Small-time Expansion of Jump Diffusions and Applications in Option Pricing

Jingzhao Yuan*, Chenxu Li (Advisor)

Peking University Guanghua School of Management

Department of Finance

February 14th, 2022

Abstract

Stochastic volatility with or without jumps models are built to better portray the dynamics of financial derivatives. Handling these complicated parametric models typically raises two concerns: the option price formula and the econometric strategy. Reviewing an expansion method for stochastic differential equations with respect to the time increment, I theoretically demonstrate how it can be applied correspondingly in the expansion of marginal transition density (which enables implementation of MLE), at-the-money option price, and its sensitivities of arbitrary orders with respect to the "moneyness". I show that this applies to at least 4 general model specifications, which are widely used in option pricing: SV (continuous diffusion), SVJ or SVJJ model with jumps in volatility only, concurrent jumps, or jumps in both asset return and volatility.

*E-mail address: `jingzhao.yuan@columbia.edu`

Contents

1 Introduction

The application of stochastic differential equation (SDE hereafter) in pricing financial options whose payoff is non-linearly related with the underlying asset price proves fruitful. Since the groundbreaking Black-Scholes-Merton model ([Black and Scholes(1973)]), researchers have developed numerous stochastic models to describe the dynamics of the underlying asset, and, motivated by certain empirical regularities, e.g. the volatility smile and skewness documented in [Bates(2000)], added additional process to characterize a time-varying volatility. Stochastic volatility (SV hereafter) models are hence constructed (e.g. [Merton(1976)], [Heston(1993)], [Dupire(1994)], [Bakshi et al.(1997)Bakshi, Cao, and Chen] and [Bates(2000)]). Except a few simpler process, e.g. the one in [Black and Scholes(1973)], [Vasicek(1977)], and [Cox et al.(1985)Cox, Ingersoll, and Ross], most SDEs do not admit closed-form solution. The option price under those models does not either. Moreover, to estimate the parameters for these parametric models, relying on their Markovian nature, we are interested in the transition density. However, we generally have little knowledge of it except the Fokker-Planck equation and the Kolmogorov backward equation.

A natural idea is to search for approximation. In this paper, I first review a small-time expansion method proposed in [Li(2014)], and discuss its applications in multiple financial scenarios, including the expansion of transition density (which enables the implementation of MLE using discretely sampled data), option price and its sensitivities. The key references of this thesis are [Li(2013)], [Yu(2019)] and [Wang(2020)].

The thesis is organized as follows. Section 2 reviews several classical papers about analytic solutions of option price under stochastic volatility with or without jumps models, and the relevant econometric works. Section 3 elaborates on the important theorems proposed in [Li(2013)] which are developed for diffusion process (i.e., a strict Markov process with continuous paths). Mimicking the process of expanding the joint transition density therein, I provide an analogous proof of its marginal counterpart. The motivation for this is that, since in a SV model the volatility process is latent and unobservable, MLE is expected to be tailored for the mere data of asset return. Section 4,5 and 6 discuss option price expansions along the moneyness-dimension for SVJ, SVJJ (concurrent jump), and SVJJ (idiosyncratic jump) model respectively, which relies on the previously established theorems substantially. Section 7 briefly outlines a future research topic. Another important econometric design for diffusions, the GMM design by Hansen, is discussed in the Appendix.

2 Literature review

Generalizations to the Black-Scholes-Merton option pricing model have been made in various ways, while a not uncommon goal is to retain at least some degree of tractability. The tractability is in different aspects, e.g. an analytic solution to the stochastic differential equations system itself, a closed-form distribution or the moments of the solution, an analytic transform, an approximation of the stock price or option price and their sensitivities with respect to many variables, etc. A popular trend is to incorporate stochastic volatility in the basic model, an extension to the constant volatility assumed by Black-Scholes. Moreover, jumps are added to further capture certain characteristics exhibited by the market price in reality.

Literature on handling parametric SVJ model can be broadly split into 2 groups: one is to derive the option price formula (or sometimes stock price distribution, implied volatility, etc.); another is to design econometric method, i.e. how to estimate the parameter and perform statistical analysis.

An early work on the analytic solution of option price under SV model is [Stein and Stein(1991)], where the volatility is assumed to follow an Ornstein-Uhlenbeck process. An analytic density of the stock price is given, using elementary mathematical functions only. Besides, if the market price of risk is supposed to be zero, or equivalently, the model is already in risk-neutral measure, then the option price can be expressed explicitly as an integration with respect to the discounted payoff using the obtained density directly. The computational cost of this model is low because it only involves numerical integration routines (e.g. Richardson extrapolation and Romberg integration). However, a restriction is that, the asset and the volatility process are allowed no direct correlation, since they are driven by two independent Brownian motions. In addition, despite the plausible property of mean-reversion of volatility that Ornstein-Uhlenbeck process incorporates, it might not be a good approximation of volatility since there is a possibility, though rather small, for it to go negative.

[Heston(1993)] extends Stein and Stein's model by allowing the volatility to correlate with spot returns. Under Heston's SV model, the option price is known up to an inversion of Fourier transform. Approaching the problem with PDE technique, he suggested that additional information about the market price of volatility risk (if not assumed to be zero) is needed, a result of the dependence of option price on the volatility process in addition to the asset return. The pricing formula is similar to BSM one, though the 2 probabilities therein are not as simple as a cumulative density of $\mathcal{N}(0,1)$, but are given by an inversion of the respective characteristic functions, which can be derived explicitly.

Some later works generalize these results. [Duffie et al.(2000)Duffie, Pan, and Singleton] solves the option pricing problem explicitly and nicely if the SDE is specified as an affine jump diffusion (AJD) process. A Fourier-like transform in form of

$$\psi(v_0, v_1, u, X(t), t, T) = \mathbb{E}[e^{-\int_t^T r(X(s))ds} \cdot (v_0 + v_1 X(T)) \cdot e^{uX(T)} \mid \mathcal{F}(t)]$$

is derived in closed-form, with associated ordinary differential equation system governing the result. The option price is obtained from an inversion of the transform (an extension of Lévy inversion). This path to the option price from the transform is an accurate one, not an approximation, though the numerical calculation typically involves solving ODE numerically (e.g. Runge-Cutta) and quadrature. Motivated by the findings in [Bates(2000)] and [Bakshi et al.(1997)Bakshi, Cao, and Chen] that mere negative jumps in asset return fails to generate the desired level of skewness in the asset price distribution and the volatility smirk observed from market data, Duffie et al. also studied the impact of a "double jump" on implied volatility smiles, nesting the existing SV, SVJ-Y, SVJ-Y-V and SVJJ model in their AJD framework. They found that for out-of-the-money calls, the introduction of a jump in volatility lowers Black-Scholes implied volatilities. However, for models beyond "affine" class, e.g. the 3-factor model (with 2 dedicated to volatility) studied in [Gallant et al.(1999)Gallant, Hsu, and Tauchen], the transform is inapplicable.

Even though the pricing formula, explicit or implicit, is obtained for the aforementioned models, one needs to devise strategies to estimate the model parameters based on empirical observations. For continuous-time models, a natural difficulty is that only discrete observations on some time-stamps are available. Although diffusion process is Markov process, which allows the log-likelihood function to be broken down into small steps, the transition density is generally not explicitly available. Implementing MLE using discretely sampled data for continuous-time models inevitably results in an inconsistent estimator.

A distinguished work that bypasses transition density is [Hansen and Scheinkman(1995)], in which a Generalized Method of Moment (GMM hereafter) approach is employed. They used infinitesimal generators to constructed moment conditions implied by stationary Markov process, therefore successully nesting the estimation of diffusion process in the GMM framework. This design is very impressive and rather simple to implement. Therefore, though as a digression, an appendix of this text is devoted a detailed discussion about their GMM tailored to diffusions.

Nevertheless, econometricians have made progress in studying the possible analytic form of transition density. A seminal paper is [Aït-Sahalia(2002)], where a closed-form asymptotic expansion of the transition density of a scalar diffusion is developed. Maximizing the truncated expansion yields a sequence of approximations to the true maximum likelihood estimator (AMLE), the distance between which and the true MLE converges to zero in probability. The process can be outlined as follows. First, we transform the diffusion process defined by

$$dX(t) = \mu(X(t))dt + \sigma(X(t))dW(t). \tag{1}$$

into $Y(t)$, one with unit diffusion, by the so-called Lamperti transform:

$$Y(t) := \gamma(X(t)) := \int^{X(t)} \frac{du}{\sigma(u)};$$

$$dY(t) = \mu_Y(Y(t))dt + dW(t) \text{ where } \mu_Y(y) = \frac{\mu(\gamma^{-1}(y))}{\sigma(\gamma^{-1}(y))} - \frac{1}{2}\sigma'(\gamma^{-1}(y)).$$

And then "normalize" Y to Z so that the density is centered around a Gaussian kernel:

$$Z(t) := \frac{Y(t) - y_0}{\sqrt{\triangle}}.$$

The transition density of Z can be expanded using Hermite polynomials $\{H_j(z)\}$ as basis functions up to any order J :

$$p_Z^{(J)}(\triangle, z \mid y_0) = \frac{\exp(-z^2/2)}{\sqrt{2\pi}} \sum_{j=0}^{J} \eta_Z^j(\triangle, y_0) H_j(z).$$

The coefficients of the expansion can be expressed in closed-form using Taylor expansion with an iterative use of the infinitesimal generator $\mathcal{G}$:

$$\begin{aligned}
\eta_Z^j(\triangle, y_0) &= \frac{1}{j!}\mathbb{E}[H_j(\triangle^{-1/2}(Y(t+\triangle) - y_0)) \mid Y(t) = y_0] \\
&= \frac{1}{j!}\sum_{k=0}^{\infty} \mathcal{G}^k(f)|_{x=y_0} \cdot \frac{\triangle^k}{k!},
\end{aligned}$$

$$\text{where } \mathcal{G}(f) = \mu_Y \frac{\partial f}{\partial y} + \frac{1}{2}\frac{\partial^2 f}{\partial y^2} \text{ and } f(x) = H_j(\triangle^{-1/2}(x - y_0)).$$

The idea is somewhat akin to Edgeworth expansion, despite that the latter is for density functions of statistics admitting the Central Limit Theorem which is not the case for transition densities. Uniform convergence of the expansion to the true transition density as $J \to \infty$ is established. In a later work of him, i.e. [Aït-Sahalia(2008)], this method was extended to include multivariate diffusion process through the introduction of the notion "reducibility", a property similar to the feasibility of Lamperti transform. Correspondingly, the Hermite expansion moves from univariate to multivariate.

[Chang and Chen(2011)] makes further contributions to this econometric method by stating clearly the conditions under which the AMLE in [Aït-Sahalia(2002)] is consistent, and the associated convergence rate. In particular, two types of asymptotics are studied: let δ be the small-time interval between successive observations, and J be the maximum order of expansion of the transition density; then, one case is that δ is fixed but $J \to \infty$, and another case is that J is fixed but $\delta \to 0$ and $n\delta \to \infty$. Let $\hat{\theta}_n^{(J)}$ denote the AMLE obtained from truncating the density expansion to order

J, then the main result regarding the consistency of AMLE is that, under the conditions provided in [Chang and Chen(2011)] (see theorem 2 therein),

$$\hat{\theta}_n^{(J)} - \theta = \begin{cases} O_p(\delta^{J+1} + (n\delta)^{-\frac{1}{2}}), \text{ if } \delta \text{ is fixed and } J \to \infty \\ O_p(\delta^J + (n\delta)^{-\frac{1}{2}}), \text{ if } J \text{ is fixed, } \delta \to 0 \text{ and } n\delta^3 \to \infty. \end{cases}$$

This has shed light on the choice of n, δ and J to ensure convergence in practice. Moreover, the asymptotic distributions of $\hat{\theta}_n^{(J)}$ under these two asymptotic scenarios are given, making statistical inference possible.

A different method of expanding the transition density of multivariate diffusions is proposed in [Li(2013)], and again in [Li(2014)], which bypasses the "reducibility" and is based on Watanabe and Yoshida's theory of Malliavin calculus ([Watanabe(1987)] and [Yoshida(1992)]). Similar to [Aït-Sahalia(2008)], AMLE is derived and the convergence of the distance between it and the true MLE to zero (in probability) is proved. The results in Chenxu's paper serve as a benchmark for this thesis, so now I review the main theorem therein in details, which is initially established for continuous diffusion. This method stands out for its closed-form nature, rapid speed for implementation, and more importantly, adaptability to more complicated and general model specifications. A distinguished feature of this method is that it does not restrict itself to certain types of model. Later, I show its applications in the option pricing problem under some more general stochastic volatility with jumps models, including SVJ and SVJJ models.

3 Starting point: continuous diffusion model

3.1 Expansion of continuous diffusion process

A continuous time-homogenous multivariate diffusion process $X(t)$ is generally specified as a solution to a SDE:

$$dX(t) = \mu(X(t))dt + \sigma(X(t))dW(t), \tag{2}$$

where $X(t)$ is a d-dimensional stochastic process, given the initial state x_0. $W(t)$ is a d-dimensional standard Brownian motion; $\mu(\cdot)$ and $\sigma(\cdot)$ are smooth vector functions $\mathbb{R}^d \to \mathbb{R}^d$ and $\mathbb{R}^d \to \mathbb{R}^{d\times d}$ respectively. In time-inhomogeneous diffusions diffusions, the coefficients are allowed to depend on time directly, as in $\mu(X(t), t)$ and $\sigma(X(t), t)$, beyond their dependence on time via the state vector. Nevertheless, treating time as an additional state variable, one can transform the time-inhomogeneous case into a homogenous one. The existence of a solution (weak or strong) to (2) requires certain conditions, e.g. see [Oksendal(2010)] Chapter 5.2 and 5.3. We assume such conditions (e.g. the linear growth condition) hold.

We consider the small-time expansion of $f(X(t))$ with f being a deterministic smooth function. The term "small-time" indicates that our expansion is with respect to the time increment, denoted as τ, and is an asymptotic expansion with $\tau \to 0$. The method of implementing such expansion is proposed and proved in [Li(2014)], which is summarized here:

Theorem 1 *Let f be a $\mathbb{R}^d \to \mathbb{R}$ function. Given the dynamics (2), $f(X(\tau))$ admits the following small-time expansion*

$$f(X(\tau)) = \sum_{m=0}^{J} F_m \varepsilon^m + \mathcal{O}(\varepsilon^{J+1}), \text{ for any } J \in \mathbb{N}, \tag{3}$$

where $\varepsilon = \sqrt{\tau}$. *The stochastic coefficient* $\{F_m\}$ *is a sequence of random variable (RV hereafter), defined as a linear combination of some iterated Stratonovich integrals except* $F_0 := f(x_0)$:

$$F_m := \sum_{||\mathbf{i}||=m} D_{\mathbf{i}}(f)(x_0) J_{\mathbf{i}}(1), \ m \in \mathbb{N}^*, \tag{4}$$

in which $\mathbf{i}$ *is an index* $\mathbf{i} = (i_1, i_2, ..., i_n) \in \{0, 1, ..., d\}^n$ *and* $||\mathbf{i}||$ *is a "norm" of* $\mathbf{i}$ *counting an index* k *with* $i_k = 0$ *twice:*

$$||\mathbf{i}|| := \sum_{k=1}^{n} [2 \cdot 1_{\{i_k=0\}} + 1_{\{i_k \neq 0\}}]. \tag{5}$$

$J_{\mathbf{i}}(t)$ *is an iterated Stratonovich integral:*

$$J_{\mathbf{i}}(t) := \int_0^t \int_0^{t_1} ... \int_0^{t_{n-1}} 1 \circ dW_{i_n}(t_n) ... \circ dW_{i_1}(t_1), \tag{6}$$

where $\{W_1(t), W_2(t), \cdots, W_d(t)\}$ *is a d-dimensional standard Brownian motion, and* $\circ$ *denotes stochastic integral in Stratonovich sense. By convention, let* $W_0(t) := t$. $D_{\mathbf{i}}(\cdot)$ *is an operator, mapping an* $\mathbb{R}^d \to \mathbb{R}$ *function to another* $\mathbb{R}^d \to \mathbb{R}$ *function, defined as*

$$D_{\mathbf{i}}(f)(x) := \mathcal{A}_{i_n}(...(\mathcal{A}_{i_2}(\mathcal{A}_{i_1}(f)))...)(x), \tag{7}$$

where $\mathcal{A}_{i_k}$ *is the differential operator defined as*

$$\mathcal{A}_0 := \sum_{i=1}^{d} b_i(x) \frac{\partial}{\partial x_i},$$

$$\text{with } b_i(x) := \mu_i(x) - \frac{1}{2} \sum_{k=1}^{d} \sum_{j=1}^{d} \sigma_{kj}(x) \frac{\partial}{\partial x_k} \sigma_{ij}(x), \ \text{for } i = 1, 2, ..., d;$$

$$\mathcal{A}_j := \sum_{i=1}^{d} \sigma_{ij}(x) \frac{\partial}{\partial x_i}, \quad \text{for } j = 1, 2, ..., d.$$

Here is another useful result regarding the pathwise small-time asymptotic expansion of $X(t)$ itself proposed in [Li(2013)].

Theorem 2 $X(\tau)$ *admits the following expansion*

$$X(\tau) = \sum_{m=0}^{J} G_m \varepsilon^m + \mathcal{O}(\varepsilon^{J+1}), \ \text{for any } J \in \mathbb{N}.$$

The stochastic coefficient $\{G_m\}$ *is defined as*

$$\begin{aligned} G_0 &:= x_0; \\ G_m &:= \sum_{||\mathbf{i}||=m} C_{\mathbf{i}}(x_0) J_{\mathbf{i}}(1), \ m \in \mathbb{N}^*, \end{aligned}$$

where $C_{\mathbf{i}}(x)$ is a function defined as

$$C_{\mathbf{i}}(x) = \mathcal{A}_{i_n}(...(\mathcal{A}_{i_3}(\mathcal{A}_{i_2}(\sigma_{\cdot i_1}(x))))...), \tag{8}$$

in which $\sigma_{\cdot i_1}(x) = (\sigma_{1i_1}(x), \cdots, \sigma_{di_1}(x))^T$ denotes the i_1th column of the diffusion coefficient matrix σ, and if $i_1 = 0$, $\sigma_{\cdot i_1}(x) = (b_1(x), \cdots, b_d(x))^T$. Other definitions are the same as those in the previous theorem.

Note that theorem 2 is an expansion of a d-dimensional stochastic process, while theorem 1 is about a $\mathbb{R}^d \to \mathbb{R}$ smooth function acting on it. Now, I show the equivalence of these two theorems, which will be applied interchangeably in the following text.

