Pricing Average and Spread Options under Local-Stochastic Volatility Jump-Diffusion Models

Kenichiro Shiraya
The University of Tokyo

Akihiko Takahashi
The University of Tokyo

Pricing Average and Spread Options under Local-Stochastic Volatility Jump-Diffusion Models [*]

Kenichiro Shiraya[†]   Akihiko Takahashi[‡]

November 30, 2017

Abstract

This paper presents a new approximation formula for pricing multi-dimensional discretely monitored average options in a local-stochastic volatility (LSV) model with jumps by applying an asymptotic expansion technique. Moreover, it provides a justification of the approximation method with some asymptotic error estimates for general payoff functions. Particularly, our model includes local volatility functions and jump components in the underlying asset price as well as its volatility processes. To the best of our knowledge, the proposed approximation is the first one which achieves analytic approximations for the average option prices in this environment.

In numerical experiments, by employing several models, we provide approximate prices for the listed average and calendar spread options on the WTI futures based on the parameters through calibration to the listed (plain-vanilla) futures options prices. Then, we compare those with the CME settlement prices, which confirms the validity of the method.

Moreover, we show that the LSV with jumps model is able to replicate consistently and precisely listed futures option, calendar spread option and average option prices with common parameters.

[*]Forthcoming in Mathematics of Operations Research.

[†]Graduate School of Economics, University of Tokyo

[‡]Graduate School of Economics, University of Tokyo

1 Introduction

It is well-known that discretely monitoring average options are mainly traded in over-the-counter (OTC) commodity markets. However, it is a tough task to compute an average option price with the computational speed fast enough for the practical purpose, mainly due to the difficulty of the analytical tractability. For instance, although the Monte Carlo method is easy to implement, it requires a substantial computational time to obtain an accurate value. On the other hand, analytic approximations for pricing discrete average options are rarely known except for models which are too simple to replicate the volatility smile and skew.

To overcome the difficulties, this paper develops a new analytical approximation formula for average option prices with the model which is able to reproduce the market volatility structures consistently.

Our contribution is summarized as follows:

- Derived asymptotic expansions for approximating average option prices under multi-dimensional general diffusion processes with jumps, which is a necessary setup for applications in practice. To obtain this formula, we need new conditional expectation formulae with some approximation techniques, which are not shown in previous papers: We note that although LSV with jumps models for basket option is considered in [34], the formula derived by [34] cannot be applied to pricing average options.

- Showed that the LSV with jumps model is able to replicate consistently and precisely listed futures option, calendar spread option and average option prices with common parameters. Although the similar analyses are done in [32] or [33], they used models without jumps and did not take account of spread options for calibration of the correlations between two futures prices underlying in target average options. In particular, we demonstrated that our method works well for SABR and Heston-type stochastic volatility with LV(local volatility) plus jumps models which are able to calibrate to the futures, spread and exotic option markets better than the models without jumps. Hence, the current work provides an important extension of our previously developed methods.

- Provided a justification of our approximation method with some asymptotic error estimates for general payoff functions in Section 3. The previous papers (e.g. [32], [33], [34] and [35]) did not provide such a consideration. Moreover, we confirmed the consistency between the theoretical asymptotic errors in Theorem 3.2 with Appendix C and the numerical errors in Figures from 3 to 8 as well as the ones in Figures 9 and 10 showing stress tests with regard to parameters.

As for pricing discretely monitored average option, Cai et al. [6] derived an approximation formula under one-dimensional general diffusion processes by the recursive method. Fusai - Meucci [17] provided a recursive algorithm as well as a control variate technique to price discretely monitored average options under the general exponential Lévy process. Fusai - Kyriakou [16] derived lower and upper bound under processes whose characteristic functions are known. For the one-dimensional time-changed Lévy process, Yamazaki [43] calculated the price by applying Gram-Charlier expansion, and Zeng - Kwok [45] derived accurate lower and upper

bounds. For more general case, Cai et al. [7] recently proposed a general framework for average option prices under general one-dimensional Markov processes. For continuous average options, Cai et al. [8] derives analytical approximations under general regime-switching Markov models.

However, in order to evaluate most of the average options traded in commodity markets, we need multi-dimensional models with multiple underlying assets to describe the concrete feature of the products. For example, in WTI average options, the underlying price is the average of the settlement prices of the first nearby WTI futures contract during the last one month prior to the maturity of the option. We also note that the expiration of an average option is the last business day of a calendar month, while trading of a WTI futures contract usually ceases on the third business day prior to the twenty-fifth calendar day. For example, we assume time-t prices of two futures contracts with maturities T_1 and T_2 as $f_1(t, T_1)$ and $f_2(t, T_2)$, respectively. In addition, let the reference dates $t_1, \cdots, t_{n_1}$ for the first contract and $t'_1, \cdots, t'_{n_2}$ for the subsequent contract with relation $t < t_1 < \cdots < t_{n_1} < t'_1 < \cdots < t'_{n_2} = T$. Then, the underlying price A_T of the WTI average option with maturity T is expressed as follows:

$$
A_T \quad = \quad \frac{1}{n_1 + n_2} \left(\sum_{i=1}^{n_1} f_1(t_i, T_1) + \sum_{i=1}^{n_2} f_2(t'_i, T_2) \right), \tag{1}
$$

and the payoff functions of average call and put options with strike price K are given by $\max\{A_T - K, 0\}$ and $\max\{K - A_T, 0\}$, respectively. Hence, the settlement prices of WTI futures contracts with two consecutive maturities are relevant for the underlying price of the average option. Similarly, calibrations to listed WTI futures option with two consecutive maturities are relevant for pricing an average option based on the market information. Clearly, these features are also true for a calendar spread option by definition of the contract.

The proposed models in the current work take the specific features embedded in commodity derivatives into consideration, which have not been incorporated in other papers for pricing average and spread options. While Shiraya - Takahashi [32], [33] approximated the option values under multi-dimensional local-stochastic volatility models, it only considered the models without jumps, which are less capable of duplicating the real markets than the models treated in this paper. We remark that there exist other interesting works on spread options such as Alòs et al. [2], Alòs - León [1], and Caldana - Fusai [9].

Although local-stochastic volatility (LSV) models are mainly used in the market practice, they are not always enough to fit a volatility smile and skew. Hence, some advanced researches investigated a local-stochastic volatility with jump model. For instance, Among them, Eraker [14] found that the models with jump components in the underlying price and volatility processes showed better performance in fitting to option prices and the underlying price returns' data simultaneously in stock markets. Sepp [30] showed the necessity of jumps on volatility to calibrate to SPX and VIX skews simultaneously. Pagliarani - Pascucci [29] derived an analytical approximation of plain-vanilla option prices by applying the adjoint expansion method. Li [26] introduced the closed-form likelihood expansions for jump-diffusion models as an application in statistics. More recently, Shiraya - Takahashi [34], [35] obtained an analytical approximation formula for basket option prices with local-stochastic volatility with jump diffusion model. However, they have not provided asymptotic error or justifications for the approximation of non-smooth functions, and their conditional expectation formulae can not be directly applied to

3

pricing the average options. This paper is the first one that derives an analytical approximation formula for the discretely monitored average option price and its asymptotic error, which means the order of the approximation error with respect to a small noise parameter ϵ, under a model which admits a local volatility function and jumps both in the underlying asset price and its volatility processes. Our closed form approximation has an advantage in making use of the better calibration to the traded listed options whose underlying assets are included in average and spread options' underlying. We also note that the formula is general enough to be applicable to pricing discretely monitored average options, which are popularly traded in practice. In addition, since continuously monitored average options can be calculated in a very similar way, and there are no continuously monitored options in real markets, we omit the details in this paper. We also remark that there exist mathematically interesting researches on continuously monitored average options such as a recent work by Cai - Kou [5]. On applications of an asymptotic expansion for Wiener functionals to continuously monitored average options, please see Kunitomo - Takahashi [24] [25], Shiraya et al. [37] and Takahashi [39].

The numerical experiments provide estimates of average and spread option prices based on the parameters obtained by calibration to the market prices of WTI futures options in local-stochastic volatility with jump models. Then, those estimated prices are compared with the settlement prices of average options and spread options traded on CME Group, and we show our model and method work well in practice.

For the 3rd order (ϵ^3) corrections, we use the ones obtained by a model without jumps rather than with jumps because of the substantial increase in the computational burdens of the 3rd order expansion for the jump components. However, the 3rd order corrections only for the diffusion terms improves approximations effectively in numerical experiments, whose details will be shown in Figures from 3 to 8. We note that deriving more conditional expectation formulae enables us to obtain the full 3rd order corrections.

We also remark that in principle, our method can be applied to general local stochastic volatility with jumps models. However, infinite activity models are not treated, because we need to develop some computational technique other than the one applied in the current work. Hence, the explicit numerical evaluation in expansions for models including jumps with infinite activity will be one of our future research topics.

The organization of the paper is as follows: The next section discusses an asymptotic expansion method for general stochastic differential equations (SDEs) with jumps. Section 3 shows the validity of our method for option pricing, which is a special case of an asymptotic expansion for a non-smooth functional of the solution to the SDEs with jumps. Particularly, we discuss in details on local stochastic volatility with compound Poisson jumps models used for the numerical analyses in the following sections. In Section 4, after describing the structure and our model specific for the average options on commodities, we derive a new approximate pricing formula based on the asymptotic expansion technique given in Section 2 and 3 for local stochastic volatility with compound Poisson jumps models. Section 5 shows numerical examples of approximate prices for average and calendar spread options on the WTI futures by using the parameters through calibration to the plain vanilla option prices, which are compared with the CME settlement prices. Section 6 concludes. Appendix A provides two propositions used in Section 3. Appendix B gives proofs of Theorem 3.2 and Corollary 3.3. Appendix C discusses error estimates of our asymptotic expansions. Appendix D shows the effect of the third-order

4

expansion for jump components.

An online appendix (Shiraya-Takahashi [31]) provides results omitted in this paper. Namely, it explains all the definitions and equations necessary for practical computations of an option pricing formula in Theorem 4.3. It also lists up the conditional expectation formulas used in the derivation of the theorem.

2   Asymptotic Expansion for SDE with Jumps

In this section, we describe an asymptotic expansion of the solution to a multi-dimensional perturbed SDE with jumps, which is an ingredient for deriving approximations of the underlying random variables for average and spread options.

On Wiener-Poisson space $(\Omega, \mathcal{F}, P)$, suppose a filtered probability space satisfying the usual conditions $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t \geq 0})$ is given. Then, let X_t $(t \in [0, T])$ be a D-dimensional process which satisfies the following SDE:

$$X_t = x + \int_0^t \mu(s, X_{s-})ds + \int_0^t \Phi(s, X_{s-})dW_s + \sum_{l=1}^{\mathcal{N}} \int_0^t \int_{E_l} \gamma_l(s, z, X_{s-})\tilde{N}_l(ds, dz), \qquad (2)$$

with $x \in \mathbf{R}^D$; $\mu : [0, T] \times \mathbf{R}^D \times \Omega \to \mathbf{R}^D$, $\Phi : [0, T] \times \mathbf{R}^D \times \Omega \to \mathbf{R}^D \times \mathbf{R}^m$, $\gamma_l : [0, T] \times E_l \times \mathbf{R}^D \times \Omega \to \mathbf{R}^D$, are predictable processes; W is a m-dimensional Brownian motion; Each $N_l(dt, dz)$ is a Poisson random measure on $[0, T] \times E_l$ where $(E_l, \underline{\mathbf{E}}_l)$ is a measurable space with $E_l \subset \mathbf{R}^{e_l}$, $e_l \in \mathbf{N}$; The intensity measure of N_l is $dt \times \nu_l(dz)$, where $\nu_l(dz)$ is a positive σ-finite measure on $(E_l, \underline{\mathbf{E}}_l)$. Then, $\tilde{N}_l(dt, dz) := N_l(dt, dz) - dt \times \nu_l(dz)$ is a compensated Poisson random measure.

Let us next introduce a perturbation parameter $\epsilon \in [0, 1]$ to the processes X, and the perturbed process $X^{(\epsilon)}$, which is the solution to the SDE:

$$X_t^{(\epsilon)} = x + \int_0^t \mu(s, X_{s-}^{(\epsilon)})ds + \epsilon \int_0^t \tilde{\Phi}(s, X_{s-}^{(\epsilon)})dW_s + \sum_{l=1}^{\mathcal{N}} \int_0^t \int_{\tilde{E}_l} \gamma_l(s, \epsilon\tilde{z}, X_{s-}^{(\epsilon)})\tilde{N}_l(ds, d\tilde{z}),$$

with $\gamma_l(s, 0, x) = 0$. Hereafter, let us use the same notations Φ, z E_l instead of $\tilde{\Phi}$, $\tilde{z}$ and $\tilde{E}_l$ as in the original process: That is,

$$X_t^{(\epsilon)} = x + \int_0^t \mu(s, X_{s-}^{(\epsilon)})ds + \epsilon \int_0^t \Phi(s, X_{s-}^{(\epsilon)})dW_s + \sum_{l=1}^{\mathcal{N}} \int_0^t \int_{E_l} \gamma_l(s, \epsilon z, X_{s-}^{(\epsilon)})\tilde{N}_l(ds, dz), \quad (3)$$

with $\gamma_l(s, 0, x) = 0$. Please see an example in Appendix C (particularly after the equation (109)) on a discussion for an error estimate of an asymptotic expansion of a perturbed SDE. Here, we introduce the notation $|A| = \sqrt{\sum_{i,j} a_{i,j}^2}$ for any matrix $A = (a_{i,j})$, $|a| = \sqrt{\sum_i a_i^2}$ for any vector $a = (a_i)$. Then, we suppose the following conditions.

Condition (A)

i $\gamma_l(t, z, x) \in C^\infty(z)$, $\mu(t, x)$, $\Phi(t, x)$, $\gamma_l(t, z, x)/\eta_l(z)$, $\frac{\partial^j}{\partial \epsilon^j}\gamma_l(t, \epsilon z, x)/\eta_l(z) \in C_b^\infty(x)$ $(j \in \mathbf{N})$, $(l = 1, \cdots, m)$ with $\eta_l : E_l \to \mathbf{R}$, $\eta_l \in \cap_{p \geq 2} L^p(E_l, \nu_l)$.

ii $|\mu(t,0)|$, $|\Phi(t,0)|$, $|\gamma_l(t,z,0)/\eta_l(z)|$, $|\frac{\partial^j}{\partial\epsilon^j}\gamma_l(t,\epsilon z,0)/\eta_l(z)| < Z_t$ for $j \in \mathbf{N}$. There exits a predictable process Z_t, such that Z_t satisfies $|||Z_T|||_p \equiv \left(\int_0^T \mathbf{E}[|Z_t|^p]dt\right)^{\frac{1}{p}} < \infty$ for $\forall p \geq 1$.

Under Condition (A), we are able to apply Proposition C.4 (or the subsequent related results) in Appendix C to obtain the L^p-boundedness for all $p \in [1,\infty)$ of the solutions to a (graded) system of the SDEs, which consists of $X_t^{(\epsilon)}$ and $\frac{\partial^j}{\partial\epsilon^j}X_t^{(\epsilon)}$, $1 \leq j \leq J$ for an arbitrary $J \in \mathbf{N}$: that is, for any $p \geq 1$, $0 \leq i \leq D$ and $0 \leq j \leq J$,

$$\mathbf{E}\left[\sup_{0 \leq t \leq T}\left|\frac{\partial^j}{\partial\epsilon^j}X_t^{i,(\epsilon)}\right|^p\right] < \infty. \tag{4}$$

We note that $X_T^{(\epsilon)}$ has a smooth asymptotic expansion in a sense that for an arbitrary $N \in \mathbf{Z}_+(N = 0$ or $N \in \mathbf{N})$,

$$X_T^{(\epsilon)} - \sum_{n=0}^N \frac{\epsilon^n}{n!}X_T^{(n)} = O(\epsilon^{N+1}C(T)),$$

with $X_T^{(n)} := (X_T^{1,(n)}, \cdots, X_T^{D,(n)})'$, $X_T^{i,(n)} = \frac{\partial^n X_T^{i,(\epsilon)}}{\partial\epsilon^n}\Big|_{\epsilon=0} \in L^p$, where $C(T)$ is an increasing function of T. It means that

$$\limsup_{\epsilon\downarrow 0} \frac{\left\|X_T^{(\epsilon)} - \sum_{n=0}^N \frac{\epsilon^n}{n!}X_T^{(n)}\right\|_{L^p}}{\epsilon^{N+1}} < \infty, \text{ for } \forall p > 1,$$

that is, for $\forall p > 1$,

$$\limsup_{\epsilon\downarrow 0} \frac{1}{\epsilon^{N+1}}\mathbf{E}\left[\left|X_T^{(\epsilon)} - \sum_{n=0}^N \frac{\epsilon^n}{n!}X_T^{(n)}\right|^p\right] = \limsup_{\epsilon\downarrow 0} \frac{1}{\epsilon^{N+1}}\mathbf{E}\left[\left|\int_0^1 \frac{(1-u)^N\epsilon^{N+1}}{N!}X_T^{(N+1),(\epsilon u)}du\right|^p\right]$$

$$\leq \frac{1}{(N+1)!}\sup_{u\in[0,1]}\mathbf{E}\left[\left|X_T^{(N+1),(\epsilon u)}\right|^p\right] < \infty. \tag{5}$$

Moreover, $\frac{\partial^n}{\partial\epsilon^n}X_t^{(\epsilon)} = \left(\frac{\partial^n}{\partial\epsilon^n}X_t^{1,(\epsilon)}, \cdots, \frac{\partial^n}{\partial\epsilon^n}X_t^{D,(\epsilon)}\right)$ is recursively determined by the following: For any multi-index $\alpha = \alpha^{(r)} := (\alpha_1, \cdots, \alpha_r) \in \{1, \cdots, D\}^r$, $r \geq 1$ with the length $|\alpha^{(r)}| = r$, and $\hat{\alpha}_l^{(r)} := (\hat{\alpha}_{l,1}, \cdots, \hat{\alpha}_{l,r}) \in \{1, \cdots, e_l\}^r$, we denote by $\partial_{\alpha^{(r)}}^r$ the partial derivative $\frac{\partial^r}{\partial x_{\alpha_1}\cdots\partial x_{\alpha_r}}$, and by $\partial_{\hat{\alpha}_l^{(\hat{r})}}^{\hat{r}}$ the partial derivative $\frac{\partial^{\hat{r}}}{\partial z_{\hat{\alpha}_{l,1}}\cdots\partial z_{\hat{\alpha}_{l,\hat{r}}}}$. Let

$$\mathcal{X}_t^{i,(\epsilon),n} := \frac{1}{n!}\frac{\partial^n}{\partial\epsilon^n}X_t^{i,(\epsilon)}. \tag{6}$$

Then,

$$\mathcal{X}_t^{i,(\epsilon),n} = \sum_{n^{(r)},\alpha^{(r)}}^{(n)} \int_0^t \left(\prod_{l=1}^r \mathcal{X}_u^{\alpha_l,(\epsilon),n_l}\right)\partial_{\alpha^{(r)}}^r \mu_i(u, X_u^{(\epsilon)})du$$

6

$$
+ \sum_{n^{(r)},\alpha^{(r)}}^{(n-1)} \int_0^t \left(\prod_{l=1}^r \mathcal{X}_u^{\alpha_l,(\epsilon),n_l} \right) \sum_{j=1}^m \partial_{\alpha^{(r)}}^r \Phi_{i,j}(u, X_u^{(\epsilon)}) dW_u^j
$$

$$
+ \epsilon \sum_{n^{(r)},\alpha^{(r)}}^{(n)} \int_0^t \left(\prod_{l=1}^r \mathcal{X}_u^{\alpha_l,(\epsilon),n_l} \right) \sum_{j=1}^m \partial_{\alpha^{(r)}}^r \Phi_{i,j}(u, X_u^{(\epsilon)}) dW_u^j
$$

$$
+ \sum_{l'=1}^{\mathcal{N}} \sum_{r'=0}^n \sum_{\hat{n}^{(\hat{r})},\hat{\alpha}_{l'}^{(\hat{r})}}^{(n-r')} \sum_{n^{(r)},\alpha^{(r)}}^{(r')} \int_0^t \int_{E_{l'}} \left(\prod_{l=1}^r \mathcal{X}_u^{\alpha_l,(\epsilon),n_l} \right)
$$

$$
\times \partial_{\hat{\alpha}_{l'}^{(\hat{r})}}^{\hat{r}} \gamma_{l',i}(u, \epsilon z, X_u^{(\epsilon)}) \left(\prod_{\hat{l}=1}^{\hat{r}} (z_{\hat{\alpha}_{l',\hat{l}}})^{\hat{n}_{\hat{l}}} \right) \tilde{N}_{l'}(du, dz), \tag{7}
$$

where $i = 1, \cdots, D$, $z_{\hat{\alpha}_{l',\hat{l}}}$ stands for the $\hat{\alpha}_{l',\hat{l}}$-th element of $z_{l'} = (z_{l',1}, \cdots, z_{l',e_{l'}})$, and

$$
\sum_{n^{(r)},\alpha^{(r)}}^{(n)} := \sum_{r=1}^n \sum_{n_1+\cdots+n_r=n, n_l \geq 1} \sum_{\alpha^{(r)} \in \{1,\cdots,D\}^r} \frac{1}{r!}, \tag{8}
$$

$$
\sum_{\hat{n}^{(\hat{r})},\hat{\alpha}_{l'}^{(\hat{r})}}^{(\hat{n})} := \sum_{\hat{r}=1}^{\hat{n}} \sum_{\hat{n}_1+\cdots+\hat{n}_{\hat{r}}=\hat{n}, \hat{n}_{l'} \geq 1} \sum_{\hat{\alpha}_{l'}^{(\hat{r})} \in \{1,\cdots,e_{l'}\}^{\hat{r}}} \frac{1}{\hat{r}!}. \tag{9}
$$

Further, let us define a notation when $\epsilon = 0$ as $\mathcal{X}_t^{(r)} := \frac{1}{r!} \frac{\partial^r X_t^{(\epsilon)}}{\partial \epsilon^r}|_{\epsilon=0}$, and $\mathcal{X}_t^{j,(r)}$, $j = 1, \cdots, D$ denote the j-th elements of $\mathcal{X}_t^{(r)}$. For $r \geq 1$, $\mathcal{X}_t^{j,(r)}$, $j = 1, \cdots, D$ is recursively determined.

