SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

April 2018, Vol. 61 042202:1–042202:17
https://doi.org/10.1007/s11432-017-9215-8

Parameter estimates of Heston stochastic volatility model with MLE and consistent EKF algorithm

Ximei WANG[1,2], Xingkang HE[1,2], Ying BAO[3] & Yanlong ZHAO[1,2*]

[1]*Key Laboratory of System and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing* 100190*, China;*
[2]*School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing* 100049*, China;*
[3]*Risk Management Department, Industrial and Commercial Bank of China, Beijing* 100033*, China*

Received 21 June 2017/Accepted 8 August 2017/Published online 5 March 2018

Abstract Heston model is the most famous stochastic volatility model in finance. This paper considers the parameter estimation problem of Heston model with both known and unknown volatilities. First, parameters in equity process and volatility process of Heston model are estimated separately since there is no explicit solution for the likelihood function with all parameters. Second, the normal maximum likelihood estimation (NMLE) algorithm is proposed based on the Itô transformation of Heston model. The algorithm can reduce the estimate error compared with existing pseudo maximum likelihood estimation. Third, the NMLE algorithm and consistent extended Kalman filter (CEKF) algorithm are combined in the case of unknown volatilities. As an advantage, CEKF algorithm can apply an upper bound of the error covariance matrix to ensure the volatilities estimation errors to be well evaluated. Numerical simulations illustrate that the proposed NMLE algorithm works more efficiently than the existing pseudo MLE algorithm with known and unknown volatilities. Therefore, the upper bound of the error covariance is illustrated. Additionally, the proposed estimation method is applied to American stock market index S&P 500, and the result shows the utility and effectiveness of the NMLE-CEKF algorithm.

Keywords Heston model, stochastic volatility model, parameter estimation, normal maximum likelihood estimation, pseudo maximum likelihood estimation, consistent extended Kalman filter

Citation Wang X M, He X K, Bao Y, et al. Parameter estimates of Heston stochastic volatility model with MLE and consistent EKF algorithm. Sci China Inf Sci, 2018, 61(4): 042202, https://doi.org/10.1007/s11432-017-9215-8

1 Introduction

Stochastic volatility models play a very important role in finance since they can describe the movements of underlying assets or risk factors, which are the essential factors of derivative pricing and risk measurement.

The actual market data show that the most important two characters of the movements of underlying assets or risk factors are drift and volatility. Hence, Black Scholes (BS) model [1] was developed, which assumed that equity prices follow geometric Brownian motions with constant parameters. The key point making BS model popular is that there is a closed-form pricing formula for European option [2]. However, amounts of empirical studies show that the actual market data has the "stylized facts" [3]: Firstly, return rates of the equity prices have the feature of "volatility clustering" [4]; Secondly, distribution of the equity price return rates is fat-tailed and with high peaks [5]; Thirdly, the equity prices are negatively

* Corresponding author (email: ylzhao@amss.ac.cn)

correlated with the volatilities [6]; Finally, the volatility of equity return rates are mean reverting [7]. These additional characteristics cannot be described by BS model.

As a further development of BS model, Hull [8], Wiggins [9] generalized BS model to allow stochastic volatilities. The stochastic volatility models were constructed as two related stochastic diffusion processes: the equity process and the volatility one. These models can explain the movements of equity prices successfully, while they have disadvantages that there are no closed form solutions for the two-dimensional partial differential equations. Heston [10] assumed that the volatility process in the stochastic volatility models follows a square root mean reverting process, which ensures the non-negative of the volatilities. Additionally, Heston model is able to explain the "stylised facts" as well as drift and volatility [10, 11]. Heston model can also present "smile effect" of the implied volatility [12, 13], which can be considered as one of the well-know discrepancies of BS model. Most importantly, there exists a closed-form formula for pricing European options. Hence, Heston model is one of the most famous stochastic volatility models.

However, there still exist big challenges to estimate the parameters of Heston model. Since there is no closed joint transition probability function of the two-dimensional diffusion process, maximum likelihood estimation (MLE) method is difficult to implement. Furthermore, if all the parameters are considered in one likelihood function, there is no explicit solution for the likelihood function and numerical solutions are unstable.

With assumption of known volatilities, Ref. [14] proposed a pseudo-MLE (PMLE) method, which approximated the volatility model of Heston model as a Gaussian process and it did not provide estimation methods for correlation coefficient ρ between noises. Ref. [15] proposed an approximate form of the transition function, which is helpful to obtain the MLE of the parameters. Ref. [16] derived the moment estimations on three parameters except correlation coefficient of Heston model. In practical market, the volatilities (i.e., the states of the system model) are generally unobservable. Hence, it is impossible to apply a straightforward MLE based on unknown volatilities.

For unknown volatilities case, there are some literatures on the estimation problem. Ref. [3] proposed efficient method of moment (EMM) algorithm, which included two steps to complete the estimation process. The first step is to estimate an auxiliary model with MLE and the second one is to estimate the true model with generalized method of moment (GMM). Yet, this method depends on the selection of the auxiliary model. Ref. [17] put up Monto Carlo Markov chain (MCMC) algorithm, which depended on the prior distribution of the parameters. Since EMM and MCMC have high computational burden, [18] put forward the quasi-maximum likelihood estimation (QMLE) method, which is widely used due to its easy implementation. However, objective functions with respect to the parameters are sometimes non-convex, then the maximum points may not be unique.

Besides estimation methods, filtering methods can be applied when the volatilities are unknown. For the nonlinear Heston model, [19] utilized extended Kalman filter (EKF) to estimate the volatilities. Ref. [20] provided an algorithm based on particle filter (PF) and Ref. [21] utilized unscented Kalman filter to estimate the volatilities. In [22], the authors considered the volatility estimation problem for nonlinear and non-Gaussian systems through utilizing the auxiliary particle filter (APF). However, computational consumption of PF and APF becomes much larger with the increasing of the particle numbers, especially in the pursuing for better estimation performance. The unscented Kalman filter (UKF) algorithm in [21] had the similar properties as EKF. Actually, they are constructed based on the second and first order expansions of the nonlinear model, respectively. In [23], a recursive filtering method was proposed for a class of nonlinear time-varying systems subject to multiplicative noises, missing measurements and quantisation effects. The algorithm can provide an optimized upper bound for the filtering error covariance through properly designing gain parameters.

In this paper, parameters in Heston model are estimated under the assumptions of unknown as well as known volatilities. As for known volatilities, joint distribution functions of Heston model are considered in the literatures. Such an approach brings difficulties to estimate the parameters. While after orthogonal transformation, the parameters in the state and measurement equations can be estimated separately. Compared with the existing PMLE method, closed form of all the parameter estimations can be solved based on the distribution of noises. The correlation coefficient between noises can also be estimated with moment estimation methods. In the condition of unknown volatilities, Kalman filter (KF) is optimal in the sense of linear minimum mean square error (MMSE) for linear systems. One can evaluate the estimation precision of KF at real time since the parameter matrix P stands for the error covariance matrix. Nevertheless, for nonlinear systems, EKF cannot provide the precision evaluation because the parameter matrix P does not have a clear relationship with the estimation error covariance matrix. Ref. [24] designed an efficient algorithm for time-varying state estimator such that an upper bound of the estimation error covariance can be guaranteed and the explicit expression of the estimator parameters is given. This algorithm is implemented based on EKF algorithm in the presence of missing measurements. Refs. [25–27] proposed the definition of consistent extended Kalman filter (CEKF) and a CEKF algorithm which can bound the estimation errors at a real time. In this paper, CEKF method is utilized to estimate the volatilities in Heston model. The contributions of this paper can be summarized as follows:

- The normal maximum likelihood estimation (NMLE) is proposed to estimate the parameters of Heston model. Parameters in equity process and volatility process are estimated separately. Through the transformation of Itô lemma, the volatility follows an exact Gaussian process instead of an approximate one. The NMLE is shown to be more efficient than the PMLE by simulations under the assumption of known volatilities. The closed form of MLE for the parameters in the state equation can be obtained by maximizing the joint likelihood function. The correlation coefficient between noises can also be estimated with moment estimation methods.
- The NMLE and CEKF algorithm are combined to estimate the parameters and the unknown volatilities. The NMLE is utilized to update the parameters and the CEKF is employed to update the volatilities alternately. The algorithm can provide an upper bound of the estimation error covariance matrix to ensure the volatilities to be well evaluated. Both the NMLE and CEKF algorithm act better than the existing methods, PMLE and EKF algorithm, respectively. Hence NMLE-CEKF can estimate parameters and volatilities efficiently.
- The proposed estimation method is applied to American stock market index S&P 500. The result shows that the volatilities estimated by the NMLE-CEKF algorithm are between history values and Chicago board options exchange volatility index (CBOE VIX index), which indicates the utility and effectiveness of the NMLE-CEKF algorithm.

