Research article

Multigrid method for pricing European options under the CGMY process

Justin W. L. Wan∗

David R. Cheriton School of Computer Science, University of Waterloo, 200 University Avenue West, Waterloo, ON N2L 3G1, Canada

∗ **Correspondence:** Email: justin.wan@uwaterloo.ca; Tel: +15198884567 ext 34468.

Abstract: We propose a fast multigrid method for solving the discrete partial integro-differential equations (PIDEs) arising from pricing European options when the underlying asset is driven by an infinite activity Lévy process. We consider the CGMY model whose kernel singularity gets worse when the parameter Y approaches two. Due to the integral term, the discretization matrix is dense. In order to obtain an efficient multigrid method, we apply a fixed point iteration as a smoother for multigrid. In each smoothing step, we only need to solve a sparse matrix corresponding to the differential operator and compute a matrix-vector product involving the integral operator by a fast Fourier transform (FFT). We prove that the fixed point iteration smoother is effective reducing the high frequency components. Moreover, we also prove a two-grid convergence of the multigrid method by a local mode analysis. We demonstrate the effectiveness of the multigrid method by solving the option pricing equation under the CGMY model with finite and infinite variation processes.

Keywords: multigrid; two-grid analysis; Black-Scholes equation; CGMY model; PIDE
Mathematics Subject Classification: 65F10, 65M55

1. Introduction

The standard Black-Scholes model for pricing options assumes the underlying asset price satisfies the geometric Brownian motion model. However, it has been observed in practice that market prices often have large jumps that do not follow the geometric Brownian motion. Jump diffusion models have been proposed to more accurately capture the actual market behavior; see e.g., [1]. Recently, models based on the Lévy process have become popular in the financial literature [1–7]. Option pricing under the exponential Lévy process with finite activity [5, 8–12] and infinite activity [13–17] has been extensively studied. Under this model, pricing of the European option requires solving a partial integro-differential equation (PIDE), which consists of differential and integral terms in the equation.

Numerical solution of PIDEs is challenging since the discretization of the integral term will give

rise to a dense matrix. In addition, the singularity of the Lévy measure inside the jump integral may lead to a less accurate numerical approximate solution. (More discussion about the singularity of the jump integral will be given in the next section.) In the literature, the jump process can be classified into finite activity, infinite activity with finite variation, and infinite activity with infinite variation, where the singularity increases from the first to the later cases. In [14], the jump integral was split into local and nonlocal parts with the local term computed using implicit time stepping and the nonlocal term computed using explicit time stepping. In [13, 18], an integration by parts technique was used to transform the integral part into a weakly singular Volterra equation. A collocation method was then applied to the CGMY (Carr-Geman-Madan-Yor) process [2] in the case of infinite activity and finite variation. In [16], a wavelet Galerkin finite element technique was applied to an infinite activity process.

In this paper, we are concerned with the numerical solution method proposed by Wang et al. [19]. This approach is able to achieve second order accuracy under an infinite activity and finite variation process, and a better than first order accurate for an infinite variation process. Moreover, it is applicable to different types of options such as European and American, non-constant coefficient PIDEs, non-constant volatility, and various types of initial and boundary conditions. In each time step, a fixed point iteration is performed to compute the option value. In this approach, an FFT is used to efficiently compute the jump integral and more importantly, it avoids inverting a full dense matrix arising from the discretization of the jump integral. However, the convergence of the fixed point iteration can be very slow. In fact, it can be proved that the convergence rate deteriorates as the number of grid points increases (Y larger than 1) or the singularity of the Lévy measure increases. It will be made more precise in the next section.

Instead of fixed point iteration, we propose an efficient multigrid method for solving the linear system arising from the discretization of the PIDE using the approach in [19]. Multigrid has shown to be a powerful and one of the most efficient numerical techniques for solving elliptic partial differential equations (PDEs) [20–22]. Its convergence rate is often independent of the mesh size [21]. Sophisticated smoothing [20, 23, 24], coarsening, and interpolation [25–28] techniques have been developed. The technique has also been extended to convection dominated problems [29–32], and hyperbolic equations [33–35]. Recently, it has been applied to solving American option equations [36]. Multigrid methods have been developed for option pricing in the literature, for instance in [37–39]. However these methods do not consider the CGMY or any jump models.

The success of a multigrid technique depends on smoothing and its interaction with the coarse grid correction. Our idea is to apply the fixed point iteration as a smoother in the multigrid framework, rather than to use it as a solver, for solving the PIDE arising from option pricing. We will perform a smoothing analysis and show that the fixed point iteration can effectively reduce high frequency errors. Moreover, using a local mode analysis, we provide a two-grid convergence analysis of the method. Direct discretization is used for the coarse grid matrices. Linear interpolation and full weighting are used for inter-grid transfer.

We remark that the proposed multigrid method can be used with the penalty method for pricing American options. The discretized linear system is almost the same as that of the European option with some of the diagonal entries modified as a result of the early exercise constraint. Due to the similarity, we will focus only on European option in this paper.

In Section 2, we describe the PIDE for option pricing under the Lévy process and derive the linear

system resulting from a finite difference method. In Section 3, we present a multigrid method for solving the linear system. Moreover, we provide a smoothing analysis for the proposed smoother and a convergence analysis for the two-grid method. Finally, numerical results are given in Section 4 to demonstrate the fast convergence of the proposed multigrid method by solving PIDEs under a Lévy process with finite activity, infinite activity with finite variation, and infinite activity with infinite variation. Compared with the fixed point iteration which may take tens to hundreds or even thousands of iterations to converge, our proposed multigrid method converges within 10 iterations as the Y parameter of the CGMY model increases from 0 (mildly singular) to close to 2 (highly singular).

2. Model equation

In Black-Scholes modelling with jumps, the value of the European option price, V, satisfies the following partial integro-differential equation (PIDE) [1]:

$$
\begin{aligned}
V_\tau \;=\; & \frac{\sigma^2}{2} S^2 V_{SS} + (r - q) S V_S - rV \\
& + \int_{-\infty}^{\infty} v(y)[V(S e^y, \tau) - V(S, \tau) - S(e^y - 1) V_S]\, dy,
\end{aligned}
\tag{2.1}
$$

where r is the risk-free interest rate, q the continuous dividend yield, and σ the volatility. The parameter τ is the time from expiry; i.e., $\tau = 0$ denotes the expiry time T and $\tau = T$ denotes the start of the option. The evolution of the underlying risky asset price S is driven by a Lévy process whose Lévy measure v satisfies

$$
\int_{|y|<1} y^2 v(dy) < \infty, \qquad \int_{|y|\geq 1} v(dy) < \infty.
\tag{2.2}
$$

Under the CGMY model [2], v is defined as follows

$$
v(y) = \frac{C e^{-My}}{y^{1+Y}} 1_{y>0} + \frac{C e^{-G|y|}}{|y|^{1+Y}} 1_{y<0},
\tag{2.3}
$$

where

$$
1_{y>0} = \begin{cases} 1, & \text{if } y > 0 \\ 0, & \text{otherwise,} \end{cases} \qquad 1_{y<0} = \begin{cases} 1, & \text{if } y < 0 \\ 0, & \text{otherwise,} \end{cases}
$$

are the indicator variables. The parameter $C > 0$ is the measure of the overall level of activity, $G \geq 0$ and $M \geq 0$ control the rate of exponential decay on the left and right of the Lévy density, and $Y < 2$ describes the behavior of the Lévy density in the neighborhood of zero where the density tends to infinity.

The parameter Y is of interest mathematically and numerically since its value determines the regularity of v. If $Y < 0$, the measure v integrates to a finite value yielding a process of finite activity. If $Y \in [0, 1]$, the process displays infinite activity but finite variation since $\int_{|y|<1} y^2 v(dy) < \infty$. If $Y \in (1, 2)$, the process is said to have infinite activity and infinite variation.

2.1. Discretization

There have been a number of discretization methods for solving PIDEs [13, 14, 16, 18]. Here, we will consider the method proposed in [19], which is easy to implement and applicable to non-constant

coefficients. To be self contained, we present the major steps here and refer the readers to [19] for details. Meanwhile, notations are introduced which will be used in the smoothing analysis in Section 3.1.

The jump integral in (2.1) defined on $[-\infty, \infty]$ is first approximated by a finite domain $[y_{min}, y_{max}]$ with appropriately chosen y_{min} and y_{max} so that the error of approximation is negligibly small. The interval is divided into N subintervals $[y_j - \Delta y/2, y_j + \Delta y/2]$, $j = 0, 1, \ldots, N - 1$, where $\Delta y = (y_{max} - y_{min})/N$. Note that the Lévy measure, $\nu(y)$ in (2.3) goes to infinity as y goes to 0, and it is only second moment integrable in general; see (2.2). The computation of the approximate integral is splitted into three parts: I_{Ω_0}, I_{Ω_1} and I_{Ω_2}, corresponding to the three regions

$$\Omega_0 = \left\{ y \middle| -\frac{\Delta y}{2} \le y \le \frac{\Delta y}{2} \right\}, \qquad \Omega_1 = \left\{ y \middle| \frac{\Delta y}{2} < |y| < 1 \right\},$$
$$\Omega_2 = \left\{ y \middle| y_{min} \le y \le -1 \text{ or } 1 \le y \le y_{max} \right\}.$$

Notice that $\nu(y)$ is smooth on Ω_2 but exhibits singular behaviours on Ω_0 and Ω_1. Using the discretization techniques introduced in [19] to handle the singularity, the integral can be approximated by

$$\frac{\bar{\sigma}}{2} S^2 V_{SS} - \kappa S V_S - \lambda V + \sum_{j=0}^{N-1} V(S e^{y_j}) \hat{\gamma}(y_j), \tag{2.4}$$

where

$$\bar{\sigma} = \int_{-\Delta y/2}^{\Delta y/2} \nu(y)(e^y - 1)^2 \, dy, \quad \lambda = \sum_{j=0}^{N-1} \hat{\gamma}(y_j), \quad \kappa = \sum_{j=0}^{N-1} (e^{y_j} - 1)\hat{\gamma}(y_j), \tag{2.5}$$

and $\hat{\gamma}(y_j)$ is the value of the following integral $\gamma(y_j)$ computed by a quadrature rule:

$$\gamma(y_j) = \begin{cases} \frac{1}{y_j^2} \int_{y_j-\Delta y/2}^{y_j+\Delta y/2} y^2 \nu(y) dy; & \text{if } y_j \in \Omega_1 \\ \int_{y_j-\Delta y/2}^{y_j+\Delta y/2} \nu(y) dy; & \text{if } y_j \in \Omega_2 \\ 0; & \text{if } y_j \in \Omega_0. \end{cases}$$

As mentioned above, $\nu(y)$ is smooth on Ω_2. The quadrature rule can be just any standard numerical integration method such as the composite trapezoidal rule. On Ω_1, a special quadrature rule is needed to capture the singular behaviour of $\nu(y)$. The idea is to choose quadrature weights such that the integral is exact for polynomials of up to degree 3. Thus the accuracy of the quadrature rule is fourth order which is needed to compensate for the $1/y_j^2$ factor when y_j is near 0.

