



# Estimating jump–diffusions using closed-form likelihood expansions



Chenxu Li<sup>a,\*</sup>, Dachuan Chen<sup>b</sup>

<sup>a</sup> Guanghua School of Management, Peking University, Beijing, 100871, PR China

<sup>b</sup> Liautaud Graduate School of Business, University of Illinois at Chicago, Chicago, 60607, USA

## ARTICLE INFO

### Article history:

Received 3 December 2013

Received in revised form

28 March 2016

Accepted 4 July 2016

Available online 18 July 2016

### JEL classification:

C13

C22

C32

G12

### Keywords:

Maximum-likelihood estimation

Jump–diffusion

Discrete observations

Transition density

Expansion

## ABSTRACT

The indispensable role of likelihood expansions in financial econometrics for continuous-time models has been established since the ground-breaking work of Aït-Sahalia (1999, 2002a, 2008). Jump–diffusions play an important role in modeling a variety of economic and financial variables. As a significant generalization of Li (2013), we propose a new closed-form expansion for transition density of Poisson-driven jump–diffusion models and its application in maximum-likelihood estimation based on discretely sampled data. Technically speaking, our expansion is obtained by perturbing paths of the underlying model; correction terms can be calculated explicitly using any symbolic software. Numerical examples and Monte Carlo evidence for illustrating the performance of density expansion and the resulting approximate MLE are provided in order to demonstrate the practical applicability of the method. Using the theoretical results in Hayashi and Ishikawa (2012), some convergence properties related to the density expansion and the approximate MLE method can be justified under some standard sufficient (but not necessary) conditions.

© 2016 Elsevier B.V. All rights reserved.

## 1. Introduction

Continuous-time jump–diffusion processes have been widely used in various fields of science and technology for providing approximations to real-world dynamics of random fluctuations involving both relatively mild diffusive evolutions and sudden discontinuities caused by significant shocks. In financial economics, jump–diffusion models were introduced in the seminal work of Merton (1976), in which asset price is modeled by a combination of the celebrated Black–Scholes–Merton model (see Black and Scholes (1973) and Merton (1973)) and a compound Poisson process. During the past few decades, they have become a natural choice for modeling financial variables.

The literature has witnessed an explosion of developments and applications of jump–diffusion models in asset pricing, risk management and portfolio consumption optimization. Various stochastic volatility models with jump were proposed and investigated in, e.g., Bates (1996), Bates (2000), Duffie et al. (2000), Pan (2002), Johannes et al. (2003), and Broadie et al. (2007). By enriching both diffusive and jump components as well as their interactions, the affine jump–diffusion models were formally proposed

in Duffie et al. (2000), which facilitate asset pricing and econometric analysis owing to their analytical tractability. For pricing various exotic options in using analytical methods, the double exponential jump–diffusion model was proposed by Kou (2002). By employing the backward induction principle based on the Hamilton–Jacobi–Bellman equations, portfolio planning problems involving jump risk were considered in, e.g., Liu et al. (2003), Pan and Liu (2003), Aït-Sahalia et al. (2009), Aït-Sahalia and Hurd (2015), and Jin and Zhang (2012). By enriching specifications of jump intensity according to the idea of Hawkes processes (see, e.g., Hawkes (1971)), self-exciting and mutual-exciting jumps are considered in, e.g., Aït-Sahalia et al. (2015), Aït-Sahalia and Hurd (2015), Errais et al. (2010), and Giesecke et al. (2011).

Econometric analysis of jump–diffusion models leads to issues that are significantly different from those typically encountered in discrete-time series analysis, e.g., the estimation of models formulated in continuous-time using data sampled at discrete-time intervals. To conduct likelihood-based inferences in this practical setting, transition densities play an important role; see, e.g., related discussions in Aït-Sahalia (2002b, 2004) and the references therein. Maximum-likelihood estimation (MLE hereafter) for jump–diffusions usually encounters challenges arising from time-consuming computation of transition densities. Closed-form expressions for transition densities cannot be obtained even for some simple jump–diffusion models,

\* Corresponding author.

E-mail addresses: [cxli@gsm.pku.edu.cn](mailto:cxli@gsm.pku.edu.cn) (C. Li), [dchen57@uic.edu](mailto:dchen57@uic.edu) (D. Chen).

e.g., a jump–diffusion mean-reverting Ornstein–Uhlenbeck model. To conduct MLE, one usually needs computationally intensive numerical methods, e.g., Monte Carlo simulation and characteristic-function-based inversion method. Even if characteristic functions of the transition distribution exist in closed-form (e.g., for the affine jump–diffusion models proposed by Duffie et al. (2000)), the Fourier inversion based density evaluation suffers from a large computational load for each parameter set searched in the numerical procedure of optimization. Given a typically large number of possible candidate parameter sets and a large number of observations in high-frequency financial datasets (see, e.g., the survey in Mykland and Zhang (2010)), this is computationally expensive (if not impractical) because of the repeated valuation of numerical Fourier inversions in the whole procedure for MLE.

Among various methods for approximating transition densities, closed-form expansions have become popular because of their fast computing time and numerical accuracy. In particular, as a result of the progressive development of modern computation technology, calculation of high-order expansions will become increasingly feasible, and thus renders arbitrary accuracy at least in principle. For diffusion models, a milestone is the ground-breaking invention of Hermite-polynomial-based density expansion and its application in MLE proposed in Ait-Sahalia (1999, 2002a, 2008), which motivated various substantial refinements and applications, see, e.g., Bakshi et al. (2006), Ait-Sahalia and Mykland (2004, 2003), Ait-Sahalia and Kimmel (2007, 2010), Egorov et al. (2003), Xiu (2014), Chang and Chen (2011), Dipietro (2001), Stramer et al. (2010), and Choi (2013, 2015a,b). Enlightened by this stream of literature, various density expansions for jump–diffusion models were proposed, see, e.g., Ait-Sahalia and Yu (2006) for the application of saddle point approximation, Yu (2007) obtained from solving for correction terms of an expansion from Kolmogorov’s forward and backward equations, Schaumburg (2001) for expanding transition density of a Levy-driven model on a related functional space, Filipović et al. (2013) for a general approximation theory in weighted Hilbert spaces for random variables, Giesecke and Schwenkler (2011) for approximating point process filters, as well as Choi (2015a) for approximating transition density function of a multivariate time-inhomogeneous jump–diffusion process in a closed-form expression.

Complementing to the existing methods, we will propose a new closed-form expansion for transition density and apply it in approximate MLE for multivariate Poisson-driven jump–diffusion models. Our method can be viewed as a significant extension of the method for diffusion models proposed in Li (2013). Because of the fundamental challenge led by adding jumps, our expansion starts from a new method of parametrization, which can be regarded as a path perturbation and is different from the small-time setting employed in Li (2013) for diffusion models. With presence of jumps, the calculation of correction terms involves various explicit computations related to both the diffusive and jump components. Following similar discussions in Li (2013) (see pp. 1351–1352), our expansion can be regarded as a jump–diffusion analogy of the celebrated Edgeworth-type expansions; see, e.g., Chapter 2 in Hall (1995) and applications to martingales in Mykland (1992, 1993). However, in contrast to the traditional Edgeworth expansions, our expansion does not require the knowledge of generally implicit moments, cumulants or characteristic function of the underlying variable, and thus it is applicable to a wide range of jump–diffusion processes.

The theoretical foundation for validity of our expansion originates in the theory of Watanabe (1987) and Yoshida (1992) for analyzing generalized Wiener functionals, as well as its theoretical generalization in Hayashi and Ishikawa (2012) for analyzing generalized Wiener–Poisson functionals, which focus on an alternative class of expansions relying on the theory of large-deviations. The

uniform convergence rate (with respect to various parameters) of our density expansion for a parameterized jump–diffusion model can be proved under some standard sufficient conditions on the drift and diffusion coefficients. This leads to convergence of the resulting approximate MLE to the true MLE; and thus, the approximate MLE inherits the asymptotic properties of the true MLE. Such theoretical results will be supported by numerical tests and Monte Carlo simulations for some representative examples.

The rest of this paper is organized as follows. In Section 2, we introduce the model with some technical assumptions. In Section 3, we propose the transition density expansion with closed-form correction terms of any arbitrary order. In Section 4, numerical performance of the density expansion and Monte Carlo evidence for the resulting approximate MLE are demonstrated through examples. In Section 5, we conclude the paper and outline some opportunities for future research. Technical details on explicit calculation of expansion terms are provided in Appendices A–D. In an online supplementary material, Li and Chen (2016), we document some examples of closed-form expansion formulas, proofs of the results in the appendices, detailed calculation regarding some alternative specifications of the jump-size distribution, some theoretical discussions on the validity of our density expansion and the resulting approximate MLE.

## 2. The model and basic setup

We focus on a Poisson-driven jump–diffusion model governed by the following stochastic differential equation (SDE hereafter):

$$dX(t) = \mu(X(t); \theta)dt + \sigma(X(t); \theta)dW(t) + dJ(t; \theta), X(0) = x_0 \quad (1)$$

where  $X(t)$  is a  $d$ -dimensional random vector;  $\{W(t)\}$  is a  $d$ -dimensional standard Brownian motion;  $\mu(x; \theta) = (\mu_1(x; \theta), \mu_2(x; \theta), \dots, \mu_d(x; \theta))^T$  is a  $d$ -dimensional vector-valued function and  $\sigma = (\sigma_{ij}(x; \theta))_{d \times d}$  is a  $d \times d$  matrix-valued function with an unknown parameter  $\theta$  belonging to a multidimensional open bounded set  $\Theta$ . Here,  $J(t; \theta)$  is a vector-valued jump process modeled by a compound Poisson process which can be specified as

$$J(t; \theta) \equiv (J_1(t; \theta), J_2(t; \theta), \dots, J_d(t; \theta))^T \\ := \sum_{n=1}^{N(t)} Z_n \equiv \sum_{n=1}^{N(t)} (Z_{n,1}, Z_{n,2}, \dots, Z_{n,d})^T,$$

where  $\{N(t)\}$  is a Poisson process with a constant intensity  $\lambda$ . For different integers  $n$ ,  $Z_n = (Z_{n,1}, Z_{n,2}, \dots, Z_{n,d})^T$  are i.i.d. multivariate random variables. Assuming  $\tau_1, \tau_2, \dots$  are the jump arrival times, the jump path can be expressed as a step function, i.e.,

$$J(t; \theta) = \sum_{n=1}^{\infty} \left( \sum_{i=1}^n (Z_{i,1}, Z_{i,2}, \dots, Z_{i,d})^T \right) 1_{[\tau_n, \tau_{n+1})}(t). \quad (2)$$

Let  $E \subset \mathbb{R}^d$  denote the state space of  $X$ .

We note that various popular jump–diffusion-based asset pricing models (see, e.g., Merton (1976), Kou (2002), Bates (2000), Duffie et al. (2000), and Broadie et al. (2007)) take or can be easily transformed into the form of (1). This model relaxes the condition on linear drift and diffusion of the affine jump–diffusion model proposed in Duffie et al. (2000). By assuming the intensity of  $\{N(t)\}$  to be a constant, the existence and uniqueness of the solution to model (1) can be guaranteed under some technical conditions, see, e.g., discussions in Yu (2007). Besides, this assumption is supported by various empirical evidences, see, e.g., Bates (2000), Andersen et al. (2002), and Chernov et al. (2003). In modeling typical financial variables using a multidimensional jump–diffusion model, the small sample problem is usually severe in the estimation of

correlations among jumps in different dimensions, since jumps are rare even with a long span (e.g., 15 or 20 years) of data, see, e.g., discussions in Johannes et al. (2003), Chernov et al. (2003) and Broadie et al. (2007) on setting and estimating the correlation between jumps in asset return and its variance under stochastic volatility with concurrent jumps models (SVCJ). As a result, from the perspective of econometric analysis, it is prevalent to assume the independence among jumps in different dimensions. Without loss of generality, we consider the following two parsimonious examples of jump-size distribution for the purpose of illustration, i.e., a normal distribution for modeling double-sided jumps and an exponential distribution for modeling single-sided jumps.

**Jump-Size Distribution 1.** The jump size  $Z_n$  has a multivariate normal distribution with mean vector  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_d)$  and covariance matrix  $\beta = \text{diag}(\beta_1^2, \beta_2^2, \dots, \beta_d^2)$ , i.e.,  $Z_n \sim \mathbf{N}(\alpha, \beta)$ .

**Jump-Size Distribution 2.**  $Z_n$  has a multivariate exponential distribution, in which  $Z_{n,j}$ 's are independent and  $Z_{n,j}$  has an exponential distribution with intensity  $\gamma_j$  for  $j = 1, 2, \dots, d$ .

Dependence among jumps in different dimensions and other specifications of the distribution can be similarly analyzed following our method in a case-by-case manner. For example, by letting some of the dimensions in the random vector  $Z_n$  be zero, we allow jumps in some but not all of the factors, see, e.g., the jump component of the stochastic volatility model with jumps in price only investigated in Bates (1996) and Duffie et al. (2000). In Sections 3 and 4 of Li and Chen (2016), we provide main techniques for performing our expansion under two alternative specifications of the jump-size distribution. Generalizations of our method for incorporating stochastic jump intensity (see, e.g., the volatility–excitation considered in Pan (2002), the self-excitation and mutual-excitation considered in, e.g., Ait-Sahalia et al. (2015), Ait-Sahalia and Hurd (2015), Errais et al. (2010), and Giesecke et al. (2011)) and even more general state-dependent jump component (see, e.g., Cinlar and Jacod (1981) and Yu (2007)) will be set as a future research project.

Before closing this section, we introduce some standard and technical assumptions, which are conventionally proposed in the study of stochastic differential equations (see, e.g., Ikeda and Watanabe (1989)). Denote by  $A(x; \theta) = \sigma(x; \theta)\sigma(x; \theta)^\top$  the diffusion matrix.

**Assumption 1.** The diffusion matrix  $A(x; \theta)$  is positive definite, i.e.  $\det A(x; \theta) > 0$ , for any  $(x, \theta) \in E \times \Theta$ .

**Assumption 2.** For each integer  $k \geq 1$ , the  $k$ th order derivatives in  $x$  of the functions  $\mu(x; \theta)$  and  $\sigma(x; \theta)$  are uniformly bounded for any  $(x, \theta) \in E \times \Theta$ .

As discussed in Section 5 of Li and Chen (2016), these assumptions provide sufficient conditions for the validity of our expansion. However, as shown momentarily in Section 4, numerical examples (e.g., the SQRJ model and the CEV-SVCJ model) suggest that the method proposed in this paper is not confined to the models strictly satisfying these sufficient (but not necessary) conditions.<sup>1</sup> Theoretical relaxation of these conditions may involve case-by-case mathematical treatment and standard approximation argument, which is beyond the scope of this paper and can be regarded as a future research topic.

<sup>1</sup> For example, as seen from the dynamics of SQRJ Model 3, the volatility function  $\sigma(x) = \sigma\sqrt{x}$  violates Assumption 2 at the point  $x = 0$ .

### 3. A closed-form expansion of transition density

#### 3.1. A general framework

By the time-homogeneity nature of jump–diffusion model (1), we denote by  $p(\Delta, x|x_0; \theta)$  its transition density corresponding to a time interval with length  $\Delta$ , i.e., the conditional density of  $X(t + \Delta)$  given  $X(t) = x_0$ :

$$\mathbb{P}(X(t + \Delta) \in dx|X(t) = x_0) = p(\Delta, x|x_0; \theta)dx. \tag{3}$$

We will propose a closed-form asymptotic expansion approximation for (3) in the following form

$$p_M(\Delta, x|x_0; \theta) = \left(\frac{1}{\sqrt{\Delta}}\right)^d \det D(x_0) \sum_{m=0}^M \Psi_m(\Delta, x|x_0; \theta),$$

where  $p_M$  denotes an expansion up to the  $M$ th order; the functions  $D(x_0)$  and  $\Psi_m(\Delta, x|x_0; \theta)$ , explicitly depending on the drift vector  $\mu$ , dispersion matrix  $\sigma$  and jump components, will be defined or calculated in what follows.

For ease of exposition in the following discussions, we drop the dependence of  $\theta$  in dynamics of the models. For computational convenience, we start from the following equivalent Stratonovich form of model (1):

$$dX(t) = b(X(t))dt + \sigma(X(t)) \circ dW(t) + dJ(t), \quad X(0) = x_0, \tag{4}$$

where  $\circ$  represents the Stratonovich integral and the new drift vector  $b(x) = (b_1(x), b_2(x), \dots, b_d(x))^\top$  satisfies that

$$b_i(x) = \mu_i(x) - \frac{1}{2} \sum_{k=1}^d \sum_{j=1}^d \sigma_{kj}(x) \frac{\partial}{\partial x_k} \sigma_{ij}(x).$$

Thus, we parameterize dynamics (4) as

$$dX^\epsilon(t) = \epsilon[b(X^\epsilon(t))dt + \sigma(X^\epsilon(t)) \circ dW(t) + dJ(t)], \quad X^\epsilon(0) = x_0, \tag{5}$$

where  $\epsilon > 0$  is an auxiliary parameter such that  $X^\epsilon(t)|_{\epsilon=1} \equiv X(t)$ . Therefore, once we obtain an expansion for the transition density of model (5)

$$p^\epsilon(\Delta, x|x_0; \theta)dx := \mathbb{P}(X^\epsilon(t + \Delta) \in dx|X^\epsilon(t) = x_0) \\ \equiv \mathbb{P}(X^\epsilon(\Delta) \in dx|X^\epsilon(0) = x_0) \tag{6}$$

as a series of  $\epsilon$ , an approximation for (3) can be directly obtained by letting  $\epsilon = 1$ .

Our main idea starts from a stochastic pathwise expansion of  $X^\epsilon(t)$  as a power series of  $\epsilon$  around  $\epsilon = 0$ . We assume that the vector  $X^\epsilon(t)$  admits the following  $M$ th order expansion

$$X^\epsilon(t) = \sum_{m=0}^M X_m(t)\epsilon^m + \mathcal{O}(\epsilon^{M+1}), \tag{7}$$

where  $X^\epsilon(t) = (X_1^\epsilon(t), X_2^\epsilon(t), \dots, X_d^\epsilon(t))^\top$  and  $X_m(t) = (X_{m,1}(t), X_{m,2}(t), \dots, X_{m,d}(t))^\top$ . Thus, the  $r$ th dimension of (7) is given by

$$X_r^\epsilon(t) = \sum_{m=0}^M X_{m,r}(t)\epsilon^m + \mathcal{O}(\epsilon^{M+1}), \quad r = 1, 2, \dots, d.$$

Without any confusion, the integer  $M$  will serve as an arbitrary order of various expansions from now on. By letting  $\epsilon = 0$  on the both sides of parameterized model (5), it is straightforward to obtain that  $dX_0(t) = 0$ . Thus, because of the initial condition  $X^\epsilon(0) = x_0$ , we have the leading term as  $X_0(t) \equiv x_0$ . By taking first-order derivatives with respect to  $\epsilon$  on the both sides of (5) and evaluating them at  $\epsilon = 0$ , we obtain the first-order correction term as

$$X_1(t) = b(x_0)t + \sigma(x_0)W(t) + J(t). \tag{8}$$

A systematical method for explicitly obtaining higher-order correction terms will be given in Section 3.3.1.

We begin by representing the transition density of  $\{X^\epsilon(t)\}$  as a conditional expectation of Dirac Delta function (see, e.g., Kanwal (2004)) acting on  $X^\epsilon(\Delta) - x$ , i.e.,

$$p^\epsilon(\Delta, x|x_0; \theta) := \mathbb{E}[\delta(X^\epsilon(\Delta) - x)|X^\epsilon(0) = x_0].$$

In the following expositions, the initial condition of  $X(0) = x_0$  will be omitted. To guarantee the convergence, our expansion starts from a standardization of  $X^\epsilon(\Delta)$  into

$$Y^\epsilon(\Delta) := \frac{D(x_0) X^\epsilon(\Delta) - x_0}{\sqrt{\Delta} \epsilon}, \tag{9}$$

where  $D(x)$  is a diagonal matrix defined by

$$D(x) := \text{diag} \left( \left( \sum_{j=1}^d \sigma_{1j}^2(x) \right)^{-\frac{1}{2}}, \left( \sum_{j=1}^d \sigma_{2j}^2(x) \right)^{-\frac{1}{2}}, \dots, \left( \sum_{j=1}^d \sigma_{dj}^2(x) \right)^{-\frac{1}{2}} \right). \tag{10}$$

Standardization (9) plays a similar role to that of the Lamperti transformation adopted by Ait-Sahalia (1999, 2002a, 2008), which leads to a normal distribution as the leading-order term and accurate higher-order correction terms for constructing the density expansion.

