{3.10}$$

The expected value on the right-hand side is generally evaluated using Monte Carlo simulation. After performing the simulation N times and obtaining N maturity asset prices S_T , these are denoted as $(S_T^{(1)}, S_T^{(2)}, \dots, S_T^{(N)})$. However, $S_T^{(i)}$ is the maturity price of the underlying asset obtained by the occurrence of the i -th path. When N is sufficiently large, by the law of large numbers, Eqs. (3.9) and (3.10) can each be evaluated by applying the following expressions:

$$C_t^{MS} = e^{-(T-t)r} \frac{1}{N} \sum_{i=1}^N \text{Max}(S_T^{(i)} - K, 0), \tag{3.11}$$

$$P_t^{MS} = e^{-(T-t)r} \frac{1}{N} \sum_{i=1}^N \text{Max}(K - S_T^{(i)}, 0). \tag{3.12}$$

Estimate the SV model parameters using the underlying asset return data from y_1 to y_T . Then, based on the estimated parameters, calculate the underlying asset price $S_T^{(i)}$ at maturity T via Monte Carlo simulation. Finally, derive the option prices C_t^{MS} , P_t^{MS} as the expected discounted present value¹⁴⁾.

3.3 SV Model and Likelihood

This paper employs Bayesian estimation methods for parameter estimation and option evaluation, so we will briefly explain the likelihood function. For example, in a simple SV model, representing the unknown parameters (ϕ, σ_η, μ) collectively as θ , the likelihood function of the SV model can be expressed as follows:

$$\begin{aligned}
L(\theta) &= \int \cdots \int f(\mathbf{y} | \mathbf{h}) f(\mathbf{h} | \theta) d\mathbf{h} \\
&= \int \cdots \int \prod_{t=1}^T \frac{1}{\sqrt{2\pi \exp(h_t)}} \exp\left[-\frac{y_t^2}{2 \exp(h_t)}\right] \\
&\quad \times \prod_{t=1}^{T-1} \frac{1}{\sqrt{2\pi \sigma_\eta^2}} \exp\left[-\frac{\{h_{t+1} - \mu - \phi(h_t - \mu)\}^2}{2\sigma_\eta^2(1 - \rho^2)}\right] \\
&\quad \times \frac{\sqrt{1 - \phi^2}}{\sqrt{2\pi \sigma_\eta^2}} \exp\left[-\frac{(1 - \phi^2)(h_1 - \mu)^2}{2\sigma_\eta^2}\right] dh_1 \cdots dh_T.
\end{aligned} \tag{3.13}$$

When deriving option prices using Bayesian estimation with HMC, the likelihood function is used for parameter estimation.

4 Extension of SV Model

Bayesian estimation using the HMC method remains applicable even when extending the SV model. Therefore, this section explains asymmetric SV models that account for fluctuations in stock prices and risk asset prices, as well as SV models where the error term follows a t -distribution.

4.1 SV Model with Asymmetry

It is known that there is asymmetry in the relationship between the rate of return and the volatility in the stock market. In particular, when the rate of return falls, the volatility tends to rise in the next period, and when the rate of return rises, the volatility tends to fall in the next period. This suggests that there is negative correlation between the rate of return and the volatility. This phenomenon is referred to as the asymmetric SV model or the leverage effect. To include such asymmetry in the model, we can consider adding the correlation between u_t and η_t to the model in Eqs. (3.6) and (3.7). If we construct the SV model assuming that u_t and η_t have correlation ρ , this is expressed as follows:

$$\begin{pmatrix} u_t \\ \eta_t \end{pmatrix} \sim i.i.d.\mathcal{N} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho\sigma_\eta \\ \rho\sigma_\eta & \sigma_\eta^2 \end{bmatrix} \right). \quad (4.1)$$