Combining (2.1) and (2.4), the equation to be solved can be written as

$$V_\tau = \frac{\sigma^2 + \bar{\sigma}}{2} S^2 V_{SS} + (r - q - \kappa) S V_S - (r + \lambda) V + \sum_{j=0}^{N-1} V(S e^{y_j}) \hat{\gamma}(y_j). \tag{2.6}$$

We see that the complex integral is simplified to a discrete sum.

A semi-Lagrangian method is used to discretize the first order term. Let $[S_1, \ldots, S_m]$ be a set of grid points for S and τ^n be the n-th time step. Also let $V_\ell^n \approx V(S_\ell, \tau^n)$ be the approximate value at (S_ℓ, τ^n). The semi-Lagrangian method traces node S_ℓ at $\tau = \tau^{n+1}$ along the characteristic trajectory

back to $\tau = \tau^n$ at a point $S_{\ell*}$, which in general is not a grid node S_j. Let $V_{\ell*}^n$ be the option value at $S_{\ell*}$, which is approximated by interpolating the nearby values of V_ℓ^n. In order to achieve second order accuracy, a quadratic interpolation Φ^{n+1} is used and the approximate value is denoted by $\Phi^{n+1}V_\ell^n$. Note that the superscript $n+1$ for Φ^{n+1} emphasizes that the interpolation operator depends on time τ^{n+1} since the location to be interpolated, $S_{\ell*}$, is derived from S_ℓ at τ^{n+1}.

Applying a fully-implicit timestepping with step size $\Delta\tau$ along the characteristic direction gives:

$$\frac{V_\ell^{n+1} - \Phi^{n+1}V_\ell^n}{\Delta\tau} = (\mathcal{L}V^{n+1})_\ell + (\mathcal{Q}V^{n+1})_\ell,$$

where

$$\mathcal{L}V = \frac{\sigma^2 + \bar{\sigma}}{2}S^2 V_{SS} - (r + \lambda)V, \qquad \mathcal{Q}V = \sum_{j=0}^{N-1} V(S\,e^{y_j}, \tau)\hat{\gamma}(y_j).$$

We note that the Crank-Nicolson scheme can also be used to obtain second order accuracy in time. However, since there is no difference in applying the multigrid method for solving the discrete linear system arising from the two timestepping schemes, we will focus on the fully-implicit discretization.

The computation of $(\mathcal{Q}V^{n+1})_\ell$ can be expensive in general. However, it can be accurately approximated by an FFT method [12], and the approximate value is denoted by $(\mathcal{B}V^{n+1})_\ell$. Applying central finite differencing to discretize $(\mathcal{L}V^{n+1})_\ell$, the fully discrete system can be written as:

$$V_\ell^{n+1}[1 + (\alpha_\ell + \beta_\ell + r + \lambda)\Delta\tau] - \Delta\tau\beta_\ell V_{\ell+1}^{n+1} - \Delta\tau\alpha_\ell V_{\ell-1}^{n+1} = (\Phi^{n+1}V_\ell^n) + \Delta\tau(\mathcal{B}V^{n+1})_\ell,$$

where α_ℓ and β_ℓ are given by

$$\alpha_\ell = \frac{(\sigma^2 + \bar{\sigma})S_\ell^2}{(S_\ell - S_{\ell-1})(S_{\ell+1} - S_{\ell-1})}, \qquad \beta_\ell = \frac{(\sigma^2 + \bar{\sigma})S_\ell^2}{(S_{\ell+1} - S_\ell)(S_{\ell+1} - S_{\ell-1})},$$

for $\ell = 2, \ldots, m-1$. When $\ell = 1$, we set $\alpha_\ell = \beta_\ell = 0$ at $S_1 = 0$; and when $\ell = m$, we set V_m^{n+1} equal to the relevant Dirichlet boundary condition.

In matrix form, the discrete system can be rewritten as:

$$[I - \Delta\tau L - \Delta\tau B]V^{n+1} = \Phi^{n+1}V^n, \tag{2.7}$$

where I is the identity matrix, L is defined such that

$$[LV^{n+1}]_\ell = -(\alpha_\ell + \beta_\ell + r + \lambda)V_\ell^{n+1} + \beta_\ell V_{\ell+1}^{n+1} + \alpha_\ell V_{\ell-1}^{n+1},$$

and B is a dense matrix whose entries satisfy the following properties [12]

$$\sum_j B_{\ell j} = \lambda, \qquad 0 \le B_{\ell j} \le 1. \tag{2.8}$$

The matrix-vector product BV^{n+1} can be computed efficiently by using an FFT.

2.2. *Fixed point iteration*

Each time step requires solving the linear system (2.7), where the coefficient matrix, denoted by A, is dense since B is dense. In [19], the authors proposed a fixed point iteration which splits A into $A = A_1 - A_2$, where

$$A_1 = I - \Delta\tau L, \qquad A_2 = \Delta\tau B.$$

The solution is computed by solving iteratively

$$A_1 V^{n+1,k+1} = A_2 V^{n+1,k} + \Phi^{n+1} V^n, \tag{2.9}$$

where $V^{n+1,k}$ is an approximate solution to V^{n+1} after k iterations. This method only involves inverting a sparse matrix A_1 and the matrix-vector product with the dense matrix B, which can be computed efficiently.

It can be proved that the fixed point iteration converges and the convergence rate, ρ, is bounded by [19]:

$$\rho \le \frac{\lambda\Delta\tau}{1 + (r + \lambda)\Delta\tau} < 1.$$

Although the fixed point iteration is convergent, the convergence rate can be very slow. Note that λ is a function of the density function $\nu(y)$. It can be estimated by

$$\lambda = O\left(\int_{\Omega_1 \cup \Omega_2} \frac{1}{|y|^{1+Y}} dy\right) = O((\Delta y)^{-Y}), \tag{2.10}$$

for $Y > 0$. As a result,

$$\frac{\lambda\Delta\tau}{1 + (r + \lambda)\Delta\tau} \approx 1 - \frac{1}{\lambda\Delta\tau} \approx 1 - O(\Delta y)^Y, \tag{2.11}$$

as Δy approaches 0.

3. Multigrid method

As shown in (2.11), the convergence of the fixed point iteration can be very slow when $Y > 0$ and Δy is small. This arises in the infinite activity case for the Lévy process. The situation gets worse when Y approaches 2 in the infinite variation case.

We propose an efficient multigrid method for solving the linear system (2.7). The idea is to apply the fixed point iteration as a smoother in the multigrid context. While the convergence rate of the fixed point iteration can be slow, especially when Y gets close to 2, it turns out that the fixed point iteration is an effective smoother which damps the high frequency.

A V-cycle multigrid method is used to solve the PIDE (2.6). We remark that other cycles such as W-cycle, F-cycle, or full multigrid can also be used. But they are generally more expensive. We find that V-cycle is the more effective option in this case.

In a two-grid setting, the algorithm consists of three major steps. We start with an initial guess for the solution, $\hat{V}_h^k \approx V^{n+1}$, on the fine grid. The first step is pre-smoothing. A small number of the fixed point iterations (typically $\nu_1 = 1$ or 2) is applied to compute an approximate solution of (2.7),

$$\tilde{V}_h = \mathcal{S}^{\nu_1}(\hat{V}_h^k, A_h, f_h),$$

where $\mathcal{S}$ denotes the fixed point iteration given by (2.9), and A_h and f_h are the fine grid matrix and right-hand side from (2.7), respectively. The role of the smoother is to damp away the high frequency components of the error, thus resulting in a smooth error. The smoothing property is analyzed in Section 3.1.

The second step is coarse grid correction. Note that the error, e_h, on the fine grid satisfies the equation,

$$A_h e_h = r_h,$$

where $r_h = f_h - A_h \tilde{V}_h$ is the residual on the fine grid. Since the error is smooth after pre-smoothing, we can solve it accurately and efficiently on the coarse grid with grid size H,

$$A_H e_H = r_H, \tag{3.1}$$

where A_H is the discretization matrix on the coarse grid and $r_H = \mathcal{R}r_h$ is the restriction of r_h given by a restriction operator $\mathcal{R}$. Here we use full weighting for $\mathcal{R}$; i.e.,

$$(\mathcal{R}r_h)_\ell = 0.25(r_h)_{2\ell-1} + 0.5(r_h)_{2\ell} + 0.25(r_h)_{2\ell+1}.$$

Note that Dirichlet conditions are used for the boundary points (see [19]). Thus the residual values are zero on the boundary points which will be used in the restriction near the boundary.

After e_H is obtained, we update the fine grid solution by

$$\check{V}_h = \tilde{V}_h + \mathcal{P}e_H,$$

where $\mathcal{P}$ is a prolongation operator which interpolates the coarse grid error e_H onto the fine grid. Linear interpolation is used for $\mathcal{P}$; i.e.,

$$(\mathcal{P}e_H)_{2\ell} = (e_H)_\ell, \quad (\mathcal{P}e_H)_{2\ell+1} = 0.5(e_H)_\ell + 0.5(e_H)_{\ell+1}.$$

The last step is to apply smoothing again,

$$\hat{V}_h^{k+1} = \mathcal{S}^{\nu_2}(\check{V}_h, A_h, f_h),$$

where ν_2 is usually 1 or 2.

