Proceedings of 2012 International Conference on
Modelling, Identification and Control, Wuhan, China, June 24-26, 2012

Calibration of Stochastic Differential Equation Models Using Implicit Numerical Methods and Particle Swarm Optimization

Tianhai Tian and Xiangyu Ge

***Abstract*— Stochastic differential equation (SDE) is a very important mathematical tool to describe complex systems in which noise plays an important role. SDEs have been widely used to study various nonlinear systems in biology, engineering, finance and economics, as well as physical sciences. Since a SDE can generate unlimited number of trajectories, it is a difficult problem to estimate model parameters based on experimental observations which may represent only one trajectory of the stochastic model. During the last decade substantial research efforts have been made to the development of effective methods for inferring parameters in SDE models. However, it is still a challenge to estimate parameters in SDE models from observations with large variations. In this work, we proposed to use the implicit numerical methods to simulate SDE models in order to generate stable trajectories for estimating parameters in stiff SDE models. In addition, we used the particle swarm optimization to search the optimal parameters from the parameter space that has a complex model error landscape. Numerical results suggested that the proposed algorithm is an effective approach to estimate parameters in SDE models.**

***Index Terms*— Stochastic differential equation; Calibration; Interest rate model; Particle swarm optimization.**

I. INTRODUCTION

A stochastic differential equation (SDE) can be defined as a deterministic differential equation perturbed by random disturbances that are not necessarily small. For example, certain key parameters in the deterministic model can be replaced by random variables to represent the varied environmental and/or experimental conditions. SDEs have gained in popularity in recent years for their ability to model systems that are subject to fluctuations, and has been widely used in a variety of disciplines including engineering, environmetrics, physical sciences, population dynamics, biology and medicine. In particular, SDEs are also central to much of modern finance theory and have been used profitably to model the behaviour of key variables such as the instantaneous short-term interest rate, asset prices, asset returns and their volatility. Consequently, the estimation of the parameters of SDEs from discretely-sampled data has received substantial attention in the financial econometrics literature, particularly in the last ten years [1], [2], [3].

However, parameter estimation in nonlinear SDEs driven by Wiener processes, when only discrete observation is available, is an inherently difficult problem because theoretically an unlimited number of solutions exist for a SDE. Although parameter estimation in deterministic systems is a relatively well-studied subject (see for example, Beck and Arnold [4]), the estimation of the parameters of stochastic systems remains a challenge [5]. One of the reasons for this is that obtaining the solution of a set of SDEs is computationally demanding in the absence of a closed-form solution for most SDEs of practical importance. Numerical methods such as the Euler methods and the Taylor schemes combined with a Monte-Carlo approach are required to calculate discrete-time trajectories of the state variables of SDEs [6]. These numerical methods require the generation of a large number of Wiener processes corresponding to different simulation trajectories, and hence accurate simulations are computationally expensive.

The methods that were developed for parameter estimation of SDEs can be classified into three different categories: maximum likelihood estimation (MLE)/simulated maximum likelihood (SML) [7], [8], the methods of moments [9], [10], and filtering (e.g. extended Kalman filter) [11]. Many of the methods have been developed in the context of financial modelling, where the systems of interest are characterised by long time horizons (of the order of months) and can often be sampled at regular but relatively infrequent intervals (e.g. on a daily basis). Among them, the maximum likelihood method is more reliable but has long been found to be difficult to apply to SDEs due to its computational cost. For many SDEs, thousands of simulation trajectories or even more must be generated to ensure a low variance of the values of state variables. Maximum likelihood estimates of the parameters of SDEs are consistent and asymptotically efficient, but unfortunately difficult to obtain if a closed-form expression for the transitional probability density function (PDF) of the process is not available. As a result, a large number of competing estimation procedures have been proposed in recent years.

There are a number of key issues that are related to the efficiency of the maximum likelihood methods. The first one is the slow convergent rate of Monte Carlo simulation. To address this issue, a number of variance-reduction techniques, including the importance sampler and random number generation techniques, have been employed to reduce the bias and thus increase the efficiency. The second issue is the numerical schemes to generate trajectories of the SDEs by using less time steps. Although attempts have been made to use high order numerical methods such as the Milstein method, the simplest Euler-Maruyara method is still dominant in

This work was supported by the Australian Research Council and the Chinese National Social Science Foundation.

