Computational Economics (2022) 60:479–527
https://doi.org/10.1007/s10614-021-10156-z

Bayesian Estimation of the Skew Ornstein-Uhlenbeck Process

Yizhou Bai[1] · Yongjin Wang[2] · Haoyan Zhang[3] · Xiaoyang Zhuo[4]

Accepted: 15 July 2021 / Published online: 30 July 2021

Abstract
In this paper, we are particularly interested in the skew Ornstein-Uhlenbeck (OU) process. The skew OU process is a natural Markov process defined by a diffusion process with symmetric local time. Motivated by its widespread applications, we study its parameter estimation. Specifically, we first transform the skew OU process into a tractable piecewise diffusion process to eliminate local time. Then, we discretize the continuous transformed diffusion by using the straightforward Euler scheme and, finally, obtain a more familiar threshold autoregressive model. The developed Bayesian estimation methods in the autoregressive model inspire us to modify a Gibbs sampling algorithm based on properties of the transformed skew OU process. In this way, all parameters including the pair of skew parameters (p, a) can be estimated simultaneously without involving complex integration. Our approach is examined via simulation experiments and empirical analysis of the Hong Kong Interbank Offered Rate (HIBOR) and the CBOE volatility index (VIX), and all of our applications show that our method performs well.

Keywords Skew OU process · Doubly skewed OU process · Bayesian estimation · Gibbs sampler

1 Introduction

The skew process was first introduced by Itô and Mckean (1965) who considered skew Brownian motion (SBM) as a natural Markov process defined by a diffusion process in which the sign of each excursion is decided using an independent Bernoulli variable of parameter p (see Lejay 2006). On the other hand, it was subsequently revealed that a

This research is supported by the National Natural Science Foundation of China (Nos. 11631004, 72001024), the Fundamental Research Funds for the Central Universities (Nos. 3122019139, 3122021120), the Scientific Research Foundation for Introduced Scholars of Civil Aviation University of China (No. 2020KYQD90), and the China Postdoctoral Science Foundation (No. 2018M641396).

Extended author information available on the last page of the article

skew process actually solves a stochastic equation that involves symmetric local time at the skew level a (see Harrison and Shepp 1981). Due to the existence of local time, a skew process coupled with a pair of parameters (p, a) exhibits the following path characteristics: it behaves like a stochastic process but without the local time component while outside the skew level a. Immediately upon reaching a, it moves upwards with the skew probability p and downwards with probability $1 - p$. Therefore, the skew process has an advantage in that it can be used to capture the controlled dynamics of asset prices or interest rates. Specifically, a large (small) value for the skew level results in a potential upper (lower) bound for the dynamics, and a large (small) value for the skew probability results in a support (resistance) force for the dynamics.

Because of the interesting path feature, study of the theoretical properties of skew diffusion has attracted much attention. For example, Walsh (1978) computed the speed density and transition probability density of SBM. Harrison and Shepp (1981) proved that SBM behaved as Brownian motion except for the origin, where it followed a fixed angular distribution from the origin. Gall (1984) extended the results of strong existence and uniqueness to the stochastic differential equations (SDEs) for a more general class of skew processes. Ouknine and Rutkowski (1995) provided a review of formulae related to the local times of functions of continuous semimartingales. Appuhamillage and Sheldon (2010) characterized the distribution of the ranked excursion heights of SBM and derived its first passage time distribution. Later, Appuhamillage et al. (2011) extended the analysis to the case of SBM with a constant drift.

In the last two decades, the skew process has been connected with several applications. For example, SBM has been applied in physics (Lang 1995), population ecology (Cantrell and Cosner 1999), multiarm bandit problems (Barlow et al. 2000), astrophysics (Zhang 2000), engineering (Lejay 2003), and geophysics (Lejay 2004 and Ramirez and Jorge 2011). Among them, we are mostly interested in the recent financial applications of skew diffusions. For instance, Zhu and He (2017) proposed a newly described “endogenous risk neutral measure” in a skew-extended Black-Scholes framework and derived its corresponding closed-form option pricing formula. Gairat and Shcherbakov (2016) gave the pricing formula for European options based on the functionals of SBM. Xu et al. (2016) used the skew CIR process to describe exchange rates in the target zone and priced two kinds of exotic options in terms of the Laplace transform and spectral expansion.

Nevertheless, the extant literature and research have mainly focused on the theoretical properties of the skew process and the corresponding derivative pricing. It is natural to ask whether the skew process can do a better job at fitting the dynamics of asset prices or interest rates. To answer this question, the first problem that we confront is parameter estimation. Among the existing studies addressing this issue, Lejay (2006, 2017) studied the distribution function of SBM and estimated the skew probability p when the skew level a is known; Bardou and Martinez (2010) gave estimates for the skew probability p and the pair (p, a), where a is the skew level, assuming that other drift and diffusion coefficients are known.

In this paper, we are particularly interested in parameter estimation for the skew Ornstein-Uhlenbeck (OU) process. The OU process is one of the most commonly used processes in finance due to its nice probability properties and ability to model a broad set of mean-reverting financial variables (see, e.g., Vasicek 1977; Hull and White 1990; Collin-Dufresne and Goldstein 2001). Based on SBM, Wang et al.

(2015) extended the skew process to the skew OU process[1] and proved its existence and the uniqueness of the solution. As a combination of the OU process and the skew local time component, the skew OU process not only possesses the advantages of the mean-reverting OU process but also incorporates the special path property of the skew process. Wang et al. (2015) priced a zero-recovery defaultable bond assuming that the asset dynamics are driven by the skew OU process. In addition, Zhuo and Menoukeupamen (2017) proposed piecewise binomial and trinomial lattice approaches for the skew OU process with discontinuous drift, which made it feasible to price nonstandard interest rate derivatives under the skew process.

Regarding the parameter estimation of the skew OU process, Wang et al. (2015) derived the transition density of the skew OU process using spectral expansion and expressed it in the form of a summation of density series. However, their results involve complex multiple integrals, which are difficult to apply to simulation or real data. Furthermore, since the skew OU process has nonconstant drift, we cannot transform it into the SBM by using the estimation approach introduced in Appuhamillage and Iresh (2011) and Lejay and Pichot (2012). In addition, neither the drift and diffusion coefficients nor the skew level a of the skew OU process is known; hence, the methods mentioned in Lejay (2006, 2017) and Bardou and Martinez (2010) are not suitable here. Furthermore, the series expansion of the transition density of the skew OU process makes the maximum likelihood estimation lose its power.

As a result, this paper seeks help from the Bayesian approach. Unlike previous attempts, our proposed Bayesian estimation is easy to implement and does not involve any complicated special functions. In this approach, the local time component is removed first by a piecewise transform. This transform technique was introduced by Harrison and Shepp (1981) and is widely used in the relevant literature (see Song et al. 2016 and Zhuo et al. 2017). The transformed process is a tractable piecewise diffusion process and resembles the self-exciting threshold autoregressive (SETAR) model in economics. Motivated by this insight, we select the Gibbs sampler, which is a popular estimation approach in threshold autoregressive models. Because of the nice probability properties of the transformed piecewise process, the distribution of an arbitrary parameter can be explicitly expressed given the fixed (ex ante) values of all the remaining parameters. Therefore, we can iteratively estimate all of the parameters one by one until their estimates converge. The convergence detection procedure is also provided. In addition, we also consider the estimation of the doubly skewed OU model, which is a natural generalization of the skew OU model and has two skew local time components.

We apply our proposed estimation approach to simulations first, and the results demonstrate that our method performs very well for both the skew OU process and the doubly skewed OU process. We also apply these processes to real data and reveal several interesting findings in Hong Kong short-term interest rates and the Chicago Board Options Exchange (CBOE) volatility index. Notably, our method is not limited to the skew OU process; it can also be generalized to other skew diffusions, such as the skew CIR process and the skew geometric Brownian motion process.

The remainder of this paper proceeds as follows: In Sect. 2, we provide the definitions of the skew OU process and the doubly skewed OU process. To facilitate

[1] It is also known as the skew Vasicek model in interest rate modeling (see Zhuo and Menoukeupamen 2017).

estimation, we transform them into piecewise processes similar to threshold autoregressive models. Section 3 proposes the Bayesian inference for the two skew OU processes. Both the simulation experiments and the empirical analysis of real data are illustrated in Sect. 4. Section 5 concludes the paper.

2 The Processes

In this section, we define the skew OU process and the doubly skewed OU process. When considering their estimation, the local time components in the processes are hard to address. Therefore, we derive their respective transformed processes such that the local time components are removed and the processes are more tractable.

2.1 The Skew OU Process

Let us fix a probability space $(\Omega, \mathcal{F}, \mathcal{P})$ and an information filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual conditions. According to Wang et al. (2015), the skew OU process is driven by the following SDE with a symmetrical local time component:

$$dX_t = \kappa(\theta - X_t)dt + \sigma dW_t + (2p-1)d\hat{L}_t^X(a), \tag{1}$$

where X_t can be a stochastic interest rate or asset price process, and κ, θ, σ, p and a are unknown structural parameters. In this setting, $\kappa > 0$ is the speed of mean reversion, θ is the unconditional long-term mean, and $\sigma > 0$ is the volatility of the process. $\hat{L}_t^X(a)$ is the symmetric local time of the continuous semimartingale $X = \{X_t, \mathcal{F}_t; 0 \leq t < +\infty\}$ at skew level a, while $p \in (0,1)$ is the skew probability capturing the possibility of upward movement after hitting skew level a.

We first recall the definition of symmetric local time from Protter (2004). Let $\mathrm{sign}(x)$ be the sign function defined by

$$\mathrm{sign}(x) = \begin{cases} 1, & \text{if } 0 < x, \\ -1, & \text{if } x \leq 0. \end{cases}$$

Then the right local time of X at level a is

$$L_t^X(a) \triangleq |X_t - a| - |X_0 - a| - \int_0^t \mathrm{sign}(X_s - a)dX_s,$$

with the left local time $L_t^X(a-) \triangleq \lim_{b\uparrow a} L_t^X(b)$.

Hence, the symmetric local time $\hat{L}_t^X(a)$ is given by

$$\hat{L}_t^X(a) \triangleq [L_t^X(a) + L_t^X(a-)]/2.$$

At first sight, the definition of local time is slightly abstract and difficult to directly address. Therefore, we do not tackle the original process X equipped with the local time component. Instead, to estimate the five unknown parameters, we transform process X into a tractable piecewise diffusion process without local time, and then the proposed estimation method in later section is applied to the transformed process. To be specific, we set a transformed process $Y_t \triangleq G(X_t)$ by

$$G(X_t) = \begin{cases} (1-p)(X_t - a) + a, & \textit{if } a \le X_t, \\ p(X_t - a) + a, & \textit{if } X_t < a. \end{cases} \tag{2}$$

Meanwhile, the inverse transform $X_t \triangleq H(Y_t)$ admits

$$H(Y_t) = \begin{cases} \dfrac{1}{1-p} Y_t - \dfrac{p}{1-p} a, & \textit{if } a \le Y_t, \\ \dfrac{1}{p} Y_t - \dfrac{(1-p)}{p} a, & \textit{if } Y_t < a. \end{cases} \tag{3}$$

Since the function $G(\cdot)$ is a combination of two convex functions and is differentiable, we apply the generalized Itô formula (see Revuz and Yor 2013) to process Y_t and obtain

$$Y_t = G(X_0) + \frac{1}{2}\int_0^t \Big(G^{'}(X_s+) + G^{'}(X_s-)\Big) dX_s + \frac{1}{2}\int_{\mathbf{R}} \hat{L}_t^X(z)\mu(dz),$$

where $G^{'}(x+)$ (resp. $G^{'}(x-)$) is the right (resp. left) derivatives of $G(x)$, μ is the signed measure (when restricted to compacts), which is the second derivative of G in the generalized function sense with the properties $\mu((a,b]) = G^{'}(b+) - G^{'}(a-)$ and $\mu(\{x\}) = G^{'}(x+) - G^{'}(x-)$. Furthermore, if $G(\cdot)$ is the bounded Borel measurable function, we have the occupation time formula

$$\int_{\mathbf{R}} \hat{L}_t^X(z)\mu(dz) = \int_0^t G^{''}(X_s) d < X >_s,$$

then we have

$$\begin{aligned} Y_t &= G(X_0) + \frac{1}{2}\int_0^t \Big(G^{'}(X_s+) + G^{'}(X_s-)\Big) 1_{\{X_s \neq a\}} dX_s \\ &\quad + \frac{1}{2}\int_0^t \Big(G^{'}(X_s+) + G^{'}(X_s-)\Big) 1_{\{X_s = a\}} dX_s \\ &\quad + \frac{1}{2}\int_{\mathbf{R}\setminus\{a\}} \hat{L}_t^X(z)\mu(dz) + \frac{1}{2}\int_{\{a\}} \hat{L}_t^X(z)\mu(dz) \\ &= G(X_0) + \int_0^t G^{'}(X_s)(\kappa(\theta - X_s)ds + \sigma dW_s) \\ &\quad + \frac{1}{2}\Big(G^{'}(a+) + G^{'}(a-)\Big)(2p-1)\hat{L}_t^X(a) \\ &\quad + \frac{1}{2}\int_0^t G^{''}(X_s) d < X >_s + \frac{1}{2}\Big(G^{'}(a+) - G^{'}(a-)\Big)\hat{L}_t^X(a) \\ &= G(X_0) + \int_0^t G^{'}(X_s)(\kappa(\theta - X_s)ds + \sigma dW_s) + \frac{1}{2}\int_0^t G^{''}(X_s)\sigma^2 ds. \end{aligned} \tag{4}$$

We finally obtain the new process Y_t without the local time component.

Remark 2.1 The last equality holds due to the fact that the skew diffusion admits the following boundary (see for example, Harrison and Shepp (1981)):

$$pG^{'}(a+) - (1-p)G^{'}(a-) = 0.$$

We rewrite the right-hand-side of Eq.(4) in terms of Y_t; then, the skew OU process X_t defined in Eq.(1) is transformed into a tractable piecewise process Y_t satisfying the following SDE:

$$dY_t = \begin{cases} \kappa[(1-p)\theta + pa - Y_t]dt + (1-p)\sigma dW_t, & \text{if } a \le Y_t, \\ \kappa[p\theta + (1-p)a - Y_t]dt + p\sigma dW_t, & \text{if } Y_t < a. \end{cases} \tag{5}$$

The process Y_t shown in Eq.(5) is much easier to deal with than the skew OU process X_t shown in Eq.(1).

When considering the parameter estimation of X_t, generally the most straightforward method is maximum likelihood estimation (MLE), which needs the transition density of X_t. Wang et al. (2015) give the transition density of X_t using spectral expansion, based on Eq.(5), as

$$p_X(t; x_1, x_2) = p_Y(t; G(x_1), G(x_2))G'(x_2), \tag{6}$$

$$p_Y(t; y_1, y_2) = m_Y(y_2) \sum_{n=1}^{+\infty} \exp(-\lambda_n t)\varphi_n(y_1)\varphi_n(y_2). \tag{7}$$

More details are given in Appendix A. The transition density is expressed in the form of a summation of density series, involving parabolic cylinder function and the Hermite function. Due to these complex multiple integrals, which are difficult to apply to simulation or real data, it is basically impossible to implement this accurate MLE method for the skew OU process.

Therefore, in Sect. 3, we will propose our Bayesian estimation approach based on the transformed process Y_t.

2.2 The Doubly Skewed OU Process

For practical purposes, we also consider the doubly skewed OU process, which has two skew local time components. Specifically, the doubly skewed OU process is governed by the following SDE:

$$dX_t = \kappa(\theta - X_t)dt + \sigma dW_t + (2p_1 - 1)d\hat{L}_t^X(a_1) + (2p_2 - 1)d\hat{L}_t^X(a_2). \tag{8}$$

As a natural generalization of the skew OU process, the doubly skewed OU process allows more flexibility in determining the potential upper (lower) bounds and support (resistance) forces of the dynamics. If we set the two local time components as an upper bound with resistance force and a lower bound with support force, the path of the process X_t will be restricted between the two bounds with a certain probability. This property means that the process X_t shown in Eq.(8) is able to more accurately describe the dynamics of many financial variables. For example, we hope that the exchange rate is stable in a certain range, and it usually is so. However, the exchange rate may also break through the bound due to the hysteresis of regulation; or some unexpected incident occurring in the market. The doubly skewed process maintains the possibility that the bounds of the dynamics can be broken.

To address the process X_t shown in Eq.(8), a three-piece transform is required to remove the local time components. Similarly, the transformed process $Y_t \triangleq \tilde{G}(X_t)$ has the form

$$\tilde{G}(X_t) = \begin{cases} p_1 p_2 (X_t - a_1) + \omega_1 a_1 + (1-\omega_1) a_2, & \text{if } X_t < a_1, \\ (\omega_1 - p_1)(X_t - a_1) + \omega_1 a_1 + (1-\omega_1) a_2, & \text{if } a_1 \le X_t < a_2, \\ (1-\omega_1)(X_t - a_2) + p_1 a_1 + (1-\omega_1) a_2, & \text{if } a_2 \le X_t, \end{cases} \tag{9}$$

where $\omega_1 = p_1 - p_1 p_2 + p_2$.