Proposition 3 *Theorem 1 and theorem 2 are equivalent.*

Proof. Theorem 1$\Longrightarrow$Theorem 2:

We prove that for any index k, the k-th component of the expansion in theorem 2 is in line with the expansion of taking $f(x_1, x_2, ..., x_d) = x_k$ in theorem 1.

For the former, the coefficient of $X_k(\tau)$ is $C_{\mathbf{i},k}(x_0) = \mathcal{A}_{i_n}(...(\mathcal{A}_{i_3}(\mathcal{A}_{i_2}(\sigma_{ki_1})))...)(x_0)$. If $i_1 = 0$, the k-th component of $\sigma_{\cdot i_1}(x) = (b_1(x), \cdots, b_d(x))^T$ is $b_k(x)$, and if $i_1 = 1, ..., d$, the k-th component of $\sigma_{\cdot i_1}(x) = (\sigma_{1i_1}(x), \cdots, \sigma_{di_1})^T$ is $\sigma_{ki_1}(x)$.

For the latter, taking $f(x_1, x_2, ..., x_d) = x_k$ yields $\mathcal{A}_0(f)(x) = b_k(x)$, $\mathcal{A}_j(f)(x) = \sigma_{kj}(x)$.

Clearly, taking f as $f(x_1, x_2, ..., x_d) = x_k$, regardless of k and i_1, function $\mathcal{A}_{i_1}(f)(x)$ is the same with $\sigma_{ki_1}(x)$.

Theorem 2$\Longrightarrow$Theorem 1:

We add an additional component, $f(X(t))$, to the SDE. First, it is worth noting that the purpose of the change of variable from $\mu(\cdot)$ to $b(\cdot)$ by $b_i(x) = \mu_i(x) - \frac{1}{2}\sum\limits_{k=1}^{d}\sum\limits_{j=1}^{d}\sigma_{kj}(x)\frac{\partial}{\partial x_k}\sigma_{ij}(x)$ is to transform the initial SDE (2) in Itô sense into one in Stratonovich sense:

$$dX(t) = b(X(t))dt + \sigma(X(t)) \circ dW(t). \tag{9}$$

The operator $\mathcal{A}_{i_k}$, mapping a $\mathbb{R}^d \to \mathbb{R}^d$ function to another $\mathbb{R}^d \to \mathbb{R}^d$ function, can hence be seen as built upon (9). For arbitrary function f in theorem 1, by the "Itô formula" corresponding to Stratonovich integral, which is similar to Taylor formula in functional form, we have the SDE of $f(X(t))$:

$$df(X(t)) = \nabla f \cdot dX(t) = \nabla f \cdot b(X(t))dt + \nabla f \cdot \sigma(X(t)) \circ dW(t), \tag{10}$$

where ∇f is the gradient of f :

$$\nabla f := (\frac{\partial f}{\partial x_1}, \cdots, \frac{\partial f}{\partial x_d}).$$

And then, we apply theorem 2 to expand the additional component $f(X(t))$. The differential operator of $\frac{\partial}{\partial x_i}$ and the subsequent summation in both $\mathcal{A}_0$ and $\mathcal{A}_j$ in theorem 1 is in line with the inner product of ∇f with the other terms in (10). And by this pathwise expansion for the added component we obtain the expansion of $f(X(t))$, which agrees with theorem 1. $\blacksquare$

Another useful result in [Li(2013)] is about the conditional expectations of iterated Stratonovich integrals in form of

$$\mathbb{E}(\int_0^1 \int_0^{t_1} \cdots \int_0^{t_{n-1}} 1 \circ dW_{i_n}(t_n) \cdots \circ dW_{i_1}(t_1) \mid W_1(1) = z_1, \cdots, W_d(1) = z_d)$$

and

$$\mathbb{E}(\int_0^1 \int_0^{t_1} \cdots \int_0^{t_{n-1}} 1 \circ dW_{i_n}(t_n) \cdots \circ dW_{i_1}(t_1) \mid W_1(1) = z_1).$$

Using the algorithm there, one can directly calculate these conditional expectations as polynomials of $\mathbf{z}$. Note that the randomness of the expansion (3) lies in the iterated Stratonovich integrals $J_{\mathbf{i}}(1)$, while the coefficient $C_{\mathbf{i}}(x_0)$, or the "weight" of those stochastic integrals, is deterministic given that $\mu(\cdot)$ and $\sigma(\cdot)$, as well as initial states x_0, are prespecified. Therefore, the calculations of those conditional expectations are once and for all, forming a "library", which is taken for granted in the following passage, and can be directly utilized.

3.2 Two applications

The above expansion theorem can be applied in multiple financial scenarios by fitting different functions f into it, as long as the dynamics of the underlying asset and the volatility are specified as a continuous diffusion process.

3.2.1 Marginal transition density

Taking f as the famous Dirac-delta function yields the expansion of the transition density. Generally speaking, transition density has at least two forms:

$$p(\tau, y_\tau \mid x_0) = \frac{\partial^d}{\partial y_{\tau 1} \partial y_{\tau 2} ... \partial y_{\tau d}} \mathbb{P}\{X_1(\tau) \leq y_{\tau 1}, X_2(\tau) \leq y_{\tau 2}, ..., X_d(\tau) \leq y_{\tau d} \mid X(0) = x_0\}; \quad (11)$$

$$p(\tau, y_\tau \mid x_0) = \frac{\partial}{\partial y_\tau} \mathbb{P}\{X_1(\tau) \leq y_\tau \mid X(0) = x_0\}, \quad (12)$$

where $X_1(t)$ and $X_2(t)$ are notations of $Y(t)$ and $Z(t)$; τ is the small time increment.

The joint transition density (11) has been developed in [Li(2013)]. However, considering that in practice we concern more with the dynamics of the underlying asset rather than the latent volatility process, it is worthwhile to extend the result from joint one to its marginal counterpart (12). Another motivation is that, since only underlying asset price is discretely observable and the state of volatility is missing, econometric design is expected to tailor to this fact. Overall, since the marginal transition density of a d-dimensional process $X(t)$ concerns only one of its components, say, $X_1(t)$, rather than that of the whole vector $X(t) = [X_1(t), X_2(t), ..., X_d(t)]^T$, the fundamental distinction between the expansion of marginal transition density and its joint counterpart is that different libraries of conditional expectations in form of $\mathbb{E}(\prod_{w=1}^{l} J_{\mathbf{i}_w}(1) \mid W(1) = z)$ are needed. For the former, $W(1) = z$ refers to a one-dimensional Brownian motion $\{W_1(1) = z_1\}$, while for the latter, $\{W_1(1) = z_1, \cdots, W_d(1) = z_d\}$.

Now, I rigorously elaborate on how to derive the expansion of marginal transition density (12). For simplicity, I reduce my scope to a common SV model, i.e., a 2-dimensional SDE specified as

$$dX(t) = \begin{bmatrix} \mu_1(Y(t), Z(t)) \\ \mu_2(Z(t)) \end{bmatrix} dt + \begin{bmatrix} \sigma_{11}(Y(t), Z(t)) & 0 \\ \sigma_{21}(Z(t)) & \sigma_{22}(Z(t)) \end{bmatrix} \cdot dW(t), \quad (13)$$

$$\text{with } X(t) = \begin{bmatrix} Y(t) \\ Z(t) \end{bmatrix}, \; W(t) = \begin{bmatrix} W_1(t) \\ W_2(t) \end{bmatrix},$$

where $(W_1(t), W_2(t))^T$ is a 2-dimensional standard Brownian motion; $Y(t)$ is the price of underlying asset; $Z(t)$ is the latent volatility process. Note that $Z(t)$ itself is driven solely by a 1-dimensional SDE without the involvement of $Y(t)$:

$$dZ(t) = \mu_2(Z(t))dt + \sigma_{21}(Z(t))dW_1(t) + \sigma_{22}(Z(t))W_2(t).$$

Instead of stating the theorem first and subsequently prove it, I progress step by step.

Transformation to Stratonovich Integral According to theorem 2, define

$$b : \mathbb{R}^2 \to \mathbb{R}^2,\ b(x) := \begin{bmatrix} b_1(x) \\ b_2(x) \end{bmatrix},\ x \in \mathbb{R}^2, \tag{14}$$

where

$$\begin{aligned} b_1(x) &:= \mu_1(x) - \frac{1}{2}[\frac{\partial \sigma_{11}(x)}{\partial x_1}\sigma_{11}(x) + \frac{\partial \sigma_{11}(x)}{\partial x_2}\sigma_{21}(x)]; \\ b_2(x) &:= \mu_2(x) - \frac{1}{2}[\frac{\partial \sigma_{21}(x)}{\partial x_2}\sigma_{21}(x) + \frac{\partial \sigma_{22}(x)}{\partial x_2}\sigma_{22}(x)]. \end{aligned}$$

The initial SDE (13) is then transformed into one in Stratonovich sense:

$$dX(t) = b(X(t))dt + \sigma(X(t)) \circ dW(t),\ X(0) = x_0 = \begin{bmatrix} y_0 \\ z_0 \end{bmatrix}. \tag{15}$$

To bring forth finer local behavior of the diffusion process, we begin with replacing t with $\varepsilon^2 t$, rescaling (15) as $X^\varepsilon(t) := X(\varepsilon^2 t)$, and $W^\varepsilon(t) := \frac{1}{\varepsilon}W(\varepsilon^2 t)$ which is a standard Brownian motion. And then, we have

$$dX^\varepsilon(t) = \varepsilon^2 \mu(X^\varepsilon(t))dt + \varepsilon\sigma(X^\varepsilon(t)) \circ dW^\varepsilon(t),\ X^\varepsilon(0) = x_0. \tag{16}$$

Pathwise asymptotic expansion By relating τ and ε with $\varepsilon = \sqrt{\tau}$, the marginal transition density for the logarithmic price Y can be expressed as

$$p(\tau, y_\tau \mid x_0) = \mathbb{E}[\delta(X_1^\varepsilon(1) - y_\tau) \mid X(0) = x_0], \tag{17}$$

where $\delta(\cdot)$ is the famous Dirac-delta function, formally defined as a distribution. According to to Lemma 1 in [Li(2013)] (and [Watanabe(1987)] also), $X^\varepsilon(1)$ admits the following pathwise asymptotic expansion:

$$X^\varepsilon(1) = \sum_{k=0}^{J} F_k \varepsilon^k + \mathcal{O}(\varepsilon^{J+1}),\ \text{for any } J \in \mathbb{N} \tag{18}$$

$$\text{with } F_0 = x_0,\ F_k = \sum_{\|\mathbf{i}\|=k,\ \mathbf{i}\in\{0,1,2\}^m} C_{\mathbf{i}}(x_0)J_{\mathbf{i}}(1), \tag{19}$$

where $\|\mathbf{i}\|$, $C_{\mathbf{i}}(x_0)$, $J_{\mathbf{i}}(1)$ is defined by (5), (8), (6).

Picking out the first component of the vector on both sides of equation (18), we immediately have the element-wise expansion for $X_1^\varepsilon(1)$, which directly enters the expression of marginal transition density (17), and is therefore a focus in the following derivation. I state it as a corollary:

Corollary 4 $X_1^\varepsilon(1)$ *admits the following small-time expansion:*

$$X_1^\varepsilon(1) = \sum_{k=0}^{J} F_{k,1}\varepsilon^k + \mathcal{O}(\varepsilon^{J+1}), \tag{20}$$

$$\textit{where } F_{k,1} := \sum_{\|\mathbf{i}\|=k} C_{\mathbf{i},1}(x_0) J_{\mathbf{i}}(1), \ \textit{for } k \in \mathbb{N}^*, \tag{21}$$

with

$$C_{\mathbf{i},1}(x_0) := \mathcal{A}_{i_n}(...(\mathcal{A}_{i_3}(\mathcal{A}_{i_2}(\sigma_{1i_1}(x))))...)|_{x=x_0}$$

and $F_{0,1} := y_0$.

Standardization A standard normal RV inside the Dirac-delta function would be convenient to be integrated. Noticing that due to (20), as $\varepsilon \to 0$, $X_1^\varepsilon(1)$ converges to a normal (but not standard normal) RV, we begin by standardizing $X_1^\varepsilon(1)$.

Introduce a one-dimensional standard Brownian motion $\{B(t)\}$:

$$B(t) := D(x_0)\sum_{j=1}^{2}\sigma_{1j}(x_0)W_j(t), \text{ where } D(x_0) := \frac{1}{\sqrt{\sum\limits_{j=1}^{2}\sigma_{1j}^2(x_0)}} = \frac{1}{\sigma_{11}(x_0)}. \tag{22}$$

In fact, by Levy's Theorem, $B(t)$ can be easily verified as being a standard 1-dimensional Brownian motion.

And we make change of variable:

$$Z^\varepsilon := D(x_0)\frac{X_1^\varepsilon(1) - y_0}{\varepsilon}. \tag{23}$$

As $\varepsilon \to 0$, Z^ε converges to $B(1)$, a standard normal RV. Plugging (20) into (23), we immediately obtain the pathwise expansion for Z^ε :

$$Z^\varepsilon = \sum_{k=0}^{J} D(x_0)F_{k+1,1}\varepsilon^k + \mathcal{O}(\varepsilon^{J+1}). \tag{24}$$

To simplify notations, denote $Z_k := D(x_0)F_{k+1,1}$, which is a one-dimensinal RV. In particular, $Z_0 = D(x_0)\sum\limits_{i=1}^{2}\sigma_{1i}(x_0)\int_0^1 1 \circ dW_i(t) = W_1(1) \sim \mathcal{N}(0,1)$. (24) is equivalent to

$$Z^\varepsilon = \sum_{k=0}^{J} Z_k\varepsilon^k + \mathcal{O}(\varepsilon^{J+1}). \tag{25}$$

The final stage of the standardization process is a change of constant term:

$$y_\tau' = D(x_0)\frac{y_\tau - y_0}{\varepsilon}.$$

Based on the equation

$$X_1^\varepsilon(1) - y_\tau = \frac{\varepsilon}{D(x_0)}(Z^\varepsilon - y'_\tau),$$

we have

$$\begin{aligned}\mathbb{E}[\delta(X_1^\varepsilon(1) - y_\tau) \quad | \quad X(0) = x_0] &= \mathbb{E}[\delta(\frac{\varepsilon}{D(x_0)}(Z^\varepsilon - y'_\tau)) \mid X(0) = x_0] \\ &= \frac{D(x_0)}{\varepsilon}\mathbb{E}[\delta(Z^\varepsilon - y'_\tau) \mid X(0) = x_0].\end{aligned} \tag{26}$$

Therefore, we only need to develop the expansion for $\mathbb{E}[\delta(Z^\varepsilon - y'_\tau) \mid X(0) = x_0]$. For simplicity, the condition $X(0) = x_0$ is omitted in following text.

A heuristic expansion Heuristically speaking, because (24) can be seen as a function of ε, applying the classical chain rule to a generalized function $G(x) = \delta(x - y'_\tau)$, one can obtain a Taylor-type expansion for $\delta(Z^\varepsilon - y'_\tau)$ as follows:

$$\delta(Z^\varepsilon - y'_\tau) = \sum_{k=0}^{J} \psi_k(y'_\tau)\varepsilon^k + \mathcal{O}(\varepsilon^{J+1}). \tag{27}$$

Taking expectations on both sides of (27) yields

$$\mathbb{E}[\delta(Z^\varepsilon - y'_\tau)] = \sum_{k=0}^{J} \Omega_k^{(0)}(y'_\tau)\varepsilon^k + \mathcal{O}(\varepsilon^{J+1}), \tag{28}$$

in which

$$\Omega_k^{(0)}(y'_\tau) := \mathbb{E}[\psi_k(y'_\tau)]. \tag{29}$$

Hence, the problem reduces to calculating the coefficient term $\Omega_k^{(0)}(y'_\tau)$ out in closed form.

Calculation of the coefficients using univariate Taylor expansion In contrast to the expansion formula for joint transition density proposed in [Li(2013)] which is based on a multivariate Taylor expansion of the Dirac-delta function acting on a $\mathbb{R}^d$ vector, the inside RV Z^ε is one-dimensional in our context. Nevertheless, by analogy, applying univariate Taylor expansion instead of its multivariate counterpart, we are able to calculate $\Omega_k^{(0)}(y'_\Delta)$ explicitly.

Univariate Taylor expansion of a generalized function $\delta(x - y'_\tau)$ at $x = Z_0$ yields

$$\delta(x - y'_\tau) = \sum_{k=0}^{n} \frac{1}{k!}\frac{d^k}{dx^k}\delta(Z_0 - y'_\tau) \cdot (x - Z_0)^k + \mathcal{O}((x - Z_0)^{n+1}).$$

Replace x with (25),

$$\delta(Z^\varepsilon - y'_\tau) = \sum_{k=0}^{n} \frac{1}{k!}\frac{d^k}{dx^k}\delta(Z_0 - y'_\tau) \cdot (\sum_{j=1}^{J} Z_j\varepsilon^j)^k + \mathcal{O}(\varepsilon^{n+1}). \tag{30}$$

In order to collect the coefficient terms of ε^k in the summation (30), we introduce the array $\mathbf{j}(m) := (j_1, j_2, ..., j_m)$ with $j_i \in \mathbb{N}^*$ and $\sum_{i=1}^{m} j_i = k$, with m taking values from 1 to k. Suppose the entry of each $Z_j\varepsilon^j$ is $(j_1, j_2, ..., j_m)$, then the corresponding coefficient for these terms that contribute ε^k term is $\frac{1}{m!}\frac{d^m}{dx^m}\delta(Z_0 - y'_\tau)\cdot Z_{j_1}Z_{j_2}...Z_{j_m}$.

Define index set $\mathcal{J}_k$:

$$\mathcal{J}_k := \{(m, \mathbf{j}(m)) \mid m \in \{1, ..., k\},\ \mathbf{j}(m) = (j_1, ..., j_m) \text{ with } j_i \in \mathbb{N}^* \text{ and } \sum_{i=1}^{m} j_i = k\}. \tag{31}$$

By letting m go through all $1, 2, ..., k$ and summing all such possible combinations of $(m, \mathbf{j}(m)) \in \mathcal{J}_k$, we obtain

$$\psi_k(y'_\tau) = \sum_{(m,\mathbf{j}(m))\epsilon\mathcal{J}_k} \frac{1}{m!}\frac{d^m}{dx^m}\delta(Z_0 - y'_\tau)\cdot Z_{j_1}Z_{j_2}...Z_{j_m}, \tag{32}$$

where $\psi_k(y'_\tau)$, $\mathcal{J}_k$ is defined in (27) and (31). And because $\Omega_k^{(0)}(y'_\tau) = \mathbb{E}[\psi_k(y'_\tau)]$, the problem reduces to the calculation of $\sum\limits_{(m,\mathbf{j}(m))\epsilon\mathcal{J}_k}\frac{1}{m!}\mathbb{E}[\delta^{(m)}(Z_0 - y'_\tau)\cdot Z_{j_1}Z_{j_2}...Z_{j_m}]$, in which $\delta^{(m)}$ denotes the m-th order derivative of Dirac-Delta function δ.