In a multigrid setting, the coarse grid equation (3.1) is solved by applying the two-grid algorithm recursively.

The success of the multigrid method hinges on the effectiveness of the smoother. In the next section, we perform a smoothing analysis for the fixed point iteration.

3.1. Smoothing analysis

To simplify the analysis, we assume that an evenly spaced grid along the $\log S$ coordinate is used together with periodic boundary conditions. By a change of variable $x = \log S$, we define $\overline{V}(x, \tau) = V(e^x, \tau)$. Then (2.6) can be written as the following PIDE with constant coefficients,

$$\overline{V}_\tau = \frac{\sigma^2 + \bar{\sigma}}{2}\overline{V}_{xx} + (r - q - \kappa - \frac{\sigma^2 + \bar{\sigma}}{2})\overline{V}_x - (r + \lambda)\overline{V} + \overline{V} \otimes \hat{\gamma}, \tag{3.2}$$

where $\overline{V} \otimes \hat{\gamma}$ denotes a correlation product; i.e. $(\overline{V} \otimes \hat{\gamma})_i = \sum_{j=0}^{N-1} \overline{V}(x_i, y_j)\hat{\gamma}(y_j)$. The PIDE is discretized similarly as described in Section 2.1 where the unknowns are now $\overline{V}^{n+1}$. More precisely, we have

$$\frac{\overline{V}_j^{n+1} - \overline{V}_{j^*}^n}{\Delta \tau} = \frac{\sigma^2 + \bar{\sigma}}{2} \frac{\overline{V}_{j+1}^{n+1} - 2\overline{V}_j^{n+1} + \overline{V}_{j-1}^{n+1}}{\Delta x^2} - (r + \lambda)\overline{V}_j^{n+1} + (\overline{V}^{n+1} \otimes \hat{\gamma})_{j^*}, \tag{3.3}$$

where the index j^* is defined such that when applying the semi-Lagrangian discretization, the grid point x_j at time τ^{n+1} is traced back to x_{j^*} at time τ^n. The value of $\overline{V}_{j^*}^n \equiv \overline{V}(x_{j^*}, \tau^n)$ is computed by an upwind quadratic interpolation. After rearranging terms, (3.3) becomes

$$[1 + 2\alpha\Delta\tau + (r + \lambda)\Delta\tau]\,\overline{V}_j^{n+1} - \alpha\Delta\tau\overline{V}_{j+1}^{n+1} - \alpha\Delta\tau\overline{V}_{j-1}^{n+1}$$
$$= \overline{V}_{j^*}^n + \Delta\tau(\overline{V}^{n+1} \otimes \hat{\gamma})_{j^*}, \tag{3.4}$$

where $\alpha = (\sigma^2 + \bar{\sigma})/(2(\Delta x)^2)$. Then, the corresponding fixed point iteration (2.9) is given in component form by

$$[1 + 2\alpha\Delta\tau + (r + \lambda)\Delta\tau]\,\overline{V}_j^{n+1,k+1} - \alpha\Delta\tau\overline{V}_{j+1}^{n+1,k+1} - \alpha\Delta\tau\overline{V}_{j-1}^{n+1,k+1}$$
$$= \overline{V}_{j^*}^n + \Delta\tau(\overline{V}^{n+1,k} \otimes \hat{\gamma})_{j^*}, \tag{3.5}$$

where $\overline{V}^{n+1,k}$ is an approximate solution to $\overline{V}^{n+1}$ after k fixed point iterations.

Note that the choice of the grid points for the upwind quadratic interpolation depends on the sign of the coefficient for $\overline{V}_x$ in (3.2). Without loss of generality, assuming the coefficient is negative, we have

$$\overline{V}_{j^*}^n = \overline{V}_{j-p-2}^n \psi_{j,j-p-2} + \overline{V}_{j-p-1}^n \psi_{j,j-p-1} + \overline{V}_{j-p}^n \psi_{j,j-p}, \tag{3.6}$$

where ψ's are the Lagrangian basis functions and x_{j^*} lies between x_{i-p-1} and x_{i-p} for some integer p. Substituting (3.6) into (3.5), we obtain

$$\overline{V}_j^{n+1,k+1} [1 + (2\alpha + r + \lambda)\Delta\tau] - \alpha\Delta\tau\overline{V}_{j+1}^{n+1,k+1} - \alpha\Delta\tau\overline{V}_{j-1}^{n+1,k+1}$$
$$= \overline{V}_{j-p-2}^n \psi_{j,j-p-2} + \overline{V}_{j-p-1}^n \psi_{j,j-p-1} + \overline{V}_{j-p}^n \psi_{j,j-p}$$
$$+ (\overline{V}^{n+1,k} \otimes \hat{\gamma})_{j-p-2}\Delta\tau\psi_{j,j-p-2} + (\overline{V}^{n+1,k} \otimes \hat{\gamma})_{j-p-1}\Delta\tau\psi_{j,j-p-1}$$
$$+ (\overline{V}^{n+1,k} \otimes \hat{\gamma})_{j-p}\Delta\tau\psi_{j,j-p}. \tag{3.7}$$

The fixed point iteration is used as smoother in our multigrid method. We will estimate the smoothing factor of the fixed point iteration by applying the local Fourier analysis [20] to (3.7). Let $E^k = \overline{V}^{n+1,k} - \overline{V}^{n+1}$ be the error of $\overline{V}^{n+1,k}$. Let C_μ^k and G_μ be the symbols of E^k and $\hat{\gamma}$, respectively, where μ is the Fourier mode; i.e. C_μ^k and G_μ are the Fourier transform of E^k and $\hat{\gamma}$, respectively [20]. Also, let $h = 1/N$ be the mesh size and $i = \sqrt{-1}$. Then (3.7) becomes

$$C_\mu^{k+1} \left[1 + (2\alpha + r + \lambda)\Delta\tau - \alpha\Delta\tau e^{\mu\pi h i} - \alpha\Delta\tau e^{-\mu\pi h i} \right]$$
$$= C_\mu^k \left[G_{-\mu} e^{-(p+2)\mu\pi h i}\Delta\tau\psi_{j,j-p-2} + G_{-\mu} e^{-(p+1)\mu\pi h i}\Delta\tau\psi_{j,j-p-1} \right.$$
$$\left. + G_{-\mu} e^{-p\mu\pi h i}\Delta\tau\psi_{j,j-p} \right]. \tag{3.8}$$

Note the negative subscript of $G_{-\mu}$ is a result of the correlation operator. The left-hand side of (3.8) can be simplified as

$$
\begin{aligned}
LHS &= C_\mu^{k+1} \left[1 + (2\alpha + r + \lambda)\Delta\tau - 2\alpha\Delta\tau \cos(\mu\pi h) \right] \\
&= C_\mu^{k+1} \left[1 + (r + \lambda)\Delta\tau + 2\alpha\Delta\tau(1 - \cos(\mu\pi h)) \right] \\
&= C_\mu^{k+1} \left[1 + (r + \lambda + 4\alpha \sin^2(\mu\pi h/2))\Delta\tau \right].
\end{aligned}
$$

Similarly, the right-hand side of (3.8) can be written as

$$
\begin{aligned}
RHS &= C_\mu^k \Big[(\Delta\tau G_{-\mu})(e^{-(p+2)\mu\pi hi}\psi_{j,j-p-2} + e^{-(p+1)\mu\pi hi}\psi_{j,j-p-1} \\
&\qquad + e^{-p\mu\pi hi}\psi_{j,j-p}) \Big] \\
&= C_\mu^k \Delta\tau G_{-\mu}\eta_j, \tag{3.9}
\end{aligned}
$$

where $\eta_j = \sum_{m=j-p-2}^{j-p} \psi_{j,m} e^{-(j-m)\mu\pi hi}$.

Let $\hat{S}_\mu$ be the symbol of the smoother. Then one can easily show that $\hat{S}_\mu = C_\mu^{k+1}/C_\mu^k$; i.e.,

$$
\hat{S}_\mu = \frac{C_\mu^{k+1}}{C_\mu^k} = \frac{\Delta\tau G_{-\mu}\eta_j}{1 + (r + \lambda)\Delta\tau + 4\alpha\Delta\tau \sin^2(\mu\pi h/2)}. \tag{3.10}
$$

It has been proved that $|\eta_j| \leq 1$ [40, Section 5.2]. Furthermore, since G_μ is the symbol of $\hat{\gamma}$, by (2.5), we have $|G_{-\mu}| \leq \lambda$. Hence, we can bound $\hat{S}_\mu$ as

$$
\begin{aligned}
|\hat{S}_\mu| &= \frac{|\Delta\tau G_{-\mu}||\eta_j|}{|1 + (r + \lambda + 4\alpha \sin^2(\mu\pi h/2))\Delta\tau|} \\
&\leq \frac{\lambda\Delta\tau}{1 + (r + \lambda + 4\alpha \sin^2(\mu\pi h/2))\Delta\tau}. \tag{3.11}
\end{aligned}
$$

Let ζ be the smoothing factor of the fixed point iteration, which is defined as

$$
\zeta \equiv \max\{|\hat{S}_\mu| : \ N/2 \leq \mu \leq N\}.
$$

Considering the high frequency Fourier modes; i.e., $N/2 \leq \mu \leq N$, it follows from (3.11) that

$$
\zeta \leq \frac{\lambda\Delta\tau}{1 + (r + \lambda + 2\alpha)\Delta\tau}.
$$

To bring more insight into the formula, we estimate the upper bound for ζ by noting that

$$
\lambda = O(\Delta x)^{-Y} \ \text{ and } \ \alpha = O(\Delta x)^{-2}.
$$

Considering the limit where $\Delta x \approx 0$ and keeping $\Delta\tau$ fixed, we have

$$
\zeta \approx \frac{O(\Delta x)^{-Y}\Delta\tau}{1 + O(\Delta x)^{-Y}\Delta\tau + O(\Delta x)^{-2}\Delta\tau} \approx O(\Delta x)^{2-Y}.
$$

In the Lévy process under the CGMY model, we always have $Y < 2$. Hence for small Δx, the smoothing factor for the fixed point iteration is small, which is fundamental to achieve an effective multigrid method.