Assuming the pathwise expansion for  $Y^\epsilon(\Delta)$  as

$$Y^\epsilon(\Delta) = \sum_{m=0}^M Y_m(\Delta) \epsilon^m + \mathcal{O}(\epsilon^{M+1}), \tag{11}$$

it follows from (7) and (9) that

$$Y_m(\Delta) = \frac{D(x_0)}{\sqrt{\Delta}} X_{m+1}(\Delta), \text{ for } m = 0, 1, 2, \dots \tag{12}$$

The corresponding elementwise form of (11) and (12) satisfies that

$$Y_r^\epsilon(\Delta) = \sum_{m=0}^M Y_{m,r}(\Delta) \epsilon^m + \mathcal{O}(\epsilon^{M+1}),$$

where  $Y_{m,r}(\Delta) = \frac{D_{rr}(x_0)}{\sqrt{\Delta}} X_{m+1,r}(\Delta),$  \tag{13}

where  $D_{rr}(x_0)$  refers to the  $r$ th diagonal element of  $D(x_0)$ , i.e.,  $D_{rr}(x_0) = \left( \sum_{j=1}^d \sigma_{rj}^2(x) \right)^{-\frac{1}{2}}$ . It is evident that, as  $\epsilon \rightarrow 0$ ,  $Y^\epsilon(\Delta)$  converges to

$$Y_0(\Delta) = \frac{B(\Delta)}{\sqrt{\Delta}} + \frac{D(x_0)}{\sqrt{\Delta}} (b(x_0)\Delta + J(\Delta)), \tag{14}$$

where

$$B(t) \equiv (B_1(t), B_2(t), \dots, B_d(t)) := D(x_0)\sigma(x_0)W(t) \tag{15}$$

is a  $d$ -dimensional correlated Brownian motion.

Simple algebra yields that

$$Y^\epsilon(\Delta) - \frac{D(x_0)}{\sqrt{\Delta}} \left( \frac{x - x_0}{\epsilon} \right) = \frac{D(x_0)}{\sqrt{\Delta} \epsilon} (X^\epsilon(\Delta) - x).$$

Thus, by the scaling property of Dirac Delta function (see, e.g., Kanwal (2004)), we obtain that

$$\begin{aligned} &\mathbb{E} \delta(X^\epsilon(\Delta) - x) \\ &= \left( \frac{1}{\sqrt{\Delta} \epsilon} \right)^d \det D(x_0) \mathbb{E} [\delta(Y^\epsilon(\Delta) - y)] \Big|_{y = \frac{D(x_0)}{\sqrt{\Delta}} \left( \frac{x - x_0}{\epsilon} \right)}. \end{aligned}$$

Heuristically speaking, by the expansion of  $Y^\epsilon(\Delta)$  and the classical rule for differentiating composition of functions, we obtain a Taylor-like expansion of  $\delta(Y^\epsilon(\Delta) - y)$  as

$$\delta(Y^\epsilon(\Delta) - y) = \sum_{m=0}^M \Phi_m(y) \epsilon^m + \mathcal{O}(\epsilon^{M+1}), \tag{16}$$

where  $\Phi_m(y)$  represents the  $m$ th expansion term. By taking expectations, it is natural to obtain that

$$\mathbb{E} [\delta(Y^\epsilon(\Delta) - y)] := \sum_{m=0}^M \Psi_m(y) \epsilon^m + \mathcal{O}(\epsilon^{M+1}),$$

where

$$\Psi_m(y) := \mathbb{E} [\Phi_m(y)].$$

Thus, the  $M$ th order expansion of the density  $p^\epsilon(\Delta, x|x_0; \theta)$  is proposed as

$$\begin{aligned} p_M^\epsilon(\Delta, x|x_0; \theta) \\ = \left( \frac{1}{\sqrt{\Delta} \epsilon} \right)^d \det D(x_0) \sum_{m=0}^M \Psi_m \left( \frac{D(x_0)}{\sqrt{\Delta}} \left( \frac{x - x_0}{\epsilon} \right) \right) \epsilon^m. \end{aligned} \tag{17}$$

By letting  $\epsilon = 1$ , we define an  $M$ th order approximation to the transition density  $p(\Delta, x|x_0; \theta)$  as

$$p_M(\Delta, x|x_0; \theta) := \left( \frac{1}{\sqrt{\Delta}} \right)^d \det D(x_0) \sum_{m=0}^M \Psi_m \left( \frac{D(x_0)}{\sqrt{\Delta}} (x - x_0) \right). \tag{18}$$

We note that  $Y_0$  is non-degenerate in the Wiener-Poisson space. Such a setting renders the validity of our expansion method, based on a generalization of the theory of Watanabe (1987) and Yoshida (1992) established in Hayashi and Ishikawa (2012). Under some technical conditions, the convergence of (16) is in the sense of distribution and Malliavin calculus and further renders the convergence of (17). In this article, we focus on the practical calculation and implementation of the density expansion; we provide the theoretical justifications of convergence in Section 5 of Li and Chen (2016).

In practice, the explicit calculation of  $\Psi_m(y)$  hinges on the total number of jump arrivals. Indeed, we deduce that

$$\Psi_m(y) = \mathbb{E} [\Phi_m(y)] = \sum_{n=0}^{\infty} \mathbb{E} [\Phi_m(y) | N(\Delta) = n] \mathbb{P}(N(\Delta) = n).$$

Thus, for  $m, n \geq 0$ , denote by

$$T_{m,n}(y) := \mathbb{E} [\Phi_m(y) | N(\Delta) = n]. \tag{19}$$

Observing

$$\mathbb{P}(N(\Delta) = n) = \exp(-\lambda \Delta) \frac{\lambda^n \Delta^n}{n!},$$

we obtain that

$$\Psi_m(y) = \sum_{n=0}^{\infty} \exp(-\lambda \Delta) \frac{\lambda^n \Delta^n}{n!} T_{m,n}(y). \tag{20}$$

In practice, to approximate the true value, such a series can be implemented by a truncation with a finite number of terms approaching to numerical stability. Thus, we define

$$\Psi_{m,N}(y) = \sum_{n=0}^N \exp(-\lambda \Delta) \frac{\lambda^n \Delta^n}{n!} T_{m,n}(y). \tag{21}$$

as an  $N$ th order approximation of  $\Psi_m(y)$ . According to the previous discussions, density expansion (17) is further approximated by the following double expansion

$$\begin{aligned}
 p_{M,N}^\epsilon(\Delta, x|x_0; \theta) &:= \left(\frac{1}{\sqrt{\Delta\epsilon}}\right)^d \det D(x_0) \sum_{m=0}^M \Psi_{m,N} \left(\frac{D(x_0)}{\sqrt{\Delta}} \left(\frac{x-x_0}{\epsilon}\right)\right) \epsilon^m \\
 &\equiv \left(\frac{1}{\sqrt{\Delta\epsilon}}\right)^d \det D(x_0) \sum_{m=0}^M \sum_{n=0}^N \exp(-\lambda \Delta) \frac{\lambda^n \Delta^n}{n!} \\
 &\quad \times T_{m,n} \left(\frac{D(x_0)}{\sqrt{\Delta}} \left(\frac{x-x_0}{\epsilon}\right)\right) \epsilon^m. \tag{22}
 \end{aligned}$$

Finally, the  $M$ th order approximation (18) of the transition density is further approximated by the following double summation

$$\begin{aligned}
 p_{M,N}(\Delta, x|x_0; \theta) &:= \left(\frac{1}{\sqrt{\Delta}}\right)^d \det D(x_0) \sum_{m=0}^M \sum_{n=0}^N \exp(-\lambda \Delta) \\
 &\quad \times \frac{\lambda^n \Delta^n}{n!} T_{m,n} \left(\frac{D(x_0)}{\sqrt{\Delta}} (x-x_0)\right). \tag{23}
 \end{aligned}$$

Theoretical discussions related to the validity of these approximations are provided in Li and Chen (2016). In what follows, we will concentrate on the explicit calculation of  $T_{m,n}(y)$  in Sections 3.2 and 3.3, an application of the density expansion in maximum-likelihood estimation in Section 3.4, and numerical examples in Section 4.

### 3.2. Calculation of the leading term $T_{0,n}(y)$

It is obvious that the leading term of pathwise expansion (16) is given by  $\Phi_0(y) = \delta(Y_0(\Delta) - y)$ . We note that

$$T_{0,n}(y) = \mathbb{E}[\delta(Y_0(\Delta) - y) | N(\Delta) = n] \tag{24}$$

is essentially a conditional density of  $Y_0$ . By plugging (14) in (24) and conditioning on the jump component  $J(\Delta) = \sum_{i=1}^n Z_i$ , we obtain that

$$\begin{aligned}
 T_{0,n}(y) &= \mathbb{E}[\mathbb{E}(\delta(Y_0(\Delta) - y) | J(\Delta), N(\Delta) = n) | N(\Delta) = n] \\
 &= \mathbb{E} \left[ \phi_{\Sigma(x_0)} \left( y - \frac{D(x_0)}{\sqrt{\Delta}} (b(x_0)\Delta + J(\Delta)) \right) | N(\Delta) = n \right], \tag{25}
 \end{aligned}$$

where  $\Sigma(x_0)$  is the correlation matrix of Brownian motion (15), i.e.,

$$\Sigma(x_0) := (\text{Corr}(B_i(t), B_j(t)))_{d \times d} = D(x_0)\sigma(x_0)\sigma(x_0)^\top D(x_0), \tag{26}$$

and  $\phi_C(y)$  denotes the probability density of a normal distribution with zero mean and covariance matrix  $C$ , i.e.,

$$\phi_C(y) := \frac{1}{(2\pi)^{\frac{d}{2}} (\det C)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}y^\top C^{-1}y\right). \tag{27}$$

The explicit calculation of (25) under the assumptions of Jump-Size Distributions 1 and 2 will be illustrated in Appendix A.

### 3.3. Calculation of high-order terms $T_{m,n}(y)$ for $m \geq 1$

#### 3.3.1. High-order pathwise expansion in (7)

We propose an iterative algorithm for obtaining any arbitrary order of expansion (7). Assume the following expansions according to  $\epsilon$

$$b(X^\epsilon(t)) := \sum_{m=0}^M b_m(t)\epsilon^m + \mathcal{O}(\epsilon^{M+1}), \tag{28a}$$

$$\sigma(X^\epsilon(t)) := \sum_{m=0}^M \sigma_m(t)\epsilon^m + \mathcal{O}(\epsilon^{M+1}), \tag{28b}$$

where elementwise forms of the correction terms are given by

$$\begin{aligned}
 b_m(t) &= (b_{m,1}(t), b_{m,2}(t), \dots, b_{m,d}(t))^\top \quad \text{and} \\
 \sigma_m(t) &= (\sigma_{m,(k,r)}(t))_{d \times d}, \quad \text{for any } m = 0, 1, 2, \dots, M.
 \end{aligned}$$

Thus, for any  $k = 1, 2, \dots, d, r = 1, 2, \dots, d$ , differentiation of composite functions  $b_k(X^\epsilon(t))$  and  $\sigma_{kr}(X^\epsilon(t))$  with respect to  $\epsilon$  yields that

$$\begin{aligned}
 b_{m,k}(t) &= \frac{1}{m!} \frac{\partial^{(m)} b_k(X^\epsilon(t))}{\partial \epsilon^m} \Big|_{\epsilon=0} \\
 &= \sum_{\substack{(\ell, (j_1, j_2, \dots, j_\ell), \\ (r_1, r_2, \dots, r_\ell)) \in \delta_m}} \frac{1}{\ell!} \frac{\partial^\ell b_k(X_0(t))}{\partial x_{r_1} \partial x_{r_2} \dots \partial x_{r_\ell}} \prod_{i=1}^\ell X_{j_i, r_i}(t), \tag{29a}
 \end{aligned}$$

$$\begin{aligned}
 \sigma_{m,(k,r)}(t) &= \frac{1}{m!} \frac{\partial^{(m)} \sigma_{kr}(X^\epsilon(t))}{\partial \epsilon^m} \Big|_{\epsilon=0} \\
 &= \sum_{\substack{(\ell, (j_1, j_2, \dots, j_\ell), \\ (r_1, r_2, \dots, r_\ell)) \in \delta_m}} \frac{1}{\ell!} \frac{\partial^\ell \sigma_{kr}(X_0(t))}{\partial x_{r_1} \partial x_{r_2} \dots \partial x_{r_\ell}} \prod_{i=1}^\ell X_{j_i, r_i}(t), \tag{29b}
 \end{aligned}$$

where the index set  $\delta_m$  is defined by

$$\begin{aligned}
 \delta_m &:= \{(\ell, \mathbf{j}(\ell), \mathbf{r}(\ell)) \mid \ell = 1, 2, \dots, \mathbf{j}(\ell) = (j_1, j_2, \dots, j_\ell) \text{ with} \\
 &\quad j_1, j_2, \dots, j_\ell \geq 1 \text{ and } j_1 + j_2 + \dots + j_\ell = m, \\
 &\quad \mathbf{r}(\ell) = (r_1, r_2, \dots, r_\ell) \text{ with } r_1, r_2, \dots, r_\ell \in \{1, 2, \dots, d\}\}. \tag{30}
 \end{aligned}$$

We note that formulas (29a) and (29b) follow from the classical differential calculus.

After plugging expansions (7), (28a) and (28b) in (5), a comparison of the coefficients of  $\epsilon^m$ , for  $m \geq 2$ , in both sides of Eq. (5) results in

$$dX_m(t) = b_{m-1}(t)dt + \sigma_{m-1}(t) \circ dW(t), \text{ for } m \geq 2.$$

According to the fact that

$$X^\epsilon(0) = \sum_{m=0}^M X_m(0)\epsilon^m + \mathcal{O}(\epsilon^{M+1}) \equiv x_0,$$

a comparison of the coefficients of each order yields that

$$X_0(0) = x_0 \text{ and } X_m(0) = 0, \text{ for all } m \geq 1. \tag{31}$$

Thus, we obtain that

$$X_m(t) = \int_0^t b_{m-1}(s)ds + \int_0^t \sigma_{m-1}(s) \circ dW(s), \text{ for } m \geq 2. \tag{32}$$

According to (29a) and (29b), all expressions involved on the right-hand side of (32) contain expansion terms of  $X(t)$  with orders at most  $m - 1$ . Therefore, iterative applications of (32) result in explicit form of  $X_m(t)$  for any  $m \geq 2$  via iterated Stratonovich integrals defined as follows.

For an arbitrary index  $\mathbf{i} = (i_1, i_2, \dots, i_l)$  with  $i_1, i_2, \dots, i_l \in \{0, 1, 2, \dots, d\}$  and an  $l$ -dimensional stochastic process  $\mathbf{f} = \{f_1(t), f_2(t), \dots, f_l(t)\}$ , we introduce an iterated Stratonovich integral

$$\begin{aligned}
 S_{\mathbf{i}, \mathbf{f}}(t) &:= \int_0^t \int_0^{t_1} \dots \int_0^{t_{l-1}} f_l(t_l) \circ dW_{i_l}(t_l) \dots \\
 &\quad f_2(t_2) \circ dW_{i_2}(t_2) f_1(t_1) \circ dW_{i_1}(t_1), \tag{33}
 \end{aligned}$$

which is iteratively defined from inside to outside according to the definition of Stratonovich integrals (see e.g., Section 3.3 in Karatzas and Shreve (1991)). Here, we let  $W_0(t) = t$ . From iteration (32), it is evident that the correction term  $X_{j_i+1, r_i}$  can be expressed by iterations and multiplications of Stratonovich integrals. The integrands involve step function (2) created by

jump arrivals. Without loss of generality and for the purpose of illustration, we provide the first three closed-form pathwise expansion terms of a one-dimensional jump–diffusion driven by a one-dimensional Brownian motion ( $d = 1$  in the general setting (1)). In the following examples, we denote by  $W(t)$  the driving one-dimensional Brownian motion. The first order (8) can be rewritten using iterated Stratonovich integrals as

$$X_1(t) \equiv b(x_0)S_{(0),(1)}(t) + \sigma(x_0)S_{(1),(1)}(t) + J(t). \tag{34}$$

For simplicity, denote by

$$b^{(k)}(x) = \frac{\partial^k b(x)}{\partial x^k} \text{ and } \sigma^{(k)}(x) = \frac{\partial^k \sigma(x)}{\partial x^k},$$

for integer  $k \geq 1$ . According to iteration (32), the second order term satisfies

$$X_2(t) = b^{(1)}(x_0) \int_0^t X_1(s) ds + \sigma^{(1)}(x_0) \int_0^t X_1(s) \circ dW(s). \tag{35}$$

By plugging (34) in (35),  $X_2(t)$  can be written as a linear combination of iterated Stratonovich integrals

$$\begin{aligned} X_2(t) = & b^{(1)}(x_0)b(x_0)S_{(0,0),(1,1)}(t) + b^{(1)}(x_0)\sigma(x_0)S_{(0,1),(1,1)}(t) \\ & + \sigma^{(1)}(x_0)b(x_0)S_{(1,0),(1,1)}(t) + \sigma^{(1)}(x_0)\sigma(x_0)S_{(1,1),(1,1)}(t) \\ & + b^{(1)}(x_0)S_{(0),(J(t))}(t) + \sigma^{(1)}(x_0)S_{(1),(J(t))}(t). \end{aligned} \tag{36}$$

Similarly, the third order satisfies

$$\begin{aligned} X_3(t) = & b^{(1)}(x_0) \int_0^t X_2(s) ds + \frac{1}{2}b^{(2)}(x_0) \int_0^t X_1(s)^2 ds \\ & + \sigma^{(1)}(x_0) \int_0^t X_2(s) \circ dW(s) \\ & + \frac{1}{2}\sigma^{(2)}(x_0) \int_0^t X_1(s)^2 \circ dW(s). \end{aligned} \tag{37}$$

Employing (34), we observe that

$$\begin{aligned} X_1(t)^2 = & b(x_0)^2S_{(0),(1)}(t)^2 + 2b(x_0)\sigma(x_0)S_{(0),(1)}(t)S_{(1),(1)}(t) \\ & + \sigma(x_0)^2S_{(1),(1)}(t)^2 + 2b(x_0)J(t)S_{(0),(1)}(t) \\ & + 2\sigma(x_0)J(t)S_{(1),(1)}(t) + J^2(t). \end{aligned} \tag{38}$$

Thus,  $X_3(t)$  involves multiplications and iterations of Stratonovich integrals with integrands involving jump path (2). It is evident from iteration (32) that such a pattern is inductively inherited by any order of the expansion.

### 3.3.2. Calculation of $T_{m,n}(y)$ for $m \geq 1$

Similar to expressions (29a) and (29b), the  $m$ th order correction term for expansion (16) follows from differentiating the composite function  $\delta(Y^\epsilon(\Delta) - y)$ , that is

$$\begin{aligned} \Phi_m(y) = & \sum_{\substack{(\ell, \mathbf{j}(\ell), \mathbf{r}(\ell)) \\ = (\ell, (j_1, j_2, \dots, j_\ell), (r_1, r_2, \dots, r_\ell)) \in \mathcal{S}_m}} \frac{1}{\ell!} \left( \frac{1}{\sqrt{\Delta}} \right)^\ell \prod_{i=1}^{\ell} D_{r_i r_i}(x_0) \\ & \times \frac{\partial^{(\ell)} \delta(Y_0(\Delta) - y)}{\partial x_{r_1} \partial x_{r_2} \dots \partial x_{r_\ell}} \prod_{i=1}^{\ell} X_{j_i+1, r_i}(\Delta). \end{aligned} \tag{39}$$

Our goal is to explicitly calculate (19).

Denote by  $\{\mathcal{F}(t)\}$  the filtration generated by the jump process, i.e.,  $\{\mathcal{F}(t)\} = \sigma(J(s), s \leq t)$ . For  $\mathbf{j}(\ell) = (j_1, j_2, \dots, j_\ell)$  and  $\mathbf{r}(\ell) = (r_1, r_2, \dots, r_\ell)$ , we define

$$\begin{aligned} P_{n, (\ell, \mathbf{j}(\ell), \mathbf{r}(\ell))}(w) \\ := \mathbb{E} \left( \prod_{i=1}^{\ell} X_{j_i+1, r_i}(\Delta) | W(\Delta) = w, N(\Delta) = n, \mathcal{F}(\Delta) \right), \end{aligned} \tag{40}$$

which will be calculated in the following sections as a polynomial in  $w$  with coefficients involving polynomials of the jump arrival times  $\tau_1, \tau_2, \dots, \tau_n$  as well as jump amplitudes  $Z_1, Z_2, \dots, Z_n$ . To systematically express some derivatives involved in our expansion terms, we introduce the following differential operator

$$\mathcal{D}_i u(z) := \frac{\partial u(z)}{\partial z_i} - u(z)(\Sigma(x_0)^{-1}z)_i, \tag{41}$$

for any index  $i \in \{1, 2, \dots, d\}$  and differentiable function  $u(z)$  with  $z \in \mathbb{R}^d$ , where  $(\Sigma(x_0)^{-1}z)_i$  denotes the  $i$ th element of the vector  $\Sigma(x_0)^{-1}z$ .