Example 2.1. *For an example to introduce the perturbation parameter into a model, we show the case of Merton jump diffusion model which is expressed as*

$$
S_t = s_0 + \int_0^t \mu S_{s-} ds + \int_0^t \sigma S_{s-} dW_s + \sum_{j=1}^M \left(e^{Y_j} - 1 \right) S_{\tau_j -} - \int_0^t \Lambda S_{s-} \mathbf{E}[e^{Y_1} - 1] ds, \tag{10}
$$

where μ and σ are constants, N_t is a Poisson process with a constant intensity Λ independent of W, τ_j is the j-th jump time of N_t, and Y_j $(j = 1, 2, \cdots)$ follow an i.i.d. normal distribution $N(m, v)$.

Then, the perturbation parameter ϵ is introduced as follows:

$$
S_t^{(\epsilon)} = s_0 + \int_0^t \mu S_{s-}^{(\epsilon)} ds + \epsilon \int_0^t \sigma S_{s-}^{(\epsilon)} dW_s + \sum_{j=1}^{N_t} \left(h^{(\epsilon)_j} \right) S_{\tau_j -}^{(\epsilon)} - \int_0^t \Lambda S_{s-}^{(\epsilon)} \mathbf{E}[h_1^{(\epsilon)}] ds, \tag{11}
$$

$$
h_j^{(\epsilon)} = e^{\epsilon Y_j} - 1. \tag{12}
$$

Here, we use the same notation of the coefficients with those of the original SDE.

In the case of $\epsilon = 0$, the jump term $h_j^{(0)} = e^{\epsilon Y_j} - 1|_{\epsilon=0} = 0$, thus, we obtain $S_t^{(0)} = s_0 + \int_0^t \mu S_s^{(0)} ds$, and its solution is expressed as $S^{(0)} = s_0 e^{\mu t}$. In a similar way, we obtain

$$
\begin{aligned}
S_T^{(1)} &= \int_0^T e^{\mu(T-t)} \sigma S_{t-}^{(0)} dW_t + \sum_{j=1}^{N_T} h_j^{(1)} e^{\mu(T-\tau_j)} S_{\tau_j-}^{(0)} - \int_0^T \Lambda \mathbf{E}\left[h_1^{(1)}\right] e^{\mu(T-t)} S_{t-}^{(0)} dt \\
&= e^{\mu T} \sigma s_0 W_T + \sum_{j=1}^{N_T} h_j^{(1)} e^{\mu T} s_0 - \Lambda \mathbf{E}\left[h_1^{(1)}\right] e^{\mu T} s_0 T, \qquad (13)
\end{aligned}
$$

$$
h_j^{(1)} = Y_j, \qquad (14)
$$

$$
\begin{aligned}
S_T^{(2)} = 2\Bigg(&\int_0^T e^{\mu(T-t)} \sigma S_{t-}^{(1)} dW_t + \sum_{j=1}^{N_T} h_j^{(2)} e^{\mu T} s_0 - \Lambda \mathbf{E}\left[h_1^{(2)}\right] e^{\mu T} s_0 T \\
&+ 2\sum_{j=1}^{N_T} h_j^{(1)} e^{\mu(T-\tau_j)} S_{\tau_j-}^{(1)} - 2\int_0^T \Lambda \mathbf{E}\left[h_1^{(1)}\right] e^{\mu(T-t)} S_{t-}^{(1)} dt \Bigg), \qquad (15)
\end{aligned}
$$

$$
h_j^{(2)} = Y_j^2. \qquad (16)
$$

3 Asymptotic Expansion for Non-smooth Function of the Solution to SDE with Jumps

This section develops an asymptotic expansion method for non-smooth functions of the solutions to SDEs with Jumps. Firstly, let us introduce a smooth function $g : \mathbf{R}^D \to \mathbf{R}^{\mathcal{M}}$ $(D, \mathcal{M} \in \mathbf{N})$ of which all derivatives have polynomial growth orders. Then, we can apply Proposition C.4 (or the subsequent related results) in Appendix C to obtain the L^p-boundedness of $g(X_T^{(\epsilon)})$ and $g_{n,T}^{(\epsilon)} := \frac{1}{n!} \frac{\partial^n g(X_T^{(\epsilon)})}{\partial \epsilon^n}$ for all $p \in [1, \infty)$, where

$$
g_{n,T}^{(\epsilon)} = \sum_{n^{(r)}, \alpha^{(r)}}^{(n)} \partial_{\alpha^{(r)}}^r g(X_T^{(\epsilon)}) \mathcal{X}_T^{\alpha_1, (\epsilon), n_1} \cdots \mathcal{X}_T^{\alpha_r, (\epsilon), n_r}, \qquad (17)
$$

with $\partial_j g(x) = \frac{\partial}{\partial x_j} g(x)$, $j = 1, \cdots, D$. Hence, $g(X_T^{(\epsilon)})$ has a smooth asymptotic expansion: for an arbitrary $M \in \mathbf{Z}_+$,

$$
g(X_T^{(\epsilon)}) = g_{0,T} + \epsilon g_{1,T} + \epsilon^2 g_{2,T} + \cdots + \epsilon^M g_{M,T} + O(\epsilon^{M+1} C(T)), \qquad (18)
$$

where $g_{n,T} := \frac{1}{n!} \frac{\partial^n g(X_T^{(\epsilon)})}{\partial \epsilon^n}\big|_{\epsilon=0}$, and $C(T)$ is an increasing function of T.

Here, $g_{n,T}$ can be written as

$$
g_{0,T} = g(X_T^{(0)}), \qquad (19)
$$

$$
g_{1,T} = \sum_{j=1}^{D} \partial_j g(X_T^{(0)}) \mathcal{X}_T^{j,(1)}, \qquad (20)
$$

8

$$g_{n,T} = \sum_{n^{(r)}, \alpha^{(r)}}^{(n)} \partial_{\alpha^{(r)}}^r g(X_T^{(0)}) \mathcal{X}_T^{\alpha_1,(n_1)} \cdots \mathcal{X}_T^{\alpha_r,(n_r)}. \tag{21}$$

Next, we put the next condition.

Condition (B)

For $\forall l$,

$$\sup_{t,x,z} \left| \left(I + \frac{\partial}{\partial x} \gamma_l(t, \epsilon z, x) \right)^{-1} \right| < \infty, \tag{22}$$

$$\Phi(t,x) \neq 0 \text{ for all } (t,x) \in \mathbf{R}_+ \times \mathbf{R}. \tag{23}$$

Under Condition (A) and (B), there exist unique processes $V_t^{(\epsilon)}$ and $(V_t^{(\epsilon)})^{-1}$ such that

$$\sup_{0 \leq t \leq T} |V_t^{(\epsilon)}|, \quad \sup_{0 \leq t \leq T} |(V_t^{(\epsilon)})^{-1}| \in L^p, \tag{24}$$

for $\forall t \geq 0$, $1 < p < \infty$. $V_t^{(\epsilon)}$ and $(V_t^{(\epsilon)})^{-1}$ satisfy the following stochastic differential equations.

$$
\begin{aligned}
V_t^{(\epsilon)} &= I + \int_0^t \frac{\partial}{\partial x} \mu(u, X_{u-}^{(\epsilon)}) V_{u-}^{(\epsilon)} du + \epsilon \int_0^t \frac{\partial}{\partial x} \Phi(u, X_{u-}^{(\epsilon)}) V_{u-}^{(\epsilon)} dW_u \\
&\quad + \sum_{l=1}^{\mathcal{N}} \int_0^t \int_{E_l} \frac{\partial}{\partial x} \gamma_l(u, \epsilon z, X_{u-}^{(\epsilon)}) V_{u-}^{(\epsilon)} \tilde{N}_l(du, dz),
\end{aligned}
\tag{25}
$$

$$
\begin{aligned}
(V_t^{(\epsilon)})^{-1} &= I - \int_0^t (V_{u-}^{(\epsilon)})^{-1} \left\{ \frac{\partial}{\partial x} \mu(u, X_{u-}^{(\epsilon)}) - \left(\epsilon \frac{\partial}{\partial x} \Phi(u, X_{u-}^{(\epsilon)}) \right)^2 \right. \\
&\quad \left. - \sum_{l=1}^{\mathcal{N}} \int_{E_l} \left(I + \frac{\partial}{\partial x} \gamma_l(u, \epsilon z, X_{u-}^{(\epsilon)}) \right)^{-1} \left(\frac{\partial}{\partial x} \gamma_l(u, \epsilon z, X_{u-}^{(\epsilon)}) \right)^2 \nu(dz) \right\} du \\
&\quad - \sum_{l=1}^{\mathcal{N}} \int_0^t \int_{E_l} (V_{u-}^{(\epsilon)})^{-1} \left(I + \frac{\partial}{\partial x} \gamma_l(u, \epsilon z, X_{u-}^{(\epsilon)}) \right)^{-1} \frac{\partial}{\partial x} \gamma_l(u, \epsilon z, X_{u-}^{(\epsilon)}) \tilde{N}_l(du, dz) \\
&\quad - \int_0^t (V_{u-}^{(\epsilon)})^{-1} \epsilon \frac{\partial}{\partial x} \Phi(u, X_{u-}^{(\epsilon)}) dW_u.
\end{aligned}
\tag{26}
$$

Then, the Malliavin derivative in the Wiener direction $D_s X_{t-}^{(\epsilon)}$ satisfies the following equation (see e.g. Cass [10] Theorem 2, Delong [12] Theorem 4.1.2).

$$D_s X_{t-}^{(\epsilon)} = V_t^{(\epsilon)} (V_{s-}^{(\epsilon)})^{-1} \epsilon \Phi(s, X_{s-}^{(\epsilon)}, \epsilon) 1_{\{s \leq t\}}. \tag{27}$$

Let us next define $\mathscr{U}_T^{(\epsilon)}$ as

$$\mathscr{U}_T^{(\epsilon)} := \frac{1}{\epsilon} \left(g(X_T^{(\epsilon)}) - g(X_T^{(0)}) \right). \tag{28}$$

9

Then, due to the asymptotic expansion of $g(X_T^{(\epsilon)})$, $\mathscr{U}_T^{(\epsilon)}$ has a smooth expansion: for an arbitrary $M \in \mathbf{Z}_+$,

$$\mathscr{U}_T^{(\epsilon)} = \mathscr{U}_T^{(0)} + \epsilon \mathscr{U}_T^{(1)} + \epsilon^2 \mathscr{U}_T^{(2)} + \cdots + \epsilon^M \mathscr{U}_T^{(M)} + O(\epsilon^{M+1} C(T)), \tag{29}$$

with $\mathscr{U}_T^{(n)} := g_{n+1,T}$.

In this setting, the Malliavin covariance $\mathscr{U}_T^{(\epsilon)}$ is defined as

$$\sigma_{\mathscr{U}_T^{(\epsilon)}} := \int_0^T \partial g(X_T^{(\epsilon)}) V_T^{(\epsilon)} (V_{u-}^{(\epsilon)})^{-1} \Phi(u, X_{u-}^{(\epsilon)}) \otimes \partial g(X_T^{(\epsilon)}) V_T^{(\epsilon)} (V_{u-}^{(\epsilon)})^{-1} \Phi(u, X_{u-}^{(\epsilon)}) du. \tag{30}$$

Remark 3.1. *When* $\det(\sigma_{\mathscr{U}_T^{(\epsilon)}})^{-1} \in L^p$ *(for* $\forall p \geq 2$*) is satisfied,* $\mathscr{U}_T^{(\epsilon)}$ *has a smooth density (see Di Nunno et al. [13] Theorem 18.5).*

In order to justify the asymptotic expansion of $f(\mathscr{U}_T^{(\epsilon)})$ with a Schwartz distribution ($f \in \mathcal{S}'$) around $\epsilon = 0$, we put the following condition.
Condition (C) For $\forall p > 1$,

$$\limsup_{\epsilon \downarrow 0} \det(\sigma_{\mathscr{U}_T^{(\epsilon)}})^{-1} \in L^p. \tag{31}$$

From Proposition A.1, A.2 and C.4, we obtain the following theorem.

Theorem 3.2. $X_t^{(\epsilon)}$ *is a perturbed process defined as*

$$X_t^{(\epsilon)} = x + \int_0^t \mu(s, X_{s-}^{(\epsilon)}) ds + \epsilon \int_0^t \Phi(s, X_{s-}^{(\epsilon)}) dW_s + \sum_{l=1}^k \int_0^t \int_{E_l} \gamma_l(s, \epsilon z, X_{s-}^{(\epsilon)}) \tilde{N}_l(ds, dz), \tag{32}$$

where $x \in \mathbf{R}^D$; $\mu : [0,T] \times \mathbf{R}^D \times \Omega \to \mathbf{R}^D$, $\Phi : [0,T] \times \mathbf{R}^D \times \Omega \to \mathbf{R}^D \times \mathbf{R}^m$, $\gamma_l : [0,T] \times E_l \times \mathbf{R}^D \times \Omega \to \mathbf{R}^D$ *are predictable processes satisfying Condition (A); W is an m-dimensional Brownian motion. Each* $N_l(dt, dz)$ *is a Poisson random measure on* $[0,T] \times E_l$*, where* $(E_l, \underline{E_l})$ *is a measurable space, which is defined in* $\mathbf{R}^{e_l}$*,* $e_l \in \mathbf{N}$*; The intensity measure of* N_l *is* $dt \times \nu_l(z)$*, where* $\nu_l(z)$ *is a positive* σ*-finite measure on* $(E_l, \underline{E_l})$*;* $\tilde{N}_l(dz, dt) := N_l(dt, dz) - dt \times \nu_l(z)$ *is a compensated Poisson random measure. We also define* $\mathscr{U}_T^{(\epsilon)} := \frac{1}{\epsilon} \left(g(X_T^{(\epsilon)}) - g(X_T^{(0)}) \right)$ *on* $[0,T]$*. Assume that g is a* $\mathbf{R}^{\mathcal{M}}$*-valued smooth, of which all derivatives have polynomial growth orders function on X. Suppose also f satisfies* $f \in \mathcal{S}'(\mathbf{R}^{\mathcal{M}})$*, Condition (B) and (C).*

Then, we obtain an approximation of $\mathbf{E}\left[f(\mathscr{U}_T^{(\epsilon)})\right]$ *as*

$$\mathbf{E}\left[f(\mathscr{U}_T^{(\epsilon)})\right] = \mathbf{E}\left[f_{0,T} + \epsilon f_{1,T} + \epsilon^2 f_{2,T} + \cdots + \epsilon^M f_{M,T}\right] + O(\epsilon^{M+1} C(T)), \tag{33}$$

where $C(T)$ *is an increasing function of T, and for* $n \in \mathbf{N}$*,*

$$\mathbf{E}[f_{0,T}] = \int_{\mathbf{R}} f(x) p^{\mathscr{U}_T^{(0)}}(x) dx, \tag{34}$$

$$\mathbf{E}[f_{n,T}] = \sum_{n^{(r)}}^{(n)} \int_{\mathbf{R}} f(x) (-\partial)^r \left(\mathbf{E}\left[\mathscr{U}_T^{(n_1)} \cdots \mathscr{U}_T^{(n_r)} | \mathscr{U}_T^{(0)} = x\right] p^{\mathscr{U}_T^{(0)}}(x) \right) dx, \tag{35}$$

10

$$\sum_{n^{(r)}}^{(n)} := \sum_{r=1}^{n} \sum_{n_1+\cdots+n_r=n,n_l\geq 1} \frac{1}{r!}. \tag{36}$$

The proof of this theorem is given in Appendix B.

Next, let us discuss a particular case that the jump term is a compound Poisson process, whence Condition (C) and Theorem 3.2 are expressed more specifically. We first consider a process $\hat{X}$ with fixed jump times and sizes.

$$
\begin{aligned}
\hat{X}_{s,T}^{\hat{x},\mathcal{Z},(\epsilon)} &= \hat{x} + \int_s^T \left\{ \mu(t,\hat{X}_{s,t-}^{\hat{x},\mathcal{Z},(\epsilon)}) - \sum_{l=1}^{\mathcal{N}} \Lambda_l \mathbf{E}[h_l(\epsilon z)]\hat{\gamma}_l(t,\hat{X}_{s,t-}^{\hat{x},\mathcal{Z},(\epsilon)}) \right\} dt \\
&\quad + \epsilon \int_s^T \Phi(t,\hat{X}_{s,t-}^{\hat{x},\mathcal{Z},(\epsilon)},\epsilon)dW_t + \sum_{l=1}^{\mathcal{N}} \sum_{j_l=k_l(s)+1}^{k_l(T)} h_l(\epsilon\mathcal{Z}_{l,j_l})\hat{\gamma}_l(\tau_{j_l},\hat{X}_{\tau_{j_l},\tau_{j_l}-}^{\hat{x},\mathcal{Z},(\epsilon)}) \\
&= \hat{x} + \int_s^T \left\{ \mu(t,\hat{X}_{s,t-}^{\hat{x},\mathcal{Z},(\epsilon)}) - \sum_{l=1}^{\mathcal{N}} \Lambda_l \mathbf{E}[h_l(\epsilon z)]\hat{\gamma}_l(t,\hat{X}_{s,t-}^{\hat{x},\mathcal{Z},(\epsilon)}) \right\} dt \\
&\quad + \epsilon \int_s^T \Phi(t,\hat{X}_{s,t-}^{\hat{x},\mathcal{Z},(\epsilon)},\epsilon)dW_t + \sum_{l=1}^{\mathcal{N}} \int_0^T \hat{\gamma}_l(t-,\hat{X}_{s,t-}^{\hat{x},\mathcal{Z},(\epsilon)})d\Upsilon_{l,t}, \tag{37}
\end{aligned}
$$

where $k_l(t)$ is jump times of the l-th jump term until time t, Λ_l is an intensity of the l-th jump, $\mathcal{Z}_{l,j_l}$ is a jump size, Υ_l is a piecewise constant finite many jump cadlag trajectory $\Upsilon_l : [0,T] \to \mathbf{R}^{\mathcal{N}}$ with $\Upsilon_{l,0} = 0$, $\hat{\gamma}_l : [0,T] \times \mathbf{R}^D \times \Omega \to \mathbf{R}^D$ and $h_l : E_l \to \mathbf{R}^D \times \mathbf{R}^D$. Hereafter, we denote $k_l = k_l(T)$ when a maturity of the option is T.

Next, let us arrange the jump times as $t_1 < t_2 < \cdots < t_{\mathcal{K}(T)}$, and the jump sizes are also arranged as $\mathcal{Z}_{l,1},\cdots,\mathcal{Z}_{l,\mathcal{K}(T)}$, where $\mathcal{K}(t) = \sum_{l=1}^{\mathcal{N}} k_l(t) - \#\{\tau_{l,j} = \tau_{l',i}, (l \neq l')\}$ and $h_l(\epsilon\mathcal{Z}_{l,j_l}) = 0$ if $\tau_{l,i} \neq t_j$ for $1 \leq \forall i \leq k_l(t)$.

Then, as in Forster et al. [15], $(\hat{X}_{s,T}^{\hat{x},\mathcal{Z},(\epsilon)})_{s\geq 0}$ can be given explicitly in terms of the jump times τ_{j_l} for $j_l \geq 0$ and the diffusion process between two consecutive jumps. We define $\hat{X}_t^{x,\mathcal{Z},(\epsilon)}$ as

$$
\begin{aligned}
\hat{X}_t^{x,\mathcal{Z},(\epsilon)} &:= \hat{X}_{0,t}^{\hat{x},\mathcal{Z},(\epsilon)} \text{ for } 0 \leq t < t_1, \\
\hat{X}_t^{x,\mathcal{Z},(\epsilon)} &:= \hat{X}_{t_1,t}^{\hat{x},\mathcal{Z},(\epsilon)}\Big|_{\hat{x}=\hat{X}_{0,t_1-}^{\hat{x},\mathcal{Z},(\epsilon)}+\sum_{l=1}^{\mathcal{N}} h_l(\epsilon\mathcal{Z}_{l,1})\hat{\gamma}_l(t_1-,\hat{X}_{0,t_1-}^{\hat{x},\mathcal{Z},(\epsilon)})} \text{ for } t_1 \leq t < t_2, \\
&\quad\vdots \\
\hat{X}_t^{x,\mathcal{Z},(\epsilon)} &:= \hat{X}_{t_{\mathcal{K}(T)},t}^{\hat{x},\mathcal{Z},(\epsilon)}\Big|_{\hat{x}=\hat{X}_{0,t_{\mathcal{K}(T)}-}^{\hat{x},\mathcal{Z},(\epsilon)}+\sum_{l=1}^{\mathcal{N}} h_l(\epsilon\mathcal{Z}_{l,n})\hat{\gamma}_l(t_{\mathcal{K}(T)}-,\hat{X}_{0,t_{\mathcal{K}(T)}-}^{\hat{x},\mathcal{Z},(\epsilon)})} \text{ for } t_{\mathcal{K}(T)} \leq t \leq T, \tag{38}
\end{aligned}
$$

Next, we define the process $\hat{X}^{x,\hat{\mathcal{Z}},(\epsilon)}$ by inserting Poisson jump times for τ_j, and i.i.d random variables $\hat{\mathcal{Z}}_{l,j}$ for $\mathcal{Z}_{l,j}$ for each l. Then, the process $\hat{X}^{x,\hat{\mathcal{Z}}}$ is indistinguishable from X which is the solution of (3) (see e.g. Forster et al. [15] Theorem 1).

