Probability in the Engineering and Informational Sciences (2024), 1–21
doi:10.1017/S0269964824000147

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Method of moments estimation for the superposition of square-root diffusions

Yan-Feng Wu and Jian-Qiang Hu

Department of Management Science, Fudan University, Shanghai 200433, China
Corresponding author: Jian-Qiang Hu; Email: hujq@fudan.edu.cn

Keywords: CIR model; Method of moments; Square-root diffusion; Superposition

Abstract
We consider the problem of parameter estimation for the superposition of square-root diffusions. We first derive the explicit formulas for the moments and auto-covariances based on which we develop our moment estimators. We then establish a central limit theorem for the estimators with the explicit formulas for the asymptotic covariance matrix. Finally, we conduct numerical experiments to validate our method.

1. Introduction

Square-root diffusion has become a popular model in the econometric area since it was first proposed by [12] for modeling the term structure of interest rates, thus it is also called Cox–Ingersoll–Ross (CIR) model. Heston [17] further extends its usage in option pricing by using it in modeling the underling asset price volatility, to improve the famous Black–Scholes option pricing formula. Among an important class of processes, affine jump-diffusions (see [14]), square-root diffusion functions as a major component. One-factor square-root diffusion is preferred for simple cases, however, superposition of several square-root diffusions or multi-factor square-root diffusion plays a more important role in many applications because of its better explanation for many scenarios in practice. Barndorff-Nielsen [5] constructs a class of stochastic processes based on the superposition of Ornstein–Uhlenbeck type processes. In this paper we consider the problem of parameter estimation for superposition of square-root diffusions.

Due to lack of closed-form transition density, maximum likelihood estimation (MLE) is not applicable to the problem of our interest without approximation. Even for one-factor square-root diffusion, this remains the case, though the transition law is known as a scaled noncentral chi-square distribution (see [12]). What is worse is the log likelihood is not globally concave, as pointed out in [22] for the one-factor model, which makes the efficiency of MLE depending on the goodness of starting values in numerical optimization algorithms. Therefore, Overbeck and Rydén [22] propose conditional least-squares estimators for the parameters of the one-factor model. Almost all the previous works on the parameter estimation for the superposition of square-root diffusions are based on various types of approximations. By using a closed-form approximation of the density, Ait-Sahalia and Kimmel [1] design an MLE for the multi-factor model, in which they first make the factors (states) observable by inverting the zero-coupon bond yields. This simplification makes it not applicable to general cases where the states are latent, such as the superposition of square-root diffusions studied in this paper. Using Kalman filter is another way of approximation. By approximating the transition density with a normal density, Geyer and Pichler [16] first infer the unobservable states with an approximate Kalman filter, and then estimate the parameters with quasi-maximum-likelihood estimation; see [9] for another Kalman filter approach. Christoffersen *et al.* [10] consider nonlinear Kalman filtering of multi-factor affine term structure models. Based on continuous time observations, Barczy *et al.* [4] estimate the parameters by MLE and least squares of a subcritical affine two-factor model in which one factor is a square-root diffusion. There are some works that only consider the drift-parameter estimation for the one-factor model, such as [2], [13], and parameter estimation for extended one-dimensional CIR model, such as [3], [20], [11], and [21]. A gradient-based simulated MLE estimate is proposed in [23, 24] for a related model. In summary, these approximations used in the MLE methods are computationally time-consuming, and some of the assumptions may be hard to satisfy in practical applications.

Method of moments (MM) offers an alternative to likelihood-based estimation, especially for cases where the likelihood does not have a closed-form expression while the moments can be easily derived analytically, see [26] and [8] for detailed description. The major issue of moment-based methods is the so-called statistical inefficiency in the sense that the higher order moments are used, the greater likelihood of estimation bias occurs. However, it is possible to overcome this problem by making use of relative low order moments and employing better computational techniques, for example, see [27–30].

In this paper, we consider the superposition of square-root diffusions and develop explicit MM estimators for the parameters in such a model. We first derive the closed-form formulas for the moments and auto-covariances, and then use them to develop our MM estimators. In fact, we only need relatively low order moments and auto-covariances. We also prove the central limit theorem for our MM estimators and derive the explicit formulas for the asymptotic covariance matrix. Additionally, numerical experiments are conducted to test the efficiency of our method. One clear advantage of our estimators is their simplicity and ease of implementation.

The rest of the paper is organized as follows. In Section 2, we present the model for the superposition of square-root diffusions and calculate the moments and auto-covariances. In Section 3, we derive our MM estimators and establish the central limit theorem. In Section 4, we provide some numerical experiments. Section 5 concludes this paper. The appendix contains some of the detailed calculations omitted in the main body of the paper.

2. Preliminaries

For ease of exposition, we focus on the superposition of two square-root diffusions which is usually used in most applications. The basic of our method can be extended to cases with multiple square-root diffusions, though as the number of square-root diffusions increases, calculations become more tedious and complex. The two-factor square-root diffusion can be described by the following stochastic differential equations:

$$x(t) = x_1(t) + x_2(t), \tag{2.1}$$

$$dx_1(t) = k_1(\theta_1 - x_1(t))dt + \sigma_{x1}\sqrt{x_1(t)}dw_1(t), \tag{2.2}$$

$$dx_2(t) = k_2(\theta_2 - x_2(t))dt + \sigma_{x2}\sqrt{x_2(t)}dw_2(t), \tag{2.3}$$

where the two square-root diffusions (also called CIR processes) $x_1(t)$ and $x_2(t)$ are independent of each other, $w_1(t)$ and $w_2(t)$ are two Wiener processes with independent initial values $x_1(0) > 0$ and $x_2(0) > 0$, respectively. The parameters $k_i > 0, \theta_i > 0, \sigma_{xi} > 0$ satisfy the condition $\sigma_{xi}^2 \leq 2k_i\theta_i$ for $i = 1, 2$, which assures $x_i(t) > 0$ for $t > 0$ (see [12]). The component processes $x_1(t)$, $x_2(t)$ and the driving processes $w_1(t)$, $w_2(t)$ are all latent, except that $x(t)$ is observable.

The square-root diffusions in Eqs. (2.2) and (2.3) can be re-written as:

$$x_i(t) = e^{-k_i(t-s)}x_i(s) + \theta_i\left[1 - e^{-k_i(t-s)}\right] + \sigma_{xi}e^{-k_it}\int_s^t e^{k_iu}\sqrt{x_i(u)}dw_i(u), \tag{2.4}$$

https://doi.org/10.1017/S0269964824000147 Published online by Cambridge University Press

for $i = 1, 2$. The process $x_i(t)$ decays at an exponential rate $e^{-k_i t}$, thus the parameter k_i is called the decay parameter. It is easy to see that $x_i(t)$ has a long-run mean θ_i and its instantaneous volatility is determined by its current value and σ_{xi}. In fact, the process $x_i(t)$ is a Markov process and has a steady-state gamma distribution with mean θ_i and variance $\theta_i \sigma_{xi}^2/(2k_i)$ (see [12]). Since we are interested in estimating the parameters based on a larger number of samples of the process, without loss of generality, throughout this paper it is assumed that $x_i(0)$ is distributed according to the steady-state distribution of $x_i(t)$, implying that $x_i(t)$ is strictly stationary and ergodic (see [22]). Therefore, the process $x(t)$ is also strictly stationary and ergodic. Actually, all the results we shall derive in the rest of this paper remain the same for any non-negative $x_1(0)$ and $x_2(0)$ as long as the observation time t is sufficiently long. Since $x_i(t)$ is stationary, it has a gamma distribution with mean θ_i and variance $\theta_i \sigma_{xi}^2/(2k_i)$ and its mth moment is given by

$$E[x_i^m(t)] = \prod_{j=0}^{m-1} \left(\theta_i + \frac{j\sigma_{xi}^2}{2k_i} \right), \quad m = 1, 2, \ldots. \tag{2.5}$$

Though modeled as a continuous time process, $x(t)$ is actually observed at discrete time points. Assume $x(t)$ is observed at equidistant time points $t = nh, n = 0, 1, \ldots, N$ and let $x_n \triangleq x(nh)$. Similarly, let $x_{in} \triangleq x_i(nh)$. We denote the mth central moment of x_n by $\mathrm{cm}_m[x_n]$, that is, $\mathrm{cm}_m[x_n] \triangleq E[(x_n - E[x_n])^m]$. Then, we have

$$\mathrm{cm}_m[x_n] = \sum_{j=0}^{m} C_m^j E[(x_{1n} - \theta_1)^j] E[(x_{2n} - \theta_2)^{m-j}], \tag{2.6}$$

due to the independence between x_{1n} and x_{2n}, where C_m^j is the combinatorial number. For notation simplicity, we introduce $\sigma_1 \triangleq \sigma_{x1}^2/(2k_1)$ and $\sigma_2 \triangleq \sigma_{x2}^2/(2k_2)$.

With Eqs. (2.4)–(2.6), we have the following moment and auto-covariance formulas

$$E[x_n] = \theta_1 + \theta_2, \tag{2.7}$$

$$\mathrm{var}(x_n) = \theta_1 \sigma_1 + \theta_2 \sigma_2, \tag{2.8}$$

$$\mathrm{cm}_3[x_n] = 2\theta_1 \sigma_1^2 + 2\theta_2 \sigma_2^2, \tag{2.9}$$

$$\mathrm{cov}(x_n, x_{n+1}) = e^{-k_1 h} \theta_1 \sigma_1 + e^{-k_2 h} \theta_2 \sigma_2, \tag{2.10}$$

$$\mathrm{cov}(x_n, x_{n+2}) = e^{-2k_1 h} \theta_1 \sigma_1 + e^{-2k_2 h} \theta_2 \sigma_2, \tag{2.11}$$

$$\mathrm{cov}(x_n, x_{n+3}) = e^{-3k_1 h} \theta_1 \sigma_1 + e^{-3k_2 h} \theta_2 \sigma_2. \tag{2.12}$$

The intermediate steps of the derivation are omitted because they are straightforward. We will use these six moments/auto-covariances to construct our estimators of the parameters in the next section.

https://doi.org/10.1017/S0269964824000147 Published online by Cambridge University Press

3. Parameter estimation

In this section, we derive our MM estimators for the six parameters in Eqs. (2.2) and (2.3) based on the moments/auto-covariances obtained in the previous section. We prove both the moment/auto-covariance estimators and the MM estimators satisfy the central limit theorem and also compute the explicit formulas for their asymptotic covariance matrices.