We propose the following theorem for explicitly calculating  $T_{m,n}(y)$ .

**Theorem 1.** For any integer  $m \geq 1$ , the correction term  $T_{m,n}(y)$  in (19) admits the following explicit expression:

$$\begin{aligned} T_{m,n}(y) = & \sum_{\substack{(\ell, \mathbf{j}(\ell), \mathbf{r}(\ell)) \\ = (\ell, (j_1, j_2, \dots, j_\ell), (r_1, r_2, \dots, r_\ell)) \in \mathcal{S}_m}} \frac{1}{\ell!} \left( -\frac{1}{\sqrt{\Delta}} \right)^\ell \prod_{i=1}^{\ell} D_{r_i r_i}(x_0) \\ & \times \mathbb{E} \left[ F_{n, (\ell, \mathbf{j}(\ell), \mathbf{r}(\ell))} \left( y - \frac{D(x_0)}{\sqrt{\Delta}} (b(x_0)\Delta + J(\Delta)) \right) | N(\Delta) = n \right], \end{aligned} \tag{42}$$

where  $F_{n, (\ell, \mathbf{j}(\ell), \mathbf{r}(\ell))}(z)$  is a polynomial explicitly calculated from

$$\begin{aligned} F_{n, (\ell, \mathbf{j}(\ell), \mathbf{r}(\ell))}(z) := & \mathcal{D}_{r_1} \left( \mathcal{D}_{r_2} \left( \dots \mathcal{D}_{r_\ell} \left( P_{n, (\ell, \mathbf{j}(\ell), \mathbf{r}(\ell))}(\sigma(x_0)^{-1} \right. \right. \right. \\ & \left. \left. \left. \times D(x_0)^{-1} \sqrt{\Delta} z \right) \dots \right) \right) \phi_{\Sigma(x_0)}(z) \end{aligned} \tag{43}$$

with coefficients involving polynomials of the jump arrival times  $\tau_1, \tau_2, \dots, \tau_n$  as well as jump amplitudes  $Z_1, Z_2, \dots, Z_n$ . Here,  $\mathcal{S}_m$  and  $\phi_{\Sigma(x_0)}(y)$  are defined in (30) and (27), respectively.

**Proof.** See Appendix B.  $\square$

This theorem provides a convenient expression for calculating the closed-form formula for  $T_{m,n}(y)$ . For illustration, we provide two examples under the aforementioned one-dimensional ( $d = 1$ ) case. For example,  $T_{1,1}(y)$  admits the following expression:

$$\begin{aligned} T_{1,1}(y) = & -\frac{D(x_0)}{\sqrt{\Delta}} \\ & \times \mathbb{E} \left( F_{1, (1, (1), (1))} \left( y - \frac{D(x_0)}{\sqrt{\Delta}} (b(x_0)\Delta + J(\Delta)) \right) | N(\Delta) = 1 \right), \end{aligned} \tag{44}$$

where

$$F_{1, (1, (1), (1))}(z) = \mathcal{D} \left( P_{1, (1, (1), (1))}(\sqrt{\Delta}z) \right) \phi(z) \tag{45}$$

and

$$P_{1, (1, (1), (1))}(w) = \mathbb{E} (X_2(\Delta) | W(\Delta) = w, N(\Delta) = 1, \mathcal{F}(\Delta)) \tag{46}$$

as well as  $J(t) = Z_1 1_{[\tau_1, \Delta]}(t)$ . Here, as a special case of (41), the differential operator satisfies that  $\mathcal{D}u(z) := \partial u(z)/\partial z - zu(z)$  for any differentiable function  $u$ ; as a special case of (27),  $\phi(z) = e^{-1/2z^2/\sqrt{2\pi}}$  is the p.d.f. of the standard normal distribution. Similarly, we have

$$\begin{aligned} T_{2,2}(y) = & -\frac{D(x_0)}{\sqrt{\Delta}} \mathbb{E} \left( F_{2, (1, (2), (1))} \right. \\ & \left. \left( y - \frac{D(x_0)}{\sqrt{\Delta}} (b(x_0)\Delta + J(\Delta)) \right) | N(\Delta) = 2 \right) \\ & + \frac{1}{2!} \left( -\frac{D(x_0)}{\sqrt{\Delta}} \right)^2 \mathbb{E} \left( F_{2, (2, (1, 1), (1, 1))} \right. \\ & \left. \left( y - \frac{D(x_0)}{\sqrt{\Delta}} (b(x_0)\Delta + J(\Delta)) \right) | N(\Delta) = 2 \right), \end{aligned}$$

where

$$F_{2,(1,(2),(1))}(z) = \mathcal{D} \left( P_{2,(1,(2),(1))}(\sqrt{\Delta}z) \right) \phi(z), \text{ and}$$

$$F_{2,(2,(1,1),(1,1))}(z) = \mathcal{D} \left( \mathcal{D} \left( P_{2,(2,(1,1),(1,1))}(\sqrt{\Delta}z) \right) \right) \phi(z),$$

and

$$P_{2,(1,(2),(1))}(w) = \mathbb{E} (X_3(\Delta) | W(\Delta) = w, N(\Delta) = 2, \mathcal{F}(\Delta)), \tag{47a}$$

$$P_{2,(2,(1,1),(1,1))}(w) = \mathbb{E} (X_2(\Delta)^2 | W(\Delta) = w, N(\Delta) = 2, \mathcal{F}(\Delta)) \tag{47b}$$

as well as  $J(t) = Z_1 1_{[\tau_1, \tau_2]}(t) + (Z_1 + Z_2) 1_{[\tau_2, \Delta]}(t)$ .

To obtain closed-form formulas, we will provide technical details on calculating conditional expectation (40) and jump-component-related expectation (42) in Appendices C and D, respectively.

### 3.4. An application in maximum-likelihood estimation (MLE)

Since the expansion of transition density proposed in (23) provides an approximation for the true but unknown transition density (3), it is natural to employ it in constructing an approximation for likelihood functions. Thus, following the idea of Ait-Sahalia (1999, 2002a, 2008), we propose a method of approximate maximum-likelihood estimation (MLE) in what follows. Based on the discrete observations of the jump-diffusion  $X$  defined in (1) at time grids  $\{\Delta, 2\Delta, \dots, n\Delta\}$  for some integer  $n$ , which correspond to the daily, weekly or monthly monitoring frequency, etc., the likelihood function is constructed as

$$l_n(\theta) = \prod_{i=1}^n p(\Delta t, X(i\Delta) | X((i-1)\Delta); \theta) \tag{48}$$

where  $p$  is the transition density defined in (3). Assuming success in identification, the true MLE  $\hat{\theta}_n$  is obtained by identifying the maximizer in  $\theta \in \Theta$  for function (48), i.e.,

$$\hat{\theta}_n = \arg \max_{\theta \in \Theta} l_n(\theta). \tag{49}$$

By analogy, we introduce the  $(M, N)$ th order approximate likelihood function as

$$l_n^{(M,N)}(\theta) = \prod_{i=1}^n p_{M,N}(\Delta t, X(i\Delta) | X((i-1)\Delta); \theta) \tag{50}$$

where  $p_{M,N}$  is the density approximation defined in (23). Thus, the approximate MLE  $\hat{\theta}_n^{(M,N)}$  is obtained by identifying the maximizer in  $\theta \in \Theta$  for function (50), i.e.,

$$\hat{\theta}_n^{(M,N)} = \arg \max_{\theta \in \Theta} l_n^{(M,N)}(\theta). \tag{51}$$

We employ  $\hat{\theta}_n^{(M,N)}$  as an approximation of  $\hat{\theta}_n$ , in particular, when the likelihood function  $l_n(\theta)$  is not easy to calculate. We demonstrate the numerical performance of this method through Monte Carlo simulation in Section 4. Theoretical discussions related to asymptotic properties of these estimators are provided in Li and Chen (2016).

## 4. Numerical performance and simulation results for MLE

Since general jump-diffusion models are rarely analytically tractable, we begin by employing an arithmetic Brownian motion with jump with closed-form transition density and three affine jump-diffusion models (see, e.g., Duffie et al. (2000)) with explicitly known characteristic functions to demonstrate the

performance of our method. To provide benchmarks for testing the accuracy of our density expansion, we truncate the infinite-series transition density for the former example and evaluate the true transition densities by Fourier transform inversion for the latter three examples. For all these examples, the expansion formulas are calculated from our general method discussed in the previous sections. For the purpose of illustration, we provide the first several expansion terms of these examples in Li and Chen (2016). All such formulas and those with higher orders are documented in the form of Mathematica notebook, which will be provided upon request. The corresponding likelihood expansions will be further used in Monte Carlo analysis for approximate MLE in Section 4.2.

The arithmetic Brownian motion with jump process (ABMJ hereafter) is specified as follows.

**Model 1.** The ABMJ model:

$$dX(t) = \mu dt + \sigma dW(t) + d \left( \sum_{n=0}^{N(t)} Z_n \right),$$

where  $\{(W(t))\}$  is a standard one-dimensional Brownian motion and the jump size has a normal distribution according to Jump-Size Distribution 1, i.e.,  $Z_n \sim \mathbf{N}(\alpha, \beta^2)$ .

It is straightforward to obtain its transition density as the following infinite series

$$\mathbb{P}(X(\Delta) \in dx | X(0) = x_0) = \sum_{n=0}^{\infty} \frac{1}{\sigma^2 \Delta + n\beta^2} \phi \left( \frac{x - x_0 - \mu \Delta - n\alpha}{\sigma^2 \Delta + n\beta^2} \right) \frac{(\lambda \Delta)^n}{n!} e^{-\lambda \Delta} dx,$$

where  $\phi$  is the probability density function of a standard normal variable, i.e.,  $\phi(x) = e^{-x^2/2} / \sqrt{2\pi}$ . In the numerical experiments, we will employ a set of parameters similar to that in Yu (2007), i.e.,  $\mu = 0.2, \sigma = 0.3, \lambda = 0.33, \alpha = 0$  and  $\beta = 0.2$ .

The mean-reverting Ornstein-Uhlenbeck with jump process (MROUJ hereafter) is specified as follows.

**Model 2.** The MROUJ model:

$$dX(t) = \kappa(\theta - X(t))dt + \sigma dW(t) + d \left( \sum_{n=0}^{N(t)} Z_n \right),$$

where  $\{(W(t))\}$  is a standard one-dimensional Brownian motion and the jump size has a normal distribution according to Jump-Size Distribution 1, i.e.,  $Z_n \sim \mathbf{N}(\alpha, \beta^2)$ .

Following standard methods (see, e.g., Chapter 5 in Singleton (2006)), the characteristic function of  $X(t)$  can be written as

$$\phi(t; \omega) = \mathbb{E} (e^{i\omega X(t)} | X(0) = x_0) = \exp(A(t; \omega) + x_0 B(t; \omega)), \tag{52}$$

with  $i = \sqrt{-1}$ ,

where

$$A(t; \omega) = i\omega\theta (1 - e^{-\kappa t}) + \frac{\omega^2 \sigma^2 (\exp(-2\kappa t) - 1)}{4\kappa} - \lambda t + \frac{\lambda}{2\kappa} \int_0^t \exp \left( i\omega \alpha e^{-\kappa s} - \frac{1}{2} \omega^2 \beta^2 e^{-2\kappa s} \right) ds,$$

and  $B(t; \omega) = i\omega \exp(-\kappa t)$ . In the numerical experiments, we will employ a set of parameters similar to that in Yu (2007), i.e.,  $\kappa = 0.5, \theta = 0, \sigma = 0.2, \lambda = 0.33, \alpha = 0$  and  $\beta = 0.28$ .

The square root diffusion with jump process (SQRJ hereafter) is specified as follows.

**Model 3.** The SQRJ model:

$$dX(t) = \kappa(\theta - X(t))dt + \sigma\sqrt{X(t)}dW(t) + d\left(\sum_{n=0}^{N(t)} Z_n\right),$$

where  $\{(W(t))\}$  is a standard one-dimensional Brownian motion and the jump size has i.i.d. exponential distribution with parameter  $\gamma$  according to [Jump-Size Distribution 2](#), i.e.,  $Z_n \sim \text{expo}(\gamma)$ .

This model is well known as the basic affine jump–diffusion model (BAJD), which generalizes the celebrated Cox–Ingersoll–Ross (CIR) model; see, e.g., its application for modeling credit default intensity in [Duffie and Gârleanu \(2001\)](#). Similar to the previous example, the characteristic function of SQRJ can be written as

$$\phi(t; \omega) = \mathbb{E}(e^{i\omega X(t)} | X(0) = x_0) = \exp(A(t; \omega) + x_0 B(t; \omega)),$$

$$\text{with } i = \sqrt{-1},$$

where

$$A(t; \omega) = -\frac{2\kappa\theta}{\sigma^2} \times \log\left(\frac{\kappa - \frac{1}{2}i\omega\sigma^2(1 - e^{-\kappa t})}{\kappa}\right) + \frac{2\lambda}{2\kappa - \omega^2\gamma} \times \log\left(\frac{\kappa - \frac{1}{2}i\omega\sigma^2 + i\omega\left(\frac{\sigma^2}{2} - \frac{\kappa}{\gamma}\right)e^{-\kappa t}}{\kappa - \frac{i\omega\kappa}{\gamma}}\right),$$

and

$$B(t; \omega) = \frac{i\omega\kappa e^{-\kappa t}}{\kappa - \frac{1}{2}i\omega\sigma^2(1 - e^{-\kappa t})}.$$

In the numerical experiments, we will employ a set of parameters similar to that in [Duffie and Gârleanu \(2001\)](#), i.e.,  $\kappa = 0.6$ ,  $\theta = 0.02$ ,  $\sigma = 0.141$ ,  $\lambda = 0.2$  and  $\gamma = 10$ .

A bivariate mean-reverting Ornstein–Uhlenbeck with jump model (BMROUJ hereafter) is specified as

**Model 4.** The BMROUJ model:

$$d\begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} = \begin{pmatrix} \kappa_{11} & 0 \\ \kappa_{21} & \kappa_{22} \end{pmatrix} \begin{pmatrix} \theta_1 - X_1(t) \\ \theta_2 - X_2(t) \end{pmatrix} dt + d\begin{pmatrix} W_1(t) \\ W_2(t) \end{pmatrix} + d\left(\sum_{n=1}^{N(t)} Z_{n,1}\right) + d\left(\sum_{n=1}^{N(t)} Z_{n,2}\right),$$

where  $\{(W_1(t), W_2(t))\}$  is a standard two-dimensional Brownian motion and the jump size has a bivariate normal distribution according to [Jump-Size Distribution 1](#), i.e.,

$$\begin{pmatrix} Z_{n,1} \\ Z_{n,2} \end{pmatrix} \sim \mathbf{N}\left(\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}, \begin{pmatrix} \beta_1^2 & 0 \\ 0 & \beta_2^2 \end{pmatrix}\right).$$

This model is an example of the multidimensional affine jump–diffusion model proposed in [Duffie et al. \(2000\)](#). Similarly to the previous examples, the characteristic function is obtained as

$$\phi(t; \omega_1, \omega_2) = \mathbb{E}(e^{i(\omega_1 X_1(t) + \omega_2 X_2(t))} | X(0) = (x_{0,1}, x_{0,2})) = \exp(A(t; \omega_1, \omega_2) + B_1(t; \omega_1, \omega_2)x_{0,1} + B_2(t; \omega_1, \omega_2)x_{0,2}), \text{ with } i = \sqrt{-1}.$$

Here, we have

$$A(t; \omega_1, \omega_2) = i\frac{\omega_1\kappa_{11}\theta_1(1 - e^{-\kappa_{11}t})}{\kappa_{11}} - \frac{\omega_1^2(1 - e^{-2\kappa_{11}t})}{4\kappa_{11}} - i\frac{\kappa_{21}\theta_1 + \kappa_{22}\theta_2}{\kappa_{22} - \kappa_{11}} \left(\frac{\omega_1\kappa_{21}(1 - e^{-\kappa_{11}t})}{\kappa_{11}}\right)$$

$$- \frac{(\omega_1\kappa_{21} + \omega_2(\kappa_{22} - \kappa_{11}))(1 - e^{-\kappa_{22}t})}{\kappa_{22}} - \frac{1}{(\kappa_{22} - \kappa_{11})^2} \left(\frac{(\omega_1\kappa_{21} + \omega_2(\kappa_{22} - \kappa_{11}))^2}{4\kappa_{22}}(1 - e^{-2\kappa_{22}t}) - \frac{\omega_1\kappa_{21}(\omega_1\kappa_{21} + \omega_2(\kappa_{22} - \kappa_{11}))(1 - e^{-(\kappa_{11} + \kappa_{22})t})}{\kappa_{11} + \kappa_{22}}\right) + \frac{\omega_1^2\kappa_{21}^2(1 - e^{-2\kappa_{11}t})}{4\kappa_{11}} - \frac{\lambda}{2} \int_0^t \exp(-c_1 e^{-2\kappa_{11}s} - c_2 e^{-2\kappa_{22}s} + c_3 e^{-(\kappa_{11} + \kappa_{22})s} + i(c_4 e^{-\kappa_{11}s} + c_5 e^{-\kappa_{22}s})) ds) - \lambda t,$$

as well as

$$B_1(t; \omega_1, \omega_2) = i\omega_1 e^{-\kappa_{11}t} \text{ and } B_2(t; \omega_1, \omega_2) = i\omega_2 e^{-\kappa_{22}t} + i\omega_1 \frac{\kappa_{21}}{\kappa_{11} - \kappa_{22}} (e^{-\kappa_{11}t} - e^{-\kappa_{22}t}),$$

where

$$c_1(\omega_1, \omega_2) = \omega_1^2(\kappa_{11}^2\beta_1^2 - 2\kappa_{11}\kappa_{22}\beta_1^2 + \kappa_{22}^2\beta_1^2 + \kappa_{21}^2\beta_2^2) \times (\kappa_{22} - \kappa_{11})^{-2} / 2, \\ c_2(\omega_1, \omega_2) = \beta_2^2(\omega_1\kappa_{21} + \omega_2(\kappa_{22} - \kappa_{11}))^2(\kappa_{22} - \kappa_{11})^{-2} / 2, \\ c_3(\omega_1, \omega_2) = \omega_1\beta_2^2\kappa_{21}(\omega_1\kappa_{21} + \omega_2(\kappa_{22} - \kappa_{11}))(\kappa_{22} - \kappa_{11})^{-2}, \\ c_4(\omega_1, \omega_2) = -\omega_1(\kappa_{22} - \kappa_{11})(\kappa_{11}\alpha_1 - \kappa_{22}\alpha_1 + \kappa_{21}\alpha_2) \times (\kappa_{22} - \kappa_{11})^{-2}, \\ c_5(\omega_1, \omega_2) = \alpha_2(\kappa_{22} - \kappa_{11})(\omega_1\kappa_{21} + \omega_2(\kappa_{22} - \kappa_{11})) \times (\kappa_{22} - \kappa_{11})^{-2}.$$

In the numerical experiments, we will employ a set of parameters in partial agreement with those employed for the bivariate Ornstein–Uhlenbeck model studied in [Cheridito et al. \(2007\)](#), i.e.,  $\kappa_{11} = 0.1570$ ,  $\kappa_{21} = 0.3279$ ,  $\kappa_{22} = 2.2883$ ,  $\theta_1 = \theta_2 = 0$ ,  $\lambda = 9$ ,  $\alpha_1 = 0.2$ ,  $\alpha_2 = 0.1$ ,  $\beta_1 = 0.3$  and  $\beta_2 = 0.5$ .