11

Let us define the first variation process for $\hat{X}^{x,\mathcal{Z},(\epsilon)}$ as $\hat{V}^{x,\mathcal{Z},(\epsilon)}$ which satisfies the following stochastic differential equation.

$$
\hat{V}_t^{\hat{x},\mathcal{Z},(\epsilon)} = I + \int_0^t \left\{ \frac{\partial}{\partial x}\mu(u-,\hat{X}_{0,u-}^{\hat{x},\mathcal{Z},(\epsilon)}) - \sum_{l=1}^{\mathcal{N}} \Lambda_l \mathbf{E}[h_l(\epsilon z)]\frac{\partial}{\partial x}\hat{\gamma}_l(u,\hat{X}_{0,u-}^{\hat{x},\mathcal{Z},(\epsilon)}) \right\} \hat{V}_{u-}^{\hat{x},\mathcal{Z},(\epsilon)} du
$$

$$
+ \int_0^t \epsilon \frac{\partial}{\partial x}\Phi(u,\hat{X}_{0,u}^{\hat{x},\mathcal{Z},(\epsilon)})\hat{V}_{u-}^{\hat{x},\mathcal{Z},(\epsilon)}dW_u + \sum_{l=1}^{\mathcal{N}}\sum_{j=1}^{\mathcal{K}(t)} \frac{\partial}{\partial x}h_l(\epsilon\mathcal{Z}_{l,j})\hat{\gamma}_l(u_j,\hat{X}_{0,u_j-}^{\hat{x},\mathcal{Z},(\epsilon)})\hat{V}_{u-}^{\hat{x},\mathcal{Z},(\epsilon)}.
$$

$$(39)$$

Then, the Malliavin derivative in the Wiener direction $D_s\hat{X}_t^{x,\mathcal{Z},(\epsilon)}$ satisfies the following equation (Poissonian trajectory-wise).

$$
D_s\hat{X}_t^{x,\mathcal{Z},(\epsilon)} = \hat{V}_t^{x,\mathcal{Z},(\epsilon)}(\hat{V}_{s-}^{x,\mathcal{Z},(\epsilon)})^{-1}\epsilon\Phi(s,\hat{X}_{s-}^{x,\mathcal{Z},(\epsilon)},\epsilon)1_{\{s\leq t\}}, \tag{40}
$$

where

$$
(\hat{V}_t^{\hat{x},\mathcal{Z},(\epsilon)})^{-1} = I - \int_0^t (\hat{V}_{u-}^{\hat{x},\mathcal{Z},(\epsilon)})^{-1}\left\{ \frac{\partial}{\partial x}\mu(u-,\hat{X}_{0,u-}^{\hat{x},\mathcal{Z},(\epsilon)}) - \sum_{l=1}^{\mathcal{N}} \Lambda \mathbf{E}[h_l(\epsilon z)]\frac{\partial}{\partial x}\hat{\gamma}_l(u,\epsilon z,\hat{X}_{0,u-}^{\hat{x},\mathcal{Z},(\epsilon)}) \right.
$$

$$
\left. - \left(\epsilon\frac{\partial}{\partial x}\Phi(u,\hat{X}_{0,u}^{\hat{x},\mathcal{Z},(\epsilon)}) \right)^2 \right\}du - \sum_{l=1}^{\mathcal{N}}\sum_{j=1}^{\mathcal{K}(t)}(\hat{V}_{u_j-}^{\hat{x},\mathcal{Z},(\epsilon)})^{-1}h_l(\epsilon\mathcal{Z}_{l,j})\frac{\partial}{\partial x}\hat{\gamma}_l(u_j,\hat{X}_{0,u_j-}^{\hat{x},\mathcal{Z},(\epsilon)})
$$

$$
+ \sum_{l=1}^{\mathcal{N}}\sum_{j=1}^{\mathcal{K}(t)}(\hat{V}_{u-}^{\hat{x},\mathcal{Z},(\epsilon)})^{-1}\left(I + h_l(\epsilon\mathcal{Z}_{l,j})\frac{\partial}{\partial x}\hat{\gamma}_l(u_j,\hat{X}_{0,u_j-}^{\hat{x},\mathcal{Z},(\epsilon)}) \right)^{-1}
$$

$$
\times \left(h_l(\epsilon\mathcal{Z}_{l,j})\frac{\partial}{\partial x}\hat{\gamma}_l(u_j,\hat{X}_{0,u_j-}^{\hat{x},\mathcal{Z},(\epsilon)}) \right)^2 - \int_0^t (\hat{V}_{u-}^{\hat{x},\mathcal{Z},(\epsilon)})^{-1}\epsilon\frac{\partial}{\partial x}\Phi(u,\hat{X}_{0,u}^{\hat{x},\mathcal{Z},(\epsilon)})dW_u. \tag{41}
$$

We also define $\hat{\mathscr{U}}_T^{(\epsilon)}$ as

$$
\hat{\mathscr{U}}_T^{(\epsilon)} := \frac{1}{\epsilon}\left(g(\hat{X}_T^{(\epsilon)}) - g(\hat{X}_T^{(0)}) \right), \tag{42}
$$

In this setting, $\hat{\mathscr{U}}$ has an asymptotic expansion, and the Malliavin covariance is defined as

$$
\sigma_{\hat{\mathscr{U}}_T^{(\epsilon)}} = \sum_{j=0}^{\mathcal{K}(T)}\int_{t_j}^{t_{j+1}} \partial g(\hat{X}_T^{(\epsilon)})\hat{V}_T(\hat{V}_s^{x,\mathcal{Z},(\epsilon)})^{-1}\Phi(s,\hat{X}_{s-}^{(\epsilon)}) \otimes \partial g(\hat{X}_T^{(\epsilon)})\hat{V}_T(\hat{V}_s^{x,\mathcal{Z},(\epsilon)})^{-1}\Phi(s,\hat{X}_{s-}^{(\epsilon)})ds, \tag{43}
$$

where $t_{\mathcal{K}(T)+1} = T$.

In order to justify the asymptotic expansion of $f(\hat{\mathscr{U}}_T^{(\epsilon)})$ around $\epsilon = 0$, we put the following condition.

Condition (C') For $\forall p > 1$,

$$
\limsup_{\epsilon\downarrow 0}\left|\sigma_{\hat{\mathscr{U}}_T^{(\epsilon)}}\right|^{-1} \in L^p. \tag{44}
$$

Then, we obtain a special case of Theorem 3.2 as the next corollary.

Corollary 3.3. *Suppose that the $\hat{\mathscr{U}}^{(\epsilon)}$ ($\epsilon \in (0,1]$) satisfies Condition (A), (B) and (C'). Moreover, the jump term is assumed as a compound Poisson process.*

Then, for every $f \in \mathcal{S}'(\mathbf{R}^{\mathcal{M}})$, $\mathbf{E}[f(\hat{\mathscr{U}}_T^{(\epsilon)})]$ has an asymptotic expansion:

$$\mathbf{E}[f(\hat{\mathscr{U}}_T^{(\epsilon)})] = \mathbf{E}\left[f_{0,T} + \epsilon f_{1,T} + \epsilon^2 f_{2,T} + \cdots + \epsilon^M f_{M,T}\right] + O(\epsilon^{M+1}C(T)), \tag{45}$$

where $C(T)$ is an increasing function of T.

$\mathbf{E}[f_{n,T}]$ *are calculated as follows:*

$$\mathbf{E}[f_{0,T}] = \sum_{k=0}^{\infty} \sum_{\sum_{l=1}^{\mathcal{N}} k_l = k} p_{\{k_l\}} \left(\prod_{l=1}^n \frac{k_l!}{T^{k_l}}\right) \int_0^{\vec{T}} \int_{\mathbf{R}} f(x) p^{\hat{\mathscr{U}}_{\{\tau_l = \vec{t_l}\},T}^{(0)}}(x) dx d\vec{t}, \tag{46}$$

$$\mathbf{E}[f_{n,T}] = \sum_{k=0}^{\infty} \sum_{\sum_{l=1}^{\mathcal{N}} k_l = k} p_{\{k_l\}} \left(\prod_{l=1}^n \frac{k_l!}{T^{k_l}}\right) \sum_{n^{(r)},\alpha^{(r)}}^{(n)} \int_0^{\vec{T}} \int_{\mathbf{R}} f(x)(-\partial_{\alpha^{(r)}})^n$$
$$\left(\mathbf{E}\left[\mathcal{X}_T^{\alpha_1,(n_1)} \cdots \mathcal{X}_T^{\alpha_r,(n_r)} | \hat{\mathscr{U}}_{\{\tau_l = \vec{t_l}\},T}^{(0)} = x, \{N_{l,T} = k_l\}, \{\tau_l = \vec{t_l}\}\right] p^{\hat{\mathscr{U}}_{\{\tau_l = \vec{t_l}\},T}^{(0)}}(x)\right) dx d\vec{t}, \tag{47}$$

where $\{N_l = k_l\} = \{N_{1,T} = k_1, \cdots, N_{n,T} = k_n\}$, $\{\tau_l = \vec{t_l}\} = \{\tau_{1,1} = t_{1,1}, \cdots, \tau_{1,k_1} = t_{1,k_1}, \cdots, \tau_{n,1} = t_{n,1}, \cdots, \tau_{n,k_n} = t_{n,k_n}\}$, $p_{\{k_l\}} = \prod_{l=1}^n \frac{(\Lambda_l T)^{k_l} e^{-\Lambda_l T}}{k_l!}$, and $\tau_{l,j}$ is the j-th jump time of the jump process of the l-th type. $p^{\hat{\mathscr{U}}_{\{\tau_l = \vec{t_l}\},T}^{(0)}}$ is a density function of $\hat{\mathscr{U}}_{\{\tau_l = \vec{t_l}\},T}^{(0)}$, and $\hat{\mathscr{U}}_{\{\tau_l = \vec{t_l}\},T}^{(0)}$ stands for $\hat{\mathscr{U}}_T^{(0)}$ conditioned on jumps at $\{\tau_l = \vec{t_l}\}$ with $\{N_{l,T} = k_l\}$. Moreover,

$$\int_0^{\vec{T}} = \prod_{l=1}^n \int_0^T \int_0^{t_{l,k_l}} \cdots \int_0^{t_{l,2}}; \quad d\vec{t} = \prod_{l=1}^n dt_{l,1} \cdots dt_{l,k_l}. \tag{48}$$

The conditional expectations in (47) are calculated based on the formulas in Section B of the online appendix [31], as well as formulas in Appendix B of Shiraya - Takahashi [34].

The proof of this corollary is provided in Appendix B.

Example 3.4. *In the case of one-dimensional Merton jump diffusion model defined in Example 2.1 with $g(x) = x$, $\hat{\mathscr{U}}_T^{(\epsilon)}$ is expanded as follows:*

$$\hat{\mathscr{U}}_T^{(\epsilon)} = S_T^{(1)} + \frac{\epsilon}{2!} S_T^{(2)} + \frac{\epsilon^2}{3!} S_T^{(3)} + \cdots, \tag{49}$$

where $S_T^{(1)}$, $S_T^{(2)}$ are given in Example 2.1, and $S_T^{(3)}$ is obtained similarly. Then, for a path-independent option with the payoff $f(\hat{\mathscr{U}}_T^{(\epsilon)})$, $\mathbf{E}[f_{0,T}]$, $\mathbf{E}[f_{1,T}]$ and $\mathbf{E}[f_{2,T}]$ are expressed as follows:

$$\mathbf{E}[f_{0,T}] = \sum_{k=0}^{\infty} \frac{(\Lambda T)^k e^{-\Lambda T}}{k!} \int_{\mathbf{R}} f(x) n(x; \xi, \Sigma_T^{\{k\}}) dx, \tag{50}$$

$$\mathbf{E}[f_{1,T}] = -\sum_{k=0}^{\infty} \frac{(\Lambda T)^k e^{-\Lambda T}}{k!} \int_{\mathbf{R}} f(x) \partial_x \left(\mathbf{E}\left[S_T^{(2)} | S_T^{(1)} = x, \{N_T = k\}\right] n(x; \xi, \Sigma_T^{\{k\}})\right) dx \tag{51}$$

13

$$
\begin{aligned}
\mathbf{E}[f_{2,T}] \;=\; & \sum_{k=0}^{\infty} \frac{(\Lambda T)^k e^{-\Lambda T}}{k!} \int_{\mathbf{R}} f(x) \Bigg\{ -\partial_x \left(\mathbf{E}\left[S_T^{(3)} | S_T^{(1)} = x, \{N_T = k\} \right] n(x; \xi, \Sigma_T^{\{k\}}) \right) \\
& + \partial_x^2 \left(\mathbf{E}\left[\frac{1}{2} \left(S_T^{(2)} \right)^2 | S_T^{(1)} = x, \{N_T = k\} \right] n(x; \xi, \Sigma_T^{\{k\}}) \right) \Bigg\} dx,
\end{aligned}
\tag{52}
$$

where $\xi = (k - \Lambda T) m e^{\mu T} s_0$, and $\Sigma_T^{\{k\}} := \mathbf{Var}\left[S_T^{(1)} \right] = (e^{\mu T} \sigma s_0)^2 T + k(v e^{\mu T} s_0)^2$.

4 Approximation Formula of Average and Spread Option Prices under Stochastic Volatility with Jumps Models

This section first describes the structure of average and spread options in practice, where the underlying assets are the averages or spreads of future prices. Let us note that the other types of average options, which don't have rollovers, are contained in our formulation. Then, based on the previous section, we present an approximate pricing formula for average and spread options under local stochastic volatility with jump diffusion models, which will be used in the next section.

To explain the structure of discrete average options, let us introduce new processes A_t^i defined by

$$
A_t^i = \sum_{j=1}^{m_i} w_j^{(i)} S_{t_j^{(i)}}^i 1_{\{t_j^{(i)} \le t\}},
\tag{53}
$$

where $0 \le t_1^{(i)} < \cdots < t_{m_i}^{(i)} \le T$, m_i denotes the number of the underlying asset price S^i to which the average option refers, each $w_j^{(i)}$ stands for the weight for the price of the contract i at date $t_j^{(i)}$, and the dynamics of each S^i is described by the stochastic differential equation. Then, we can deal with an average option whose underlying asset price $g(A_t)$ is given by the following:

$$
g(A_t) = \sum_{i=1}^{d} A_t^i.
\tag{54}
$$

Next, let us define the underlying asset of a spread option as the spread of futures prices with two different maturities ($T_1 < T_2$). More precisely, the spread consists of a long position in the first expiring futures (S^1) in the spread and a short position in the second expiring futures (S^2). Then, the underlying of a spread option is expressed as:

$$
\begin{aligned}
A_t^i \;=\;& w^i S_{T_i}^i 1_{\{T_i \le t\}}, \;\; (i = 1, 2, \; w^1 = 1, \; w^2 = -1), \\
g(A_t) \;=\;& A_t^1 - A_t^2.
\end{aligned}
\tag{55}
$$

In this way, we are able to treat pricing problems for average and spread options in a unified manner, since the spread option is a special case of discrete average options. Thus, in the remaining part of this section, we concentrate on the discrete average option. Then, we define the payoff function of an average option with maturity T and strike K as $(g(A_T) - K)^+ (:= \max\{g(A_T) - K, 0\})$ for a call option and $(K - g(A_T))^+ (:= \max\{K - g(A_T), 0\})$ for a put option.

Remark 4.1. *We are able to obtain an approximation formula for pricing the continuously monitored average options by simply changing the process A and the function g as*

$$A_t = \int_0^t S_u du, \tag{56}$$

$$g(A_t) = A_t. \tag{57}$$

However, since the continuously monitored options are not traded in the real world, and the derivation is very similar as in the discretely monitored case, we omit the details.

In the following, we show the model of the underlying asset prices and its volatility processes, which is used for pricing discrete average options. In particular, suppose that the filtered probability space $(\Omega, \mathbf{F}, \mathbf{P}, \{\mathbf{F}_t\}_{t\geq 0})$ is given, where $\mathbf{P}$ is an equivalent martingale measure and the filtration satisfies the usual conditions. Then, $(S_t^i)_{t\in[0,T]}$ and $(\sigma_t^i)_{t\in[0,T]}$, $i = 1, \cdots, d$ represent the underlying asset prices and their volatilities, respectively. Particularly, let us assume that S_T^i and σ_T^i are given by the solutions to the next stochastic integral equations:

$$S_T^i = s_0^i + \int_0^T \alpha_{t-}^i S_{t-}^i dt + \int_0^T \phi_{S^i}(\sigma_{t-}, S_{t-}) dW_t$$
$$+ \sum_{l=1}^n \left(\sum_{j=1}^{N_{l,T}} h_{S^i,l,j} S_{\tau_{j,l}-}^i - \int_0^T \Lambda_l S_{t-}^i \mathbf{E}[h_{S^i,l,1}] dt \right), \tag{58}$$

$$\sigma_T^i = \sigma_0^i + \int_0^T \mu_{\sigma^i}(\sigma_{t-}, t-) dt + \int_0^T \phi_{\sigma^i}(\sigma_{t-}) dW_t$$
$$+ \sum_{l=1}^n \left(\sum_{j=1}^{N_{l,T}} h_{\sigma^i,l,j} \sigma_{\tau_{j,l}-}^i - \int_0^T \Lambda_l \sigma_{t-}^i \mathbf{E}[h_{\sigma^i,l,1}] dt \right), \tag{59}$$

where s_0^i and σ_0^i, $i = 1, \cdots, d$ are given constants. α_t^i is a deterministic function of t. $\mu_{\sigma^i}(\sigma^i, t)$, $\phi_{S^i}(x, y)$, $\phi_{\sigma^i}(x)$ are some deterministic functions with appropriate regularity conditions in Section 2. W is a $2d$-dimensional Brownian motion. Each N_l, $(l = 1, \cdots, n)$ is a Poisson process with constant intensity Λ_l, which is independent of each other. N_l, $(l = 1, \cdots, n)$ are also independent of all W; $\tau_{j,l}$ stands for the j-th jump time of N_l; For each $l = 1, \cdots, n$ and $i = 1, \cdots, d$, both $\left(\sum_{j=1}^{N_{l,t}} h_{S^i,l,j} \right)_{t\geq 0}$ and $\left(\sum_{j=1}^{N_{l,t}} h_{\sigma^i,l,j} \right)_{t\geq 0}$ are compound Poisson processes. $(\sum_{j=1}^{N_{l,t}} \equiv 0$ when $N_{l,t} = 0)$.

For each l and x^i, $h_{x^i,l,j}$ $(j \in \mathbf{N})$ are independent, identically distributed random variables, where x^i stands for one of S^i, σ^i $(i = 1, \cdots, d)$. For instance, we can consider a log-normal jump case as $h_{x^i,l,j} = e^{Y_{x^i,l,j}} - 1$, where $Y_{x^i,l,j}$ is a random variable which follows a normal distribution $N(m_{x^i,l}, v_{x^i,l}^2)$. Also, $h_{x^i,l,j}$ and $h_{x^{i'},l',j'}$ $(l \neq l')$ are independent. N_l and $h_{x^i,l',j}$ are independent, too. On the other hand, for the same l and j, $h_{S^i,l,j}$ and $h_{\sigma^{i'},l,j}$ $(i, i' = 1, \cdots, d)$ are allowed to be dependent, that is $Y_{S^i,l,j}$ and $Y_{\sigma^{i'},l,j}$ $(i, i' = 1, \cdots, d)$ are generally correlated.

Remark 4.2. *We briefly comment on the relation of the model above to the SDE (2) in Section 2. Firstly, the dimension $2d$ above corresponds to D and m (i.e. $2d = D = m$) in (2). Also, the*

jump term in (2) is specified as follows: for $X = S^i$ or σ^i,

$$\int_0^T \int_{E_l} \gamma_l(t, z, X_{t-}) \tilde{N}_l(ds, dz) = \int_0^T \int_{\mathbf{R}\backslash\{0\}} X_{t-} \tilde{\gamma}_{x,l}(z) \tilde{N}_l(ds, dz), \tag{60}$$

$$(\text{That is, } E_l = \mathbf{R} \backslash \{0\}, \ \gamma_l(t, z, X_{t-}) = X_{t-}\tilde{\gamma}_{x,l}(z)) \, . \tag{61}$$

Moreover, for the log-normal jump case,

$$\tilde{N}_l(ds, dz) = N_l(ds, dz) - \Lambda_l ds \ \mu(dz) \, , \tag{62}$$

$$N_l([0,t], A) = \sum_{j=1}^{N_{l,t}} I_A \left(e^{\xi_{x,l,j}} - 1 \right), \ A \in B(\mathbf{R} \backslash \{0\}), \ \xi_{x,l,j} \sim N(m_{x,l}, v_{x,l}^2), \ i.i.d., \tag{63}$$

where $N_{l,t}$ is a Poisson process with the intensity Λ_l.