Assume that we are given a sample path of $x(t)$, $X_1, \ldots, X_N$, that is, samples of the stochastic process $x(t)$ observed at $t = h, \ldots, Nh$, which will be used to calculate the sample moments and auto-covariances of x_n as the estimates for their population counterparts:

$$
\begin{aligned}
E[x_n] &\approx \overline{X} \triangleq \frac{1}{N}\sum_{n=1}^{N} X_n,\\
\mathrm{var}(x_n) &\approx S^2 \triangleq \frac{1}{N}\sum_{n=1}^{N} (X_n - \overline{X})^2,\\
\mathrm{cm}_3[x_n] &\approx \hat{\mathrm{cm}}_3[x_n] \triangleq \frac{1}{N}\sum_{n=1}^{N} (X_n - \overline{X})^3,\\
\mathrm{cov}(x_n, x_{n+1}) &\approx \hat{\mathrm{cov}}(x_n, x_{n+1}) \triangleq \frac{1}{N-1}\sum_{n=1}^{N-1} (X_n - \overline{X})(X_{n+1} - \overline{X}),\\
\mathrm{cov}(x_n, x_{n+2}) &\approx \hat{\mathrm{cov}}(x_n, x_{n+2}) \triangleq \frac{1}{N-2}\sum_{n=1}^{N-2} (X_n - \overline{X})(X_{n+2} - \overline{X}),\\
\mathrm{cov}(x_n, x_{n+3}) &\approx \hat{\mathrm{cov}}(x_n, x_{n+3}) \triangleq \frac{1}{N-3}\sum_{n=1}^{N-3} (X_n - \overline{X})(X_{n+3} - \overline{X}).
\end{aligned}
$$

Let

$$
\begin{aligned}
\boldsymbol{\gamma} &= (E[x_n], \mathrm{var}(x_n), \mathrm{cm}_3[x_n], \mathrm{cov}(x_n, x_{n+1}), \mathrm{cov}(x_n, x_{n+2}), \mathrm{cov}(x_n, x_{n+3}))^T,\\
\hat{\boldsymbol{\gamma}} &= (\overline{X}, S^2, \hat{\mathrm{cm}}_3[x_n], \hat{\mathrm{cov}}(x_n, x_{n+1}), \hat{\mathrm{cov}}(x_n, x_{n+2}), \hat{\mathrm{cov}}(x_n, x_{n+3}))^T,
\end{aligned}
$$

where T denotes the transpose of a vector or matrix. We also define

$$
\begin{aligned}
z_n^1 &\triangleq (x_n - E[x_n])^2, & z_n^2 &\triangleq (x_n - E[x_n])^3,\\
z_n^3 &\triangleq (x_n - E[x_n])(x_{n+1} - E[x_n]), & z_n^4 &\triangleq (x_n - E[x_n])(x_{n+2} - E[x_n]),\\
z_n^5 &\triangleq (x_n - E[x_n])(x_{n+3} - E[x_n]), & z_n &\triangleq (z_n^1, z_n^2, z_n^3, z_n^4, z_n^5).
\end{aligned}
$$

Let

$$
\begin{aligned}
\sigma_{11} &= \lim_{N\to\infty} \frac{1}{N} \mathbf{1}_N^T [\mathrm{cov}(x_r, x_c)]_{N\times N} \mathbf{1}_N,\\
\sigma_{1m} &= \lim_{N\to\infty} \frac{1}{N} \mathbf{1}_N^T [\mathrm{cov}(x_r, z_c^{m-1})]_{N\times N_m} \mathbf{1}_{N_m}, \qquad 2 \le m \le 6,\\
\sigma_{lm} &= \lim_{N\to\infty} \frac{1}{N} \mathbf{1}_{N_l}^T [\mathrm{cov}(z_r^{l-1}, z_c^{m-1})]_{N_l\times N_m} \mathbf{1}_{N_m}, \quad 2 \le l \le m \le 6,
\end{aligned}
$$

https://doi.org/10.1017/S0269964824000147 Published online by Cambridge University Press

where $(N_2, N_3, N_4, N_5, N_6) = (N, N, N-1, N-2, N-3)$ and $\mathbf{1}_p = (1, \cdots, 1)^T$ with p elements. We further introduce the following notations:

$$\begin{aligned} e_{ij} &\triangleq \frac{e^{-jk_i h}}{1-e^{-jk_i h}}, \qquad i=1,2, j=1,2,\ldots, \\ e_{p,q} &\triangleq \frac{e^{-(pk_1+qk_2)h}}{1-e^{-(pk_1+qk_2)h}}, \quad p=0,1,\cdots, q=0,1,\ldots, \end{aligned}$$

and $\bar{x}_{1n} \triangleq x_{1n} - \theta_1$ and $\bar{x}_{2n} \triangleq x_{2n} - \theta_2$. We are now ready to prove the following central limit theorem for our moment/auto-covariance estimators.

Theorem 3.1

$$\lim_{N\to\infty} \sqrt{N}(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}) \overset{d}{=} \mathcal{N}(\mathbf{0}, \Sigma), \tag{3.1}$$

where $\overset{d}{=}$ denotes equal in distribution and $\mathcal{N}(\mathbf{0}, \Sigma)$ is a multivariate normal distribution with mean $\mathbf{0}$ and asymptotic covariance matrix $\Sigma = [\sigma_{lm}]_{6\times 6}$, with entries

$$\sigma_{lm} = c_{lm,1} + c_{lm,2}, \quad l, m = 1, \ldots, 6,$$

where $c_{lm,1}$, $c_{lm,2}$ are some constants which are provided in the Appendix.

Proof. The square-root diffusion $x_i(t)$ $(i=1,2)$ is a Markov process, so is the discrete-time process x_{in}. $\{(x_{1n}, x_{2n}, x_n)\}$ is also a Markov process. Under the condition $\sigma_{xi}^2 \le 2k_i\theta_i$, the strictly stationary square-root diffusion $x_i(t)$ is ρ-mixing; see Proposition 2.8 in [15]. Hence, $x_i(t)$ is ergodic and strong mixing (also known as α-mixing) with an exponential rate; see [6]. The discrete-time process $\{x_{in}\}$ is also ergodic and inherits the exponentially fast strong mixing properties of $x_i(t)$. It can be verified that $\{(x_{1n}, x_{2n}, x_n)\}$ is also ergodic and strong mixing with an exponential rate. Applying the central limit theorem for strictly stationary strong mixing sequences (see Theorem 1.7 in [18]), we can prove Eq. (3.1) as Corollary 3.1 in [15].

Next, we present the calculations for the diagonal entries of the asymptotic covariance matrix, however omit those for the off-diagonal entries since they are very similar. First, we point out the following formulas

$$\sum_{1\le r<c\le N} e^{-(c-r)kh} = \frac{e^{-kh}}{1-e^{-kh}}(N-1) - \frac{e^{-2kh-e^{-(N+1)kh}}}{(1-e^{-kh})^2},$$

which will be used frequently to calculate infinite series in these entries. We now calculate the six diagonal entries.

- σ_{11}: We have

$$\begin{aligned} \sigma_{11} &= \lim_{N\to\infty} \frac{1}{N}\left(N\mathrm{var}(x_n) + 2\sum_{1\le r<c\le N} \mathrm{cov}(x_r, x_c)\right) \\ &= \mathrm{var}(x_n) + 2\lim_{N\to\infty}\frac{1}{N}\sum_{1\le r<c\le N}(\mathrm{cov}(x_{1r}, x_{1c}) + \mathrm{cov}(x_{2r}, x_{2c})) \\ &= \mathrm{var}(x_{1n}) + \mathrm{var}(x_{2n}) + \frac{2e^{-k_1h}}{1-e^{-k_1h}}\mathrm{var}(x_{1n}) + \frac{2e^{-k_2h}}{1-e^{-k_2h}}\mathrm{var}(x_{2n}) \\ &= (1+2e_{11})\frac{\theta_1\sigma_{x1}^2}{2k_1} + (1+2e_{21})\frac{\theta_2\sigma_{x2}^2}{2k_2}, \end{aligned}$$

where we have used the results of $\mathrm{cov}(x_{1r}, x_{1c}) = e^{-(c-r)k_1 h}\mathrm{var}(x_{1r})$ $(r < c)$, $\mathrm{var}(x_{1r}) = \mathrm{var}(x_{1n})$, $\forall r = 1, \cdots, N$ (similarly for x_{2r}).

- σ_{22}: We have

$$\begin{aligned}\sigma_{22} &= \lim_{N\to\infty}\frac{1}{N}\left(N\mathrm{var}(z_n^1) + 2\sum_{1\le r<c\le N}\mathrm{cov}(z_r^1, z_c^1)\right)\\ &= \mathrm{var}(\bar{x}_{1n}^2) + \mathrm{var}(\bar{x}_{2n}^2) + 4\mathrm{var}(x_{1n})\mathrm{var}(x_{2n})\\ &\quad + \lim_{N\to\infty}\frac{2}{N}\sum_{1\le r<c\le N}(\mathrm{cov}(\bar{x}_{1r}^2, \bar{x}_{1c}^2) + \mathrm{cov}(\bar{x}_{2r}^2, \bar{x}_{2c}^2) + 4\mathrm{cov}(\bar{x}_{1r}\bar{x}_{2r}, \bar{x}_{1c}\bar{x}_{2c}))\\ &= c_{22,1} + c_{22,2},\end{aligned}$$

where $\mathrm{cov}(\bar{x}_{1r}^2, \bar{x}_{1c}^2)$ $(r < c)$ can be calculated as

$$\begin{aligned}&\mathrm{cov}(\bar{x}_{1r}^2, \bar{x}_{1c}^2)\\ &= e^{-2(c-r)k_1 h}\mathrm{var}(\bar{x}_{1r}^2) + \frac{\sigma_{x1}^2}{k_1}[e^{-(c-r)k_1 h} - e^{-2(c-r)k_1 h}]\mathrm{cov}(\bar{x}_{1r}^2, \bar{x}_{1r}),\end{aligned}$$

in which we have used the decay formula (A.2) for $\bar{x}_{1c}^2$ in the Appendix, similarly for $\mathrm{cov}(\bar{x}_{2r}^2, \bar{x}_{2c}^2)$ $(r < c)$, and $\mathrm{cov}(\bar{x}_{1r}\bar{x}_{2r}, \bar{x}_{1c}\bar{x}_{2c})$ $(r < c)$ as

$$\mathrm{cov}(\bar{x}_{1r}\bar{x}_{2r}, \bar{x}_{1c}\bar{x}_{2c}) = e^{-(c-r)(k_1+k_2)h}\mathrm{var}(x_{1r})\mathrm{var}(x_{2r}).$$

- σ_{33}: We have

$$\begin{aligned}\sigma_{33} &= \lim_{N\to\infty}\frac{1}{N}\left(N\mathrm{var}(z_n^2) + 2\sum_{1\le r<c\le N}\mathrm{cov}(z_r^2, z_c^2)\right)\\ &= c_{33,1} + c_{33,2},\end{aligned}$$

where $\mathrm{cov}(z_r^2, z_c^2)$ can be expanded as

$$\begin{aligned}\mathrm{cov}(z_r^2, z_c^2) &= (\mathrm{cov}(\bar{x}_{1r}^3, \bar{x}_{1c}^3) + \mathrm{cov}(\bar{x}_{2r}^3, \bar{x}_{2c}^3)) + 3(\mathrm{cov}(\bar{x}_{1r}^3, \bar{x}_{1c}\bar{x}_{2c}^2) + \mathrm{cov}(\bar{x}_{2r}^3, \bar{x}_{1c}^2\bar{x}_{2c}))\\ &\quad + 9(\mathrm{cov}(\bar{x}_{1r}^2\bar{x}_{2r}, \bar{x}_{1c}^2\bar{x}_{2c}) + \mathrm{cov}(\bar{x}_{1r}\bar{x}_{2r}^2, \bar{x}_{1c}\bar{x}_{2c}^2))\\ &\quad + 9(\mathrm{cov}(\bar{x}_{1r}^2\bar{x}_{2r}, \bar{x}_{1c}\bar{x}_{2c}^2) + \mathrm{cov}(\bar{x}_{1r}\bar{x}_{2r}^2, \bar{x}_{1c}^2\bar{x}_{2c}))\\ &\quad + 3(\mathrm{cov}(\bar{x}_{1r}^2\bar{x}_{2r}, \bar{x}_{2c}^3) + \mathrm{cov}(\bar{x}_{1r}\bar{x}_{2r}^2, \bar{x}_{1c}^3)).\end{aligned}$$

- σ_{44} : We have

$$\begin{aligned}\sigma_{44} &= \lim_{N\to\infty}\frac{1}{N}\left[(N-1)\mathrm{var}(z_n^3) + 2\sum_{1\le r<c\le N-1}\mathrm{cov}(z_r^3, z_c^3)\right]\\ &= c_{44,1} + c_{44,2},\end{aligned}$$

where $\mathrm{cov}(z_r^3, z_c^3)$ is expanded as

$$\begin{aligned}\mathrm{cov}(z_r^3, z_c^3) &= (\mathrm{cov}(\bar{x}_{1r}\bar{x}_{1(r+1)}, \bar{x}_{1c}\bar{x}_{1(c+1)}) + \mathrm{cov}(\bar{x}_{2r}\bar{x}_{2(r+1)}, \bar{x}_{2c}\bar{x}_{2(c+1)}))\\ &\quad + (\mathrm{cov}(\bar{x}_{1r}\bar{x}_{2(r+1)}, \bar{x}_{1c}\bar{x}_{2(c+1)}) + \mathrm{cov}(\bar{x}_{2r}\bar{x}_{1(r+1)}, \bar{x}_{2c}\bar{x}_{1(c+1)}))\\ &\quad + (\mathrm{cov}(\bar{x}_{1r}\bar{x}_{2(r+1)}, \bar{x}_{2c}\bar{x}_{1(c+1)}) + \mathrm{cov}(\bar{x}_{2r}\bar{x}_{1(r+1)}, \bar{x}_{1c}\bar{x}_{2(c+1)})).\end{aligned}$$

https://doi.org/10.1017/S0269964824000147 Published online by Cambridge University Press