The inverse transform $X_t \triangleq \tilde{H}(Y_t)$ admits

$$\tilde{H}(Y_t) = \begin{cases} \dfrac{Y_t}{p_1 p_2} - \dfrac{\omega_1}{p_1 p_2} a_1 - \dfrac{1-\omega_1}{p_1 p_2} a_2, \\ \qquad \text{if } Y_t < \omega_1 a_1 + (1-\omega_1) a_2, \\ \dfrac{Y_t}{\omega_1 - p_1} - \dfrac{p_1}{\omega_1 - p_1} a_1 - \dfrac{1-p_2}{p_2} a_2, \\ \qquad \text{if } \omega_1 a_1 + (1-\omega_1) a_2 \le Y_t < p_1 a_1 + (1-p_1) a_2, \\ \dfrac{Y_t}{(1-\omega_1)} - \dfrac{p_1}{(1-\omega_1)} a_1 - \dfrac{p_2}{1-p_2} a_2, \\ \qquad \text{if } p_1 a_1 + (1-p_1) a_2 \le Y_t. \end{cases} \tag{10}$$

Based on Eqs.(9) and (10), we apply the generalized Itô formula (see Revuz and Yor 2013) to process Y_t, which yields

$$dY_t = \begin{cases} \kappa[p_1 p_2 \theta + (\omega_1 - p_1 p_2) a_1 + (1-\omega_1) a_2 - Y_t] dt + p_1 p_2 \sigma dW_t, \\ \qquad \text{if } Y_t < \omega_1 a_1 + (1-\omega_1) a_2, \\ \kappa[(\omega_1 - p_1)\theta + p_1 a_1 + (1-\omega_1) a_2 - Y_t] dt + (\omega_1 - p_1) \sigma dW_t, \\ \qquad \text{if } \omega_1 a_1 + (1-\omega_1) a_2 \le Y_t < p_1 a_1 + (1-p_1) a_2, \\ \kappa[(1-\omega_1)\theta + p_1 a_1 + (\omega_1 - p_1) a_2 - Y_t] dt + (1-\omega_1) \sigma dW_t, \\ \qquad \text{if } p_1 a_1 + (1-p_1) a_2 \le Y_t. \end{cases} \tag{11}$$

It is interesting to note that, as shown in the next section, the discretized version of the transformed processes Y_t in Eq.(5) for the skew OU process and in Eq.(11) for the doubly skewed OU process resembles the self-exciting threshold autoregressive (SETAR) model.[2] The SETAR model, sometimes called the threshold autoregressive (TAR) model, proposed by Tong (1978), has been applied in a wide range of fields, such as economics, econometrics, and finance. SETAR models are popular because they can exhibit many nonlinear phenomena, such as limit cycles, chaos, harmonic distortion, jump phenomena, and time irreversibility (see, e.g., Gonzalo

[2] A typical two-regime SETAR model is defined as

and Wolf 2005). A comprehensive survey of SETAR models is available in Tong (1990). Before we move on, it is worth noting that what we can observe in the market is X_t, satisfying the skew-extended process Eq.(1) or Eq.(8). After the transform by $G(x)$, X_t becomes the SETAR model Y_t. However, since the skew levels and skew probabilities are unknown, Y_t is unobservable.

The discretized versions of the skew-extended OU processes are similar to the SETAR model. For the SETAR model, it is common to use the Bayesian method for estimation purposes; see, for example, Broemeling and Cook (1992); Geweke and Terui (1993); Pfann et al. (1996) and Chen and Lee (2010). Motivated by this insight, we also seek help from the Bayesian approach to solve the estimation problem of the two skew-extended OU processes.

3 The Bayesian Estimation

In this section, we propose Bayesian estimation techniques for both the skew OU process and the doubly skewed OU process. First, we make use of the discretized versions of their transformed processes to compute these processes' likelihood functions. Then, to apply the Gibbs sampler, we derive the conditional posterior distribution of an arbitrary parameter. Finally, we obtain the parameter estimates by averaging over the samples generated from their posterior distributions.

For the Bayesian methods, if we assume Θ to be the set of all parameters $(\theta_1, \theta_2, \ldots, \theta_n)$, the joint prior density function of Θ is $\pi(\Theta)$, the likelihood function is $l(X|\Theta)$, and then the posterior density function is

$$\pi(\Theta|X) = \frac{l(X|\Theta)\pi(\Theta)}{\int l(X|\Theta)\pi(\Theta)d\Theta} \quad \propto \quad l(X|\Theta)p(\Theta). \tag{13}$$

In the Bayesian approach, the inference that we want to conduct can be evaluated from the expectation of a certain function $g(\Theta)$, i.e.,

$$E[g(\Theta)] = \int g(\Theta)\pi(\Theta|X)d\Theta. \tag{14}$$

To avoid the complexity of multiple integrations, the Monte Carlo method is adopted. Let $\left\{\Theta^{(1)}, \ldots, \Theta^{(m)}\right\}$ be the samples generated from the posterior density function $\pi(\Theta|X)$; then, Eq.(14) is approximated by

Footnote 2 continued

$$X_t = \begin{cases} \phi_{10} + \sum_{i=1}^{l} \phi_{1i}X_{t-i} + \sigma_1\epsilon_t, & \text{if } X_{t-d} \le r, \\ \phi_{20} + \sum_{i=1}^{l} \phi_{2i}X_{t-i} + \sigma_2\epsilon_t, & \text{if } r < X_{t-d}, \end{cases} \tag{12}$$

where $d \le l$ is the threshold lag, r is the threshold level and $\{\epsilon_t\}$ is the white noise with mean zero and unit variance.

$$E[g(\Theta)] = \frac{1}{m}\sum_{i=1}^{m} g\left(\Theta^{(i)}\right). \tag{15}$$

However, due to the difficulty in generating samples $\left\{\Theta^{(1)}, \ldots, \Theta^{(m)}\right\}$ directly, we employ the Gibbs sampler, which provides a feasible way to generate these samples. In fact, the Gibbs sampler always uses the full set of univariate conditionals to define the iteration. In our case, instead of generating $\Theta^{(i+1)} = \left\{\theta_1^{(i+1)}, \theta_2^{(i+1)}, \ldots, \theta_n^{(i+1)}\right\}$ from $\Theta^{(i)} = \left\{\theta_1^{(i)}, \theta_2^{(i)}, \ldots, \theta_n^{(i)}\right\}$ by $\pi(\Theta|X)$ directly, we obtain $\Theta^{(i+1)}$ with the conditional probability densities as follows:

$$\begin{aligned} \theta_1^{(i+1)} &\sim \pi\left(\theta_1|X, \theta_2 = \theta_2^{(i)}, \theta_3 = \theta_3^{(i)}, \ldots, \theta_n = \theta_n^{(i)}\right), \\ \theta_2^{(i+1)} &\sim \pi\left(\theta_2|X, \theta_1 = \theta_1^{(i+1)}, \theta_3 = \theta_3^{(i)}, \ldots, \theta_n = \theta_n^{(i)}\right), \\ &\cdots \\ \theta_n^{(i+1)} &\sim \pi\left(\theta_n|X, \theta_1 = \theta_1^{(i+1)}, \theta_2 = \theta_2^{(i+1)}, \ldots, \theta_{n-1} = \theta_{n-1}^{(i+1)}\right). \end{aligned} \tag{16}$$

To make the Gibbs sampler computationally efficient, the priors are chosen so that the conditional posterior distributions are easy to simulate. By convention, conjugate priors are used to obtain simple analytical forms for the resulting posterior distributions; we refer the reader to Chen and Li (1995).

3.1 Estimation of the Skew OU Process

We assume that the data $\{X_{t_i}, i = 1, \ldots, N+1\}$ are sampled at discrete times $\{t_i, i = 1, \ldots, N+1\}$, subject to $t_1 < \ldots < t_{N+1}$. The times $\{t_i, i = 1, \ldots, N+1\}$ are equally spaced and the time step is Δt, i.e., $t_{i+1} = t_i + \Delta t$.

For the skew OU process, applying the transform to $\{X_{t_i}, i = 1, \ldots, N+1\}$ as in Eq.(2), we can obtain the data $\{Y_{t_i}, i = 1, \ldots, N+1\}$. Then, according to the independent increment property of Brownian motion, for $i = 1, \ldots, N+1$, the discretized version of Eq.(5) can be expressed as

$$\triangle Y_{t_i} = \begin{cases} \kappa[(1-p)\theta + pa]\triangle t - \kappa\triangle t Y_{t_i} + (1-p)\sigma\sqrt{\triangle t}\epsilon_{t_i}, & \textit{if } a \leq Y_{t_i}, \\ \kappa[p\theta + (1-p)a]\triangle t - \kappa\triangle t Y_{t_i} + p\sigma\sqrt{\triangle t}\epsilon_{t_i}, & \textit{if } Y_{t_i} < a, \end{cases} \tag{17}$$

where $\triangle Y_{t_i} = Y_{t_{i+1}} - Y_{t_i}$ and $\{\epsilon_{t_i}\}_{i=1}^{N}$ are independent and standard normally distributed. By setting $y_{t_i} \triangleq \triangle Y_{t_i}$ and $x_{t_i} \triangleq Y_{t_i}$, Eq.(17) becomes

$$y_{t_i} = \begin{cases} \kappa[(1-p)\theta + pa]\Delta t - \kappa\Delta t x_{t_i} + (1-p)\sigma\sqrt{\Delta t}\epsilon_{t_i}, & \text{if } a \leq x_{t_i}, \\ \kappa[p\theta + (1-p)a]\Delta t - \kappa\Delta t x_{t_i} + p\sigma\sqrt{\Delta t}\epsilon_{t_i}, & \text{if } x_{t_i} < a. \end{cases} \tag{18}$$

Denote by N_1 (resp. N_2) the set in which the sample value is larger (resp. smaller) than the skew level a, i.e., $N_1 \triangleq \{i : i = 1, \ldots, N, a \leq x_{t_i}\}$ and $N_2 \triangleq \{i : i = 1, \ldots, N, x_{t_i} < a\}$. Let n_1 and n_2 be the amounts of i included in the sets N_1 and N_2, respectively. Obviously, $n_1 + n_2 = N$. Then the likelihood function of Eq.(18) is

$$\begin{aligned} l(X|\Theta) = & \left(\frac{1}{(1-p)\sqrt{2\pi\sigma^2\Delta t}}\right)^{n_1} \exp\left\{-\frac{1}{2(1-p)^2\sigma^2\Delta t}\right. \\ & \left.\times \sum_{i\in N_1}[y_{t_i} + \kappa\Delta t x_{t_i} - \kappa\Delta t((1-p)\theta + pa)]^2\right\} \\ & \times \left(\frac{1}{p\sqrt{2\pi\sigma^2\Delta t}}\right)^{n_2} \exp\left\{-\frac{1}{2p^2\sigma^2\Delta t}\right. \\ & \left.\times \sum_{i\in N_2}[y_{t_i} + \kappa\Delta t x_{t_i} - \kappa\Delta t(p\theta + (1-p)a)]^2\right\}, \end{aligned} \tag{19}$$

where Θ is the set of five parameters $\{\kappa, \theta, \sigma^2, a, p\}$ for the skew OU process, and X represents the data set.

Even though the accurate MLE method using Eqs.(6) and (7) is infeasible for the original data set X_t, we can still have a two-step MLE method revised from Su and Chan (2015). The first step selects a pair of $(a, p) \in (a_d, a_u) \times [0, 1]$, where (a_d, a_u) is a predetermined interval from past observations. Given the pair (a, p), we can obtain a transformed data set Y_t; and the maximum likelihood estimators for the drift and diffusion components, labeled as $(\hat{\kappa}, \hat{\theta}, \hat{\sigma}^2)$, can be obtained using Eq.(19). Then, for that group of parameters, a likelihood function value $l(a, p) = l(\hat{\kappa}, \hat{\theta}, \hat{\sigma}^2, a, p)$ is obtained. Hence, the second step finds the final estimator of the pair $(\hat{a}, \hat{p})$ by choosing the maximum likelihood value $l(a, p)$ from all possibilities of $(a, p) \in (a_d, a_u) \times [0, 1]$. However, this two-step method suffers from dimension explosion when we consider more skew levels in the diffusion process. Specifically, for the doubly skewed OU process, it results in searching for all possibilities of $(a_1, p_1, a_2, p_2) \in (a_{1d}, a_{1u}) \times [0, 1] \times (a_{2d}, a_{2u}) \times [0, 1]$, which is computationally intensive and inefficient. Therefore, to compare with our Bayesian method, we only perform the results of this two-step MLE method for the skew OU process later in this paper.

Regarding the procedure of our proposed Bayesian method, according to Eq.(13), we know that the posterior density function $\pi(\Theta|X)$ is proportional to the product of likelihood function $l(X|\Theta)$ in Eq.(19) and prior density $\pi(\Theta)$. Since there are five parameters in Θ for the skew OU process, the joint posterior density $\pi(\Theta|X)$ is hard to know. Hence, we employ the Gibbs sampler and only need to find the conditional posterior density of each parameter with given values for the remaining four parameters. Then, the parameters can be simulated one by one according to Eq.(16). Now, we will look for the conditional distribution of each parameter.

3.1.1 Conditional distribution of the instantaneous return

Conditional on θ, σ, p and a, the proper prior distribution of mean-reversion speed κ should be a normal distribution $\mathcal{N}\left(\mu_\kappa, \sigma_\kappa^2\right)$. Combined with the likelihood function Eq.(19), and after a rearrangement of terms, we obtain the posterior distribution of κ as follows:

$$\kappa | X, \theta, \sigma^2, p, a \sim \mathcal{N}\left(\widehat{\mu}_\kappa, \widehat{\sigma}_\kappa^2\right), \tag{20}$$

where

$$\begin{aligned}
\widehat{\mu}_\kappa = & \left\{ \frac{1}{(1-p)^2\sigma^2} \sum_{i \in N_1} y_{t_i} \{[-x_{t_i} + ((1-p)\theta + pa)]\} \right. \\
& \left. + \frac{1}{p^2\sigma^2} \sum_{i \in N_2} y_{t_i} \{[-x_{t_i} + (p\theta + (1-p)a)]\} + \frac{\mu_\kappa}{\sigma_\kappa^2} \right\} \widehat{\sigma}_\kappa^2, \\
\widehat{\sigma}_\kappa^{-2} = & \frac{1}{(1-p)^2\sigma^2} \sum_{i \in N_1} \{[x_{t_i} - ((1-p)\theta + pa)]\}^2 \triangle t \\
& + \frac{1}{p^2\sigma^2} \sum_{i \in N_2} \{[x_{t_i} - (p\theta + (1-p)a)]\}^2 \triangle t + \frac{1}{\sigma_\kappa^2}.
\end{aligned}$$

Similarly, given κ, σ, p and a, the proper prior distribution of long-term mean θ should be a normal distribution $\mathcal{N}(\mu_\theta, \sigma_\theta^2)$. Again, by the likelihood function Eq.(19), the posterior distribution of θ is

$$\theta | X, \kappa, \sigma^2, p, a \sim \mathcal{N}\left(\widehat{\mu}_\theta, \widehat{\sigma}_\theta^2\right), \tag{21}$$

where

$$\begin{aligned}
\widehat{\mu}_\theta = & \left\{ \frac{1}{(1-p)\sigma^2} \sum_{i \in N_1} [\kappa(y_{t_i} + \kappa \triangle t x_{t_i} - \kappa pa \triangle t)] \right. \\
& \left. + \frac{1}{p\sigma^2} \sum_{i \in N_2} \{\kappa[y_{t_i} + \kappa \triangle t x_{t_i} - \kappa(1-p)a\triangle t]\} + \frac{\mu_\theta}{\sigma_\theta^2} \right\} \widehat{\sigma}_\theta^2, \\
\widehat{\sigma}_\theta^{-2} = & \frac{\kappa^2 \triangle t N}{\sigma^2} + \frac{1}{\sigma_\theta^2}.
\end{aligned}$$

Therefore, we use Eqs.(20) and (21) as the conditional posterior density to estimate κ and θ in Eq.(5) by the approach introduced in Eq.(16).

3.1.2 Conditional Distribution of the Volatility

Different from κ and θ, the prior distribution of instantaneous variance σ^2 is an inverted Gamma distribution $\mathcal{IG}(\alpha_\sigma, \lambda_\sigma)$ conditional on κ, θ, p and a. Its density function is denoted by

$$\pi(\sigma^2) = \frac{\lambda_\sigma}{\Gamma(\alpha_\sigma)} \left(\frac{1}{\sigma^2}\right)^{\alpha_\sigma+1} e^{\frac{-\lambda_\sigma}{\sigma^2}}. \tag{22}$$

Thus, the posterior distribution of σ^2 is

$$\sigma^2 | X, \kappa, \theta, p, a \sim \mathcal{IG}\left(\widehat{\alpha}_\sigma, \widehat{\lambda}_\sigma\right), \tag{23}$$

where

$$\begin{aligned}
\widehat{\alpha}_\sigma =& \frac{n_1}{2} + \frac{n_2}{2} + \alpha_\sigma, \\
\widehat{\lambda}_\sigma =& \frac{1}{2(1-p)^2 \Delta t} \sum_{i \in N_1} \{y_{t_i} + \kappa \Delta t x_{t_i} - \kappa[(1-p)\theta + pa]\Delta t\}^2 \\
&+ \frac{1}{2p^2 \Delta t} \sum_{i \in N_2} \{y_{t_i} + \kappa \Delta t x_{t_i} - \kappa[p\theta + (1-p)a]\Delta t\}^2 + \lambda_\sigma.
\end{aligned}$$

Based on Eq.(23), we can obtain the estimation of σ^2.

