Accepted Manuscript

A multi-level dimension reduction Monte-Carlo method for jump-diffusion models

Duy-Minh Dang

PII: S0377-0427(17)30169-3
DOI: http://dx.doi.org/10.1016/j.cam.2017.04.014
Reference: CAM 11093

To appear in: *Journal of Computational and Applied Mathematics*

Received date: 3 May 2016

A multi-level dimension reduction Monte-Carlo method for jump-diffusion models *

Duy-Minh Dang†

April 12, 2017

Abstract

This paper develops and analyses convergence properties of a novel multi-level Monte-Carlo (mlMC) method for computing prices and hedging parameters of plain-vanilla European options under a very general b-dimensional jump-diffusion model, where b is arbitrary. The model includes stochastic variance and multi-factor Gaussian interest short rate(s). The proposed mlMC method is built upon (i) the powerful dimension and variance reduction approach developed in Dang et al. (2017) for jump-diffusion models, which, for certain jump distributions, reduces the dimensions of the problem from b to 1, namely the variance factor, and (ii) the highly effective multi-level MC approach of Giles (2008) applied to that factor. Using the first-order strong convergence Lamperti-Backward-Euler scheme, we develop a multi-level estimator with variance convergence rate $\mathcal{O}(h^2)$, resulting in an overall complexity $\mathcal{O}(\epsilon^{-2})$ to achieve a root-mean-square error of ϵ. The proposed mlMC can also avoid potential difficulties associated with the standard multi-level approach in effectively handling simultaneously both multi-dimensionality and jumps, especially in computing hedging parameters. Furthermore, it is considerably more effectively than existing mlMC methods, thanks to a significant variance reduction associated with the dimension reduction. Numerical results illustrating the convergence properties and efficiency of the method with jump sizes following normal and double-exponential distributions are presented.

Keywords: Monte Carlo, variance reduction, dimension reduction, multi-level, jump-diffusions, Lamperti-Backward-Euler, Milstein

AMS Classification 65C05, 78M31, 80M31, 42A38, 37M05

1 Introduction

In mathematical finance, Monte-Carlo (MC) is a very popular computational approach, especially for high-dimensional stochastic models. This is primarily due to the fact that the complexity of MC methods increases linearly with respect to the number of dimensions. However, it is also well-known that MC methods typically converge at a rate proportional to $M^{-\frac{1}{2}}$, where M is the number of paths in the MC simulation. As a result, the main challenge in developing an efficient MC method is often to find an effective variance reduction technique. We refer the reader to

*This research was supported in part by a University of Queensland Early Career Researcher (ECR) Grant [Grant number 006301-01-298-21-609775].

†School of Mathematics and Physics, The University of Queensland, St Lucia, Brisbane 4072, Australia, email: duyminh.dang@uq.edu.au

Glasserman (2003) and relevant references therein for a detailed discussion on various variance reduction techniques. Using an ordinary MC approach with a (time) discretization scheme having first-order weak convergence, such as the Euler-Maruyama scheme, the computational complexity to achieve a root-mean-square error (RMSE) of ϵ is $\mathcal{O}(\epsilon^{-3})$ (Duffie and Glynn, 1995).

The multi-level MC (mlMC) approach, developed in Giles (2008), is based on the multi-grid idea for iterative solutions of partial differential equations (PDEs), but applied to MC *path* calculations. More specifically, the mlMC approach combines simulations with different numbers of timestep sizes to achieve the same level of accuracy obtained by the ordinary MC approach at the finest timestep size, but at a much lower computational cost. It is well-known that the efficiency of a mlMC method primarily depends on the strong convergence of the scheme used to discretize the underlying processes (see, for example, Giles et al. (2013); Giles and Szpruch (2014), among several others). More specifically, with a time discretization scheme that has first-order strong convergence, such as the Milstein (Kloeden and Platen, 1992) or the Lamperti-Backward-Euler (LBE) (Neuenkirch and Szpruch, 2014) schemes, to achieve a RMSE of ϵ, the computational complexity is reduced to $\mathcal{O}(\epsilon^{-2})$ for European options with Lipschitz continuous payoffs. This significant complexity reduction can also be achieved for discontinuous and path-dependent payoffs, but requires careful treatment and special estimators, as discussed in Giles (2006). This reduction a significant computational complexity saving compared to the Euler-Maruyama scheme which has only half-order strong convergence, and hence $\mathcal{O}(\epsilon^{-2}(\log(\epsilon))^2)$ computational complexity (Giles, 2008).

There is much interest in the computational finance community in using mlMC with the Milstein scheme. See, for example, the series of works by Giles and coauthors in Giles (2006); Giles et al. (2013); Giles and Szpruch (2014). The popularity of the Milstein scheme is primarily due to its well-established first-order strong convergence results (Kloeden and Platen, 1992). However, a disadvantage of the Milstein scheme is that, for multi-dimensional models, except in some special cases, to achieve an overall complexity $\mathcal{O}(\epsilon^{-2})$ for a RMSE of ϵ, it usually requires simulation of iterated Itô integrals, also known as Lévy areas, and this is usually very slow. In Giles and Szpruch (2014), it is shown that, through the construction of a suitable antithetic mlMC estimator, it is possible to avoid simulating Lévy areas, but still achieve an overall complexity $\mathcal{O}(\epsilon^{-2})$ for a RMSE of ϵ. To the best of our knowledge, this is the only mlMC method that can *effectively* deal with multi-dimensional models. Nonetheless, this method still requires multi-dimensional MC simulations. In addition to the Milstein scheme, the LBE scheme, recently studied in Neuenkirch and Szpruch (2014), also has first-order strong convergence and positivity preserving properties. Applications of this scheme in a context of mlMC setting, however, have not been studied.

All above mlMC methods are developed for pure-diffusion models. However, from a modelling point of view, a jump-diffusion model combined with stochastic volatility, and possibly (multi-factor) interest rate(s), can capture more faithfully important empirical phenomena, such as the observed volatility smile/skew for both short and long maturities. See discussions in, for example, Alizadeh et al. (2002); Andersen et al. (2002); Bakshi et al. (2000, 1997); Bates (1996), among many others. The implied volatility smile/skew phenomena are present in various asset classes, such as equity and foreign exchange (FX). Moreover, from a risk-management point of view, it is important to model jumps in the underlying asset prices to account for "crash" effects. However, the current literature on mlMC methods for jump-diffusion processes is rather under-developed, with focus on only *one-dimensional* jump-diffusion models (Xia, 2011, 2013; Xia and Giles, 2012). Furthermore, in all of these works, only the normal jump distribution of Merton (1976) is considered, with virtually no discussions of other popular jump distributions, such as the double-exponential distribution of Kou (2002).

The common thread in the solution techniques proposed in the above mlMC works for one-dimensional jump-diffusion models is to develop a jump-adapted Milstein scheme. It appears possible to extend this approach to multi-factor jump-diffusion models; however, the major challenge would be to develop a multi-dimensional version of the jump-adapted Milstein scheme in combination of the antithetic mlMC method developed in Giles and Szpruch (2014) so that simulation of the Lévy areas can be avoided. Based on the current mlMC literature, this possible extension appears to be the only way that can effectively handle simultaneously both multi-dimensionality and jumps. Nonetheless, this approach still requires multi-dimensional MC simulations. In addition, as well-noted in the mlMC literature, this approach may have difficulties in computing hedging parameters for jump-diffusion models, especially high-order ones, such as Gamma, due to lack of smoothness in the payoff (Burgos and Giles, 2012).

Along a different line of MC research, in Dang et al. (2015a), we develop a powerful and easy-to-implement dimension reduction approach for MC methods, referred to as drMC, for plain-vanilla European options under a very general b-dimensional pure-diffusion model, where b is arbitrary. This general model includes stochastic variance/volatility and (multi-factor) Gaussian interest short rate(s). The underlying idea of the drMC approach of Dang et al. (2015a) is to combine (i) the conditional MC technique applied to the variance factor, and (ii) a derivation of a Black-Scholes-Merton type closed-form solution of an associated conditional Partial Differential Equation (PDE) via a Fourier transform technique. Results of Dang et al. (2015a) show that the option price can be computed simply by taking the expectation of this closed-form solution. Hence, the drMC approach results in a powerful dimension reduction from b to only one, namely the variance factor. This dimension reduction often results in a significant variance reduction as well, since the variance associated with the other $(b-1)$ factors in the original model are completely removed from the drMC simulation.

In Dang et al. (2017), we extend the drMC framework developed in Dang et al. (2015a) to handle jumps in the underlying asset. One of the major findings of Dang et al. (2017) is that the analytical tractability of the associated conditional Partial Integro-Differential Equation (PIDE) is fully determined by that of the (well-studied) Black-Scholes-Merton model augmented with the *same* jump components as the model under investigation. As a result, for certain jump distributions, such as the normal (Merton, 1976) and the double-exponential (Kou, 2002) distributions, the option price under the above-mentioned very general jump-diffusion model can be simply expressed as an expectation of an analytical solution to the conditional PIDE, which depends only on the variance path. The option's hedging parameters can also be computed very efficiently in the same fashion as the option price.

In this paper, we propose and analyse the convergence properties of a novel mlMC method for computing the price and hedging parameters for plain-vanilla European options under the above-described general jump-diffusion model. The proposed method essentially consists of two stages. In the first stage, by applying the drMC method of Dang et al. (2017), we reduce the dimension of the pricing problem from b to only one, namely the variance factor. In the second stage, we apply the mlMC technique with a first-order strong convergence scheme, such as the Milstein or the LBE schemes, to the stochastic variance factor on which we condition in the first stage. We refer to the proposed MC method as multi-level drMC (ml-drMC).

The main contributions of this paper are

- The proposed ml-drMC method is the *first* multi-level based MC method reported in the literature that can effectively handle simultaneously both multi-dimensionality of the pricing problem and jumps in the underlying asset, especially in computing hedging parameters.

The ml-drMC method naturally avoids the above-mentioned difficulties of the standard mlMC approach in this case by handling effectively these issues in a separate stage using the drMC technique. Moreover, the proposed method is easy to implement, and can readily handle different jump distributions.

- We show that the closed-form solution of the conditional PIDE, i.e. the payoff, is a Lipschitz function of the values of its variables. We then construct a multi-level estimator based on the first-order strong convergence LBE scheme (Neuenkirch and Szpruch, 2014), and show that the multi-level variance converges at rate $\mathcal{O}(h^2)$. By a general complexity result in Giles (2008), the proposed ml-drMC method requires only an overall complexity $\mathcal{O}(\epsilon^{-2})$ to achieve a RMSE of ϵ. These convergence and complexity results hold for both price and hedging parameters, such as Delta and Gamma.

- Since the application of the drMC technique in first stage of the ml-drMC method often results in a significant variance reduction, it is expected that the ml-drMC approach is significantly more efficient than the antithetic mlMC based approach of Giles and Szpruch (2014) when applied to pricing plain-vanilla European options under (j ump-) diffusion models with stochastic variance and (multi-factor) Gaussian interest rates.

The remainder of the paper is organized as follows. We start by introducing a general pricing model and reviewing the drMC approach in Sections 2 and 3, respectively. In Section 4, we discuss the ml-drMC method in detail. The convergence results are proven in Section 5. In Section 6, numerical results with a 3-factor equity model and a 6-factor FX mode are presented to illustrate the convergence properties of the ml-drMC method and its efficiency. Section 7 concludes the paper and outlines possible future work.

2 A general pricing model

We consider an (international) economy consisting of $c+1$ markets (currencies), $c \in \{0,1\}$, indexed by $i \in \{d, f\}$, where "d" stands for the domestic market (Dang et al., 2017). We consider a complete probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{Q})$, with sample space Ω, sigma-algebra $\mathcal{F}$, filtration $\{\mathcal{F}_t\}_{t\geq 0}$, and "$d$" risk-neutral measure $\mathbb{Q}$ defined on $\mathcal{F}$. We denote by $\mathbb{E}$ the expectation taken under $\mathbb{Q}$ measure. Let the underlying asset $S(t)$, its instantaneous variance $\nu(t)$, and the two short rates $r_d(t)$ and $r_f(t)$ be governed by the following SDEs under the measure $\mathbb{Q}$:

$$\frac{\mathrm{d}S(t)}{S(t^-)} = (r_d(t) - c\, r_f(t) - \lambda\delta)\ \mathrm{d}t + \sqrt{\nu(t)}\,\mathrm{d}W_s(t) + \mathrm{d}J(t), \tag{2.1a}$$

$$r_d(t) = \sum_{i=1}^{m} X_i(t) + \gamma_d(t),$$
$$\text{with } \mathrm{d}X_i(t) = -\kappa_{d_i}(t)\,X_i(t)\,\mathrm{d}t + \sigma_{d_i}(t)\,\mathrm{d}W_{d_i}(t), \quad X_i(0) = 0, \tag{2.1b}$$

$$r_f(t) = \sum_{i=1}^{l} Y_i(t) + \gamma_f(t),$$
$$\text{with } \mathrm{d}Y_i(t) = -\kappa_{f_i}(t)\,Y_i(t)\,\mathrm{d}t + \sigma_{f_i}(t)\,\mathrm{d}W_{f_i}(t) - \rho_{s,f_i}\sigma_{f_i}(t)\sqrt{\nu(t)}\,\mathrm{d}t, \quad Y_i(0) = 0, \tag{2.1c}$$

$$\mathrm{d}\nu(t) = \kappa_\nu\,(\bar{\nu} - \nu(t))\ \mathrm{d}t + \sigma_\nu\sqrt{\nu(t)}\,\mathrm{d}W_\nu(t)\ . \tag{2.1d}$$

We work under the following assumptions for model (2.1).

- Processes $W_s(t)$, $W_{d_i}(t)$, $i = 1, \ldots, m$, $W_{f_i}(t)$, $i = 1, \ldots, l$, and $W_\nu(t)$ are correlated Brownian motions (BMs) with a constant correlation coefficient $\rho_{(\cdot)(\cdot)} \in [-1, 1]$ between each BM pair.

- The process $J(t) = \sum_{j=1}^{\pi(t)} (y_j - 1)$ is a compound Poisson process. Specifically, $\pi(t)$ is a Poisson process with a constant finite jump intensity $\lambda > 0$, and y_j, $j = 1, 2, \ldots$, are independent and identically distributed (i.i.d.) positive random variables representing the jump amplitude, and having the density $g(\cdot)$.

  Several popular cases for $g(\cdot)$ are (i) the log-normal distribution given in Merton (1976), and (ii) the log-double-exponential distribution given in Kou (2002). When a jump occurs at time t^-, we have $S(t) = yS(t^-)$, where t^- is the instant of time just before the time t. In (2.1a), $\delta = \mathbb{E}[y - 1]$ represents the expected percentage change in the underlying asset price.

- The Poisson process $\pi(t)$, and the sequence of random variables $\{y_j\}_{j=1}^{\infty}$ are mutually independent, as well as independent of the BMs $W_s(t)$, $W_{d_i}(t)$, $i = 1, \ldots, m$, $W_{f_i}(t)$, $i = 1, \ldots, l$, and $W_\nu(t)$.

The functions $\kappa_{d_i}(t)$, $\sigma_{d_i}(t)$, $i = 1, \ldots, m$, $m \geq 1$, $\kappa_{f_i}(t)$, and $\sigma_{f_i}(t)$, $i = 1, \ldots, l$, $l \geq 1$, are strictly positive deterministic functions of t, with $\kappa_{d_i}(t)$, and $\kappa_{f_i}(t)$ being the positive mean-reversion rates. The functions $\gamma_d(t)$ and $\gamma_f(t)$ are also deterministic, and they, respectively, capture the "*d*" and "*f*" current term structures. They are defined as

$$\gamma_i(t) = r_i(0)\, e^{-\kappa_{i_1} t} + \kappa_{i_1} \int_0^t e^{-\kappa_{i_1}(t-s)}\, \theta_i(s)\, \mathrm{d}s, \quad i \in \{d, f\}, \tag{2.2}$$

where θ_i are deterministic, and represent the interest rates' mean levels. In addition, κ_ν, σ_ν and $\bar{\nu}$ are also positive constants.

The constant c takes on the value of either zero or one, and essentially serves as an on/off switch of the "*f*" economy. That is, by setting $c = 0$, the model (2.1) reduces to an option pricing model in a single market. It can be used for stock options, in which case, $S(t)$ denotes the underlying stock price. When $c = 1$, the model (2.1) becomes a FX model, with indexes "*d*" and "*f*" respectively denoting the domestic and foreign markets (currencies). In this case, $S(t)$ denotes the spot FX rate, which is defined as the number of units of "*d*" currency per one unit of "*f*" currency.