T. Tian is with the School of Mathematical Sciences, Monash University, Melbourne VIC 3800, Australia. tianhai.tian@monash.edu

X. Ge is with the School of Statistics and Mathematics, Zhongnan University of Economics and Law, Wuhan 430073, China. xiangyu_ge@163.com

literature for estimating parameters in SDEs. In addition, nearly all the methods used in the literature are explicit methods that do not have satisfactory stability properties. The third issue is associated with the optimization methods for searching the optimal parameters when the search space is associated with a complex error landscape. The machine-learning techniques, including the Markov Chain Monte Carlo (MCMC) method, genetic algorithm (GA) and particle swarm optimization (PSO) algorithm, have been utilized for searching the optimal parameters of the SDE models in order to realize experimental observations [2], [12], [13]. Although it was reported that the genetic algorithms converge faster than the MCMC [14], our research results demonstrated that the genetic algorithm produced a wide range of estimates that all faithfully reproduced the experimental observations [15].

In this paper, we propose an algorithm to estimate parameters for stochastic models by advocating the usage of implicit methods to simulate SDEs. The semi-implicit Milstein method is particularly recommended in this work. In addition, we use the PSO algorithm to search the optimal solution in the maximum likelihood method. The paper is laid out as follows. Section 2 gives the stochastic models for term structure of interest rates and numerical algorithms for simulating these stochastic models. In Section 3, the algorithms used to obtain parameter estimates are described, and we give the optimization algorithm for searching the optimal estimates. Section 4 reports the numerical results for estimating parameters in the stochastic models for the term structure of interest rates.

II. Stochastic model

We first introduce numerical methods for simulating SDEs

$$dX = f(t,X)dt + g(t,X)dW(t), \tag{1}$$

where $f(t,X)$ is the drift term, $g(t,X)$ the diffusion term and $W(t)$ the Wiener process whose increment follows the Gaussian distribution, namely $\Delta W_n = W(t_{n+1}) - W(t_n) \sim N(0, t_{n+1} - t_n)$. The widely used method in computational finance is the Euler-Maruyama method whose strong convergence order is just 0.5, given by

$$X_{n+1} = X_n + hf(t_n, X_n) + g(t_n, X_n)\Delta W_n, \tag{2}$$

where X_n is the numerical solution at time point t_n, and $h = t_{n+1} - t_n$. Although this method is easy to implement, its stability property is not good enough to simulate SDEs with relatively large diffusion term. In order to obtain stable simulations, a very small stepsize is required, which may lead to large computing time.

The Milstein scheme uses a higher order Taylor expansion and thus has a strong convergence order 1.0, given by

$$\begin{aligned} X_{n+1} &= X_n + hf(t_n, X_n) + g(t_n, X_n)\Delta W_n \\ &\quad + \frac{1}{2} g(t_n, y_n) g'(t_n, y_n)((\Delta W_n)^2 - h). \end{aligned} \tag{3}$$

In order to improve the stability of the Milstein method, the semi-implicit and fully implicit Milstein methods were proposed. In the semi-implicit Milstein method only the drift term is implicit, namely

$$\begin{aligned} X_{n+1} &= X_n + hf(t_{n+1}, X_{n+1}) + g(t_n, X_n)\Delta W_n \\ &\quad + \frac{1}{2} g(t_n, y_n) g'(t_n, y_n)((\Delta W_n)^2 - h); \end{aligned} \tag{4}$$

while the fully implicit method has implicit drift and diffusion terms [16], given by

$$\begin{aligned} X_{n+1} &= X_n + hf(t_{n+1}, X_{n+1}) + g(t_{n+1}, X_{n+1})\Delta W_n \\ &\quad + \frac{1}{2} g(t_{n+1}, y_{n+1}) g'(t_{n+1}, y_{n+1})((\Delta W_n)^2 + h). \end{aligned}$$

In this paper we used two models of the term structure of interest rate as the test systems to examine the accuracy of inference methods [2]. The first stochastic model is the CIR model (Cox, Ingersoll and Ross), which is a linear mean reversion model and uses a diffusion process [17]. This stochastic model has been widely used to model the short interest rate [18], [19]. The CIR model states that the short rate follows a square root diffusion process, which has the following continuous-time representation:

$$dX = \alpha(\beta - X)dt + \sigma\sqrt{X}dW(t), \tag{5}$$

where α is the speed of adjustment (or mean reversion), β represents the long run mean of the short-term interest rate, and σ is a constant volatility. Under this model, both the drift and volatility change with the level of the short rate.