We do integration by part:

$$\begin{aligned}
\mathbb{E}[\delta^{(m)}(Z_0 - y'_\tau)\cdot Z_{j_1}Z_{j_2}...Z_{j_m}] &= \int_{-\infty}^{+\infty}\mathbb{E}(Z_{j_1}Z_{j_2}...Z_{j_m} \mid Z_0 = z)\cdot\delta^{(m)}(z - y'_\tau)\phi(z)dz &(33)\\
&= \int_{-\infty}^{+\infty}[\mathbb{E}(Z_{j_1}Z_{j_2}...Z_{j_m} \mid Z_0 = z)\phi(z)]d(\delta^{(m-1)}(z - y'_\tau)) &(34)\\
&= (-1)\cdot\int_{-\infty}^{+\infty}\frac{d}{dz}[\mathbb{E}(Z_{j_1}Z_{j_2}...Z_{j_m} \mid Z_0 = z)\phi(z)]\cdot\delta^{(m-1)}(z - y'_\tau)dz\\
&= ...\\
&= (-1)^m\int_{-\infty}^{+\infty}\frac{d^m}{dz^m}[\mathbb{E}(Z_{j_1}Z_{j_2}...Z_{j_m} \mid Z_0 = z)\phi(z)]\cdot\delta(z - y'_\tau)dz\\
&= (-1)^m[\frac{d^m}{dz^m}(\mathbb{E}(Z_{j_1}Z_{j_2}...Z_{j_m} \mid Z_0 = z)\phi(z))]|_{z=y'_\tau},
\end{aligned}$$

where $\phi(z) = \frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}$ is the probability density function of $\mathcal{N}(0,1)$.

Plug $Z_k = D(x_0)F_{k+1,1}$ into (33),

$$\begin{aligned}
\mathbb{E}(Z_{j_1}Z_{j_2}...Z_{j_m} \quad | \quad Z_0 = z) &= \mathbb{E}(\prod_{t=1}^{m}D(x_0)F_{j_t+1,1} \mid Z_0 = z)\\
&= D(x_0)^m\mathbb{E}(\prod_{t=1}^{m}\sum_{\|\mathbf{i}\|=j_t+1}C_{\mathbf{i},1}(x_0)J_{\mathbf{i}}(1) \mid Z_0 = z).
\end{aligned}$$

Summing up the product of terms in form of $C_{\mathbf{i},1}(x_0)J_{\mathbf{i}}(1)$ over $(\mathbf{i}_1, \mathbf{i}_2, ..., \mathbf{i}_m)$ satisfying $||\mathbf{i}_w|| = j_w+1$, $w = 1, 2, ..., m$, we can rearrange the expression by

$$\prod_{t=1}^{m}\sum_{\|\mathbf{i}\|=j_t+1}C_{\mathbf{i},1}(x_0)J_{\mathbf{i}}(1) = \sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,...,\mathbf{i}_m),\\ ||\mathbf{i}_w||=j_w+1,\\ w=1,2,...,m}}\prod_{w=1}^{m}C_{\mathbf{i}_w,1}(x_0)J_{\mathbf{i}_w}(1).$$

Taking expectations on both sides yields

$$\mathbb{E}(\prod_{t=1}^{m}\sum_{\|\mathbf{i}\|=j_t+1} C_{\mathbf{i},1}(x_0)J_{\mathbf{i}}(1) \mid Z_0=z) = \sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_m),\\ ||\mathbf{i}_w||=j_w+1,\\ w=1,2,\ldots,m}} \prod_{w=1}^{m} C_{\mathbf{i}_w,1}(x_0)\mathbb{E}[\prod_{w=1}^{m} J_{\mathbf{i}_w}(1) \mid B(1)=z]. \tag{35}$$

Define differential operator $\mathcal{D}$:

$$\text{for any function } f:\ \mathbb{R}\to\mathbb{R},\ \mathcal{D}(f)(x) := \frac{df(x)}{dx} - xf(x). \tag{36}$$

Because the differential operator $\mathcal{D}$ satisfies that

$$\text{for any function } g,\ \frac{d[g(x)\phi(x)]}{dx} = g'(x)\phi(x) - xg(x)\phi(x) = \mathcal{D}(g)(x)\phi(x).$$

Differentiating (35) multiplied by $\phi(z)$ using $\mathcal{D}$ repeatedly results in

$$[\frac{d^m}{dz^m}(\mathbb{E}(Z_{j_1}Z_{j_2}...Z_{j_m} \quad | \quad Z_0=z)\phi(z))]|_{z=y'_\tau} = \tag{37}$$
$$\sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_m),\\ ||\mathbf{i}_w||=j_w+1,\\ w=1,2,\ldots,m}} \prod_{w=1}^{m} C_{\mathbf{i}_w,1}(x_0)\mathcal{D}^m(\mathbb{E}[\prod_{w=1}^{m} J_{\mathbf{i}_w}(1) \quad | \quad B(1)=y'_\tau])\phi(y'_\tau).$$

As building blocks, further define

$$P_{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_l)}(z) := \mathbb{E}(\prod_{w=1}^{l} J_{\mathbf{i}_w}(1) \mid W(1)=z), \tag{38}$$

which can be directly obtained from the library, though this library is different from that proposed in [Li(2013)] since $W(1)=z$ here is one-dimensional.

Combining (32), (33), (37), we have achieved the goal of calculating the coefficient in closed form, as stated in the following lemma.

Lemma 5 *For any $k\in\mathbb{N}$, the correction term $\Omega_k^{(0)}(y'_\tau)$ admits the following explicit expression:*

$$\Omega_k^{(0)}(y'_\tau) = \phi(y'_\tau)\sum_{(m,\mathbf{j}(m))\epsilon\mathcal{J}_k} Q_{(m,\mathbf{j}(m))}(x_0,y'_\tau)$$

where

$$Q_{(m,\mathbf{j}(m))}(x_0,y'_\tau) = \frac{(-1)^m}{m!}D(x_0)^m \sum_{||\mathbf{i}_w||=j_w+1,\ w=1,2,\ldots,m} (\prod_{w=1}^{m} C_{\mathbf{i}_w,1}(x_0))\mathcal{D}^m(P_{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_m)}(y'_\tau)).$$

The final expansion formula of marginal transition density We ultimately arrive at the closed-form small-time expansion of marginal transition density of SDE, summarized as the following theorem.

Theorem 6 *The marginal transition density $p(\tau, y_\tau|x_0)$ admits the following small-time expansion:*

$$p(\tau, y_\tau|x_0) = \frac{D(x_0)}{\varepsilon}\sum_{k=0}^{J}\Omega_k^{(0)}(y'_\tau)\varepsilon^k + \mathcal{O}(\varepsilon^{J+1})$$

with

$$y'_\tau = D(x_0)\frac{y_\tau - y_0}{\varepsilon}$$
$$\Omega_k^{(0)}(y'_\tau) = \phi(y'_\tau)\sum_{(m,\mathbf{j}(m))\epsilon\mathcal{J}_k} Q_{(m,\mathbf{j}(m))}(x_0, y'_\tau)$$

where

$$Q_{(m,\mathbf{j}(m))}(x_0, y'_\tau) = \frac{(-1)^m}{m!}D(x_0)^m \sum_{\shortparallel\mathbf{i}_w\shortparallel=j_w+1,\ w=1,2,\ldots,m} (\prod_{w=1}^{m} C_{\mathbf{i}_w,1}(x_0))\mathcal{D}^m(P_{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_m)}(y'_\tau))$$

with $D(x_0)$, $\mathcal{J}_k$, $C_{\mathbf{i}_w,1}(x_0)$, $P_{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_m)}$ defined in (22), (31), (21), (38) and $\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$.

3.2.2 Option price

Assuming that the SV model (13) is in risk-neutral measure, one can also generate the at-the-money expansion of option price, by writing it as $\mathbb{E}(\max\{Y(t) - y_0, 0\})$ and fitting f as the Heaviside function (or its derivatives of various orders, for option price sensitivities), based on an interchange of integration and differentiation. To avoid repetition, I leave this later because an extended version of it will be developed.

4 Volatility jump model (with jumps in volatility only)

Now, I move from SV to SVJ model.

4.1 Model specification

To begin with, we specify the following SVJ model with jumps in volatility only:

$$\begin{aligned}\frac{dS(t)}{S(t)} &= \mu_1(X(t))dt + \sigma_{11}(X(t))dW_1(t) \qquad (39)\\ dX(t) &= \mu_2(X(t))dt + \sigma_{21}(X(t))dW_1(t) + \sigma_{22}(X(t))dW_2(t) + J(t)dN(t)\end{aligned}$$

where $S(t)$ is the underlying asset price and $X(t)$ the latent volatility process; $(W_1(t), W_2(t))^T$ is a 2-dimensional standard Brownian motion; $N(t)$ is a Poisson process with constant intensity λ. $J(t)$ characterizes the jump size. Suppose that (39) is in risk-neutral measure.

Still, our goal is to develop the pathwise expansion of $S(\tau)$, and then the small-time expansion of option price under this model specification. Since the option price is the expectation of some functions of $S(\tau)$, which is determined by what kind of option it is (e.g. $f(x) = \max\{x - K, 0\}$ for call option and $f(x) = \max\{K - x, 0\}$ for put option):

$$P = \mathbb{E}[f(S(\tau))],$$

compositing the pathwise expansion of $S(\tau)$ with the outer function f, and taking expectations suffice. This will become our routine route in other scenarios as well.

4.2 Pathwise expansion of $S(\tau)$

The basic idea of handling volatility jumps is to treat diffusion and jump separately, or, in other words, to adapt our method for continuous diffusion with due adjustment to jumps.

In a pathwise expansion, the jump arrival times $\boldsymbol{\tau} = (\tau_1, \tau_2, \cdots, \tau_n)$ and the jump sizes $\mathbf{j} = (j_1, j_2, \cdots, j_n)$ are supposed to be known. In general, regarding the period between 2 jumps, say, (τ_i, τ_{i+1}), as a pure diffusion process, we apply our existing method to expand $S(t)$ as a series of $\sqrt{\tau_{i+1} - \tau_i}$, the coefficients of which are functions of $S(\tau_i)$ and $X(\tau_i)$, because $(S(\tau_i), X(\tau_i))$ is the initial state in this expansion. At an exact jump time, say, τ_i, only $X(\tau_i)$ is influenced, by being added with a constant equal to the jump size j_i. Hence, the coefficient is nothing but a function of $S(\tau_i)$ and $X(\tau_i)$. As long as it remains as a function of $S(\tau_i)$ and $X(\tau_i)$, the existing method can be applied for expansion, though we should be careful that the functional form then includes j_i as well as the inherited form. Iterative conditioning and repeated expansion of the coefficients, which are nothing but different functions of $S(\tau_i)$ and $X(\tau_i)$, suffice.

For example, starting from $(\tau_n, \tau]$ (i.e., conditional on $\mathcal{F}(\tau_n)$), during which (39) reduces to a continuous diffusion process:

$$\begin{aligned} \frac{dS(t)}{S(t)} &= \mu_1(X(t))dt + \sigma_{11}(X(t))dW_1(t) \\ dX(t) &= \mu_2(X(t))dt + \sigma_{21}(X(t))dW_1(t) + \sigma_{22}(X(t))dW_2(t), \end{aligned} \tag{40}$$

we apply corollary 4 to obtain a conditional expansion:

$$S(\tau) \mid \mathcal{F}(\tau_n) = \sum_{M_{n+1}=0}^{J_{n+1}} S^{(M_{n+1})}(S(\tau_n), X(\tau_n))(\sqrt{\tau - \tau_n})^{M_{n+1}}, \tag{41}$$

in which

$$S^{(M_{n+1})}(S(\tau_n), X(\tau_n)) := \sum_{||\mathbf{i}_{n+1}||=M_{n+1}} C_{\mathbf{i}_{n+1},1}(S(\tau_n), X(\tau_n)) J^{(n+1)}_{\mathbf{i}_{n+1}}(1).$$

The subscript $(n+1)$ all along (41) marks that the relevant term corresponds to the expansion in the interval $(\tau_n, \tau]$. Moreover, J_{n+1} denotes the maximum order of expansion in this step; M_{n+1} denotes the running index; $(S(\tau_n), X(\tau_n))$ is the conditional "initial" state; $S^{(M_{n+1})}(S(\tau_n), X(\tau_n))$ is the coefficient to M_{n+1}-th order coefficient; $\sqrt{\tau - \tau_n}$ is the square root of the small time increment; $C_{\mathbf{i}_{n+1},1}(S(\tau_n), X(\tau_n))$ follows the definition of (21); $J^{(n+1)}_{\mathbf{i}_{n+1}}(1)$ is a Stratonovich integral constructed by Brownian motion

$$\{\frac{W_1(t(\tau - \tau_n) + \tau_n) - W_1(\tau_n)}{\sqrt{\tau - \tau_n}}, \frac{W_2(t(\tau - \tau_n) + \tau_n) - W_2(\tau_n)}{\sqrt{\tau - \tau_n}}\}$$

and an index $\mathbf{i}_{n+1}$, as defined in (6).

Because τ_n is a jump time with jump size j_n, moving from τ_n to τ_n- involves a shift in the coefficient:

$$\begin{aligned} S(\tau) \quad | \quad \mathcal{F}(\tau_n-) &= \sum_{M_{n+1}=0}^{J_{n+1}} S^{(M_{n+1})}(S(\tau_n-), X(\tau_n-)+j_n)(\sqrt{\tau-\tau_n})^{M_{n+1}} \qquad (42) \\ &= \sum_{M_{n+1}=0}^{J_{n+1}} \sum_{||\mathbf{i}_{n+1}||=M_{n+1}} C_{\mathbf{i}_{n+1},1}(S(\tau_n-), X(\tau_n-)+j_n)(\sqrt{\tau-\tau_n})^{M_{n+1}} J^{(n+1)}_{\mathbf{i}_{n+1}}(1). \end{aligned}$$

Notice that $C_{\mathbf{i}_{n+1},1}(S(\tau_n-), X(\tau_n-)+j_n)$ is nothing but a function of $(S(\tau_n-), X(\tau_n-))$, and we might well denote this function as $C^{j_n}_{\mathbf{i}_{n+1},1}(\cdot)$:

$$C^{j_n}_{\mathbf{i}_{n+1},1}(S(\tau_n-), X(\tau_n-)) := C_{\mathbf{i}_{n+1},1}(S(\tau_n-), X(\tau_n-)+j_n).$$

Therefore, replacing the function $f(\cdot)$ in theorem 1 with $C^{j_n}_{\mathbf{i}_{n+1},1}(\cdot)$ gives the expansion of the random (from an overall rather than a conditional perspective, $S(\tau_n)$ and $X(\tau_n)$ are random) coefficient in (42).

Applying the above technique iteratively with the time period moving from $(\tau_n-, \tau]$ to $(\tau_{n-1}-, \tau_n)$, subsequently to $(\tau_{n-2}-, \tau_{n-1}), ..., (0, \tau_1)$, we finally arrive at the pathwise expansion of $S(\tau)$ under this single jump model stated as follow.

Theorem 7 *Given the arrival times and sizes vector* $\boldsymbol{\tau}$ *and* $\mathbf{j}$, $S(\tau)$ *driven by (39) admits the following closed form expansion:*

$$\begin{aligned} S(\tau)^{(J)} &= \sum_{|\mathbf{M}_{n+1}|\leq J} [\sum_{||\mathbf{i}_1||=M_1} \cdots \sum_{||\mathbf{i}_{n+1}||=M_{n+1}} D_{\mathbf{i}_1} C^{(j_1,j_2,\cdots,j_n)}_{(\mathbf{i}_2,\cdots,\mathbf{i}_{n+1}),1}(S_0, X_0) \qquad (43) \\ &\quad \cdot (\prod_{p=1}^{n+1} J^{(p)}_{\mathbf{i}_p}(1))] \boldsymbol{\tau}^{\mathbf{M}_{n+1}} + \mathcal{O}(\sum_{|\mathbf{M}_{n+1}|=J+1} \boldsymbol{\tau}^{\mathbf{M}_{n+1}}) \end{aligned}$$

where

$$\begin{aligned} \mathbf{M}_{n+1} \quad &: \quad = (M_1, ..., M_{n+1}),\ M_k \in \mathbb{N}^*;\ |\mathbf{M}_{n+1}| := \sum_{i=1}^{n+1} M_i; \\ \boldsymbol{\tau}^{\mathbf{M}_{n+1}} \quad &: \quad = \sqrt{\tau_1 - 0}^{M_1} \sqrt{\tau_2-\tau_1}^{M_2} ... \sqrt{\tau_n - \tau_{n-1}}^{M_n} \sqrt{\tau-\tau_n}^{M_{n+1}}. \end{aligned}$$

$C^{(j_1,j_2,\cdots,j_n)}_{(\mathbf{i}_2,\cdots,\mathbf{i}_{n+1}),1}(\cdot)$ *is recursively defined:*

$$\begin{aligned} C^{j_n}_{\mathbf{i}_{n+1},1}(S(\tau_n-), X(\tau_n-)) \quad &: \quad = C_{\mathbf{i}_{n+1},1}(S(\tau_n-), X(\tau_n-)+j_n) \\ C^{(j_{n-1},j_n)}_{\mathbf{i}_n,\mathbf{i}_{n+1},1}(S(\tau_{n-1}-), X(\tau_{n-1}-)) \quad &: \quad = D_{\mathbf{i}_n} C^{j_n}_{\mathbf{i}_{n+1},1}(S(\tau_{n-1}-), X(\tau_{n-1}-)+j_{n-1}) \\ &\quad\ \ \cdots \\ C^{(j_k,\ldots,j_n)}_{\mathbf{i}_{k+1},\ldots,\mathbf{i}_{n+1},1}(S(\tau_k-), X(\tau_k-)) \quad &: \quad = D_{\mathbf{i}_{k+1}} C^{(j_{k+1},\ldots,j_n)}_{\mathbf{i}_{k+2},\ldots,\mathbf{i}_{n+1},1}(S(\tau_k-), X(\tau_k-)+j_k) \\ &\quad\ \ \cdots \end{aligned}$$

with the last term being $D_{\mathbf{i}_1} C^{(j_1,j_2,\cdots,j_n)}_{(\mathbf{i}_2,\cdots,\mathbf{i}_{n+1}),1}(S_0, X_0)$. *Operator* $D_{\mathbf{i}}$ *is given by*

$$D_{\mathbf{i}}(f)(x_1, x_2) = \mathcal{A}_{i_l}(...(\mathcal{A}_{i_2}(\mathcal{A}_{i_1}(f)))...)(x_1, x_2),\ \textit{for}\ \mathbf{i} = (i_1, ..., i_l) \in \{0, 1, 2\}^l.$$

There are 4 remarks regarding this important theorem.