We emphasize the generality of the model. A number of widely used pricing models are a special case of (2.1). For example, for stock options, (2.1) covers the Heston model due to Heston (1993), its jump-extension, or the Bates model (Bates, 1996), as well as the popular (3D) Heston-Hull-White (HHW) equity model used in Grzelak and Oosterlee (2012b); Haentjens and in 't Hout (2012). For FX options, the widely used four-factor model with stochastic volatility and one-factor Gaussian interest rates is also a special case of (2.1) (see, for example, Grzelak and Oosterlee (2011, 2012a); Haastrecht et al. (2009); Haastrecht and Pelsser (2011)).

3 Review of the dimension reduction MC method

Denote by $b = m + 2 + c\, l$, where $c \in \{0, 1\}$, the total number of stochastic factors in the model. As the first step, we decompose the (correlated) BM processes into a linear combination of independent

BM processes $\widetilde{W}_i(t)$, $i = 1, \ldots, b$. The decomposition is as follows

$$
\begin{aligned}
c = 0: \quad & \Big(W_s(t), W_{d_1}(t), \ldots W_{d_m}(t), W_\nu(t)\Big)^\top \\
& \qquad = \mathbf{A}\left(\widetilde{W}_1(t), \widetilde{W}_2(t), \ldots, \widetilde{W}_{b-1}(t), \widetilde{W}_b(t) \right)^\top, \\
c = 1: \quad & \Big(W_s(t), W_{d_1}(t), \ldots W_{d_m}(t), W_{f_1}(t), \ldots, W_{f_l}(t), W_\nu(t)\Big)^\top \\
& \qquad = \mathbf{A}\left(\widetilde{W}_1(t), \widetilde{W}_2(t), \ldots, \widetilde{W}_{m+1}(t), \widetilde{W}_{m+2}(t), \ldots, \widetilde{W}_{b-1}(t), \widetilde{W}_b(t) \right)^\top.
\end{aligned}
\tag{3.1}
$$

Here, $\mathbf{A} \equiv [a_{ij}] \in \mathbb{R}^{b\times b}$, obtained using a Cholesky factorization, is an upper triangular matrix with $a_{b,b} = 1$. The normalization condition on the correlation matrix requires $\sum_{j=1}^{b} a_{i,j}^2 = 1$ for each row.

We denote by

$$V(S(t), t, \cdot) \equiv V(S(t), t, r_d(t), r_f(t), \nu(t))$$

the price at time t of a plain-vanilla European option under the model (2.1) with payoff $\Phi(S(T))$. We further assume that the payoff $\Phi(x)$ is a continuous function of its argument having at most polynomial (sub-exponential) growth, which is satisfied in the case of call and put options.

In the following, we briefly review the dimension reduction MC approach for the jump-diffusion model (2.1). The reader is referred to Dang et al. (2015a, 2017) for detailed discussions of the approach and relevant proofs. Using standard arbitrage theory (Delbaen and Schachermayer, 1994), and the "tower property" of the conditional expectation, the option price under the general model (2.1) can be expressed as two-level nested expectation, with the inner expectation being conditioned on the filtration associated with $\widetilde{W}_i(t)$, $i = 2, \ldots, b$. More specifically,

$$V(S(0), 0, \cdot) = \mathbb{E}\left[\mathrm{e}^{-\int_0^T r_d(t)\,\mathrm{d}t}\Phi(S(T))\right] = \mathbb{E}\left[\mathbb{E}\left[\mathrm{e}^{-\int_0^T r_d(t)\,\mathrm{d}t}\Phi(S(T))\,\middle|\, \left\{\widetilde{W}_i(\tau)\right\}_{i=2}^{b}\right]\right], \tag{3.2}$$

where $\left\{\widetilde{W}_i(\tau)\right\}_{i=2}^{b} \equiv \left\{\widetilde{W}_i(\tau; 0 \le \tau \le T)\right\}_{i=2}^{b}$ denotes the filtration generated by the corresponding BMs. The focus of the drMC method developed in Dang et al. (2015a, 2017) is primarily on the development of an analytical evaluation of the inner expectation, whereas the outer expectation is approximated by the usual means of MC simulation. The application of the multi-level technique is on the outer expectation, and this is the focus of the next section.

3.1 Step 1: conditional PIDE and solution via Fourier transform

Under certain regularity conditions, which are satisfied in the present case, by the Feynman-Kac theorem for jump-diffusion processes (Cont and Tankov, 2004), the inner expectation of (3.2) can be shown to be equal to the unique solution to an associated (conditional) PIDE. Specifically, under log variables $z = \ln(S)$ and $\omega = \ln(y)$, and letting $v(z, 0, \cdot) = V(S, 0, \cdot)$, it can be shown that

$$v\left(z(0), 0, \cdot\right) = \mathbb{E}\left[u\left(z(0), 0; \left\{\widetilde{W}_i\right\}_{i=2}^{b}\right)\right], \tag{3.3}$$

where $u\left(z, t; \left\{\widetilde{W}_i\right\}_{i=2}^{b}\right)$ is the time-t solution of an associated (conditional) PIDE.

To solve the conditional PIDE, we first transform it into the Fourier space to obtain an ordinary differential equation in $\hat{u}(\xi, t, \cdot)$, which is the Fourier transform of $u(z, t, \cdot)$. This ordinary differential equation can then be easily solved in closed-form from maturity $t = T$ to time $t = 0$ to obtain $\hat{u}(\xi, 0; \cdot)$. It turns out that

$$\hat{u}\left(\xi, 0; \left\{\widetilde{W}_i(\tau)\right\}_{i=2}^{b}\right) = \hat{\phi}(\xi)\exp\Bigg(-\xi^2 \int_0^T \frac{a_{11}^2}{2}\nu(t)\,\mathrm{d}t + i\xi \int_0^T \left(r_d(t) - cr_f(t) - \lambda\delta - \frac{\nu(t)}{2}\right)\mathrm{d}t + i\xi \sum_{j=2}^{b} a_{1j} \int_0^T \sqrt{\nu(t)}\,\mathrm{d}\widetilde{W}_j(t) - \int_0^T (r_d(t) + \lambda)\,\mathrm{d}t + \int_0^T \lambda\Gamma(\xi)\mathrm{d}t\Bigg), \tag{3.4}$$

where $\hat{\phi}(\xi)$ is the Fourier transform of $\phi(z) = \Phi(\mathrm{e}^z)$, and $\Gamma(\xi)$ the characteristic function of $\ln(y)$.

3.2 Step 2: dimension reduction

The next step in our dimension reduction MC approach is to express $\mathbb{E}[\hat{u}(\xi, 0; \cdot)]$ as an expectation of a quantity that depends *only* on the $\{\widetilde{W}_b(\tau)\} \equiv \{W_\nu(\tau)\}$, which is the filtration generated by the BM associated with the variance factor. First, we apply iterated conditional expectation to obtain

$$\mathbb{E}[\hat{u}(\xi, 0; \cdot)] = \mathbb{E}\left[\mathbb{E}\left[\hat{u}(\xi, 0; \cdot)\middle|\{\widetilde{W}_b(\tau)\}\right]\right], \tag{3.5}$$

where $\hat{u}(\xi, 0; \cdot)$ is defined in (3.4). Then, we handle the terms $\exp\left(\int_0^T r_i(t)\mathrm{d}t\right)$, $i = d, f$, present in $\hat{u}(\xi, 0; \cdot)$, see (3.4), as follows. Using the Gaussian dynamics of the interest rates and the decomposition (3.1), we express $\int_0^T r_i(t)\mathrm{d}t$, $i = d, f$, as a sum of of Itô integrals involving independent BMs $\widetilde{W}_j$, $j = 2, \ldots, b$. As a result, the expectation of exponential terms involves these Itô integrals in $\mathbb{E}\left[\hat{u}(\xi, 0; \cdot)|\{\widetilde{W}_b(\tau)\}\right]$ can be factored out and evaluated in closed-form. The step results in the following expression for the transformed option price $\hat{v}(\xi, 0, \cdot)$

$$\hat{v}(\xi, 0, \cdot) = \mathbb{E}[\hat{u}(\xi, 0; \cdot)] = \mathbb{E}\left[\hat{\phi}(\xi)\exp\left(-G\,\xi^2 + iF\xi + H + \lambda T\Gamma(\xi)\right)\right], \tag{3.6}$$

where the coefficients G, F, and H are given by

$$G = \frac{a_{11}^2}{2}\int_0^T \nu(t)\,\mathrm{d}t + \frac{1}{2}\sum_{k=2}^{b-1}\int_0^T \left(\sum_{j=1}^{m} a_{(j+1),k}\,\beta_{d_j}(t) - c\sum_{j=1}^{l} a_{(j+m+1),k}\,\beta_{f_j}(t) + a_{1,k}\sqrt{\nu(t)}\right)^2 \mathrm{d}t, \tag{3.7a}$$

$$\begin{aligned} F = &-\frac{1}{2}\int_0^T \nu(t)\,\mathrm{d}t + \int_0^T (\gamma_d(t) - c\gamma_f(t))\,\mathrm{d}t \\ &- \sum_{k=2}^{b-1}\int_0^T \left(\sum_{j=1}^{m} a_{(j+1),k}\,\beta_{d_j}(t)\left(\sum_{j=1}^{m} a_{(j+1),k}\,\beta_{d_j}(t) - c\sum_{j=1}^{l} a_{(j+m+1),k}\,\beta_{f_j}(t)\right)\right)\mathrm{d}t \\ &+ \sum_{j=1}^{m} a_{(j+1),h}\int_0^T \beta_{d_j}(t)\,\mathrm{d}W_\nu(t) - c\sum_{j=1}^{l} a_{(j+m+1),h}\int_0^T \beta_{f_j}(t)\,\mathrm{d}W_\nu(t) \\ &+ a_{1,h}\int_0^T \sqrt{\nu(t)}\,\mathrm{d}W_\nu(t) + c\sum_{j=1}^{l}\rho_{s,f_j}\int_0^T \beta_{f_j}(t)\sqrt{\nu(t)}\,\mathrm{d}t - \sum_{k=2}^{b-1}\sum_{j=1}^{m}\int_0^T a_{1,k}a_{(j+1),k}\,\beta_{d_j}(t)\sqrt{\nu(t)}\,\mathrm{d}t \\ &- \lambda\delta T, \end{aligned} \tag{3.7b}$$

$$H = -\sum_{j=1}^{m} a_{(j+1),h} \int_0^T \beta_{d_j}(t)\,\mathrm{d}W_\nu(t) - \int_0^T \gamma_d(t)\,\mathrm{d}t + \frac{1}{2}\sum_{k=2}^{b-1}\int_0^T \left(\sum_{j=1}^{m} a_{(j+1),k}\beta_{d_j}(t)\right)^2 \mathrm{d}t - \lambda T, \tag{3.7c}$$

In (3.7a)-(3.7c), $\beta_{d_i}(t)$, $i = 1, \ldots, m$, and $\beta_{f_i}(t)$, $i = 1, \ldots, l$, are defined as

$$\beta_{d_i}(t) = \sigma_{d_i}(t)\int_t^T \mathrm{e}^{-\int_t^{t'} \kappa_{d_i}(t'')\,\mathrm{d}t''}\,\mathrm{d}t', \quad \beta_{f_i}(t) = \sigma_{f_i}(t)\int_t^T \mathrm{e}^{-\int_t^{t'} \kappa_{f_i}(t'')\,\mathrm{d}t''}\,\mathrm{d}t'\ . \tag{3.8}$$

We emphasize that the quantities F, G, H are conditional on the variance path only. The variance coming from the r_d's BMs and the r_f's BMs, if any, is completely removed from the computation. Thus, the drMC method not only offers a powerful dimension reduction from b factors to at most two, namely the S and ν factors, but it also significantly reduces the variance in the simulated results in many cases.

3.3 Step 3: inverse Fourier transform

The final step in the approach is to inverse the result in (3.6) back to the real space to obtain the option price. When $\lambda = 0$, i.e. the pricing model (2.1) reduces to a pure-diffusion model, a closed-form solution to the conditional PDE for a plain-vanilla European option can be obtained. More specifically, results in (Dang et al., 2015a) show that, for a European call option, we have

$$V(S(0), 0, \cdot) = \mathbb{E}[P], \quad \text{where} \quad P = S(0)\mathrm{e}^{(G+F+H)}\mathcal{N}(d_1) - K\mathrm{e}^{H}\mathcal{N}(d_2)\,. \tag{3.9}$$

Here,

$$d_1 = \frac{\ln\left(\frac{S(0)}{K}\right) + F}{\sqrt{2G}} + \sqrt{2G}, \quad d_2 = d_1 - \sqrt{2G}, \quad \mathcal{N}(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \mathrm{e}^{-v^2/2}\mathrm{d}v\ . \tag{3.10}$$

When $\lambda > 0$, the analytical tractability of the conditional PIDE depends on the distribution of the jump amplitude y, or equivalently, on that of $w = \ln(y)$. It is shown in Dang et al. (2017) that the analytical tractability of the conditional PIDE is fully determined by that of the (well-studied) Black-Scholes-Merton model augmented with the *same* jump component $\mathrm{d}J(t)$ as in model (2.1). In particular, in the case $w = \ln(y) \sim \text{Normal}(\tilde{\mu}, \tilde{\sigma}^2)$ (Merton, 1976), the European call option value is given by (Dang et al., 2017)[Corollary 3.2]

$$V(S(0), 0, \cdot) = \mathbb{E}\left[\sum_{n=0}^{\infty} \frac{(\lambda T)^n}{n!}\left\{\exp\left(n\tilde{\mu} + \frac{n\tilde{\sigma}^2}{2}\right) S(0)\mathrm{e}^{(G+F+H)}\mathcal{N}(d_{1,n}) - K\mathrm{e}^{H}\mathcal{N}(d_{2,n})\right\}\right], \tag{3.11}$$

where

$$d_{1,n} = \frac{\ln\left(\frac{S(0)}{K}\right) + n\tilde{\mu} + F}{\sqrt{2\left(G + \frac{n\tilde{\sigma}^2}{2}\right)}} + \sqrt{2\left(G + \frac{n\tilde{\sigma}^2}{2}\right)}, \quad d_{2,n} = d_{1,n} - \sqrt{2\left(G + \frac{n\tilde{\sigma}^2}{2}\right)}\ . \tag{3.12}$$

The Delta and Gamma of the option respectively are (Dang et al., 2017)[Corollary 4.2]

$$\begin{aligned} \left.\frac{\partial V}{\partial S}\right|_{(S(0),0,\cdot)} &= \mathbb{E}\left[\sum_{n=0}^{\infty} \frac{(\lambda T)^n}{n!}\left\{\exp\left(n\tilde{\mu} + \frac{n\tilde{\sigma}^2}{2} + G + F + H\right)\mathcal{N}(d_1)\right\}\right], \\ \left.\frac{\partial^2 V}{\partial S^2}\right|_{(S(0),0,\cdot)} &= \mathbb{E}\left[\sum_{n=0}^{\infty} \frac{(\lambda T)^n}{n!}\left\{\exp\left(n\tilde{\mu} + \frac{n\tilde{\sigma}^2}{2} + G + F + H\right)\frac{\mathcal{N}'(d_1)}{S(0)\sqrt{2\left(G + \frac{n\tilde{\sigma}^2}{2}\right)}}\right\}\right]. \end{aligned} \tag{3.13}$$

In our analysis, for simplicity, we focus on the normal jump case. For the case of double-exponential distribution (Kou, 2002), the analytical solution to the conditional PIDE is presented in Dang et al. (2017)[Corrolary 3.1], and is repeated in Appendix D.

4 Multi-level drMC

The previous results show that, for a jump-distribution of $\ln(y)$ such that the conditional PIDE is analytically tractable, i.e. the inner expectation of (3.2) can be evaluated analytically, the option price can be expressed as an expectation of this analytical solution. This solution involves only the variance factor. The application of the multi-level technique is on the outer expectation of (3.2), and this is the focus of this section.

In the ml-drMC method, we apply the multi-level technique to the variance factor $\nu(t)$, which is driven by the BM $\widetilde{W}_b(t)$. For simplicity, for the rest of the paper, let $W(t) \equiv \widetilde{W}_b(t)$. In this paper, to simulate $\nu(t)$, we use the so-called Lamperti-Backward-Euler (LBE) discretization scheme, studied in Neuenkirch and Szpruch (2014). Given a timestep size $h = T/N$, the LBE discretization scheme for the variance process (2.1d) is given by (Neuenkirch and Szpruch, 2014)

$$\hat{\nu}_{n+1} = \quad (\hat{z}_{n+1})^2, \tag{4.1}$$

where

$$\hat{z}_{n+1} = \frac{1}{2+\kappa_\nu\, h}\left(\hat{z}_n + \frac{1}{2}\sigma_\nu \Delta W_n + \sqrt{\left(\hat{z}_n + \frac{1}{2}\sigma_\nu \Delta W_n\right)^2 + \kappa_\nu\left(\bar{\nu} - \frac{\sigma_\nu^2}{4\kappa_\nu}\right)h}\right), \quad \hat{z}_0 = \sqrt{v(0)}\ .$$

Here, $\hat{\nu}_n$ denotes the discrete approximation to the exact value $\nu(t_n)$, where $t_n = nh$, $n = 0, \ldots, N-1$, $\Delta W_n = W_{n+1} - W_n = \text{Normal}(0, h)$. As shown in Neuenkirch and Szpruch (2014), we have the following result on the strong convergence with order one of the LBE scheme.