3.1.3 Conditional Distribution of the Skew Level

Because it is hard to find the conjugate prior for skew level a, the Griddy-Gibbs sampler is chosen here following Ritter and Tanner (1992). Assume that a is uniform on a predetermined interval (a_d, a_u) from past observations. Conditional on κ, θ, σ and p, the density of a is developed

$$\begin{aligned}
\pi\left(a | X, \kappa, \theta, \sigma^2, p\right) \quad \propto & \left(\frac{1}{(1-p)\sqrt{2\pi\sigma^2 \Delta t}}\right)^{n_1} \exp\left\{ - \frac{1}{2(1-p)^2 \sigma^2 \Delta t} \right. \\
& \left. \times \sum_{i \in N_1} [y_{t_i} + \kappa \Delta t x_{t_i} - \kappa \Delta t((1-p)\theta + pa)]^2 \right\} \\
& \times \left(\frac{1}{p\sqrt{2\pi\sigma^2 \Delta t}}\right)^{n_2} \exp\left\{ - \frac{1}{2p^2 \sigma^2 \Delta t} \right. \\
& \left. \times \sum_{i \in N_2} [y_{t_i} + \kappa \Delta t x_{t_i} - \kappa \Delta t(p\theta + (1-p)a)]^2 \right\} \frac{1}{a_u - a_d}.
\end{aligned} \tag{24}$$

The skew level a in Eq.(24) takes values at the n grid points $\{a_1, \ldots, a_n\}$ on the interval (a_d, a_u). Based on the right-hand-side of Eq.(24) and the grid points, we can compute the cumulative density F_a by numerical integration. Then together with a uniform random number U_a, we can produce $a^{(i+1)}$ from $F_a^{-1}(U_a)$ and use it in the next iteration. At the same time, the new grid points used in the next iteration are obtained by computing the empirical quantiles $F_a^{-1}(\xi_i)$ for preselected values $0<\xi_1<\ldots<\xi_n<1$.

3.1.4 Conditional Distribution of the Skew Probability

Analogously, we assume the skew probability p to be uniform on the interval $[0,1]$ and the grid points to be $\{p_1,\ldots,p_n\}$. For each $p\in\{p_1,\ldots,p_n\}$, conditional on κ, θ, σ and a, we calculate the density of skew probability p as follows

$$\begin{aligned}\pi\left(p|X,\kappa,\theta,\sigma^2,a\right) \quad \propto & \left(\frac{1}{(1-p)\sqrt{2\pi\sigma^2\Delta t}}\right)^{n_1}\exp\left\{-\frac{1}{2(1-p)^2\sigma^2\Delta t}\right.\\ & \left.\times\sum_{i\in N_1}\left[y_{t_i}+\kappa\Delta t x_{t_i}-\kappa\Delta t((1-p)\theta+pa)\right]^2\right\}\\ & \times\left(\frac{1}{p\sqrt{2\pi\sigma^2\Delta t}}\right)^{n_2}\exp\left\{-\frac{1}{2p^2\sigma^2\Delta t}\right.\\ & \left.\times\sum_{i\in N_2}\left[y_{t_i}+\kappa\Delta t x_{t_i}-\kappa\Delta t(p\theta+(1-p)a)\right]^2\right\}.\end{aligned} \tag{25}$$

The cumulative density F_p can be obtained similarly. The remaining procedures are the same as those in the estimation of the skew level. Finally, the $(i+1)$-th sample $p^{(i+1)}$ equals $F_p^{-1}\left(U_p\right)$, and the new grid points are $\left\{F_p^{-1}(\xi_1),\ldots,F_p^{-1}(\xi_n)\right\}$.

3.2 Estimation of the Doubly Skewed OU Process

For the doubly skewed OU process, the discretized version of Eq.(11) can be expressed as

$$\Delta Y_{t_i}=\begin{cases}\kappa[p_1p_2\theta+(\omega_1-p_1p_2)a_1+(1-\omega_1)a_2]\Delta t-\kappa\Delta tY_{t_i}+p_1p_2\sigma\sqrt{\Delta t}\epsilon_{t_i}, \\ \qquad \textit{if}\ Y_{t_i}<\omega_1a_1+(1-\omega_1)a_2, \\ \kappa[(\omega_1-p_1)\theta+p_1a_1+(1-\omega_1)a_2]\Delta t-\kappa\Delta tY_{t_i}+(\omega_1-p_1)\sigma\sqrt{\Delta t}\epsilon_{t_i}, \\ \qquad \textit{if}\ \omega_1a_1+(1-\omega_1)a_2\le Y_{t_i}<p_1a_1+(1-p_1)a_2, \\ \kappa[(1-\omega_1)\theta+p_1a_1+(\omega_1-p_1)a_2]\Delta t-\kappa\Delta tY_{t_i}+(1-\omega_1)\sigma\sqrt{\Delta t}\epsilon_{t_i}, \\ \qquad \textit{if}\ p_1a_1+(1-p_1)a_2\le Y_{t_i}.\end{cases} \tag{26}$$

where $\Delta Y_{t_i}=Y_{t_{i+1}}-Y_{t_i}$ and $\{\epsilon_{t_i}\}_{i=1}^N$ are independent and standard normally distributed. By setting $y_{t_i}\triangleq\Delta Y_{t_i}$ and $x_{t_i}\triangleq Y_{t_i}$, Eq.(26) becomes

$$y_{t_i} = \begin{cases} \kappa[p_1p_2\theta + (\omega_1 - p_1p_2)a_1 + (1-\omega_1)a_2]\triangle t - \kappa\triangle t x_{t_i} + p_1p_2\sigma\sqrt{\triangle t}\epsilon_{t_i}, \\ \quad \text{if } x_{t_i} < \omega_1 a_1 + (1-\omega_1)a_2, \\ \kappa[(\omega_1 - p_1)\theta + p_1a_1 + (1-\omega_1)a_2]\triangle t - \kappa\triangle t x_{t_i} + (\omega_1 - p_1)\sigma\sqrt{\triangle t}\epsilon_{t_i}, \\ \quad \text{if } \omega_1 a_1 + (1-\omega_1)a_2 \le x_{t_i} < p_1a_1 + (1-p_1)a_2, \\ \kappa[(1-\omega_1)\theta + p_1a_1 + (\omega_1 - p_1)a_2]\triangle t - \kappa\triangle t x_{t_i} + (1-\omega_1)\sigma\sqrt{\triangle t}\epsilon_{t_i}, \\ \quad \text{if } p_1a_1 + (1-p_1)a_2 \le x_{t_i}. \end{cases} \tag{27}$$

Denote by N_1, N_2, and N_3 the sets as follows

$$N_1 \triangleq \{i : i = 1, \ldots, N, x_{t_i} < \omega_1 a_1 + (1-\omega_1)a_2\},$$
$$N_2 \triangleq \{i : i = 1, \ldots, N, \omega_1 a_1 + (1-\omega_1)a_2 \le x_{t_i} < p_1a_1 + (1-p_1)a_2\},$$
$$N_3 \triangleq \{i : i = 1, \ldots, N, p_1a_1 + (1-p_1)a_2 \le x_{t_i}\}.$$

Let n_1, n_2, and n_3 be the amounts of i included in the sets N_1, N_2, and N_3, respectively. We have $n_1 + n_2 + n_3 = N$. Then, the likelihood function of Eq.(27) is

$$\begin{aligned} l(X|\Theta) = & \left(\frac{1}{p_1p_2\sqrt{2\pi\sigma^2\triangle t}}\right)^{n_1} \exp\left\{-\frac{1}{2p_1^2p_2^2\sigma^2\triangle t}\right. \\ & \left.\times \sum_{i\in N_1}[y_{t_i} + \kappa\triangle t x_{t_i} - \kappa\triangle t(p_1p_2\theta + (\omega_1 - p_1p_2)a_1 + (1-\omega_1)a_2)]^2\right\} \\ & \times \left(\frac{1}{(\omega_1 - p_1)\sqrt{2\pi\sigma^2\triangle t}}\right)^{n_2} \exp\left\{-\frac{1}{2(\omega_1 - p_1)^2\sigma^2\triangle t}\right. \\ & \left.\times \sum_{i\in N_2}[y_{t_i} + \kappa\triangle t x_{t_i} - \kappa\triangle t((\omega_1 - p_1)\theta + p_1a_1 + (1-\omega_1)a_2)]^2\right\} \\ & \times \left(\frac{1}{(1-\omega_1)\sqrt{2\pi\sigma^2\triangle t}}\right)^{n_3} \exp\left\{-\frac{1}{2(1-\omega_1)^2\sigma^2\triangle t}\right. \\ & \left.\times \sum_{i\in N_3}[y_{t_i} + \kappa\triangle t x_{t_i} - \kappa\triangle t((1-\omega_1)\theta + p_1a_1 + (\omega_1 - p_1)a_2)]^2\right\}, \end{aligned} \tag{28}$$

where Θ is the set of seven parameters $\{\kappa, \theta, \sigma^2, a_1, p_1, a_2, p_2\}$ for the doubly skewed OU process, and X represents the data set. Next, we find the conditional distribution of all seven parameter to estimate them by Eq.(16).

3.2.1 Conditional Distribution of the Instantaneous Return

Similar to the skew OU process, conditional on θ, σ, a_1, a_2, p_1 and p_2, the conjugate prior distribution of mean-reversion speed κ is chosen to be a normal distribution $\mathcal{N}\left(\mu_\kappa, \sigma_\kappa^2\right)$. Together with the prior distribution and the likelihood function Eq.(28), we have the posterior distribution of κ as follows:

$$\kappa|X, \theta, \sigma^2, a_1, a_2, p_1, p_2 \sim \mathcal{N}\left(\widehat{\mu}_\kappa, \widehat{\sigma}_\kappa^2\right), \tag{29}$$

where

$$\begin{aligned}
\widehat{\mu}_\kappa =& \left\{ \frac{1}{p_1^2 p_2^2 \sigma^2} \sum_{i \in N_1} y_{t_i} \{[-x_{t_i} + p_1 p_2 \theta + (\omega_1 - p_1 p_2) a_1 + (1-\omega_1) a_2]\} \right. \\
&+ \frac{1}{(\omega_1 - p_1)^2 \sigma^2} \sum_{i \in N_2} y_{t_i} \{[-x_{t_i} + (\omega_1 - p_1)\theta + p_1 a_1 + (1-\omega_1) a_2]\} \\
&\left. + \frac{1}{(1-\omega_1)^2 \sigma^2} \sum_{i \in N_3} y_{t_i} \{[-x_{t_i} + (1-\omega_1)\theta + p_1 a_1 + (\omega_1 - p_1) a_2]\} + \frac{\mu_\kappa}{\sigma_\kappa^2} \right\} \widehat{\sigma}_\kappa^2, \\
\widehat{\sigma}_\kappa^{-2} =& \frac{1}{p_1^2 p_2^2 \sigma^2} \sum_{i \in N_1} \{[-x_{t_i} + p_1 p_2 \theta + (\omega_1 - p_1 p_2) a_1 + (1-\omega_1) a_2]\}^2 \triangle t \\
&+ \frac{1}{(\omega_1 - p_1)^2 \sigma^2} \sum_{i \in N_2} \{[-x_{t_i} + (\omega_1 - p_1)\theta + p_1 a_1 + (1-\omega_1) a_2]\}^2 \triangle t \\
&+ \frac{1}{(1-\omega_1)^2 \sigma^2} \sum_{i \in N_3} \{[-x_{t_i} + (1-\omega_1)\theta + p_1 a_1 + (\omega_1 - p_1) a_2]\}^2 \triangle t + \frac{1}{\sigma_\kappa^2}.
\end{aligned}$$

Given κ, σ, a_1, a_2, p_1 and p_2, the proper prior distribution of the long-term mean θ is also a normal distribution $\mathcal{N}(\mu_\theta, \sigma_\theta^2)$. The posterior distribution of θ is then

$$\theta|X, \kappa, \sigma^2, a_1, a_2, p_1, p_2 \sim \mathcal{N}\left(\widehat{\mu}_\theta, \widehat{\sigma}_\theta^2\right), \tag{30}$$

where

$$\begin{aligned}
\widehat{\mu}_\theta =& \left\{ \frac{1}{p_1 p_2 \sigma^2} \sum_{i \in N_1} [\kappa(y_{t_i} + \kappa \triangle t x_{t_i} - \kappa \triangle t((\omega_1 - p_1 p_2) a_1 + (1-\omega_1) a_2))] \right. \\
&+ \frac{1}{(\omega_1 - p_1)\sigma^2} \sum_{i \in N_2} [\kappa(y_{t_i} + \kappa \triangle t x_{t_i} - \kappa \triangle t(p_1 a_1 + (1-\omega_1) a_2))] \\
&\left. + \frac{1}{(1-\omega_1)\sigma^2} \sum_{i \in N_3} [\kappa(y_{t_i} + \kappa \triangle t x_{t_i} - \kappa \triangle t(p_1 a_1 + (\omega_1 - p_1) a_2))] + \frac{\mu_\theta}{\sigma_\theta^2} \right\} \widehat{\sigma}_\theta^2, \\
\widehat{\sigma}_\theta^{-2} =& \frac{\kappa^2 \triangle t N}{\sigma^2} + \frac{1}{\sigma_\theta^2}.
\end{aligned}$$

Next, we obtain the conditional posterior distribution of κ and θ in Eq.(11) and can estimate them by the method shown in Eq.(16).

3.2.2 Conditional Distribution of the Volatility

Conditional on κ, θ, a_1, a_2, p_1 and p_2, the prior distribution of instantaneous variance σ^2 is still an inverted Gamma distribution $\mathcal{IG}(\alpha_\sigma, \lambda_\sigma)$. Thus, the posterior distribution of σ^2 is

$$\sigma^2 | X, \kappa, \theta, a_1, a_2, p_1, p_2 \sim \mathcal{IG}\left(\widehat{\alpha}_\sigma, \widehat{\lambda}_\sigma\right), \tag{31}$$

where

$$\begin{aligned}
\widehat{\alpha}_\sigma &= \frac{N}{2} + \alpha_\sigma, \\
\widehat{\lambda}_\sigma &= \frac{1}{2p_1^2 p_2^2 \Delta t} \sum_{i \in N_1} \{y_{t_i} + \kappa \Delta t x_{t_i} - \kappa[p_1 p_2 \theta + (\omega_1 - p_1 p_2) a_1 + (1 - \omega_1) a_2] \Delta t\}^2 \\
&\quad + \frac{1}{2(\omega_1 - p_1)^2 \Delta t} \sum_{i \in N_2} \{y_{t_i} + \kappa \Delta t x_{t_i} - \kappa[(\omega_1 - p_1)\theta + p_1 a_1 + (1 - \omega_1) a_2] \Delta t\}^2 \\
&\quad + \frac{1}{2(1 - \omega_1)^2 \Delta t} \sum_{i \in N_3} \{y_{t_i} + \kappa \Delta t x_{t_i} - \kappa[(1 - \omega_1)\theta + p_1 a_1 + (\omega_1 - p_1) a_2] \Delta t\}^2 + \lambda_\sigma.
\end{aligned}$$

Then, we can easily estimate σ^2 in Eq.(11).

3.2.3 Conditional Distribution of the Skew Level

The Griddy-Gibbs sampler is used to estimate the skew level a_1 (resp. a_2) as before. Assume that a_1 is uniform on a predetermined interval (a_{1d}, a_{1u}) from past observations. Then, the skew level a_1 takes values at the n grid points $\{a_{11}, \ldots, a_{1n}\}$ on the interval. Conditional on κ, θ, σ, a_2, p_1 and p_2, the posterior density of the skew level a_1 satisfies

$$
\begin{aligned}
\pi\big(a_1|X,\kappa,\theta,\sigma^2,a_2,p_1,p_2\big) \quad &\propto \\
&\left(\frac{1}{p_1p_2\sqrt{2\pi\sigma^2\triangle t}}\right)^{n_1}\exp\Bigg\{-\frac{1}{2p_1^2p_2^2\sigma^2\triangle t} \\
&\times\sum_{i\in N_1}[y_{t_i}+\kappa\triangle tx_{t_i}-\kappa\triangle t(p_1p_2\theta+(\omega_1-p_1p_2)a_1+(1-\omega_1)a_2)]^2\Bigg\} \\
&\times\left(\frac{1}{(\omega_1-p_1)\sqrt{2\pi\sigma^2\triangle t}}\right)^{n_2}\exp\Bigg\{-\frac{1}{2(\omega_1-p_1)^2\sigma^2\triangle t} \\
&\times\sum_{i\in N_2}[y_{t_i}+\kappa\triangle tx_{t_i}-\kappa\triangle t((\omega_1-p_1)\theta+p_1a_1+(1-\omega_1)a_2)]^2\Bigg\} \\
&\times\left(\frac{1}{(1-\omega_1)\sqrt{2\pi\sigma^2\triangle t}}\right)^{n_3}\exp\Bigg\{-\frac{1}{2(1-\omega_1)^2\sigma^2\triangle t} \\
&\times\sum_{i\in N_3}[y_{t_i}+\kappa\triangle tx_{t_i}-\kappa\triangle t((1-\omega_1)\theta+p_1a_1+(\omega_1-p_1)a_2)]^2\Bigg\}\frac{1}{a_{1u}-a_{1d}}.
\end{aligned}
\tag{32}
$$

The generation of the $(i+1)$-th sample $a_1^{(i+1)}$ for the doubly skewed OU process is the same as that for the skew OU process. Finally, the procedures for skew level a_2 are completely analogous to those for skew level a_1.

