CIRJE-F-1007

A General Control Variate Method for Multi-dimensional SDEs: An Application to Multi-asset Options under Local Stochastic Volatility with Jumps Models in Finance

Kenichiro Shiraya
The University of Tokyo

Akihiko Takahashi
The University of Tokyo

March 2016; Revised in September 2016

CIRJE Discussion Papers can be downloaded without charge from:
http://www.cirje.e.u-tokyo.ac.jp/research/03research02dp.html

Discussion Papers are a series of manuscripts in their draft form. They are not intended for circulation or distribution except as indicated by the author. For that reason Discussion Papers may not be reproduced or distributed without the written consent of the author.
A General Control Variate Method for Multi-dimensional SDEs: An Application to Multi-asset Options under Local Stochastic Volatility with Jumps Models in Finance *

Kenichiro Shiraya
Graduate School of Economics, the University of Tokyo.
Akihiko Takahashi
Graduate School of Economics, the University of Tokyo.

Abstract

This paper presents a new control variate method for general multi-dimensional stochastic differential equations (SDEs) including jumps in order to reduce the variance of Monte Carlo method. Our control variate method is based on an asymptotic expansion technique, and does not require an explicit characteristic function of SDEs. This is an extension of previous researches using asymptotic expansions to obtain the control variates for such general models. Moreover, in our control variate method, the regression estimators can be chosen for each number of jump times with a stratified sampling, and improve the efficiency of the variance reduction. This paper also provides the asymptotic bias and variance of our method in terms of its terminal time and a small noise parameter used in an asymptotic expansion method.

For an application of our method, we evaluate multi-asset options whose payoffs are expressed as linear functions of the underlying asset price processes in general local stochastic volatility with jumps models, and show a calculation scheme of control variates for Greeks.

In numerical experiments, we apply the new control variate method to pricing basket, spread, and average options and Delta of basket options under a ZABR-type local stochastic volatility model with jumps, and confirm that our method works very well.

Keywords: Control variate, Asymptotic expansion, Multi-asset options, Monte Carlo simulation, Stratified sampling

*Forthcoming in European Journal of Operational Research. This research is supported by CARF (Center for Advanced Research in Finance) and JSPS KAKENHI (Grant Number 25380389).

We are very grateful to the editor and anonymous referees for their precious comments, which have improved the previous version of our paper.
1 Introduction

Monte Carlo (MC) simulation is a very powerful method to obtain target values in complex numerical problems, especially for the ones involving multi-dimensional SDEs. Although PDE methods or tree methods may be alternatives, they are not suitable to be applied to high dimensional problems. On the other hand, MC methods need a lot of computational time for its convergence to a value sufficiently close to the target one due to their large variances. Hence, in order to reduce the computational time, control variate (CV) methods are widely used in MC simulation methods.

Especially in finance, to obtain an accurate value of multi-assets (corresponding to multi-dimensional SDEs) derivatives is one of the most important problems, and there are various approaches for achieving it. For example, analytical approximation formulas for basket option prices are studied by Brigo et al. [4] with a moment matching method, Deelstra et al. [9] with a comonotonic approach, and Bayer and Laurence [3] with a heat kernel expansion and the Laplace approximation method. For mean-reverting models, Li and Wu [24] derived an approximation formula by a moment matching method. As for pricing spread options, which are the well-known example of multi-asset options, Caldana and Fusai [6] derived tight bounds, and Alós et al. [2] studied three-asset spread options.


Their methods are useful for traders to obtain an approximate value quickly, but they have biases which cannot be eliminated. From this reason, many practitioners use MC methods to obtain accurate values stably. However, MC methods have the computational cost problems, and the variance reduction of MC methods for pricing multi-asset options is very important to reduce the substantial computational burden in practice. See Glasserman [17] for a general reference of Monte Carlo methods in finance.

The key point of CV methods is to find a control variate, i.e. a composite function of a target function and the underlying stochastic process, which has a computable expectation and a high correlation with objective random variables. Moreover, it is not enough to find only an approximation formula for the target expectation such as option prices (and implied volatilities), because the method needs a control variate process of which distribution duplicates the distribution of the approximate random variable for the objective one. In particular, our method in the current work applies approximation formulas and control variates based on an asymptotic expansion technique.

Although there exist many studies on financial applications of variance reduction methods, there are few studies on variance reduction methods for evaluating multi-asset options (i.e. for multi-dimensional SDEs). Dahl and Benth [10] studied variance reduction with Quasi-Monte Carlo methods and singular value decomposition. Weng et al. [34] proposed an auto-realignment
method for discontinuous functions, and demonstrated the performance of the QMC method for some digital average and digital multi-asset options.

On CV methods, Dingeç and Hörmann [12] found a very efficient simulation method in the Black-Scholes model. Black-Scholes model with Merton jump diffusion (MJD) case was discussed by Lai et al. [22]. Caldana et al. [7] used pricing bounds under models of which characteristic function of the vector of log-prices are known. On average options, there are successful CV methods using characteristic functions techniques. Dingeç et al. [13] suggested a variance reduction method using a combination of a control variate and a conditional Monte Carlo technique, and showed the pathwise derivative methods for deriving Delta and Gamma under their variance reduction method. Fusai and Kyriakou [15] derived the pricing bounds of average options under exponential Lévy model, affine stochastic volatility models, and the CEV model, and they demonstrated their lower bound are able to be used as the control variate in their numerical examples.

We remark that most of the papers on variance reduction methods for the basket, spread, and average options pricing consider a geometric weighted average value as the control variate and calculate the value with the characteristic functions or closed form expressions of geometric weighted average distributions. However, the general multi-dimensional SDEs with jumps models (corresponding to local stochastic volatility (LSV, e.g. SABR [18]) with jumps models in finance) don’t have the closed form expressions of characteristic functions. Thus, there had been no efficient CV methods for general multi-dimensional SDEs.

On the other hand, since practitioners in finance need to adjust their model parameters to fit the implied volatility smile, but simple models do not fit the real markets, many use local stochastic volatility (LSV) models. However, they are not yet enough to duplicate the market prices, and hence, LSV with jump-diffusion (LSVJ) models has been proposed. Also in time-series analysis, 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.

In order to deal with such general models, we propose a new control variate method, which does not need explicit characteristic functions nor closed form probability density functions. It is also noted that our method is an extension of previous researches using asymptotic expansions in the following sense: it derives a CV method by the second order expansions with its asymptotic errors for pricing multi-asset derivatives such as the basket, spread and average options under multi-dimensional SDEs with jumps, particularly, nonlinear asset return models with jumps in the underlying prices and their volatilities (e.g. LSVJ models). Moreover, our method is able to choose the regression estimators for each number of jump times and can be combined with a stratified sampling technique. (We remark that a combination of a CV method with a stratified sampling was studied, for instance by L’Ecuyer and Buist [23] in an application to a telephone call center model.)

More concretely, in order to obtain an efficient control variate, we apply an asymptotic expansion technique, which allows us to find it in a unified manner for the broad class of stochastic processes, and provide an asymptotic variance of our method in terms of a small noise parameter in the expansion and the terminal time of SDEs. Although control variates with an asymptotic expansion were studied by Takahashi and Yoshida [33] on estimating the asymptotic variance with the first order expansions under diffusion models, and Matsuoka et al.
on computing Greeks under one-dimensional diffusion models, their method is not applicable to general multi-dimensional models with jumps. While Kunitomo and Takahashi [21] studied a CV method based on the asymptotic expansions under jump diffusion models, their method cannot be applied to non-linear asset return models (e.g. LSVJ models), and they only derived control variates corresponding to simple first order expansions. On the other hand, our new method is able to apply more effective the second order expansions to deriving the control variates. Moreover, [21] did not show the asymptotic error estimates for jump diffusion models.

We also remark that the CV method developed in this work is directly applicable to pricing options whose payoffs are expressed as some linear functions of the underlying asset prices (e.g. basket, spread, average options), but not to pricing those whose payoffs depend on the maximums or/and minimums of the underlying assets prices: Although asymptotic expansion schemes for pricing under stochastic volatility models were presented by Shiraya et al. [30], [31], Kato et al. [19], [20] (continuous or discrete monitoring single or double barrier options), and Yamamoto et al. [35] (drawdown (lookback) options), CV methods with the expansions have not been developed mainly due to the complexity of the schemes, which remains to be solved in our future researches. As for pricing rainbow options, to the best of our knowledge, there have no previous works on approximations based on asymptotic expansions.

Also, our method is not applicable to infinite activity models (e.g. Lévy and time-changed Lévy processes), mainly because explicit approximation formulas based on asymptotic expansions have not been obtained to such models. On the other hand, there exist previously cited works based on different approaches such as Caldana and Fusai [6], Caldana et al. [7], Dingeç et al. [13] and Fusai and Kyriakou [15], where CV methods using known characteristic functions (CFs) are successfully applied to those models. As a related study within our approach under some stochastic volatility model with jumps, Takahashi and Takehara [32] introduced a control variable method with a characteristic-function-based Monte Carlo simulation by applying the asymptotic expansion for a diffusion component and a known characteristic function for a jump component. However, they have the same limitation as in Kunitomo and Takahashi [21]: their method cannot be applied to non-linear models (e.g. LSVJ models) nor to pricing multi-asset options such as the basket, spread and average options. We also note that Dingeç and Hörmann [11] studied CV methods using probability density functions for infinite activity Lévy processes in various types of single-asset options such as lookback, barrier and average options. Consequently, one of our next research topics includes an extension of the current method to pricing derivatives in the infinite activity Lévy processes, where asymptotic expansion formulas and the corresponding proxy processes have not been derived.

The organization of the paper is as follows: The next section explains the basic principles of control variate methods, LSVJ models and concrete examples of multi-asset options. In Section 3, we introduce a new control variate method, and Section 4 shows numerical examples of pricing basket options under ZABR type LSV [1] with jump diffusion model. Section 5 concludes. In Appendix A, we provide the proof of the theorem to evaluate asymptotic errors of our control variate method, and in Appendix B, we show a calculation scheme with numerical examples of our control variates for Greeks.
2 Preliminaries

In this section, we briefly explain the control variate method and multi-asset options under local stochastic volatility with jumps (LSVJ) models.

2.1 Control Variate Method

First, we explain the control variate method. Let \( f(X) \) and \( g(\hat{X}) \) be payoff functions with respect to random variables \( X \) and \( \hat{X} \), where \( X \) and \( \hat{X} \) are simulated values of the original random variable and a control variate, respectively. We remark that it needs to choose a control variate \( g(\hat{X}) \) which highly correlates with the original variable \( f(X) \), and has an analytically computable expectation \( \mathbb{E}[g(\hat{X})] \).

Then, we introduce a new random variable \( A \) as follows:

\[
A := f(X) - c \left( g(\hat{X}) - \mathbb{E}[g(\hat{X})] \right),
\]

where \( c \) is a regression estimator, and

\[
\text{Var}(A) = \text{Var}(f(X)) - 2c \text{Cov}(f(X), g(\hat{X})) + c^2 \text{Var}(g(\hat{X})).
\]

Therefore, the optimal value of \( c \) is calculated as:

\[
c_* = \frac{\text{Cov}(f(X), g(\hat{X}))}{\text{Var}(g(\hat{X}))}.
\]

For the optimal \( c_* \), the variance of \( A \) turns out to be

\[
\text{Var}(A) = \text{Var}(f(X))(1 - \rho^2),
\]

where \( \rho \) is the correlation between \( f(X) \) and \( g(\hat{X}) \). Then, in the case of \( \rho \neq 0 \), \( \text{Var}(A) < \text{Var}(f(X)) \). Since \( \mathbb{E}[A] = \mathbb{E}[f(X)] \), we can obtain the option value \( \mathbb{E}[f(X)] \) by calculating \( \mathbb{E}[A] \) of which variance is smaller than the original one.

2.2 Multi-asset Options under Local Stochastic Volatility with Jumps Models in Finance

For an application of the control variate methods, we explain a local stochastic volatility with jumps (LSVJ) model and multi-asset options in finance.

