Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2012, Article ID 342705, 15 pages
doi:10.1155/2012/342705

Research Article

Asymptotic Parameter Estimation for a Class of Linear Stochastic Systems Using Kalman-Bucy Filtering

Xiu Kan,[1] Huisheng Shu,[2] and Yan Che[1]

[1] *School of Information Science and Technology, Donghua University, Shanghai 200051, China*
[2] *School of Science, Donghua University, Shanghai 200051, China*

Correspondence should be addressed to Huisheng Shu, hsshu@dhu.edu.cn

Received 21 June 2012; Accepted 21 July 2012

Academic Editor: Jun Hu

The asymptotic parameter estimation is investigated for a class of linear stochastic systems with unknown parameter $\theta : dX_t = (\theta\alpha(t) + \beta(t)X_t)dt + \sigma(t)dW_t$. Continuous-time Kalman-Bucy linear filtering theory is first used to estimate the unknown parameter θ based on Bayesian analysis. Then, some sufficient conditions on coefficients are given to analyze the asymptotic convergence of the estimator. Finally, the strong consistent property of the estimator is discussed by comparison theorem.

1. Introduction

Stochastic differential equations (SDEs) are a natural choice to model the time evolution of dynamic systems which are subject to random influences. Such models have been used with great success in a variety of application areas, including biology, mechanics, economics, geophysics, oceanography, and finance. For instance, refer to [1–8]. In reality, it is unavoidable that a stochastic system contains unknown parameters. Since 1962, Arato et al. [10] first applied parameter estimation to geophysical problem. Parameter estimation for SDEs has attracted the close attention of many researchers, and many parameter estimation methods for various advanced models have been studied, such as maximum likelihood estimation (MLE), Bayes estimation (BE), maximum probability estimation (MPE), minimum distance estimation (MDE), minimum contrast estimation (MCE), and M-estimation (ME). See [10–15] for details.

In practice, most stochastic systems cannot be observed completely, but the development of filtering theory provides an effective method to solve this problem. Over the past few decades, a lot of effective approaches have been proposed to overcome the difficulties in parameter estimation for stochastic models by filtering methods. It turns out to be helpful both in computability and asymptotic studies. See [9, 16–26]. In particular, the parameter estimation has been studied based on filtering observation, and the strong consistency property has also been shown in [27, 28]. In [29], a large deviation inequality has been obtained which implies the strong consistency, local asymptotic normality, and the convergence of moments. The asymptotic properties of estimators have been studied for a class of special Gaussian Itô processes with noisy observations in [30]. It should be pointed out that, so far, although the parameter estimation problem has been widely investigated for SDEs, the parameter estimation problem for stock price model has gained much less research attention due probably to the mathematical complexity.

Stock return volatility process is an important topic in options pricing theory. During the past decades, many SDEs have been modeled to solve the financial problems. For instance, refer to [2, 31–35]. Particularly, the so-called Hull-White model has been established by Hull and White [34] to analyze European call options prices under stochastic volatility at 1987. Using Taylor series expansion, an accurate formula for call options has been derived where stock returns and stock volatilities are uncorrelated. In addition, the Hull-White model readily lends itself to the estimation of underlying stochastic process parameters. Since the Hull-White formula is an effective options pricing model, it has been widely used to model the practice stock price problem. Therefore, it is reasonable to study the parameter estimation problem for Hull-White model with unknown parameter. Unfortunately, to the best of the authors' knowledge, the parameter estimation for Hull-White model with unknown parameter based on Kalman-Bucy linear filtering theory has not been fully studied despite its potential in practical application, and this situation motivates our present investigation.

Summarizing the above discussions, in this paper, we aim to investigate the parameter estimation problem for a general class of linear stochastic systems. The main contributions of this paper lie in the following aspects. (1) *Kalman-Bucy linear filtering is used to solve the parameter estimation problem.* (2) *The asymptotic convergence of the estimator is investigated by analyzing Riccati equation.* (3) *The strong consistent property is studied by comparison theorem*. The rest of this paper is organized as follows. In Section 2, we formulate the problem and state the well-known fact which would be used later. In Section 3, we study the asymptotic convergence of the estimator. In Section 4, the strong consistent of estimator is given. In Section 5, some conclusions are drawn.

Notation. The notation used here is fairly standard except where otherwise stated. $\mathbb{R} = (-\infty, +\infty)$ and $\mathbb{R}_+ = [0, +\infty)$. For a vector $x =\in \mathbb{R}$, $|x|$ is the Euclidean norm (or L^2 norm) with $|x| = \sqrt{x \cdot x}$. M^T and M^{-1} represent the transpose and inverse of the matrix M. $\det(M)$ denotes the determinant of the matrix M. I denotes the identity matrix of compatible dimension. Moreover, let $(\Omega, \mathcal{F}, \mathbf{P})$ be a complete probability space with a natural filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual conditions (i.e., it is right continuous, and $\mathcal{F}_0$ contains all $\mathbf{P}$-null sets). $\mathbb{E}[x]$ stands for the expectation of the stochastic variable x with respect to the given probability measure $\mathbf{P}$. $C(\mathbb{R}_+)$ denotes the class of all continuous time on $t \in \mathbb{R}_+$.

2. Problem Statement

Hull-White model is a continuous-time, real stochastic process as follows:

$$X_t = X_0 + \int_0^t (\alpha(s) + \beta(s)X_s)ds + \int_0^t \sigma(s)dW_s \tag{2.1}$$

with initial value X_0 as a Gaussian random variable, where α, β, σ are deterministic continuous functions on time t, W_t is a Brownian motion independent of the initial value X_0. Obviously, Hull-White model (2.1) is a general continuous-time linear SDE for X_t, and we assume that the coefficient α contains an unknown parameter $\theta \in R$ as follows:

$$dX_t = \left(\theta\alpha(t) + \beta(t)X_t\right)dt + \sigma(t)dW_t \quad t \geq 0, \tag{2.2}$$

and we observe the process X_t by the following filtering observations:

$$dY_t = \mu(t)X_t dt + \gamma(t)dV_t \quad t \geq 0, \tag{2.3}$$

where μ, γ are deterministic bounded continuous functions on time t, and V_t is a Brownian motion independent of W_t.