Remark 8 *To write (43) compactly, we may denote*

$$S_{\mathbf{j},\mathbf{M}_{n+1}} := [\sum_{||\mathbf{i}_1||=M_1} \cdots \sum_{||\mathbf{i}_{n+1}||=M_{n+1}} D_{\mathbf{i}_1} C^{(j_1,j_2,\cdots,j_n)}_{(\mathbf{i}_2,\cdots,\mathbf{i}_{n+1}),1}(S_0, X_0) \cdot (\prod_{p=1}^{n+1} J^{(p)}_{\mathbf{i}_p}(1))], \tag{44}$$

and as a result,

$$S(\tau)^{(J)} = \sum_{|\mathbf{M}_{n+1}|\leq J} S_{\mathbf{j},\mathbf{M}_{n+1}} \boldsymbol{\tau}^{\mathbf{M}_{n+1}} + \mathcal{O}(\sum_{|\mathbf{M}_{n+1}|=J+1} \boldsymbol{\tau}^{\mathbf{M}_{n+1}}). \tag{45}$$

The subscript $\mathbf{j}$ *of* $S_{\mathbf{j},\mathbf{M}_{n+1}}$ *indicates that this coefficient involves randomness originated from the jump sizes, which needs to be integrated when taking expectations.*

Remark 9 *What facilitates computation is that, as the* 2*-dimensional Brownian motion sequence*

$$\{B_1^{(i)}(t), B_2^{(i)}(t)\} = \{\frac{W_1(t(\tau_i - \tau_{i-1}) + \tau_{i-1}) - W_1(\tau_{i-1})}{\sqrt{\tau_i - \tau_{i-1}}}, \frac{W_2(t(\tau_i - \tau_{i-1}) + \tau_{i-1}) - W_2(\tau_{i-1})}{\sqrt{\tau_i - \tau_{i-1}}}\}$$

with i *ranging from* 1 *to* $n+1$ *is independent, the corresponding iterated Stratonovich integrals constructed by these Brownian motions are hence independent as well. This indicates that* $\{J^{(p)}_{\mathbf{i}_p}(1)\}$ *with* p *ranging from* 1 *to* $n+1$ *are independent, a very useful property when we take expectations in the future.*

Remark 10 *It is worthwhile to find out the term when* $|\mathbf{M}_{n+1}| = 0$ *or* 1 *in (45). Clearly, when* $|\mathbf{M}_{n+1}| = 0$, $\mathbf{M}_{n+1} = (0, \cdots, 0)$, $S_{\mathbf{j},\mathbf{M}_{n+1}} = S_0$*. When* $|\mathbf{M}_{n+1}| = 1$, $\mathbf{M}_{n+1}$ *has a component* 1 *while the others are* 0*. If the* p*-th component is* 1 *(with* p *ranging from* 1 *to* $n+1$*), then the corresponding* $S_{\mathbf{j},\mathbf{M}_{n+1}} = B^{(p)}_{11}\sigma_{11}(S_0, X_0 + \sum_{k=1}^{p-1} j_k)$*, where* $B^{(p)}_{11}$ *is a notation for* $B^{(p)}_1(1)$*. Therefore, if* $J = 1$*, then the expansion (45) reduces to*

$$S(\tau) = S_0 + \sum_{p=1}^{n+1} B^{(p)}_{11}\sigma_{11}(S_0, X_0 + \sum_{k=1}^{p-1} j_k)\sqrt{\tau_p - \tau_{p-1}} + \mathcal{O}(\sum_{|\mathbf{M}_{n+1}|=2} \boldsymbol{\tau}^{\mathbf{M}_{n+1}}). \tag{46}$$

Remark 11 *Clearly, when* $n = 0$*, the expansion reduces to its continuous counterpart, which is exactly what we have studied in section 3. In fact, (45) is an extension of the pathwise expansion of* $S(\tau)$ *to SVJ model with jumps in volatility only.*

4.3 Expansion of at-the-money option price

Following the risk-neutral dynamics (39), the fair price of an at-the-money (i.e., $K = S_0$) call option can be expressed as an expectation:

$$P(\tau, S_0, S_0, X_0) = \mathbb{E}(\max\{S(\tau) - S_0, 0\}). \tag{47}$$

We see it as a function of τ, and develop the expansion of $P(\tau, S_0, S_0, X_0)$ as $\tau \to 0$. Given that

$$
\begin{aligned}
P(\tau, S_0, S_0, X_0) &= \sum_{n=0}^{\infty} \mathbb{E}[(S(\tau) - S_0)^+ \mid N(\tau) = n] \cdot \mathbb{P}\{N(\tau) = n\} && (48) \\
&= \sum_{n=0}^{\infty} \mathbb{E}[(S(\tau) - S_0)^+ \mid N(\tau) = n] \cdot \frac{(\lambda\tau)^n}{n!} e^{-\lambda\tau}, && (49)
\end{aligned}
$$

we only need to derive the expansion of $\mathbb{E}[(S(\tau) - S_0)^+ \mid N(\tau) = n]$.

In addition, to include the j-th order price sensitivities with respect to the call option's "moneyness" $k = S_0 - K$, which is the earnings of excercising the option at present, we consider

$$
\begin{aligned}
\frac{\partial^j}{\partial k^j} P(\tau, k, S_0, X_0)|_{k=0} &= \sum_{n=0}^{\infty} \frac{\partial^j}{\partial k^j} \mathbb{E}[(S(\tau) - S_0 + k)^+ \mid N(\tau) = n]|_{k=0} \cdot \mathbb{P}\{N(\tau) = n\} && (50) \\
&= \sum_{n=0}^{\infty} \mathbb{E}[g^{(j)}(S(\tau) - S_0) \mid N(\tau) = n] \cdot \mathbb{P}\{N(\tau) = n\},
\end{aligned}
$$

where

$$
g^{(0)}(x) = \max\{x, 0\}, g^{(1)}(x) = \mathbf{1}_{\{x>0\}}, g^{(2)}(x) = \delta(x), g^{(j)}(x) = \delta^{(j-2)}(x) \tag{51}
$$

with δ the Dirac-delta function. It appears that the key is to expand $\mathbb{E}[g^{(j)}(S(\tau) - S_0) \mid N(\tau) = n]$. For $j = 0$, it reduces to the option price. Clearly, the Dirac-delta function facilitates the derivation a lot by allowing us to interchange the order of differentiation and integration without detailed rigorous proof.

4.3.1 Composite expansion

Our idea is to composite the expansion. On one hand,

$$
S(\tau)^{(J)} = S_0 + \sum_{l=1}^{J} \sum_{|\mathbf{M}_{n+1}|=l} S_{\mathbf{j}, \mathbf{M}_{n+1}} \boldsymbol{\tau}^{\mathbf{M}_{n+1}};
$$

on the other hand, we have the Taylor expansion of a generalized function:

$$
g^{(j)}(x_0 + a) = g^{(j)}(x_0) + \sum_{v=1}^{H} \frac{1}{v!} \delta^{(v+j-2)}(x_0) a^v + o(a^H), \ a \to 0. \tag{52}
$$

To ensure convergence, we make change of variable:

$$
Z := \frac{S(\tau) - S_0}{\varepsilon} = \sum_{l=1}^{J} \sum_{|\mathbf{M}_{n+1}|=l} \frac{1}{\varepsilon} S_{\mathbf{j}, \mathbf{M}_{n+1}} \boldsymbol{\tau}^{\mathbf{M}_{n+1}}.
$$

And then

$$
g^{(j)}(S(\tau) - S_0) = g^{(j)}(\varepsilon Z) = \varepsilon^{1-j} g^{(j)}(Z).
$$

To simplify notations, denote $z_0 := \sum_{|\mathbf{M}_{n+1}|=1} \frac{1}{\varepsilon} S_{\mathbf{j},\mathbf{M}_{n+1}} \boldsymbol{\tau}^{\mathbf{M}_{n+1}}$, which is also the center of the upcoming expansion, and denote the remaining term $z' := \sum_{|\mathbf{M}_{n+1}|\geq 2} \frac{1}{\varepsilon} S_{\mathbf{j},\mathbf{M}_{n+1}} \boldsymbol{\tau}^{\mathbf{M}_{n+1}}$, so that $Z = z_0 + z'$. Referring to (46), we also have the explicit expression for z_0 :

$$z_0 = \sum_{i=1}^{n+1} B_{11}^{(i)} \sigma_{11}(S_0, X_0 + \sum_{k=1}^{i-1} j_k) \frac{\sqrt{\tau_i - \tau_{i-1}}}{\varepsilon}.$$

In (52), replacing x_0 with z_0 (i.e. expanding at z_0), and a with z', yields

$$g^{(j)}(Z) = g^{(j)}(z_0) + \sum_{v=1}^{H} \frac{1}{v!} \delta^{(v+j-2)}(z_0) \cdot (\sum_{|\mathbf{M}_{n+1}|\geq 2} \frac{1}{\varepsilon} S_{\mathbf{j},\mathbf{M}_{n+1}} \boldsymbol{\tau}^{\mathbf{M}_{n+1}})^v,$$

where the remainder is omitted. Now, we organize this composite expansion, the key to which is to organize

$$(\sum_{|\mathbf{M}_{n+1}|\geq 2} \frac{1}{\varepsilon} S_{\mathbf{j},\mathbf{M}_{n+1}} \boldsymbol{\tau}^{\mathbf{M}_{n+1}})^v. \tag{53}$$

The rule that we follow when arranging (53) is the order of $\boldsymbol{\tau}$, i.e., collecting terms of $\boldsymbol{\tau}^{\mathbf{M}_{n+1}}$. As building blocks, define

$$\mathcal{R}(\mathbf{M}_{n+1}, v) = \{(\mathbf{M}_{n+1}^{(1)}, ..., \mathbf{M}_{n+1}^{(v)}) \mid \mathbf{M}_{n+1}^{(1)} + \cdots + \mathbf{M}_{n+1}^{(v)} = \mathbf{M}_{n+1},\ \mathbf{M}_{n+1}^{(i)} \in \mathbb{N}^{n+1},\ |\mathbf{M}_{n+1}^{(i)}| \geq 2\}. \tag{54}$$

The constraints $|\mathbf{M}_{n+1}^{(i)}| \geq 2$ and $\mathbf{M}_{n+1}^{(i)} \in \mathbb{N}^{n+1}$ are the requirements of (53). To contribute some $\boldsymbol{\tau}^{\mathbf{M}_{n+1}}$ terms, considering its definition, the entries $\mathbf{M}_{n+1}^{(1)}, ..., \mathbf{M}_{n+1}^{(v)}$ should sum up to $\mathbf{M}_{n+1}$ element-wise, which is conveyed by $\mathbf{M}_{n+1}^{(1)} + \cdots + \mathbf{M}_{n+1}^{(v)} = \mathbf{M}_{n+1}$. These discussions, altogether, result in the following organized expansion:

$$\sum_{v=1}^{H} \frac{1}{v!} \delta^{(v+j-2)}(z_0) \cdot (\sum_{|\mathbf{M}_{n+1}|\geq 2} \frac{1}{\varepsilon} S_{\mathbf{j},\mathbf{M}_{n+1}} \boldsymbol{\tau}^{\mathbf{M}_{n+1}})^v =$$

$$\sum_{l=2}^{J} \sum_{|\mathbf{M}_{n+1}|=l} \sum_{v=1}^{H} \frac{\delta^{(v+j-2)}(z_0)}{\varepsilon^v v!} \sum_{\mathcal{R}(\mathbf{M}_{n+1}, v)} S_{\mathbf{j},\mathbf{M}_{n+1}^{(1)}} \cdots S_{\mathbf{j},\mathbf{M}_{n+1}^{(v)}} \cdot \boldsymbol{\tau}^{\mathbf{M}_{n+1}}. \tag{55}$$

Summary 12 *The composite expansion is*

$$g^{(j)}(S(\tau) - S_0) = \varepsilon^{1-j}[g^{(j)}(z_0) + (55)]. \tag{56}$$

4.3.2 Integration

In this section, we take expectations on both sides of (56) to obtain the option price (or sensitivities), which is our ultimate goal:

$$\mathbb{E}[g^{(j)}(S(\tau) - S_0)] = \varepsilon^{1-j}\mathbb{E}[g^{(j)}(z_0)] + \varepsilon^{1-j}\mathbb{E}[(55)]. \tag{57}$$

The randomness originated from jump arrival times $\boldsymbol{\tau}$, sizes $\mathbf{j}$, and the Brownian motion related RVs, which are independent with each other. We integrate them in steps, i.e. first conditioning on $\boldsymbol{\tau}$ and $\mathbf{j}$, integrate the Brownian motion related RVs, and then integrate with respect to $\boldsymbol{\tau}$ and $\mathbf{j}$, the integral operators of which are denoted as $\mathcal{E}$ and $\mathcal{I}$, respectively.

Integrate Brownian motion related RVs Conditional on $\boldsymbol{\tau}$ and $\mathbf{j}$, consider the expectation of (55). Plug in the definition of $S_{\mathbf{j},\mathbf{M}_{n+1}^{(i)}}$, which is (44) (and for $|\mathbf{M}_{n+1}| = 0$ or 1, (46)), and utilize the linearity property of expectation:

$$\mathbb{E}_{n,\boldsymbol{\tau},\mathbf{j}}[S_{\mathbf{j},\mathbf{M}_{n+1}^{(1)}} \cdots S_{\mathbf{j},\mathbf{M}_{n+1}^{(v)}} \delta^{(v+j-2)}(z_0)] =$$

$$\sum_{||\mathbf{i}_1^{(1)}||=M_1^{(1)}} \cdots \sum_{||\mathbf{i}_{n+1}^{(1)}||=M_{n+1}^{(1)}} \sum_{||\mathbf{i}_1^{(2)}||=M_1^{(2)}} \cdots \sum_{||\mathbf{i}_{n+1}^{(2)}||=M_{n+1}^{(2)}} \cdots \sum_{||\mathbf{i}_1^{(v)}||=M_1^{(v)}} \cdots\cdots \sum_{||\mathbf{i}_{n+1}^{(v)}||=M_{n+1}^{(v)}} \tag{58}$$
$$\prod_{m=1}^{v} (D_{\mathbf{i}_1^m} C^{(j_1,j_2,\cdots,j_n)}_{(\mathbf{i}_2^m,\cdots,\mathbf{i}_{n+1}^m),1}(S_0,X_0)) \cdot \mathbb{E}_{n,\boldsymbol{\tau},\mathbf{j}}[\delta^{(v+j-2)}(\sum_{i=1}^{n+1} B_{11}^{(i)} \sigma_{11}(S_0, X_0 + \sum_{k=1}^{i-1} j_k) \frac{\sqrt{\tau_i - \tau_{i-1}}}{\varepsilon}) \prod_{r=1}^{v} (\prod_{p=1}^{n+1} J_{\mathbf{i}_p}^{(p)}(1))],$$

where the subscripts $n, \boldsymbol{\tau}, \mathbf{j}$ under $\mathbb{E}$ refer to conditioning on them. Next, we show how to express this expectation in closed form.

Here is a general discussion of the algorithm of dealing with the integration of Dirac-delta function, which typically entails integration by part.

Algorithm 13 *Integration of Dirac-delta function: let Z be a standard normal RV with probability density function $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$, and X be a RV. $\delta^{(j)}(\cdot)$ is the j-th order derivative of Dirac-delta function. Then*

$$\begin{aligned}
\mathbb{E}[\delta^{(j)}(Z)X] &= \mathbb{E}[\mathbb{E}(X \mid Z)\delta^{(j)}(Z)] \\
&= \int_{-\infty}^{+\infty} \mathbb{E}(X \mid Z = z)\delta^{(j)}(z)\phi(z)dz \\
&= \int_{-\infty}^{+\infty} \mathbb{E}(X \mid Z = z)\phi(z)d(\delta^{(j-1)}(z)) \\
&= \int_{-\infty}^{+\infty} -\delta^{(j-1)}(z) \cdot \frac{\partial}{\partial z}[\mathbb{E}(X \mid Z = z)\phi(z)]dz \\
&= \cdots \\
&= (-1)^j \frac{\partial^j}{\partial z^j}[\mathbb{E}(X \mid Z = z)\phi(z)]|_{z=0}.
\end{aligned}$$

Back to our context, to calculate

$$\mathbb{E}_{n,\boldsymbol{\tau},\mathbf{j}}[\delta^{(v+j-2)}(\sum_{i=1}^{n+1} B_{11}^{(i)} \sigma_{11}(S_0, X_0 + \sum_{k=1}^{i-1} j_k) \frac{\sqrt{\tau_i - \tau_{i-1}}}{\varepsilon}) \prod_{r=1}^{v} (\prod_{p=1}^{n+1} J_{\mathbf{i}_p}^{(p)}(1))], \tag{59}$$

we hope to use the algorithm above. Thus, to get a standard normal RV inside the Dirac-delta function, we make change of variable:

$$\begin{aligned}
Z &= \sum_{i=1}^{n+1} B_{11}^{(i)} \sigma_{11}(S_0, X_0 + \sum_{k=1}^{i-1} j_k) \frac{\sqrt{\tau_i - \tau_{i-1}}}{\sigma\varepsilon}, \qquad (60) \\
\text{where } \sigma &= \sqrt{\sum_{i=1}^{n+1} \sigma_{11}^2(S_0, X_0 + \sum_{k=1}^{i-1} j_k) \frac{\tau_i - \tau_{i-1}}{\tau}}.
\end{aligned}$$

Because $\{B_1^{(i)}(t)\}$ is an independent Brownian motion sequence, it is easy to verify that Z is a standard normal RV. Replace X by $\prod_{r=1}^{v}(\prod_{p=1}^{n+1} J_{\mathbf{i}_p}^{(p)}(1))$ and Z by (60) in the above algorithm, and (59) is equivalent to

$$-(-\sigma)^{1-v-j}(\frac{\partial^{v+j-2}}{\partial z^{v+j-2}}\{\mathbb{E}_{n,\boldsymbol{\tau},\mathbf{j}}[X \mid Z=z]\phi(z)\})|_{z=0}. \tag{61}$$

Hence, the problem reduces to calculate

$$\mathbb{E}[\prod_{r=1}^{v}(\prod_{p=1}^{n+1} J_{\mathbf{i}_p}^{(p)}(1)) \mid \sum_{i=1}^{n+1} B_{11}^{(i)}\sigma_{11}(S_0, X_0+\sum_{k=1}^{i-1} j_k)\frac{\sqrt{\tau_i-\tau_{i-1}}}{\sigma\varepsilon}=z]. \tag{62}$$