3.2.4 Conditional Distribution of the Skew Probability

The skew probability p_1 is assumed to be uniform on the interval $[0,1]$ and the grid points are assumed to be $\{p_1^1,\ldots,p_1^n\}$. For each $p_1\in\{p_1^1,\ldots,p_1^n\}$, given κ, θ, σ, a_1, a_2 and p_2, we obtain the following relation:

$$
\begin{aligned}
\pi\big(p_1|X,\kappa,\theta,\sigma^2,a_1,a_2,p_2\big) \quad &\propto \\
&\left(\frac{1}{p_1p_2\sqrt{2\pi\sigma^2\triangle t}}\right)^{n_1}\exp\Bigg\{-\frac{1}{2p_1^2p_2^2\sigma^2\triangle t} \\
&\times\sum_{i\in N_1}[y_{t_i}+\kappa\triangle t x_{t_i}-\kappa\triangle t(p_1p_2\theta+(\omega_1-p_1p_2)a_1+(1-\omega_1)a_2)]^2\Bigg\} \\
&\times\left(\frac{1}{(\omega_1-p_1)\sqrt{2\pi\sigma^2\triangle t}}\right)^{n_2}\exp\Bigg\{-\frac{1}{2(\omega_1-p_1)^2\sigma^2\triangle t} \\
&\times\sum_{i\in N_2}[y_{t_i}+\kappa\triangle t x_{t_i}-\kappa\triangle t((\omega_1-p_1)\theta+p_1a_1+(1-\omega_1)a_2)]^2\Bigg\} \\
&\times\left(\frac{1}{(1-\omega_1)\sqrt{2\pi\sigma^2\triangle t}}\right)^{n_3}\exp\Bigg\{-\frac{1}{2(1-\omega_1)^2\sigma^2\triangle t} \\
&\times\sum_{i\in N_3}[y_{t_i}+\kappa\triangle t x_{t_i}-\kappa\triangle t((1-\omega_1)\theta+p_1a_1+(\omega_1-p_1)a_2)]^2\Bigg\}.
\end{aligned} \tag{33}
$$

The remaining steps are quite similar to those for the skew OU process. The sample generating procedures of the other skew probability p_2 are analogous to those of the skew probability p_1.

3.3 Detecting Convergence

We also care whether the Markov chain of parameters generated from the Gibbs sampler converges to the real posterior density function $\pi(\Theta|X)$.

Generally, there are two ways to detect convergence. Gelfand and Smith (1990) and Gelfand et al. (1990) suggest choosing from m independent Gibbs sequences $\{G^{(1)},\ldots,G^{(m)}\}$, and then a proper length l is selected so that all l-th runs in each sequence form the Markov chain $\{G^{(i)}(l), i=1,\ldots,m\}$. In this way, the Gibbs sequence l does not need to be very long. However, Gelman and Rubin (1991) stated that this approach is not foolproof.

An alternative method is to generate a long Gibbs sequence $\{G(i), i=1,\ldots,d+m\}$. Geweke (1991) and Nakatsuma (2000) provide a practical perspective on the convergence problem. In their work, after discarding the first d runs in the sequence, the convergence can be tested by comparing the first m_1 values in the remainder with the last m_2 values. Formally, the convergence diagnostic (CD) is defined as

$$
CD=\left(m_1^{-1}\sum_{i=1+d}^{m_1+d}G(i)-m_2^{-1}\sum_{i=m+d-m_2+1}^{m+d}G(i)\right)\Big/\left[\widehat{S}_G^1(0)/m_1+\widehat{S}_G^2(0)/m_2\right]^{\frac{1}{2}},
$$

where $\widehat{S}_G^i(\cdot)$ is the spectrum density estimate for m_i runs. Let $(m_1+m_2)/m<1$ and fix the ratios m_1/m and m_2/m. If the sequence $\{G(i), i = d+1, \ldots, d+m\}$ is stationary, then applying the central limit theorem, we have

$$CD \to \mathcal{N}(0,1) \quad \text{when} \quad m \to \infty.$$

In this paper, we adopt the latter method and set the ratios m_1/m and m_2/m to be 0.1 and 0.5, as in Geweke (1991).

3.4 Error Analysis

The standard error is also important for evaluating the estimation. Since it is difficult to derive the error directly, we give three proxy errors in this paper.

The first error (Error1) is similar to that in Franke et al. (2002). Recalling the estimates Θ and $\{x_{t_i}, i = 1\ldots N\}$, we generate the bootstrap samples $\left\{y^*_{t_i}, i = 1\ldots N\right\}$ with new random noise $\left\{\epsilon^*_{t_i}, i = 1\ldots N\right\}$ from Eq.(18) for the skew OU process, or Eq.(27) for the doubly skewed OU process. Then we obtain estimates Θ^* by using the Gibbs sampler based on the bootstrap samples $\left\{(x_{t_i}, y^*_{t_i}), i = 1, \ldots, N\right\}$. After a sufficient number of simulations, the standard errors of the parameters, Error1, are found.

The second error (Error2) is used to assess the accuracy for the Gibbs sampler. For the same samples $\{(x_{t_i}, y_{t_i}), i = 1, \ldots, N\}$, if we generate a sufficient number of Gibbs samplers $\left\{G^{(i)}\right\}$, then $\Theta^{(i)}$ is estimated, and the standard deviation can be evaluated.

The last error (Error3) is the simplest way to measure the accuracy for the Gibbs sampler. For the Gibbs sampler $\{G(i), i = 1, \ldots, d+m\}$, Error3 is the sample standard deviation of $\{G(i), i = d+1, \ldots, d+m\}$.

4 Simulations and Empirical Analysis

4.1 Simulation Experiments

In this subsection, we show how the Bayesian inference works and then test the effectiveness of the proposed method by simulating data sets generated from the skew OU process or the doubly skewed OU process.

Without loss of generality, we set $\kappa = 0.15$, $\theta = 8$, $\sigma = 0.2$, and $a = 8.44$ for the skew OU process defined in Eq.(1) and generate 1000 weekly observations for different values of skew probability p. We first show how the value of skew probability p influences the behavior of the sample path of the skew OU process. As exhibited in Fig. 1, when p is equal to 0 (resp. 1), the process becomes a reflected OU process. This means that when the sample path hits the skew level, it will move downwards (resp. upwards) with certainty. When p is equal to 0.2 (resp. 0.8), it is obvious that it is hard for the sample path to move upwards (resp. downwards) when hitting the skew level. When p is equal to 0.5, the skew OU process reduces to the

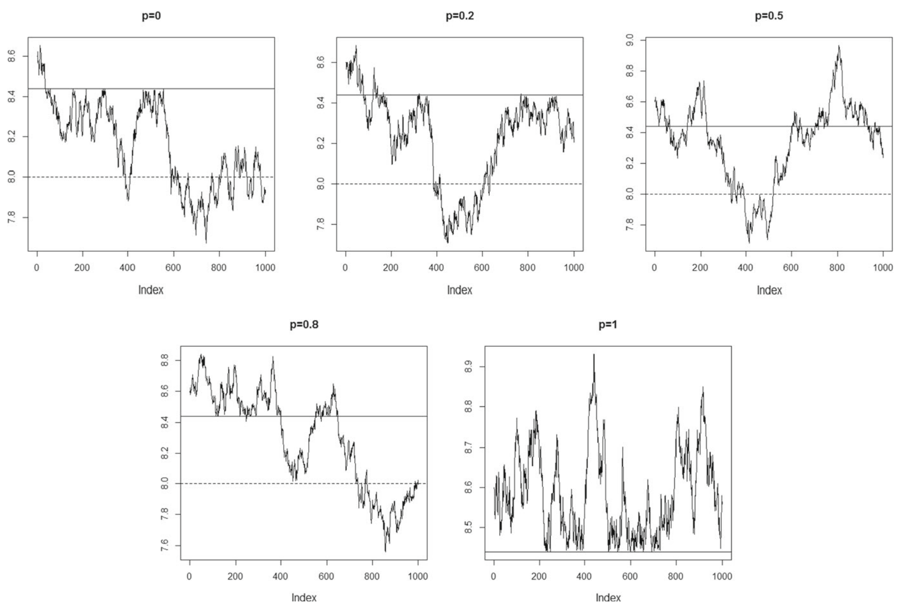

Fig. 1 Simulation plots of the skew OU process. Solid - the skew level, dotted - the long-term mean

standard OU process, and no obvious trend can be found after the sample path hits the skew level.

Here, we let p be 0.2 to show the application of our proposed algorithm. In our estimation, the parameter samples are iterated from the Gibbs sampler. To see how the parameter samples converge and to decide the number of iteration runs, Fig. 2 presents the trace plots of the five parameters for 5000 iteration runs in the skew OU process. We can see that after 200 iterations, the samples of the parameter values are quite stable and change very little, indicating that the Markov chains have attained convergence. Then, Table 1 displays the Bayesian estimates for two alternative iteration runs: 4000 iterations after 1000 burn-in cycles and 800 iterations after 200 burn-in cycles. Table 1 shows that all of the parameter estimates are very close to their true values, and the sample standard deviations (Error3 defined in Sect. 3.4) of the sequences are very small. The CD statistic suggests that convergence is attained for both Markov chains. Therefore, to improve efficiency, all the results shown below are based on 800 iterations after 200 burn-in cycles.

Table 2 reports our main Bayesian estimation results for both the skew OU process and the doubly skewed OU process. In addition, maximum likelihood estimation (MLE) inferences are also provided. In the MLE estimates[1], the simulated samples are assumed to follow an OU process; hence, only the parameters κ, θ and σ are estimated. In the MLE estimates[2], the simulated samples are assumed to follow the skew OU process; we can obtain all five parameters by the revised method of Su and Chan (2015) introduced in Sect. 3.1. Overall, our proposed Bayesian method gives satisfactory estimates for both of these two skew-extended OU processes. A comparison between the Baysian estimates and MLE estimates[1]

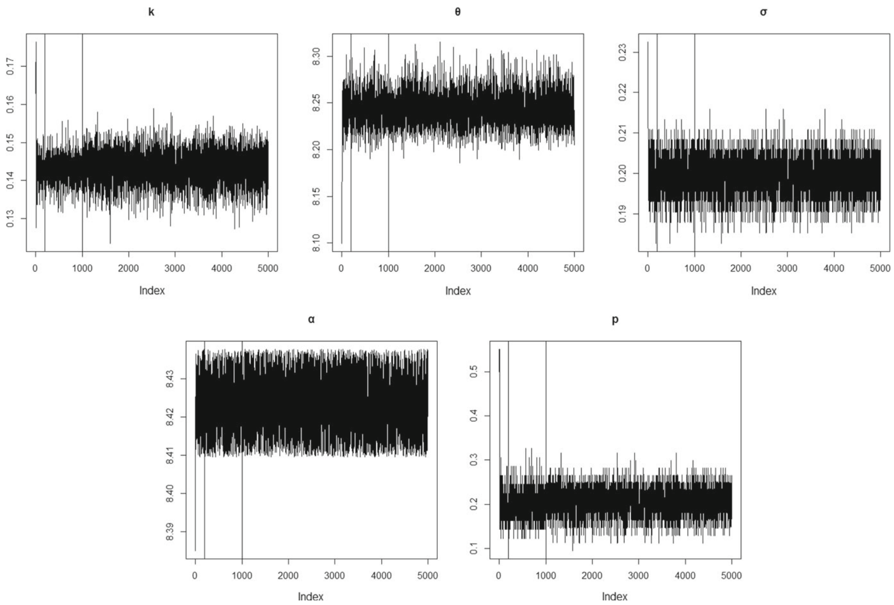

Fig. 2 Trace plots of the parameters for the skew OU process. Two vertical solid lines—200 burn-in cycle and 1000 burn-in cycle

Table 1 The Bayesian estimates of the skew OU process for different iteration runs

Parameter	True values	Bayesian estimates (5000 runs) Mean	Median	CD	Error3	Bayesian estimates (1000 runs) Mean	Median	CD	Error3
κ	0.15	0.1425	0.1427	0.3575	0.0036	0.1428	0.1428	0.3767	0.0036
θ	8	8.2332	8.2321	− 0.9574	0.0134	8.2383	8.2363	− 0.5022	0.0133
σ	0.2	0.1998	0.1997	− 0.0916	0.0046	0.1993	0.1991	− 0.0020	0.0044
a	8.44	8.4235	8.4235	0.3655	0.0156	8.4377	8.4251	0.4479	0.0152
p	0.2	0.2064	0.2041	0.8849	0.0346	0.1989	0.2041	0.5133	0.0339

This table reports the Bayesian estimates for all five parameters in the skew OU process. CD is the convergence diagnostic statistic defined in Sect. 3.3. Error3 is the sample standard deviation of the Gibbs sampler $\{G(i), i = d+1, \ldots, d+m\}$, where d is 1000 and m is 4000 for 5000 runs, and d is 200 and m is 800 for 1000 runs

shows that the skew OU process indeed fits the simulated samples better than the OU process. On the other hand, The Bayesian estimates work as well as MLE estimates[2] for the traditional drift and volatility coefficients, i.e., κ, θ and σ; but the Bayesian estimates are better than MLE estimates[2], for the pair of skew parameters (p, a), especially for the skew probability p.

When the data follow a standard OU process, we would like to know if the proposed estimation approaches for the skew-extended OU processes can still give the correct parameter values, in other words, if the skew probability estimates are

Table 2 Estimation results for the skew-extended OU simulated samples

Parameter	True values	MLE estimates[1]	MLE estimates[2]	Bayesian estimates Mean	Median	CD	Error1	Error2	Error3
Panel A: Using the skew OU process to estimate the skew OU simulated samples									
κ	0.15	0.1632	0.1456	0.1428	0.1428	0.3767	0.0236	0.0016	0.0036
θ	8	8.1203	8.2145	8.2383	8.2363	− 0.5022	0.1803	0.0134	0.0133
σ	0.2	0.2067	0.1997	0.1993	0.1991	− 0.0020	0.0051	0.0001	0.0044
a	8.44	–	8.4383	8.4377	8.4251	0.4479	0.0784	0.0002	0.0152
p	0.2	–	0.2561	0.1989	0.2041	0.5133	0.0945	0.0011	0.0339
Panel B: Using the doubly skewed OU process to estimate the doubly OU samples									
κ	0.15	0.1643	–	0.1465	0.1505	0.6699	0.0462	0.0023	0.0295
θ	8	7.7398	–	7.9319	7.9310	− 1.2092	0.6629	0.0216	0.1010
σ	0.2	0.2095	–	0.2590	0.2582	0.8137	0.0242	0.0006	0.0190
a_1	7.5	–	–	7.4851	7.4862	− 1.0602	0.1117	0.0512	0.2696
p_1	0.8	–	–	0.6728	0.6624	1.3448	0.0067	0.0037	0.0278
a_2	8.2	–	–	8.1909	8.1913	1.1435	0.0047	0.0001	0.0041
p_2	0.2	–	–	0.2749	0.2768	− 0.6158	0.0200	0.0012	0.0387

This table reports the Bayesian estimates of both the skew OU process and the doubly skewed OU process. MLE estimates[1] are the maximum likelihood estimation results assuming that the skew OU simulated samples follow an OU process. MLE estimates[2] are the maximum likelihood estimation results by using revised method of Su and Chan (2015) introduced in Sect. 3.1 assuming that the skew OU simulated samples follow the skew OU process. CD is the convergence diagnostic statistic defined in Sect. 3.3. Error1, Error2 and Error3 are three proxies for the sample standard deviation defined in Sect. 3.4

approximately 0.5. Table 3 shows the estimation results in this case. It is obvious that the Bayesian estimates are as good as the MLE estimates[1]. Moreover, the Bayesian inferences of the skew probabilities p, p_1, and p_2 are very close to the true value 0.5. Therefore, our skew-extended OU estimation method also works well for the standard OU samples and would not mistakenly regard an OU process as the skew-extended OU process. Notably, the MLE estimates[2] of the skew probability p for the skew OU process, 0.4286, deviates a bit far from the true value 0.5, indicating that we should be cautious when interpreting an non-extreme value of skew probability from MLE estimates[2].