3.2. Two-grid convergence analysis

In the previous section, we have shown by a local Fourier analysis that the fixed point iteration smoother is effective damping high frequency errors. In this section, we will perform a two-grid convergence analysis based on the local Fourier analysis. For easy exposition, we assume there is no post-smoothing.

The iteration matrix of the two-grid iteration can be written as

$$\mathcal{M} = (I - P(L^H)^{-1}RL^h)S,$$

where L^h and L^H are the fine and coarse grid operator, respectively, P is the prolongation, R is the restriction, and S is the smoother. For two-grid analysis [21], it is customary to pair up the low and high frequency modes such that the Fourier transformed matrix, $\hat{\mathcal{M}}$, is a block diagonal matrix with each block a 2×2 matrix, $\hat{\mathcal{M}}_\mu$, which is given by

$$\hat{\mathcal{M}}_\mu = (I - \hat{P}_\mu(\hat{L}_\mu^H)^{-1}\hat{R}_\mu\hat{L}_\mu^h)\hat{S}_\mu \qquad 1 \le \mu \le N/2, \tag{3.12}$$

where

$$\hat{P}_\mu = \sqrt{2}\begin{bmatrix} c_\mu^2 \\ -s_\mu^2 \end{bmatrix}, \qquad \hat{R}_\mu = \frac{1}{\sqrt{2}}\begin{bmatrix} c_\mu^2 & -s_\mu^2 \end{bmatrix}. \tag{3.13}$$

Here $c_\mu = \cos(\mu\pi h)/2$, $s_\mu = \sin(\mu\pi h)/2$, and $h = \Delta x$.

By (3.10), we have

$$
\begin{aligned}
\hat{S}_\mu &= \begin{bmatrix} \hat{S}_\mu & 0 \\ 0 & \hat{S}_{\mu-N} \end{bmatrix} \\
&= \begin{bmatrix} \frac{h^2\Delta\tau G_{-\mu}\eta_j}{h^2+\lambda h^2\Delta\tau+4\tilde{\sigma}\Delta\tau s_\mu^2} & 0 \\ 0 & \frac{h^2\Delta\tau G_\mu\eta_j}{h^2+\lambda h^2\Delta\tau+4\tilde{\sigma}\Delta\tau c_\mu^2} \end{bmatrix},
\end{aligned} \tag{3.14}
$$

where $\tilde{\sigma} = (\sigma^2 + \bar{\sigma})/2$. Also, we have assumed the interest rate r to be zero, just to simplify the algebraic expression.

A similar calculation yields

$$\hat{L}_\mu^h = \begin{bmatrix} 1 + (\lambda + \frac{4\tilde{\sigma}s_\mu^2}{h^2} - G_{-\mu}\eta_j)\Delta\tau & 0 \\ 0 & 1 + (\lambda + \frac{4\tilde{\sigma}c_\mu^2}{h^2} - G_\mu\eta_j)\Delta\tau \end{bmatrix}, \tag{3.15}$$

$$\hat{L}_\mu^H = 1 + (\lambda + \frac{4\tilde{\sigma}s_\mu^2 c_\mu^2}{h^2} - G_\mu\eta_{j'})\Delta\tau, \tag{3.16}$$

where $\eta_{j'}$ is defined similarly as η_j on the coarse grid. Details can be found in Appendix A.

In order to prove convergence, we assume the following inequalities hold:

$$\frac{h^2}{\Delta\tau} < \frac{2\tilde{\sigma}^2}{3\lambda h^2 + 6\tilde{\sigma}}, \tag{3.17}$$

$$\lambda h^2 < \frac{2\sqrt{3}-3}{3}\tilde{\sigma}. \tag{3.18}$$

Note that $\tilde{\sigma} = O(1)$ and hence the right-hand side of assumption (3.17) is also $O(1)$. In practice, $\Delta\tau$ is typically chosen as $O(h)$. Thus assumption (3.17) can be easily satisfied. By (2.10), $\lambda h^2 = O(h^{2-Y})$. Since $Y < 2$, for sufficiently small h, λh^2 will satisfy the bound in assumption (3.18).

Lemma 3.1. *Let* $\hat{\mathcal{M}}_\mu$ *be defined as (3.12). Assume the inequalities (3.17) and (3.18) hold. Then*

$$\|\hat{\mathcal{M}}_\mu\|_\infty < 1$$

for any μ.

Proof. The proof of the lemma is very technical. Here we will outline the main steps of the proof. We refer the interested reader to Appendix A for details.

The idea is to derive a bound for $|\hat{\mathcal{M}}_\mu(1, 1)| + |\hat{\mathcal{M}}_\mu(1, 2)|$ and $|\hat{\mathcal{M}}_\mu(2, 1)| + |\hat{\mathcal{M}}_\mu(2, 2)|$, where $\hat{\mathcal{M}}_\mu(i, j)$ is the (i, j) entry of the 2×2 matrix $\hat{\mathcal{M}}_\mu$. Combining (3.13), (3.14), (3.15), and (3.16), we have

$$\hat{\mathcal{M}}_\mu(1, 1) = \left[1 - c_\mu^4 \frac{1 + (\lambda + 4\tilde{\sigma} s_\mu^2/h^2 - G_{-\mu}\eta_j)\Delta\tau}{1 + (\lambda + 4\tilde{\sigma} s_\mu^2 c_\mu^2/h^2 - G_\mu\eta_{j'})\Delta\tau}\right]\left[\frac{\Delta\tau G_{-\mu}\eta_j}{1 + (\lambda + 4\tilde{\sigma} s_\mu^2/h^2)\Delta\tau}\right]$$

$$\hat{\mathcal{M}}_\mu(1, 2) = \left[s_\mu^2 c_\mu^2 \frac{1 + (\lambda + 4\tilde{\sigma} c_\mu^2/h^2 - G_{-\mu}\eta_j)\Delta\tau}{1 + (\lambda + 4\tilde{\sigma} s_\mu^2 c_\mu^2/h^2 - G_\mu\eta_{j'})\Delta\tau}\right]\left[\frac{\Delta\tau G_{-\mu}\eta_j}{1 + (\lambda + 4\tilde{\sigma} c_\mu^2/h^2)\Delta\tau}\right].$$

By direct computation and that facts that $|G_\mu| = |G_{-\mu}| \le \lambda$ and $|\eta_j| \le 1$, one can show that

$$|\hat{\mathcal{M}}_\mu(1, 1)| < \left[\frac{h^2}{4\tilde{\sigma}\Delta\tau}\left(\frac{1 + c_\mu^2}{c_\mu^2}\right) + \frac{\lambda h^2}{2\tilde{\sigma}}\left(\frac{1 + c_\mu^2}{c_\mu^2}\right) + s_\mu^2\right]\left[\frac{\lambda h^2}{\lambda h^2 + 4\tilde{\sigma} s_\mu^2}\right] \qquad (3.19)$$

$$|\hat{\mathcal{M}}_\mu(1, 2)| < \left[\frac{h^2}{4\tilde{\sigma}\Delta\tau} + \frac{\lambda h^2}{2\tilde{\sigma}} + c_\mu^2\right]\left[\frac{\lambda h^2}{\lambda h^2 + 4\tilde{\sigma} c_\mu^2}\right]. \qquad (3.20)$$

By assumptions (3.17) and (3.18), one can further show that

$$|\hat{\mathcal{M}}_\mu(1, 1)| + |\hat{\mathcal{M}}_\mu(1, 2)| < \frac{1 + 1/3\xi}{1 + 1/2\xi} < 1,$$

where $\xi = 4\tilde{\sigma}/\lambda h^2$. The case for $|\hat{\mathcal{M}}_\mu(2, 1)| + |\hat{\mathcal{M}}_\mu(2, 2)| < 1$ is similar and hence is omitted here. $\quad\square$

Theorem 3.2. *Assume the inequalities (3.17) and (3.18) hold. Let* $\rho(\mathcal{M})$ *be the spectral radius of the iteration matrix* $\mathcal{M}$. *Then*

$$\rho(\mathcal{M}) < 0.75.$$

Hence, the two-grid method converges and the upper bound is independent of the grid size and the value Y.

Proof. By lemma 3.1, $\|\hat{\mathcal{M}}_\mu\|_\infty < 1$ for any μ. Furthermore, the proof of the lemma shows that

$$|\hat{\mathcal{M}}_\mu(1, 1)| + |\hat{\mathcal{M}}_\mu(1, 2)| < \frac{1 + 1/3\xi}{1 + 1/2\xi}$$

$$= \frac{3\lambda h^2 + 4\tilde{\sigma}}{3\lambda h^2 + 6\tilde{\sigma}}$$

$$= \frac{3\frac{\lambda h^2}{\tilde{\sigma}} + 4}{3\frac{\lambda h^2}{\tilde{\sigma}} + 6}$$

$$< \frac{3\frac{\lambda h^2}{\tilde{\sigma}} + 4}{6}$$

$$< \frac{(2\sqrt{3} - 3) + 4}{6} < 0.75.$$

The last inequality comes from assumption (3.18). The same bound for $|\hat{\mathcal{M}}_\mu(2, 1)| + |\hat{\mathcal{M}}_\mu(2, 2)|$ can be derived similarly. Thus

$$\|\hat{\mathcal{M}}\|_\infty = \max_{1 \le \mu \le N/2} \|\hat{\mathcal{M}}_\mu\|_\infty < 0.75.$$

As a result, $\rho(\mathcal{M}) = \rho(\hat{\mathcal{M}}) \le \|\hat{\mathcal{M}}\|_\infty < 0.75.$ □

4. Numerical results

In this section, we present results of numerical experiments to demonstrate the effectiveness of the proposed multigrid method for pricing European options under the CGMY model. The PIDE is discretized by the semi-Lagrangian method in the spatial domain and implicit Euler's method for time. The discrete linear system of equations is solved by the proposed multigrid method. We also include numerical results using a fixed point iteration for comparison.