Now, our aim is to estimate θ in (2.2) based on the observation of (2.3). First, we can use Bayesian analysis to deal with the unknown parameter θ. We model θ as a random variable and denoted it as θ_0. We assume θ_0 normally distributed and independent of $\sigma(W_t, V_t, t \geq 0)$. Then, we can rewrite (2.2) as a two-component system for (X_t, θ_t) as follows:

$$\begin{pmatrix} dX_t \\ d\theta_t \end{pmatrix} = \begin{pmatrix} \beta(t) & \alpha(t) \\ 0 & 0 \end{pmatrix} \begin{pmatrix} X_t \\ \theta_t \end{pmatrix} dt + \begin{pmatrix} \sigma(t) \\ 0 \end{pmatrix} dW_t \quad t \geq 0. \tag{2.4}$$

Similarly, filtering observations system (2.3) can be expressed as follows:

$$dY_t = \begin{pmatrix} \mu(t) & 0 \end{pmatrix} \begin{pmatrix} X_t \\ \theta_t \end{pmatrix} dt + \gamma(t)dV_t \quad t \geq 0. \tag{2.5}$$

Therefore, we can use the Kalman-Bucy linear filtering theory to estimate θ_0 as follows:

$$\widehat{\theta}_t = \mathbb{E}[\theta_0 \mid Y_s, 0 \leq s \leq t], \tag{2.6}$$

and moreover, we also have $\widehat{X}_t = \mathbb{E}[X_t | Y_s, 0 \leq s \leq t]$.

For given Gaussian initial conditions X_0 and θ_0, it is well known from Kalman-Bucy linear filtering theory that error covariance matrix $S(t)$ satisfies the following Riccati equation:

$$\dot{S}(t) = AS + SA^T - SC^T\left(DD^T\right)^{-1}CS + BB^T, \tag{2.7}$$

where $A = \begin{pmatrix} \beta(t) & \alpha(t) \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} \sigma(t) \\ 0 \end{pmatrix}$, $C = (\mu(t)\ 0)$, $D = \gamma(t)$, and as we all know the error covariance matrix $S(t)$ is defined as follows:

$$S(t) = \begin{pmatrix} S_{xx}(t) & S_{x\theta}(t) \\ S_{\theta x}(t) & S_{\theta\theta}(t) \end{pmatrix} = \begin{pmatrix} \mathbb{E}\left[\left(X_t - \widehat{X}_t\right)^2\right] & \mathbb{E}\left[\left(X_t - \widehat{X}_t\right)\left(\theta_0 - \widehat{\theta}_t\right)\right] \\ \mathbb{E}\left[\left(X_t - \widehat{X}_t\right)\left(\theta_0 - \widehat{\theta}_t\right)\right] & \mathbb{E}\left[\left(\theta_0 - \widehat{\theta}_t\right)^2\right] \end{pmatrix}. \tag{2.8}$$

Set $a = S_{xx}$, $b = S_{x\theta} = S_{\theta x}$, and $c = S_{\theta\theta}$. From Riccati equation (2.7), one can get the following system:

$$\begin{aligned}\dot{a} &= 2\beta a + 2\alpha b + \sigma^2 - \frac{\mu^2}{\gamma^2} a^2,\\ \dot{b} &= \beta b + \alpha c - \frac{\mu^2}{\gamma^2} ab,\\ \dot{c} &= -\frac{\mu^2}{\gamma^2} b^2.\end{aligned} \tag{2.9}$$

Remark 2.1. Equation (2.9) is a nontrivial nonlinear ordinary differential equation system, and it is well known from the Kalman-Bucy linear filtering theory that such Riccati equations have unique solutions for all $t \in \mathbb{R}_+$.

Remark 2.2. From the equation $\dot{c} = -(\mu^2/\gamma^2)b^2$, we can see that the error variance $\mathbb{E}[(\theta_0 - \widehat{\theta}_t)^2]$ is monotonically decreasing.

3. Asymptotic Convergence Analysis

Assume that the initial conditions X_0 and θ_0 are independent and have nonvariances, so that $b(0) = 0$ and $a(0) = \mathbb{E}[X_0^2] > 0$, $c(0) = \mathbb{E}[\theta_0^2] > 0$; thus, $S(0)$ is a regular matrix. For the property of continuity of $S(t)$, $S^{-1}(t)$ exists at least for small times. In order to obtain the rate of convergence of the estimator, $S(t)$ should satisfy the regularity conditions. The following Theorem certifies the regularity of $S(t)$.

Theorem 3.1. *(a1) Assume the initial conditions X_0 and θ_0 for system* (2.2) *are independent and have nonvanishing variances.*

(a2) Let $\alpha(t), \beta(t), \sigma(t), \mu(t), \gamma(t) \in C(\mathbb{R}_+)$.

Then, the error covariance matrix $S(t)$ satisfies $\det(S(t)) > 0$ *for all $t \geq 0$, and*

$$S_{xx}(t) > 0, \quad S_{\theta\theta}(t) > 0 \quad \forall t \geq 0. \tag{3.1}$$

Proof. By Kalman-Bucy linear filtering theory, we know that $\det(S(t)) > 0$ for all $t \geq 0$. Furthermore, it is not difficult to show that (3.1) holds for all $t \geq 0$.

Since $\det(S(t)) > 0$, it follows that $S^{-1}(t)$ exists. Set

$$R(t) = S^{-1}(t) = \begin{pmatrix} e(t) & f(t) \\ f(t) & g(t) \end{pmatrix}. \tag{3.2}$$

As we know that $R = 1/S$ implies that $\dot{R} = -(1/S^2)\dot{S}$, one can easily have that

$$\dot{R} = -R\dot{S}R. \tag{3.3}$$

It follows readily form (2.9) and (3.3) that

$$\dot{R} = -RA - A^T R + C^T \left(DD^T\right)^{-1} C - RBB^T R. \tag{3.4}$$

Using a similar computation as (2.9), we can get

$$\begin{aligned} \dot{e} &= \frac{\mu^2}{\gamma^2} - 2\beta e - \sigma^2 e^2, \\ \dot{f} &= -\alpha e - \beta f - \sigma^2 e f, \\ \dot{g} &= -2\alpha f - \sigma^2 f^2. \end{aligned} \tag{3.5}$$