Given the explicit-form characteristic functions, we obtain the density functions via numerical inversion of Fourier transform. We choose a widely used algorithm proposed in [Abate and Whitt \(1992\)](#) to accomplish this goal. Without loss of generality, we employ MROUJ [Model 2](#) as an example to briefly outline the algorithm. According to [Abate and Whitt \(1992\)](#), the transition density function can be obtained by

$$\mathbb{P}(X(\Delta) \in dx | X(0) = x_0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ix\omega} \phi(\Delta; \omega) d\omega \equiv \frac{1}{\pi} \int_0^{\infty} [\cos x\omega \text{Re}(\phi)(\Delta; \omega) + \sin x\omega \text{Im}(\phi)(\Delta; \omega)] d\omega.$$

Thus, an efficient approximation for this density function is given by the following Euler summation

$$E(m, n, x) = \sum_{k=1}^m \binom{m}{k} 2^{-m} s_{n+k}(x),$$

where the truncated series is defined by

$$s_n(x) := \frac{h}{2\pi} + \frac{h}{\pi} \sum_{k=1}^n [\text{Re}(\phi)(\Delta; kh) \cos khx + \text{Im}(\phi)(\Delta; kh) \sin khx].$$

#### 4.1. Accuracy of the density expansion

For the sampling increment  $\Delta$ , we denote by  $e_{M,N}(\Delta, x|x_0; \theta) := p_{M,N}(\Delta, x|x_0; \theta) - p(\Delta, x|x_0; \theta)$



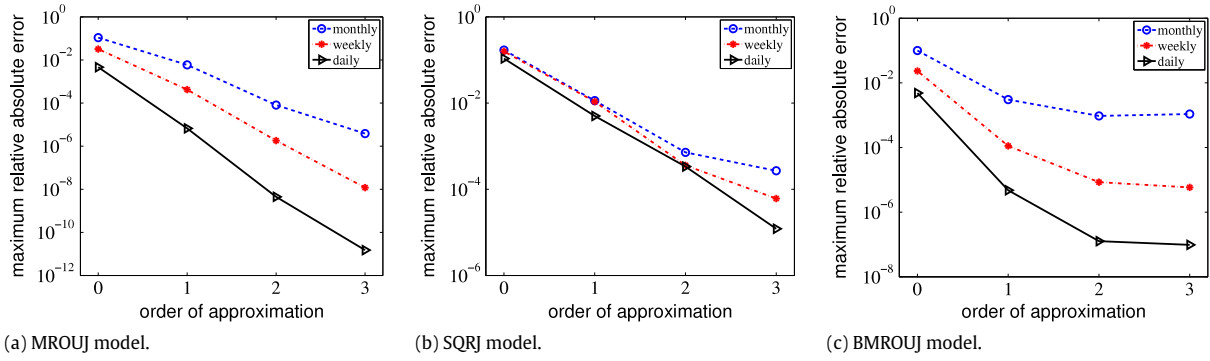


Fig. 1. Maximum relative absolute error of density approximation for Models 2, 3 and 4 with orders  $M = 0, 1, 2, 3$  and fixed  $N = 3$ .

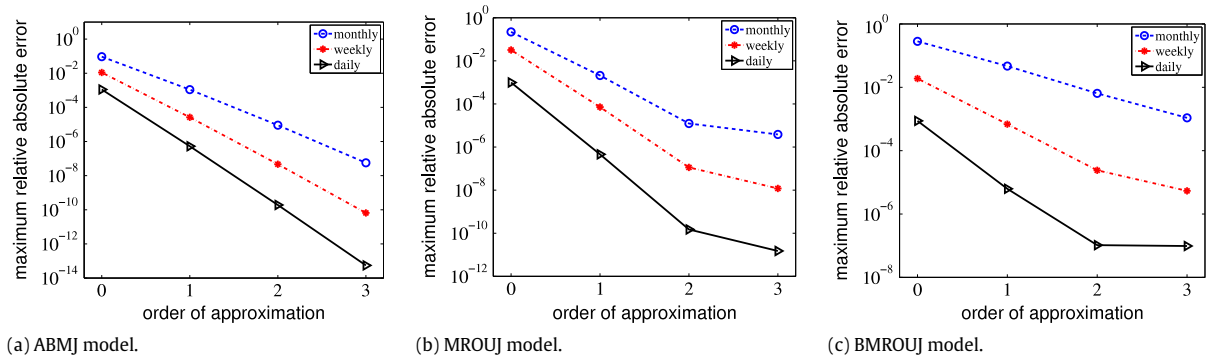


Fig. 2. Maximum relative absolute error of density approximation for Models 1, 2 and 4 with orders  $N = 0, 1, 2, 3$  and fixed  $M = 3$ .

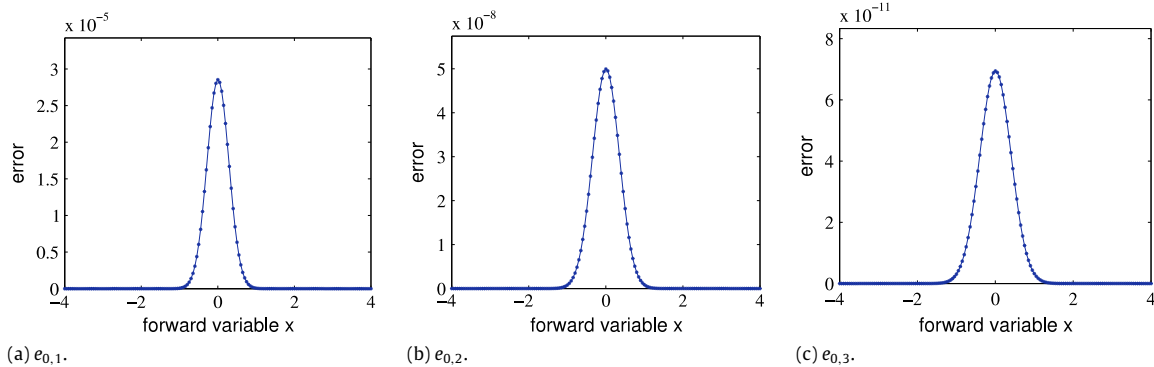


Fig. 3. Errors of density approximation for the Arithmetic Brownian Motion with Jump (ABMJ) Model, i.e.,  $e_{M,N}(\Delta, x|x_0; \theta)$ , for  $(M, N) = (0, 1), (0, 2), (0, 3)$  and  $\Delta = 1/52$ .

the error of the  $(M, N)$ th order approximation, that is (23). To demonstrate the overall accuracy of our expansion, we consider the maximum relative error  $\max_{x \in \mathcal{D}} |e_{M,N}(\Delta, x|x_0; \theta)/p(\Delta, x|x_0; \theta)|$  over a region  $\mathcal{D}$ , which is several standard deviations around the mean of the forward position (i.e.,  $\mathbb{E}(X(\Delta)|X(0) = x_0)$ ). Without loss of generality, for the ABMJ, MROUJ, and BMROUJ models, the initial positions  $x_0$  are chosen are 0, 0, and (0, 0), respectively; for the SQRJ model,  $x_0$  is chosen as 1.5 in order to keep it away from 0, the boundary of this jump–diffusion process. Considering monthly, weekly, and daily monitoring frequencies ( $\Delta = 1/12, 1/52, 1/252$ ), we plot the maximum relative errors of different orders  $M = 0, 1, 2, 3$  with fixed  $N = 3$  for the MROUJ, SQRJ, and BMROUJ models in Fig. 1 and plot the maximum relative errors of different orders  $N = 0, 1, 2, 3$  with fixed  $M = 3$  for the ABMJ, MROUJ, and BMROUJ models in Fig. 2. It is evident that the approximation errors tend to decrease as more correction terms are included ( $M$  and/or  $N$  increase).

In Figs. 3–5, we plot approximation errors of the density expansion in detail for the ABMJ, MROUJ, and SQRJ models, respectively. In Fig. 6, we plot the contours of the errors of the density expansion for the BMROUJ model. Using the ABMJ model, Fig. 3 demonstrates the correction effect resulting from increasing the order  $N$  in the approximation (23). By comparing Figs. 4(a) and 4(c), Figs. 5(a) and 5(c), Figs. 6(a) and 6(c), it is easy to observe the correction effects resulting from increasing the order  $M$ . Similarly, by comparing Figs. 4(b) and 4(c), Figs. 5(b) and 5(c), Figs. 6(b) and 6(c), it is easy to observe the correction effects resulting from increasing the order  $N$ .

4.2. Monte Carlo simulation evidence

In this section, we conduct Monte Carlo simulations to demonstrate the performance of the approximate maximum-likelihood estimation method proposed in Section 3.4. For the

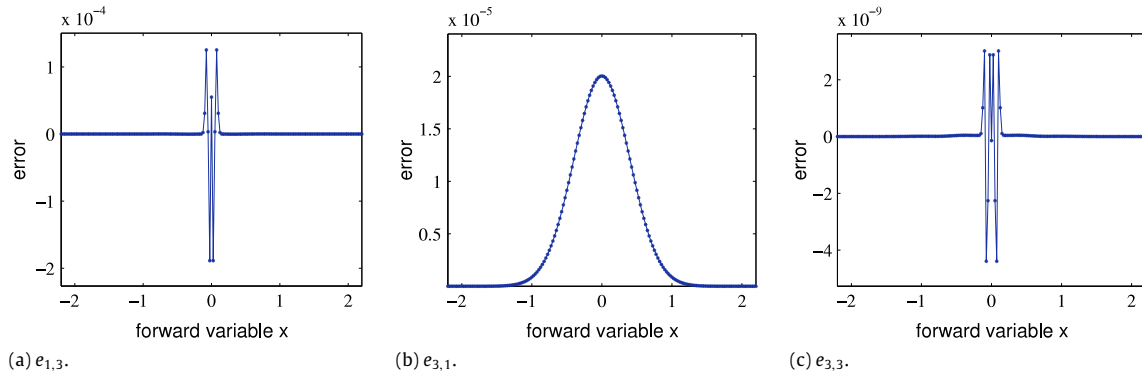


Fig. 4. Errors of density approximation for the Mean-reverting Ornstein–Uhlenbeck with Jump (MROUJ) Model, i.e.  $e_{M,N}(\Delta, x|x_0; \theta)$ , for  $(M, N) = (1, 3), (3, 1), (3, 3)$  and  $\Delta = 1/52$ .

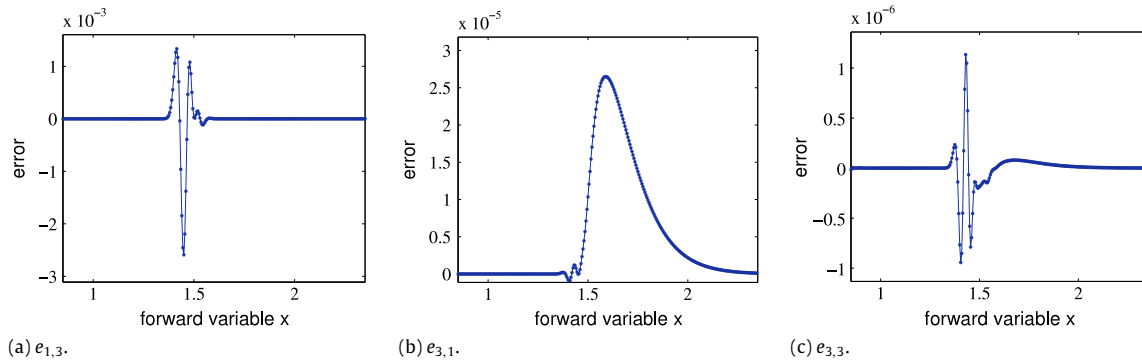


Fig. 5. Errors of density approximation for Square Root Diffusion with Jump (SQRJ) Model, i.e.  $e_{M,N}(\Delta, x|x_0; \theta)$ , for  $(M, N) = (1, 3), (3, 1), (3, 3)$  and  $\Delta = 1/52$ .

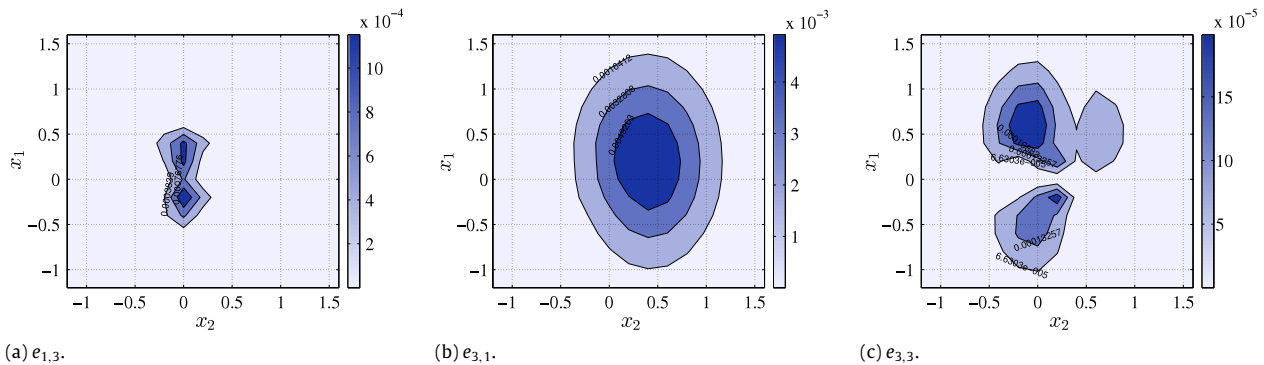


Fig. 6. Errors of density approximation for Bivariate Mean-reverting Ornstein–Uhlenbeck with Jump (BMROUJ) Model, i.e.  $e_{M,N}(\Delta, x|x_0; \theta)$ , for  $(M, N) = (1, 3), (3, 1), (3, 3)$  and  $\Delta = 1/52$ .

purpose of illustration, we begin by investigating the four models discussed in the previous section. We employ exact simulation methods (according to the true transition distribution) to generate the sample paths of the jump-diffusion models, see, e.g., the methods discussed in Chapter 3 of Glasserman (2004) and an alternative efficient method proposed in Giesecke and Smelov (2013).<sup>2</sup> In our experiments, true parameters are set as those employed in Section 4.1. In these and subsequent experiments, the total number of simulation trials is set as 1000 and the total number of observations on each sample path is set as  $n = 1000$  for each model; we consider three typical choices of the monitoring

increment:  $\Delta = 1/252$  (daily),  $\Delta = 1/52$  (weekly), and  $\Delta = 1/12$  (monthly). As seen from the Monte Carlo simulation results in Tables 1–4, we report the mean and standard deviation of the discrepancy between the true MLE and the true parameter value,  $\hat{\theta}_n - \theta^{\text{True}}$ , and the discrepancy between the approximate MLE and the true MLE,  $\hat{\theta}_n^{(M,N)} - \hat{\theta}_n$ . For Model 1 as shown in Table 1, our choices of the orders are  $(M, N) = (0, 1)$  and  $(0, 3)$ , since the simplicity of this model renders zero correction terms as  $M$  increases to  $M > 0$ . For the other three models as shown in Tables 2–4, our choices of the orders are  $(M, N) = (1, 3), (3, 1),$  and  $(3, 3)$ , for illustrating the effect of changing  $M$  and  $N$ .

We further demonstrate the practical applicability of our method by employing a CEV stochastic volatility with concurrent jump (CEV-SVCJ) model specified as follows.

<sup>2</sup> We thank Kay Giesecke for generously providing the code, which was directly applied in our simulation studies.

**Table 1**  
Monte Carlo Evidence for the ABMJ Model.

Parameters $\theta^{\text{True}}$	Finite sample $\hat{\theta}_n - \theta^{\text{True}}$		Finite sample $\hat{\theta}_n^{(0,1)} - \hat{\theta}_n$		Finite sample $\hat{\theta}_n^{(0,3)} - \hat{\theta}_n$	
	Mean	Stddev	Mean	Stddev	Mean	Stddev
$\Delta = 1/252$						
$\mu = 0.2$	0.021476	0.141770	0.003296	0.019200	-0.000022	0.000099
$\sigma = 0.3$	0.007785	0.041724	0.000062	0.000251	-0.000001	0.000004
$\lambda = 0.33$	0.167720	0.427378	-0.015509	0.058959	-0.000044	0.000196
$\alpha = 0$	0.000001	0.000003	0.004311	0.019276	0.000035	0.000155
$\beta = 0.2$	-0.049043	0.080835	0.005396	0.024911	0.000016	0.000072
$\Delta = 1/52$						
$\mu = 0.2$	0.002138	0.069304	-0.000865	0.012370	-0.000049	0.001601
$\sigma = 0.3$	-0.000867	0.007273	0.000196	0.001677	-0.000001	0.000230
$\lambda = 0.33$	0.146825	0.458778	-0.046288	0.293532	-0.002605	0.068776
$\alpha = 0$	-0.000773	0.009835	-0.000265	0.014892	0.000072	0.001582
$\beta = 0.2$	-0.016549	0.076476	0.001956	0.018169	0.000020	0.000723
$\Delta = 1/12$						
$\mu = 0.2$	0.002379	0.035149	-0.000732	0.006223	0.000014	0.000347
$\sigma = 0.3$	-0.001931	0.009824	0.000535	0.002727	0.000000	0.000117
$\lambda = 0.33$	0.166236	0.496944	-0.055653	0.260738	-0.000054	0.010354
$\alpha = 0$	-0.000692	0.011974	-0.000086	0.008871	-0.000007	0.000188
$\beta = 0.2$	-0.005491	0.066906	0.001428	0.011922	0.000008	0.000229

Notes. The number of simulation trials is set as 1000 and the number of observations on each path is  $n = 1000$ .

**Table 2**  
Monte Carlo Evidence for the MROUJ Model.

Parameters $\theta^{\text{True}}$	Finite sample $\hat{\theta}_n - \theta^{\text{True}}$		Finite sample $\hat{\theta}_n^{(1,3)} - \hat{\theta}_n$		Finite sample $\hat{\theta}_n^{(3,1)} - \hat{\theta}_n$		Finite sample $\hat{\theta}_n^{(3,3)} - \hat{\theta}_n$	
	Mean	Stddev	Mean	Stddev	Mean	Stddev	Mean	Stddev
$\Delta = 1/252$								
$\kappa = 0.5$	0.030645	0.061289	0.018137	0.032763	0.012603	0.039911	0.001266	0.002531
$\theta = 0$	-0.000104	0.000208	0.000415	0.000486	-0.000040	0.001132	-0.000076	0.000152
$\sigma = 0.2$	0.000106	0.000212	0.001667	0.003584	0.000050	0.000460	-0.000007	0.000014
$\lambda = 0.33$	-0.013829	0.027658	0.028869	0.061288	0.001541	0.054470	-0.000552	0.001104
$\alpha = 0$	-0.000723	0.001445	0.000345	0.000635	-0.000208	0.001131	0.000012	0.000024
$\beta = 0.28$	0.068028	0.136055	-0.062129	0.121034	0.001449	0.009289	-0.000112	0.000224
$\Delta = 1/52$								
$\kappa = 0.5$	0.226511	0.076686	0.004611	0.001503	0.006773	0.002436	-0.000697	0.000986
$\theta = 0$	0.001394	0.001029	-0.000408	0.001137	-0.000150	0.000571	0.000019	0.000027
$\sigma = 0.2$	0.003059	0.001773	-0.000065	0.000021	0.000022	0.000403	0.000062	0.000088
$\lambda = 0.33$	0.257111	0.222929	-0.009779	0.005662	-0.005719	0.004870	-0.000463	0.000655
$\alpha = 0$	-0.000234	0.001390	0.000267	0.000648	0.000077	0.000589	0.000006	0.000009
$\beta = 0.28$	-0.091571	0.079626	-0.000028	0.001381	0.002702	0.007062	-0.000053	0.000075
$\Delta = 1/12$								
$\kappa = 0.5$	0.018959	0.115585	0.012132	0.008716	0.000775	0.004838	0.000649	0.002034
$\theta = 0$	0.000009	0.000027	0.000095	0.000302	0.000065	0.000453	-0.000006	0.000019
$\sigma = 0.2$	0.004006	0.005580	0.000231	0.000450	0.000075	0.000413	0.000122	0.000287
$\lambda = 0.33$	0.079698	0.108969	0.001533	0.004054	-0.008227	0.005262	-0.000033	0.000104
$\alpha = 0$	0.000002	0.000007	0.000041	0.000131	-0.000257	0.000117	0.000004	0.000012
$\beta = 0.28$	0.000910	0.049361	0.000335	0.002419	-0.009980	0.003020	0.000338	0.000713

Notes. The number of simulation trials is set as 1000 and the number of observations on each path is  $n = 1000$ .