We are also interested in the relation between the skew OU process and the doubly skewed OU process. If the simulated data are based on the skew OU process, Table 4 gives the estimation results of the doubly skewed OU process. We can see that the skew OU process can be considered as a special case of the doubly skewed OU process. The estimation of the pair of skew parameters (p_1, a_1) in the doubly skewed OU process is very close to the pair (p, a) in the skew OU process, while the other skew probability p_2 is almost equal to 0.5, indicating that the (single) skew OU process is sufficient to capture the dynamics. By contrast, Table 5 shows that if the data are generated by the doubly skewed OU process, the skew OU process does not perform well. The estimates of the pair (p, a) in the skew OU process seem to

Table 3 Estimation results of the skew-extended OU processes for the OU simulated samples

Parameter	True values	MLE estimates[1]	MLE estimates[2]	Bayesian estimates Mean	Median	CD	Error1	Error2	Error3
Panel A: Using the skew OU process to estimate the OU simulated samples									
κ	0.15	0.1398	0.1477	0.1409	0.1407	− 0.2313	0.0841	0.0004	0.0085
θ	8	8.0997	8.1596	8.1046	8.0987	− 1.2167	0.2485	0.0009	0.0211
σ	0.2	0.2034	0.2037	0.2030	0.2030	0.7393	0.0045	0.0006	0.0045
a	–	–	8.2615	8.6718	8.8442	1.3777	0.2096	0.0214	0.4741
p	0.5	–	0.4286	0.5122	0.4898	1.5554	0.0649	0.0050	0.1492
Panel B: Using the doubly skewed OU process to estimate the OU simulated samples									
κ	0.15	0.1634	–	0.2088	0.1950	0.6580	0.0428	0.0031	0.1098
θ	8	7.9735	–	8.3967	8.3305	− 1.5918	0.0580	0.0219	0.3894
σ	0.2	0.2054	–	0.2582	0.0.2841	0.2468	0.0236	0.0016	0.1086
a_1	–	–	–	8.5655	8.6730	− 1.0481	0.2203	0.0158	0.7207
p_1	0.5	–	–	0.4928	0.5055	− 1.5396	0.0417	0.0030	0.0928
a_2	–	–	–	8.9081	9.1272	1.2538	0.0296	0.0173	0.5365
p_2	0.5	–	–	0.4518	0.4951	0.3951	0.0419	0.0038	0.0970

This table reports the Bayesian estimates of the skew-extended OU process for the standard OU simulated samples. MLE estimates[1] are the maximum likelihood estimation results assuming that the OU simulated samples follow the OU process. MLE estimates[2] are the maximum likelihood estimation results by using revised method of Su and Chan (2015) introduced in Sect. 3.1 assuming that the OU simulated samples follow a skew OU process. CD is the convergence diagnostic statistic defined in Sect. 3.3. Error1, Error2 and Error3 are three proxies for the sample standard deviation defined in Sect. 3.4

Table 4 Estimation results of the doubly skewed OU process for the skew OU simulated samples

Parameter	True values	Skew OU process		Doubly skewed OU process					
		Mean	Median	Mean	Median	CD	Error1	Error2	Error3
κ	0.15	0.1428	0.1428	0.1826	0.1844	0.9179	0.0386	0.0042	0.0510
θ	8	8.2383	8.2363	8.2374	8.2505	0.9409	0.5454	0.0178	0.1418
σ	0.2	0.1993	0.1991	0.2888	0.2856	0.9833	0.0227	0.0028	0.0387
a_1	8.44	8.4377	8.4251	8.4213	8.4204	0.4872	0.0044	0.00276	0.0177
p_1	0.2	0.1989	0.2041	0.2544	0.2465	− 0.2282	0.0196	0.0071	0.1049
a_2	–	–	–	8.3778	8.4226	− 0.7663	0.1179	0.0248	0.1740
p_2	–	–	–	0.5186	0.5106	− 0.1885	0.0070	0.0183	0.1451

This table reports the Bayesian estimates of the doubly skewed OU processes for the skew OU simulated samples. CD is the convergence diagnostic statistic defined in Sect. 3.3. Error1, Error2 and Error3 are three proxies for the sample standard deviation defined in Sect. 3.4

exhibit the combined effect of the pairs (p_1, a_1) and (p_2, a_2) in the doubly skewed OU process. Therefore, one implication is that we can always first use the doubly skewed OU process to estimate parameters. If one skew probability in the doubly skewed OU process is approximately 0.5, we can then apply the skew OU process to that data. However, the skew components estimated by the MLE estimates[2] are a bit confusing. This MLE method gives a skew level, 7.5819, which is close to the (true) support skew level 7.5, hence we anticipate that the MLE estimates[2] for the skew probability should be larger than 0.5. But the resulting estimated skew probability is less than 0.5, such that the skew level becomes a counterintuitive resistance line. As a result, one cannot count on MLE estimates[2] too much for the doubly skewed OU process.

4.2 Real Data Applications

We now illustrate our method by analyzing real market data, which consist of the Hong Kong Interbank Offered Rate (HIBOR) and the CBOE volatility index (VIX).

4.2.1 Overnight HIBOR

The yields of the overnight HIBOR are used, and the data sets include 2,830 daily observations, from January 4, 2006, to December 29, 2017. Since the interest rates are changing over time, we estimated the skew-extended processes for each year. Tables 6 and 7 present the estimates of the skew-extended OU processes for the overnight HIBOR. The CD statistics, reported in square brackets, show that the Markov chains attained convergence after 200 iterations. As can be seen, it is common for the skew probability to be quite different from 0.5 for the overnight HIBOR. We name such observations the *skew phenomenon*. Therefore, in a market that exhibits skew phenomena, skew-extended OU processes are very useful. When

Table 5 Estimation results of the skew OU process for the doubly skewed OU simulated samples

Parameter	True	Doubly skewed OU		MLE	Skew OU process					
	values	Mean	Median	estimates^2	Mean	Median	CD	Error1	Error2	Error3
κ	0.15	0.1465	0.1505	0.1697	0.1492	0.1491	− 0.2293	0.0128	0.0003	0.0073
θ	8	7.9319	7.9310	7.7797	7.7999	7.7991	− 0.4465	0.2872	0.0003	0.0238
σ	0.2	0.2590	0.2582	0.2097	0.2086	0.2084	− 0.5692	0.0066	0.0003	0.0048
a_1	7.5	7.4851	7.4862	7.5819	8.1560	8.1549	1.0331	0.2685	0.0003	0.0089
p_1	0.8	0.6728	0.6624	0.4153	0.3933	0.3908	− 0.1739	0.1737	0.0008	0.0357
a_2	8.2	8.1909	8.1913	–	–	–	–	–	–	–
p_2	0.2	0.2749	0.2768	–	–	–	–	–	–	–

This table reports the Bayesian estimates of the skew OU processes for the doubly skewed OU simulated samples. MLE estimates^2 are the maximum likelihood estimation results by using revised method of Su and Chan (2015) introduced in Sect. 3.1 assuming that the doubly skewed OU simulated samples follow a skew OU process. CD is the convergence diagnostic statistic defined in Sect. 3.3. Error1, Error2 and Error3 are three proxies for the sample standard deviation defined in Sect. 3.4

Table 6 Estimation results of the drift and diffusion components for the overnight HIBOR

Year	Skew OU-MLE			Skew OU-Bayesian			Doubly skewed OU-Bayesian		
	κ	θ	σ	κ	θ	σ	κ	θ	σ
2006	10.2233	3.7724	1.4699	46.4759	3.7769	2.6797	34.2393	3.8815	3.3185
	(0.7266)	(0.3794)	(0.2356)	(0.5241)	(0.0105)	(0.1335)	(17.2572)	(0.1538)	(0.2772)
				[0.7236]	[− 0.5724]	[1.5471]	[− 1.1634]	[1.1826]	[− 0.8897]
2007	2.0969	3.5615	2.1261	10.4372	3.7218	4.1298	8.7486	3.8455	4.2965
	(0.1714)	(0.3506)	(0.1244)	(0.3669)	(0.0241)	(0.1978)	(0.8952)	(0.4317)	(0.1744)
				[0.7236]	[− 0.5724]	[1.5471]	[− 0.4662]	[1.5164]	[0.3639]
2008	3.4576	1.2417	1.8791	11.1261	1.2675	3.6844	13.3643	1.0537	4.9442
	(0.1062)	(0.0260)	(0.3031)	(1.9167)	(0.0483)	(0.1662)	(1.9100)	(0.2390)	(0.5731)
				[− 1.1303]	[1.0633]	[0.0603]	[0.7430]	[− 0.2953]	[− 0.0402]
2009	16.9304	0.0602	0.2726	53.4416	0.0616	0.4259	4.0858	0.1176	0.4522
	(2.9598)	(0.0085)	(0.0370)	(3.7007)	(0.0005)	(0.0191)	(3.9587)	(0.0379)	(0.0365)
				[− 1.1303]	[1.0633]	[0.0603]	[− 1.0432]	[− 0.6134]	[0.2059]
2010	9.1983	0.0547	0.1299	48.303	0.0499	0.1978	2.5108	0.0655	0.2271
	(0.7597)	(0.0129)	(0.0301)	(1.9030)	(0.0007)	(0.0142)	(1.5967)	(0.0378)	(0.0647)
				[0.3231]	[− 0.5760]	[− 1.0944]	[0.2337]	[− 1.4980]	[− 1.8296]
2011	13.3052	0.0624	0.2884	92.4979	0.0575	0.4418	40.5394	0.0017	0.2454
	(3.7538)	(0.0078)	(0.0502)	(8.1715)	(0.0017)	(0.1269)	(2.8433)	(0.0007)	(0.0181)
				[− 1.6478]	[1.4062]	[0.9136]	[1.0531]	[− 1.3316]	[1.3243]
2012	10.8012	0.1039	0.0838	14.4829	0.099	0.0619	12.1195	0.1081	0.0792
	(0.4150)	(0.0183)	(0.0104)	(0.2348)	(0.0005)	(0.0042)	(5.3242)	(0.0076)	(0.0057)
				[1.1983]	[− 0.4957]	[0.1621]	[− 0.0531]	[1.0993]	[− 0.9382]

Table 6 continued

Year	Skew OU-MLE			Skew OU-Bayesian			Doubly skewed OU-Bayesian		
	κ	θ	σ	κ	θ	σ	κ	θ	σ
2013	3.2002	0.0813	0.0211	11.3366	0.083	0.0398	11.9582	0.0881	0.0492
	(0.7056)	(0.0026)	(0.0062)	(0.7083)	(0.0005)	(0.0019)	(4.1877)	(0.0074)	(0.0066)
				[− 0.7238]	[0.7303]	[0.2195]	[0.8845]	[− 0.0842]	[− 0.5577]
2014	11.2168	0.0575	0.1051	54.6503	0.0587	0.1817	5.3371	0.0383	0.1504
	(0.8059)	(0.0053)	(0.0004)	(2.1552)	(0.0007)	(0.0176)	(5.8207)	(0.0266)	(0.0525)
				[0.8438]	[− 1.444]	[− 0.0541]	[0.4318]	[− 0.5799]	[0.3759]
2015	19.3444	0.0550	0.0652	88.9726	0.0532	0.0937	3.388	0.0635	0.1067
	(3.0590)	(0.0006)	(0.0200)	(3.7104)	(0.0002)	(0.0066)	(1.5214)	(0.0245)	(0.0050)
				[− 0.3644]	[0.8354]	[− 0.5178]	[− 1.5004]	[0.4158]	[− 0.5525]
2016	3.9308	0.0476	0.2625	5.9661	0.2456	0.4914	3.4354	0.196	0.3738
	(0.2446)	(0.0049)	(0.0245)	(4.2544)	(0.2111)	(0.0612)	(3.2283)	(0.0927)	(0.0691)
				[0.2415]	[− 1.2504]	[0.1902]	[0.0748]	[− 0.0596]	[− 0.3690]
2017	25.9713	0.1686	2.1240	33.0972	0.0744	0.5404	41.4209	0.0037	0.5422
	(3.7091)	(0.0274)	(0.2927)	(4.5247)	(0.0006)	(0.0264)	(25.7965)	(0.0128)	(0.0286)
				[0.6573]	[− 0.4940]	[0.0684]	[− 1.2943]	[1.5050]	[− 1.3029]

The standard errors (Error3) are reported in parenthesis, the convergence diagnostic (CD) statistics are reported in square brackets

Table 7 Estimation results of the skew local time components for the overnight HIBOR

Year	Skew OU-MLE		Skew OU-Bayesian		Doubly skewed OU-Bayesian				No. of
	a	p	a	p	a_1	p_1	a_2	p_2	skew levels
2006	3.7182	0.7469	3.7245	0.7224	3.7733	0.674	4.7891	0.3949	2
	(0.0931)	(0.0651)	(0.0142)	(0.0324)	(0.0432)	(0.0476)	(0.3733)	(0.0883)	
			[0.9309]	[0.6795]	[0.7897]	[0.1377]	[0.4714]	[− 0.9148]	
2007	3.7311	0.3724	2.6843	0.2817	4.2546	0.5517	4.5436	0.3986	1
	(0.0375)	(0.0331)	(0.0754)	(0.0605)	(0.5245)	(0.1888)	(0.5195)	(0.1878)	
			[0.9309]	[0.6795]	[− 1.7973]	[0.4014]	[− 0.1719]	[− 1.7767]	
2008	2.4917	0.2694	2.4917	0.3542	2.8634	0.3419	3.0852	0.4326	1
	(0.1584)	(0.0279)	(0.6539)	(0.0912)	(0.0308)	(0.1402)	(0.0316)	(0.2003)	
			[0.1938]	[− 0.1433]	[− 0.6072]	[− 1.4178]	[0.3347]	[1.4916]	
2009	0.4449	0.1633	0.3315	0.3227	0.3376	0.4738	0.3385	0.2690	1
	(0.0951)	(0.0580)	(0.0031)	(0.0360)	(0.0035)	(0.1337)	(0.0028)	(0.1005)	
			[0.1938]	[− 0.1433]	[0.3805]	[− 1.1301]	[− 1.5663]	[1.4316]	
2010	0.0793	0.8102	0.0969	0.6899	0.0504	0.5962	0.0992	0.6793	1
	(0.0070)	(0.2375)	(0.0072)	(0.0545)	(0.0207)	(0.1058)	(0.0009)	(0.0589)	
			[− 0.8475]	[− 0.1585]	[− 1.4289]	[− 0.5497]	[1.4688]	[1.3028]	
2011	0.05	0.4949	0.1251	0.6866	0.1403	0.8298	0.5622	0.1316	2
	(0.0163)	(0.1208)	(0.0773)	(0.1269)	(0.0031)	(0.0894)	(0.0033)	(0.0533)	
			[1.3308]	[− 1.7764]	[0.5267]	[0.3122]	[− 0.8205]	[− 0.2480]	
2012	0.0982	0.8367	0.0998	0.777	0.0979	0.6837	0.1439	0.3392	2
	(0.0063)	(0.1196)	(0.0004)	(0.0338)	(0.0009)	(0.0816)	(0.0041)	(0.0639)	
			[1.7155]	[0.8903]	[− 0.4494]	[0.3220]	[− 1.7188]	[0.2825]	

Table 7 continued

Year	Skew OU-MLE		Skew OU-Bayesian		Doubly skewed OU-Bayesian				No. of
	a	p	a	p	a_1	p_1	a_2	p_2	skew levels
2013	0.0932	0.7602	0.0808	0.4607	0.0822	0.5564	0.0987	0.4295	0
	(0.0051)	(0.0618)	(0.0117)	(0.0976)	(0.0012)	(0.0558)	(0.0020)	(0.1396)	
			[− 0.1584]	[− 0.4159]	[− 0.4724]	[0.5394]	[− 1.3597]	[0.9248]	
2014	0.0594	0.7337	0.0802	0.6643	0.0471	0.5325	0.0784	0.6508	1
	(0.0116)	(0.0298)	(0.0038)	(0.0377)	(0.0191)	(0.1156)	(0.0010)	(0.0302)	
			[0.2986]	[1.6711]	[− 0.5207]	[− 0.2483]	[0.0198]	[− 0.5213]	
2015	0.0494	0.3224	0.0618	0.6415	0.0527	0.593	0.1077	0.2812	1
	(0.0015)	(0.0527)	(0.0019)	(0.0367)	(0.0004)	(0.0254)	(0.0007)	(0.0529)	
			[0.0916]	[− 0.9431]	[1.5929]	[− 0.5560]	[0.9545]	[0.2770]	
2016	0.0487	0.7867	0.332	0.6729	0.0478	0.4593	0.3336	0.5581	1
	(0.0044)	(0.1509)	(0.0496)	(0.1839)	(0.0004)	(0.1169)	(0.0031)	(0.0091)	
			[1.0582]	[0.0641]	[− 1.0000]	[− 1.3787]	[− 0.8905]	[1.4549]	
2017	0.6978	0.3092	0.6296	0.3263	0.6421	0.3232	0.6431	0.4153	1
	(0.0958)	(0.0347)	(0.0043)	(0.0311)	(0.0047)	(0.1003)	(0.0037)	(0.0974)	
			[− 0.1634]	[0.6163]	[− 0.1513]	[− 0.1403]	[− 0.4836]	[0.9582]	

The standard errors (Error3) are reported in parenthesis, the convergence diagnostic (CD) statistics are reported in square brackets. The last column counts the number of skew levels that the corresponding skew probabilities are far different from 0.5

the skew probability is obviously different from 0.5, we label the corresponding skew level in Fig. 3. If the skew probability is larger (smaller) than 0.5, we treat its corresponding skew level, marked in red (green) in Fig. 3, as a support (resistance) line.

Table 7 also gives the number of skew levels that have corresponding skew probabilities far different from 0.5. During the sample periods, the skew phenomenon can be commonly observed in the Hong Kong overnight interest rate market; only in 2013 there is no obvious skew level and no skew phenomena. In 2006, 2011 and 2012, there are two skew levels, and in the remaining eight years, one skew level is found; in other words, one of the skew probabilities p_1 and p_2 is approximately 0.5. In most cases, the skew level of the doubly skewed OU process is close to that of the skew OU process, except for 2007, in which the skew levels are quite different from each other. Additionally, in 2012, the skew level in the skew OU process is 0.0998, which can be considered to be similar to 0.0979, one of the two skew levels in the doubly skewed OU process. Furthermore, according to the corresponding skew probabilities, the two skew levels can be treated as a support line and a resistance line, respectively. Hence, the doubly skewed OU process can describe the path properties of the overnight HIBOR more accurately.