The model consisting of Eqs. (3.6), (3.7), and (4.1) is called the SV with leverage (SVL) model in the present paper. When the unknown parameters $(\phi, \sigma_\eta, \rho, \mu)$ of the SVL model are collectively represented by θ , the likelihood function $f(\mathbf{y}|\theta)$ of the SVL model can be expressed as follows:

$$\begin{aligned} f(\mathbf{y}|\theta) &= \int \cdots \int \prod_{t=1}^n \frac{1}{\sqrt{2\pi \exp(h_t)}} \exp \left[-\frac{y_t^2}{2 \exp(h_t)} \right] \\ &\times \prod_{t=1}^{n-1} \frac{1}{\sqrt{2\pi\sigma_\eta^2} \sqrt{1-\rho^2}} \exp \left[-\frac{\{h_{t+1} - \mu - \phi(h_t - \mu) - \rho\sigma_\eta y_t \exp(-h_t/2)\}^2}{2\sigma_\eta^2(1-\rho^2)} \right] \\ &\times \frac{\sqrt{1-\phi^2}}{\sqrt{2\pi\sigma_\eta^2}} \exp \left[-\frac{(1-\phi^2)(h_1 - \mu)^2}{2\sigma_\eta^2} \right] dh_1 \cdots dh_n. \end{aligned} \quad (4.2)$$

When deriving option prices using Bayesian estimation with HMC, the above expression is used for parameter estimation.

4.2 SV Model with Fat-Tails Distribution

The distribution of the rate of return on risky assets has long been known to be thicker than the normal distribution, as pointed out by Mandelbrot (1963) and Fama (1965). In order to consider an error distribution with a thick base, we introduce a random variable that follows the t -distribution. In this case, the process of the rate of return and volatility can be expressed as follows:

$$y_t = \exp\left(\frac{h_t}{2}\right) \sqrt{z_t} u_t, \quad t = 1, \dots, n, \quad (4.3)$$

$$h_{t+1} = \mu + \phi(h_t - \mu) + \eta_t, \quad t = 0, \dots, n-1, \quad (4.4)$$

$$z_t \sim i.i.d.IG\left(\frac{\nu}{2}, \frac{\nu}{2}\right). \quad (4.5)$$

Here, IG denotes the inverse gamma distribution, and the error term (u_t, η_t) follows Eq. (3.3) or (4.1). The $\sqrt{z_t}u_t$ term follows the t -distribution $z_t \sim i.i.d.IG(\nu/2, \nu/2)$, which has ν degrees of freedom. In the present paper, the model consisting of Eqs. (3.8) and (5.9)–(5.11) is called the SV- t model, and the model consisting of Eqs. (4.1) and (4.3)–(4.5) is called the SVL- t model.

When the unknown parameters of the SVL- t model $(\phi, \sigma_\eta, \rho, \mu, \nu)$ are collectively represented by θ , the likelihood function of the SVL- t model can be expressed as follows:

$$\begin{aligned} L(\theta) &= \int \cdots \int f(\mathbf{y}|\mathbf{h}, \mathbf{z}) f(\mathbf{h}|\mathbf{z}, \theta) f(\mathbf{z}|\theta) d\mathbf{h} d\mathbf{z} \\ &= \int \cdots \int \prod_{t=1}^n \frac{1}{\sqrt{2\pi z_t} \exp(h_t)} \exp\left[-\frac{y_t^2}{2z_t \exp(h_t)}\right] \\ &\quad \times \prod_{t=1}^{n-1} \frac{1}{\sqrt{2\pi\sigma_\eta^2} \sqrt{1-\rho^2}} \exp\left[-\frac{\{h_{t+1} - \mu - \phi(h_t - \mu) - \rho\sigma_\eta y_t \exp(-h_t/2)/\sqrt{z_t}\}^2}{2\sigma_\eta^2(1-\rho^2)}\right] \\ &\quad \times \frac{\sqrt{1-\phi^2}}{\sqrt{2\pi\sigma_\eta^2}} \exp\left[-\frac{(1-\phi^2)(h_1 - \mu)^2}{2\sigma_\eta^2}\right] \\ &\quad \times \prod_{t=1}^n \frac{(\nu/2)^{\frac{\nu}{2}}}{\Gamma(\nu/2)} z_t^{-(\nu/2+1)} \exp\left(-\frac{\nu}{2z_t}\right) dh_1 \cdots dh_n dz_1 \cdots dz_n. \end{aligned} \quad (4.6)$$

When deriving option prices using Bayesian estimation with HMC, like in the SVL model case, the above likelihood function is used for parameter estimation.