Suppose that the filtered probability space \( (\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t \geq 0}) \) is given, where \( \mathbb{P} \) is an equivalent martingale measure and the filtration satisfies the usual conditions. \( (S^i_t)_{t \in [0,T]} \) and \( (\sigma^i_t)_{t \in [0,T]} \), \( i = 1, \ldots, d \) represent the underlying asset prices and their volatilities for \( t \in [0,T] \), respectively. Particularly, let us assume that \( S_T = (S^1_T, \ldots, S^d_T) \) and \( \sigma = (\sigma^1_T, \ldots, \sigma^d_T) \) are given by the solutions of the following stochastic integral equations:

\[
S^i_T = s^i_0 + \int_0^T \mu_{S^i}(S^i_{t-}) \, dt + \int_0^T \phi_{S^i}(\sigma^i_{t-}, S^i_{t-}) \, d\tilde{W}^S_{t-}.
\]
\[\sum_{i=1}^{L} \sum_{j=1}^{N_{i,t}} h_{i,j} S_{i,j} S_{i,j} - \int_{0}^{T} \Lambda_{i} S_{i,j} \mathbf{E}[h_{i,j},1] dt,\]  
(5)

\[\sigma_{T}^{i} = \sigma_{0}^{i} + \int_{0}^{T} \mu_{\sigma_{i}} (\sigma_{-}^{i}) dt + \int_{0}^{T} \phi_{\sigma_{i}} (\sigma_{-}^{i}) d\bar{W}_{t}^{\sigma_{i}}\]

\[+ \sum_{i=1}^{L} \sum_{j=1}^{N_{i,t}} h_{i,j} \sigma_{i,j}^{i} \sigma_{i,j}^{i} - \int_{0}^{T} \Lambda_{i} \sigma_{i,j}^{i} \mathbf{E}[h_{i,j},1] dt,\]  
(6)

where \(s_{0}^{i}\) and \(\sigma_{0}^{i}\), \(i = 1, \ldots, d\) are given constants. \(\mu_{S_{i}}(x), \mu_{\sigma_{i}}(x), \phi_{S_{i}}(x,y)\) and \(\phi_{\sigma_{i}}(x)\) are some deterministic functions with appropriate regularity conditions. \(\bar{W}_{t}^{S_{i}}\) and \(\bar{W}_{t}^{\sigma_{i}}\), are correlated Brownian motions, and each \(N_{i}, (l = 1, \ldots, L)\) is a Poisson process with constant intensity \(\Lambda_{l}\), and \(N_{i}, l = 1, \ldots, L\) are independent, and also independent of all \(\bar{W}_{t}^{S_{i}}\) and \(\bar{W}_{t}^{\sigma_{i}}\). \(\tau_{j,l}\) stands for the \(j\)-th jump time of \(N_{l}\), and for each \(l = 1, \ldots, L\) and \(i = 1, \ldots, d\), both \(\left(\sum_{j=1}^{N_{i,t}} h_{i,j}\right)_{t \geq 0}\) and \(\left(\sum_{j=1}^{N_{i,t}} h_{i,j}\right)_{t \geq 0}\) are compound Poisson processes. \(\sum_{i=1}^{N_{i,t}} = 0\) when \(N_{i,t} = 0\). For each \(l\) and \(x_{l}, h_{x_{l},j}\) \((j \in \mathbb{N})\) are independent and identically distributed random variables, where \(x_{l}\) stands for one of \(S_{i}\) \((i = 1, \ldots, d)\). For the same \(l\) and \(j\), \(h_{S_{i},j}, h_{\sigma_{i},j}\) \((i, i' = 1, \ldots, d)\) are allowed to be dependent.

Next, we define the discretized processes \(\tilde{S}_{i,n}^{i}\) and \(\tilde{\sigma}_{i,n}^{i}\) of (5) and (6) with \(n\) time steps as:

\[\tilde{S}_{((i+1)T)}^{i,n} = \tilde{S}_{iT}^{i,n} + \mu_{S_{i}} \left(\tilde{S}_{iT}^{i,n} \frac{(i+1)T}{n} - iT \right) + \phi_{S_{i}} \left(\tilde{S}_{iT}^{i,n} \frac{(i+1)T}{n} - iT \right) \left(\tilde{W}_{(i+1)T}^{S_{i}} - \tilde{W}_{iT}^{S_{i}}\right)\]

\[+ \sum_{i=1}^{L} \sum_{j=1}^{N_{i,t}} \left(\sum_{l=1}^{L} \sum_{j=1}^{N_{i,t}} \mathbf{1}_{\{t_{j,l} \in \left[\frac{iT}{n}, \frac{(i+1)T}{n}\right]\}} h_{S_{i},j} \tilde{S}_{i,n}^{i} \frac{(i+1)T}{n} - \Lambda_{l} \mathbf{E}[h_{S_{i},j},1] \frac{(i+1)T}{n} - iT \right),\]

\[\tilde{\sigma}_{((i+1)T)}^{i,n} = \tilde{\sigma}_{iT}^{i,n} + \mu_{\sigma_{i}} \left(\tilde{\sigma}_{iT}^{i,n} \frac{(i+1)T}{n} - iT \right) + \phi_{\sigma_{i}} \left(\tilde{\sigma}_{iT}^{i,n} \frac{(i+1)T}{n} - iT \right) \left(\tilde{W}_{(i+1)T}^{\sigma_{i}} - \tilde{W}_{iT}^{\sigma_{i}}\right)\]

\[+ \sum_{i=1}^{L} \sum_{j=1}^{N_{i,t}} \left(\sum_{l=1}^{L} \sum_{j=1}^{N_{i,t}} \mathbf{1}_{\{t_{j,l} \in \left[\frac{iT}{n}, \frac{(i+1)T}{n}\right]\}} h_{\sigma_{i},j} \tilde{\sigma}_{i,n}^{i} \frac{(i+1)T}{n} - \Lambda_{l} \mathbf{E}[h_{\sigma_{i},j},1] \frac{(i+1)T}{n} - iT \right),\]

This discretization scheme is called as Euler-Maruyama discretization. In this study, we use this discretization scheme to obtain simulated values.

Next, we show the payoffs of multi-asset options. Let us introduce new processes \(\mathcal{S}_{t}^{i}\) defined by:

\[\mathcal{S}_{t}^{i} = \sum_{j=1}^{m_{i}} w_{j}^{(i)} S_{t_{j}^{(i)}}^{i} \mathbf{1}_{\{t_{j}^{(i)} \leq T\}},\]

where \(w_{j}^{(i)}\) is a constant and \(0 \leq t_{1}^{(i)} < \cdots < t_{m_{i}}^{(i)} \leq T\). \(m_{i}\) 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
Then, we can deal with an average option whose underlying asset price \( q(\mathcal{A}_t) \) is given by the following:

\[
q(\mathcal{A}_t) = \sum_{i=1}^{d} \mathcal{A}_t^i,
\]

where \( \mathcal{A}_t = (\mathcal{A}_t^1, \ldots, \mathcal{A}_t^d) \). The payoff of a call option with the strike price \( K \) and the maturity \( T \) is expressed as \( (q(\mathcal{A}_T) - K)^+ = \max\{q(\mathcal{A}_T) - K, 0\} \).

Several well-known option products are expressed by the function.

1. Spread Options
   
   The underlying asset prices of spread options are the difference of futures or spot prices, or interest rates with different maturities, or the difference between the prices of different assets. In this case, by setting \( d = 2 \), \( m_i = 1 \), and \( w_1 = 1 \), \( w_2 = -1 \), we have
   
   \[
   q(\mathcal{A}_t) = S_t^1 - S_t^2.
   \]

2. Basket Options
   
   The underlying asset of a basket option is the weighted average of the prices of different assets, where the weights are typically some prespecified (positive) constants, that is \( m_i = 1 \) and \( w_i t = w_i > 0 \) for all \( i \):
   
   \[
   q(\mathcal{A}_t) = \sum_{i=1}^{d} w_i S_t^i.
   \]

3. Average Options
   
   The underlying with several consecutive maturities may become the underlying assets of an average option as in OTC (over the counter) oil market. In the case of OTC average options traded mainly in commodity markets, \( t^{(i)}_j \) has the relationship; \( 0 < t^{(i)}_1 < \cdots < t^{(i)}_{m_i} < t^{(i+1)}_1 < \cdots \leq T \), and for the case of basket average options; \( 0 < t^{(i)}_1 = t^{(i')}_1 < \cdots < t^{(i)}_{m_i} = t^{(i')}_{m_i'} \leq T \). The underlying asset of the average options is expressed as:
   
   \[
   q(\mathcal{A}_t) = \sum_{i=1}^{d} \mathcal{A}_t^i.
   \]

Hereafter, we concentrate on basket options, and express the underlying as \( q(S_T), S_T = (S_T^1, \ldots, S_T^d) \) for simplicity. The other types of options can be obtained in a similar way.

**Remark 2.1.** The asymptotic bias of Euler-Maruyama scheme is expressed as:

\[
|E[q(S_T)] - E[q(S_T)]| = O\left(\frac{\mathcal{K}(T)}{n}\right),
\]

where \( \mathcal{K}(T) \) is an increasing function on \( T \). Please see [27] for the details.


3 New Control Variate Method

In this section, we introduce a new control variate method for multi-dimensional SDEs with jumps by using a class of LSVJ models, which is frequently applied in finance. Please see Appendix A.1 for a discussion on an asymptotic expansion under the SDEs with jumps related to our method.

3.1 Asymptotic Expansion based Control Variate Method

In order to obtain a control variate based on an asymptotic expansion method (e.g. [28]), let us introduce perturbation parameter \( \epsilon \in (0, 1] \) to the model (5) and (6). We consider the following stochastic integral equations: for \( i = 1, \ldots, d \),

\[
S^i_T = s^i_0 + \int_0^T \mu^i_s \left(S^i_{t-}\right) dt + \epsilon \int_0^T \phi^i_s \left(\sigma^i_{t-}, S^i_{t-}\right) dW^i_t \\
+ \sum_{l=1}^L \sum_{j=1}^{N_{i,l,T}} h^i_{S^j_{t,j-}} - \int_0^T \Delta_i^i s^i_{t-} E[h^i_{S^j_{1,t-},1}] dt, \\
\sigma^i_T = \sigma^i_0 + \int_0^T \mu^{\sigma^i} s \left(S^i_{t-}\right) dt + \epsilon \int_0^T \phi^{\sigma^i} s \left(\sigma^i_{t-}\right) dW^{\sigma^i}_t \\
+ \sum_{l=1}^L \sum_{j=1}^{N_{i,l,T}} h^{\sigma^i}_{s^{j,1}_T} \sigma^i_{t,j-} - \int_0^T \Delta_i^{\sigma^i} s^{i,1}_T E[h^{\sigma^i}_{s^{j,1}_T}] dt, \\
\]

(15)

where \( h^i_{x,j} = e^{iY_{x,j}-1} \), \( \epsilon Y_{x,j} \sim N(\epsilon \mu_{x,j}, \epsilon^2 \gamma_{x,j}) \). In the case of general jump sizes, \( h^i_{x,j} \) of original processes (5) and (6) are approximated by log-normal jumps \( h^i_{x,j} \) by using a moment matching method. The corresponding discretized processes \( S^i_{n,t} \) and \( \sigma^i_{n,t} \) of (15) and (16) with \( n \) time steps are written as:

\[
S^i_{(i+1)/n} = S^i_{i/n} + \int_0^{i/n} \mu^i_s \left(S^i_{t-}\right) dt + \int_0^{i/n} \phi^i_s \left(\sigma^i_{t-}, S^i_{t-}\right) dW^i_t \\
+ \sum_{l=1}^L \sum_{j=1}^{N_{i,l,T,n}} h^i_{S^j_{T,j-}} - \int_0^{i/n} \Delta_i^i S^i_{T,t-} E[h^i_{S^j_{1,t-},1}] \left(\frac{(i+1)/T - i/T}{n} \right), \\
\sigma^i_{(i+1)/n} = \sigma^i_0 + \int_0^{i/n} \mu^{\sigma^i} s \left(S^i_{t-}\right) dt + \int_0^{i/n} \phi^{\sigma^i} s \left(\sigma^i_{t-}\right) dW^{\sigma^i}_t \\
+ \sum_{l=1}^L \sum_{j=1}^{N_{i,l,T,n}} h^{\sigma^i}_{s^{j,1}_T} \sigma^i_{t,j-} - \int_0^{i/n} \Delta_i^{\sigma^i} s^{i,1}_T E[h^{\sigma^i}_{s^{j,1}_T}] \left(\frac{(i+1)/T - i/T}{n} \right), \\
\]

(17)

In financial applications, for the function \( q \) and a call option with the maturity \( T \) and the
strike price $K = q(S_{T}^{(e)}) - \epsilon y$ for an arbitrary $y \in \mathbb{R}$,

$$
(q(S_{T}^{(e)}) - K)^+ = \epsilon \left( \frac{q(S_{T}^{(e)}) - q(S_{T}^{[0]})}{\epsilon} + y \right)^+ \\
= \epsilon \left( q(S_{T}^{[1]}) + y \right)^+ + \frac{\epsilon^2}{2} 1_{q(S_{T}^{[1]})+y>0} q(S_{T}^{[2]}) + O(\epsilon^3 c_1(T) c_2(T))
$$

(19)

where $c_1(T)$ and $c_2(T)$ are increasing functions of $T$ with at most polynomial growth, and

$$
S_{t}^{i,[r]} = \frac{\partial^{r} f_{S_t}^{(i)}}{\partial \epsilon^{r}} S_{t}^{i,(e)} \bigg|_{\epsilon = 0}.
$$

(20)

$S_{t}^{i,[r]}$ is recursively determined and $S_{t}^{i,[r]} = (S_{t}^{1,[r]}, \ldots, S_{t}^{d,[r]})$.