**Model 5.** The CEV-SVCJ model:

$$d \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} = \begin{pmatrix} \mu - \frac{1}{2} X_2(t) \\ \kappa (\theta - X_2(t)) \end{pmatrix} dt + \begin{pmatrix} \sqrt{X_2(t)} & 0 \\ \sigma \rho X_2(t)^\alpha & \sigma \sqrt{1 - \rho^2} X_2(t)^\alpha \end{pmatrix} d \begin{pmatrix} W_1(t) \\ W_2(t) \end{pmatrix} + d \begin{pmatrix} \sum_{n=1}^{N(t)} Z_{n,1} \\ \sum_{n=1}^{N(t)} Z_{n,2} \end{pmatrix},$$

where  $\{(W_1(t), W_2(t))\}$  is a standard two-dimensional Brownian motion. According to Duffie et al. (2000), we assume that the jump size in variance  $Z_{n,2}$  has an exponential with parameter  $\gamma_2$ . Conditional on a jump realization  $Z_{n,2}$ , the jump size in  $X_1(t)$

is normally distributed according to the conditional distribution  $Z_{n,1} | Z_{n,2} \sim \mathbf{N}(\mu_1 + \rho_j Z_{n,2}, \sigma_1^2)$ .

This model generalizes the affine SVCJ model investigated in Duffie et al. (2000) by allowing nonaffine specifications and extends the CEV stochastic volatility model investigated in Ait-Sahalia and Kimmel (2007) by adding jumps. As suggested by Duffie et al. (2000),  $X_1(t)$  can be used to model asset return; accordingly,  $X_2(t)$  can be used to model its stochastic variance. Without loss of generality, we employ two parameter sets in the Monte Carlo simulations. Corresponding to the affine SVCJ model, the first parameter set is given by  $\mu = 0.03, \kappa = 3, \theta = 0.1, \sigma = 0.25, \alpha = 0.5, \rho = -0.8, \lambda = 0.47, \mu_1 = -0.10, \sigma_1 = 0.0001, \rho_j = -0.38$ , and  $\gamma_2 = 20$ . As a nonaffine case, the second parameter set is given by  $\mu = 0.03, \kappa = 4, \theta = 0.05, \sigma = 0.75, \alpha = 0.8, \rho = -0.75, \lambda = 0.47, \mu_1 = -0.10, \sigma_1 = 0.0001, \rho_j = -0.38$ , and  $\gamma_2 = 20$ . These parameter sets are in partial agreement with the estimated parameters from Ait-Sahalia and Kimmel (2007) as well as the calibrated parameters

**Table 3**  
Monte Carlo Evidence for the SQRJ Model.

Parameters $\theta^{\text{True}}$	Finite sample $\hat{\theta}_n - \theta^{\text{True}}$		Finite sample $\hat{\theta}_n^{(1,3)} - \hat{\theta}_n$		Finite sample $\hat{\theta}_n^{(3,1)} - \hat{\theta}_n$		Finite sample $\hat{\theta}_n^{(3,3)} - \hat{\theta}_n$	
	Mean	Stddev	Mean	Stddev	Mean	Stddev	Mean	Stddev
$\Delta = 1/252$								
$\kappa = 0.6$	-0.073254	0.004977	-0.001686	0.000662	0.048160	0.000923	0.000009	0.000013
$\theta = 0.02$	0.005587	0.002711	0.002867	0.003614	0.038388	0.003893	-0.000337	0.000477
$\sigma = 0.141$	-0.000132	0.000208	-0.000003	0.000005	-0.000188	0.000000	-0.000002	0.000003
$\lambda = 0.2$	0.076182	0.228058	0.007174	0.003860	-0.049411	0.001088	-0.000046	0.000064
$\gamma = 10$	0.196001	0.277187	-0.071938	0.176927	0.000739	0.002646	-0.000269	0.000839
$\Delta = 1/52$								
$\kappa = 0.6$	0.059112	0.016394	-0.000252	0.000489	0.000904	0.001132	0.000051	0.000350
$\theta = 0.02$	0.012609	0.024885	0.000541	0.000848	0.000600	0.000514	0.000078	0.000442
$\sigma = 0.141$	-0.000242	0.000382	-0.000110	0.000036	-0.000003	0.000016	-0.000008	0.000019
$\lambda = 0.2$	-0.033253	0.087899	0.015980	0.017179	-0.003744	0.007501	0.000104	0.003477
$\gamma = 10$	0.161996	0.212702	-0.174539	0.217127	0.000051	0.014237	-0.001943	0.003887
$\Delta = 1/12$								
$\kappa = 0.6$	-0.004761	0.013056	0.000056	0.000962	-0.000323	0.001168	0.000013	0.000034
$\theta = 0.02$	0.001308	0.002733	0.004804	0.008235	0.000451	0.001139	0.000036	0.000082
$\sigma = 0.141$	-0.000198	0.000391	-0.000075	0.000141	-0.000152	0.000054	0.000002	0.000005
$\lambda = 0.2$	-0.065673	0.182982	-0.014253	0.078604	0.012678	0.014106	0.000483	0.001080
$\gamma = 10$	0.082496	0.116678	0.079719	0.346084	-0.019243	0.010672	0.000208	0.000294

Notes. The number of simulation trials is set as 1000 and the number of observations on each path is  $n = 1000$ .

**Table 4**  
Monte Carlo Evidence for the BMROUJ Model.

Parameters $\theta^{\text{True}}$	Finite sample $\hat{\theta}_n - \theta^{\text{True}}$		Finite sample $\hat{\theta}_n^{(1,3)} - \hat{\theta}_n$		Finite sample $\hat{\theta}_n^{(3,1)} - \hat{\theta}_n$		Finite sample $\hat{\theta}_n^{(3,3)} - \hat{\theta}_n$	
	Mean	Stddev	Mean	Stddev	Mean	Stddev	Mean	Stddev
$\Delta = 1/252$								
$\kappa_{11} = 0.1570$	0.004660	0.011640	0.000544	0.012311	0.000744	0.042146	0.000098	0.006109
$\kappa_{21} = 0.3279$	0.007827	0.024166	0.025917	0.032805	-0.050358	0.269770	-0.000300	0.012732
$\kappa_{22} = 2.2883$	0.033425	0.176910	0.242949	0.249903	0.089786	0.187985	-0.000439	0.095596
$\theta_1 = 0$	-0.000093	0.000292	-0.000100	0.000528	0.004543	0.001590	0.000009	0.000255
$\theta_2 = 0$	-0.000093	0.000250	-0.000109	0.000493	0.008386	0.007982	-0.000041	0.000262
$\lambda = 9$	0.928836	0.942508	0.299744	0.879494	-0.118305	0.325743	-0.006459	0.437318
$\alpha_1 = 0.2$	-0.013176	0.016091	-0.004234	0.013029	0.004232	0.054184	0.000418	0.007730
$\alpha_2 = 0.1$	-0.008912	0.019549	-0.005194	0.013036	0.050591	0.087059	-0.000246	0.006902
$\beta_1 = 0.3$	-0.001822	0.008944	-0.003823	0.009860	0.040831	0.049751	0.000701	0.005080
$\beta_2 = 0.5$	0.045888	0.029032	-0.006084	0.021357	0.025910	0.030929	0.000728	0.011999
$\Delta = 1/52$								
$\kappa_{11} = 0.1570$	0.001438	0.012208	-0.000376	0.003714	0.056400	0.002522	0.000043	0.000947
$\kappa_{21} = 0.3279$	0.002863	0.025816	-0.000037	0.006722	-0.000472	0.009470	-0.000054	0.001365
$\kappa_{22} = 2.2883$	0.022048	0.173931	0.004892	0.044678	0.008020	0.020741	-0.000681	0.010617
$\theta_1 = 0$	-0.000004	0.000218	-0.000035	0.000830	-0.000061	0.000905	-0.000001	0.000025
$\theta_2 = 0$	0.000006	0.000217	-0.000001	0.000093	0.000002	0.000086	-0.000001	0.000018
$\lambda = 9$	0.20399	0.633815	-0.003968	0.171036	-0.003919	0.065517	0.002565	0.052984
$\alpha_1 = 0.2$	-0.002022	0.019001	-0.000243	0.005043	-0.000781	0.007676	0.000022	0.000992
$\alpha_2 = 0.1$	-0.004155	0.021997	-0.000111	0.004160	-0.000514	0.002707	0.000085	0.001285
$\beta_1 = 0.3$	0.000692	0.007507	0.000185	0.002927	-0.000362	0.004451	-0.000008	0.000480
$\beta_2 = 0.5$	0.007557	0.036205	0.000121	0.007023	-0.000208	0.004114	-0.000035	0.001533
$\Delta = 1/12$								
$\kappa_{11} = 0.1570$	0.000366	0.004175	0.000222	0.003316	-0.000321	0.004436	0.000027	0.000546
$\kappa_{21} = 0.3279$	0.001168	0.010937	0.000140	0.008242	0.000559	0.003467	-0.000031	0.001324
$\kappa_{22} = 2.2883$	0.015837	0.110700	0.000413	0.036098	0.000585	0.012586	0.000515	0.018027
$\theta_1 = 0$	-0.000000	0.000119	0.000085	0.002728	0.000645	0.002084	0.000003	0.000093
$\theta_2 = 0$	-0.000001	0.000121	0.000036	0.001012	0.000059	0.007691	0.000003	0.000075
$\lambda = 9$	0.081659	0.448389	-0.002326	0.132204	0.000992	0.104276	-0.004515	0.095104
$\alpha_1 = 0.2$	0.001418	0.009973	0.000192	0.003463	0.001673	0.002303	-0.000013	0.000217
$\alpha_2 = 0.1$	0.003280	0.016796	0.000075	0.005034	0.000877	0.005291	0.000007	0.000571
$\beta_1 = 0.3$	0.000319	0.003235	0.000308	0.010172	0.000412	0.002681	0.000023	0.000785
$\beta_2 = 0.5$	0.003356	0.030656	-0.000839	0.009294	0.000733	0.006212	0.000062	0.003930

Notes. The number of simulation trials is set as 1000 and the number of observations on each path is  $n = 1000$ .

from Duffie et al. (2000). We employ Euler discretizations to generate the sample paths of this model, see, for example, Chapter 6 of Glasserman (2004). In the numerical experiments exhibited in Tables 5 and 6, we report the mean and standard deviation of the discrepancy between the approximate MLE and the true parameter value,  $\hat{\theta}_n^{(M,N)} - \theta^{\text{True}}$ , for  $(M, N) = (1, 3), (3, 1),$  and  $(3, 3)$ .

As shown from the numerical results in Tables 1–4, when the order of approximation increases, the approximate MLEs get closer

to the exact (but usually incomputable) MLEs, and thus get closer to the true parameter, if the sample size  $n$  is large enough. Thus, the approximate MLEs obtained by maximizing the approximate likelihood function approach the asymptotic efficiency of the true MLEs. This phenomenon agrees with the accuracy of our density expansion investigated in Section 5 of Li and Chen (2016) and reconciles our theoretical discussions on the asymptotic property of the approximate MLE in Section 6 of Li and Chen (2016). As shown in

**Table 5**  
Monte Carlo Evidence for the Affine SVCJ Model.

Parameters $\theta^{\text{True}}$	Finite sample $\hat{\theta}_n^{(1,3)} - \theta^{\text{True}}$		Finite sample $\hat{\theta}_n^{(3,1)} - \theta^{\text{True}}$		Finite sample $\hat{\theta}_n^{(3,3)} - \theta^{\text{True}}$	
	Mean	Stddev	Mean	Stddev	Mean	Stddev
$\Delta = 1/252$						
$\mu = 0.03$	0.000364	0.000706	-0.000156	0.001610	0.000219	0.000451
$\kappa = 3$	0.328481	0.143916	0.171773	0.383091	0.038922	0.060455
$\theta = 0.1$	0.000619	0.020642	0.001073	0.006309	0.001626	0.002313
$\sigma = 0.25$	-0.003245	0.055997	0.022458	0.048356	-0.001101	0.016275
$\rho = -0.8$	0.001641	0.008685	0.009610	0.059919	0.008991	0.043841
$\lambda = 0.47$	0.001044	0.007557	-0.020646	0.050806	0.000105	0.000616
$\mu_1 = -0.10$	-0.002672	0.001487	0.015115	0.057429	-0.002282	0.004846
$\sigma_1 = 0.0001$	0.000001	0.000004	-0.000001	0.000003	0.000000	0.000003
$\rho_1 = -0.38$	-0.000639	0.001487	0.000116	0.006628	-0.000420	0.000753
$\gamma_2 = 20$	0.208234	0.517280	-0.174754	0.655111	0.067385	0.021390
$\Delta = 1/52$						
$\mu = 0.03$	0.000295	0.000945	0.000115	0.000312	-0.000046	0.000885
$\kappa = 3$	0.124675	0.398034	0.034114	0.074008	0.001789	0.001494
$\theta = 0.1$	0.000558	0.005378	0.001832	0.003928	0.000641	0.005608
$\sigma = 0.25$	0.006401	0.028605	-0.000509	0.009599	0.000045	0.000258
$\rho = -0.8$	0.003375	0.054602	0.018891	0.057645	0.014652	0.047471
$\lambda = 0.47$	-0.001855	0.010300	0.005485	0.016746	0.000382	0.004148
$\mu_1 = -0.10$	0.003248	0.023227	-0.003413	0.012143	-0.002008	0.005823
$\sigma_1 = 0.0001$	-0.000001	0.000008	0.000001	0.000003	0.000001	0.000002
$\rho_1 = -0.38$	0.001835	0.008196	-0.000521	0.001289	0.000129	0.000383
$\gamma_2 = 20$	0.149504	0.620797	0.139752	0.302418	0.053390	0.111969
$\Delta = 1/12$						
$\mu = 0.03$	0.000149	0.000275	0.000262	0.000491	0.000007	0.000010
$\kappa = 3$	0.026804	0.067132	0.030145	0.055030	-0.007464	0.001053
$\theta = 0.1$	0.001014	0.004378	0.002082	0.002431	-0.003325	0.004703
$\sigma = 0.25$	0.001807	0.005018	-0.001046	0.016289	0.026106	0.036919
$\rho = -0.8$	0.018800	0.052283	0.012469	0.044603	0.001919	0.008317
$\lambda = 0.47$	0.005566	0.016069	-0.021294	0.030114	0.000888	0.005609
$\mu_1 = -0.10$	-0.002509	0.005310	-0.002365	0.004889	0.002329	0.003293
$\sigma_1 = 0.0001$	0.000001	0.000002	0.000001	0.000003	-0.000003	0.000004
$\rho_1 = -0.38$	-0.000700	0.001376	-0.000442	0.000778	-0.003554	0.005026
$\gamma_2 = 20$	0.074488	0.218087	0.071733	0.214253	0.004023	0.005699

Notes. The number of simulation trials is set as 1000 and the number of observations on each path is  $n = 1000$ .

Tables 1–4, while holding the length of sampling interval  $\Delta$  fixed, the approximation error  $\hat{\theta}_n^{(M,N)} - \theta_n$  decreases and is dominated by the intrinsic sampling error  $\hat{\theta}_n - \theta^{\text{True}}$  as  $M$  and/or  $N$  increase. According to Ait-Sahalia (1999, 2002a, 2008), once the approximation error resulted from replacing the true likelihood (48) by its approximation (50), is dominated by the sampling error due to the true MLE, approximate MLE  $\hat{\theta}_n^{(M,N)}$  is appropriate in practice; such a proper replacement has an effect that is statistically indiscernible from the sampling variation of the true but incomputable MLE  $\hat{\theta}_n$  around  $\theta$ . Thus, following the discussions in Ait-Sahalia (1999, 2002a, 2008), a small-order approximation is adequate for replacing the true MLE  $\hat{\theta}_n$  for the purpose of estimating unknown parameter  $\theta$ .

Similarly, as exhibited in Tables 5 and 6, the performance the approximate MLE measured by  $\hat{\theta}_n^{(M,N)} - \theta^{\text{True}}$  is improved as  $M$  and/or  $N$  increase for  $n$  large enough. This phenomenon reconciles the consistency of the incomputable true MLE  $\hat{\theta}_n$  as well as the improvement of the approximation error of  $\hat{\theta}_n^{(M,N)} - \hat{\theta}_n$  as orders increase. As a result of the fast development of modern computation technology, calculation of high-order likelihood approximations will become increasingly feasible; thus the performance of our approximate MLEs can be arbitrarily improved at least in principle.

### 5. Concluding remarks

In this paper, we propose a closed-form expansion for transition density of Poisson-driven jump–diffusion models, for which any arbitrary order of correction terms can be systematically obtained through a generally implementable algorithm. As an application, likelihood function is approximated explicitly and

thus employed in approximate maximum-likelihood estimation (MLE) for jump–diffusion models from discretely sampled data. Numerical examples and Monte Carlo evidence for illustrating the performance of our density expansion and the resulting approximate MLE are provided in order to demonstrate the wide applicability of the method. The convergence related to the density expansion and the approximate MLE are theoretically justified under some standard (but not necessary) sufficient conditions. Owing to the limited space of this paper, which focuses on introducing a method of estimation, investigations on more asymptotic properties related to the approximate MLE can be regarded as a future research topic. One may also apply the idea of explicitly approximating transition density in various other aspects of financial econometrics, for which explicit asymptotic expansions of certain quantities, e.g., option price, are helpful. An extension of the current method in order to incorporate more general jump–diffusion models (see, e.g., Cinlar and Jacod (1981) and Yu (2007)) can be set as another important direction for future research.

### Acknowledgments

We are grateful to Yacine Ait-Sahalia (the editor) and two anonymous referees for very helpful comments and suggestions. Chenxu Li also benefited from discussions with Jose Blanchet, Mark Broadie, Peter Carr, Kay Giesecke, Ioannis Karatzas, Arturo Kohatsu-higa, Per Mykland, Philip Protter, Yasutaka Shimizu, Masayuki Uchida, Dacheng Xiu, Nakahiro Yoshida, Lan Zhang, and seminar participants at Morgan Stanley, Columbia University, Shanghai Jiaotong University, Bernoulli Society Satellite Meeting to the ISI World Statistics Congress 2013 in Tokyo, International

**Table 6**  
Monte Carlo Evidence for the CEV-SVCJ Model.

Parameters $\theta^{\text{True}}$	Finite sample $\hat{\theta}_n^{(1,3)} - \theta^{\text{True}}$		Finite sample $\hat{\theta}_n^{(3,1)} - \theta^{\text{True}}$		Finite sample $\hat{\theta}_n^{(3,3)} - \theta^{\text{True}}$	
	Mean	Stddev	Mean	Stddev	Mean	Stddev
$\Delta = 1/252$						
$\mu = 0.03$	0.000341	0.000096	0.000169	0.000817	0.000046	0.000024
$\kappa = 4$	0.045455	0.012856	0.026933	0.089693	0.001388	0.004946
$\theta = 0.05$	0.000568	0.000161	0.000148	0.002660	0.000032	0.000195
$\sigma = 0.75$	0.008523	0.002411	0.002415	0.019433	0.000343	0.002240
$\alpha = 0.8$	-0.014545	0.035998	0.016998	0.022219	0.000503	0.003305
$\rho = -0.75$	0.034091	0.057854	0.019868	0.073066	0.000202	0.005833
$\lambda = 0.47$	0.005341	0.001511	0.003360	0.012329	-0.000059	0.000743
$\mu_1 = -0.10$	-0.004318	0.001221	-0.001009	0.005007	-0.000185	0.000502
$\sigma_1 = 0.0001$	0.000001	0.000000	0.000001	0.000005	0.000000	0.000002
$\rho_j = -0.38$	-0.001136	0.000321	-0.000349	0.001433	-0.000433	0.000107
$\gamma_2 = 20$	0.227273	0.064282	0.147132	0.396738	0.002961	0.026454
$\Delta = 1/52$						
$\mu = 0.03$	0.000514	0.000726	0.000052	0.000153	0.000012	0.000066
$\kappa = 4$	0.126993	0.179595	0.017026	0.066626	0.005203	0.010674
$\theta = 0.05$	0.001811	0.002561	0.000347	0.001060	0.000016	0.000125
$\sigma = 0.75$	0.025743	0.016627	0.007148	0.018427	-0.000479	0.003095
$\alpha = 0.8$	-0.007838	0.011085	0.013408	0.029640	0.001489	0.002908
$\rho = -0.75$	0.071619	0.101285	-0.003591	0.023346	0.001294	0.004984
$\lambda = 0.47$	0.000121	0.000171	-0.000045	0.002416	0.000187	0.000801
$\mu_1 = -0.10$	-0.016252	0.022983	-0.000228	0.001075	-0.000292	0.000361
$\sigma_1 = 0.0001$	-0.000000	0.000000	-0.000000	0.000001	0.000000	0.000001
$\rho_j = -0.38$	0.000484	0.000685	-0.000136	0.001347	-0.000086	0.000171
$\gamma_2 = 20$	0.183647	0.259717	0.095744	0.266145	0.000824	0.026568
$\Delta = 1/12$						
$\mu = 0.03$	0.000046	0.000619	0.000151	0.000337	-0.000010	0.000047
$\kappa = 4$	-0.069939	0.190072	0.013383	0.046579	0.001695	0.005898
$\theta = 0.05$	0.001592	0.001112	0.000257	0.000798	0.000019	0.000154
$\sigma = 0.75$	0.025906	0.019544	0.004547	0.010606	0.000472	0.002314
$\alpha = 0.8$	-0.057450	0.065819	0.001519	0.025132	0.001733	0.002930
$\rho = -0.75$	0.014489	0.082946	-0.011847	0.025319	0.000551	0.005274
$\lambda = 0.47$	-0.003380	0.015492	0.000747	0.002201	-0.000031	0.000697
$\mu_1 = -0.10$	-0.009448	0.004701	0.000400	0.005663	-0.000208	0.000456
$\sigma_1 = 0.0001$	0.000002	0.000000	0.000000	0.000000	0.000000	0.000001
$\rho_j = -0.38$	-0.002040	0.000606	-0.000260	0.000548	-0.000058	0.000178
$\gamma_2 = 20$	0.177732	0.204470	0.038872	0.088285	0.003001	0.003154

Notes. The number of simulation trials is set as 1000 and the number of observations on each path is  $n = 1000$ .