5 Bayesian Inference with HMC Method

5.1 Bayesian Estimation via HMC Method

Given n data points $\mathbf{y} = \{y_t\}_{t=1}^n$, the Bayesian estimation of the model parameters θ , the latent variables $\mathbf{h} = \{h_t\}_{t=1}^n$, and $\mathbf{z} = \{z_t\}_{t=1}^n$ is performed using Bayesian inference based on the following equation:

$$f(\mathbf{h}, \mathbf{z}, \theta|\mathbf{y}) = \frac{f(\mathbf{y}|\mathbf{h}, \mathbf{z}, \theta) f(\mathbf{h}) f(\mathbf{z}) f(\theta)}{Z}. \quad (5.1)$$

Here, $f(\mathbf{h}, \mathbf{z}, \theta|\mathbf{y})$ and $f(\mathbf{y}|\theta)$ are the posterior probability density function and the likelihood function, respectively. $f(\mathbf{h})$, $f(\mathbf{z})$, and $f(\theta)$ are the prior probability density functions for \mathbf{h} , \mathbf{z} , and θ , respectively. Furthermore, Z represents the normalization factor¹⁵⁾ and is defined as follows:

$$Z = \int f(\mathbf{y}|\theta) f(\mathbf{h}) f(\mathbf{z}) f(\theta) d\mathbf{h} d\mathbf{z} d\theta. \quad (5.2)$$

The prior probability density function is set based on information from past data. However, when no information from past data is available, an uninformative prior distribution is used, indicating the absence of prior information. Here, we introduce the conjugate momentum of \mathbf{h} , \mathbf{z} , and $\boldsymbol{\theta}$: $\mathbf{p}_h = \{p_{ht}\}_{t=1}^n$, $\mathbf{p}_z = \{p_{zt}\}_{t=1}^n$, and \mathbf{p}_θ as the conjugate momenta of \mathbf{h} , \mathbf{z} , and $\boldsymbol{\theta}$. The joint distribution of these conjugate momenta's standard normal distribution and the posterior probability density distribution $f(\mathbf{h}, \mathbf{z}, \boldsymbol{\theta}|\mathbf{y})$ is given by the following equations:

$$f(\mathbf{p}_h)f(\mathbf{p}_z)f(\mathbf{p}_\theta) = \frac{1}{\sqrt{(2\pi)^{3n}}} \exp\left(-\frac{\mathbf{p}_h^2}{2} - \frac{\mathbf{p}_z^2}{2} - \frac{\mathbf{p}_\theta^2}{2}\right). \quad (5.3)$$

$$f(\mathbf{h}, \mathbf{z}, \boldsymbol{\theta}, \mathbf{p}_h, \mathbf{p}_z, \mathbf{p}_\theta|\mathbf{y}) = f(\mathbf{h}, \mathbf{z}, \boldsymbol{\theta}|\mathbf{y})f(\mathbf{p}_h)f(\mathbf{p}_z)f(\mathbf{p}_\theta). \quad (5.4)$$

The HMC method generates random numbers from this joint distribution. Based on Eq. (5.4), the following relation holds:

$$\begin{aligned} f(\mathbf{h}, \mathbf{z}, \boldsymbol{\theta}, \mathbf{p}_h, \mathbf{p}_z, \mathbf{p}_\theta|\mathbf{y}) &\propto \exp\left[-\frac{\mathbf{p}_h^2}{2} - \frac{\mathbf{p}_z^2}{2} - \frac{\mathbf{p}_\theta^2}{2} + \log f(\mathbf{y}|\mathbf{h}, \mathbf{z}, \boldsymbol{\theta})\right] \\ &= \exp[-H(\mathbf{h}, \mathbf{z}, \boldsymbol{\theta}, \mathbf{p}_h, \mathbf{p}_z, \mathbf{p}_\theta)]. \end{aligned} \quad (5.5)$$

Here, the Hamiltonian is defined as follows:

$$H(\mathbf{h}, \mathbf{z}, \boldsymbol{\theta}, \mathbf{p}_h, \mathbf{p}_z, \mathbf{p}_\theta) = \frac{\mathbf{p}_h^2}{2} + \frac{\mathbf{p}_z^2}{2} + \frac{\mathbf{p}_\theta^2}{2} - \log f(\mathbf{y}|\mathbf{h}, \mathbf{z}, \boldsymbol{\theta}), \quad (5.6)$$

where \mathbf{h} , \mathbf{z} , $\boldsymbol{\theta}$ and their conjugate momenta \mathbf{p}_h , \mathbf{p}_z , \mathbf{p}_θ obey the Hamilton's equations of motion below. Note that τ represents imaginary time.