The condition (a1) shows that $a(0) > 0,\ b(0) = 0$, and $c(0) > 0$, which implies that $e(0) > 0$, $f(0) = 0$, and $g(0) > 0$. Since the Riccati equations (2.9) have unique solutions on R_+, thus the nonlinear system (3.5) has a unique solution on $\mathbb{R}_+$. Furthermore, the first equation $\dot{e} = \mu^2/\gamma^2 - 2\beta e - \sigma^2 e^2$ with initial condition $e(0) > 0$ has a unique solution on a maximal time interval $[0, T)$, where $T \in \mathbb{R}_+$. Assume that there exists a smallest time $\bar{t} \in (0, T)$ such that $e(\bar{t}) = 0$. By the property of continuity of $e(t)$, we have $e(t) > 0$, for $0 \le t < \bar{t}$. Thus,

$$\dot{e}(t) = \lim_{\Delta t \to 0} \frac{e\left(\bar{t}\right) - e\left(\bar{t} - \Delta t\right)}{\Delta t} < 0, \tag{3.6}$$

this contradicts with $\dot{e}(t) = \mu^2(\bar{t})/\gamma^2(\bar{t}) - 2\beta(\bar{t})e(\bar{t}) - \sigma^2(\bar{t})e^2(\bar{t}) \le \mu^2(\bar{t})/\gamma^2(\bar{t})$ for all $t \in [0, T)$. Therefore, $e(t) > 0$, for $t \in [0, T)$.

As long as $\dot{e}(t) = \mu^2(\bar{t})/\gamma^2(\bar{t}) - 2\beta(\bar{t})e(\bar{t}) - \sigma^2(\bar{t})e^2(\bar{t}) \le \mu^2(\bar{t})/\gamma^2(\bar{t})$ for all $t \in [0, T)$ and $\mu(t), \gamma(t)$ are bounded, we have $\dot{e}(t) \le C$, where C is a constant. So that $e(t)$ is bounded from below by 0 and from above by $e(0) + t$, which implies that $e(t)$ cannot explode in finite time, thus $T = +\infty$. This shows that system (3.5) has a unique solution on $\mathbb{R}_+$ because the second equation is a linear equation for f which can be solved analytically on $\mathbb{R}_+$, and g can get by integration.

Define $h(t) := \det(R(t)) = e(t)g(t) - f^2(t)$. Since $\det(S(t)) > 0$ for all $t \ge 0$, thus $h(t) = \det(R(t)) = 1/\det(S(t)) > 0$ for all $t \ge 0$, moreover, $S_{\theta\theta} > 0$ for all $t \ge 0$. Finally, we assume that there exists t_0 such that, $S_{xx}(t_0) = 0$, then $g(t_0) = S_{xx}(t_0)h(t_0) = 0$, so that $h(t_0) = e(t_0)g(t_0) - f^2(t_0) \le 0$, and this contradicts $h(t_0) > 0$. Hence, $S_{xx} > 0$ for all $t \ge 0$.

The proof is complete. □

In order to obtain the convergence rate, the Riccati equation must be solved, and we just need the solution of (3.5). Now, we solve the equation $\dot{e} = \mu^2/\gamma^2 - 2\beta e - \sigma^2 e^2$ when $\beta, \sigma, \mu, \gamma$ are equal to constants.

In the case $e(0) \ne l_2$, we get

$$e(t) = \frac{l_1 + l_2 L \exp\left[(l_1 + l_2)\sigma^2 t\right]}{L \exp\left[(l_1 + l_2)\sigma^2 t\right] - 1}, \tag{3.7}$$

where $L = (e(0) + l_1)/(e(0) - l_2)$, $l_1 = (2\beta/\sigma^2 + \sqrt{4\beta^2/\sigma^4 + 4\mu^2/\sigma^2\gamma^2})/2$, $l_2 = (-(2\beta/\sigma^2) + \sqrt{4\beta^2/\sigma^4 + 4\mu^2/\sigma^2\gamma^2})/2$.

In the other case $e(0) = l_2$, the solution shows that $e(t) = l_2$ for all $t \geq 0$.

Thus, for each $\alpha > 0$, $\beta > 0$, $\sigma > 0$, $\mu > 0$, $\gamma > 0$, the solution $e(t)$ obviously satisfies

$$e(t) \longrightarrow l_2 \quad \text{as}\, t \longrightarrow +\infty. \tag{3.8}$$

The convergence rate of the estimator is given by following theorem.

Theorem 3.2. *Assume that $\alpha, \beta, \sigma, \mu, \gamma \in C(\mathbb{R}_+)$, are all bounded, and there are constants α_1, α_2, β_1, β_2, σ_1, σ_2, μ_1, μ_2, γ_1, γ_2, and t_0, such that*

(b1) *: $0 < \alpha_1 \leq |\alpha(t)| \leq \alpha_2$ for all $t \geq t_0$;*

(b2) *: $0 < \beta_1 \leq |\beta(t)| \leq \beta_2$ for all $t \geq t_0$;*

(b3) *: $0 < \sigma_1 \leq |\sigma(t)| \leq \sigma_2$ for all $t \geq t_0$;*

(b4) *: $0 < \mu_2 \leq |\mu(t)| \leq \mu_1$ for all $t \geq t_0$;*

(b5) *: $0 < \gamma_1 \leq |\gamma(t)| \leq \gamma_2$ for all $t \geq t_0$;*

(b6) *: $2\alpha_1(\beta_1 + \sigma_1^2 l_{22}) > \sigma_2^2 l_{21}$ where $l_{2i} = (-2\beta_i/\sigma_i^2 + \sqrt{(4\beta_i^2)/(\sigma_i^4) + (4\mu_i^2)/(\sigma_i^2\ \gamma_i^2)})/2,\ i = 1, 2$.*

Then, for arbitrary $\epsilon > 0$ and $T > 0$, we have

$$P\left(\left|\theta_0 - \widehat{\theta}_t\right| > \epsilon\right) \leq \frac{1}{\epsilon^2} C T^{-1}, \tag{3.9}$$

where C is a positive constant independent of ϵ and T.

Proof. Let e_i be the solution to $\dot{e}_i = \mu_i^2/\gamma_i^2 - 2\beta_i e_i - \sigma_i^2 e_i^2,\ i = 1, 2$, and $e_i(t_0) = e(t_0)$.