Then, we introduce perturbations to the models (58), (59) with the same notation of the coefficients as those of the original SDEs, as in Section 2. That is, for a known parameter $\epsilon \in [0, 1]$ we consider the following stochastic integral equations: for $i = 1, \cdots, d$,

$$\begin{aligned}
S_T^{i,(\epsilon)} &= s_0^i + \int_0^T \alpha_{t-}^i S_{t-}^{i,(\epsilon)} dt + \epsilon \int_0^T \phi_{S^i} \left(\sigma_{t-}^{(\epsilon)}, S_{t-}^{(\epsilon)} \right) dW_t \\
&\quad + \sum_{l=1}^n \left(\sum_{j=1}^{N_{l,T}} h_{S^i,l,j}^{(\epsilon)} S_{\tau_{j,l}-}^{i,(\epsilon)} - \int_0^T \Lambda_l S_{t-}^{i,(\epsilon)} \mathbf{E}[h_{S^i,l,1}^{(\epsilon)}] dt \right),
\end{aligned} \tag{64}$$

$$\begin{aligned}
\sigma_T^{i,(\epsilon)} &= \sigma_0^i + \int_0^T \mu_{\sigma^i} \left(\sigma_{t-}^{(\epsilon)}, t- \right) dt + \epsilon \int_0^T \phi_{\sigma^i} \left(\sigma_{t-}^{(\epsilon)} \right) dW_t \\
&\quad + \sum_{l=1}^n \left(\sum_{j=1}^{N_{l,T}} h_{\sigma^i,l,j}^{(\epsilon)} \sigma_{\tau_{j,l}-}^{i,(\epsilon)} - \int_0^T \Lambda_l \sigma_{t-}^{i,(\epsilon)} \mathbf{E}[h_{\sigma^i,l,1}^{(\epsilon)}] dt \right),
\end{aligned} \tag{65}$$

where s_0, σ_0 are given constants, and α_t is a deterministic function of t. We assume that $\mu_{x^i}^{(\epsilon)}$, $\phi_{x^i}^{(\epsilon)}$, $h_{x^i,l,j}^{(\epsilon)}$ satisfy Condition (A) and (B). We also suppose that $h_{x^i,l,j}^{(\epsilon)} = \epsilon m_{x^i,l}$ ($m_{x^i,l}$: a constant), or $h_{x^i,l,j}^{(\epsilon)} = e^{\epsilon Y_{x^i,l,j}} - 1$ where $Y_{x^i,l,j} \sim N(m_{x^i,l}, v_{x^i,l}^2)$.

Next, by applying Corollary 3.3 with approximations for reduction of computational burdens, we derive a pricing formula of an average call option with strike K and maturity T, whose payoff function is given by $f(g(A_T^{(\epsilon)}))$ with $f(x) = (x - K)^+$. Also, for some fixed constants $w_j^{(i)}$ and reference dates $t_j^{(i)}$ ($i = 1, \cdots, d$, $j = 1, \cdots, m_i$) with the function (54), $g(A_T^{(\epsilon)})$ is specified as

$$g(A_T^{(\epsilon)}) = \sum_{i=1}^d A_T^{i,(\epsilon)}, \quad A_T^{i,(\epsilon)} = \sum_{j=1}^{m_i} w_j^{(i)} S_{T_j^{(i)}}^{i,(\epsilon)} 1_{\{t_j^{(i)} \leq T\}}. \tag{66}$$

Then, the call payoff is expanded as follows:

$$f(g(A_T^{(\epsilon)})) = \left(g(A_T^{(\epsilon)}) - K \right)^+$$

$$= \epsilon \left(\frac{g(A_T^{(\epsilon)}) - g(A_T^{(0)})}{\epsilon} + \mathscr{Y} \right)^+$$

$$= \epsilon \left(g(A_T^{(1)}) + \mathscr{Y} \right)^+ + \frac{\epsilon^2}{2} 1_{\left\{ g(A_T^{(1)}) > -\mathscr{Y} \right\}} g\left(A_T^{(2)} \right)$$

$$+ \epsilon^3 \left(\frac{1}{6} 1_{\left\{ g(A_T^{(1)}) > -\mathscr{Y} \right\}} g\left(A_T^{(3)} \right) + \frac{1}{8} \delta_{\left\{ g(A_T^{(1)}) = -\mathscr{Y} \right\}} g\left(A_T^{(2)} \right)^2 \right) + o(\epsilon^3), \quad (67)$$

where

$$g(A_T^{(n)}) = g\left(\frac{\partial^n}{\partial \epsilon^n} A_T^{(\epsilon)} \Big|_{\epsilon=0} \right) = \sum_{i=1}^{d} \sum_{j=1}^{m_i} w_j^{(i)} \frac{\partial^n}{\partial \epsilon^n} S_{t_j^{(i)}}^{i,(\epsilon)} \Big|_{\epsilon=0} 1_{\{t_j^{(i)} \leq T\}} = \sum_{i=1}^{d} \sum_{j=1}^{m_i} w_j^{(i)} S_{t_j^{(i)}}^{i,(n)} 1_{\{t_j^{(i)} \leq T\}}.$$
(68)

Hence, the calculation of $g(A_T^{(n)})$ is reduced to that of $S_t^{(n)}$. Here, the strike price is expressed as $K = g(A_T^{(0)}) - \epsilon \mathscr{Y}$ for an arbitrary $\mathscr{Y} \in \mathbf{R}$. We also note that $\frac{g(A_T^{(\epsilon)}) - g(A_T^{(0)})}{\epsilon}$ and $g(A_T^{(1)})$ with $A_T^{(1)} = \frac{\partial A_T^{(\epsilon)}}{\partial \epsilon}|_{\epsilon=0}$ in (67) corresponds to $\hat{\mathscr{U}}_T^{(\epsilon)}$ and $\hat{\mathscr{U}}_T^{(0)}$ in Corollary 3.3, respectively. We suppose that Condition (C') is satisfied for $\hat{\mathscr{U}}_T^{(\epsilon)}$.

In particular, given $\{N_l = k_l\} = \{N_{1,T} = k_1, \cdots, N_{n,T} = k_n\}$ (i.e. given the number of jumps until T), let us obtain a concrete expression of $g(A_T^{(1)})$ and its approximate distribution, which is a basis in derivation of our formula for average option prices. We set $k_{l,t} := \sum_{p=1}^{\infty} \mathbf{1}_{\{\tau_{p,l} \leq t\}}$, $k_l = k_{l,T}$, and $\check{S} := (\check{S}^1, \cdots, \check{S}^d)$ which is defined as

$$\check{S}_t^i := \sum_{q=1}^{2d} \int_0^t e^{\int_u^t \alpha_s ds} \phi_{S^i,q} \left(\sigma_{u-}^{(0)}, S_{u-}^{(0)} \right) dW_u^q + \sum_{l=1}^{n} \sum_{p=1}^{k_{l,t}} \upsilon_{S^i,l} \zeta_{S^i,p,l} e^{\int_0^t \alpha_s^i ds} s_0^i, \quad (69)$$

where each $\zeta_{S^i,p,l}$ follows $N(0,1)$ which is independent among p, l, but possibly correlated for among S^i.

In this setting, we have $g(A_T^{(1)}) = g(\check{A}_T) + g(\xi_{\{k_l\},T})(= \sum_{i=1}^{d} \check{A}_T^i + \sum_{i=1}^{d} \xi_{\{k_l\},T}^i)$, where

$$\check{A}_T^i := \sum_{j=1}^{m_i} w_j^{(i)} \check{S}_{t_j^{(i)}}^i 1_{\{t_j^{(i)} \leq T\}},$$

$$\xi_{\{k_l\},T}^i := \sum_{j=1}^{m_i} w_j^{(i)} \sum_{l=1}^{n} \left(k_{l,t_j^{(i)}} - \Lambda_l t_j^{(i)} \right) m_{S^i,l} e^{\int_0^{t_j^{(i)}} \alpha_s^i ds} s_0^i. \quad (70)$$

Then, for ease of numerical computation, i.e. to avoid the multiple time-integrals appearing in Corollary 3.3, given $\{N_l = k_l\}$ we approximate $\xi_{\{k_l\},t}^i$ and the distribution of $g(\check{A}_T)$ so as to make them independent of timing of jumps: With new notations $\bar{w}_i(t) := \sum_{j=1}^{m_i} w_j^{(i)} 1_{\{t \leq t_j^{(i)}\}} e^{\int_t^{t_j^{(i)}} \alpha_s ds}$ and $\tilde{w}_i := \frac{1}{T} \sum_{j=1}^{m_i} w_j^{(i)} t_j^{(i)} e^{\int_0^{t_j^{(i)}} \alpha_s ds}$, let us approximate $\xi_{\{k_l\},t}^i$ by $\xi_{\{k_l\}}^i$, which is defined as

$$\xi_{\{k_l\}}^i := \sum_{l=1}^{n} \tilde{w}_i (k_l - \Lambda_l T) m_{S^i,l} s_0^i, \quad (71)$$

17

where the number of jumps until $t_j^{(i)}$ (i.e. $k_{l,t_j^{(i)}}$) is replaced by the number of jumps until T (i.e. $k_l = k_{l,T}$). Moreover, the distribution of $g(\check{A}_T)$ is approximated as $N\left(0, \Sigma_T^{\{k_l\}}\right)$, that is the normal distribution with mean zero and variance $\Sigma_T^{\{k_l\}}$ defined as

$$
\begin{aligned}
\Sigma_T^{\{k_l\}} \quad := \quad & \sum_{i=1}^{d}\sum_{q=1}^{2d}\sum_{i'=1}^{d}\sum_{q'=1}^{2d}\int_0^T \left(\bar{w}_i(t)\phi_{S^i,q}\left(\sigma_t^{(0)}, S_t^{(0)}\right)\right)\left(\bar{w}_{i'}(t)\phi_{S^{i'},q'}\left(\sigma_t^{(0)}, S_t^{(0)}\right)\right)dt \\
& + \sum_{i=1}^{d}\sum_{i'=1}^{d}\sum_{l=1}^{n} k_l\left(\tilde{w}_i v_{S^i,l} s_0^i\right)\vartheta_{S^i,S^{i'}}\left(\tilde{w}_{i'} v_{S^{i'},l} s_0^{i'}\right),
\end{aligned}
\tag{72}
$$

where $\vartheta_{S^i,S^{i'}}$ denotes the correlation between $\zeta_{S^i,j,l}$ and $\zeta_{S^{i'},j,l}$ for given j and l. In the case of $d = 1$ and $n = 1$, $\Sigma_T^{\{k_l\}}$ corresponds to $\Sigma_T^{\{k\}}(T)$ in Remark B.2 of the online appendix [31]. In sum, given $\{N_l = k_l\}$, we obtain an approximation for the distribution of $g(A_T^{(1)})$ as $N(g(\xi_{\{k_l\}}), \Sigma_T^{\{k_l\}})$ with $g(\xi_{\{k_l\}}) = \sum_{i=1}^{d}\xi_{\{k_l\}}^i$.

Next, we need to calculate $\mathbf{E}[f_{i,T}]$ in Corollary 3.3, where we apply the approximations in Remark B.2 of the online appendix [31] to the multiple time-integrals. Furthermore, we use the 3rd order (ϵ^3) corrections obtained by a no-jump model, which is defined as

$$
S_T^{i,LSV(\epsilon)} \quad = \quad s_0^i + \int_0^T \alpha_t^i S_t^{i,LSV(\epsilon)}dt + \epsilon\int_0^T \phi_{S^i}\left(\sigma_t^{i,LSV(\epsilon)}, S_t^{i,LSV(\epsilon)}\right)dW_t,
\tag{73}
$$

$$
\sigma_T^{i,LSV(\epsilon)} \quad = \quad \sigma_0^i + \int_0^T \mu_{\sigma^i}\left(\sigma_t^{i,LSV(\epsilon)}, t\right)dt + \epsilon\int_0^T \phi_{\sigma^i}\left(\sigma_t^{i,LSV(\epsilon)}\right)dW_t.
\tag{74}
$$

Hereafter, we call this approximation the partial 3rd order (asymptotic) expansion. Then, for the numerical evaluation of conditional expectations appearing in $\mathbf{E}[f_{i,T}]$, we can use formulas in Appendix B of Shiraya - Takahashi [34], as well as Lemma B.1 of the online appendix [31] which is an extension of [34] and Takahashi et al. [42].

Finally, we summarize an approximate pricing formula for average call options as the following theorem.

Theorem 4.3. *An approximation formula for the initial value $C(K,T)$ of an average call option with maturity T and strike price K is given as follows:*

$$
\begin{aligned}
C(K,T) \quad \approx \quad & \sum_{k=0}^{\infty}\sum_{\sum_{l=1}^{n}k_l=k} p_{\{k_l\}}e^{-rT}\left\{\epsilon\left\{y_{k_l}N\left(\frac{y_{k_l}}{\sqrt{\Sigma_T^{\{k_l\}}}}\right) + \Sigma_T^{\{k_l\}}n(y_{k_l};0,\Sigma_T^{\{k_l\}})\right\}\right. \\
& +\epsilon^2\left\{C_{1,k_l}N\left(\frac{y_{k_l}}{\sqrt{\Sigma_T^{\{k_l\}}}}\right) + \left(C_{2,k_l}\Sigma_T^{\{k_l\}} + C_{3,k_l}\frac{H_1(y_{k_l};\Sigma_T^{\{k_l\}})}{\Sigma_T^{\{k_l\}}}\right)n(y_{k_l};0,\Sigma_T^{\{k_l\}})\right\} \\
& \left. +\epsilon^3\left\{\left(C_4\frac{H_2(\mathscr{Y};\Sigma_T)}{(\Sigma_T)^2} + C_5\frac{H_4(\mathscr{Y};\Sigma_T)}{(\Sigma_T)^4} + C_6\right)n(\mathscr{Y};0,\Sigma_T)\right\}\right\},
\end{aligned}
\tag{75}
$$

18

where $p_{\{k_l\}} = \prod_{l=1}^{n} \frac{(\Lambda_l T)^{k_l} e^{-\Lambda_l T}}{k_l!}$, $\Sigma_T = \Sigma_T^{\{0\}}$, r is a constant risk-free rate, K is expressed as $K = g(A_T^{(0)}) - \epsilon \mathscr{Y}$ for an arbitrary $\mathscr{Y} \in \mathbf{R}$, $y_{\{k_l\}} = g(\xi_{\{k_l\}}) + \mathscr{Y}$, $N(x)$ denotes the standard normal distribution function and $n(x; 0, \Sigma) = \frac{1}{\sqrt{2\pi\Sigma}} \exp\left(\frac{-x^2}{2\Sigma}\right)$. Here, g, $\xi_{\{k_l\}} = (\xi_{\{k_l\}}^1, \cdots, \xi_{\{k_l\}}^d)$ and $\Sigma_T^{\{k_l\}}$ are given by (54), (71) and (72), respectively. The coefficients $C_{1,k_l}, \cdots, C_{3,k_l}, C_4, \cdots, C_6$ are some constants, and $H_k\left(x; \Sigma_T^{\{k_l\}}\right)$ denotes the k-th order Hermite polynomial.

The concrete expressions of coefficients $C_{1,k_l}, \cdots, C_{3,k_l}, C_4, \cdots, C_6$ and the details of the derivation are provided in Section A of the online appendix [31].

Remark 4.4. *When we do not have closed-forms for the multiple integrals on the time parameter that appear in the calculation of the conditional expectation formulas, some numerical method is necessary.*

However, all the multiple integrals necessary for the evaluation of $C_{1,k}, \cdots, C_{3,k}, C_4, \cdots, C_6$ are computed by the program code with only one loop against the time parameter. For instance, a multiple integral is approximated for the numerical integration as follows:

$$\int_0^T f(s) \int_0^t g(u) \int_0^s h(v) dv du ds$$

$$\approx \sum_{i=1}^{M} \Delta_{t_i} f(t_i) \sum_{j=1}^{i} \Delta_{t_j} g(t_j) \sum_{k=1}^{j} \Delta_{t_k} h(t_k)$$

$$= \sum_{i=1}^{M} \Delta_{t_i} f(t_i) \left(G(t_{i-1}) + \Delta_{t_i} g(t_i) \left(H(t_{i-1}) + \Delta_{t_i} h(t_i)\right)\right),$$

where $\Delta_{t_i} = (t_i - t_{i-1})$, $H(t_i) = H(t_{i-1}) + \Delta_{t_i} h(t_j)$ and $G(t_i) = G(t_{i-1}) + \Delta_{t_i} g(t_i) H(t_i)$.

Hence, the order of the computational effort is at most M, where M is the number of time-steps for the discretization in the numerical integral. Note that we have no problems in terms of computational complexity and speed since various fast numerical integration methods are available such as the extrapolation method.

Remark 4.5. *To calculate the option premium with our approximation method numerically, we need to truncate the number of jumps. In the numerical examples, we ignore the jumps whose probability is less than $1/100,000$.*

5 Numerical Examples on WTI Average and Spread Options

This section shows concrete numerical examples based on our method developed in the previous sections. Firstly, to calculate the option values, we adopt two types of stochastic volatility models for the numerical experiments under the risk-neutral probability measure. In particular, each underlying asset price process has a CEV (constant elasticity of variance)-type diffusion term with compound Poisson component. On volatility processes, one has a log-normal diffusion term with no drift and compound Poisson component, which is an extension of the SABR model (see [19]), and we call it as extended SABR ((76), (77)). The other is the Heston type variance

process, which is an extension of Bates model (see [3]), and we call it as extended Bates ((78), (79)). The explicit form of the processes are as follows:

$$S_T^i = S_0^i + \int_0^T v^i \sigma_{t-}^i (S_{t-}^i)^{\beta_i} dW_t^{S^i} + \sum_{l=1}^n \left(\sum_{j=1}^{N_{l,T}} h_{S^i,l,j} S_{\tau_{j,l}-}^i - \int_0^T \Lambda_l S_{t-}^i \mathbf{E}[h_{S^i,l,1}] dt \right) \quad (76)$$

$$\sigma_T^i = \sigma_0^i + \int_0^T \nu^i \sigma_{t-}^i dW_t^{\sigma^i} + \sum_{l=1}^n \left(\sum_{j=1}^{N_{l,T}} m_{\sigma^i,l} \sigma_{\tau_{j,l}-}^i - \int_0^T \Lambda_l \sigma_{t-}^i m_{\sigma^i,l} dt \right), \quad (77)$$

$$S_T^i = S_0^i + \int_0^T v^i \sqrt{\sigma_{t-}^i} (S_{t-}^i)^{\beta_i} dW_t^{S^i} + \sum_{l=1}^n \left(\sum_{j=1}^{N_{l,T}} h_{S^i,l,j} S_{\tau_{j,l}-}^i - \int_0^T \Lambda_l S_{t-}^i \mathbf{E}[h_{S^i,l,1}] dt \right), \quad (78)$$

$$\sigma_T^i = \sigma_0^i + \int_0^T \kappa^i (\theta^i - \sigma_{t-}^i) dt + \int_0^T \nu^i \sqrt{\sigma_{t-}^i} dW_t^{\sigma^i}, \quad (79)$$

where the jump size $m_{\sigma^i,l}$ in the volatility process is a constant, while the jump size in the futures price process is a constant or log-normally distributed, that is $h_{S^i,l,j} = e^{Y_{S^i,l,j}} - 1$ with $Y_{S^i,l,j}$ following a normal distribution $N(m_{S^i,l}, v_{S^i,l}^2)$ for all j. We remark that in the diffusion component, adopting a local volatility with Heston model rather than the original Heston is due to its well-known advantage in calibration to the option prices in practice. (For the CEV-Heston model, see [38] for instance.)

We apply the formula with $\epsilon = 1$ in Theorem 4.3, and calculate the average options on WTI futures, which are regarded as examples of the average option described in Section 4. In particular, we set $d = 2$, $w_j^{(i)} = \frac{1}{M}$, $M = d_1 + d_2$ for all $w_j^{(i)}$ in (54) where d_1, d_2 express the numbers of the reference dates of each underlying asset, because two WTI futures are relevant for one average option in a listed contract in CME.

Although we are able to treat more general cases, we consider a systematic jump (i.e. all the jumps of the underlying asset prices and their volatilities occur at the same time) and perfectly correlated jump size case, that is $n = 1$ and $\vartheta_{S^i,S^{i'}} = 1$. Also, the intensity parameter Λ is fixed as 1.

Then, we set the calculation date for average option prices as of November 28, 2014, and report the target average option contracts in CME with their underlying average prices on the date, the time to the maturities and the relevant interest rates in Table 1. We also show the WTI futures contracts relevant for each average contract, their prices, and the time to the maturities with interest rates on November 28, 2014 in Table 2:

20

	Average Price	Maturity	Interest Rate
MAY 15	67.16	0.499	0.3%
NOV 15	68.26	1.005	0.4%

Table 1: Market data for average options

Average Price is the underlying asset price of the average option. Maturity is the time to maturity of the option.

| Contract Month	Futures Price	Maturity	Interest Rate
JUN 15	67.1	0.458	0.3%
JUL 15	67.26	0.551	0.3%
DEC 15	68.22	0.970	0.4%
JAN 16	68.38	1.049	0.4%

Table 2: Market data for futures options

Futures Price is the underlying asset price of the listed option. Maturity is the time to maturity of the option.

Here, JUN 15 and JUL 15 are the relevant futures contracts for the MAY 15 average option, and DEC 15 and JAN 16 for the NOV 15 average option.
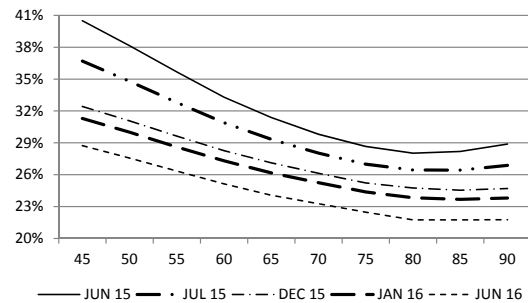
Moreover, we compute the average option prices by our approximation formula to compare those with the corresponding CME prices. Hence, we first need to obtain the model parameters through calibration to the relevant WTI futures options prices. Figure 1 and Figure 2 show the implied volatilities of the WTI futures options.

Figure 1: Original Implied Volatilities
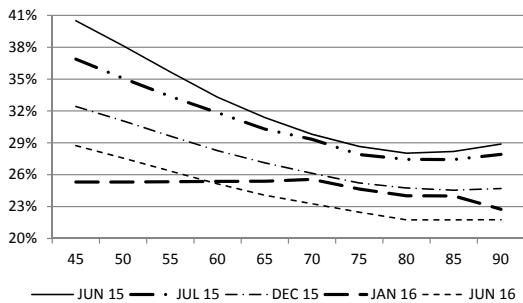
Figure 2: Interpolated Implied Volatilities

X axis is strike prices of the listed options, and Y axis is the value of the implied volatility.

X axis is strike prices of the listed options, and Y axis is the value of the implied volatility.