The second benchmark model is the Ornstein-Uhlenbeck (OU) model. The OU process proposed by Vasicek evolves according to the SDE

$$dX = \alpha(\beta - X)dt + \sigma dW(t), \tag{6}$$

where α is the speed of adjustment, β represents the mean of interest rate, and σ is the volatility control parameter. The OU process also exhibits mean reversion of X to the state $X = \beta$, but unlike the CIR process, the domain of the state variable is unrestricted, that is, the sample space of the model is the real line.

It was reported that the explicit Milstein method (3) had increased the estimation accuracy over the Euler method (2). In this work we compare the accuracy of the Milstein method (3) and the semi-implicit Milstein method (4). For the two stochastic models of interest rate, the application of the Milstein method is straightforward. Due to the linear feature of the drift term in the interest rate models, the semi-implicit method can be written in an explicit formula.

For the two-benchmark models, the Milstein Scheme is

(1) the CIR model

$$\begin{aligned} X_{n+1} &= \frac{1}{1+\alpha h}\left(X_n + \alpha\beta h + \sigma\sqrt{X_n}\Delta W_n \right. \\ &\quad \left. + \frac{\sigma^2}{4}(\Delta W_n^2 - h)\right); \end{aligned} \tag{7}$$

(2) the OU model

$$X_{n+1} = \frac{1}{1+\alpha h}\left(X_n + \alpha\beta h + \sigma\Delta W_n\right). \tag{8}$$

III. METHOD OF MOMENT

The task in this paper is to estimate parameter θ in the one-dimensional SDE

$$dX = f(X,\theta)dt + g(X,\theta)dW(t), \tag{9}$$

For the simplicity of notations, time t is not explicitly represented in functions $f(X,\theta)$ and $g(X,\theta)$. If we have a sample of $(N+1)$ observations $X_0, X_1, \ldots, X_n$ of the process at known time points $t_0, t_1, \ldots, t_N$, the maximum-likelihood (ML) estimate of θ is generated by maximizing the likelihood function

$$L(\theta) = f_0(X_0|\theta)\prod_{k=0}^{N-1} f(X_{k+1}|X_k;\theta). \tag{10}$$

which is equivalent to the minimization of the negative log-likelihood function

$$-\log L(\theta) = -\log[f_0(X_0|\theta)] - \sum_{k=0}^{N-1}\log[f(X_{k+1}|X_k;\theta)]. \tag{11}$$

Here $f_0(X_0|\theta)$ is the density of the initial state and

$$f(X_{k+1}|X_k;\theta) \equiv f((t_{k+1},X_{k+1})|(t_k,X_k);\theta)$$

is the value of the transitional probability density function (PDF) at (t_{k+1},X_{k+1}) for the process starting at (t_k,X_k) and evolving to (t_{k+1},X_{k+1}). Note that the Markovian property of Equation (9) ensured that the transitional PDF satisfies the Fokker-Planck equation. Unfortunately, exact maximum likelihood estimation is rare because in most cases the closed-form solution of the Fokker-Planck equation is not available.

However, based on the sample of observations, the transitional PDF can be approximated by the numerical solution of the original SDE (9). For example, when using the Euler-Maruyama method to solve (9), given by

$$X_{k+1} = X_k + \Delta f(X_k,\theta) + g(X_k,\theta)\Delta W_n, \tag{12}$$

where Δ is the stepsize of experimental observations, the transitional PDF of X can be approximated by the normal PDF with mean $X_k + \Delta f(X_k,\theta)$ and variance $g^2(X_k,\theta)\Delta$. Thus the simplest version of discrete maximum likelihood, namely the method of moment, replaces the true transitional PDF $f(X_{k+1}|X_k;\theta)$ in (11) by using

$$\frac{1}{g(X_k,\theta)\sqrt{2\pi\Delta}}\exp\left[-\frac{(X_{k+1}-X_k-\Delta f(X_k,\theta))^2}{2g^2(X_k,\theta)\Delta}\right]. \tag{13}$$