In the multigrid procedure, a V-cycle is used with one pre- and one post- smoothing where the smoother is a fixed point iteration. The coarse grid operators are obtained from direct discretization with linear interpolation and full-weighting restriction for intergrid transfer. The stopping criterion is when the relative residual l_2-norm was less than 10^{-8}.

As explained in Section 2, the fixed point iteration (cf. Section 2.1) used in the literature converges slowly when the grid size is small and becomes worse for $Y \approx 2$. We will show the multigrid convergence results where Y changes from 0 to 1.98.

Example 1: In this example, $Y = 0$. In this case, the CGMY process is also known as variance gamma. The input parameters for the PIDE and CGMY model are given in Table 1.

Table 1. The input parameters are obtained from [41] for pricing a European call option under a variance gamma process.

S	K	T	r	q	σ	C	G	M	Y
90	98	0.5	0.0	0.0	0.0	5.9311	20.2648	39.784	0.0

Table 2 presents the convergence of the proposed multigrid method (MG) for different grid sizes. In this example, the total number of time steps is 500, and so $\Delta\tau = 10^{-3}$, which is typical for pricing option values with reasonable accuracy. Multigrid converges very fast in only 2 iterations. The number of fixed point iterations (FP) is relatively constant which shows that the convergence is independent of the grid size under a Variance Gamma process, as explained in (2.11).

Table 2. Number of multigrid V-cycles for pricing a European call option under CGMY when $Y = 0$ with different grid sizes and $\Delta\tau = 10^{-3}$.

			$Y = 0, \Delta\tau = 10^{-3}$		
h	1/64	1/128	1/256	1/512	1/1024
FP	5	6	6	7	8
MG	2	2	2	2	2

Example 2: In this example, $Y = 0.6442$. When $0 < Y < 1$, the Lévy process has infinite activity but finite variation. The input parameters are given in Table 3. The convergence results are shown in Table 4. In this case, the number of fixed point iteration increases for smaller grid size, supported by the formula in (2.11) with $Y > 0$. The multigrid convergence remains constant, independent of the grid size.

Table 3. The input parameters are obtained from [2] for pricing a European call option when $Y = 0.6442$.

S	K	T	r	q	σ	C	G	M	Y
90	98	0.25	0.06	0.0	0.0	16.97	7.08	29.97	0.6442

Table 4. Number of multigrid V-cycles for pricing a European call option under CGMY when $Y = 0.6442$ for different grid sizes with (top) $\Delta\tau = 10^{-3}$ and (bottom) $\Delta\tau = 10^{-2}$.

			$Y = 0.6442, \Delta\tau = 10^{-3}$		
h	1/64	1/128	1/256	1/512	1/1024
FP	22	35	56	90	143
MG	4	5	6	6	6
			$Y = 0.6442, \Delta\tau = 10^{-2}$		
h	1/64	1/128	1/256	1/512	1/1024
FP	107	207	372	643	1075
MG	7	7	8	9	10

In this example, we also use a larger $\Delta\tau$ to test its effect on the convergence. In practice, a large time step size may be necessary when pricing options with long expiry dates; e.g., $T = 10$. However, a larger time step size will generally result in a more ill-condition discrete linear system. Table 4 shows that the fixed point iteration numbers increase dramatically, by a factor between 5 to 7. The number of multigrid iteration only increases mildly.

Example 3: When $1 < Y < 2$, this corresponds to a Lévy process with infinite activity and infinite variation. Table 6 shows the convergence results for pricing a call option for $Y = 1.4$ and $Y = 1.98$, with input model parameters given in Table 5.

Table 5. The input parameters are obtained from [16] when $Y = 1.4$ and from [42] when $Y = 1.98$ for pricing a European call option.

S	K	T	r	q	σ	C	G	M	Y
500	500	0.25	0.4	0.0	0.2	1.0	1.4	2.5	1.4
100	100	0.25	0.1	0.0	0.0	1.0	5.0	5.0	1.98

Table 6. Number of multigrid V-cycles for pricing a European call option under CGMY for (top) $Y = 1.4$ and (bottom) $Y = 1.98$ with different grid sizes and $\Delta\tau = 10^{-3}$.

|					
		$Y = 1.4$, $\Delta\tau = 10^{-3}$			
h	1/64	1/128	1/256	1/512	1/1024
FP	82	204	527	1394	3691
MG	3	4	4	4	4
		$Y = 1.98$, $\Delta\tau = 10^{-3}$			
h	1/64	1/128	1/256	1/512	1/1024
FP	>5000	>5000	>5000	>5000	>5000
MG	5	6	7	8	9

When $Y = 1.4$, the discrete problem becomes much more difficult to solve, as can be seen from the number of fixed point iteration which increases from hundreds to thousands as the grid size decreases. The multigrid iteration remains very efficient with mesh independent convergence rate. The case where $Y = 1.98$ is numerically challenging since the kernel function of the Lévy process becomes very singular. The fixed point iteration converges very slowly. The multigrid iteration deteriorates slightly from 4 to 9, but still converges very quickly.

Example 4: In this example, we illustrate the convergence behaviour of multigrid, fixed point iteration and BiCGstab with the fixed point iteration as preconditioner [19]. The results are shown in Figure 1. Here we consider the case where $Y = 1.4$. The convergence behaviour is similar for other values of Y and they are not shown here. We can see that multigrid converges quickly with a constant rate and the convergence is essentially unchanged for different grid sizes. Fixed point iteration converges at a much slower rate and the rate of convergence deteriorates as the grid size h decreases. The convergence of preconditioned BiCGstab is somewhere between the two. It is much better than fixed point iteration but not as good as multigrid. Also, the convergence is not as smooth and uniform as the other two, and it is mesh dependent.

Table 7 shows the rates of convergence for multigrid and fixed point iteration, which are computed by taking the average rates of the last 3 iterations. The convergence of BiCGstab is not very uniform and we do not include its rate of convergence here.

In Sections 3.1 and 3.2, we presented a local Fourier analysis to analyze the smoother and the two-grid convergence. It would be interesting to show how the theoretical analysis results compare with the actual convergence. However, both the smoothing factor and two-grid convergence factor include a quantity η_j (3.9) which comes from the interpolation operator during the semi-Lagrangian discretization. This quantity is difficult to compute and it varies at different grid points, grid sizes and

time steps. As such, we were unable to compute the smoothing factor and the spectral radius of $\hat{\mathcal{M}}$. Instead, we compute an upper bound for the smoothing factor and the infinity norm of $\hat{\mathcal{M}}$ as described in (3.11), (3.19) and (3.20). They are shown in columns 3 and 4 in Table 7. Since they are upper bounds, they do not match very well with the actual convergence of multigrid. Nevertheless, we can see that the smoothing factor decreases with h as discussed in Section 3.1 and the two-grid factor is bounded by 0.75 as shown in Theorem 3.2. We also remark that the discrepancy may be due to the fact that the underlying PIDE is a nonlocal operator where the local mode analysis may not be the right tool for analysis.
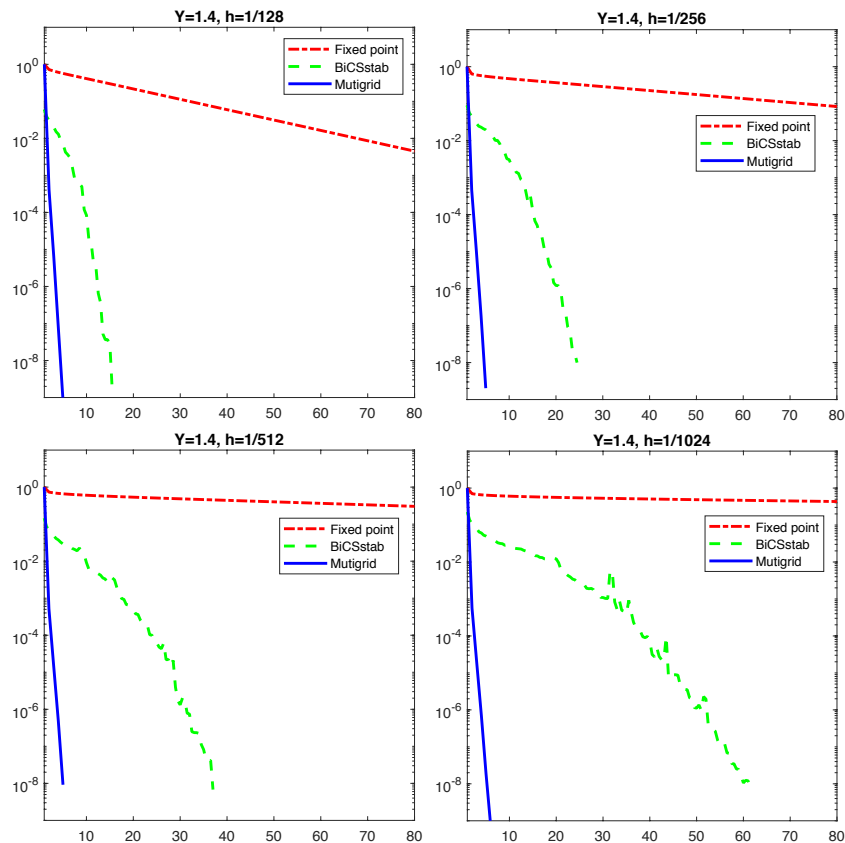
Figure 1. Convergence of multigrid, fixed point iteration and preconditioned BiCGstab for $Y = 1.4$ with grid size h=1/128, 1/256, 1/512, and 1/1024.