The HIBOR is greatly affected by the U.S. financial market. For 2006, the Federal Reserve was still during a period of rate increases. As a result, the HIBOR exerted more upward pressure. The skew process shows that there is a support level near 3.7. A similar case also occured in 2015 and 2016. During these two years, our results also show the existence of support levels. After 2007, following the outbreak of the U.S. subprime mortgage crisis, the Federal Reserve announced quantitative easing and entered a cycle of rate cuts. Therefore, from 2007 to 2009, resistance levels appeared continuously in our results. Then, from 2010 to 2014, although the four rounds of quantitative easing still worked, confidence in the global financial

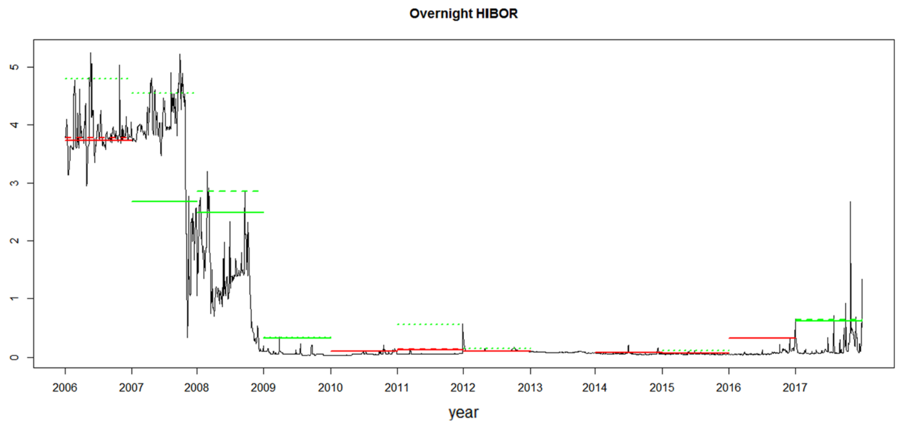

Fig. 3 Trace plots of the overnight HIBOR and the annual skew levels. Solid - the skew level of skew OU process, dashed—the lower skew level of doubly skewed OU process, dotted—the higher skew level of doubly skewed OU process. Red—the skew probability is larger than 0.5, green - the skew probability is smaller than 0.5

system began to be restored. Affected by the slowdown in the Chinese economy at the same time, Hong Kong local shares presented better profit ability than China shares listed in Hong Kong. This leads to the existence of support levels of interest rates during that time. In 2017, after the RMB exchange rate reform, the Bank of China prevailed against the attempt to short RMB. The real depreciation of USD against RMB leads to the resistance level in Fig. 3.

Next, we divided the total duration into four parts: 2006, 2007–2009, 2010–2016, and 2017. The results for the long duration, shown in Fig. 4 and Table 8, are aligned with the results of the annual estimation in Fig. 3 and Table 7. During the U.S. subprime mortgage crisis, we obtain a resistance level. It is obvious that 2007 and 2008 were the major intervals representing the breakout of the U.S. subprime mortgage crisis. Therefore, the resistance level over 2007 to 2009 in Fig. 4 is quite near the level during 2007 or 2008 in Fig. 3. A similar case can also be observed for 2010 to 2016.

Table 9 uses the measures of mean square error (MSE) and mean absolute error (MAE) to show the comparison of in-sample estimation performance among different processes, while Table 10 evaluates the corresponding out-of-sample estimation performance. The out-of-sample estimation is the forecasting results for one month after each year. We can see that the two skew-extended OU processes estimated by our proposed Bayesian methods outperform the other two alternatives, and fit the overnight HIBOR best basically all the time. Between these two Bayesian estimated processes, the skew OU process works a bit better than the doubly skewed OU process, especially when measured by MSEs.

4.2.2 VIX Index

Someone may want to know whether the skew phenomenon exists in a developed market, hence, we consider the CBOE Market Volatility Index (VIX) index. The

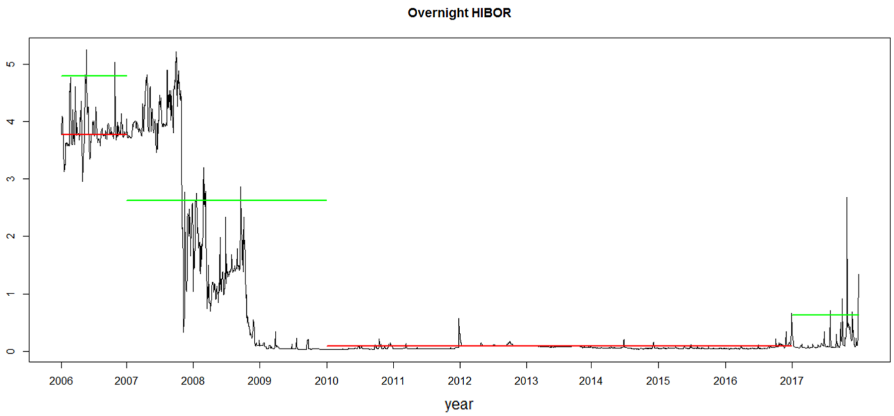

Fig. 4 Trace plots of the overnight HIBOR and the long duration skew levels. Solid - the skew level for long duration. Red - the skew probability is larger than 0.5, green - the skew probability is smaller than 0.5

Table 8 Estimation results of the skew local time components for the long duration of overnight HIBOR

Year	Skew OU-MLE		Skew OU-Bayesian		Doubly skewed OU-Bayesian				No. of skew levels
	a	p	a	p	a_1	p_1	a_2	p_2	
2006	3.7182	0.7469	3.7245	0.7224	3.7733	0.6740	4.7891	0.3949	2
	(0.0931)	(0.0651)	(0.0142)	(0.0324)	(0.0432)	(0.0476)	(0.3733)	(0.0883)	
			[0.9309]	[0.6795]	[0.7897]	[0.1377]	[0.4714]	[− 0.9148]	
2007–2009	2.7860	0.2133	2.6231	0.3527	1.1612	0.4617	2.8411	0.2609	1
	(0.3316)	(0.0163)	(0.0538)	(0.0317)	(0.9136)	(0.2032)	(0.2310)	(0.0605)	
			[− 1.3354]	[0.9754]	[− 0.7537]	[0.2237]	[0.0677]	[1.3449]	
2010–2016	0.0945	0.7602	0.1446	0.6400	0.0645	0.4816	0.0796	0.4929	1
	(0.0070)	(0.0813)	(0.0569)	(0.0400)	(0.0010)	(0.0221)	(0.0011)	(0.0215)	
			[− 1.5629]	[− 0.0916]	[0.3500]	[− 0.5522]	[0.6088]	[− 0.1897]	
2017	0.6978	0.3092	0.6296	0.3263	0.6421	0.3232	0.6431	0.4153	1
	(0.0958)	(0.0347)	(0.0043)	(0.0311)	(0.0047)	(0.1003)	(0.0037)	(0.0974)	
			[− 0.1634]	[0.6163]	[− 0.1513]	[− 0.1403]	[− 0.4836]	[0.9582]	

The standard errors (Error3) are reported in parenthesis, the convergence diagnostic (CD) statistics are reported in square brackets. The last column counts the number of skew levels that the corresponding skew probabilities are far different from 0.5

Table 9 In-sample estimation performance for the overnight HIBOR

Year	Error	OU MLE	Skew OU MLE	Skew OU Bayesian	Doubly skewed OU Bayesian	Best
2006	MSE	0.0656	0.0924	0.0397	0.0395	4
	MAE	0.1790	0.1934	0.1061	0.1087	3
2007	MSE	0.0948	0.1022	0.0781	0.0795	3
	MAE	0.1809	0.1947	0.1459	0.1438	4
2008	MSE	0.0799	0.0731	0.0589	0.0603	3
	MAE	0.2030	0.1897	0.1504	0.1417	4
2009	MSE	0.0009	0.0021	0.0008	0.0010	3
	MAE	0.0121	0.0276	0.0126	0.0085	4
2010	MSE	0.0002	0.0005	0.0002	0.0002	3
	MAE	0.0067	0.0144	0.0072	0.0071	1
2011	MSE	0.0051	0.0030	0.0012	0.0012	3
	MAE	0.0455	0.0202	0.0072	0.0072	4
2012	MSE	0.0001	0.0001	0.0001	0.0001	3
	MAE	0.0046	0.0032	0.0041	0.0026	4
2013	MSE	0.0000	0.0000	0.0000	0.0000	3
	MAE	0.0025	0.0023	0.0018	0.0018	4
2014	MSE	0.0002	0.0004	0.0002	0.0002	3
	MAE	0.0055	0.0098	0.0051	0.0057	3
2015	MSE	0.0001	0.0001	0.0000	0.0001	3
	MAE	0.0041	0.0083	0.0039	0.0042	3
2016	MSE	0.0013	0.0016	0.0013	0.0013	4
	MAE	0.0151	0.0254	0.0112	0.0116	3
2017	MSE	0.0764	0.0475	0.0398	0.0403	3
	MAE	0.1388	0.0725	0.0805	0.0804	2

This table reports the mean square error (MSE) and mean absolute error (MAE) of different processes using in-sample data of overnight HIBOR. The last column chooses the process that has the smallest value of MSE or MAE

VIX data sets consist of 7,053 daily observations from January 2, 1990, to December 29, 2017.

Similarly, the estimation results for the VIX index are shown in Table 11 and Fig. 5. We can observe that, in a developed market, the skew phenomenon did not appear as frequently as it did in the HIBOR market. However, an interesting finding is that the skew phenomenon generally occurs in a turbulent economy. The skew probabilities are different from 0.5 regularly during the 1990s, when the Japanese financial crisis broke out and the Asian financial crisis occurred. The same situation is observed after the global financial crisis in 2007. In the remaining years, the world's economic environment is relatively peaceful, and the results illustrate that there is no significant skew phenomenon during these periods.

Table 10 Out-of-sample estimation performance for overnight HIBOR

Year	Error	OU MLE	Skew OU MLE	Skew OU Bayesian	Doubly skewed OU Bayesian	Best
2006	MSE	159.2486	157.5833	25.1184	27.3300	3
	MAE	1131.3054	1082.5664	357.6188	335.3188	4
2007	MSE	1757.8938	1944.0835	907.7673	912.9007	3
	MAE	3454.4973	3581.0704	2002.2573	2018.9366	3
2008	MSE	596.5901	618.1698	1.7510	60.1514	3
	MAE	2398.0267	2441.3332	97.2832	753.2628	3
2009	MSE	0.4580	10.1954	0.0008	0.9346	3
	MAE	66.5357	313.9351	2.8432	95.0490	3
2010	MSE	0.0084	0.0002	0.0040	0.0070	2
	MAE	8.7953	1.0859	6.0260	8.0780	2
2011	MSE	51.2341	73.5194	3.6371	3.6368	4
	MAE	676.5471	789.8392	138.0472	138.0299	4
2012	MSE	0.1548	0.0253	0.0134	0.0589	3
	MAE	36.4956	12.8791	8.3877	21.2987	3
2013	MSE	0.4764	0.3375	0.1837	0.1905	3
	MAE	61.1413	50.5235	33.0744	33.2463	3
2014	MSE	0.2762	0.6147	0.2814	0.2428	4
	MAE	39.2083	67.1287	40.8699	36.7110	4
2015	MSE	0.2418	1.5465	0.2468	0.2512	1
	MAE	39.5921	105.3276	40.4136	38.7380	4
2016	MSE	5.6388	8.8609	6.3211	3.3080	4
	MAE	153.8378	249.8586	132.2659	91.3216	4
2017	MSE	56.8035	52.5670	17.1024	129.0681	4
	MAE	908.9276	428.4096	360.8912	1578.9554	3

This table reports the mean square error (MSE) and mean absolute error (MAE) of different processes using out-of-sample data of overnight HIBOR. The last column chooses the process that has the smallest value of MSE or MAE. Since the MSEs and MAEs are quit small for the out-of-sample data, all the results in this table are multiplied by 10^4

The VIX is generally used to measure the market volatility. When a financial panic happens, the VIX will be faced with upward pressure, as shown in the result for 1996 to 1997 during the Asian financial crisis and 2007 to 2008 during the global financial crisis. In addition, after the peak of the panic, the VIX should then fall back. Hence, after the Japanese financial crisis, resistance levels existed during 1991 to 1994. Analogous phenomena also appear in 1998 and 2009. Therefore, we divided the whole duration into eight parts: 1990, 1991-1995, 1996-1997, 1998, 1999–2006, 2007–2008, 2009 and 2010–2017. The results for the long duration are shown in Table 12 and Fig. 6. Similar to the results for the annual estimation, during the financial crisis, we obtain support levels, while after the financial crisis,

Table 11 Estimation results of the skew local time components for the VIX index

Year	Skew OU-MLE		Skew OU-Bayesian		Doubly skewed OU-Bayesian				No. of skew levels
	a	p	a	p	a_1	p_1	a_2	p_2	
1990	19.0380	0.2827	20.564	0.5972	20.5259	0.5668	27.5107	0.5542	0
	(1.7070)	(0.0040)	(3.8547)	(0.0688)	(0.7923)	(0.0542)	(4.6761)	(0.0762)	
			[− 0.9387]	[0.2699]	[0.6503]	[− 0.2680]	[0.0011]	[− 0.3472]	
1991	15.8622	0.4551	15.8153	0.6545	16.0995	0.6414	30.3876	0.3519	2
	(4.0941)	(0.0702)	(0.4986)	(0.0507)	(0.8024)	(0.0507)	(3.4847)	(0.0889)	
			[0.5589]	[− 1.2941]	[− 0.3083]	[1.2022]	[− 0.4193]	[1.2625]	
1992	13.6002	0.2694	16.1845	0.489	14.3862	0.4981	19.2433	0.4332	0
	(0.3884)	(0.0148)	(2.8898)	(0.0846)	(0.9114)	(0.0709)	(1.4546)	(0.0958)	
			[0.7516]	[0.9548]	[0.2002]	[0.7100]	[− 1.5151]	[0.7577]	
1993	11.6118	0.4020	14.3418	0.4888	11.985	0.4453	17.0067	0.2723	1
	(1.7916)	(0.0151)	(1.9817)	(0.1353)	(2.3667)	(0.1689)	(0.2426)	(0.0770)	
			[− 1.6085]	[0.6778]	[− 0.1089]	[0.7559]	[0.9118]	[0.7055]	
1994	12.0461	0.3490	22.917	0.3246	14.1261	0.4629	23.5003	0.2498	1
	(1.3556)	(0.0089)	(0.7516)	(0.0443)	(4.5022)	(0.2068)	(0.1708)	(0.0699)	
			[0.0423]	[0.9585]	[− 0.6743]	[− 0.0322]	[− 1.2800]	[0.5096]	
1995	11.5431	0.4684	12.7002	0.5439	11.3546	0.5543	14.4383	0.4293	0
	(0.0479)	(0.0944)	(1.3778)	(0.0673)	(0.6086)	(0.0988)	(1.3332)	(0.1251)	
			[0.0488]	[− 1.2483]	[0.0459]	[0.8589]	[1.4472]	[0.5004]	
1996	15.0608	0.3357	17.3689	0.6273	17.3157	0.6329	20.2987	0.3016	2
	(1.5692)	(0.0437)	(0.0997)	(0.0274)	(0.0815)	(0.0183)	(0.2304)	(0.0624)	
			[− 0.7492]	[− 0.5480]	[− 0.4209]	[1.2406]	[− 1.6793]	[1.1170]	

Table 11 continued

Year	Skew OU-MLE		Skew OU-Bayesian		Doubly skewed OU-Bayesian				No. of skew levels
	a	p	a	p	a_1	p_1	a_2	p_2	
1997	22.9771	0.5214	24.1457	0.617	21.4902	0.5075	25.6215	0.6116	1
	(4.6083)	(0.0782)	(2.7308)	(0.0817)	(2.4807)	(0.0616)	(3.2957)	(0.0844)	
			[0.8915]	[− 0.5458]	[0.1976]	[− 0.9104]	[1.4167]	[− 1.1657]	
1998	21.4141	0.2163	37.0939	0.4691	35.847	0.5833	42.8642	0.3176	1
	(2.0669)	(0.0157)	(7.2413)	(0.1104)	(6.0051)	(0.0986)	(0.9545)	(0.0652)	
			[− 0.9204]	[0.4439]	[− 0.9581]	[0.0961]	[− 0.7739]	[0.8296]	
1999	22.6757	0.2827	24.4771	0.4774	22.6844	0.464	28.4525	0.4244	0
	(0.6471)	(0.0511)	(4.8848)	(0.0774)	(4.0304)	(0.0904)	(3.8301)	(0.0880)	
			[0.2557]	[0.5092]	[0.7953]	[− 0.5553]	[− 0.6116]	[0.5135]	
2000	20.5508	0.1765	23.6821	0.5004	18.9308	0.5111	26.9053	0.4886	0
	(3.7557)	(0.0277)	(5.5345)	(0.1108)	(2.1359)	(0.1553)	(4.7874)	(0.0667)	
			[0.5642]	[− 0.7335]	[− 1.1676]	[0.0352]	[0.6027]	[− 0.6099]	
2001	21.9104	0.2827	31.8409	0.5735	25.2559	0.4907	31.9419	0.5848	0
	(0.5866)	(0.0531)	(4.7842)	(0.1141)	(4.1689)	(0.1114)	(0.4627)	(0.0607)	
			[− 1.2634]	[0.0998]	[− 0.7936]	[− 0.6430]	[− 1.0816]	[− 0.6036]	
2002	20.9492	0.2163	30.9679	0.5002	25.0856	0.5156	35.6849	0.4762	0
	(2.7269)	(0.0028)	(6.9645)	(0.0837)	(5.4377)	(0.1109)	(6.1854)	(0.7088)	
			[− 0.6491]	[− 1.1202]	[− 0.8540]	[0.7189]	[0.1127]	[− 1.3306]	
2003	16.8348	0.3967	25.1034	0.4847	22.4894	0.4377	28.9506	0.4484	0
	(2.0779)	(0.0396)	(3.9924)	(0.1659)	(3.9648)	(0.1628)	(4.6628)	(0.0803)	
			[− 0.0681]	[0.5752]	[− 0.8295]	[0.8887]	[− 0.6216]	[− 0.9584]	