Proposition 4.1 (Proposition 3.1 of Neuenkirch and Szpruch (2014))**.** *Let $T > 0$ and $2 \le p < \frac{4\kappa_\nu\bar{\nu}}{3\sigma_\nu^2}$, there exists a bounded constant C_p such that*

$$\mathbb{E}\left[\sup_{n=0,\ldots,\lceil T/h\rceil} |v(t_n) - \hat{v}_n|^p\right] \le C_p h^p.$$

In our context, we are primarily interested in the above result for the case $p = 2$. For this special case, as required in the above proposition, the condition $p = 2 < \frac{4\kappa_\nu\bar{\nu}}{3\sigma_\nu^2}$ must hold.

Assumption 4.1. *We assume that the parameters of the process $\nu(t)$, defined in* (2.1d)*, are such that $2\kappa_\nu\bar{\nu} > 3\sigma_\nu^2$.*

We note that this assumption is slightly stricter than the Feller's condition $2\kappa_\nu\bar{\nu} > \sigma_\nu^2$ which guarantees that $\nu(t) > 0$ and is bounded, as shown in Andersen and Piterbarg (2007).

4.1 Preliminaries

We illustrate the idea of the ml-drMC method via the pure-diffusion case. Consider multiple sets of simulations of $\nu(t)$ with different timesteps sizes $h_\ell = \frac{T}{N_\ell}$, $N_\ell = 2^\ell$, $\ell = 0, \ldots, L$, and so the level ℓ has 2 times more timesteps than the level $(\ell - 1)$. For a given simulated BM path $W(t)$, we denote by $\hat{P}_\ell$, $\ell = 0, \ldots, L$, an approximation to the payoff P, defined in (3.9), using the discretization scheme (4.1) with timestep size h_ℓ. Note the key identity underlying the mlMC method

$$\mathbb{E}(\hat{P}_L) = \mathbb{E}(\hat{P}_0) + \sum_{\ell=1}^{L} \mathbb{E}[\hat{P}_\ell - \hat{P}_{\ell-1}]. \tag{4.2}$$

We denote by $\hat{Y}_0$ an estimator for $\mathbb{E}(\hat{P}_0)$, and by $\hat{Y}_\ell$, $\ell = 1, \ldots, L$, an estimator for $\mathbb{E}[\hat{P}_\ell - \hat{P}_{\ell-1}]$ using M_ℓ simulation paths. In the simplest scheme, the estimator $\hat{Y}_\ell$ is a mean of M_ℓ paths, i.e.

$$\hat{Y}_\ell = \frac{1}{M_\ell} \sum_{m=1}^{M_\ell} \left(\hat{P}_\ell^{(m)} - \hat{P}_{\ell-1}^{(m)} \right). \tag{4.3}$$

A key point in the mlMC approach is that the quantity $\hat{P}_\ell^{(m)} - \hat{P}_{\ell-1}^{(m)}$ comes from two discrete approximations with different timestep sizes, but are based on the same BM path. We denote by $\hat{Y}$ the combined estimator, defined as $\hat{Y} = \sum_{\ell=0}^{L} \hat{Y}_\ell$. The idea of mlMC is to *independently* estimate each $\hat{Y}_\ell$, $\ell = 1, \ldots, L$, in such a way that, for a given computational cost, the variance of the combined estimator, namely $\mathbb{V}(\hat{Y})$, is minimized. As showed in Giles (2008), this can be achieved by choosing M_ℓ proportional to $\sqrt{V_\ell h_\ell}$, where $\mathbb{V}_\ell \equiv \mathbb{V}\left[\hat{P}_\ell - \hat{P}_{\ell-1}\right]$. Thus, the convergence of the sample variance $\mathbb{V}_\ell$ as $\ell \to \infty$ is very important to the efficiency of the methods, since it determines an optimal choice of M_ℓ, i.e. the number of sample paths used the ℓ-th level.

In the remainder of this section, we show that it is possible to construct an ml-drMC estimator that can achieve $\mathbb{V}_\ell = \mathcal{O}(h_\ell^2)$. Following from Giles (2008)[Theorem 3.1], the computational complexity required by the ml-drMC method to obtain a RMSE of ϵ is $\mathcal{O}(\epsilon^{-2})$. We primarily focus on the case that $\ln(y)$ follows a normal distribution (Merton, 1976), for simplicity reasons. The proof techniques for the case of normal distribution can be extended to the case of double-exponential distribution (Kou, 2002).

For simplicity, in our analysis as, well as in the numerical experiments, we consider the case where κ_{d_i}, and σ_{d_i}, $i = 1, \ldots, m$, and κ_{f_i}, σ_{f_i}, $i = 1, \ldots, l$, are constants. In this case, (3.8) reduces to the following form

$$\beta_{(\cdot)}(t) = \sigma_{(\cdot)} \int_t^T \mathrm{e}^{\kappa_{(\cdot)}(t-t')}\, \mathrm{d}t' = \frac{\sigma_{(\cdot)}}{\kappa_{(\cdot)}} \left(1 - \mathrm{e}^{-\kappa_{(\cdot)}(T-t)}\right), \tag{4.4}$$

for some positive constant $\kappa_{(\cdot)}$ and $\sigma_{(\cdot)}$.

For the rest of the paper, the super-scripts *"f"* and *"c"* are used to denote the dependence of the quantities on fine and coarse levels, respectively. This is not to be confused with the sub-script *"f"* used to indicate association with the *"f"* interest rate factor.

4.2 Approximation schemes for integrals

Define the following stochastic variables

$$\begin{aligned}
x_1 &= \int_0^T \nu(t)\mathrm{d}t, & x_2 &= \int_0^T \sqrt{\nu(t)}\, \mathrm{d}W(t), & & \\
x_{d_i,1} &= \int_0^T \beta_{d_i}(t)\sqrt{\nu(t)}\, \mathrm{d}t, & x_{f_i,1} &= \int_0^T \beta_{f_i}(t)\sqrt{\nu(t)}\, \mathrm{d}t, & i &= 1, \ldots, m, \\
x_{d_i,2} &= \int_0^T \beta_{d_i}(t)\, \mathrm{d}W(t), & x_{f_i,2} &= \int_0^T \beta_{f_i}(t)\, \mathrm{d}W(t), & i &= 1, \ldots, l\,.
\end{aligned} \tag{4.5}$$

We note that the option price and hedging parameters are functions of these random variables only.

In the analysis, the discrete paths of the variance $\nu(t)$ are simulated using the LBE scheme (4.1), with the ℓ-th level having twice as many number of timesteps as the $(\ell-1)$-th level. In the following discussion, we denote by $\hat{x}^f_{(\cdot),\ell}$ an approximation to $x_{(\cdot)}$ on a fine-path using $N_\ell = 2^\ell$ timesteps, and by $\hat{x}^c_{(\cdot),\ell-1}$ the *corresponding* coarse-path approximation to $x_{(\cdot)}$ using $N_{\ell-1} = 2^{\ell-1}$ timesteps. That is, $\hat{x}^f_{(\cdot),\ell}$ is and $\hat{x}^c_{(\cdot),\ell-1}$ are two discrete approximations to $x_{(\cdot)}$ with T/N_ℓ and $T/N_{\ell-1}$ timestep sizes, respectively, but are based on the same BM path.

Frequently in our analysis, we use the following inequality.

Proposition 4.2. *For random variables a_i, $i = 1, \ldots, n$, we have*

$$\mathbb{E}\left[\left(\sum_{i=1}^{n} a_i\right)^2\right] \leq n\left(\sum_{i=1}^{n} \mathbb{E}\left[(a_i)^2\right]\right) .$$

4.2.1 An approximation scheme for $x_1 = \int_0^T \nu(t)\mathrm{d}t$

Following Giles et al. (2013), given $N_\ell = 2^\ell$, we define the following piecewise linear interpolant (PLI)

$$\hat{\nu}_{\mathrm{PLI},\ell}(t) = \hat{\nu}_n + \frac{t-t_n}{h_\ell}\left(\hat{\nu}_{n+1} - \hat{\nu}_n\right), \quad t_n \leq t \leq t_{n+1}, \quad n = 0, \ldots, N_\ell - 1 . \tag{4.6}$$

Furthermore, by approximating the drift and diffusion coefficient of the $\mathrm{d}\nu$ as being constant within each timestep, we define the following Brownian motion interpolant (BMI)

$$\begin{aligned}\hat{\nu}_{\mathrm{BMI},\ell}(t) = \hat{\nu}_n + \frac{t-t_n}{h_\ell}\left(\hat{\nu}_{n+1} - \hat{\nu}_n\right) + \sigma_\nu\sqrt{\hat{\nu}_n}\left(W(t) - W_n - \frac{t-t_n}{h_\ell}\left(W_{n+1} - W_n\right)\right),\\ t_n \leq t \leq t_{n+1}, \quad n = 0, \ldots, N_\ell - 1 .\end{aligned} \tag{4.7}$$

Note that, $\hat{\nu}_{\mathrm{BMI},\ell}(t)$ deviates from $\hat{\nu}_{\mathrm{PLI},\ell}(t)$ if and only if $W(t)$ deviates from the BM piecewise linear interpolant $W_n + \frac{t-t_n}{h_\ell}\left(W_{n+1} - W_n\right)$.

We present two schemes for computing $\hat{x}^f_{1,\ell}$. In the first scheme, we integrate the Brownian motion interpolant $\hat{\nu}_{\mathrm{BMI},\ell}(t)$ from 0 to T. More specifically,

$$\hat{x}^f_{1,\ell} = \int_0^T \hat{\nu}_{\mathrm{BMI},\ell}(t)\,\mathrm{d}t = \sum_{n=0}^{N_{\ell-1}} \frac{h_\ell}{2}\left(\hat{v}^f_n + \hat{v}^f_{n+1}\right) + \sigma_\nu\sqrt{\hat{\nu}_n} I^f_{n,\ell}, \tag{4.8}$$

where $I^f_{n,\ell}$ are independent Normal$(0, h_\ell^3/12)$. The corresponding coarse-path approximation to x_1, i.e. $\hat{x}^c_{1,\ell-1}$, is defined similarly as (4.8), and it turns out that, for $n = 0, \ldots, \frac{N_\ell}{2} - 1$, we have

$$\begin{aligned}I^c_{n,\ell-1} &= \int_{t_n}^{t_{n+2}} \left(W(t) - W_n - \frac{t-t_n}{2h_\ell}\left(W_{n+2} - W_n\right)\right)\mathrm{d}t\\ &= I^f_{n,\ell} + I^f_{n+1,\ell} - \frac{h_\ell}{2}\left(W_{n+2} - 2W_{n+1} + W_n\right),\end{aligned}$$

which can be obtained using the BM information utilized for the fine path. An alternative approximation scheme is the same as the first one, but with the terms $I^f_{n,\ell}$ and $I^c_{n,\ell-1}$ omitted. This approximation can be viewed as being obtained by integrating the PLI $\hat{\nu}_{\mathrm{PLI},\ell}(t)$ from 0 to T. More specifically,

$$\hat{x}^f_{1,\ell} = \int_0^T \hat{\nu}_{\mathrm{PLI},\ell}(t)\,\mathrm{d}t = \sum_{n=0}^{N_{\ell-1}} \frac{h_\ell}{2}\left(\hat{v}^f_n + \hat{v}^f_{n+1}\right). \tag{4.9}$$

Lemma 4.1. *Both approximations* (4.8)-(4.9) *give* $\mathbb{E}\left[\left(\hat{x}^f_{1,\ell} - \hat{x}^c_{1,\ell-1}\right)^2\right] = \mathcal{O}(h_\ell^2)$.

Proof. See Appendix A. □

For the rest of the analysis and in the numerical experiments, we use the approximation (4.9).

4.2.2 An approximation scheme for $x_2 = \int_0^T \sqrt{\nu(t)}\,\mathrm{d}W(t)$

We note that, by first integrating (2.1d) from t_n to t_{n+1} for $\nu(t)$, and then rearranging, we obtain

$$\int_{t_n}^{t_{n+1}} \sqrt{\nu(t)}\,\mathrm{d}W(t) = \frac{\nu(t_{n+1}) - \nu(t_n) - \kappa_\nu \bar{\nu} h_\ell + \kappa_\nu \int_{t_n}^{t_{n+1}} \nu(t)\mathrm{d}t}{\sigma_\nu}. \tag{4.10}$$

Thus, (4.9) and (4.10) gives rise to the following scheme for $\hat{x}^f_{2,\ell}$:

$$\hat{x}^f_{2,\ell} = \frac{\hat{\nu}^f_{N_\ell} - \nu(0) - \kappa_\nu \bar{\nu} T + \kappa_\nu \sum_{n=0}^{N_\ell - 1} \frac{h_\ell}{2}\left(\hat{\nu}^f_n + \hat{\nu}^f_{n+1}\right)}{\sigma_\nu}. \tag{4.11}$$

The corresponding coarse-path approximation to x_2, namely $\hat{x}^c_{2,\ell-1}$, is defined similarly.

Lemma 4.2. *The approximation* (4.11) *gives* $\mathbb{E}\left[\left(\hat{x}^f_{2,\ell} - \hat{x}^c_{2,\ell-1}\right)^2\right] = \mathcal{O}(h_\ell|^2)$.

Proof. First, note that

$$\begin{aligned}\mathbb{E}\left[\left(\hat{\nu}^f_{N_\ell} - \hat{\nu}^c_{N_{\ell-1}}\right)^2\right] &= \mathbb{E}\left[\left(\hat{\nu}^f_{N_\ell} - \nu(T) + \nu(T) - \hat{\nu}^c_{N_{\ell-1}}\right)^2\right] \\ &\le 2\left(\mathbb{E}\left[\left(\hat{\nu}^f_{N_\ell} - \nu(T)\right)^2\right] + \mathbb{E}\left[\left(\nu(T) - \hat{\nu}^c_{N_{\ell-1}}\right)^2\right]\right) = \mathcal{O}(h_\ell^2).\end{aligned} \tag{4.12}$$

Here, the inequality follows from Proposition 4.2, and the $\mathcal{O}(h_\ell^2)$ bound follows from Proposition 4.1. The desired result follows from (4.11), (4.12) and Lemma 4.1. □

4.2.3 An approximation scheme for $x_{d_i,1} = \int_0^T \beta_{d_i}(t)\sqrt{\nu(t)}\,\mathrm{d}t$, $i = 1,\ldots,m$, and $x_{f_i,1} = \int_0^T \beta_{f_i}(t)\sqrt{\nu(t)}\,\mathrm{d}$, $i = 1,\ldots,l$

All of these integrals are of the form $y_1 = \int_0^T \beta(t)\sqrt{\nu(t)}\,\mathrm{d}t$, where $\beta(t)$ is define in (4.4). On the fine-path of the ℓ-th level, we approximate these integrals by

$$\hat{y}^f_{1,\ell} = \sum_{n=0}^{N_\ell - 1} \frac{h_\ell}{2}\left(\beta(t_n)\sqrt{\hat{\nu}^f_n} + \beta(t_{n+1})\sqrt{\hat{\nu}^f_{n+1}}\right), \tag{4.13}$$

Lemma 4.3. *The approximation* (4.13) *has* $\mathbb{E}\left[\left(\hat{y}^f_{1,\ell} - \hat{y}^c_{1,\ell-1}\right)^2\right] = \mathcal{O}(h_\ell^2)$.

Proof. See Appendix B. □

4.2.4 An approximation scheme for $x_{d_2,i} = \int_0^T \beta_{d_i}(t)\,\mathrm{d}W(t)$, $i = 1,\ldots,m$, and $x_{f_2,i} = \int_0^T \beta_{f_i}(t)\,\mathrm{d}W(t)$, $i = 1,\ldots,l$

All of these integrals are of the form $y_2 = \int_0^T \beta(t)\,\mathrm{d}W(t)$, where $\beta(t)$ is defined in (4.4). On the fine path of the ℓ-th level, we use the following approximation

$$\hat{y}^f_{2,\ell} = \sum_{n=0}^{N_\ell - 1} \beta(t_n)\left(W_{n+1} - W_n\right). \tag{4.14}$$

The scheme for $\hat{y}^c_{2,\ell-1}$ is defined similarly.

Lemma 4.4. *The approximation* (4.14) *has* $\mathbb{E}\left[\left(\hat{y}^f_{2,\ell} - \hat{y}^c_{2,\ell-1}\right)^2\right] = \mathcal{O}(h_\ell^2)$.