- σ_{55} : We have

$$\sigma_{55} = \lim_{N\to\infty} \frac{1}{N}\left[(N-2)\mathrm{var}(z_n^4) + 2\sum_{1\le r<c\le N-2} \mathrm{cov}(z_r^4, z_c^4)\right]$$
$$= c_{55,1} + c_{55,2},$$

where $\mathrm{cov}(z_r^4, z_c^4)$ can be calculated through

$$\begin{aligned}\mathrm{cov}(z_r^4, z_c^4) = {} & (\mathrm{cov}(\bar{x}_{1r}\bar{x}_{1(r+2)}, \bar{x}_{1c}\bar{x}_{1(c+2)}) + \mathrm{cov}(\bar{x}_{2r}\bar{x}_{2(r+2)}, \bar{x}_{2c}\bar{x}_{2(c+2)}))\\ & + (\mathrm{cov}(\bar{x}_{1r}\bar{x}_{2(r+2)}, \bar{x}_{1c}\bar{x}_{2(c+2)}) + \mathrm{cov}(\bar{x}_{2r}\bar{x}_{1(r+2)}, \bar{x}_{2c}\bar{x}_{1(c+2)}))\\ & + (\mathrm{cov}(\bar{x}_{1r}\bar{x}_{2(r+2)}, \bar{x}_{2c}\bar{x}_{1(c+2)}) + \mathrm{cov}(\bar{x}_{2r}\bar{x}_{1(r+2)}, \bar{x}_{1c}\bar{x}_{2(c+2)})).\end{aligned}$$

- σ_{66} : We have

$$\sigma_{66} = \lim_{N\to\infty} \frac{1}{N}\left[(N-3)\mathrm{var}(z_n^5) + 2\sum_{1\le r<c\le N-3} \mathrm{cov}(z_r^5, z_c^5)\right]$$
$$= c_{66,1} + c_{66,2},$$

where $\mathrm{cov}(z_r^5, z_c^5)$ can be calculated through

$$\begin{aligned}\mathrm{cov}(z_r^5, z_c^5) = {} & (\mathrm{cov}(\bar{x}_{1r}\bar{x}_{1(r+3)}, \bar{x}_{1c}\bar{x}_{1(c+3)}) + \mathrm{cov}(\bar{x}_{2r}\bar{x}_{2(r+3)}, \bar{x}_{2c}\bar{x}_{2(c+3)}))\\ & + (\mathrm{cov}(\bar{x}_{1r}\bar{x}_{2(r+3)}, \bar{x}_{1c}\bar{x}_{2(c+3)}) + \mathrm{cov}(\bar{x}_{2r}\bar{x}_{1(r+3)}, \bar{x}_{2c}\bar{x}_{1(c+3)}))\\ & + (\mathrm{cov}(\bar{x}_{1r}\bar{x}_{2(r+3)}, \bar{x}_{2c}\bar{x}_{1(c+3)}) + \mathrm{cov}(\bar{x}_{2r}\bar{x}_{1(r+3)}, \bar{x}_{1c}\bar{x}_{2(c+3)})).\end{aligned}$$

The off-diagonal entries of Σ can be derived similarly. This completes our proof. □

With these moment and auto-covariance estimates, we are ready to construct our estimators for the parameters. In theory, we can obtain our parameter estimates by solving the system of moment equations using any general nonlinear root-finding algorithms. However, based on our extensive numerical experiments, we find that such an approach often leads to estimates with very large errors due to various reasons. Fortunately, we can overcome this problem by exploring the special characteristics of the equations of our interest and developing better numerical methods. In what follows, we first estimate the decay parameters k_1 and k_2, and then estimate the parameters in the marginal distribution of the process, that is, $\theta_1, \sigma_{x1}, \theta_2$ and σ_{x2}. Details of our analysis are presented in the following two subsections.

3.1. Estimation of the decay parameters

In order to estimate the decay rates $e^{-k_1 h}$ and $e^{-k_2 h}$ of the two component processes, we make use of system of equations consisting of Eqs. (2.8) and (2.10)–(2.12). In principal, we can solve this system of equations to obtain $e^{-k_1 h}, e^{-k_2 h}, \theta_1\sigma_1$ and $\theta_2\sigma_2$. However, it is difficult to find the true roots by using general numerical algorithms, such as Newton–Raphson method. In what follows, we propose a numerical procedure to solve the system of equations which can achieve very good precision.

For notational simplicity, we rewrite the above system of equations as following:

$$\begin{aligned} v_1 + v_2 &= b,\\ d_1 v_1 + d_2 v_2 &= b_1,\\ d_1^2 v_1 + d_2^2 v_2 &= b_2,\\ d_1^3 v_1 + d_2^3 v_2 &= b_3,\end{aligned}$$

https://doi.org/10.1017/S0269964824000147 Published online by Cambridge University Press

where $v_1 \triangleq \theta_1\sigma_1, v_2 \triangleq \theta_2\sigma_2$, $b \triangleq \mathrm{var}(x_n)$, $b_j \triangleq \mathrm{cov}(x_n, x_{n+j}), j = 1, 2, 3$, $d_1 \triangleq e^{-k_1 h}$ and $d_2 \triangleq e^{-k_2 h}$. By using the first two equations, we have

$$v_1 = \frac{-d_2 b + b_1}{d_1 - d_2}, \quad v_2 = \frac{d_1 b - b_1}{d_1 - d_2}.$$

Further,

$$d_2^2 = \frac{d_1 b_2 - b_3}{d_1 b - b_1}, \quad d_2 = \frac{d_1^2 b_1 - b_3}{d_1^2 b - b_2}.$$

Thus,

$$\frac{d_1 b_2 - b_3}{d_1 b - b_1} = \left(\frac{d_1^2 b_1 - b_3}{d_1^2 b - b_2}\right)^2,$$

which leads to the following quintic equation with unknown variable d_1

$$\begin{aligned}(bb_1^2 - b^2 b_2) \cdot d_1^5 + (b^2 b_3 - b_1^3) \cdot d_1^4 + 2(bb_2^2 - bb_1 b_3) \cdot d_1^3 \\ + 2(b_1^2 b_3 - bb_2 b_3) \cdot d_1^2 + (bb_3^2 - b_2^3) \cdot d_1 + (b_2^2 b_3 - b_1 b_3^2) = 0.\end{aligned} \tag{3.2}$$

The Jenkins–Traub algorithm for polynomial root-finding (see [19]) can be used to solve the quintic equation, an implementation of which is a function named *polyroot* in R programming language (see [25]). The roots returned by the algorithm are five complex numbers, among which we only need to keep those real-number roots that are between 0 and 1. Because of the symmetry between d_1 and d_2 (due to the symmetry of k_1 and k_2) d_2 must also be another root. Therefore, there should be at least two real-number roots that satisfy $d_2 = (d_1^2 b_1 - b_3)/(d_1^2 b - b_2)$. We will use this property to find the correct roots for d_1 and d_2. In our extensive numerical experiments, we are able to find a pair of roots for d_1 and d_2 satisfying the required conditions.

We now are ready to derive the estimators for the decay parameters. Denote the sample version of Eq. (3.2) by

$$\begin{aligned}(\hat{b}\hat{b}_1^2 - \hat{b}^2 \hat{b}_2) \cdot \mathrm{x}^5 + (\hat{b}^2 \hat{b}_3 - \hat{b}_1^3) \cdot \mathrm{x}^4 + 2(\hat{b}\hat{b}_2^2 - \hat{b}\hat{b}_1 \hat{b}_3) \cdot \mathrm{x}^3 \\ + 2(\hat{b}_1^2 \hat{b}_3 - \hat{b}\hat{b}_2 \hat{b}_3) \cdot \mathrm{x}^2 + (\hat{b}\hat{b}_3^2 - \hat{b}_2^3) \cdot \mathrm{x} + (\hat{b}_2^2 \hat{b}_3 - \hat{b}_1 \hat{b}_3^2) = 0,\end{aligned} \tag{3.3}$$

where $\hat{b}, \hat{b}_1, \hat{b}_2, \hat{b}_3$ are S^2, $\widehat{\mathrm{cov}}(x_n, x_{n+1})$, $\widehat{\mathrm{cov}}(x_n, x_{n+2})$, $\widehat{\mathrm{cov}}(x_n, x_{n+3})$, respectively. Let us denote the root obtained based on Eq. (3.3) for the estimate of d_1 by x^*, that is, $\hat{d}_1 = \mathrm{x}^*$, then our estimators for d_2, v_1 $(= \theta_1\sigma_1)$, v_2 $(= \theta_2\sigma_2)$ are given by

$$\hat{d}_2 = \frac{\hat{d}_1^2 \hat{b}_1 - \hat{b}_3}{\hat{d}_1^2 \hat{b} - \hat{b}_2}, \tag{3.4}$$

$$\hat{v}_1 = \frac{-\hat{d}_2 \hat{b} + \hat{b}_1}{\hat{d}_1 - \hat{d}_2}, \tag{3.5}$$

$$\hat{v}_2 = \frac{\hat{d}_1 \hat{b} - \hat{b}_1}{\hat{d}_1 - \hat{d}_2}. \tag{3.6}$$

Before closing this subsection, we point out that there may exist further improvement as stated in what follows. In estimating of the decay rates, we have just used the first four auto-covariances $\mathrm{cov}(x_n, x_n), \ldots, \mathrm{cov}(x_n, x_{n+3})$. Actually, $\forall j \geq 1$, we have

$$\mathrm{cov}(x_n, x_{n+j}) = e^{-jk_1 h}\theta_1\sigma_1 + e^{-jk_2 h}\theta_2\sigma_2.$$

One possible improvement is to make use of more these lagged auto-covariances. For instance, we can use $\mathrm{cov}(x_n, x_{n+1}), \mathrm{cov}(x_n, x_{n+2}), \mathrm{cov}(x_n, x_{n+3}), \mathrm{cov}(x_n, x_{n+4})$ to construct another system of equations based on which we can produce another estimate of (d_1, d_2). Then we average this estimate with the previous one, produced by the first four auto-covariances $\mathrm{cov}(x_n, x_n), \ldots, \mathrm{cov}(x_n, x_{n+3})$, to obtain the final estimate, which should be more accurate. We can use more auto-covariances to improve the accuracy of our estimates.