Table 7. Rates of convergence for fixed point iteration and multigrid with different grid sizes, $Y = 1.4$. The last two columns show an upper bound for the smoothing factor and the two-grid convergence factor.

h	FP	MG	Smoothing	Two-grid
1/128	0.938	0.0108	0.514	0.415
1/256	0.975	0.0181	0.350	0.241
1/512	0.991	0.0314	0.234	0.165
1/1024	0.996	0.0262	0.155	0.134

5. Conclusion

In this paper, we have proposed a fast multigrid method for solving the discrete equations from a finite difference discretization of the partial integro-differential equations arising from pricing European options under the CGMY process. The fixed point iteration originally proposed in [19] converges well when the density kernel is not singular ($Y \approx 0$) but very slowly when the kernel becomes more singular ($Y \approx 2$). While the fixed point iteration is not an efficient solver by itself, we have proved that it is effective removing high frequency error components. By using the fixed point iteration as a smoother for multigrid, we have obtained a fast multigrid solver which converges within 10 iterations for different values of Y ranging from 0 to 1.98. In addition, we have provided a two-grid convergence analysis for the multigrid method.

The current paper focuses primarily on European options. For American options under the CGMY process, the corresponding PIDE is nonlinear with a linear complimentarity condition. The proposed multigrid method would not be directly applicable since it is designed for solving linear systems. An interesting idea is to combine the techniques in this paper with a nonlinear multigrid method such as Full Approximation Scheme (FAS), which will be a future work. Another future work is to perform better local mode analysis of the method, such as three-grid analysis, in order to obtain improved quantitative results.

Acknowledgments

Conflict of interest

The author declares no conflict of interest.

References

1. R. Cont, P. Tankov, *Financial Modelling With Jump Processes*, Chapman & Hall/CRC Press, 2003.

2. P. Carr, H. Geman, D. B. Madan, et al. *The fine structure of asset returns: An empirical investigation*, J. Bus., **75** (2002), 305–332.

3. P. Carr, H. Geman, D. B. Madan, et al. *Stochastic volatility for Lévy processes*, Math. Financ., **13** (2003), 345–382.

4. E. Eberlein, *Application of generalized hyperbolic Lévy motion to finance*, In: O. Barndorff-Nielsen, S. Resnick, T. Mikosch, editors, *Lévy Processes – Theory and Application*, Birkhäuser, Boston, 2001, 319–337.

5. D. B. Madan, P. Carr, E. Change, *The Variance Gamma process and option pricing*, Rev. Financ., **2** (1998), 79–105.

6. D. Madan, *Financial modeling with discontinuous price processes*, In: O. Barndorff-Nielsen, S. Resnick, T. Mikosch, editors, *Lévy Processes – Theory and Application*, Birkhäuser, Boston, 2001.

7. W. Schoutens, *Lévy Processes in Finance: Pricing Financial Derivatives*, Wiley, New York, 2003.

8. D. B. Madan, E. Seneta, *The Variance Gamma (VG) model for share market returns*, J. Bus., **63** (1990), 511–524.

9. A. Almendral, C. W. Oosterlee, *On American options under the Variance Gamma process*, Appl. Math. Financ., **14** (2007), 131–152.

10. Y. D'Halluin, P. Forsyth, G. Labahn, *A penalty method for American options with jump-diffusion processes*, Numer. Math., **97** (2004), 321–352.

11. A. Hirsa, D. B. Madan, *Pricing American options under Variance Gamma*, J. Comput. Financ., **7** (2004), 63–80.

12. Y. d'Halluin, P. A. Forsyth, K. Vetzal, *Robust numerical methods for contingent claims under jump diffusion processes*, IMA J. Numer. Anal., **25** (2005), 87–112.

13. A. Almendral, *Numerical valuation of American options under the CGMY process*, In: A. Kyprianou, W. Schoutens, P. Wilmott, editors, *Exotic Option Pricing and Advanced Lévy Models*, Wiley, UK, 2005.

14. R. Cont, E. Voltchkova, *A finite difference scheme for option pricing in jump diffusion and exponential Lévy models*, SIAM J. Numer. Anal., **43** (2005), 1596–1626.

15. H. Geman, *Pure jump Lévy processes for asset price modeling*, J. Bank. Financ., **26** (2002), 1297–1316.

16. A. M. Matache, P. A. Nitsche, C. Schwab, *Wavelet Galerkin pricing of American options on Lévy driven assets*, Quant. Financ., **5** (2005), 403–424.

17. S. Levendorskiǐ, O. Kudryavtsev, V. Zherder, *The relative efficiency of numerical methods for pricing American options under Lévy processes*, J. Comput. Financ., **9** (2005), 69–97.

18. A. Almendral, C. W. Oosterlee, *Numerical valuation of options with jumps in the underlying*, Appl. Numer. Math., **53** (2005), 1–18.

19. I. R. Wang, J. W. L. Wan, P. A. Forsyth, *Robust numerical valuation of European and American options under the CGMY process*, J. Comput. Financ., **10** (2007), 31–69.

20. A. Brandt, *Multi-level adaptive solutions to boundary-value problems*, Math. Comput., **31** (1977), 333–390.

21. W. Hackbusch, *Multi-Grid Methods and Applications*, Springer-Verlag, Berlin, 1985.

22. T. F. Chan, W. L. Wan, *Robust multigrid methods for elliptic linear systems*, J. Comput. Appl. Math., **123** (2000), 323–352.

23. W. P. Tang, W. L. Wan, *Sparse approximate inverse smoother for multi-grid*, SIAM J. Matrix Anal. Appl., **21** (2000), 1236–1252.

24. G. Wittum, *On the robustness of ilu smoothing*, SIAM J. Sci. Stat. Comput., **10** (1989), 699–717.

25. R. E. Alcouffe, A. Brandt, Jr. J. E. Dendy, et al. *The multi-grid method for the diffusion equation with strongly discontinuous coefficients*, SIAM J. Sci. Stat. Comput., **2** (1981), 430–454.

26. Jr. J. E. Dendy, *Black box multigrid*, J. Comput. Phys., **48** (1982), 366–386.

27. A. Reusken, *Multigrid with matrix-dependent transfer operators for a singular perturbation problem*, Computing, **50** (1993), 199–211.

28. W. L. Wan, T. F. Chan, B. Smith, *An energy-minimizing interpolation for robust multigrid*, SIAM J. Sci. Comput., **21** (2000), 1632–1649.

29. R. E. Bank, J. Wan, Z. Qu, *Kernel preserving multigrid methods for convection-diffusion equations*, SIAM J. Matrix Anal. Appl., **27** (2006), 1150–1171.

30. J. Bey, G. Wittum, *Downwind numbering: Robust multigrid for convection-diffusion problems*, Appl. Numer. Math., **23** (1997), 177–192.

31. C. W. Oosterlee, F. J. Gaspar, T. Washio, et al. *Multigrid line smoothers for high order upwind discretization of convection-dominated problems*, J. Comput. Phys., **139** (1998), 274–307.

32. I. Yavneh, *Coarse-grid correction for nonelliptic and singular perturbation problems*, SIAM J. Sci. Comput., **19** (1998), 1682–1699.

33. A. Jameson, *Solution of the Euler equations for two dimensional transonic flow by a multigrid method*, Appl. Math. Comput., **13** (1983), 327–355.

34. P. W. Hemker, G. M. Johnson, *Multigrid approaches to the Euler equations*, In: S. McCormick, editor, *Multigrid Methods*, SIAM, 1987, 57–72.

35. J. W. L. Wan, A. Jameson, *Monotonicity preserving multigrid time stepping schemes for conservation laws*, Comput. Visual. Sci., **11** (2007), 41–58.

36. C. W. Oosterlee, *On multigrid for linear complementarity problems with application to American-style options*, ETNA, **15** (2003), 165–185.

37. N. Clarke, K. Parrott, *Multigrid for American option pricing with stochastic volatility*, Appl. Math. Financ., **6** (1999), 177–195.

38. C. Reisinger, G. Wittum, *On multigrid for anisotropic equations and variational inequalities "pricing multi-dimensional European and American options"*, Comput. Visual. Sci., **7** (2004), 189–197.

39. H. B. Zubair, C. W. Oosterlee, R. Wienands, *Multigrid for high-dimensional elliptic partial differential equations on non-equidistant grids*, SIAM J. Sci. Comput., **29** (2007), 1613–1636.

40. M. Falcone, R. Ferretti, *Convergence analysis for a class of high-order semi-Lagrangian avection schemes*, SIAM J. Numer. Anal., **35** (1998), 909–940.

41. R. W. Lee, *Option pricing by transform methods: Extensions, unification, and error control*, J. Comput. Financ., **7** (2004), 51–86.

42. F. Fang, C. W. Oosterlee, *A novel pricing method for European options based on Fourier-cosine series expansions*, SIAM J. Sci. Comput., **31** (2008), 826–848.

Appendix

A. Two-grid convergence analysis

Here we provide the technical details for the analysis in Section 3.2.

Lemma A.1. *The Fourier symbols for the fine and coarse grid operators are given by*

$$
\hat{L}_\mu^h = \left[\begin{array}{cc} 1 + (\lambda + \frac{4\tilde{\sigma} s_\mu^2}{h^2} - G_{-\mu}\eta_j)\Delta\tau & 0 \\ 0 & 1 + (\lambda + \frac{4\tilde{\sigma} c_\mu^2}{h^2} - G_\mu\eta_j)\Delta\tau \end{array} \right],
$$

$$
\hat{L}_\mu^H = 1 + (\lambda + \frac{4\tilde{\sigma} s_\mu^2 c_\mu^2}{h^2} - G_\mu\eta_{j'})\Delta\tau.
$$

Proof. By (3.4), the fine grid operator can be written as

$$
(L^h\overline{V}^{n+1})_j = [1 + (2\alpha + \lambda)\Delta\tau]\,\overline{V}_j^{n+1} - \alpha\Delta\tau\overline{V}_{j+1}^{n+1} - \alpha\Delta\tau\overline{V}_{j-1}^{n+1} - \Delta\tau(\overline{V}^{n+1} \otimes \hat{\gamma})_{j^*},
$$

with the assumption that $r = 0$. Applying the local Fourier analysis similar to (3.8), we have

$$
\begin{aligned}
\hat{L}_\mu^h &= 1 + (2\alpha + \lambda)\Delta\tau - \alpha\Delta\tau e^{\mu\pi hi} - \alpha\Delta\tau e^{-\mu\pi hi} - \Delta\tau G_{-\mu}\eta_j \\
&= 1 + (\lambda + 2\alpha - 2\alpha\cos(\mu\pi h) + G_{-\mu}\eta_j)\Delta\tau \\
&= 1 + (\lambda + \frac{4\tilde{\sigma} s_\mu^2}{h^2} - G_{-\mu}\eta_j)\Delta\tau,
\end{aligned}
$$

noting that $\alpha = \tilde{\sigma})/h^2$. The formula for $\hat{L}_{N-\mu}^h$ can be easily derived from $\hat{L}_\mu^h$ which is omitted here.

