Comput Visual Sci (2019) 22:27–41
https://doi.org/10.1007/s00791-017-0284-8

SPECIAL ISSUE IMG 2016

Multigrid methods for convergent mixed finite difference scheme for Monge–Ampère equation

Yangang Chen[1] · **Justin W. L. Wan**[2]

Received: 1 April 2017 / Accepted: 12 September 2017 / Published online: 5 October 2017

Abstract We propose multigrid methods for convergent mixed finite difference discretization for the two dimensional Monge–Ampère equation. We apply mixed standard 7-point stencil and semi-Lagrangian wide stencil discretization, such that the numerical solution is guaranteed to converge to the viscosity solution of the Monge–Ampère equation. We investigate multigrid methods for two scenarios. The first scenario considers applying standard 7-point stencil discretization on the entire computational domain. We use full approximation scheme with four-directional alternating line smoothers. The second scenario considers the more general mixed stencil discretization and is used for the linearized problem. We propose a coarsening strategy where wide stencil points are set as coarse grid points. Linear interpolation is applied on the entire computational domain. At wide stencil points, injection as the restriction yields a good coarse grid correction. Numerical experiments show that the convergence rates of the proposed multigrid methods are mesh-independent.

Communicated by Gabriel Wittum.

✉ Yangang Chen
y493chen@uwaterloo.ca

Justin W. L. Wan
justin.wan@uwaterloo.ca

[1] Department of Applied Mathematics, University of Waterloo, 200 University Avenue West, Waterloo, ON N2L 3G1, Canada

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

1 Introduction

We consider multigrid methods for solving a two dimensional elliptic partial differential equation (PDE), called Monge–Ampère equation:

$$\begin{aligned} u_{xx}u_{yy} - u_{xy}^2 &= f, && \text{in } \Omega, \\ u &= g, && \text{on } \partial\Omega, \\ u &\text{ is convex.} \end{aligned} \tag{1}$$

Here Ω is a bounded convex domain in $\mathbb{R}^2$, $\partial\Omega$ is its boundary, $u : \Omega \to \mathbb{R}$ is the unknown function, and $f : \Omega \to \mathbb{R}$ and $g : \partial\Omega \to \mathbb{R}$ are given functions. The Monge–Ampère equation has a wide range of applications, including differential geometry, optimal mass transport (or Monge–Kantorovich) problem, image registration, mesh generation, etc. [13].

In this paper, we compute the viscosity solution [15,16,20] of the Monge–Ampère equation. Designing a numerical scheme that converges to the viscosity solution turns out to be challenging. A number of numerical schemes has been proposed [6,20–22,33]. However, the number of the stencil points grows to infinity, or as high as 48, as $h \to 0$, which results in high computational costs. Our approach is to first convert the Monge–Ampère equation to an equivalent Hamilton–Jacobi–Bellman (HJB) equation [27,29], and then apply a mixed standard 7-point stencil and semi-Lagrangian wide stencil finite difference discretization. Solving the equivalent HJB equation is numerically more manageable. Mixed finite difference scheme guarantees the convergence to the viscosity solution, and meanwhile achieves an optimal accuracy of $O(h^2)$. Moreover, the mixed stencil discretization has at most 17 stencil points for any h. Our recent investigations in [14,28] elaborate the mixed

stencil discretization and provide a theoretical proof for the convergence.

To solve the resulting nonlinear discretized system, we consider using fast solvers, especially when the grid size N is large. Multigrid methods [40] are considered as the optimal solvers for many elliptic problems, in the sense that the convergence rates are independent of the grid size. Considering that the Monge–Ampère equation is elliptic, our primary objective is to develop multigrid methods for the Monge–Ampère equation, or the equivalent HJB equation, and yield mesh-independent convergence rates.

In general, there are two families of multigrid methods for solving nonlinear HJB equations. The first family is to use multigrid for the linearized problem nested inside the policy iteration [19,25], such as the multigrid in [1,2]. The second family is the full approximation scheme (FAS), proposed in [11] and is further studied in [9,23,24]. In this paper, we investigate FAS for the Monge–Ampère equation where standard 7-point stencil discretization is applied on the entire computational domain. We show that it yields a more effective multigrid solver than the first family of multigrid. However, for most of the paper, in particular for the general mixed stencil discretization, we consider first family of multigrid. That is, multigrid methods are used to solve the linearized discrete system.

Standard multigrid methods turn out to have poor convergence. There are two major factors behind the poor convergence. One is that the PDE may become anisotropic along various directions. Standard pointwise smoothers fail to smooth the error along the weakly connected directions.

The other factor that leads to the poor convergence is the semi-Lagrangian wide stencil discretization. An immediate challenge associated with the wide stencil discretization is that the resulting matrix A_h is non-symmetric. We note that standard discretization of elliptic PDEs usually gives rise to symmetric matrices, and most of the multigrid theories are developed for symmetric matrices. The existing literature of multigrid methods for non-symmetric matrices are mostly restricted to convection–diffusion equations [4,8,12,26].

Very few investigations, such as [34,36], are related to non-symmetric matrices beyond convection–diffusion equations. In particular, to the best of our knowledge, [34] is the only reference that investigates multigrid methods in the context of semi-Lagrangian discretization. Both [36] and [34] use existing algebraic multigrid (AMG) methods [31,32,35,38,39] as preconditioners. We refer readers to [38,40,42] for substantial reviews of AMG. The basic idea of AMG is to perform coarsening along the strongly connected grid points, where the strength of connection between grid points i and j is defined by the magnitude of the matrix entry $A_{i,j}$. AMG methods assess geometric information indirectly though the strength of connections. Hence, it may not be the optimal design for the mixed stencil discretization on a square grid. As pointed out in [34], although AMG methods as preconditioners give approximately mesh-independent convergence, AMG methods themselves are not efficient if used as stand-alone solvers. Unfortunately, no new multigrid methods are proposed to address the issue.

Our contribution is to propose a multigrid method as an effective stand-alone solver. We observe that wide stencils introduce oscillations (or local singularities) to the error, and such oscillations cannot be eliminated by smoothers, including the alternating line smoothers. However, the oscillations are restricted at the wide stencil points. One possible solution to capture the oscillations is to use a sophisticated interpolation, such as biased AMG interpolation in [35]. However, the interpolation matrix can be dense, and expensive to set up. Alternatively, our solution is to change the coarsening strategy. We propose setting wide stencil points as coarse grid points. The purpose is to directly use the coarse grid points to capture the oscillations. As the wide stencils are mainly restricted to the singular points or singular lines, setting wide stencil points as coarse grid points does not significantly increase the number of the coarse grid points. In our numerical experiments, we illustrate that the proposed multigrid method has mesh-independent convergence rate even as a standalone linear solver.

This paper is organized as follows. We first introduce the mixed finite difference discretization in Sect. 2. In Sects. 3–4, we propose multigrid methods for the standard 7-point stencil discretization and the more general mixed stencil discretization separately. Section 5 includes smoothing analysis on the four-directional alternating line smoother. Section 6 shows that the proposed multigrid method as a standalone solver can achieve mesh-independent convergence.

2 Mixed finite difference discretization

In this section, we describe a mixed finite difference discretization for the Monge–Ampère Eq. (1). The significance of the mixed finite difference discretization is that it is guaranteed to converge to the viscosity solution of the Monge–Ampère equation [15,16,20]. We refer the interested readers to [14] for more details.

2.1 HJB formulation of Monge–Ampère equation

In order to design a finite difference scheme that converges to the viscosity solution, we first convert the Monge–Ampère equation into an equivalent HJB equation. The equivalence of the two PDEs is first established in [27] and [29]. Here we present the equivalent HJB equation as follows:

Theorem 1 *Let Ω be a convex open set in $\mathbb{R}^2$. Let $u \in C^2(\Omega)$ be a convex function, and $f \in C(\Omega)$ be a non-*

negative function. Then u solves the Monge–Ampère Eq. (1) if and only if it solves the following HJB equation,

$$\max_{(a(\mathbf{x}),\theta(\mathbf{x}))\in\Gamma}\Big\{-\alpha_{11}(a(\mathbf{x}),\theta(\mathbf{x}))u_{xx}(\mathbf{x}) - 2\alpha_{12}(a(\mathbf{x}),\theta(\mathbf{x}))u_{xy}(\mathbf{x}) - \alpha_{22}(a(\mathbf{x}),\theta(\mathbf{x}))u_{yy}(\mathbf{x}) + 2\sqrt{a(\mathbf{x})(1-a(\mathbf{x}))f(\mathbf{x})}\Big\} = 0, \tag{2}$$

where $(a(\mathbf{x}),\theta(\mathbf{x}))$ *is the pair of controls at point* $\mathbf{x}$, $\Gamma = [0,1]\times\left[-\frac{\pi}{4},\frac{\pi}{4}\right)$ *is the set of admissible controls, and the coefficients are*

$$\begin{aligned}\alpha_{11}(a(\mathbf{x}),\theta(\mathbf{x})) &= \frac{1}{2}[1-(1-2a(\mathbf{x}))\cos 2\theta(\mathbf{x})],\\ \alpha_{22}(a(\mathbf{x}),\theta(\mathbf{x})) &= \frac{1}{2}[1+(1-2a(\mathbf{x}))\cos 2\theta(\mathbf{x})],\\ \alpha_{12}(a(\mathbf{x}),\theta(\mathbf{x})) &= \frac{1}{2}(1-2a(\mathbf{x}))\sin 2\theta(\mathbf{x}).\end{aligned} \tag{3}$$

Proof See [37]. □

The reasons for choosing the HJB formulation over the Monge–Ampère equation are threefold. One is that for a fixed given control pair (a,θ), the differential operator of the HJB Eq. (2) is linear, which makes the development of finite difference schemes more tractable. Another reason is that the convexity constraint of (1) is already implicitly enforced in the HJB formulation and can thus be removed. Last but not least, numerous literature has been devoted to convergent numerical schemes for HJB equations [3,10,17,19,25,30,41].