This can still be challenging, but not insoluble. Remember that $\{B_1^{(i)}(t), B_2^{(i)}(t)\}$ is an independent Brownian motion sequence with i ranging from 1 to $n+1$. Since the condition in (62) is a weighted sum of some standard normal variables $B_{11}^{(i)}$ and the RV to be taken expectation is a product of some iterated Stratonovich integrals constructed by $\{B_1^{(i)}(t), B_2^{(i)}(t)\}$, we wish to utilize again our experience in generating the conditional expectation library in form of

$$\mathbb{E}(\int_0^1\int_0^{t_1}\cdots\int_0^{t_{n-1}} 1\circ dW_{i_n}(t_n)\cdots\circ dW_{i_1}(t_1) \mid W_1(1)=z_1,\cdots,W_d(1)=z_d).$$

Denote the weights $\sigma_{11}(S_0, X_0+\sum_{k=1}^{i-1} j_k)\frac{\sqrt{\tau_i-\tau_{i-1}}}{\sigma\varepsilon}$ as λ_i, and denote our goal (62) as

$$G := \mathbb{E}[\prod_{r=1}^{v}(\prod_{p=1}^{n+1} J_{\mathbf{i}_p}^{(p)}(1)) \mid \sum_{i=1}^{n+1} B_{11}^{(i)}\lambda_i = z]. \tag{63}$$

By doing more conditioning,

$$\begin{aligned} G &= \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\mathbb{E}[\prod_{r=1}^{v}(\prod_{p=1}^{n+1} J_{\mathbf{i}_p}^{(p)}(1)) \mid \sum_{i=1}^{n+1} B_{11}^{(i)}\lambda_i = z, \qquad (64)\\ B_{11}^{(2)} &= z_2,\cdots,B_{11}^{(n+1)}=z_{n+1}, B_{21}^{(1)}=z_{n+2},\cdots,B_{21}^{(n+1)}=z_{2n+2}]\cdot\phi(z_2,\cdots,z_{2n+2}\mid z)dz_2\cdots dz_{2n+2}, \end{aligned}$$

where $\phi(z_2,\cdots,z_{2n+2}\mid z)$ is the conditional density of

$$(B_1^{(2)}(1),\cdots,B_1^{(n+1)}(1),B_2^{(n+1)}(1)) \mid \sum_{i=1}^{n+1} B_{11}^{(i)}\lambda_i = z. \tag{65}$$

This conditional law can be explicitly specified. As $\{B_1^{(1)}(1), B_2^{(1)}(1),\cdots,B_1^{(n+1)}(1),B_2^{(n+1)}(1)\}$ constructs a $(2n+2)$-dimensional standard normal RV, by the conditioning law of multivariate normal RV, one can formulate (65) as a $(2n+1)$-dimensional standard normal RV, with mean vector μ and covariance matrix $\sum\in\mathbb{R}^{(2n+1)\times(2n+1)}$:

$$\mu = (z\lambda_2,\cdots,z\lambda_{n+1},0,\cdots,0)^T;$$

$$\sum = \begin{bmatrix} 1-\lambda_2^2 & -\lambda_2\lambda_3 & \cdots & -\lambda_2\lambda_{n+1} & 0 & \cdots & 0 \\ -\lambda_3\lambda_2 & 1-\lambda_3^2 & \cdots & -\lambda_3\lambda_{n+1} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ -\lambda_{n+1}\lambda_2 & -\lambda_{n+1}\lambda_3 & \cdots & 1-\lambda_{n+1}^2 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

And thus $\phi(z_2, \cdots, z_{2n+2} \mid z)$ is explicit:

$$\phi(z_2, \cdots, z_{2n+2} \mid z) = \frac{1}{(\sqrt{2\pi})^{2n+1}\det(\sum)} \exp\{-\frac{1}{2}(z-\mu)^T {\sum}^{-1}(z-\mu)\};$$

as well as the moment-generating function:

$$M(\theta) := \mathbb{E}(e^{\theta^T z}) = \exp\{\theta\mu + \frac{1}{2}\theta^T \sum \theta\},\ \theta \in \mathbb{R}^{2n+1}. \tag{66}$$

Back to (64), the condition is equivalent with

$$B_{11}^{(1)} = z_1 = \frac{1}{\lambda_1}(z - \sum_{i=1}^{n+1}\lambda_i z_i), B_{11}^{(2)} = z_2, \cdots, B_{11}^{(n+1)} = z_{n+1}, B_{21}^{(1)} = z_{n+2}, \cdots, B_{21}^{(n+1)} = z_{2n+2}.$$

Relying on the algorithm in ([Li(2014)]), we can calculate $\mathbb{E}[\prod_{r=1}^{v}(\prod_{p=1}^{n+1} J_{\mathbf{i}_p}^{(p)}(1)) \mid B(1) = \mathbf{z}]$ explicitly as a polynomial of $\mathbf{z}$. As $z_1 = \frac{1}{\lambda_1}(z - \sum_{i=1}^{n+1}\lambda_i z_i)$ is a polynomial of $\{z, z_2, \cdots, z_{2n+2}\}$, the initial conditional expectation is also a polynomial of $\{z, z_2, \cdots, z_{2n+2}\}$. Assume this polynomial is

$$\sum_{s_1,\cdots,s_{2n+2}\in\mathbb{N}} c(s_1, \cdots, s_{2n+2}) z^{s_1} z_2^{s_2} \cdots z_{2n+2}^{s_{2n+2}}. \tag{67}$$

Plug (67) into (64),

$$\begin{aligned} G &= \sum_{s_1,\cdots,s_{2n+2}\in\mathbb{N}} c(s_1, \cdots, s_{2n+2}) \cdot \int_{(z_1,\cdots,z_{2n+2})\in\mathbb{R}^{2n+1}} z^{s_1} z_2^{s_2} \cdots z_{2n+2}^{s_{2n+2}} \phi(z_2, \cdots, z_{2n+2} \mid z) dz_2 \cdots dz_{2n+2} \qquad (68) \\ &= \sum_{s_1,\cdots,s_{2n+2}\in\mathbb{N}} c(s_1, \cdots, s_{2n+2}) z^{s_1} \mathbb{E}[(B_{11}^{(2)})^{s_2} \cdots (B_{21}^{(n+1)})^{s_{2n+2}} \mid \sum_{i=1}^{n+1} B_{11}^{(i)}\lambda_i = z]. \end{aligned}$$

And the above cross moment can be obtained by differentiating the moment generating function (66):

$$\mathbb{E}[(B_{11}^{(2)})^{s_2} \cdots (B_{21}^{(n+1)})^{s_{2n+2}} \mid \sum_{i=1}^{n+1} B_{11}^{(i)}\lambda_i = z] = \frac{\partial^{s_2+\cdots+s_{2n+2}}}{\partial\theta_1^{s_2}\cdots\partial\theta_{2n+1}^{s_{2n+2}}} M(\theta)|_{\theta=0}.$$

These insights render (62) solvable, and it is a polynomial of z and λ_i. After the differentiation process in (61), the result turns out to be a polynomial of λ_i, or, equivalently, $\sigma_{11}(S_0, X_0 + \sum_{k=1}^{i-1} j_k)\frac{\sqrt{\tau_i - \tau_{i-1}}}{\sigma\varepsilon}$. An example term could be

$$\prod_{i=1}^{n+1}[\sigma_{11}(S_0, X_0 + \sum_{k=1}^{i-1} j_k)\frac{\sqrt{\tau_i - \tau_{i-1}}}{\sigma\varepsilon}]^{M_i}, \ M_i \in \mathbb{N}. \tag{69}$$

Integrate jump arrival times $\boldsymbol{\tau}$ Back to (57) and (58), after the integration of Brownian motion related RVs, we now integrate the jump arrival times $\boldsymbol{\tau}$. Notice that the randomness brought about by $\boldsymbol{\tau}$ appears only in the $\mathbb{E}_{n,\boldsymbol{\tau},\mathbf{j}}$ term in (58) and the $\boldsymbol{\tau}^{\mathbf{M}_{n+1}}$ term in (55), not in $D_{\mathbf{i}_1^m} C^{(j_1, j_2, \cdots, j_n)}_{(\mathbf{i}_2^m, \cdots, \mathbf{i}_{n+1}^m), 1}(S_0, X_0)$. Therefore, along with the polynomial result of (69), we only need to consider integration in the following form with respect to $\boldsymbol{\tau}$:

$$\mathbb{E}_{n,\mathbf{j}}[(\sum_{i=1}^{n+1}\sigma_{11}^2(S_0, X_0 + \sum_{k=1}^{i-1} j_k)\frac{\tau_i - \tau_{i-1}}{\tau})^{-\frac{r}{2}} \cdot \prod_{p=1}^{n+1}(\frac{\sqrt{\tau_p - \tau_{p-1}}}{\varepsilon})^{M_p}], \tag{70}$$

where I have plugged in the definition of the "constant" σ separated during the standardization process (60), and $\tau_0 := 0$; $\tau_{n+1} := \tau$; $M_i, r \in \mathbb{N}$. The subscripts $n, \mathbf{j}$ under $\mathbb{E}$ refer to conditioning on them.

A good observation is that, $\tau_1, \tau_2, \cdots, \tau_n$ are the ordered statistics of n independent RVs on $[0, \tau]$. Let

$$t_i = \frac{\tau_i - \tau_{i-1}}{\tau}, \ i = 1, \ldots, n+1.$$

Then $(t_1, t_2, \ldots, t_n)$ is uniformly distributed in the area of $\sum_{i=1}^{n} t_i \le 1$ with the density being $n!$[1]. Denote $a_i = \sigma_{11}^2(S_0, X_0 + \sum_{k=1}^{i-1} j_k)$ for $i = 1, \ldots, n+1$. Then, we simplify (70) as:

$$\begin{aligned}
(\sum_{i=1}^{n+1} a_i \frac{\tau_i - \tau_{i-1}}{\tau})^{-\frac{r}{2}} \cdot \prod_{p=1}^{n+1}(\frac{\sqrt{\tau_p - \tau_{p-1}}}{\varepsilon})^{M_p} &= (\prod_{p=1}^{n+1} t_p^{\frac{M_p}{2}}) \cdot (\sum_{i=1}^{n+1} a_i t_i)^{-\frac{r}{2}} \\
&= (\prod_{p=1}^{n+1} t_p^{\frac{M_p}{2}}) \cdot a_{n+1}^{-\frac{r}{2}}(\sum_{i=1}^{n+1}\frac{a_i t_i}{a_{n+1}})^{-\frac{r}{2}} \\
&= (\prod_{p=1}^{n+1} t_p^{\frac{M_p}{2}}) \cdot a_{n+1}^{-\frac{r}{2}}(1 - \sum_{i=1}^{n+1}\frac{a_{n+1} - a_i}{a_{n+1}} t_i)^{-\frac{r}{2}} \\
&= (\prod_{p=1}^{n+1} t_p^{\frac{M_p}{2}}) \cdot a_{n+1}^{-\frac{r}{2}}(1 - \sum_{i=1}^{n}\frac{a_{n+1} - a_i}{a_{n+1}} t_i)^{-\frac{r}{2}},
\end{aligned}$$

in which I use that $\sum_{i=1}^{n+1} t_i = 1$ and $\varepsilon = \sqrt{\tau}$.

[1] Because the volume of this area is $\frac{1}{n!}$

Taking expectations on both sides yields

$$a_{n+1}^{-\frac{r}{2}} \int\limits_{t_1+\cdots+t_n\leq 1} n! t_1^{\frac{M_1}{2}} t_2^{\frac{M_2}{2}} \cdots t_{n+1}^{\frac{M_{n+1}}{2}} (1-\sum_{i=1}^{n} u_i t_i)^{-\frac{r}{2}} dt_1 dt_2 \cdots dt_n,$$

where $u_i = \frac{a_{n+1}-a_i}{a_{n+1}}$. Existing literature such as [Lijoi and Regazzini(2004)] provides the result of the integral above as

$$n! a_{n+1}^{-\frac{r}{2}} \frac{\prod\limits_{i=1}^{n+1} \Gamma(\frac{M_i}{2}+1)}{\Gamma(\sum\limits_{i=1}^{n+1} \frac{M_i}{2}+n+1)} F_D(\frac{r}{2}, \frac{M_1}{2}+1, \ldots, \frac{M_n}{2}+1; \sum_{i=1}^{n+1} \frac{M_i}{2}+n+1, u_1, \ldots, u_n)$$

with F_D the Lauricella function of the fourth kind and $\Gamma(\cdot)$ the Gamma function: $\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt$.

To simplify notations, denote the integral operator with respect to τ as $\mathcal{E}$. Then, the final result after the integration with respect to jump times is

$$\mathcal{E}(\boldsymbol{\tau}^{\mathbf{M}_{n+1}} \cdot (61)). \tag{71}$$

Integrate jump sizes j Lastly, it is the jump sizes that remain to be integrated. Examining (58), we see that $\mathbf{j}$ is included in $D_{\mathbf{i}_1} C^{(j_1,j_2,\cdots,j_n)}_{(\mathbf{i}_2,\cdots,\mathbf{i}_{n+1}),1}(S_0, X_0)$ as well as the integration result of $\mathcal{E}(\cdot)$, or more specifically, $\sigma_{11}(S_0, X_0 + \sum\limits_{k=1}^{i-1} j_k)$, but not the Lauricella function F_D. To obtain explicit formula, we need additional assumption on the distribution of $\mathbf{j}$, e.g. $j_i \sim Exp(\lambda)$. Denote the integral operator with respect to $\mathbf{j}$ as $\mathcal{I}$, and the final result after all integrations is

$$\mathcal{I}\{\prod_{m=1}^{v} (D_{\mathbf{i}_1^m} C^{(j_1,j_2,\cdots,j_n)}_{(\mathbf{i}_2^m,\cdots,\mathbf{i}_{n+1}^m),1}(S_0, X_0)) \cdot (71)\}.$$

4.4 Summary of computation procedure

Below are the key expressions for expanding the option price or sensitivities.

$$\frac{\partial^j}{\partial k^j} P(\tau, k, S_0, X_0)|_{k=0} = \sum_{n=0}^{\infty} \mathbb{E}_n[g^{(j)}(S(\tau)-S_0)] \frac{(\lambda\tau)^n}{n!} e^{-\lambda\tau};\ k = S_0 - K.$$

$$\begin{aligned} \mathbb{E}_n[g^{(j)}(S(\tau)-S_0)] &= \varepsilon^{1-j}\mathbb{E}_n[g^{(j)}(z_0)] + \varepsilon^{1-j}\mathbb{E}_n[(55)], \\ \text{where } \varepsilon &= \sqrt{\tau},\ z_0 = \sum_{i=1}^{n+1} B_{11}^{(i)} \sigma_{11}(S_0, X_0 + \sum_{k=1}^{i-1} j_k) \frac{\sqrt{\tau_i - \tau_{i-1}}}{\varepsilon}. \end{aligned}$$

And,

$$\mathbb{E}_n[(55)] = \sum_{l=2}^{J} \sum_{|\mathbf{M}_{n+1}|=l} \sum_{v=1}^{H} \sum_{\mathcal{R}(\mathbf{M}_{n+1},v)} \frac{1}{\varepsilon^v v!} \mathbb{E}_n[(58) \cdot \boldsymbol{\tau}^{\mathbf{M}_{n+1}}], \tag{72}$$

where $\boldsymbol{\tau}^{\mathbf{M}_{n+1}} = \tau_1^{\frac{M_1}{2}}(\tau_2-\tau_1)^{\frac{M_2}{2}}\cdots(\tau_n-\tau_{n-1})^{\frac{M_n}{2}}(\tau-\tau_n)^{\frac{M_{n+1}}{2}}$. One part of (58) is the coefficients $D_{\mathbf{i}_1^m}C^{(j_1,j_2,\cdots,j_n)}_{(\mathbf{i}_2^m,\cdots,\mathbf{i}_{n+1}^m),1}(S_0,X_0)$, which are directly obtained by the definition in Theorem 7. Another part is (59), which is

$$-(-\sigma)^{1-v-j}\frac{\partial^{v+j-2}}{\partial z^{v+j-2}}\{G\cdot\phi(z)\}|_{z=0},\ \text{where } \sigma=\sqrt{\sum_{i=1}^{n+1}\sigma_{11}^2(S_0,X_0+\sum_{k=1}^{i-1}j_k)\frac{\tau_i-\tau_{i-1}}{\tau}} \tag{73}$$

with G defined in (63). As given by (68),

$$G=\sum_{s_1,\cdots,s_{2n+2}\in\mathbb{N}}c(s_1,\cdots,s_{2n+2})z^{s_1}\cdot\frac{\partial^{s_2+\cdots+s_{2n+2}}}{\partial\theta_1^{s_2}\cdots\partial\theta_{2n+1}^{s_{2n+2}}}M(\theta)|_{\theta=0},$$

with $M(\theta)$ defined in (66) and the coefficients $c(s_1,\cdots,s_{2n+2})$ calculated in [Li(2014)], as described in (67). The expression (73) up to now is a polynomial of λ_i, with an example term being (69). Following the 2 integration procedures in the last 2 parts of section 4.3.2,

$$(72)=\mathcal{I}\{\prod_{m=1}^{v}D_{\mathbf{i}_1^m}C^{(j_1,j_2,\cdots,j_n)}_{(\mathbf{i}_2^m,\cdots,\mathbf{i}_{n+1}^m),1}(S_0,X_0)\cdot\mathcal{E}[\boldsymbol{\tau}^{\mathbf{M}_{n+1}}\cdot(73)]\},$$

which is the final result.

5 Stochastic volatility concurrent jump model

5.1 Model specification

Now we add jumps in asset return, in addition to the volatility factor, but restrict the jump time to be the same. This is generally specified as stochastic volatility concurrent jump model:

$$\begin{aligned}\frac{dS(t)}{S(t-)} &= \mu_1(X(t))dt+v(X(t))dW_1(t)+dJ_1(t) \\ dX(t) &= \mu_2(X(t))dt+\sigma_{21}(X(t))dW_1(t)+\sigma_{22}(X(t))dW_2(t)+dJ_2(t)\end{aligned} \tag{74}$$

with $J_1(t), J_2(t)$ being a pure jump process and a compound Poisson process respectively:

$$J_1(t)=\sum_{i=1}^{N(t)}(e^{Z_i^s}-1),\ J_2(t)=\sum_{i=1}^{N(t)}Z_i^v,$$

where Z_i^s and Z_i^v are jump sizes for S and X respectively. Furthermore, we assume the Poisson process $N(t)$ in this model to be with time-varying intensity $\lambda=\lambda(X(t))$ rather than a constant.

Note that $J_1(t)$ and $J_2(t)$ are driven by the same Poisson process $N(t)$, indicating the exactly same jump arrival times, even though the jump sizes Z_i^s and Z_i^v may follow different distributions, which are assumed to be known. This is what "concurrent jump" means. Assume (74) is in risk-neutral measure $\mathbb{Q}$.