For the CIR process (5), it has been established that the optimal values $\overline{\alpha}$ and $\overline{\beta}$ of parameters α and β satisfy the equation [2]

$$\overline{\alpha}\left(\overline{\beta}\sum_{k=0}^{N-1}\Delta - \sum_{k=0}^{N-1}X_k\Delta\right) = X_N - X_0, \tag{14}$$

$$\overline{\alpha}\left(\overline{\beta}\sum_{k=0}^{N-1}\frac{\Delta}{X_k} - \sum_{k=0}^{N-1}\Delta\right) = \sum_{k=0}^{N-1}\frac{X_{k+1}-X_k}{X_k}, \tag{15}$$

and the optimal value of σ is

$$\overline{\sigma}^2 = \frac{1}{N}\sum_{k=0}^{N-1}\frac{(X_{k+1}-X_k-\overline{\alpha}(\overline{\beta}-X_k)\Delta)^2}{X_k\Delta}, \tag{16}$$

For the OU process (6), the optimal values $\overline{\alpha}$ and $\overline{\beta}$ of parameters α and β satisfy the equation [2]

$$\overline{\alpha}\left(\overline{\beta}\sum_{k=0}^{N-1}\Delta - \sum_{k=0}^{N-1}X_k\Delta\right) = X_N - X_0,$$

$$\overline{\alpha}\left(\overline{\beta}\sum_{k=0}^{N-1}X_k\Delta - \sum_{k=0}^{N-1}X_k^3\Delta\right) = \sum_{k=0}^{N-1}(X_{k+1}-X_k)X_k$$

and the optimal value of σ is

$$\overline{\sigma}^2 = \frac{1}{N}\sum_{k=0}^{N-1}\frac{(X_{k+1}-X_k-\overline{\alpha}(\overline{\beta}-X_k)\Delta)^2}{X_k\Delta}. \tag{17}$$

Although the Milstein variant has been proposed to increase the accuracy of the discrete maximum likelihood method [2], [14], it is difficult to derive an analytical expression of the parameter estimate from the transitional PDF. In that case the simulated maximum likelihood method is needed to estimate the transitional PDF from stochastic simulations.

IV. SIMULATED MAXIMUM LIKELIHOOD METHOD

Parameter estimation in deterministic models can be achieved by the best fit of numerical simulations to experimental observations. However, this method is not valid for SDE models because we can generate a unlimited number of trajectories from a single SDE model. Here we used methods based on stochastic models, such as the simulated maximum likelihood (SML) method [20], [21]. Based on a sequence of $N+1$ observations $\{\mathbf{X}_0,\mathbf{X}_1,\ldots,\mathbf{X}_N\}$ at time points $\{t_0,t_1,\ldots,t_N\}$, we define the joint transitional density or likelihood function of these observations as

$$f_0[(t_0,\mathbf{X}_0)|\theta]\prod_{i=1}^{N} f[(t_i,\mathbf{X}_i)|(t_{i-1},\mathbf{X}_{i-1}),\ldots,(t_0,\mathbf{X}_0);\theta], \tag{18}$$

where $\theta = (\theta_1,\ldots,\theta_s)$ are the undetermined parameters in model (9), $f_0[(t_0,\mathbf{X}_0)|\theta]$ is the density of the initial state, and

$$f[(t_i,\mathbf{X}_i)|(t_{i-1},\mathbf{X}_{i-1}),\ldots,(t_0,\mathbf{X}_0);\theta]$$

is the transitional density starting from $(t_{i-1},\mathbf{X}_{i-1})$ and evolving to $(t_i,\mathbf{X}_i)$. When the financial system is described by the stochastic model (9), the stochastic process $\mathbf{X}$ is Markov [22], and the transitional density can be simplified as

$$\begin{aligned} & f[(t_i,\mathbf{X}_i)|(t_{i-1},\mathbf{X}_{i-1}),\ldots,(t_0,x_0);\theta] \\ = \; & f[(t_i,\mathbf{X}_i)|(t_{i-1},\mathbf{X}_{i-1});\theta]. \end{aligned} \tag{19}$$

An equivalent form of the maximum of the joint transitional density (18) is the minimum of the negative log-likelihood function (11) when time t is not explicitly presented in the formula.