2.2 Standard 7-point stencil discretization

Next, we will construct a convergent finite difference discretization for the HJB Eq. (2). Let us consider an $n\times n$ square grid $\{\mathbf{x}_{i,j}=(x_i,y_j)\,|\,i,j=1,\ldots,n\}$. Let $N=n^2$ be the total number of the grid points. Let h be the mesh size. Also, let $u_{i,j}, a_{i,j}, \theta_{i,j}$ and $f_{i,j}$ be the grid functions of $u(\mathbf{x}_{i,j}), a(\mathbf{x}_{i,j}), \theta(\mathbf{x}_{i,j})$ and $f(\mathbf{x}_{i,j})$ respectively. Our goal is to solve the set of the unknowns $\{u_{i,j}\,|\,i,j=1,\ldots,n\}$.

Reference [5] has proved that consistency, stability and monotonicity are sufficient conditions for a numerical solution to converge to the viscosity solution of a nonlinear PDE as $h\to 0$. For the equivalent HJB Eq. (2), ensuring monotonicity is the major issue.

Consider discretizing the differential operator of the HJB Eq. (2) at a grid point $\mathbf{x}_{i,j}$. When $\theta_{i,j}=0$, $u_{xy}(\mathbf{x}_{i,j})$ term disappears. The discretization is given by a 5-point stencil:

$$L_h^{[0]} \equiv \frac{1}{h^2}\begin{bmatrix} 0 & -\alpha_{22} & 0\\ -\alpha_{11} & 2(\alpha_{11}+\alpha_{22}) & -\alpha_{11}\\ 0 & -\alpha_{22} & -0\end{bmatrix}_h. \tag{4}$$

Here the discretization is written in terms of the widely-used stencil notations in multigrid literature [40]. For simplicity, we have suppressed the dependency of $\alpha_{11},\alpha_{22},\alpha_{12}$ and $L_h^{[0]}$ on the control $(a_{i,j},\theta_{i,j})$.

However, $\theta_{i,j}$ is not necessarily 0 and $u_{xy}(\mathbf{x}_{i,j})$ appears in the HJB equation in general. It can be shown that the standard 7-point stencil discretization can lead to a monotone discretization in the following two cases [5]:

(i) When the coefficients α_{11}, α_{22} and α_{12}, given by (3), satisfy

$$\begin{aligned}&\alpha_{11}(a_{i,j},\theta_{i,j}) \ge |\alpha_{12}(a_{i,j},\theta_{i,j})|,\\ &\alpha_{22}(a_{i,j},\theta_{i,j}) \ge |\alpha_{12}(a_{i,j},\theta_{i,j})|,\\ \text{and}\quad &\alpha_{12}(a_{i,j},\theta_{i,j}) \ge 0 \text{ at the grid point } \mathbf{x}_{i,j},\end{aligned} \tag{5}$$

the HJB differential operator is monotonically discretized by

$$L_h^{[1]} \equiv \frac{1}{h^2}\begin{bmatrix} 0 & -(\alpha_{22}-\alpha_{12}) & -\alpha_{12}\\ -(\alpha_{11}-\alpha_{12}) & 2(\alpha_{11}+\alpha_{22}-\alpha_{12}) & -(\alpha_{11}-\alpha_{12})\\ -\alpha_{12} & -(\alpha_{22}-\alpha_{12}) & 0\end{bmatrix}_h. \tag{6}$$

(ii) When the coefficients α_{11}, α_{22} and α_{12} in (3) satisfy

$$\begin{aligned}&\alpha_{11}(a_{i,j},\theta_{i,j}) \ge |\alpha_{12}(a_{i,j},\theta_{i,j})|,\\ &\alpha_{22}(a_{i,j},\theta_{i,j}) \ge |\alpha_{12}(a_{i,j},\theta_{i,j})|,\\ \text{and}\quad &\alpha_{12}(a_{i,j},\theta_{i,j}) \le 0 \text{ at the grid point } \mathbf{x}_{i,j},\end{aligned} \tag{7}$$

the HJB differential operator is monotonically discretized by

$$L_h^{[2]} \equiv \frac{1}{h^2}\begin{bmatrix} \alpha_{12} & -(\alpha_{22}+\alpha_{12}) & 0\\ -(\alpha_{11}+\alpha_{12}) & 2(\alpha_{11}+\alpha_{22}+\alpha_{12}) & -(\alpha_{11}+\alpha_{12})\\ 0 & -(\alpha_{22}+\alpha_{12}) & \alpha_{12}\end{bmatrix}_h. \tag{8}$$

Figure 1a, b provide graphic representations for the 7-point stencils (6) and (8).

2.3 Semi-Lagrangian wide stencil discretization

However, if neither (5) nor (7) is fulfilled at the grid point $\mathbf{x}_{i,j}$, then it is unclear how to directly discretize $u_{xy}(\mathbf{x}_{i,j})$ monotonically. Following [17,30], we eliminate the cross derivative $u_{xy}(\mathbf{x}_{i,j})$ by a local coordinate transformation. Let $\{(\mathbf{e}_z)_{i,j},(\mathbf{e}_w)_{i,j}\}$ be a rotated orthogonal basis at $\mathbf{x}_{i,j}$; see Fig. 1c. One can prove that if the rotation angle is

$$\phi_{i,j} = \frac{1}{2}\arctan\frac{2\alpha_{12}\left(a_{i,j},\theta_{i,j}\right)}{\alpha_{11}\left(a_{i,j},\theta_{i,j}\right)-\alpha_{22}\left(a_{i,j},\theta_{i,j}\right)} = -\theta_{i,j},$$

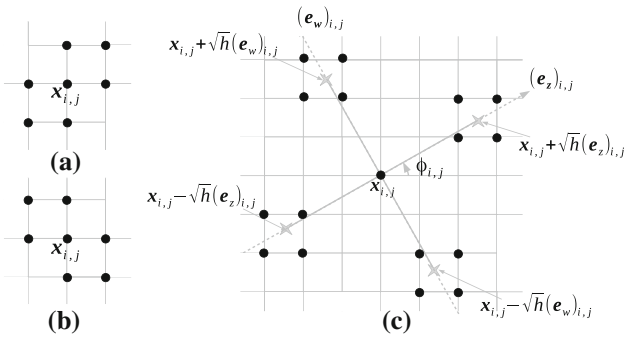

Fig. 1 **a** Graphic representations of the 7-point stencils (6). **b** Graphic representations of the 7-point stencils (8). **c** Graphic representations of a semi-Lagrangian wide stencil. The rotation angle is $\phi_{i,j}$. The grey dashed lines are the orthogonal axes $\{(\mathbf{e}_z)_{i,j}, (\mathbf{e}_w)_{i,j}\}$. The stencil length is $\sqrt{h}$. The grey stars are the points $\mathbf{x}_{i,j} \pm \sqrt{h}(\mathbf{e}_z)_{i,j}$ and $\mathbf{x}_{i,j} \pm \sqrt{h}(\mathbf{e}_w)_{i,j}$. The unknowns at these points are approximated by the bilinear interpolation from the neighboring points (black dots)

then the cross derivative vanishes under the rotated basis, and the HJB Eq. (2) becomes

$$\max_{(a_{i,j},\theta_{i,j})\in\Gamma} \left\{ -a_{i,j}\, u_{zz}(\mathbf{x}_{i,j}) - \left(1 - a_{i,j}\right) u_{ww}(\mathbf{x}_{i,j}) + 2\sqrt{a_{i,j}\left(1 - a_{i,j}\right) f_{i,j}} \right\} = 0, \tag{9}$$

where $u_{zz}(\mathbf{x}_{i,j})$ and $u_{ww}(\mathbf{x}_{i,j})$ are the directional derivatives along the basis $(\mathbf{e}_z)_{i,j}$ and $(\mathbf{e}_w)_{i,j}$.

Figure 1c explains the discretization of (9). The standard central differencing $u_{zz}(\mathbf{x}_{i,j})$, namely,

$$\frac{1}{h^2}\left[u(\mathbf{x}_{i,j} + h(\mathbf{e}_z)_{i,j}) - 2u_{i,j} + u(\mathbf{x}_{i,j} - h(\mathbf{e}_z)_{i,j})\right],$$

is guaranteed to be monotone. However, since the stencil is rotated, the stencil points $\mathbf{x}_{i,j} \pm h(\mathbf{e}_z)_{i,j}$ may no longer coincide with any grid points. We can consider approximating $u(\mathbf{x}_{i,j} \pm h(\mathbf{e}_z)_{i,j})$ using bilinear interpolation from its four neighboring grid points. However, under bilinear interpolation, the truncation error becomes $O(1)$, which is not consistent. Instead, we choose the stencil length to be $\sqrt{h}$, which yields $O(h)$ truncation error. As a result, the finite difference discretization for $u_{zz}(\mathbf{x}_{i,j})$ is given by

$$(\delta_{zz}u)_{i,j} \equiv \frac{\mathcal{I}_h u|_{\mathbf{x}_{i,j}+\sqrt{h}(\mathbf{e}_z)_{i,j}} - 2u_{i,j} + \mathcal{I}_h u|_{\mathbf{x}_{i,j}-\sqrt{h}(\mathbf{e}_z)_{i,j}}}{h},$$

where $\mathcal{I}_h u|_{\mathbf{x}_{i,j}\pm\sqrt{h}(\mathbf{e}_z)_{i,j}}$ represents the bilinear interpolation from the four neighboring unknown values at the points $\mathbf{x}_{i,j} \pm \sqrt{h}(\mathbf{e}_z)_{i,j}$. The other directional derivative, $u_{ww}(\mathbf{x}_{i,j})$, can be discretized in a similar fashion. Note that when h is small, $\sqrt{h} > h$, which means the stencil length appears to be wide. Hence this approach is called "semi-Lagrangian wide stencil discretization" [17,30].