Then, we set a control variate for continuous processes $\tilde{S}_{T}^{(e)}$ as:

$$
\tilde{g}(q(\tilde{S}_{T}^{(e)})) = \left( q(\tilde{S}_{T}^{(e)}) - K \right) 1_{q(S_{T}^{[0]} + S_{T}^{[1]}) > K},
$$

(21)

and for discretized processes $\tilde{S}_{T}^{(e)}$ as:

$$
\tilde{g}(q(\tilde{S}_{T}^{(e)})) = \left( q(\tilde{S}_{T}^{(e)}) - K \right) 1_{q(S_{T}^{[0]} + S_{T}^{[1]}) > K},
$$

(22)

where we note that $q(S_{T}^{(e)})$ or $q(\tilde{S}_{T}^{(e)})$ corresponds to $\tilde{X}$ in Section 2.1.

Here, $\tilde{S}_{T}^{(e)} = (\tilde{S}_{T}^{1,(e)}, \ldots, \tilde{S}_{T}^{d,(e)})$ is a proxy process for $S_{T}^{(e)}$, and is defined as:

$$
\tilde{S}_{t}^{i,(e)} = S_{t}^{i,[0]} + \epsilon S_{t}^{i,[1]} + \frac{\epsilon^2}{2} S_{t}^{i,[2]},
$$

(23)

For a particular model, those expressions of $S_{T}^{i,[n]}$ and $\sigma_{T}^{i,[n]}$ are explicitly given in Section 4.1, and $\tilde{S}_{T}^{(e)}$ is discretized process of $S_{T}^{(e)}$.

These random variables $\tilde{g}(q(S_{T}^{(e)}))$ and $\tilde{g}(q(\tilde{S}_{T}^{(e)}))$ correspond to the target random variables

$$
\hat{f}(q(S_{T}^{(e)})) = (q(S_{T}^{(e)}) - K)^+,
$$

(24)

$$
\hat{f}(q(\tilde{S}_{T}^{(e)})) = (q(\tilde{S}_{T}^{(e)}) - K)^+,
$$

(25)

respectively.

In order to use the control variate in the MC method, we need to obtain the expectation of the above composite function on the proxy process (that is, $\mathbb{E}[\tilde{g}(X)]$ in (1)). In the above setting, the expectation corresponds to the non-discounted initial value of the multi-asset option, whose payoff function is given by $\tilde{g}(q(S_{T}^{(e)}))$. Here, the expectation is calculated by the asymptotic expansion technique as follows:

$$
\mathbb{E}[\tilde{g}(q(\tilde{S}_{T}^{(e)}))] = \sum_{k=0}^{\infty} \sum_{\sum_{l=1}^{d} k_l = k} p_{(k_l)} \mathbb{E}[\tilde{g}(q(\tilde{S}_{T}^{(e)}))|\{N_{t} = k_{l}\}]
$$
\[ \sum_{k=0}^{\infty} \sum_{\sum_{i=1}^{L} k_i = k} p(k) \left\{ \epsilon \left( y_{k_1} \frac{y_{k_1}}{\sqrt{\Sigma(k_1)}} + \Sigma_T^{(k_1)} n(y_{k_1}; 0, \Sigma_T^{(k_1)}) \right) \right\} \]

\[ + \epsilon^2 \left\{ C_{1k_1} N \left( \frac{y_{k_1}}{\sqrt{\Sigma(k_1)}} \right) + \left( C_{2k_1} \Sigma_T^{(k_1)} + C_{3k_1} \frac{y_{k_1}}{\sqrt{\Sigma(k_1)}} \right) n(y_{k_1}; 0, \Sigma_T^{(k_1)}) \right\}, \]

where the probability of \( \{ N_t = k_i \} := \{ N_{1, T} = k_1, \ldots, N_{L, T} = k_L \} \) is defined as:

\[ p(k) := \prod_{l=1}^{L} \frac{(\Lambda_l T)^{k_l} e^{-\Lambda_l T}}{k_l!}. \]  

(27)

Thanks to the independence of \( N_{l, T} \) \( (l = 1, \ldots, L) \), this is the product of the \( k_l \) times of the jump probabilities of \( N_{l, T} \) \( (l = 1, \ldots, L) \), that is \( \prod_{l=1}^{L} P(\{ N_{l, T} = k_l \}) \). \( \lambda_l \) is a deterministic risk-free rate, \( \mathcal{V} = q(S_T^{(l)}) - K, \ y_{l+k} = q(\xi_{l+k}) + \mathcal{V}, \ N(x) \) 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, \( \xi_{l+k} = (\xi_{1, k_1}, \ldots, \xi_{d, k_1}) \) are given by:

\[ \xi_{l+k} := \sum_{l=1}^{L} (k_l - \Lambda_l T) m_{S_l, k_l} S_l^{(0)} \]  

(28)

We define that \( \varrho \cdot B := (W^{S_1}, \ldots, W^{S_d}, W^{V_1}, \ldots, W^{V_d})' \) where \( B \) is a 2d-dimensional (independent) Brownian motion, \( \varrho \) is the \( 2d \times 2d \) Cholesky factor of the correlation matrix of \( B \). Then,

\[ \Sigma_T^{(k_1)} := \sum_{i=1}^{d} \int_{0}^{T} \left( \sum_{j=1}^{2d} w_{i,j} S_t^{(0)} (S_t^{(0)})^{-1} (\varrho)_{i,j} \Phi_{S_t, j} (\sigma_t^{(0)), S_t^{(0)})} \right)^2 dt \]

\[ + \sum_{i=1}^{d} \sum_{j=1}^{d} k_l (w_{i,j} \gamma_{S_t^{(0)}, S_t^{(0)}}) \varrho_{i,j} (w_{i,j} \gamma_{S_t^{(0)}, S_t^{(0)}}), \]  

(29)

where \( \varrho_{i,j} = (\varrho_{i,j})_{1\leq i,j\leq 2d} \) stands for the correlation matrix of \( \zeta_{S_t, j} = (\zeta_{S_t^{(0), j}}, \ldots, \zeta_{S_t^{(0), j}}) \), which is a vector of random variables for jump sizes appearing in \( S_t^{(0)} \) such that \( \zeta_{S_t^{(0), j}} \sim N(0, 1) \), i.e. \( Y_{S_t, j} = u_{S_t, j} (\zeta_{S_t^{(0), j}} + m_{S_t, j}) \). \( \Phi_{S_t, j} := \Phi_{S_t} (\sigma_t; S_t^{(0)})_{i,j} \), where \( (\varrho)_{i,j} \) denotes the \( (i, j) \)-element of \( \varrho \), and \( \Phi_{S_t} := (\Phi_{S_t, 1}, \ldots, \Phi_{S_t, 2d}) \) is a 2d-dimensional vector.

The coefficients \( C_{1k_1}, C_{2k_1}, C_{3k_1} \) are some constants which are obtained by (96) in [28] with integrating them with respect to \( x \) from \(-y_{k_1}\) to \( \infty \) and collecting the terms with the same order of Helmite polynomial. The parenthesized numbers in the following equations stand for the equation numbers listed in Appendix A of [28]:

\[ C_{1k_1} N(\frac{y_{k_1}}{\Sigma_T^{(k_1)}}) = 1 + \int_{-y_{k_1}}^{\infty} \left\{ (117) + (120) + (129) + \text{the first term of (133) + (137)} \right\} n(x; 0, \Sigma_T^{(k_1)}) dx, \]  

(30)
\[
C_{2,1\Sigma_{T}^{(k_{1})}} n[y_{k}; 0, \Sigma_{T}^{(k_{1})}] = \int_{-y_{k_{1}}}^{\infty} \{(108) + (110) + \text{the first term of (112)} + (113) + (118) + (123) + (130) + (131) + (134) + (138)\} \{x; 0, \Sigma_{T}^{(k_{1})}\} \, dx,
\]

\[
C_{3,1\Sigma_{T}^{(k_{1})}} n[y_{k}; 0, \Sigma_{T}^{(k_{1})}] = \int_{-y_{k_{1}}}^{\infty} \{(105) + (109) + (111) + \text{the second term of (112)} + (119) + (124) + (132)\} \{x; 0, \Sigma_{T}^{(k_{1})}\} \, dx.
\]

Please see [28] for the details.

We can obtain control variates for Greeks in a similar way. An example of calculation scheme and its validity are given in Appendix B.1.

**Remark 3.1.** When \( \Sigma \) and \( C_{i} (i = 1, \cdots, 3) \) are obtained as closed-forms, we have obviously no problems in terms of computational complexity and speed. Thus, let us discuss the cases that their closed-forms are not available, and numerical integrations are necessary. All the multiple integrals appearing in \( C_{i} (i = 1, \cdots, 3) \) 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:

\[
\int_{0}^{T} f(s) \int_{0}^{t} g(u) \, du \, ds \approx \sum_{i=1}^{I} \Delta t_{i} f(t_{i}) \sum_{j=1}^{I} \Delta t_{j} g(t_{j})
\]

\[
= \sum_{i=1}^{I} \Delta t_{i} f(t_{i}) \left( G(t_{i-1}) + \Delta t_{i} g(t_{i}) \right),
\]

where \( \Delta t_{i} = (t_{i} - t_{i-1}) \), \( G(t_{i}) = G(t_{i-1}) + \Delta t_{i} g(t_{i}) \) and \( G(t_{0}) = 0 \).

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 3.2.** It needs to truncate the infinite summation in (26) to calculate the value. We decide the level of truncation by the probability of jump times. In our numerical examples, we truncate the summation at the number of jumps whose probability is less than 1/100,000.

The asymptotic bias and variance of our control variate method in terms of a perturbation parameter \( \epsilon \) and an option maturity \( T \) is obtained as the next theorem, whose proof is shown in Appendix A.2.

**Theorem 3.3.** Suppose that the coefficients in SDEs (15) and (16) satisfy Condition A in Appendix A.1. Then, the asymptotic bias \( \mathbf{B}(n, 2) \) and variance \( \mathbf{Var}(\mathcal{M}) \) of our control variate method with \( n \) time steps and \( \mathcal{M} \) sample paths with the second order expansion and \( c \equiv 1 \) (i.e. for \( \tilde{f}(q(S_{T}^{(c)}; n)) - \tilde{f}(q(S_{T}^{(c)}; n)) - \mathbf{E}[\tilde{f}(q(S_{T}^{(c)}; n))] \)) are estimated as:

\[
\mathbf{B}(n, 2) = O \left( \frac{\epsilon^{3}K(T)}{n} \right),
\]

11
\[
\text{Var}(M) = O\left(\epsilon^6 \mathcal{K}(T) \right),
\]
where \( \mathcal{K}(T) \) are increasing functions of \( T \).

### 3.2 Multi-regression Estimators with Stratified Sampling

In order to obtain our new control variate method, adapting the discussions in Section 4.3 of Glasserman [17], we introduce multi-regression estimators with stratified sampling to combine those with our control variate method based on the asymptotic expansion in the previous subsection. Hereafter in this subsection, we omit \( \epsilon \), and consider the case of \( L = 1 \) with setting \( N = N_1 \) for simplicity. We also abbreviate the discretized notation (\( \bar{\)\) on random variables, since we suppose that the number of time steps is enough for the bias of discretization to be negligible.