Still, our goal is to expand the at-the-money option price (and its various sensitivities). Following the definition in the previous section, e.g. (51),

$$P(\tau,S_0,S_0,X_0)=\mathbb{E}^{\mathbb{Q}}(\max\{S(\tau)-S_0,0\});$$

$$
\begin{aligned}
\frac{\partial^j}{\partial k^j} P(\tau, k, S_0, X_0)|_{k=0} &= \frac{\partial^j}{\partial k^j} \mathbb{E}^{\mathbb{Q}}[g^{(0)}(S(\tau) - S_0 + k)]|_{k=0} \\
&= \mathbb{E}^{\mathbb{Q}}[g^{(j)}(S(\tau) - S_0)],
\end{aligned}
$$

where $k = S_0 - K$.

5.2 A change of measure

To deal with the time-varying intensity $\lambda = \lambda(X(t))$ of the Poisson process, we begin by a change of measure from the original $\mathbb{Q}$ to $\mathbb{P}$:

$$
\frac{d\mathbb{P}}{d\mathbb{Q}} = \Lambda(t)^{-1} \text{ with } \Lambda(t) := \prod_{k=1}^{N(t)} \lambda(X(\tau_k -)) \cdot \exp\{t - \int_0^t \lambda(X(s)) ds\},
$$

where $\Lambda(t)$ is the Radon-Nikodym derivative. $N(t)$ is a Poisson process with constant intensity $\lambda = 1$ under the new measure $\mathbb{P}$.

Following the Bayesian Rule,

$$
\begin{aligned}
\mathbb{E}^{\mathbb{Q}}[g^{(j)}(S(\tau) - S_0)] &= \mathbb{E}^{\mathbb{P}}[\Lambda(\tau) g^{(j)}(S(\tau) - S_0)] \\
&= \sum_{n=0}^{\infty} \mathbb{E}^{\mathbb{P}}[\Lambda(\tau) g^{(j)}(S(\tau) - S_0) \mid N(\tau) = n] \cdot \mathbb{P}(N(\tau) = n) \\
&= \sum_{n=0}^{\infty} \mathbb{E}^{\mathbb{P}}[\Lambda(\tau) g^{(j)}(S(\tau) - S_0) \mid N(\tau) = n] \cdot \frac{\tau^n e^{-\tau}}{n!}.
\end{aligned}
$$

The problem reduces to expand $\mathbb{E}^{\mathbb{P}}[\Lambda(\tau) g^{(j)}(S(\tau) - S_0) \mid N(\tau) = n]$.

5.3 Separate expansion of $S(\tau)$ and $\Lambda(\tau)$

Compared with (48), there are 2 more complicating factors. First, a Radon-Nikodym derivative $\Lambda(\tau)$ is added, suggesting that a corresponding expansion of it has to be developed. Second, the dynamic of $S(\tau)$ has changed due to the additional jumps in asset return, suggesting that the original expansion of $S(\tau)$ has to be adapted. Nevertheless, our basic idea remains the same: to derive the expansion separately for some rather simple constituents, and then combine into a final complicated one.

5.3.1 Decomposition of continuous part and jump part

We first make a decomposition of continuous part and jump part in both $S(t)$ and $\Lambda(t)$, a usual technique for handling jump:

$$
S(t) = S_0 S^C(t) S^J(t),
$$

where $S^C(t)$ and $S^J(t)$ are the continuous and jump part of S respectively:

$$
\begin{aligned}
S^C(t) &= \exp\{\int_0^t [\mu_1(X(s)) - \frac{1}{2} v^2(X(s))] ds + \int_0^t v(X(s)) dW_1(s)\} \\
S^J(t) &= \exp(\sum_{i=1}^{N(t)} Z_i^s),
\end{aligned}
$$

and

$$\Lambda(t) = \Lambda^C(t)\Lambda^J(t)$$

where $\Lambda^C(t)$ and $\Lambda^J(t)$ are the continuous and jump part of Λ respectively:

$$\begin{aligned}\Lambda^C(t) &= \exp\{t - \int_0^t \lambda(X(s))ds\} \\ \Lambda^J(t) &= \prod_{k=1}^{N(t)} \lambda(X(\tau_k -)).\end{aligned}$$

Correspondingly, our target is reformulated as

$$\begin{aligned}\mathbb{E}_n^{\mathbb{P}}[\Lambda(\tau)g^{(j)}(S(\tau) - S_0)] &= \mathbb{E}_n^{\mathbb{P}}[\Lambda^C(\tau)\Lambda^J(\tau)g^{(j)}(S_0 S^C(\tau)S^J(\tau) - S_0)] \\ &= S_0^{1-j}\mathbb{E}_n^{\mathbb{P}}[\Lambda^C(\tau)\Lambda^J(\tau)g^{(j)}(S^C(\tau)S^J(\tau) - 1)].\end{aligned} \tag{75}$$

After this decomposition, both the continuous part and the jump part can be regarded as a variant of SVJ model with jumps in volatility only, which has been studied in section 4. The decomposition transforms the concurrent jump to single jump, a significant simplification.

5.3.2 Continuous part

By Itô formula, the continuous part $S^C(t)$ follows

$$\begin{aligned}\frac{dS^C(t)}{S^C(t)} &= \mu_1(X(t))dt + v(X(t))dW_1(t) \\ dX(t) &= \mu_2(X(t))dt + \sigma_{21}(X(t))dW_1(t) + \sigma_{22}(X(t))dW_2(t) + J_2(t)dN(t),\end{aligned} \tag{76}$$

as well as the continuous part $\Lambda^C(t)$:

$$\begin{aligned}\frac{d\Lambda^C(t)}{\Lambda^C(t)} &= [1 - \lambda(X(t))]dt \\ dX(t) &= \mu_2(X(t))dt + \sigma_{21}(X(t))dW_1(t) + \sigma_{22}(X(t))dW_2(t) + J_2(t)dN(t)\end{aligned} \tag{77}$$

with $N(t)$ a Poisson process with constant intensity 1 under $\mathbb{P}$.

Obviously, $S^C(t)$ and $\Lambda^C(t)$ both follow the form of (39). Straightforward application of theorem 7 gives the expansion:

$$S^C(\tau) = \sum_{|\mathbf{M}_{n+1}|\leq J} S^C_{\mathbf{j},\mathbf{M}_{n+1}} \boldsymbol{\tau}^{\mathbf{M}_{n+1}} + \mathcal{O}(\sum_{|\mathbf{M}_{n+1}|=J+1} \boldsymbol{\tau}^{\mathbf{M}_{n+1}}); \tag{78}$$

$$\Lambda^C(\tau) = \sum_{|\mathbf{M}^{(1)}_{n+1}|\leq J} \Lambda^C_{\mathbf{j},\mathbf{M}^{(1)}_{n+1}} \boldsymbol{\tau}^{\mathbf{M}^{(1)}_{n+1}} + \mathcal{O}(\sum_{|\mathbf{M}^{(1)}_{n+1}|=J+1} \boldsymbol{\tau}^{\mathbf{M}^{(1)}_{n+1}}), \tag{79}$$

where the definitions of $S^C_{\mathbf{j},\mathbf{M}_{n+1}}$ and $\Lambda^C_{\mathbf{j},\mathbf{M}^{(1)}_{n+1}}$ are specified by theorem 7 in accordance with the specific functional form that governs the dynamics of $S^C(t)$ and $\Lambda^C(t)$, given by (76) and (77).

5.3.3 Jump part

Now, we consider the expansion of jump part $S^J(t) = \exp(\sum_{i=1}^{N(t)} Z_i^s)$ and $\Lambda^J(t) = \prod_{k=1}^{N(t)} \lambda(X(\tau_k-))$. $S^J(t)$ has already been in its simplest form, because conditional on $N(\tau) = n$, $S^J(\tau) = \exp(\sum_{i=1}^{n} Z_i^s)$, while $\Lambda^J(t)$ has yet to be dealt with, because each τ_k is random. The technique is to expand $\lambda(X(\tau_k-))$ for each fixed k and then mutiply them. Indeed, the idea follows directly from Section 4.2 where we treat diffusion and jump separately for the first time. $X(t)$ is governed by a one-dimensional SDE solely:

$$dX(t) = \mu_2(X(t))dt + \sigma_{21}(X(t))dW_1(t) + \sigma_{22}(X(t))dW_2(t) + J_2(t)dN(t). \tag{80}$$

Now, the reasoning here is a restatement of the theories developed in section 4. We analogously see the period between jump arrival times as a univariate continuous diffusion process and apply the corresponding method developed in section 3.1. At an exact jump arrival time, we add the jump to the coefficient term, hence constructing a new RV, which can further be expanded in a similar technique used in Section 4.2. Do this iteratively. This simple analogy yields

$$\lambda^{(J)}(X(\tau_k-)) = \sum_{|\mathbf{M}_k| \le J} \boldsymbol{\tau}^{\mathbf{M}_k} \lambda_{\mathbf{M}_k}, \text{ for each fixed } k \in \{1, \cdots, N(\tau)\}, \tag{81}$$

where $\mathbf{M}_k \in \{(M_{k1}, \cdots, M_{kk}, 0) \mid M_{ki} \in \mathbb{N}\}$. The last component is fixed at 0 because now our expansion ends at a jump time τ_k- rather than the previous τ when no jump occurs in probability 1. The coefficient $\lambda_{\mathbf{M}_k}$ is similarly defined as

$$\lambda_{\mathbf{M}_k} := [\sum_{||\mathbf{i}_1|| = M_{k1}} \cdots \sum_{||\mathbf{i}_k|| = M_{kk}} \mathcal{L}_{\mathbf{i}_1} \mathcal{L}^{(j_1, j_2, \cdots, j_{k-1})}_{(\mathbf{i}_2, \cdots, \mathbf{i}_k), 1}(\lambda)(X_0) \cdot (\prod_{p=1}^{k} J^{(p)}_{\mathbf{i}_p}(1))]$$

with the differential operator $\mathcal{L}$, mapping a $\mathbb{R} \to \mathbb{R}$ function to another $\mathbb{R} \to \mathbb{R}$ function, defined as

$$\mathcal{L}_{\mathbf{i}}(\lambda)(x) := \mathcal{D}_{i_l}(...(\mathcal{D}_{i_2}(\mathcal{D}_{i_1}(\lambda)))...)(x) \text{ for } \mathbf{i} = (i_1, \cdots, i_l) \in \{0, 1, 2\}^l,$$

where

$$\mathcal{D}_0 := b(x)\frac{\partial}{\partial x}, \ \mathcal{D}_j := \sigma_{2j}(x)\frac{\partial}{\partial x},$$

with $b(x)$ defined for the transformation of Itô integral to Stratonovich integral:

$$b(x) = \mu_2(x) - \frac{1}{2}(\sigma_{21}(x)\frac{\partial \sigma_{21}(x)}{\partial x} + \sigma_{22}(x)\frac{\partial \sigma_{22}(x)}{\partial x}).$$

(80) is equivalent with

$$dX(t) = b(X(t))dt + \sigma_{21}(X(t)) \circ dW_1(t) + \sigma_{22}(X(t)) \circ dW_2(t) + J_2(t)dN(t).$$

Clearly, when $N(\tau) = 0$, i.e. no jumps occur until τ, the problem reduces to the simplest version (continuous diffusion) developed in Section 3.1.

5.3.4 Combination and composite expansion

Plugging all expansion formulae of the constituent (78), (79) and (81) into (75) yields

$$\Lambda^C(\tau)\Lambda^J(\tau)g^{(j)}(S^C(\tau)S^J(\tau)-1) = [\sum_{|\mathbf{M}^{(1)}_{n+1}|\leq J}\Lambda^C_{\mathbf{M}^{(1)}_{n+1}}\boldsymbol{\tau}^{\mathbf{M}^{(1)}_{n+1}}]\cdot[\prod_{i=1}^{n}(\sum_{|\mathbf{M}_i|\leq J}\lambda_{\mathbf{M}_i}\boldsymbol{\tau}^{\mathbf{M}_i})]\cdot g^{(j)}(S^C(\tau)S^J(\tau)-1) \tag{82}$$

and

$$\begin{aligned} S^C(\tau)S^J(\tau)-1 &= \sum_{|\mathbf{M}_{n+1}|\leq J} S^C_{\mathbf{M}_{n+1}}S^J(\tau)\boldsymbol{\tau}^{\mathbf{M}_{n+1}} - 1 \\ &= S^J(\tau) - 1 + \sum_{1\leq|\mathbf{M}_{n+1}|\leq J} S^C_{\mathbf{M}_{n+1}}S^J(\tau)\boldsymbol{\tau}^{\mathbf{M}_{n+1}}, \end{aligned}$$

in which I use $S^C(0) = 1$.

Now, let us develop the expansion of $g^{(j)}(S^C(\tau)S^J(\tau)-1)$. The process here is similar to that in the "composite expansion" part in the previous section, or, more specifically, (55). Review the Taylor expansion of $g^{(j)}(x)$:

$$g^{(j)}(x_0+a) = g^{(j)}(x_0) + \sum_{v=1}^{H}\frac{1}{v!}\delta^{(v+j-2)}(x_0)a^v + o(a^H), a\to 0.$$

To simplify notations, we denote $z_0 := S^J(\tau)-1$, which is also the center of the upcoming composite expansion, and the remaining term $z' := \sum\limits_{1\leq|\mathbf{M}_{n+1}|\leq J} S^C_{\mathbf{M}_{n+1}}S^J(\tau)\boldsymbol{\tau}^{\mathbf{M}_{n+1}}$, so that $S^C(\tau)S^J(\tau)-1 = z_0 + z'$. And obviously, $z_0 = \exp(\sum\limits_{i=1}^{n} Z^s_i) - 1$.

Expanding $g^{(j)}(S^C(\tau)S^J(\tau)-1)$ at z_0 yields

$$g^{(j)}(S^C(\tau)S^J(\tau)-1) = g^{(j)}(z_0) + \sum_{v=1}^{H}\frac{1}{v!}\delta^{(v+j-2)}(z_0)(\sum_{1\leq|\mathbf{M}_{n+1}|\leq J} S^C_{\mathbf{M}_{n+1}}S^J(\tau)\boldsymbol{\tau}^{\mathbf{M}_{n+1}})^v.$$

We organize this composite expansion, the key to which is to organize

$$(\sum_{1\leq|\mathbf{M}_{n+1}|\leq J} S^C_{\mathbf{M}_{n+1}}S^J(\tau)\boldsymbol{\tau}^{\mathbf{M}_{n+1}})^v. \tag{83}$$

Similar to (55), what we follow is the order of $\boldsymbol{\tau}$ when collecting terms. To contribute some $\boldsymbol{\tau}^{\mathbf{M}_{n+1}}$ terms, considering its definition, the entries $\mathbf{M}^{(1)}_{n+1}, ..., \mathbf{M}^{(v)}_{n+1}$ should sum up to $\mathbf{M}_{n+1}$ element-wise. Analogous to (54), define

$$\mathcal{R}(\mathbf{M}_{n+1}, v) := \{(\mathbf{M}^{(1)}_{n+1}, ..., \mathbf{M}^{(v)}_{n+1}) \mid \mathbf{M}^{(1)}_{n+1} + \cdots + \mathbf{M}^{(v)}_{n+1} = \mathbf{M}_{n+1},\ \mathbf{M}^{(i)}_{n+1} \in \mathbb{N}^{n+1},\ |\mathbf{M}^{(i)}_{n+1}| \geq 1\}.$$

The only difference appears in the constraint $|\mathbf{M}^{(i)}_{n+1}| \geq 1$. This is because each $\mathbf{M}_{n+1} \in \mathbb{N}^{n+1}$ and $|\mathbf{M}^{(i)}_{n+1}| \geq 1$ in (83).

These discussions, altogether, result in the following organized expansion:

$$\sum_{v=1}^{H}\frac{1}{v!}\delta^{(v+j-2)}(z_0)(\sum_{1\le|\mathbf{M}_{n+1}|\le J}S^C_{\mathbf{M}_{n+1}}S^J(\tau)\boldsymbol{\tau}^{\mathbf{M}_{n+1}})^v =$$

$$\sum_{1\le|\mathbf{M}_{n+1}|\le J}\sum_{v=1}^{H}\frac{S^J(\tau)^v\cdot\delta^{(v+j-2)}(z_0)}{v!}\sum_{\mathcal{R}(\mathbf{M}_{n+1},v)}S^C_{\mathbf{M}^{(1)}_{n+1}}\cdots S^C_{\mathbf{M}^{(v)}_{n+1}}\cdot\boldsymbol{\tau}^{\mathbf{M}_{n+1}}. \tag{84}$$

In contrast to (55), The two additional terms $\Lambda^C(\tau)$ and $\Lambda^J(\tau)$ in (82) render that (84) is not an end. We have to consider

$$[\sum_{|\mathbf{M}^{(1)}_{n+1}|\le J}\Lambda^C_{\mathbf{M}^{(1)}_{n+1}}\boldsymbol{\tau}^{\mathbf{M}^{(1)}_{n+1}}]\cdot[\prod_{i=1}^{n}(\sum_{|\mathbf{M}_i|\le J}\lambda_{\mathbf{M}_i}\boldsymbol{\tau}^{\mathbf{M}_i})]\cdot(84). \tag{85}$$

As usual, we collect terms with respect to $\boldsymbol{\tau}^{\mathbf{M}_{n+1}}$. Define a new index set $\mathcal{S}$:

$$\begin{aligned}\mathcal{S}(\mathbf{M}_{n+1},n+2) &:= \{(M^{(1)}_{n+1},M^{(2)}_{n+1},M_1,\dots,M_n)\mid M^{(1)}_{n+1}+M^{(2)}_{n+1}+M_1+\cdots+M_n=\mathbf{M}_{n+1};\\ M^{(1)}_{n+1},M^{(2)}_{n+1},M_i &\in \mathbb{N}^{n+1};\ \text{the last } (n+1-i) \text{ elements of } M_i \text{ is } 0\}.\end{aligned}$$

The requirement of M_i is due to that in (85) $\mathbf{M}_i\in\mathbb{N}^i$, and to extend it to a vector of length $n+1$ in order to be in line with $\mathbf{M}^{(1)}_{n+1}$ and $\mathbf{M}_{n+1}$ in (84), we manually set the last $(n+1-i)$ elements of it to be 0. $M^{(2)}_{n+1}$ is a substitute for the notation $\mathbf{M}_{n+1}$ in $\mathcal{R}(\mathbf{M}_{n+1},v)$.