2.4 Mixed discretization

The advantage of the semi-Lagrangian wide stencil discretization is that it is unconditionally monotone [14,18]. However, it is only first order accurate, while the standard 7-point stencil discretization is second order accurate. In order to combine the advantages of the two discretization schemes, we apply the following "mixed stencil discretization" at each grid point $\mathbf{x}_{i,j}$:

(i) When the control pair $(a_{i,j}, \theta_{i,j})$ satisfies Condition (5) or (7), the discrete equation is given by the standard 7-point stencil discretization:

$$\max_{(a_{i,j},\theta_{i,j})\in\Gamma} \left\{ L_h^{[disc]}(a_{i,j}, \theta_{i,j})u_{i,j} + 2\sqrt{a_{i,j}(1 - a_{i,j})\, f_{i,j}} \right\} = 0, \tag{10}$$

where $disc = 1$ (or 2) if Condition (5) (or Condition (7)) is satisfied respectively. We note that $L_h^{[disc]}(a_{i,j}, \theta_{i,j})$, given by (4), (6) and (8), depends on the control $(a_{i,j}, \theta_{i,j})$.

(ii) Otherwise, the discrete equation is given by semi-Lagrangian wide stencil discretization:

$$\max_{(a_{i,j},\theta_{i,j})\in\Gamma} \left\{ -a_{i,j}\,(\delta_{zz}u)_{i,j} - \left(1 - a_{i,j}\right)(\delta_{ww}u)_{i,j} + 2\sqrt{a_{i,j}\left(1 - a_{i,j}\right) f_{i,j}} \right\} = 0. \tag{11}$$

Similar to (6) and (8), we can use a graph to represent the wide stencil (11); see the black dots in Fig. 1c. Since each bilinear interpolation introduces 4 stencil points, there are in total 17 stencil points in a wide stencil.

The significance of mixed stencil discretization is that monotonicity is strictly maintained at every grid point, and meanwhile, the numerical scheme is as accurate as possible.

2.5 The nonlinear discrete system

The mixed discretization scheme gives rise to a nonlinear discrete system that contains $N = n^2$ discrete equations. If we define a vector of the unknowns $u_h \equiv (u_{1,1}, u_{1,2}, \ldots, u_{1,n}, u_{2,1}, \ldots, u_{n,n})^T \in \mathbb{R}^N$, and similarly, vectors of controls $a_h \in \mathbb{R}^N, \theta_h \in \mathbb{R}^N$, then the entire nonlinear discrete system can be written into the following matrix form:

$$\max_{(a_h,\theta_h)\in\Gamma} \{A_h(a_h, \theta_h)\, u_h - b_h(a_h, \theta_h)\} = 0, \tag{12}$$

where $A_h(a_h, \theta_h) \in \mathbb{R}^{N\times N}$ is a matrix that consists of the coefficients of u^h, and $b_h(a_h, \theta_h) \in \mathbb{R}^N$ is a vector that does not explicitly contain u^h.

3 Multigrid methods for standard 7-point stencil discretization

The next objective is to solve the complete nonlinear discretized system (12). When N is large, effective solvers are required. In this paper, we consider using multigrid methods. The motivation is that multigrid methods have been proved efficient for many elliptic problems, in the sense that the number of iterations is often independent of the grid size N. Since our discrete matrix A_h comes from the elliptic Monge–Ampère equation, we expect multigrid to be an effective solver for (12) as well.

3.1 Policy-MG iteration

One family of multigrid methods for solving the discretized HJB Eq. (12) is based on a global Newton-like iteration for the nonlinear system (12), called policy iteration (or Howard's algorithm) [19,25]. At each policy iteration, a linear multigrid solver is applied to solve the linearized system. The algorithm can be written as follows:

Start with an initial guess of the solution $u_h^{(0)}$.
For $k = 0, 1, \ldots$ until convergence:

1. Solve for the optimal control pair $(a_h^{(k)}, \theta_h^{(k)})$ under the current solution $u_h^{(k)}$:

$$\left(a_{i,j}^{(k)}, \theta_{i,j}^{(k)}\right) = \underset{(a_{i,j},\theta_{i,j})\in\Gamma_{i,j}}{\arg\max} \left\{ A_h(a_h, \theta_h)u_h^{(k)} - b_h(a_h, \theta_h) \right\}_{i,j}, \tag{13}$$

for all $\mathbf{x}_{i,j} \in \Omega$. Here $\Gamma_{i,j} = [0, 1] \times [-\frac{\pi}{4}, \frac{\pi}{4})$ is the control set at $\mathbf{x}_{i,j}$.
Meanwhile, obtain the residual

$$r_h^{(k)} = A_h\left(a_h^{(k)}, \theta_h^{(k)}\right) u_h^{(k)} - b_h\left(a_h^{(k)}, \theta_h^{(k)}\right). \tag{14}$$

2. If $\|r_h^{(k)}\| \le$ tolerance: break
Else, *use the multigrid V-cycle* to solve the following linear system for the solution $u_h^{(k+1)}$ under the current optimal control pair $(a_h^{(k)}, \theta_h^{(k)})$:

$$A_h\left(a_h^{(k)}, \theta_h^{(k)}\right) u_h^{(k+1)} = b_h\left(a_h^{(k)}, \theta_h^{(k)}\right) \Rightarrow u_h^{(k+1)}. \tag{15}$$

To summarize, in order to solve (12), the inner multigrid V-cycle iteration for linearized problems is nested in an outer policy iteration. For convenience, we refer this type of multigrid methods as "policy-MG iteration".

The advantage of using this approach is that policy iteration is guaranteed to converge for any initial guess $u_h^{(0)}$, if HJB equation is monotonically discretized [3,10]. Policy iteration consists of two sub-steps. The first sub-step is to solve the optimization problem at each grid point $\mathbf{x}_{i,j}$; see (13). Our recent work [14] discusses speeding up computation of the optimization problem in details. The second sub-step of the policy iteration is to solve the linear system under a given control pair; see (15). The second sub-step is our focus of developing multigrid methods.

We start with multigrid methods for the standard 7-point stencil discretization. More precisely, we consider the case where the standard 7-point stencil discretization can be applied on the entire computational domain and still results in a monotone scheme. We will leave the discussion of multigrid for more general mixed stencil discretization to Sect. 4.

3.2 Four-directional alternating line smoother

The components of standard multigrid include pointwise smoother, full coarsening, full-weighting restriction, bilinear interpolation and coarse grid operator (i.e. Galerkin coarse grid operator or direct discretization). However, there are some scenarios where standard multigrid may give poor convergence. We note that the equivalent HJB Eq. (2) can sometimes become anisotropic. For instance, when $a = \epsilon$ is a small constant close to 0 and $\theta = 0$, the linearized HJB equation becomes

$$\epsilon u_{xx} + (1 - \epsilon)u_{yy} - 2\sqrt{\epsilon(1 - \epsilon)}f = 0,$$

which is an anisotropic Poisson equation. It is well-known that when solving the anisotropic Poisson equation, the typical pointwise smoothers do not smooth the error along the weakly connected axis, which causes poor convergence rate.

To address anisotropy, we consider using line smooth-ers [40]. More specifically, instead of updating the unknowns point by point, we update strongly-connected grid points collectively. Fourier analysis has proved that line smoothers can effectively smooth the error.

For the 7-point discretization, the directions of strong connections can vary from patches to patches. In some patches, strong connections are aligned to x-axis, while in some others, aligned to y-axis, or even diagonal axes. Considering this, we apply four-directional alternating Gauss–Seidel line smoother. More specifically, the line smoother is applied four times - along x-axis (left to right), y-axis (top to bottom), diagonal axis (top left to bottom right) and eventually transpose diagonal axis (top right to bottom left).

3.3 Seven-point restriction and interpolation

Once the error becomes smooth along x, y and diagonal axes under the four-directional alternating line smoother, standard full coarsening can be applied. We note that some literature, such as [40], has discussed the multigrid methods for the 7-point discretization of $-u_{xx} - u_{yy} - \tau u_{xy}$, $|\tau| < 2$. In order to capture the directional feature of the 7-point discretization, the literature suggests the use of 7-point restriction operators. We would adopt the 7-point restriction operators to the HJB Eq. (2). More specifically, if the discretization of the HJB operator is given by (4), (6), or (8), then the corresponding 7-point restriction operator is given by

$$R^{[0]} = \frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}, \quad R^{[1]} = \frac{1}{8}\begin{bmatrix} 0 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 0 \end{bmatrix},$$
$$R^{[2]} = \frac{1}{8}\begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

respectively. Meanwhile, the corresponding interpolation operator is scaled transposed / adjoint with respect to grid size scaled scalar product, or more precisely,

$$P^{[disc]} = 4(R^{[disc]})^T, \quad disc = 0, 1, 2.$$

Galerkin coarse grid operator can be used as the coarse grid operator A_{2h}.

3.4 Full approximation scheme (FAS)

So far we have discussed solving the nonlinear discretized system (12) using policy-MG iteration. We will show in Sect. 6 that the convergence rate of the inner multigrid V-cycle is independent of mesh size. However, the convergence rate of the outer policy iteration is in general not bounded by a mesh-independent constant.

One approach to extend from linear systems to nonlinear systems is Full Approximation Scheme (FAS) [11,23,40]. The key of the FAS is the construction of the coarse grid problem. In particular, the coarse grid problem should be nonlinear. Denote the discretized nonlinear operator in (12) as

$$\mathcal{N}_h(u_h) \equiv \max_{(a_h,\theta_h)\in\Gamma} \{A_h(a_h,\theta_h)u_h - b_h(a_h,\theta_h)\}.$$

We use direct discretization to obtain $\mathcal{N}_{2h}$. Then we define the following coarse grid problem with respect to w_{2h}

$$\mathcal{N}_{2h}(w_{2h}) = \mathcal{N}_{2h}(u_{2h}) + Rr_h,$$

where u_{2h} is the injection of the solution u_h. The coarse grid error is then given by $e_{2h} = w_{2h} - u_{2h}$.