The motivation of this paper is to propose a straightforward and effective algorithm to estimate parameters in Heston model. The existing EMM, GMM, and MCMC are model-dependent or with heavy computation. Although QMLE method is easy to implement, it usually shows poor performance on the convergence of the estimations. Besides, EKF method can be utilized to estimate the unknown volatilities, but it cannot evaluate the estimation errors in real time. Thus we propose the NMLE-CEKF algorithm to estimate parameters as well as volatilities. Through the proposed algorithm, the amount of calculation can be effectively reduced since the estimations of parameters have closed forms. The estimation error covariance matrix can be bounded with CEKF algorithm.

The rest of this paper is organized as follows. Section 2 describes problem formulation. Section 3 presents the NMLE of the parameters. Section 4 analyzes the CEKF algorithm. Section 5 introduces the NMLE-CEKF estimation algorithm of Heston model. Section 6 carries out several simulation results including comparing different discretization methods, evaluating the PMLE and NMLE algorithm with known volatilities, and applying NMLE-CEKF algorithm with unknown volatilities. Section 7 shows the application of the algorithm to the S&P 500 data. The conclusion of this paper is presented in Section 8.

2 Problem formulation

In empirical studies, volatilities of the equity prices are mean reverting, i.e., the volatilities will be pulled back to their long time average value. Ref. [10] assumed that the volatility process of the equity follows a square root mean reverting process. The Heston model in a risk-neutral probability space $\mathbb{Q}$[1)] is presented

1) The risk neutral probability measure $\mathbb{Q}$ is the equivalent martingale measure of the real world probability $\mathbb{P}$ [28].

as follows:

$$\begin{cases} \mathrm{d}S(t) = rS(t)\mathrm{d}t + \sqrt{V(t)}S(t)\mathrm{d}W_1(t), & (1) \\ \mathrm{d}V(t) = \kappa(\theta - V(t))\mathrm{d}t + \sigma\sqrt{V(t)}\mathrm{d}W_2(t), & (2) \\ \mathrm{d}W_1(t)\mathrm{d}W_2(t) = \rho\mathrm{d}t, \quad t \in [0, T], & (3) \end{cases}$$

where $S(t)$, $V(t)$ are equity price and the variance of the price at time t, respectively. In the subsequents, equation $V(t)$ is called as the volatility instead of the variance intuitively. In this model, Eq. (1) is called equity process and Eq. (2) is called volatility process. r is the average yield of the equity. In risk neutral world, r is risk free rate. If dividends are considered, the risk free rate is adjusted by subtracting the dividends. $[0, T]$ is the time interval. In the volatility process (2), θ is the long run mean of the volatilities. Wherever the initial value is, the volatility will go around θ eventually. κ is the mean reversion rate towards θ, and it reveals the speed of the volatilities going back to their average value. $W_1(t)$ and $W_2(t)$ are Brownian motions with correlation coefficient ρ in the risk neutral world $\mathbb{Q}$. ρ also affects the heavy tails of the stock return rate distribution. σ is called the volatility of volatility, which influences the variation of the volatilities and the kurtosis of the stock return rate distribution. Additionally, if κ, θ satisfy the condition $2\kappa\theta > \sigma^2$, volatilities generated by (2) are positive [29]. More detail studies of the effectiveness of the parameters in Heston model can turn to [11, 13, 30]. What makes Heston model a charming one and highly utilized in financial institutions is that Heston model has a closed form solution for pricing the European option, which provids a relatively simple method to price financial derivatives [10].

Consider the Heston model as state-space representation of a dynamic system [31], which consists of a state equation and a measurement equation. Usually, stock price $S(t)$ can be observed from the market, but $V(t)$ is unavailable. Hence (1) and (2) can be regarded as measurement and state equation, respectively. Since in real market, equity prices return rates are more concerned, the logarithms of the prices are conducted by Itô equation as follows:

$$\begin{aligned} \mathrm{d}\ln(S(t)) &= \frac{1}{S(t)}\mathrm{d}S(t) - \frac{1}{2}\frac{1}{S^2(t)}\mathrm{d}S(t)\mathrm{d}S(t) \\ &= \left(r - \frac{1}{2}V(t)\right)\mathrm{d}t + \sqrt{V(t)}\mathrm{d}W_1(t). \end{aligned} \tag{4}$$

Continuous form of the Heston model (1)–(3) can be written as the following state and measurement equations:

$$\begin{cases} \mathrm{d}V(t) = \kappa(\theta - V(t))\mathrm{d}t + \sigma\sqrt{V(t)}\mathrm{d}W_v, & (5) \\ \mathrm{d}\ln(S(t)) = \left(r - \frac{1}{2}V(t)\right)\mathrm{d}t + \sqrt{(1-\rho^2)V(t)}\mathrm{d}W_s(t) + \rho\sqrt{V(t)}\mathrm{d}W_v(t), \quad t \in [0, T], & (6) \end{cases}$$

where $\mathrm{d}W_s(t)\mathrm{d}W_v(t) = 0$, and the following factorization holds:

$$\begin{cases} W_1(t) = \sqrt{1-\rho^2}W_s(t) + \rho W_v(t), \\ W_2(t) = W_v(t). \end{cases}$$

The parameters $\kappa, \theta, \sigma, \rho$ in the system are always unknown in practice. To employ the Heston model, one should estimate these parameters firstly. Generally, since the volatilities are unknown, filtering methods are utilized to estimate the volatilities in real time. Meanwhile, the implementation of the filtering algorithm needs known parameters. Therefore, the parameter estimation and volatility tracking constitute a coupled issue. In this paper, we will design an algorithm to estimate the parameters and track the volatilities of the Heston model synchronously.

3 Parameter estimation of Heston model

In this section, an NMLE for Heston model is proposed. Noticing that the volatility process of Heston model is the main different point compared with other stochastic volatility models in [8, 9, 32], we focus on the parameter estimation problems of the volatility model. In general, GMM, least square (LS), and MLE methods are carried out to estimate the parameters in volatility model. Ref. [33] proposed a numerical optimization method to minimize the objective function. Ref. [14] proposed a PMLE method to estimate the parameters. They assumed the parameters of the diffusion process as fixed coefficients, thus the recursive form of volatility can be approximated with a Gaussian process. However, it is known that conditioned on the previous value $V(t-\delta)$ (where δ is an arbitrarily small positive number), the distribution of volatility $V(t)$ is noncentral chi-square with the degrees of freedom $\mathrm{d}f = \frac{4\kappa\theta}{\sigma^2}$ and noncentrality parameter $\mathrm{nc} = \frac{4\kappa \mathrm{e}^{-\kappa\delta}}{\sigma^2(1-\mathrm{e}^{-\kappa\delta})}V(t-\delta)$ [29]. It will be more appropriate to find an exact Gaussian distribution than an approximate one. According to the Itô lemma, there is

$$\mathrm{d}\sqrt{V(t)} = \frac{1}{2\sqrt{V(t)}}\mathrm{d}V(t) - \frac{1}{8\sqrt{V^3(t)}}\mathrm{d}V(t)\mathrm{d}V(t). \tag{7}$$

Substituting (5) into (7), we can obtain

$$\begin{aligned}\mathrm{d}\sqrt{V(t)} &= \frac{1}{2\sqrt{V(t)}}[\kappa(\theta - V(t))\mathrm{d}t + \sigma\sqrt{V(t)}\mathrm{d}W_v(t)] - \frac{1}{8\sqrt{V^3(t)}}V(t)\sigma^2\mathrm{d}t \\ &= \frac{1}{2\sqrt{V(t)}}\left[\kappa\theta - \kappa V(t) - \frac{1}{4}\sigma^2\right]\mathrm{d}t + \frac{1}{2}\sigma\mathrm{d}W_v(t). \end{aligned}\tag{8}$$

Thus we know that conditioned on $\sqrt{V(t-\delta)}$, $\sqrt{V(t)}$ is subject to a Gaussian distribution. Since the measurements of states are obtained only at discrete time instants, $V(t)$ can be discretized as $\{V_0, V_\delta, \ldots, V_{n\delta}\}$, where $n = \frac{T}{\delta}$. Omitting δ in the subscript, the volatility sequence can be denoted as $\{V_k\}_{k=0}^n$. Then Eq. (8) can be written as follows:

$$\sqrt{V_k} - \sqrt{V_{k-1}} = \frac{1}{2\sqrt{V_{k-1}}}\left[\kappa\theta - \kappa V_{k-1} - \frac{1}{4}\sigma^2\right]\delta + \frac{1}{2}\sigma\Delta W_{v_k},$$

where $\Delta W_{v_k} = W_{v_k} - W_{v_{k-1}}$. Thus,

$$\sqrt{V_k} = \sqrt{V_{k-1}} + \frac{1}{2\sqrt{V_{k-1}}}\left[\kappa\theta - \kappa V_{k-1} - \frac{1}{4}\sigma^2\right]\delta + \frac{1}{2}\sigma\Delta W_{v_k}. \tag{9}$$