Figure 1 ("original implied volatilities") reports the implied volatilities of JUN15, JUL15, DEC15, JAN16, JUN16 contracts of the WTI futures options.

However, as it is well-known that the odd term options, JUL15 and JAN16 are illiquid. Thus we do not use these implied volatilities calculated from their quoted prices but employ the volatilities based on the interpolations of the implied volatilities for the liquid term options, that is, JUN15, DEC15 and JUN16. In Figure 2 ("interpolated implied volatilities"), the volatilities of JUN15 and JAN16 contracts are the interpolated volatilities, where JUN15 and DEC15 contracts are used for JUL15, while DEC15 and JUN16 for JAN16. That is, we get the implied volatility σ for the illiquid option with the time to maturity T by the following formula: for $T_1 < T < T_2$,

$$\sigma^2 T = ((T_2 - T)\sigma_1^2 T_1 + (T - T_1)\sigma_2^2 T_2)/(T_2 - T_1), \qquad (80)$$

where σ_1 and σ_2 are the implied volatilities of the liquid options with the time to maturities T_1 and T_2, respectively.

It is common in practice that pricing an average option is based on the calibration to the prices of the liquid vanilla options whose underlying assets are the same spot or futures contracts as those of the average option. For the calibration, the parameter β_i in the diffusion term of the futures price process in (76) is set as 0.5, and $v \equiv 1$ for extended SABR for the stability of calibration. On the other hand, for extended Bates model, we need to estimate β_i by calibration, because the original Bates model ($\beta_i = 1$) is too restrictive to fit the model to the market skews. Moreover, we fix the parameters $\sigma_0 = \theta \equiv 1$ and $\kappa \equiv 1$ to stabilize the other parameters. In the jump component, the parameters are assumed to take common values for the two relevant futures price processes used for the calculation of an average option price and are obtained by calibration to the market futures options prices. This is true for the volatility processes of the two futures prices.

The correlations and local volatility parameters between the futures prices and their volatilities are assumed to take common values for the two relevant futures, which are used to calculate an average option. Then, the correlations are obtained by calibration to the market futures options prices, which are shown in Table 4. In order to avoid a degenerate correlation matrix, we implement a joint calibration for the two associated futures options contracts.

Further, for the two relevant futures contracts, the correlation between a futures price and the futures price's volatility of the different contract is assumed to be the same as the correlation between the price and its volatility.

Moreover, the correlations between the two futures price processes are estimated by the past one month historical data of future prices. The correlations between the corresponding volatility processes of the two futures prices are assumed to be the same as the correlations of the futures prices. Then, we obtain the following estimates: the correlation between JUN 15 and JUL 15 is 0.99986, and the correlation between DEC 15 and JAN 16 is 0.99992.

For computational efficiency, the settlement prices of American options are transformed to those of the European options before calibration: More precisely, after an implied volatility of each American option price is estimated under a binomial version of the Black-Scholes model, the corresponding European option price is computed. Hereafter, this European option price is called the "transformed CME" option price. Then, calibration is implemented against the "transformed CME" option prices with different strikes simultaneously, where out-of-the-money (OTM) prices are used for the calibration; the strikes of the options range USD (us dollar) 35 to USD 75 with every 5 dollars for JUN15 and JUL15 futures options, and those of range USD 40 to USD 80 with every 5 dollars for DEC15 and JAN16 futures options.

Given the above assumptions in the calibration, we compare the following six specifications of the model:

(i) Local stochastic volatility model without jump (SABR model, i.e. $h_{S^i,l,j} = 0$, $m_{\sigma^i,l,j} = 0$ in (76), (77))

(ii) Local stochastic volatility with constant jumps in the futures prices model (SABR + constant jump model, i.e. $h_{S^i,l,j} = m_{S^i,l,j}$ (a constant), $m_{\sigma^i,l,j} = 0$ in (76), (77))

(iii) Local stochastic volatility with log-normal jumps in the futures prices model (SABR +

log-normal jump model i.e. $h_{S^i,l,j} = e^{Y_{S^i,l,j}} - 1$, $m_{\sigma^i,l,j} = 0$ in (76), (77))

(iv) Local stochastic volatility with constant jumps in the futures prices and volatilities model (SABR + constant jumps in the futures prices and volatilities model, i.e. $h_{S^i,l,j} = m_{S^i,l,j}$, $m_{\sigma^i,l,j}(\neq 0)$ in (76), (77))

(v) Local stochastic volatility with log-normal jumps in the futures prices and constant jumps in the volatilities model (SABR + log-normal jumps in the futures prices and constant jumps in the volatilities model, i.e. $h_{S^i,l,j} = e^{Y_{S^i,l,j}} - 1$, $m_{\sigma^i,l,j}(\neq 0)$ in (76), (77))

(vi) Local stochastic volatility with log-normal jumps in the futures prices model (Bates + local volatility model, i.e. $h_{S^i,l,j} = e^{Y_{S^i,l,j}} - 1$ in (78), (79))

We choose these models to demonstrate that our method is well applicable to models such as SABR and Heston with local volatility models, which are very popular and frequently used in practice, but have no explicit characteristic functions nor closed form probability density functions.

Due to limitations of space, the results of the 6 month maturity cases are given upon request.

The parameters obtained by calibration to the WTI futures options with 1 year maturity are shown in Table 3.

		v	$\sigma(0)$	β	ν	ρ	m_S	v_S	m_σ
DEC 15	(i)	1	215.5%	0.5	44.9%	-0.215	-	-	-
	(ii)	1	118.2%	0.5	148.5%	0.072	-0.171	-	-
	(iii)	1	159.5%	0.5	44.1%	-0.801	-0.024	0.192	-
	(iv)	1	128.9%	0.5	151.9%	-0.377	-0.151	-	-0.685
	(v)	1	158.8%	0.5	31.6%	-1.000	-0.036	0.199	0.153
	(vi)	5.50	100.0%	0.206	107.1%	-0.706	-0.026	0.199	-
JAN 16	(i)	1	210.3%	0.5	43.6%	-0.215	-	-	-
	(ii)	1	114.4%	0.5	139.5%	0.072	-0.171	-	-
	(iii)	1	151.5%	0.5	46.4%	-0.801	-0.024	0.192	-
	(iv)	1	115.9%	0.5	163.8%	-0.377	-0.151	-	-0.685
	(v)	1	150.9%	0.5	33.4%	-1.000	-0.036	0.199	0.153
	(vi)	5.25	100.0%	0.206	117.7%	-0.706	-0.026	0.199	-

Table 3: Parameters obtained by calibration to WTI futures options(DEC15, JAN16)

Each (i),$\cdots$,(vi) shows a model number.

The results of the calibration to the WTI futures options are shown in Table 4. It is observed that the models with jumps (ii)-(v) mostly provide the better fitting than the one without jumps (i), especially much better in OTM. Moreover, among the models with jumps, Model (ii) is the worst in terms of calibration.

	Strike	Transformed CME Price	Price (i)	(ii)	(iii)	(iv)	(v)	(vi)	Difference (i)	(ii)	(iii)	(iv)	(v)	(vi)
DEC 15	45	23.92	23.99	23.95	23.93	23.94	23.93	23.93	0.06	0.02	0.00	0.02	0.00	0.00
	50	19.58	19.59	19.58	19.57	19.58	19.57	19.57	0.00	-0.00	-0.01	-0.00	-0.01	-0.01
	55	15.54	15.51	15.53	15.54	15.53	15.54	15.54	-0.03	-0.01	-0.00	-0.01	-0.00	-0.00
	60	11.90	11.86	11.90	11.91	11.90	11.91	11.91	-0.04	-0.00	0.01	0.00	0.01	0.01
	65	8.77	8.74	8.76	8.77	8.76	8.77	8.77	-0.03	-0.01	0.00	-0.00	0.00	0.00
	70	6.20	6.20	6.18	6.19	6.18	6.19	6.19	0.00	-0.02	-0.01	-0.02	-0.01	-0.01
	75	4.18	4.24	4.19	4.19	4.19	4.19	4.19	0.06	0.01	0.01	0.01	0.01	0.01
	80	2.76	2.80	2.76	2.75	2.76	2.76	2.76	0.04	-0.00	-0.00	-0.00	-0.00	-0.00
	85	1.80	1.80	1.81	1.80	1.81	1.80	1.80	-0.00	0.01	-0.00	0.01	-0.00	-0.00
	90	1.20	1.13	1.20	1.19	1.20	1.20	1.20	-0.07	0.00	-0.01	-0.00	-0.00	-0.00
JAN 16	45	24.10	24.17	24.12	24.11	24.12	24.11	24.11	0.07	0.01	0.01	0.01	0.00	0.01
	50	19.77	19.78	19.76	19.77	19.76	19.77	19.77	0.01	-0.01	-0.01	-0.01	-0.01	-0.01
	55	15.74	15.71	15.73	15.74	15.73	15.74	15.74	-0.03	-0.01	-0.00	-0.01	-0.00	-0.00
	60	12.11	12.07	12.11	12.11	12.11	12.11	12.11	-0.04	0.01	0.01	0.01	0.01	0.01
	65	8.97	8.94	8.98	8.97	8.98	8.97	8.97	-0.02	0.01	0.00	0.01	0.00	0.00
	70	6.39	6.38	6.38	6.37	6.38	6.37	6.37	-0.00	-0.00	-0.02	-0.00	-0.02	-0.02
	75	4.34	4.40	4.36	4.35	4.35	4.35	4.35	0.05	0.02	0.00	0.01	0.00	0.00
	80	2.88	2.93	2.88	2.89	2.88	2.89	2.89	0.05	0.00	0.00	0.00	0.00	0.00
	85	1.91	1.90	1.90	1.91	1.90	1.90	1.91	-0.00	-0.01	0.00	-0.00	-0.00	-0.00
	90	1.28	1.21	1.26	1.29	1.27	1.28	1.28	-0.07	-0.02	0.01	-0.01	0.00	0.00

Table 4: Calibration Errors on WTI futures options(DC15, JAN16)

Each (i),$\cdots$,(vi) shows a model number. Transformed CME Price is the corresponding European option price calculated from the settlement price of an American option traded on CME. Difference columns show the difference between the approximated prices and the transformed CME prices.

5.1 Numerical Results of Average Options

Using the parameters obtained through the calibration, Table 5 shows the comparison of the 1 year average option prices given by our approximation (AE) and CME settlement (CME).

			55(Put)	60(Put)	65(Put)	70(Put)	75(Call)	80(Call)
Price	CME MAY15 (Average)		2.27	3.6	5.44	7.86	4.11	2.69
	AE	(i)	2.24	3.56	5.40	7.84	4.14	2.72
		(ii)	2.21	3.53	5.35	7.75	4.03	2.62
		(iii)	2.27	3.60	5.43	7.83	4.10	2.68
		(iv)	2.22	3.53	5.35	7.74	4.01	2.61
		(v)	2.27	3.60	5.43	7.83	4.11	2.69
		(vi)	2.27	3.60	5.43	7.83	4.11	2.68
Difference	AE - CME	(i)	-0.03	-0.04	-0.04	-0.02	0.03	0.03
		(ii)	-0.06	-0.07	-0.09	-0.11	-0.08	-0.07
		(iii)	-0.00	0.00	-0.01	-0.03	-0.01	-0.01
		(iv)	-0.05	-0.07	-0.09	-0.12	-0.10	-0.08
		(v)	-0.00	0.00	-0.01	-0.03	-0.00	-0.00
		(vi)	0.00	0.00	-0.01	-0.03	-0.00	-0.01

Table 5: Comparison of average option prices of our approximation with CME listed prices (1Y)

AE shows the price with the partial 3rd order approximation formula. Each (i),$\cdots$,(vi) shows a model number. Difference rows show the difference between the approximated prices and the settlement prices of average options in CME.

It is observed that in terms of the deviations of our approximations from the CME settlement

prices, Model (v) is the best, (iii) and (vi) follow.

Next table and figures show the difference of the computational time and accuracies in each order of expansion. Table 6 presents the comparison of the 1 year average option prices given by our approximation (AE) with the partial third order expansion and Monte Carlo simulations (MC) with a control variate method (the value in parenthesis is the 95% confidence interval (e.g. (1.5), (1.6) in pp 5-6 of [18])). This control variate method introduced by [36] can be applied to the models whose characteristic functions are unknown as in these models, and its calculation speed ranges from several tens to hundreds of times faster than that of the crude Monte Carlo method to reduce the variance. Figures from 3 to 8 show the difference between each order approximation values and MC values.
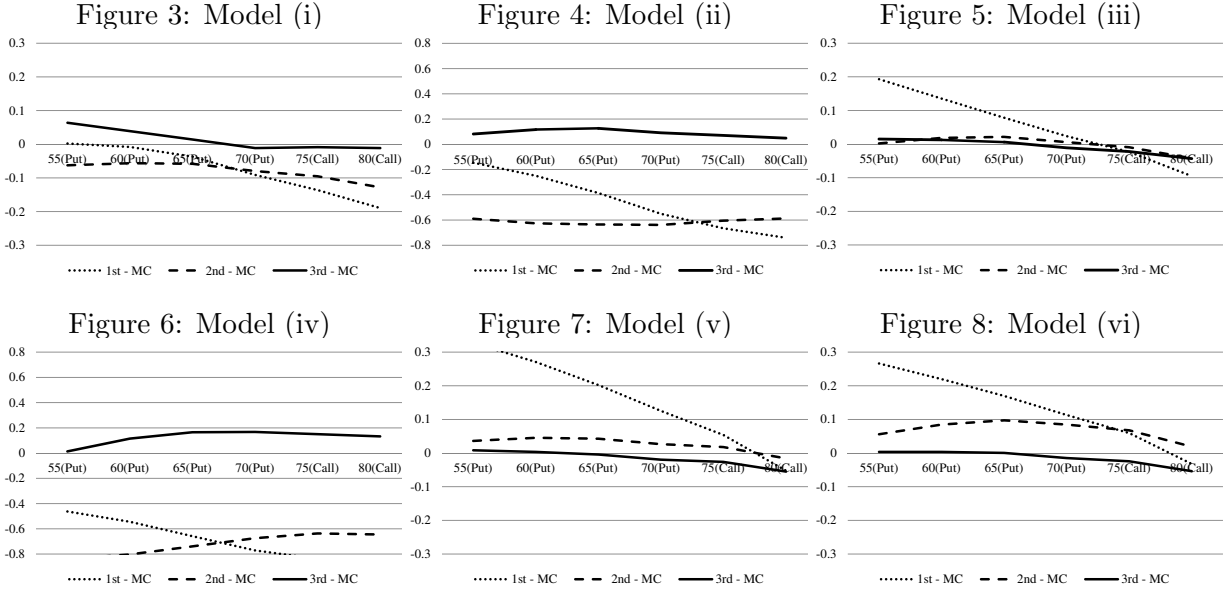
The program is implemented in C++, and the computational time is calculated with one core of Intel Core(TM) i7-3960X CPU @ 3.30GHz 32GB RAM. To get the accurate computational time, it is evaluated by the average of 6 times calculations for Monte Carlo simulation and the average of 6,000 times calculation for the approximation method.

In Monte Carlo simulations, the number of time partitions is set to 10 per month (in the case of the average reference month, the time steps are the same as the numbers of business days in the month). The numbers of simulations are set to 100,000 for the average options and 10,000 for spread options.

		55(Put)	60(Put)	65(Put)	70(Put)	75(Call)	80(Call)	Time(sec)
AE	(i)	2.24	3.56	5.40	7.84	4.14	2.72	0.005
3rd	(ii)	2.21	3.53	5.35	7.75	4.03	2.62	0.005
(Partial)	(iii)	2.27	3.60	5.43	7.83	4.10	2.68	0.005
	(iv)	2.22	3.53	5.35	7.74	4.01	2.61	0.005
	(v)	2.27	3.60	5.43	7.83	4.11	2.69	0.005
	(vi)	2.27	3.60	5.43	7.83	4.11	2.68	0.005
MC	(i)	2.17 (0.027)	3.52 (0.026)	5.39 (0.025)	7.84 (0.024)	4.15 (0.023)	2.72 (0.022)	13.8
	(ii)	2.13 (0.06)	3.41 (0.058)	5.22 (0.056)	7.65 (0.054)	3.97 (0.052)	2.58 (0.05)	13.8
	(iii)	2.25 (0.012)	3.59 (0.011)	5.43 (0.011)	7.84 (0.01)	4.13 (0.009)	2.73 (0.009)	13.8
	(iv)	2.2 (0.063)	3.42 (0.061)	5.18 (0.06)	7.57 (0.058)	3.88 (0.056)	2.49 (0.054)	13.8
	(v)	2.26 (0.005)	3.6 (0.005)	5.44 (0.004)	7.85 (0.004)	4.13 (0.004)	2.74 (0.004)	13.8
	(vi)	2.26 (0.014)	3.6 (0.014)	5.43 (0.013)	7.85 (0.012)	4.13 (0.011)	2.74 (0.01)	15.7
Difference	(i)	0.07	0.04	0.02	-0.01	-0.00	-0.01	
	(ii)	0.09	0.12	0.13	0.09	0.06	0.04	
	(iii)	0.02	0.01	0.01	-0.01	-0.02	-0.05	
	(iv)	0.01	0.11	0.17	0.16	0.13	0.12	
	(v)	0.01	0.00	-0.00	-0.02	-0.03	-0.05	
	(vi)	0.01	0.00	0.00	-0.02	-0.03	-0.06	

Table 6: Comparison of accuracies and computational times of our approximation with those of MC (1Y average options)

AE shows the price with our approximation formula, and 3rd shows the order of ϵ in our formula (the 3rd order approximation is applied only to the diffusion terms, and the jump terms are approximated with the 2nd order formula). MC shows the price with Monte Carlo simulation with a control variate method. Each (i),···,(vi) shows a model number. Difference rows show the difference between the approximated prices and the prices calculated with Monte Carlo simulations. The value in parenthesis is the 95% confidence interval. Time is the calculation time.

Figure 3: Model (i) Figure 4: Model (ii) Figure 5: Model (iii)

Figure 6: Model (iv) Figure 7: Model (v) Figure 8: Model (vi)

X axis is strike prices, and Y axis is the difference between the approximated values and Monte Carlo values. "1st" is the first order values, "2nd" is the first order values, and "3rd" is the partial third order values.

As for discussions below on the validity of our expansion method in practice, since models with jumps work better than the one without jumps (model (i)) in calibration as seen in Table 4, let us concentrate on models with jumps (ii)-(vi).

The 1st and 2nd order values in the models (ii), (iv) are worse than the models (iii), (v), (vi). In fact, the constant jump parameters in the models (ii), (iv) may be too large in approximations with our method, which seems the reason for the worse results in those models. This is consistent with an observation (explained in Remark C.5) based on a theoretical analysis in Theorem 3.2 and Appendix C. We also note that the confidence interval of a Monte Carlo simulation becomes wider when the accuracy of our approximation becomes worse.

Then, examining overall results, we see that the 2nd order approximations improve the 1st order ones in the models (iii), (v), (vi). Thus, we can conclude at least from the numerical experiments that the increasing order of expansion for the jump parts improves the precision of the approximation under such reasonable jump parameters in the models (iii), (v), (vi). Moreover, it is observed that the partial 3rd order expansions generally improve the approximations than the lower order expansions.

As for the calculation time, our approximation is about $2800 \sim 3100$ times as fast as the Monte Carlo simulation with control variate method even if the approximation is the partial 3rd order. The computational time for the lower orders is around 0.0007 seconds for the 1st order, around 0.003 seconds for the 2nd order. The computational times for the lower orders are around 0.0007 seconds for the 1st order and around 0.003 seconds for the 2nd order. Although those for the 1st order cases are about $4 \sim 8$ times faster than the ones for the 2nd or the partial 3rd order, it is hard to use the 1st order approximation in practice because of its large approximation errors.

We finally remark that the partial 3rd order expansion is able to calculate the prices very

fast with good accuracy, which is useful when pricing quickly is necessary while very precise values are not required (e.g. when we need to indicate prices for many customers in a short time).

Next, we implement stress tests with respect to the parameters to see the robustness of our method. We set v or σ_0, ν, m_S, v_S, T twice as much as those in Table 3 for the case (iii) and (vi), which are the extended SABR and the extended Bates models, respectively.

Figure 9: Extended SABR model	Figure 10: Extended Bates model
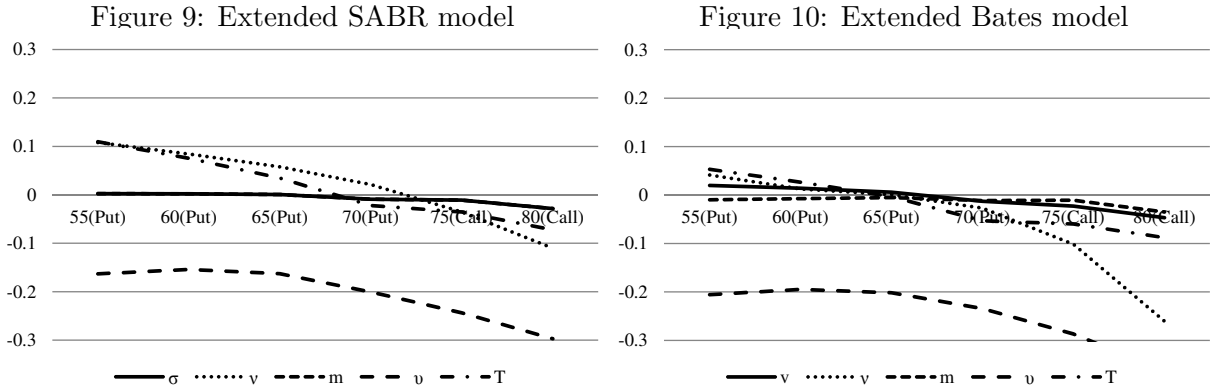
X axis is strike prices, and Y axis is the difference between the partial third order approximation values and Monte Carlo values. Each parameter shows the results with the parameter which is set as twice as much as that in Table 3.