FAS uses nonlinear smoothers, which are convergent nonlinear relaxations that solve the nonlinear system (12) directly. One convergent nonlinear relaxation is the following:

Start with an initial guess of the solution $u_h^{(0)}$.
For $k = 0, 1, \ldots$ until convergence:

1. Use (13) to solve for the optimal control pair $(a_h^{(k)}, \theta_h^{(k)})$ under the current solution $u_h^{(k)}$. Meanwhile, obtain the residual $r_h^{(k)}$ by (14).
2. If $\|r_h^{(k)}\| \le$ tolerance: break
   Else, apply *one step four-directional alternating Gauss-Seidel line relaxation* to the linear system

$$A_h\left(a_h^{(k)}, \theta_h^{(k)}\right) u = b_h\left(a_h^{(k)}, \theta_h^{(k)}\right), \tag{16}$$

   which updates the solution from $u_h^{(k)}$ to $u_h^{(k+1)}$.

We will use this iteration as the nonlinear smoother for the FAS. For the coarse grid correction process, we use the standard coarsening, and 7-point restriction and interpolation.

The main advantage of the FAS is that it does not require the outer–inner multi-layer iteration as policy-MG iteration does. Instead, it contains only one layer of iteration.

4 Multigrid methods for mixed stencil discretization

In this section, we will discuss multigrid methods for mixed stencil discretization. We will only consider policy-MG iteration, as defined in Sect. 3.1. The reason is that for our proposed approach, the coarse grids are no longer square grids, which poses difficulties for defining the coarse grid problem under the FAS framework.

In the mixed stencil discretization, semi-Lagrangian wide stencils are applied to the grid points that do not satisfy Condition (5) or (7). An immediate challenge arising from the wide stencils is that the discrete matrix A_h becomes non-symmetric. Figure 2 shows the arising asymmetry of A_h under the mixed stencil discretization. Suppose wide stencil discretization is applied at the grid point i, and at one of its stencil points (denoted by j), standard 7-point stencil is applied. Then $A_{ij} \neq 0$ and $A_{ji} = 0$, resulting in asymmetry. This asymmetry also leads to non-smooth errors.

4.1 Failure of the standard multigrid method

We first investigate the performance of the standard multigrid method for mixed stencil discretization. To start with a simple

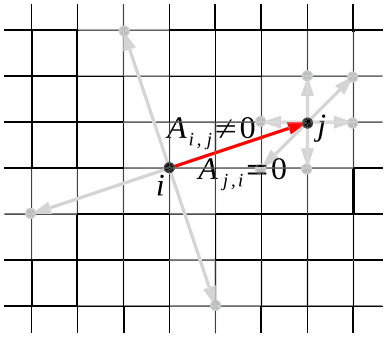

Fig. 2 Suppose wide stencil discretization is applied at the grid point i, and at one of its stencil points (denoted by j), standard 7-point stencil is applied. Then $A_{ij} \neq 0$ and $A_{ji} = 0$, resulting in asymmetry

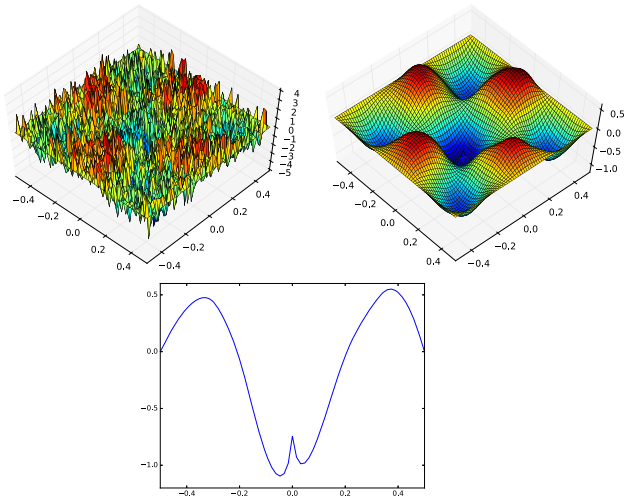

Fig. 3 The error after 2-step four-directional alternating Gauss–Seidel line smoothing. (top-left) Initial error. (top-right) Smoothed error. (bottom) Cross section of the smoothed error along x axis. A kink appears at the center of Ω

scenario, we assume that the control is given as $a = \frac{1}{2}$ and $\theta = 0$ on the entire computational domain, where standard 7-point stencil is applied, except that at the origin, $a = 1$ and $\theta = \frac{\pi}{8}$, where wide stencil is applied. In other words, we consider solving the following linearized HJB equation

$$\begin{aligned} &\frac{1}{2}u_{xx} + \frac{1}{2}u_{yy} = \sqrt{f}, \quad \text{in } \Omega\backslash\{(0,0)\}, \\ &\frac{2+\sqrt{2}}{4}u_{xx} + \frac{2-\sqrt{2}}{4}u_{yy} + \frac{1}{\sqrt{2}}u_{xy} = 0, \quad \text{at } (0,0), \\ &u = g, \quad \text{on } \partial\Omega. \end{aligned} \tag{17}$$

A monotone discretization is to apply standard 5-point stencil discretization everywhere and wide stencil at the origin. Figure 3 shows the error after applying the four-directional alternating line smoother. A kink appears at the center of Ω. In general, a kink appears wherever wide stencil discretization is applied at a grid point. This introduces an oscillatory perturbation to the smoothed error. We note that such oscillation also occurs for the other types of smoothers, such as (symmetric) pointwise Gauss–Seidel smoother.

4.2 Coarsening strategy

Choosing alternating line smoother is not sufficient to handle the oscillations from semi-Lagrangian wide stencils, and thus more sophisticated multigrid method is needed. One idea is to use algebraic multigrid (AMG) methods [35,38,39], which coarsens the grid only along the strongly connected grid points. However, AMG coarsening may lead to a dense coarse grid operator [38]. We observe that if we apply the four-directional alternating line smoother, the error becomes smooth on those standard 7-point stencil patches; see Fig. 3. This motivates us to still apply standard coarsening on the standard 7-point stencil patches, but consider other type of coarsening strategy to the wide stencil patches.

To motivate our coarsening strategy, we first consider the one-dimensional cross section of the error; see Fig. 4. We notice that if a wide stencil point i is an F-point, which requires an interpolation, then a naive bilinear interpolated error will fail to represent the kink accurately. Instead, we simply reset the wide stencil F-point as a coarse grid point. As a result, interpolation at the wide stencil point is no longer needed. The error at the wide stencil point is simply copied from the coarse grid to fine grid. This yields a more accurate estimated error.

The proposed coarsening strategy can be extended to two dimensions. Figure 5 illustrates the coarsening process. On the fine grid, the black dots are selected as coarse grid points, or C-points. The hollow dots are F-points. Suppose wide stencils are applied on the three red dots. Then these three dots are all assigned as C-points. The resulting 1st coarse grid is a combination of a square grid that comes from geometric coarsening, and some additional coarse grid points that come from wide stencils. We can continue to coarsen the square sub-grid and meanwhile keep all the wide stencil points as C-points, which generates the 2nd coarse grid. Such coarsening strategy can be applied recursively until the coarsest level.

One may argue that by setting all the wide stencil points as coarse grid points, the number of coarse grid points, and thus the computational complexity, will increase. However, it is observed in the numerical simulation that wide stencils are usually restricted around the grid points where the right hand side f or the solution u is singular. In other words, it is sufficient to apply wide stencils only at singularities and still maintain monotonicity on the entire computational domain. In practice, such singular points account for only a small portion of the total grid points. As a result, setting the wide stencil points as coarse grid points does not result in a significant increase of the number of the coarse grid points.

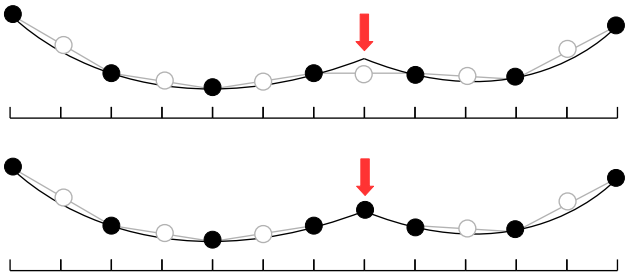

Fig. 4 Coarsening strategy for a wide stencil point (red arrow), where a kink appears. Black dots are coarse grid points (C-points), while hollow dots are non-coarse grid points (F-points). (top) Standard coarsening with linear interpolation. The interpolated error (grey piecewise straight lines) has a big error at the wide stencil point. (bottom) By setting the wide stencil point as a coarse grid point, the error at the point no longer needs to be interpolated, resulting in a more accurate estimated error

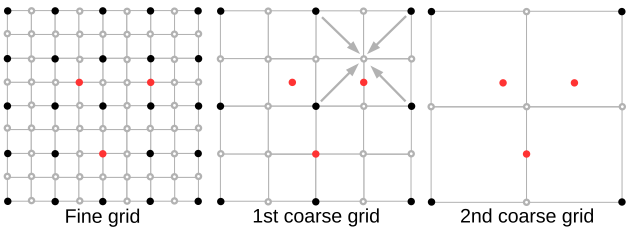

Fig. 5 Proposed coarsening strategy. On the fine grid, the black dots are selected as coarse grid points, or C-points. The hollow dots are F-points. Suppose wide stencils are applied on the three red dots. Then these three dots are all assigned as C-points. The resulting 1st coarse grid is a combination of a square grid that comes from geometric coarsening, and some additional coarse grid points that come from wide stencils. We can continue to coarsen the square sub-grid and meanwhile keep all the wide stencil points as C-points, which generates the 2nd coarse grid

4.3 Interpolation

Under the proposed coarsening strategy, all the wide stencil points are excluded from the set of F-points. In other words, F-points must be the standard 7-point stencils. Hence, the 7-point interpolation, as described in Sect. 3.3, can be used for interpolating the errors at these F-points.

We have seen in Fig. 5 that the coarse grids are no longer square grids. However, each of these coarse grids can be seen as a combination of a square grid and some additional wide-stencil C-points. Then all the F-points can still be interpolated from the C-points on the square sub-grid. The arrows in Fig. 5 show how F-points can be interpolated on a coarse grid.

4.4 Restriction

In both the standard multigrid and classical AMG methods, restriction is simply the transpose of interpolation. However, it does not result in mesh-independent convergence for the non-symmetric matrices A_h arising from the mixed stencil discretization. We will show such poor convergence in Sect. 6.2. In this paper, we propose a restriction operator R that is different from the transpose of the interpolation P.