Symposium on Financial Engineering and Risk Management 2014 in Beijing, and Tsinghua-PKU-Stanford conference on quantitative finance 2015. Chenxu Li's research was supported by the Guanghua School of Management, the Center for Statistical Sciences, and the Key Laboratory of Mathematical Economics and Quantitative Finance (Ministry of Education) at Peking University, as well as the National Natural Science Foundation of China (Grant 11201009). Dachuan Chen is grateful to the scholarship from the University of Illinois at Chicago.

**Appendix A. Calculation of (A.1)**

To explicitly calculate  $T_{0,n}(y)$  from formula (25) under the assumptions of Jump-Size Distributions 1 and 2, it is sufficient to consider the explicit calculation of the following expectation

$$\mathbb{E} \left[ \phi_{\Sigma(x_0)}(A + BJ(\Delta)) \prod_{l=1}^d J_l(\Delta)^{n_l} | N(\Delta) = n \right], \tag{A.1}$$

where  $A = (A_1, A_2, \dots, A_d)^\top$  and  $B = \text{diag}(B_1, B_2, \dots, B_d)$  for some constants  $A_1, A_2, \dots, A_d, B_1, B_2, \dots, B_d$  and non-negative integers  $n_1, n_2, \dots, n_d$ . As shown momentarily in Appendix D, the explicit calculation of (42) for constructing higher-order correction terms also hinges on (A.1).

Under Jump-Size Distribution 1, conditioning on  $N(\Delta) = n$ ,  $J_l(\Delta) = \sum_{i=1}^{N(\Delta)} Z_{i,l}$  has a univariate normal distribution, i.e.,  $J_l(\Delta) | N(\Delta) = n \sim \mathbf{N}(n\alpha_l, n\beta_l^2)$ . We have the following Lemma.

**Lemma 1.** *Let*

$$\widehat{\Sigma}(x_0) := (B^\top \Sigma(x_0)^{-1} B + (n\beta)^{-1})^{-1} \quad \text{and} \\ Q = B^\top \Sigma(x_0)^{-1} A - \beta^{-1} \alpha, \tag{A.2}$$

where  $\alpha$  and  $\beta$  are the mean vector and the covariance matrix introduced in Jump-Size Distribution 1, respectively. We have

$$\mathbb{E} \left[ \phi_{\Sigma(x_0)}(A + BJ(\Delta)) \prod_{l=1}^d J_l(\Delta)^{n_l} | N(\Delta) = n \right] \\ = \frac{(\det \widehat{\Sigma}(x_0))^{\frac{1}{2}}}{(2\pi)^{\frac{d}{2}} (\det \Sigma(x_0) \det n\beta)^{\frac{1}{2}}} \\ \times \exp \left( -\frac{1}{2} A^\top \Sigma(x_0)^{-1} A - \frac{n}{2} \alpha^\top \beta^{-1} \alpha + \frac{1}{2} Q^\top \widehat{\Sigma}(x_0) Q \right) \\ \times \frac{\partial^{(n_1)} \partial^{(n_2)} \dots \partial^{(n_d)}}{\partial \theta_1^{n_1} \partial \theta_2^{n_2} \dots \partial \theta_d^{n_d}} \\ \times \exp \left( -\theta^\top \widehat{\Sigma}(x_0) Q + \frac{1}{2} \theta^\top \widehat{\Sigma}(x_0) \theta \right) \Big|_{\theta_1, \theta_2, \dots, \theta_d = 0},$$

where  $\theta = (\theta_1, \theta_2, \dots, \theta_d)^\top$ .

**Proof.** The proof of the lemma follows from straightforward calculations; see Section 2.1 of Li and Chen (2016).  $\square$

Thus, denoting by

$$A = y - D(x_0)b(x_0)\sqrt{\Delta} \text{ and } B = -\frac{D(x_0)}{\sqrt{\Delta}} \tag{A.3}$$

and letting  $n_1, n_2, \dots, n_d = 0$ , it is easy to obtain the leading term  $T_{0,n}(y)$  as

$$T_{0,n}(y) = \phi_{\Sigma(x_0) + \frac{n}{d}D(x_0)\beta D(x_0)} \left( y - D(x_0)b(x_0)\sqrt{\Delta} - \frac{n}{\sqrt{\Delta}}D(x_0)\alpha \right). \tag{A.4}$$

Under **Jump-Size Distribution 2**, conditioning on  $N(\Delta) = n$ ,  $J_l(\Delta)$  has an Erlang distribution (see, e.g., Chapter 17 in [Johnson et al. \(1996\)](#)) with p.d.f.

$$P(J_l(\Delta) \in dv | N(\Delta) = n) = \frac{\gamma_l^n}{(n-1)!} \exp(-\gamma_l v) v^{n-1} dv, \tag{A.5}$$

for each  $l = 1, 2, \dots, d$ . Similar to [Lemma 1](#), we have the following formula.

**Lemma 2.** Let

$$\widehat{A} = (\widehat{A}_1, \widehat{A}_2, \dots, \widehat{A}_d)^\top = \Sigma(x_0) (B^{-1})^\top (B^\top \Sigma(x_0)^{-1}A + \Gamma) \text{ with } \Gamma = (\gamma_1, \gamma_2, \dots, \gamma_d)^\top. \tag{A.6}$$

We have

$$\begin{aligned} & \mathbb{E} \left[ \phi_{\Sigma(x_0)} (A + BJ(\Delta)) \prod_{i=1}^d J_i(\Delta)^{n_i} | N(\Delta) = n \right] \\ &= \frac{1}{\det B} \prod_{i=1}^d \frac{\gamma_i^n}{(n-1)!} (-1)^{(n-1)d + \sum_{i=1}^d n_i} \\ & \times \frac{\partial^{n_1+n-1} \partial^{n_2+n-1} \dots \partial^{n_d+n-1}}{\partial \gamma_1^{n_1+n-1} \partial \gamma_2^{n_2+n-1} \dots \partial \gamma_d^{n_d+n-1}} \left( \exp \left( -\frac{1}{2} A^\top \Sigma(x_0)^{-1} A \right. \right. \\ & \left. \left. + \frac{1}{2} \widehat{A}^\top \Sigma(x_0)^{-1} \widehat{A} \right) \mathcal{N}_{\Sigma(x_0)}(-\widehat{A}) \right). \tag{A.7} \end{aligned}$$

Here,  $\mathcal{N}_C(\cdot)$  denotes the cumulative distribution function of a normal distribution with zero mean and covariance matrix  $C$ .

**Proof.** The proof of the lemma follows from straightforward calculations; see Section 2.2 of [Li and Chen \(2016\)](#).  $\square$

Thus, it is easy to obtain the leading term  $T_{0,n}(y)$  by plugging  $n_1, n_2, \dots, n_d = 0$  in formula (A.7).

**Appendix B. Proof of Theorem 1**

**Proof.** Following (19) and (39), we have

$$\begin{aligned} T_{m,n}(y) &= \sum_{\substack{(\ell, \mathbf{j}(\ell), \mathbf{r}(\ell)) \\ = (\ell, (j_1, j_2, \dots, j_\ell), (r_1, r_2, \dots, r_\ell)) \in \mathcal{S}_m}} \frac{1}{\ell!} \left( \frac{1}{\sqrt{\Delta}} \right)^\ell \prod_{i=1}^\ell D_{r_i r_i}(x_0) \\ & \times \mathbb{E} \left( \frac{\partial^{(\ell)} \delta(Y_0 - y)}{\partial x_{r_1} \partial x_{r_2} \dots \partial x_{r_\ell}} \prod_{i=1}^\ell X_{j_i+1, r_i} | N(\Delta) = n \right). \end{aligned}$$

We note that

$$\begin{aligned} & \mathbb{E} \left( \frac{\partial^{(\ell)} \delta(Y_0 - y)}{\partial x_{r_1} \partial x_{r_2} \dots \partial x_{r_\ell}} \prod_{i=1}^\ell X_{j_i+1, r_i} | N(\Delta) = n \right) \\ &= \mathbb{E} \left[ \mathbb{E} \left( \frac{\partial^{(\ell)} \delta(Y_0 - y)}{\partial x_{r_1} \partial x_{r_2} \dots \partial x_{r_\ell}} \right. \right. \\ & \left. \left. \times \prod_{i=1}^\ell X_{j_i+1, r_i} | N(\Delta) = n, \mathcal{J}(\Delta) \right) | N(\Delta) = n \right]. \end{aligned}$$

Recall that  $E_n(\cdot)$  denotes the conditional expectation operator  $\mathbb{E}(\cdot | N(\Delta) = n, \mathcal{J}(\Delta))$ . Moreover, we denote by

$$z := y - \frac{D(x_0)}{\sqrt{\Delta}} (b(x_0)\Delta + J(\Delta)).$$

Using the integration-by-parts formula of Dirac Delta function (see, e.g., Section 2.6 in [Kanwal \(2004\)](#)) and recalling (14), we arrive at

$$\begin{aligned} & E_n \left( \frac{\partial^{(\ell)} \delta(Y_0 - y)}{\partial x_{r_1} \partial x_{r_2} \dots \partial x_{r_\ell}} \prod_{i=1}^\ell X_{j_i+1, r_i} \right) \\ &= \int_{-\infty}^{\infty} \frac{\partial^{(\ell)} \delta(w - z)}{\partial x_{r_1} \partial x_{r_2} \dots \partial x_{r_\ell}} E_n \left( \prod_{i=1}^\ell X_{j_i+1, r_i} \middle| \frac{B(\Delta)}{\sqrt{\Delta}} = w \right) \\ & \times \phi_{\Sigma(x_0)}(w) dw = (-1)^\ell \frac{\partial^{(\ell)}}{\partial z_{r_1} \partial z_{r_2} \dots \partial z_{r_\ell}} \\ & \left[ E_n \left( \prod_{i=1}^\ell X_{j_i+1, r_i} \middle| \frac{B(\Delta)}{\sqrt{\Delta}} = z \right) \phi_{\Sigma(x_0)}(z) \right]. \end{aligned}$$

Recalling the definitions of (15) and (40) and noticing that, for any smooth function  $u(z)$ ,

$$\begin{aligned} \frac{\partial}{\partial z_i} (u(z) \phi_{\Sigma(x_0)}(z)) &= \left( \frac{\partial u(z)}{\partial z_i} - u(z) (\Sigma(x_0)^{-1} z)_i \right) \phi_{\Sigma(x_0)}(z) \\ &\equiv \mathcal{D}_i u(z), \end{aligned}$$

we obtain formulas (42) and (43) immediately.  $\square$

**Appendix C. Calculation of (40)**

*C.1. Conversion from multiplication to linear combination*

The first technical issue is to convert  $X_{j_1+1, r_1} X_{j_2+1, r_2} \dots X_{j_\ell+1, r_\ell}$  into a linear combination of iterated Stratonovich integrals using the following lemma. For an arbitrary index  $\mathbf{i} = (i_1, i_2, \dots, i_\ell)$ , we denote by  $-\mathbf{i}$  the index obtained from deleting the first component of index  $\mathbf{i}$ , i.e.,  $-\mathbf{i} = (i_2, i_3, \dots, i_\ell)$ . Similarly, let  $-\mathbf{f} = \{f_2(t), f_3(t), \dots, f_\ell(t)\}$ . Thus, it is obvious that

$$S_{-\mathbf{i}, -\mathbf{f}}(t) := \int_0^t \dots \int_0^{t_{i-1}} f_i(t_i) \circ dW_{i_i}(t_i) \dots f_2(t_2) \circ dW_{i_2}(t_2).$$

**Lemma 3.** For two indices  $\mathbf{i} = (i_1, i_2, \dots, i_\ell)$  and  $\mathbf{j} = (j_1, j_2, \dots, j_k)$  as well as stochastic processes  $\mathbf{f} = \{f_1(t), f_2(t), \dots, f_\ell(t)\}$  and  $\mathbf{g} = \{g_1(t), g_2(t), \dots, g_k(t)\}$ , we have

$$\begin{aligned} S_{\mathbf{i}, \mathbf{f}}(t) S_{\mathbf{j}, \mathbf{g}}(t) &= \int_0^t S_{\mathbf{i}, \mathbf{f}}(u) S_{-\mathbf{j}, -\mathbf{g}}(u) g_1(u) \circ dW_{j_1}(u) \\ & \quad + \int_0^t S_{\mathbf{j}, \mathbf{g}}(u) S_{-\mathbf{i}, -\mathbf{f}}(u) f_1(u) \circ dW_{i_1}(u). \tag{C.1} \end{aligned}$$

**Proof.** See Section 2.3 of [Li and Chen \(2016\)](#).  $\square$

For example, applications of (C.1) to the products generated by (38) yields that

$$\begin{aligned} S_{(0),(1)}(t)^2 &= 2S_{(0,0),(1,1)}(t), \\ S_{(0),(1)}(t) S_{(1),(1)}(t) &= S_{(0,1),(1,1)}(t) + S_{(1,0),(1,1)}(t), \\ S_{(1),(1)}(t)^2 &= 2S_{(1,1),(1,1)}(t). \end{aligned}$$

Thus, we obtain a linear combination form of  $X_1(t)^2$  as

$$\begin{aligned} X_1(t)^2 &= 2b(x_0)^2 S_{(0,0),(1,1)}(t) \\ & \quad + 2\sigma(x_0)b(x_0) (S_{(0,1),(1,1)}(t) + S_{(1,0),(1,1)}(t)) \\ & \quad + 2\sigma(x_0)^2 S_{(1,1),(1,1)}(t) + 2b(x_0)J(t)S_{(0),(1)}(t) \\ & \quad + 2\sigma(x_0)J(t)S_{(1),(1)}(t) + J^2(t). \end{aligned}$$

By plugging this expression and (36) in (37), it is straightforward to write  $X_3(t)$  as a linear combination of iterated Stratonovich integrals. Due to the length of this paper, we omit such a tedious formula. Similarly, to simplify  $X_2(\Delta)^2$  for calculating (47b), we need the following conversion among others:

$$S_{(1),J(t)}(t)^2 = 2S_{(1,1),J(t)J(t)}(t).$$

C.2. Conversion from iterated Stratonovich integrals to Itô integrals

In this section, we discuss the conversion from iterated Stratonovich integral (33) to a linear combination of iterated Itô integrals defined in a similar way except for changing the Stratonovich integrals in (33) into that of Itô sense, i.e.,

$$I_{i,f}(t) = \int_0^t \int_0^{t_1} \cdots \int_0^{t_{l-1}} f_1(t_1) dW_{i_1}(t_1) \cdots f_2(t_2) dW_{i_2}(t_2) f_1(t_1) \times dW_{i_1}(t_1). \tag{C.2}$$

Since the conditioned jump path is stepwise constant, we concentrate on the case where  $f_i$  are step functions. By using the fact that, for two continuous semimartingales  $X$  and  $Y$ ,

$$\int_0^t X(s) \circ dY(s) = \int_0^t X(s) dY(s) + \frac{1}{2} \langle X, Y \rangle (t),$$

see, e.g., Section II.7 in Protter (1990), we will generalize the relation between iterated Stratonovich and Itô integrals investigated in Section 5.2 of Kloeden and Platen (1992). Define the length of  $\mathbf{i} = (i_1, i_2, \dots, i_l)$  with  $i_1, i_2, \dots, i_l \in \{0, 1, 2, \dots, d\}$  by

$$l(\mathbf{i}) := l((i_1, i_2, \dots, i_l)) = l. \tag{C.3}$$

Let  $\zeta$  denote the index with zero length, i.e.,  $l(\zeta) = 0$ . We propose the following lemma.

**Lemma 4.** For an arbitrary index  $\mathbf{i} = (i_1, i_2, \dots, i_l)$  with  $i_1, i_2, \dots, i_l \in \{0, 1, 2, \dots, d\}$ , if  $l(\mathbf{i}) = 0, 1$ , we have  $S_{i,f}(t) = I_{i,f}(t)$ ; if  $l(\mathbf{i}) \geq 2$ , we have

$$S_{i,f}(t) = \int_0^t S_{-i,-f}(t_1) f_1(t_1) dW_{i_1}(t_1) + \frac{1}{2} 1_{\{i_1=i_2 \neq 0\}} \times \int_0^t S_{(-i),(-f)}(t_2) f_2(t_2) f_1(t_2) dt_2.$$

**Proof.** See Section 2.4 of Li and Chen (2016). □

For example, to calculate (46), we employ the following blocks to convert (36) to a linear combination of iterated Itô combination of iterated Itô integrals

$$S_{(0,1),(1,1)}(t) = I_{(0,1),(1,1)}(t), \tag{C.4a}$$

$$S_{(1,0),(1,1)}(t) = I_{(1,0),(1,1)}(t), \tag{C.4b}$$

$$S_{(1,1),(1,1)}(t) = I_{(1,1),(1,1)}(t) + \frac{1}{2} I_{(0),(1)}(t), \tag{C.4c}$$

$$S_{(1),J(t)}(t) = I_{(1),J(t)}(t). \tag{C.4d}$$

Similarly, to calculate (47b), the following conversion is needed among others:

$$S_{(1,1),J(t)J(t)}(t) = I_{(1,1),J(t)J(t)}(t) + \frac{1}{2} I_{(0),J(t)^2}(t). \tag{C.5}$$

C.3. Conditional expectation of iterated Itô integrals

Following the previous discussions, we finally focus on the conditional expectation of iterated Itô integrals in the

following form:  $E_n(I_{i,f}(\Delta)|W(\Delta) = w)$ , where  $E_n(\cdot)$  denotes the conditional expectation operator  $\mathbb{E}(\cdot|N(\Delta) = n, \mathcal{J}(\Delta))$  for simplicity. A natural start is to remove the conditioning by an explicit construction of Brownian bridge (see, e.g., Section 5.6 in Karatzas and Shreve (1991)), i.e., for any  $r = 1, 2, \dots, d$ ,

$$(W_r(s)|W(\Delta) = w) \stackrel{\text{in law}}{=} (W_r(s)|W_r(\Delta) = w_r) \stackrel{\text{in law}}{=} \mathfrak{B}_r^w(s) := \mathcal{B}_r(s) - \frac{s}{\Delta} \mathcal{B}_r(\Delta) + \frac{s}{\Delta} w_r,$$

where  $\mathcal{B}_r$ 's are independent Brownian motions and

$$\mathfrak{B}_r^w(s) := \mathcal{B}_r(s) - \frac{s}{\Delta} \mathcal{B}_r(\Delta) + \frac{s}{\Delta} w_r \tag{C.6}$$

is distributed as a Brownian bridge starting from 0 and ending at  $w_r$  at time  $\Delta$ . We assume  $\mathfrak{B}_0^w(s) \equiv s$  and  $\mathcal{B}_0(s) \equiv s$ . Thus, it follows that

$$E_n(I_{i,f}(\Delta)|W(\Delta) = w) = E\left(\int_0^\Delta \int_0^{t_1} \cdots \int_0^{t_{l-1}} f_1(t_1) d\mathfrak{B}_{i_1}^w(t_1) \cdots f_2(t_2) d\mathfrak{B}_{i_2}^w(t_2) f_1(t_1) d\mathfrak{B}_{i_1}^w(t_1)\right). \tag{C.7}$$

To articulate some useful tools, we introduce an important type of iterated Itô integral as follows. For an index  $\mathbf{i} = (i_1, i_2, \dots, i_l)$  with an ascending order  $0 \leq i_1 \leq i_2 \leq \dots \leq i_l \leq d$  and an integrand  $\mathbf{f} = \{f_1(t), f_2(t), \dots, f_l(t)\}$ , denote by

$$\mathcal{J}_{\mathbf{v}_i, \mathbf{f}}(t) := \int_0^t \int_0^{t_1} \cdots \int_0^{t_{l-1}} f_1(t_1) d\mathcal{B}_{i_1}(t_1) \cdots f_2(t_2) d\mathcal{B}_{i_2}(t_2) f_1(t_1) d\mathcal{B}_{i_1}(t_1), \tag{C.8}$$

where  $\mathbf{v}_i = (v_0, v_1, v_2, \dots, v_d)$  is a vector with

$$v_j := \#\{k = 1, 2, \dots, l : i_k = j\}, \text{ for } j = 0, 1, \dots, d, \tag{C.9}$$

i.e., the total number of  $k$  from the set  $\{1, 2, \dots, l\}$  such that  $i_k = j$ . It is obvious that  $\sum_{j=0}^d v_j = l$ .