3.2. *Estimation of the marginal distribution parameters*

In order to estimate the marginal distribution parameters, we make use of Eqs. (2.7) and (2.9). Since $\theta_1\sigma_1$ $(= v_1)$ and $\theta_2\sigma_2$ $(= v_2)$ can be estimated by Eqs. (3.5) and (3.6) respectively, we rewrite $\mathrm{cm}_3[x_n]$ as $\mathrm{cm}_3[x_n] = 2v_1^2/\theta_1 + 2v_2^2/\theta_2$. Then we have the following quadratic equation

$$a_1\theta_1^2 + a_2\theta_1 + a_3 = 0, \tag{3.7}$$

with coefficiencies

$$a_1 \triangleq \mathrm{cm}_3[x_n], \quad a_2 \triangleq 2v_2^2 - 2v_1^2 - \mathrm{cm}_3[x_n]E[x_n], \quad a_3 \triangleq 2v_1^2 E[x_n].$$

Eq. (3.7) has two solutions

$$\mathrm{x}_1 = \frac{-a_2 + \sqrt{\Delta}}{2a_1}, \quad \mathrm{x}_2 = \frac{-a_2 - \sqrt{\Delta}}{2a_1},$$

where $\Delta = a_2^2 - 4a_1a_3 = 4(\theta_1\theta_2)^2(\sigma_1^2 - \sigma_2^2)^2$. We can show that only one equals θ_1. If $\sigma_1 > \sigma_2$, we have

$$\mathrm{x}_1 = \frac{(\theta_1\sigma_1)^2 + \theta_1\theta_2\sigma_1^2}{\theta_1\sigma_1^2 + \theta_2\sigma^2}, \quad \mathrm{x}_2 = \theta_1.$$

Otherwise, we have

$$\mathrm{x}_1 = \theta_1, \quad \mathrm{x}_2 = \frac{(\theta_1\sigma_1)^2 + \theta_1\theta_2\sigma_1^2}{\theta_1\sigma_1^2 + \theta_2\sigma^2}.$$

In determining whether $\sigma_1 > \sigma_2$, we only need to verify if $v_1/\mathrm{x}_2 < v_2/\theta_2$. Assuming $\sigma_1 > \sigma_2$, then $\mathrm{x}_2 = \theta_1$ and

$$\theta_2 = E[x_n] - \mathrm{x}_2, \quad \sigma_1 = v_1/\mathrm{x}_2, \quad \sigma_2 = v_2/\theta_2.$$

Denote the sample version of Eq. (3.7) as

$$\hat{a}_1\mathrm{x}^2 + \hat{a}_2\mathrm{x} + \hat{a}_3 = 0, \tag{3.8}$$

https://doi.org/10.1017/S0269964824000147 Published online by Cambridge University Press

with $\hat{a}_1$, $\hat{a}_2$, $\hat{a}_3$ being the sample estimates of a_1, a_2, a_3, respectively. Let us denote the root obtained based on Eq. (3.8) for θ_1 by x^*, that is, $\hat{\theta}_1 = \mathrm{x}^*$, then the other estimators are given by

$$\hat{\theta}_2 = \overline{X} - \hat{\theta}_1, \quad \hat{\sigma}_1 = \hat{v}_1/\hat{\theta}_1, \quad \hat{\sigma}_2 = \hat{v}_2/\hat{\theta}_2.$$

Finally, we have

$$\hat{k}_1 = -\log \hat{d}_1/h, \quad \hat{k}_2 = -\log \hat{d}_2/h, \quad \hat{\sigma}_{x1} = \sqrt{2\hat{\sigma}_1\hat{k}_1}, \quad \hat{\sigma}_{x2} = \sqrt{2\hat{\sigma}_2\hat{k}_2}.$$

We can define moment Eqs. (2.7)–(2.12) as a mapping $f : \mathbb{R}^6 \to \mathbb{R}^6$, that is,

$$f(\theta_1, \theta_2, \sigma_1, \sigma_2, d_1, d_2) = \boldsymbol{\gamma},$$

where $\boldsymbol{\gamma}$, by definition, represents the vector of the first moment, the second central moment and the third central moment and the first three auto-covariances of the process x_n.

Let J_f be the Jacobian of f, which is a 6×6 matrix of partial derivatives of f with respect to the entries of f. Taking the first row of J_f as an example, it is the gradient of $E[x_n]$ with respect to $(\theta_1, \theta_2, \sigma_1, \sigma_2, d_1, d_2)^T$, that is,

$$\left(\frac{\partial E[x_n]}{\partial \theta_1}, \ldots, \frac{\partial E[x_n]}{\partial d_2}\right).$$

With some calculations, we have

$$J_f = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ \sigma_1 & \sigma_2 & \theta_1 & \theta_2 & 0 & 0 \\ 2\sigma_1^2 & 2\sigma_2^2 & 4\theta_1\sigma_1 & 4\theta_2\sigma_2 & 0 & 0 \\ d_1\sigma_1 & d_2\sigma_2 & d_1\theta_1 & d_2\sigma_2 & \theta_1\sigma_1 & \theta_2\sigma_2 \\ d_1^2\sigma_1 & d_2^2\sigma_2 & d_1^2\theta_1 & d_2^2\sigma_2 & 2d_1\theta_1\sigma_1 & 2d_2\theta_2\sigma_2 \\ d_1^3\sigma_1 & d_2^3\sigma_2 & d_1^3\theta_1 & d_2^3\sigma_2 & 3d_1^2\theta_1\sigma_1 & 3d_2^2\theta_2\sigma_2 \end{bmatrix}.$$

Though the inverse function $f^{-1}(\boldsymbol{\gamma})$ does not have an explicit expression, we will show that the Jacobian J_f is invertible, thus $f^{-1}(\boldsymbol{\gamma})$ exists by the inverse function theorem.

We further define our final parameter estimators as another mapping $g : \mathbb{R}^6 \to \mathbb{R}^6$, that is,

$$g(\hat{\theta}_1, \hat{\theta}_2, \hat{\sigma}_1, \hat{\sigma}_2, \hat{d}_1, \hat{d}_2) = (\hat{\theta}_1, \hat{\theta}_2, \hat{\sigma}_{x1}, \hat{\sigma}_{x2}, \hat{k}_1, \hat{k}_2)^T,$$

whose Jacobian, denoted by J_g, is given as the following

$$J_g = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{-\log d_1/h}{\sqrt{-2\sigma_1 \log d_1/h}} & 0 & \frac{-\sigma_1/(d_1 h)}{\sqrt{-2\sigma_1 \log d_1/h}} & 0 \\ 0 & 0 & 0 & \frac{-\log d_2/h}{\sqrt{-2\sigma_2 \log d_2/h}} & 0 & \frac{-\sigma_2/(d_2 h)}{\sqrt{-2\sigma_2 \log d_2/h}} \\ 0 & 0 & 0 & 0 & -\frac{1}{d_1 h} & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{1}{d_2 h} \end{bmatrix}.$$

We first present the following central limit theorem for the parameter estimators $(\hat{\theta}_1, \hat{\theta}_2, \hat{\sigma}_1, \hat{\sigma}_2, \hat{d}_1, \hat{d}_2)^T$.

https://doi.org/10.1017/S0269964824000147 Published online by Cambridge University Press

Theorem 3.2 *Suppose the superposition of two square-root diffusions described by Eqs. (2.1)–(2.3) does not degenerate into a one square-root diffusion, that is, $d_1 \neq d_2$, $\theta_1 \neq \theta_2$, and $\sigma_1 \neq \sigma_2$. Then the parameter estimators $(\hat{\theta}_1, \hat{\theta}_2, \hat{\sigma}_1, \hat{\sigma}_2, \hat{d}_1, \hat{d}_2)^T$ exist with probability tending to one and satisfy*

$$\lim_{N\to\infty} \sqrt{N}((\hat{\theta}_1, \hat{\theta}_2, \hat{\sigma}_1, \hat{\sigma}_2, \hat{d}_1, \hat{d}_2)^T - (\theta_1, \theta_2, \sigma_1, \sigma_2, d_1, d_2)^T) \stackrel{d}{=} \mathcal{N}(\mathbf{0}, J_f^{-1}\Sigma(J_f^{-1})^T). \tag{3.9}$$

Proof. We first prove the Jacobian J_f is invertible. With some elementary row operations, J_f is equivalent to the following matrix

$$J_1 \triangleq \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & \frac{\theta_1}{\sigma_2-\sigma_1} & \frac{\theta_2}{\sigma_2-\sigma_1} & 0 & 0 \\ 0 & 0 & 1 & -\frac{\theta_2}{\theta_1} & 0 & 0 \\ 0 & 0 & 0 & a_{44} & \theta_1\sigma_1 & \theta_2\sigma_2 \\ 0 & 0 & 0 & a_{54} & 2d_1\theta_1\sigma_1 & 2d_2\theta_2\sigma_2 \\ 0 & 0 & 0 & a_{64} & 3d_1^2\theta_1\sigma_1 & 3d_2^2\theta_2\sigma_2 \end{bmatrix},$$

where

$$\begin{aligned} a_{44} &\triangleq (d_1+d_2)\theta_2 - \frac{2\theta_2}{\sigma_2-\sigma_1}(d_2\sigma_2 - d_1\sigma_1), \\ a_{54} &\triangleq (d_1^2+d_2^2)\theta_2 - \frac{2\theta_2}{\sigma_2-\sigma_1}(d_2^2\sigma_2 - d_1^2\sigma_1), \\ a_{64} &\triangleq (d_1^3+d_2^3)\theta_2 - \frac{2\theta_2}{\sigma_2-\sigma_1}(d_2^3\sigma_2 - d_1^3\sigma_1). \end{aligned}$$

Therefore, the invertibility of J_f depends on that of the bottom right submatrix in J_1, that is,

$$J_2 \triangleq \begin{bmatrix} a_{44} & \theta_1\sigma_1 & \theta_2\sigma_2 \\ a_{54} & 2d_1\theta_1\sigma_1 & 2d_2\theta_2\sigma_2 \\ a_{64} & 3d_1^2\theta_1\sigma_1 & 3d_2^2\theta_2\sigma_2 \end{bmatrix}.$$

With elementary row operations, J_2 is equivalent to the following matrix

$$J_3 \triangleq \begin{bmatrix} a_{11} & 3(d_2^2-d_1^2)\theta_1\sigma_1 & 0 \\ a_{21}(d_1+d_2)/d_1 - a_{11} & 0 & 0 \\ a_{64} & 3d_1^2\theta_1\sigma_1 & 3d_2^2\theta_2\sigma_2 \end{bmatrix},$$

where

$$\begin{aligned} a_{11} &\triangleq (3d_1d_2^2 + 2d_2^3 - d_1^3)\theta_2 - \frac{2\theta_2}{\sigma_2-\sigma_1}(2d_2^3\sigma_2 - 3d_1d_2^2\sigma_1 + d_1^3\sigma_1), \\ a_{21} &\triangleq \left(\frac{3}{2}d_1^2d_2 + \frac{1}{2}d_2^3 - d_1^3\right)\theta_2 - \frac{2\theta_2}{\sigma_2-\sigma_1}\left(\frac{1}{2}d_2^3\sigma_2 - \frac{3}{2}d_1^2d_2\sigma_1 + d_1^3\sigma_1\right). \end{aligned}$$

https://doi.org/10.1017/S0269964824000147 Published online by Cambridge University Press

Thus, the invertibility of J_f reduces to whether the term $a_{21}(d_1+d_2)/d_1 - a_{11}$ equals zero or not. After some calculations, we have

$$\begin{aligned} a_{21}(d_1+d_2)/d_1 - a_{11} &= \theta_2 \frac{\sigma_1+\sigma_2}{\sigma_2-\sigma_1} \frac{d_2^4}{2d_1} \left[\left(\frac{d_1}{d_2}\right)^3 - 3\left(\frac{d_1}{d_2}\right)^2 + 3\left(\frac{d_1}{d_2}\right) - 1 \right] \\ &= \theta_2 \frac{\sigma_1+\sigma_2}{\sigma_2-\sigma_1} \frac{d_2^4}{2d_1} \left(\frac{d_1}{d_2} - 1\right)^3 . \end{aligned}$$

Note that $d_1 \neq d_2$. Therefore, $a_{21}(d_1+d_2)/d_1 - a_{11}$ does not equal zero, consequently the Jacobian J_f is invertible.

With the covariance matrix Σ in Theorem 3.1, it is apparently

$$E[x_n^2 + \bar{x}_n^4 + \bar{x}_n^6 + \bar{x}_n^2\bar{x}_{n+1}^2 + \bar{x}_n^2\bar{x}_{n+2}^2 + \bar{x}_n^2\bar{x}_{n+3}^2] < \infty, \tag{3.10}$$

where $\bar{x}_{n+j} \triangleq x_{n+j} - E[x_n], j = 0, 1, 2, 3$.

In summary, we have verified: (1) the mapping f from parameters $(\theta_1, \theta_2, \sigma_1, \sigma_2, d_1, d_2)^T$ to moments $\boldsymbol{\gamma}$, are continuously differentiable with nonsingular Jacobian J_f, and (2) the summation of the double order moments is finite, that is, Eq. (3.10). Therefore, by applying Theorem 4.1 in [26], we have Theorem 3.2. This completes the proof. □

We now present the central limit theorem for our parameter estimators $(\hat{\theta}_1, \hat{\theta}_2, \hat{\sigma}_{x1}, \hat{\sigma}_{x2}, \hat{k}_1, \hat{k}_2)^T$.