Similarly, the coarse grid operator can be written as

$$
\begin{aligned}
(L^H\overline{V}^{n+1})_j &= \left[1 + (2\frac{\tilde{\sigma}}{(2h)^2} + \lambda)\Delta\tau\right]\overline{V}_j^{n+1} - \frac{\tilde{\sigma}}{(2h)^2}\Delta\tau\overline{V}_{j+1}^{n+1} - \frac{\tilde{\sigma}}{(2h)^2}\Delta\tau\overline{V}_{j-1}^{n+1} \\
&\quad -\Delta\tau(\overline{V}^{n+1} \otimes \hat{\gamma})_{j'^*},
\end{aligned}
$$

where j'^* is defined similarly as j^* on the coarse grid. Applying the local Fourier analysis, we obtain

$$
\begin{aligned}
\hat{L}_\mu^H &= 1 + (\frac{\tilde{\sigma}}{2h^2} + \lambda)\Delta\tau - \frac{\tilde{\sigma}}{4h^2}\Delta\tau e^{2\mu\pi hi} - \frac{\tilde{\sigma}}{4h^2}\Delta\tau e^{-2\mu\pi hi} - \Delta\tau G_{-\mu}\eta_{j'} \\
&= 1 + (\lambda + \frac{\tilde{\sigma}}{2h^2} - \frac{\tilde{\sigma}}{2h^2}\cos(2\mu\pi h) + G_{-\mu}\eta_{j'})\Delta\tau \\
&= 1 + (\lambda + \frac{4\tilde{\sigma} s_\mu^2 c_\mu^2}{h^2} - G_{-\mu}\eta_{j'})\Delta\tau.
\end{aligned}
$$

$\square$

Lemma A.2. *Let $\hat{\mathcal{M}}_\mu$ be defined as (3.12). Assume $r = 0$ and the inequalities (3.17) and (3.18) hold. Then*

$$
\|\hat{\mathcal{M}}_\mu\|_\infty < 1
$$

for any μ.

Proof. By (3.12), (3.13), (3.14), (3.15), and (3.16), we have

$$
\begin{aligned}
\hat{\mathcal{M}}_\mu &= (I - \hat{P}_\mu (\hat{L}_\mu^H)^{-1} \hat{R}_\mu \hat{L}_\mu^h) \hat{S}_\mu \qquad 1 \le \mu \le N/2 \\
&= \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{1}{\hat{L}_\mu^H} \begin{bmatrix} c_\mu^2 \\ -s_\mu^2 \end{bmatrix} \begin{bmatrix} c_\mu^2 & -s_\mu^2 \end{bmatrix} \begin{bmatrix} \hat{L}_\mu^h(1,1) & 0 \\ 0 & \hat{L}_\mu^h(2,2) \end{bmatrix} \right) \begin{bmatrix} \hat{S}_\mu & 0 \\ 0 & \hat{S}_{\mu-N} \end{bmatrix} \\
&= \begin{bmatrix} 1 - \frac{1}{\hat{L}_\mu^H} c_\mu^4 \hat{L}_\mu^h(1,1) & \frac{1}{\hat{L}_\mu^H} c_\mu^2 s_\mu^2 \hat{L}_\mu^h(2,2) \\ \frac{1}{\hat{L}_\mu^H} c_\mu^2 s_\mu^2 \hat{L}_\mu^h(1,1) & 1 - \frac{1}{\hat{L}_\mu^H} s_\mu^4 \hat{L}_\mu^h(2,2) \end{bmatrix} \begin{bmatrix} \hat{S}_\mu & 0 \\ 0 & \hat{S}_{\mu-N} \end{bmatrix}.
\end{aligned}
\tag{A.1}
$$

Thus, the (1,1) and (1,2) entries are given by

$$
\begin{aligned}
\hat{\mathcal{M}}_\mu(1,1) &= (1 - \frac{1}{\hat{L}_\mu^H} c_\mu^4 \hat{L}_\mu^h(1,1)) \hat{S}_\mu, \\
\hat{\mathcal{M}}_\mu(1,2) &= \frac{1}{\hat{L}_\mu^H} c_\mu^2 s_\mu^2 \hat{L}_\mu^h(2,2) \hat{S}_{\mu-N}.
\end{aligned}
\tag{A.2}
$$

We first consider $\hat{\mathcal{M}}_\mu(1,1)$. Using the formulas (3.15), (3.16), and (3.14) for $\hat{L}_\mu^h(1,1)$, $\hat{L}_\mu^H$, and $\hat{S}_\mu$, respectively, we obtain

$$
\begin{aligned}
&\hat{\mathcal{M}}_\mu(1,1) \\
&= \left[1 - c_\mu^4 \frac{1 + (\lambda + 4\tilde{\sigma} s_\mu^2/h^2 - G_{-\mu}\eta_j)\Delta\tau}{1 + (\lambda + 4\tilde{\sigma} s_\mu^2 c_\mu^2/h^2 - G_\mu \eta_{j'})\Delta\tau} \right] \left[\frac{\Delta\tau G_{-\mu}\eta_j}{1 + (\lambda + 4\tilde{\sigma} s_\mu^2/h^2)\Delta\tau} \right] \\
&= \left[\frac{h^2(1 - c_\mu^4) + (\lambda - G_{-\mu}\eta_j)h^2\Delta\tau(1 - c_\mu^4) + 4\tilde{\sigma}\Delta\tau s_\mu^4 c_\mu^2}{h^2 + (\lambda - G_\mu\eta_j)h^2\Delta\tau + 4\tilde{\sigma}\Delta\tau s_\mu^2 c_\mu^2} \right] \left[\frac{h^2\Delta\tau G_{-\mu}\eta_j}{h^2 + \lambda h^2\Delta\tau + 4\tilde{\sigma}\Delta\tau s_\mu^2} \right]
\end{aligned}
$$

where $\tilde{\sigma} = (\sigma^2 + \bar{\sigma})/2$. Since G_μ is the Fourier symbol of $\hat{\gamma}$, by (2.5), we have $|G_\mu| \le \lambda$. Also, $|\eta_j| \le 1$ as proved in [40, Section 5.2]. Hence

$$
\begin{aligned}
&|\hat{\mathcal{M}}_\mu(1,1)| \\
&\le \frac{h^2(1 - c_\mu^4) + 2\lambda h^2\Delta\tau(1 - c_\mu^4) + 4\tilde{\sigma}\Delta\tau s_\mu^4 c_\mu^2}{|h^2 + (\lambda - G_\mu\eta_{j'})h^2\Delta\tau + 4\tilde{\sigma}\Delta\tau s_\mu^2 c_\mu^2|} \frac{\lambda h^2\Delta\tau}{h^2 + \lambda h^2\Delta\tau + 4\tilde{\sigma}\Delta\tau s_\mu^2}.
\end{aligned}
\tag{A.3}
$$

Write the complex number $G_\mu\eta_{j'}$ as $a + bi$ where $i = \sqrt{-1}$. Then we have

$$
\begin{aligned}
&|h^2 + (\lambda - G_\mu\eta_{j'})h^2\Delta\tau + 4\tilde{\sigma}\Delta\tau s_\mu^2 c_\mu^2|^2 \\
&= (h^2 + 4\tilde{\sigma}\Delta\tau s_\mu^2 c_\mu^2 + (\lambda - a)h^2\Delta\tau)^2 + (h^2\Delta\tau b)^2 \\
&= (h^2 + 4\tilde{\sigma}\Delta\tau s_\mu^2 c_\mu^2)^2 + 2(\lambda - a)h^2\Delta\tau(h^2 + 4\tilde{\sigma}\Delta\tau s_\mu^2 c_\mu^2) \\
&\quad + h^4\Delta\tau^2((\lambda - a)^2 + b^2).
\end{aligned}
\tag{A.4}
$$

Note that $a^2 + b^2 \le \lambda^2$ since $|G_\mu\eta_{j'}| \le \lambda$. Thus $\lambda - a \ge 0$. Since all the terms on the right-hand side of (A.4) are non-negative, it can be lower bounded by

$$
|h^2 + (\lambda - G_\mu\eta_{j'})h^2\Delta\tau + 4\tilde{\sigma}\Delta\tau s_\mu^2 c_\mu^2| \ge 4\tilde{\sigma}\Delta\tau s_\mu^2 c_\mu^2.
\tag{A.5}
$$

Combining (A.3) and (A.5), we have

$$
\begin{aligned}
&|\hat{\mathcal{M}}_\mu(1,1)| \\
&\leq \frac{h^2(1-c_\mu^4) + 2\lambda h^2 \Delta\tau(1-c_\mu^4) + 4\tilde{\sigma}\Delta\tau s_\mu^4 c_\mu^2}{4\tilde{\sigma}\Delta\tau s_\mu^2 c_\mu^2} \frac{\lambda h^2 \Delta\tau}{h^2 + \lambda h^2 \Delta\tau + 4\tilde{\sigma}\Delta\tau s_\mu^2} \\
&= \left[\frac{h^2}{4\tilde{\sigma}\Delta\tau}\left(\frac{1+c_\mu^2}{c_\mu^2}\right) + \frac{\lambda h^2}{2\tilde{\sigma}}\left(\frac{1+c_\mu^2}{c_\mu^2}\right) + s_\mu^2 \right] \frac{\lambda h^2 \Delta\tau}{h^2 + \lambda h^2 \Delta\tau + 4\tilde{\sigma}\Delta\tau s_\mu^2} \\
&< \left[\frac{h^2}{4\tilde{\sigma}\Delta\tau}\left(\frac{1+c_\mu^2}{c_\mu^2}\right) + \frac{\lambda h^2}{2\tilde{\sigma}}\left(\frac{1+c_\mu^2}{c_\mu^2}\right) + s_\mu^2 \right]\left[\frac{\lambda h^2}{\lambda h^2 + 4\tilde{\sigma}s_\mu^2} \right].
\end{aligned} \tag{A.6}
$$