$$\frac{d\boldsymbol{\theta}}{d\tau} = \frac{\partial H(\mathbf{h}, \mathbf{z}, \boldsymbol{\theta}, \mathbf{p}_h, \mathbf{p}_z, \mathbf{p}_\theta)}{\partial \mathbf{p}_\theta} = \mathbf{p}_\theta, \quad (5.7)$$

$$\frac{d\mathbf{p}_\theta}{d\tau} = -\frac{\partial H(\mathbf{h}, \mathbf{z}, \boldsymbol{\theta}, \mathbf{p}_h, \mathbf{p}_z, \mathbf{p}_\theta)}{\partial \boldsymbol{\theta}} = \frac{\partial}{\partial \boldsymbol{\theta}} \log f(\mathbf{y}|\mathbf{h}, \mathbf{z}, \boldsymbol{\theta}). \quad (5.8)$$

Here, we show only Hamilton's equations of motion for $\boldsymbol{\theta}$. Eqs. (5.7) and (5.8) cannot be solved analytically, so we generate random numbers by solving them using numerical integration methods that satisfy time reversal and volume conservation, such as the leapfrog method, and then perform Bayesian estimation using the Metropolis method.

5.2 Leapfrog Integration

As stated in the previous subsection, Eqs. (5.7) and (5.8) of the HMC method cannot be solved analytically, so they are obtained numerically using the leapfrog method. The leapfrog method has two arbitrary parameters (ϵ , τ), where ϵ represents the step size, which is a virtual infinitesimal time step, and τ represents the number of time steps. In the HMC method, it is necessary to set (ϵ , τ). Empirically, it is believed that setting ϵ small and τ large ensures that the sample paths sufficiently and uniformly visit the state space.

Equations (5.7) and (5.8) are differentiated as follows:

$$\mathbf{p}\left(\tau + \frac{\epsilon}{2}\right) = \mathbf{p}(\tau) - \left(\frac{\epsilon}{2}\right) \frac{\partial H}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}(\tau)}, \quad (5.9)$$

$$\boldsymbol{\theta}\left(\tau + \frac{\epsilon}{2}\right) = \boldsymbol{\theta}(\tau) + \epsilon \mathbf{p}\left(\tau + \frac{\epsilon}{2}\right), \quad (5.10)$$

$$\mathbf{p}(\tau + \epsilon) = \mathbf{p}\left(\tau + \frac{\epsilon}{2}\right) - \left(\frac{\epsilon}{2}\right) \frac{\partial H}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}(\tau+\epsilon)}. \quad (5.11)$$

Here, the time interval T and the number of divisions N satisfy $T = N \times \epsilon$. The HMC algorithm involves repeating the following steps:

1. Generate \mathbf{p} as a random variable following $N(0, \sigma^2)$.
2. Using Eqs. (5.9)–(5.11), evolve the parameters over time.
3. Using the Hamiltonians before and after the update, the Metropolis test is employed to determine whether to accept or reject the parameters.

Reducing the time interval T increases the frequency of momentum updates, reproducing short random walks. Conversely, increasing T prolongs the time spent evolving under the Hamiltonian equation, delaying transitions to different values of H and increasing the likelihood of becoming trapped in locally stable orbits. Therefore, in the Leapfrog method for HMC, the time interval is not only a parameter determining computational time but also a crucial parameter for estimation accuracy.

5.3 Convergence Diagnosis

To assess the validity of the HMC method, it is necessary to determine whether values obtained through sampling converge to a certain posterior distribution. While several methods exist for judging convergence, we recommend a combination of visual inspection using plots of the sample autocorrelation function, sample paths, and posterior probability density function of the sampled parameters, alongside the Gelman-Rubin statistic¹⁶⁾. When the Gelman-Rubin statistic is close to 1, it can be judged that the chain has converged to a steady state¹⁷⁾.