Because the closed-form expression of the transitional density (19) is usually unavailable, we use a nonparametric kernel density function

$$\overline{f}_M\left[(t,\mathbf{X})|(t_{i-1},\mathbf{X}_{i-1});\theta\right]=\frac{1}{MB}\sum_{j=1}^{M}K\left(\frac{\mathbf{X}-\mathbf{Y}_j}{B}\right) \quad (20)$$

to evaluate the transitional density based on the M realizations $\mathbf{Y}_1,\ldots,\mathbf{Y}_M$ of $\mathbf{X}_i$ at t_i given the initial condition $(t_{i-1},\mathbf{X}_{i-1})$. Here B is the kernel bandwidth and $K(\cdot)$ is a non-negative kernel function enclosing unit probability mass. In the case of SDE models with a single variable, the normal kernel is widely used and the bandwidth can be chosen as

$$B=0.9\sigma M^{-1/5},$$

where σ is the sample standard deviation of the M realizations [20]. For SDE models with multiple variables, we can either assume the independence of random variables or use the theory of multivariate density estimation [23]. Note that the same increments of the Wiener process should be used in numerical simulations with different values of parameter θ. Finally the optimal value of parameter θ can be estimated by minimizing the log-likelihood function (11) over θ. Thus we can derive the SML method for estimating unknown parameters in the SDEs model (9).

Algorithm 1

Step 1. Input the system states $\{\mathbf{X}_0,\mathbf{X}_1,\ldots,\mathbf{X}_N\}$ and time points $\{t_0,t_1,\ldots,t_N\}$.

Step 2. Take $\mathbf{X}_{i-1}$ at time t_{i-1} $(i=1,\ldots,N)$ as the starting value and use a numerical method to generate M realizations $\mathbf{Y}_1,\ldots,\mathbf{Y}_M$ of $\mathbf{X}$ at t_i. A random seed is specified for generating samples of the Gaussian random variables.

Step 3. Use the nonparametric density (20) with the normal kernel or multivariate density functions to evaluate the transitional density (19).

Step 4. Steps 2 and 3 are repeated for each time point $t_0,\ldots,t_{N-1}$, and results are used to construct the log-likelihood function (11).

Step 5. Search the optimal kinetic rates by a genetic optimisation algorithm based on the minimum of $L(\theta)$ in (11).

Note that in order to reduce the variation of the estimated parameters, the same random seeds (namely the same random samples) in Step 2 should be used in different candidate estimates of parameters.

The PSO algorithm is a population based stochastic optimization technique developed by Dr. Eberhart and Dr. Kennedy in 1995, inspired by social behaviour of bird flocking or fish schooling [12]. PSO shares many similarities with evolutionary computation techniques such as the GA. The system is initialized with a population of random solutions and searches for optima by updating generations. However, unlike GA, PSO has no evolution operators such as crossover and mutation. In PSO, the potential solutions, called particles, fly through the problem space by following the current optimum particles.

In past several years, PSO has been successfully applied in many research and application areas. There are a number of research papers that compared the accuracy and efficiency of the PSO method with other evolutionary algorithms such as the GA [13], [24]. General speaking, PSO and the GA were able to arrive at solutions with the same quality. However, PSO offers a less expensive approach than the GA in general, and the expected computational saving offered by PSO over the GA is problem dependent. Numerical tests suggested that PSO offered more computational saving for unconstrained nonlinear problems with continuous design variables whereas the computational saving is lower for constrained and mixed integer nonlinear problems [13]. Another reason that PSO is attractive is that there are few parameters to adjust. In recent years PSO has been used for approaches that can be used across a wide range of applications, as well as for specific applications focused on a specific requirement.

In this work we used a PSO MATLAB toolbox that was downloaded from the MATLAB File Exchange Central [25]. This is a carefully designed software system and can be implemented in estimating a wide range of optimization problems. It has the same syntax as the GA Toolbox and can be easily used by researchers with the experience of using the GA. This toolbox was used in this work to estimate parameters in SDE models.

V. Estimation of parameters in the interest models

Using the methods discussed in the previous sections, now we estimate parameters in the CIR model and OU model. For each model, we used a given set of parameters (namely exact parameters) to generate 20 trajectories by using the semi-implicit Milstein method with a very small stepsize ($h=0.001$). For each trajectory, we used the method of moment, the SML method with the explicit Milstein method, and the SML method with the semi-implicit Milstein method to estimate the model parameters. In the PSO algorithm, the population size is 40 and it was carried out over 200 generations. For each trajectory of the 20 trajectories of each model, we used the PSO algorithm to obtain 20 estimated sets of model parameters, from which we calculated the mean and standard deviation (STD) of the errors to the exact parameters. Numerical results were presented in Tables 1 and 2. It is clear the method of moment cannot generate reliable estimates of model parameters when the length of the consecutive observation time points (Δ), namely the stepsize of observations, is large. We obtained negative coefficients, which is meaningless in finance, even when the strength of noise is small. In some cases the relative error is more than 100%. In addition, there is substantial difference between the estimates of the same parameter in different implementations. However, our further tests suggested that the accuracy of the moment method depends on the length of observation points (Δ) and the number of observation time points. When the length is moderate and the number of observation points is