Since $\mu_2^2/\gamma_2^2 - 2\beta_2 e - \sigma_2^2 e^2 \leq \dot{e} = \mu^2/\gamma^2 - 2\beta e - \sigma^2 e^2 \leq \mu_1^2/\gamma_1^2 - 2\beta_1 e - \sigma_1^2 e^2$ for all $t \geq t_0$, by the comparison theorem [2, 36], we obtain that

$$e_2(t) \leq e(t) \leq e_1(t) \quad \forall t \geq t_0. \tag{3.10}$$

It follows from (3.7) that e is bounded, and for any given $\delta \in (0, 1)$, there is a $t_1 \geq t_0$ such that

$$0 < l_{22}(1 - \delta) \leq e(r) \leq l_{21}(1 + \delta) \quad \forall r \geq t_1. \tag{3.11}$$

For $t \geq t_1$, we can obtain from (3.5) and $f(0) = 0$ that

$$\begin{aligned} f(t) = & -\int_0^t \exp\left[-\int_s^t \left(\beta(r) + \sigma^2(r)e(r)\right)dr\right]\alpha(s)e(s)ds \\ = & -\exp\left[-\int_0^t \left(\beta(r) + \sigma^2(r)e(r)\right)dr\right]\int_0^{t_1} \exp\left[\int_0^s \left(\beta(r) + \sigma^2(r)e(r)\right)dr\right]\alpha(s)e(s)ds \\ & -\int_{t_1}^t \exp\left[-\int_s^t \left(\beta(r) + \sigma^2(r)e(r)\right)dr\right]\alpha(s)e(s)ds. \end{aligned} \tag{3.12}$$

As $\beta(r) + \sigma^2(r)e(r) \geq \beta_1 + \sigma_1^2 l_{22}(1-\delta)$ holds for all $t \geq t_1$, thus, the first term in (3.12) goes to 0 as $t \to \infty$. For the second term in (3.12), we have

$$\begin{aligned} & \left|\int_{t_1}^t \exp\left[-\int_s^t \left(\beta(r) + \sigma^2(r)e(r)\right)dr\right]\alpha(s)e(s)ds\right| \\ & \quad \leq \int_0^t \exp\left[-\left(\beta_1 + \sigma_1^2 l_{22}(1-\delta)\right)(t-s)\right]\, l_{21}(1+\delta)ds \\ & \quad = \frac{l_{21}(1+\delta)}{\beta_1 + \sigma_1^2 l_{22}(1-\delta)} \int_0^t \exp\left[-\left(\beta_1 + \sigma_1^2 l_{22}(1-\delta)\right)(t-s)\right] d\left(\beta_1 + \sigma_1^2 l_{22}(1-\delta)\right)s \\ & \quad = \frac{l_{21}(1+\delta)}{\beta_1 + \sigma_1^2 l_{22}(1-\delta)}\left(1 - \exp\left[-\left(\beta_1 + \sigma_1^2 l_{22}(1-\delta)\right)t\right]\right) \\ & \quad \leq \frac{l_{21}(1+\delta)}{\beta_1 + \sigma_1^2 l_{22}(1-\delta)}. \end{aligned} \tag{3.13}$$

By similar arguments, we obtain that

$$\left|\int_{t_1}^t \exp\left[-\int_s^t \left(\beta(r) + \sigma^2(r)e(r)\right)dr\right]\alpha(s)e(s)ds\right| \geq \frac{l_{22}(1-\delta)}{\beta_2 + \sigma_2^2 l_{21}(1+\delta)}. \tag{3.14}$$

Therefore, for any $\xi > 0$, there exists $t(\xi) > 0$ such that

$$\frac{l_{22}(1-\delta)}{\beta_2 + \sigma_2^2 l_{21}(1+\delta)} \leq |f(t)| \leq \frac{l_{21}(1+\delta)}{\beta_1 + \sigma_1^2 l_{22}(1-\delta)} \quad \forall t \geq t(\xi). \tag{3.15}$$

For all $t \geq t(\xi)$, we can get from (3.5) that

$$\begin{aligned}
\dot{g} &= \left(2|\alpha| - \sigma^2 |f|\right)|f| \\
&\geq \left(2\alpha_1 - \sigma_2^2 \frac{l_{21}(1+\delta)}{\beta_1 + \sigma_1^2 l_{22}(1-\delta)}\right) \frac{l_{22}(1-\delta)}{\beta_2 + \sigma_2^2 l_{21}(1+\delta)} \\
&= \left(\frac{2\alpha_1(\beta_1 + \sigma_1^2 l_{22}) - \sigma_2^2(l_{21}(1+\delta))}{\beta_1 + \sigma_1^2 l_{22}(1-\delta)}\right) \frac{l_{22}(1-\delta)}{\beta_2 + \sigma_2^2 l_{21}(1+\delta)}.
\end{aligned} \tag{3.16}$$

By assumption (b6), we get $\dot{g} > 0$ for a sufficiently small $\xi > 0$. This implies that $g(t)$ goes to infinity at least as a linear function. Thus, there exists a constant $C > 0$, such that

$$\mathbb{E}\left(\theta_0 - \widehat{\theta}_t\right)^2 = S_{\theta\theta} = \frac{e}{h} \leq Ct^{-1}. \tag{3.17}$$

Hence, for arbitrary $\epsilon > 0$ and all $T > 0$, it follows from Chebyshev's inequality that

$$P\left(\left|\theta_0 - \widehat{\theta}_t\right| > \epsilon\right) \leq \frac{1}{\epsilon^2} C T^{-1}. \tag{3.18}$$

The proof is complete. □

Remark 3.3. From the proof of Theorem 3.2, we can see that $\theta_0 - \widehat{\theta}_t$ goes to 0 in L^2-sense under the given conditions. In other words, $\widehat{\theta}_t$ is asymptotically unbiased.

Remark 3.4. It is well known that Kalman-Bucy linear filtering theory remains valid if one replaces the Brownian motion (W_t, V_t) in systems (2.2) and (2.3) by an arbitrary centered orthogonal increment process of the same covariance structure. Thus, Theorem 3.2 remains valid under this replacement.

4. Strong Consistency

In last section, we give the conditions for the convergence rate of the estimator. Furthermore, we use the comparison theorem to proof the strong consistency in this section. As we all know, if the parameter θ is, a genuine Gaussian random variable, then we can have a clear statistical interpretation for the convergence rate. Firstly, we pick θ_0 at random; secondly, let system (2.2) run up to time t and simultaneously observe Y by system (2.3); finally, compute $\widehat{\theta}_t$ as the following form.