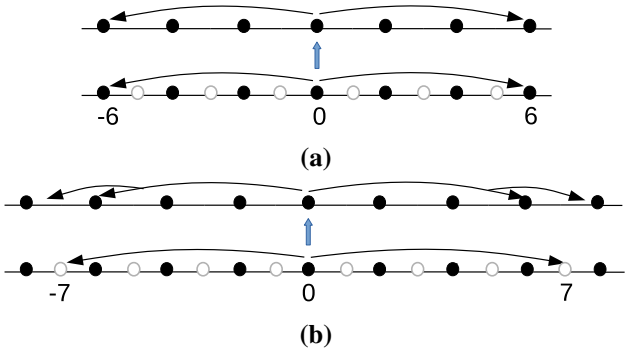

Fig. 6 Restriction for one-dimensional Poisson equation. The black points are C-points, while the hollow points are F-points. **a** $h = \frac{1}{36}$ and $\sqrt{h} = 6h$. **b** $h = \frac{1}{49}$ and $\sqrt{h} = 7h$.

Our approach is simply to use *injection* on wide stencil points. To motivate the use of injection, let us simplify our problem and start with the one-dimensional Poisson equation

$$-u_{xx} = 0, \quad x \in [-0.5, 0.5].$$

Under the standard coarsening, the C-points are $i = -\frac{N}{2}, \ldots, 0, 2, \ldots, \frac{N}{2}$, and F-points are $i = -\frac{N}{2} + 1, \ldots, -1, 1, \ldots, \frac{N}{2} - 1$. We apply wide stencil discretization at $i = 0$ and standard finite difference discretization on the rest of the computational domain. A naive choice of restriction at $i = 0$ would be the transpose of interpolation, or the full-weighting restriction. However, this leads to a poor coarse grid estimated error. In order to find the correct restriction, we investigate two cases.

Case 1: $h = \frac{1}{36}$ and $\sqrt{h} = 6h$. Figure 6a shows the discretization and coarsening. In this case, the wide stencil discretization at $x = 0$ reads

$$\frac{-u_{-6} + 2u_0 - u_6}{(6h)^2} = 0.$$

The residual at $i = 0$ is the given by

$$r_0 = \frac{-e_{-6} + 2e_0 - e_6}{(6h)^2}. \tag{18}$$

We notice that $i = 0$, $i = -6$ and $i = 6$ are all C-points. Then a natural construction of the coarse grid problem at $i = 0$ is to discretized the Poisson equation using these three points, or more precisely,

$$\frac{-e^H_{-6} + 2e^H_0 - e^H_6}{(6h)^2} = r^H_0, \tag{19}$$

where the left hand side is a discretization of the Poisson equation on the coarse grid with the stencil length $6h$, and the right hand side is the coarse grid residual r_0^H. Comparing (18) and (19), we can see that the restriction at $i = 0$ is a simple injection:

$$r_0^H \equiv r_0.$$

Case 2: $h = \frac{1}{49}$ and $\sqrt{h} = 7h$. Figure 6b shows the discretization and coarsening. Unlike the previous case, here the points that $i = 0$ connects to, namely, $i = 7$ and $i = -7$, are both F-points. To discretize the Poisson equation on the coarse grid, we interpolate the errors at $i = 7$ and $i = -7$ from their neighboring C-points, which gives

$$\frac{-\frac{1}{2}\left(e_{-8}^H + e_{-6}^H\right) + 2e_0^H - \frac{1}{2}\left(e_6^H + e_8^H\right)}{(7h)^2} = r_0^H. \tag{20}$$

We want to find a restriction such that the right hand side r_0^H matches the left hand side of (20). One scheme is to use the linear combination of the following fine grid residuals:

$$r_0 = \frac{-e_{-7} + 2e_0 - e_7}{(7h)^2},$$
$$r_7 = \frac{-e_6 + 2e_7 - e_8}{h^2}, \quad r_{-7} = \frac{-e_{-6} + 2e_{-7} - e_{-8}}{h^2}.$$

If we combine r_0, r_7 and r_{-7} as follows

$$r_0 + \frac{1}{98}r_7 + \frac{1}{98}r_{-7} = \frac{-\frac{1}{2}(e_{-8} + e_{-6}) + 2e_0 - \frac{1}{2}(e_6 + e_8)}{(7h)^2}, \tag{21}$$

then it matches the left hand side of (20) in the exact sense. Equation (21) defines a possible restriction. We note that the restriction (21) makes use of the residuals r_7 and r_{-7}, which are the points that the wide stencil point $i = 0$ connects to. This is different from standard full-weighting restriction that uses the neighboring points r_1 and r_{-1}. Since the coefficients of r_7 and r_{-7} are small, we simply drop them and yield again an injection:

$$r_0^H \equiv r_0.$$

More generally, given a wide stencil C-point $i \in C$ with a stencil length $\sqrt{h}$, the non-zero restriction weights come from the set of the F-points that it connects to, denoted as $\{j | j \in F, A_{i,j} \neq 0\}$. We can show that the restriction weights are

$$w_{i,j} = -\frac{A_{i,j}}{A_{j,j}} = -\frac{-\frac{1}{(\sqrt{h})^2}}{\frac{2}{h^2}} = \frac{h}{2}. \tag{22}$$

When h is small, the restriction (22) can be left out. In other words, injection is sufficient for a good coarse grid problem.

We extend the injection at the wide stencil C-points from the one-dimensional Poisson equation to the two-dimensional HJB equation. This gives rise to a restriction operator R that is no longer the transpose of the interpolation. Once the restriction operator is specified, we define the coarse grid operator as

$$A_{2h} \equiv R A_h P. \tag{23}$$

Since $R \neq P^T$, it results in the Petrov–Galerkin coarse grid operator.

We will show in Sect. 6.2 that injection for the wide stencil C-points would lead to a good approximation of the coarse grid error estimate and eventually a mesh-independent convergence rate for the proposed multigrid method.

Another benefit of using injection is its simplicity and the sparsity of the resulting restriction operator. For AMG, the coarse grid operator A_{2h} could become much denser than A_h, which causes a high computational complexity. When we use injection for wide stencils, combined with the 7-point restriction and interpolation for the standard 7-point stencils, the resulting coarse grid operator (23) will still remain sparse.

5 Smoothing analysis

In this section, we analyze the smoothing property of the four directional alternating line smoother by local Fourier analysis (LFA) [40]. For each policy iteration, the control $(a_{i,j}, \theta_{i,j})$ is fixed, and a linear four directional alternating line smoother is applied. As the smoothing property for the mixed stencil discretization is complicated, this section only analyzes the standard 7-point stencil discretization. We first analyze the x-line smoother for the 7-point stencil operator $L_h^{[1]}$, defined by (6). The splitting of $L_h^{[1]}$ reads

$$\left(L_h^{[1]}\right)_{x+} \equiv \frac{1}{h^2}\begin{bmatrix} 0 & 0 & 0 \\ -(\alpha_{11} - \alpha_{12}) & 2(\alpha_{11} + \alpha_{22} - \alpha_{12}) & -(\alpha_{11} - \alpha_{12}) \\ -\alpha_{12} & -(\alpha_{22} - \alpha_{12}) & 0 \end{bmatrix}_h,$$
$$\left(L_h^{[1]}\right)_{x-} \equiv \frac{1}{h^2}\begin{bmatrix} 0 & -(\alpha_{22} - \alpha_{12}) & -\alpha_{12} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}_h,$$

where we have suppressed the dependencies of α_{11}, α_{22} and α_{12} on $(a_{i,j}, \theta_{i,j})$. This gives rise to the symbol of the x-line smoother:

$$\widetilde{\left(S_h^{[1]}\right)}_x(\kappa; a_{i,j}, \theta_{i,j})$$
$$= \left[(\alpha_{22} - \alpha_{12})e^{i\kappa_2} + \alpha_{12}e^{i(\kappa_1+\kappa_2)}\right] \Big/$$
$$\left[2(\alpha_{11} + \alpha_{22} - \alpha_{12}) - (\alpha_{22} - \alpha_{12})e^{-i\kappa_2}\right.$$

$$- (\alpha_{11} - \alpha_{12})(e^{-i\kappa_1} + e^{i\kappa_1}) - \alpha_{12} e^{-i(\kappa_1+\kappa_2)}\Big].$$

One can obtain the symbols for y-line smoother, first diagonal line smoother (top left to bottom right) and second diagonal line smoother (top right to bottom left) in a similar fashion. The symbol for the four-directional alternating line smoother $\widetilde{S_h^{[1]}}(\kappa; a_{i,j}, \theta_{i,j})$ is then the product of the symbols of the four line smoothers. Define the smoothing factor of $\widetilde{S_h^{[1]}}(\kappa; a_{i,j}, \theta_{i,j})$ as

$$\mu_{loc}^{[1]}(a_{i,j}, \theta_{i,j}) \equiv \sup_{\kappa} \left\{ |\widetilde{S_h^{[1]}}(\kappa; a_{i,j}, \theta_{i,j})| : \kappa \in \text{high frequency mode} \right\}.$$

Furthermore, we note that $L_h^{[1]}$ and $L_h^{[2]}$ are applied under Condition (5) and (7) respectively. Hence, we define the smoothing factor of the standard 7-point discretization as

$$\mu_{loc}(a_{i,j}, \theta_{i,j}) \equiv \begin{cases} \mu_{loc}^{[1]}(a_{i,j}, \theta_{i,j}), & \text{if } (a_{i,j}, \theta_{i,j}) \text{ satisfies (5)}, \\ \mu_{loc}^{[2]}(a_{i,j}, \theta_{i,j}), & \text{if } (a_{i,j}, \theta_{i,j}) \text{ satisfies (7)}. \end{cases} \tag{24}$$

We evaluate the smoothing factor (24) for each of the four individual line smoothers and their combination. Figure 7 shows the following

- x line smoother has a good smoothing property except at $a \to 0$, or $(a, \theta) \to \left(1, \frac{\pi}{4}\right), \left(1, -\frac{\pi}{4}\right)$.
- y line smoother has a good smoothing property except at $a \to 1$, or $(a, \theta) \to \left(0, \frac{\pi}{4}\right), \left(0, -\frac{\pi}{4}\right)$.
- First-diagonal line smoother has a good smoothing property except at $a \to 0$ or $a \to 1$. However, the smoother is still effective at $(a, \theta) \to \left(0, -\frac{\pi}{4}\right), \left(1, \frac{\pi}{4}\right)$.
- Second-diagonal line smoother has a good smoothing property except at $a \to 0$ or $a \to 1$. However, the smoother is still effective at $(a, \theta) \to \left(0, \frac{\pi}{4}\right), \left(1, -\frac{\pi}{4}\right)$.