Consider the first term on the right-hand side of (A.6). Since $0.5 \leq c_\mu^2 \leq 1$ and $4\tilde{\sigma}s_\mu^2 \geq 0$, we have

$$
\begin{aligned}
\frac{1}{4\tilde{\sigma}}\frac{h^2}{\Delta\tau}\left(\frac{1+c_\mu^2}{c_\mu^2}\right)\frac{\lambda h^2}{\lambda h^2 + 4\tilde{\sigma}s_\mu^2} &\leq \frac{1}{4\tilde{\sigma}}\frac{h^2}{\Delta\tau}\left(1+\frac{1}{c_\mu^2}\right) \\
&\leq \frac{3}{4\tilde{\sigma}}\frac{h^2}{\Delta\tau}.
\end{aligned} \tag{A.7}
$$

By assumption (3.17), equation (A.7) becomes

$$
\begin{aligned}
\frac{3}{4\tilde{\sigma}}\frac{h^2}{\Delta\tau} &< \frac{3}{4\tilde{\sigma}}\frac{2\tilde{\sigma}^2}{3\lambda h^2 + 6\tilde{\sigma}} \\
&= \frac{\tilde{\sigma}}{2\lambda h^2 + 4\tilde{\sigma}} \\
&= \frac{\frac{\tilde{\sigma}}{2\lambda h^2}}{1 + \frac{4\tilde{\sigma}}{2\lambda h^2}}.
\end{aligned}
$$

Let $\xi = 4\tilde{\sigma}/\lambda h^2$. Then, we have

$$
\frac{3}{4\tilde{\sigma}}\frac{h^2}{\Delta\tau} < \frac{\frac{1}{8}\xi}{1 + \frac{1}{2}\xi}. \tag{A.8}
$$

For the second term on the right-hand side of (A.6), we have

$$
\begin{aligned}
\frac{\lambda h^2}{2\tilde{\sigma}}\left(\frac{1+c_\mu^2}{c_\mu^2}\right)\frac{\lambda h^2}{\lambda h^2 + 4\tilde{\sigma}s_\mu^2} &\leq \frac{\lambda h^2}{2\tilde{\sigma}}\left(1+\frac{1}{c_\mu^2}\right) \\
&\leq \frac{3}{2}\frac{\lambda h^2}{\tilde{\sigma}}.
\end{aligned} \tag{A.9}
$$

By assumption (3.18), we have

$$
3\lambda h^2 < (2\sqrt{3}-3)\tilde{\sigma}. \tag{A.10}
$$

Also by assumption (3.18), we have

$$
\lambda h^2 + 2\tilde{\sigma} < \frac{2\sqrt{3}+3}{3}\tilde{\sigma}. \tag{A.11}
$$

Multiplying (A.10) and (A.11) yields

$$(3\lambda h^2)(\lambda h^2 + 2\tilde{\sigma}) < \tilde{\sigma}^2,$$

which implies

$$\frac{3\lambda h^2}{2\tilde{\sigma}} < \frac{\frac{1}{2}\tilde{\sigma}}{\lambda h^2 + 2\tilde{\sigma}} = \frac{\frac{1}{8}\xi}{1 + \frac{1}{2}\xi}. \tag{A.12}$$

Combining (A.9) and (A.12), we obtain

$$\frac{\lambda h^2}{2\tilde{\sigma}} \left(\frac{1 + c_\mu^2}{c_\mu^2} \right) \frac{\lambda h^2}{\lambda h^2 + 4\tilde{\sigma} s_\mu^2} < \frac{\frac{1}{8}\xi}{1 + \frac{1}{2}\xi}. \tag{A.13}$$

We will bound the third term on the right-hand side of (A.6) later. Now, consider $\hat{\mathcal{M}}_\mu(1, 2)$, the (1,2) entry of $\hat{\mathcal{M}}_\mu$. By (A.2), (3.15), (3.16), and (3.14), we have

$$
\begin{aligned}
\hat{\mathcal{M}}_\mu(1, 2) &= \left[s_\mu^2 c_\mu^2 \frac{1 + (\lambda + 4\tilde{\sigma} c_\mu^2/h^2 - G_\mu \eta_j)\Delta\tau}{1 + (\lambda + 4\tilde{\sigma} s_\mu^2 c_\mu^2/h^2 - G_{-\mu}\eta_{j'})\Delta\tau} \right] \left[\frac{\Delta\tau G_\mu \eta_j}{1 + (\lambda + 4\tilde{\sigma} c_\mu^2/h^2)\Delta\tau} \right] \\
&= \left[\frac{h^2 s_\mu^2 c_\mu^2 + (\lambda - G_\mu \eta_j)h^2 \Delta\tau s_\mu^2 c_\mu^2 + 4\tilde{\sigma}\Delta\tau s_\mu^2 c_\mu^4}{h^2 + (\lambda - G_{-\mu}\eta_{j'})h^2\Delta\tau + 4\tilde{\sigma}\Delta\tau s_\mu^2 c_\mu^2} \right] \left[\frac{\Delta\tau G_\mu \eta_j}{1 + (\lambda + 4\tilde{\sigma} c_\mu^2/h^2)\Delta\tau} \right]
\end{aligned}
$$

By a similar calculation as above, we can bound $|\hat{\mathcal{M}}_\mu(1, 2)|$ as:

$$
\begin{aligned}
|\hat{\mathcal{M}}_\mu(1, 2)| &< \left[\frac{h^2 s_\mu^2 c_\mu^2 + 2\lambda h^2 \Delta\tau s_\mu^2 c_\mu^2 + 4\tilde{\sigma}\Delta\tau s_\mu^2 c_\mu^4}{4\tilde{\sigma}\Delta\tau s_\mu^2 c_\mu^2} \right] \left[\frac{\lambda h^2}{\lambda h^2 + 4\tilde{\sigma} c_\mu^2} \right] \\
&= \left[\frac{h^2}{4\tilde{\sigma}\Delta\tau} + \frac{\lambda h^2}{2\tilde{\sigma}} + c_\mu^2 \right] \left[\frac{\lambda h^2}{\lambda h^2 + 4\tilde{\sigma} c_\mu^2} \right]. \tag{A.14}
\end{aligned}
$$

The first term on the right-hand side can be bounded as:

$$\frac{1}{4\tilde{\sigma}} \frac{h^2}{\Delta\tau} \frac{\lambda h^2}{\lambda h^2 + 4\tilde{\sigma} c_\mu^2} \leq \frac{1}{4\tilde{\sigma}} \frac{h^2}{\Delta\tau} < \frac{\frac{1}{24}\xi}{1 + \frac{1}{2}\xi}, \tag{A.15}$$

where the last inequality follows from (A.8). The second term on the right-hand side of (A.14) can be bounded similarly as:

$$\frac{\lambda h^2}{2\tilde{\sigma}} \frac{\lambda h^2}{\lambda h^2 + 4\tilde{\sigma} c_\mu^2} \leq \frac{\lambda h^2}{2\tilde{\sigma}} < \frac{\frac{1}{24}\xi}{1 + \frac{1}{2}\xi}, \tag{A.16}$$

where the last inequality follows from (A.12). Combining (A.6)-(A.16), we obtain

$$
\begin{aligned}
|\hat{\mathcal{M}}_\mu(1, 1)| + |\hat{\mathcal{M}}_\mu(1, 2)| &< \frac{\frac{1}{8}\xi}{1 + \frac{1}{2}\xi} + \frac{\frac{1}{8}\xi}{1 + \frac{1}{2}\xi} + \frac{\frac{1}{24}\xi}{1 + \frac{1}{2}\xi} + \frac{\frac{1}{24}\xi}{1 + \frac{1}{2}\xi} + \frac{\lambda h^2 s_\mu^2}{\lambda h^2 + 4\tilde{\sigma} s_\mu^2} \\
&\quad + \frac{\lambda h^2 c_\mu^2}{\lambda h^2 + 4\tilde{\sigma} c_\mu^2}
\end{aligned}
$$

$$
= \quad \frac{\frac{1}{3}\xi}{1 + \frac{1}{2}\xi} + \frac{s_\mu^2}{1 + \xi s_\mu^2} + \frac{c_\mu^2}{1 + \xi c_\mu^2}
$$

$$
= \quad \frac{\frac{1}{3}\xi}{1 + \frac{1}{2}\xi} + \frac{1 + 2\xi s_\mu^2 c_\mu^2}{(1 + \xi s_\mu^2)(1 + \xi c_\mu^2)}
$$

$$
= \quad \frac{\frac{1}{3}\xi}{1 + \frac{1}{2}\xi} + \frac{1 + \frac{1}{2}\xi \sin^2(\mu\pi h)}{1 + \xi + \frac{1}{4}\xi^2 \sin^2(\mu\pi h)}.
$$

By simple calculus, the maximum of the second term occurs when $\sin^2(\mu\pi h) = 1$. Hence it can be bounded by $1/(1 + \frac{1}{2}\xi)$. As a result, we have

$$
\begin{aligned}
|\hat{\mathcal{M}}_\mu(1, 1)| + |\hat{\mathcal{M}}_\mu(1, 2)| \quad &< \quad \frac{\frac{1}{3}\xi}{1 + \frac{1}{2}\xi} + \frac{1}{1 + \frac{1}{2}\xi} \\
&= \quad \frac{1 + \frac{1}{3}\xi}{1 + \frac{1}{2}\xi} < 1.
\end{aligned}
$$

The case for the second row of $\hat{\mathcal{M}}_\mu$ is similar and we omit the corresponding calculations here. □