Theorem 1. The MLE of the parameters κ, σ, θ in (9) are denoted as

$$\begin{cases}\hat{\kappa} = \dfrac{2}{\delta}\left(1 + \dfrac{\hat{P}\delta}{2}\dfrac{1}{n}\displaystyle\sum_{k=1}^n \dfrac{1}{V_{k-1}} - \dfrac{1}{n}\displaystyle\sum_{k=1}^n\sqrt{\dfrac{V_k}{V_{k-1}}}\right), \\ \hat{\sigma} = \sqrt{\dfrac{4}{\delta}\dfrac{1}{n}\displaystyle\sum_{k=1}^n\left[\sqrt{V_k} - \sqrt{V_{k-1}} - \dfrac{\delta}{2\sqrt{V_{k-1}}}(\hat{P} - \hat{\kappa}V_{k-1})\right]^2}, \\ \hat{\theta} = \dfrac{\hat{P} + \frac{1}{4}\hat{\sigma}^2}{\hat{\kappa}}, \end{cases}\tag{10}$$

where

$$\hat{P} = \frac{\frac{1}{n}\sum_{k=1}^n\sqrt{V_{k-1}V_k} - \frac{1}{n^2}\sum_{k=1}^n\sqrt{\frac{V_k}{V_{k-1}}}\sum_{k=1}^n V_{k-1}}{\frac{\delta}{2} - \frac{\delta}{2}\frac{1}{n^2}\sum_{k=1}^n\frac{1}{V_{k-1}}\sum_{k=1}^n V_{k-1}}.$$

Proof. Conditioned on $\sqrt{V_{k-1}}$, $\sqrt{V_k}$ is subject to the following Gaussian distribution:

$$\sqrt{V_k} \sim N\left(m_{k-1}, \frac{\sigma^2}{4}\delta\right), \tag{11}$$

where

$$m_{k-1} = \sqrt{V_{k-1}} + \frac{1}{2\sqrt{V_{k-1}}}\left[\kappa\theta - \kappa V_{k-1} - \frac{1}{4}\sigma^2\right]\delta.$$

Denote the likelihood function as

$$\ell(\Theta)=\prod_{k=1}^{n}\frac{1}{\sqrt{2\pi}\sqrt{\frac{\sigma^2}{4}\delta}}\exp\left(-\frac{(\sqrt{V_k}-m_{k-1})^2}{\frac{\sigma^2}{2}\delta}\right), \tag{12}$$

where $\Theta=(\kappa,\kappa\theta,\sigma)$. Taking the logarithm of the above likelihood function (12), omitting the constant variables, the logarithm likelihood function can be written as

$$\ell(\Theta)=-\sum_{k=1}^{n}\frac{1}{2}\log\sigma^2-\frac{(\sqrt{V_k}-m_{k-1})^2}{\frac{\sigma^2}{2}\delta}. \tag{13}$$

Taking the derivative of the function $\ell(\Theta)$ with respect to the parameters, $\kappa\theta$, κ, σ^2, respectively, we can obtain

$$\begin{cases}\dfrac{\partial\ell(\Theta)}{\partial(\kappa\theta)}=-\delta\displaystyle\sum_{k=1}^{n}\left(\sqrt{V_k}-m_{k-1}\right)\frac{1}{\sqrt{V_{k-1}}},\\ \dfrac{\partial\ell(\Theta)}{\partial\kappa}=\delta\displaystyle\sum_{k=1}^{n}\left(\sqrt{V_k}-m_{k-1}\right)\sqrt{V_{k-1}},\\ \dfrac{\partial\ell(\Theta)}{\partial\sigma^2}=-\dfrac{n}{2\sigma^2}+\dfrac{1}{\delta}\displaystyle\sum_{k=1}^{n}\frac{2\left(\sqrt{V_k}-m_{k-1}\right)^2}{\sigma^4}.\end{cases} \tag{14}$$

To obtain the MLE of the parameter Θ, the derivatives of the parameters are set as follows:

$$\begin{cases}\dfrac{\partial\ell(\Theta)}{\partial(\kappa\theta)}=0,\\ \dfrac{\partial\ell(\Theta)}{\partial\kappa}=0,\\ \dfrac{\partial\ell(\Theta)}{\partial\sigma^2}=0.\end{cases} \tag{15}$$

Then the equations in (15) implies that

$$\sum_{k=1}^{n}\left(\sqrt{V_k}-m_{k-1}\right)\frac{1}{\sqrt{V_{k-1}}}=0, \tag{16}$$

$$\sum_{k=1}^{n}\left(\sqrt{V_k}-m_{k-1}\right)\sqrt{V_{k-1}}=0, \tag{17}$$

$$\frac{n}{2\sigma^2}-\frac{1}{\delta}\sum_{k=1}^{n}\frac{2\left(\sqrt{V_k}-m_{k-1}\right)^2}{\sigma^4}=0. \tag{18}$$

With an ingenious technique, setting $P=\kappa\theta-\frac{1}{4}\sigma^2$, sorting (16) and (17), we can get

$$\frac{n\kappa\delta}{2}=n+\frac{P\delta}{2}\sum_{k=1}^{n}\frac{1}{V_{k-1}}-\sum_{k=1}^{n}\sqrt{\frac{V_k}{V_{k-1}}},$$

$$\frac{nP\delta}{2}=\sum_{k=1}^{n}\sqrt{V_{k-1}V_k}-\sum_{k=1}^{n}V_{k-1}+\frac{\kappa\delta}{2}\sum_{k=1}^{n}V_{k-1}.$$

The above equations can be solved as

$$\hat{P}=\frac{\frac{1}{n}\sum_{k=1}^{n}\sqrt{V_{k-1}V_k}-\frac{1}{n^2}\sum_{k=1}^{n}\sqrt{\frac{V_k}{V_{k-1}}}\sum_{k=1}^{n}V_{k-1}}{\frac{\delta}{2}-\frac{\delta}{2}\frac{1}{n^2}\sum_{k=1}^{n}\frac{1}{V_{k-1}}\sum_{k=1}^{n}V_{k-1}},$$

$$\hat{\kappa} = \frac{2}{\delta}\left(1 + \frac{\hat{P}\delta}{2}\frac{1}{n}\sum\nolimits_{k=1}^{n}\frac{1}{V_{k-1}} - \frac{1}{n}\sum\nolimits_{k=1}^{n}\sqrt{\frac{V_k}{V_{k-1}}}\right), \tag{19}$$

then according to (18),

$$\hat{\sigma}^2 = \frac{4}{\delta}\frac{1}{n}\sum_{k=1}^{n}\left[\sqrt{V_k} - \sqrt{V_{k-1}} - \frac{\delta}{2\sqrt{V_{k-1}}}(\hat{P} - \hat{\kappa}V_{k-1})\right]^2,$$

thus

$$\hat{\sigma} = \sqrt{\frac{4}{\delta}\frac{1}{n}\sum_{k=1}^{n}\left[\sqrt{V_k} - \sqrt{V_{k-1}} - \frac{\delta}{2\sqrt{V_{k-1}}}(\hat{P} - \hat{\kappa}V_{k-1})\right]^2}. \tag{20}$$

With $\hat{P},\ \hat{\kappa},\ \hat{\sigma}^2$, we can get that

$$\hat{\theta} = \frac{\hat{P} + \frac{1}{4}\hat{\sigma}^2}{\hat{\kappa}}. \tag{21}$$

Then Eqs. (19)–(21) are the estimations of $\kappa,\ \sigma^2,\ \sigma,\ \theta$ by MLE, respectively. Since the MLE of the parameters are derived through the Gaussian distribution of $\sqrt{V_k}$ rather than the approximation of chi-square distribution, the estimation method is called NMLE.

Regarding the correlation coefficient of the Heston Model ρ, there is

$$E\mathrm{d}W_1(t)\mathrm{d}W_2(t) = \rho\mathrm{d}t,$$

where

$$\begin{cases} \mathrm{d}W_1(t) = \dfrac{\mathrm{dln}S(t) - (r - \frac{1}{2}V(t))\mathrm{d}t}{\sqrt{V(t)}}, \\ \mathrm{d}W_2(t) = \dfrac{\mathrm{d}V(t) - \kappa(\theta - V(t))\mathrm{d}t}{\sigma\sqrt{V(t)}}. \end{cases} \tag{22}$$

The discrete form of (22) is as follows:

$$\begin{cases} \Delta W_{1_k} = \dfrac{\mathrm{ln}S_k - \mathrm{ln}S_{k-1} - (r - \frac{1}{2}V_{k-1})\delta}{\sqrt{V_{k-1}}}, \\ \Delta W_{2_k} = \dfrac{V_k - V_{k-1} - \kappa(\theta - V_{k-1})\delta}{\sigma\sqrt{V_{k-1}}}, \end{cases}$$

where $\Delta W_{i_k} = W_{i_k} - W_{i_{k-1}}$, $i = 1, 2$. Then the moment estimation of ρ can be given as

$$\widehat{\rho} = \frac{1}{n\delta}\sum_{k=1}^{n}\Delta W_{1_k}\Delta W_{2_k}. \tag{23}$$

Intuitively, V_k follows chi-square the distribution, the square-root process of V_k, i.e., $\sqrt{V_k}$ is Gaussian. Basically, the MLE based on $\sqrt{V_k}$ is more accurately than V_k, since the MLE of V_k simply includes the discretization errors yet the MLE of $\sqrt{V_k}$ includes the discretization error as well as the approximate error.