Theorem 3.3 *Suppose the same assumption as in Theorem 3.2 holds. Then,*

$$\lim_{N\to\infty} \sqrt{N}((\hat{\theta}_1, \hat{\theta}_2, \hat{\sigma}_{x1}, \hat{\sigma}_{x2}, \hat{k}_1, \hat{k}_2)^T - (\theta_1, \theta_2, \sigma_{x1}, \sigma_{x2}, k_1, k_2)^T) \overset{d}{=} \mathcal{N}(\mathbf{0}, J_g J_f^{-1} \Sigma (J_f^{-1})^T J_g^T). \tag{3.11}$$

Proof. Note that the mapping $g(\theta_1, \theta_2, \sigma_1, \sigma_2, d_1, d_2)$ takes the following expression

$$\begin{bmatrix} \theta_1 \\ \theta_2 \\ \sigma_{x1} \\ \sigma_{x2} \\ k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \sqrt{-2\sigma_1 \log d_1/h} \\ \sqrt{-2\sigma_2 \log d_2/h} \\ -\log d_1/h \\ -\log d_2/h \end{bmatrix} \equiv g(\theta_1, \theta_2, \sigma_1, \sigma_2, d_1, d_2),$$

which is differentiable and has the Jacobian J_g. Meanwhile, we have verified the convergence of the estimators $(\hat{\theta}_1, \hat{\theta}_2, \hat{\sigma}_1, \hat{\sigma}_2, \hat{d}_1, \hat{d}_2)^T$ in Theorem 3.2. Therefore, by applying the delta method (Theorem 3.1 in [26]), we have the convergence in Eq. (3.11). This completes the proof. □

4. Numerical experiments

In this section, we conduct numerical experiments to test the estimators derived in the previous section. First, we validate the accuracy of our estimators under different parameter value settings. Then, we analyze the asymptotic performances of our estimators as the sample length increases.

In the first set of experiments, we consider four different combinations of parameter values, with S0 as the base combination in which $k_1 = 2, \theta_1 = 1.5, \sigma_{x1} = 1.6, k_2 = 0.2, \theta_2 = 0.5, \sigma_{x2} = 0.2$. Each of the

https://doi.org/10.1017/S0269964824000147 Published online by Cambridge University Press

Table 1. *Numerical results under different parameter settings.*

Setting	k_1	θ_1	σ_{x1}	k_2	θ_2	σ_{x2}
S0	2	1.5	1.6	0.2	0.5	0.2
	2.00 ± 0.03	1.50 ± 0.03	1.60 ± 0.02	0.20 ± 0.07	0.50 ± 0.03	0.20 ± 0.05
S1	3	1.5	1.6	0.2	0.5	0.2
	3.00 ± 0.07	1.50 ± 0.02	1.60 ± 0.02	0.20 ± 0.04	0.50 ± 0.02	0.20 ± 0.02
S2	2	3	1.6	0.2	1	0.2
	2.00 ± 0.03	3.00 ± 0.05	1.60 ± 0.02	0.20 ± 0.06	1.00 ± 0.05	0.20 ± 0.04
S3	2	1.5	0.8	0.2	0.5	0.1
	2.00 ± 0.03	1.50 ± 0.03	0.80 ± 0.01	0.20 ± 0.06	0.50 ± 0.03	0.10 ± 0.02

Table 2. *Asymptotic behavior as* N *increases.*

	k_1	θ_1	σ_{x1}	k_2	θ_2	σ_{x2}
N	2	1.5	1.6	0.2	0.5	0.2
100K	2.04 ± 0.09	1.47 ± 0.08	1.62 ± 0.06	0.24 ± 0.16	0.53 ± 0.08	0.23 ± 0.12
400K	2.01 ± 0.05	1.49 ± 0.04	1.61 ± 0.03	0.21 ± 0.10	0.51 ± 0.04	0.21 ± 0.07
1600K	2.00 ± 0.02	1.50 ± 0.02	1.60 ± 0.01	0.19 ± 0.05	0.50 ± 0.02	0.20 ± 0.04
6400K	2.00 ± 0.01	1.50 ± 0.01	1.60 ± 0.01	0.20 ± 0.03	0.50 ± 0.01	0.20 ± 0.02

other three combinations differs from S0 in only one pair of parameter values: S1 increases (k_1, k_2)'s values to (3,0.2), S2 increases (θ_1, θ_2)'s values to (3,1), S3 decreases $(\sigma_{x1}, \sigma_{x2})$'s values to (0.8,0.1).

The transition distribution of $x_i(t)$ given $x_i(s)$ $(i = 1, 2)$ for $s < t$ is a noncentral chi-squared distribution multiplied with a scale factor (see [7, 12]), that is,

$$x_i(t) = \frac{\sigma_{xi}^2 [1 - e^{-k_i(t-s)}]}{4k_i} \chi_d'^2 \left(\frac{4k_i e^{-k_i(t-s)}}{\sigma_{xi}^2 [1 - e^{-k_i(t-s)}]} x_i(s) \right), \quad s < t,$$

where $\chi_d'^2(\lambda)$ represents the noncentral chi-squared random variable with d degrees of freedom and non-centrality parameter λ, and $d = 4k_i\theta_i/\sigma_{xi}^2$. Therefore, we use this transition function to generate sample observations of $x(t)$ and set the time interval, between any two points, $h = 1$. For each parameter setting, we run 400 replications with 1000K samples for each replication. The numerical results are presented in Table 1 with the format "mean ± standard deviation" based on these 400 replications (the format remains the same for all numerical results). The results illustrate our estimators are fairly accurate.

To test the effect of N on the performance of our estimators, we run the second set of experiments in which we increase N from 100K to 6400K for S0, while all other settings remain the same. The results are given in Table 2, which demonstrates as N increases the accuracy of our estimators improve approximately at the rate of $1/\sqrt{N}$.

5. Conclusion

In this paper, we consider the problem of parameter estimation for the superposition of two square-root diffusions. We first derive their moments and auto-covariances based on which we develop our MM estimators. A major contribution of our method is that we only employ low order moments and auto-covariances and find an efficient way to solve the system of moment equations which produces very good estimates. Our MM estimators are accurate and easy to implement. We also establish the central limit theorem for our estimators in which the explicit formulas for the asymptotic covariance matrix are given. Finally, we conduct numerical experiments to test and validate our method. The MM proposed in this paper can be potentially applied to other extensions, such as including jumps in the component process or the superposed process. This is a possible future research direction.

Acknowledgments. This work is supported in part by the National Nature Science Foundation of China (NSFC) under grants 72033003 and 72350710219.

Conflict of interest. The authors declare that they have no conflict of interest.

References

[1] Ait-Sahalia, Y. & Kimmel, R.L. (2010). Estimating affine multifactor term structure models using closed-form likelihood expansions. *Journal of Financial Economics* 98(1): 113–144.
[2] Alaya, B.M. & Kebaier, A. (2012). Parameter estimation for the square-root diffusions: Ergodic and nonergodic cases. *Stochastic Models* 28(4): 609–634.
[3] Barczy, M., Alaya, B.M., Kebaier, A., & Pap, G. (2018). Asymptotic properties of maximum likelihood estimator for the growth rate for a jump-type CIR process based on continuous time observations. *Stochastic Processes and their Applications* 128(4): 1135–1164.
[4] Barczy, M., Do, L., Li, Z., Pap, G., et al. (2014). Parameter estimation for a subcritical affine two factor model. *Journal of Statistical Planning and Inference* 151: 37–59.
[5] Barndorff-Nielsen, O.E. (2001). Superposition of Ornstein–Uhlenbeck type processes. *Theory of Probability and Its Applications* 45(2): 175–194.
[6] Bradley, R.C. (2005). Basic properties of strong mixing conditions. A survey and some open questions. *Probability Surveys* 2: 107–144.
[7] Broadie, M. & Kaya, Ö. (2006). Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research* 54(2): 217–231.
[8] Casella, G. & Berger, R. (2024). *Statistical inference*. Boca Raton: CRC Press.
[9] Chen, R.-R. & Scott, L. (2003). Multi-factor Cox–Ingersoll–Ross models of the term structure: Estimates and tests from a Kalman filter model. *The Journal of Real Estate Finance and Economics* 27: 143–172.
[10] Christoffersen, P., Dorion, C., Jacobs, K., & Karoui, L. (2014). Nonlinear Kalman filtering in affine term structure models. *Management Science* 60(9): 2248–2268.
[11] Cleur, E.M. (2001). Maximum likelihood estimates of a class of one-dimensional stochastic differential equation models from discrete data. *Journal of Time Series Analysis* 22(5): 505–515.
[12] Cox, J.C., Ingersoll, Jr., J.E., & Ross, S.A. (1985). A theory of the term structure of interest rates. *Econometrica* 53: 385–407.
[13] Dehtiar, O., Mishura, Y., & Ralchenko, K. (2022). Two methods of estimation of the drift parameters of the Cox–Ingersoll–Ross process: Continuous observations. *Communications in Statistics-Theory and Methods* 51(19): 6818–6833.
[14] Duffie, D., Pan, J., & Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica* 68(6): 1343–1376.
[15] Genon-Catalot, V., Jeantheau, T., & Larédo, C. (2000). Stochastic volatility models as hidden Markov models and statistical applications. *Bernoulli* 6(6): 1051–1079.
[16] Geyer, A.L. and Pichler, S. (1999). A state-space approach to estimate and test multifactor Cox–Ingersoll–Ross models of the term structure. *Journal of Financial Research* 22(1): 107–130.
[17] Heston, S.L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies* 6(2): 327–343.
[18] Ibragimov, I.A. (1962). Some limit theorems for stationary processes. *Theory of Probability and Its Applications* 7(4): 349–382.
[19] Jenkins, M.A. & Traub, J.F. (1972). Algorithm 419: Zeros of a complex polynomial [C2]. *Communications of the ACM* 15(2): 97–99.
[20] Li, Z. & Ma, C. (2015). Asymptotic properties of estimators in a stable Cox–Ingersoll–Ross model. *Stochastic Processes and their Applications* 125(8): 3196–3233.
[21] Mishura, Y., Ralchenko, K., & Dehtiar, O. (2022). Parameter estimation in CKLS model by continuous observations. *Statistics and Probability Letters* 184: 109391.
[22] Overbeck, L. & Rydén, T. (1997). Estimation in the Cox–Ingersoll–Ross model. *Econometric Theory* 13(3): 430–461.
[23] Peng, Y., Fu, M.C., & Hu, J.-Q. (2016). Gradient-based simulated maximum likelihood estimation for stochastic volatility models using characteristic functions. *Quantitative Finance* 16(9): 1393–1411.
[24] Peng, Y.-J., Fu, M.C., & Hu, J.-Q. (2014). Gradient-based simulated maximum likelihood estimation for Lévy-driven Ornstein–Uhlenbeck stochastic volatility models. *Quantitative Finance* 14(8): 1399–1414.
[25] R Core Team. (2022). R: A Language and Environment for Statistical Computing : https://www.R-project.org/ Accessed 10 10 2024.
[26] Van der Vaart, A.W. (2000). *Asymptotic statistics*, vol. 3. Cambridge, United Kindom: Cambridge University Press.

[27] Wu, Y., Hu, J., & Yang, X. (2022). Moment estimators for parameters of Lévy-driven Ornstein–Uhlenbeck processes. *Journal of Time Series Analysis* 43(4): 610–639.
[28] Wu, Y., Hu, J., & Zhang, X. (2019). Moment estimators for the parameters of Ornstein–Uhlenbeck processes driven by compound Poisson processes. *Discrete Event Dynamic Systems* 29: 57–77.
[29] Wu, Y., Yang, X., & Hu, J.-Q. (2023). Method of moments estimation for affine stochastic volatility models. Working paper, Fudan University.
[30] Yang, X., Wu, Y., Zheng, Z., & Hu, J.-Q. (2021). Method of moments estimation for Lévy-driven Ornstein–Uhlenbeck stochastic volatility models. *Probability in the Engineering and Informational Sciences* 35(4): 975–1004.