Figures 9 and 10 is consistent with asymptotic error estimates in Theorem 3.2 and in Appendix C, since the large ϵ and the large T mean the large volatility, the large jump size and the long maturity. Particularly, the large ν, v_S which create large curvatures in the volatility smile, make the approximation less accurate and make the confidence intervals of Monte Carlo simulation wider.

5.2   Numerical Results of Calendar Spread Options

This subsection shows numerical analysis for the calendar spread options whose underlying assets are the same futures as those of the average options in the previous section. Hereafter, we concentrate on (i), (iii) models (which are called LSV model and LSV with log-normal jump model, respectively).

Since the calculation date and the underlying futures of the spread options are the same as those of the average options in the previous section, the calibrated parameters in the both models are the same as for (i), (iii) in Table 3.

The maturity of calendar spread option of DEC15 is 19 Nov. 2015. The correlation between DEC 15 and JAN 16 is 0.99992, which is estimated from the past one month historical data of the futures prices as in the average option cases.

The results of the approximate prices based on the formula in Theorem 4.3 with Monte Carlo benchmarks are given in Table 7 for the 1 year calendar spread option.

		55(Put)	60(Put)	65(Put)	70(Put)	75(Call)	80(Call)	Time
AE	(i)	0.00	0.02	0.08	0.13	0.04	0.01	0.003
	(iii)	0.00	0.02	0.12	0.20	0.06	0.01	0.003
MC	(i)	0 (0.006)	0.01 (0.006)	0.07 (0.005)	0.13 (0.005)	0.04 (0.004)	0.01 (0.003)	1.38
	(iii)	0 (0.002)	0.02 (0.002)	0.12 (0.001)	0.2 (0.001)	0.06 (0.001)	0.02 (0.001)	1.39
AE - MC	(i)	0.00	0.01	0.00	0.00	0.00	-0.01	
	(iii)	0.00	-0.00	-0.00	0.00	-0.00	-0.00	

Table 7: Comparison of accuracies and computational times of our approximation with those of MC (1Y spread options)

AE shows the price with the partial 3rd order approximation formula. MC shows the price with Monte Carlo simulation with a control variate method. Each (i), (iii) shows a model number. "AE - MC" rows show the difference between the approximated prices and the prices calculated with Monte Carlo simulations. The value in parenthesis is the 95% confidence interval. Time is the calculation time.

It is observed that the errors for the LSV model (i) are larger than those for the LSV jump model (iii), which seems mainly because the levels of the (calibrated) initial volatilities in the LSV model are larger than those in the LSV jump model. (See $\sigma(0)$ in Table 3.)

We also note that since the correlations are estimated from the past one month historical data of the underlying futures prices, the estimated prices in Table 7 are slightly different from the CME settlement prices. In fact, we also estimate the correlations between the two underlying futures prices through the calibration to the calendar spread options, which are reported in Table 8.

		DEC15 and JAN16
Historical correlation		0.99992
Calibrated correlation	(i)	0.99907
	(iii)	0.99876
Diff	(i)	-0.00084
	(iii)	-0.00115

Table 8: Calibrated correlations and Historical correlations

Historical correlation shows the correlation calculated by the time series data. Calibrated correlation shows the calibrated correlation from the spread option. Each (i), (iii) shows a model number. Diff rows show the difference between the historical correlation and the calibrated correlation.

Moreover, Table 9 shows the approximate spread option prices with historically estimated or calibrated correlations.

		-1.5(Put)	-1.0(Put)	-0.5(Put)	0(Call)	0.5(Call)	1(Call)
CME		0.03	0.09	0.22	0.28	0.14	0.05
AE	(i)	0.00	0.02	0.08	0.13	0.04	0.01
Historical	(iii)	0.00	0.02	0.12	0.20	0.06	0.01
AE	(i)	0.03	0.09	0.21	0.29	0.13	0.06
Calibrated	(iii)	0.03	0.09	0.22	0.29	0.13	0.05
AE - CME	(i)	-0.03	-0.07	-0.14	-0.15	-0.10	-0.04
Historical	(iii)	-0.03	-0.07	-0.10	-0.08	-0.08	-0.04
AE - CME	(i)	0.00	-0.00	-0.01	0.01	-0.01	0.01
Calibrated	(iii)	0.00	-0.00	0.00	0.01	-0.01	-0.00

Table 9: Comparison of spread option prices using calibrated correlation and historical correlation with CME listed prices (1Y)

CME shows the settlement spread option price in CME. "AE Historical" shows the spread option price with the partial 3rd order approximation formula using the historical correlation. "AE Calibrated" shows the spread option price with the partial 3rd order approximation formula using the calibrated correlation. "AE - CME Historical" shows the difference between the price of "AE Historical" and that of "CME." "AE - CME Calibrated" shows the difference between the price of "AE Calibrated" and that of "CME." Each (i), (iii) shows a model number in.

We can see the high sensitivity of the prices to the changes in the correlations. Thus, it seems difficult to obtain the accurate correlation from the historical data in order for calibration to the long term spread option prices.

On the other hand, the effect of the correlation on the average option prices seems very little, which is observed in Table 10 below.

			55(Put)	60(Put)	65(Put)	70(Put)	75(Call)	80(Call)
Price	Historical Correlation	(i)	2.24	3.56	5.40	7.83	4.14	2.71
		(iii)	2.27	3.60	5.43	7.83	4.10	2.68
	Calibrated Correlation	(i)	2.24	3.56	5.40	7.83	4.14	2.71
		(iii)	2.27	3.60	5.43	7.83	4.10	2.68
Diff	Calibrated corr.	(i)	0.000	0.000	0.000	0.000	0.000	0.000
	- Historical corr.	(iii)	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001

Table 10: Comparison of average option prices using calibrated correlation and historical correlation with CME listed prices (1Y)

"Historical Correlation" shows the average option price with the partial 3rd order approximation formula using the historical correlation. "Calibrated Correlation" shows the average option price with the partial 3rd order approximation formula using the calibrated correlation. Diff shows the difference between the result of "Historical Correlation" and that of "Calibrated Correlation". Each (i), (iii) shows a model number.

5.3 Practical Relevance

This subsection briefly explains the practical relevance of the numerical examples in this section.

Firstly, it is important in practice that all derivatives are evaluated with the same model and parameters to hedge these derivatives consistently. If the model or parameters are different in each product, an arbitrage among these products may exist, and combining these options may make manipulated P/Ls. Thus, LSV models are frequently used in practice (see e.g. pp.162-163

in [11], Section 2 in [22], and p.3 in [28]) because only LV, SV or simple affine models have some difficulties in fitting to both market plain-vanilla and exotic option prices simultaneously.

Hence, we have provided an approximation formula for pricing the futures, average and spread options, which are very popular and traded mainly in oil derivatives markets, and confirmed its validity by comparing approximate prices with the listed option prices in CME. Moreover, we have shown that LSV models with jumps are more suitable for calibrating market option prices than LSV models without jumps. We note that the model needs to be chosen carefully because some models do not fit the multiple markets in the same parameters (e.g. constant jumps models did not fit in our examples).

Next, we note that if we could obtain all option prices from the listed derivatives market, it would be easy to evaluate P/Ls of traders' positions by using the listed prices as the present values of derivatives. However, because average options are mainly traded on OTC markets, and the strikes and maturities of average options traded as the listed option are very limited, it is so hard to obtain individual average option prices without a common option pricing model among all contracts. The numerical examples above demonstrated that our method is very effective for this purpose.

Finally, for estimating these OTC option premiums consistently with listed option prices, the results in the numerical examples indicate an appropriate calibration and pricing procedure in practice, which is summarized as follows:

1. calibrate model parameters for the underlying asset prices and their volatilities to the listed futures options prices.

2. calibrate the correlation parameters between two futures prices (and their volatilities) to the market spread option prices.

3. pricing the average and the spread options in OTC contracts for the counterparties and customers.

Here, we remark that since the sensitivity of spread options on the correlation between two futures prices is larger than that in the average option, as observed in the numerical examples, spread options need to be used for calibration before pricing average options.

Accordingly, one can manage positions of average options with relevant futures and spread options in a unified and consistent manner based on a common underlying model with our method.

6   Conclusion

We have shown a new approximation formula for pricing average and spread options in local-stochastic volatility with jumps models, and have provided a justification of our approximation method with some asymptotic error estimates for general payoff functions. Especially, our model admits local volatility functions and jumps in both the underlying asset price and its volatility processes. Thanks to the closed-form formula, the computational speed of the method is very fast, which makes sensitivity analyses and calibration procedures much more efficient than the other numerical schemes.

Moreover, in numerical experiments, we firstly calibrate the model to the WTI futures options by applying our approximation formula to pricing the plain-vanilla option prices. Then, valuation of average and spread options on the WTI futures based on the parameters and comparing those with the CME settlement prices have demonstrated the effectiveness of our approximation scheme.

Further, we have obtained the implied correlations embedded in the spread options and found the importance of the jump component in the futures price processes for duplicating consistently and precisely listed futures option, calendar spread option and average option prices with common parameters.

7 Acknowledgments

References

[1] Alòs, E. and León, J.A. On the short-maturity behaviour of the implied volatility skew for random strike options and applications to option pricing approximation. *Quantitative Finance*, pages 31–42, 16(1) 2016.

[2] Alòs, E., Eydeland, A. and Laurence, P. A kirk's and a bachelier's formula for three-asset spread options. *Energy Risk*, pages 52–57, 9 2011.

[3] Bates, D. S. Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *The Review of Financial Studies*, pages 69–107, 9(1) 1996.

[4] Bichteler, K., Gravereaux, J.B. and Jacod, J. *Malliavin Calculus for Processes with Jumps, Stochastic Monographs, vol.2*. London, Gordon and Breach / M. Davis, ed., 1987.

[5] Cai, N., Kou, S. Pricing asian options under a hyper-exponential jump diffusion model. *Operations Research*, pages 64–77, 60(2) 2012.

[6] Cai, N., Li, C., Shi, C. Closed-form expansions of discretely monitored asian options in diffusion models. *Mathematics of Operations Research*, pages 789–822, 39(3) 2014.

[7] Cai, N., Song, Y., Kou, S. A general framework for pricing asian options under markov processes. *Operations Research*, pages 540–554, 63(3) 2015.

[8] Cai, N., Song, Y., Kou, S. A unified framework for options pricing under regime switching models. *working paper*, 2016.

[9] Caldana, R. and Fusai, G. A general closed-form spread option pricing formula. *Journal of Banking and Finance*, pages 4893–4906, 37 2013.

[10] Cass, T. Smooth densities for solutions to stochastic differential equations with jumps. *Stochastic Processes and their Applications*, pages 1416–1435, 119 2009.

[11] Dadachanji, Z. *Fx barrier options: a comprehensive guide for industry quants*. Palgrave Macmillan, 2015.

[12] Delong, L. *Backward SDEs with Jumps and Their Actuarial and Financial Applications*. Springer, 2013.

[13] Di Nunno, G., Oksendal, B., Proske, F. *Malliavin Calculus for Lévy Processes with Applications to Finance*. Springer, 2009.

[14] Eraker, B. Do stock prices and volatility jump? reconciling evidence from spot and option prices. *The Journal of Finance*, pages 1367–1404., LIX(3) 2004.

[15] Forster, B., Lütkebohmert, E., Teichmann, J. Absolutely continuous laws of jump-diffusions in finite and infinite dimensions with applications to mathematical finance. *SIAM Journal of Mathematical Analysis*, pages 2132–2153, 40(5) 2009.

[16] Fusai, G., and Kyriakou, I. General optimized lower and upper bounds for discrete and continuous arithmetic asian options. *Mathematics of Operations Research*, pages 531–559, 41(2) 2016.

[17] Fusai, G., and Meucci, A. Pricing discretely monitored asian options under lévy processes. *Journal of Banking and Finance*, pages 2076–2088, 32 2008.

[18] Glasserman, P. *Monte Carlo Methods in Financial Engineering*. Springer, New York, 2004.

[19] Hagan, P., Kumar, D., Lesniewski, A., and Woodward, D. Managing smile risk. *Wilmott*, pages 84–108, September 2002.

[20] Hayashi, M. Asymptotic expansions for functionals of a poisson random measure. *Journal of Mathematics of Kyoto University*, pages 91–132, 48-1 2008.

[21] Hayashi, M. Coefficients of asymptotic expansions of sde with jumps. *Asia-Pacific Financial Markets*, pages 373–380, 17(4) 2010.

[22] Homescu, C. Local stochastic volatility models: calibration and pricing. *working paper, http://ssrn.com/abstract=2448098*, 2014.

[23] Jacod, J., Kurtz, T.G., Méléard, and Protter, P. The approximate euler method for lévy driven stochastic differential equations. *Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques*, pages 523–558, 41 2005.

[24] Kunitomo, N. and Takahashi, A. The asymptotic expansion approach to the valuation of interest rate contingent claims. *Mathematical Finance*, 11, 117-151, 2001. 2001.

[25] Kunitomo, N., and Takahashi, A. On validity of the asymptotic expansion approach in contingent claim analysis. *Annals of Applied Probability*, 13(3):914–952, 2003.

[26] Li, C., and Chen, D. Estimating jump-diffusions using closed-form likelihood expansions. *Journal of Econometrics*, pages 51–70, 195(1) 2016.

[27] Lipster, R.Sh. and Shiryayev, A.N. *Theory of Martingales*. Kluwer Academic, Dordrecht, 1989.

[28] Mercurio, F. Swaption smile and cms adjustment. *presentation paper, http://www.fabiomercurio.it/RiskDRME.pdf*, 2006.

[29] Pagliarani S., Pascucci A. Local stochastic volatility with jumps. *International Journal of Theoretical and Applied Finance*, 16 2013.

[30] Sepp, A. Achieving consistent modeling of vix and equities derivatives. *presentation paper, Global Derivatives Conference in Barcelona*, 2012.

[31] Shiraya, K. and Takahashi, A. Pricing average and spread options under local-stochastic volatility jump-diffusion models (online appendix). *CARF working paper F-412*, 2017.

[32] Shiraya, K. Takahashi, A. Pricing average options on commodity. *Journal of Futures Markets*, pages 407–439, 31(5) 2011.

[33] Shiraya, K. Takahashi, A. Pricing multi-asset cross currency options. *Journal of Futures Markets*, pages 1–19, 34(1) 2014.

[34] Shiraya, K. Takahashi, A. An approximation formula for basket option prices under local stochastic volatility with jumps, an application to commodity markets. *Journal of Computational and Applied Mathematics*, pages 230–256, 292(15) 2016.

[35] Shiraya, K. Takahashi, A. An asymptotic expansion for local-stochastic volatility with jump models. *Stochastics: An International Journal of Probability and Stochastic Processes*, pages 65–88, 89(1) 2017.

[36] Shiraya, K. Takahashi, A. A general control variate method for multi-dimensional sdes: An application to multi-asset options under local stochastic volatility with jumps models in finance. *European Journal of Operational Research*, pages 358–371, 258(1) 2017.

[37] Shiraya, K., Takahashi, A. and Toda, M. Pricing barrier and average options under stochastic volatility environment. *Journal of Computational Finance*, pages 111–148, 15(2) 2012.

[38] Shiraya, K., Takahashi, A. and Yamazaki, A. Pricing swaptions under the libor market model of interest rates with local-stochastic volatility models. *Willmott Magazine*, pages 48–63, 61 2012.

[39] Takahashi, A. An asymptotic expansion approach to pricing contingent claims. *Asia-Pacific Financial Markets*, 6:115–151, 1999.

[40] Takahashi, A. and Yamada, T. On error estimates for asymptotic expansions with malliavin weights: Application to stochastic volatility model. *Mathematics of Operations Research*, 40(3):513–541, 2015.

[41] Takahashi, A. and Yamada, T. A weak approximation with asymptotic expansion and multidimensional malliavin weights. *Annals of Applied Probability*, 26(2):818–856, 2016.

[42] Takahashi, A., Takehara, K., Toda, M. Computation in an asymptotic expansion method. *working paper F-149*, 2009.

[43] Yamazaki, A. Pricing average options under time-changed lévy processes. *Review of Derivatives Research*, 17(1):79–111, 2014.

[44] Yoshida, N. Conditional expansions and their applications. *Stochastic Processes and their Applications*, 107:53–81, 2003.

[45] Zeng, P., and Kwok, Y.K. Pricing bounds and approximation for discrete arithmetic asian options under time-changed lévy processes. *Quantitative Finance*, 16(9):1375–1391, 2016.

A   Propositions used in Section 3

This appendix shows two propositions with the same notations as in Section 3, that are used for the proof of Theorem 3.2. Firstly, we introduce a version of Theorem 2 in Yoshida [44], which is simplified to fit our purpose. The notations are the same as those in Section 3.

Proposition A.1. *Suppose that the d-dimensional sequence $(F_\epsilon)_{\epsilon \in (0,1]}$ has a smooth asymptotic expansion:*

$$F_\epsilon \sim f_0 + \epsilon f_1 + \epsilon^2 f_2 + \cdots , (\epsilon \downarrow 0). \tag{81}$$

Suppose also that for every $p > 1$,

$$\limsup_{\epsilon \downarrow 0} \mathbf{E}[\Delta_{F_\epsilon}^{-p}] < \infty, \tag{82}$$

where $\Delta_{F_\epsilon} := \det \sigma_{F_\epsilon}$ and σ_{F_ϵ} denotes the Malliavin covariance of F_ϵ. Then, for every $g \in \mathcal{S}'(\mathbf{R}^D)$, $g(F_\epsilon)$ has an asymptotic expansion:

$$\mathbf{E}[g(F_\epsilon)] \sim \mathbf{E}[\Phi_0] + \epsilon \mathbf{E}[\Phi_1] + \epsilon^2 \mathbf{E}[\Phi_2] + \cdots , \tag{83}$$

where Φ_i are determined by the formal Taylor expansion of $g(F_\epsilon)$ around f_0. In particular,

$$\Phi_0 = g(F_0), \tag{84}$$

$$\Phi_1 = \sum_{a=1}^{D} \partial_a g(F_0) f_1^{(a)}. \tag{85}$$

where $f_i^{(a)}$ is the a-th element of f_i.

We remark that F_ϵ has a smooth asymptotic expansion in a sense that for every $p > 1$ and $n \in \mathbf{Z}_+$, $F_\epsilon - \sum_{i=0}^{n} \epsilon^i f_i = O(\epsilon^{n+1})$, that is,

$$\limsup_{\epsilon \downarrow 0} \frac{\left\| F_\epsilon - \sum_{i=0}^{n} \epsilon^i f_i \right\|_{L^p}}{\epsilon^{(n+1)}} < \infty.$$

Moreover, in order to evaluate the coefficients in the expansion, $\mathbf{E}[\Phi_i]$ ($\mathbf{E}[f_{i,T}]$ in our case), we apply the next proposition, a specific version of Proposition 3 in [44] to fit our setup.

Proposition A.2. *Suppose that F is smooth in a Malliavin sense, and*

$$\mathbf{E}\left[\Delta_F^{-p}\right] < \infty, \ \forall p > 1. \tag{86}$$

Then, F has a smooth density p^F. Moreover, for $g \in \mathcal{F}_\uparrow(\mathbf{R}^D)$ and a smooth G in a Malliavin sense, where $\mathcal{F}_\uparrow(\mathbf{R}^D)$ being the set of measurable functions on $\mathbf{R}^D$ of at most polynomial growth,

$$\mathbf{E}[\partial^n g(F)G] = \int_{\mathbf{R}^D} g(z)(-\partial)^n \{\mathbf{E}[G|F = z]p^F(z)\}dz \ (n \in \mathbf{Z}_+^D). \tag{87}$$

We note that in our situation, smooth F and G in a Malliavin sense mean that F and G are Malliavin-differentiable in the Wiener direction up to any orders and the derivatives have finite L_p-moments for every $p > 1$.

B   Proofs of Theorem 3.2 and Corollary 3.3

This appendix gives proofs of Theorem 3.2 and Corollary 3.3, where notations are the same as those in Section 3.