Proof. Note that

$$\mathbb{E}\left[\left(\hat{y}^f_{2,\ell} - \hat{y}^c_{2,\ell-1}\right)^2\right] = \mathbb{E}\left[\left(\sum_{n=0}^{\frac{N_\ell}{2}-1} \left(\beta(t_{2n+1}) - \beta(t_{2n})\right)\left(W_{2n+2} - W_{2n+1}\right)\right)^2\right]. \tag{4.15}$$

Since, $(\beta(t+h_\ell) - \beta(t))^2 = \mathcal{O}(h_\ell^2)$, for each $n = 0, \ldots, \frac{N_\ell}{2} - 1$, we have

$$\begin{aligned}\mathbb{E}\left[\left(\left(\beta(t_{2n+1}) - \beta(t_{2n})\right)\left(W_{2n+2} - W_{2n+1}\right)\right)^2\right] &= \left(\beta(t_{2n+1}) - \beta(t_{2n})\right)^2 \mathbb{E}\left[\left(W_{2n+2} - W_{2n+1}\right)^2\right] \\ &= \left(\beta(t_{2n+1}) - \beta(t_{2n})\right)^2 h_\ell = \mathcal{O}(h_\ell^3)\ .\end{aligned} \tag{4.16}$$

The result follows from using (4.16), and noting that the cross terms in (4.15) have expectation zero. □

5 Variance convergence results

5.1 Option price, pure-diffusion

We consider ml-drMC method applied to computing option price under a pure-diffusion model, i.e. when $\lambda = 0$. In this case, the payoff is P defined in (3.9).

5.1.1 Lipschitz payoff

Analyses of multi-level MC methods are typically built upon the Lipschitz property of the payoff function. In our case, however, the presence of the stochastic variables $x_{f_i,2}$, $i = 1, \ldots, l$, in the payoff gives rise to a non Lipschitz payoff. This is because (i) these stochastic variables are Gaussian, and hence unbounded, and (ii) they appear only in the F (see (3.7)). As a result, the payoff has $P \to \pm\infty$, as $x_{f_i,2} \to \pm\infty$, due to the term e^{G+F+H}. Inspection of the F in (3.7) shows that these stochastic variables disappear if the correlations between the BMs associated with factors of the *"f"* interest rate and the BM of the variance, i.e. between $W_{f_i}(t)$, $i = 1, \ldots, l$, and $W_\nu(t) \equiv W(t)$, are zero. We establish the convergence analysis of the ml-drMC method under the modelling assumption that these afore-mentioned correlations are zero.

Assumption 5.1. *The correlations between the BMs* $W_{f_i}(t)$, $i = 1, \ldots, l$, *and* $W_\nu(t) \equiv W(t)$ *are zero.*

Lemma 5.1. *Suppose Assumptions 4.1 and 5.1 hold and* $\lambda = 0$*. Then, the payoff function*

$$P = \mathcal{F}\left(x_1, x_2, x_{d_1,1}, \ldots, x_{d_m,1}, x_{f_1,1}, \ldots, x_{f_l,1}, x_{d_1,2}, \ldots, x_{d_m,2}\right)$$

defined in (3.9) *is a Lipschitz function of the values of variables* x_1, x_2, $x_{d_i,1}$, $i = 1, \ldots, m$, $x_{f_i,1}$, $i = 1, \ldots, l$, *and* $x_{d_i,2}$, $i = 1, \ldots, m$, *with the Lipschitz bound*

$$\begin{aligned}&\Big|\mathcal{F}\left(x_1^{(1)}, x_2^{(1)}, x_{d_1,1}^{(1)}, \ldots, x_{d_m,1}^{(1)}, x_{f_1,1}^{(1)}, \ldots, x_{f_l,1}^{(1)}, x_{d_1,2}^{(1)}, \ldots, x_{d_m,2}^{(1)}\right) \\ &\quad - \mathcal{F}\left(x_1^{(2)}, x_2^{(2)}, x_{d_1,1}^{(2)}, \ldots, x_{d_m,1}^{(2)}, x_{f_1,1}^{(2)}, \ldots, x_{f_l,1}^{(2)}, x_{d_2,2}^{(2)}, \ldots, x_{d_m,2}^{(2)}\right)\Big| \\ &\quad \le C\left(\sum_{i=1}^{2}\left|x_i^{(1)} - x_i^{(2)}\right| + \sum_{i=1}^{m}\left|x_{d_i,1}^{(1)} - x_{d_i,1}^{(2)}\right| + \sum_{i=1}^{l}\left|x_{f_i,1}^{(1)} - x_{f_i,1}^{(2)}\right| + \sum_{i=1}^{m}\left|x_{d_i,2}^{(1)} - x_{d_i,2}^{(2)}\right|\right)\end{aligned} \tag{5.1}$$

for some $C < \infty$.

Proof. See Appendix C. □

Given a fine-path of $\nu(t)$ simulated using timestep size $h_\ell = T/N_\ell$, where $N_\ell = 2^\ell$, the corresponding fine-path estimate of the payoff is defined by

$$\hat{P}^f_\ell \equiv \mathcal{F}\left(\hat{x}^f_{1,\ell}, \hat{x}^f_{2,\ell}, \hat{x}^f_{d_1,1,\ell}, \ldots, \hat{x}^f_{d_m,1,\ell}, \hat{x}^f_{f_1,1,\ell}, \ldots, \hat{x}^f_{f_l,1,\ell}, \hat{x}^f_{d_1,2,\ell}, \ldots, \hat{x}^f_{d_m,2,\ell}, \hat{x}^f_{f_1,2,\ell}, \ldots, \hat{x}^f_{f_l,2,\ell}\right),$$

where each $\hat{x}^f_{(\cdot),\ell}$ is defined as in the previous subsection. The corresponding coarse-path estimate of the payoff using timestep size $2h_\ell$, namely $\hat{P}^c_{\ell-1}$, is constructed similarly. We now state the main result of the convergence analysis for the pure-diffusion case.

Theorem 5.1. *Suppose Assumptions 4.1 and 5.1 hold and and $\lambda = 0$. Approximations* (4.9), (4.11), (4.13) *and* (4.14) *result in a ml-drMC estimator for the option price that has $\mathbb{V}_\ell = \mathcal{O}(h^2_\ell)$.*

Proof. We have

$$\begin{aligned}
\mathbb{V}\left[\hat{P}^f_\ell - \hat{P}^c_{\ell-1}\right] &\le \mathbb{E}\left[\left(\hat{P}^f_\ell - \hat{P}^c_{\ell-1}\right)^2\right] \\
&\le C^2\mathbb{E}\left(\sum_{i=1}^{2}\left|\hat{x}^f_{i,\ell} - \hat{x}^c_{i,\ell-1}\right| + \sum_{i=1}^{m}\left|\hat{x}^f_{d_i,1,\ell} - \hat{x}^c_{d_i,1,\ell-1}\right| + \sum_{i=1}^{l}\left|\hat{x}^f_{f_i,1,\ell} - \hat{x}^c_{f_i,1,\ell-1}\right| + \sum_{i=1}^{m}\left|\hat{x}^f_{d_i,2,\ell} - \hat{x}^c_{d_i,2,\ell-1}\right|\right)^2 \\
&\le bC^2\left(\sum_{i=1}^{2}\mathbb{E}\left[\left(\hat{x}^f_{i,\ell} - \hat{x}^c_{i,\ell-1}\right)^2\right] + \sum_{i=1}^{m}\mathbb{E}\left[\left(\hat{x}^f_{d_i,1,\ell} - \hat{x}^c_{d_i,1,\ell-1}\right)^2\right]\right. \\
&\qquad\left. + \sum_{i=1}^{l}\mathbb{E}\left[\left(\hat{x}^f_{f_i,1,\ell} - \hat{x}^c_{f_i,1,\ell-1}\right)^2\right] + \sum_{i=1}^{m}\mathbb{E}\left[\left(\hat{x}^f_{d_i,2,\ell} - \hat{x}^c_{d_i,2,\ell-1}\right)^2\right]\right),
\end{aligned}$$

for some bounded constant C, and b is the number of stochastic factors in the model. Here, the second inequality comes from the Lipschitz bound (5.1), and the third inequality comes from Proposition 4.2. Applying Lemmas 4.1, 4.2, 4.3, and 4.4 gives the desired result. □

Remark 5.1. *We note that when the Assumption 5.1 is not satisfied, the* extreme *path technique in Giles et al. (2009) may be used to show that $\mathbb{V}_\ell$ is probably still $\mathcal{O}(h^2_\ell)$. Specifically, this technique involves (i) partitioning the set of $\nu(t)$ paths into two subsets, namely the sets of extreme paths, i.e. paths along which $\hat{x}_{f_i,2}$ satisfies certain extreme conditions, and* non-extreme *paths, and (ii) showing that the contribution of the set of extreme paths to $\mathbb{E}\left[\left(\hat{P}^f_\ell - \hat{P}^c_{\ell-1}\right)^2\right]$ is negligible. We plan to investigate this issue in the near future. Nonetheless, as shown in numerical experiments, we observe that the presence of these stochastic variables does not have any impact on the expected optimal convergence rate of $\mathbb{V}_\ell$.*

5.2 Option price, normal jump

Recall that in this case, the option price can be expressed as

$$V(S(0), 0, \cdot) = \mathbb{E}\left[\sum_{n=0}^{\infty}\frac{(\lambda T)^n}{n!}P_n\right], \quad P_n = \exp\left(n\tilde{\mu} + \frac{n\tilde{\sigma}^2}{2}\right)S(0)\mathrm{e}^{(G+F+H)}\mathcal{N}(d_{1,n}) - K\mathrm{e}^{H}\mathcal{N}(d_{2,n}). \tag{5.2}$$

Here, the relevant quantities $d_{i,n}$, $i = 1, 2$, are defined in (3.12). Typically, in a numerical implementation, the (quickly converging) infinite series (5.2) is truncated to a finite number of terms, if a certain tolerance, denoted by `tol` > 0, has been met.

For a given simulated BM path $W(t)$, and a value of n, $n = 1, 2, \ldots$, we denote by $\hat{P}^f_{n,\ell}$ an approximation to the *conditional* payoff P_n, defined in (5.2), on a fine-path using $N_\ell = 2^\ell$ timesteps, and by $\hat{P}^f_\ell$ the corresponding fine-path approximation to the payoff. We have

$$\hat{P}^f_\ell = \sum_{n=0}^{N_{\mathtt{tol},\ell}} \frac{(\lambda T)^n}{n!} \hat{P}^f_{n,\ell} = \sum_{n=0}^{N_{\mathtt{tol},\ell}} \frac{(\lambda T)^n}{n!} \mathcal{F}_n \left(\hat{x}^f_{1,\ell}, \hat{x}^f_{2,\ell}, \hat{x}^f_{d_1,1,\ell}, \ldots, \hat{x}^f_{d_m,1,\ell}, \hat{x}^f_{f_1,1,\ell}, \ldots, \hat{x}^f_{f_l,1,\ell}, \ldots, \hat{x}^f_{f_l,2} \right). \tag{5.3}$$

In (5.3), $\mathcal{F}_n(\cdot)$ is defined in (5.2) as a function of stochastic variables $x_{(\cdot)}$. We note that in (5.3)

$$N_{\mathtt{tol},\ell} = \max \left(N^f_{\mathtt{tol},\ell}, N^c_{\mathtt{tol},\ell-1,} \right), \tag{5.4}$$

where $N^f_{\mathtt{tol},\ell}$ and $N^c_{\mathtt{tol},\ell-1,}$ are the finite number of terms required to achieve the tolerance $\mathtt{tol}$ on corresponding the fine- and coarse-path, respectively.

Theorem 5.2. *Suppose that Assumptions 4.1 and 5.1 hold, and that* $\ln(y) \sim$ *Normal*$(\tilde{\mu}, \tilde{\sigma}^2)$. *Approximations* (4.9), (4.11), (4.13) *and* (4.14) *result in an ml-drMC estimator for the option price that has* $\mathbb{V}_\ell = \mathcal{O}(h_\ell^2)$.

Proof. The result follows from Theorem 5.1 and the fact that $N_{\mathtt{tol}}$ is finite. □

5.3 Hedging parameters

We consider the Delta and Gamma of the option. We start with the Delta and Gamma for the pure-diffusion case, which can be obtained by setting $n = 0$ in (3.13). It is straightforward to show that the payoffs in these cases are also satisfied a Lipschitz bound. The fine- and coarse-path payoffs for the Delta and Gamma can be constructed the same way as the option price. Following the steps used previously, we can show that the pure-diffusion case, the ml-drMC estimator for the option's Delta and Gamma has $\mathbb{V}_\ell = \mathcal{O}(h_\ell^2)$. For the jump case, the convergence results of the ml-drMC estimator for option's Delta and Gamma can be obtained in the same fashion as previously for the option price.

6 Numerical results

In the experiments, we consider the following two models: (i) a 3-factor Heston-Hull-White (HHW) jump-diffusion model for stock options, and (ii) a 6-factor jump-diffusion model for FX options. The models for these two cases respectively are

$$\begin{aligned}
\frac{\mathrm{d}S(t)}{S(t^-)} &= (r_d(t) - \lambda\delta)\,\mathrm{d}t + \sqrt{\nu(t)}\,\mathrm{d}W_s(t) + \mathrm{d}J(t), \quad J(t) = \sum_{j=1}^{\pi(t)} (y_j - 1), \\
r_d(t) &= r_d(0)\,\mathrm{e}^{-\kappa_d t} + \kappa_d \int_0^t \mathrm{e}^{-\kappa_d (t-t')}\,\theta_d(t')\,\mathrm{d}t' + X(t), \\
&\quad \text{with } \mathrm{d}X(t) = -\kappa_d\,X(t)\,\mathrm{d}t + \sigma_d\,\mathrm{d}W_d(t), \quad X(0) = 0, \\
\mathrm{d}\nu(t) &= \kappa_\nu\,(\bar{\nu} - \nu(t))\,\mathrm{d}t + \sigma_\nu \sqrt{\nu(t)}\,\mathrm{d}W_\nu(t),
\end{aligned} \tag{6.1}$$

and

$$\begin{aligned}
&\frac{\mathrm{d}S(t)}{S(t^-)} = (r_d(t) - r_f(t) - \lambda\delta)\ \mathrm{d}t + \sqrt{\nu(t)}\,\mathrm{d}W_s(t) + \mathrm{d}J(t), \quad J(t) = \sum_{j=1}^{\pi(t)} (y_j - 1),\\
&r_d(t) = X_1(t) + X_2(t) + \gamma_d(t),\\
&\qquad \text{with } \mathrm{d}X_i(t) = -\kappa_{d_i} X_i(t)\,\mathrm{d}t + \sigma_{d_i}\,\mathrm{d}W_{d_i}(t), \quad X_i(0) = 0, \quad i = 1, 2,\\
&r_f(t) = Y_1(t) + Y_2(t) + \gamma_f(t),\\
&\qquad \text{with } \mathrm{d}Y_i(t) = -\kappa_{f_i} Y_i(t)\,\mathrm{d}t + \sigma_{f_i}\,\mathrm{d}W_{f_i}(t) - \rho_{s,f_i}\sigma_{f_i}\sqrt{\nu(t)}\,\mathrm{d}t, \quad Y_i(0) = 0, \quad i = 1, 2,\\
&\mathrm{d}\nu(t) = \kappa_\nu\,(\bar{\nu} - \nu(t))\ \mathrm{d}t + \sigma_\nu\sqrt{\nu(t)}\,\mathrm{d}W_\nu(t).
\end{aligned} \tag{6.2}$$

For the jump components, we consider two distributions, namely (i) $\ln(y_j) \sim \text{Normal}(\tilde{\mu}, \tilde{\sigma}^2)$, and (ii) $\ln(y_j) \sim \text{double-exponential}(p, \eta_1, \eta_2)$, $j = 1, 2, \ldots$, where $\ln(y_j)$ are i.i.d. Note that, as stated earlier, in these models, all coefficients $\kappa_{(\cdot)}$, $\sigma_{(\cdot)}$, κ_ν, σ_ν and $\bar{\nu}$ are also constant. Furthermore, for simplicity, for the interest rate model, we assume θ_i, $i = \{d, f\}$, defined in (2.2), are constant. As a result, all the deterministic integrals in G, F and H can be computed analytically. The quantities G, F and H defined in (3.7) can further be reduced for the above two cases. For brevity, we omit these reduced formulas, which can be found in Dang et al. (2017).

Since we compare the efficiency of various MC methods, it is important to determine the computational complexity of each MC method. Following Giles (2008), for a pure mlMC method, we define the computational complexity of a MC method as the *total* number of random numbers generated for *all* factors in the model. More specifically, due to presence of jumps, the computational cost is approximated by $\sum_{\ell=1}^{L}\sum_{m=1}^{M_\ell}\left(J_{[0,T]}^{(m)} + N_\ell\right)$, where $J_{[0,T]}^{(m)}$ is the number of jumps along the m-th path from time 0 to time T.