4 Consistent extended Kalman filter

In this section, we will introduce an effective filtering algorithm, the CEKF to estimate the unknown volatilities in the Heston model. Then we can estimate the parameters with the NMLE method introduced in Section 3. The discrete form of the Heston model of (5) and (6) can be written as

$$\begin{cases} V_k = V_{k-1} + \kappa(\theta - V_{k-1})\delta + \sigma\sqrt{V_{k-1}}\Delta W_{v_k}, \\ z_k = \left(r - \dfrac{1}{2}V_k\right)\delta + \sqrt{(1-\rho^2)V_k}\Delta W_{s_k} + \rho\sqrt{V_k}\Delta W_{v_k}, \end{cases} \tag{24}$$

where V_k is the variance (volatility) of the equity price at step k, δ is the sampling step, $\Delta W_{s_k} = W_{s_k} - W_{s_{k-1}}$, $\Delta W_{v_k} = W_{v_k} - W_{v_{k-1}}$, and $z_k = \log(S_{k+1}) - \log(S_k)$. W_{v_k}, W_{s_k} are independent Brownian motion, $k = 1, 2, \ldots$.

Since Eq. (24) is a nonlinear system, one can utilize the EKF algorithm which linearizes a nonlinear system to a linear one [34]. To evaluate the estimation precision, the consistent EKF is introduced, which compensates the linearization error of EKF. Consider the general discrete system

$$\begin{cases} X_k = f(X_{k-1}, w_k), \\ Y_k = h(X_k, w_k), \end{cases} \tag{25}$$

where $f(\cdot,\cdot)$ and $h(\cdot,\cdot)$ are first order continuously differentiable, X_k is the n-dimensional state vector at time k, Y_k is the m-dimensional measurement output, and $\{w_k\}$ follows an *i.i.d* Gaussian process with zero mean and variance given by $Q_k > 0$, $k = 1, 2, \ldots$.

In linear systems with Gaussian noise, Kalman filter can directly give the filtering error covariance matrix. For nonlinear systems, however, in traditional EKF, the P_k matrix is no longer the filtering error covariance matrix. The CEKF in [25] can give the upper bound of the filtering error covariance matrix by designing proper parameter matrices $\varDelta Q_k$, $\varDelta R_k$. According to [25], the consistency of the EKF is defined as follows.

Definition 1. As for EKF algorithm, if the mean square error matrix satisfies the following property:

$$E(X_k - \hat{X}_k)(X_k - \hat{X}_k)^{\mathrm{T}} \leqslant P_k, \quad k = 0, 1, \ldots,$$

then, we say that the EKF algorithm have the consistency. The CEKF for (25) is as follows:

$$\begin{cases} \bar{X}_{k+1} = f(\hat{X}_k, 0), \\ \bar{P}_{k+1} = F_k P_k F_k^{\mathrm{T}} + L_k Q_k L_k^{\mathrm{T}} + \varDelta Q_k, \\ K_{k+1} = (\bar{P}_{k+1} H_{k+1}^{\mathrm{T}} + L_k Q_k M_{k+1}^{\mathrm{T}})(H_{k+1} \bar{P}_k H_{k+1}^{\mathrm{T}} \\ \qquad\qquad + M_{k+1} Q_k M_{k+1}^{\mathrm{T}} + H_{k+1} L_k Q_k M_{k+1}^{\mathrm{T}} + M_{k+1} Q_k L_k^{\mathrm{T}} H_{k+1}^{\mathrm{T}})^{-1}, \\ \hat{X}_{k+1} = \bar{X}_{k+1} + K_{k+1}(Y_{k+1} - h(\bar{X}_{k+1})), \\ P_{k+1} = \bar{P}_{k+1} - K_{k+1}(H_{k+1} \bar{P}_{k+1} + M_{k+1} L_k^{\mathrm{T}}) + \varDelta R_{k+1}, \end{cases} \tag{26}$$

where

$$F_k = \left.\frac{\partial f}{\partial X}\right|_{\hat{X}_k}, \quad L_k = \left.\frac{\partial f}{\partial w}\right|_{\hat{X}_k}, \quad H_{k+1} = \left.\frac{\partial h}{\partial X}\right|_{\bar{X}_{k+1}}, \quad M_{k+1} = \left.\frac{\partial h}{\partial w}\right|_{\bar{X}_{k+1}},$$

and $\varDelta Q_k$, $\varDelta R_{k+1}$ are the undetermined coefficients, satisfying the following lemma. Note that one can use the pseudo-inverse to replace the inverse if the term of K_{k+1} is irreversible.

Lemma 1. The EKF algorithm has consistency if $\varDelta Q_k$, $\varDelta R_k$ satisfy the following property:

$$\begin{cases} \varDelta Q_k \geqslant E[F_k^*(\hat{X}_k - X_k){\varphi_k}^{\mathrm{T}}] + E[\varphi_k(\hat{X}_k - X_k)^{\mathrm{T}} {F_k^*}^{\mathrm{T}}] + E[\varphi_k {\varphi_k}^{\mathrm{T}}] \\ \qquad\quad + E[F_k^*(\hat{X}_k - X_k)(\hat{X}_k - X_k)^{\mathrm{T}} {F_k^*}^{\mathrm{T}}] - F_k E[(\hat{X}_k - X_k)(\hat{X}_k - X_k)^{\mathrm{T}}] {F_k}^{\mathrm{T}}, \\ \varDelta R_k \geqslant -E[(I - K_k H_k^*)(\bar{X}_k - X_k){\psi_k}^{\mathrm{T}} {K_k}^{\mathrm{T}}] - E[K_k \psi_k (\bar{X}_k - X_k)^{\mathrm{T}} (I - K_k H_k^*)^{\mathrm{T}}] \\ \qquad\quad + E[K_k \psi_k {\psi_k}^{\mathrm{T}} K_k^{\mathrm{T}}] + E[K_k w_k w_k^{\mathrm{T}} {K_k}^{\mathrm{T}}] - K_k E[w_k w_k^{\mathrm{T}}] {K_k}^{\mathrm{T}} \\ \qquad\quad - (I - K_k H_k) E[(\bar{X}_k - X_k)(\bar{X}_k - X_k)^{\mathrm{T}}](I - K_k H_k)^{\mathrm{T}} \\ \qquad\quad + E[(I - K_k H_k^*)(\bar{X}_k - X_k)(\bar{X}_k - X_k)^{\mathrm{T}} (I - K_k H_k^*)^{\mathrm{T}}], \end{cases}$$

where $k = 0, 1, 2, \ldots$, $F_k^* = \frac{\partial f}{\partial X}|_{X_k}$, $H_k^* = \frac{\partial h}{\partial X}|_{X_k}$, φ_k and ψ_k respectively stand for the linearization errors in the process equation and measurement equation.

Remark 1. In the case that $\varDelta Q_k = 0$ and $\varDelta R_k = 0$, the CEKF (26) becomes the traditional EKF.

5 Algorithm of estimations

In this section, the NMLE method in Section 3 and the CEKF method in Section 4 are combined to estimate the parameters and the volatilities of the system (24). We call this algorithm as NMLE-CEKF algorithm. In this system, the volatility V_k and the price z_k correspond to the state variable X_k and the observation variables Y_k mentioned in (25), respectively. We update the parameters and the volatilities synchronously in each step.

The algorithm is presented as Algorithm 1.

Algorithm 1 NMLE-CEKF

Initialization:

$\hat{V}_0 = E(V_0)$,

$P_0 = E[(V_0 - \hat{V}_0^+)(V_0 - \hat{V}_0^+)^{\mathrm{T}}]$,

$\Theta_0 = (\hat{\kappa}_0, \hat{\kappa}_0\hat{\theta}_0, \hat{\sigma}_0,)^{\mathrm{T}}, \quad \rho = \hat{\rho}_0$.

Linearization matrices of the state function:

$F_k = 1 - \hat{\kappa}_k \Delta t$,

$L_k = \left[0, \hat{\sigma}_k \sqrt{\hat{V}_k \Delta t}\right]$.

Updating the state prediction estimation and the prediction estimation-error covariance:

$\bar{V}_{k+1} = \hat{V}_k + \hat{\kappa}_k\hat{\theta}_k\Delta t - \hat{\kappa}_k\hat{V}_k\Delta t$,

$\bar{P}_{k+1} = F_k P_k F_k^{\mathrm{T}} + L_k Q_k L_k^{\mathrm{T}} + \varDelta Q_k$.