TABLE I
ESTIMATION RESULTS OF THE PARAMETERS IN THE CIR MODEL

	Moment method			Milstein			Semi-implicit		
	Mean	Bias	STD	Mean	Bias	STD	Mean	Bias	STD
Exact parameter ($\alpha = 0.2$, $\beta = 0.08$, $\sigma = 0.1$, $\Delta = 5$)									
α	0.0501	0.1499	0.0103	0.2052	0.0052	6.3E-4	0.2091	0.0091	1.0E-3
β	-0.0118	0.0918	0.0193	0.0814	0.0014	1.1E-4	0.0787	0.0013	1.6E-6
σ	0.1067	0.0933	0.0148	0.0955	0.0045	3.3E-5	0.0933	0.0067	3.6E-5
Exact parameter ($\alpha = 0.2$, $\beta = 0.08$, $\sigma = 0.2$, $\Delta = 5$)									
α	0.1013	0.0987	0.0253	0.2184	0.0184	0.0037	0.2116	0.0116	0.0028
β	0.0722	0.0078	0.0271	0.0812	0.0012	0.0005	0.0808	0.0008	0.0005
σ	0.3650	0.1650	0.2647	0.1835	0.0165	0.0006	0.1850	0.0150	0.0006
Exact parameter ($\alpha = 0.2$, $\beta = 0.08$, $\sigma = 0.3$, $\Delta = 5$)									
α	0.0907	0.1093	0.0490	0.3269	0.1269	0.0158	0.3023	0.1023	0.0096
β	0.0746	0.0054	0.0230	0.0496	0.0304	0.0005	0.0534	0.0266	0.0006
σ	0.3319	0.1319	0.2128	0.2434	0.0566	0.0015	0.2375	0.0625	0.0007

relatively large, the moment method could provide estimates with acceptable accuracy.

To examine the influence of noise in determining the estimation accuracy, we tested three values of the volatility control parameter σ in each SDE model. Numerical results suggested that, when the fluctuations are not large in the SDE model (the CIR model with $\sigma = 0.1$), the SML method with either the explicit Milstein method or the semi-implicit Milstein method provided reliable estimates of parameters with small estimation errors and standard deviations. For two of the three variables, the SML method with the explicit Milstein method actually provided more accurate estimates. The reason may be that for non-stiff SDEs the explicit Milstein method has better accuracy of simulations than the semi-implicit Milstein method.

When the fluctuations in the SDE models are moderate (the CIR model with $\sigma = 0.2$ and the OU model with $\sigma = 0.05$), numerical results suggested that, compared with the Milstein method, the semi-implicit Milstein method led to better parameter estimates that have smaller errors as well as smaller standard deviations. In this case, the better stability property of the semi-implicit method is more important than the slightly better accuracy property of the explicit method. However, if noise in the interest rate models is large (the CIR model with $\sigma = 0.3$ and the OU model with $\sigma = 0.1$), the estimated parameters by using the semi-implicit Milstein method have acceptable accuracy which is better than that of the parameters obtained by using the Milstein method. If the noise components in the SDE models are very large (the OU model with $\sigma = 0.2$), even the semi-implicit Milstein method could not produce reliable estimated model parameters. In that case, we may need to use a smaller stepsize in simulating the SDE models or use the fully implicit Milstein method [16] in order to guarantee the stability property of the numerical simulation.

VI. CONCLUSIONS AND FUTURE WORKS

This work presented an effective algorithm for the estimation of parameters in the SDE models. The proposed approach is based on the usage of implicit numerical scheme in the Monte-Carlo simulation for the integration of the discretised SDEs. In particular, we advocated the usage of the semi-implicit Milstein method which has both higher convergence order and better stability property than the widely used Euler-Maruyama method. Numerical results suggested that the semi-implicit Milstein method led to estimates with better accuracy than those obtained by the explicit Milstein method when the SDE model is moderate stiff. To accelerate the convergence rate of the stochastic searching method, we used the PSO method to find the optimal model parameters. Numerical results suggested that the PSO algorithm can produce reliable estimates which are nearly independent to the implementation details.