Table 11 continued

Year	Skew OU-MLE		Skew OU-Bayesian		Doubly skewed OU-Bayesian				No. of skew levels
	a	p	a	p	a_1	p_1	a_2	p_2	
2004	13.9680	0.2959	16.1614	0.5341	14.0878	0.4979	19.004	0.5546	0
	(0.9758)	(0.0700)	(3.1512)	(0.1187)	(2.4340)	(0.1520)	(1.2281)	(0.0845)	
			[0.2014]	[1.0148]	[0.9684]	[− 1.1256]	[1.1079]	[0.4887]	
2005	13.7290	0.7867	14.6866	0.5848	13.1263	0.5462	15.3633	0.5442	0
	(1.6028)	(0.0587)	(0.7557)	(0.0397)	(1.5597)	(0.0775)	(1.0365)	(0.0942)	
			[− 0.1579]	[0.0025]	[− 0.2565]	[− 0.9520]	[− 0.5444]	[0.8187]	
2006	11.3169	0.2429	20.5031	0.2721	14.6558	0.6247	20.6872	0.1898	2
	(0.7993)	(0.0047)	(0.1282)	(0.0502)	(0.8951)	(0.0644)	(0.4227)	(0.0669)	
			[1.3604]	[− 1.2797]	[0.3583]	[− 0.7328]	[− 0.1785]	[0.9718]	
2007	13.4084	0.2163	22.6232	0.5683	21.1945	0.6142	30.6053	0.2341	2
	(0.6437)	(0.0133)	(3.7170)	(0.1177)	(1.2598)	(0.0391)	(0.1941)	(0.0698)	
			[− 0.1696]	[0.5947]	[− 0.0178]	[− 1.1405]	[0.1170]	[− 0.5650]	
2008	22.5939	0.1633	55.1299	0.6445	55.1983	0.6846	67.1213	0.3221	2
	(2.1234)	(0.0117)	(0.8083)	(0.0379)	(0.3695)	(0.0367)	(1.6393)	(0.0370)	
			[− 0.6086]	[0.0381]	[0.7295]	[− 0.7046]	[− 0.7612]	[0.2073]	
2009	40.5306	0.7867	46.6967	0.5514	43.2243	0.5731	55.889	0.2583	1
	(2.2094)	(0.0852)	(7.1024)	(0.1204)	(7.1599)	(0.0963)	(1.0228)	(0.0762)	
			[− 0.6718]	[− 0.3557]	[1.0391]	[0.2664]	[− 0.2402]	[0.3049]	
2010	18.2369	0.3357	29.1261	0.5887	25.1874	0.6035	36.797	0.4641	0
	(2.5233)	(0.0545)	(5.0897)	(0.0727)	(4.6930)	(0.5017)	(6.7733)	(0.0931)	
			[1.1338]	[− 0.4396]	[1.0009]	[1.5083]	[0.7886]	[1.3082]	

Table 11 continued

Year	Skew OU-MLE		Skew OU-Bayesian		Doubly skewed OU-Bayesian				No. of skew levels
	a	p	a	p	a_1	p_1	a_2	p_2	
2011	17.6884	0.2031	32.0465	0.6717	31.9923	0.6813	47.2329	0.2688	2
	(0.2168)	(0.0345)	(0.3016)	(0.0351)	(0.5695)	(0.0227)	(0.4949)	(0.0417)	
			[− 0.8051]	[− 0.6924]	[− 0.2129]	[0.9523]	[0.3260]	[− 1.1852]	
2012	16.2502	0.2163	20.7326	0.535	18.0787	0.4806	22.4965	0.5466	0
	(2.2747)	(0.0001)	(3.1001)	(0.1117)	(3.0775)	(0.0537)	(2.2567)	(0.1442)	
			[− 0.8516]	[0.0699]	[− 0.0886]	[1.4485]	[− 0.0399]	[− 0.4862]	
2013	12.8755	0.4418	13.4523	0.6193	12.9873	0.6273	16.8045	0.5182	1
	(1.6397)	(0.0800)	(1.4053)	(0.0626)	(1.3158)	(0.0724)	(2.5692)	(0.0849)	
			[0.0193]	[− 0.2027]	[− 1.3104]	[0.9458]	[0.1783]	[1.0139]	
2014	13.2663	0.2296	15.172	0.5939	14.7423	0.5656	19.4413	0.5023	0
	(0.7304)	(0.0223)	(1.7537)	(0.0503)	(1.1964)	(0.0636)	(3.3710)	(0.1023)	
			[0.1345]	[− 0.5507]	[− 0.4525]	[− 0.3255]	[0.5291]	[− 0.6229]	
2015	13.5108	0.1633	27.0278	0.5444	24.0425	0.5227	31.132	0.531	0
	(1.8198)	(0.0325)	(5.8252)	(0.1430)	(3.5623)	(0.1435)	(4.4442)	(0.1440)	
			[− 0.1667]	[− 0.2991]	[1.1288]	[− 0.3611]	[− 1.2604]	[0.5986]	
2016	13.0700	0.3888	18.7528	0.5459	16.7828	0.5221	21.9612	0.4939	0
	(1.7131)	(0.0581)	(3.1011)	(0.0701)	(3.6201)	(0.1124)	(3.9811)	(0.1079)	
			[− 0.2492]	[1.3166]	[− 0.3999]	[1.0219]	[1.5360]	[− 0.9094]	
2017	10.0071	0.3357	11.6304	0.6195	11.6604	0.6267	14.5007	0.3908	1
	(2.0787)	(0.0979)	(0.2555)	(0.0238)	(0.4919)	(0.0466)	(0.7011)	(0.0631)	
			[− 1.0324]	[1.3141]	[0.4522]	[0.1793]	[0.0407]	[0.1389]	

The standard errors (Error3) are reported in parenthesis, the convergence diagnostic (CD) statistics are reported in square brackets. The last column counts the number of skew levels that the corresponding skew probabilities are far different from 0.5

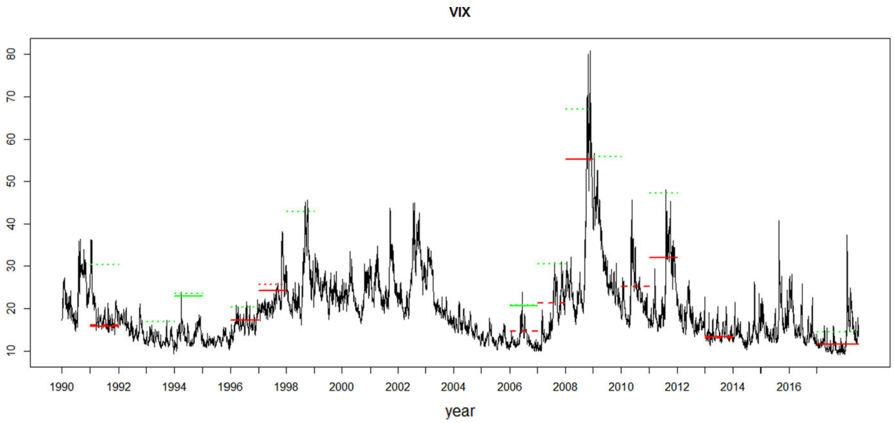

Fig. 5 Trace plots of the VIX index and the annual skew levels. Solid - the skew level of the skew OU process, dashed—the first skew level of the doubly skewed OU process, dotted—the second skew level of the doubly skewed OU process. Red—the skew probability is larger than 0.5, green - the skew probability is smaller than 0.5

resistance levels appear. From 2010 to 2017, stimulated by quantitative easing and other economic policies, a support level exists a long period of time.

Table 13 shows the in-sample estimation performance for each process, and Table 14 shows a comparison of the forecasting results for one month after each year. For the in-sample case, the skew OU process estimated by the Bayesian method works best in almost all years. For the out-of-sample case, the skew OU process estimated by the Bayesian method also outperforms the other processes for more than half of all durations. Notably, the doubly skewed OU process cannot be ignored for out-of-sample data as in Table 14. In particular, in the years when the Asian financial crisis and the global financial crisis broke out, the doubly skewed OU process works best. It is able to capture more characteristics of the VIX than other processes when a panic occurs.

5 Conclusion

In this paper, we have derived a Markov chain Monte Carlo method with the Gibbs sampler algorithm for Bayesian inference of the skew OU process and the doubly skewed OU process. Through the simulations, we have shown that our proposed method can work for the traditional OU process and performs better than the MLE method for the OU process. Applications with real data show that our method is also feasible in practice and can explain real economic phenomena. For instance, the classical process is not capable of capturing the dynamics of a market that exhibits the skew phenomenon. As a result, these two skew processes are useful, and our proposed estimation method can effectively capture these dynamics. It would next be interesting to see how the estimation for the local time component is influenced by the diffusion component, which is the object of future research.

Table 12 Estimation results of the skew local time components for the long duration of VIX index

Year	Skew OU-MLE		Skew OU-Bayesian		Doubly skewed OU-Bayesian				No. of skew levels
	a	p	a	p	a_1	p_1	a_2	p_2	
1990	19.0380	0.2827	20.5640	0.5972	20.5259	0.5668	27.5107	0.5542	0
	(1.7070)	(0.0040)	(3.8547)	(0.0688)	− 0.7923	− 0.0542	− 4.6761	− 0.0762	
			[− 0.9387]	[0.2699]	[0.6503]	[− 0.2680]	[0.0011]	[− 0.3472]	
1991–1995	16.1247	0.3224	18.8465	0.4432	14.0067	0.5518	15.3106	0.5658	1
	(1.3060)	(0.0310)	(1.6980)	(0.0414)	(1.3759)	(0.0618)	(1.5153)	(0.0576)	
			[− 0.2774]	[− 1.1833]	[− 0.6666]	[− 1.3506]	[− 0.9369]	[− 0.6170]	
1996–1997	16.0541	0.5694	23.1729	0.5744	17.0527	0.5448	18.4454	0.5344	1
	1.4774	0.0265	2.5938	0.0538	1.8514	0.0599	1.9014	0.0578	
			[− 0.3811]	[0.1358]	[− 1.8567]	[− 0.3394]	[− 0.6713]	[− 0.1623]	
1998	21.4141	0.2163	37.0939	0.4691	35.8470	0.5833	42.8642	0.3176	1
	(2.0669)	(0.0157)	(7.2413)	(0.1104)	(6.0051)	(0.0986)	(0.9545)	− 0.0652	
			[− 0.9204]	[0.4439]	[− 0.9581]	[0.0961]	[− 0.7739]	[0.8296]	
1999–2006	13.9335	0.4765	12.5713	0.5427	15.9038	0.5416	19.9891	0.4891	0
	(1.4081)	(0.0160)	(2.9378)	(0.0533)	(1.6135)	(0.0482)	(2.1260)	(0.0458)	
			[0.7640]	[− 1.7312]	[0.1942]	[0.2674]	[0.9054]	[0.1080]	
2007–2008	16.5514	0.5633	55.5028	0.6422	18.5333	0.5566	22.2130	0.5691	2
	(1.5662)	(0.0172)	(5.5052)	(0.0656)	(2.0049)	(0.0500)	(2.1715)	(0.0471)	
			[− 1.1303]	[− 0.8888]	[− 1.2717]	[− 0.4657]	[0.6871]	[0.5918]	
2009	40.5306	0.7867	46.6967	0.5514	43.2243	0.5731	55.8890	0.2583	1
	(2.2094)	(0.0852)	(7.1024)	(0.1204)	(7.1599)	(0.0963)	(1.0228)	− 0.0762	
			[− 0.6718]	[− 0.3557]	[1.0391]	[0.2664]	[− 0.2402]	[0.3049]	

Table 12 continued

Year	Skew OU-MLE		Skew OU-Bayesian		Doubly skewed OU-Bayesian				No. of skew levels
	a	p	a	p	a_1	p_1	a_2	p_2	
2010–2017	13.3000	0.6959	15.7456	0.6093	15.3116	0.5379	19.3797	0.5526	1
	(1.1786)	(0.0317)	(3.4662)	(0.0567)	(1.6370)	(0.0515)	(1.8163)	(0.0588)	
			[0.1714]	[1.1636]	[1.5299]	[− 0.7064]	[− 0.3434]	[0.8839]	

The standard errors (Error3) are reported in parenthesis, the convergence diagnostic (CD) statistics are reported in square brackets. The last column counts the number of skew levels that the corresponding skew probabilities are far different from 0.5

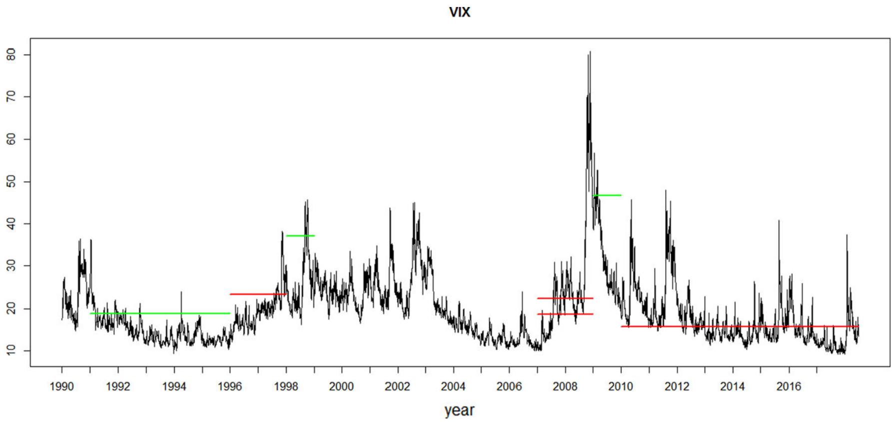

Fig. 6 Trace plots of the VIX index and the long duration skew levels. Solid - the skew level for the long duration. Red—the skew probability is larger than 0.5, green—the skew probability is smaller than 0.5

Table 13 In-sample estimation performance for the VIX index

Year	Error	OU MLE	Skew OU MLE	Skew OU Bayesian	Doubly skewed OU Bayesian	Best
1990	MSE	3.9691	2.7581	2.7558	5.0119	3
	MAE	1.5355	1.1645	1.1562	1.8104	3
1991	MSE	2.1056	1.4287	1.4285	1.6163	3
	MAE	0.9444	0.8038	0.8011	0.8907	3
1992	MSE	0.7174	0.4880	0.4875	0.7046	3
	MAE	0.6646	0.5266	0.5230	0.6788	3
1993	MSE	1.0090	0.4754	0.4753	1.1534	3
	MAE	0.7785	0.4812	0.4787	0.9085	3
1994	MSE	1.1327	0.8420	0.8425	0.8097	4
	MAE	0.8364	0.6549	0.6503	0.6662	3
1995	MSE	0.8003	0.3133	0.3132	0.5119	3
	MAE	0.7077	0.4163	0.4160	0.5601	3
1996	MSE	2.0720	0.8934	0.8929	1.0576	3
	MAE	1.0798	0.6904	0.6867	0.8048	3
1997	MSE	2.0760	1.5142	1.5142	3.9386	3
	MAE	0.9758	0.8636	0.8636	1.4753	2
1998	MSE	3.8835	3.3651	3.3558	3.5034	3
	MAE	1.4726	1.3374	1.3225	1.3938	3
1999	MSE	3.3947	1.7971	1.7894	4.3487	3
	MAE	1.4549	1.0496	1.0449	1.7423	3
2000	MSE	2.8992	1.8218	1.8167	3.4204	3
	MAE	1.3368	1.0405	1.0259	1.4955	3

Table 13 continued

Year	Error	OU MLE	Skew OU MLE	Skew OU Bayesian	Doubly skewed OU Bayesian	Best
2001	MSE	3.3241	2.2737	2.2707	5.0595	3
	MAE	1.3120	1.1049	1.0970	1.7500	3
2002	MSE	3.6164	2.9165	2.9062	3.6639	3
	MAE	1.4730	1.2931	1.2847	1.5392	3
2003	MSE	0.8634	0.8167	0.7646	0.8261	3
	MAE	0.6846	0.6701	0.6543	0.6854	3
2004	MSE	1.0456	0.6118	0.6117	1.4289	3
	MAE	0.7781	0.5969	0.5944	0.9414	3
2005	MSE	1.0134	0.4748	0.4750	1.3596	2
	MAE	0.7674	0.4991	0.5025	0.9527	2
2006	MSE	1.0581	0.7660	0.7646	0.7794	3
	MAE	0.6847	0.5604	0.5531	0.5496	4
2007	MSE	3.3166	2.4047	2.3998	2.4467	3
	MAE	1.4261	1.1056	1.0848	1.1315	3
2008	MSE	16.5940	12.5692	12.5315	13.7690	3
	MAE	2.6968	2.1812	2.1344	2.5840	3
2009	MSE	5.1622	4.0612	4.0558	4.0558	3
	MAE	1.5645	1.3652	1.3694	1.3580	4
2010	MSE	6.1281	3.8857	3.8870	4.7876	2
	MAE	1.7420	1.2272	1.2197	1.6284	3
2011	MSE	10.0958	6.1504	6.1414	6.8185	3
	MAE	2.2483	1.6214	1.5875	1.8121	3
2012	MSE	2.2260	1.2943	1.2901	2.7622	3
	MAE	1.1656	0.8532	0.8490	1.3717	3
2013	MSE	2.3431	0.9642	0.9639	1.5056	3
	MAE	1.1666	0.6609	0.6557	0.8543	3
2014	MSE	2.4978	1.3715	1.3690	1.4945	3
	MAE	1.0639	0.8228	0.8072	0.9205	3
2015	MSE	5.3413	3.0286	3.0184	4.1463	3
	MAE	1.5979	1.1631	1.1428	1.5679	3
2016	MSE	2.8788	1.8554	1.8545	2.5781	3
	MAE	1.1944	0.9041	0.9000	1.1488	3
2017	MSE	1.6465	0.6331	0.6328	0.9245	3
	MAE	0.9177	0.5005	0.4940	0.7803	3