Linearization matrices of the measurement function:

$H_{k+1} = -\frac{1}{2}\Delta t$,

$M_{k+1} = [\sqrt{(1-\hat{\rho}_k^2)\bar{V}_{k+1}\Delta t}, \hat{\rho}_k\sqrt{\bar{V}_{k+1}\Delta t}]$.

Updating the state estimate and error covariance:

$K_{k+1} = (\bar{P}_{k+1}H_{k+1}^{\mathrm{T}} + L_kQ_kM_{k+1}^{\mathrm{T}})(H_{k+1}\bar{P}_kH_{k+1}^{\mathrm{T}} + M_{k+1}Q_kM_{k+1}^{\mathrm{T}} + H_{k+1}L_kQ_kM_{k+1}^{\mathrm{T}} + M_{k+1}Q_kL_k^{\mathrm{T}}H_{k+1}^{\mathrm{T}})^{-1}$,

$\hat{V}_{k+1} = \bar{V}_{k+1} + K_{k+1}[z_{k+1} - \frac{1}{2}\bar{V}_{k+1}]$,

$P_{k+1} = \bar{P}_{k+1} - K_{k+1}(H_{k+1}\bar{P}_{k+1} + M_{k+1}L_k^{\mathrm{T}}) + \varDelta R_{k+1}$.

Parameters estimations:

Conduct the parameter estimation with (10), (23).

In this algorithm, $\varDelta Q_k$ and $\varDelta R_{k+1}$ can be designed as follows:

$$\varDelta Q_k = P_k|1-\hat{\kappa}\Delta t|^2_{\max} + \Delta t^2|\hat{\kappa}\hat{\theta}|^2_{\max} + |\hat{\sigma}_k|^2_{\max}\Delta t\hat{V}_kQ_k(2,2) - F_kP_kF_k^{\mathrm{T}} + L_kQ_kL_k^{\mathrm{T}},$$

$$\begin{aligned}\varDelta R_{k+1} =&\bar{P}_{k+1}\left(1+\frac{K_{k+1}\Delta t}{2}\right)^2 + 2K_{k+1}^2\Delta t\bar{V}_k\big((1-\hat{\rho}^2)Q_k(1,1)+\hat{\rho}^2Q_k(2,2)\big) - \bar{P}_{k+1}\\ &+ K_{k+1}(H_{k+1}\bar{P}_{k+1} + M_{k+1}L_k^{\mathrm{T}}).\end{aligned}$$

6 Simulation implementations

The simulation results in this section are designed to (i) reduce the discretization errors of Heston model, (ii) evaluate the effectiveness of PMLE and NMLE method with known volatilities, and (iii) utilize NMLE-CEKF algorithm in Section 5 to obtain the estimation results of parameters and volatilities. All the simulations are with the same initial settings unless otherwise stated, i.e., $V_0 = 0.2$, $S_0 = 100$, $r = 0.005$, $\kappa = 1$, $\theta = 0.250$, $\sigma = 0.5$, $\delta = 0.01$, $n = 1000$.

6.1 Euler discretization of Heston model

In this subsection, we compare three schemes to discrete Heston model. Note that the volatility process is the mainly difference between Heston model and other stochastic volatility models. The three discretization methods are especially distinct in the volatility process.

(i) A basic Euler method for Heston model would take the form of (24). However there is a problem that V_k can become negative values with non-zero probability. The discretization with smallest bias [35]

Table 1 Statistics of the three discretization methods

Statistics	Method (i)	Method (ii)	Method (iii)
Size	25000	25000	25000
Mean	0.220	0.220	0.222
Std	0.130	0.130	0.130
Skew	0.990	1.016	0.998
Kurtosis	1.352	1.391	1.321
Time (ms)	77	56	120

is proposed as

$$\begin{cases} \widetilde{V}_k = \widetilde{V}_{k-1} + \kappa(\theta - \widetilde{V}_{k-1}^{+})\delta + \sigma\sqrt{\widetilde{V}_{k-1}^{+}}\sqrt{\delta}\varepsilon_2, \\ V_k = \widetilde{V}_k^{+}, \end{cases} \tag{27}$$

$$z_k = \left(r - \frac{1}{2}V_k\right)\delta + \sqrt{(1-\rho^2)V_k}\sqrt{\delta}\varepsilon_1 + \rho\sqrt{V_k}\sqrt{\delta}\varepsilon_2, \tag{28}$$

where $\widetilde{V}_k^{+} = \max(V_k, 0)$, $\varepsilon_i \sim N(0,1)$, $i = 1, 2$. This scheme is called "full truncation" [36]. The full truncation ensures the positive of V_k.

(ii) The second discretization scheme is based on NMLE method. Discretization of the volatility model is transformed into (9) with Itô lemma, then discretization of (27) can be substituted as follows:

$$\begin{cases} \sqrt{\widetilde{V}_k} = \sqrt{\widetilde{V}_{k-1}} + \dfrac{1}{2\sqrt{\max(\widetilde{V}_{k-1}, \eta)}}\left[\kappa\theta - \kappa V_{k-1} - \dfrac{1}{4}\sigma^2\right]\delta + \dfrac{1}{2}\sigma\sqrt{\delta}\varepsilon_2, \\ \sqrt{V_k} = \sqrt{\widetilde{V}_k^{+}}, \end{cases} \tag{29}$$

where η is a positive real number which is close enough to zero.

(iii) The last method to simulate the volatility model is based on the exact distribution of V_k. As we mentioned in Section 3, defining the degrees of freedom as $\mathrm{d}f = \frac{4\kappa\theta}{\sigma^2}$, the noncentrality parameter as $\mathrm{nc} = \frac{4\kappa \mathrm{e}^{-\kappa\delta}}{\sigma^2(1-\mathrm{e}^{-\kappa\delta})}V_{k-1}$. the distribution of V_k is $\frac{\mathrm{e}^{-\kappa\delta}}{\mathrm{nc}}V_{k-1}$ times the noncentral chi-square with $\mathrm{d}f$ and nc [35]. Then V_k can be written as follows:

$$V_k = \frac{\sigma^2(1-\mathrm{e}^{-\kappa\delta})}{4\kappa}\chi'^2(\mathrm{d}f, \mathrm{nc}),$$

where $\chi'^2(\cdot,\cdot)$ is a noncentral chi-square random number. Nevertheless, this exact way to simulate volatilities is time-consuming compared to the Euler discretization method. The reason is that the computational requirements for taking a random sample from non-central chi-square distributions are higher than from the standard normal distribution. Table 1 shows the statistic results of the three simulation methods. The identical first-order moments and the second order moments indicate that the Euler method (Methods (i) and (ii)) can approximate the exact method (Method (iii)) precisely. The differences appear in the higher moments. Meanwhile, the exact method will take almost twice as long as the Euler methods. Therefore considering the computing speed of the Monte Carlo, the Euler method (i) or (ii) is a good substitution of the exact method to simulate the data.

Then we compare the simulation results of Euler methods (i) and (ii). It is well known that the discretization errors are reduced as the time step shrinks. In Figure 1, the sample number of n is 1000, the step size of δ is 0.01, and the maximum absolute error $\max(|V_k - (\sqrt{V_k})^2|)$ $(k = 1, 2, \ldots, n)$ of volatilities simulated by the Euler methods (i) (Eq. (27)) and (ii) (Eq. (29)) is 2.94×10^{-2}. In Figure 2, the sample number and step size are changed with $\widetilde{n} = n \times s$, $\widetilde{\delta} = \frac{\delta}{s}$, where $s = 100$, and the sample is taken as V_{1+ks} and $(\sqrt{V_{1+ks}})^2$, $k = 0, 1, 2, \ldots, n-1$. The maximum absolute error is 1.26×10^{-3}. The maximum error is almost ten times declined compared to the former one. The simulation results show that with the decreasing of the step size, the errors between the two Euler methods are decreasing. In this paper, we use Euler method (ii) with (28) and (29) to simulate Heston model.

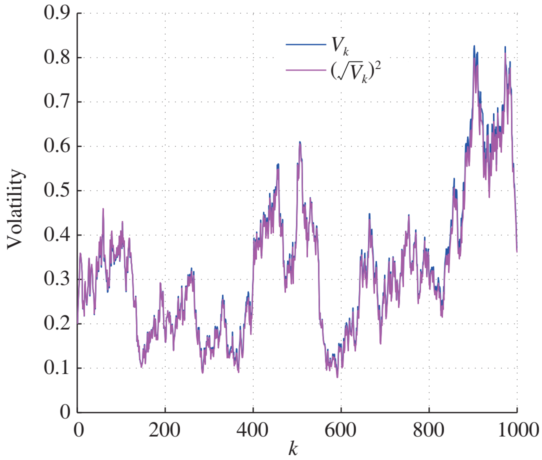

Figure 1 (Color online) Values of V_k and $(\sqrt{V_k})^2$ in the condition that $\delta = 0.01$, $n = 1000$, where $\max(|V_k - (\sqrt{V_k})^2|) = 2.94 \times 10^{-2}$, $k = 1, 2, \ldots, n$.