We define the set \( \Omega_i \), \( i = 1, \cdots, J + 1 \) as follows:

\[
\Omega_i = \{ \omega \in \Omega | N(\omega) = i \} \text{ for } i \leq J,
\]
\[
\Omega_{J+1} = \{ \omega \in \Omega | N(\omega) > J \}. \tag{36} \]

Next, let us briefly explain our scheme for Monte Carlo simulations in the numerical examples. Firstly, we define a sample mean, an unbiased estimator of \( E[Y] \) for a given random variable \( Y \) as:

\[
\hat{Y} := \frac{1}{M} \sum_{i=1}^{M} Y_i, \tag{38} \]

where \( Y_i, i = 1, \cdots, M \) are independent draws from the distribution of \( Y \). We can write \( E[\hat{Y}] \) and \( \text{Var}(\hat{Y}) \) by using \( p_k := P(\Omega_k) \), \( \mu_k := E[Y|\Omega_k] \) and \( \sigma_k^2 := \text{Var}(Y|\Omega_k) = E[(Y - \mu_k)^2|\Omega_k] \) as follows:

\[
E[\hat{Y}] = E[Y] = \sum_{k=0}^{J+1} p_k \mu_k, \tag{39} \]
\[
\text{Var}(\hat{Y}) = \frac{1}{M} \sum_{k=0}^{J+1} p_k \sigma_k^2 + \frac{1}{M} \left[ \sum_{k=0}^{J+1} p_k \mu_k^2 - \left( \sum_{k=0}^{J+1} p_k \mu_k \right)^2 \right]. \tag{40} \]

Since \( \left[ \sum_{k=0}^{J+1} p_k \mu_k^2 - \left( \sum_{k=0}^{J+1} p_k \mu_k \right)^2 \right] \geq 0 \) due to Jensen’s inequality, we have

\[
\text{Var}(\hat{Y}) \geq \frac{1}{M} \sum_{k=0}^{J+1} p_k \sigma_k^2. \tag{41} \]

*We really appreciate an anonymous reviewer’s comments and suggestions, which have substantially improved the previous version of Section 3.2.*
On the other hand, in the case of stratified sampling, let $Y_{k,i}$, $i = 1, \ldots, M_k$ with $M = \sum_{k=0}^{J+1} M_k$ be independent draws from the conditional distribution of $Y$ conditioned on $\Omega_k$, and $\bar{Y}_k$ be the sample mean for each $k$:

$$\bar{Y}_k := \frac{1}{M_k} \sum_{i=1}^{M_k} Y_{k,i}. \quad (42)$$

Then, let us define $Y'$, another estimator of $E[Y]$ as:

$$Y' := \sum_{k=0}^{J+1} p_k \bar{Y}_k, \quad (43)$$

which is clearly an unbiased estimator of $E[Y]$.

Moreover, because of independence of $\bar{Y}_k$, $k = 0, 1, \ldots, J + 1$, its variance is given by:

$$\text{Var}(Y') = \sum_{k=0}^{J+1} p_k^2 \text{Var}(\bar{Y}_k)$$

$$= \sum_{k=0}^{J+1} p_k^2 \frac{1}{M_k} \sigma_k^2$$

$$= \frac{1}{M} \sum_{k=0}^{J+1} p_k^2 \sigma_k^2, \quad (44)$$

where we define $\tilde{p}_k$ as the fraction of samples drawn from the stratum $k$, that is:

$$\tilde{p}_k := \frac{M_k}{M}. \quad (45)$$

Particularly, if we apply the proportional allocation, that is $\tilde{p}_k = p_k$,

$$\text{Var}(Y') = \frac{1}{M} \sum_{k=0}^{J+1} p_k \sigma_k^2 \leq \text{Var}(\bar{Y}), \quad (46)$$

as noted in (41).

Moreover, if we set $\tilde{p}_k = \frac{p_k \sigma_k}{\sum_{k=0}^{J+1} p_k \sigma_k}$ which minimizes $\text{Var}(Y')$ with respect to $\tilde{p}_k$, $k = 0, 1, \ldots, J + 1$ with $(\tilde{p}_k)_{k=0,1,\ldots,J+1}$ being a probability, we obtain:

$$\text{Var}(Y') = \frac{1}{M} \left( \sum_{k=0}^{J+1} p_k \sigma_k \right)^2, \quad (47)$$

which is the optimal allocation for the stratified sampling.

In the numerical examples below, we apply the discussion above by replacing $Y$ with specific random variables: In the classical CV method, we define a random variable $A$ as:

$$A = \bar{f}(q(S_T)) - c \left( \bar{g}(q(S_T)) - E[\bar{g}(q(S_T))] \right) \quad (48)$$
with a common regression coefficient,

\[ c = \frac{Cov(\tilde{f}(q(S_T)), \tilde{g}(q(\hat{S}_T)))}{Var(\tilde{g}(q(\hat{S}_T)))}. \]  

(49)

Then, we use a sample mean estimator for \( \mathbb{E}[A] \) that is,

\[ A = \frac{1}{M} \sum_{i=1}^{M} A_i, \]

(50)

where \( A_i, i = 1, \ldots, M \) are independent draws from the distribution of \( A \). Clearly, \( \mathbb{E}[A] = \mathbb{E}[\tilde{A}] = \mathbb{E}[\tilde{f}(q(S_T))] \).

Next, we discuss on a stratified sampling estimator in the CV method with multiple regressions. Let \( \tilde{f}_{k,i} \) and \( \tilde{g}_{k,i} \) be independent draws from the conditional distributions of \( \tilde{f}(q(S_T)) \) and \( \tilde{g}(q(\hat{S}_T)) \) conditioned on \( \Omega_k \), respectively. Then, a stratified sampling estimator is defined as:

\[ X' := \sum_{k=0}^{J+1} p_k \tilde{X}_k, \]

(51)

where

\[ \tilde{X}_k := \frac{1}{\mathcal{M}_k} \sum_{i=1}^{\mathcal{M}_k} X_{k,i}, \]

\[ X_{k,i} := \tilde{f}_{k,i} - c_k (\tilde{g}_{k,i} - \mathbb{E}[\tilde{g}_{k,i}]). \]

Here, \( c_k \) is the CV coefficient for stratum \( k \), which is defined as:

\[ c_k := \frac{Cov(\tilde{f}_{k,1}, \tilde{g}_{k,1})}{Var(\tilde{g}_{k,1})} = \frac{Cov(\tilde{f}(q(S_T)), \tilde{g}(q(\hat{S}_T))|\Omega_k)}{Var(\tilde{g}(q(\hat{S}_T))|\Omega_k)}, \]

(52)

and \( \mathbb{E}[\tilde{g}_{k,1}] = \mathbb{E}[\tilde{g}(q(\hat{S}_T))|\Omega_k] \) is the conditional expectation of the CV for stratum \( k \).

Then, \( \mathbb{E}[X'] = \mathbb{E}[\tilde{f}(q(S_T))] \), and the variance of \( X' \) is given by:

\[ Var(X') = \frac{1}{\mathcal{M}} \sum_{k=0}^{J+1} p_k Var(X_{k,1}). \]

(53)

(54)

Thus, it is clear that for any \( k \),

\[ Var(X_{k,1}) \leq Var(\tilde{f}_{k,1} - c_k (\tilde{g}_{k,1} - \mathbb{E}[\tilde{g}])). \]

(55)

as \( c_k \) is non-optimal where \( \mathbb{E}[\tilde{g}] := \mathbb{E}[\tilde{g}(q(\hat{S}_T))] \). Therefore, using multi-regression estimators should always yield a smaller variance than using a common CV coefficient \( c \), and a common mean \( \mathbb{E}[\tilde{g}] = \mathbb{E}[\tilde{g}(q(\hat{S}_T))] \) in stratification.

Further, in real simulations for numerical examples, by following a similar rule as in Remark 3.2, we omit the \( \Omega_{J+1} = \{ \omega : N(\omega) > J \} \)-related terms.

### 4 Numerical Examples

In this section, we confirm the effectiveness of our control variate method for pricing basket options under an LSVJ model, to which other control variate methods cannot be applied.
4.1 Settings

For numerical examples, we consider ZABR [1] local stochastic volatility with Merton [26] jumps model on a risk-neutral probability measure. The details of the model are that each underlying asset price process has a CEV (constant elasticity of variance)-type diffusion term with compound Poisson jump component, and each volatility process also has a CEV-type diffusion term and compound Poisson jump component such that

\[
S^i_T = S^i_0 + \int_0^T \alpha^i S^i_t dt + \int_0^T \sigma^i_t (S^i_t)^{\beta^i} dW^S_t^i \\
+ \sum_{l=1}^{N_{1,T}} \sum_{j=1}^{N_{i,T}} h^i_{S^i,l,j} S^i_{\tau^i_{l,j},l} - \int_0^T \Lambda_l S^i_{l-1} \mathbb{E}[h^i_{S^i,l,1}] dt,
\]

(56)

\[
\sigma^i_T = \sigma^i_0 + \int_0^T \nu^i (\sigma^i_t)^{\beta^i} dW^\sigma_t^i \\
+ \sum_{l=1}^{N_{1,T}} \sum_{j=1}^{N_{i,T}} h_{\sigma^i,l,j} \sigma^i_{\tau^i_{l,j},l} - \int_0^T \Lambda_l \sigma^i_{l-1} \mathbb{E}[h_{\sigma^i,l,1}] dt,
\]

(57)

where the jump size \( h^i_{x^i,l,j} \) is given by \( h^i_{x^i,l,j} = e^{Y^i_{x^i,l,j}} - 1 \) with \( Y^i_{x^i,l,j} \) following a normal distribution \( N(m_{x^i,l,j}, \gamma^2_{x^i,l,j}) \) for all \( j \).

The corresponding discretized processes \( \hat{S}^i_{(n+1)T/n} \) and \( \hat{\sigma}^i_{(n+1)T/n} \) of (56) and (57) with \( n \) time steps are written as:

\[
\frac{\hat{S}^i_{(n+1)T/n}}{n} = S^i_{(n+1)T/n} + \alpha^i S^i_{(n+1)T/n} \left( \frac{(i+1)T}{n} - iT \right) + \sigma^i \left( \frac{S^i_{(n+1)T/n} - S^i_{iT/n}}{n} \right) \\
+ \sum_{l=1}^{N_{1,T}} \sum_{j=1}^{N_{i,T}} 1_{(n,j \in \{\frac{iT}{n}, \frac{(i+1)T}{n}\})} h^i_{S^i,l,j} S^i_{\tau^i_{l,j},l} - \sum_{l=1}^{N_{1,T}} \sum_{j=1}^{N_{i,T}} 1_{(n,j \in \{\frac{iT}{n}, \frac{(i+1)T}{n}\})} \Lambda_l S^i_{l-1} \mathbb{E}[h^i_{S^i,l,1}] \left( \frac{(i+1)T}{n} - iT \right),
\]

(58)

\[
\frac{\hat{\sigma}^i_{(n+1)T/n}}{n} = \sigma^i_{(n+1)T/n} + \nu^i \left( \frac{\hat{\sigma}^i_{(n+1)T/n}}{n} \right) \\
+ \sum_{l=1}^{N_{1,T}} \sum_{j=1}^{N_{i,T}} 1_{(n,j \in \{\frac{iT}{n}, \frac{(i+1)T}{n}\})} h_{\sigma^i,l,j} \sigma^i_{\tau^i_{l,j},l} - \sum_{l=1}^{N_{1,T}} \sum_{j=1}^{N_{i,T}} 1_{(n,j \in \{\frac{iT}{n}, \frac{(i+1)T}{n}\})} \Lambda_l \sigma^i_{l-1} \mathbb{E}[h_{\sigma^i,l,1}] \left( \frac{(i+1)T}{n} - iT \right).
\]

(59)

For this model, the proxy processes of continuous version are expressed as follows:

\[
\hat{S}^i_T = S^i_{T/2} + S^i_{T/2} + \frac{1}{2} S^i_{T/2},
\]

(60)

\[
S^i_{T/2} = e^{\int_0^T \sigma^i_t dt} s_0,
\]

(61)

\[
\sigma^i_{T/2} = \sigma^i_0,
\]

(62)

\[
h^i_{x^i,l,j} = 0,
\]