The Kalman-Bucy linear filtering theory shows us

$$\begin{aligned}\begin{pmatrix} dX_t \\ d\theta_t \end{pmatrix} &= \left(A(t) - \frac{C^T(t)C(t)}{D^2(t)} S(t) \right) \begin{pmatrix} X_t \\ \theta_t \end{pmatrix} dt + \frac{C(t)}{D^2(t)} S(t) dY_t \\ &= \begin{pmatrix} \beta(t) - \dfrac{\mu^2(t)}{\gamma^2(t)} S_{xx}(t) & \alpha(t) \\ -\dfrac{\mu^2(t)}{\gamma^2(t)} S_{\theta x}(t) & 0 \end{pmatrix} \begin{pmatrix} X_t \\ \theta_t \end{pmatrix} dt + \frac{\mu^2(t)}{\gamma^2(t)} \begin{pmatrix} S_{xx}(t) \\ S_{\theta x}(t) \end{pmatrix} dY_t\end{aligned} \tag{4.1}$$

with initial conditions $\widehat{X}_0 = \mathbb{E}[X_0]$ and $\widehat{\theta}_0 = \mathbb{E}[\theta_0]$. If we denote that $\Phi(t)$ is the matrix fundamental solution of the deterministic linear system

$$\begin{pmatrix} \dot{x}_t \\ \dot{y}_t \end{pmatrix} = \begin{pmatrix} \beta(t) - \dfrac{\mu^2(t)}{\gamma^2(t)} S_{xx}(t) & \alpha(t) \\ -\dfrac{\mu^2(t)}{\gamma^2(t)} S_{\theta x}(t) & 0 \end{pmatrix} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}, \tag{4.2}$$

then the solution to (4.1) is given by

$$\begin{pmatrix} \widehat{X}_t \\ \widehat{\theta}_t \end{pmatrix} = \Phi(t)\Phi^{-1}(0) \begin{pmatrix} \mathbb{E}[X_0] \\ \mathbb{E}[\theta_0] \end{pmatrix} + \int_0^t \Phi(t)\Phi^{-1}(s) \begin{pmatrix} S_{xx}(t) \\ S_{\theta x}(t) \end{pmatrix} dY_s. \tag{4.3}$$

And for every particular experiment ω, the quantity $(\theta_0(\omega) - \widehat{\theta}_t(\omega))^2$ would be the squared estimation error.

But in this paper θ is a fixed parameter, so we can only choose $\theta_0(\omega) = \theta$, and then the statistical mean over different values of $\theta_0(\omega)$ has no experimental meaning. The true estimation error is given by $\theta - \widehat{\theta}_t$, not $\theta_0 - \widehat{\theta}_t$. It is therefore desirable that estimator $\widehat{\theta}_t$ converges to θ_0 for "all fixed values $v = \theta_0$" a.s. To establish such an assertion we work with a product space $(R \times \Omega, \mathcal{B}(R) \otimes \mathcal{F}, \eta \otimes P)$, where η denotes the law of θ_0, and $(\Omega, \mathcal{F}, P)$ is the underlying probability space for Brownian motion $(W_t, V_t)_{t \geq 0}$. This space is most appropriate because one can make P a.s. statements for fixed $v \in \mathbb{R}$. Notice that in this representation we have $\theta_0(v, \omega) = v$ for all $(v, \omega) \in \mathbb{R} \times \Omega$. Assuming this underlying probability space, we use the comparison theorem to get the following consistency result.

In the proof of Theorem 3.2, we know that e, f is bonded and g is monotonically increasing, moreover, $S_{xx}(t) = a = g/h = g/(eg - f^2) = (g - f^2/e + f^2/e)/(eg - f^2) = 1/e + f^2/e(eg - f^2)$ and $S_{\theta x}(t) = b = f/h = f/(eg - f^2)$. Thus, there exist positive constants a_1, a_2, b_1, and b_2 such that $a_1 \leq a \leq a_2$ and $b_1 \leq b \leq b_2$.

Theorem 4.1. *Assume that the following two conditions are satisfied:*

(c1) : *$\widehat{\theta}_t$ converges to θ_0 in $L^2(\eta \otimes P)$;*

(c2) : *$\beta_2 - \mu_2^2/\gamma_2^2 < 0$;*

(c3) : *$(\beta_2 - (\mu_2^2/\gamma_2^2) a_2)^2 - 4\alpha_2 (\mu_2^2/\gamma_2^2) b_2 < 0$.*

Then, for all fixed $\upsilon \in \mathbb{R}$, we have

$$\widehat{\theta}_t(\upsilon, \cdot) \longrightarrow \upsilon, \quad P\text{-}a.s., \quad as\, t \longrightarrow \infty. \tag{4.4}$$

Proof. We will show that (4.4) holds for all $\upsilon \in N^c$, where $\eta(N) = 0$.

By Kalman-Bucy linear filtering theory, we know

$$\begin{aligned}\begin{pmatrix} dX_t \\ d\theta_t \end{pmatrix} &= \left(A(t) - \frac{C^T(t)C(t)}{D^2(t)} S(t) \right) \begin{pmatrix} X_t \\ \theta_t \end{pmatrix} dt + \frac{C(t)}{D^2(t)} S(t) dY_t \\ &= \begin{pmatrix} \beta(t) - \dfrac{\mu^2(t)}{\gamma^2(t)} S_{xx}(t) & \alpha(t) \\ -\dfrac{\mu^2(t)}{\gamma^2(t)} S_{\theta x}(t) & 0 \end{pmatrix} \begin{pmatrix} X_t \\ \theta_t \end{pmatrix} dt + \frac{\mu^2(t)}{\gamma^2(t)} \begin{pmatrix} S_{xx}(t) \\ S_{\theta x}(t) \end{pmatrix} dY_t\end{aligned} \tag{4.5}$$

with initial conditions $\widehat{X}_0 = \mathbb{E}[X_0]$ and $\widehat{\theta}_0 = \mathbb{E}[\theta_0] = \mathbb{E}[\upsilon] = \upsilon$.