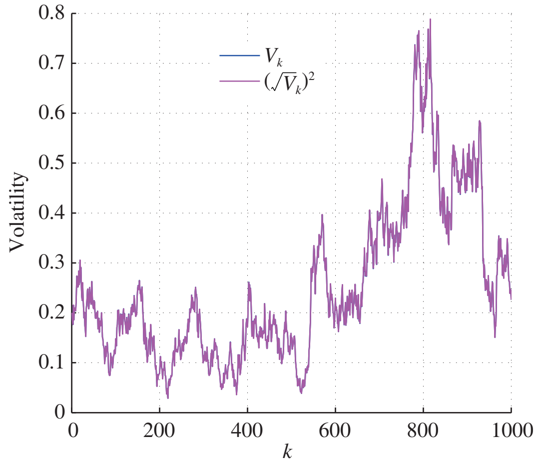

Figure 2 (Color online) Values of V_{1+ks} and $(\sqrt{V_{1+ks}})^2$ in the condition that $\tilde{n} = 1000 \times 100$, $\tilde{\delta} = \frac{\delta}{100}$, where $\max(|V_k - (\sqrt{V_k})^2|) = 1.26 \times 10^{-3}$, $k = 0, 1, 2, \ldots, n-1$.

6.2 Parameter estimation results with known volatilities

In this subsection, we evaluate the effectiveness of NMLE method under the assumption of known volatilities. To make a clear illustration, the NMLE method is compared with the PMLE method proposed by [37], which is

$$\hat{\kappa} = -\frac{1}{\delta}\log(\hat{\beta}_1), \quad \hat{\theta} = \hat{\beta}_2, \quad \hat{\sigma}^2 = \frac{2\hat{\kappa}\hat{\beta}_3}{1 - \hat{\beta}_1^2},$$

where

$$\hat{\beta}_1 = \frac{n^{-2}\sum_{k=1}^n V_k \sum_{k=1}^n V_{k-1}^{-1} - n^{-1}\sum_{k=1}^n V_k V_{k-1}^{-1}}{n^{-2}\sum_{k=1}^n V_{k-1}\sum_{k=1}^n V_{k-1}^{-1} - 1},$$

$$\hat{\beta}_2 = \frac{n^{-1}\sum_{k=1}^n V_k V_{k-1}^{-1} - \hat{\beta}_1}{(1 - \hat{\beta}_1)n^{-1}\sum_{k=1}^n V_{k-1}^{-1}},$$

$$\hat{\beta}_3 = n^{-1}\sum_{k=1}^n \{V_k - V_{k-1}\hat{\beta}_1 - \hat{\beta}_2(1 - \hat{\beta}_1)^2\}V_{k-1}^{-1}.$$

Figure 3 shows the parameters estimated by NMLE and PMLE algorithm. The bias and standard deviation of NMLE method are shown in Table 2. Although some extraordinary values may appear at the beginning, the estimations of parameters κ, θ, σ converge to their true values as k gets larger. In the process of estimating, the bias of $\hat{\kappa}$, $\hat{\theta}$, $\hat{\sigma}$ with PMLE are bigger than NMLE as shown in Figure 3. Also, from Table 2 we can see that, as k gets larger, the bias and the standard deviation become smaller with NMLE method. Additionally, $\hat{\rho}$ is estimated by (23) with $\hat{\kappa}$, $\hat{\theta}$, $\hat{\sigma}$. In Figure 4 we can see that $\hat{\rho}$ converges to its true value.

6.3 Parameter estimation results with unknown volatilities

In practical, the volatilities are unknown. Thus, filtering methods are implemented to track the volatilities (states) based on the stock prices (observations). Since the Heston model (24) we studied is a nonlinear system, the EKF method can be utilized. While the EKF method cannot evaluate the errors of the nonlinear model, the CEKF method is preferred to estimate the volatilities. In the subsequent simulations, the comparison of CEKF and EKF algorithm with true parameters are simulated firstly, then the NMLE and PMLE are combined with CEKF (NMLE-CEKF, PMLE-CEKF), respectively to show the effectiveness of NMLE-CEKF algorithm.

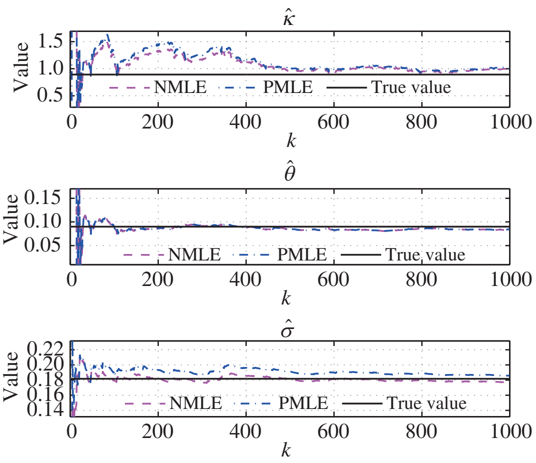

Figure 3 (Color online) Estimation of κ, θ, σ with NMLE and PMLE.

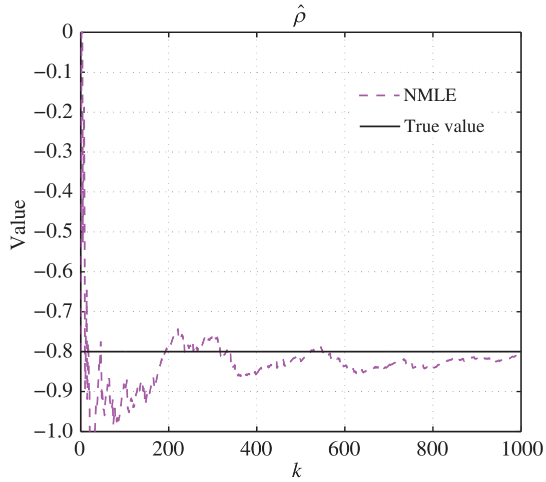

Figure 4 (Color online) Estimation of ρ with NMLE.

Table 2 Bias and standard deviation (SD) of NMLE for volatility model

Condition		κ	θ	σ
True value		0.8920	0.0900	0.182
$k = 100$	Bias	0.3068	0.0032	−0.0094
	SD	0.4683	0.0225	0.0126
$k = 300$	Bias	0.0921	0.0006	−0.0074
	SD	0.2497	0.0121	0.0074
$k = 1000$	Bias	0.0074	−0.0003	−0.0072
	SD	0.1449	0.0066	0.0038

The volatility tracking results as well as the square errors with CEKF and EKF algorithm are presented in Figure 5. The relationships between the estimation covariance matrix and P_k are present in Figure 6. These figures indicate that not only the CEKF algorithm is more accurate than EKF algorithm, but also P_k can upper bound the estimation covariance in each step. The simulation results conform to the formal theoretical analysis. The estimation covariance is calculated as the mean square error.

As for the unknown parameters and volatilities, NMLE for the parameters and filtering algorithms are combined to achieve estimations synchronously. Since the parameter ρ is sensitive to the states, and the estimation of ρ depends on the estimation of the other three parameters. The accumulated errors of both the states and parameters will lead to larger estimation bias for ρ. Thus in the following simulations, ρ is set as a fixed prior value. The volatility tracking results and square errors of NMLE-CEKF and PMLE-CEKF are shown in Figure 7. The parameter estimation results are shown in Figure 8[2)]. We can see that both the volatilities and parameters estimated with NMLE-CEKF are more accurate than the PMLE-EKF algorithm[3)].

7 Empirical applications

In this section, we apply NMLE-CEKF algorithm to S&P 500 index to testify the performance of the algorithm.

2) The parameters of the previous 1000 steps are set fixed as $\kappa_0 = 3$, $\theta_0 = 0.15$, $\sigma_0 = 0.35$.

3) Before $k = 1000$, parameters are assumed to be constant and then to be estimated when $k > 1000$, hence there is a sudden jump at $k = 1000$ in Figure 8.

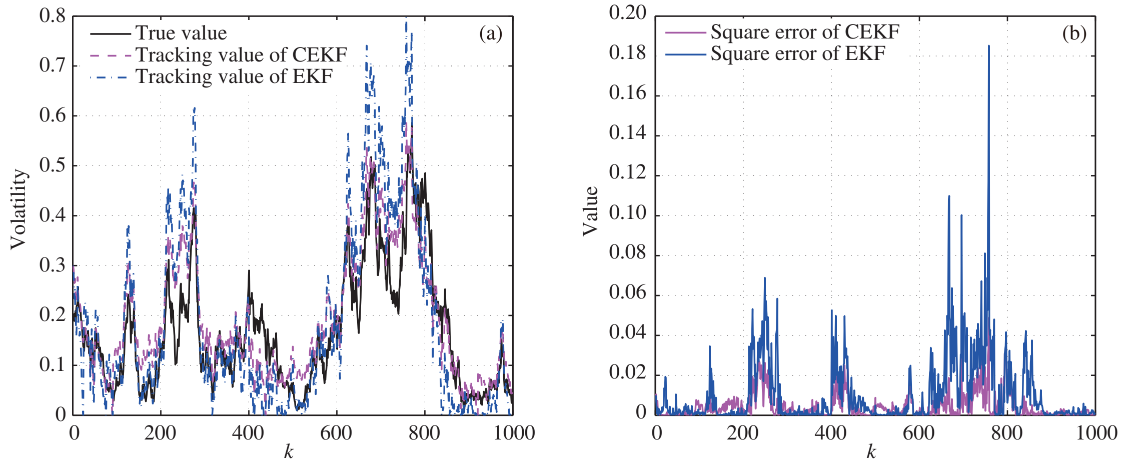

Figure 5 (Color online) (a) Volatility tracking with CEKF and EKF; (b) square errors of CEKF and EKF.