Four-directional alternating line smoother combines these four smoothers together. Figure 7 shows that the smoothing factor of the combined smoother is bounded by 0.5 for any $(a, \theta) \in [0, 1] \times \left[-\frac{\pi}{4}, \frac{\pi}{4}\right)$. Hence, the combined smoother has a good smoothing property.

6 Numerical results

In this section, we demonstrate the mesh-independent convergence rates of the proposed multigrid methods for Monge–Ampère equation. The examples are mainly from [7,20].

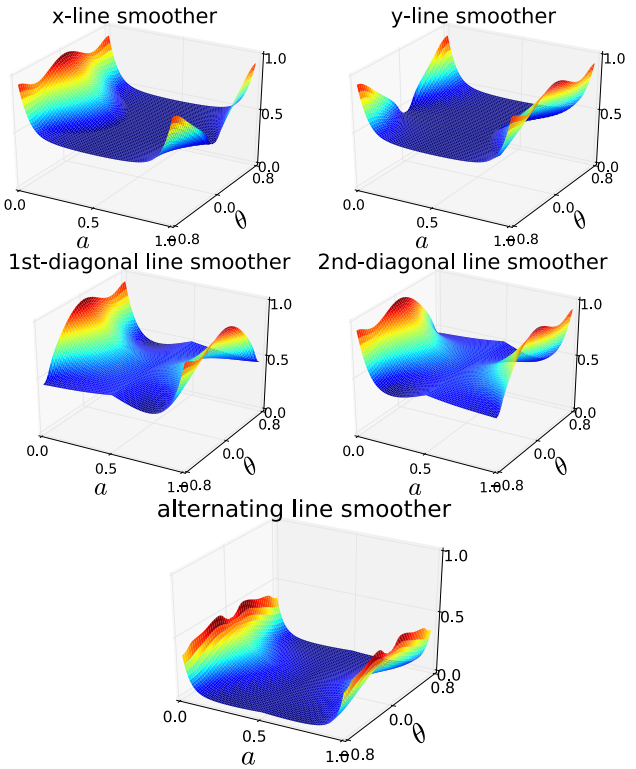

Fig. 7 Smoothing factor $\mu_{loc}(a_{i,j}, \theta_{i,j})$ for **a** x line smoother, **b** y line smoother, **c** first diagonal line smoother, **d** second diagonal line smoother, **e** four-directional alternating line smoother

6.1 Multigrid for standard 7-point stencil discretization

In this subsection, we consider examples where the standard 7-point stencil discretization is applied on the entire computational domain. We compare the performance of two families of multigrid methods. One family of multigrid methods is "policy-MG iteration"; see Sect. 3.1. More specifically, policy iteration is used to solve the nonlinear discretized system (12). At the k-th policy iteration, multigrid V(1,1)-cycle is applied to solve the linearized system under the fixed control $(a^{(k)}, \theta^{(k)})$. The tolerances for outer policy iteration and inner multigrid V-cycle are 10^{-6} and 10^{-7} respectively. We apply pointwise Gauss–Seidel smoother in Example 1 and four-directional alternating Gauss–Seidel line smoother in Example 2. We apply standard full coarsening, and the 7-point restriction and interpolation, as described in Sect. 3.3. Galerkin coarse grid operator is used to construct coarse grid problems.

The other family of multigrid methods is the full approximation scheme (FAS). The multigrid components are basically the same as the policy-MG iteration, except that the value iteration is used as a nonlinear smoother, and direct discretization is used for coarse grid operator; see Sect. 3.4 for details.

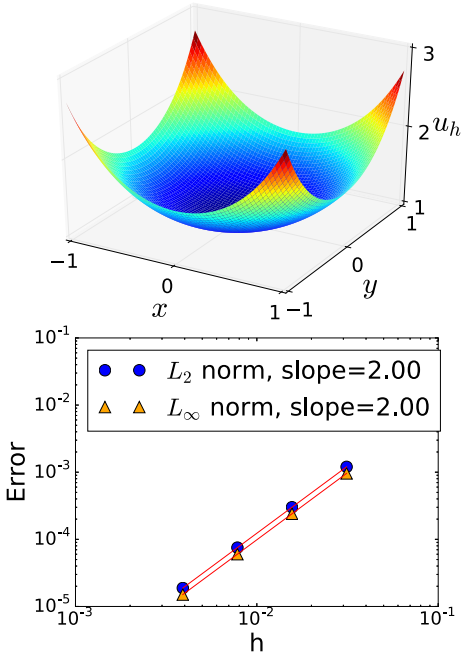

Fig. 8 (1) Numerical solution for Example 1. (2) Norms of the error $\|u - u_h\|$. The convergence rates, indicated by the slopes, are $O(h^2)$ in both L_2 and L_∞ norms

Example 1 We consider the multigrid methods for solving

$$f(x, y) = (1+x^2+y^2)e^{x^2+y^2}, \quad g(x, y) = e^{\frac{1}{2}(x^2+y^2)} \tag{25}$$

on $\Omega = [-1, 1] \times [-1, 1]$, where the viscosity solution $u(x, y) = e^{\frac{1}{2}(x^2+y^2)}$ is smooth. Figure 8 shows that the numerical solution converges at the optimal theoretical convergence rate $O(h^2)$.

Next we show the convergence rate of the policy-MG iteration for solving the nonlinear discretized system (12). This example turns out to be isotropic, so it suffices to apply the less expensive pointwise Gauss-Seidel smoother. Table 1 reports the convergence rates. The number of policy iterations is 3. The number of multigrid V-cycles within each policy iteration ranges from 2 to 9. The total number of multigrid V-cycles for the entire nonlinear system is 17–19, independent of mesh size. As a side remark, the number of multigrid V-cycles decreases as policy iteration approaches the final solution. The reason is that we let the k-th solution of the policy iteration, $u_h^{(k)}$, be the initial guess of the $(k+1)$-th multigrid V-cycle for solving $u_h^{(k+1)}$. As policy iteration converges, $u_h^{(k)} \to u_h^{(k+1)}$, which means that the initial guess of each multigrid V-cycle becomes more and more precise.

Table 2 Convergence of the full approximation scheme for Example 1

$n \times n$	Total number of multigrid V-cycles
32×32	8
64×64	8
128×128	9
256×256	9

Pointwise Gauss–Seidel smoother is applied

We compare the policy-MG iteration with the FAS iteration. Table 2 shows that the total number of the FAS V-cycles is 8–9 and is independent of mesh size. We note that for both policy-MG and FAS iterations, the computational cost per multigrid iteration is approximately the same. Hence, full approximation scheme is less expensive and converges faster.

Example 2 We consider multigrid method for solving

$$f(x, y) = 1+24(x+y)^4, \quad g(x, y) = \frac{1}{2}(x^2+y^2)+(x+y)^4$$

on $\Omega = [-1, 1] \times [-1, 1]$, where the viscosity solution $u(x, y) = \frac{1}{2}(x^2 + y^2) + (x + y)^4$ is smooth. Similar to Example 1, as the mesh size $h \to 0$, the numerical solution converges towards the viscosity solution at the optimal theoretical convergence rate $O(h^2)$; see Fig. 9.

Table 3 reports the convergence of the policy-MG iteration. The number of policy iterations is approximately 6. The multigrid V-cycle with the alternating line smoother converges at 20–32 iterations in total, which is approximately independent of mesh size. Conversely, standard multigrid V-cycle converges at more than 70 iterations, and the number of iterations is not mesh-independent. This is because the example is an-isotropic along the diagonal direction. We note that in general, the direction of the anisotropy may change among different patches of the computational domain, so

Table 1 Convergence of the policy-MG iteration for Example 1

$n \times n$	Number of policy iterations	Number of multigrid V-cycles within each policy iteration	Total number of multigrid V-cycles
32×32	3	8, 7, 2	17
64×64	3	9, 7, 3	19
128×128	3	9, 7, 3	19
256×256	3	9, 7, 3	19

Pointwise Gauss–Seidel smoother is applied

Table 3 Convergence of the policy-MG iteration for Example 2

$n \times n$	Number of policy iterations	Proposed policy-MG		Standard policy-MG
		Number of multigrid V-cycles within each policy iteration	Total number of multigrid V-cycles	Total number of multigrid V-cycles
32×32	5	5, 5, 5, 3, 2	20	73
64×64	6	5, 6, 6, 4, 2, 1	24	94
128×128	6	6, 6, 7, 5, 3, 1	28	129
256×256	6	7,7,7,6,3,1	32	161

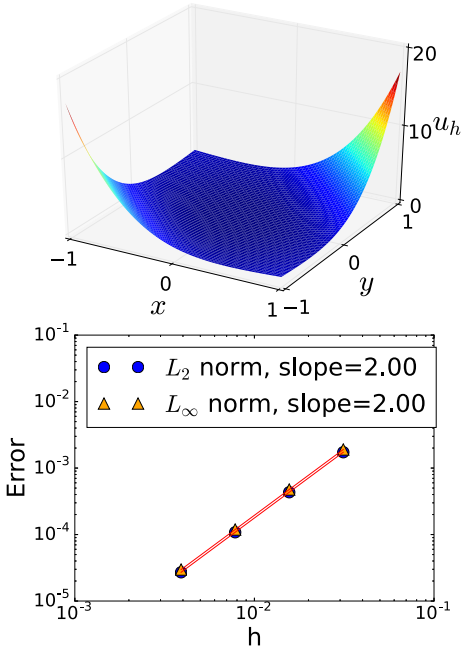

Fig. 9 (1) Numerical solution for Example 2. (2) Norms of the error $\|u - u_h\|$. The convergence rates, indicated by the slopes, are $O(h^2)$ in both L_2 and L_∞ norms

we implement four-directional alternating line smoother in a non-discriminatory fashion.