(63)
\[ S_{T}^{i,[1]} = \int_{0}^{T} e^{\int_{0}^{t} \alpha_{t}^{i} \sigma_{t}^{i,[0]} (S_{t}^{i,[0]} \beta_{s}^{i}) dW_{t}^{i} + \sum_{l=1}^{L} \left( \sum_{j=1}^{N_{i,T}} h_{S_{t},l,j} - \Lambda_{l} T m_{S_{t},l} \right) S_{t}^{i,[0]}, \]  
\[ \sigma_{T}^{i,[1]} = \int_{0}^{T} \left( \sigma_{t}^{i,[0]} \right)^{\beta_{s}^{i}} dW_{t}^{i+\delta} + \sum_{l=1}^{L} \left( \sum_{j=1}^{N_{i,T}} h_{S_{t},l,j} \sigma_{t}^{i,[0]} - \Lambda_{l} m_{\sigma_{t},l} \int_{0}^{T} \sigma_{t}^{i,[0]} dt \right), \]  
\[ h_{x_{t},l,j}^{[1]} = Y_{x_{t},l,j} := (Y_{x_{t},l,j}, \ldots, Y_{x_{t},l,j}), \]  
\[ S_{T}^{i,[2]} = 2 \int_{0}^{T} e^{\int_{0}^{t} \alpha_{t}^{i} \sigma_{S_{t}}^{i,[0]} \beta_{s}^{i} (S_{t}^{i,[0]} \beta_{s}^{i} - 1) S_{t}^{i,[1]} dW_{t}^{i} + 2 \int_{0}^{T} e^{\int_{0}^{t} \alpha_{t}^{i} \sigma_{S_{t}}^{i,[0]} \beta_{s}^{i} \sigma_{t}^{i,[1]} dW_{t}^{i} + \sum_{l=1}^{L} \left( \sum_{j=1}^{N_{i,T}} h_{S_{t},l,j}^{[2]} - \Lambda_{l} T (m_{S_{t},l}^{2} + \gamma_{S_{t},l}^{2}) \right) S_{T}^{i,[0]} \right) + 2 \sum_{j=1}^{N_{i,T}} h_{S_{t},l,j}^{[1]} e^{\int_{0}^{T} \alpha_{t}^{i} \sigma_{S_{t}}^{i,[1]} - 2 \Lambda_{l} m_{S_{t},l} e^{\int_{0}^{T} \alpha_{t}^{i} \sigma_{S_{t}}^{i,[1]} dt} \int_{0}^{T} e^{-\int_{0}^{t} \alpha_{t}^{i} \sigma_{S_{t}}^{i,[1]} dt} \right), \]  
\[ h_{x_{t},l,j}^{[2]} = Y_{x_{t},l,j} Y_{x_{t},l,j}. \]  

The discretized version of them are also expressed as:

\[ \tilde{S}_{T}^{i} = \tilde{S}_{T}^{i,[0]} + \tilde{S}_{T}^{i,[1]} + \frac{1}{2} \tilde{S}_{T}^{i,[2]}, \]  
\[ A_{(i+1)T}^{i} = A_{(i+1)T}^{i,[0]} + \alpha_{(i+1)T}^{i,[1]} \left( \frac{(i+1)T}{n} - \frac{iT}{n} \right), \]  
\[ \tilde{S}_{(i+1)T}^{i,[0]} = \tilde{S}_{(i+1)T}^{i,[0]} = e^{\frac{A_{(i+1)T}^{i,[1]}}{n} - m_{0}}, \]  
\[ \tilde{\sigma}_{(i+1)T}^{i,[0]} = \sigma_{0}, \]  
\[ h_{(i+1)T}^{[0]} = 0, \]  
\[ \tilde{S}_{(i+1)T}^{i,[1]} = \tilde{S}_{(i+1)T}^{i,[1]} + e^{\frac{A_{(i+1)T}^{i,[1]} - m_{0}}{n} \left( \tilde{S}_{(i+1)T}^{i,[0]} \beta_{s}^{i} \right) \left( W_{(i+1)T}^{i,n} - W_{i,n}^{i} \right)} \]  
\[ + \sum_{l=1}^{L} \left( \sum_{j=1}^{N_{i,T}} h_{S_{t},l,j}^{[1]} - \Lambda_{l} T m_{S_{t},l} \right) \tilde{S}_{(i+1)T}^{i,[0]} \beta_{s}^{i} \left( \frac{(i+1)T}{n} - \frac{iT}{n} \right) \right) \tilde{S}_{(i+1)T}^{i,[0]} n, \]  
\[ \tilde{\sigma}_{(i+1)T}^{i,[1]} = \tilde{\sigma}_{(i+1)T}^{i,[1]} + \left( \tilde{\sigma}_{(i+1)T}^{i,[0]} \beta_{s}^{i} \right) \left( W_{(i+1)T}^{i+\delta} - W_{i+\delta}^{i+\delta} \right) \]  
\[ + \sum_{l=1}^{L} \left( \sum_{j=1}^{N_{i,T}} h_{S_{t},l,j}^{[1]} \tilde{\sigma}_{(i+1)T}^{i,[0]} n - \Lambda_{l} m_{\sigma_{t},l} \tilde{\sigma}_{(i+1)T}^{i,[0]} n \right) \left( \frac{(i+1)T}{n} - \frac{iT}{n} \right) \right). \]
\[ h_{x,l,j}^{[1]} = Y_{x,l,j} := (Y_{x,1,l,j}, \cdots, Y_{x,5,l,j}), \]  
\[ s_{l;j}^{[2]} = S_l^{[2]} + 2e^{A_{l;j}^{(i+1)T}} - A_{l;j}^{(i+1)T} \beta_S \sigma_{l;j}^{[0],n} (S_l^{[0],n})^{\beta_S - 1} S_l^{[1],n} \left( \frac{W_{i+1}^l - W_i^l}{n} \right) \]  
\[ + 2e^{A_{l;j}^{(i+1)T}} - A_{l;j}^{(i+1)T} \sigma_l^{[1],n} S_l^{[1],n} \left( \frac{W_{i+1}^j - W_i^j}{n} \right) \]  
\[ + \sum_{l=1}^{L} \left\{ \sum_{j=1}^{N_{i,l}} 1 \{ \tau_{l,j} \in (\frac{l}{n}, \frac{(l+1)}{n}) \} h_{S,l,j}^{[2]} - A_l^{(m_{S,l}^2 + \gamma_{S,l}^2)} \left( \frac{(i+1)T}{n} - \frac{iT}{n} \right) \right\} \sigma_l^{[0],n} \]  
\[ + 2h_{S,l,j}^{[1]} e^{A_{l;j}^{(i+1)T}} - A_{l;j}^{(i+1)T} \sigma_l^{[1],n} \]  
\[ - 2A_l^{(m_{S,l}^2 + \gamma_{S,l}^2)} \left( \frac{(i+1)T}{n} - \frac{iT}{n} \right), \]  
\[ h_{x,l,j}^{[2]} = Y_{x,l,j} Y_{x,l,j}. \]  

Here, the multiple stochastic integrals in above equations are simulated by the program code with only one loop against the time parameter due to a similar argument as in Remark 3.1.

We consider five-asset basket options as a base case. For illustrative purpose, we consider a systematic jump case, that is all the jumps of the underlying asset prices and their volatilities occur at the same time (i.e. \( n = 1 \) and \( \sigma_{x,y} = 1 \) where \( \sigma \) denotes the 10 \times 10 correlation matrix among \( \zeta_{S,l,j} \) and \( \zeta_{S',l,j} \) \((i = 1, \cdots, 5)\)), though we can treat more general cases.

As the base case, We apply the same parameters, which are listed in Table 1 and Table 2, to all assets and volatilities.

<table>
<thead>
<tr>
<th>( S_l^0 )</th>
<th>( \sigma_l^0 )</th>
<th>( \alpha_l )</th>
<th>( \beta_S )</th>
<th>( \beta_{S_l} )</th>
<th>( \nu_l )</th>
<th>( \Lambda )</th>
<th>( L )</th>
<th>( m_{S_l} )</th>
<th>( \gamma_{S_l} )</th>
<th>( m_{S_l^*} )</th>
<th>( \gamma_{S_l^*} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>2</td>
<td>0</td>
<td>0.5</td>
<td>0.75</td>
<td>0.2</td>
<td>1</td>
<td>1</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
</tr>
</tbody>
</table>

**Table 1: Common Parameters**

<table>
<thead>
<tr>
<th>( S_l )</th>
<th>( S_2 )</th>
<th>( S_3 )</th>
<th>( S_4 )</th>
<th>( S_5 )</th>
<th>( \sigma_1 )</th>
<th>( \sigma_2 )</th>
<th>( \sigma_3 )</th>
<th>( \sigma_4 )</th>
<th>( \sigma_5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S_1 )</td>
<td>1</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>-0.4</td>
<td>-0.4</td>
<td>-0.4</td>
<td>-0.4</td>
</tr>
<tr>
<td>( S_2 )</td>
<td>0.4</td>
<td>1</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>-0.4</td>
<td>-0.4</td>
<td>-0.4</td>
<td>-0.4</td>
</tr>
<tr>
<td>( S_3 )</td>
<td>0.4</td>
<td>0.4</td>
<td>1</td>
<td>0.4</td>
<td>0.4</td>
<td>-0.4</td>
<td>-0.4</td>
<td>-0.4</td>
<td>-0.4</td>
</tr>
<tr>
<td>( S_4 )</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>1</td>
<td>0.4</td>
<td>-0.4</td>
<td>-0.4</td>
<td>-0.4</td>
<td>-0.4</td>
</tr>
<tr>
<td>( S_5 )</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>1</td>
<td>-0.4</td>
<td>-0.4</td>
<td>-0.4</td>
<td>-0.4</td>
</tr>
</tbody>
</table>

| \( \sigma_1 \) | -0.4 | -0.4 | -0.4 | -0.4 | -0.4 | 1 | 0.4 | 0.4 | 0.4 |
| \( \sigma_2 \) | -0.4 | -0.4 | -0.4 | -0.4 | -0.4 | 0.4 | 1 | 0.4 | 0.4 |
| \( \sigma_3 \) | -0.4 | -0.4 | -0.4 | -0.4 | -0.4 | 0.4 | 0.4 | 1 | 0.4 |
| \( \sigma_4 \) | -0.4 | -0.4 | -0.4 | -0.4 | -0.4 | 0.4 | 0.4 | 0.4 | 1 |
| \( \sigma_5 \) | -0.4 | -0.4 | -0.4 | -0.4 | -0.4 | 0.4 | 0.4 | 0.4 | 0.4 |

**Table 2: Correlations**
The weight of each asset is set as $w_i = 1/d$ ($d = 5$ in the base case). When we don’t mention about parameters, we use these values for numerical examples.

4.2 Numerical Results

First, we show the results of base parameters, which are shown in Table 1 and Table 2, with 256 time steps and 50,000 trials for each simulation on the basket, spread, and average options. Spread options are calculated from the price of asset 1 minus that of asset 2. Each underlying reference price of an average option is calculated from the average of Asset 1 & 2’s prices during the last 1 month (for the 1 year maturity in this example), which reflects the convention in the oil markets: Precisely, Asset 1’s prices from $237/256$ (around 0.926 years) to $246/256$ (around 0.961 years) and Asset 2’s prices from $247/256$ (around 0.965 years) to 1 year with $1/256$ time step. In order to prevent bias of paths, we calculate regression estimators if the number of samples is more than five. When the number of samples is less than or equal to five (e.g. the number of jumps is $k$), we set the regression estimator as 1 (i.e. $c_k = 1$), and does not adjust the probability.

The results are shown in Table 3.

<table>
<thead>
<tr>
<th>Price</th>
<th>Diff</th>
<th>Error</th>
<th>VRF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K</td>
<td>NCV</td>
<td>AE</td>
</tr>
<tr>
<td>Basket</td>
<td>80</td>
<td>21.95</td>
<td>21.98</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>10.51</td>
<td>10.43</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>5.10</td>
<td>4.99</td>
</tr>
<tr>
<td>Spread</td>
<td>-20</td>
<td>22.36</td>
<td>22.15</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>8.81</td>
<td>8.74</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>2.34</td>
<td>2.15</td>
</tr>
<tr>
<td>Average</td>
<td>80</td>
<td>22.26</td>
<td>22.28</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>10.73</td>
<td>10.65</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>5.18</td>
<td>5.05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time (sec)</th>
<th>EF</th>
</tr>
</thead>
<tbody>
<tr>
<td>K</td>
<td>MC</td>
</tr>
<tr>
<td>Basket</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>120</td>
</tr>
<tr>
<td>Spread</td>
<td>-20</td>
</tr>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>20</td>
</tr>
<tr>
<td>Average</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>120</td>
</tr>
</tbody>
</table>

Table 3: Results of Basket, Average and Spread Options

CV means the control variate method with the classical regression estimator (using only one
regression estimator), and NCV means our new control variate method using the multi-regression estimators and stratified sampling. AE denotes the approximation value which is calculated in (26) with numerical integral. VRF means variance reduction factor which is calculated as \( \frac{\sigma_{MC}^2}{\sigma_{CV}^2} \) or \( \frac{\sigma_{MC}^2}{\sigma_{NCV}^2} \). Here, \( \sigma_{MC} \) is a standard deviation of crude Monte Carlo simulations, and \( \sigma_{CV} \) and \( \sigma_{NCV} \) are standard deviations of CV method and NCV method, respectively. Diff stands for a difference between AE and NCV (AE – NCV). Error means 95 % error bound of a crude Monte Carlo method and our new control variate methods. EF means the efficiency factor which is defined as \( EF = VRF \times \frac{t_{MC}}{t_{CV}} \) where \( t_{MC} \) is a computational time of the crude Monte Carlo, and \( t_{CV} \) is that of the new method. 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.