Ultimately, we expand $\Lambda^C(\tau)\Lambda^J(\tau)g^{(j)}(S^C(\tau)S^J(\tau)-1)$ as

$$\sum_{\mathcal{S}(\mathbf{M}_{n+1},n+2)}\sum_{v=0}^{|M^{(2)}_{n+1}|}\sum_{\mathcal{R}(\mathbf{M}^{(2)}_{n+1},v)}\frac{S^J(\tau)^v\cdot\delta^{(v+j-2)}(z_0)}{v!}\Lambda^C_{\mathbf{M}^{(1)}_{n+1}}S^C_{\mathbf{M}^{(2)(1)}_{n+1}}\cdots S^C_{\mathbf{M}^{(2)(v)}_{n+1}}\cdot(\prod_{i=1}^{n}\lambda_{\mathbf{M}_i})\boldsymbol{\tau}^{\mathbf{M}_{n+1}} \tag{86}$$

where $(M^{(1)}_{n+1},M^{(2)}_{n+1},M_1,\dots,M_n)$ is an example term from $\mathcal{S}(\mathbf{M}_{n+1},n+2)$, and $(\mathbf{M}^{(2)(1)}_{n+1},...,\mathbf{M}^{(2)(v)}_{n+1})$ is an example term from $\mathcal{R}(\mathbf{M}^{(2)}_{n+1},v)$ correspondingly.

5.4 Integration

The at-the-money option price (or its sensitivities) needs integration:

$$S_0^{1-j}\mathbb{E}^{\mathbb{P}}_n[\Lambda^C(\tau)\Lambda^J(\tau)g^{(j)}(S^C(\tau)S^J(\tau)-1)]=S_0^{1-j}\mathbb{E}^{\mathbb{P}}_n[(86)].$$

Now, let us have an examination of the randomness in (86), and integrate them. Again, we take expectations in steps.

5.4.1 Integrate jump size Z^s_i and Z^v_i

Z^s_i appears in $z_0=\exp(\sum_{i=1}^{n}Z^s_i)-1$ and $S^J(\tau)=\exp(\sum_{i=1}^{n}Z^s_i)$ only, thus

$$\mathbb{E}^{\mathbb{P}}_n[S^J(\tau)^v\cdot\delta^{(v+j-2)}(z_0)]=\int_{\mathcal{J}^n}\delta^{(j-2)}(e^{u_1+\cdots+u_n}-1)e^{v(u_1+\cdots+u_n)}f(u_1)\cdots f(u_n)du_1\cdots du_n, \tag{87}$$

in which we assume that Z_i^s is i.i.d. with f being the probability density function of the jump size. However, only a few distributions admit an explicit solution. For example, if we further assume this distribution to be $\mathcal{N}(\mu_S, \sigma_S^2)$, then this integration is explicitly solvable, the method of which could be found in the appendix in [Wang(2020)].

Z_i^v appears in $\Lambda^C_{\mathbf{M}_{n+1}}$, $S^C_{\mathbf{M}_{n+1}}$ and $\lambda_{\mathbf{M}_i}$. we may assume it follows an exponential distribution $Exp(\lambda)$, and use operators $\mathcal{I}$ in section 4.3.2 for integration.

5.4.2 Integrate jump arrival times τ

The randomness brought by $\boldsymbol{\tau}$ is only through $\boldsymbol{\tau}^{\mathbf{M}_{n+1}}$, thus its integration is actually easier than that in volatility jump model, such as (70), because the center of expansion $z_0 = \exp(\sum_{i=1}^{n} Z_i^s) - 1$ does not include $\boldsymbol{\tau}$ here, in contrast to the center $z_0 = \sum_{i=1}^{n+1} B_{11}^{(i)} \sigma_{11}(S_0, X_0 + \sum_{k=1}^{i-1} j_k) \frac{\sqrt{\tau_i - \tau_{i-1}}}{\varepsilon}$ there. It is for this reason that we only need to consider expectations of the following type:

$$\mathbb{E}_n^{\mathbb{P}}(\boldsymbol{\tau}^{\mathbf{M}_{n+1}}) = \mathbb{E}_n^{\mathbb{P}}[(\tau - \tau_n)^{\frac{M_{n+1}}{2}} (\tau_n - \tau_{n-1})^{\frac{M_n}{2}} \cdots \tau_1^{\frac{M_1}{2}}], \tag{88}$$

where the subscripts n refers to conditional on $N(\tau) = n$. No Lauricella function is needed now, while the same argument that $\tau_1, \tau_2, \cdots, \tau_n$ are the ordered statistics of n independent RVs on $[0, \tau]$ applies equally. Let

$$t_i = \frac{\tau_i - \tau_{i-1}}{\tau}, \ i = 1, \ldots, n+1.$$

Then $(t_1, t_2, \ldots, t_n)$ is uniformly distributed in the area of $\sum_{i=1}^{n} t_i \le 1$ with the density being $n!$. Therefore,

$$\begin{aligned}(88) &= \frac{n!}{\tau^n} \int_0^{\tau} \cdots \int_0^{\tau_3} \int_0^{\tau_2} (\tau - \tau_n)^{\frac{M_{n+1}}{2}} (\tau_n - \tau_{n-1})^{\frac{M_n}{2}} \cdots \tau_1^{\frac{M_1}{2}} d\tau_1 \cdots d\tau_n \\ &= n! \frac{\prod_{i=1}^{n+1} \Gamma(\frac{M_i}{2} + 1)}{\Gamma(\sum_{i=1}^{n+1} \frac{M_i}{2} + n + 1)} \tau^{\frac{M_1 + \cdots + M_{n+1}}{2}}.\end{aligned}$$

Notice that no operator like $\mathcal{E}$ is needed.

5.4.3 Integrate Brownian motion related RVs

We follow the same algorithm developed in section 4.3.2. The only difference turns out to be that, because of the nature of (86), there are more iterated Stratonovich integrals $J_{\mathbf{i}}(1)$ (originated from $\Lambda^C_{\mathbf{M}^{(1)}_{n+1}}$, $S^C_{\mathbf{M}^{(2)(i)}_{n+1}}$ and $\lambda_{\mathbf{M}_i}$) multiplied with each other, but this does not interfere with the algorithm of calculating these conditional expectations proposed in [Li(2014)]. Denote the operator to integrate Brownian motion related RVs as $\mathcal{W}$. And then,

$$\mathbb{E}^{\mathbb{P}}_{n,\mathbf{j}}[\Lambda^C_{\mathbf{M}^{(1)}_{n+1}} S^C_{\mathbf{M}^{(2)(1)}_{n+1}} \cdots S^C_{\mathbf{M}^{(2)(v)}_{n+1}} \cdot (\prod_{i=1}^{n} \lambda_{\mathbf{M}_i})] = \mathcal{W}[\Lambda^C_{\mathbf{M}^{(1)}_{n+1}} S^C_{\mathbf{M}^{(2)(1)}_{n+1}} \cdots S^C_{\mathbf{M}^{(2)(v)}_{n+1}} \cdot (\prod_{i=1}^{n} \lambda_{\mathbf{M}_i})].$$

5.5 Summary of computation procedure

Below are the key expressions for expanding the option price or sensitivities.

$$\frac{\partial^j}{\partial k^j}P(\tau,k,S_0,X_0)|_{k=0}=\sum_{n=0}^{\infty}\mathbb{E}_n^{\mathbb{P}}[\Lambda(\tau)g^{(j)}(S(\tau)-S_0)]\cdot\frac{\tau^n e^{-\tau}}{n!};\ k=S_0-K.$$

$$\mathbb{E}_n^{\mathbb{P}}[\Lambda(\tau)g^{(j)}(S(\tau)-S_0)]=S_0^{1-j}\mathbb{E}_n^{\mathbb{P}}[\Lambda^C(\tau)\Lambda^J(\tau)g^{(j)}(S^C(\tau)S^J(\tau)-1)]=S_0^{1-j}\mathbb{E}_n^{\mathbb{P}}[(86)].$$

(86) is

$$\sum_{\mathcal{S}(\mathbf{M}_{n+1},n+2)}\sum_{v=0}^{|M_{n+1}^{(2)}|}\sum_{\mathcal{R}(\mathbf{M}_{n+1}^{(2)},v)}\frac{S^J(\tau)^v\cdot\delta^{(v+j-2)}(z_0)}{v!}\Lambda^C_{\mathbf{M}_{n+1}^{(1)}}S^C_{\mathbf{M}_{n+1}^{(2)(1)}}\cdots S^C_{\mathbf{M}_{n+1}^{(2)(v)}}\cdot(\prod_{i=1}^{n}\lambda_{\mathbf{M}_i})\boldsymbol{\tau}^{\mathbf{M}_{n+1}}$$

where $S^J(\tau)=\exp(\sum_{i=1}^{n}Z_i^s)$, $z_0=\exp(\sum_{i=1}^{n}Z_i^s)-1$, $\Lambda^C_{\mathbf{M}_{n+1}}$, $S^C_{\mathbf{M}_{n+1}}$ and $\lambda_{\mathbf{M}_i}$ are stochastic coefficients corresponding to the respective expansion of continuous diffusions Λ^C, S^C and the 1-dimensional jump diffusion $\lambda^J(X(t))$, as defined in (78), (79) and (81).

The randomness in $S^J(\tau)^v\cdot\delta^{(v+j-2)}(z_0)$ is completely eliminated by (87), although a limitation is that only a few distributions, e.g. $\mathcal{N}(\mu_S,\sigma_S^2)$, are allowed for Z_i^s for an explicit solution of this integration:

$$\mathbb{E}_n^{\mathbb{P}}[S^J(\tau)^v\cdot\delta^{(v+j-2)}(z_0)]=\int_{\mathcal{J}^n}\delta^{(j-2)}(e^{u_1+\cdots+u_n}-1)e^{v(u_1+\cdots+u_n)}f(u_1)\cdots f(u_n)du_1\cdots du_n.$$

The randomness in $\boldsymbol{\tau}^{\mathbf{M}_{n+1}}$ is completely eliminated by (88):

$$\mathbb{E}_n^{\mathbb{P}}(\boldsymbol{\tau}^{\mathbf{M}_{n+1}})=n!\frac{\prod_{i=1}^{n+1}\Gamma(\frac{M_i}{2}+1)}{\Gamma(\sum_{i=1}^{n+1}\frac{M_i}{2}+n+1)}\tau^{\frac{M_1+\cdots+M_{n+1}}{2}}.$$

The remaining part of the expectation is solved by operators $\mathcal{I}$ and $\mathcal{W}$:

$$\mathbb{E}_n^{\mathbb{P}}[\Lambda^C_{\mathbf{M}_{n+1}^{(1)}}S^C_{\mathbf{M}_{n+1}^{(2)(1)}}\cdots S^C_{\mathbf{M}_{n+1}^{(2)(v)}}\cdot(\prod_{i=1}^{n}\lambda_{\mathbf{M}_i})]=\mathcal{I}\{\mathcal{W}[\Lambda^C_{\mathbf{M}_{n+1}^{(1)}}S^C_{\mathbf{M}_{n+1}^{(2)(1)}}\cdots S^C_{\mathbf{M}_{n+1}^{(2)(v)}}\cdot(\prod_{i=1}^{n}\lambda_{\mathbf{M}_i})]\}.$$

6 Idiosyncratic volatility jump model (with jumps in both asset return and volatility)

6.1 Model specification

Now, we go one step further by adding a second jump in volatility, which can be generally specified as an idiosyncratic volatility jump model with jumps in both asset return and volatility:

$$\begin{aligned}\frac{dS(t)}{S(t-)} &= \mu_1(X(t))dt+v(X(t))dW_1(t)+dJ_1(t) \qquad (89)\\ dX(t) &= \mu_2(X(t))dt+\sigma_{21}(X(t))dW_1(t)+\sigma_{22}(X(t))dW_2(t)+dJ_2(t)+dJ_3(t),\end{aligned}$$

where the concurrent jump $J_1(t)$ and $J_2(t)$ remain unchanged: $J_1(t) = \sum_{i=1}^{N_1(t)}(e^{Z_i^s}-1)$, $J_2(t) = \sum_{i=1}^{N_1(t)} Z_{1i}^v$, but $J_3(t) = \sum_{i=1}^{N_2(t)} Z_{2i}^v$. $N_1(t)$ and $N_2(t)$ are assumed to be independent Poisson process with constant intensity λ_1 and λ_2 respectively, for simplicity's sake. The additional jump $dJ_3(t)$ allows the volatility process to jump more frequently and intensively than the concurrent jump model (74) suggests.

6.2 Expansion of at-the-money option price and sensitivities

Because the at-the-money option price or sensitivities can be written as

$$\frac{\partial^j}{\partial k^j}P(\tau,k,S_0,X_0)|_{k=0} = \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty}\mathbb{E}[g^{(j)}(S(\tau)-S_0) \mid N_1(\tau)=n_1, N_2(\tau)=n_2]\cdot\mathbb{P}\{N_1(\tau)=n_1, N_2(\tau)=n_2\}$$

$$\text{with } \mathbb{P}\{N_1(\tau)=n_1, N_2(\tau)=n_2\} = \frac{(\lambda_1\tau)^{n_1}e^{-\lambda_1\tau}}{n_1!}\frac{(\lambda_2\tau)^{n_2}e^{-\lambda_2\tau}}{n_2!},$$

the problem reduces to the expansion of

$$\mathbb{E}[g^{(j)}(S(\tau)-S_0) \mid N_1(\tau)=n_1, N_2(\tau)=n_2]. \tag{90}$$

We still follow the routine process: expand $S(\tau)$, composite it with $g^{(j)}$, and take expectations. Without time-varying intensity, a change of measure, which entails the expansion of the Radon-Nikodym derivative $\Lambda(\tau)$ and hence necessitates huge calculations, is unnecessary.

6.2.1 Inheritance

A first glance suggests that if $n_1 = 0$ or $n_2 = 0$, this model inherits from the ones that have already been studied.

$$\text{It inherits from } \begin{cases} \text{continuous diffusion model (2), if } n_1 = n_2 = 0. \\ \text{volatility jump model (39), if } n_1 = 0,\ n_2 > 0. \\ \text{concurrent jump model (74), if } n_1 > 0,\ n_2 = 0. \end{cases}$$

6.2.2 Analogy

If $n_1 > 0, n_2 > 0$, by analogy, we make decomposition $S(t) = S_0 S^C(t) S^J(t)$, where

$$\begin{aligned} S^C(t) \quad &: = \exp\{\int_0^t [\mu_1(X(s)) - \frac{1}{2}v^2(X(s))]ds + \int_0^t v(X(s))dW_1(s)\} \\ S^J(t) \quad &: = \exp(\sum_{i=1}^{N_1(t)} Z_i^s), \end{aligned}$$

By Itô formula, $S^C(t)$ follows:

$$\begin{aligned} \frac{dS^C(t)}{S^C(t)} &= \mu_1(X(t))dt + v(X(t))dW_1(t) \\ dX(t) &= \mu_2(X(t))dt + \sigma_{21}(X(t))dW_1(t) + \sigma_{22}(X(t))dW_2(t) + J_2(t)dN_1(t) + J_3(t)dN_2(t). \end{aligned}$$

According to theorem 7, we can immediately write the expansion of $S^C(\tau)$ as

$$\begin{aligned} S^C(\tau) &= \sum_{|\mathbf{M}_{n_1+n_2+1}|\leq J} \boldsymbol{\tau}^{\mathbf{M}_{n_1+n_2+1}} \cdot [\sum_{||\mathbf{i}_1||=M_1} \cdots \sum_{||\mathbf{i}_{n_1+n_2+1}||=M_{n_1+n_2+1}} \\ & D_{\mathbf{i}_1} C^{(j_1,j_2,\cdots,j_{n_1+n_2})}_{(\mathbf{i}_2,\cdots,\mathbf{i}_{n_1+n_2+1}),1}(S_0, X_0) \cdot (\prod_{p=1}^{n_1+n_2+1} J^{(p)}_{\mathbf{i}_p}(1))] + \mathcal{O}(\sum_{|\mathbf{M}_{n_1+n_2+1}|=J+1} \boldsymbol{\tau}^{\mathbf{M}_{n_1+n_2+1}}). \end{aligned} \tag{91}$$

Thus

$$\begin{aligned} \mathbb{E}[g^{(j)}(S(\tau)-S_0) \quad | \quad N_1(\tau)=n_1, N_2(\tau)=n_2] &= S_0^{1-j}\mathbb{E}_{n_1,n_2}[g^{(j)}(S^C(\tau)S^J(\tau)-1)] \\ &= S_0^{1-j}\mathbb{E}_{n_1,n_2}[g^{(j)}((91)\cdot \exp(\sum_{i=1}^{n_1} Z_i^s)-1)]. \end{aligned}$$

And then, we make composition of the expansion by expanding $g^{(j)}$ at $z_0 = S^J(\tau)-1$. Finally, using the aforementioned integration method and the integral operator $\mathcal{E}$ and $\mathcal{I}$, we are able to obtain the small-time expansion of at-the-money option price or sensitivities.

7 Future topic

[Aït-Sahalia et al.(2017)Aït-Sahalia, Li, and Li] develops a framework to derive the implied volatility expansion from the expansion of option price by matching terms. Given the expansion results of option price under the aforementioned 4 general types of stochastic models, one may nest it in this framework, i.e, to analyze the implied volatility properties under these more complicated models with jumps.

8 Appendix: A detailed discussion about GMM design for diffusions

This appendix is devoted to the GMM design proposed by [Hansen and Scheinkman(1995)] for diffusion process, with the original idea of GMM coming from [Hansen(1982)].

8.1 Derivation of moment conditions

Consider scalar diffusion of the type (1). To include parameter set θ, we may write it as

$$dX(t) = \mu(X(t);\theta)dt + \sigma(X(t);\theta)dW(t).$$

Assume that certain restrictions (e.g. growth condition, existence of a proper initialization, etc.) to ensure this SDE admits a unique stationary solution hold. A solution $X(t)$ is called stationary if $X(t)$ and $W(t)$ are strictly stationarily correlated, i.e. if the probability law of the system

$$(X, dW) := (X(t), -\infty < t < \infty, B(v)-B(u), -\infty < u < v < \infty)$$

is invariant under the time shift.