Appendix.

Asymptotic covariance matrix entries calculation

With slight abuse of notation, we use x_n to denote either x_{1n} or x_{2n} and $\bar{x}_n$ for either $\bar{x}_{1n}$ or $\bar{x}_{1n}$ to simplify the notations, and similarly, use k, θ, σ for either $k_1, \theta_1, \sigma_{x1}$ or $k_2, \theta_2, \sigma_{x2}$. In order to calculate the asymptotic variances and covariances, we need the following approximate equations (we also call them as the decay equations), $\forall n \geq 0$:

$$\bar{x}_n \approx e^{-nkh}\bar{x}_0, \tag{A.1}$$

$$\bar{x}_n^2 \approx e^{-2nkh}\bar{x}_0^2 + (e^{-nkh} - e^{-2nkh})\bar{x}_0\frac{\sigma^2}{k} + c_2, \tag{A.2}$$

$$\begin{aligned}\bar{x}_n^3 \approx{}& e^{-3nkh}\bar{x}_0^3 + (e^{-2nkh} - e^{-3nkh})\bar{x}_0^2\frac{3\sigma^2}{k} + (e^{-nkh} - e^{-3nkh})\bar{x}_0\frac{3\sigma^2\theta}{2k} \\ &+ (e^{-nkh} - 2e^{-2nkh} + e^{-3nkh})\bar{x}_0\frac{3\sigma^4}{2k^2} + c_3,\end{aligned} \tag{A.3}$$

by eliminating the martingale parts, where

$$c_2 \triangleq (1 - e^{-2nkh})\frac{\sigma^2\theta}{2k}, \quad c_3 \triangleq (1 - 3e^{-2nkh} + 2e^{-3nkh})\frac{\sigma^4\theta}{2k^2}.$$

The derivation of above equations is given later.

Let us use

$$dx(t) = k(\theta - x(t))dt + \sigma\sqrt{x(t)}dw(t)$$

to denote the component process either Eq. (2.2) or (2.3), for notation simplification. Then, by introducing the following notation

$$IE_t \triangleq \int_0^t e^{ks}\sqrt{x(s)}dw(s), \quad IE_n \triangleq IE_{nh},$$

we have

$$\begin{aligned} x(t) &= e^{-kt}x(0) + \theta\left[1 - e^{-kt}\right] + \sigma e^{-kt}IE_t, \quad && \forall t \geq 0, \\ \bar{x}_n &= e^{-knh}\bar{x}_0 + \sigma e^{-knh}IE_n, && \forall n = 0, 1, 2, \ldots. \end{aligned}$$

Therefore,

$$\bar{x}_n^m = \sum_{j=0}^{m} C_m^j \left(e^{-knh}\bar{x}_0\right)^j \left(\sigma e^{-knh}IE_n\right)^{m-j}, m = 1, 2, 3, 4.$$

https://doi.org/10.1017/S0269964824000147 Published online by Cambridge University Press

Taking $\bar{x}_n^2$ as an example, we show how to obtain its approximation. By Itô calculus,

$$dIE_t^2 = 2IE_t e^{kt}\sqrt{x(t)}dw(t) + e^{2kt}x(t)dt,$$
$$IE_t^2 = 2\int_0^t IE_s e^{ks}\sqrt{x(s)}dw(s) + \int_0^t e^{2ks}x(s)ds.$$

Note that

$$E\left[\int_0^t IE_s e^{ks}\sqrt{x(s)}dw(s)\right] = 0,$$

thus, by deleting this term, we have following approximation

$$IE_t^2 \approx \int_0^t e^{2ks}x(s)ds.$$

Furthermore,

$$\int_0^t e^{2ks}x(s)ds = \frac{e^{kt}-1}{k}\bar{x}_0 + \frac{e^{2kt}-1}{2k}\theta + \int_0^t \sigma e^{ks} IE_s ds$$
$$\approx \frac{e^{kt}-1}{k}\bar{x}_0 + \frac{e^{2kt}-1}{2k}\theta,$$

by deleting the third term in the right hand of the first equation because its expectation is also zero, that is,

$$E\left[\int_0^t \sigma e^{ks} IE_s ds\right] = 0.$$

Letting $t = nh$, we have

$$IE_n^2 \approx \frac{e^{knh}-1}{k}\bar{x}_0 + \frac{e^{2knh}-1}{2k}\theta.$$

Therefore,

$$\bar{x}_n^2 \approx e^{-2nkh}\bar{x}_0^2 + \left(e^{-nkh} - e^{-2nkh}\right)\bar{x}_0\frac{\sigma^2}{k} + \left(1 - e^{-2nkh}\right)\frac{\sigma^2\theta}{2k}$$

by deleting all those terms having expectation 0. The results for all other cases can be derived similarly.

With these approximation formulas, we can easily compute the following auto-covariances

$$\mathrm{cov}(x_r, z_c^m), \mathrm{cov}(z_r^{m_1}, z_c^{m_2}), m, m_1, m_2 = 1, 2, 3, 4,$$

for example, $\forall c \geq r$,

$$\mathrm{cov}(x_r, z_c^1) = e^{-2(c-r)kh}\mathrm{cov}(x_r, \bar{x}_r^2) + (e^{-(c-r)kh} - e^{-2(c-r)kh})\frac{\sigma^2}{k}\mathrm{cov}(x_r, \bar{x}_r),$$

noting that $z_c^1 = \bar{x}_c^2$ by definition.

By omitting the tedious intermediate steps, we present the results in following two subsections.

On-diagonal entries

$$
\begin{aligned}
c_{11,1} &\triangleq (1+2e_{11})\mathrm{var}(x_{1n}) = (1+2e_{11})\frac{\theta_1\sigma_{x1}^2}{2k_1},\\
c_{22,1} &\triangleq (1+2e_{12})\mathrm{var}(\bar{x}_{1n}^2) + 2(e_{11}-e_{12})\frac{\sigma_{x1}^2}{k_1}\mathrm{cm}_3[x_{1n}]\\
&\quad + 2(1+2e_{1,1})\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}),\\
c_{33,1} &\triangleq (1+2e_{13})\mathrm{var}(\bar{x}_{1n}^3) + 2(e_{12}-e_{13})\frac{3\sigma_{x1}^2}{k_1}\mathrm{cov}(\bar{x}_{1n}^3,\bar{x}_{1n}^2)\\
&\quad + 2\left[(e_{11}-e_{13})\frac{3\sigma_{x1}^2\theta_1}{2k_1} + (e_{11}-2e12+e_{13})\frac{3\sigma_{x1}^4}{2k_1^2}\right]\mathrm{cm}_4[x_{1n}]\\
&\quad + 3(5+2e_{11}+6e_{2,1})\mathrm{cm}_4[x_{1n}]\mathrm{var}(x_{2n}) + 18(1+e_{1,2})\mathrm{cm}_3[x_{1n}]\mathrm{cm}_3[x_{2n}]\\
&\quad + 18\left[(e_{1,1}-e_{2,1})\frac{\sigma_{x1}^2}{k_1} + (e_{1,1}-e_{1,2})\right]\mathrm{cm}_3[x_{1n}]\mathrm{var}(x_{2n})\\
&\quad + 18(e_{21}-e_{2,1})\mathrm{var}^2(x_{1n})\mathrm{var}(x_{2n})\\
&\quad + 6e_{23}\mathrm{var}(x_{1n})\mathrm{cm}_4[x_{2n}] + 6(e_{22}-e_{23})\frac{3\sigma_{x2}^2}{k_2}\mathrm{var}(x_{n1})\mathrm{cm}_3[x_{2n}]\\
&\quad + 6\left[(e_{21}-e_{23})\frac{3\sigma_{x2}^2\theta_2}{2k_2} + (e_{21}-2e_{22}+e_{23})\frac{3\sigma_{x2}^4}{2k_2^2}\right]\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}),\\
c_{44,1} &\triangleq \mathrm{var}(\bar{x}_{1n}\bar{x}_{1(n+1)}) + 2e_{12}e^{k_1h}\mathrm{cov}(\bar{x}_{1n}\bar{x}_{1(n+1)},\bar{x}_{1(n+1)}^2)\\
&\quad + 2(e_{11}-e_{12}e^{k_1h})\frac{\sigma_{x1}^2}{k_1}\mathrm{cov}(\bar{x}_{1n}\bar{x}_{1(n+1)},\bar{x}_{1(n+1)})\\
&\quad + \left[1+2e_{1,1}+e^{-(k_1+k_2)h}\right]\mathrm{var}(x_{1n})\mathrm{var}(x_{2n})\\
&\quad + \left[\left(e^{-2k_1h}+e^{-2k_1h}e_{1,1}\right) + \left(e^{-2k_2h}+e^{-2k_2h}e_{1,1}\right)\right]\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}),\\
c_{55,1} &\triangleq \mathrm{var}(\bar{x}_{1n}\bar{x}_{1(n+2)}) + 2\mathrm{cov}(\bar{x}_{1n}\bar{x}_{1(n+2)},\bar{x}_{1(n+1)}\bar{x}_{1(n+3)})\\
&\quad + 2e_{12}\mathrm{cov}(\bar{x}_{1n}\bar{x}_{1(n+2)},\bar{x}_{1(n+2)}^2)\\
&\quad + 2(e_{11}e^{-k_1h}-e_{12})\frac{\sigma_{x1}^2}{k_1}\mathrm{cov}(\bar{x}_{1n}\bar{x}_{1(n+2)},\bar{x}_{1(n+2)})\\
&\quad + \left[1+2e_{1,1}+e^{-2(k_1+k_2)h}\right]\mathrm{var}(x_{1n})\mathrm{var}(x_{2n})\\
&\quad + \left[e^{-(3k_1+k_2)h}+e^{-4k_1h}+e^{-4k_1h}e_{1,1}\right]\mathrm{var}(x_{1n})\mathrm{var}(x_{2n})\\
&\quad + \left[e^{-(k_1+3k_2)h}+e^{-4k_2h}+e^{-4k_2h}e_{1,1}\right]\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}),\\
c_{66,1} &\triangleq \mathrm{var}(\bar{x}_{1n}\bar{x}_{1(n+3)}) + 2\mathrm{cov}(\bar{x}_{1n}\bar{x}_{1(n+3)},\bar{x}_{1(n+1)}\bar{x}_{1(n+4)})\\
&\quad + 2\mathrm{cov}(\bar{x}_{1n}\bar{x}_{1(n+3)},\bar{x}_{1(n+2)}\bar{x}_{1(n+5)}) + 2e_{12}e^{-k_1h}\mathrm{cov}(\bar{x}_{1n}\bar{x}_{1(n+3)},\bar{x}_{1(n+3)}^2)\\
&\quad + 2(e_{11}e^{-2k_1h}-e_{12}e^{-k_1h})\frac{\sigma_{x1}^2}{k_1}\mathrm{cov}(\bar{x}_{1n}\bar{x}_{1(n+3)},\bar{x}_{1(n+3)})\\
&\quad + \left[1+2e_{1,1}+e^{-3(k_1+k_2)h}\right]\mathrm{var}(x_{1n})\mathrm{var}(x_{2n})\\
&\quad + \left[e^{-(4k_1+2k_2)h}+e^{-(5k_1+k_2)h}+e^{-6k_1h}+e^{-6k_1h}e_{1,1}\right]\mathrm{var}(x_{1n})\mathrm{var}(x_{2n})\\
&\quad + \left[e^{-(2k_1+4k_2)h}+e^{-(k_1+5k_2)h}+e^{-6k_2h}+e^{-6k_2h}e_{1,1}\right]\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}).
\end{aligned}
$$

https://doi.org/10.1017/S0269964824000147 Published online by Cambridge University Press

The terms $\mathrm{cm}_3[x_{1n}], \mathrm{cm}_4[x_{1n}], \mathrm{var}(\bar{x}_{1n}^3)$ and so on are easy to compute with Eq. (2.5). It is also straightforward to compute terms like $\mathrm{cov}(\bar{x}_{1n}\bar{x}_{1(n+3)}, \bar{x}_{1(n+1)}\bar{x}_{1(n+4)})$ by employing Eqs. (A.1)–(A.3). We omit their explicit formulas here.