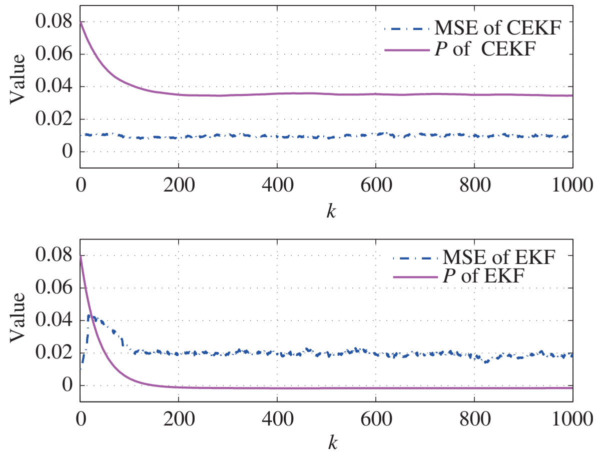

Figure 6 (Color online) The relationship of MSE and P_k for CEKF and EKF.

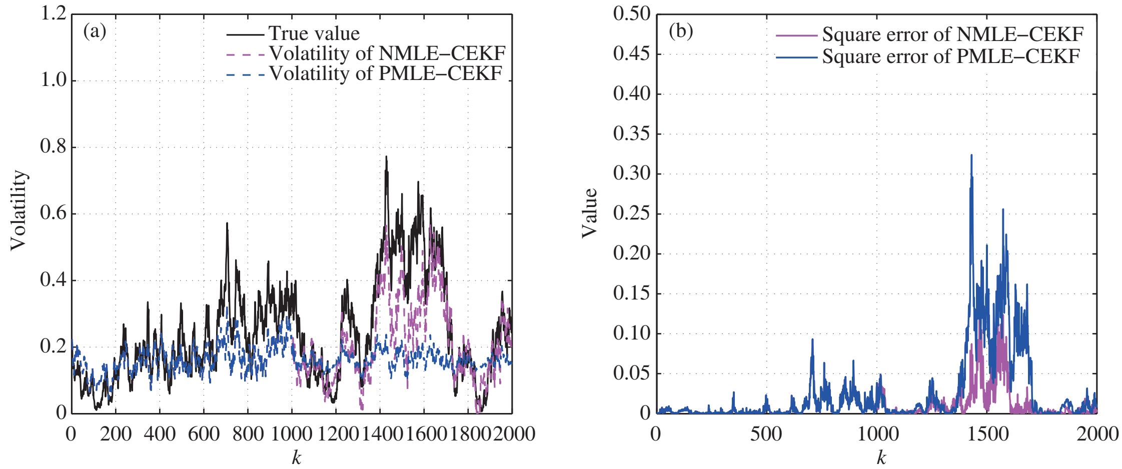

Figure 7 (Color online) (a) Volatility tracking results with NMLE-CEKF and PMLE-CEKF; (b) volatility errors with NMLE-CEKF and PMLE-CEKF.

7.1 Data description

The algorithm is applied to daily S&P 500 index ranging from December 1, 2006 to November 4, 2016 (regardless of non-trading days), 2500 observations. S&P 500 is an American stock market index recording

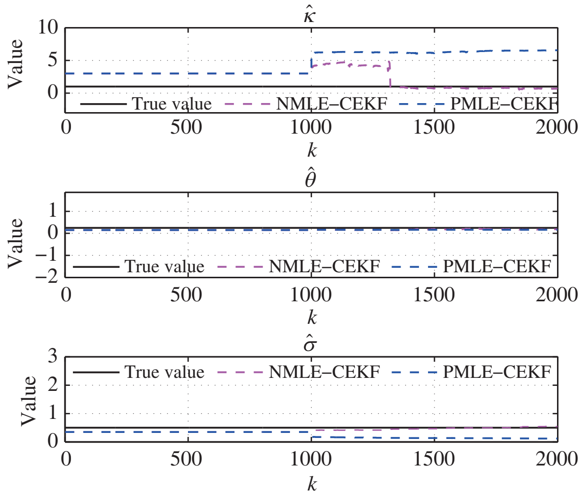

Figure 8 (Color online) Estimation of κ, θ, σ with NMLE-CEKF and PMLE-CEKF.

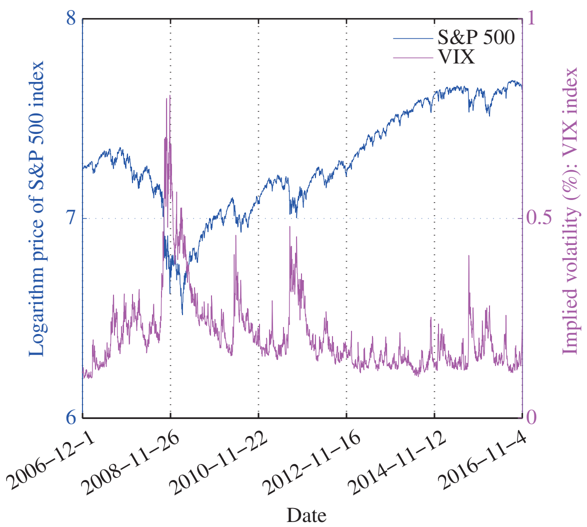

Figure 9 (Color online) The logarithm price of S&P 500 (SPX) index value and VIX.

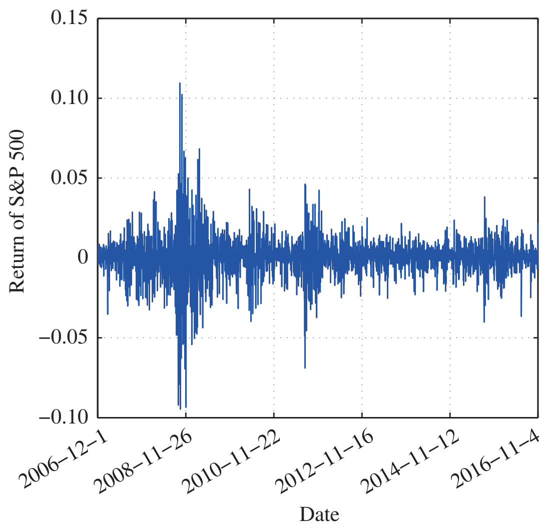

Figure 10 (Color online) Return of S&P 500 (SPX) index value.

500 public companies. It covers all the large market capitalization companies listed on the New York stock exchange (NYSE) or national association of securities dealers automated quotations (NASDAQ). Therefore the risk of S&P 500 is more diversing. People consider S&P 500 as one of the best indications of the U.S. stock market, and it can reveal the U.S. economy. Figure 9 shows the logarithm price of S&P 500 index value. In the following estimations, we use the average value of LIBOR (London Interbank Offered Rate) during 12 months in the years observed and adjust it with appropriate dividend yield.

We compare the volatilities tracked by NMLE-CEKF algorithm with VIX and history volatilities during the same period. The VIX is calculated by Chicago board options exchange (CBOE). It is an annualised percentage value measuring 30-day expected volatility of the S&P 500 Index. The new method to calculate VIX was proposed on September 22, 2003, including near-term and next-term put and call options with more than 23 days and less than 37 days to expiration [38]. VIX is an approximate value of the implied volatility of the S&P 500. History volatilities are the real volatilities of the past return rates, the standard deviation intuitively. VIX reflects the future expectation of the volatilities and the history volatilities reflect the past. They are simply reference values in this subsection (see more details in [39]).