This table reports the mean square error (MSE) and mean absolute error (MAE) of different processes using in-sample data of VIX index. The last column chooses the process that has the smallest value of MSE or MAE

Table 14 Out-of-sample estimation performance for the VIX index

Year	Error	OU MLE	Skew OU MLE	Skew OU Bayesian	Doubly skewed OU Bayesian	Best
1990	MSE	4.7372	2.9870	2.9805	5.1123	3
	MAE	1.5007	1.2863	1.2771	1.6335	3
1991	MSE	0.3515	0.3418	0.3404	0.3734	3
	MAE	0.4608	0.4465	0.4438	0.4681	3
1992	MSE	0.7388	0.2542	0.2345	1.1180	3
	MAE	0.7815	0.3988	0.3729	0.9858	3
1993	MSE	1.4264	0.8884	0.8883	2.1013	3
	MAE	1.0207	0.5812	0.5761	1.2973	3
1994	MSE	0.9685	0.3594	0.3339	0.4505	3
	MAE	0.8515	0.4659	0.4410	0.5406	3
1995	MSE	2.3058	0.6579	0.6594	1.0159	2
	MAE	1.2046	0.5909	0.5931	0.7485	2
1996	MSE	5.4005	0.7773	0.7810	0.6611	4
	MAE	2.1520	0.6871	0.6898	0.6664	4
1997	MSE	1.1593	1.0928	1.0932	1.5515	2
	MAE	0.8255	0.7838	0.7843	0.9899	2
1998	MSE	2.6299	2.6843	2.7293	2.6962	1
	MAE	1.3437	1.3502	1.3640	1.3410	4
1999	MSE	2.4273	2.0719	2.0546	5.1490	3
	MAE	1.2458	1.0877	1.0595	1.9800	3
2000	MSE	1.1761	0.9105	0.8974	1.5140	3
	MAE	0.7727	0.6889	0.6709	1.0188	3
2001	MSE	1.8601	1.7040	1.6822	3.5436	3
	MAE	1.1509	0.9945	0.9791	1.6620	3
2002	MSE	1.8994	1.6466	1.6833	1.6790	2
	MAE	1.0033	0.8559	0.8681	0.9443	2
2003	MSE	0.6348	0.5876	0.5753	0.7585	3
	MAE	0.6181	0.5943	0.5968	0.6816	2
2004	MSE	1.4233	0.2678	0.2421	3.3244	3
	MAE	1.0572	0.4313	0.4108	1.6918	3
2005	MSE	0.5322	0.3886	0.3927	1.0432	2
	MAE	0.5671	0.4325	0.4401	0.8384	2
2006	MSE	0.6209	0.3170	0.2896	0.2712	4
	MAE	0.6728	0.4636	0.4440	0.4232	4
2007	MSE	5.7237	2.4835	2.4874	2.3701	4
	MAE	1.8837	1.1809	1.1894	1.2169	2
2008	MSE	15.6659	12.5651	12.7047	12.1526	4
	MAE	2.8785	2.4451	2.4591	2.4268	4
2009	MSE	2.9841	2.5827	2.5949	2.5878	2
	MAE	1.4274	1.1075	1.1102	1.0880	4

Table 14 continued

Year	Error	OU MLE	Skew OU MLE	Skew OU Bayesian	Doubly skewed OU Bayesian	Best
2010	MSE	4.3868	1.1040	1.0548	4.0061	3
	MAE	1.9895	0.7984	0.7580	1.8946	3
2011	MSE	1.6252	0.9691	0.7988	1.4318	3
	MAE	1.1107	0.7607	0.6907	1.0232	3
2012	MSE	4.9416	0.7321	0.5708	8.4151	3
	MAE	2.1315	0.7491	0.6406	2.8109	3
2013	MSE	5.3831	1.8065	1.8045	3.3664	3
	MAE	1.7277	0.8988	0.8946	1.2212	3
2014	MSE	7.1700	2.2796	2.2917	2.3760	2
	MAE	2.1886	1.2887	1.2962	1.2473	4
2015	MSE	14.1617	4.0750	4.0853	4.0431	4
	MAE	3.2493	1.5867	1.5938	1.5821	4
2016	MSE	1.6741	0.3714	0.3383	1.6979	3
	MAE	1.1971	0.4751	0.4438	1.2048	3
2017	MSE	87.3199	19.5194	19.2175	22.2642	3
	MAE	5.4028	1.7833	1.7442	2.2370	3

This table reports the mean square error (MSE) and mean absolute error (MAE) of different processes using out-of-sample data of VIX index. The last column chooses the process that has the smallest value of MSE or MAE

Appendix

Details of Eqs.(6) and (7)

The transition density of Y_t can be constructed by the eigenvalues and eigenfunctions of the Strum–Liouville equation

$$(\mathcal{G}u)(x) = -\lambda u(x), \quad x \in (-\infty, +\infty),$$

with u satisfying proper conditions at the boundaries. When the spectrum is simple and purely discrete, the spectral expansion of the density reduces to the series Eq.(7). Based on the scale density $\zeta_Y(x)$ of the diffusion Y_t, $m_Y(x)$ is the speed density of the diffusion Y_t, defined as

$$\zeta_Y(x) = \begin{cases} \exp(\frac{\kappa(x-(1-p)\theta - pa)^2}{(1-p)^2\sigma^2}), & \textit{if } a \le x, \\ \exp(\frac{\kappa(x-p\theta-(1-p)a)^2}{p^2\sigma^2}), & \textit{if } x < a. \end{cases}$$
$$m_Y(x) = \begin{cases} \exp(\frac{2}{\zeta_Y(x)(1-p)^2\sigma^2}), & \textit{if } a \le x, \\ \exp(\frac{2}{\zeta_Y(x)p^2\sigma^2}), & \textit{if } x < a. \end{cases} \tag{34}$$

Let

$$z_1 \triangleq \frac{\sqrt{2\kappa}}{(1-p)\sigma}(x-(1-p)\theta - pa), \quad \alpha_1 \triangleq -\frac{\sqrt{2\kappa}}{(1-p)\sigma}((1-p)\theta + pa), \quad v \triangleq \frac{\lambda}{\kappa},$$
$$z_2 \triangleq \frac{\sqrt{2\kappa}}{p\sigma}(x - p\theta - (1-p)a), \quad \alpha_2 \triangleq -\frac{\sqrt{2\kappa}}{p\sigma}(p\theta + (1-p)a), \quad \varrho \triangleq -\frac{\kappa\theta^2}{\sigma^2},$$

then the eigenvalues $0 \le \lambda_1 < \lambda_2 < \ldots < \lambda_n \to \infty$ as $n \uparrow \infty$ in Eq.(7) are the simple discrete zeros of the Wronskian

$$\omega(\lambda) = e^{\varrho} 2^{1-v} v \sqrt(\kappa) \left[\frac{H_v(-\frac{\alpha_1}{\sqrt{2}}) H_{v-1}(\frac{\alpha_2}{\sqrt{2}})}{(1-p)\sigma} + \frac{H_{v-1}(-\frac{\alpha_1}{\sqrt{2}}) H_v(\frac{\alpha_2}{\sqrt{2}})}{p\sigma} \right], \tag{35}$$

and the normalized eigenfunction $\varphi_n(x)$ follows

$$\varphi_n(x) = \begin{cases} \mathrm{sign}(\xi(a,\lambda_n)\eta(a,\lambda_n)) \sqrt{\frac{\eta(a,\lambda_n)}{\omega'(\lambda_n)\xi(a,\lambda_n)}} \xi(x,\lambda_n), & \textit{if } a \le x, \\ \sqrt{\frac{\eta(a,\lambda_n)}{\omega'(\lambda_n)\xi(a,\lambda_n)}} \xi(x,\lambda_n), & \textit{if } x < a, \end{cases} \tag{36}$$

where

$$\xi(x,\lambda) = e^{\frac{z_1^2}{4}} D_v(-z_1), \quad \eta(x,\lambda) = e^{\frac{z_2^2}{4}} D_v(z_2),$$

the special functions $D_v(z)$ and $H_v(z)$ declared above are the parabolic cylinder function and the Hermite function respectively (see, e.g., Buchholz et al. 1970).

References

Appuhamillage, T., Bokil, V., Thomann, E., Waymire, E., & Wood, B. (2011). Occupation and local times for skew Brownian motion with applications to dispersion across an interface. *Annals of Applied Probability, 21*(1), 183–214.

Appuhamillage, T., & Iresh, T. (2011). Skew diffusion with drift: A new class of stochastic processes with applications to parabolic equations with piecewise smooth coefficients. Dissertations & Theses—Gradworks.

Appuhamillage, T., & Sheldon, D. (2010). First passage time of skew Brownian motion. *Journal of Applied Probability, 49*(49), 685–696.

Bardou, O., & Martinez, M. (2010). Statistical estimation for reflected skew processes. *Statistical Inference for Stochastic Processes, 13*(3), 231–248.

Barlow, M., Burdzy, K., Kaspi, H., & Mandelbaum, A. (2000). Variably skewed Brownian motion. *Institute of Mathematical Statistics, 5,* 57–66.

Broemeling, L. D., & Cook, P. (1992). Bayesian analysis of threshold autoregressions. *Communications in Statistics-Theory and Methods, 21*(9), 2459–2482.

Buchholz, H., Lichtblau, H., & Wetzel, K. (1970). *The confluent hypergeometric function, with special emphasis on its applications*. Berlin: Springer.

Cantrell, R. S., & Cosner, C. (1999). Diffusion models for population dynamics incorporating individual behavior at boundaries: Applications to refuge design. *Theoretical Population Biology, 55*(2), 189.

Chan, K. C., Karolyi, G. A., Longstaff, F. A., & Sanders, A. B. (1992). An empirical comparison of alternative models of the short-term interest rate. *The Journal of Finance, 47*(3), 1209–1227.

Chan, K. S., & Tsay, R. S. (1998). Limiting properties of the least squares estimator of a continuous threshold autoregressive model. *Biometrika,* 413–426.

Chen, C. W. S., & Lee, J. C. (2010). Bayesian inference of threshold autoregressive models. *Journal of Time, 16*(5), 483–492.

Chen, R., & Li, T. H. (1995). Blind restoration of linearly degraded discrete signals by gibbs sampling. *IEEE Transactions on Signal Processing, 43*(10), 2410–2413.

Collin-Dufresne, P., & Goldstein, R. S. (2001). Do credit spreads reflect stationary leverage ratios? *The Journal of Finance, 56*(5), 1929–1957.

Franke, J., Kreiss, J. P., & Mammen, E. (2002). Bootstrap of kernel smoothing in nonlinear time series. *Bernoulli, 8*(1), 1–37.

Gairat, A., & Shcherbakov, V. (2016). Density of skew Brownian motion and its functionals with application in finance. *Mathematical Finance*.

Gall, J. F. L. (1984). *One-dimensional stochastic differential equations involving the local times of the unknown process*. Springer, Berlin Heidelberg: Stochastic Analysis and Applications.

Gelfand, A. E., Hills, S. E., Racine-Poon, A., & Smith, A. F. (1990). Illustration of Bayesian inference in normal data models using Gibbs sampling. *Journal of the American Statistical Association, 85*(412), 972–985.

Gelfand, A. E., & Smith, A. F. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association, 85*(410), 398–409.

Gelman, A., & Rubin, D. (1991). An overview and approach to inference from iterative simulation. Technical Report, University of California-Berkeley, Dept. of Statistics.

Geweke, J. (1991). *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments*. Minneapolis, MN, USA: Federal Reserve Bank of Minneapolis, Research Department.

Geweke, J., & Terui, N. (1993). Bayesian threshold autoregressive models for nonlinear time series. *Journal of Time Series Analysis, 14*(5), 441–454.

Gonzalo, J., & Wolf, M. (2005). Subsampling inference in threshold autoregressive models. *Journal of Econometrics, 127*(2), 201–224.

Harrison, J. M., & Shepp, L. A. (1981). On skew Brownian motion. *Annals of Probability, 9*(2), 309–313.

Hull, J., & White, A. (1990). Pricing interest-rate-derivative securities. *The Review of Financial Studies, 3*(4), 573–592.

Itô, K., & Mckean, H. P. (1965). Diffusion processes and their sample paths.

Karatzas, I., & Shreve, S. E. (1984). Trivariate density of Brownian motion, its local and occupation times, with application to stochastic control. *Annals of Probability, 12*(3), 819–828.

Lang, R. (1995). Effective conductivity and skew Brownian motion. *Journal of Statistical Physics, 80*(1–2), 125–146.

Lejay, A. (2003). Simulating a diffusion on a graph. Application to reservoir engineering. *Mcma, 9*(3), 241–255.

Lejay, A. (2004). Monte carlo methods for fissured porous media: A gridless approach. *Mcma, 10*(3–4), 385–392.

Lejay, A. (2006). On the constructions of the skew Brownian motion. *Probability Surveys, 3,* 413–466.

Lejay, A. (2017). Estimation of the bias parameter of the skew random walk and application to the skew Brownian motion. Statistical Inference for Stochastic Processes (1), 1–13.

Lejay, A., & Pichot, G. (2012). Simulating diffusion processes in discontinuous media: A numerical scheme with constant time steps. *Journal of Computational Physics, 231*(21), 7299–7314.

Nakatsuma, T. (2000). Bayesian analysis of arma-garch models: A Markov chain sampling approach. *Journal of Econometrics, 95*(1), 57–69.

Ouknine, Y., & Rutkowski, M. (1995). Local times of functions of continuous semimartingales. *Stochastic Analysis & Applications, 12*(13), 211–231.

Pfann, G. A., Schotman, P. C., & Tschernig, R. (1996). Nonlinear interest rate dynamics and implications for the term structure. *Journal of Econometrics, 74*(1), 149–176.

Protter, P. (2004). *Stochastic integration and differential equations*. Berlin: Springer.

Ramirez, J. M. (2011). Multi-skewed Brownian motion and diffusion in layered media. *Proceedings of the American Mathematical Society, 139*(10), 3739–3752.

Revuz, D., & Yor, M. (2013). *Continuous martingales and Brownian motion* (Vol. 293). Berlin: Springer Science & Business Media.

Ritter, C., & Tanner, M. A. (1992). Facilitating the Gibbs sampler: The Gibbs stopper and the Griddy-Gibbs sampler. *Journal of the American Statistical Association, 87*(419), 861–868.

Schwartz, E. S. (1997). The stochastic behavior of commodity prices: Implications for valuation and hedging. *The Journal of Finance, 52*(3), 923–973.

Song, S., Xu, G., & Wang, Y. (2016). On first hitting times for skew CIR processes. *Methodology and Computing in Applied Probability, 18*(1), 169.

Su, F., & Chan, K. S. (2015). Quasi-likelihood estimation of a threshold diffusion process. *Journal of Econometrics, 189*(2), 473–484.

Tiao, G. C., & Tsay, R. S. (1994). Some advances in non-linear and adaptive modelling in time-series. *Journal of Forecasting, 13*(2), 109–131.

Tong, H. (1978). *On a threshold model in pattern recognition and signal processing*. Amsterdam: Sijthoff & Noordhoff.

Tong, H. (1990). *Non-linear time series: A dynamical system approach*. Oxford: Oxford University Press.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial and Quantitative Analysis, 5*(4), 177–188.

Walsh, J. B. (1978). A diffusion with a discontinuous local time. *Astérisque, 52*(53), 37–45.

Wang, S., Song, S., & Wang, Y. (2015). Skew Ornstein-Uhlenbeck processes and their financial applications. *Journal of Computational & Applied Mathematics, 273,* 363–382.

Xu, G., Song, S., & Wang, Y. (2016). The valuation of options on foreign exchange rate in a target zone. *International Journal of Theoretical and Applied Finance, 19*(03), 1650020.

Zhang, M. (2000). Calculation of diffusive shock acceleration of charged particles by skew Brownian motion. *Astrophysical Journal, 541*(1), 428–435.

Zhu, S. P., & He, X. J. (2017). A new closed-form formula for pricing European options under a skew Brownian motion. *The European Journal of Finance,* 1–13.

Zhuo, X., & Menoukeupamen, O. (2017). Efficient piecewise trees for the generalized skew Vasicek model with discontinuous drift. *International Journal of Theoretical & Applied Finance, 20,* 1750028.

Zhuo, X., Xu, G., & Zhang, H. (2017). A simple trinomial lattice approach for the skew-extended CIR models. *Mathematics and Financial Economics, 11*(4), 499–526.

Authors and Affiliations

Yizhou Bai[1] · Yongjin Wang[2] · Haoyan Zhang[3] · Xiaoyang Zhuo[4]

✉ Xiaoyang Zhuo
zhuoxy@bit.edu.cn

Yizhou Bai
yzbai@cauc.edu.cn

Yongjin Wang
yjwang@nankai.edu.cn

Haoyan Zhang
zhang-hy@cauc.edu.cn

[1] College of Science, Civil Aviation University of China, Tianjin 300300, China

[2] Business School, Nankai University, Tianjin 300071, China

[3] College of Science, Civil Aviation University of China, Tianjin 300300, China

[4] School of Management and Economics, Beijing Institute of Technology, Beijing 100081, China