The proposed algorithm has been applied to two important stochastic models of the term structure of interest rate, which is a fundamental issue in the research areas of financial mathematics. By using different strength of fluctuations in stochastic models, we compared the effectiveness of stochastic simulation methods in estimating model parameters. Our results suggested that the semi-implicit Milstein method can lead to accurate estimates of the model parameters when the fluctuations are small or moderate. It is still a challenging problem for estimating parameters in stiff stochastic models. For the large errors of the estimates for still SDEs, the reason may be the stability property of the numerical schemes for simulating the SDEs. It is also possible that the current approaches for estimating model parameters may fail to generate reliable estimates if the experimental observations have large variations and the length between the consecutive observations is large, such as the case of biological experimental data [15]. Thus more effective calibration methods should be designed for estimating parameters in stiff SDEs.

Efficiency is another major issue of the calibration methods for stochastic models. Since thousands of trajectories are needed to calculate the transitional PDF, any improvement over the numerical efficiency will significantly reduce the total computational time. Sophisticated simulation schemes such as the variable-stepsize simulation methods are needed in stochastic simulations. In addition, the variance reduction techniques are important to reduce the trajectories numbers for calculating the transitional PDF and thus increase the

TABLE II

ESTIMATION RESULTS OF THE PARAMETERS IN THE OU MODEL

	Moment method			Milstein			Semi-implicit		
	Mean	Bias	STD	Mean	Bias	STD	Mean	Bias	STD
Exact parameter ($\alpha = 0.2$, $\beta = 0.08$, $\sigma = 0.05$, $\Delta = 5$)									
α	-0.1714	0.3714	0.0895	0.3232	0.1232	0.2005	0.2561	0.0561	0.1106
β	0.0928	0.0128	0.0234	0.0817	0.0017	0.0129	0.0792	0.0008	0.0138
σ	0.0400	0.1600	0.0051	0.0564	0.0064	0.0173	0.0499	0.0001	0.0102
Exact parameter ($\alpha = 0.2$, $\beta = 0.08$, $\sigma = 0.1$, $\Delta = 5$)									
α	0.2555	0.0555	8.4096	0.3184	0.1184	0.2342	0.2872	0.0872	0.2397
β	0.0829	0.0029	0.0394	0.0861	0.0061	0.0424	0.0905	0.0105	0.0457
σ	0.1028	0.0972	0.0718	0.1116	0.0116	0.0388	0.1066	0.0066	0.0355
Exact parameter ($\alpha = 0.2$, $\beta = 0.08$, $\sigma = 0.2$, $\Delta = 5$)									
α	0.8796	0.6796	3.3409	0.3799	0.1799	0.2918	0.3304	0.1304	0.2572
β	0.0924	0.0124	0.0568	0.0879	0.0079	0.0474	0.0985	0.0185	0.0631
σ	0.1589	0.0411	0.0539	0.2353	0.0353	0.0931	0.2169	0.0169	0.0839

computing efficiency. Finally, the optimization methods for searching the optimal estimate from the space with complex error landscape are particularly important in the calibration of stochastic models. All of these issues are interesting future research topics.

VII. ACKNOWLEDGMENTS

This research work was supported by the Australian Research Council (ARC) (DP120104460, FT100100748, and DP1094181) and the Chinese National Social Science Foundation (SSF) (10BJY104). The authors also gratefully acknowledge the contribution of reviewers' comments.

REFERENCES

[1] M.W. Brandt, P. Santa-Clara, Simulated Likelihood Estimation of Diffusions with an Application to Exchange Rate Dynamics in Incomplete Markets. *J. Finan. Econo.* vol. 63, 2002, pp. 161-210.

[2] A.S. Hurn, J.I. Jeisman and K.A. Lindsay, Seeing the Wood for the Tree: A Critical Evaluation of Methods to Estimate the Parameters of Stochastic Differential Equations, *J. Finan. Econo.*, vol. 5, 2007, pp390-455.

[3] G. De Rossi, Maximum Likelihood Estimation of the Cox-Ingersoll-Ross Model Using Particle Filter, *Comput. Econ*, vol. 36, 2010, pp. 1-16.

[4] J.V. Beck and K.J. Arnold, *Parameter Estimation in Engineering and Science*, Wiley, New York, 1977.

[5] J. Timmer, Parameter Estimation in Nonlinear Stochastic Differential Equations, *Chaos Solitons Fractals*, vol. 11, 2000, pp. 2571-2578.