Proof of Theorem 3.2. Since Condition (A) is satisfied, $\mathscr{U}_T^{(\epsilon)}$ has a smooth asymptotic expansion by Proposition C.4. By this result with Condition (B) and (C), the conditions in Proposition A.1 are satisfied. Moreover, in an application of Proposition A.1, the validity of the expansion is proved through an integration-by-parts(IBP) formula to have an expansion for the product of the so called Malliavin weight (e.g. Takahashi - Yamada [40], [41]) with smooth asymptotic expansion and a composite functional of a smooth function with $\mathscr{U}_T^{(\epsilon)}$. (Please see the proof for Theorem 3 in [44] for the details.) Then, with an application of Proposition C.4 in Appendix C to the current setup, $\mathbf{E}[f(\mathscr{U}_T^{(\epsilon)})]$ has an expansion:

$$\mathbf{E}[f(\mathscr{U}_T^{(\epsilon)})] = \mathbf{E}\left[f_{0,T} + \epsilon f_{1,T} + \epsilon^2 f_{2,T} + \cdots + \epsilon^M f_{M,T}\right] + O(\epsilon^{M+1}C(T)). \tag{88}$$

In addition, Condition (C) ensures that $\mathbf{E}[f_{n,T}]$ has the following expression by Proposition A.2:

$$
\begin{aligned}
\mathbf{E}[f_{n,T}] &= \mathbf{E}\left[\frac{\partial^n}{\partial \epsilon^n} f(\mathscr{U}_T^{(\epsilon)})\Big|_{\epsilon=0}\right] \\
&= \sum_{n^{(r)}}^{(n)} \int_{\mathbf{R}} f(x)(-\partial)^r \left(\mathbf{E}\left[\mathscr{U}_T^{(n_1)} \cdots \mathscr{U}_T^{(n_r)} | \mathscr{U}_T^{(0)} = x\right] p^{\mathscr{U}_T^{(0)}}(x)\right) dx.
\end{aligned} \tag{89}
$$

$\square$

Proof of Corollary 3.2. The probability of $\{N_l = k_l\}$ is expressed as $p_{\{k_l\}} = \prod_{l=1}^n \frac{(\Lambda_l T)^{k_l} e^{-\Lambda_l T}}{k_l!}$, which is the product of the k_l times of the jump probabilities of $N_{l,T}$ $(l = 1, \cdots, n)$, that is $\prod_{l=1}^n P(\{N_{l,T} = k_l\})$, thanks to the independence of $N_{l,T}$ $(l = 1, \cdots, n)$. Thus, $\mathbf{E}[f(\hat{\mathscr{U}}_T^{(\epsilon)})]$ is expressed as

$$\mathbf{E}[f(\hat{\mathscr{U}}_T^{(\epsilon)})] = \sum_{k=0}^\infty \sum_{\sum_{l=1}^n k_l = k} p_{\{k_l\}} \mathbf{E}[f(\hat{\mathscr{U}}_T^{(\epsilon)})|\{N_{l,T} = k_l\}]$$

$$= \sum_{k=0}^{\infty} \sum_{\sum_{l=1}^{n} k_l = k} p_{\{k_l\}} \left(\prod_{l=1}^{n} \frac{k_l!}{T^{k_l}} \right) \int_0^{\vec{T}} \mathbf{E}[f(\hat{\mathscr{U}}_T^{(\epsilon)})|\{N_{l,T} = k_l\}, \{\tau_l = \vec{t_l}\}]\vec{dt}. \tag{90}$$

Here, we have used that the joint density of $(\tau_{l,1}, \cdots, \tau_{l,k_l})$ conditioned on $N_{l,T} = k_l$ is given by $g(t_{l,1}, \cdots, t_{l,k_l}|N_{l,T} = k_l) = \frac{k_l!}{T^{k_l}}$ $(0 < t_{l,1} < \cdots < t_{l,k_l} \le T)$, and hence the joint density $\hat{g}$ of $(\tau_{1,1}, \cdots, \tau_{n,k_n})$ conditioned on $\{N_{l,T} = k_l\}$ is provided as:

$$\hat{g}(t_{1,1}, \cdots, t_{n,k_n}|\{N_{l,T} = k_l\}) = \prod_{l=1}^{n} \frac{k_l!}{T^{k_l}}. \tag{91}$$

Moreover, because of Condition (A) and (B), the same argument as in Theorem 3.2 derive next expressions:

$$\mathbf{E}[f(\hat{\mathscr{U}}_T^{(\epsilon)})|\{N_{l,T} = k_l\}, \{\tau_l = \vec{t_l}\}]$$
$$= \mathbf{E}\left[\tilde{f}_{0,T,\{k_l\}} + \epsilon\tilde{f}_{1,T,\{k_l\}} + \epsilon^2 \tilde{f}_{2,T,\{k_l\}} + \cdots + \epsilon^M \tilde{f}_{M,T,\{k_l\}} \} \right] + O(\epsilon^{M+1}C(T)), \tag{92}$$

$$\mathbf{E}[\tilde{f}_{0,T\{k_l\}}] = \int_{\mathbf{R}} f(x) p^{\hat{\mathscr{U}}_{\{\tau_l = \vec{t_l}\},T}^{(0)}}(x)dx, \tag{93}$$

$$\mathbf{E}[\tilde{f}_{n,T\{k_l\}}] = \sum_{n^{(r)},\alpha^{(r)}}^{(n)} \int_{\mathbf{R}} f(x)(-\partial_{\alpha^{(r)}})^n$$
$$\times \left(\mathbf{E}\left[\mathcal{X}_T^{\alpha_1,(n_1)} \cdots \mathcal{X}_T^{\alpha_r,(n_r)} | \hat{\mathscr{U}}_{\{\tau_l = \vec{t_l}\},T}^{(0)} = x, \{N_{l,T} = k_l\}, \{\tau_l = \vec{t_l}\} \right] p^{\hat{\mathscr{U}}_{\{\tau_l = \vec{t_l}\},T}^{(0)}}(x) \right) dx, \tag{94}$$

where $p^{\hat{\mathscr{U}}_{\{\tau_l = \vec{t_l}\},T}^{(0)}}$ is a density function of $\hat{\mathscr{U}}_{\{\tau_l = \vec{t_l}\},T}^{(0)}$, and $\hat{\mathscr{U}}_{\{\tau_l = \vec{t_l}\},T}^{(0)}$ stands for $\hat{\mathscr{U}}_T^{(0)}$ conditioned on jumps at $\{\tau_l = \vec{t_l}\}$ with $\{N_{l,T} = k_l\}$.

Finally, we obtain the lemma by putting $\mathbf{E}[f_{n,T}]$ as

$$\mathbf{E}[f_{n,T}] = \sum_{k=0}^{\infty} \sum_{\sum_{l=1}^{n} k_l = k} p_{\{k_l\}} \left(\prod_{l=1}^{n} \frac{k_l!}{T^{k_l}} \right) \int_0^{\vec{T}} \mathbf{E}[\tilde{f}_{n,T,\{k_l\}}]\vec{dt}. \tag{95}$$

□

C   Error Estimates of Asymptotic Expansions

This appendix discusses error estimates of our asymptotic expansions, which are obtained from moment estimates for some graded stochastic differential equations(SDEs). In fact, an asymptotic expansion of an SDE creates a new system of SDEs, whose coefficients are not globally Lipschitz, but have a "Lipschitz lower triangular structure" (p.45 in Bichteler et al.[4]).

First, we consider a I-dimensional stochastic differential equation of the form: for $t \in [0, T]$,

$$Y_t^\epsilon = \hat{y}_t + \int_0^t A(s, Y_{s-}^\epsilon)ds + \epsilon \sum_{j=1}^K \int_0^t B_j(s, Y_{s-}^\epsilon)dW_s^j + \sum_{l=1}^k \int_0^t \int_{E_l} \Gamma_l(\epsilon, s, z, Y_{s-}^\epsilon)\tilde{N}_l(ds, dz),$$

(96)

where $\epsilon \in (0, 1]$, $\hat{y}_t$ is $\mathbf{R}^I$-valued càdlàg process, and $A : [0, T] \times \mathbf{R}^I \times \Omega \to \mathbf{R}^I$, $B_j : [0, T] \times \mathbf{R}^I \times \Omega \to \mathbf{R}^I$, $\Gamma_l : (0, 1] \times [0, T] \times E_l \times \mathbf{R}^I \times \Omega \to \mathbf{R}^I$ are predictable processes, and W is a K-dimensional Brownian motion. Each $N_l(dt, dz)$ is a Poisson random measure on $[0, T] \times E_l$, where $(E_l, \underline{E}_l)$ is a measurable space with $E_l \subset \mathbf{R}^{e_l}$, $e_l \in \mathbf{N}$, and the intensity measure of N_l is $dt \times \nu_l(z)$, where $\nu_l(z)$ is a positive σ-finite measure on $(E_l, \underline{E}_l)$. Then, $\tilde{N}_l(dt, dz) := N_l(dt, dz) - dt \times \nu_l(z)$ is a compensated Poisson random measure. Also, a norm $|||Z_T|||_p$ is defined as

$$|||Z_T|||_p = \left(\int_0^T \mathbf{E}[|Z_t|^p]dt \right)^{\frac{1}{p}}.$$

Then, let us provide the following definition for graded SDEs from Definition 5-5 [4].

Definition C.1. *A grading of $\mathbf{R}^I$ is a decomposition $\mathbf{R}^I = \mathbf{R}^{I_1} \times \cdots \times \mathbf{R}^{I_q}$ with $I = I_1 + \cdots + I_q$. The coordinates of a point in $\mathbf{R}^I$ are always arranged in an increasing order along the subspace $\mathbf{R}^{I_i}$, and set $M_0 = 0$ and $M_m = I_1 + \cdots + I_m$ for $1 \leq m \leq q$. The coefficients A, B and Γ are graded according to the grading, that is, $\mathbf{R}^I = \mathbf{R}^{I_1} \times \cdots \times \mathbf{R}^{I_q}$ if $A_i(t, y)$, $B_{j,i}(t, y)$, $j = 1, \cdots, K$ and $\Gamma_{l,i}(\epsilon, t, z, y)$, which are i-th element of A, B_j and Γ_l, depend upon only through the coordinates $(y_k)_{1 \leq k \leq M_m}$ when $M_{m-1} < i \leq M_m$.*

We note that when we define Π_m as an orthogonal projection on $\mathbf{R}^{I_1} \times \cdots \times \mathbf{R}^{I_m}$, graded coefficients mean that

$$\Pi_m A(t, y) = \Pi_m A(t, \Pi_m y),$$
$$\Pi_m B_j(t, y) = \Pi_m B_j(t, \Pi_m y),$$
$$\Pi_m \Gamma_l(\epsilon, t, z, y) = \Pi_m \Gamma_l(\epsilon, t, z, \Pi_m y).$$

Moreover, we put the next assumptions for the coefficients in the SDE. (Hereafter, in order to avoid the notational complexity we may omit ϵ from the notation, and may use the notation, $|X|_T^*$ defined by $|X|_T^* := \sup_{0 \leq t \leq T} |X_t|$.)

Assumption C.2. *the coefficients are graded according to $\mathbf{R}^I = \mathbf{R}^{I_1} \times \cdots \times \mathbf{R}^{I_q}$ in a sense of **Definition** above. Moreover, let $F(t, z, y) = A(t, y)$, or $B_j(t, y)$ (j-th column of the matrix $B(t, y)$, $j = 1, \cdots, K$), or $\frac{\Gamma_l(\epsilon, t, z, y)}{\eta_l(\epsilon, z)}$ with $\eta_l(\epsilon, z) \in \cap_{2 \leq p < \infty} L^p(E_l, \nu_l)$. We assume F is differentiable in y on R^I, and*

1. $|F_i(t, z, 0)| \leq Z_t$ *for $i = 1, \cdots, I$,*

2. $|\frac{\partial}{\partial y_j} F_i(t, z, y)| \leq \zeta$ *if $M_{m-1} < i, j \leq M_m$ for some $m \leq q$,*

   i.e. $|\frac{\partial}{\partial y_j} F_i(t, z, y)| \leq \zeta$ for $F_i \in Q_m F$ and $y_j \in Q_m y$ with $Q_m := \Pi_m - \Pi_{m-1}$,

3. $|\frac{\partial}{\partial y}F_i(t,z,y)| \leq \hat{Z}_t(1+|y|^\theta)$ for all i,

where F_i is the i-th element in F, $\zeta, \theta \geq 0$, and Z, $\hat{Z}$ are $\mathbf{R}$-valued predictable processes such that $|||Z_T|||_p$, $|||\hat{Z}_T|||_p$ are finite for $\forall p \in [0, \infty)$.

Then, the existence of the unique solution and its L^p integrability for $\forall p > 1$ to the SDE (96) are proved in a similar way of Theorem 5-10 in [4]. Thus, we concentrate to show (103) below with specifying the associated constant terms explicitly.

Firstly, by applying Burkholder-Davis-Gundy inequalities (e.g. for the Brownian integral, Lemma 11 in p.79 of Lipster - Shiryayev [27] and for the compensated Poisson integral, Lemma 7 in p.77 with Corollary 1 in p.76 of [27]), a slight modification of Lemma 5.1 in [23] and Hölder's inequalities, we can show the following useful lemma, which is applied to obtain the error estimate of our asymptotic expansion method.

Lemma C.3. *For $p \geq 2$, we assume the above conditions 1, 2 and 3. Then,*

$$\mathbf{E}\left[\sup_{0 \leq t \leq T}\left|\int_0^t A(s, Y_{s-}^\epsilon)ds\right|^p\right] \leq c_A(p)\int_0^T \mathbf{E}\left[|A(s, Y_{s-}^\epsilon)|^p\right]ds, \tag{97}$$

$$\mathbf{E}\left[\sup_{0 \leq t \leq T}\left|\int_0^t B_j(s, Y_{s-}^\epsilon)dW_s^j\right|^p\right] \leq c_B(p)\int_0^T \mathbf{E}\left[|B_j(s, Y_{s-}^\epsilon)|^p\right]ds, \tag{98}$$

$$\mathbf{E}\left[\sup_{0 \leq t \leq T}\left|\int_0^t \int_{E_l}\Gamma_l(\epsilon, s, z, Y_{s-}^\epsilon)\tilde{N}_l(ds, dz)\right|^p\right] \leq c_{\Gamma_l}(p)\int_0^T \mathbf{E}\left[\left|\frac{\Gamma_l(\epsilon, s, z, Y_{s-}^\epsilon)}{\eta_l(\epsilon, z)}\right|^p\right]ds, \tag{99}$$

where

$$c_A(p) = T^{p-1}, \tag{100}$$

$$c_B(p) = \left(\frac{p^{p+1}}{2(p-1)^{p-1}}\right)^{\frac{p}{2}}T^{\frac{p}{2}-1}, \tag{101}$$

$$c_{\Gamma_l}(p) = 2^{2p-1}(1+p^p)\left(2^{\frac{p^2}{4}}\left(\int_{E_l}|\eta_l(\epsilon, z)|^2\nu_l(dz)\right)^{\frac{p}{2}} + 2^{\frac{p}{2}}\int_{E_l}|\eta_l(\epsilon, z)|^p\nu_l(dz)\right). \tag{102}$$

Then, under Assumption C.2, we apply this lemma and a Gronwall's inequality with an induction argument for the grading, $m = 1, \cdots, q$ by defining $\Pi_0 y = 0$ to show the next proposition.

Proposition C.4. *Under Assumption C.2 with $|\hat{y}|_T^* \in \cap_{p<\infty}L^p$, (96) has a unique solution Y^ϵ, and for every $p \geq 1$, there exist a constant ι_p and an increasing function of T, $c_p(T)$ such that*

$$\big\| |Y^\epsilon|_T^* \big\|_{L^p} \leq c_p(T)(\big\| |\hat{y}|_T^* \big\|_{L^{\iota_p}} + |||Z_T|||_{\iota_p}). \tag{103}$$

Here, we set $c_p(T)$ as

$$c_p(T) = \delta_p^{\frac{1}{p}}\exp(\zeta^p I^p \delta_p T/p),$$

38

where I is the dimension of Y and

$$\delta_p = 4^{p-1}\max\left\{1, \left(c_A(p) + K^{\frac{p}{2}}c_B(p) + k^{\frac{p}{2}-1}\sum_{l=1}^{k}c_{\Gamma_l}(p)\right)2^{p-1}\right\}, \qquad (104)$$

with $c_A(p)$, $c_B(p)$ $c_{\Gamma_l}(p)$ given by 100), (101) and (102), respectively.

In particular, for $p = 2$(L^2-norm) since it is not necessary to use Burkholder-Davis-Gundy inequalities, those are simplified as follows:

$$c_2(T) = \sqrt{\delta_2}\exp(\zeta^2 I^2\delta_2 T/2),$$

$$\delta_2 = 4\max\left\{1, \left(c_A(2) + c_B(2)K + \sum_{l=1}^{k}c_{\Gamma_l}(2)\right)2\right\} \qquad (105)$$

$$= 8\left(T + K + \sum_{l=1}^{k}\int_{E_l}\eta_l(\epsilon, z)^2\nu_l(dz)\right), \quad (K \geq 1, k \geq 1),$$

$$c_A(2) = T, \ c_B(2) = 1, \ c_{\Gamma_l}(2) = \int_{E_l}\eta_l(\epsilon, z)^2\nu_l(dz).$$

Next, let us proceed to a general scheme for error estimates for asymptotic expansions. Firstly, since $\Pi_{m-1}A(t, \Pi_{m-1}y) = \Pi_{m-1}A(t, \Pi_m y)$ and hence (with $Q_m = \Pi_m - \Pi_{m-1}$),

$$Q_m A(t, y) = \Pi_m A(t, \Pi_m y) - \Pi_{m-1}A(t, \Pi_{m-1}y) = Q_m A(t, \Pi_m y),$$

similarly for B_j and Γ_l, the SDE corresponding to the m-th grade is expressed as follows:

$$Q_m Y_t^\epsilon = Q_m \hat{y}_t + \int_0^t \bar{A}^m(s, Q_m Y_{s-}^\epsilon)ds + \epsilon\sum_{j=1}^{K}\int_0^t \bar{B}_j^m(s, Q_m Y_{s-}^\epsilon)dW_s^j$$

$$+ \sum_{l=1}^{k}\int_0^t\int_{E_l}\bar{\Gamma}_l^m(\epsilon, s, z, Q_m Y_{s-}^\epsilon)\tilde{N}_l(ds, dz),$$

where $\bar{A}^m(t, x) := Q_m A(t, x + \Pi_{m-1}Y_t)$, $\bar{B}_j^m(t, x) := Q_m B_j(t, x + \Pi_{m-1}Y_t)$ and $\bar{\Gamma}_l^m(\epsilon, t, z, x) := Q_m\Gamma_l(\epsilon, t, z, x + \Pi_{m-1}Y_t)$.

Then, we slightly elaborate Assumption C.2 as the one for the coefficients in each graded SDE as follows:

Assumption C.2.'

1. $|F_i^m(t, z, 0)| \leq Z_{m,t}$,

2. $\left|\dfrac{\partial}{\partial y_j}F_i^m(t, z, y)\right| \leq \zeta_m, \ y_j \in Q_m y$,

3. $\left|\dfrac{\partial}{\partial y}F_i^m(t, z, \Pi_m y)\right| \leq \hat{Z}_{m,t}\left(1 + |\Pi_m y|^{\theta_m}\right)$,

where $\zeta_m, \theta_m \geq 0$, and $Z_m, \hat{Z}_m$ are $\mathbf{R}$-valued predictable processes such that $|||Z_{m,T}|||_p$, $|||\hat{Z}_{m,T}|||_p$ are finite for $\forall p \in [0, \infty)$. Here, we use a similar notation as before: $F^m(t, z, y) = Q_m A(t, y)$, $Q_m B_j(t, y)$ or $\frac{Q_m\Gamma_l(\epsilon, t, z, y)}{\eta_l(\epsilon, z)}$, and F_i^m denotes the i-th element in F^m.

39

Next, because $\bar{A}^m(t,0) = Q_m A(t, \Pi_{m-1}Y_t)$, $\bar{B}_j^m(t,0) = Q_m B_j(t, \Pi_{m-1}Y_t)$ and $\bar{\Gamma}_l^m(\epsilon, t, z, 0) = Q_m \Gamma_l(\epsilon, t, z, \Pi_{m-1}Y_t)$, we have

$$\left| \bar{F}_i^m(\epsilon, t, z, 0) \right| \leq |F_i^m(\epsilon, t, z, 0)| + \left| \frac{\partial}{\partial y} F_i^m(\epsilon, t, z, \Pi_{m-1}y) \right| |\Pi_{m-1}y|$$

$$\leq Z_{m,t} + \hat{Z}_{m,t} \left(1 + |\Pi_{m-1}y|^{\theta_m} \right) |\Pi_{m-1}y|.$$

Hence, by setting $Q_0 Y = \Pi_0 Y \equiv 0$ and applying above estimates with Hölder's inequality, the similar argument as in deriving Proposition C.4 gives the following recursive inequalities:

$$\text{For } m = 1, \ \left\| \, |Q_1 Y|_T^* \, \right\|_{L^p} = \left\| \, |\Pi_1 Y|_T^* \, \right\|_{L^p} \leq c_{1,p}(T) \left(\left\| \, |\Pi_1 \hat{y}|_T^* \, \right\|_{L^p} + |||Z_{1,T}|||_p \right),$$
$$\text{For } m \geq 2, \ \left\| \, |\Pi_m Y|_T^* \, \right\|_{L^p} \leq \left\| \, |\Pi_{m-1} Y|_T^* \, \right\|_{L^p} + \left\| \, |Q_m Y|_T^* \, \right\|_{L^p}, \tag{106}$$

with

$$\left\| \, |Q_m Y|_T^* \, \right\|_{L^p}$$
$$\leq c_{m,p}(T) \Bigg\{ \left(\left\| \, |Q_m \hat{y}|_T^* \, \right\|_{L^p} + |||Z_{m,T}|||_p \right) + |||\hat{Z}_{m,T}|||_{2p} \left(\left\| \, |\Pi_{m-1} Y|_T^* \, \right\|_{L^{2p}} + \left\| \, |\Pi_{m-1} Y|_T^* \, \right\|_{L^{2p(\theta_m+1)}} \right) \Bigg\}, \tag{107}$$

where for $m = 1, \cdots, q$,

$$c_{m,p}(T) = \delta_{m,p}^{\frac{1}{p}} \exp(\zeta_m^p I_m^p \delta_{m,p} T/p),$$

with some $\delta_{m,p}$, similarly given by (104) or (105).

Moreover, for the case of the graded SDE created by an asymptotic expansion of $X_T^{(\epsilon)}$ in Section 2 (and $g(X_T^{(\epsilon)})$ in Section 3) with respect to ϵ, all the initial values for the SDEs regarding $\frac{\partial^k X_T^{(\epsilon)}}{\partial \epsilon^k}$ with $k \geq 1$ are zero, and hence, $|Q_m \hat{y}|_T^* = 0$ for all m except for the ones corresponding to the original underlying SDEs for $X_T^{(\epsilon)}$. Thus, for all m related to the expansion and its error estimates (i.e. $\frac{1}{k!} \frac{\partial^k X_T^{(\epsilon)}}{\partial \epsilon^k}$, $\frac{1}{k!} \frac{\partial^k g(X_T^{(\epsilon)})}{\partial \epsilon^k}$, $k \geq 1$), we replace (107) by the following equation:

$$\left\| \, |Q_m Y|_T^* \, \right\|_{L^p} \tag{108}$$
$$\leq c_{m,p}(T) \left\{ |||Z_{m,T}|||_p + |||\hat{Z}_{m,T}|||_{2p} \left(\left\| \, |\Pi_{m-1} Y|_T^* \, \right\|_{L^{2p}} + \left\| \, |\Pi_{m-1} Y|_T^* \, \right\|_{L^{2p(\theta+1)}} \right) \right\}.$$

Consequently, the recursive inequalities (108) and (106) provide a general scheme to obtain error estimates for asymptotic expansions. Further, given particular forms of the coefficients in the graded SDEs (e.g.(7), (17)) arising from an asymptotic expansion and the assumptions (e.g. Condition (A), a smooth function g with all derivatives of polynomial growth orders) for the original SDE for $X_t^{(\epsilon)}$ (and $g(X_t^{(\epsilon)})$), we are able to express those inequalities more explicitly. (Please see an example below for the detail.)