6 Estimation of SV Options Using HMC Method

When performing option valuation using the SVL model, Eqs. (3.6), (3.7), and (4.1) are used to compute $\{y_{t+1}, \dots, y_T\}$, whereas when performing option valuation using the SVL-t model, Eqs. (4.3) to (4.5) are used. The calculated $\{y_{t+1}, \dots, y_T\}$ are converted to the underlying asset price at maturity as follows:

$$S_T = S_t \exp \left[\sum_{i=t+1}^T y_i \right]. \quad (6.1)$$

Therefore, under these formulations, using a sample of N parameters denoted by $\{\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_N\}$, the European call option price and European put option price can be evaluated by averaging the discounted present values as follows:

$$C_t^{HMC} = e^{-(T-t)r} \frac{1}{N} \sum_{j=1}^N [\text{Max}(S_T(\boldsymbol{\theta}_j) - K, 0)], \quad (6.2)$$

$$P_t^{HMC} = e^{-(T-t)r} \frac{1}{N} \sum_{j=1}^N \left[\text{Max} (K - S_T(\theta_j), 0) \right]. \quad (6.3)$$

Furthermore, using the N samples corresponding to $\{\theta_1, \dots, \theta_N\}$, the expressions in Eqs. (6.2) and (6.3) can be calculated as follows:

$$C_t^{HMC} = e^{-(T-t)r} \frac{1}{N} \sum_{j=\theta_1}^{\theta_N} \left[\text{Max} \left(S_t \exp \left(\sum_{s=t+1}^T y_s^j \right) - K, 0 \right) \right], \quad (6.4)$$

$$P_t^{HMC} = e^{-(T-t)r} \frac{1}{N} \sum_{j=\theta_1}^{\theta_N} \left[\text{Max} \left(K - S_t \exp \left(\sum_{s=t+1}^T y_s^j \right), 0 \right) \right]. \quad (6.5)$$

Using HMC Bayesian estimation, option prices can be derived without separating the estimation of SV model parameters from the option valuation via Monte Carlo simulation, as in previous studies.

7 Conclusion

This paper proposes option valuation using Bayesian inference via the HMC method. Employing Bayesian inference via the HMC method enables the derivation of option prices using asymmetric SV models and SV models with heavy-tailed distributions based on the t -distribution. Future challenges include improving computational accuracy and convergence¹⁸⁾. To improve computational accuracy, it is necessary to introduce a no U-turn algorithm¹⁹⁾ that automatically determines the parameters (number of updates and update size) when performing the Leapfrog method. Furthermore, as an attempt to improve computational convergence, methods such as the Riemannian manifold HMC method²⁰⁾ have been proposed that leverage Riemannian geometry in high-dimensional parameter spaces to enable efficient sampling. It is considered necessary to investigate these methods in the future.

Notes

- 1) The HMC method was originally devised by Duane *et al.* (1987) as an efficient method for generating gauge configurations in lattice quantum chromodynamics (QCD) calculations for particle physics. For details, see Mackay (2003, Chapter 30), Neal (1994), and Neal (2011).
- 2) Molecular dynamics is a method for solving numerically the motions of interacting molecules or atoms using classical mechanics, such as Newtonian mechanics or analytical mechanics. For details, see Alder and Wainwright (1956).
- 3) Empirical studies using Bayesian estimation of the SV model via MCMC for risk asset returns include Yu (2005), Omori *et al.* (2007), Omori and Watanabe (2008), Takahashi *et al.* (2009), Nakajima and Omori (2012), and Takahashi *et al.* (2023).
- 4) For details, see Girolami and Calderhead (2011).
- 5) This equation is for modeling the dynamic behavior of a continuous-time random variable.
- 6) In finance theory, σ is often referred to as volatility, whereas in econometric analysis, the term volatility refers to σ^2 .
- 7) The square root process, unlike the Ornstein-Uhlenbeck process, does not become negative. This makes it extremely useful for modeling interest rate fluctuations.