[6] K. Burrage , P.M. Burrage and T. Tian, Numerical Methods for Strong Solutions of Stochastic Differential Equations: an Overview, *Proc. R. Soc. A-Math. Phys. Eng. Sci.*, vol. 460, 2004, pp. 373-402.

[7] G.B. Durham and A.R. Gallant, Numerical Techniques for Maximum Likelihood Estimation of Continuous-time Diffusion Processes, *J. Busi. Econo. Stat.*, vol. 20, 2002, pp. 297-316.

[8] D. Florens-Zmirou, Approximate Discrete-time Schemes for Statistics of Diffusion Processes, *Statistics*, vol. 20, 1989, pp. 547557.

[9] A.R. Gallant and G. Tauchen, The Relative Efficiency of Method of Moments Estimators, *J Econometrics*, vol. 92, 1999, pp. 149-172.

[10] B. Pereira Ló, A.J. Haslam and C.S. Adjiman, An Algorithm for the Estimation of Parameters in Models with Stochastic Differential Equations, *Chem. Engin. Sci.*, vol. 63, 2008, pp. 4820-4833.

[11] J.N. Nielsen and H. Madsen, Applying the EKF to Stochastic Differential Equations with Level Effects, *Automatica*, vol. 37, 2001, pp. 107-112.

[12] R.C. Eberhart and J. Kennedy, A New Optimizer Using Particles Swarm Theory, *in Proc. Sixth International Symposium on Micro Machine and Human Science (Nagoya, Japan)*, IEEE Service Center, Piscataway, NJ, 1995, pp. 39-43.

[13] R. Hassan, B. Cohanim, O.L. DeWeck and G. Venter, A Comparison of Particle Swarm Optimization and the Genetic Algorithm, *Proceeedings of the First AIAA Multidisciplinary Design Optimization Specialist Conference*, 2005, pp. 1821

[14] O. Elerian, A Note on the Existence of a Closed Form Conditional Transition Density for the Milstein Scheme, *Working paper*, Nuffield College, Oxford University, 1998.

[15] T. Tian, S. Xu, J. Gao and K. Burrage, Simulated Maximum Likelihood Method for Estimating Kinetic Rates in Gene Expression, *Bioinformatics*, vol. 23, 2007, pp84-91.

[16] T. Tian and K. Burrage, Implicit Taylor Methods for Stiff Stochastic Differential Equations, *Applied Numer. Maths.*, vol. 38, 2001, pp 167-185.

[17] J.C. Cox, J.E. Ingersoll and S.A. Ross, A Theory of the Term Structure of Interest Rates, *Econometrica*, vol. 53, 1985, pp. 385-407.

[18] K.C. Chan, G.A. Karolyi, F.A. Longstaff and A.B. Sanders, An Empirical Comparison of Alternative Models of the Short-Term Interest Rate, *Journal of Finance*, vol. 47, 1992, pp. 1209-1227.

[19] C.S. Jones, Nonlinear Mean Reversion in the Short-Term Interest Rate. *Review of Financial Studies* vol. 16, 2003, pp. 793-843.

[20] A.S. Hurn and K.A. Lindsay, Estimating the Parameters of Stochastic Differential Equations, *Math Comput Simulat*, vol. 48, 1999, pp. 373-384.

[21] A.S. Hurn, K.A. Lindsay and V.L. Martin, On the Efficacy of Simulated Maximum Likelihood for Estimating the Parameters of Stochastic Differential Equations, *J. Time Series Anal* vol. 24, 2003, pp. 45-63.

[22] D.T. Gillespie, Approximate Accelerated Stochastic Simulation of Chemically Reacting Systems, *J. Chem. Phys.*, vol. 115, 2001, pp. 1716-1733.

[23] D.W. Scott, *Multivariate Density Estimate: Theory, Practiceand Vasualization*, Wiley, NY; 1992.

[24] S. Panda and N.P. Padhy, Comparison of Particle Swarm Optimization and Genetic Algorithm for TACTS-based Controller Design, *Appl. Soft Comput.*, vol. 8, 2008, pp. 1418-1427.

[25] Particle Swarm Toolbox, http://www.mathworks.com/matlabcentral /fileexchange/25986-another-particle-swarm-toolbox.

[26] T. Tian, Estimation of Kinetic Rates of MAP Kinase Activation from Experimental Data, *in Proceedings of IJCBS*, Shanghai, China, 2009, pp. 457-462, IEEE Press.