Although the model used in this experiment is very complex, our control variate seems quite successful, and the results are stable regardless of strikes. Moreover, the cases of multi-regression estimators work better than that of the one regression estimator method. We also note that the second order AE method is able to calculate the prices very fast, which is useful when pricing quickly is necessary while very accurate prices are not required (e.g. when we need to indicate prices for many customers in a short time). However, as we need to estimate P/Ls (profit and losses) very accurately in practice (e.g. when we evaluate daily P/Ls for traders), it is not appropriate to use only the second order AE approximations due to the error level. Thus, our new control variate method is very useful so as to estimate derivatives values fast and accurately. We remark that the similar results and observation hold in the subsequent numerical examples.

Next, we show the relationship between the number of assets and accuracy. Hereafter, we concentrate on basket options. We change the number of assets from 3 to 30, and the results are shown in Table 4.

<table>
<thead>
<tr>
<th>Num of Assets</th>
<th>K</th>
<th>Price</th>
<th>Error</th>
<th>VRF</th>
<th>Time (sec)</th>
<th>EF</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>80</td>
<td>22.18</td>
<td>22.19</td>
<td>0.245</td>
<td>0.016</td>
<td>240</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>10.75</td>
<td>10.66</td>
<td>0.201</td>
<td>0.015</td>
<td>178</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>5.26</td>
<td>5.13</td>
<td>0.154</td>
<td>0.013</td>
<td>146</td>
</tr>
<tr>
<td>5</td>
<td>80</td>
<td>21.95</td>
<td>21.98</td>
<td>0.242</td>
<td>0.014</td>
<td>303</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>10.51</td>
<td>10.43</td>
<td>0.198</td>
<td>0.014</td>
<td>206</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>5.10</td>
<td>4.99</td>
<td>0.150</td>
<td>0.011</td>
<td>176</td>
</tr>
<tr>
<td>10</td>
<td>80</td>
<td>21.79</td>
<td>21.81</td>
<td>0.240</td>
<td>0.013</td>
<td>322</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>10.33</td>
<td>10.25</td>
<td>0.196</td>
<td>0.013</td>
<td>222</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>4.99</td>
<td>4.88</td>
<td>0.149</td>
<td>0.011</td>
<td>179</td>
</tr>
<tr>
<td>20</td>
<td>80</td>
<td>21.70</td>
<td>21.73</td>
<td>0.236</td>
<td>0.011</td>
<td>455</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>10.24</td>
<td>10.17</td>
<td>0.192</td>
<td>0.011</td>
<td>294</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>4.92</td>
<td>4.82</td>
<td>0.144</td>
<td>0.009</td>
<td>242</td>
</tr>
<tr>
<td>30</td>
<td>80</td>
<td>21.68</td>
<td>21.70</td>
<td>0.238</td>
<td>0.011</td>
<td>448</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>10.22</td>
<td>10.14</td>
<td>0.195</td>
<td>0.011</td>
<td>291</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>4.91</td>
<td>4.81</td>
<td>0.147</td>
<td>0.010</td>
<td>235</td>
</tr>
</tbody>
</table>

Table 4: VRF and EF with different number of assets

19
This result shows that VRF and EF tend to be higher as the number of assets is larger.

Next, we examine VRFs and EFs associated with the number of simulations. The computational time is proportional to the number of simulations or the number of time steps. Thus, we omit the computational time from the tables, hereafter. The results are listed in Table 5.

<table>
<thead>
<tr>
<th>Num of Simulations</th>
<th>( K )</th>
<th>Price</th>
<th>Error</th>
<th>VRF</th>
<th>EF</th>
</tr>
</thead>
<tbody>
<tr>
<td>5000</td>
<td>80</td>
<td>21.97</td>
<td>21.98</td>
<td>0.746</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>10.52</td>
<td>10.43</td>
<td>0.604</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>5.13</td>
<td>4.99</td>
<td>0.450</td>
<td>0.035</td>
</tr>
<tr>
<td>10000</td>
<td>80</td>
<td>21.97</td>
<td>21.98</td>
<td>0.538</td>
<td>0.030</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>10.52</td>
<td>10.43</td>
<td>0.440</td>
<td>0.029</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>5.11</td>
<td>4.99</td>
<td>0.333</td>
<td>0.024</td>
</tr>
<tr>
<td>50000</td>
<td>80</td>
<td>21.95</td>
<td>21.98</td>
<td>0.242</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>10.51</td>
<td>10.43</td>
<td>0.198</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>5.10</td>
<td>4.99</td>
<td>0.150</td>
<td>0.011</td>
</tr>
<tr>
<td>100000</td>
<td>80</td>
<td>21.96</td>
<td>21.98</td>
<td>0.171</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>10.51</td>
<td>10.43</td>
<td>0.140</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>5.11</td>
<td>4.99</td>
<td>0.106</td>
<td>0.008</td>
</tr>
<tr>
<td>500000</td>
<td>80</td>
<td>21.96</td>
<td>21.98</td>
<td>0.076</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>10.51</td>
<td>10.43</td>
<td>0.062</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>5.11</td>
<td>4.99</td>
<td>0.048</td>
<td>0.004</td>
</tr>
</tbody>
</table>

Table 5: VRF and EF with different number of simulations

The results show that the number of simulations does not need so many to obtain the accurate value of the basket option by using our control variate method. VRFs of 5,000 samples are smaller than that of the others, because the number of samples to calculate each regression estimator is small, and it causes some bias of regression estimators. However, the cases of more than 10,000 samples seem enough to obtain the stable efficiency of the variance reduction in this setting.

For the reference, Figure 1 shows the convergence of simulations with the case of the strike price \( K = 100 \) and 50,000 samples.
Next, we provide a parameter sensitivity analysis to examine our new control variate method with changes in the model parameters. "High" means the twice value of the base parameters given in Table 1 and Table 2, and "Low" means the half value of the base parameters. Especially as for the maturity, we also test the 4 times ($4T$) and 8 times ($8T$) cases. The results are shown in Table 6.
### Table 6: Sensitivity Analysis

The low parameter cases, which means that the volatility of the basket price is small, show better performance. This is because the asymptotic expansion method expands the original processes around the small volatility and hence, the low parameters cases tend to approximate accurately the original processes (see Theorem 3.3).
Next, we compare the bias of MC method with that of NCV method with the basket options. To compare them, we set the number of partitions from 8 to 256, and run 100 million sample paths for MC and 500 thousand for NCV. As observed in Table 5 with Error for MC method transformed into the 100 million samples case, we note that all the 95% error bounds for NCV and MC are much smaller than 0.01, so that the bias estimates are accurate up to two digits after the decimal point. The results are shown in Table 7.

<table>
<thead>
<tr>
<th>K</th>
<th>Number of Steps</th>
<th>8</th>
<th>16</th>
<th>32</th>
<th>64</th>
<th>128</th>
<th>256</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>MC</td>
<td>10.44</td>
<td>10.48</td>
<td>10.49</td>
<td>10.50</td>
<td>10.51</td>
<td>10.51</td>
</tr>
<tr>
<td></td>
<td>NCV</td>
<td>10.57</td>
<td>10.54</td>
<td>10.53</td>
<td>10.52</td>
<td>10.51</td>
<td>10.51</td>
</tr>
<tr>
<td>120</td>
<td>MC</td>
<td>4.89</td>
<td>5.00</td>
<td>5.05</td>
<td>5.08</td>
<td>5.09</td>
<td>5.10</td>
</tr>
<tr>
<td></td>
<td>NCV</td>
<td>5.18</td>
<td>5.14</td>
<td>5.13</td>
<td>5.12</td>
<td>5.11</td>
<td>5.11</td>
</tr>
</tbody>
</table>

Table 7: Testing Bias for Basket Option Prices

We observe that in the case of OTM (out of the money: strike price = 120), the bias of CV method is much smaller than that of MC method. Although the reduction of bias depends on the models and parameters, it is expected to be achieved in many cases, mainly because of the difference in orders of $\epsilon$ between (102) and (103) in Appendix A.2.

Finally, we report the results of optimal allocation (See Glasserman [17]) with our NCV method. We allocate the number of samples for each number of jumps to satisfy

$$M_i = M \frac{p_i \sigma_i}{\sum_{i=0}^{I} p_i \sigma_i},$$

where $M_i$ is the number of samples of $i$ times jumps occurring in the path, $\sigma_i$ is the standard deviation obtained by 256 pilot samples of $i$ times jump paths, $I$ is the first number of jumps whose probability is less than 1/100,000, and $M$ is the number of samples of NCV method. Please see e.g. [17] for the details. We test the cases of 5 assets basket options, spread and average options whose conditions are the same as in the previous tests. The results are in Table 8.
<table>
<thead>
<tr>
<th>$K$</th>
<th>Price</th>
<th>VRF</th>
<th>Time (sec)</th>
<th>EF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NCV</td>
<td>NCV</td>
<td>OA</td>
<td>MC</td>
</tr>
<tr>
<td>Basket</td>
<td>80</td>
<td>21.95</td>
<td>303</td>
<td>461</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>10.51</td>
<td>206</td>
<td>345</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>5.10</td>
<td>176</td>
<td>435</td>
</tr>
<tr>
<td>Spread</td>
<td>-20</td>
<td>22.36</td>
<td>42</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>8.81</td>
<td>33</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>2.34</td>
<td>12</td>
<td>16</td>
</tr>
<tr>
<td>Average</td>
<td>80</td>
<td>22.26</td>
<td>203</td>
<td>250</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>10.73</td>
<td>154</td>
<td>198</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>5.18</td>
<td>126</td>
<td>232</td>
</tr>
</tbody>
</table>

Table 8: Optimal Allocation

OA means the result of optimal allocation with our new control variate method. The computational time of OA is a little bit smaller than that of NCV because OA method does not need to generate random variables to judge jumps. VRF and EF of OA are better than those of NCV method, especially in OTM cases. In sum, this numerical experiment confirmed that our new control variate method is able to reduce the variance of Monte Carlo very much.

**Remark 4.1.** Appendix B.1 shows numerical examples of Delta for basket options.

## 5 Conclusion

We have introduced a new control variate method using an asymptotic expansion technique for general multi-dimensional SDEs with jumps, which can be applied to multi-asset options under local stochastic volatility (LSV) with jumps (LSVJ) models in finance. Moreover, since our method is able to make use of multi-regression estimators with stratified sampling, it reduces the variance of simulations, further. Moreover, we have provided an asymptotic variance of our method in terms of a small noise parameter $\epsilon$ in the asymptotic expansion and a terminal time $T$. In addition, we have shown a calculation scheme and validity of our control variate for Greeks. In numerical examples, we have demonstrated that our method works efficiently for pricing basket options with Delta, and spread and average options under a ZABR-type LSVJ model. Particularly, we use the second order expansions for pricing options and the first order expansions for computing Delta.

We remark that our method is currently not applicable to pricing options whose payoffs depend on the maximums or minimums of the underlying assets prices under LSV and LSVJ models, and not to infinite activity models (e.g. Lévy and time-changed Lévy processes), mainly because of no explicit approximation formulas or complexity of the schemes based on asymptotic expansions. Hence, our next research topics include extensions of our method to pricing derivatives, where asymptotic expansion formulas and the corresponding proxy processes have not been derived or the approximation schemes based on the expansions are more complex.
References


A Error Estimate of Our Control Variate Method

In this section, we estimate the asymptotic bias and variance of our control variate method in terms of a small noise parameter $\epsilon$ and an option maturity $T$.