For ml-drMC methods, however, it is not appropriate to use just the number of random numbers generated for the variance factor, as this does not reflect the fact that each ml-drMC sample requires additional computations. Inspection of the analytical solution (5.2) indicates that, for each level ℓ, the extra costs are primarily for (i) approximations of integrals and computation of the terms F, H, and G (see (3.7)), which is done only once per path, and (ii) evaluations of a total of $N_{\mathtt{tol},\ell} + 1$ terms in the sum (5.3). (For pure-diffusion case, $N_{\mathtt{tol},\ell} = 0$.) Based on operation counts and timing results of the drMC and ordinary MC methods (see Dang et al. (2015a, 2017)), our estimate is that, on average, given the same number of timestepping, for the 3-factor HHW model, the cost *per path* of the drMC is approximately 1.5 times that of the ordinary MC, while for the 6-factor model (6.2), the difference is about 2 times. These factors are taken into account in the complexity comparisons between ml-drMC and mlMC methods in this section.

The computational cost of a non-multi-level method is computed as $\sum_{\ell=0}^{L} M_\ell^* N_\ell$, where $M_\ell^* = 2\epsilon^{-2}\mathbb{V}[\hat{P}_\ell]$, so that the variance bound is also $\epsilon^2/2$ as with its multi-level counterpart (Giles, 2008). We also note that in all of the experiments reported below, Assumption 5.1 is not satisfied. Nonetheless, as noted in Remark 5.1, the ml-drMC method with LBE scheme performs well, requiring only an overall complexity $\mathcal{O}(\epsilon^{-2})$ to achieve a RMSE of ϵ.

6.1 Pure-diffusion: a 6-factor model

First, we illustrate the the efficiency of the ml-drMC method when applied to a pure-diffusion model. For this experiment, we consider a European option under the 6-factor model (6.2) with the jump intensity $\lambda = 0$. For the numerical experiments, we use the following parameters (Dang et al., 2015b): $r_d(0) = 0.02$, $\kappa_{d_1} = 0.03$, $\kappa_{d_2} = 0.03$, $\sigma_{d_1} = 0.03$, $\sigma_{d_2} = 0.03$, $\theta_d = 0.02$, and

$r_f(0) = 0.05$, $\kappa_{f_1} = 0.03$, $\kappa_{f_2} = 0.03$, $\sigma_{f_1} = 0.012$, $\sigma_{f_2} = 0.012$, and $\theta_f = 0.05$. The correlations are from Dang et al. (2015a): $\rho_{S,d_1} = 0.08$, $\rho_{S,d_2} = 0.08$, $\rho_{S,f_1} = 0.08$, $\rho_{S,f_2} = 0.08$, $\rho_{S,\nu} = -0.02$, $\rho_{d_1,d_2} = 0.12$, $\rho_{d_1,f_1} = 0.12$, $\rho_{d_1,f_2} = 0.12$, $\rho_{d_1,\nu} = 0.15$, $\rho_{d_2,f_1} = 0.12$, $\rho_{d_2,f_2} = 0.12$, $\rho_{d_2,\nu} = 0.15$, $\rho_{f_1,f_2} = -0.70$, $\rho_{f_1,\nu} = 0.15$, $\rho_{f_2,\nu} = 0.15$. For the variance factor, we use the parameters $\kappa_\nu = 0.5$, $\bar{\nu} = 0.9$, $\sigma_\nu = 0.05$, $\nu(0) = 0.9$, which are taken from Giles and Szpruch (2014). We also use $S(0) = 10$, $K = 10$, and $T = 20$ (years). The parameters above are highly challenging for practical applications, due to long maturity.

For comparison purposes, we also implement an antithetic mlMC method combined with a Milstein discretization scheme, as developed in Giles and Szpruch (2014). We refer to this method as anti-mlMC. To the best of our knowledge, anti-mlMC is currently the most efficient mlMC method for multi-dimensional pure-diffusion models, since it requires only an overall complexity $\mathcal{O}(\epsilon^{-2})$ to achieve a RMSE of ϵ without simulating Lévy areas. For this method, due to the non-linearity of the diffusion coefficient in the price process $S(t)$, we work with $\log(S(t))$ instead, as suggested by Giles and Szpruch (2014). Given a timestep size $h = T/N$, the Milstein scheme for the 6-factor model under consideration with the Lévy area terms set to zero is given by

$$
\begin{aligned}
\log(\hat{S}_{n+1}) &= \log(\hat{S}_n) + \big(\hat{r}_{d,n} - \hat{r}_{f,n} - 0.5\hat{\nu}_n\big)h + \sqrt{\hat{\nu}_n^+}\,\Delta W_{s,n} + 0.5\hat{\nu}_n\big((\Delta W_{s,n})^2 - h\big) \\
&\qquad + 0.25\sigma_\nu\big(\Delta W_{s,n}\Delta W_{\nu,n} - \rho_{s,\nu}h\big), \\
\hat{r}_{d,n+1} &= \sum_{i=1}^{2}\hat{X}_{i,n+1} + \gamma_{d,n+1}, \quad \hat{X}_{i,n+1} = \hat{X}_{i,n} - \kappa_{d_i}\hat{X}_{i,n}h + \sigma_{d_i}\Delta W_{d_i,n}, \quad \hat{X}_{i,0} = 0, \quad i = 1,2, \\
\hat{r}_{f,n+1} &= \sum_{i=1}^{2}\hat{Y}_{i,n+1} + \gamma_{f,n+1}, \quad \hat{Y}_{i,n+1} = \hat{Y}_{i,n} - \big(\kappa_{f_i}\hat{Y}_{i,n} + \rho_{S,f_i}\sigma_{f_i}\sqrt{\hat{\nu}_n^+}\big)h + \sigma_{f_i}\Delta W_{f_i,n}, \\
&\qquad Y_{i,0} = 0, \quad i = 1,2, \\
\hat{\nu}_{n+1} &= \frac{\hat{\nu}_n + \kappa_\nu\,\bar{\nu}h + \sigma_\nu\sqrt{\hat{\nu}_n^+}\,\Delta W_{\nu,n} + 0.25\,\sigma_\nu^2\big((\Delta W_{\nu,n})^2 - h\big)}{1 + h\,\kappa_\nu}\,.
\end{aligned}
\tag{6.3}
$$

Here, $\Delta W_{(\cdot),n} = W_{(\cdot),n+1} - W_{(\cdot),n}$, and $\gamma_{i,n} = (r_i(0) - \theta_i)e^{(-\kappa_{i_1}nh)} + \theta_i$, $i \in \{d, f\}$. Details of the antithetic mlMC technique for multi-dimensional pure-diffusion problems discretized by the Milstein scheme, such as (6.3), are discussed in Giles and Szpruch (2014), and hence omitted here. We also note that, although the coefficients of the variance process are not Lipschitz continuous, and hence the assumptions in Giles and Szpruch (2014) are not satisfied, the numerical tests show that the anti-mlMC performs well, and is able to achieve $\mathbb{V}_\ell = \mathcal{O}(h_\ell^2)$. Similar convergence results are reported in Giles and Szpruch (2014) for the Heston model.

For the 6-factor pure-diffusion model (6.2), we compare three MC methods, namely ml-drMC, drMC, anti-mlMC. Here, drMC with the Lamperti-Backward-Euler (LBE) scheme is the non-multi-level counterpart of ml-drMC. The non-multi-level counterpart of the anti-mlMC is essentially the ordinary MC, and hence is skipped for brevity. The plots in the experiments are produced using Matlab code adapted from the code freely available from Giles (2008).

6.1.1 Accuracy

In Table D.1, to illustrate the accuracy of the ml-drMC method, we present the option prices obtained by the three methods, and the corresponding standard derivation (in brackets) for the case $\epsilon = 10^{-3}$. We observed that the option prices obtained by all methods agree well. Also, the standard deviation for each method is $\leq \frac{\epsilon}{\sqrt{2}} \approx 0.000707$. This indicates that the variance bound $\epsilon^2/2$ is satisfied by all methods, as expected by analysis of mlMC methods.

In the above test, the ml-drMC and anti-mlMC method respectively requires $L = 4$ and $L = 14$ to achieve the variance bound $\epsilon^2/2$. The drMC method with the LBE scheme for the variance factor requires $16 = 2^4$ timesteps and about 46×10^6 samples to achieve the same variance bound. For ordinary MC method, although the results are not presented here, we note that the timesteps and samples required to achieve the same variance bound respectively are $16384 = 2^{14}$ and 845×10^6.

6.1.2 Convergence properties and efficiency

We present numerical results to show the convergence properties and compare the efficiency of the three methods, namely ml-drMC, drMC, anti-mlMC, in computing the option price. In Figure D.1 (a), we investigate the convergence behavior of $\mathbb{V}_\ell = \mathbb{V}[P_\ell - P_{\ell-1}]$ as a function of the level of approximation when $\epsilon = 10^{-3}$. These values were estimated using 10^6 samples, so the sampling error is negligible.

We make following observations. The variance of the (non-multi-level) drMC varies very little with level ℓ. Both ml-drMC and anti-mlMC methods result in lines having slope -2, which indicates that $\mathbb{V}_\ell = \mathcal{O}(h_\ell^2)$, as expected from the complexity analysis. Moreover, the $\mathbb{V}_\ell$ of the ml-drMC method is about 50 times smaller than that of the anti-mlMC method, which is expected, due to the a significant variance reduction offered by the drMC approach. We also note that the multi-level-based methods are substantially more accurate than their non-multi-level-based counterparts. In particular, on level $\ell = 2$, which has just 4 timesteps, $\mathbb{V}_\ell$ of ml-drMC is already more than 1000 times smaller than that of drMC. (Compare $\mathbb{V}_\ell = \mathbb{V}[P_\ell - P_{\ell-1}]$ of ml-drMC and $\mathbb{V}[P_\ell]$ of drMC at level $\ell = 2$ on Figure D.1 (a)).

In Figure D.1 (b), the mean value for the multi-level correction is shown. Both multi-level based methods' estimators result in approximately a first-order convergence for $\mathbb{E}[P_\ell - P_{\ell-1}]$, as indicated by the slope -1.

Next, we investigate the computational complexity of the three methods. Figure D.1 (c) show the dependence of the computational complexity `Cost`, defined as the total of random numbers generated, as a function of the desired accuracy ϵ. Here, we plot ϵ^2`Cost` versus ϵ. As observed from Figure D.1 (c), for the drMC method, the quantity ϵ^2`Cost` exhibits the well-known "staircase" effect of non-multi-level MC methods (Giles, 2008). For both anti-mlMC and ml-drMC, the quantity ϵ^2`Cost` appears to be *independent* of ϵ. This result indicates that the first-order strong convergence of the Milstein and LBE discretization techniques results in a computational complexity `Cost` $= \mathcal{O}(\epsilon^{-2})$. This result is expected from the complexity analysis of multi-level methods in Giles (2008)[Theorem 3.1].

Furthermore, we also observe that the ml-drMC is significantly more efficient than the anti-mlMC method, about 40 times more efficient than the anti-mlMC method for this example. These results from Figure D.1 indicate that the ml-drMC estimator can achieve the same second-order rate of convergence for $\mathbb{V}_\ell$ as that of the anti-mlMC method of Giles and Szpruch (2014), but is significantly more efficient.

6.2 Jump-diffusion: 3-factor HHW with normal jumps

In the remaining experiments, we consider the popular 3-factor HHW model (6.1) with $\ln(y_j)$ following the normal (Merton, 1976) and the double-exponential (Kou, 2002) distributions. For validation purposes, we extend the anti-mlMC method of Giles and Szpruch (2014) to handle jumps. Specifically, since the option is not path-dependent, the overall jump effects on the underlying asset can be evaluated separately at time T, and be taken into account at that time. The main focus of this section is to demonstrate the convergence results of LBE scheme, and its benefit over the Euler-Maruyama scheme. The Euler-Maruyama scheme for (2.1d) is given by $\hat{\nu}_{n+1} = \hat{\nu}_n + \kappa_\nu (\bar{\nu} + \hat{\nu}_n) h + \sigma_\nu \sqrt{\hat{\nu}_n^+} \Delta W_n$.

6.2.1 Accuracy

In Table D.2, to illustrate the accuracy of the ml-drMC methods, we present the option prices obtained by ml-drMC methods with the Lamperti-Backward-Euler and the Euler-Maruyama schemes, as well as by the anti-mlMC, and the drMC method with the Milstein scheme of Dang et al. (2017), as well as the corresponding standard derivation (in brackets) for the case of $\epsilon = 10^{-3}$. We observed that the option prices obtained by all methods agree well. Also, as in the pure-diffusion case, the standard deviation for each method is $\leq \frac{\epsilon}{\sqrt{2}} \approx 0.000707$. This indicates that the variance bound $\epsilon^2/2$ is satisfied, as expected by analysis of mlMC methods.

In the above test, the ml-drMC method with Lamperti-Backward-Euler and Euler-Maruyama schemes respectively requires $L = 7$ and $L = 9$ to achieve the variance bound $\epsilon^2/2$, whereas the anti-mlMC method requires $L = 20$. The drMC method with Milstein scheme for the variance factor requires $128 = 2^7$ timesteps and about 8×10^6 samples to achieve the same variance bound.

6.2.2 Convergence properties and efficiency - price

We price a European call with initial spot price $S(0) = 10$, strike price $K = 10$, and maturity of $T = 1$ (years). We use the following parameters taken from Dang et al. (2017): $r_d(0) = 0.05$, $\theta_d = 0.05$, $\kappa_d = 1.5$, $\sigma_d = 0.1$, $\nu(0) = 0.04$, $\bar{\nu} = 0.0225$, $\kappa_\nu = 2.5$, $\sigma_\nu = 0.2$. The correlations are $\rho_{s,d} = 0.4$, $\rho_{s,\nu} = 0.1$, $\rho_{d,\nu} = 0.35$. The parameters for the normal jump amplitude w are $\lambda = 1$, $\tilde{\mu} = -0.08$, $\tilde{\sigma} = 0.3$.

Figure D.2 present our results for this test case obtained by various methods. In Figure D.2 (a), we investigate the convergence behavior of $\mathbb{V}_\ell$ as a function of the level of approximation when $\epsilon = 10^{-3}$. As in the pure-diffusion case, these $\mathbb{V}_\ell$ values were estimated using 10^6 samples, so the sampling error is negligible.

We observe that both drMC estimators, i.e. non-multi-level, result in variances that vary very little with level. The ml-drMC estimator built upon the Euler-Maruyama scheme results in approximately first-order of convergence for $\mathbb{V}_\ell$ (slope ≈ -1). When the LBE is employed, the resulting ml-drMC estimator achieves second-order of convergence for $\mathbb{V}_\ell$ (slope ≈ -2), same as the anti-mlMC method, as expected.

Figure D.2 (b) shows the mean value and correction at each level. As expected, all methods' estimators result in approximately a first-order convergence for $\mathbb{E}[P_\ell - P_{\ell-1}]$, as indicated by the slope -1. We note that the strong and weak convergence of the Euler-Maruyama scheme observed in Figures D.2 (a) and (b) are respectively slightly more and less than the half-order strong and first-order weak convergence of the Euler-Maruyama scheme reported in Giles (2008) in the context of European options under Heston model.

Figure D.2 (c) show the dependence of the computational complexity `Cost` as a function of the desired accuracy ϵ. As in the 6-factor pure-diffusion case, we observe that while the quantity ϵ^2`Cost` is weakly dependent on ϵ for the Euler-Maruyama scheme, it is *independent* of ϵ for the LBE scheme and for the anti-mlMC method. These results again highlight the advantage of the first-order strong convergence of the LBE technique. To achieve a RMSE of ϵ, the computational complexity required by the ml-drMC built upon the LBE technique is only $\mathcal{O}(\epsilon^{-2})$, which is expected from the complexity analysis of multi-level methods in Giles (2008)[Theorem 3.1]. Also from Figure D.2 (c), we observe that using the LBE scheme results in much lower computational complexity for the ml-drMC than using the Euler-Maruyama scheme, about 7-8 times smaller. Furthermore, the ml-drMC methods are significantly more efficient than the anti-mlMC, about 50 times.

6.2.3 Hedging parameters

We now illustrate that the ml-drMC can also be readily applied to computing hedging parameters. We focus on the Delta and Gamma of the option obtained by the ml-drMC method. Figure D.3 present plots showing the convergence order for $\mathbb{V}[P_\ell - P_{\ell-1}]$ and for $\mathbb{E}[P_\ell - P_{\ell-1}]$. We observe that these plots have the same structure to the results presented in Figure D.2 for the option price. In particular, $\mathbb{V}_\ell$ obtained by the LBE scheme is $\mathcal{O}(h_\ell^2)$, whereas the variance obtained by the Euler-Maruyama technique is $\mathcal{O}(h_\ell)$. The computational complexity of the ml-drMC methods in this case have the same behaviour as in Figure D.3 (c), and hence omitted.