6.2 Multigrid for mixed stencil discretization

In this section, we illustrate the multigrid convergence rates for the mixed stencil discretization. The examples are solved by the policy-MG iterations. Four-directional alternating line smoother is applied. Regarding the coarse grid correction, we treat standard 7-point stencil patches and semi-Lagrangian wide stencil patches separately. For standard 7-point stencil patches, we apply standard full coarsening, and the 7-point restriction and interpolation. For wide stencil patches, we set all the wide stencil points as coarse grid points, and use injection for restriction. Petrov-Galerkin coarse grid operator is used for constructing coarse grid problems.

Example 3 We consider solving the linearized HJB equation (17), where f and g are given in (25). A monotone discretization is to apply the standard 5-point stencil discretization everywhere and wide stencil at the origin. We compare the performance of (i) proposed multigrid method, (ii) standard multigrid with four-directional alternating line smoother, and (iii) standard multigrid with pointwise Gauss–Seidel smoother. The only difference between Method (i) and (ii) is that injection is applied at the wide stencil point for Method (i), while full-weighting restriction is applied for Method (ii). Table 4 shows that Method (iii) has poor convergence. Method (ii) changes the smoother to the alternating line smoother, but the convergence rate still grows as N increases. Method (i) uses injection at the wide stencil points. The convergence rate is around 5–6, independent of N.

Figure 10 explains the reason by plotting the cross section of the error for a one-step two-grid algorithm. After pre-smoothing, the error (blue lines) becomes smooth everywhere, except that a kink appears at the wide stencil point. Figure 10 (top) uses injection at the wide stencil point. The coarse grid problem yields an accurate coarse grid estimated errors (red lines), where all the C-points are marked by the red dots. This eliminates the error effectively. Conversely, under the same smoother, if full-weighting is used at the wide stencil, as shown in Fig. 10 (bottom), the coarse grid estimated error is no longer a good approximation of the smoothed error.

Example 4 We consider

$$f(x, y) = \max\left(1 - \frac{0.15}{\sqrt{x^2 + y^2}}, 0\right),$$
$$g(x, y) = \frac{1}{2}(\sqrt{x^2 + y^2} - 0.15)^2$$

on $\Omega = [-0.5, 0.5] \times [-0.5, 0.5]$. The viscosity solution is given by $u(x, y) = \frac{1}{2}\max\left(\sqrt{x^2 + y^2} - 0.15, 0\right)^2$. This is a C^1 function where the solution is not smooth at the ring

Table 4 Convergence of linear multigrid V-cycle for Example 3

$n \times n$	(i) Proposed MG	(ii) Standard MG (alternating line smoother)	(iii) Standard MG (pointwise smoother)
32×32	5	7	23
64×64	5	9	46
128×128	6	12	198
256×256	6	17	*

The tested multigrid methods are (i) proposed multigrid method, where wide stencil is restricted by injection, (ii) standard multigrid with four-directional alternating line smoother, and (iii) standard multigrid with pointwise Gauss–Seidel smoother. The star represents 200+ iterations

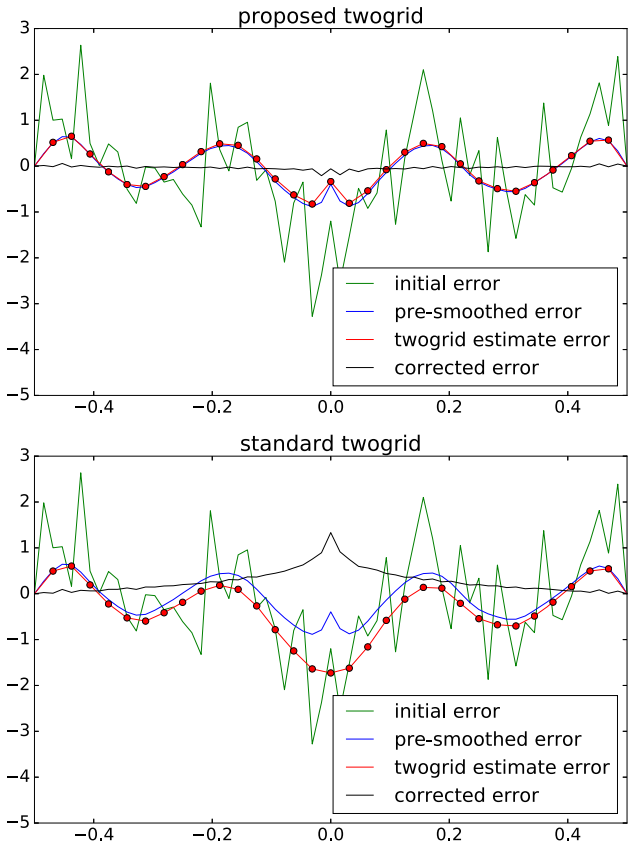

Fig. 10 Cross section of the error along x-axis for a one-step two-grid algorithm. (top) Proposed two-grid algorithm, where injection is used at the origin (the wide stencil point). The coarse grid problem yields an accurate coarse grid estimated error, where all the C-points are marked by the red dots. (bottom) Standard two-grid algorithm, where full-weighting restriction is used

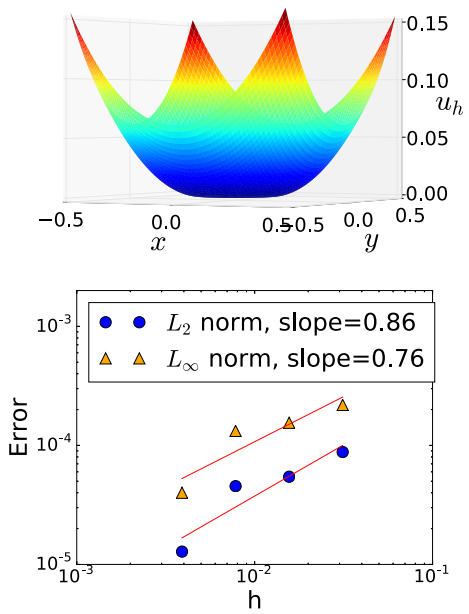

Fig. 11 (1) Numerical solution for Example 4. The solution is not smooth at the ring $x^2 + y^2 = 0.15^2$. (2) Norms of the error $\|u - u_h\|$. The convergence rates are indicated by the slopes in both L_2 and L_∞ norms

$x^2 + y^2 = 0.15^2$. Semi-Lagrangian wide stencils are applied near the ring. First we check the convergence of the numerical solution u_h towards the viscosity solution u as $h \to 0$. References [14,18] have proved that the convergence rate is at most $O(h)$ if wide stencil discretization is applied. Figure 11 shows that the convergence rate is approximately 0.76–0.86, although convergence can be clearly observed as $h \to 0$. Indeed, slow convergence for singular solutions is observed in most of the existing numerical schemes in the literature. However, we note that unlike the other existing schemes, the mixed discretization is proved to converge to the viscosity solution (instead of the other weak solutions) [14,18].

Table 5 reports the convergence of the policy-MG iteration. We remark that the number of the outer policy iterations increases from 5 to 10 as n increases from 32 to 256. Such increase of outer iteration is related to nonlinearity. Also, this example contains a singular S^1 manifold, which in general worsens the outer convergence behavior.

To compare the number of multigrid V-cycles among different mesh sizes fairly, we compute the average number of multigrid V-cycles per policy iteration. Table 5 reports the convergence rate. The number of the multigrid V-cycles is approximately a constant ranging from 3.0 to 4.2 as n increases from 32 to 256. Hence, the inner multigrid V-cycle for solving linearized systems is close to mesh-independence.

Table 5 Convergence of the policy-MG iteration for Example 4

$n \times n$	Number of policy iterations	Number of multigrid V-cycles within each policy iteration	Average number of multigrid V-cycles per policy iteration
32×32	5	4, 5, 3, 2, 1	3.0
64×64	5	4, 6, 3, 2, 1	3.2
128×128	6	5, 6, 4, 3, 3, 2	3.8
256×256	10	6, 6, 6, 6, 5, 4, 3, 3, 2, 1	4.2

Table 6 Convergence of the policy-MG iteration for Example 5

$n \times n$	Number of policy iterations	Number of multigrid V-cycles within each policy iteration	Average number of multigrid V-cycles per policy iteration
32×32	7	4, 5, 5, 4, 4, 4, 2	4.0
64×64	10	5, 6, 6, 7, 5, 4, 4, 3, 2, 1	4.3
128×128	12	5, 7, 8, 7, 7, 6, 5, 5, 4, 2, 1, 1	4.8
256×256	17	6, 7, 8, 8, 8, 9, 8, 7, 6, 5, 5, 4, 4, 4, 2, 1, 1	5.5

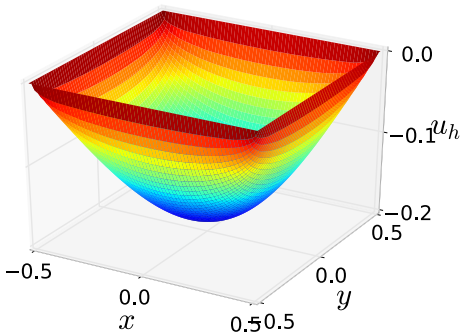

Fig. 12 Numerical solution for Example 5. u_{xx} or u_{yy} approaches ∞ near the boundary

Example 5 We consider an example from [7]:

$$f(x, y) = 1, \quad g(x, y) = 0$$

on $\Omega = [-0.5, 0.5] \times [-0.5, 0.5]$. For this example, the exact solution is not available. However, it can be inferred that either $u_{xx} \to \infty$ or $u_{yy} \to \infty$ along $\partial\Omega$ [7]. As a result, wide stencils are applied near $\partial\Omega$. Figure 12 shows the numerical solution.