A.1 Asymptotic Expansion under SDEs with jumps

For the preparation of deriving error estimates, we consider an asymptotic expansion for the next general process.

\[
Z_t^{(\epsilon)} = \int_0^t \int_E h(\epsilon, z)N(dt, dz), \tag{80}
\]

\[
Z_t = \int_0^t \int_E h(1, z)N(dt, dz), \tag{81}
\]

\[
X_t = x + \int_0^t \mu(s, X_{s-})ds + \int_0^t \Phi(s, X_{s-})dW_s + \int_0^t \gamma(s, X_{s-})dZ_s, \tag{82}
\]

where $x \in \mathbb{R}^D$, $\mu : [0, T] \times \mathbb{R}^D \times \Omega \rightarrow \mathbb{R}^D$, $\Phi : [0, T] \times \mathbb{R}^D \times \Omega \rightarrow \mathbb{R}^D \times \mathbb{R}^m$, $h : (0, 1] \times \mathbb{R} \times \Omega \rightarrow \mathbb{R}^D \times \mathbb{R}^D$, $\gamma : [0, T] \times \mathbb{R}^D \times \Omega \rightarrow \mathbb{R}^D$ are predictable processes, and $W$ is an $m$-dimensional
Brownian motion. Each \( N(dt, dz) \) is a Poisson random measure on \([0, T] \times E\) where \((E, \mathcal{E})\) is a measurable space with \( E \subset \mathbb{R}^e, e \in \mathbb{N}\), and the intensity measure of \( N \) is \( dt \times \nu(dz) \), where \( \nu(dz) \) is a positive \( \sigma \)-finite measure on \((E, \mathcal{E})\). Then, \( \tilde{N}(dt, dz) := N(dt, dz) - dt \times \nu(dz) \) is a compensated Poisson random measure.

Next, we introduce a perturbation parameter \( \epsilon \in (0, 1) \) to the processes \( X \), and the perturbed process \( X^{(\epsilon)} \) is defined as:

\[
X^{(\epsilon)}_t = x + \int_0^t \mu(s, X^{(\epsilon)}_s)ds + \epsilon \int_0^t \Phi(s, X^{(\epsilon)}_s)dW_s + \int_0^t \gamma(s, X^{(\epsilon)}_s)dZ^{(\epsilon)}_s, \tag{83}
\]

and assume the following conditions:

**Condition A**

1. \( h(\epsilon, z) \in C^\infty(\epsilon), \mu(t, x), \Phi(t, x), h(\epsilon, z)\gamma(t, x)/\eta(z), \frac{\partial^2}{\partial z^2} h(\epsilon, z)\gamma(t, x)/\eta(z) \in C^\infty_0(x) \text{ (} j \in \mathbb{N}\), with \( \eta, \eta_j : E \rightarrow \mathbb{R} \) and \( \eta, \eta_j \in \cap_{p \geq 2} L^p(E, \nu) \).

2. \( |\mu(t, 0)|, |\Phi(t, 0)|, |h(\epsilon, z)\gamma(t, x)/\eta(z)|, |\frac{\partial^2}{\partial z^2} h(\epsilon, z)\gamma(t, x)/\eta(z)| < Z_f \) for \( j \in \mathbb{N}\) and there exists a predictable process \( Z_t \), such that \( Z_t \) satisfies \( \int_0^T E[|Z_t|^p]dt < \infty \) for \( \forall p \geq 1 \).

We also assume that \( \tilde{\psi} : \mathbb{R}^D \rightarrow \mathbb{R} \) is a Borel measurable function such that \( |\tilde{\psi}(x)| \leq C(1 + |x|^p) \) for \( \exists C > 0 \) and \( \forall p \geq 0 \), and has a finite number of discontinuous points.

Here, there exists \( C^\infty_0\)-function \( \{\psi_n\} \) such that \( ||\psi_l(x)|| \leq C(1 + |x|^p) \) and \( \lim_{l \rightarrow \infty} \psi_l(x) = \tilde{\psi}(x) \) for \( \forall x \). Then, there exists \( \exists \mathcal{N} \) such that

\[
|E[\psi_l(X^{(\epsilon)}_T)] - E[\tilde{\psi}(X^{(\epsilon)}_T)]| \leq E[|\psi_l(X^{(\epsilon)}_T) - \tilde{\psi}(X^{(\epsilon)}_T)|] < \delta. \tag{84}
\]

for \( \forall \delta > 0 \) and \( l \geq \mathcal{N} \). For a small \( \delta > 0 \), some \( l \geq \mathcal{N} \), we take \( \psi = \psi_l \) \( (l \geq \mathcal{N}) \).

Then,

\[
E \left[ \psi_l(X^{(\epsilon)}_T) \right] = E \left[ a_{0, T}^{\psi} + \epsilon a_{1, T}^{\psi} + \epsilon^2 a_{2, T}^{\psi} + \cdots + \epsilon^M a_{M, T}^{\psi} + \epsilon^{M+1} \int_0^1 \frac{(1 - u)^M}{M!} a^{\psi, (\epsilon u)}_{M+1, T} du \right], \tag{85}
\]

where

\[
a_{0, T}^{\psi} = \psi(X^{[0]}_{T}), \tag{86}
\]

\[
a_{1, T}^{\psi} = \sum_{j=1}^{D} \partial_j \psi(X^{[0]}_{T}) F_T^{j, [1]}, \tag{87}
\]

\[
a_{m, T}^{\psi} = \sum_{m(r), \alpha(r)} \partial_{\alpha(r)}^{r} \psi(X^{[0]}_{T}) F_T^{\alpha_1, [m_1]} \cdots F_T^{\alpha_r, [m_r]}, \tag{88}
\]

\[
a_{m, T}^{\psi, (\epsilon)} = \sum_{m(r), \alpha(r)} \partial_{\alpha(r)}^{r} \psi_l(X^{(\epsilon)}_{T}) F_T^{\alpha_1, (\epsilon)_1, m_1} \cdots F_T^{\alpha_r, (\epsilon), m_r}, \tag{89}
\]

where \( \partial_j \psi(x) = \frac{\partial}{\partial x_j} \psi(x), j = 1, \cdots, D, a^{\psi, (\epsilon)}_{m, T} := \frac{1}{m!} \frac{\partial^m \psi_l(X^{(\epsilon)}_{T})}{\partial x^m}, F_{T}^{(\epsilon), r} = \frac{1}{r!} \frac{\partial^r X^{(\epsilon)}_{T}}{\partial x^r}, \text{ and } F_{T}^{j, (\epsilon), r}, j = 1, \cdots, D \) denotes the \( j \)-th element of \( F_{T}^{(\epsilon), r} \).
denote the \( j \)-th elements of \( F^{(r)}_t \). For \( r \geq 1 \), \( F^{(r)}_t \), \( j = 1, \ldots, D \) is recursively determined, and

\[
\sum_{m^{(r)}, \alpha^{(r)}}^{[m]} := \sum_{r=1}^{m} \sum_{m_1 + \cdots + m_r = m, m_i \geq 1, \alpha^{(r)} \in [1, \ldots, D]^r} \frac{1}{r!}.
\]  

(89)

From Condition A, we can apply Theorem 2.1 in [29],

\[
\mathbb{E} \left[ \tilde{\psi}(X_T^{(c)}) \right] = \mathbb{E} \left[ a_{0,T}^{\psi} + \epsilon a_{1,T}^{\psi} + \epsilon^2 a_{2,T}^{\psi} + \cdots + \epsilon^M a_{M,T}^{\psi} \right] + O(\epsilon^{M+1} c_1(T) c_2(T)),
\]

(90)

where \( c_1(T) \) and \( c_2(T) \) are increasing functions of \( T \) with at most polynomial growth.

Next, let the discretization step on \([0, T]\) be \( \frac{T}{n} \), and define a discretized process of (83) as:

\[
\bar{X}_t^{(c), n} = \bar{X}_t^{(c), n} + \mu \left( \frac{t}{n}, \bar{X}_t^{(c), n} \right) \left( \frac{(i + 1)T}{n} - \frac{t}{n} \right) + \Phi \left( \frac{t}{n}, \bar{X}_t^{(c), n} \right) \left( \frac{W_{(i+1)T} - W_t}{n} \right)
\]

\[+ \gamma \left( \frac{t}{n}, \bar{X}_t^{(c), n} \right) \left( \frac{Z_{(i+1)T} - Z_t}{n} \right),
\]

(91)

where \( \frac{(i+1)T}{n} \leq t < \frac{(i+2)T}{n} \).

In a similar way, we obtain

\[
\mathbb{E} \left[ \psi(\bar{X}_T^{(c), n}) \right] = \mathbb{E} \left[ b_{0,T}^{\psi,n} + \epsilon b_{1,T}^{\psi,n} + \epsilon^2 b_{2,T}^{\psi,n} + \cdots + \epsilon^M b_{M,T}^{\psi,n} + M+1 \int_0^1 \left( 1 - u \right)^M M! b_{M+1,T}^{\psi(u),n} du \right],
\]

(92)

where

\[
b_{0,T}^{\psi,n} = \psi(\bar{X}_T^{(c), n}),
\]

(93)

\[
b_{1,T}^{\psi,n} = \sum_{j=1}^{D} \partial_j \psi(\bar{X}_T^{(c), n}) G_T^{j,[1],n},
\]

(94)

\[
b_{m,T}^{\psi,n} = \sum_{m^{(r)}, \alpha^{(r)}}^{[m]} \partial_{\alpha^{(r)}} \psi(\bar{X}_T^{(c), n}) G_T^{\alpha^{(r)},[m],n} \cdots G_T^{\alpha^{(r)},[m_r],n},
\]

(95)

\[
b_{m,T}^{\psi, (c), n} = \sum_{m^{(r)}, \alpha^{(r)}}^{[m]} \partial_{\alpha^{(r)}} \psi(\bar{X}_T^{(c), n}) G_T^{\alpha^{(r)},[m_1],n} \cdots G_T^{\alpha^{(r)},[m_r],n},
\]

\[
G_t^{(c), r, n} = \frac{1}{r!} \partial_{\alpha^{(r)}} \bar{X}_t^{(c), n},
\]

(96)

\( G_t^{(c), r, n} \), \( j = 1, \ldots, D \) denotes the \( j \)-th element of \( G_t^{(c), r, n} \). \( G_t^{(r), n} = \frac{1}{r!} \partial_{\alpha^{(r)}} \bar{X}_t^{(r), n} \big|_{k=0} \), and \( G_t^{\psi, (c), n} \), \( j = 1, \ldots, D \) denote the \( j \)-th elements of \( G_t^{(c), n} \).
A.2 Asymptotic Bias and Variance of Our Control Variate Method (Proof of Theorem 3.3)

Hereafter, let us concentrate on getting the asymptotic bias and variance in the setting of Section 2 and 3. For a payoff function \( f : \mathbb{R} \to \mathbb{R} \), we can take \( f \in C^\infty(\mathbb{R}) \) satisfying

\[
|\mathbb{E}[f(q(S_T^{(c)}))] - \mathbb{E}[\tilde{f}(q(S_T^{(c)}))]| < \delta, \tag{97}
\]

for all \( \delta > 0 \). Since \( q \) is a linear function defined in (13), Taylor expansion of \( f \left(q \left(S_T^{(c)}\right)\right) \) around \( q \left(S_T^{[0]} + \epsilon S_T^{[1]}\right) \) leads to

\[
f \left(q \left(S_T^{(c)}\right)\right) = f \left(q \left(S_T^{[0]} + \epsilon S_T^{[1]}\right) + q \left(\frac{\epsilon^2}{2!} S_T^{[2]} + \frac{\epsilon^3}{3!} S_T^{[3]} + \cdots\right)\right) \nonumber
\]

\[
= f \left(q \left(S_T^{[0]} + \epsilon S_T^{[1]}\right)\right) + \frac{\epsilon^2}{2} \partial f \left(q \left(S_T^{[0]} + \epsilon S_T^{[1]}\right)\right) q \left(S_T^{[2]}\right) + O(\epsilon^3 c_1(T)e^{c_2(T)}). \tag{98}
\]

On the other hand, control variate \( g \left(q \left(\hat{S}_T\right)\right) \) is expressed as:

\[
g \left(q \left(\hat{S}_T^{(c)}\right)\right) = g \left(q \left(S_T^{[0]} + \epsilon S_T^{[1]}\right) + q \left(\frac{\epsilon^2}{2!} S_T^{[2]}\right)\right) \nonumber
\]

\[
= g \left(q \left(S_T^{[0]} + \epsilon S_T^{[1]}\right)\right) + \frac{\epsilon^2}{2} \partial g \left(q \left(S_T^{[0]} + \epsilon S_T^{[1]}\right)\right) q \left(S_T^{[2]}\right) + O(\epsilon^3 c_1(T)e^{c_2(T)}). \tag{99}
\]

When we take limits as \( f \to \tilde{f} \) defined in (24) and \( g \to \tilde{g} \) defined in (21), we obtain

\[
\tilde{f} \left(q \left(S_T^{(c)}\right)\right) = \left(q \left(S_T^{[0]} + \epsilon S_T^{[1]}\right) - K\right)^+ + \frac{\epsilon^2}{2} 1_{\{q \left(S_T^{[0]} + \epsilon S_T^{[1]}\right) > K\}} q \left(S_T^{[2]}\right) + O(\epsilon^3 c_1(T)e^{c_2(T)}). \tag{100}
\]

\[
\tilde{g} \left(q \left(S_T^{(c)}\right)\right) = \left(q \left(S_T^{[0]} + \epsilon S_T^{[1]}\right) - K\right)^+ + \frac{\epsilon^2}{2} 1_{\{q \left(S_T^{[0]} + \epsilon S_T^{[1]}\right) > K\}} q \left(S_T^{[2]}\right) + O(\epsilon^3 c_1(T)e^{c_2(T)}). \tag{101}
\]

where \( c_1(T) \) and \( c_2(T) \) are increasing functions of \( T \) with at most polynomial growth. Thus, we can choose \( f \) and \( g \) as \( a_{m,T}^f = a_{m,T}^g \) and \( b_{m,T}^f = b_{m,T}^g \), \( (m = 0, 1, 2) \).