Since the following linear equations:

$$\begin{pmatrix} \dot{x}_t \\ \dot{y}_t \end{pmatrix} = \begin{pmatrix} \beta(t) - \dfrac{\mu^2(t)}{\gamma^2(t)} S_{xx}(t) & \alpha(t) \\ -\dfrac{\mu^2(t)}{\gamma^2(t)} S_{\theta x}(t) & 0 \end{pmatrix} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} \tag{4.6}$$

equal to

$$\begin{aligned}\dot{x}_t &= \left[\beta(t) - \frac{\mu^2(t)}{\gamma^2(t)} S_{xx}(t) \right] x(t) + \alpha(t) Y(t), \\ \dot{y}_t &= -\frac{\mu^2(t)}{\gamma^2(t)} S_{\theta x}(t) x(t),\end{aligned} \tag{4.7}$$

it follows from (c1)–(c3) that

$$\begin{gathered}\beta_1 - \frac{\mu_1^2}{\gamma_1^2} a_1 \le \beta(t) - \frac{\mu^2(t)}{\gamma^2(t)} S_{xx}(t) \le \beta_2 - \frac{\mu_2^2}{\gamma_2^2} a_2 < 0, \\ \alpha_1 \le \alpha(t) \le \alpha_2, \\ -\frac{\mu_1^2}{\gamma_1^2} b_1 \le -\frac{\mu^2(t)}{\gamma^2(t)} S_{\theta x}(t) \le -\frac{\mu_2^2}{\gamma_2^2} b_2.\end{gathered} \tag{4.8}$$

For linear equations:

$$\begin{pmatrix} \dot{x}_t \\ \dot{y}_t \end{pmatrix} = \begin{pmatrix} \beta_1 - \dfrac{\mu_1^2}{\gamma_1^2} a_1 & \alpha_1 \\ -\dfrac{\mu_1^2}{\gamma_1^2} b_1 & 0 \end{pmatrix} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix},$$
$$\begin{pmatrix} \dot{x}_t \\ \dot{y}_t \end{pmatrix} = \begin{pmatrix} \beta_2 - \dfrac{\mu_2^2}{\gamma_2^2} a_2 & \alpha_2 \\ -\dfrac{\mu_2^2}{\gamma_2^2} b_2 & 0 \end{pmatrix} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}, \tag{4.9}$$

if we set $\Phi_1(t)$ and $\Phi_2(t)$ that are the matrix fundamental solution of (4.9), we can obtain from the comparison theorem that

$$\Phi_1(t) \le \Phi(t) \le \Phi_2(t). \tag{4.10}$$

It is not difficult to explore (4.9), and get

$$\Phi_1(t) = \begin{pmatrix} -\dfrac{\lambda_1'}{N_{21}} e^{\lambda_1' t} & -\dfrac{\lambda_2'}{N_{21}} e^{\lambda_2' t} \\ e^{\lambda_1' t} & e^{\lambda_2' t} \end{pmatrix}, \qquad \Phi_2(t) = \begin{pmatrix} -\dfrac{\lambda_1}{M_{21}} e^{\lambda_1 t} & -\dfrac{\lambda_2}{M_{21}} e^{\lambda_2 t} \\ e^{\lambda_1 t} & e^{\lambda_2 t} \end{pmatrix},$$
$$\Phi_1^{-1}(t) = \begin{pmatrix} -\dfrac{N_{21}}{\lambda_1' - \lambda_2} e^{-\lambda_1' t} & -\dfrac{\lambda_2'}{\lambda_1' - \lambda_2'} e^{-\lambda_1' t} \\ \dfrac{N_{21}}{\lambda_1' - \lambda_2'} e^{-\lambda_2' t} & \dfrac{\lambda_1'}{\lambda_1' - \lambda_2'} e^{-\lambda_2' t} \end{pmatrix}, \qquad \Phi_2^{-1}(t) \begin{pmatrix} -\dfrac{M_{21}}{\lambda_1 - \lambda_2} e^{-\lambda_1 t} & -\dfrac{\lambda_2}{\lambda_1 - \lambda_2} e^{-\lambda_1 t} \\ \dfrac{M_{21}}{\lambda_1 - \lambda_2} e^{-\lambda_2 t} & \dfrac{\lambda_1}{\lambda_1 - \lambda_2} e^{-\lambda_2 t} \end{pmatrix}, \tag{4.11}$$

where $N_{11} = \beta_1 - (\mu_1^2/\gamma_1^2)a_1$, $N_{12} = \alpha_1$, $N_{21} = (\mu_1^2/\gamma_1^2)b_1$, $\lambda_1' = (N_{11} + \sqrt{N_{11}^2 - 4N_{12}N_{21}})/2$, $\lambda_2' = (N_{11} - \sqrt{N_{11}^2 - 4N_{12}N_{21}})/2$, $M_{11} = \beta_2 - (\mu_2^2/\gamma_2^2)a_2$, $M_{12} = \alpha_2$, $M_{21} = (\mu_2^2/\gamma_2^2)b_2$, $\lambda_1 = (M_{11} + \sqrt{M_{11}^2 - 4M_{12}M_{21}})/2$, $\lambda_2 = (M_{11} - \sqrt{M_{11}^2 - 4M_{12}M_{21}})/2$.

By assumption (c2) and (c3), we know that $\lambda_1' < 0$, $\lambda_2' < 0$, $\lambda_1 < 0$, and $\lambda_2 < 0$.

By the ODE theory [37, 38] and above discussion, we know that the solution of (4.1) is given by

$$\begin{pmatrix} \widehat{X}_t \\ \widehat{\theta}_t \end{pmatrix} = \Phi(t)\Phi^{-1}(0) \begin{pmatrix} \mathbb{E}[X_0] \\ \mathbb{E}[\theta_0] \end{pmatrix} + \int_0^t \Phi(t)\Phi^{-1}(s) \begin{pmatrix} S_{xx}(t) \\ S_{\theta x}(t) \end{pmatrix} dY_s. \tag{4.12}$$

Using the similar method, we can also obtain the solutions for the following two equations:

$$\begin{pmatrix} d\widehat{X}_t \\ d\widehat{\theta}_t \end{pmatrix} = \begin{pmatrix} \beta_1 - \dfrac{\mu_1^2}{\gamma_1^2} a_1 & \alpha_1 \\ -\dfrac{\mu_1^2}{\gamma_1^2} b_1 & 0 \end{pmatrix} \begin{pmatrix} \widehat{X}_t \\ \widehat{\theta}_t \end{pmatrix} dt + \frac{\mu_1}{\gamma_1} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} dY_t, \tag{4.13}$$

$$\begin{pmatrix} d\widehat{X}_t \\ d\widehat{\theta}_t \end{pmatrix} = \begin{pmatrix} \beta_2 - \dfrac{\mu_2^2}{\gamma_2^2} a_2 & \alpha_2 \\ -\dfrac{\mu_2^2}{\gamma_2^2} b_2 & 0 \end{pmatrix} \begin{pmatrix} \widehat{X}_t \\ \widehat{\theta}_t \end{pmatrix} dt + \frac{\mu_2}{\gamma_2} \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} dY_t, \tag{4.14}$$

where $\widehat{X}_0 = \mathbb{E}[X_0]$ and $\widehat{\theta}_0 = \mathbb{E}[\theta_0] = \mathbb{E}[\upsilon] = \upsilon$.