Define operators $\mathcal{P}$:

$$\mathcal{P}_t\phi(y) := \mathbb{E}[\phi(X(t)) \mid X(0)=y]$$

and its "derivative" $\mathcal{G}$:

$$\mathcal{G}\phi := \lim_{t\to 0+} \frac{\mathcal{P}_t\phi - \phi}{t}.$$

$\mathcal{G}$ is referred as the infinitesimal generator of $X(t)$, which is shown to be equivalent with the definition below:

$$\mathcal{G}\phi := \mu\frac{\partial\phi}{\partial y} + \frac{1}{2}\sigma^2\frac{\partial^2\phi}{\partial y^2}. \tag{92}$$

The domain $\mathcal{D}$ of this generator is the family of functions ϕ for which $\mathcal{G}\phi$ is well defined. Given the stationarity of $\{X(t)\}$, $\mathbb{E}[X(t)]$ is independent of t, implying that its derivative with respect to t is zero. Therefore, the first set of moment conditions is constructed as

$$\mathbb{E}[\mathcal{G}\phi(X(t))] = 0 \text{ for all } \phi \in \mathcal{D}. \tag{93}$$

Another set of moment conditions is constructed by an interchange of $\mathcal{G}$ and conditional expectation:

$$\mathbb{E}[\mathcal{G}\phi(X(t+\triangle)) \mid X(t) = y] = \mathcal{G}\{\mathbb{E}[\phi(X(t+\triangle)) \mid X(t) = y]\} \text{ for all } \phi \in \mathcal{D}.$$

Using these moment conditions, one can perform GMM (see [Hansen(1982)]) on diffusions. Now, I elaborate on the procedure and some practical concerns of GMM.

8.2 GMM formulation

Suppose $\theta \in \Theta \subset \mathbb{R}^p$. Denote θ_0 as the true parameter. A r-dimensional vector function g specifies the moment conditions:

$$g(x,\theta) := \begin{bmatrix} g_1(x,\theta) \\ \vdots \\ g_r(x,\theta) \end{bmatrix} \text{ that satisfies } \mathbb{E}(g(X,\theta_0)) = 0.$$

In our context, we can simply specify the function g_i as $\mathcal{G}\phi$ where $\phi \in \mathcal{D}$, as in (93).

An empirical approximation of $\mathbb{E}(g(X,\theta_0))$ is, of course, $\frac{1}{n}\sum_{i=1}^{n} g(X_i,\theta)$. The usual moment estimator requires $r = p$ in order to ensure that the equation system with r equations and p unknowns

$$\frac{1}{n}\sum_{i=1}^{n} g(X_i,\theta) = 0$$

admits a unique solution $\hat{\theta}_n$, which serves as an estimator of θ_0.

However, it is often the case that the number of equations is larger than the number of parameters, i.e. $r > p$, the case of which is called "over-identified". Some equations have to make sacrifice, and the problem is therefore transformed into an optimization problem:

$$\hat{\theta}_n = \arg\min_{\theta\in\Theta}[\frac{1}{n}\sum_{i=1}^{n} g(X_i,\theta)]^T \hat{W}_n [\frac{1}{n}\sum_{i=1}^{n} g(X_i,\theta)] \tag{94}$$

where $\hat{W}_n \in \mathbb{R}^{r\times r}$ is a semi-positive definite matrix, $\hat{W}_n \xrightarrow{p} W_0$, W_0 is a deterministic semi-positive definite matrix. $\hat{W}_n$ denotes the weight put on each of the r equations. The choice of $\hat{W}_n$ depends on econometricians. However, there exists an optimal limit weight matrix W_0 that minimizes the asymptotic variance of the estimator $\hat{\theta}_n$, as shown later. Suppose we take $\hat{W}_n = W_0$ and obtain a $\hat{\theta}_{n0}$, then $\hat{\theta}_n$ and $\hat{\theta}_{n0}$ are asymptotically equivalent, i.e. $d(\hat{\theta}_n, \hat{\theta}_{n0}) \xrightarrow{p} 0$.

8.3 GMM identification

To ensure the GMM estimator $\hat{\theta}_n$ given by (94) is well-defined and unique, the condition below is often assumed.

Theorem 14 *Let $g_0(\theta) := \mathbb{E}(g(X,\theta))$, and then θ_0 is a root of $g_0(\theta)$. If $\forall\theta \neq \theta_0, W_0 g_0(\theta) \neq 0$, then $g_0(\theta)^T W_0 g_0(\theta)$ is minimized uniquely at θ_0.*
Proof. *Since W_0 is semi-positive definite, let $W_0 = CC^T$ where $C \in \mathbb{R}^{r\times r}$ is a non-singular matrix. If there exists a root of $g_0(\theta)^T W_0 g_0(\theta)$ other than θ_0, which I denote as θ_1, then $g_0(\theta_1)^T CC^T g_0(\theta_1) = 0$. Therefore, $C^T g_0(\theta_1) = 0$, and $W_0 g_0(\theta_1) = 0$, hence $\theta_1 = \theta_0$.* ■

8.4 The consistency and asymptotic normality of GMM estimator

In practice, we hope that the GMM estimator $\hat{\theta}_n$ defined by (94) is consistent, i.e. $\hat{\theta}_n \xrightarrow{p} \theta_0$, and is asymptotically normal, i.e. $\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathcal{N}(0, V)$. Several mild conditions are required to ensure these, as shown below.

8.4.1 Consistency

Theorem 15 *Suppose $g(x,\theta)$ is continuous with respect to θ on a compact set Θ, and $\mathbb{E}(\sup_{\theta\in\Theta}\|g(X,\theta)\|) < \infty$, then $\hat{\theta}_n \xrightarrow{p} \theta_0$.*
Proof. *Since $\hat{\theta}_n$ is an M-estimator, we only need to verify the conditions for an M-estimator to be consistent (e.g. see [Vaart(1998)] chapter 5 theorem 5.7 for those conditions). The condition of identification and compactification naturally hold. let*

$$g_0(\theta) := \mathbb{E}(g(X,\theta)),\ \ M(\theta) := g_0(\theta)^T W_0 g_0(\theta)$$

By the continuity of $g(x,\theta)$,

$$g(x,\theta_1) - g(x,\theta_2) \xrightarrow{p} 0,\ \ as\ \|\theta_1 - \theta_2\| \to 0.$$

Thus

$$\mathbb{E}[g(x,\theta_1) - g(x,\theta_2)] \to 0,\ \ as\ \|\theta_1 - \theta_2\| \to 0,$$

which yields the continuity of $g_0(\theta)$,and hence $M(\theta)$. Finally, we prove the uniform convergence:

$$\sup_{\theta\in\Theta} \mid g_n(\theta)^T \hat{W}_n g_n(\theta) - g_0(\theta)^T W_0 g_0(\theta) \mid \xrightarrow{p} 0.$$

By Markov inequality, to prove the convergence below suffices:

$$\mathbb{E}(\sup_{\theta\in\Theta} \mid g_n(\theta)^T \hat{W}_n g_n(\theta) - g_0(\theta)^T W_0 g_0(\theta) \mid) \to 0.$$

By triangular inequality,

$$\begin{aligned}\sup_{\theta\in\Theta} \mid & g_n(\theta)^T \hat{W}_n g_n(\theta) - g_0(\theta)^T W_0 g_0(\theta) \mid \leq \sup_{\theta\in\Theta} \mid g_n(\theta)^T \hat{W}_n g_n(\theta) - g_n(\theta)^T \hat{W}_n g_0(\theta) \mid + \\ \sup_{\theta\in\Theta} \mid & g_n(\theta)^T \hat{W}_n g_0(\theta) - g_0(\theta)^T \hat{W}_n g_0(\theta) \mid + \sup_{\theta\in\Theta} \mid g_0(\theta)^T \hat{W}_n g_0(\theta) - g_0(\theta)^T W_0 g_0(\theta) \mid .\end{aligned}$$

Taking expectations on RHS, the first and the second term converge to zero because $\hat{W}_n = O_p(1)$, $g_n(\theta) - g_0(\theta) \xrightarrow{p} 0$, and $\mathbb{E}(\sup_{\theta\in\Theta}\|g(X,\theta)\|) < \infty$. The third term also converges to zero because $\hat{W}_n - W_0 = o_p(1)$ and $\mathbb{E}(\sup_{\theta\in\Theta}\|g(X,\theta)\|) < \infty$. ■

8.4.2 Asymptotic normality

Theorem 16 *Suppose $g(x,\theta)$ is once differentiable with respect to θ, and*

$$G := \mathbb{E}\begin{bmatrix} \frac{\partial g_1}{\partial \theta_1} & \cdots & \frac{\partial g_1}{\partial \theta_p} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_r}{\partial \theta_1} & \cdots & \frac{\partial g_r}{\partial \theta_p} \end{bmatrix}|_{\theta=\theta_0}$$

is a full rank Jacobian matrix of g at θ_0. Denote

$$\Omega := Var \begin{pmatrix} g_1(X,\theta_0) \\ \vdots \\ g_r(X,\theta_0) \end{pmatrix}$$

as the variance-covariance matrix of g at θ_0. Then

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathcal{N}(0, (G^T W_0 G)^{-1} G^T W_0 \Omega W_0^T G (G^T W_0 G)^{-1}). \tag{95}$$

Proof. The proof is based on a Taylor expansion of $g(X_i, \hat{\theta}_n)$ at θ_0. Since $\frac{\partial g^T \hat{W}_n g}{\partial \theta} = 2 \frac{\partial g^T}{\partial \theta} \hat{W}_n g$ and $\hat{\theta}_n$ is a maximizer of $g^T \hat{W}_n g$,

$$(\frac{1}{n}\sum_{i=1}^{n} \frac{\partial g^T(X_i, \hat{\theta}_n)}{\partial \theta}) \hat{W}_n (\frac{1}{n}\sum_{i=1}^{n} g(X_i, \hat{\theta}_n)) = 0.$$

Since $\hat{\theta}_n \xrightarrow{p} \theta_0$, we can expand $g(X_i, \hat{\theta}_n)$ at θ_0 with Lagrangian remaining term:

$$(\frac{1}{n}\sum_{i=1}^{n} \frac{\partial g^T(X_i, \hat{\theta}_n)}{\partial \theta}) \hat{W}_n (\frac{1}{n}\sum_{i=1}^{n} g(X_i, \theta_0) + \frac{1}{n}\sum_{i=1}^{n} \frac{\partial g(X_i, \hat{\theta}_n^*)}{\partial \theta} (\hat{\theta}_n - \theta_0)) = 0$$

where $\hat{\theta}_n^*$ is a point between $\hat{\theta}_n$ and θ_0. Now that we have

$$\frac{1}{n}\sum_{i=1}^{n} \frac{\partial g^T(X_i, \hat{\theta}_n)}{\partial \theta} \xrightarrow{p} G^T, \ \sum_{i=1}^{n} \frac{\partial g(X_i, \hat{\theta}_n^*)}{\partial \theta} \xrightarrow{p} G, \ \hat{W}_n \xrightarrow{p} W_0,$$

and

$$\sqrt{n}(\frac{1}{n}\sum_{i=1}^{n} g(X_i, \theta_0) - 0) \xrightarrow{d} \mathcal{N}(0, \Omega),$$

it follows that

$$G^T W_0 \mathcal{N}(0, \Omega) + G^T W_0 G \sqrt{n}(\hat{\theta}_n - \theta_0) = o_p(1).$$

The asymptotic normality indicated by (95) is immediately obtained from this. ■

8.5 Some practical concerns

A remaining question is how to specify the limit weight matrix W_0.

Theorem 17 *The asymptotic variance* $(G^T W_0 G)^{-1} G^T W_0 \Omega W_0^T G (G^T W_0 G)^{-1}$ *is minimized at* $W_0 = \Omega^{-1}$*, an optimal choice of* W_0.

Proof. *Denote* $V(W) := (G^T W G)^{-1} G^T W \Omega W^T G (G^T W G)^{-1}$*. Assume* $\Omega = CC^T$ *where* C *is a non-singular matrix. Let* $D := (G^T W G)^{-1} G^T W C - (G^T (CC^T)^{-1} G)^{-1} G^T (C^{-1})^T$*. It is easy to verify that* $V(W) - V(\Omega^{-1}) = DD^T$*, which indicates that* $V(W) - V(\Omega^{-1})$ *is semi-positive definite.* ■

However, in practice we don't know Ω either. Then we can complete in two steps by plugging in an estimated Ω. For example, first take $\hat{W}_n = I$ and obtain a corresponding $\hat{\theta}_n$. Take θ_0 as this $\hat{\theta}_n$ to get an estimated $\Omega = \hat{\Omega}$. In the second step, take $\hat{W}_n = \hat{\Omega}$ to obtain the final $\hat{\theta}_n$. If we follow this route, then the asymptotic normality reduces to

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathcal{N}(0, (G^T \Omega^{-1} G)^{-1}),$$

which allows statistical inference under GMM.

9 Acknowledgement

I am truly grateful to the excellent professors at Peking University Guanghua School of Management and other departments for guiding my 4 years of study in the picturesque campus. In particular, I hope to extend my very enthusiastic thankfulness to my advisor Professor Chenxu Li. It was he who exposed me to the splendid world of stochastic calculus and derivative pricing, and who mentored my own research in this area. At times when I experienced hardships in getting through the intricate symbolic calculations, he was always kind and patient to me, without whose encouraging and caring words I couldn't have even stepped into quantitative finance.

Besides, I am thankful to Professor Michael Pettis, who in my sophomore year extricated me from my very vexations about future by inviting me to an intriguing Sunday seminar. I still remember the time when I for the first time in my life visited a professor's house, and my nervousness for the meeting, which turned out to be inspiring and convivial.

Time elapses without notice, but friendship lasts long. Now, it is time to say goodbye to my dear peers who offered me the best companionship. Of course, it was my parents who assisted me most for my undergraduate study, maintaining a warm hearth and home. Thank you all for contributing to my growth, mentally and physically.

References

[Aït-Sahalia(2002)] Aït-Sahalia, Y., 2002. Maximum likelihood estimation of discretely sampled diffusions: A closed-form approximation approach. Econometrica 70 (1), 223–262.

[Aït-Sahalia(2008)] Aït-Sahalia, Y., 2008. Closed-form likelihood expansions for multivariate diffusions. Annals of Statistics 36, 906–937.

[Aït-Sahalia et al.(2017)Aït-Sahalia, Li, and Li] Aït-Sahalia, Y., Li, C., Li, C. X., 2017. Closed-form implied volatility surfaces for stochastic volatility models. Tech. rep., Princeton University and Peking University.

[Bakshi et al.(1997)Bakshi, Cao, and Chen] Bakshi, G., Cao, C., Chen, Z., 1997. Empirical performance of alternative option pricing models. Journal of Finance 52 (5), 2003–2049.

[Bates(2000)] Bates, D. S., 2000. Post-'87 crash fears in the S&P 500 futures option market. Journal of Econometrics 94 (1-2), 181–238.

[Black and Scholes(1973)] Black, F., Scholes, M. S., May-June 1973. The pricing of options and corporate liabilities. Journal of Political Economy 81 (3), 637–54.

[Chang and Chen(2011)] Chang, J., Chen, S., 2011. On the approximate maximum likelihood estimation for diffusion processes. Annals of Statistics 39 (6), 2820–2851.

[Cox et al.(1985)Cox, Ingersoll, and Ross] Cox, J. C., Ingersoll, J. E., Ross, S. A., March 1985. A theory of the term structure of interest rates. Econometrica 53 (2), 385–407.

[Duffie et al.(2000)Duffie, Pan, and Singleton] Duffie, D., Pan, J., Singleton, K., 2000. Transform analysis and asset pricing for affine jump-diffusions. Econometrica 68 (6), 1343–1376.

[Dupire(1994)] Dupire, B., February 1994. Pricing with a smile. Risk , 18–20.

[Gallant et al.(1999)Gallant, Hsu, and Tauchen] Gallant, A. R., Hsu, C.-T., Tauchen, G., November 1999. Using daily range data to calibrate volatility diffusions and extract the forward integrated variance. The Review of Economics and Statistics 81 (4), 617–631.

[Hansen(1982)] Hansen, L. P., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50, 1029–1054.

[Hansen and Scheinkman(1995)] Hansen, L. P., Scheinkman, J. A., 1995. Back to the future: Generating moment implications for continuous-time Markov processes. Econometrica 63, 767–804.

[Heston(1993)] Heston, S. L., 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. Review of Financial Studies 6 (2), 327–343.

[Hull and White(1987)] Hull, J., White, A., 1987. The pricing of options on assets with stochastic volatilities. The Journal of Finance 42 (2), 281–300.

[Ito and Nisio(1964)] Ito, K., Nisio, M., 1964. On stationary solutions of a stochastic differential equation. J. Math. Kyoto Univ. 4 (1), 1–75.

[Karatzas and Shreve(1991)] Karatzas, I., Shreve, S. E., 1991. Brownian Motion and Stochastic Calculus, 2nd Edition. Vol. 113 of Graduate Texts in Mathematics. Springer-Verlag, New York.

[Li(2013)] Li, C., 2013. Maximum-likelihood estimation for diffusion processes via closed-form density expansions. Annals of Statistics 41, 1350–1380.

[Li(2014)] Li, C., 2014. Closed-form expansion, conditional expectation, and option valuation. Mathematics of Operations Research 39, 487–516.

[Lijoi and Regazzini(2004)] Lijoi, A., Regazzini, E., 2004. Means of a dirichlet process and multiple hypergeometric functions. The Annals of Probability 32 (2), 1469–1495.

[Merton(1976)] Merton, R. C., 1976. Option pricing when underlying stock returns are discontinuous. Journal of financial economics 3, 125–144.

[Oksendal(2010)] Oksendal, B., 2010. Stochastic Differential Equations: An Introduction with Applications, 6th Edition. Springer-Verlag, Berlin Heidelberg.

[Sargan(1958)] Sargan, J. D., 1958. The estimation of economic relationships using instrumental variables. Econometrica 26, 393–415.

[Shreve(2004)] Shreve, S., 2004. Stochastic Calculus for Finance II: Continuous-Time Models. Springer-Verlag, New York, U.S.

[Stein and Stein(1991)] Stein, E. M., Stein, J. C., 1991. Stock price distributions with stochastic volatility: An analytic approach. Review of Financial Studies 4, 727–752.

[Sundaresan(2000)] Sundaresan, S., 2000. Continuous-time methods in finance: A review and an assessment. Journal of Finance 55, 1569–1622.

[Vaart(1998)] Vaart, A. V. D., 1998. Asymptotic Statistics. Cambridge University Press, Cambridge, U.K.

[Vasicek(1977)] Vasicek, O., November 1977. An equilibrium characterization of the term structure. Journal of Financial Economics 5 (2), 177–188.

[Wang(2020)] Wang, B., 2020. A method for handling volatility jumps. Ph.D. thesis, Peking University.

[Watanabe(1987)] Watanabe, S., 1987. Analysis of Wiener functionals (Malliavin calculus) and its applications to heat kernels. Annals of Probability 15, 1–39.

[Yoshida(1992)] Yoshida, N., 1992. Asymptotic expansions for statistics related to small diffusions. Journal of the Japan Statistical Society 22, 139–159.

[Yoshida(1997)] Yoshida, N., 1997. Malliavin calculus and asymptotic expansion for martingales. Probability Theory and Related Fields 109, 301–342.

[Yu(2019)] Yu, T., 2019. An expansion method in dealing with option price and implied volatility under stochastic volatility jump model. Ph.D. thesis, Peking University.