- 8) The tendency to return to the long-term mean is called mean reversion.
- 9) The price C_T^{BS} of a European call option with strike price K and remaining time τ at time T is given by the Black-Scholes model. By analogy with the Black-Scholes formula, we can infer the following closed-form solution:

$$C_T^{BS} = S_T N(d_1) - K e^{-r^* \tau} N(d_2),$$

$$d_1 = \frac{\ln(S_T/K) + (r^* + \sigma^2/2)\tau}{\sigma \sqrt{\tau}}, \quad d_2 = d_1 - \sigma \sqrt{\tau},$$

$$N(d_i) = \int_{-\infty}^{d_i} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx, \quad i = 1, 2.$$

Here, $N(\cdot)$ denotes the cumulative distribution function of the standard normal distribution, and r^* represents the risk-free interest rate for continuous compounding. For details, see Black and Scholes (1973).

- 10) The following two expressions in the text are each the Fourier transform of a probability measure on Euclidean space. Therefore, it possesses all information about the corresponding probability measure and so enables verification of independence among random variables and estimation of probabilities.
- 11) The transformation of $P_j(x, \sigma^2, T; \ln[K])$ to the characteristic function $f_j(x, \sigma^2, T; \phi)$ is given by

$$f_j(x, \sigma^2, T; \phi) = \int e^{i\phi x} P_j(x, \sigma^2, T; \ln[K]) dx.$$

Furthermore, $f_j(x, \sigma^2, T; \phi)$ satisfies the same diffusion equation as $P_j(x, \sigma^2, T; \ln[K])$.

- 12) For details, see the Appendix in Heston (1993).
- 13) At time t when valuing the option price, the risk-free interest rate r_{t+1} from time t to time $t+1$ is known, but the interest rates beyond that are unknown. Therefore, we assume that the interest rate remains constant from time t until maturity at time T , and calculate the present value using $r = r_{t+1}, r_{t+2}, \dots, r_T$.
- 14) For details on option valuation via Monte Carlo simulation, refer to Barraquand (1995), Broadie and Glasserman (1996), Boyle *et al.* (1997), and Ross (2002).
- 15) Also known as the marginal likelihood.
- 16) For an example in empirical research, see Totsuka and Mitsui (2022) for details.
- 17) For details, see Gelman and Rubin (1992), Gelman (1996).
- 18) Convergence refers to reaching the steady-state distribution; improving convergence means shortening the operational inspection period.
- 19) For details, see Hoffman and Gelman (2014).
- 20) For details, see Girolami and Calderhead (2011).

References

- [1] Alder, B. J. and Wainwright, T. E. (1959), "Studies in Molecular Dynamics. I. General Method," *Journal of Chemical Physics*, 31(2), pp. 459–466.
- [2] Barraquand, J. (1995), "Numerical Valuation of High Dimensional Multivariate European Securities," *Management Science*, 41, pp.1882–1891.
- [3] Black, F. and M. Scholes (1973), "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81, pp.637–654.