Figures 9, 10, and Table 3 show the stylised facts [3, 7]. From Figure 9 we can see that the returns

Table 3 Statistics of the S&P data

Mean	Variance	Skewness	Kurtosis
1.69×10^{-4}	1.75×10^{-4}	-0.32	12.86

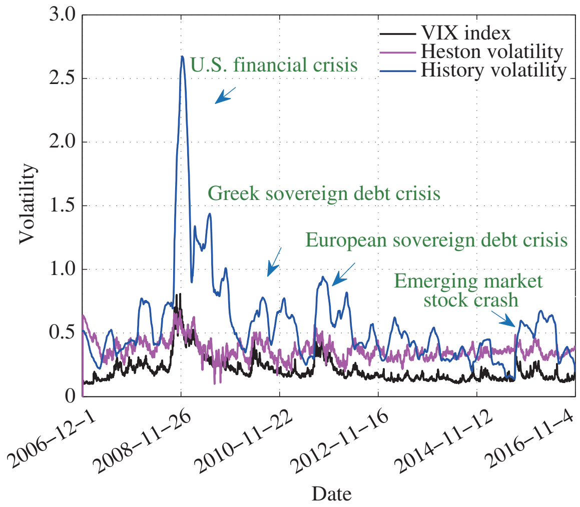

Figure 11 (Color online) The volatility trackings of S&P 500 of Heston model compared with VIX index and history volatilities.

of equity prices are negatively correlated with the volatilities, and the volatilities have a character of mean reverting. Thus there is an average level for the volatilities in the long run. Figure 10 shows the volatility clustering effect, which means that the large prices changes are often followed by large ones, and vice versa. The statistic features of price returns are shown in Table 3. The S&P index has a small positive average return, and the daily variance is 1.75×10^{-4}. The skewness coefficient indicates that the return distribution is negatively skewed and the kurtosis coefficient is very high. The kurtosis measures the thickness of distribution tails.

7.2 Estimation results

With the PMLE-CEKF algorithm, the volatility tracking results is shown in Figure 11. The VIX index and the 90-day short time history volatilities are also presented in Figure 11 as reference values. We can see that, the volatilities estimated by the NMLE-CEKF algorithm are between history values and VIX index. The estimation results can also capture some big events in the market. As the change of time, the return rate of mean reversion converges to 8.31, the long run mean converges to 0.13, and the volatility of volatility converges to 0.66.

8 Conclusion

Heston model provides us a solid theory basis to study the movement of equity market. In this paper, we propose the NMLE method to estimate the parameters of Heston model and combine NMLE with CEKF algorithm to estimate the parameters and volatilities under the condition of unknown volatilities. The equity and volatility equations are considered separately with the NMLE method, which reduces the complexity of solving the joint likelihood function of the system function. The closed form of the MLE method is obtained. It is shown to converge faster than PMLE by simulations. An upper bound for the estimation error in each step is designed by the CEKF algorithm, which compensates the linearization error of the EKF for nonlinear systems. For illustration, we compare the estimation results of NMLE and PMLE with the known and unknown states, respectively. Through the numerical simulation results, the parameter and volatilities estimations act efficiently with NMLE and NMLE-CEKF algorithm. Eventually, the NMLE-CEKF is implemented to S&P 500 index to estimate the parameters in the market.

Heston model is widely used in simulating the scenarios of equity movement. If the parameters is estimated with NMLE-CEKF precisely in the Heston model, it can be used to price the option prices as well as to calculate risks of some risk factors, such as foreign exchange rate and rate of interest. Specifically, if the parameter in Heston model is estimated with historical market equities by NMLE-CEKF method, this model can be used to calculate counterparty credit risk [40], which has attracted more and more interests after world wide financial crisis.

Acknowledgements This work was in part supported by National Key Research and Development Program of China (Grant No. 2016YFB0901902), National Natural Science Foundation of China (Grant No. 61622309), and National Basic Research Program of China (973 Program) (Grant No. 2014CB845301).

References

1 Black F, Scholes M. The pricing of options and corporate liabilities. J Polit Econ, 1973, 81: 637–654
2 Hull J. Options, futures, and other derivatives. In: Asset Pricing. Berlin: Springer, 2009. 9–26
3 Gallant A R, Tauchen G. Which moments to match? Economet Theory, 1996, 12: 657–681
4 Fama E F. The behavior of stock-market prices. J Bus, 1965, 38: 34–105
5 Cont R. Empirical properties of asset returns: stylized facts and statistical issues. Quant Financ, 2001, 1: 223–236
6 Engle R F, Ng V K. Measuring and testing the impact of news on volatility. J Financ, 1993, 48: 1749–1778
7 Engle R F, Patton A J. What good is a volatility model. Quant Financ, 2001, 1: 237–245
8 Hull J, White A. The pricing of options on assets with stochastic volatilities. J Financ, 1987, 42: 281–300
9 Wiggins J B. Option values under stochastic volatility: theory and empirical estimates. J Financ Econ, 1987, 19: 351–372
10 Heston S L. A closed-form solution for options with stochastic volatilities with applications to bond and currency options. Rev Financ Stud, 1993, 6: 327–343
11 Moodley N. The Heston model: a practical approach with matlab code. Johannesburg: University of the Witwatersrand, 2005
12 Dupire B. Pricing and hedging with smiles. In: Mathematics of Derivative Securities. Cambridge: Cambridge University Press, 1997. 103–111
13 Weron R, Wystup U. Heston's Model and the Smile. Berlin: Springer, 2005. 161–181
14 Tang C Y, Chen S X. Parameter estimation and bias correction for diffusion processes. J Economet, 2009, 149: 65–81
15 Aït-Sahalia Y, Kimmel R. Maximum likelihood estimation of stochastic volatility models. J Financ Econ, 2007, 83: 413–452
16 Ren P. Parametric estimation of the Heston model under the indirect observability framework. Dissertation for Ph.D. Degree. Houston: University of Houston, 2014
17 Jacquier E, Polson N G, Rossi P E. Bayesian analysis of stochastic volatility models. J Bus Econ Stat, 1994, 20: 69–87
18 Nelson D B. The time series behavior of stock market volatility and returns. Dissertation for Ph.D. Degree. Cambridge: MIT, 1988
19 Javaheri A, Lautier D, Galli A. Filtering in finance. Wilmott, 2003, 3: 67–83
20 Aihara S I, Bagchi A, Saha S. On parameter estimation of stochastic volatility models from stock data using particle filter-application to AEX index. Int J Innov Comput Inf Control, 2009, 5: 17–27
21 Li J. An unscented Kalman smoother for volatility extraction: evidence from stock prices and options. Comput Stat Data Anal, 2013, 58: 15–26
22 Pitt M K, Shephard N. Filtering via simulation: auxiliary particle filters. J Am Stat Assoc, 1999, 94: 590–599
23 Hu J, Wang Z D, Shen B, et al. Quantised recursive filtering for a class of nonlinear systems with multiplicative noises and missing measurements. Int J Control, 2013, 86: 650–663
24 Hu J, Wang Z D, Liu S, et al. A variance-constrained approach to recursive state estimation for time-varying complex networks with missing measurements. Automatica, 2016, 64: 155–162
25 Jiang Y G, Xue W C, Huang Y, et al. The consistent extended Kalman filter. In: Proceedings of Chinese Control Conference (CCC), Chengdu, 2014. 6838–6843
26 He X K, Jing Y G, Xue W C, et al. Track correlation based on the quasi-consistent extended Kalman filter. In: Proceedings of the 34th Chinese Control Conference (CCC), Hangzhou, 2015. 2150–2155
27 Jiang Y G, Huang Y, Xue W C, et al. On designing consistent extended Kalman filter. J Syst Sci Complex, 2017, 30: 751–764
28 Shreve S E. Stochastic Calculus for Finance II: Continuous-time Models. New York: Springer Science & Business Media, 2004
29 Cox J C, Ingersoll J E, Ross S A. A theory of the term structure of interest rates. Econometrica, 1985, 53: 385–407
30 Nykvist J. Time consistency in option pricing models. Dissertation for Master Degree. Stockholm: Royal Institute of Technology, 2009

31 Hamilton J D. Time Series Analysis. Princeton: Princeton University Press, 1994
32 Stein E M, Stein J C. Stock price distributions with stochastic volatility: an analytic approach. Rev Financ Stud, 1991, 4: 727–752
33 Kladívko K. Maximum likelihood estimation of the Cox-Ingersoll-Ross process: the matlab implementation. Technical Computing Prague, 2007. http://www2.humusoft.cz/www/papers/tcp07/kladivko.pdf
34 Simon D. Optimal State Estimation: Kalman, H_∞, and Nonlinear Approaches. Hoboken: John Wiley & Sons Inc, 2006. 395–407
35 Andersen L B G. Simple and efficient simulation of the Heston stochastic volatility model. J Comput Financ, 2008, 11
36 Lord R, Koekkoek R, Dijk D V. A comparison of biased simulation schemes for stochastic volatility models. Quant Financ, 2010, 10: 177–194
37 Tang C Y, Chen S X. Parameter estimation and bias correction for diffusion processes. J Economet, 2009, 149: 65–81
38 Exchange C B O. The CBOE volatility index-VIX. White Paper, 2009. www.cboe.com/micro/vix/vixwhite.pdf
39 Christensen B J, Prabhala N R. The relation between implied and realized volatility. J Financ Econ, 1998, 50: 125–150
40 Zhu S H, Pykhtin M. A guide to modeling counterparty credit risk. Social Sci Elect Pub, 2008, 1: 16–22