Finally, in order to obtain a more explicit form of an error estimate for an asymptotic expansion, which depends on the functional properties of the associated graded SDEs' coefficients,

such as their values at $Y = 0$ and the bounds with growth rates of their derivatives, let us consider the following one-dimensional local volatility model as an example:

$$S_t^{(\epsilon)} = s_0 + \int_0^t \epsilon \Phi\left(S_s^{(\epsilon)}\right) dW_s + \int_{\mathbf{R}} \Psi\left(S_s^{(\epsilon)}\right) (e^{\epsilon z} - 1) \tilde{N}(ds, dz), \tag{109}$$

with $\epsilon \in (0, 1)$. (Here, we start with a perturbed SDE corresponding to (3) in the general case.)

Let us rewrite the equation by introducing a new perturbation parameter $\epsilon_1 > \epsilon$ with $\alpha := \frac{\epsilon}{\epsilon_1} \in (0, 1)$, and expand the equation around $\epsilon_1 = 0$:

$$S_t^{(\epsilon_1)} = s_0 + \int_0^t \epsilon_1 \alpha \Phi\left(S_s^{(\epsilon_1)}\right) dW_s + \int_{\mathbf{R}} \alpha \Psi\left(S_s^{(\epsilon_1)}\right) \frac{e^{\epsilon_1 \alpha z} - 1}{\alpha} \tilde{N}(ds, dz). \tag{110}$$

Here, we suppose that $\Phi(t, x), \frac{\Psi(t,x)(e^{\epsilon_1 z} - 1)}{\alpha \eta(\epsilon_1, \alpha, z)} \in C_b^\infty(x)$ with $\eta(\epsilon_1, \alpha, z) = \max\left\{\left|\frac{e^{\epsilon_1 \alpha z} - 1}{\alpha}\right|, \left|e^{\epsilon_1 \alpha z} z\right|, \left|e^{\epsilon_1 \alpha z} z^2\right|\right\}$. In this example, we use an abbreviated notation, $\eta(\epsilon_1, \alpha, z) \equiv \eta_1(\epsilon_1, \alpha, z)$.

Let us remark that this transformation includes the one in numerical examples as in Section 5, where we frequently put $\epsilon_1 = 1$ for $(\epsilon_1)^k$ in expansions, that is,

$$S_T^{(\epsilon_1)} \sim S_T^{(0)} + \epsilon_1 S_T^{(1)} + \frac{\epsilon_1^2}{2!} S_T^{(2)} + \cdots = S_T^{(0)} + S_T^{(1)} + \frac{1}{2!} S_T^{(2)} + \cdots, \tag{111}$$

with $S_T^{(k)} = \left.\frac{\partial^k S_T^{(\epsilon_1)}}{\partial \epsilon_1^k}\right|_{\epsilon_1 = 0}$.

As an example, in order to evaluate the error of the first order expansion of $S_T^{(\epsilon)}$, we set $Y_t = (Y_{1,t}, Y_{2,t}, Y_{3,t})$, where Y satisfy the following graded SDE:

$$
\begin{aligned}
Y_{1,t} &= S_t^{(\epsilon_1)} \\
&= s_0 + \int_0^t \epsilon_1 \alpha \Phi(Y_{1,s}) dW_s + \int_0^t \int_{\mathbf{R}} \alpha \Psi(Y_{1,s}) \frac{e^{\epsilon_1 \alpha z} - 1}{\alpha} \tilde{N}(ds, dz), \\
Y_{2,t} &= \frac{\partial}{\partial \epsilon_1} S_t^{(\epsilon_1)} \\
&= \int_0^t \alpha \Phi(Y_{1,s}) dW_s + \epsilon_1 \alpha \int_0^t \Phi'(Y_{1,s}) Y_{2,s} dW_s \\
&\quad + \int_0^t \int_{\mathbf{R}} \alpha \Psi(Y_{1,s}) e^{\epsilon_1 \alpha z} z \tilde{N}(ds, dz) + \int_0^t \int_{\mathbf{R}} \alpha \Psi'(Y_{1,s}) Y_{2,s} \frac{e^{\epsilon_1 \alpha z} - 1}{\alpha} \tilde{N}(ds, dz), \\
Y_{3,t} &= \frac{\partial^2}{\partial \epsilon_1^2} S_t^{(\epsilon_1)} \\
&= 2 \int_0^t \alpha \Phi'(Y_{1,s}) Y_{2,s} dW_s + \epsilon_1 \alpha \int_0^t \Phi''(Y_{1,s}) (Y_{2,s})^2 dW_s \\
&\quad + \epsilon_1 \alpha \int_0^t \Phi'(Y_{1,s}) Y_{3,s} dW_s + \int_0^t \int_{\mathbf{R}} \alpha^2 \Psi(Y_{1,s}) e^{\epsilon_1 \alpha z} z^2 \tilde{N}(ds, dz) \\
&\quad + 2 \int_0^t \int_{\mathbf{R}} \alpha \Psi'(Y_{1,s}) Y_{2,s} e^{\epsilon_1 \alpha z} z \tilde{N}(ds, dz) + \int_0^t \int_{\mathbf{R}} \alpha \Psi''(Y_{1,s}) (Y_{2,s})^2 \frac{e^{\epsilon_1 \alpha z} - 1}{\alpha} \tilde{N}(ds, dz) \\
&\quad + \int_0^t \int_{\mathbf{R}} \alpha \Psi'(Y_{1,s}) Y_{3,s} \frac{e^{\epsilon_1 \alpha z} - 1}{\alpha} \tilde{N}(ds, dz).
\end{aligned}
$$

41

We need to evaluate $\frac{\epsilon_1^2}{2} \left\| \, |Y_3|_T^* \, \right\|_{L^p}$ as an error estimate for the first order asymptotic expansion.

First, let us note that $I = 3$, $I_i = 1$, $i = 1, 2, 3$ and $K = k = 1$. In addition, $\hat{y} = (s_0, 0, 0)'$ and

$$
\begin{aligned}
Q_1 Y &= \Pi_1 Y_t = Y_{1,t}, \\
Q_2 Y &= (\Pi_2 - \Pi_1) Y_t = Y_{2,t}, \\
Q_3 Y &= (\Pi_3 - \Pi_2) Y_t = Y_{3,t}.
\end{aligned}
$$

Then as the error estimate, we obtain $\frac{\epsilon_1^2}{2} \left\| \, |Y_3|_T^* \, \right\|_{L^p}$ by using the lower order estimates, recursively as follows.

Firstly, $\left\| \, |Y_1|_T^* \, \right\|_{L^p}$ is given by

$$
\left\| \, |Y_1|_T^* \, \right\|_{L^p} \le c_{1,p}(T) \left[|s_0| + |||Z_{1,T}|||_p \right] = c_{1,p}(T) \left[|s_0| + T^{\frac{1}{p}} Z_1 \right], \tag{112}
$$

where

$$
c_{1,p}(T) = \delta_{1,p}^{\frac{1}{p}} \exp\left(\frac{\delta_{1,p} \zeta_1^p T}{p} \right), \; \delta_{1,p} = 3^{p-1} \max\{1, [c_B(p) + c_\Gamma(p)] 2^{p-1}\},
$$

$$
\zeta_1 = \alpha \max\{\epsilon_1 \zeta_{1,\Phi}, \zeta_{1,\Psi}\}, \; \zeta_{1,\Phi} = \sup_{y_1} |\Phi'(y_1)|, \; \zeta_{1,\Psi} = \sup_{y_1} |\Psi'(y_1)|,
$$

$$
Z_{1,t} = Z_1 = \alpha \max\{\epsilon_1 |\Phi(0)|, |\Psi(0)|\}, \; |||Z_{1,T}|||_p = T^{\frac{1}{p}} Z_1.
$$

Then, $\left\| \, |Y_2|_T^* \, \right\|_{L^p}$ is shown as

$$
\begin{aligned}
\left\| \, |Y_2|_T^* \, \right\|_{L^p} &\le c_{2,p}(T) 2^{1-\frac{1}{p}} \left[|||Z_{2,T}|||_p + \alpha T^{\frac{1}{p}} \hat{\zeta}_1 \left\| \, |Y_1|_T^* \, \right\|_{L^p} \right] \\
&= \alpha c_{2,p}(T) 2^{1-\frac{1}{p}} T^{\frac{1}{p}} \left[\max\{|\Phi(0)|, |\Psi(0)|\} + \hat{\zeta}_1 \left\| \, |Y_1|_T^* \, \right\|_{L^p} \right], \tag{113}
\end{aligned}
$$

where

$$
c_{2,p}(T) = \delta_{2,p}^{\frac{1}{p}} \exp\left(\frac{\delta_{2,p} \zeta_2^p T}{p} \right), \; \delta_{2,p} = 4^{p-1} \{c_B(p) + c_\Gamma(p)\}, \; \hat{\zeta}_1 = \max\{\zeta_{1,\Phi}, \zeta_{1,\Psi}\},
$$

$$
\zeta_2 = \alpha \max\{\epsilon_1 \zeta_{2,\Phi}, \zeta_{2,\Psi}\}, \; \zeta_{2,\Phi} = \sup_{y_1} |\Phi''(y_1)|, \; \zeta_{2,\Psi} = \sup_{y_1} |\Psi''(y_1)|,
$$

$$
Z_{2,t} = Z_2 = \alpha \max\{|\Phi(0)|, |\Psi(0)|\}, \; |||Z_{2,T}|||_p = T^{\frac{1}{p}} Z_2.
$$

Finally, we obtain an error estimate for the first order expansion (e.g. with $\epsilon_1 = 1$, i.e. $\alpha = \epsilon$) as:

$$
\begin{aligned}
\frac{\epsilon_1^2}{2} \left\| \, |Y_3|_T^* \, \right\|_{L^p} &\le \epsilon_1^2 c_{3,p}(T) \left[\alpha \hat{\zeta}_1 T^{\frac{1}{p}} \left\| \, |Y_2|_T^* \, \right\|_{L^p} + \frac{\alpha}{2} \hat{\zeta}_2 T^{\frac{1}{p}} \left\| \, |Y_2|_T^* \, \right\|_{L^p}^2 \right. \\
&\quad \left. + \frac{1}{2} \left(\frac{8^{p-1} c_\Gamma(p)}{\hat{\delta}_{3,p}} \right)^{\frac{1}{p}} |||Z_{3,T}|||_p + \frac{\alpha^2}{2} \left(\frac{8^{p-1} c_\Gamma(p)}{\hat{\delta}_{3,p}} \right)^{\frac{1}{p}} \zeta_{1,\Psi} T^{\frac{1}{p}} \left\| \, |Y_1|_T^* \, \right\|_{L^p} \right] \\
&= \alpha \epsilon_1^2 c_{3,p}(T) T^{\frac{1}{p}} \left[\frac{\alpha}{2} \left(\frac{8^{p-1} c_\Gamma(p)}{\hat{\delta}_{3,p}} \right)^{\frac{1}{p}} \left(|\Psi(0)| + \zeta_{1,\Psi} \left\| \, |Y_1|_T^* \, \right\|_{L^p} \right) \right.
\end{aligned}
$$

$$+ \hat{\zeta}_1 \left\| \, |Y_2|_T^* \, \right\|_{L^p} + \frac{1}{2} \hat{\zeta}_2 \left\| \, |Y_2|_T^* \, \right\|_{L^p}^2 \Bigg], \tag{114}$$

where

$$c_{3,p}(T) = \delta_{3,p}^{\frac{1}{p}} \exp\left(\frac{\delta_{3,p} \zeta_3^p T}{p} \right), \ \ \delta_{3,p} = 2^{p-1} \hat{\delta}_{3,p}, \ \ \hat{\delta}_{3,p} = \left\{ 3^{p-1} c_B(p) + 4^{p-1} c_\Gamma(p) \right\},$$

$$\hat{\zeta}_2 = \max\{\zeta_{2,\Phi}, \zeta_{2,\Psi}\}, \ \ \zeta_3 = \alpha \max\{\epsilon_1 \zeta_{3,\Phi}, \zeta_{3,\Psi}\}, \ \ \zeta_{3,\Phi} = \sup_{y_1} |\Phi''(y_1)|, \ \ \zeta_{3,\Psi} = \sup_{y_1} |\Psi''(y_1)|,$$

$$Z_{3,t} = Z_3 = \alpha^2 |\Psi(0)|, \ \ |||Z_{3,T}|||_p = T^{\frac{1}{p}} Z_3.$$

Here, let us recall $c_B(p)$ and $c_\Gamma(p)$ from (101) and (102) as follows:

$$c_B(p) = \left(\frac{p^{p+1}}{2(p-1)^{p-1}} \right)^{\frac{p}{2}} T^{\frac{p}{2}-1},$$

$$c_\Gamma(p) = 2^{2p-1}(1+p^p) \left(2^{\frac{p^2}{4}} \left(\int_E |\eta(\epsilon_1,\alpha,z)|^2 \nu(dz) \right)^{\frac{p}{2}} + 2^{\frac{p}{2}} \int_E |\eta(\epsilon_1,\alpha,z)|^p \nu(dz) \right).$$

Particularly, since it's not necessary to use Burkholder-Davis-Gundy inequalities for $p = 2(L^2$-norm), those are simplified as $c_B(2) = 1$ and $c_\Gamma(2) = \int_E \eta(\epsilon_1,\alpha,z)^2 \nu(dz)$.

We also note that because $\sup_{u \in [0,1]} \left\| \frac{\partial^2}{\partial \epsilon_1^2} S_T^{(\epsilon_1 u)} \right\| = \left\| \frac{\partial^2}{\partial \epsilon_1^2} S_T^{(\epsilon_1)} \right\|$, in this example, the bound of $\frac{\epsilon_1^2}{2} \left\| \, |Y_3|_T^* \, \right\|_{L^p}$ derived above is an error estimate for the first order expansion.

Remark C.5. *Error estimates of higher order expansions are obtained in a similar manner based on a recursive scheme (108) with (106), though explicit expressions are more involved. Also, the observations made below hold for higher order expansions, and are consistent with the results for numerical examples in Section 5, shown in Table 6 and Figures from 3 to 10.*

In terms of the orders of $\alpha \in (0,1)$, that is $O(\alpha^k)$, $k = 0,1,2$ (except $\eta(\epsilon_1,\alpha,z)$ in $c_\Gamma(p)$ which is common in each expansion), we observe that $\left\| \, |Y_1|_T^* \, \right\|_{L^p} = O(1)$, $\left\| \, |Y_2|_T^* \, \right\|_{L^p} = O(\alpha)$ and $\frac{\epsilon_1^2}{2} \left\| \, |Y_3|_T^* \, \right\|_{L^p} = O(\alpha^2)$. Hence, as α is smaller, the error becomes smaller in the first order expansion.($\eta(\epsilon_1,\alpha,z)$ also becomes smaller as α goes smaller.) This is true of the higher order expansions, because the error of the k-th order asymptotic expansion is $O(\alpha^{k+1})$. This means that as long as the diffusion and jump coefficients in the original SDE (109) is smaller, the higher order expansions become more effective.

This example also shows a validity of an expansion with $\epsilon_1 = 1$ as in (111) after appropriate transformation (scaling) as in (110). In fact, since $\alpha = \epsilon \in (0,1)$ when $\epsilon_1 = 1$, the error order of k-th order expansion is $O(\epsilon^{k+1})$ such as $\frac{1}{2} \left\| \, |Y_3|_T^* \, \right\|_{L^p} = O(\epsilon^2)$ for the first order expansion.

Moreover, the error depends on time-to-maturity T, and the functional properties of the associated graded SDEs' coefficients, such as $|\Phi(0)|$, $|\Psi(0)|$, the bounds with growth rates of the derivatives $|\Phi'|$, $|\Psi'|$, $|\Phi''|$, $|\Psi''|$, and jump size(reflected in $\eta(\epsilon_1,\alpha,z)$) and intensity. In fact, we observe in this example that as those are smaller, the error becomes smaller and vice versa.

That is, for the error of the k-th order expansion, in addition to the effect of $\alpha^{k+1}(= \epsilon^{k+1})$, the point $\epsilon_1 = 1$ is indeed not "far away" from $\epsilon_1 = 0$, if the function evolves in a mild way from $\epsilon_1 = 0$ to $\epsilon_1 = 1$. (e.g. in our example, if $|\Phi(0)|$, $|\Psi(0)|$, the bounds with growth rates of

the derivatives such as $|\Phi'|$, $|\Psi'|$, $|\Phi''|$, $|\Psi''|$ ($|\Phi'''|$, $|\Psi'''|$ etc. in higher order expansions) and jump intensity and size are small enough.) That is, if the function is good enough, practical applications do not require the point, at which the function is evaluated, to be restricted to the neighborhood (of $\epsilon_1 = 0$). *

Remark C.6. *The error estimates for an asymptotic expansion of the underlying asset price $g(S)$ depend on the functional properties of g such as the growth rates of the function g and its derivatives. Particularly, in the case of an average option, the error is given by a weighted average of the error of each asset price consisting of the average price.*

When a derivatives payoff function f on $g(S)$ is not smooth, (e.g. in a plain-vanilla case, the payoff is a continuous function with only one non-smooth point at the strike), it is approximated by a smooth function as close as possible, by use of an integration-by-parts(IBP) formula if necessary as in Theorem 3.2. In fact, for pricing a plain-vanilla option, we do not need an IBP formula. (Please see Section 2 with Theorem 2 in [35] for the detail.) Thus, the similar discussion as above is true of error estimates for an asymptotic expansion of the derivatives price.

D   Effect of the Third-order Expansion for Jump Components

In this appendix, we compare the accuracy of the full third-order expansion with that of the partial third-order expansion introduced in Section 4. In particular, we take the target payoff as a quadratic function of the average payoff, $\left((A_T^{(\epsilon)} - K)^+ \right)^2$ with $A_T^{(\epsilon)} = \int_0^T S_t^{(\epsilon)} dt$, since it enables us to calculate the full third-order approximation without dealing with a delta function, which would arise from the second order differentiation of the average payoff $(A_T^{(\epsilon)} - K)^+$ in expansion on ϵ as in (67). Hence, we are able to compute its value $\mathbf{E}\left[\left((A_T^{(\epsilon)} - K)^+ \right)^2 \right]$ based on a simple Monte Carlo method with expansion of the payoff. While our analytical approximation is applicable to delta functions and the calculation speed is considered fast enough, the full third-order expansion needs a number of new conditional expectaition formulas. Also, we adopt the following Merton jump-diffusion model as the underlying asset price process:

$$S_T^{(\epsilon)} = s_0 + \int_0^T \epsilon \sigma S_{t-}^{(\epsilon)} dW_t + \sum_{j=0}^{N_T} S_{t-}^{(\epsilon)} (e^{\epsilon Y_j} - 1) - \int_0^T \Lambda (\mathbf{E}[e^{\epsilon Y_1}] - 1) S_{t-}^{(\epsilon)} dt, \tag{115}$$

where $Y_j \sim N(m, \upsilon^2)$.

Then, we use two parameters sets with $\epsilon = 1$: In the "standard parameters" case, we set $s_0 = 100$, $\sigma = 10\%$, $\Lambda = 1$, $Y_j \sim N(-0.05, (0.1)^2)$, which are similar to (iii) in Table 3. Here, in order to adjust the effect of the quadraric payoff function, we put the volatiliy for the diffusion and the standard deviation for the jump size as about a half values of the calibrated ones in Section 5. In the "stressed parameters" case, we increase the standard deviation of the jump size to 20%, that is, $s_0 = 100$, $\sigma = 10\%$, $\Lambda = 1$, $Y_j \sim N(-0.05, (0.2)^2)$. The numbers of time

---

steps and trials in the Monte Carlo method are 100 and 100 million, respectively. The results appear in the figure 11.

Figure 11: Errors of the partial 3rd and the full 3rd approximations
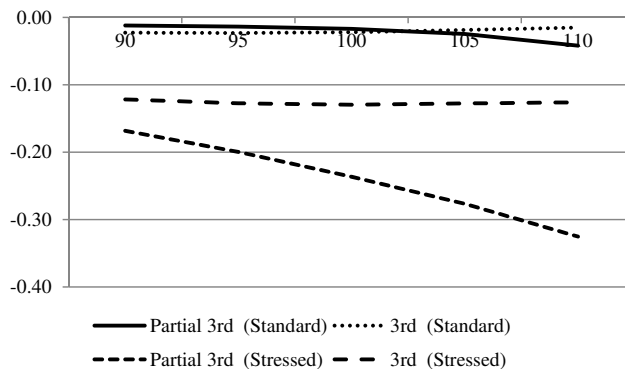
X axis is strike price, and Y axis is the error between the partial third approximation and the true value or the error between the full third approximation and the true one. (Standard) and (Stressed) stand for the results of the standard parameters and the stressed parameters, respectively.

The figure shows the results of the standard and stressed parameters, respectively, where the errors are calculated for the square-root of the values, $\sqrt{\mathbf{E}\left[\left((A_T^{(\epsilon)} - K)^+\right)^2\right]}$, which makes those comparable to the errors for the average options $\mathbf{E}\left[(A_T^{(\epsilon)} - K)^+\right]$ in Section 5. It is observed that while the accuracies of the partial third-order and the full third-order approximations are similar in the standard parameters case, the full third-order approximation improves the partial one in the stressed parameters case.