- [4] Boyle, P. (1977), "Options: A Monte Carlo Approach," *Journal of Financial Economics*, 4, pp.323–338.
- [5] Boyle, P., Broadie, M. and Glasserman, P. (1997), "Monte Carlo Methods for Security Pricing," *Journal of Economic Dynamics and Control*, 21, pp.1267–1321.
- [6] Breeden, D. T. (1979), "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities," *Journal of Financial Economics*, 7, pp. 265–296.
- [7] Broadie, M. and P. Glasserman (1996), "Estimating Security Price Derivatives Using Simulation," *Management Science*, 42, pp. 269–285.
- [8] Cox, J. C., Ingersoll, J. E. and Ross, S. A. (1985), "An Intertemporal General Equilibrium Model of Asset Prices," *Econometrica*, 53, pp.363-384.
- [9] Duane, S., Kennedy, A. D., Pendleton, B. J., and Roweth, D. (1997), "Hybrid Monte Carlo," *Physics Letters*, B195(2), pp. 216–222.
- [10] Fama, E. (1965), "The Behavior of Stock Prices," *Journal of Business*, 38, pp. 34–105.
- [11] Gelman, A. (1996), "Inference and Monitoring Convergence," in Gilks, W. R., Richardson, S. and Spiegelhalter, D. J. eds., *Markov Chain Monte Carlo in Practice*, pp.131–143, Chapman & Hall.
- [12] Gelman, A. and Rubin, D. B. (1992), "Inference from Iterative Simulation Using Multiple Sequences (with discussion)," *Statistical Science*, 7(4), pp. 457–511.
- [13] Girolami, M. and Calderhead, B. (2011), "Riemann Manifold Langevin and Hamiltonian Monte Carlo Methods," *Journal of the Royal Statistical Society, Series B* 73, pp. 873–889.
- [14] Heston, S. L. (1993), "A Closed Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," *Review of Financial Studies*, 6, pp.327–343.
- [15] Hoffmani, M. D. and Gelman, A. (2014), "The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo," *Journal of Machine Learning Research*, 15, pp. 1593–1623.
- [16] MacKay, D. J. C. (2003), *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press.
- [17] Mandelbrot, B. (1963), "The Variation of Certain Speculative Prices," *Journal of Business*, 36, pp. 394–416.
- [18] Mitsui, H. and Totsuka, H. (2022), "Bayesian Estimation of Stochastic Volatility Model Using Hamiltonian Monte Carlo Method and its Application to Nikkei 225 Data," *Bulletin of Research Institute of Economic Science College of Economic Nihon University*, No.52, pp. 17–25.
- [19] Nakajima, J. and Omori, Y. (2012), "Stochastic Volatility Model with Leverage and Asymmetrically Heavy Tailed Error Using GH Skew Student's *t*-Distribution," *Computational Statistics & Data Analysis*, 56(11), pp. 3690–3704.
- [20] Neal, R. M. (1994), "An Improved Acceptance Procedure for the Hybrid Monte Carlo Algorithm," *Journal of Computational Physics*, 111, pp. 194-203.
- [21] Neal, R. M. (2011), "MCMC Using Hamiltonian Dynamics," in Brooks, S., Gelman, A., Jones, G. L. and Meng, X. - L. eds., *Handbook of Markov Chain Monte Carlo*, pp. 113–162, Chapman & Hall.
- [22] Newton, M. A. and Raftery, A. E. (1994), "Approximate Bayesian Inference by the Weighted Likelihood Bootstrap," *Journal of the Royal Statistical Society, Series B* 56, pp. 3–26.
- [23] Nugroho, D. B., and Morimoto, T. (2015), "Estimation of Realized Stochastic Volatility Models Using Hamiltonian Monte Carlo-Based Methods," *Computational Statistics*, 30, pp. 491–516.
- [24] Omori, Y., Chib, S., Shephard, N. and Nakajima, J. (2007), "Stochastic Volatility with Leverage: Fast Likelihood Inference," *Journal of Econometrics*, 140, pp. 425–449.

- [25] Omori, Y. and Watanabe, T. (2008), "Block Sampler and Posterior Mode Estimation for Asymmetric Stochastic Volatility Models," *Computational Statistics & Data Analysis*, 52, pp. 2892–2910.
- [26] Ross, S. M. (2002), *Simulation*, 3rd ed. Academic Press.
- [27] Takahashi, M., Omori, Y. and Watanabe, T. (2009), "Estimating Stochastic Volatility Models Using Daily Returns and Realized Volatility Simultaneously," *Computational Statistics & Data Analysis*, 53, pp. 2404–2426.
- [28] Takahashi, M., Omori, Y. and Watanabe, T. (2023), *Stochastic Volatility and Realized Stochastic Volatility Models*, Springer.
- [29] Takaishi, T. (2008), "Financial Time Series Analysis of SV Model by Hybrid Monte Carlo," *Lecture Notes in Computer Science*, 5226, pp. 929–936.
- [30] Takaishi, T. (2009), "Bayesian Inference of Stochastic Volatility Model by Hybrid Monte Carlo," *Journal of Circuits, Systems, and Computers*, 18, pp. 1381–1396.
- [31] Takaishi, T. (2013), "Empirical Analysis of Stochastic Volatility Model by Hybrid Monte Carlo Algorithm," *Journal of Physics: Conference Series*, 423(1), 012021.
- [32] Totsuka, H. and Mitsui, H. (2022), "Bayesian Estimation of a Stable Distribution Using the Hamiltonian Monte Carlo Method with Application to Stock Indices," *Bulletin of Research Institute of Economic Science College of Economic Nihon University*, No.52, pp. 3–15.
- [33] Yu, J. (2005), "On Leverage in a Stochastic Volatility Model," *Journal of Econometrics*, 127, pp. 165–178.