The solutions of the two equations are explored as the following form:

$$\begin{aligned} \begin{pmatrix} \widehat{X}_t \\ \widehat{\theta}_t \end{pmatrix} &= \Phi_1(t)\Phi_1^{-1}(0) \begin{pmatrix} \mathbb{E}[X_0] \\ \mathbb{E}[\theta_0] \end{pmatrix} + \int_0^t \Phi_1(t)\Phi_1^{-1}(s) \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} dY_s, \\ \begin{pmatrix} \widehat{X}_t \\ \widehat{\theta}_t \end{pmatrix} &= \Phi_2(t)\Phi_2^{-1}(0) \begin{pmatrix} \mathbb{E}[X_0] \\ \mathbb{E}[\theta_0] \end{pmatrix} + \int_0^t \Phi_2(t)\Phi_2^{-1}(s) \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} dY_s. \end{aligned} \tag{4.15}$$

For (4.14), we have that

$$\begin{pmatrix} \widehat{X}_t \\ \widehat{\theta}_t \end{pmatrix} = \Phi_2(t)\Phi_2^{-1}(0) \begin{pmatrix} \mathbb{E}[X_0] \\ \mathbb{E}[\theta_0] \end{pmatrix} + \int_0^t \Phi_2(t)\Phi_2^{-1}(s) \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} dY_s \tag{4.16}$$

yields that

$$\begin{aligned} \widehat{\theta}_t = {} & \int_0^t \left[a_2 \left(\frac{M_{21}}{\lambda_1 - \lambda_2} e^{-\lambda_2(t-s)} - \frac{M_{21}}{\lambda_1 - \lambda_2} e^{-\lambda_2(t-s)} \right) + b_2 \left(\frac{\lambda_1}{\lambda_1 - \lambda_2} e^{-\lambda_2(t-s)} - \frac{\lambda_2}{\lambda_1 - \lambda_2} e^{-\lambda_2(t-s)} \right) \right] dY_s \\ & + \left(\frac{M_{21}}{\lambda_1 - \lambda_2} e^{-\lambda_2 t} - \frac{M_{21}}{\lambda_1 - \lambda_2} e^{-\lambda_2 t} \right) X_0 + \left(\frac{\lambda_1}{\lambda_1 - \lambda_2} e^{-\lambda_2 t} - \frac{\lambda_2}{\lambda_1 - \lambda_2} e^{-\lambda_2 t} \right) \theta_0. \end{aligned} \tag{4.17}$$

Since $\lambda_1 < 0$ and $\lambda_2 < 0$, it is easy to get

$$\widehat{\theta}_t(\upsilon, \cdot) \longrightarrow \upsilon, P\text{-a.s.}, \quad \text{as } t \longrightarrow \infty. \tag{4.18}$$

For (4.13), we can also get

$$\widehat{\theta}_t(\upsilon, \cdot) \longrightarrow \upsilon, P\text{-a.s.}, \quad \text{as } t \longrightarrow \infty. \tag{4.19}$$

Hence, for (4.1), we can get the following result:

$$\widehat{\theta}_t(\upsilon,\cdot) \longrightarrow \upsilon, P\text{-a.s.}, \quad \text{as } t \longrightarrow \infty. \tag{4.20}$$

The proof is complete. □

Remark 4.2. Under the probability space used in this paper, we can see that Theorem 3.2 is the particular form of Theorem 4.1 if we use Chebyshev's inequality on the result of Theorem 4.1.

Remark 4.3. The strong consistency in Deck [30] requires that $\widehat{\theta}_t$ is a martingale, while, in our result, $\widehat{\theta}_t$ can be not a martingale. Furthermore, when $\widehat{\theta}_t$ is a martingale, our result is more strong than Deck's, so in that case we can relax the conditions as Deck.

5. Conclusions

In this paper, we have investigated the parameter estimation problem for a class of linear stochastic systems called Hull-White stochastic differential equations which are important models in finance. Firstly, Bayesian viewpoint is first chosen to analyze the parameter estimation problem based on Kalman-Bucy linear filtering theory. Secondly, some sufficient conditions on coefficients are given to study the asymptotic convergence problem. Finally, the strong consistent property of estimator is discussed by Kalman-Bucy linear filtering theory and comparison theorem.

Acknowledgments

This work was supported by the National Nature Science Foundation of China under Grant no. 60974030 and the Science and Technology Project of Education Department in Fujian Province JA11211.

References

[1] B. M. Bibby and M. Sørensen, "Martingale estimation functions for discretely observed diffusion processes," *Bernoulli*, vol. 1, no. 1-2, pp. 17–39, 1995.

[2] C. Corrado and T. Su, "An empirical test of the Hull-White option pricing model," *Journal of Futures Markets*, vol. 18, no. 4, pp. 363–378, 1998.

[3] H. Dong, Z. Wang, and H. Gao, "H_∞ fuzzy control for systems with repeated scalar nonlinearities and random packet losses," *IEEE Transactions on Fuzzy Systems*, vol. 17, no. 2, pp. 440–450, 2009.

[4] J. Hu, Z. Wang, and H. Gao, "A delay fractioning approach to robust sliding mode control for discrete-time stochastic systems with randomly occurring non-linearities," *IMA Journal of Mathematical Control and Information*, vol. 28, no. 3, pp. 345–363, 2011.

[5] J. Hu, Z. Wang, H. Gao, and L. K. Stergioulas, "Robust sliding mode control for discrete stochastic systems with mixed time delays, randomly occurring uncertainties, and randomly occurring non-linearities," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 7, pp. 3008–3015, 2012.

[6] J. Hu, Z. Wang, and H. Gao, "Robust H_∞ sliding mode control for discrete time-delay systems with stochastic nonlinearities," *Journal of the Franklin Institute*, vol. 349, no. 4, pp. 1459–1479, 2012.

[7] E. Ince, *Ordinary Differential Equations*, Dover, New York, NY, USA, 1956.