6.3 Jump-diffusion: 3-factor HHW with double-exponential jumps

Next, we present the convergence results for the case of double-exponential distribution. In this example, the parameters for the w are taken from Kou (2002): $\lambda = 1$, $p = 0.4$, $\eta_1 = 10$, $\eta_2 = 5$. Figure D.4 presents plots showing approximate orders of convergence of $\mathbb{V}[P_\ell - P_{\ell-1}]$ and $\mathbb{E}[P_\ell - P_{\ell-1}]$ for ml-drMC methods with the LBE and Euler-Maruyama schemes applied to computing option's price, Delta and Gamma. Again, we observe that these plots have the same structure to those presented earlier for the normal jump case.

We conclude this section by emphasize the ml-drMC method can naturally compute very efficiently the hedging parameters under jump-diffusion models, especially high-order ones, such as Gamma. This is a significant advantage over existing mlMC methods, which typically encounter difficulties in this case, due to lack of smoothness in the payoff Burgos and Giles (2012). We also note that, although we focus on ml-drMC built-upon the LBE scheme for the variance factor, we can also use the Milstein scheme, which also have the same strong and weak convergence orders, as well as the positivity preserving property, as the LBE scheme (Neuenkirch and Szpruch, 2014). Numerical results, which are not presented herein, for brevity, confirm that the two schemes have similar convergence and efficiency advantages over the Euler-Maruyama scheme in the context of drMC.

7 Summary and conclusions

In this paper, we develop a highly efficient multi-level and dimension reduction MC method, referred to as ml-drMC, for pricing plain-vanilla European options under a very general b-dimensional jump-diffusion model, where b is arbitrary. The model includes stochastic variance and multi-factor Gaussian interest short rate(s), and is highly suitable for options having a wide range of maturities in various asset classes, such as equity and foreign exchange. To the best of our knowledge, the proposed ml-drMC method is the first multi-level based MC method reported in the literature that can effectively handle both multi-dimensionality and jumps in the underlying asset in computing the option price and hedging parameters.

The proposed ml-drMC method is based on two steps. First, by applying the drMC method of Dang et al. (2017), we can reduce the number of dimensions of the pricing problem from b to only 1, namely the variance factor. In the second step, we apply the multi-level technique with the Lamperti-Backward-Euler scheme of Neuenkirch and Szpruch (2014) on the variance factor, and this step is essentially an application of the multi-level technique on a one-dimensional problem. We show that the proposed ml-drMC method requires only an overall complexity $\mathcal{O}(\epsilon^{-2})$ to achieve a RMSE of ϵ. These complexity results hold for both price and hedging parameters, such as Delta and Gamma. Moreover, due to a (possible) significant variance reduction offered by the drMC method, it is expected that the ml-drMC method is significantly more efficient than the antithetic mlMC based approach of Giles and Szpruch (2014) when applied to pricing plain-vanilla European options under jump-diffusion models.

Major research directions of the ml-drMC approach go in parallel with the developments of the drMC approach. Current research shows that drMC approach can be extended to effectively deal with exotic features, such as early exercise or barrier, as well as multi-asset options with stochastic volatility and interest rates. Preliminary results indicate that the ml-drMC approach will also work very effectively for options with early exercise features. It is expected that the theoretical analysis developed in this paper will serve as a building block for future work on ml-drMC. Finally, we note that a Shannon wavelet based approach is proposed in Dang and Ortiz-Gracia (2017) as an alternative to the multi-level approach in effectively handling the outer expectation.

Appendix

A Proof of Lemma 4.1

A.1 Preliminaries

First, we present the following bound for $|\hat{\nu}_{\text{BMI},h}(t) - \hat{\nu}_{\text{PLI},h}(t)|$.

Lemma A.1. *Consider $\hat{\nu}_{PLI,h}(t)$ and $\hat{\nu}_{BMI,h}(t)$, respectively defined in* (4.6) *and* (4.7)*, with stepsize $h = T/N$. Then*

$$\mathbb{E}\left[\left(\int_0^T \hat{\nu}_{BMI,h}(t) - \hat{\nu}_{PLI,h}(t)\mathrm{d}t\right)^2\right] = \mathcal{O}\left(h^3\right). \tag{A.1}$$

Proof. Let

$$x_n = \int_{t_n}^{t_{n+1}} y(t)\mathrm{d}t, \quad t_{n+1} - t_n = h = T/N,$$

where

$$y(t) = W(t) - W_n - \frac{t - t_n}{h}\left(W_{n+1} - W_n\right).$$

For simplicity, let $b_n = \sigma_\nu \sqrt{\hat{\nu}_n}$. We have that

$$\mathbb{E}\left[\left(\int_0^T \hat{\nu}_{\text{BMI},h}(t) - \hat{\nu}_{\text{PLI},h}(t)\mathrm{d}t\right)^2\right] = \mathbb{E}\left[\left(\sum_{n=0}^{N-1} b_n x_n\right)^2\right] = \mathbb{E}\left[\sum_{n=0}^{N-1} b_n^2 x_n^2\right] + 2\mathbb{E}\left[\sum_{n=0,m>n}^{N-1} b_n b_m x_n x_m\right]$$
$$= \mathbb{E}\left[\sum_{n=0}^{N-1} b_n^2 x_n^2\right] + 2\sum_{n=0,m>n}^{N-1} \mathbb{E}[x_n]\mathbb{E}\left[b_n b_m x_m\right] = \mathbb{E}\left[\sum_{n=0}^{N-1} b_n^2 x_n^2\right],$$

where the third equality is due to the independence between x_n and x_m, for $m > n$, and the fourth equality is due to the fact that $\mathbb{E}[x_n] = 0$. Next, we consider $\mathbb{E}\left[\sum_{n=0}^{N-1} b_n^2 x_n^2\right]$. By noting that all x_n, $n = 0, \ldots, N-1$, are i.i.d., it follows that

$$\mathbb{E}\left[\left(\int_0^T \hat{\nu}_{\text{BMI},h}(xt) - \hat{\nu}_{\text{PLI},h}(t),\ \mathrm{d}t\right)^2\right] = \mathbb{E}\left[\sum_{n=0}^{N-1} b_n^2 x_n^2\right] = \mathbb{E}\left[x_0^2\right]\mathbb{E}\left[\sum_{n=0}^{N-1} b_n^2\right].$$

We note that the quantity $\mathbb{E}\left[\sum_{n=0}^{N-1} b_n^2\right]$ is bounded, due to the boundedness of $\hat{\nu}_n$, $n = 0, \ldots, N-1$ (see Neuenkirch and Szpruch (2014)[Lemma 2.5]).

Next, let $\chi_1(h) = \int_0^h W(t)\,\mathrm{d}t$ and $\chi_2(h) = W_1 \int_0^h \frac{t}{h}\,\mathrm{d}t$. Note that $\chi_2(h) \sim \text{Normal}(0, h^3/4)$, and hence $\mathbb{E}[(\chi_2(h))^2] = h^3/4$. We have

$$\mathbb{E}\left[x_0^2\right] = \mathbb{E}\left[(\chi_1(h) - \chi_2(h))^2\right] = \mathbb{E}\left[(\chi_1(h))^2 - 2\chi_1(h)\chi_2(h) + (\chi_2(h))^2\right] = \mathbb{E}\left[(\chi_1(h))^2\right] + \mathbb{E}\left[(\chi_2(h))^2\right],$$

where the third equality comes from linearity of expectation, and the facts that $\chi_1(h)$ and $\chi_2(h)$ are independent, and that $\mathbb{E}[\chi_2(h)] = 0$. To compute $\mathbb{E}\left[(\chi_1(h))^2\right]$, note that

$$\begin{aligned}\mathbb{E}\left[(\chi_1(h))^2\right] &= \mathbb{E}\left[\int_0^h W(s)\,\mathrm{d}s \int_0^h W(t)\,\mathrm{d}t\right] = \mathbb{E}\left[\int_0^h \int_0^h W(s)W(t)\,\mathrm{d}s\,\mathrm{d}t\right] \\ &= \int_0^h \int_0^h \mathbb{E}\left[W(s)W(t)\right]\,\mathrm{d}s\,\mathrm{d}t = \int_0^h \int_0^h \mathbb{E}[\min(s,t)]\,\mathrm{d}s\,\mathrm{d}t = \frac{h^3}{3}. \end{aligned} \tag{A.2}$$

Here, in the third equality, Fubini's theorem is applied. The result of (A.2), together with $\mathbb{E}[(\chi_2(h))^2] = h^3/4$, concludes the proof. □

A.2 Proof of Lemma 4.1

We are now in a position to prove Lemma 4.1. First, we show the desired result for scheme (4.8). We have

$$\begin{aligned}\mathbb{E}\left[\left(\hat{x}^f_{1,\ell} - \hat{x}^c_{1,\ell-1}\right)^2\right] = \mathbb{E}\Bigg[\Bigg(&\left(\hat{x}^f_{1,\ell} - \int_0^T \hat{\nu}^f_{\mathrm{PLI},\ell}(t)\,\mathrm{d}t\right) - \left(\hat{x}^c_{1,\ell-1} - \int_0^T \hat{\nu}^c_{\mathrm{PLI},\ell-1}(t)\,\mathrm{d}t\right) \\ &+ \left(\int_0^T \hat{\nu}^f_{\mathrm{PLI},\ell}(t)\,\mathrm{d}t - \int_0^T \nu(t)\,\mathrm{d}t\right) + \left(\int_0^T \nu(t)\,\mathrm{d}t - \int_0^T \hat{\nu}^c_{\mathrm{PLI},\ell-1}(t)\,\mathrm{d}t\right)\Bigg)^2\Bigg] \\ = \mathbb{E}\Bigg[\Bigg(&\left(\int_0^T \hat{\nu}^f_{\mathrm{BMI},\ell}(t)\,\mathrm{d}t - \int_0^T \hat{\nu}^f_{\mathrm{PLI},\ell}(t)\,\mathrm{d}t\right) - \left(\int_0^T \hat{\nu}^c_{\mathrm{BMI},\ell-1}(t)\,\mathrm{d}t - \int_0^T \hat{\nu}^c_{\mathrm{PLI},\ell-1}(t)\,\mathrm{d}t\right) \\ &+ \left(\int_0^T \hat{\nu}^f_{\mathrm{PLI},\ell}(t)\,\mathrm{d}t - \int_0^T \nu(t)\,\mathrm{d}t\right) + \left(\int_0^T \nu(t)\,\mathrm{d}t - \int_0^T \hat{\nu}^c_{\mathrm{PLI},\ell-1}(t)\,\mathrm{d}t\right)\Bigg)^2\Bigg] \\ \le 4\Bigg(&\mathbb{E}\left[\left(\int_0^T \hat{\nu}^f_{\mathrm{BMI},\ell}(t)\,\mathrm{d}t - \int_0^T \hat{\nu}^f_{\mathrm{PLI},\ell}(t)\,\mathrm{d}t\right)^2\right] + \mathbb{E}\left[\left(\int_0^T \hat{\nu}^c_{\mathrm{BMI},\ell-1}(t)\,\mathrm{d}t - \int_0^T \hat{\nu}^c_{\mathrm{PLI},\ell-1}(t)\,\mathrm{d}t\right)^2\right] \\ &+ \mathbb{E}\left[\left(\int_0^T \hat{\nu}^f_{\mathrm{PLI},\ell}(t)\,\mathrm{d}t - \int_0^T \nu(t)\,\mathrm{d}t\right)^2\right] + \mathbb{E}\left[\left(\int_0^T \nu(t)\,\mathrm{d}t - \int_0^T \hat{\nu}^c_{\mathrm{PLI},\ell-1}(t)\,\mathrm{d}t\right)^2\right]\Bigg), \end{aligned} \tag{A.3}$$

where the inequality is obtained by applying Proposition 4.2. From Lemma A.1, it follows that the first and second expectations on the right-side of the inequality are $\mathcal{O}\left(h_\ell^3\right)$. From Proposition 4.1, the third and fourth expectations are $\mathcal{O}\left(h_\ell^2\right)$. This concludes the proof for scheme (4.8).

We now show that scheme (4.9) also has $\mathbb{E}\left[\left(\hat{x}^f_{1,\ell} - \hat{x}^c_{1,\ell-1}\right)^2\right] = \mathcal{O}\left(h_\ell^2\right)$. Under this scheme, we have

$$\begin{aligned}\mathbb{E}\left[\left(\hat{x}^f_{1,\ell} - \hat{x}^c_{1,\ell-1}\right)^2\right] &= \mathbb{E}\left[\left(\int_0^T \hat{\nu}^f_{\mathrm{PLI},\ell}(t)\,\mathrm{d}t - \int_0^T \hat{\nu}^c_{\mathrm{PLI},\ell-1}(t)\,\mathrm{d}t\right)^2\right] \\ &= \mathbb{E}\Bigg[\Bigg(\left(\int_0^T \hat{\nu}^f_{\mathrm{BMI},\ell}(t)\,\mathrm{d}t - \int_0^T \hat{\nu}^c_{\mathrm{BMI},\ell-1}(t)\,\mathrm{d}t\right) - \left(\int_0^T \hat{\nu}^f_{\mathrm{BMI},\ell}(t)\,\mathrm{d}t - \int_0^T \hat{\nu}^f_{\mathrm{PLI},\ell}(t)\,\mathrm{d}t\right)\end{aligned}$$

$$
\begin{aligned}
&+ \left(\int_0^T \hat{\nu}^c_{\mathrm{BMI},\ell-1}(t)\,\mathrm{d}t - \int_0^T \hat{\nu}^c_{\mathrm{PLI},\ell-1}(t)\,\mathrm{d}t\right)\Bigg)^2\Bigg] \\
\leq 3\Bigg(&\mathbb{E}\left[\left(\int_0^T \hat{\nu}^f_{\mathrm{BMI},\ell}(t)\,\mathrm{d}t - \int_0^T \hat{\nu}^c_{\mathrm{BMI},\ell-1}(t)\,\mathrm{d}t\right)^2\right] \\
&+ \mathbb{E}\left[\left(\int_0^T \hat{\nu}^f_{\mathrm{BMI},\ell}(t)\,\mathrm{d}t - \int_0^T \hat{\nu}^f_{\mathrm{PLI},\ell}(t)\,\mathrm{d}t\right)^2\right] \\
&+ \mathbb{E}\left[\left(\int_0^T \hat{\nu}^c_{\mathrm{BMI},\ell-1}(t)\,\mathrm{d}t - \int_0^T \hat{\nu}^c_{\mathrm{PLI},\ell-1}(t)\,\mathrm{d}t\right)^2\right]\Bigg),
\end{aligned}
$$

where the inequality is obtained by applying Proposition 4.2. The desired result follows from the bound for scheme (4.8), as shown previously, and Lemma A.1.

B Proof of Lemma 4.3

We first recall an useful result from Neuenkirch and Szpruch (2014)(see page 120, Section 3.1). Let $z(t) = \sqrt{\nu(t)}$, the dynamics of which can be obtained by applying Itô's rule to (2.1d). Under Assumption 4.1, there exists a bounded constant C such that

$$
\mathbb{E}\left[\sup_{n=0,\ldots,\lceil T/h\rceil} |z(t_n) - \hat{z}_n|^2\right] \leq Ch^2,
$$

where $\hat{z}_n$ denotes the discrete approximation to the exact value $z(t_n)$ at time t_n obtained by the Backward-Euler-Maruyama scheme (Neuenkirch and Szpruch, 2014). Using the above result, the proof of Lemma 4.3 can be obtained by closely following the steps of proof of Lemma 4.1, presented in Appendix A, using the idea of piecewise linear interpolant and Brownian motion interpolant, and noting that function $\beta(t)$ is bounded on $[0, T]$.

C Proof of Lemma 5.1

Without loss of generality, we can express G, H, and F as

$$
\begin{aligned}
G &= G_1 x_1 + \sum_{i=1}^{m} G_{d_i,1} x_{d_i,1} + \sum_{i=1}^{l} G_{f_i,1} x_{f_i,1} + G_c, \\
F &= F_1 x_1 + F_2 x_2 + \sum_{i=1}^{m} F_{d_i,1} x_{d_i,1} + \sum_{i=1}^{l} F_{f_i,1} x_{f_i,1} + \sum_{i=1}^{m} F_{d_i,2} x_{d_i,2} + \sum_{i=1}^{l} F_{f_i,2} x_{f_i,2} + F_c, \\
H &= \sum_{i=1}^{m} H_{d_i,2} x_{d_i,2} + H_c,
\end{aligned}
\tag{C.1}
$$

where all the coefficients $G_{(\cdot)}$, $F_{(\cdot)}$, and $H_{(\cdot)}$ are (deterministic) bounded constants. Under Assumption 5.1, the coefficient $F_{f_i,2}$, $i = 1, \ldots, l$, are zero.