**Proposition 1.** The expectation (C.7) admits the following representation

$$E_n(I_{i,f}(\Delta)|W(\Delta) = w) = \sum_{p_1=0}^{v_1} \binom{v_1}{p_1} \left(\frac{w_1}{\Delta}\right)^{p_1} \sum_{q_1=0}^{v_1-p_1} \binom{v_1-p_1}{q_1} \left(-\frac{1}{\Delta}\right)^{v_1-p_1-q_1} \times \sum_{p_2=0}^{v_2} \binom{v_2}{p_2} \left(\frac{w_2}{\Delta}\right)^{p_2} \sum_{q_2=0}^{v_2-p_2} \binom{v_2-p_2}{q_2} \left(-\frac{1}{\Delta}\right)^{v_2-p_2-q_2} \times \cdots \times \sum_{p_d=0}^{v_d} \binom{v_d}{p_d} \left(\frac{w_d}{\Delta}\right)^{p_d} \sum_{q_d=0}^{v_d-p_d} \binom{v_d-p_d}{q_d} \left(-\frac{1}{\Delta}\right)^{v_d-p_d-q_d} \times E_n\left(\prod_{j=1}^d \mathcal{B}_j^{v_j-p_j-q_j}(\Delta) \times \mathcal{J}_{(q_0, q_1, q_2, \dots, q_d), \mathbf{f}}(\Delta)\right) \tag{C.10}$$

where  $q_0 = l - \sum_{k=1}^d q_k$ .

**Proof.** See Section 2.5 of Li and Chen (2016). □

To illustrate for Proposition 1, we employ the following two conditional expectations for calculating (46) based on



(C.4a)–(C.4d):

$$\begin{aligned}
 & E_1(I_{(0,1),(1,1)}(\Delta)|W(\Delta) = w) \\
 &= -\frac{1}{\Delta} E_1(\mathcal{B}(\Delta) \times \mathcal{I}_{(2,0),(1,1)}(\Delta)) + E_1(\mathcal{I}_{(1,1),(1,1)}(\Delta)) \\
 &\quad + \frac{w}{\Delta} E_1(\mathcal{I}_{(2,0),(1,1)}(\Delta)), \tag{C.11a}
 \end{aligned}$$

$$\begin{aligned}
 & E_1(I_{(1),(J(t))}(\Delta)|W(\Delta) = w) \\
 &= -\frac{1}{\Delta} E_1(\mathcal{B}(\Delta) \times \mathcal{I}_{(1,0),(J(t))}(\Delta)) + E_1(\mathcal{I}_{(0,1),(J(t))}(\Delta)) \\
 &\quad + \frac{w}{\Delta} E_1(\mathcal{I}_{(1,0),(J(t))}(\Delta)), \tag{C.11b}
 \end{aligned}$$

as well as an example for calculating (47b) based on (C.5):

$$\begin{aligned}
 & E_2(I_{(1,1),(J(t)J(t))}(\Delta)|W(\Delta) = w) \\
 &= \frac{1}{\Delta^2} E_2(\mathcal{B}^2(\Delta) \times \mathcal{I}_{(2,0),\mathbf{f}}(\Delta)) - \frac{2}{\Delta} E_2(\mathcal{B}(\Delta) \times \mathcal{I}_{(1,1),\mathbf{f}}(\Delta)) \\
 &\quad - \frac{2w}{\Delta^2} E_2(\mathcal{B}(\Delta) \times \mathcal{I}_{(2,0),\mathbf{f}}(\Delta)) + E_2(\mathcal{I}_{(0,2),\mathbf{f}}(\Delta)) \\
 &\quad + \frac{2w}{\Delta} E_2(\mathcal{I}_{(1,1),\mathbf{f}}(\Delta)) + \left(\frac{w}{\Delta}\right)^2 E_2(\mathcal{I}_{(2,0),\mathbf{f}}(\Delta)), \tag{C.12}
 \end{aligned}$$

where  $\mathbf{f} = \{J(t), J(t)\}$ .

Next, it suffices to calculate the expectation in (C.10). We propose the following iteration-based algorithm.

**Proposition 2.** For non-negative integers  $k_1, k_2, \dots, k_d = 0, 1, 2, \dots, \mathbf{v}_i = (v_0, v_1, v_2, \dots, v_d)$  and  $\mathbf{f} = \{f_1(t), f_2(t), \dots, f_i(t)\}$ , if there exists  $i \in \{1, 2, \dots, d\}$  such that  $v_i > k_i$ , we have

$$E_n\left(\prod_{j=1}^d \mathcal{B}_j^{k_j}(\Delta) \times \mathcal{I}_{\mathbf{v}_i, \mathbf{f}}(\Delta)\right) = 0; \tag{C.13}$$

if there exists  $i \in \{1, 2, \dots, d\}$  such that  $v_i \leq k_i$ , we have

$$\begin{aligned}
 & E_n\left(\prod_{j=1}^d \mathcal{B}_j^{k_j}(\Delta) \times \mathcal{I}_{\mathbf{v}_i, \mathbf{f}}(\Delta)\right) \\
 &= v_i \times E_n\left(\mathcal{B}_i^{k_i-1}(\Delta) \prod_{j=1, j \neq i}^d \mathcal{B}_j^{k_j}(\Delta) \right. \\
 &\quad \times \mathcal{I}_{(v_0+1, v_1, \dots, v_{i-1}, v_i-1, v_{i+1}, \dots, v_d), \mathbf{f}}(\Delta)) \\
 &\quad + \sum_{j=1}^{i+1} E_n\left(\mathcal{B}_i^{k_i-1}(\Delta) \prod_{q=1, q \neq i}^d \mathcal{B}_q^{k_q}(\Delta) \right. \\
 &\quad \times \mathcal{I}_{(v_0, v_1, \dots, v_{i-1}, v_i+1, v_{i+1}, \dots, v_d), (f_1, f_2, \dots, f_{j-1}, 1, f_j, \dots, f_i)}(\Delta)); \tag{C.14}
 \end{aligned}$$

in particular, if  $v_i = k_i$ , we have

$$\begin{aligned}
 & E_n\left(\prod_{j=1}^d \mathcal{B}_j^{k_j}(\Delta) \times \mathcal{I}_{\mathbf{v}_i, \mathbf{f}}(\Delta)\right) \\
 &= k_i! \times E_n\left(\prod_{j=1, j \neq i}^d \mathcal{B}_j^{k_j}(\Delta) \times \mathcal{I}_{(v_0+v_i, v_1, \dots, v_{i-1}, 0, v_{i+1}, \dots, v_d), \mathbf{f}}(\Delta)\right). \tag{C.15}
 \end{aligned}$$

**Proof.** See Section 2.6 of Li and Chen (2016).  $\square$

Following (C.11a) and (C.11b), we illustrate Proposition 2 using the following examples among others:

$$\begin{aligned}
 & E_1(\mathcal{B}(\Delta) \times \mathcal{I}_{(2,0),(1,1)}(\Delta)) = 3E_1(\mathcal{I}_{(2,1),(1,1,1)}(\Delta)) = 0, \\
 & E_1(\mathcal{B}(\Delta) \times \mathcal{I}_{(1,0),(J(t))}(\Delta)) = E_1(\mathcal{I}_{(1,1),(1,J(t))}(\Delta)) \\
 &\quad + E_1(\mathcal{I}_{(1,1),(J(t),1)}(\Delta)) = 0, \\
 & E_1(\mathcal{I}_{(2,0),(1,1)}(\Delta)) = \int_0^\Delta \int_0^{t_1} dt_2 dt_1 = \frac{1}{2} \Delta^2 \\
 & E_1(\mathcal{I}_{(1,0),(J(t))}(\Delta)) = \int_0^\Delta J(t_1) dt_1 = \int_0^\Delta Z_1 1_{[\tau_1, \Delta]}(t_1) dt_1 \\
 &\quad = Z_1(\Delta - \tau_1).
 \end{aligned}$$

Similarly, following (C.12), we provide the following example:

$$\begin{aligned}
 & E_2(\mathcal{B}^2(\Delta) \times \mathcal{I}_{(2,0),(J(t)J(t))}(\Delta)) \\
 &= E_2(\mathcal{B}(\Delta) \times \mathcal{I}_{(2,1),(1J(t)J(t))}(\Delta)) \\
 &\quad + E_2(\mathcal{B}(\Delta) \times \mathcal{I}_{(2,1),(J(t),1J(t))}(\Delta)) \\
 &\quad + E_2(\mathcal{B}(\Delta) \times \mathcal{I}_{(2,1),(J(t),J(t),1)}(\Delta)),
 \end{aligned}$$

where, for instance,

$$\begin{aligned}
 & E_2(\mathcal{B}(\Delta) \times \mathcal{I}_{(2,1),(J(t),1J(t))}(\Delta)) \\
 &= \int_0^\Delta \int_0^{t_1} \int_0^{t_2} J(t_3) dt_3 dt_2 J(t_1) dt_1 \\
 &= -\frac{1}{6} Z_1^2 (\tau_1 - \tau_2)^3 - \frac{1}{2} Z_1 (Z_1 + Z_2) (T - \tau_1) (T - \tau_2) (\tau_1 - \tau_2) \\
 &\quad + \frac{1}{6} (Z_1 + Z_2)^2 (\Delta - \tau_2)^3.
 \end{aligned}$$

#### C.4. Summary of the algorithm

Before closing this section, we summarize the aforementioned three-stage algorithm as follows:

- Algorithm 1.** • Convert the multiplications of iterated Stratonovich integrals to linear combinations.  
 • Convert each iterated Stratonovich integral resulting from the previous step into a linear combination of iterated Itô integrals.  
 • Compute conditional expectation of iterated Itô integrals.

According to the above discussions in this section, conditional expectation (40) can be calculated as a polynomial in  $w$  with coefficients involving polynomials in the jump arrival times  $\tau_1, \tau_2, \dots, \tau_n$  as well as jump sizes  $Z_1, Z_2, \dots, Z_n$ . For example, conditional expectation (46) admits the following closed-form

$$\begin{aligned}
 P_{1,(1),(1)}(w) &= \frac{1}{2} b^{(1)}(x_0) \sigma(x_0) \Delta w + \frac{1}{2} b(x_0) \sigma^{(1)}(x_0) \Delta w \\
 &\quad + \frac{1}{2} b(x_0) b^{(1)}(x_0) \Delta^2 - b^{(1)}(x_0) \tau_1 Z_1 \\
 &\quad + b^{(1)}(x_0) \Delta Z_1 + \frac{1}{2} \sigma(x_0) \sigma^{(1)}(x_0) w^2 \\
 &\quad + \sigma^{(1)}(x_0) Z_1 w - \frac{1}{\Delta} \sigma^{(1)}(x_0) \tau_1 Z_1 w.
 \end{aligned}$$

#### Appendix D. Calculation of (42)

Following the investigation from the previous section, it is evident that (43) can be calculated as a product of the normal

p.d.f.  $\phi_{\Sigma(x_0)}(z)$  and a polynomial in  $z$  with coefficients involving polynomials in the jump arrival times  $\tau_1, \tau_2, \dots, \tau_n$  as well as jump sizes  $Z_1, Z_2, \dots, Z_n$ . For example, (45) admits the following explicit form

$$\begin{aligned}
 &F_{1,(1),(1),(1)}(z) \\
 &= \phi(z) \left( \frac{1}{2} b^{(1)}(x_0) \sigma(x_0) \Delta^{3/2} + \frac{1}{2} b(x_0) \sigma^{(1)}(x_0) \Delta^{3/2} \right. \\
 &\quad - \frac{1}{2} b^{(1)}(x_0) \sigma(x_0) \Delta^{3/2} z^2 - \frac{1}{2} b(x_0) \sigma^{(1)}(x_0) \Delta^{3/2} z^2 \\
 &\quad + b^{(1)}(x_0) \tau_1 Z_1 z - \frac{1}{2} b(x_0) b^{(1)}(x_0) \Delta^2 z - b^{(1)}(x_0) \Delta Z_1 z \\
 &\quad - \frac{1}{2} \sigma(x_0) \sigma^{(1)}(x_0) \Delta z^3 - \sigma^{(1)}(x_0) \sqrt{\Delta} Z_1 z^2 \\
 &\quad + \sigma(x_0) \sigma^{(1)}(x_0) \Delta z + \sigma^{(1)}(x_0) \sqrt{\Delta} Z_1 \\
 &\quad \left. + \frac{\sigma^{(1)}(x_0) \tau_1 Z_1 z^2}{\sqrt{\Delta}} - \frac{\sigma^{(1)}(x_0) \tau_1 Z_1}{\sqrt{\Delta}} \right). \tag{D.1}
 \end{aligned}$$

To obtain correction term (42) under any arbitrary specifications of jump-size distribution, we need to consider the following type of expectation

$$\begin{aligned}
 &\mathbb{E} \left( \prod_{i=1}^n \tau_i^{a_i} \prod_{l=1}^d \prod_{k=1}^n Z_{k,l}^{b_{k,l}} \right. \\
 &\quad \left. \times \phi_{\Sigma(x_0)} \left( y - \frac{D(x_0)}{\sqrt{\Delta}} (b(x_0) \Delta + J(\Delta)) \right) \mid N(\Delta) = n \right), \tag{D.2}
 \end{aligned}$$

where  $a_i$  and  $b_{k,l}$  are all non-negative integers for  $i, k = 1, 2, \dots, n$  and  $l = 1, 2, \dots, d$ ;  $\phi_{\Sigma(x_0)}$  is the p.d.f. of a normal distribution with zero mean and covariance matrix  $\Sigma(x_0)$  defined in (26). By the independence between  $(\tau_1, \tau_2, \dots, \tau_n)$  and  $(Z_1, Z_2, \dots, Z_n)$ , the calculation of (D.2) boils down to calculating expectations of monomial of jump arrival times as well as expectations involving jump sizes in the following form, i.e., (D.2) equals to

$$\begin{aligned}
 &\mathbb{E} \left( \prod_{i=1}^n \tau_i^{a_i} \mid N(\Delta) = n \right) \\
 &\quad \times \mathbb{E} \left( \prod_{l=1}^d \prod_{k=1}^n Z_{k,l}^{b_{k,l}} \phi_{\Sigma(x_0)}(A + BJ(\Delta)) \mid N(\Delta) = n \right),
 \end{aligned}$$

where  $A$  and  $B$  are defined in (A.3).

### D.1. Expectation of monomial of jump arrival times

We note that the conditional distribution of arrival times follows an order statistics associated with a sample of independent uniformly distributed random variables (see, e.g., Chapter 13 in Karlin and Taylor (1981)), i.e.,

$$\begin{aligned}
 &\mathbb{P} \left( \bigcap_{i=1}^n \{\tau_i = t_i\} \mid N(\Delta) = n \right) \\
 &= \frac{n!}{\Delta^n} 1\{0 \leq t_1 \leq \dots \leq t_n \leq \Delta\} dt_1 dt_2 \dots dt_n, \tag{D.3}
 \end{aligned}$$

Thus, we provide a general formula for the conditional expectation of a monomial of jump arrival times  $\tau_1^{p_1} \tau_2^{p_2} \dots \tau_r^{p_r}$  in the following lemma, which follows directly from classical calculus.

**Lemma 5.** For  $r \leq n$  and  $1 \leq i_1 < i_2 < \dots < i_r \leq n$ , we have:

$$\begin{aligned}
 &\mathbb{E} \left( \tau_{i_1}^{p_1} \tau_{i_2}^{p_2} \dots \tau_{i_r}^{p_r} \mid N(\Delta) = n \right) = \left( \prod_{k=1}^r (n+k) \right)^{-1} \\
 &\quad \times \prod_{k=2}^r \prod_{h=1}^{p_k} \left( i_k - 1 + h + \sum_{j=1}^{k-1} p_j \right) \times \prod_{h=1}^{p_1} (i_1 - 1 + h) \Delta^P,
 \end{aligned}$$

where  $P = \sum_{j=1}^r p_j$ .

### D.2. Expectation involving jump sizes

In this section, we resort to the technique of conditioning in the calculation of

$$\mathbb{E} \left( \prod_{l=1}^d \prod_{k=1}^n Z_{k,l}^{b_{k,l}} \phi_{\Sigma(x_0)}(A + BJ(\Delta)) \mid N(\Delta) = n \right). \tag{D.4}$$

For example, by conditioning on  $J(\Delta)$ , expectation (D.4) is equal to

$$\begin{aligned}
 &\mathbb{E} \left[ \phi_{\Sigma(x_0)}(A + BJ(\Delta)) \right. \\
 &\quad \left. \times \mathbb{E} \left( \prod_{l=1}^d \prod_{k=1}^n Z_{k,l}^{b_{k,l}} \mid J(\Delta), N(\Delta) = n \right) \mid N(\Delta) = n \right]. \tag{D.5}
 \end{aligned}$$

To calculate the inside conditional expectation, we note that, for each  $l$ ,

$$\begin{aligned}
 &\mathbb{E} \left( \prod_{k=1}^n Z_{k,l}^{b_{k,l}} \mid J_l(\Delta), N(\Delta) = n \right) \\
 &= \prod_{k=1}^n \mathbb{E} \left( Z_{k,l}^{b_{k,l}} \mid J_l(\Delta), N(\Delta) = n \right), \tag{D.6}
 \end{aligned}$$

where  $J_l(\Delta) = \sum_{i=1}^{N(\Delta)} Z_{i,l}$ . The conditional distribution  $(Z_{1,l}, Z_{2,l}, \dots, Z_{n,l} \mid J_l(\Delta), N(\Delta) = n)$  can be explicitly calculated under the assumptions of Jump-Size Distributions 1 and 2. For other specifications of jump-size distribution, case-by-case analysis on how conditioning techniques can be applied in the calculation of (D.4) is needed. In Sections 3 and 4 of Li and Chen (2016), as illustrations, we document the main technical details on handling a correlated multivariate normal distribution and a correlated mixture of normal and exponential distributions for modeling jumps in Model 5 (the CEV-SVCJ model).

#### D.2.1. An example under Jump-Size Distribution 1

Under Jump-Size Distribution 1, we have

$$\begin{aligned}
 &((Z_{1,l}, Z_{2,l}, \dots, Z_{n,l}) \mid J_l(\Delta), N(\Delta) = n) \\
 &\stackrel{\text{in law}}{=} \left( \mathbf{N} \left( \frac{J_l(\Delta)}{n} \mathbf{1}_n, \beta_l^2 \left( I_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^\top \right) \right) \mid J_l(\Delta), N(\Delta) = n \right), \tag{D.7}
 \end{aligned}$$

where  $I_n = \text{diag}\{1, 1, \dots, 1\}$  and  $\mathbf{1}_n = (1, 1, \dots, 1)^\top$ . Thus, it is straightforward to obtain the expectation in (D.6) in closed-form as a polynomial of  $J_l(\Delta)$  by differentiating the moment generating function of conditional distribution (D.7), that is,

$$\begin{aligned}
 &\mathbb{E} \left( \prod_{k=1}^n Z_{k,l}^{b_{k,l}} \mid J_l(\Delta), N(\Delta) = n \right) \\
 &= \frac{\partial^{(b_{1,l})} \partial^{(b_{2,l})} \dots \partial^{(b_{n,l})}}{\partial \theta_1^{b_{1,l}} \partial \theta_2^{b_{2,l}} \dots \partial \theta_n^{b_{n,l}}} \varphi(\theta_1, \theta_2, \dots, \theta_n) \mid_{\theta_1, \theta_2, \dots, \theta_n = 0},
 \end{aligned}$$

where

$$\varphi(\theta_1, \theta_2, \dots, \theta_n) = \exp\left(\frac{J_l^{(n)}(\Delta)}{n} \sum_{i=1}^n \theta_i + \frac{1}{2} \beta_l^2 \left(\sum_{i=1}^n \theta_i^2 - \frac{1}{n} \left(\sum_{i=1}^n \theta_i\right)^2\right)\right)$$

with  $J_l^{(n)}(\Delta) = \sum_{i=1}^n Z_{i,l}$ . Thus, expectation (D.4) can be further calculated in closed-form using Lemma 1.