[8] C. R. Rao, *Linear Statistical Inference and its Applications*, John Wiley & Sons, New York, NY, USA, 1973.

[9] M. H. A. Davis, *Linear Estimation and Stochastic Control*, Chapman and Hall, London, UK, 1977.

[10] M. Arato, A. Kolmogrorov, and Y. Sinai, *On Parameter Estimation of a Complex Stationary Gaussian Process*, vol. 146, Doklady Academy, USSR, 1962.

[11] H. Frydman and P. Lakner, “Maximum likelihood estimation of hidden Markov processes,” *The Annals of Applied Probability*, vol. 13, no. 4, pp. 1296–1312, 2003.

[12] L. Galtchouk and V. Konev, “On sequential estimation of parameters in semimartingale regression models with continuous time parameter,” *The Annals of Statistics*, vol. 29, no. 5, pp. 1508–1536, 2001.

[13] N. R. Kristensen, H. Madsen, and S. B. Jørgensen, “Parameter estimation in stochastic grey-box models,” *Automatica*, vol. 40, no. 2, pp. 225–237, 2004.

[14] B. L. S. Prakasa Rao, *Statistical Inference for Diffusion Type Processes*, vol. 8, Edward Arnold, London, UK, 1999.

[15] I. Shoji and T. Ozaki, “Comparative study of estimation methods for continuous time stochastic processes,” *Journal of Time Series Analysis*, vol. 18, no. 5, pp. 485–506, 1997.

[16] H. Dong, Z. Wang, and H. Gao, “Robust H_∞ filtering for a class of nonlinear networked systems with multiple stochastic communication delays and packet dropouts,” *IEEE Transactions on Signal Processing*, vol. 58, no. 4, pp. 1957–1966, 2010.

[17] H. Dong, Z. Wang, and H. Gao, “Distributed filtering for a class of time-varying systems over sensor networks with quantization errors and successive packet dropouts,” *IEEE Transactions on Signal Processing*, vol. 60, no. 6, pp. 3164–3173, 2012.

[18] J. Hu, Z. Wang, H. Gao, and L. K. Stergioulas, “Probability-guaranteed H finite-horizon filtering for a class of nonlinear time-varying systems with sensor saturations,” *Systems and Control Letters*, vol. 61, no. 4, pp. 477–484, 2012.

[19] J. Hu, Z. Wang, Y. Niu, and L. K. Stergioulas, “H_∞ sliding mode observer design for aclass of nonlinear discrete time-delay systems: a delay-fractioning approach,” *International Journal of Robust and Nonlinear Control*. In press.

[20] J. Picard, “Asymptotic study of estimation problems with small observation noise,” in *Stochastic Modelling and Filtering (Rome, 1984)*, vol. 91 of *Lecture Notes in Control and Inference Science*, Springer, 1987.

[21] B. Shen, Z. Wang, and Y. S. Hung, “Distributed H_∞-consensus filtering in sensor networks with multiple missing measurements: the finite-horizon case,” *Automatica*, vol. 46, no. 10, pp. 1682–1688, 2010.

[22] B. Shen, Z. Wang, and X. Liu, “Bounded H_∞ synchronization and state estimation for discrete time-varying stochastic complex networks over a finite-horizon,” *IEEE Transactions on Neural Networks*, vol. 22, no. 1, pp. 145–157, 2011.

[23] B. Shen, Z. Wang, Y. Hung, and G. Chesi, “Distributed H_∞ filtering for polynomial nonlinear stochastic systems in sensor networks,” *IEEE Transactions on Industrial Electronics*, vol. 58, no. 5, pp. 1971–1979, 2011.

[24] Z. Wang, J. Lam, and X. Liu, “Filtering for a class of nonlinear discrete-time stochastic systems with state delays,” *Journal of Computational and Applied Mathematics*, vol. 201, no. 1, pp. 153–163, 2007.

[25] Z. Wang, X. Liu, Y. Liu, J. Liang, and V. Vinciotti, “An extended Kalman filtering approach to modelling nonlinear dynamic gene regulatory networks via short gene expression time series,” *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 6, no. 3, pp. 410–419, 2009.

[26] G. Wei and H. Shu, “H_∞ filtering on nonlinear stochastic systems with delay,” *Chaos, Solitons and Fractals*, vol. 33, no. 2, pp. 663–670, 2007.

[27] M. R. James and F. Le Gland, “Consistent parameter estimation for partially observed diffusions with small noise,” *Applied Mathematics and Optimization*, vol. 32, no. 1, pp. 47–72, 1995.

[28] R. S. Liptser and A. N. Shiryayev, *Statistics of Random Processes. I*, Springer, New York, NY, USA, 1977.

[29] Yu. A. Kutoyants, “Parameter estimation for diffusion type processes of observation,” *Mathematische Operationsforschung und Statistik Series Statistics*, vol. 15, no. 4, pp. 541–551, 1984.

[30] T. Deck, “Asymptotic properties of Bayes estimators for Gaussian Itô-processes with noisy observations,” *Journal of Multivariate Analysis*, vol. 97, no. 2, pp. 563–573, 2006.

[31] J. P. N. Bishwal, *Parameter Estimation in Stochastic Differential Equations*, Springer, Berlin, Germany, 2008.

[32] A. Brace, D. Gatarek, and M. Musiela, “The market model of interest rate dynamics,” *Mathematical Finance*, vol. 7, no. 2, pp. 127–155, 1997.

[33] D. Heath, R. Jarrow, and A. Morton, “Bond pricing and the term structure of interest-rates: a new methodology,” *Econometrica*, vol. 60, no. 1, pp. 77–105, 1992.

[34] J. Hull and A. White, "The pricing of option on assets with stochastic volatilities," *The Journal of Finance*, vol. 42, no. 2, pp. 281–300, 1987.
[35] J. Hull and A. White, "The general Hull-White model and super calibration," *Financial Analysis Journal*, vol. 57, no. 6, pp. 34–43, 2001.
[36] G. Kallianpur and R. S. Selukar, "Parameter estimation in linear filtering," *Journal of Multivariate Analysis*, vol. 39, no. 2, pp. 284–304, 1991.
[37] V. I. Arnold, *Geometrical Methods in the Theory of Ordinary Differential Equations*, Springer, New York, NY, USA, 1983.
[38] E. L. Ince, *Ordinary Differential Equations*, Dover, New York, NY, USA, 1944.