First we consider the pure-diffusion case. Recall that the payoff in this case is given by

$$
\begin{aligned}
\mathcal{F}\Big(x_1, x_2, x_{d_1,1}, \ldots, x_{d_m,1}, x_{f_1,1}, \ldots, x_{f_l,1}, x_{d_1,2}, \ldots, x_{d_m,2}\Big)& \\
= S(0)\mathrm{e}^{G+F+H}\mathcal{N}(d_1) - K\mathrm{e}^{H}\mathcal{N}(d_2)&,
\end{aligned}
\tag{C.2}
$$

where

$$d_1 = \frac{\ln\left(\frac{S(0)}{K}\right) + F}{\sqrt{2G}} + \sqrt{2G}, \quad d_2 \;= d_1 - \sqrt{2G}.$$

First, we show that $\frac{\partial \mathcal{F}}{\partial x_1}$ is bounded. By Andersen and Piterbarg (2007), under Feller's condition $2\kappa_\nu \bar{\nu} > \sigma_\nu^2$, we have that $0 < \nu(t) < \infty$, $t \in [0, T]$. As a result, we have $0 < x_1 = \int_0^T \nu(t)\mathrm{d}t < \infty$. We also note that x_1 appears only in F and G. Furthermore, by inspecting (3.7a), if $G_1 \neq 0$, then $0 < G < \infty$. Now, for $G_1 \neq 0$ (and hence $G \neq 0$), we have

$$\begin{aligned}\frac{\partial \mathcal{F}}{\partial x_1} = S(0)(G_1 + F_1)\mathrm{e}^{G+F+H}\mathcal{N}(d_1) + S(0)\mathrm{e}^{G+F+H}\frac{\mathrm{e}^{-\frac{d_1^2}{2}}}{\sqrt{2\pi}}\left(\frac{F_1\sqrt{2G} - F\frac{1}{2\sqrt{2G}}}{2G} + \frac{1}{2\sqrt{2G}}\right)\\ - Ke^H\frac{\mathrm{e}^{-\frac{d_2^2}{2}}}{\sqrt{2\pi}}\left(\frac{F_1\sqrt{2G} - F\frac{1}{2\sqrt{2G}}}{2G}\right),\end{aligned}$$

which is bounded, noting $G \neq 0$. For $G_1 = 0$, then x_1 appears only in F, and the proof is similar in this case.

Next, we show that $\frac{\partial \mathcal{F}}{\partial x_2}$ is bounded. First, we note that, using (4.10) for the period $[0, T]$, we have

$$x_2 = \frac{\nu(T) - \nu(0) - \kappa_\nu \bar{\nu} T + \kappa_\nu x_1}{\sigma_\nu}.$$

Because $\nu(0)$, κ_ν, $\bar{\nu}$, and σ_ν, are constant, as well as x_1 is bounded, together with the boundedness of $\nu(T)$ (Andersen and Piterbarg, 2007), it follows that x_2 is bounded. We also note that x_2 only appears in F. We can compute $\frac{\partial \mathcal{F}}{\partial x_2}$ explicitly and it is straightforward to show that $\frac{\partial \mathcal{F}}{\partial x_2}$ is also bounded.

For the case of $\frac{\partial \mathcal{F}}{\partial x_{d_i,1}}$, $i = 1, \ldots, m$, and $\frac{\partial \mathcal{F}}{\partial x_{f_i,1}}$, $i = 1, \ldots, l$, as noted earlier, all of the variables are of the form $\int_0^T \beta(t)\sqrt{\nu(t)}\,\mathrm{d}t$ for positive bounded function $\beta(t)$, defined in (4.4). Since $\nu(t)$ is positive and bounded for $0 \le t \le T$, it follows that $x_{d_i,1}$, $i = 1, \ldots, m$, and $x_{f_i,1}$, $i = 1, \ldots, l$, are bounded and non-zero. We also note that, similar to x_1, these variables appear only in G and F. We can then compute the derivatives of f with respect to these variables explicitly, and show that they are are bounded, as we did for $\frac{\partial \mathcal{F}}{\partial x_1}$.

For the case $\frac{\partial \mathcal{F}}{\partial x_{d_i,2}}$, $i = 1, \ldots, m$, we first note that all of the variables are of the form $\int_0^T \beta(t)\mathrm{d}W(t)$, and hence, is unbounded. First, we consider $\frac{\partial \mathcal{F}}{\partial x_{d_i,2}}$, $i = 1, \ldots, m$. By inspection of (3.7), we see that $x_{d_i,2}$ appears only in F and H, and not in G, with

$$F_{d_i,2} + H_{d_i,2} = 0 \quad \Leftrightarrow \quad F_{d_i,2} = -H_{d_i,2}, \quad i = 1, \ldots, m. \tag{C.3}$$

By (C.3), we also have e^{G+F+H} does not depends on $x_{d_i,2}$. We have

$$\begin{aligned}\frac{\partial \mathcal{F}}{\partial x_{d_i,2}} &= S(0)\mathrm{e}^{G+F+H}\frac{\mathrm{e}^{-\frac{d_1^2}{2}}}{\sqrt{2\pi}}\frac{F_{d_i,2}}{\sqrt{2G}} - K\,H_{d_i,2}\,\mathrm{e}^H\mathcal{N}(d_2) - K\,\mathrm{e}^H\frac{\mathrm{e}^{-\frac{d_2^2}{2}}}{\sqrt{2\pi}}\frac{F_{d_i,2}}{\sqrt{2G}}\\ &= S(0)\,F_{d_i,2}\,\mathrm{e}^{G+F+H}\frac{\mathrm{e}^{-\frac{d_1^2}{2}}}{2\sqrt{\pi G}} - K\,H_{d_i,2}\,\mathrm{e}^H\left(\mathcal{N}(d_2) - \frac{\mathrm{e}^{-\frac{d_2^2}{2}}}{2\sqrt{\pi G}}\right).\end{aligned}$$

We consider the following two limit cases:

- As $F_{d_i,2}x_{d_i,2} \to \infty$, by (C.3), we have $H_{d_i,2}x_{d_i,2} \to -\infty$. In this case, from the formulas for d_1 and d_2, we have both d_1 and $d_2 \to \infty$ and thus, $\mathcal{N}(d_2) \to 1$. We also have $\mathrm{e}^H \to 0$, $\mathrm{e}^{-\frac{d_1^2}{2}} \to 0$, $\mathrm{e}^{-\frac{d_2^2}{2}} \to 0$. Thus, $\displaystyle\lim_{F_{d_i,2}x_{d_i,2}\to\infty}\frac{\partial \mathcal{F}}{\partial x_{d_i,2}} = 0$.

- As $F_{d_i,2}x_{d_i,2} \to -\infty$, by (C.3), we have $H_{d_i,2}x_{d_i,2} \to \infty$. In this case, from the formulas for d_1 and d_2, we have both d_1 and $d_2 \to -\infty$, and thus $\mathcal{N}(d_2) \to 0$. Also, we have $\mathrm{e}^H \to \infty$ and both $\mathrm{e}^{-\frac{d_1^2}{2}} \to 0$, and $\mathrm{e}^{-\frac{d_2^2}{2}} \to 0$. We have

$$\begin{aligned}\lim_{F_{d_i,2}x_{d_i,2}\to-\infty} \frac{\partial \mathcal{F}}{\partial x_{d_i,2}} &= \lim_{F_{d_i,2}x_{d_i,2}\to-\infty} S(0)F_{d_i,2}\mathrm{e}^{G+F+H}\frac{\mathrm{e}^{-\frac{d_1^2}{2}}}{2\sqrt{\pi G}} \\ &- \lim_{F_{d_i,2}x_{d_i,2}\to-\infty} KH_{d_i,2}\mathrm{e}^H\left(\mathcal{N}(d_2) - \frac{\mathrm{e}^{-\frac{d_2^2}{2}}}{2\sqrt{\pi G}}\right) \\ &= -\lim_{F_{d_i,2}x_{d_i,2}\to\infty} KH_{d_i,2}\mathrm{e}^H\mathcal{N}(d_2) = 0,\end{aligned}$$

where the last equality can be obtained by L'Hopital rule.

Furthermore, it is straightforward to see that $\frac{\partial \mathcal{F}}{\partial x_{d_i,2}}$ is bounded for $-\infty < F_{d_i,2}x_{d_i,2} < +\infty$. We can conclude that in this case $\frac{\partial \mathcal{F}}{\partial x_{d_i,2}}$ is bounded.

Finally, we show that, given all partial derivatives of $\mathcal{F}(\cdot)$ with respect to the variables x_1, x_2, $x_{d_i,1}$, $i = 1,\ldots,m$, $x_{f_i,1}$, $i = 1,\ldots,l$, and $x_{d_i,2}$, $i = 1,\ldots,m$, are bounded, $\mathcal{F}(\cdot)$ is Lipschitz , satisfying the Lipschitz bound (5.1). We note that the boundedness of $\frac{\partial \mathcal{F}}{x_{(\cdot)}}$ implies that

$$\left|\mathcal{F}(\ldots,x_{(\cdot)}^{(1)},\ldots) - \mathcal{F}(\ldots,x_{(\cdot)}^{(2)},\ldots)\right| \le C_{(\cdot)}\left|x_{(\cdot)}^{(1)} - x_{(\cdot)}^{(2)}\right|, \tag{C.4}$$

for some constant $C_{(\cdot)}$. Now, using a telescoping sum, we have

$$\begin{aligned}&\left|\mathcal{F}\left(x_1^{(1)},x_2^{(1)},x_{d_1,1}^{(1)},\ldots,x_{d_m,1}^{(1)},x_{f_1,1}^{(1)},\ldots,x_{f_l,1}^{(1)},x_{d_1,2}^{(1)},\ldots,x_{d_m,2}^{(1)}\right)\right. \\ &\qquad\left. - \mathcal{F}\left(x_1^{(2)},x_2^{(2)},x_{d_1,1}^{(2)},\ldots,x_{d_m,1}^{(2)},x_{f_1,1}^{(2)},\ldots,x_{f_l,1}^{(2)},x_{d_2,2}^{(2)},\ldots,x_{d_m,2}^{(2)}\right)\right| \\ &= \left|\mathcal{F}\left(x_1^{(1)},x_2^{(1)},x_{d_1,1}^{(1)},\ldots\right) - \mathcal{F}\left(x_1^{(2)},x_2^{(1)},x_{d_1,1}^{(1)},\ldots\right)\right. \\ &\qquad + \mathcal{F}\left(x_1^{(2)},x_2^{(1)},x_{d_1,1}^{(1)},\ldots\right) - \mathcal{F}\left(x_1^{(2)},x_2^{(2)},x_{d_1,1}^{(1)},\ldots\right) \\ &\qquad \left.+ \mathcal{F}\left(x_1^{(2)},x_2^{(2)},x_{d_1,1}^{(1)},\ldots\right) - \ldots\right| \\ &\le \left|\mathcal{F}\left(x_1^{(1)},x_2^{(1)},x_{d_1,1}^{(1)},\ldots\right) - \mathcal{F}\left(x_1^{(2)},x_2^{(1)},x_{d_1,1}^{(1)},\ldots\right)\right| \\ &\qquad + \left|\mathcal{F}\left(x_1^{(2)},x_2^{(1)},x_{d_1,1}^{(1)},\ldots\right) - \mathcal{F}\left(x_1^{(2)},x_2^{(2)},x_{d_1,1}^{(1)},\ldots\right)\right| + \ldots \\ &\le C\left(\sum_{i=1}^{2}\left|x_i^{(1)} - x_i^{(2)}\right| + \sum_{i=1}^{m}\left|x_{d_i,1}^{(1)} - x_{d_i,1}^{(2)}\right| + \sum_{i=1}^{l}\left|x_{f_i,1}^{(1)} - x_{f_i,1}^{(2)}\right| + \sum_{i=1}^{m}\left|x_{d_i,2}^{(1)} - x_{d_i,2}^{(2)}\right|\right),\end{aligned} \tag{C.5}$$

where in the last inequality, we use (C.4) and $C = \max C_{(\cdot)}$. This completes the proof.

D Double-exponential (Kou, 2002)

In the case $w = \ln(y) \sim$ Double-Exponential(p, η_1, η_2), where $0 \leq p \leq 1, \eta_1 > 1, \eta_2 > 0$, the European call option value is given by (Dang et al., 2017)[Corollary 3.1]

$$V(S(0), 0, \cdot) = \mathbb{E}\left[\sum_{n=0}^{\infty} \frac{(\lambda T)^n}{n!} \left\{S(0)\mathrm{e}^{(G+F+H)} A_n - K\mathrm{e}^{H} B_n\right\}\right], \tag{D.1}$$

where

$$A_n = \frac{1}{\sqrt{2\pi}}\left[\sum_{k=1}^{n} P_{n,k}\left(\eta_1\sqrt{2G}\right)^k \mathrm{e}^{G\,(1-\eta_1)^2} I_{k-1}\left(-d_1, (1-\eta_1)\sqrt{2G}, -1, (1-\eta_1)\sqrt{2G}\right)\right. \tag{D.2a}$$

$$\left. + Q_{n,k}\left(\eta_2\sqrt{2G}\right)^k \mathrm{e}^{G\,(1+\eta_2)^2} I_{k-1}\left(-d_1, (1+\eta_2)\sqrt{2G}, 1, -(1+\eta_2)\sqrt{2G}\right)\right],$$

$$B_n = \frac{1}{\sqrt{2\pi}}\left[\sum_{k=1}^{n} P_{n,k}\left(\eta_1\sqrt{2G}\right)^k \mathrm{e}^{G\,(\eta_1)^2} I_{k-1}\left(-d_2, -\eta_1\sqrt{2G}, -1, -\eta_1\sqrt{2G}\right)\right. \tag{D.2b}$$

$$\left. + Q_{n,k}\left(\eta_2\sqrt{2G}\right)^k \mathrm{e}^{G\,(\eta_2)^2} I_{k-1}\left(-d_2, \eta_2\sqrt{2G}, 1, -\eta_2\sqrt{2G}\right)\right].$$

Here,

$$\begin{aligned} P_{n,k} &= \sum_{i=k}^{n-1}\binom{n-k-1}{i-k}\binom{n}{i}\left(\frac{\eta_1}{\eta_1+\eta_2}\right)^{i-k}\left(\frac{\eta_2}{\eta_1+\eta_2}\right)^{n-i} p^i\, q^{n-i}, \qquad 1 \leq k \leq n-1, \\ Q_{n,k} &= \sum_{i=k}^{n-1}\binom{n-k-1}{i-k}\binom{n}{i}\left(\frac{\eta_1}{\eta_1+\eta_2}\right)^{n-i}\left(\frac{\eta_2}{\eta_1+\eta_2}\right)^{i-k} p^{n-i}\, q^{i}, \qquad 1 \leq k \leq n-1, \end{aligned} \tag{D.3}$$

with $P_{n,n} = p^n$ and $Q_{n,n} = q^n$, and d_1 and d_2 are defined in (3.10). Also, $\mathrm{Hh}_k(\cdot)$, $I_k(\cdot;\cdot)$ are defined as

$$\begin{aligned} \mathrm{Hh}_k(x) &= \frac{1}{k!}\int_x^{\infty}(t-x)^k\,\mathrm{e}^{-\frac{1}{2}t^2}\mathrm{d}t, \quad k = 0, 1, 2, \ldots \\ &\qquad \text{with } \mathrm{Hh}_{-1}(x) = \mathrm{e}^{-x^2/2}, \text{ and } \mathrm{Hh}_0(x) = \sqrt{2\pi}\mathcal{N}(-x), \\ I_k(c; \alpha, \beta, \delta) &= \int_c^{\infty} \mathrm{e}^{\alpha x}\,\mathrm{Hh}_k(\beta x - \delta)\,\mathrm{d}x, \end{aligned} \tag{D.4}$$

for arbitrary constant α, c, β, and δ.

Acknowledgment

The author would like to thank Professor Mike Giles of Oxford University for very useful comments on earlier drafts of this paper. This research was supported in part by a University of Queensland Early Career Researcher Grant (Grant number: 1006301-01-298-21-609775).

References

Alizadeh, S., M. Brandt, and F. Diebold (2002). Range-based estimation of stochastic volatility models. *Journal of Finance* 57, 1047–1091.

Andersen, L. and V. Piterbarg (2007). Moment explosions in stochastic volatility models. *Finance and Stochastics* 11, 29–50.

Andersen, T. G., L. Benzoni, and J. Lund (2002). An empirical investigation of continuous-time equity return models. *Journal of Finance* 57, 1239–1284.