Here, the bias of Euler-Maruyama scheme is expressed as:

\[
\left| \mathbb{E} \left[ \tilde{f} \left(q(S_T^{(c)})\right) - \tilde{f} \left(q(\hat{S}_T^{(c)})\right) \right] \right| \nonumber
\]

\[
\leq \left| \mathbb{E} \left[ f \left(\sum_{m=0}^{2} e^m a_{m,T}^f(\epsilon u)\right) - f \left(\sum_{m=0}^{2} e^m b_{m,T}^f(\epsilon u)\right) + \epsilon^3 \int_0^1 \frac{(1 - u)^2}{2!} \left(a_{3,T}^f(\epsilon u) - b_{3,T}^f(\epsilon u)\right) du \right] \right| + \delta \nonumber
\]

\[
= O \left(\frac{K(T)}{n}\right), \tag{102}
\]
where $\hat{K}(T)$ is an increasing function on $T$, From Condition (A) and assumptions of the function $f$, which satisfy the assumptions of Theorem 2.2. of [27], we obtain the estimate of an absolute value of the bias $(B(n, 2))$ of our control variate with $n$ time steps, the second order expansion and $c \equiv 1$ as follows:

\[
B(n, 2) = \left| \mathbb{E} \left[ \left\{ \tilde{f} \left( q(S_T^{(e)}) \right) - \left( \tilde{g} \left( q(S_T^{(e)}) \right) - \mathbb{E} \left[ \tilde{g} \left( q(S_T^{(e)}) \right) \right] \right) \right\} \right| \\
- \left\{ \tilde{f} \left( q(S_T^{(e), n}) \right) - \left( \tilde{g} \left( q(S_T^{(e), n}) \right) - \mathbb{E} \left[ \tilde{g} \left( q(S_T^{(e), n}) \right) \right] \right) \right\} \bigg| \\
= \left| \mathbb{E} \left[ \left\{ \tilde{f} \left( q(S_T^{(e)}) \right) - \tilde{f} \left( q(S_T^{(e), n}) \right) \right\} \right| \\
- \left\{ \tilde{g} \left( q(S_T^{(e)}) \right) - \tilde{g} \left( q(S_T^{(e), n}) \right) \right\} \bigg| \\
\leq \left| \mathbb{E} \left[ f \left( \sum_{m=0}^{2} \epsilon_m a^f_{m,T}(u) \right) - \left( \sum_{m=0}^{2} \epsilon_m b^f_{m,T}(u) \right) \right] + \delta \right| \\
- \left\{ g \left( \sum_{m=0}^{2} \epsilon_m a^g_{m,T}(u) \right) - g \left( \sum_{m=0}^{2} \epsilon_m b^g_{m,T}(u) \right) \right\} \bigg| + \delta \\
= \mathbb{E} \left[ \left\{ \int_0^1 \left( \frac{1-u}{2 \epsilon(u^2)} \right) \left( a^f_{3,T} - b^f_{3,T} \right) \right\} \bigg| + \delta \bigg| \\
- \left\{ \int_0^1 \left( \frac{1-u}{2 \epsilon(u^2)} \right) \left( a^g_{3,T} - b^g_{3,T} \right) \right\} \bigg| + \delta \\
= O \left( \frac{\epsilon^3 \mathcal{K}(T)}{n} \right), \tag{103} \right.
\]

where $\mathcal{K}(T)$ is an increasing function on $T$, and we applied $|a^f_{3,T} - b^f_{3,T}| = O \left( \frac{\mathcal{K}(T)}{n} \right)$ in Theorem 2.2 of [27]. Thus, the bias of our control variate method is expected to be reduced from the crude Monte Carlo method.

The variance of our control variate method $\text{Var}(M)$ with $c \equiv 1$ and $M$ sample paths is also estimated as:

\[
\text{Var}(M) = \text{Var} \left( \frac{1}{M} \sum_{j=1}^{M} \left( \tilde{f}(q(S_T^{(e), n})) - \tilde{g}(q(S_T^{(e), n})) - \mathbb{E}[\tilde{g}(q(S_T^{(e), n}))] \right) \right) \\
= \frac{1}{M} \text{Var} \left( \tilde{f}(q(S_T^{(e), n})) - \tilde{g}(q(S_T^{(e), n})) - \mathbb{E}[\tilde{g}(q(S_T^{(e), n}))] \right) \\
= \frac{1}{M} \text{Var} \left( \epsilon^3 \int_0^1 \left( \frac{(1-u)^2}{2 \epsilon(u^2)} \right) \left( b^f_{3,T} - b^g_{3,T} \right) \right) \\
= O \left( \frac{\epsilon^3 \mathcal{K}(T)}{M} \right). \tag{104} \right.
\]

where $\left( \tilde{f}(q(S_T^{(e), n})) - g(q(S_T^{(e), n})) - \mathbb{E}[g(q(S_T^{(e), n}))] \right)_j$ is the $j$-th sample of the distribution $\left( \tilde{f}(q(S_T^{(e), n})) - g(q(S_T^{(e), n})) - \mathbb{E}[g(q(S_T^{(e), n}))] \right)$. In the case of multi-regression estimator with stratified sampling, the variance is more improved because of the arguments in Section 3.2.
B Control Variates for Greeks

In this section, we provide a calculation scheme and validity on our control variate method for Greeks. As an example of asymptotic expansions of Greeks for European vanilla options, under Condition (A) in Appendix A.1, we illustrate an expansion of Delta ($\Delta_i$) with regard to the $i$-th initial data of $X_T$, which is defined as:

$$\Delta_i = \frac{\partial}{\partial x_i} E[\tilde{f}(X_T)],$$

where $\tilde{f} : \mathbb{R}^D \to \mathbb{R}$ is a payoff of European Option.

In order for that, we consider $\frac{\partial}{\partial x_i} E[f(X_T)]$ with a smooth function $f \in C^\infty_b$, instead of the original one $\tilde{f}$, because a non-smooth function is approximated by a smooth function, as well as (84) in Appendix A.1.

First, let us define $X^i_t := \frac{\partial}{\partial x_i} X^i_t$. Then, it is easily seen that the coefficients of the SDE $\dot{X}$ and SDE (82) have a graded structure and that under Condition (A),

$$dX_t = \mu(t, X_{t-})dt + \Phi(t, X_{t-})dW_t + \gamma(t, X_{t-})dZ_t; \quad X_0 = x$$

(106)

$$d\dot{X}^i_t = \frac{\partial}{\partial x_i} \mu(t, X_{t-})\dot{X}^i_{t-} dt + \frac{\partial}{\partial x_i} \Phi(s, X_{t-})\dot{X}^i_{t-} dW_t$$

$$+ \frac{\partial}{\partial x_i} \gamma(t, X_{t-})\dot{X}^i_{t-} dZ_t; \quad X^i_0 = 1.$$  

(107)

Hence, with the assumption $f \in C^\infty_b$, there exists a random variable $\tilde{Z}$ such that $E[\|\tilde{Z}\|] < \infty$ and

$$|f^{[1]}(X_T)\tilde{X}^i_T| < |\tilde{Z}|.$$

(108)

Therefore, we can interchange the order of the differentiation and expectation operators for Delta as:

$$\Delta_i = \frac{\partial}{\partial x_i} E[f(X_T)] = E[f^{[1]}(X_T)\tilde{X}^i_T].$$

(109)

Thus, all we need to do is evaluate the value $E[f^{[1]}(X_T)\tilde{X}^i_T]$.

By introducing a perturbation parameter $\epsilon \in (0, 1]$ to the processes $\dot{X}^i$ as in (83),

$$d\dot{X}^i_{t-} = \frac{\partial}{\partial x_i} \mu(t, X^i_{t-})\dot{X}^i_{t-} dt + \epsilon \frac{\partial}{\partial x_i} \Phi(t, X^i_{t-})\dot{X}^i_{t-} dW_t$$

$$+ \int_E \frac{\partial}{\partial x_i} \gamma(t, X^i_{t-})\dot{X}^i_{t-} d\tilde{Z}_t; \quad X^i_{0-} = 1.$$  

(110)

Since all coefficients of (110) satisfy Condition (A), we can expand $E[f^{[1]}(X_T^i)\tilde{X}^i_T]$ with Theorem 2.1 in [29] as follows:

$$E[f^{[1]}(X_T^i)\tilde{X}^i_T] = E\left[\sum_{n=0}^M \epsilon^n c_n, T + R(\epsilon^{M+1}, T) \left( \sum_{n=0}^{N} \frac{\epsilon^n}{n!} \frac{X^i_{t-}[^{[n]}]}{T} + \hat{R}(\epsilon^{N}, T) \right) \right],$$

(111)
where
\[ c_{n,T} := \sum_{n^{(r)}, r^{(r)}} \partial^{r}_{\alpha^{(r)}} f^{[1]}(X_T^{[0]} F_{T}^{d_1,[1]} \ldots F_{T}^{d_r,[r]}), \tag{112} \]
as in (87). Then, collecting the terms with the same order of \( \epsilon \), we obtain an asymptotic expansion of \( \Delta_i \):
\[ \Delta_i = E \left[ \sum_{m=0}^{M} \epsilon^m \hat{f}_{m,T} \right] + O(\epsilon^{M+1} c_1(T) \epsilon c_2(T)), \tag{113} \]
where \( c_1(T) \) and \( c_2(T) \) are increasing functions of \( T \) with at most polynomial growth, and
\[ \hat{f}_{m,T} := \sum_{n=0}^{m} \epsilon^n c_{n,T} \frac{\epsilon^{m-n}}{(m-n)!} X_T^{[m-n]}. \tag{114} \]

Therefore, the proxy process for Delta can be set as:
\[ \hat{X}_T^i = X_T^{i,[0]} + X_T^{i,[1]} + \frac{1}{2} X_T^{i,[2]}, \tag{115} \]
like (23), and the expectation of control variate is calculated as the same form of (26) (however, constants \( C_1, \ldots, C_3 \) are different from (26)).

We finally remark that asymptotic expansions of the other Greeks such as vega, theta and rho can be obtained in similar ways.

### B.1 Numerical Examples of Delta for Basket Options

In this subsection, we show the numerical examples of Delta of basket options for the same models (56) and (57) with parameters in Tables 1 and 2 in Section 4.1. Here, we apply the first order expansion (i.e. \( M = 1 \) in (113)) for the control variate because of the limitation of the conditional expectation formulas currently available in [28]: We note that for the first order expansion of Delta, we need similar types of conditional expectation formulas for the second order expansion of the option price derived by [28]. Table 9 shows the results.

<table>
<thead>
<tr>
<th>( K )</th>
<th>Price</th>
<th>Error</th>
<th>VRF</th>
</tr>
</thead>
<tbody>
<tr>
<td>( NC )</td>
<td>( AE )</td>
<td>( MC )</td>
<td>( NC )</td>
</tr>
<tr>
<td>80</td>
<td>0.1639</td>
<td>0.1669</td>
<td>0.0009</td>
</tr>
<tr>
<td>100</td>
<td>0.0961</td>
<td>0.1047</td>
<td>0.0011</td>
</tr>
<tr>
<td>120</td>
<td>0.0527</td>
<td>0.0530</td>
<td>0.0010</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( K )</th>
<th>Time (sec)</th>
<th>( EF )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( MC )</td>
<td>( NC )</td>
<td>( OA )</td>
</tr>
<tr>
<td>80</td>
<td>21</td>
<td>32</td>
</tr>
<tr>
<td>100</td>
<td>21</td>
<td>32</td>
</tr>
<tr>
<td>120</td>
<td>21</td>
<td>32</td>
</tr>
</tbody>
</table>

Table 9: Delta
The notations in the table are the same as those in Section 4.2. The absolute value of the variance of the delta is smaller than that of the premium because the range of delta is only from 0 to 1. However, our control variate method can reduce the variance.