Terms $c_{nn,2}$ ($n = 1, \ldots, 6$) are defined similarly by replacing e_{1j} ($j = 1, \ldots, 4$) in $c_{nn,1}$ with e_{2j}, x_{1n} with x_{2n}, and $k_1, \theta_1, \sigma_{x1}$ with $k_2, \theta_2, \sigma_{x2}$, respectively, and vice versa.

Off-diagonal entries

Definitions of $c_{12,1}, c_{13,1}, c_{14,1}, c_{15,1}, c_{16,1}$ are given as following:

$$
\begin{aligned}
c_{12,1} &\triangleq (1 + e_{11} + e_{12})\mathrm{cm}_3[x_{1n}] + (e_{11} - e_{12})\frac{\sigma_{x1}^2}{k_1}\mathrm{var}(x_{1n}),\\
c_{13,1} &\triangleq (1 + e_{11} + e_{13})\mathrm{cm}_4[x_{1n}] + (e_{12} - e_{13})\frac{3\sigma_{x1}^2}{k_1}\mathrm{cm}_3[x_{1n}]\\
&\quad + \left[(e_{11} - e_{13})\frac{3\sigma_{x1}^2\theta_1}{2k_1} + (e_{11} - 2e_{12} + e_{13})\frac{3\sigma_{x1}^4}{2k_1^2}\right]\mathrm{var}(x_{1n})\\
&\quad + 3(1 + 2e_{11})\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}),\\
c_{14,1} &\triangleq (1 + e_{11} + e_{12})e^{-k_1 h}\mathrm{cm}_3[x_{1n}] + (e_{11} - e_{12})e^{-k_1 h}\mathrm{var}(x_{1n})\\
&\quad + e_{11}(1 - e^{-k_1 h})\frac{\sigma_{x1}^2}{k_1}\mathrm{var}(x_{1n}),\\
c_{15,1} &\triangleq \left[1 + e_{12} + (1 + e_{11})e^{-k_1 h}\right]e^{-2k_1 h}\mathrm{cm}_3[x_{1n}] + (e_{11} - e_{12})e^{-2k_1 h}\mathrm{var}(x_{1n})\\
&\quad + \left[e_{11}(e^{-k_1 h} - e^{-3k_1 h}) + e^{-2k_1 h} - e^{-3k_1 h})\frac{\sigma_{x1}^2}{k_1}\right]\mathrm{var}(x_{1n}),
\end{aligned}
$$

$$
\begin{aligned}
c_{16,1} &\triangleq (1 + e_{12})e^{-3k_1 h}\mathrm{cov}(\bar{x}_{1n}, \bar{x}_{1n}^2) + (e_{11} - e_{12})e^{-3k_1 h}\frac{\sigma_{x1}^2}{k_1}\mathrm{var}(\bar{x}_{1n})\\
&\quad + \mathrm{cov}(\bar{x}_{1(n+1)}, \bar{x}_{1n}\bar{x}_{1(n+3)}) + \mathrm{cov}(\bar{x}_{1(n+2)}, \bar{x}_{1n}\bar{x}_{1(n+3)})\\
&\quad + (1 + e_{11})\mathrm{cov}(\bar{x}_{1(n+3)}, \bar{x}_{1n}\bar{x}_{1(n+3)}).
\end{aligned}
$$

Again, terms $c_{1n,2}$ ($n = 2, \ldots, 6$) are defined similarly by substituting e_{1j} ($j = 1, \ldots, 4$) in $c_{1n,1}$ with e_{2j}, x_{1n} with x_{2n}, x_{2n} with x_{1n}, and $k_1, \theta_1, \sigma_{v1}$ with $k_2, \theta_2, \sigma_{v2}$, respectively, and vice versa. The other terms $c_{lm,2}$ ($l \neq 1, l \neq m$) are defined similarly according to $c_{lm,1}$ in the following formulas.

$$
\begin{aligned}
c_{23,1} &\triangleq (1 + e_{12} + e_{13})\mathrm{cov}(\bar{x}_{1n}^2, \bar{x}_{1n}^3) + (e_{11} - e_{12})\frac{\sigma_{x1}^2}{k_1}\mathrm{cm}_4[x_{1n}]\\
&\quad + (e_{12} - e_{13})\frac{3\sigma_{x1}^2}{k_1}\mathrm{cov}(\bar{x}_{1n}^2, \bar{x}_{1n}^2)\\
&\quad + \left[(e_{11} - e_{13})\frac{3\sigma_{x1}^2\theta_1}{2k_1} + (e_{11} - 2e_{12} + e_{13})\frac{3\sigma_{x1}^4}{2k_1^2}\right]\mathrm{cm}_3[x_{1n}]\\
&\quad + (9 + 3e_{11} + 3e_{12} + 6e_{1,1} + 6e_{2,1})\mathrm{cm}_3[x_{1n}]\mathrm{var}(x_{2n})\\
&\quad + (3e_{11} - 3e_{12} + 6e_{1,1} - 6e_{2,1})\frac{\sigma_{x1}^2}{k_1}\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}),
\end{aligned}
$$

https://doi.org/10.1017/S0269964824000147 Published online by Cambridge University Press

$$
\begin{aligned}
c_{24,1} \triangleq\ & (1+e_{12})e^{-k_1h}\mathrm{var}(\bar{x}_{1n}^2) + (e_{11}-e_{12})\frac{\sigma_{x1}^2}{k_1}e^{-k_1h}\mathrm{cm}_3[x_{1n}] \\
& + e_{11}e^{k_1h}\mathrm{cov}(\bar{x}_{1(n+1)}^2, \bar{x}_{1n}\bar{x}_{1(n+1)}) \\
& + (e_{11}e^{k_1h} - e_{12}e^{2k_1h})\frac{\sigma_{x1}^2}{k_1}E[\bar{x}_{1n}\bar{x}_{1(n+1)}^2] \\
& + 2(e_{1,1}e^{-k_2h} + e^{-k_2h} + e_{1,1}e^{k_2h})\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}), \\
c_{25,1} \triangleq\ & (1+e_{12})e^{-2k_1h}\mathrm{var}(\bar{x}_{1n}^2) + (e_{11}-e_{12})e^{-2k_1h}\frac{\sigma_{x1}^2}{k_1}\mathrm{cm}_3[x_{1n}] \\
& + e^{-k_1h}\mathrm{cov}(\bar{x}_{1(n+1)}^2, \bar{x}_{1n}\bar{x}_{1(n+1)}) \\
& + e_{11}e^{k_1h}\mathrm{cov}(\bar{x}_{1(n+2)}^2, \bar{x}_{1n}\bar{x}_{1(n+2)}) + (e_{11}e^{k_1h} - e_{12}e^{2k_1h})\frac{\sigma_{x1}^2}{k_1}E[\bar{x}_{1n}\bar{x}_{1(n+2)}^2] \\
& + 2(e_{1,1}e^{-2k_2h} + e^{-2k_2h} + e^{-(k_1+k_2)h} + e_{1,1}e^{2k_2h})\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}), \\
c_{26,1} \triangleq\ & (1+e_{12})e^{-3k_1h}\mathrm{var}(\bar{x}_{1n}^2) + (e_{11}-e_{12})e^{-3k_1h}\frac{\sigma_{x1}^2}{k_1}\mathrm{cm}_3[x_{1n}] \\
& + \mathrm{cov}(\bar{x}_{1(n+1)}^2, \bar{x}_{1n}\bar{x}_{1(n+3)}) + \mathrm{cov}(\bar{x}_{1(n+2)}^2, \bar{x}_{1n}\bar{x}_{1(n+3)}) \\
& + e_{12}e^{2k_1h}\mathrm{cov}(\bar{x}_{1(n+3)}^2, \bar{x}_{1n}\bar{x}_{1(n+3)}) + (e_{11}e^{k_1h} - e_{12}e^{2k_1h})\frac{\sigma_{x1}^2}{k_1}E[\bar{x}_{1n}\bar{x}_{1(n+3)}^2] \\
& + 2\left[(1+e_{1,1})e^{-3k_2h} + e^{-(k_1+2k_2)h} + e^{-(2k_1+k_2)h}\right]\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}) \\
& + 2e_{1,1}e^{(k_2-2k_1)h}\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}), \\
\sigma_{34,1} \triangleq\ & (1+e_{12})e^{-k_1h}\mathrm{cov}(\bar{x}_{1n}^3, \bar{x}_{1n}^2) + (e_{11}-e_{12})e^{-k_1h}\frac{\sigma_{x1}^2}{k_1}\mathrm{cm}_4[x_{1n}] \\
& + e_{13}e^{3k_1h}\mathrm{cov}(\bar{x}_{1(n+1)}^3, \bar{x}_{1n}\bar{x}_{1(n+1)}) \\
& + (e_{12}e^{2k_1h} - e_{13}e^{3k_1h})\frac{3\sigma_{x1}^2}{k_1}\mathrm{cov}(\bar{x}_{1(n+1)}^2, \bar{x}_{1n}\bar{x}_{1(n+1)}) \\
& + (e_{11}e^{k_1h} - e_{13}e^{3k_1h})\frac{3\sigma_{x1}^2\theta_1}{2k_1}E[\bar{x}_{1n}\bar{x}_{1(n+1)}^2] \\
& + (e_{11}e^{k_1h} - 2e_{12}e^{2k_1h} + e_{13}e^{3k_1h})\frac{3\sigma_{x1}^4}{2k_1^2}E[\bar{x}_{1n}\bar{x}_{1(n+1)}^2] \\
& + 3\left[(1+e_{1,1})e^{-k_2h} + (1+e_{1,1})e^{-k_1h} + e_{2,1}(e^{k_2h} + e^{2k_1h})\right]\mathrm{cm}_3[x_{1n}]\mathrm{var}(x_{2n}) \\
& + 3\left[(e_{1,1}-e_{2,1})e^{k_2h} + (e_{1,1}e^{k_1h} - e_{2,1}e^{2k_1h})\right]\frac{\sigma_{x1}^2}{k_1}\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}) \\
& + 3(1+e_{22})e^{-k_2h}\mathrm{var}(x_{1n})\mathrm{cm}_3[x_{2n}] + 3(e_{21}-e_{22})e^{-k_2h}\frac{\sigma_{x2}^2}{k_2}\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}) \\
& + 3e_{21}e^{k_2h}\mathrm{var}(x_{1n})E[\bar{x}_{2n}\bar{x}_{2(n+1)}^2], \\
c_{35,1} \triangleq\ & (1+e_{12})e^{-2k_1h}\mathrm{cov}(\bar{x}_{1n}^3, \bar{x}_{1n}^2) + (e_{11}-e_{12})e^{-2k_1h}\frac{\sigma_{x1}^2}{k_1}\mathrm{cm}_4[x_{1n}] \\
& + e^{-k_1h}\mathrm{cov}(\bar{x}_{1(n+1)}^3, \bar{x}_{1n}\bar{x}_{1(n+1)}) + e_{13}e^{6k_1h}\mathrm{cov}(\bar{x}_{1(n+2)}^3, \bar{x}_{1n}\bar{x}_{1(n+2)}) \\
& + (e_{12}-e_{13})\frac{3\sigma_{x1}^2}{k_1}\mathrm{cov}(\bar{x}_{1(n+2)}^2, \bar{x}_{1n}\bar{x}_{1(n+2)})
\end{aligned}
$$

https://doi.org/10.1017/S0269964824000147 Published online by Cambridge University Press