Bakshi, G., C. Cao, and Z. Chen (2000). Pricing and hedging long-term options. *Journal of Econometrics* 94, 277–318.

Bakshi, G., C. Cao, and C. Zhiwu (1997). Empirical performance of alternativee option pricing models. *Journal of Finance* 52, 2003–2049.

Bates, D. (1996). Jumps and stochastic volatility: Exchange rate process implicit in Deutsche Mark options. *Review of Financial Studies* pp. 69–107.

Burgos, S. and M. B. Giles (2012). Computing Greeks using multilevel path simulation. In L. Plaskota and H. Woźniakowski, eds., *Monte Carlo and Quasi-Monte Carlo Methods 2010*. Springer-Verlag, Berlin Heidelberg New York.

Cont, R. and P. Tankov (2004). *Financial Modelling with Jump Processes*. Chapman and Hall.

Dang, D. M., K. R. Jackson, and M. Mohammadi (2015a). Dimension and variance reduction for Monte-Carlo methods for high-dimensional models in finance. *Applied Mathematical Finance* 22(6), 522–552.

Dang, D. M., K. R. Jackson, and S. Sues (2017). A dimension and variance reduction Monte-Carlo method for option pricing under jump-diffusion models. *Applied Mathematical Finance* To appear.

Dang, D. M. and L. Ortiz-Gracia (2017). A dimension reduction Shannon-wavelet based method for option pricing. *Journal of Scientific Computing* To appear.

Dang, D. M., Q. Xu, and S. Wu (2015b). Multilevel dimension reduction Monte-Carlo simulation for high-dimensional stochastic models in finance. In *Proceedings of the 15th International Conference in Computational Science (ICCS)*, volume 51, pp. 1583–1592. Procedia Computer Science.

Delbaen, F. and W. Schachermayer (1994). A general version of the fundamental theorem of asset pricing. *Mathematische Annalen* 300, 463–520.

Duffie, D. and P. Glynn (1995). Efficient Monte-Carlo simulation of security prices. *The Annals of Applied Probability* 5(4), 897–905.

Giles, M. B. (2006). Improved multilevel Monte Carlo convergence using the Milstein scheme. In H. N. A. Keller, S. Heinrich, ed., *Monte Carlo and Quasi-Monte Carlo Methods 2006*. Springer-Verlag, Berlin Heidelberg New York.

Giles, M. B. (2008). Multi-level Monte Carlo path simulation. *Operations Research* 56, 607–617.

Giles, M. B., K. Debrabant, and A. Rößler (2013). Numerical analysis of multilevel Monte Carlo path simulation using the Milstein discretisation. Working paper, available at http://arxiv.org/abs/1302.4676.

Giles, M. B., D. Higham, and X. Mao (2009). Analysing multilevel Monte Carlo for options with non-globally Lipschitz payoff. *Finance and Stochastics* 13, 403–413.

Giles, M. B. and L. Szpruch (2014). Antithetic multilevel Monte Carlo esitmation for multi-dimensional SDEs without lévy area simulation. *The Annals of Applied Probability* 24, 1585–1620.

Glasserman, P. (2003). *Monte Carlo methods in financial engineering*. Springer, first edition.

Grzelak, L. A. and C. W. Oosterlee (2011). The affine Heston model with correlated Gaussian interest rates for pricing hybrid derivatives. *Quantitative Finance* 11, 1647–1663.

Grzelak, L. A. and C. W. Oosterlee (2012a). On cross-currency models with stochastic volatility and correlated interest rates. *Applied Mathematical Finance* 19, 1–35.

Grzelak, L. A. and C. W. Oosterlee (2012b). On the Heston model with stochastic interest rates. *SIAM Journal of Fianancial Mathematics* 2, 255–286.

Haastrecht, A. V., R. Lord, A. Pelsseri, and D. Schrager (2009). Generic pricing of FX, inflation and stock options under stochastic interest rates and stochastic. *Insurance: Math. Econ.* 45, 436–448.

Haastrecht, A. V. and A. Pelsser (2011). Generic pricing of FX, inflation and stock options under stochastic interest rates and stochastic volatility. *Quantitative Finance* 11, 665–691.

Haentjens, T. and K. J. in 't Hout (2012). Alternating direction implicit finite difference schemes for the Heston-Hull-White partial differential equation. *Journal of Computational Finance* 16(1), 83–110.

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

Kloeden, P. and E. Platen (1992). *Numerical solution of stochastic differential equations*. Springer, first edition.

Kou, S. G. (2002). A jump diffusion model for option pricing. *Management Science* 48, 1086–1101.

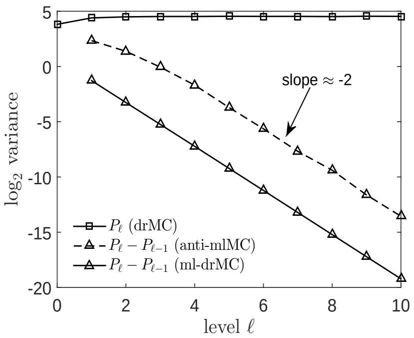
Merton, R. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics* 3, 125–144.

Neuenkirch, A. and L. Szpruch (2014). First order strong approximations of scalar SDEs with values in a domain. *Numerische Mathematik* 128, 103–136.

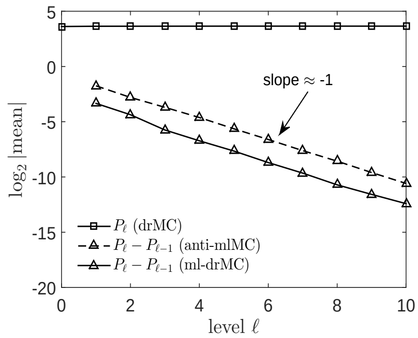
Xia, Y. (2011). Multilevel Monte Carlo method for jump-diffusion SDEs. Working paper, available at http://arxiv.org/abs/1106.4730.

Xia, Y. (2013). *Multilevel Monte Carlo for jump processes*. Ph.D. thesis, University of Oxford.

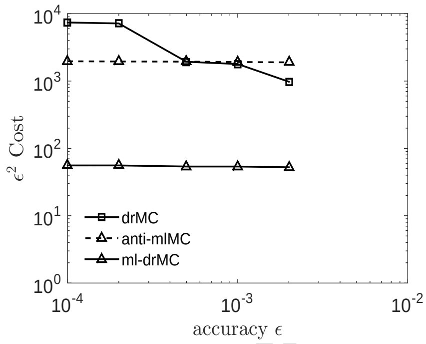
Xia, Y. and M. B. Giles (2012). Multilevel Monte Carlo method for jump-diffusion SDEs. In L. Plaskota and H. Woźniakowski, eds., *Monte Carlo and Quasi-Monte Carlo Methods 2010*. Springer-Verlag, Berlin Heidelberg New York.



(a) **Price** - variance ($\epsilon = 10^{-3}$)

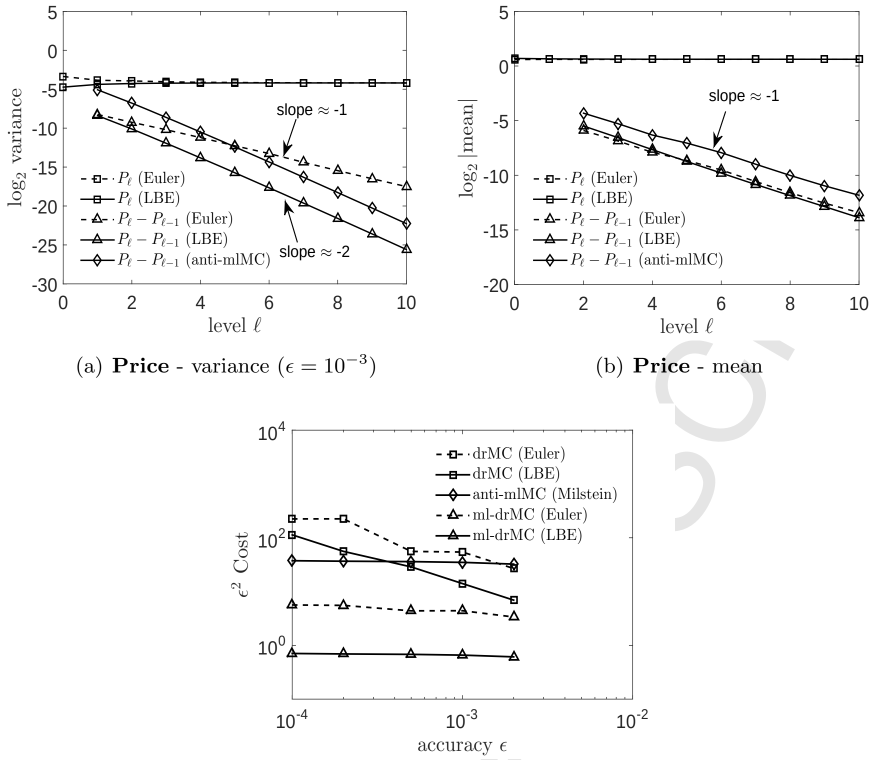


(b) **Price** - mean



(c) **Price** - efficiency

FIGURE D.1: *Plots for price under the 6-factor pure-diffusion model.*

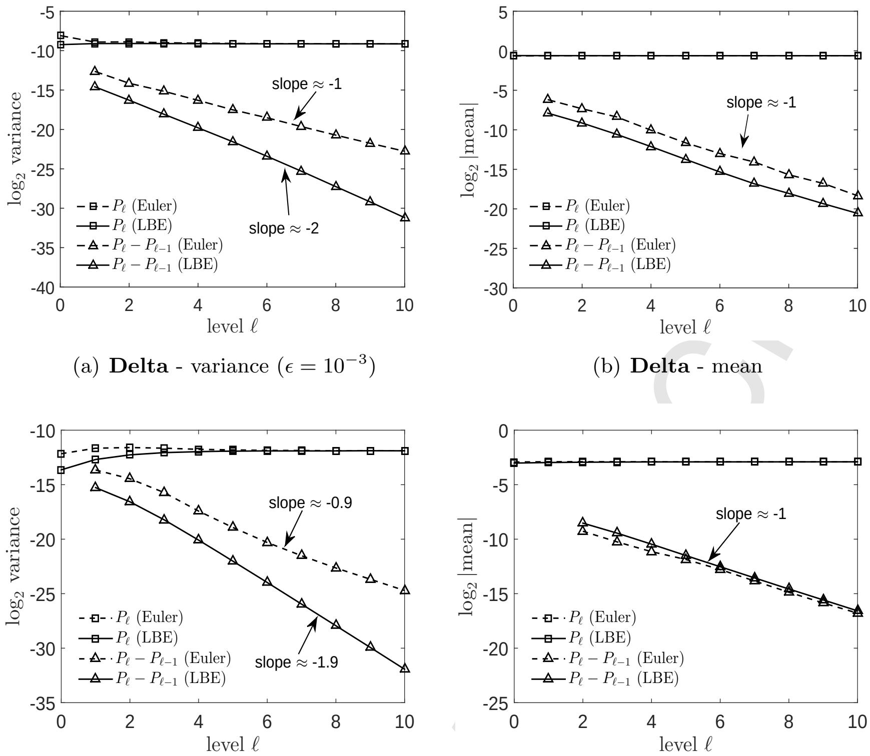


(a) **Price** - variance ($\epsilon = 10^{-3}$)

(b) **Price** - mean

(c) **Price** - efficiency

FIGURE D.2: *Plots for price under the 3-factor HHW jump-diffusion model with normal jumps. Call option's price* ≈ 1.535.



(a) **Delta** - variance ($\epsilon = 10^{-3}$)

(b) **Delta** - mean

(c) **Gamma** - variance ($\epsilon = 10^{-3}$)

(d) **Gamma** - mean

FIGURE D.3: *Plots for Delta and Gamma under the 3-factor HHW jump-diffusion model with normal jumps. Call option's Delta* ≈ 0.648, *Gamma* ≈ 0.133.

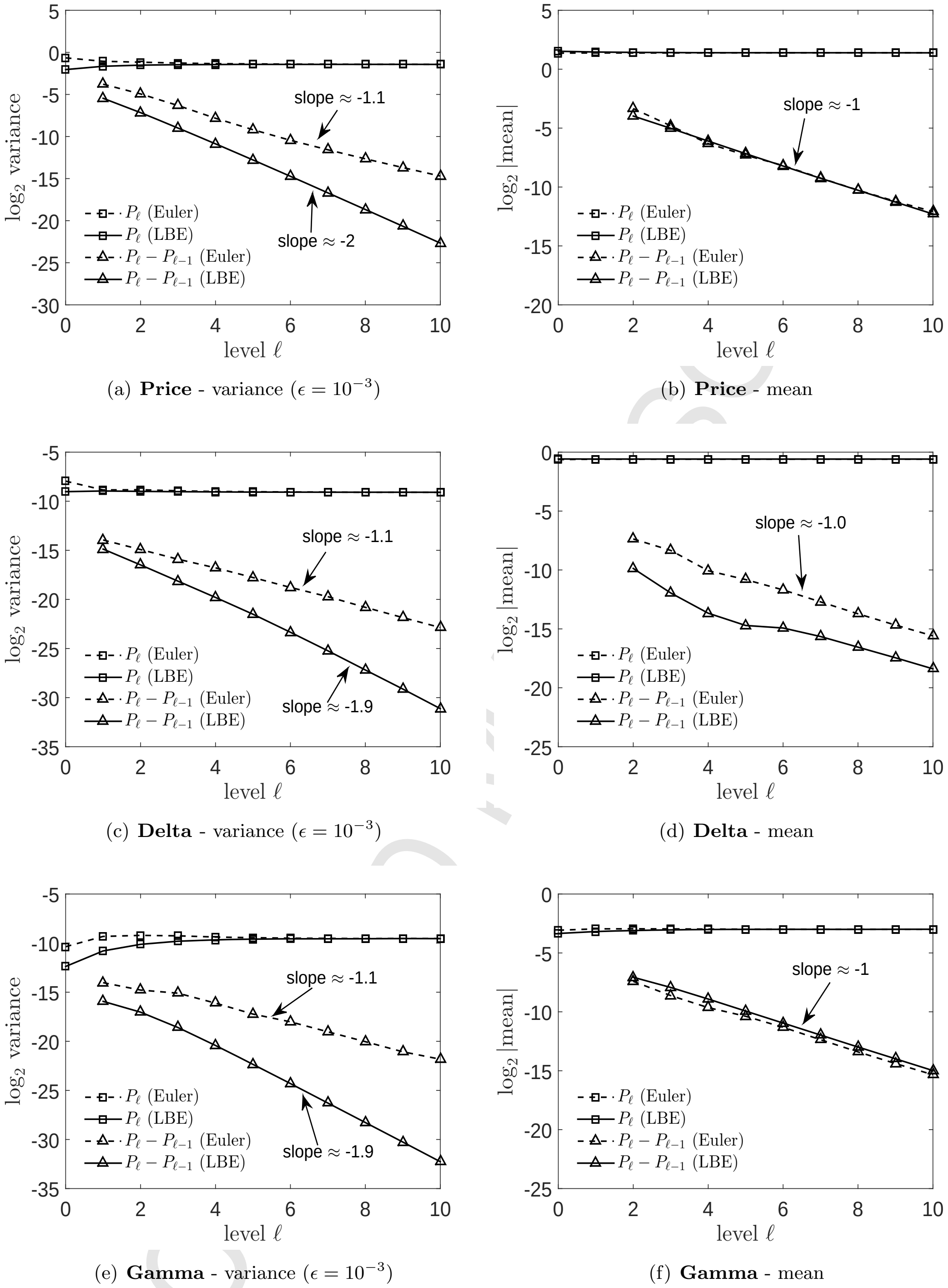


FIGURE D.4: *Variance and mean plots for the option price, Delta, and Gamma, under the 3-factor HHW jump-diffusion model with* <u>*double-exponential jumps*</u>*. Call option price* ≈ 1.302*, Delta* ≈ 0.664*, Gamma* ≈ 0.125.

ml-drMC (LBE)	drMC (LBE)	anti-mlMC (Milstein)
12.563512(0.000701)	12.563405 (0.000705)	12.563221 (0.000705)

TABLE D.1: *Option prices obtained by different methods under the 6-factor pure-diffusion model* (6.2). *For the anti-mlMC and ml-drMC methods,* $\epsilon = 10^{-3}$.

ml-drMC (Euler)	ml-drMC (LBE)	drMC (Milstein)	anti-mlMC (Milstein)
1.535023(0.000706)	1.535145 (0.000703)	1.535381 (0.000704)	1.535233 (0.000704)

Table D.2: *Call option's prices obtained by different methods under the 3-factor HHW jump-diffusion model* (6.1) *with normal jump. For the ml-drMC and anti-mlMC methods,* $\epsilon = 10^{-3}$.