D.2.2. An example under Jump-Size Distribution 2

Under Jump-Size Distribution 2, it is known from Section 13.1 of Karlin and Taylor (1981) that

$$(Z_{1,l}, Z_{2,l}, \dots, Z_{n-1,l}, Z_{n,l} | J_l(\Delta), N(\Delta) = n) \stackrel{\text{in law}}{=} (U_{1,l} J_l^{(n)}(\Delta), U_{2,l} J_l^{(n)}(\Delta), \dots, U_{n-1,l} J_l^{(n)}(\Delta), U_{n,l} J_l^{(n)}(\Delta)),$$

where  $U_{1,l} + U_{2,l} + \dots + U_{n-1,l} + U_{n,l} = 1$  and  $U_{1,l}, U_{2,l}, \dots, U_{n-1,l}$  are uniformly distributed over the region

$$\mathcal{U}_{n-1} := \left\{ (u_1, u_2, \dots, u_{n-1}) : u_i \geq 0, \text{ for } i = 1, 2, \dots, n-1, \text{ and } \sum_{i=1}^{n-1} u_i \leq 1 \right\},$$

i.e., the distribution of  $(U_{1,l}, U_{2,l}, \dots, U_{n-1,l})$  has a probability density function as

$$P(U_{1,l} \in du_1, U_{2,l} \in du_2, \dots, U_{n-1,l} \in du_{n-1}) = (n-1)! 1_{\{(u_1, u_2, \dots, u_{n-1}) \in \mathcal{U}_{n-1}\}}; \tag{D.8}$$

moreover, the random vector  $(U_{1,l}, U_{2,l}, \dots, U_{n-1,l})$  is independent of  $J_l(\Delta)$ .

Thus, we find that

$$\begin{aligned} & \mathbb{E} \left( \prod_{k=1}^n Z_{k,l}^{b_{k,l}} | J_l(\Delta), N(\Delta) = n \right) \\ &= J_l^{(n)}(\Delta) \sum_{k=1}^n b_{k,l} \mathbb{E} \left( \prod_{k=1}^n U_{k,l}^{b_{k,l}} | J_l(\Delta), N(\Delta) = n \right) \\ &= J_l^{(n)}(\Delta) \sum_{k=1}^n b_{k,l} \mathbb{E} \left( \prod_{k=1}^n U_{k,l}^{b_{k,l}} | N(\Delta) = n \right), \end{aligned}$$

which simplifies (D.4) to

$$\begin{aligned} & \mathbb{E} \left[ \phi_{\Sigma(x_0)}(A + BJ(\Delta)) \prod_{l=1}^d J_l(\Delta)^{\sum_{k=1}^n b_{k,l}} | N(\Delta) = n \right] \\ & \times \prod_{l=1}^d \mathbb{E} \left( \prod_{k=1}^n U_{k,l}^{b_{k,l}} | N(\Delta) = n \right). \tag{D.9} \end{aligned}$$

By letting  $n_l = \sum_{k=1}^n b_{k,l}$  for  $l = 1, 2, \dots, d$ , it is easy to obtain the first conditional expectation in (D.9) in closed-form using Lemma 2. Through straightforward calculations relying on distribution (D.8), we obtain that

$$\mathbb{E} \left( \prod_{k=1}^n U_{k,l}^{b_{k,l}} | N(\Delta) = n \right) = \frac{(n-1)! \prod_{k=1}^n b_{k,l}}{\left(\sum_{k=1}^n b_{k,l} + n - 1\right)!}.$$

D.3. An explicit illustration of correction term (42)

As an illustration, we obtain the correction term  $T_{1,1}(y)$  for a one-dimensional jump-diffusion model by integrating (D.1) using the aforementioned method:

$$\begin{aligned} T_{1,1}(y) &= -\frac{D(x_0)}{\sqrt{\Delta}} \left( \left( \frac{1}{2} \Delta^{3/2} (1 - A^2) (b^{(1)}(x_0) \sigma(x_0) + b(x_0) \sigma^{(1)}(x_0)) \right. \right. \\ & \quad \left. \left. + A \left( \sigma(x_0) \sigma^{(1)}(x_0) \Delta - \frac{1}{2} b(x_0) b^{(1)}(x_0) \Delta^2 \right) \right. \right. \\ & \quad \left. \left. - \frac{1}{2} \sigma(x_0) \sigma^{(1)}(x_0) A^3 \Delta \right) \mathbb{E}[\phi(A + BJ(\Delta)) | N(\Delta) = 1] \right. \\ & \quad \left. - \left( \frac{1}{\sqrt{\Delta}} \left( (1 - A^2) \sigma^{(1)}(x_0) + \sqrt{\Delta} A b^{(1)}(x_0) \right) \right. \right. \\ & \quad \left. \left( \mathbb{E}[\tau_1 | N(\Delta) = 1] - \Delta \right) - 2A b^{(1)}(x_0) \right. \\ & \quad \left. \left( \mathbb{E}[\tau_1 | N(\Delta) = 1] - \Delta \right) + AB (b^{(1)}(x_0) \sigma(x_0) \right. \right. \\ & \quad \left. \left. + b(x_0) \sigma^{(1)}(x_0) \right) \Delta^{3/2} - \left( 1 - \frac{3}{2} A^2 \right) B \sigma(x_0) \sigma^{(1)}(x_0) \Delta \right. \\ & \quad \left. + \frac{1}{2} B b(x_0) b^{(1)}(x_0) \Delta^2 \right) \mathbb{E}[\phi(A + BJ(\Delta)) J(\Delta) | N(\Delta) = 1] \\ & \quad + \left( \left( b^{(1)}(x_0) + \frac{2}{\sqrt{\Delta}} A \sigma^{(1)}(x_0) \right) B \left( \mathbb{E}[\tau_1 | N(\Delta) = 1] - \Delta \right) \right. \\ & \quad \left. - \frac{1}{2} \Delta^{3/2} B^2 (b^{(1)}(x_0) \sigma(x_0) + b(x_0) \sigma^{(1)}(x_0)) \right. \\ & \quad \left. - \frac{1}{2} \sigma(x_0) \sigma^{(1)}(x_0) 3AB^2 \Delta \right) \\ & \quad \times \mathbb{E}[\phi(A + BJ(\Delta)) J(\Delta)^2 | N(\Delta) = 1] \\ & \quad + \frac{1}{\sqrt{\Delta}} B^2 \sigma^{(1)}(x_0) \left( \mathbb{E}[\tau_1 | N(\Delta) = 1] - \Delta \right) \\ & \quad \left. - \frac{1}{2} \sigma(x_0) B \Delta^{3/2} \right) \mathbb{E}[\phi(A + BJ(\Delta)) J(\Delta)^3 | N(\Delta) = 1], \end{aligned}$$

with  $A$  and  $B$  defined in (A.3). Here, the expectation of the jump arrival time is given by  $\mathbb{E}[\tau_1 | N(\Delta) = 1] = \Delta/2$ . Under Jump-Size Distribution 1, the expectations involving jump size components can be calculated as

$$\begin{aligned} \mathbb{E}[\phi(A + BJ(\Delta)) | N(\Delta) = 1] &= \frac{1}{\sqrt{1 + B^2 \beta^2}} \phi\left(\frac{A + B\alpha}{\sqrt{1 + B^2 \beta^2}}\right), \\ \mathbb{E}[\phi(A + BJ(\Delta)) J(\Delta) | N(\Delta) = 1] &= \frac{\alpha - AB\beta^2}{(1 + B^2 \beta^2)^{3/2}} \phi\left(\frac{A + B\alpha}{\sqrt{1 + B^2 \beta^2}}\right), \\ \mathbb{E}[\phi(A + BJ(\Delta)) J(\Delta)^2 | N(\Delta) = 1] &= \frac{\alpha^2 + \beta^2 - 2AB\alpha\beta^2 + (1 + A^2) B^2 \beta^4}{(1 + B^2 \beta^2)^{5/2}} \phi\left(\frac{A + B\alpha}{\sqrt{1 + B^2 \beta^2}}\right), \\ \mathbb{E}[\phi(A + BJ(\Delta)) J(\Delta)^3 | N(\Delta) = 1] &= \frac{(\alpha - AB\beta^2)(\alpha^2 + (3 - 2AB\alpha)\beta^2 + (3 + A^2) B^2 \beta^4)}{(1 + B^2 \beta^2)^{7/2}} \\ & \quad \times \phi\left(\frac{A + B\alpha}{\sqrt{1 + B^2 \beta^2}}\right), \end{aligned}$$

where  $\alpha$  and  $\beta$  are defined via  $\alpha = EZ_1$  and  $\beta^2 = \text{Var}Z_1$  according to Jump-Size Distribution 1. Under Jump-Size

**Distribution 2**, the expectations involving jump size components can be calculated as

$$\mathbb{E}[\phi(A + BJ(\Delta)) | N(\Delta) = 1] = \frac{\gamma}{B} e^{\frac{A\gamma}{B} + \frac{\gamma^2}{2B^2}} \mathcal{N}(-\widehat{A}),$$

$$\begin{aligned} \mathbb{E}[\phi(A + BJ(\Delta))J(\Delta) | N(\Delta) = 1] \\ = -\frac{\gamma}{B^2} e^{\frac{A\gamma}{B} + \frac{\gamma^2}{2B^2}} (\widehat{A}\mathcal{N}(-\widehat{A}) - \phi(-\widehat{A})), \end{aligned}$$

$$\begin{aligned} \mathbb{E}[\phi(A + BJ(\Delta))J(\Delta)^2 | N(\Delta) = 1] \\ = \frac{\gamma}{B^3} e^{\frac{A\gamma}{B} + \frac{\gamma^2}{2B^2}} ((\widehat{A}^2 + 1)\mathcal{N}(-\widehat{A}) - \widehat{A}\phi(-\widehat{A})), \end{aligned}$$

$$\begin{aligned} \mathbb{E}[\phi(A + BJ(\Delta))J(\Delta)^3 | N(\Delta) = 1] \\ = -\frac{\gamma}{B^4} e^{\frac{A\gamma}{B} + \frac{\gamma^2}{2B^2}} (3(\widehat{A}^3 + \widehat{A})\mathcal{N}(-\widehat{A}) - (\widehat{A}^2 + 2)\phi(-\widehat{A})), \end{aligned}$$

where  $\widehat{A} = A + \gamma/B$  and  $\gamma$  is the intensity parameter of the exponential distribution. Here,  $\mathcal{N}(\cdot)$  denotes the cumulative distribution function of a standard normal distribution.

## References

- Abate, J., Whitt, W., 1992. The Fourier-series method for inverting transforms of probability distributions. *Queueing Syst. Theory Appl.* 10, 5–87.
- Ait-Sahalia, Y., 1999. Transition densities for interest rate and other nonlinear diffusions. *J. Finance* 54, 1361–1395.
- Ait-Sahalia, Y., 2002a. Maximum-likelihood estimation of discretely-sampled diffusions: a closed-form approximation approach. *Econometrica* 70, 223–262.
- Ait-Sahalia, Y., 2002b. Telling from discrete data whether the underlying continuous-time model is a diffusion. *J. Finance* 57, 2075–2112.
- Ait-Sahalia, Y., 2004. Disentangling diffusion from jumps. *J. Financ. Econ.* 74, 487–528.
- Ait-Sahalia, Y., 2008. Closed-form likelihood expansions for multivariate diffusions. *Ann. Statist.* 36, 906–937.
- Ait-Sahalia, Y., Cacho-Diaz, J., Hurd, T., 2009. Portfolio choice with a jumps: A closed form solution. *Ann. Appl. Probab.* 19, 556–584.
- Ait-Sahalia, Y., Cacho-Diaz, J., Laeven, R.J., 2015. Modeling financial contagion using mutually exciting jump processes. *J. Financ. Econ.* 117, 585–606.
- Ait-Sahalia, Y., Hurd, T., 2015. Portfolio choice in markets with contagion. *J. Financ. Econ.* 14, 1–28.
- Ait-Sahalia, Y., Kimmel, R., 2007. Maximum likelihood estimation of stochastic volatility models. *J. Financ. Econ.* 83, 413–452.
- Ait-Sahalia, Y., Kimmel, R., 2010. Estimating affine multifactor term structure models using closed-form likelihood expansions. *J. Financ. Econ.* 98, 113–144.
- Ait-Sahalia, Y., Mykland, P., 2003. The effects of random and discrete sampling when estimating continuous-time diffusions. *Econometrica* 71, 483–549.
- Ait-Sahalia, Y., Mykland, P., 2004. Estimators of diffusions with randomly spaced discrete observations: a general theory. *Ann. Statist.* 32, 2186–2222.
- Ait-Sahalia, Y., Yu, J., 2006. Saddlepoint approximations for continuous-time markov processes. *J. Econometrics* 134, 507–551.
- Andersen, T., Benzoni, L., Lund, J., 2002. An empirical investigation of continuous-time equity return models. *J. Finance* 57, 1239–1284.
- Bakshi, G., Ju, N., Ou-Yang, H., 2006. Estimation of continuous-time models with an application to equity volatility dynamics. *J. Financ. Econ.* 82, 227–249.
- Bates, D.S., 1996. Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche mark options. *Rev. Financ. Stud.* 9, 69–107.
- Bates, D.S., 2000. Post-'87 crash fears in the S&P 500 futures option market. *J. Econometrics* 94, 181–238.
- Black, F., Scholes, M.S., 1973. The pricing of options and corporate liabilities. *J. Polit. Econ.* 81, 637–654.
- Broadie, M., Chernov, M., Johannes, M., 2007. Model specification and risk premia: Evidence from futures options. *J. Finance* 62, 1453–1490.
- Chang, J., Chen, S., 2011. On the approximate maximum likelihood estimation for diffusion processes. *Ann. Statist.* 39, 2820–2851.
- Cheridito, P., Filipović, D., Kimmel, R.L., 2007. Market price of risk specifications for affine models: Theory and evidence. *J. Financ. Econ.* 83, 123–170.
- Chernov, M., Gallant, R., Ghysels, E., Tauchen, G., 2003. Alternative models for stock price dynamics. *J. Econometrics* 116, 225–257.
- Choi, S., 2013. Closed-form likelihood expansions for multivariate time-inhomogeneous diffusions. *J. Econometrics* 174, 45–65.
- Choi, S., 2015a. Approximate Transition Probability Density Function of a Multivariate Time-inhomogeneous Jump Diffusion Process in a Closed-form Expression. Working Paper, University of Seoul.
- Choi, S., 2015b. Explicit form of approximate transition probability density functions of diffusion processes. *J. Econometrics* 187, 57–73.
- Cinlar, E., Jacod, J., 1981. Representation of semimartingale Markov processes in terms of Wiener processes and Poisson random measures. In: *Seminar on Stochastic Processes*, Birkhauser, Boston.
- Dipietro, M., 2001. Bayesian Inference for Discretely Sampled Diffusion Processes with Financial Applications (Ph.D. thesis), Carnegie Mellon University.
- Duffie, D., Gârleanu, N., 2001. Risk and valuation of collateralized debt obligations. *Financ. Anal. J.* 57, 41–59.
- Duffie, D., Pan, J., Singleton, K., 2000. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica* 68, 1343–1376.
- Egorov, A.V., Li, H., Xu, Y., 2003. Maximum likelihood estimation of time-inhomogeneous diffusions. *J. Econometrics* 114, 107–139.
- Errais, E., Giesecke, K., Goldberg, L., 2010. Affine point processes and portfolio credit risk. *SIAM J. Financial Math.* 1, 642–665.
- Filipović, D., Mayerhofer, E., Schneider, P., 2013. Density approximations for multivariate affine jump-diffusion processes. *J. Econometrics* 176, 93–111.
- Giesecke, K., Kakavand, H., Mousavi, M., 2011. Exact simulation of point processes with stochastic intensities. *Oper. Res.* 59, 1233–1245.
- Giesecke, K., Schwenkler, G., Filtered Likelihood for Point Processes. Working Paper, Stanford University, 2011.
- Giesecke, K., Smelov, D., 2013. Exact sampling of jump diffusions. *Oper. Res.* 61, 894–907.
- Glasserman, P., 2004. *Monte Carlo Methods in Financial Engineering*. Springer-Verlag, New York.
- Hall, P., 1995. *The Bootstrap and Edgeworth Expansion*. Springer-Verlag, New York.
- Hawkes, A., 1971. Spectra of some self-exciting and mutually exciting point processes. *Biometrika* 58, 83–90.
- Hayashi, M., Ishikawa, Y., 2012. Composition with distributions of Wiener-Poisson variables and its asymptotic expansion. *Math. Nachr.* 285, 619–658.
- Ikeda, N., Watanabe, S., 1989. *Stochastic Differential Equations and Diffusion Processes*, second ed. Elsevier, North-Holland.
- Jin, X., Zhang, A.X., 2012. Decomposition of optimal portfolio weight in a jump-diffusion model and its applications. *Rev. Financ. Stud.* 25, 2877–2919.
- Johannes, M.S., Eraker, B., Polson, N., 2003. The impact of jumps in volatility and returns. *J. Finance* 53, 1269–1300.
- Johnson, N.L., Kotz, S., Balakrishnan, N., 1996. *Continuous Univariate Distributions* 1. Wiley, New York.
- Kanwal, R.P., 2004. *Generalized Functions: Theory and Applications*, third ed. Birkhäuser, Boston.
- Karatzas, I., Shreve, S.E., 1991. *Brownian Motion and Stochastic Calculus*, second edition. Springer-Verlag, New York.
- Karlin, S., Taylor, H., 1981. *A Second Course in Stochastic Processes*. Academic Press, Boston.
- Kloeden, P.E., Platen, E., 1992. *The Numerical Solution of Stochastic Differential Equations*. Springer-Verlag, Berlin Heidelberg.
- Kou, S., 2002. A jump diffusion model for option pricing. *Manage. Sci.* 48, 1086–1101.
- Li, C., 2013. Maximum-likelihood estimation for diffusion processes via closed-form density expansions. *Ann. Statist.* 41, 1350–1380.
- Li, C., Chen, D., Estimating jump-diffusions using closed-form likelihood expansions—online supplementary material. SSRN eLibrary, 2016, <http://ssrn.com/abstract=2807257>.
- Liu, J., Longstaff, F., Pan, J., 2003. Dynamic asset allocation with event risk. *J. Finance* 58, 231–259.
- Merton, R., 1976. Option pricing when underlying stock returns are discontinuous. *J. Financ. Econ.* 3, 125–144.
- Merton, R.C., 1973. Theory of rational option pricing. *Bell J. Econ.* 4, 141–183.
- Mykland, P., 1992. Asymptotic expansions and bootstrapping-distributions for dependent variables: a martingale approach. *Ann. Statist.* 20, 623–654.
- Mykland, P., 1993. Asymptotic expansions for martingales. *Ann. Probab.* 21, 800–818.
- Mykland, P., Zhang, L., 2010. The econometrics of high frequency data. In: Kessler, M., Lindner, A., Sørensen, M. (Eds.), *Statistical Methods for Stochastic Differential Equations*. Chapman & Hall/CRC Press, New York, pp. 109–190.
- Pan, J., 2002. The jump-risk premia implicit in options: evidence from an integrated time-series study. *J. Financ. Econ.* 63, 3–50.
- Pan, J., Liu, J., 2003. Dynamic derivative strategies. *J. Financ. Econ.* 69, 401–430.
- Protter, P., 1990. *Stochastic Integration and Differential Equations*. Springer-Verlag, Berlin Heidelberg.
- Schaumburg, E., 2001. Maximum Likelihood Estimation of Jump Processes with Applications to Finance (Ph.D. thesis), Princeton University.
- Singleton, K.J., 2006. *Empirical Dynamic Asset Pricing: Model Specification and Econometric Assessment*. Princeton University Press, Princeton.
- Stramer, O., Bognar, M., Schneider, P., 2010. Bayesian inference for discretely sampled markov processes with closed-form likelihood expansions. *J. Financ. Econ.* 8, 450–480.
- Watanabe, S., 1987. Analysis of Wiener functionals (Malliavin calculus) and its applications to heat kernels. *Ann. Probab.* 15, 1–39.
- Xiu, D., 2014. Hermite polynomial based expansion of european option prices. *J. Econometrics* 179, 158–177.
- Yoshida, N., 1992. Asymptotic expansions for statistics related to small diffusions. *J. Japan Stat. Soc.* 22, 139–159.
- Yu, J., 2007. Closed-form likelihood approximation and estimation of jump-diffusions with an application to the realignment risk of the chinese yuan. *J. Econometrics* 141, 1245–1280.