Table 6 shows that the number of the outer policy iterations increases from 7 to 17 as n increases from 32 to 256. Here we once again focus on the mesh-independent convergence for the inner multigrid V-cycle. The number of the multigrid V-cycles is approximately a constant ranging from 4.0 to 5.5 as n increases from 32 to 256, close to mesh-independence.

7 Conclusion

We propose multigrid methods for solving the mixed stencil discretization of the Monge–Ampère equation. We investigate two scenarios. One scenario is when the standard 7-point stencil discretization is applied on the entire computational domain. Full approximation scheme gives the optimal mesh-independent convergence. The other scenario we investigate is the general mixed stencil discretization. Policy-MG iteration is used. We set all wide stencil points as coarse grid points and propose injection for wide stencil points. We use Fourier analysis to show that four-directional alternating line smoother has a good smoothing property, especially for anisotropic equation. The resulting multigrid methods can converge at the mesh-independent rate.

Extending the full approximation scheme to the mixed stencil discretization and doing corresponding Fourier analysis would be interesting future works.

References

1. Akian, M., Séquier, P., Sulem, A.: A finite horizon multidimensional portfolio selection problem with singular transactions. In: Decision and Control, 1995., Proceedings of the 34th IEEE Conference on, vol. 3, pp. 2193–2198. IEEE (1995)
2. Akian, M., Sulem, A., Taksar, M.I.: Dynamic optimization of long-term growth rate for a portfolio with transaction costs and logarithmic utility. Math. Finance **11**(2), 153–188 (2001). doi:10.1111/1467-9965.00111
3. Azimzadeh, P., Forsyth, P.A.: Weakly Chained Matrices, Policy Iteration, and Impulse Control. SIAM J. Numer. Anal. **54**(3), 1341–1364 (2016). doi:10.1137/15M1043431

4. Bank, R.E., Wan, J.W.L., Qu, Z.: Kernel preserving multigrid methods for convection–diffusion equations. SIAM J. Matrix Anal. Appl. **27**(4), 1150–1171 (2006). doi:10.1137/040619533
5. Barles, G., Souganidis, P.E.: Convergence of approximation schemes for fully nonlinear second order equations. Asymptot. Anal. **4**(3), 271–283 (1991)
6. Benamou, J.D., Collino, F., Mirebeau, J.M.: Monotone and consistent discretization of the Monge–Ampère operator. Math. Comput. **85**(302), 2743–2775 (2016). doi:10.1090/mcom/3080
7. Benamou, J.D., Froese, B.D., Oberman, A.M.: Two numerical methods for the elliptic Monge–Ampère equation. M2AN. Math. Model. Numer. Anal. **44**(4), 737–758 (2010). doi:10.1051/m2an/2010017
8. Bey, J., Wittum, G.: Downwind numbering: Robust multigrid for convection–diffusion problems. Appl. Numer. Math. **23**(1), 177–192 (1997). doi:10.1016/S0168-9274(96)00067-0
9. Bloss, M., Hoppe, R.H.W.: Numerical computation of the value function of optimally controlled stochastic switching processes by multi-grid techniques. Numer. Funct. Anal. Optim. **10**(3–4), 275–304 (1989). doi:10.1080/01630568908816304
10. Bokanowski, O., Maroso, S., Zidani, H.: Some convergence results for Howard's algorithm. SIAM J. Numer. Anal. **47**(4), 3001–3026 (2009). doi:10.1137/08073041X
11. Brandt, A.: Multi-level adaptive solutions to boundary-value problems. Math. Comput. **31**(138), 333–390 (1977)
12. Brandt, A., Yavneh, I.: On multigrid solution of high-Reynolds incompressible entering flows. J. Comput. Phys. **101**(1), 151–164 (1992). doi:10.1016/0021-9991(92)90049-5
13. Caffarelli, L.A., Milman M. (eds): Monge–Ampère equation: applications to geometry and optimization. In: Contemporary Mathematics, vol. 226. American Mathematical Society, Providence, RI (1999). doi:10.1090/conm/226
14. Chen, Y., Wan, J.W.: Monotone mixed narrow/wide stencil finite difference scheme for Monge–Ampère equation. ArXiv preprint arXiv:1608.00644 (2016)
15. Crandall, M.G., Ishii, H., Lions, P.L.: User's guide to viscosity solutions of second order partial differential equations. Bull. Am. Math. Soc. (N.S.) **27**(1), 1–67 (1992). doi:10.1090/S0273-0979-1992-00266-5
16. Crandall, M.G., Lions, P.L.: Viscosity solutions of Hamilton–Jacobi equations. Trans. Am. Math. Soc. **277**(1), 1–42 (1983). doi:10.2307/1999343
17. Debrabant, K., Jakobsen, E.R.: Semi-Lagrangian schemes for linear and fully non-linear diffusion equations. Math. Comput. **82**(283), 1433–1462 (2013). doi:10.1090/S0025-5718-2012-02632-9
18. Feng, X., Jensen, M.: Convergent semi-Lagrangian methods for the Monge–Ampère equation on unstructured grids. ArXiv preprint arXiv:1602.04758 (2016)
19. Forsyth, P.A., Labahn, G.: Numerical methods for controlled Hamilton–Jacobi–Bellman PDEs in finance. J. Comput. Finance **11**(2), 1 (2007)
20. Froese, B.D., Oberman, A.M.: Convergent finite difference solvers for viscosity solutions of the elliptic Monge–Ampère equation in dimensions two and higher. SIAM J. Numer. Anal. **49**(4), 1692–1714 (2011). doi:10.1137/100803092
21. Froese, B.D., Oberman, A.M.: Fast finite difference solvers for singular solutions of the elliptic Monge–Ampère equation. J. Comput. Phys. **230**(3), 818–834 (2011). doi:10.1016/j.jcp.2010.10.020
22. Froese, B.D., Oberman, A.M.: Convergent filtered schemes for the Monge–Ampère partial differential equation. SIAM J. Numer. Anal. **51**(1), 423–444 (2013). doi:10.1137/120875065
23. Han, D., Wan, J.W.L.: Multigrid methods for second order Hamilton–Jacobi–Bellman and Hamilton–Jacob–Bellman–Isaacs equations. SIAM J. Sci. Comput. **35**(5), S323–S344 (2013). doi:10.1137/120881476
24. Hoppe, R.H.W.: Multigrid methods for Hamilton–Jacobi–Bellman equations. Numer. Math. **49**(2–3), 239–254 (1986). doi:10.1007/BF01389627
25. Howard, R.A.: Dynamic Programming and Markov Processes. The Technology Press of M.I.T., Cambridge, Mass.; John Wiley & Sons, Inc., New York, London (1960)
26. Jameson, A.: Solution of the Euler equations for two-dimensional transonic flow by a multigrid method. Appl. Math. Comput. **13**(3–4), 327–355 (1983). doi:10.1016/0096-3003(83)90019-X
27. Krylov, N.V.: The control of the solution of a stochastic integral equation. Teor. Verojatnost. i Primenen. **17**, 111–128 (1972)
28. Lin, J.: Wide Stencil for the Monge–Ampère Equation. Technical Report, University of Waterloo Master Essay, Supervised by Justin WL Wan, Available on https://uwaterloo.ca/computational-mathematics/sites/ca.computational-mathematics/files/uploads/files/cmmain1.pdf (2014)
29. Lions, P.L.: Hamilton–Jacobi–Bellman equations and the optimal control of stochastic systems. In: Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Warsaw, 1983), pp. 1403–1417. PWN, Warsaw (1984)
30. Ma, K., Forsyth, P.A.: An unconditionally monotone numerical scheme for the two-factor uncertain volatility model. IMA J. Numer. Anal. **37**(2), 905–944 (2016)
31. Napov, A., Notay, Y.: An algebraic multigrid method with guaranteed convergence rate. SIAM J. Sci. Comput. **34**(2), A1079–A1109 (2012)
32. Notay, Y.: An aggregation-based algebraic multigrid method. Electron. Trans. Numer. Anal. **37**, 123–146 (2010)
33. Oberman, A.M.: Wide stencil finite difference schemes for the elliptic Monge–Ampère equation and functions of the eigenvalues of the Hessian. Discrete Contin. Dyn. Syst. Ser. B **10**(1), 221–238 (2008). doi:10.3934/dcdsb.2008.10.221
34. Reisinger, C., Arto, J.R.: Boundary treatment and multigrid preconditioning for semi-Lagrangian schemes applied to Hamilton–Jacobi–Bellman equations. ArXiv preprint arXiv:1605.04821 (2016)
35. Ruge, J.W., Stüben, K.: Algebraic multigrid. In: Multigrid Methods, Frontiers in Applied Mathematics, vol. 3, pp. 73–130. SIAM, Philadelphia, PA (1987)
36. Seibold, B.: Performance of algebraic multigrid methods for non-symmetric matrices arising in particle methods. Numer. Linear Algebra Appl. **17**(2–3), 433–451 (2010). doi:10.1002/nla.710
37. Smears, I.: Hamilton–Jacobi–Bellman Equations Analysis and Numerical Analysis. Technical Report, Research Report Available on www.math.dur.ac.uk/Ug/projects/highlights/PR4/Smears_HJB_report.pdf
38. Stüben, K.: An introduction to algebraic multigrid. In: Trottenberg, U., Oosterlee, C.W., Schüller, A. (eds.) Multigrid, pp. 413–532. Academic Press Inc, San Diego, CA (2001)
39. Stüben, K.: A review of algebraic multigrid. J. Comput. Appl. Math. **128**(1–2), 281–309 (2001). doi:10.1016/S0377-0427(00)00516-1. Numerical analysis 2000, Vol. VII, Partial differential equations
40. Trottenberg, U., Oosterlee, C.W., Schüller, A.: Multigrid. Academic Press Inc, San Diego, CA (2001)
41. Wang, J., Forsyth, P.A.: Maximal use of central differencing for Hamilton–Jacobi–Bellman PDEs in finance. SIAM J. Numer. Anal. **46**(3), 1580–1601 (2008). doi:10.1137/060675186
42. Xu, J., Zikatanov, L.T.: Algebraic multigrid methods. ArXiv preprint arXiv:1611.01917 (2016)