$$
\begin{aligned}
&+\left[(e_{11}-e_{13})\frac{3\sigma_{x1}^2\theta_1}{2k_1}+(e_{11}-2e_{12}+e_{13})\frac{3\sigma_{x1}^4}{2k_1^2}\right]E[\bar{x}_{1n}\bar{x}_{1(n+2)}^2]\\
&+3(1+e_{1,1})e^{-2k_2h}\mathrm{cm}_3[x_{1n}]\mathrm{var}(x_{2n})+3E[\bar{x}_{1n}\bar{x}_{1(n+1)}^2]e^{-k_2h}\mathrm{var}(x_{2n})\\
&+3e_{2,1}e^{-2k_1h}e^{k_2h}\mathrm{cm}_3[x_{1n}]\mathrm{var}(x_{2n})\\
&+3\left[(1+e_{1,1})e^{-2k_1h}+(e^{-k_1h}+e_{2,1}e^{2k_1h})e^{-k_2h}\right]\mathrm{cm}_3[x_{1n}]\mathrm{var}(x_{2n})\\
&+3(1+e_{22}+e^{-k_2h}+e_{21}e^{-k_2h})e^{-2k_2h}\mathrm{var}(x_{1n})\mathrm{cm}_3[x_{2n}]\\
&+3(e_{1,1}e^{-k_1h}e^{k_2h}-e_{2,1}e^{-2k_1h}e^{k_2h})\frac{\sigma_{x1}^2}{k_1}\mathrm{var}(x_{1n})\mathrm{var}(x_{2n})\\
&+3(e_{1,1}-e_{2,1}e^{k_1h})e^{k_1h}e^{-k_2h}\frac{\sigma_{x1}^2}{k_1}\mathrm{var}(x_{1n})\mathrm{var}(x_{2n})\\
&+3\left[e_{21}-e_{22}+1-e^{-k_2h}+e_{21}(e^{k_2h}-e^{-k_2h})\right]e^{-2k_2h}\frac{\sigma_{x2}^2}{k_2}\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}),
\end{aligned}
$$

$$
\begin{aligned}
c_{36,1}\triangleq{}&(1+e_{12})e^{-3k_1h}\mathrm{cov}(\bar{x}_{1n}^3,\bar{x}_{1n}^2)+(e_{11}-e_{12})e^{-3k_1h}\frac{\sigma_{x1}^2}{k_1}\mathrm{cm}_4[x_{1n}]\\
&+\mathrm{cov}(\bar{x}_{1(n+1)}^3,\bar{x}_{1n}\bar{x}_{1(n+3)})+\mathrm{cov}(\bar{x}_{1(n+2)}^3,\bar{x}_{1n}\bar{x}_{1(n+3)})\\
&+e_{13}e^{3k_1h}\mathrm{cov}(\bar{x}_{1(n+3)}^3,\bar{x}_{1n}\bar{x}_{1(n+3)})\\
&+(e_{12}e^{2k_1h}-e_{13}e^{3k_1h})\mathrm{cov}(\bar{x}_{1(n+3)}^2,\bar{x}_{1n}\bar{x}_{1(n+3)})\\
&+(e_{11}e^{k_1h}-e_{13}e^{3k_1h})\mathrm{cov}(\bar{x}_{1(n+1)},\bar{x}_{1n}\bar{x}_{1(n+3)})\\
&+(e_{11}e^{k_1h}-2e_{12}e^{2k_1h}+e_{13}e^{3k_1h})\mathrm{cov}(\bar{x}_{1(n+1)},\bar{x}_{1n}\bar{x}_{1(n+3)})\\
&+3(1+e_{1,1})e^{-3k_2h}\mathrm{cm}_3[x_{1n}]\mathrm{var}(x_{2n})\\
&+3\mathrm{cov}(\bar{x}_{1(n+1)}^2\bar{x}_{2(n+1)},\bar{x}_{1n}\bar{x}_{2(n+3)})+3\mathrm{cov}(\bar{x}_{1(n+2)}^2\bar{x}_{2(n+2)},\bar{x}_{1n}\bar{x}_{2(n+3)})\\
&+3e_{2,1}e^{-4k_1h}\mathrm{cm}_3[x_{1n}]\mathrm{var}(x_{2n})\\
&+3(e_{1,1}e^{-2k_1h}-e_{2,1}e^{-4k_1h})\frac{\sigma_{x1}^2}{k_1}\mathrm{var}(x_{1n})\mathrm{var}(x_{2n})\\
&+3(1+e_{1,1})e^{-3k_1h}\mathrm{cm}_3[x_{1n}]\mathrm{var}(x_{2n})\\
&+3\mathrm{cov}(\bar{x}_{1(n+1)}^2\bar{x}_{2(n+1)},\bar{x}_{2n}\bar{x}_{1(n+3)})+3\mathrm{cov}(\bar{x}_{1(n+2)}^2\bar{x}_{2(n+2)},\bar{x}_{2n}\bar{x}_{1(n+3)})\\
&+3e_{2,1}e^{2k_1h}e^{-2k_2h}\mathrm{cm}_3[x_{1n}]\mathrm{var}(x_{2n})\\
&+3(e_{1,1}e^{k_1h}e^{-2k_2h}-e_{2,1}e^{2k_1h}e^{-2k_2h})\frac{\sigma_{x1}^2}{k_1}\mathrm{var}(x_{1n})\mathrm{var}(x_{2n})\\
&+3(1+e_{22})e^{-3k_2h}\mathrm{var}(x_{1n})\mathrm{cm}_3[x_{2n}]\\
&+3(e_{21}-e_{22})e^{-3k_2h}\frac{\sigma_{x2}^2}{k_2}\mathrm{var}(x_{1n})\mathrm{var}(x_{2n})\\
&+3\mathrm{var}(x_{1n})\mathrm{cov}(\bar{x}_{2(n+1)},\bar{x}_{2n}\bar{x}_{2(n+3)})+3\mathrm{var}(x_{1n})\mathrm{cov}(\bar{x}_{2(n+2)},\bar{x}_{2n}\bar{x}_{2(n+3)})\\
&+3e_{21}e^{k_2h}\mathrm{var}(x_{1n})\mathrm{cov}(\bar{x}_{2(n+3)},\bar{x}_{2n}\bar{x}_{2(n+3)}),
\end{aligned}
$$

$$
\begin{aligned}
c_{45,1}\triangleq{}&e^{-3k_1h}\mathrm{cm}_4[x_{1n}]+(e^{-2k_1h}-e^{-3k_1h})\frac{\sigma_{x1}^2}{k_1}\mathrm{cm}_3[x_{1n}]+(e^{-k_1h}-e^{-3k_1h})\mathrm{var}(x_{1n})\\
&+E[\bar{x}_{1n}\bar{x}_{1(n+1)}\bar{x}_{1(n+2)}^2]-e^{-3k_1h}\mathrm{var}^2(x_{1n})+e_{12}\mathrm{cov}(\bar{x}_{1n}\bar{x}_{1(n+1)},\bar{x}_{1(n+1)}^2)
\end{aligned}
$$

$$
\begin{aligned}
&+ (e_{11}e^{-k_1 h} - e_{12})\frac{\sigma_{x1}^2}{k_1}E[\bar{x}_{1n}\bar{x}_{1(n+1)}^2] + e_{12}e^{k_1 h}\mathrm{cov}(\bar{x}_{1(n+2)}^2, \bar{x}_{1n}\bar{x}_{1(n+2)}) \\
&+ (e_{11} - e_{12}e^{k_1 h})\frac{\sigma_{x1}^2}{k_1}E[\bar{x}_{1n}\bar{x}_{1(n+2)}^2] \\
&+ \left[(1+e_{1,1})e^{-k_2 h} + (1+e_{1,1})e^{-k_1 h}\right]\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}) \\
&+ \left[e^{-(2k_1+k_2)h} + e^{-(k_1+2k_2)h}\right]\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}) \\
&+ e_{1,1}\left[e^{(-2k_1+k_2)h} + e^{(k_1-2k_2)h}\right]\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}),
\end{aligned}
$$

$$
\begin{aligned}
c_{46,1} \triangleq\ & e^{-2k_1 h}\mathrm{var}(\bar{x}_{1n}\bar{x}_{1(n+1)}) + e_{12}e^{-k_1 h}\mathrm{cov}(\bar{x}_{1n}\bar{x}_{1(n+1)}, \bar{x}_{1(n+1)}^2) \\
&+ (e_{11}e^{k_1 h} - e_{12}e^{2k_1 h})e^{-3k_1 h}\frac{\sigma_{x1}^2}{k_1}\mathrm{cov}(\bar{x}_{1n}\bar{x}_{1(n+1)}, \bar{x}_{1(n+1)}) \\
&+ \mathrm{cov}(\bar{x}_{1(n+1)}\bar{x}_{1(n+2)}, \bar{x}_{1n}\bar{x}_{1(n+3)}) + \mathrm{cov}(\bar{x}_{1(n+2)}\bar{x}_{1(n+3)}, \bar{x}_{1n}\bar{x}_{1(n+3)}) \\
&+ e_{12}e^{k_1 h}\mathrm{cov}(\bar{x}_{1(n+3)}^2, \bar{x}_{1n}\bar{x}_{1(n+3)}) \\
&+ (e_{11}e^{k_1 h} - e_{12}e^{2k_1 h})e^{-k_1 h}\frac{\sigma_{x1}^2}{k_1}\mathrm{cov}(\bar{x}_{1(n+3)}, \bar{x}_{1n}\bar{x}_{1(n+3)}) \\
&+ (1+e_{1,1})e^{-2k_2 h}\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}) + e^{-k_1 h}e^{-k_2 h}\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}) \\
&+ e_{1,1}e^{-k_1 h}e^{k_2 h}\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}) \\
&+ e^{-3k_1 h}e^{-k_2 h}\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}) + e_{1,1}e^{-3k_1 h}e^{k_2 h}\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}) \\
&+ \left[e^{-2(k_1+k_2)h} + e^{-(k_1+3k_2)h} + e_{1,1}e^{(k_1-3k_2)h}\right]\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}),
\end{aligned}
$$

$$
\begin{aligned}
c_{56,1} \triangleq\ & e^{-k_1 h}\mathrm{var}(\bar{x}_{1n}\bar{x}_{1(n+2)}) + \mathrm{cov}(\bar{x}_{1n}\bar{x}_{1(n+2)}, \bar{x}_{1(n+1)}\bar{x}_{1(n+4)}) \\
&+ e_{12}e^{-k_1 h}\mathrm{cov}(\bar{x}_{1n}\bar{x}_{1(n+2)}, \bar{x}_{1(n+2)}^2) \\
&+ (e_{11}e^{k_1 h} - e_{12}e^{2k_1 h})e^{-3k_1 h}\frac{\sigma_{x1}^2}{k_1}\mathrm{cov}(\bar{x}_{1n}\bar{x}_{1(n+2)}, \bar{x}_{1(n+2)}) \\
&+ \mathrm{cov}(\bar{x}_{1(n+1)}\bar{x}_{1(n+3)}, \bar{x}_{1n}\bar{x}_{1(n+3)}) + \mathrm{cov}(\bar{x}_{1(n+2)}\bar{x}_{1(n+4)}, \bar{x}_{1n}\bar{x}_{1(n+3)}) \\
&+ e_{12}\mathrm{cov}(\bar{x}_{1(n+3)}^2, \bar{x}_{1n}\bar{x}_{1(n+3)}) \\
&+ (e_{11}e^{k_1 h} - e_{12}e^{2k_2 h})e^{-2k_1 h}\frac{\sigma_{x1}^2}{k_1}\mathrm{cov}(\bar{x}_{1(n+3)}, \bar{x}_{1n}\bar{x}_{1(n+3)}) \\
&+ (1+e_{1,1})e^{-k_2 h}\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}) + e_{1,1}e^{k_2 h}\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}) \\
&+ \left[e^{-3k_1 h}e^{-2k_1 h} + e^{-(3k_1+k_2)h} + e_{1,1}e^{-4k_1 h}e^{k_2 h}\right]\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}) \\
&+ \left[e^{-2k_1 h}e^{-3k_2 h} + e^{-k_1 h}e^{-4k_2 h} + e_{1,1}e^{k_1 h}e^{-4k_2 h}\right]\mathrm{var}(x_{1n})\mathrm{var}(x_{2n}).
\end{aligned}
$$

Cite this article: Wu YF. and Hu JQ. (2024). Method of moments estimation for the superposition of square-root diffusions. *Probability in the Engineering and Informational Sciences* 1–21. https://doi.org/10.1017/S0269964824000147